

Implementing the IBM Storwize V5000 Gen2 (including the Storwize V5010, V5020, and V5030) with IBM Spectrum Virtualize V8.2.1

Jon Tate

Jack Armstrong

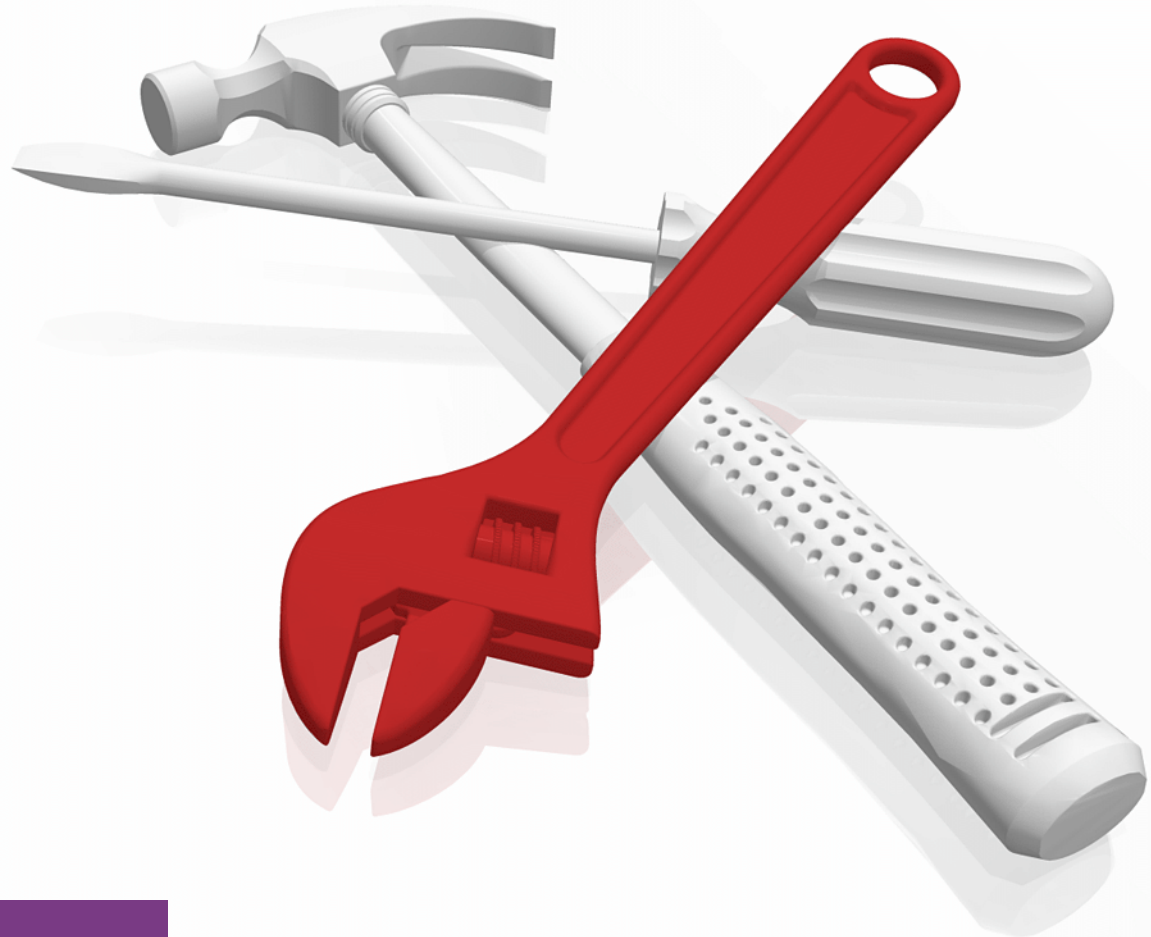
Tiago Bastos

Sergey Kubin

Hartmut Lonzer

Danilo Miyasiro

Rodrigo Suzuki



Storage



International Technical Support Organization

**Implementing the IBM Storwize V5000 Gen2 with IBM
Spectrum Virtualize V8.2.1**

April 2019

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

Fifth Edition (April 2019)

This edition applies to the IBM Storwize V5000 Gen2 hardware and software V8.2.1 and was based on pre-GSA code. Note that since this book was produced, several panels might have changed.

© Copyright International Business Machines Corporation 2019. All rights reserved.

Note to U.S. Government Users Restricted Rights -- Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

Contents

Notices	xiii
Trademarks	xiv
Preface	xv
Authors	xv
Now you can become a published author, too	xvii
Comments welcome	xviii
Stay connected to IBM Redbooks	xviii
Summary of changes	xix
April 2019, Fifth Edition	xix
Chapter 1. Overview of the IBM Storwize V5000 Gen2 system.	1
1.1 IBM Storwize V5000 Gen2 overview	2
1.2 IBM Storwize V5000 Gen2 terminology	4
1.3 IBM Storwize V5000 Gen2 models	6
1.3.1 IBM Storage Utility Offerings	10
1.4 IBM Storwize V5000 Gen1 and Gen2 compatibility	11
1.5 IBM Storwize V5000 Gen2 hardware	12
1.5.1 Control enclosure	13
1.5.2 Storwize V5010	14
1.5.3 Storwize V5020	15
1.5.4 Storwize V5030	16
1.5.5 Expansion enclosure	17
1.5.6 Host interface cards	18
1.5.7 Disk drive types	19
1.6 IBM Storwize V5000 Gen2 terms	19
1.6.1 Hosts	20
1.6.2 Node canister	20
1.6.3 I/O groups	20
1.6.4 Clustered system	21
1.6.5 RAID	21
1.6.6 Managed disks	22
1.6.7 Quorum disks	22
1.6.8 Storage pools	23
1.6.9 Volumes	25
1.6.10 iSCSI	27
1.6.11 Serial-attached SCSI	28
1.6.12 Fibre Channel	28
1.7 IBM Storwize V5000 Gen2 features	28
1.7.1 Mirrored volumes	28
1.7.2 Thin provisioning	30
1.7.3 Real-time Compression	31
1.7.4 Deduplication	32
1.7.5 Easy Tier	33
1.7.6 Storage Migration	33
1.7.7 FlashCopy	33
1.7.8 Remote Copy	34
1.7.9 IP replication	35

1.7.10 External virtualization	36
1.7.11 Encryption	36
1.8 Problem management and support	36
1.8.1 IBM Support assistance	36
1.8.2 Event notifications	36
1.8.3 SNMP traps	37
1.8.4 Syslog messages	37
1.8.5 Call Home email	37
1.9 More information resources	37
1.9.1 Useful IBM Storwize V5000 Gen2 websites	38
Chapter 2. Initial configuration	39
2.1 Hardware installation planning	40
2.1.1 Procedure to install the SAS cables	41
2.2 SAN configuration planning	44
2.3 FC direct-attach planning	46
2.4 SAS direct-attach planning	48
2.5 LAN configuration planning	50
2.5.1 Management IP address considerations	51
2.5.2 Service IP address considerations	52
2.6 Host configuration planning	52
2.6.1 Fibre Channel connection	52
2.6.2 iSCSI configuration	52
2.7 Miscellaneous configuration planning	53
2.8 System management	54
2.8.1 Graphical user interface	54
2.8.2 Command-line interface	55
2.9 First-time setup	56
2.10 Initial configuration	61
2.10.1 Adding enclosures after the initial configuration	76
2.10.2 Service Assistant Tool	82
Chapter 3. Graphical user interface overview	85
3.1 Overview of IBM Spectrum Virtualize management software	86
3.1.1 Accessing the storage management software	86
3.1.2 System pane layout	88
3.1.3 Navigation	92
3.1.4 Multiple selection	93
3.1.5 Status indicators area	94
3.2 Monitoring menu	95
3.2.1 System overview	97
3.2.2 System details	101
3.2.3 Events option	104
3.2.4 Performance pane	105
3.2.5 Background Task	107
3.3 Pools menu	107
3.3.1 Pools view	110
3.3.2 Child pools	114
3.3.3 Volumes by pool	115
3.3.4 Internal storage	117
3.3.5 External storage	117
3.3.6 MDisks by pools	121
3.3.7 System migration	122

3.4 Volumes menu	123
3.4.1 All volumes	125
3.4.2 Volumes by pool	128
3.4.3 Volumes by host	128
3.5 Hosts menu	128
3.5.1 Hosts	129
3.5.2 Host clusters	131
3.5.3 Ports by host	132
3.5.4 Host mappings	133
3.5.5 Volumes by host	134
3.6 Copy services	134
3.6.1 IBM FlashCopy	135
3.6.2 Consistency groups	136
3.6.3 FlashCopy mappings	139
3.6.4 Remote copy	141
3.6.5 Partnerships	142
3.7 Access menu	143
3.7.1 Users	144
3.7.2 Audit Log option	146
3.8 Settings menu	147
3.8.1 Notifications	147
3.8.2 Network	148
3.8.3 Security features	149
3.8.4 System menu	152
3.8.5 Call Home notifications	156
3.8.6 GUI preferences menu	158
Chapter 4. Storage pools	159
4.1 Working with internal drives	160
4.1.1 Internal Storage window	160
4.1.2 Actions on internal drives	163
4.2 Working with storage pools	172
4.2.1 Creating storage pools	172
4.2.2 Actions on storage pools	176
4.2.3 Child storage pools	184
4.3 Working with managed disks	188
4.3.1 Assigning managed disks to storage pools	189
4.3.2 RAID configuration	195
4.3.3 Distributed RAID	196
4.3.4 RAID configuration presets	199
4.3.5 Actions on external MDisks	205
4.3.6 More actions on MDisks	213
4.4 Working with external storage controllers	215
Chapter 5. Host configuration	217
5.1 Host attachment overview	218
5.2 Planning for direct-attached hosts	218
5.2.1 FC direct attachment to host systems	219
5.2.2 FC direct attachment between nodes in a Storwize V5000 system	219
5.3 Preparing the host operating system	219
5.3.1 Windows 2008 R2 and 2012 R2: Preparing for FC attachment	219
5.3.2 Windows 2008 R2 and Windows 2012 R2: Preparing for iSCSI attachment	225
5.3.3 Windows 2012 R2: Preparing for SAS attachment	232

5.3.4	VMware ESXi: Preparing for Fibre Channel attachment	233
5.3.5	VMware ESXi: Preparing for iSCSI attachment	236
5.3.6	VMware ESXi: Preparing for SAS attachment	245
5.4	N-Port ID Virtualization support.	246
5.4.1	NPIV prerequisites	248
5.4.2	Enabling NPIV on a new system.	249
5.4.3	Enabling NPIV on an existing system	252
5.5	Creating hosts by using the GUI	255
5.5.1	Creating FC hosts	258
5.5.2	Configuring the IBM Storwize V5000 for FC connectivity	266
5.5.3	Creating iSCSI hosts.	270
5.5.4	Configuring the IBM Storwize V5000 for iSCSI host connectivity	273
5.5.5	Creating SAS hosts.	278
5.6	Host Clusters.	280
5.6.1	Creating a host cluster	282
5.6.2	Adding a member to a host cluster	285
5.6.3	Listing a host cluster member	287
5.6.4	Assigning a volume to a Host Cluster	289
5.6.5	Removing volume mapping from a host cluster	292
5.6.6	Removing a host cluster member	295
5.6.7	Removing a host cluster	298
5.6.8	I/O throttling for hosts and host clusters	300
5.7	Proactive Host Failover	306
Chapter 6. Volume configuration.		309
6.1	Introduction to volumes	310
6.1.1	Image mode volumes	311
6.1.2	Managed mode volumes.	312
6.1.3	Cache mode for volumes	313
6.1.4	Mirrored volumes	314
6.1.5	Thin-provisioned volumes.	317
6.1.6	Compressed volumes	319
6.1.7	Volumes for various topologies.	320
6.2	Create Volumes menu	321
6.3	Creating volumes by using the Volume Creation	325
6.3.1	Creating Basic volumes by using Volume Creation	325
6.3.2	Creating Mirrored volumes by using Volume Creation	327
6.4	Mapping a volume to the host.	329
6.5	Creating Custom volumes.	331
6.5.1	Creating a custom thin-provisioned volume	332
6.5.2	Creating Custom Compressed volumes	334
6.5.3	Custom Mirrored Volumes	336
6.6	HyperSwap and the mkvolume command	338
6.6.1	Volume manipulation commands	340
6.7	Mapping Volumes to Host after volume creation	342
6.7.1	Mapping newly created volumes to the host using the wizard	342
6.8	Migrating a volume to another storage pool	347
6.9	Migrating volumes using the volume copy feature	350
6.10	I/O throttling.	353
6.10.1	Defining throttle on a volume	353
6.10.2	Removing a throttle from a volume	354
Chapter 7. Storage migration.		357

7.1 Storage migration wizard overview	358
7.2 Interoperation and compatibility	358
7.3 Storage migration wizard	359
7.3.1 External virtualization capability	359
7.3.2 Model and adapter card considerations	359
7.3.3 Overview of the storage migration wizard	360
7.3.4 Storage migration wizard tasks	361
Chapter 8. Advanced host and volume administration	383
8.1 Advanced host administration	384
8.1.1 Modifying volume mappings	385
8.1.2 Unmapping volumes from a host	388
8.1.3 Renaming a host	391
8.1.4 Removing a host	393
8.1.5 Host properties	395
8.2 Adding and deleting host ports	400
8.2.1 Adding host port	400
8.2.2 Deleting a host port	404
8.3 Advanced volume administration	407
8.3.1 Advanced volume functions	407
8.3.2 Mapping a volume to a host	409
8.3.3 Unmapping volumes from private hosts	410
8.3.4 Viewing which host is mapped to a volume	410
8.3.5 Renaming a volume	411
8.3.6 Shrinking a volume	412
8.3.7 Expanding a volume	413
8.3.8 Migrating a volume to another storage pool	414
8.3.9 Exporting to an image mode volume	414
8.3.10 Deleting a volume	416
8.3.11 Duplicating a volume	417
8.3.12 Adding a volume copy	418
8.4 Volume properties and volume copy properties	420
8.5 Advanced volume copy functions	422
8.5.1 Volume copy: Make Primary	423
8.5.2 Splitting into a new volume	424
8.5.3 Validate Volume Copies option	426
8.5.4 Delete volume copy option	428
8.5.5 Migrating volumes using the volume copy features	429
8.6 Volumes by storage pool	430
8.7 Volumes by host	432
Chapter 9. Advanced features for storage efficiency	435
9.1 Easy Tier	436
9.1.1 Easy Tier concepts	436
9.1.2 Implementing and tuning Easy Tier	441
9.1.3 Monitoring Easy Tier activity	446
9.2 Thin provisioned volumes	448
9.2.1 Concepts	448
9.2.2 Implementation	449
9.3 Unmap	450
9.3.1 SCSI unmap command	450
9.3.2 Back-end SCSI Unmap	450
9.3.3 Host SCSI Unmap	451

9.3.4 Offload IO throttle	451
9.4 Data Reduction Pools	452
9.4.1 Introduction to DRP	452
9.4.2 Data Reduction Pools benefits	453
9.4.3 Implementing DRP with Compression and Deduplication	454
9.5 Compression with standard pools	460
9.5.1 Real-time Compression concepts	460
9.5.2 Implementing RtC compression	461
9.6 Saving estimation for compression and deduplication	462
9.6.1 Evaluate compression savings by using IBM Comprestimator	462
9.6.2 Evaluating compression and deduplication	463
Chapter 10. Copy Services	465
10.1 IBM FlashCopy	466
10.1.1 Business requirements for FlashCopy	466
10.1.2 Backup improvements with FlashCopy	466
10.1.3 Restore with FlashCopy	467
10.1.4 Moving and migrating data with FlashCopy	467
10.1.5 Application testing with FlashCopy	468
10.1.6 Host and application considerations to ensure FlashCopy integrity	468
10.1.7 FlashCopy attributes	468
10.1.8 Reverse FlashCopy	469
10.1.9 IBM Spectrum Protect Snapshot	470
10.2 FlashCopy functional overview	471
10.3 Implementing FlashCopy	472
10.3.1 FlashCopy mappings	472
10.3.2 Multiple Target FlashCopy	473
10.3.3 Consistency Groups	474
10.3.4 FlashCopy indirection layer	476
10.3.5 Grains and the FlashCopy bitmap	477
10.3.6 Interaction and dependency between multiple target FlashCopy mappings	478
10.3.7 Summary of the FlashCopy indirection layer algorithm	480
10.3.8 Interaction with the cache	480
10.3.9 FlashCopy and image mode volumes	481
10.3.10 FlashCopy mapping events	482
10.3.11 FlashCopy mapping states	485
10.3.12 Thin-provisioned FlashCopy	486
10.3.13 Background copy	487
10.3.14 Serialization of I/O by FlashCopy	488
10.3.15 Event handling	489
10.3.16 Asynchronous notifications	490
10.3.17 Interoperation with Metro Mirror and Global Mirror	490
10.3.18 FlashCopy presets	491
10.4 Managing FlashCopy by using the GUI	493
10.4.1 Creating a FlashCopy mapping	494
10.4.2 Single-click snapshot	502
10.4.3 Single-click clone	503
10.4.4 Single-click backup	504
10.4.5 Creating a FlashCopy Consistency Group	505
10.4.6 Creating FlashCopy mappings in a Consistency Group	506
10.4.7 Showing related volumes	511
10.4.8 Moving a FlashCopy mapping to a Consistency Group	511
10.4.9 Removing a FlashCopy mapping from a Consistency Group	512

10.4.10	Modifying a FlashCopy mapping	513
10.4.11	Renaming FlashCopy mapping	515
10.4.12	Renaming a Consistency Group	515
10.4.13	Deleting FlashCopy mapping	517
10.4.14	Deleting FlashCopy Consistency Group	518
10.4.15	Starting FlashCopy process	519
10.4.16	Stopping FlashCopy process	520
10.5	Volume mirroring and migration options	521
10.6	Native IP replication	522
10.6.1	Native IP replication technology	523
10.6.2	IBM Storwize System Layers	524
10.6.3	IP partnership limitations	526
10.6.4	VLAN support	527
10.6.5	IP partnership and terminology	528
10.6.6	States of IP partnership	529
10.6.7	Remote copy groups	530
10.7	Remote Copy services	531
10.7.1	Multiple IBM Storwize V5000 system mirroring	531
10.7.2	Importance of write ordering	534
10.7.3	Remote copy intercluster communication	536
10.7.4	Metro Mirror overview	537
10.7.5	Synchronous remote copy	538
10.7.6	Metro Mirror features	539
10.7.7	Metro Mirror attributes	539
10.7.8	Practical use of Metro Mirror	540
10.7.9	Global Mirror overview	541
10.7.10	Asynchronous remote copy	541
10.7.11	Global Mirror features	543
10.7.12	Using Change Volumes with Global Mirror	545
10.7.13	Distribution of work among nodes	547
10.7.14	Background copy performance	548
10.7.15	Thin-provisioned background copy	548
10.7.16	Methods of synchronization	548
10.7.17	Practical use of Global Mirror	549
10.7.18	Valid combinations of FlashCopy, Metro Mirror, and Global Mirror	549
10.7.19	Remote Copy configuration limits	550
10.7.20	Remote Copy states and events	551
10.8	Consistency protection for Remote and Global mirror	558
10.9	Remote Copy commands	559
10.9.1	Remote Copy process	559
10.9.2	Listing available system partners	560
10.9.3	Changing the system parameters	561
10.9.4	System partnership	562
10.9.5	Creating a Metro Mirror/Global Mirror consistency group	563
10.9.6	Creating a Metro Mirror/Global Mirror relationship	563
10.9.7	Changing Metro Mirror/Global Mirror relationship	564
10.9.8	Changing Metro Mirror/Global Mirror consistency group	564
10.9.9	Starting Metro Mirror/Global Mirror relationship	564
10.9.10	Stopping Metro Mirror/Global Mirror relationship	565
10.9.11	Starting Metro Mirror/Global Mirror consistency group	565
10.9.12	Stopping Metro Mirror/Global Mirror consistency group	565
10.9.13	Deleting Metro Mirror/Global Mirror relationship	566
10.9.14	Deleting Metro Mirror/Global Mirror consistency group	566

10.9.15	Reversing Metro Mirror/Global Mirror relationship	566
10.9.16	Reversing Metro Mirror/Global Mirror consistency group	566
10.10	Managing Remote Copy using the GUI	567
10.10.1	Creating Fibre Channel partnership	567
10.10.2	Creating stand-alone remote copy relationships	569
10.10.3	Creating a Consistency Group	577
10.10.4	Renaming Consistency Group	578
10.10.5	Renaming remote copy relationship	579
10.10.6	Moving stand-alone remote copy relationship to Consistency Group	580
10.10.7	Removing remote copy relationship from Consistency Group	581
10.10.8	Starting remote copy relationship	582
10.10.9	Starting remote copy Consistency Group	583
10.10.10	Switching copy direction	584
10.10.11	Switching the copy direction for a Consistency Group	585
10.10.12	Stopping a remote copy relationship	586
10.10.13	Stopping Consistency Group	587
10.10.14	Deleting stand-alone remote copy relationships	588
10.10.15	Deleting Consistency Group	590
10.11	Troubleshooting remote copy	590
10.11.1	1920 error	591
10.11.2	1720 error	593
10.12	HyperSwap	593
10.12.1	Introduction to HyperSwap volumes	595
10.12.2	Failure scenarios	601
10.12.3	Current HyperSwap limitations	604
Chapter 11. External storage virtualization	607
11.1	Planning for external storage virtualization	608
11.1.1	License for external storage virtualization	608
11.1.2	Configuration planning for external virtualization	608
11.1.3	External storage configuration planning	610
11.1.4	Guidelines for virtualizing external storage	611
11.2	Working with external storage	611
11.2.1	Adding external FC controllers	612
11.2.2	Adding external iSCSI controllers	613
11.2.3	Working with MDisks	613
11.2.4	Importing image mode volumes	615
11.2.5	Managing external storage controllers	619
11.2.6	Removing external storage	621
Chapter 12. RAS, monitoring, and troubleshooting.	623
12.1	Reliability, availability, and serviceability features	624
12.2	System components	625
12.2.1	Enclosure midplane	625
12.2.2	Node canisters	625
12.2.3	Expansion canisters	634
12.2.4	Disk subsystem	636
12.2.5	Power supply units	640
12.3	Configuration backup	642
12.3.1	Generating a manual configuration backup by using the CLI	643
12.3.2	Downloading a configuration backup by using the GUI	644
12.4	System update	647
12.4.1	Updating node canister software	647

12.4.2	Updating the drive firmware	665
12.5	Monitoring	668
12.5.1	Email notifications and Call Home	669
12.6	Audit log	679
12.7	Event log	681
12.7.1	Managing the event log.	682
12.7.2	Alert handling and recommended actions.	686
12.8	Support assistance	691
12.8.1	Configuring support assistance.	692
12.8.2	Setting up support assistance	692
12.8.3	Disabling support assistance	704
12.9	Collecting support information.	705
12.9.1	Collecting support information by using the GUI.	705
12.9.2	Automatic upload of support packages.	705
12.9.3	Manual upload of Support Packages	712
12.9.4	Collecting support information by using the Service Assistant Tool	717
12.10	Powering off the system and shutting down the infrastructure	719
12.10.1	Powering off	719
12.10.2	Shutting down and starting up the infrastructure.	724
Chapter 13.	Encryption	725
13.1	Planning for encryption	726
13.2	Defining encryption of data-at-rest	726
13.2.1	Encryption methods	727
13.2.2	Encrypted data	727
13.2.3	Encryption keys.	730
13.2.4	Encryption licenses.	731
13.3	Activating encryption	731
13.3.1	Obtaining an encryption license	732
13.3.2	Start activation process during initial system setup	732
13.3.3	Start activation process on a running system	736
13.3.4	Activate the license automatically.	737
13.3.5	Activating the license manually.	740
13.4	Enabling encryption.	742
13.4.1	Starting the Enable Encryption wizard	743
13.4.2	Enabling encryption using USB flash drives	745
13.4.3	Enabling encryption using key servers	750
13.4.4	Enabling encryption by using both providers	765
13.5	Configuring more providers.	770
13.5.1	Adding key servers as a second provider.	771
13.5.2	Adding USB flash drives as a second provider.	774
13.6	Migrating between providers.	776
13.6.1	Migrating from USB flash drive provider to encryption key server	777
13.6.2	Migrating from encryption key server to USB flash drive provider	777
13.6.3	Migrating between different key server types	778
13.7	Recovering from a provider loss	780
13.8	Using encryption	780
13.8.1	Encrypted pools	781
13.8.2	Encrypted child pools	782
13.8.3	Encrypted arrays.	783
13.8.4	Encrypted MDisks	784
13.8.5	Encrypted volumes	787
13.8.6	Restrictions	789

13.9 Rekeying an encryption-enabled system	789
13.9.1 Rekeying using a key server	790
13.9.2 Rekeying using USB flash drives	792
13.10 Disabling encryption	795
Appendix A. CLI setup and SAN Boot.	797
Command-line interface	798
Basic setup	798
SAN Boot	812
Enabling SAN Boot for Windows.	812
Enabling SAN Boot for VMware	813
Windows SAN Boot migration.	813
Appendix B. Terminology.	815
Commonly encountered terms	816
Related publications	837
IBM Redbooks	837
IBM Storwize V5000 publications and support.	837
Help from IBM	837

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](http://www.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.

AIX®	IBM SmartCloud®	PowerHA®
DS8000®	IBM Spectrum™	Real-time Compression™
Easy Tier®	IBM Spectrum Control™	Redbooks®
FlashCopy®	IBM Spectrum Protect™	Redbooks (logo)  ®
Global Technology Services®	IBM Spectrum Scale™	Storwize®
HyperSwap®	IBM Spectrum Virtualize™	System Storage®
IBM®	Informix®	Tivoli®
IBM FlashSystem®	Insight®	XIV®

The following terms are trademarks of other companies:

SoftLayer, are trademarks or registered trademarks of SoftLayer, Inc., an IBM Company.

Celeron, Intel logo, Intel Inside logo, and Intel Centrino logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

ITIL is a Registered Trade Mark of AXELOS Limited.

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.

Java, and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

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

Organizations of all sizes face the challenge of managing massive volumes of increasingly valuable data. But storing this data can be costly, and extracting value from the data is becoming more difficult. IT organizations have limited resources, but must stay responsive to dynamic environments and act quickly to consolidate, simplify, and optimize their IT infrastructures. The IBM® Storwize® V5000 Gen2 system provides a smarter solution that is affordable, easy to use, and self-optimizing, which enables organizations to overcome these storage challenges.

The Storwize V5000 Gen2 delivers efficient, entry-level configurations that are designed to meet the needs of small and midsize businesses. Designed to provide organizations with the ability to consolidate and share data at an affordable price, the Storwize V5000 Gen2 offers advanced software capabilities that are found in more expensive systems.

This IBM Redbooks® publication is intended for pre-sales and post-sales technical support professionals and storage administrators.

It applies to the Storwize V5030, V5020, and V5010, and IBM Spectrum™ Virtualize V8.2.1.

Authors

This book was produced by a team of specialists from around the world working at the International Technical Support Organization, San Jose Center.



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



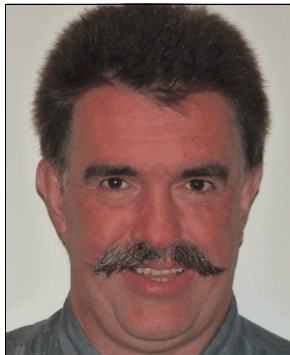
Jack Armstrong is a Storage Support Specialist for IBM Systems Group based in Hursley, UK. He joined IBM as part of the Apprenticeship Scheme in 2012 and has built up 6 years of experience working with Storage, providing support to thousands of customers across Europe and beyond. He also provides value-add work for IBM Enhanced Technical Support Services, helping clients to expand and improve their storage environments.



Tiago Bastos is a SAN and Storage Disk specialist for IBM Brazil. He has over 17 years experience in the IT arena. He is an IBM Certified Master IT Specialist, and certified on the Storwize portfolio. He works on Storage as a Service (SaaS) implementation projects and his areas of expertise include planning, configuring, and troubleshooting IBM DS8000®, Storwize V5000 and V7000, FlashSystem 900, IBM San Volume Controller, and XIV®.



Sergey Kubin is a subject-matter expert (SME) for IBM Storage and SAN support in IBM Russia. He has worked with IBM Technology Support Services for 12 years, providing L1 and L2 support on Spectrum Virtualize, SAN, DS4000/DS5000 and N Series storage for IBM customers in Russia, CEE, and EMEA. He is IBM Certified Specialist for Storwize Family Technical Solutions.



Hartmut Lonzer is an OEM Alliance Manager for IBM Storage. Before this position, he was a Client Technical Specialist for IBM Germany. He works in the IBM Germany headquarters in Ehningen. His main focus is on the IBM SAN Volume Controller, IBM Storwize Family, and IBM VersaStack. His experience with the IBM SAN Volume Controller and Storwize products goes back to the beginning of these products. Hartmut has been with IBM in various technical roles for 40 years.



Danilo Miyasiro is a SAN and Disk Storage Specialist for IBM Global Technology Services® in Brazil. He graduated in Computer Engineering at State University of Campinas, Brazil, and has more than 10 years of experience in IT. As a storage subject matter expert for several international customers, he works on designing, implementing, and supporting storage solutions. He is an IBM Certified Specialist for DS8000 and the Storwize family, and holds certifications from the ITIL Foundation and other storage products.



Rodrigo Suzuki is a SAN Storage specialist at IBM Brazil Global Technology Services in Hortolandia. Currently, Rodrigo is a subject matter expert account focal. He has been working on projects and support for international customers.

He has 24 years of IT Industry experience with the last 9 years in the SAN Storage Disk area. He also has a background in UNIX and IBM Informix® databases. He holds a bachelor's degree in Computer Science from Universidade Paulista in Sao Paulo, Brazil, and is an IBM Certified IT Specialist, NetApp NCDA, IBM Storwize V7000 Technical Solutions V2, and ITIL certified.

Thanks to the authors of the previous edition of this book:

Dharmesh Kamdar
Hartmut Lonzer
Gustavo Tinelli Martins

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

Christopher Bulmer
Debbie Butts
Carlos Fuente
Evelyn Perez
Matt Smith
IBM Hursley, UK

James Whitaker
Imran Imtiaz
Adam Lyon-Jones
IBM Manchester, UK

Jordan Fincher
Karen Brown
Mary Connell
Navin Manohar
Terry Niemeyer
IBM US

Special thanks to the Broadcom Inc. staff in San Jose, California for their support of this residency in terms of equipment and support in many areas:

Sangam Racherla
Brian Steffler
Marcus Thordal
Broadcom Inc.

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 at:

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>

Summary of changes

This section describes the technical changes made in this edition of the book and in previous editions. This edition might also include minor corrections and editorial changes that are not identified.

Summary of Changes
for SG24-8162-04
for Implementing the IBM Storwize V5000 Gen2 with IBM Spectrum Virtualize V8.2.1
as created or updated on April 26, 2019.

April 2019, Fifth Edition


This revision includes the following substantial new and changed information.

New information

- ▶ New GUI
- ▶ Storage migration
- ▶ Advanced features

Changed information

- ▶ Screen captures for new GUI



Overview of the IBM Storwize V5000 Gen2 system

This chapter provides an overview of the IBM Storwize V5000 Gen2 architecture and includes a brief explanation of storage virtualization.

This chapter includes the following topics:

- ▶ 1.1, “IBM Storwize V5000 Gen2 overview” on page 2
- ▶ 1.2, “IBM Storwize V5000 Gen2 terminology” on page 4
- ▶ 1.3, “IBM Storwize V5000 Gen2 models” on page 6
- ▶ 1.4, “IBM Storwize V5000 Gen1 and Gen2 compatibility” on page 11
- ▶ 1.5, “IBM Storwize V5000 Gen2 hardware” on page 12
- ▶ 1.6, “IBM Storwize V5000 Gen2 terms” on page 19
- ▶ 1.7, “IBM Storwize V5000 Gen2 features” on page 28
- ▶ 1.8, “Problem management and support” on page 36
- ▶ 1.9, “More information resources” on page 37

1.1 IBM Storwize V5000 Gen2 overview

The IBM Storwize V5000 Gen2 solution is a modular entry-level and midrange storage solution. The IBM Storwize V5000 Gen2 includes the capability to virtualize its own internal Redundant Array of Independent Disk (RAID) storage and existing external Storage Area Network (SAN)-attached storage (only the Storwize V5030 is capable to virtualize external SAN devices).

The three IBM Storwize V5000 Gen2 models (Storwize V5010, Storwize V5020, and Storwize V5030) offer a range of performance scalability and functional capabilities. Table 1-1 lists a summary of the features of these models.

Table 1-1 IBM Storwize V5000 Gen2 models

	Storwize V5010	Storwize V5020	Storwize V5030
CPU cores	2	2	6
Cache	16 GB	Up to 32 GB	Up to 64 GB
Supported expansion enclosures	10	10	20
External storage virtualization	No	No	Yes
Compression	No	No	Yes
Deduplication	No	No	Yes
Encryption	No	Yes	Yes

For a more detailed comparison, see Table 1-3 on page 6.

IBM Storwize V5000 Gen2 features the following benefits:

- ▶ Enterprise technology is available to entry and midrange storage
- ▶ Expert administrators are not required
- ▶ Easy client setup and service
- ▶ Simple integration into the server environment
- ▶ Ability to grow the system incrementally as storage capacity and performance needs change

The IBM Storwize V5000 Gen2 addresses the block storage requirements of small and midsize organizations. The IBM Storwize V5000 Gen2 consists of one 2U control enclosure and, optionally, up to 10 2U expansion enclosures on the Storwize V5010 and Storwize V5020 systems and up to 20 2U expansion enclosures on the Storwize V5030 systems. The Storwize V5030 systems are connected by serial-attached Small Computer Systems Interface (SCSI) (SAS) cables that make up one system that is called an *I/O group*.

With the Storwize V5030 systems, two I/O groups can be connected to form a cluster, which provides a maximum of two control enclosures and 40 expansion enclosures. With the High Density expansion drawers, you can attach up to 16 expansion enclosures to a cluster.

The control and expansion enclosures are available in the following form factors, and they can be intermixed within an I/O group:

- ▶ 12 x 3.5 inch (8.89-centimeter) drives in a 2U unit
- ▶ 24 x 2.5 inch (6.35-centimeter) drives in a 2U unit
- ▶ 92 x 2.5 inch in carriers or 3.5 inch drives in a 5U unit

Two canisters are in each enclosure. Control enclosures contain two node canisters, and expansion enclosures contain two expansion canisters.

The IBM Storwize V5000 Gen2 supports up to 1,520 x 2.5-inch drives or 3.5-inch drives, or a combination of both drive form factors for the internal storage in a two I/O group Storwize V5030 cluster.

SAS, Nearline (NL)-SAS, and flash drive types are supported.

The IBM Storwize V5000 Gen2 is designed to accommodate the most common storage network technologies to enable easy implementation and management. It can be attached to hosts through a Fibre Channel (FC) SAN fabric, an Internet Small Computer System Interface (iSCSI) infrastructure, or SAS. Hosts can be attached directly or through a network.

Important: For more information about supported environments, configurations, and restrictions, see the [IBM System Storage Interoperation Center \(SSIC\)](#).

For more information, see this [IBM Knowledge Center web page](#).

The IBM Storwize V5000 Gen2 is a virtualized storage solution that groups its internal drives into RAID arrays, which are called *managed disks* (MDisks). MDisks can also be created on the Storwize V5030 systems by importing logical unit numbers (LUNs) from external FC SAN-attached storage. These MDisks are then grouped into *storage pools*. Volumes are created from these storage pools and provisioned out to hosts.

Storage pools are normally created with MDisks of the same drive type and drive capacity. *Volumes* can be moved non-disruptively between storage pools with differing performance characteristics. For example, a volume can be moved between a storage pool that is made up of NL-SAS drives to a storage pool that is made up of SAS drives to improve performance.

The IBM Storwize V5000 Gen2 system also provides several configuration options to simplify the implementation process. It also provides configuration presets and automated wizards that are called *Directed Maintenance Procedures* (DMPs) to help resolve any events that might occur.

Included with an IBM Storwize V5000 Gen2 system is a simple and easy to use graphical user interface (GUI) to enable storage to be deployed quickly and efficiently. The GUI runs on any supported browser. The management GUI contains a series of preestablished configuration options that are called *presets* that use commonly used settings to quickly configure objects on the system. Presets are available for creating volumes and IBM FlashCopy® mappings and for setting up a RAID configuration.

You can also use the command-line interface (CLI) to set up or control the system.

1.2 IBM Storwize V5000 Gen2 terminology

The IBM Storwize V5000 Gen2 system uses terminology that is consistent with the entire IBM Storwize family and the IBM SAN Volume Controller. The terms are defined in Table 1-2. More terms can be found in Appendix B, “Terminology” on page 815.

Table 1-2 IBM Storwize V5000 Gen2 terminology

IBM Storwize V5000 Gen2 term	Definition
Battery	Each control enclosure node canister in an IBM Storwize V5000 Gen2 contains a battery.
Chain	Each control enclosure has one or two chains, which are used to connect expansion enclosures to provide redundant connections to the inside drives.
Clone	A copy of a volume on a server at a particular point. The contents of the copy can be customized and the contents of the original volume are preserved.
Control enclosure	A hardware unit that includes a chassis, node canisters, drives, and power sources.
Data migration	IBM Storwize V5000 Gen2 can migrate data from existing external storage to its internal volumes.
Distributed RAID (DRAID)	No dedicated spare drives are in an array. The spare capacity is distributed across the array, which allows faster rebuild of the failed disk.
Drive	IBM Storwize V5000 Gen2 supports a range of hard disk drives (HDDs) and Flash Drives.
Event	An occurrence that is significant to a task or system. Events can include the completion or failure of an operation, a user action, or the change in the state of a process.
Expansion canister	A hardware unit that includes the SAS interface hardware that enables the control enclosure hardware to use the drives of the expansion enclosure. Each expansion enclosure has two expansion canisters.
Expansion enclosure	A hardware unit that includes expansion canisters, drives, and power supply units.
External storage	MDisks that are SCSI logical units (LUs) that are presented by storage systems that are attached to and managed by the clustered system.
Fibre Channel port	Fibre Channel ports are connections for the hosts to get access to the IBM Storwize V5000 Gen2.
Host mapping	The process of controlling which hosts can access specific volumes within an IBM Storwize V5000 Gen2.
Internal storage	Array MDisks and drives that are held in enclosures that are part of the IBM Storwize V5000 Gen2.
iSCSI (Internet Small Computer System Interface)	Internet Protocol (IP)-based storage networking standard for linking data storage facilities.

IBM Storwize V5000 Gen2 term	Definition
Managed disk (MDisk)	A component of a storage pool that is managed by a clustered system. An MDisk is part of a RAID array of internal storage or a SCSI LU for external storage. An MDisk is not visible to a host system on the SAN.
Node canister	A hardware unit that includes the node hardware, fabric, and service interfaces, SAS expansion ports, and battery. Each control enclosure contains two node canisters.
PHY	A single SAS lane. Four PHYs are in each SAS cable.
Power Supply Unit	Each enclosure has two power supply units (PSU).
Quorum disk	A disk that contains a reserved area that is used exclusively for cluster management. The quorum disk is accessed when it is necessary to determine which half of the cluster continues to read and write data.
Serial-Attached SCSI (SAS) ports	SAS ports are connections for expansion enclosures and direct attachment of hosts to access the IBM Storwize V5000 Gen2.
Snapshot	An image backup type that consists of a point-in-time view of a volume.
Storage pool	An amount of storage capacity that provides the capacity requirements for a volume.
Strand	The SAS connectivity of a set of drives within multiple enclosures. The enclosures can be control enclosures or expansion enclosures.
Thin provisioning or thin provisioned	The ability to define a storage unit (full system, storage pool, or volume) with a logical capacity size that is larger than the physical capacity that is assigned to that storage unit.
Traditional RAID (TRAIID)	Traditional RAID uses the standard RAID levels.
Volume	A discrete unit of storage on disk, tape, or other data recording medium that supports a form of identifier and parameter list, such as a volume label or input/output control.
Worldwide port names	Each Fibre Channel port and SAS port is identified by its physical port number and worldwide port name (WWPN).

1.3 IBM Storwize V5000 Gen2 models

The IBM Storwize V5000 Gen2 platform consists of different models. Each model type supports a different set of features, as listed in Table 1-3.

Table 1-3 IBM Storwize V5000 feature comparison

Feature	V5000 Gen1	V5010	V5020	V5030
Cache	16 GB	16 GB	16 GB or 32 GB	32 GB or 64 GB
CPU	4-core Ivy Bridge Xeon CPU 2 GHz	2-core Broadwell-DE Celeron CPU 1.2 GHz	2-core Broadwell-DE Xeon CPU 2.2 GHz Hyper-threading	6-core Broadwell-DE Xeon CPU 1.9 GHz Hyper-threading
Compression	None	None	None	Licensed (with 64 GB cache only)
Deduplication	None	None	None	Yes
DRAID	Yes	Yes	Yes	Yes
SAS HW Encryption	None	None	Licensed	Licensed
External Virtualization	Licensed	Data Migration Only	Data Migration Only	Licensed
IBM Easy Tier®	Licensed	Licensed	Licensed	Licensed
FlashCopy	Licensed	Licensed	Licensed	Licensed
Hyperswap	Yes	No	No	Yes
Remote Copy	Licensed	Licensed	Licensed	Licensed
Thin Provisioning	Yes	Yes	Yes	Yes
Traditional RAID	Yes	Yes	Yes	Yes
Volume Mirroring	Yes	Yes	Yes	Yes
VMware Virtual Volumes (VVols)	Yes	Yes	Yes	Yes

More information: For more information about the features, benefits, and specifications of IBM Storwize V5000 Gen2 models, see [this website](#).

The information in this book is accurate at the time of this writing. However, as the IBM Storwize V5000 Gen2 matures, expect to see new features and enhanced specifications.

The IBM Storwize V5000 Gen2 models are listed in Table 1-4. All control enclosures have two node canisters. XXF models are expansion enclosures.

Table 1-4 IBM Storwize V5000 Gen2 models

Model	Description	Cache	Drive Slots
One-year warranty			
2077-112	IBM Storwize V5010 large form factor (LFF) Control Enclosure	16 GB	12 x 3.5 inch
2077-124	IBM Storwize V5010 small form factor (SFF) Control Enclosure	16 GB	24 x 2.5 inch
2077-212	IBM Storwize V5020 LFF Control Enclosure	16 GB or 32 GB	12 x 3.5 inch
2077-224	IBM Storwize V5020 SFF Control Enclosure	16 GB or 32 GB	24 x 2.5 inch
2077-312	IBM Storwize V5030 LFF Control Enclosure	32 GB or 64 GB	12 x 3.5 inch
2077-324	IBM Storwize V5030 SFF Control Enclosure	32 GB or 64 GB	24 x 2.5 inch
2077-AF3	IBM Storwize V5030F All-Flash Array Control Enclosure	64 GB	24 x 2.5 inch
2077-12F	IBM Storwize V5000 LFF Expansion Enclosure	N/A	12 x 3.5 inch
2077-24F	IBM Storwize V5000 SFF Expansion Enclosure	N/A	24 x 2.5 inch
2077-AFF	IBM Storwize V5030F SFF Expansion Enclosure	N/A	24 x 2.5 inch
2077-A9F	IBM Storwize V5030F High Density LFF Expansion Enclosure	N/A	92 x 3.5 inch
Three-year warranty			
2078-112	IBM Storwize V5010 LFF Control Enclosure	16 GB	12 x 3.5 inch
2078-124	IBM Storwize V5010 SFF Control Enclosure	16 GB	24 x 2.5 inch
2078-212	IBM Storwize V5020 LFF Control Enclosure	16 GB or 32 GB	12 x 3.5 inch
2078-224	IBM Storwize V5020 SFF Control Enclosure	16 GB or 32 GB	24 x 2.5 inch

Model	Description	Cache	Drive Slots
2078-312	IBM Storwize V5030 LFF Control Enclosure	32 GB or 64 GB	12 x 3.5 inch
2078-324	IBM Storwize V5030 SFF Control Enclosure	32 GB or 64 GB	24 x 2.5 inch
2078-AF3	IBM Storwize V5030F All-Flash Array Control Enclosure	64 GB	24 x 2.5 inch
2078-12F	IBM Storwize V5000 LFF Expansion Enclosure	N/A	12 x 3.5 inch
2078-24F	IBM Storwize V5000 SFF Expansion Enclosure	N/A	24 x 2.5 inch
2078-AFF	IBM Storwize V5030F SFF Expansion Enclosure	N/A	24 x 2.5 inch
2078-A9F	IBM Storwize V5030F High Density LFF Expansion Enclosure	N/A	92 x 3.5 inch

Storwize V5030F control enclosures support only the attachment of Storwize V5030F expansion enclosures (Models AFF and A9F). Storwize V5000 expansion enclosures (Models 12E, 24E, 12F, 24F, and 92F) are not supported with Storwize V5030F control enclosures.

Storwize V5030F expansion enclosures are supported for attachment to Storwize V5030F control enclosures only. Storwize V5000 control enclosures (Models 12C, 24C, 112, 124, 212, 224, 312, and 324) do not support the attachment of Storwize V5030F expansion enclosures.

Table 1-5 shows the 2U expansion enclosures and 5U expansion enclosure mix rules. Shown are the maximum numbers of Drive Slots per SAS expansion string without disks in the controller.

Table 1-5 2U expansion enclosures and 5U expansion enclosure mix rules

	5U											
2U		0	1	2	3	4	5	6	7	8	9	10
0	0	0	24	48	72	96	120	144	168	192	216	240
1	92	116	140	164	188	212	236	260	-	-	-	-
2	184	208	232	256	280	304	-	-	-	-	-	-
3	276	300	324	-	-	-	-	-	-	-	-	-
4	368	-	-	-	-	-	-	-	-	-	-	-

The Storwize V5030 systems can be added to an existing IBM Storwize V5000 Gen1 cluster to form a two-I/O group configuration. This configuration can be used as a migration mechanism to upgrade from the IBM Storwize V5000 Gen1 to the IBM Storwize V5000 Gen2.

The IBM Storwize V5000 Gen1 models are listed in Table 1-6 for completeness.

Table 1-6 IBM Storwize V5000 Gen1 models

Model	Cache	Drive slots
One-year warranty		
2077-12C	16 GB	12 x 3.5 inch
2077-24C	16 GB	24 x 2.5 inch
2077-12E	N/A	12 x 3.5 inch
2077-24E	N/A	24 x 2.5 inch
Three-year warranty		
2078-12C	16 GB	12 x 3.5 inch
2078-24C	16 GB	24 x 2.5 inch
2078-12E	N/A	12 x 3.5 inch
2078-24E	N/A	24 x 2.5 inch

Figure 1-1 shows the front view of the 2077/2078-12 and 12F enclosures.



Figure 1-1 IBM Storwize V5000 Gen2 front view for 2077/2078-12 and 12F enclosures

The drives are positioned in four columns of three horizontally mounted drive assemblies. The drive slots are numbered 1 - 12, starting at the upper left and moving left to right, top to bottom.

Figure 1-2 shows the front view of the 2077/2078-24 and 24F enclosures.

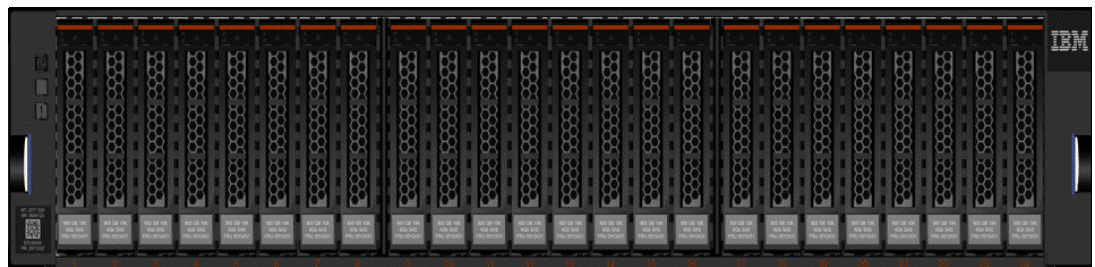


Figure 1-2 IBM Storwize V5000 Gen2 front view for 2077/2078-24 and 24F enclosure

The drives are positioned in one row of 24 vertically mounted drive assemblies. The drive slots are numbered 1 - 24, starting from the left. A vertical center drive bay molding is between slots 8 and 9 and another between slots 16 and 17.

1.3.1 IBM Storage Utility Offerings

The IBM 2078 Model U5A is the IBM Storwize V5030 with a three-year warranty, to be used in the Storage Utility Offering space. These models are physically and functionally identical to the Storwize V5030 model 324 with the exception of target configurations and variable capacity billing. The variable capacity billing uses IBM Spectrum Control™ Storage Insights to monitor the system usage, which allows allocated storage usage above a base subscription rate to be billed per TB, per month.

Allocated storage is identified as storage that is allocated to a specific host (and unusable to other hosts), whether data is written or not. For thin-provisioning, the data that is written is considered *used*. For thick provisioning, total allocated volume space is considered used.

IBM Storage Utility Offerings include several IBM storage products. For more information these products, see [this website](#).

IBM Storage Utility Offering models enable a variable capacity usage and billing. These models provide a fixed total capacity, with a base and variable usage subscription of that total capacity. IBM Spectrum Control Storage Insights is used to monitor the system capacity usage. It is used to report on capacity used beyond the base subscription capacity, which is referred to as *variable usage*. The variable capacity usage is billed on a quarterly basis. This enables customers to grow or shrink their usage, and pay for configured capacity only.

IBM Storage Utility Offering models are provided for customers who can benefit from a variable capacity system, where billing is based on actual provisioned space above the base. The base subscription is covered by a three-year lease that entitles the customer to use the base capacity at no additional cost. If storage needs increase beyond the base capacity, usage is billed based on the average daily provisioned capacity per TB, per month, on a quarterly basis.

Example

A customer has a Storwize V5030 utility model with 2 TB nearline disks, for a total system capacity of 48 TB. The base subscription for such a system is 16.8 TB. During the months where the average daily usage is below 16.8 TB, there is no additional billing.

The system monitors daily provisioned capacity and averages those daily usage rates over the month term. The result is the average daily usage for the month.

If a customer uses 25 TB, 42.6 TB, and 22.2 TB in three consecutive months, Storage Insights calculates the overage as follows (rounding to the nearest terabyte) as listed in Table 1-7.

Table 1-7 Overage calculation

Average daily	Base	Overage	To be billed
25.0	16.8	8.2	8
42.6	16.8	25.8	26
22.2	16.8	5.4	5

The capacity that is billed at the end of the quarter is a total of 39 TB-months in this example.

Disk expansions (2076-24F for the Storwize V7000 and 2078-24F for the Storwize V5030) can be ordered with the system at the initial purchase, but cannot be added through MES. The expansions must have like-type and capacity drives, and must be fully populated.

For example, on a Storwize V7000 utility model with 24 7.68 TB flash drives in the controller, a 2076-24F with 24 7.68 TB drives can be configured with the initial system. Expansion drawers do not apply to FlashSystem 900 (9843-UF3). Storwize V5030 and Storwize V7000 utility model systems support up to 760 drives in the system.

The usage data collected by Storage Insights is used by IBM to determine the actual physical data provisioned in the system. This data is compared to the base system capacity subscription, and any provisioned capacity beyond that base subscription is billed per TB, per month, on a quarterly basis. The calculated usage is based on the average use over a specific month.

In a highly variable environment, such as managed or cloud service providers, this enables the system to be used only as much as is necessary during any month. Usage can increase or decrease, and is billed accordingly. Provisioned capacity is considered capacity that is reserved by the system. In thick-provisioned environments (available on FlashSystem 900 and Storwize), this is the capacity that is allocated to a host whether it has data written or not.

For thin-provisioned environments (available on the Storwize system), this is the data that is actually written and used. This is because of the different ways in which thick and thin provisioning use disk space.

These systems are available worldwide, but there are specific client and program differences by location. Consult your IBM Business Partner or IBM sales person for specifics.

1.4 IBM Storwize V5000 Gen1 and Gen2 compatibility

The Storwize V5030 systems can be added into existing Storwize V5000 Gen1 clustered systems. All systems within a cluster must use the same version of Storwize V5000 software, which is version 7.6.1 or later.

Restriction: The Storwize V5010 and Storwize V5020 are not compatible with V5000 Gen1 systems because they cannot join an existing I/O group.

A single Storwize V5030 control enclosure can be added to a single Storwize V5000 cluster to bring the total number of I/O groups to two. They can be clustered by using Fibre Channel (FC) or Fibre Channel over Ethernet (FCoE). The possible I/O group configuration options for all Storwize V5000 models are listed in Table 1-8.

Table 1-8 IBM Storwize V5000 I/O group configurations

I/O group 0	I/O group 1
V5010	N/A
V5020	N/A
V5030	N/A
V5030	V5030
V5030	V5000 Gen1
V5000 Gen 1	V5030
V5000 Gen1	N/A
V5000 Gen1	V5000 Gen1

1.5 IBM Storwize V5000 Gen2 hardware

The IBM Storwize V5000 Gen2 solution is a modular storage system that is built on a common enclosure platform that is shared by the control enclosures and expansion enclosures.

Figure 1-3 shows an overview of hardware components of the IBM Storwize V5000 Gen2 solution.

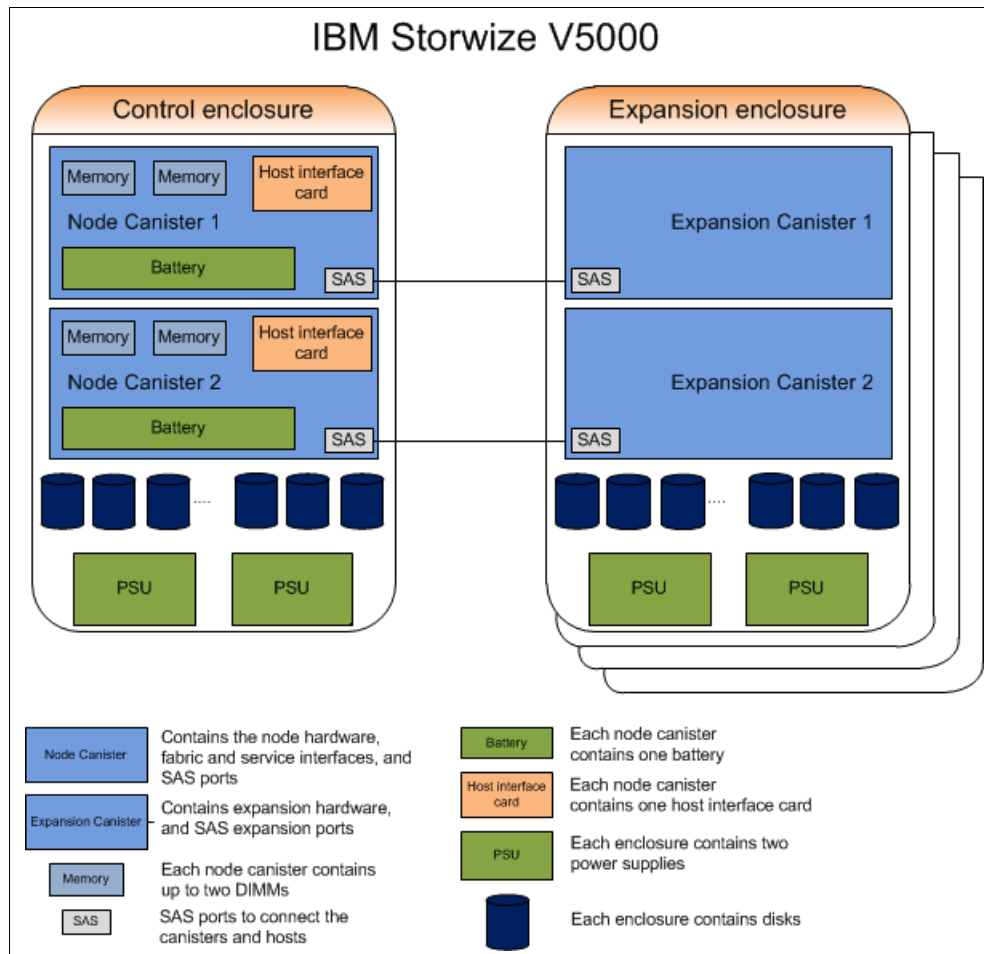


Figure 1-3 IBM Storwize V5000 Gen2 hardware components

Figure 1-4 shows the control enclosure rear view of an IBM Storwize V5000 Gen2 enclosure (the Storwize V5020).



Figure 1-4 Storwize V5020 control enclosure rear view

In Figure 1-4, you can see two power supply slots at the bottom of the enclosure. The power supplies are identical and exchangeable. The two canister slots are at the top of the chassis.

In Figure 1-5, you can see the rear view of an IBM Storwize V5000 Gen2 expansion enclosure.

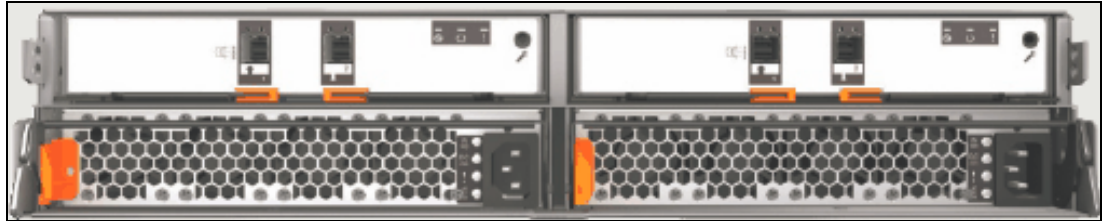


Figure 1-5 IBM Storwize V5000 Gen2 expansion enclosure rear view

You can see that the only difference between the control enclosure and the expansion enclosure is the canister. The canisters of the expansion enclosure have only two SAS ports.

For more information about the expansion enclosure, see 1.5.5, “Expansion enclosure” on page 17.

1.5.1 Control enclosure

Each IBM Storwize V5000 Gen2 system has one control enclosure that contains two node canisters (nodes), disk drives, and two power supplies.

The two node canisters act as a single processing unit and form an I/O group that is attached to the SAN fabric, an iSCSI infrastructure, or that is directly attached to hosts through FC or SAS. The pair of nodes is responsible for serving I/O to a volume. The two nodes provide a highly available fault-tolerant controller so that if one node fails, the surviving node automatically takes over. Nodes are deployed in pairs that are called *I/O groups*.

One node is designated as the configuration node, but each node in the control enclosure holds a copy of the control enclosure state information.

The Storwize V5010 and Storwize V5020 support a single I/O group. The Storwize V5030 supports two I/O groups in a clustered system.

The terms *node canister* and *node* are used interchangeably throughout this book.

The battery is used if power is lost. The IBM Storwize V5000 Gen2 system uses this battery to power the canister while the cache data is written to the internal system flash. This memory dump is called a *fire hose memory dump*.

Note: The batteries of the IBM Storwize V5000 Gen2 can process two fire hose memory dumps in a row. After this, you cannot power up the system immediately. There is a need to wait until the batteries are charged over a level that allows them to run the next fire hose memory dump.

After the system is up again, the data from the internal system flash is loaded back to the cache for destaging to the disks.

1.5.2 Storwize V5010

Figure 1-6 shows a single Storwize V5010 node canister.



Figure 1-6 Storwize V5010 node canister

Each Storwize V5010 node canister contains the following hardware:

- ▶ Battery
- ▶ Memory: 8 GB
- ▶ One 12 Gbps SAS port for expansions
- ▶ Two 10/100/1000 Mbps Ethernet ports
- ▶ One USB 2.0 port that is used to gather system information
- ▶ System flash
- ▶ Host interface card (HIC) slot (different options are possible)

Figure 1-6 shows the following features that are provided by the Storwize V5010 node canister:

- ▶ Two 10/100/1000 Mbps Ethernet ports are available for I/O. Additional Port 1 is used for management, and port 2 can optionally be used for management. Port 2 serves as a technician port (as denoted by the white box with “T” in it) for system initialization and service.

Note: All three models use a technician port to perform initial setup. The implementation of the technician port varies between models. On Storwize V5010/20 the second 1 GbE port (labelled T) is initially enabled as a technician port. After cluster creation, this port is disabled and can then be used for I/O or management.

On Storwize V5030, the onboard 1 GbE port (labelled T) is permanently enabled as a technician port. Connecting the technician port to the LAN disables the port. The Storwize V5010/20 technician port can be re-enabled after initial setup.

The following commands are used to enable or disable the technical port:

```
satask chserviceip -techport enable -force
satask chserviceip -techport disable
```

- ▶ Both ports can be used for iSCSI traffic and IP replication. For more information, see Chapter 5, “Host configuration” on page 217, and Chapter 10, “Copy Services” on page 465.
- ▶ One USB port for gathering system information.

System initialization: Unlike the Storwize V5000 Gen1, you must perform the system initialization of the Storwize V5010 by using the technician port instead of the USB port.

- ▶ One 12 Gbps serial-attached SCSI (SAS 3.0) port to connect to the optional expansion enclosures. The Storwize V5010 supports up to 10 expansion enclosures.

Important: The canister SAS port on the Storwize V5010 does not support SAS host attachment. The Storwize V5010 only supports SAS hosts by using an optional host interface card. For more information, see 1.5.6, “Host interface cards” on page 18.

Do not use the port that is marked with a wrench. This port is a service port only.

1.5.3 Storwize V5020

Figure 1-7 shows a single Storwize V5020 node canister.



Figure 1-7 Storwize V5020 node canister

Each node canister contains the following hardware:

- ▶ Battery
- ▶ Memory: 8 GB, upgradable to 16 GB
- ▶ Three 12 Gbps SAS ports (two for Host attachment, one for expansions)
- ▶ Two 10/100/1000 Mbps Ethernet ports
- ▶ One USB 2.0 port that is used to gather system information
- ▶ System flash
- ▶ HIC slot (different options are possible)

Figure 1-7 on page 15 shows the following features that are provided by the Storwize V5020 node canister:

- ▶ Two 10/100/1000 Mbps Ethernet ports are available for I/O. Additional Port 1 is used for management, and port 2 can optionally be used for management. Port 2 serves as a technician port (as denoted by the white box with “T” in it) for system initialization and service.

Note: All three models use a technician port to perform initial setup. The implementation of the technician port varies between models: On Storwize V5010/20 the second 1 GbE port (labelled T) is initially enabled as a technician port. After cluster creation, this port is disabled and can then be used for I/O or management.

On Storwize V5030 the onboard 1 GbE port (labelled T) is permanently enabled as a technician port. Connecting the technician port to the LAN will disable the port. The Storwize V5010/20 technician port can be re-enabled after initial setup.

The following commands are used to enable or disable the techport:

```
satask chserviceip -techport enable -force
satask chserviceip -techport disable
```

- ▶ Both ports can be used for iSCSI traffic and IP replication. For more information, see Chapter 5, “Host configuration” on page 217, and Chapter 10, “Copy Services” on page 465.
- ▶ One USB port for gathering system information.

System initialization: Unlike the Storwize V5000 Gen1, you must perform the system initialization of the Storwize V5020 by using the technician port instead of the USB port.

- ▶ Three 12 Gbps serial-attached SCSI (SAS 3.0) ports. The ports are numbered 1 - 3 from left to right. Port 1 is used to connect to the optional expansion enclosures. Ports 2 and 3 can be used to connect directly to SAS hosts. (Both 6 Gb and 12 Gb hosts are supported.) The Storwize V5020 supports up to 10 expansion enclosures.

Service port: Do not use the port that is marked with a wrench. This port is a service port only.

1.5.4 Storwize V5030

Figure 1-8 shows a single Storwize V5030 node canister.



Figure 1-8 Storwize V5030 node canister

Each node canister contains the following hardware:

- ▶ Battery
- ▶ Memory: 16 GB, upgradable to 32 GB
- ▶ Two 12 Gbps SAS ports for Expansions
- ▶ One 10/100/1000 Mbps Ethernet technician port
- ▶ Two 1/10 Gbps Ethernet ports
- ▶ One USB 2.0 port that is used to gather system information
- ▶ System flash
- ▶ HIC slot (different options are possible)

Figure 1-8 shows the following features that are provided by the Storwize V5030 node canister:

- ▶ One Ethernet technician port (as denoted by the white box with “T” in it). This port can be used for system initialization and service only. For more information, see Chapter 1, “Overview of the IBM Storwize V5000 Gen2 system” on page 1. It cannot be used for anything else.
- ▶ Two 1/10 Gbps Ethernet ports are available for I/O. These ports are Copper 10GBASE-T with RJ45 connectors. Additional Port 1 must be used for management. Port 2 can optionally be used for management. Both ports can be used for iSCSI traffic and IP replication. For more information, see Chapter 5, “Host configuration” on page 217, and Chapter 10, “Copy Services” on page 465.

Important: The 1/10 Gbps Ethernet ports do not support speeds less than 1 Gbps (100 Mbps is not supported).

Ensure that you use the correct port connectors. The Storwize V5030 canister 10 Gbps connectors appear the same as the 1 Gbps connectors on the other Storwize V5000 models. These RJ45 connectors differ from the optical small form-factor pluggable (SFP+) connectors on the optional 10 Gbps HIC. When you plan to implement the Storwize V5030, ensure that any network switches provide the correct connector type and speed.

- ▶ One USB port to gather system information.

System initialization: Unlike the Storwize V5000 Gen1, you must perform the system initialization of the Storwize V5030 by using the technician port instead of the USB port.

- ▶ Two 12 Gbps serial-attached SCSI (SAS 3.0) ports. The ports are numbered 1 and 2 from left to right to connect to the optional expansion enclosures. The Storwize V5030 supports up to 20 expansion enclosures. A total of 10 expansion enclosures can be connected to each port.

Important: The canister SAS ports on the Storwize V5030 do not support SAS host attachment. The Storwize V5030 supports SAS hosts by using a HIC. For more information, see 1.5.6, “Host interface cards” on page 18.

Do not use the port that is marked with a wrench. This port is a service port only.

1.5.5 Expansion enclosure

The optional IBM Storwize V5000 Gen2 expansion enclosure contains two expansion canisters, disk drives, and two power supplies. Four types of expansion enclosures are available: large form factor 2u (LFF) Expansion Enclosure Model 12F, a small form factor 2U (SFF) Expansion Enclosure Model 24F, a small form factor 2U (SFF) Expansion Enclosure for flash drives Model AFF, and the high density drawers 5U LFF Model 92F or the flash version A9F. They are available with one or three-year warranty.

Figure 1-9 shows the rear of the 2U expansion enclosure.

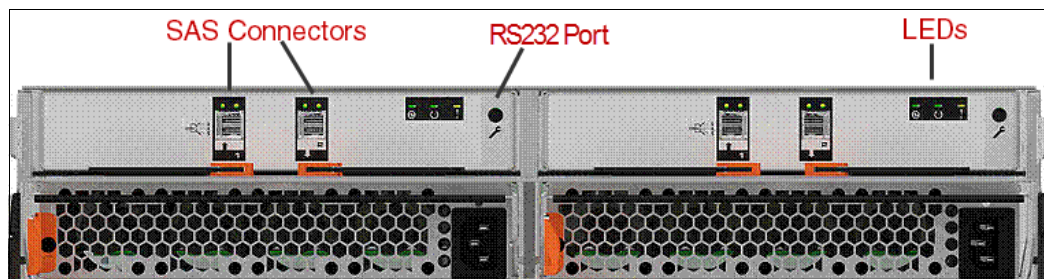


Figure 1-9 2u expansion enclosure of the IBM Storwize V5000 Gen2

The expansion enclosure power supplies are the same as the control enclosure power supplies. A single power lead connector is on each power supply unit.

Each expansion canister provides two SAS interfaces that are used to connect to the control enclosure and any further optional expansion enclosures. The ports are numbered 1 on the left and 2 on the right. SAS port 1 is the IN port, and SAS port 2 is the OUT port.

The use of SAS connector 1 is mandatory because the expansion enclosure must be attached to a control enclosure or another expansion enclosure further up in the chain. SAS connector 2 is optional because it is used to attach to further expansion enclosures down the chain.

The Storwize V5010 and Storwize V5020 support a single chain of up to 10 expansion enclosures that attach to the control enclosure. The Storwize V5030 supports up to 40 expansion enclosures in a configuration that consists of two control enclosures, which are each attached to 20 expansion enclosures in two separate chains.

Table 1-9 lists the maximum number of supported expansion enclosures and the drive limits for each model.

Table 1-9 Expansion enclosure and drive limits

	V5010	V5020	V5030
Maximum number of supported expansion enclosures	10	10	40
Maximum number of supported drives	392	392	1520

Each port includes two LEDs to show the status. The first LED indicates the link status and the second LED indicates the fault status.

For more information about LED and ports, see [this website](#).

Restriction: The IBM Storwize V5000 Gen2 expansion enclosures can be used with an IBM Storwize V5000 Gen2 control enclosure only. The IBM Storwize V5000 Gen1 expansion enclosures cannot be used with an IBM Storwize V5000 Gen2 control enclosure.

1.5.6 Host interface cards

All IBM Storwize V5000 Gen2 support Ethernet ports as standard for iSCSI connectivity. For the Storwize V5010 and Storwize V5020, these Ethernet ports are 1 GbE ports. For the Storwize V5030, these Ethernet ports are 10 GbE ports. The Storwize V5020 also includes 12 Gb SAS ports for host connectivity as a standard.

Additional host connectivity options are available through an optional adapter card. Table 1-10 lists the available configurations for a single control enclosure.

Table 1-10 IBM Storwize V5000 Gen2 configurations available

	1 Gb Ethernet (iSCSI)	10 Gb Ethernet Copper 10GBASE-T (iSCSI)	12 Gb SAS	16 Gb FC	10 Gb Ethernet Optical SFP+ iSCSI/FCoE
V5030	8 ports (with optional adapter card).	4 ports (standard).	8 ports (with optional adapter card).	8 ports (with optional adapter card).	8 ports (with optional adapter card).

V5020	4 ports (standard). Additional 8 ports (with optional adapter card)	N/A	4 ports (standard). Additional 8 ports (with optional adapter card).	8 ports (with optional adapter card).	8 ports (with optional adapter card).
V5010	4 ports (standard). Additional 8 ports (with optional adapter card).	N/A	8 ports (with optional adapter card).	8 ports (with optional adapter card).	8 ports (with optional adapter card).

Optional adapter cards: One pair of identical adapter cards is allowed for each control enclosure.

1.5.7 Disk drive types

IBM Storwize V5000 Gen2 enclosures support Flash Drives, SAS, and Nearline SAS drive types. Each drive has two ports (two PHYs) to provide fully redundant access from each node canister. I/O can be issued down both paths simultaneously.

Table 1-11 lists the IBM Storwize V5000 Gen2 disk drive types, disk revolutions per minute (RPMs), and sizes that are available at the time of writing.

Table 1-11 IBM Storwize V5000 Gen2 disk drive types

Drive type		RPM	Size
2.5 inch form factor	Flash Drive	N/A	400 GB, 800 GB, 1.6 TB, and 3.2 TB
2.5 inch form factor	Read Intensive (RI) Flash Drive	N/A	1.92 TB, 3.84 TB, 7.68 TB, and 15.36 TB
2.5 inch form factor	SAS	10,000	900 GB, 1.2 TB, 1.8 TB, and 2.4 TB
2.5 inch form factor	SAS	15,000	300 GB, 600 GB, and 900 GB
2.5 inch form factor	Nearline SAS	7,200	2 TB
3.5 inch form factor	SAS	10,000	900 GB, 1.2 TB, and 1.8 TB ^a
3.5 inch form factor	SAS	15,000	300 GB, 600 GB, and 900 GB ^a
3.5 inch form factor	Nearline SAS	7,200	4 TB, 6 TB, 8 TB, and 10 TB

a. 2.5 inch drive in a 3.5 inch drive carrier

1.6 IBM Storwize V5000 Gen2 terms

In this section, we introduce the terms that are used for the IBM Storwize V5000 Gen2 throughout this book.

1.6.1 Hosts

A *host* system is a server that is connected to IBM Storwize V5000 Gen2 through a Fibre Channel connection, an iSCSI connection, or an SAS connection.

Hosts are defined on IBM Storwize V5000 Gen2 by identifying their WWPNs for Fibre Channel and SAS hosts. The iSCSI hosts are identified by using their iSCSI names. The iSCSI names can be iSCSI qualified names (IQNs) or extended unique identifiers (EUIs). For more information, see Chapter 5, “Host configuration” on page 217.

Hosts can be Fibre Channel-attached through an existing Fibre Channel network infrastructure or direct-attached, iSCSI-attached through an existing IP network, or directly attached through SAS.

1.6.2 Node canister

A *node canister* provides host interfaces, management interfaces, and SAS interfaces to the control enclosure. A node canister has the cache memory, the internal storage to store software and logs, and the processing power to run the IBM Storwize V5000 Gen2 virtualization and management software. A clustered system consists of one or two node pairs. Each node pair forms one I/O group. For more information about I/O groups, see 1.6.3, “I/O groups” on page 20.

One of the nodes within the system, which is known as the *configuration node*, manages configuration activity for the clustered system. If this node fails, the system nominates the other node to become the configuration node.

1.6.3 I/O groups

Within IBM Storwize V5000 Gen2, one or two pairs of node canisters are known as *I/O groups*. The IBM Storwize V5000 Gen2 supports two-node or four-node canisters in a clustered system, which provides one or two I/O groups, depending on the model. For more information, see Table 1-8 on page 11.

When a host server performs I/O to one of its volumes, all of the I/Os for a specific volume are directed to the I/O group. Also, under normal conditions, the I/Os for that specific volume are always processed by the same node within the I/O group.

When a host server performs I/O to one of its volumes, all of the I/O for that volume is directed to the I/O group where the volume was defined. Under normal conditions, these I/Os are also always processed by the same node within that I/O group.

Both nodes of the I/O group act as preferred nodes for their own specific subset of the total number of volumes that the I/O group presents to the host servers (a maximum of 2,048 volumes for each host). However, both nodes also act as a failover node for the partner node within the I/O group. Therefore, a node takes over the I/O workload from its partner node (if required) without affecting the server’s application.

In an IBM Storwize V5000 Gen2 environment (which uses active-active architecture), the I/O handling for a volume can be managed by both nodes of the I/O group. The I/O groups must be connected to the SAN so that all hosts can access all nodes. The hosts must use multipath device drivers to handle this capability.

Up to 256 host server objects can be defined to one-I/O group or 512 host server objects can be defined in a two-I/O group system. For more information about I/O groups, see Chapter 6, “Volume configuration” on page 309.

Important: The active/active architecture provides the availability to process I/Os for both controller nodes and allows the application to continue to run smoothly, even if the server has only one access route or path to the storage controller. This type of architecture eliminates the path/LUN thrashing that is typical of an active/passive architecture.

1.6.4 Clustered system

A *clustered system* consists of one or two pairs of node canisters. Each pair forms an I/O group. All configuration, monitoring, and service tasks are performed at the system level. The configuration settings are replicated across all node canisters in the clustered system. To facilitate these tasks, one or two management IP addresses are set for the clustered system. By using this configuration, you can manage the clustered system as a single entity.

A process exists to back up the system configuration data on to disk so that the clustered system can be restored in a disaster. This method does not back up application data. Only IBM Storwize V5000 Gen2 system configuration information is backed up.

System configuration backup: After the system configuration is backed up, save the backup data on your local hard disk (or at the least outside of the SAN). If you cannot access the IBM Storwize V5000 Gen2, you do not have access to the backup data if it is on the SAN. Perform this configuration backup after each configuration change as a precaution.

The system can be configured by using the IBM Storwize V5000 Gen2 management software (GUI), CLI, or USB key.

1.6.5 RAID

The IBM Storwize V5000 Gen2 contains several internal drive objects, but these drives cannot be directly added to the storage pools. Drives must be included in a Redundant Array of Independent Disks (*RAID*) to provide protection against the failure of individual drives.

These drives are referred to as *members* of the array. Each array has a RAID level. RAID levels provide various degrees of redundancy and performance. The maximum number of members in the array varies based on the RAID level.

Traditional RAID (TRAIID) has the concept of hot spare drives. When an array member drive fails, the system automatically replaces the failed member with a hot spare drive and rebuilds the array to restore its redundancy. Candidate and spare drives can be manually exchanged with array members.

Apart from traditional disk arrays, IBM Spectrum Virtualize™ V7.6 introduced Distributed RAIDs. Distributed RAID improves recovery time of failed disk drives in an array by the distribution of spare capacity between primary disks, rather than dedicating a whole spare drive for replacement.

For more information about traditional and distributed RAID arrays, see Chapter 4, “Storage pools” on page 159.

1.6.6 Managed disks

A *managed disk* (MDisk) refers to the unit of storage that IBM Storwize V5000 Gen2 virtualizes. This unit can be a logical volume on an external storage array that is presented to the IBM Storwize V5000 Gen2 or a (traditional or distributed) RAID array that consists of internal drives. The IBM Storwize V5000 Gen2 can then allocate these MDisks into storage pools.

An MDisk is invisible to a host system on the storage area network because it is internal to the IBM Storwize V5000 Gen2 system. An MDisk features the following modes:

- ▶ Array

Array mode MDisks are constructed from internal drives by using the RAID functionality. Array MDisks are always associated with storage pools.

- ▶ Unmanaged

LUNs that are presented by external storage systems to IBM Storwize V5000 Gen2 are discovered as unmanaged MDisks. The MDisk is not a member of any storage pools, which means that it is not used by the IBM Storwize V5000 Gen2 storage system.

- ▶ Managed

Managed Disks are LUNs, which are presented by external storage systems to an IBM Storwize V5000 Gen2, that are assigned to a storage pool and provide extents so that volumes can use them. Any data that might be on these LUNs when they are added is lost.

- ▶ Image

Image MDisks are LUNs that are presented by external storage systems to an IBM Storwize V5000 Gen2 and assigned directly to a volume with a one-to-one mapping of extents between the MDisk and the volume. For more information, see Chapter 6, “Volume configuration” on page 309.

1.6.7 Quorum disks

A *quorum disk* is an MDisk that contains a reserved area for use exclusively by the system. In the IBM Storwize V5000 Gen2, internal drives can be considered as quorum candidates. The clustered system uses quorum disks to break a tie when exactly half the nodes in the system remain after a SAN failure.

The clustered system automatically forms the quorum disk by taking a small amount of space from an MDisk. It allocates space from up to three different MDisks for redundancy, although only one quorum disk is active.

To avoid the possibility of losing all of the quorum disks because of a failure of a single storage system if the environment has multiple storage systems, you must allocate the quorum disk on different storage systems. You can manage the quorum disks by using the CLI.

IP quorum base support provides an alternative for Storwize V5000 IBM HyperSwap® implementations. Instead of Fibre Channel storage on a third site, the IP network is used for communication between the IP quorum application and node canisters in the system to cope with tie-break situations if the inter-site link fails. The IP quorum application is a Java application that runs on a host at the third site. The IP quorum application enables the use of a lower-cost IP network-attached host as a quorum disk for simplified implementation and operation.

Note: IP Quorum allows the user to replace a third-site Fibre Channel-attached quorum disk with an IP Quorum application. The Java application runs on a Linux host and is used to resolve split-brain situations. Quorum disks are still required in sites 1 and 2 for cookie crumb and metadata. The application can also be used with clusters in a standard topology configuration, but the primary use case is a customer with a cluster split over two sites (stretched or HyperSwap).

You need Java to run the IP quorum. Your Network must provide as least < 80 ms round-trip latency. All nodes need a service IP address, and all service IP addresses must be pingable from the quorum host. The maximum number of IP quorum applications that can be deployed is five. Applications can be deployed on multiple hosts to provide redundancy.

1.6.8 Storage pools

A *storage pool* (up to 1024 per system) is a collection of MDisks (up to 128) that are grouped to provide capacity for volumes. All MDisks in the pool are split into extents of the same size. Volumes are then allocated out of the storage pool and are mapped to a host system.

MDisks can be added to a storage pool at any time to increase the capacity of the pool. MDisks can belong in only one storage pool. For more information, see Chapter 4, “Storage pools” on page 159.

Each MDisk in the storage pool is divided into a number of extents. The size of the extent is selected by the administrator when the storage pool is created and cannot be changed later. The size of the extent ranges from 16 MB - 8 GB.

Default extent size: The GUI of IBM Storwize V5000 Gen2 has a default extent size value of 1024 MB when you define a new storage pool.

The extent size directly affects the maximum volume size and storage capacity of the clustered system.

A system can manage 2^{22} (4,194,304) extents. For example, with a 16 MB extent size, the system can manage up to 16 MB x 4,194,304 = 64 TB of storage.

The effect of extent size on the maximum volume and cluster size is shown in Table 1-12.

Table 1-12 Maximum volume and cluster capacity by extent size

Extent size (MB)	Maximum volume capacity for normal volumes (GB)	Maximum storage capacity of cluster
16	2,048 (2 TB)	64 TB
32	4,096 (4 TB)	128 TB
64	8,192 (8 TB)	256 TB
128	16,384 (16 TB)	512 TB
256	32,768 (32 TB)	1 PB
512	65,536 (64 TB)	2 PB
1024	131,072 (128 TB)	4 PB
2048	262,144 (256 TB)	8 PB
4096	262,144 (256 TB)	16 PB
8192	262,144 (256 TB)	32 PB

Use the same extent size for all storage pools in a clustered system. This rule is a prerequisite if you want to migrate a volume between two storage pools. If the storage pool extent sizes are not the same, you must use volume mirroring to copy volumes between storage pools, as described in Chapter 4, “Storage pools” on page 159.

You can set a threshold warning for a storage pool that automatically issues a warning alert when the used capacity of the storage pool exceeds the set limit.

Child storage pools

Instead of being created directly from MDisks, *child pools* are created from existing capacity that is allocated to a parent pool. As with parent pools, volumes can be created that specifically use the capacity that is allocated to the child pool. Parent pools grow automatically as more MDisks are allocated to them. However, child pools provide a fixed capacity pool of storage. You can use a child pool to manage a quota of storage for a particular purpose.

Child pools can be created by using the management GUI, CLI, or IBM Spectrum Control when you create VMware vSphere virtual volumes. For more information about child pools, see Chapter 4, “Storage pools” on page 159.

Single-tiered storage pool

MDisks that are used in a single-tiered storage pool must have the following characteristics to prevent performance problems and other problems:

- ▶ They have the same hardware characteristics; for example, the same RAID type, RAID array size, disk type, and disk revolutions per minute (RPMs).
- ▶ The disk subsystems that provide the MDisks have similar characteristics; for example, maximum input/output operations per second (IOPS), response time, cache, and throughput.
- ▶ Use MDisks of the same size and ensure that the MDisks provide the same number of extents. If this configuration is not feasible, you must check the distribution of the volumes’ extents in that storage pool.

Multi-tiered storage pool

A *multi-tiered storage pool* has a mix of MDisks with more than one type of disk; for example, a storage pool that contains a mix of generic_hdd *and* generic_ssd MDisks.

A multi-tiered storage pool contains MDisks with different characteristics, unlike the single-tiered storage pool. MDisks with similar characteristics then form the tiers within the pool. However, each tier must have MDisks of the same size and that provide the same number of extents.

A multi-tiered storage pool is used to enable automatic migration of extents between disk tiers by using the IBM Storwize V5000 Gen2 IBM Easy Tier function, as described in Chapter 9, “Advanced features for storage efficiency” on page 435.

This functionality can help improve the performance of host volumes on the IBM Storwize V5000.

1.6.9 Volumes

A *volume* is a logical disk that is presented to a host system by the clustered system. In our virtualized environment, the host system has a volume that is mapped to it by IBM Storwize V5000 Gen2. IBM Storwize V5000 Gen2 translates this volume into a number of extents, which are allocated across MDisks. The advantage with storage virtualization is that the host is decoupled from the underlying storage, so the virtualization appliance can move around the extents without affecting the host system.

The host system cannot directly access the underlying MDisks in the same manner as it can access RAID arrays in a traditional storage environment.

The following types of volumes are available:

- ▶ **Striped**

A striped volume is allocated one extent in turn from each MDisk in the storage pool. This process continues until the space that is required for the volume is satisfied.

It also is possible to supply a list of MDisks to use.

Figure 1-10 on page 26 shows how a striped volume is allocated, assuming that 10 extents are required.

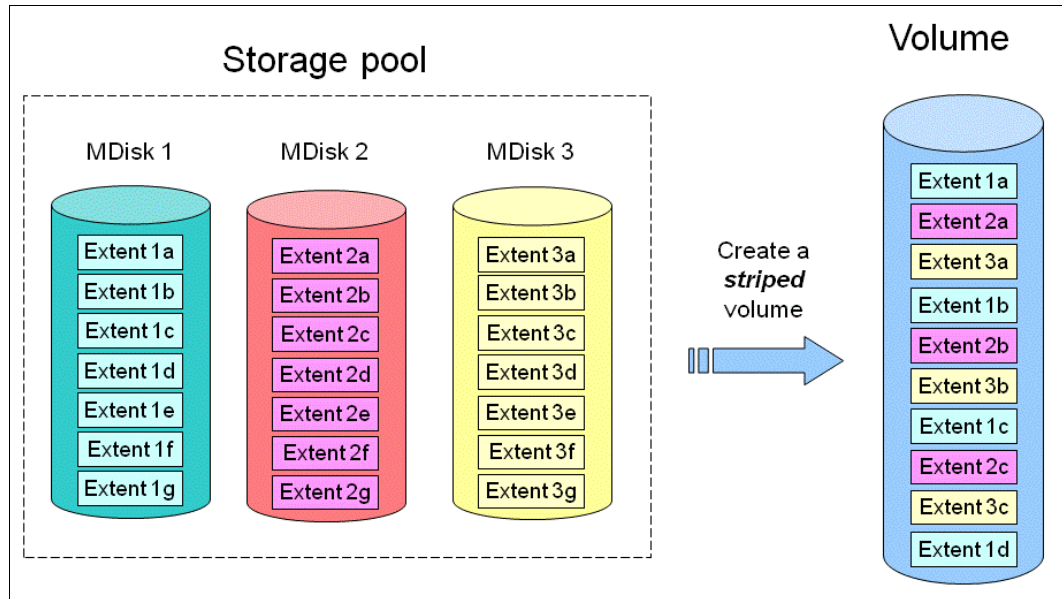


Figure 1-10 Striped volume

► Sequential

A sequential volume is a volume in which the extents are allocated one after the other from one MDisk to the next MDisk, as shown in Figure 1-11.

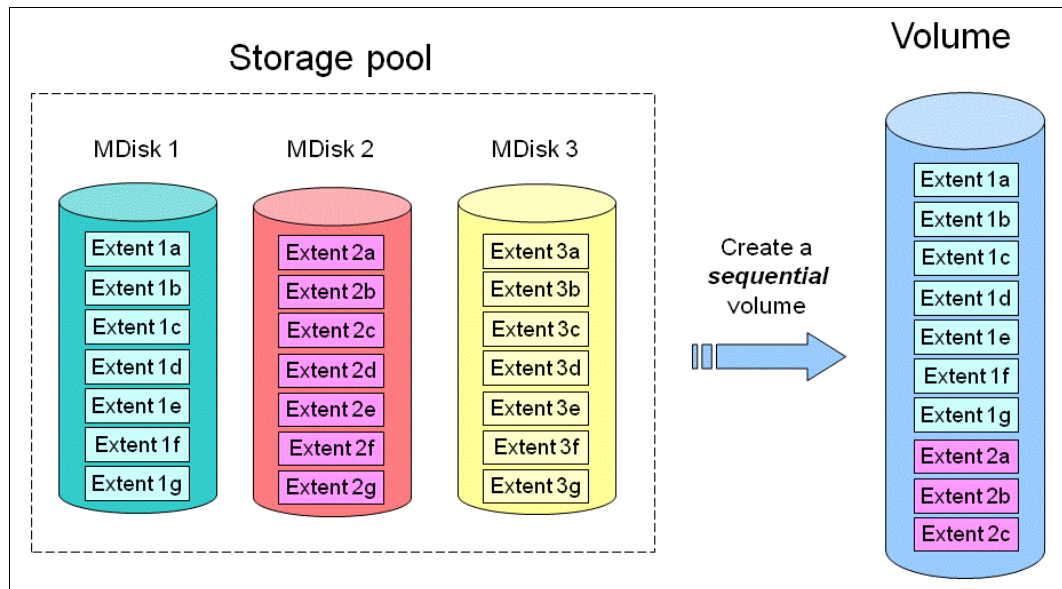


Figure 1-11 Sequential volume

► Image mode

Image mode volumes are special volumes that have a direct relationship with one MDisk. They are used to migrate data into and out of the clustered system to or from external FC SAN-attached storage.

When the image mode volume is created, a direct mapping is made between extents that are on the MDisk and the extents that are on the volume. The logical block address (LBA) x on the MDisk is the same as the LBA x on the volume, which ensures that the data on the MDisk is preserved as it is brought into the clustered system, as shown in Figure 1-12.

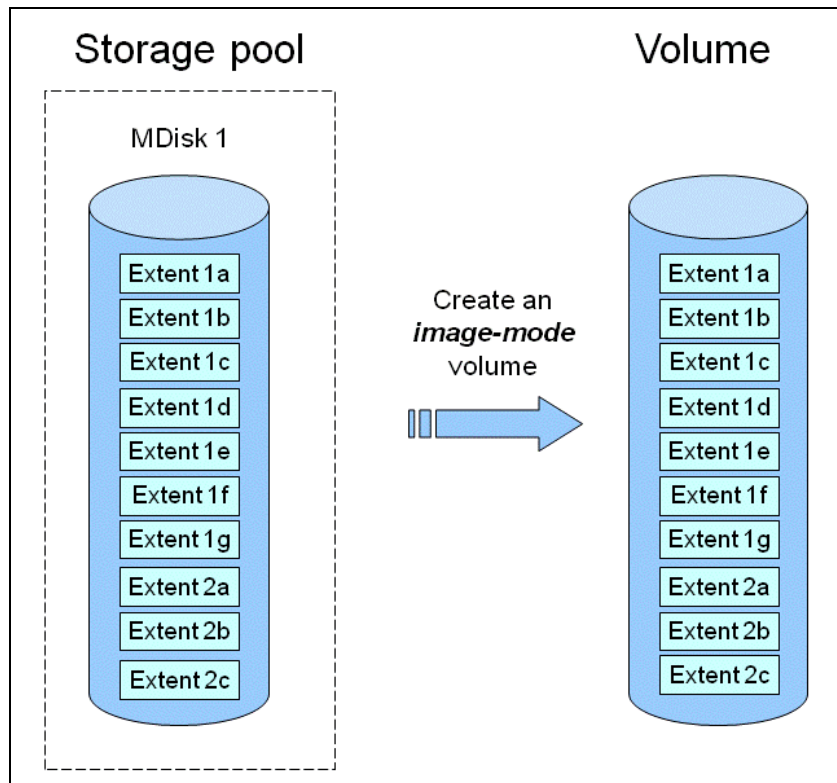


Figure 1-12 Image mode volume

Certain virtualization functions are not available for image mode volumes, so it is often useful to migrate the volume into a new storage pool. After it is migrated, the MDisk becomes a managed MDisk.

If you want to migrate data from a storage subsystem, use the storage migration wizard, which guides you through the process.

For more information, see Chapter 7, “Storage migration” on page 357.

If you add an MDisk that contains data to a storage pool, any data on the MDisk is lost. If you are presenting externally virtualized LUNs that contain data to an IBM Storwize V5000 Gen2, import them as image mode volumes to ensure data integrity or use the migration wizard.

1.6.10 iSCSI

iSCSI is an alternative method of attaching hosts to the IBM Storwize V5000 Gen2. The iSCSI function is a software function that is provided by the IBM Storwize V5000 Gen2 code, not hardware. In the simplest terms, iSCSI allows the transport of SCSI commands and data over an Internet Protocol network that is based on IP routers and Ethernet switches.

iSCSI is a block-level protocol that encapsulates SCSI commands into TCP/IP packets and uses an IP network instead of requiring FC host bus adapters (HBAs) and a SAN fabric infrastructure. Concepts of names and addresses are carefully separated in iSCSI.

An iSCSI name is a location-independent, permanent identifier for an iSCSI node. An iSCSI node has one iSCSI name, which stays constant for the life of the node. The terms *initiator name* and *target name* also refer to an iSCSI name.

An iSCSI address specifies the iSCSI name of an iSCSI node and a location of that node. The address consists of a host name or IP address, a TCP port number (for the target), and the iSCSI name of the node. An iSCSI node can have any number of addresses, which can change at any time, particularly if they are assigned by way of Dynamic Host Configuration Protocol (DHCP). An IBM Storwize V5000 node represents an iSCSI node and provides statically allocated IP addresses.

Each iSCSI node that is an initiator or target has a unique IQN, which can have a size of up to 255 bytes. The IQN is formed according to the rules that were adopted for Internet nodes. The IQNs can be abbreviated by using a descriptive name, which is known as an *alias*. An alias can be assigned to an initiator or a target.

For more information about configuring iSCSI, see Chapter 4, “Storage pools” on page 159.

1.6.11 Serial-attached SCSI

The serial-attached SCSI (SAS) standard is an alternative method of attaching hosts to the IBM Storwize V5000 Gen2. The IBM Storwize V5000 Gen2 supports direct SAS host attachment to address easy-to-use, affordable storage needs. Each SAS port device has a worldwide unique 64-bit SAS address and operates at 12 Gbps.

1.6.12 Fibre Channel

Fibre Channel (FC) is the traditional method that is used for data center storage connectivity. The IBM Storwize V5000 Gen2 supports FC connectivity at speeds of 4, 8, and 16 Gbps. Fibre Channel Protocol is used to encapsulate SCSI commands over the FC network. Each device in the network has a unique 64-bit worldwide port name (WWPN). The IBM Storwize V5000 Gen2 supports FC connections directly to a host server or to external FC switched fabrics.

1.7 IBM Storwize V5000 Gen2 features

In this section, we describe the features of the IBM Storwize V5000 Gen2. Different models offer a different range of features. See Table 1-3 on page 6 for a comparison.

1.7.1 Mirrored volumes

IBM Storwize V5000 Gen2 provides a function that is called *storage volume mirroring*, which enables a volume to have two physical copies. Each volume copy can belong to a different storage pool and be on a different physical storage system to provide a high-availability (HA) solution. Each mirrored copy can be a generic, thin-provisioned, or compressed volume copy.

When a host system issues a write to a mirrored volume, IBM Storwize V5000 Gen2 writes the data to both copies. When a host system issues a read to a mirrored volume, IBM Storwize V5000 Gen2 requests it from the primary copy.

If one of the mirrored volume copies is temporarily unavailable, the IBM Storwize V5000 Gen2 automatically uses the alternative copy without any outage for the host system. When the mirrored volume copy is repaired, IBM Storwize V5000 Gen2 synchronizes the data again.

A mirrored volume can be converted into a non-mirrored volume by deleting one copy or by splitting away one copy to create a non-mirrored volume.

The use of mirrored volumes can also assist with migrating volumes between storage pools that have different extent sizes. Mirrored volumes can also provide a mechanism to migrate fully allocated volumes to thin-provisioned or compressed volumes without any host outages.

The Volume Mirroring feature is included as part of the base software, and no license is required.

Figure 1-13 shows an example of a mirrored volume.

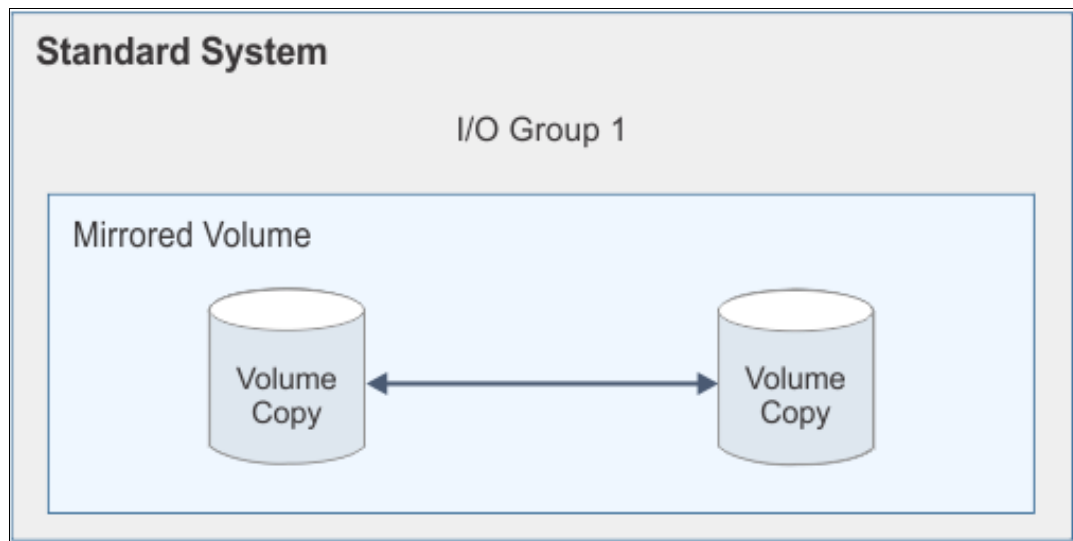


Figure 1-13 Example of a mirrored Volume

HyperSwap volumes are where copies of a single volume are in different storage pools that are at different sites. As Figure 1-14 on page 30 shows, the volume is cached in two I/O groups that are on different sites. These volumes can be created only on Storwize V5030 and Storwize V5030F systems when the system topology is HyperSwap.

Figure 1-14 on page 30 also shows an example of a HyperSwap Volume.

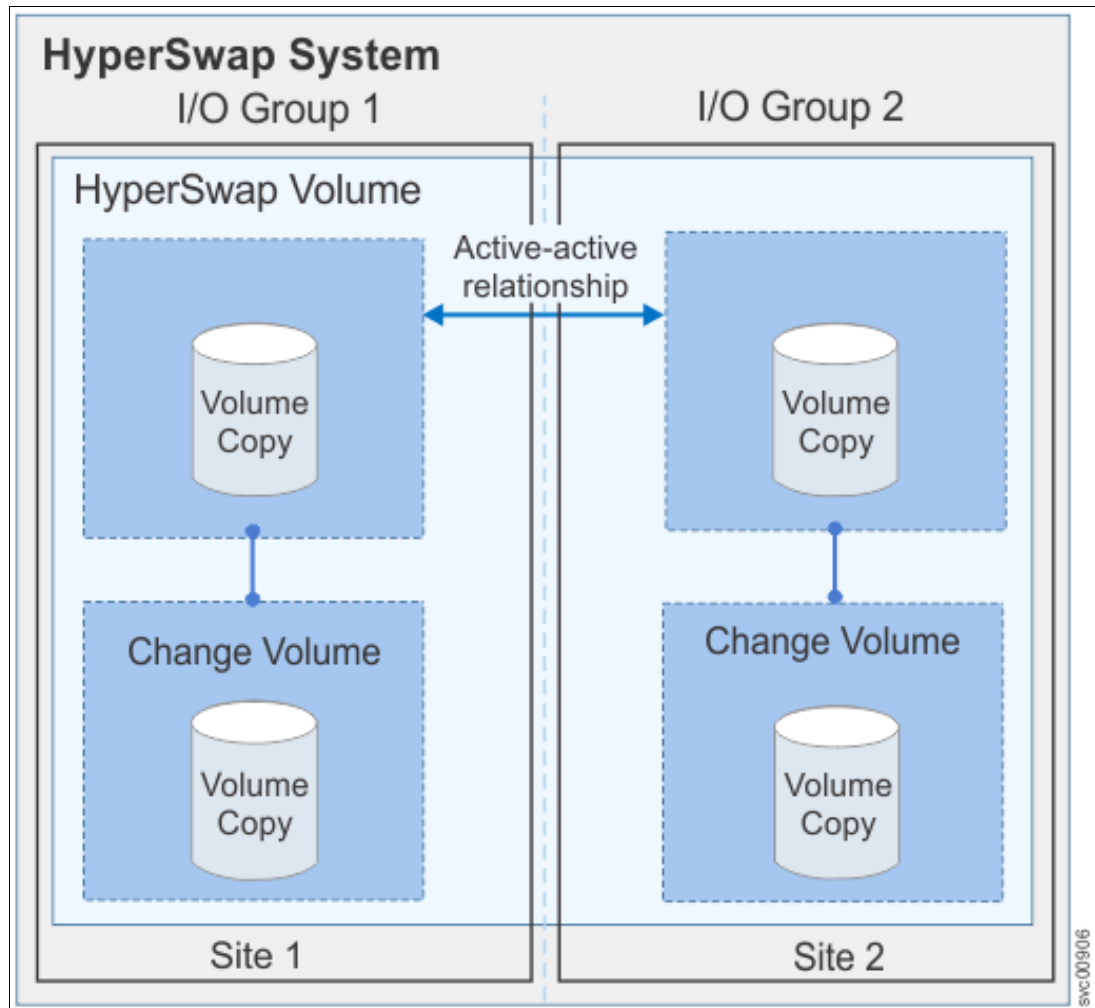


Figure 1-14 Example of a HyperSwap Volume

1.7.2 Thin provisioning

Volumes can be configured to be *thin-provisioned* or *fully allocated*. A thin-provisioned volume behaves as though it were a fully allocated volume in terms of read/write I/O. However, when a volume is created, the user specifies two capacities: the real capacity of the volume and its virtual capacity.

The *real capacity* determines the quantity of MDisk extents that are allocated for the volume. The *virtual capacity* is the capacity of the volume that is reported to IBM Storwize V5000 Gen2 and to the host servers.

The real capacity is used to store the user data and the metadata for the thin-provisioned volume. The real capacity can be specified as an absolute value or a percentage of the virtual capacity.

The Thin Provisioning feature can be used on its own to create over-allocated volumes, or it can be used with FlashCopy. Thin-provisioned volumes also can be used with the mirrored volume feature.

A thin-provisioned volume can be configured to *auto expand*, which causes the IBM Storwize V5000 Gen2 to automatically expand the real capacity of a thin-provisioned volume as it gets used. This feature prevents the volume from going offline. Auto expand attempts to maintain a fixed amount of unused real capacity on the volume. This amount is known as the *contingency capacity*.

When the thin-provisioned volume is initially created, the IBM Storwize V5000 Gen2 initially allocates only 2% of the virtual capacity in real physical storage. The contingency capacity and auto expand features seek to preserve this 2% of free space as the volume grows.

If the user modifies the real capacity, the contingency capacity is reset to be the difference between the used capacity and real capacity. In this way, the autoexpand feature does not cause the real capacity to grow much beyond the virtual capacity.

A volume that is created with a zero contingency capacity goes offline when it must expand. A volume with a non-zero contingency capacity stays online until it is used up.

To support the auto expansion of thin-provisioned volumes, the volumes themselves have a configurable warning capacity. When the used free capacity of the volume exceeds the warning capacity, a warning is logged.

For example, if a warning of 80% is specified, the warning is logged when 20% of the free capacity remains. This approach is similar to the capacity warning that is available on storage pools.

A thin-provisioned volume can be converted to a fully allocated volume or compressed volume by using volume mirroring (and vice versa).

The Thin Provisioning feature is included as part of the base software, and no license is required.

1.7.3 Real-time Compression

The Storwize V5030 model can create compressed volumes, which allows more data to be stored in the same physical space. IBM Real-time Compression™ (RtC) can be used for primary active volumes and with mirroring and replication (FlashCopy/Remote Copy). RtC is available on the Storwize V5030 model only.

Existing volumes can take advantage of Real-time Compression to result in an immediate capacity saving. A volume can be converted to a compressed volume by creating a compressed volume copy of the original volume followed by deleting the original volume.

No changes to the environment are required to take advantage of RtC. It is transparent to hosts while the compression occurs within the IBM Storwize V5000 Gen2 system.

Software-only compression: The use of RtC on the Storwize V5030 requires dedicated CPU resources from the node canisters. If more performance is required for deploying RtC, consider purchasing the Storwize V7000 system. The Storwize V7000 system uses dedicated hardware options for compression acceleration.

The Storwize V5030 model has the additional memory upgrade (32 GB for each node canister). When the first compressed volume is created 4 of the 6 CPU cores are allocated to RtC. Of the 32 GB of memory on each node canister, roughly 9 - 10 GB is allocated to RtC. There are no hardware compression accelerators as in the Storwize V7000 Gen2. The actual LZ4 compression is done by the CPUs as was the case with the Storwize V7000 Gen1.

Table 1-13 shows how the cores are used with RtC.

Table 1-13 Cores usage with RtC

Model	Compression Disabled		Compression Enabled	
	Normal Processing	RtC	Normal Processing	RtC
V5010	2 cores	NA	NA	NA
V5020	2 cores	NA	NA	NA
V5030	6 cores	0 cores	2 cores	4 cores + HT

The faster CPU with more cores, the extra memory and the hyper-threading capability of the Storwize V5030, and improvements to RtC software results in good performance for smaller customer configurations common to the market segment this product is intended to serve. The feature is licensed per enclosure. Conversely, RtC is not available on the Storwize V5010 or Storwize V5020 models.

1.7.4 Deduplication

A deduplicated volume or volume copy can be created in a data reduction pool. When you implement deduplication, you must consider 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; that is, when Random Access Compression Engine (RACE) compression is in use on that 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 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: The following software and hardware requirements are needed for deduplication. There are also update and performance considerations:

- ▶ Code level 8.1.2 or higher is needed for data reduction pools.
- ▶ Code level 8.1.3 or higher 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.

RACE compression and deduplication are not supported in the same I/O group. However, data reduction compression and deduplication might be supported on certain platforms. Figure 1-15 on page 33 shows the features that are supported on each platform.

Product	Platform	Node/canister memory (GBs)	Supported features			
			RACE	DRP	Compression	Deduplication
Storwize® V5000 Gen2	V5030 (32 GB)	32	Yes ¹	Yes	Yes ¹	Yes
	V5030 (16 GB)	16	No	Yes	No	No
	V5020	8/16	No	Yes	No	No
	V5010	8	No	Yes	No	No

Figure 1-15 Supported compression features

1 - Does not support data reduction compression and RACE compression in the same I/O group at the same time.nop

1.7.5 Easy Tier

IBM Easy Tier provides a mechanism to seamlessly migrate extents to the most appropriate tier within the IBM Storwize V5000 Gen2 solution. This migration can be to different tiers of internal drives within IBM Storwize V5000 Gen2 or to external storage systems that are virtualized by IBM Storwize V5000 Gen2; for example, an IBM FlashSystem® 900.

The Easy Tier function can be turned on or turned off at the storage pool and volume level.

You can demonstrate the potential benefit of Easy Tier in your environment before you install Flash drives by using the IBM Storage Advisor Tool. For more information about Easy Tier, see Chapter 9, “Advanced features for storage efficiency” on page 435.

The IBM Easy Tier feature is licensed per enclosure.

1.7.6 Storage Migration

By using the IBM Storwize V5000 Gen2 Storage Migration feature, you can easily move data from other existing Fibre Channel-attached external storage to the internal capacity of the IBM Storwize V5000 Gen2. You can migrate data from other storage to the IBM Storwize V5000 Gen2 storage system to realize the benefits of the IBM Storwize V5000 Gen2 with features, such as the easy-to-use GUI, internal virtualization, thin provisioning, and copy services.

The Storage Migration feature is included in the base software, and no license is required.

1.7.7 FlashCopy

The FlashCopy feature copies a source volume on to a target volume. The original contents of the target volume is lost. After the copy operation starts, the target volume has the contents of the source volume as it existed at a single point in time. Although the copy operation completes in the background, the resulting data at the target appears as though the copy was made instantaneously. FlashCopy is sometimes described as an instance of a *time-zero* (T0) copy or *point-in-time* (PiT) copy technology.

FlashCopy can be performed on multiple source and target volumes. FlashCopy permits the management operations to be coordinated so that a common single point in time is chosen for copying target volumes from their respective source volumes.

IBM Storwize V5000 Gen2 also permits multiple target volumes to be FlashCopies from the same source volume. This capability can be used to create images from separate points in time for the source volume, and to create multiple images from a source volume at a common point in time. Source and target volumes can be any volume type (generic, thin-provisioned, or compressed).

Reverse FlashCopy enables target volumes to become restore points for the source volume without breaking the FlashCopy relationship and without waiting for the original copy operation to complete. IBM Storwize V5000 Gen2 supports multiple targets and multiple rollback points.

The FlashCopy feature is licensed per enclosure.

For more information about FlashCopy copy services, see Chapter 10, “Copy Services” on page 465.

1.7.8 Remote Copy

Remote Copy can be implemented in one of two modes: synchronous or asynchronous.

With the IBM Storwize V5000 Gen2, Metro Mirror and Global Mirror are the IBM branded terms for the functions that are synchronous Remote Copy and asynchronous Remote Copy.

By using the Metro Mirror and Global Mirror copy services features, you can set up a relationship between two volumes so that updates that are made by an application to one volume are mirrored on the other volume. The volumes can be in the same system or on two different systems.

For both Metro Mirror and Global Mirror copy types, one volume is designated as the primary and the other volume is designated as the secondary. Host applications write data to the primary volume, and updates to the primary volume are copied to the secondary volume. Normally, host applications do not perform I/O operations to the secondary volume.

The Metro Mirror feature provides a synchronous copy process. When a host writes to the primary volume, it does not receive confirmation of I/O completion until the write operation completes for the copy on the primary and secondary volumes. This design ensures that the secondary volume is always up-to-date with the primary volume if a failover operation must be performed.

The Global Mirror feature provides an asynchronous copy process. When a host writes to the primary volume, confirmation of I/O completion is received before the write operation completes for the copy on the secondary volume. If a failover operation is performed, the application must recover and apply any updates that were not committed to the secondary volume. If I/O operations on the primary volume are paused for a brief time, the secondary volume can become a match of the primary volume.

Global Mirror can operate with or without cycling. When it is operating without cycling, write operations are applied to the secondary volume as soon as possible after they are applied to the primary volume. The secondary volume is less than 1 second behind the primary volume, which minimizes the amount of data that must be recovered in a failover. However, this approach requires that a high-bandwidth link is provisioned between the two sites.

When Global Mirror operates with cycling mode, changes are tracked and where needed copied to intermediate change volumes. Changes are transmitted to the secondary site periodically. The secondary volumes are much further behind the primary volume, and more data must be recovered in a failover. Because the data transfer can be smoothed over a longer time, lower bandwidth is required to provide an effective solution.

For more information about the IBM Storwize V5000 Gen2 copy services, see Chapter 10, “Copy Services” on page 465.

The IBM Remote Copy feature is licensed for each enclosure.

1.7.9 IP replication

IP replication enables the use of lower-cost Ethernet connections for remote mirroring. The capability is available as a chargeable option on all Storwize family systems.

The function is transparent to servers and applications in the same way that traditional Fibre Channel-based mirroring is transparent. All remote mirroring modes (Metro Mirror, Global Mirror, and Global Mirror with Change Volumes) are supported.

Configuration of the system is straightforward. The Storwize family systems normally find each other in the network, and they can be selected from the GUI.

IP replication includes Bridgeworks SANSlide network optimization technology, and it is available at no additional charge. Remember, Remote Mirror is a chargeable option but the price does not change with IP replication. Existing Remote Mirror users have access to the function at no additional charge.

IP connections that are used for replication can have long *latency* (the time to transmit a signal from one end to the other), which can be caused by distance or by many “hops” between switches and other appliances in the network. Traditional replication solutions transmit data, wait for a response, and then transmit more data, which can result in network usage as low as 20% (based on IBM measurements). Also, this scenario gets worse the longer the latency.

Bridgeworks SANSlide technology that is integrated with the IBM Storwize family requires no separate appliances, no additional cost, and no configuration steps. It uses artificial intelligence (AI) technology to transmit multiple data streams in parallel, adjusting automatically to changing network environments and workloads.

SANSlide improves network bandwidth usage up to 3x so clients can deploy a less costly network infrastructure or take advantage of faster data transfer to speed up replication cycles, improve remote data currency, and enjoy faster recovery.

IP replication can be configured to use any of the available 1 GbE or 10 GbE Ethernet ports (apart from the technician port) on the IBM Storwize V5000 Gen2. For more information about port configuration options, see Table 1-10 on page 18.

Copy services configuration limits

For the most up-to-date list of these limits, see [this website](#).

1.7.10 External virtualization

By using this feature, you can consolidate FC SAN-attached disk controllers from various vendors into pools of storage. In this way, the storage administrator can manage and provision storage to applications from a single user interface and use a common set of advanced functions across all of the storage systems under the control of the IBM Storwize V5000 Gen2. External virtualization is only available for the IBM Storwize V5030.

The External Virtualization feature is licensed per disk enclosure.

1.7.11 Encryption

IBM Storwize V5000 Gen2 provides optional encryption of data-at-rest functionality, which protects against the potential exposure of sensitive user data and user metadata that is stored on discarded, lost, or stolen storage devices. Encryption can be enabled and configured only on the Storwize V5020 and Storwize V5030 enclosures that support encryption. The Storwize V5010 does not offer encryption functionality.

Encryption is a licensed feature that requires a license key to enable it before it can be used.

1.8 Problem management and support

In this section, we introduce problem management and support topics.

1.8.1 IBM Support assistance

To use IBM Support assistance, you must have internet access. Support assistance enables support personnel to access the system to complete troubleshooting and maintenance tasks. You can configure local support assistance, where support personnel visit your site to fix problems with the system, or remote support assistance.

Both local and remote support assistance, uses secure connections to protect data exchange between the support center and system. More access controls can be added by the system administrator. You can use the management GUI or the command-line interface to view support assistance settings.

1.8.2 Event notifications

IBM Storwize V5000 Gen2 can use Simple Network Management Protocol (SNMP) traps, syslog messages, and email to notify you and the IBM Support Center when significant events are detected. Any combination of these notification methods can be used simultaneously.

You can configure IBM Storwize V5000 Gen2 to send different types of notification to specific recipients and choose the alerts that are important to you. When you configure Call Home to the IBM Support Center, all events are sent through email only.

1.8.3 SNMP traps

SNMP is a standard protocol for managing networks and exchanging messages. IBM Storwize V5000 Gen2 can send SNMP messages that notify personnel about an event. You can use an SNMP manager to view the SNMP messages that IBM Storwize V5000 Gen2 sends. You can use the management GUI or the IBM Storwize V5000 Gen2 CLI to configure and modify your SNMP settings.

You can use the Management Information Base (MIB) file for SNMP to configure a network management program to receive SNMP messages that are sent by the IBM Storwize V5000 Gen2. This file can be used with SNMP messages from all versions of IBM Storwize V5000 Gen2 software.

1.8.4 Syslog messages

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. IBM Storwize V5000 Gen2 can send syslog messages that notify personnel about an event. IBM Storwize V5000 Gen2 can transmit syslog messages in expanded or concise format. You can use a syslog manager to view the syslog messages that IBM Storwize V5000 Gen2 sends. IBM Storwize V5000 Gen2 uses the User Datagram Protocol (UDP) to transmit the syslog message. You can use the management GUI or the CLI to configure and modify your syslog settings.

1.8.5 Call Home email

The Call Home feature transmits operational and error-related data to you and IBM through a Simple Mail Transfer Protocol (SMTP) server connection in the form of an event notification email. When configured, this function alerts IBM service personnel about hardware failures and potentially serious configuration or environmental issues. You can use the Call Home function if you have a maintenance contract with IBM or if the IBM Storwize V5000 Gen2 is within the warranty period.

To send email, you must configure at least one SMTP server. You can specify as many as five other SMTP servers for backup purposes. The SMTP server must accept the relaying of email from the IBM Storwize V5000 Gen2 clustered system IP address. You can then use the management GUI or the CLI to configure the email settings, including contact information and email recipients.

Set the reply address to a valid email address. Send a test email to check that all connections and infrastructure are set up correctly. You can disable the Call Home function at any time by using the management GUI or the CLI.

1.9 More information resources

This section describes resources that are available for more information.

1.9.1 Useful IBM Storwize V5000 Gen2 websites

For more information about IBM Storwize V5000 Gen2, see the following websites:

- ▶ The IBM Storwize V5000 Gen2 [home page](#)
- ▶ IBM Storwize V5000 Gen2 Knowledge Center [web page](#)
- ▶ IBM Storwize V5000 Gen2 Online Announcement Letters:
 - [Storwize V5000 general](#)
 - [Storwize V5000 Model 2077 1 year warranty](#)
 - [hStorwize V5000 Model 2078 3 year warranty](#)

The Online Information Center also includes a Learning and Tutorial section where you can obtain videos that describe the use and implementation of the IBM Storwize V5000 Gen2.



Initial configuration

This chapter describes the initial configuration steps for the IBM Storwize V5000 Gen2 and includes the following topics:

- ▶ 2.1, “Hardware installation planning” on page 40
- ▶ 2.2, “SAN configuration planning” on page 44
- ▶ 2.3, “FC direct-attach planning” on page 46
- ▶ 2.4, “SAS direct-attach planning” on page 48
- ▶ 2.5, “LAN configuration planning” on page 50
- ▶ 2.6, “Host configuration planning” on page 52
- ▶ 2.7, “Miscellaneous configuration planning” on page 53
- ▶ 2.8, “System management” on page 54
- ▶ 2.9, “First-time setup” on page 56
- ▶ 2.10, “Initial configuration” on page 61

2.1 Hardware installation planning

After you verify that you have all of the hardware components that you purchased, it is important to carefully plan for physical installation. The following checklist of requirements can be used to plan your installation:

1. Install the hardware as described in Chapter 2 of *IBM Storwize V5000 Gen2 Quick Installation Guide*, [GC27-8581](#).
2. An appropriate 19-inch rack must be available. Depending on the number of enclosures to install, more than one might be required. Each enclosure measures 2 U. A single Storwize V5010 or Storwize V5020 control enclosure supports up to 10 expansion enclosures. A single Storwize V5030 control enclosure supports up to 20 expansion enclosures.
3. Redundant power outlets must be in the rack for each of the two power cords that are required for each enclosure to be installed. Several power outlets are required, depending on the number of enclosures to be installed. The power cords must conform to the IEC320 C13/C14 standards.
4. A minimum of four Fibre Channel ports that are attached to redundant fabrics are required. For dual I/O group systems (two V5030), a minimum of eight Fibre Channel ports are required.

Fibre Channel ports: Fibre Channel (FC) ports are required only if you are using FC hosts or clustered systems that are arranged as two I/O groups. You can use the IBM Storwize V5000 Gen2 with Ethernet-only cabling for Internet Small Computer System Interface (iSCSI) hosts or use serial-attached SCSI (SAS) cabling for hosts that are directly attached. Also HyperSwap is supported over Ethernet.

5. For the Storwize V5020 systems, up to two hosts can be directly connected by using SAS ports 2 and 3 on each node canister, with SFF-8644 mini SAS HD cabling.
6. You must have a minimum of two Ethernet ports on the LAN, with four preferred for more redundancy or iSCSI host access.
7. You must have a minimum of two Ethernet cable connected, four for more redundancy or iSCSI host access. If you have two I/O groups, you must have a minimum of four Ethernet cable connected. Ethernet port 1 on each node canister must be connected to the LAN (for Management purposes), with port two as optional.

LAN connectivity: Port 1 on each node canister must be connected to the same physical local area network (LAN) or be configured in the same virtual LAN (VLAN) and be on the same subnet or set of subnets.

Technician port: On the Storwize V5010 and V5020 models, Port 2 is the *technician port*, which is used for system initialization and service. Port 2 must not be connected to the LAN until the system initialization or service is complete.

The Storwize V5030 model features a dedicated technician port.

8. The 10 Gb Ethernet (copper) ports of a Storwize V5030 system require a Category 6A shielded cable that is terminated with an 8P8C modular connector (RJ45 compatible connector) to function at 10 Gb.
9. Verify that the default IP addresses that are configured on Ethernet port 1 on each of the node canisters (192.168.70.121 on node 1 and 192.168.70.122 on node 2) do not conflict with existing IP addresses on the LAN. The default mask that is used with these IP addresses is 255.255.255.0, and the default gateway address that is used is 192.168.70.1.

10. You need a minimum of three IPv4 or IPv6 IP addresses for systems that are arranged as one I/O group and minimum of five if you have two I/O groups. One is for the clustered system and is used by the administrator for management, and one for each node canister for service access as needed.

IP addresses: An additional IP address must be used for backup configuration access. This other IP address allows a second system IP address to be configured on port 2 of either node canister, which the storage administrator can also use for management of the IBM Storwize V5000 Gen2 system.

11. A minimum of one and up to eight IPv4 or IPv6 addresses are needed if iSCSI-attached hosts access volumes from the IBM Storwize V5000 Gen2.
12. At least two 0.6-meter (1.96 feet), 1.5-meter (4.9 feet), or 3-meter (9.8 feet) 12 Gb mini-SAS cables are required for each expansion enclosure. The length of the cables depends on the physical rack location of the expansion enclosure relative to the control enclosures or other expansion enclosures.

2.1.1 Procedure to install the SAS cables

In this section, we describe the procedures to install the SAS cables for the different models.

Storwize V5010 and Storwize V5020

The Storwize V5010 and Storwize V5020 support up to 10 expansion enclosures in a single chain. To install the cables, complete the following steps:

1. By using the supplied SAS cables, connect the control enclosure to the first expansion enclosure:
 - a. Connect SAS port 1 of the left node canister in the control enclosure to SAS port 1 of the left expansion canister in the first expansion enclosure.
 - b. Connect SAS port 1 of the right node canister in the control enclosure to SAS port 1 of the right expansion canister in the first expansion enclosure.
2. To connect a second expansion enclosure, use the supplied SAS cables to connect it to the previous enclosure in the chain:
 - a. Connect SAS port 2 of the left canister in the previous expansion enclosure to SAS port 1 of the left expansion canister in the next expansion enclosure.
 - b. Connect SAS port 2 of the right canister in the previous expansion enclosure to SAS port 1 of the right expansion canister in the next expansion enclosure.
3. Repeat steps 1 and 2 until all expansion enclosures are connected.

Figure 2-1 shows how to connect the Expansions to a Storwize V5010.

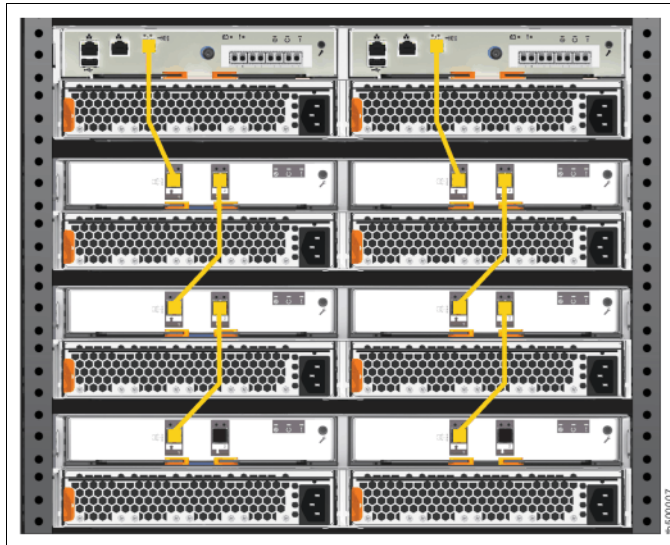


Figure 2-1 Storwize V5010 SAS cabling

Figure 2-2 shows how to connect the Expansions to a Storwize V5020.

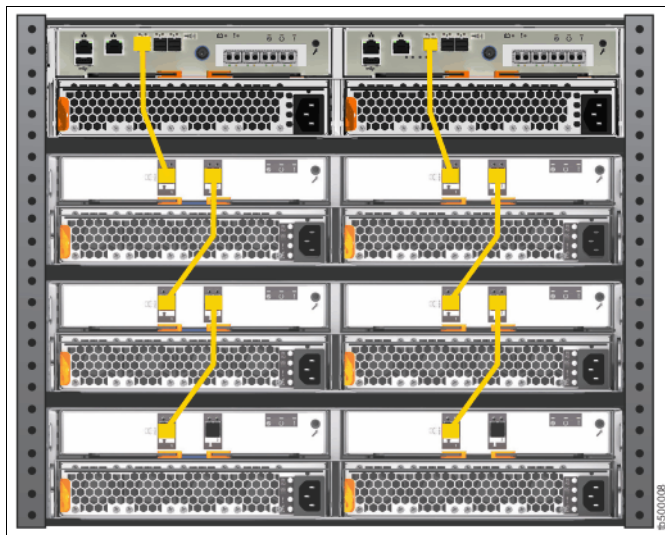


Figure 2-2 Storwize V5020 SAS cabling

Storwize V5030

The Storwize V5030 supports up to 20 expansion enclosures per I/O group in two SAS chains of 10. Up to 40 expansion enclosures can be supported in a two I/O group configuration. To install the cables, complete the following steps:

1. By using the supplied SAS cables, connect the control enclosure to first expansion enclosure by using the first chain:
 - a. Connect SAS port 1 of the left node canister in the control enclosure to SAS port 1 of the left expansion canister in the first expansion enclosure.
 - b. Connect SAS port 1 of the right node canister in the control enclosure to SAS port 1 of the right expansion canister in the first expansion enclosure.
2. To connect a second expansion enclosure, use the supplied SAS cables to connect the control enclosure to expansion enclosure by using the second chain:
 - a. Connect SAS port 2 of the left node canister in the control enclosure to SAS port 1 of the left expansion canister in the second expansion enclosure.
 - b. Connect SAS port 2 of the right node canister in the control enclosure to SAS port 1 of the right expansion canister in the second expansion enclosure.
3. To connect additional expansion enclosures, alternate connecting them between chain one and chain two to keep the configuration balanced:
 - a. Connect SAS port 2 of the left canister in the previous expansion enclosure to SAS port 1 of the left expansion canister in the next expansion enclosure.
 - b. Connect SAS port 2 of the right canister in the previous expansion enclosure to SAS port 1 of the right expansion canister in the next expansion enclosure.
4. Repeat these steps until all expansion enclosures are connected.

Note: The controller should be placed in the middle of the rack. The first expansion should be placed above the controller, the second below, the third again above the controller, and so on. The reason for this configuration is that the SAS cables have a limited length.

Figure 2-3 shows how to connect the expansions to a Storwize V5030.

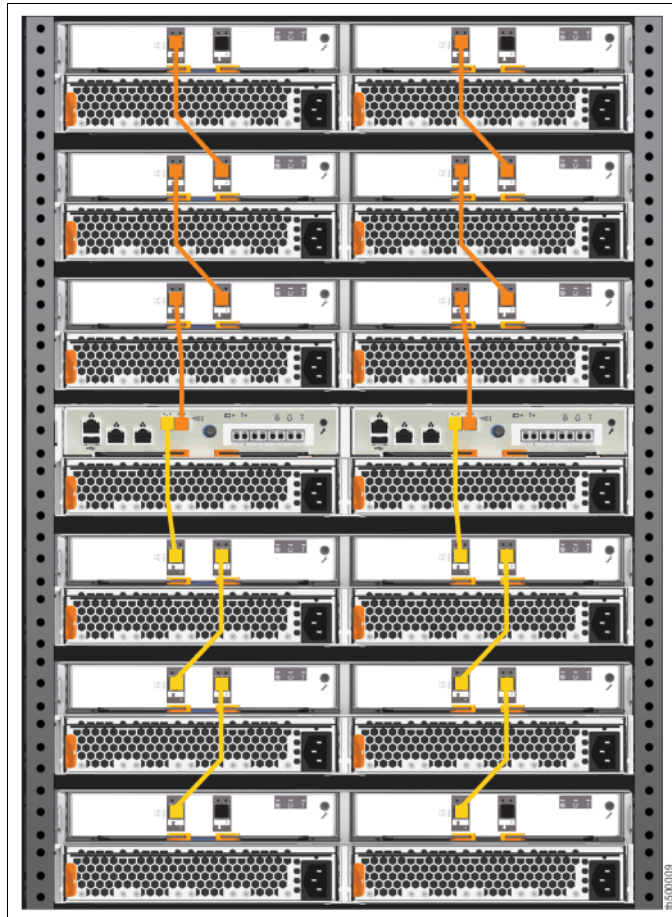


Figure 2-3 Storwize V5030 cabling

2.2 SAN configuration planning

To connect a Storwize V5000 Gen2 system to a Fibre Channel (FC) SAN, you must first install the optional 16 Gb FC adapters in every node canister that you connect. Ensure that you use the correct fibre cables to connect the Storwize V5000 Gen2 to the FC SAN. With the FC cards installed, the Storwize V5000 Gen2 can be connected to FC hosts, external storage controllers, and other Storwize systems that are visible on the SAN fabric.

The Storwize V5010 and V5020 models support a single I/O group only and can migrate from external storage controllers only. The Storwize V5030 supports up to two I/O groups that form a cluster over the network. In addition, the Storwize V5030 supports full virtualization of external storage controllers.

The advised SAN configuration consists of a minimum of two fabrics that encompass all host ports and any ports on external storage systems that are to be virtualized by the IBM Storwize V5000 Gen2. The IBM Storwize V5000 Gen2 ports must be identically cabled. They also must be evenly split between the two fabrics to provide redundancy if one of the fabrics goes offline (planned or unplanned).

Appropriated Zoning must be implemented before the IBM Storwize V5000 Gen2, hosts, and optional external storage systems are connected to the SAN fabrics. To enable the node canisters to communicate in band, create a zone with only the IBM Storwize V5000 Gen2 WWPNs (two from each node canister) on each of the two fabrics.

If an external storage system is to be virtualized, create a zone in each fabric with the IBM Storwize V5000 Gen2 worldwide port names (WWPNs) (two from each node canister) with up to a maximum of eight WWPNs from the external storage system. Assume that every host has a Fibre Channel connection to each fabric. Create a zone with the host WWPN and one WWPN from each node canister in the IBM Storwize V5000 Gen2 system in each fabric.

Important: It is mandatory that only one initiator host bus adapter (HBA) is in any zone.

For load balancing between the node ports on the IBM Storwize V5000 Gen2, alternate the host Fibre Channel ports between the ports of the IBM Storwize V5000 Gen2.

A maximum of eight SAN paths are allowed from each host to the IBM Storwize V5000 Gen2. Hosts that exceeded this number are not supported. The restriction limits the number of paths that the multipathing driver must resolve. A host with only two HBAs cannot exceed this limit with the correct zoning in a dual fabric SAN.

Maximum ports or WWPNs: The IBM Storwize V5000 Gen2 supports a maximum of 16 ports or WWPNs from a virtualized external storage system.

Figure 2-4 shows how to cable devices to the SAN. Optionally, ports 3 and 4 can be also connected to the SAN fabric to provide more redundancy and throughput. Refer to the example that is shown in Figure 2-4 when the zoning is described.

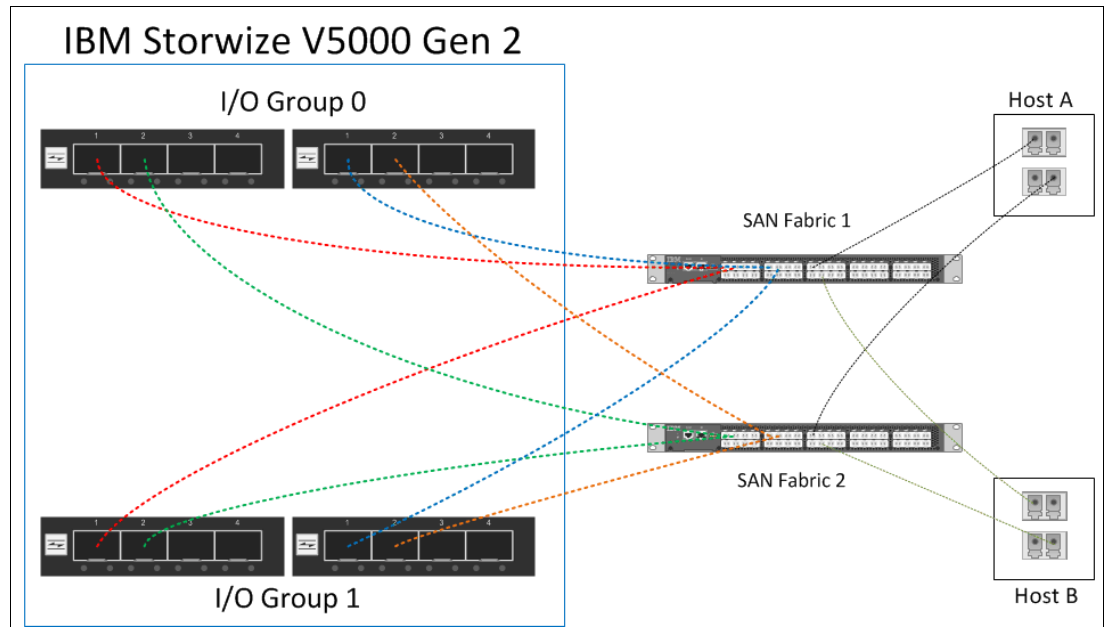


Figure 2-4 SAN cabling and zoning diagram

Create a host/IBM Storwize V5000 Gen2 zone for each server that has volumes mapped to and from the clustered system, as shown in the following examples in Figure 2-4 on page 45:

- ▶ Zone Host A port 1 (HBA 1) with all node canister ports 1
- ▶ Zone Host A port 2 (HBA 2) with all node canister ports 2
- ▶ Zone Host B port 1 (HBA 1) with all node canister ports 3
- ▶ Zone Host B port 2 (HBA 2) with all node canister ports 4

Similar zones must be created for all other hosts with volumes on the IBM Storwize V5000 Gen2 I/O groups.

Verify the interoperability of a IBM Storwize V5000 Gen2, which is connected to SAN switches or directors by following the requirements that are provided at [this website](#).

Switches or directors must be at the required firmware level, which are supported by the IBM Storwize V5000 Gen2. For more information about firmware levels, see [this website](#).

Important: The IBM Storwize V5000 Gen2 port login maximum could not be exceeded. For more information about restrictions, see [this document](#).

Connectivity issues: If you recognize connectivity issues between the IBM Storwize V5000 Gen2 ports and Brocade SAN switches or directors running at 8 Gbps, see [this website](#). It describes the correct setting of the `fillword port config` parameter in the Brocade Switch OS.

2.3 FC direct-attach planning

It is possible to connect Hosts directly to the IBM Storwize V5000 Gen2. There must be at least one FC cable from the Host to each node of the IBM Storwize V5000 Gen2. This ensures redundancy in case one of the controller nodes of the IBM Storwize V5000 Gen2 goes offline (see Figure 2-5 on page 47).

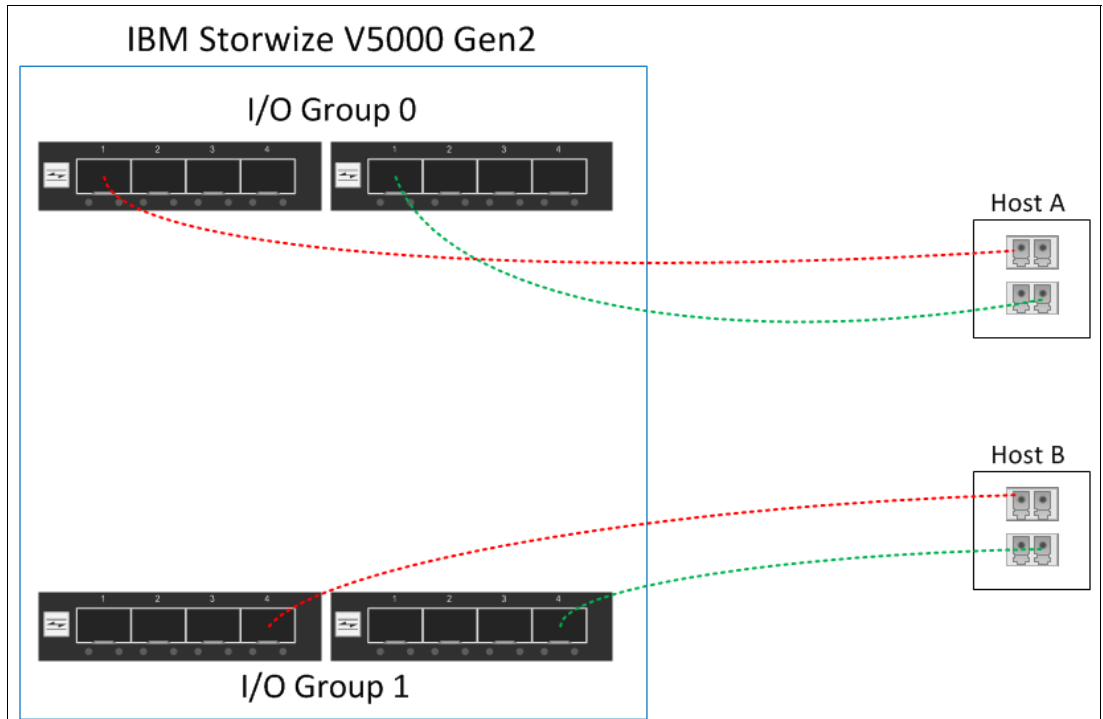


Figure 2-5 FC direct-attach host configuration

If you plan to connect the host directly in a clustered environment, it is recommended that you connect each controller of a I/O group with the host. At least four connections are needed (see Figure 2-6 for an example). This suggestion also applies to a cluster where one I/O group is a Storwize V5000 Gen1 and the other I/O group is a Storwize V5030.

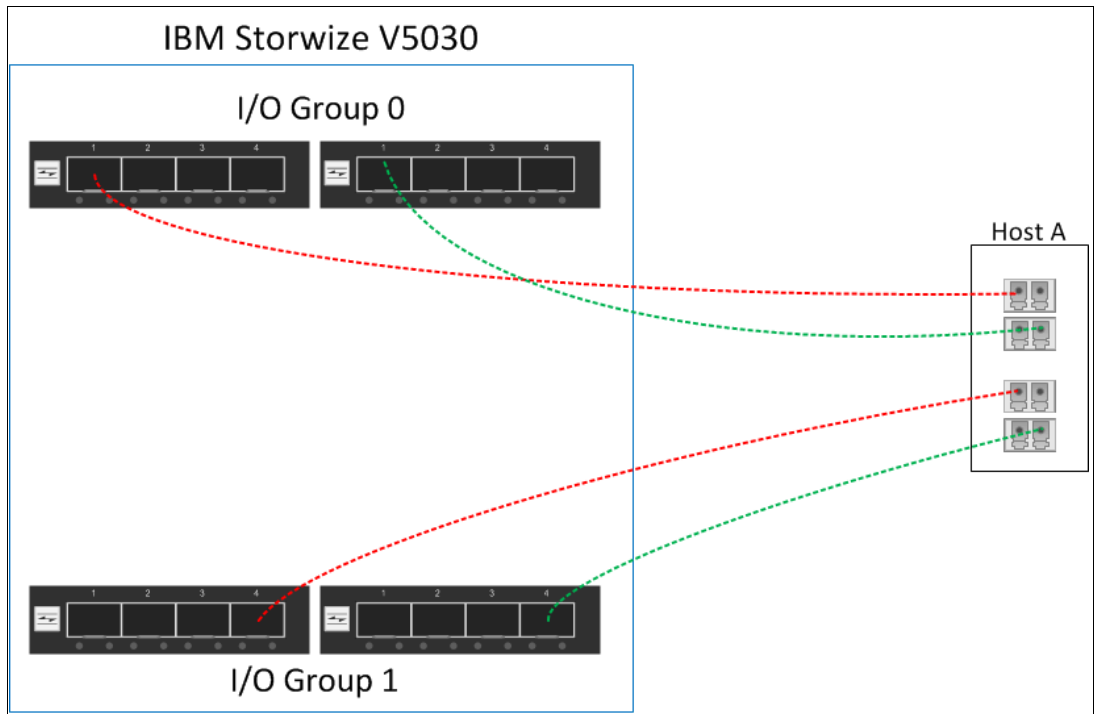


Figure 2-6 FC direct attach host configuration to I/O groups

For more information about the requirements for direct host attachment, see [this website](#).

2.4 SAS direct-attach planning

The Storwize V5000 Gen2 allows SAS host attachment by using an optional SAS card or at the Storwize V5020 with two onboard SAS ports. The SAS expansion ports cannot be used for host attachment because they are used for the connection to the Expansions only. Figure 2-7, Figure 2-8 on page 49, and Figure 2-9 on page 49 show the SAS ports connectivity for host attachments.

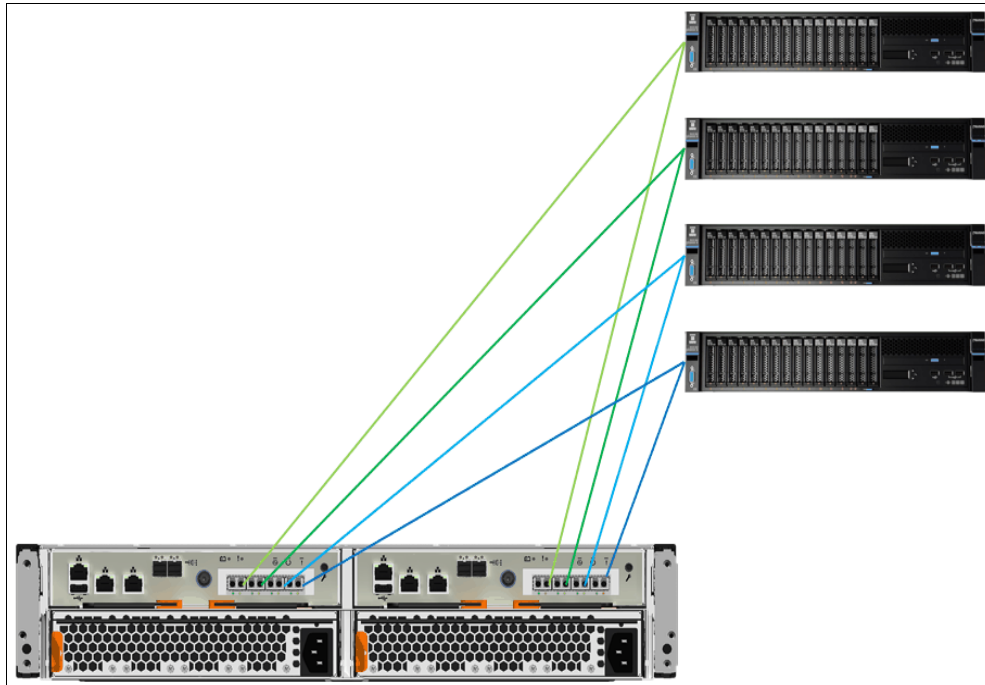


Figure 2-7 Storwize V5010 SAS host attachment



Figure 2-8 Storwize V5020 SAS host attachment

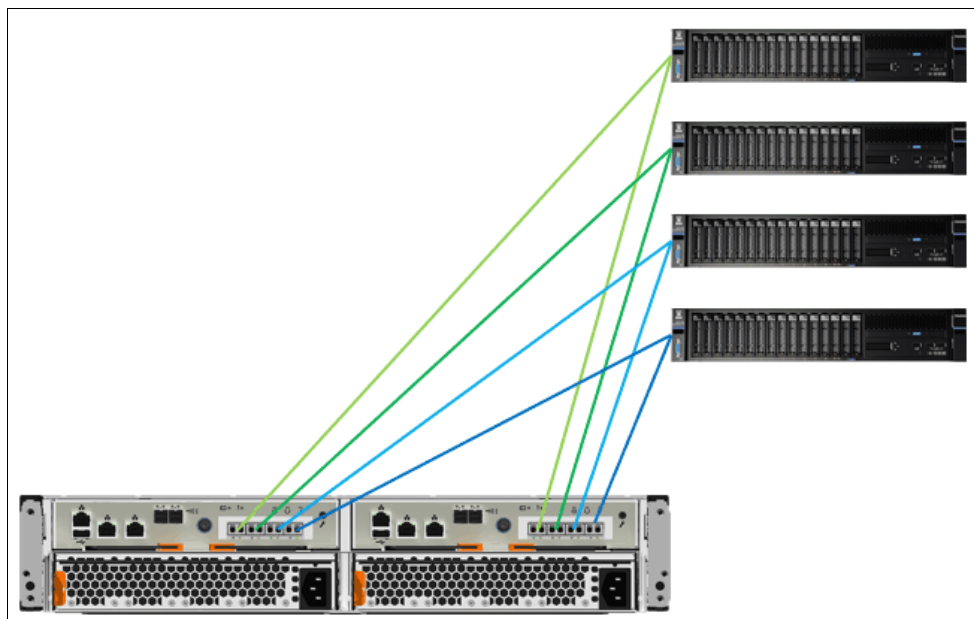


Figure 2-9 Storwize V5030 SAS host attachment

Inserting cables: You can insert the cables upside down, despite the keyway. Ensure that the blue tag on the SAS connector is underneath when you insert the cables.

We suggest that each SAS host is connected to both node canisters, because this approach provides redundancy in a path or canister failure.

2.5 LAN configuration planning

Two Ethernet ports per node canister are available for connection to the LAN. Use Ethernet port 1 to access the management graphical user interface (GUI), the service assistant GUI for the node canister, and iSCSI host attachment. Port 2 can be used for the management GUI and iSCSI host attachment.

Each node canister in a control enclosure connects over an Ethernet cable from Ethernet port 1 of the canister to an enabled port on your Ethernet switch or router. Optionally, you can attach an Ethernet cable from Ethernet port 2 to your Ethernet network.

Configuring IP addresses: Management IP addresses that are assigned to a system are different from iSCSI IP addresses and are used for different purposes. If iSCSI is used, iSCSI addresses are assigned to node ports. On the configuration node, a port has multiple IP addresses active at the same time. However, you cannot use the same IP address for management and iSCSI host use.

Table 2-1 lists possible IP configuration options of the Ethernet ports on the IBM Storwize V5000 Gen2 system.

Table 2-1 Storwize V5000 Gen2 IP address configuration options per node canister

Storwize V5000 Gen2 management node canister 1		Storwize V5000 Gen2 partner node canister 2	
IPv4/6 management address	Ethernet port 1	IPv4/6 service address	Ethernet port 1
IPv4/6 service address		IPv4/6 iSCSI address	
IPv4/6 iSCSI address			
IPv4/6 management address	Ethernet port 2	IPv4/6 iSCSI address	Ethernet port 2
IPv4/6 iSCSI address			

IP management addresses: The IP management address that is shown on node canister 1 in Table 2-1 on page 50 is an address on the configuration node. If a failover occurs, this address transfers to node canister 2, and this node canister becomes the new configuration node. The management addresses are managed by the configuration node canister only (1 or 2, and in this case, node canister 1).

Technician port: On the Storwize V5010 and V5020 models, port 2 serves as the technician port, which is used for system initialization and service. Do not connect port 2 to a network switch until the system initialization or service is complete. After the system initializes, the technician port is automatically disabled and port 2 can be used for Ethernet connectivity. However, when port 2 is used to perform system service, you must first run the `satask chserviceip -techport disable` command to disable the technician port. You can then use port 2 to provide additional Ethernet connectivity.

The Storwize V5030 model has a dedicated technician port. *Never* use the technician port to provide an Ethernet connection to the system. Do not connect the Ethernet technician port to a network switch. The technician port must be directly connected only to a personal computer when initializing a system or servicing a node.

2.5.1 Management IP address considerations

Because Ethernet port 1 from each node canister must connect to the LAN, a single management IP address for the clustered system is configured as part of the initial setup of the IBM Storwize V5000 Gen2 system.

The management IP address is associated with one of the node canisters in the clustered system and that node then becomes the configuration node. If this node goes offline (planned or unplanned), the management IP address fails over to the other node's Ethernet port 1.

For more clustered system management redundancy, you need to connect Ethernet port 2 on each of the node canisters to the LAN, which allows for a backup management IP address to be configured for access, if necessary.

Figure 2-10 shows a logical view of the Ethernet ports that are available for the configuration of the one or two management IP addresses. These IP addresses are for the clustered system and associated with only one node, which is then considered the configuration node.

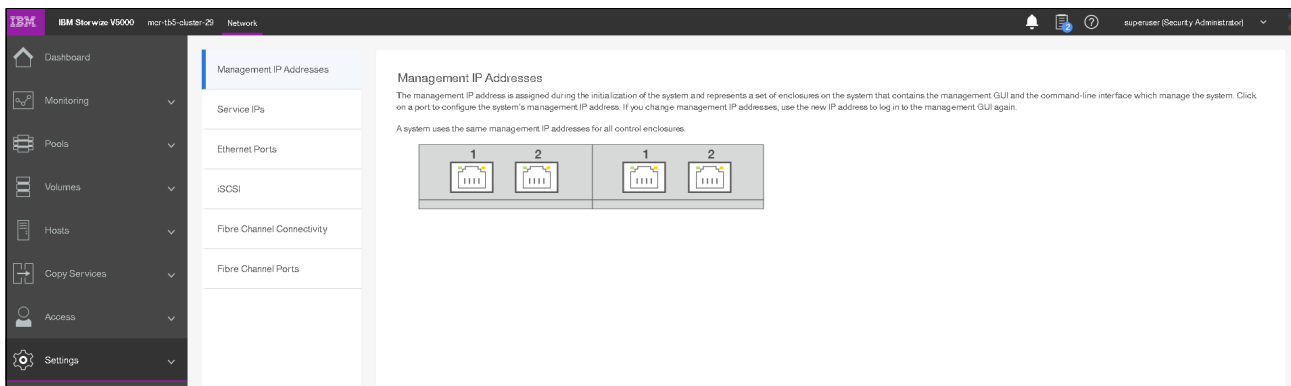


Figure 2-10 Ethernet ports that are available for configuration

2.5.2 Service IP address considerations

Ethernet port 1 on each node canister is used for system management and, when required, for service access. In normal operation, the service IP addresses are not needed. However, if a node canister problem occurs, it might be necessary for service personnel to log on to the node to perform service actions.

Figure 2-11 shows a logical view of the Ethernet ports that are available for the configuration of the service IP addresses. Only port one on each node can be configured with a service IP address.

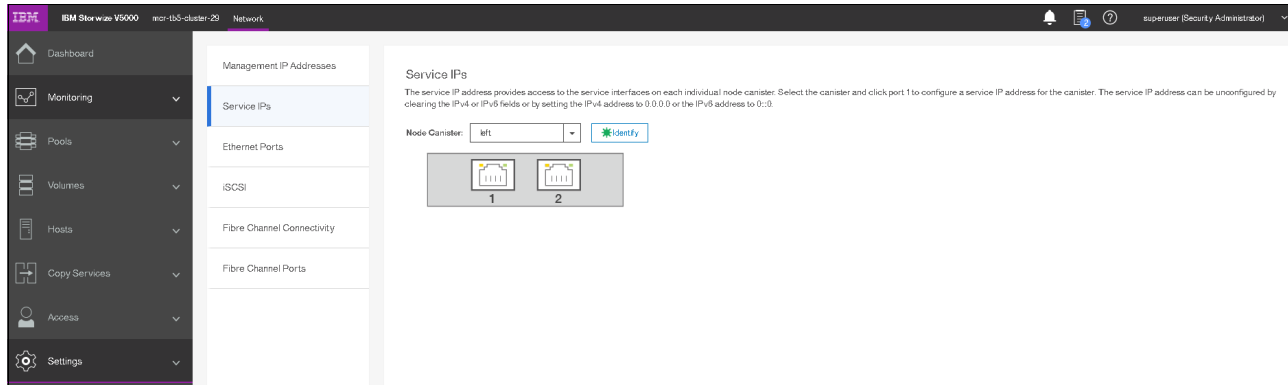


Figure 2-11 Service IP addresses that are available for configuration

2.6 Host configuration planning

In this section, we describe host configuration planning.

2.6.1 Fibre Channel connection

Configuring your SAN with at least two independent switches, or networks of switches, ensures a redundant fabric with no single point of failure. If one of the two SAN fabrics fails, the configuration is in a degraded mode, but is still valid. Maintain separate fabrics for FCoE and FC. If you attempt to combine these fabrics, you might risk adding paths to the volumes. Supported configurations allow a maximum of eight paths. A SAN with only one fabric is a valid configuration, but risks loss of access to data if the fabric fails. SANs with only one fabric are exposed to a single point of failure.

2.6.2 iSCSI configuration

iSCSI hosts connect to the system through the node-port IP addresses, which can be assigned to any Ethernet port of the node. If the node fails, the address becomes unavailable and the host loses communication with the system via that node. To allow hosts to maintain access to data, the node-port IP addresses for the failed node are transferred to the partner node in the I/O group. The partner node handles requests for its own node-port IP addresses and also for node-port IP addresses on the failed node. This process is known as *node-port IP failover*.

In addition to node-port IP addresses, the iSCSI name and iSCSI alias for the failed node are transferred to the partner node. After the failed node recovers, the node-port IP address and the iSCSI name and alias are returned to the original node.

Various operating systems are supported by the IBM Storwize V5000 Gen2. For more information about various supported configurations, see the following IBM System Storage Interoperation Center (SSIC) [web page](#).

For more information, see Chapter 5, “Host configuration” on page 217.

2.7 Miscellaneous configuration planning

During the initial setup of the IBM Storwize V5000 Gen2 system, the installation wizard asks for various information that needs to be available during the installation process. Several of these fields are mandatory to complete the initial configuration.

Collect the information in the following checklist *before* the initial setup is performed. Although the date and time can be manually entered, use a Network Time Protocol (NTP) service to keep the clock synchronized:

1. Document the LAN NTP server IP address that is used for the synchronization of devices.
2. To send alerts to storage administrators and to set up Call Home to IBM for service and support, you need the following information:
 - a. Name of the primary storage administrator for IBM to contact, if necessary.
 - b. Email address of the storage administrator for IBM to contact, if necessary.
 - c. Phone number of the storage administrator for IBM to contact, if necessary.
 - d. Physical location of the IBM Storwize V5000 Gen2 system for IBM service (for example, Building 22, first floor).
 - e. Simple Mail Transfer Protocol (SMTP) or email server address to direct alerts to and from the IBM Storwize V5000 Gen2.
 - f. For the Call Home service to work, the IBM Storwize V5000 Gen2 system must have access to an SMTP server on the LAN that can forward emails to the default IBM service address. 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.

To test connections to the support center, select **Settings** → **Support** → **Support Assistance**. On the Support Assistance page, select **Test Connection** to verify connectivity between the system and the support center. Enter the email address of local administrators that must be notified of alerts.

- g. IP address of Simple Network Management Protocol (SNMP) server to direct alerts to, if required (for example, operations or Help desk).

After the IBM Storwize V5000 Gen2 initial configuration, you might want to add users who can manage the system. You can create as many users as you need, but the following roles generally are configured for users:

- ▶ Security Admin
- ▶ Administrator
- ▶ CopyOperator
- ▶ Service
- ▶ Monitor

The user in the Security Admin role can perform any function on the IBM Storwize V5000 Gen2.

The user in the Administrator role can perform any function on the IBM Storwize V5000 Gen2 system, except manage users.

User creation: The Security Admin role is the only role that has the create users function. Limit this role to as few users as possible.

The user in the CopyOperator role can view anything in the system, but the user can configure and manage only the copy functions of the FlashCopy capabilities.

The user in the Monitor role can view object and system configuration information but cannot configure, manage, or modify any system resource.

The only other role that is available is the service role, which is used if you create a user ID for the IBM service support representative (SSR). With this user role, IBM service personnel can view anything on the system (as with the Monitor role) and perform service-related commands, such as adding a node back to the system after it is serviced or including disks that were excluded.

2.8 System management

The graphical user interface (GUI) is used to configure, manage, and troubleshoot the IBM Storwize V5000 Gen2 system. It is used primarily to configure Redundant Array of Independent Disks (RAID) arrays and logical drives, assign logical drives to hosts, replace and rebuild failed disk drives, and expand the logical drives.

It allows for troubleshooting and management tasks, such as checking the status of the storage server components, updating the firmware, and managing the storage server.

The GUI also offers advanced functions, such as FlashCopy, Volume Mirroring, Remote Mirroring, and Easy Tier. A command-line interface (CLI) for the IBM Storwize V5000 Gen2 system also is available.

This section describes system management by using the GUI and CLI.

2.8.1 Graphical user interface

A web browser is used for GUI access. You must use a supported web browser to access the management GUI. At the time of this writing, the Storwize V5000 Gen2 management GUI supports the following web browsers:

- ▶ Mozilla Firefox 59
- ▶ Mozilla Firefox Extended Support Release (ESR) 52
- ▶ Microsoft Internet Explorer (IE) 11 and Microsoft Edge 40
- ▶ Google Chrome 65

Supported web browsers: For more information about supported browsers and to check the latest supported levels, see [this website](#).

Complete the following steps to open the management GUI from any web browser:

1. Browse to one of the following locations:
 - `http(s)://host name of your cluster/`
 - `http(s)://cluster IP address of your cluster/`
(for example, `https://192.168.70.120.`)

2. Use the password that you created during system setup to authenticate with the superuser or any other accounts that you created. The default user name and password for the management GUI is shown:

- User name: **superuser**
- Password: **passw0rd**

Note: The 0 character in the password is the number zero, not the letter O.

For more information, see Chapter 3, “Graphical user interface overview” on page 85.

After you complete the initial configuration that is described in 2.10, “Initial configuration” on page 61, the IBM Storwize V5000 Gen2 System overview window opens, as shown in Figure 2-12.

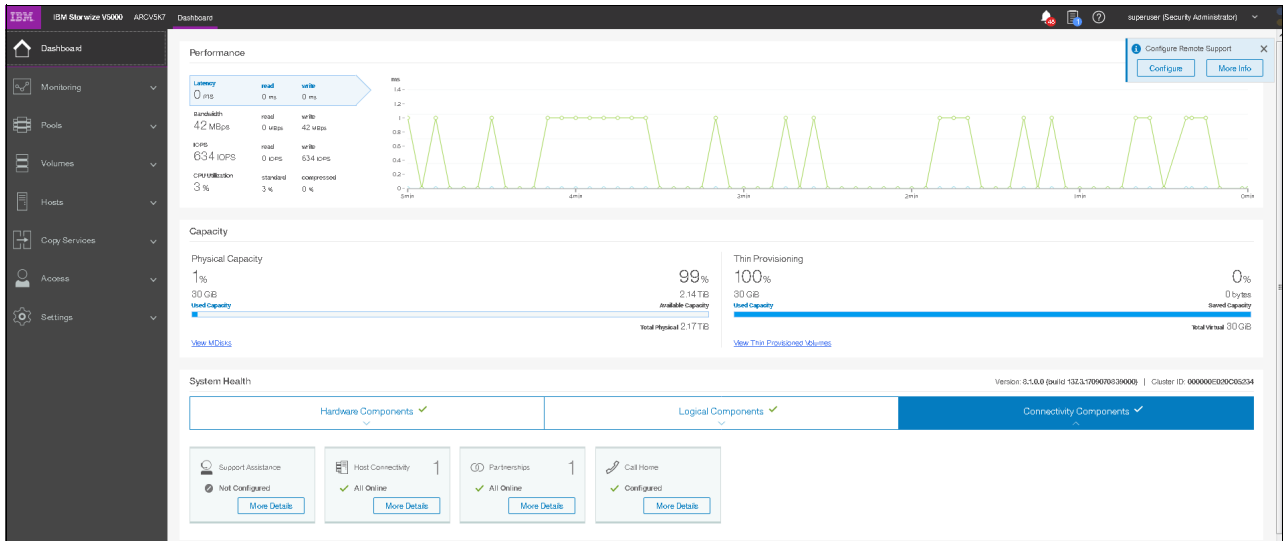


Figure 2-12 Setup wizard: Overview window

2.8.2 Command-line interface

The command-line interface (CLI) is a flexible tool for system management that uses the Secure Shell (SSH) protocol. A public/private SSH key pair is optional for SSH access. The storage system can be managed by using the CLI, as shown in Example 2-1.

Example 2-1 System management by using the command-line interface

```
IBM_Storwize:ITS0-V5000:superuser>svcinfolenclosureslot
enclosure_id slot_id port_1_status port_2_status drive_present drive_id
1 1 online online yes 10
1 2 online online yes 11
1 3 online online yes 15
1 4 online online yes 16
1 5 online online yes 12
1 6 online online yes 4
1 7 online online yes 7
1 8 online online yes 8
1 9 online online yes 9
1 10 online online yes 5
1 11 online online yes 18
1 12 online online yes 14
```

1	13	online	online	yes	13
1	14	online	online	yes	2
1	15	online	online	yes	6
1	16	online	online	yes	3
1	17	online	online	yes	1
1	18	online	online	yes	0
1	19	online	online	yes	20
1	20	online	online	no	
1	21	online	online	yes	19
1	22	online	online	yes	21
1	23	online	online	yes	22
1	24	online	online	yes	17
2	1	online	online	yes	25
2	2	online	online	yes	27
2	3	online	online	no	
2	4	online	online	yes	31
2	5	online	online	yes	24
2	6	online	online	yes	26
2	7	online	online	yes	33
2	8	online	online	yes	32
2	9	online	online	yes	23
2	10	online	online	yes	28
2	11	online	online	yes	29
2	12	online	online	yes	30

IBM_Storwize:ITS0-V5000:superuser>

You can set up the initial IBM Storwize V5000 Gen2 system by using the process and tools that are described in 2.9, “First-time setup” on page 56.

2.9 First-time setup

This section describes how to set up a first-time IBM Storwize V5000 Gen2 service and system.

Before you set up the initial IBM Storwize V5000 Gen2 system, ensure that the system is powered on.

Power on: For more information and to check the power status of the system, see [this website](#).

Complete the following steps to set up the IBM Storwize V5000 Gen2 system by using the technician Ethernet port:

1. Configure an Ethernet port on the personal computer to enable the Dynamic Host Configuration Protocol (DHCP) configuration of its IP address and Domain Name System (DNS) settings.

If you do not use DHCP, you must manually configure the personal computer. Specify the static IPv4 address 192.168.0.2, subnet mask 255.255.255.0, gateway 192.168.0.1, and DNS 192.168.0.1.

2. Locate the Ethernet port that is labeled T on the rear of the node canister.

On the Storwize V5010 and Storwize V5020 systems, the second Ethernet port is also used as the technician port, as shown in Figure 2-13 and Figure 2-14.



Figure 2-13 Storwize V5010 technician port

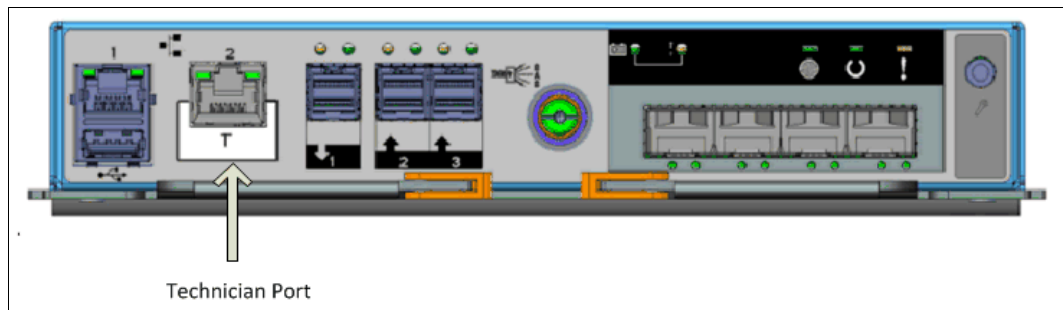


Figure 2-14 Storwize V5020 technician port

The Storwize V5030 systems use a dedicated technician port, which is shown in Figure 2-15.

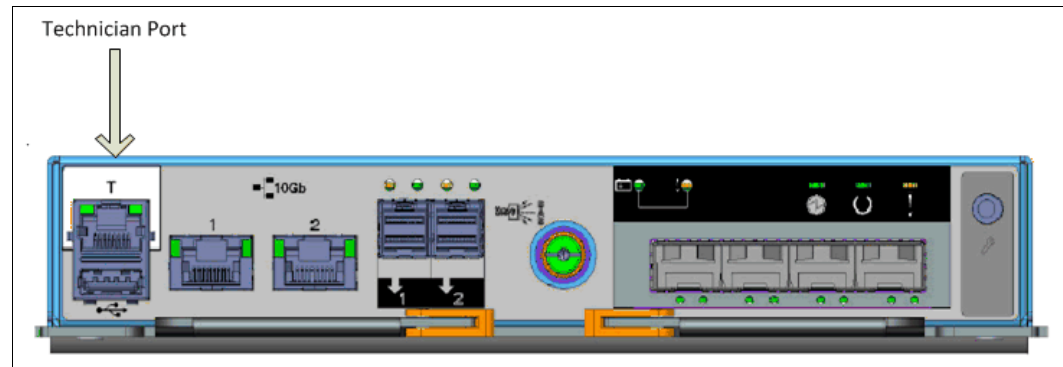


Figure 2-15 Storwize V5030 technician port

3. Connect an Ethernet cable between the port of the personal computer that is configured in step 2 and the technician port. After the connection is made, the system automatically configures the IP address and DNS settings for the personal computer if DHCP is available. If it is not available, the system uses the values that you provided in step 1.

4. After the Ethernet port of the personal computer connects, open a supported browser and browse to the address `http://install`. (If you do not have DHCP, open a supported browser and go to this static IP address: 192.168.0.1.) The browser is automatically directed to the initialization tool, as shown in Figure 2-16.

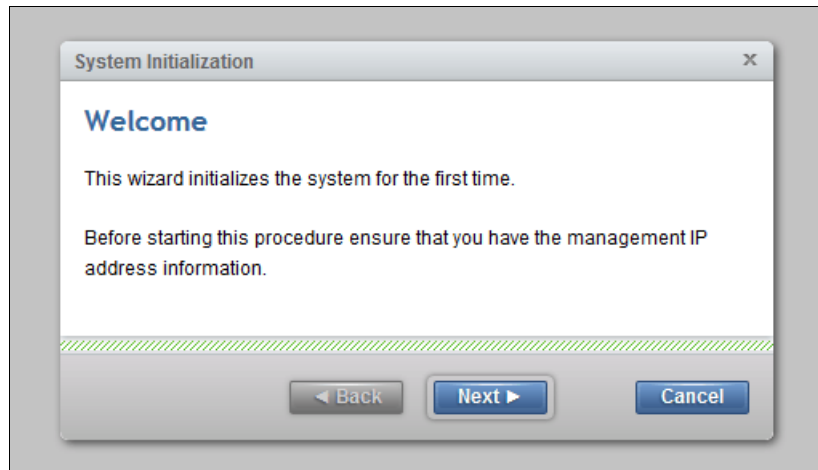


Figure 2-16 System initialization: Welcome

5. If you experience a problem when you try to connect due to a change in system states, wait 5 - 10 seconds and try again.
6. Click **Next**, as shown in Figure 2-17.

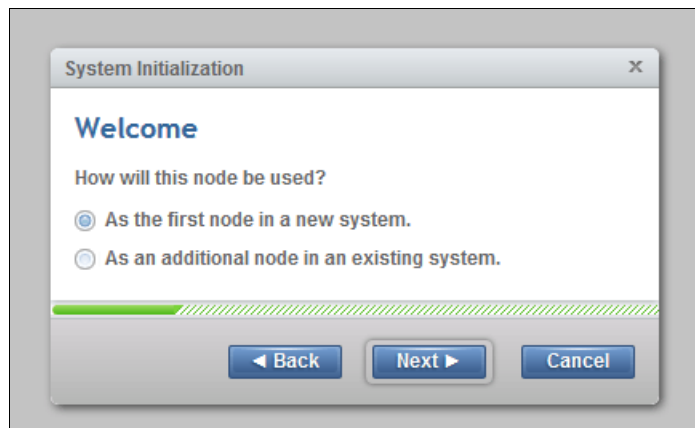


Figure 2-17 System initialization node usage

7. Choose the first option to set up the node as a new system and click **Next** to continue to the window that is shown in Figure 2-18.

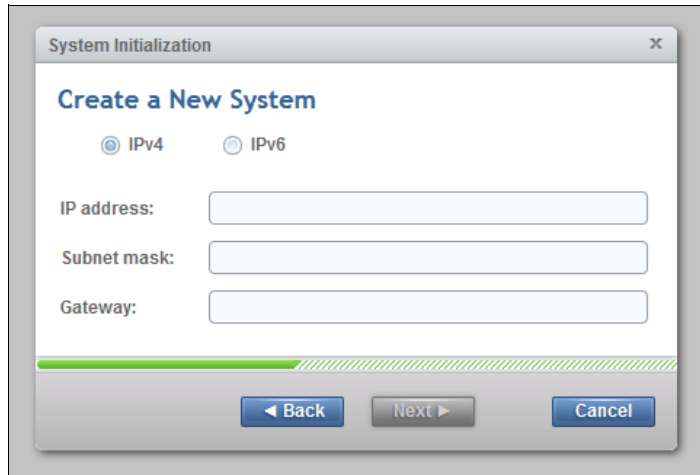


Figure 2-18 System initialization: Create a New System

8. Complete all of the fields with the networking details for managing the system and click **Next**. When the task completes, as shown in Figure 2-19, click **Close**.

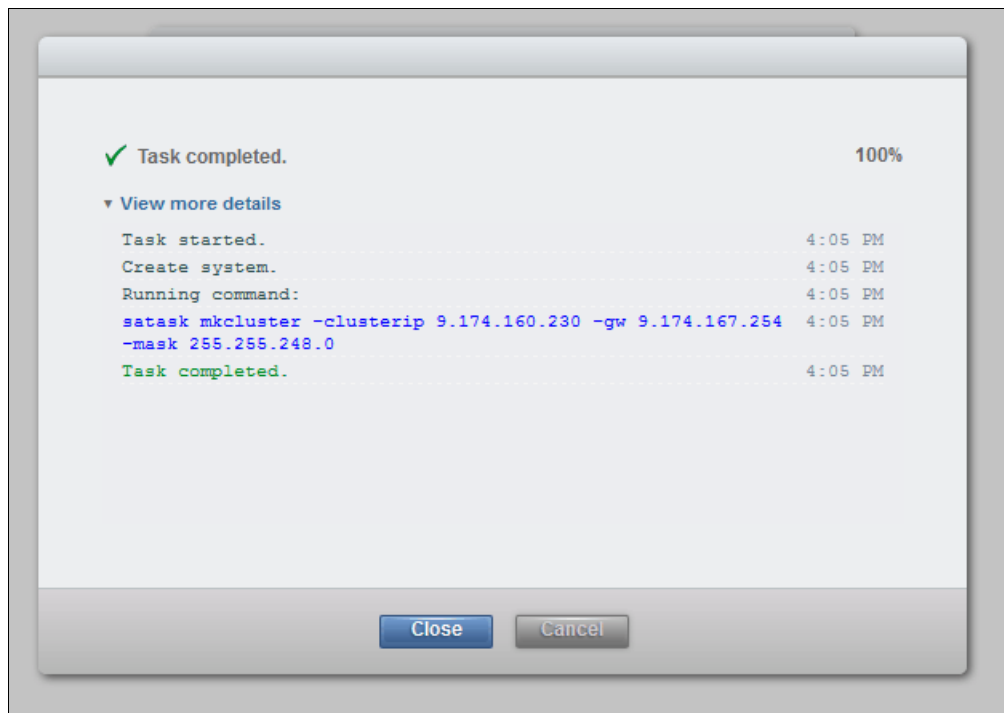


Figure 2-19 System initialization: Cluster creation

Note: The IBM Storwize V5000 Gen2 GUI shows the CLI as you go through the configuration steps.

9. The system takes approximately 10 minutes to reboot and reconfigure the Web Server, as shown in Figure 2-20. After this time, click **Next** to proceed to the final step.

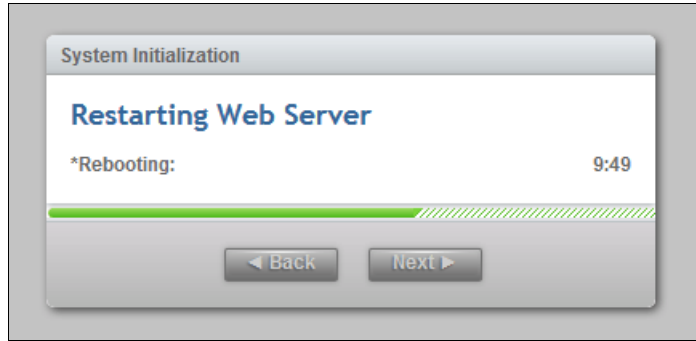


Figure 2-20 System Initialization: Restarting Web Server

10. After you complete the initialization process, disconnect the cable between the personal computer and the technician port, as shown in Figure 2-21. Reestablish the connection to the customer network and click **Next** to be redirected to the management address that you provided to configure the system initially.

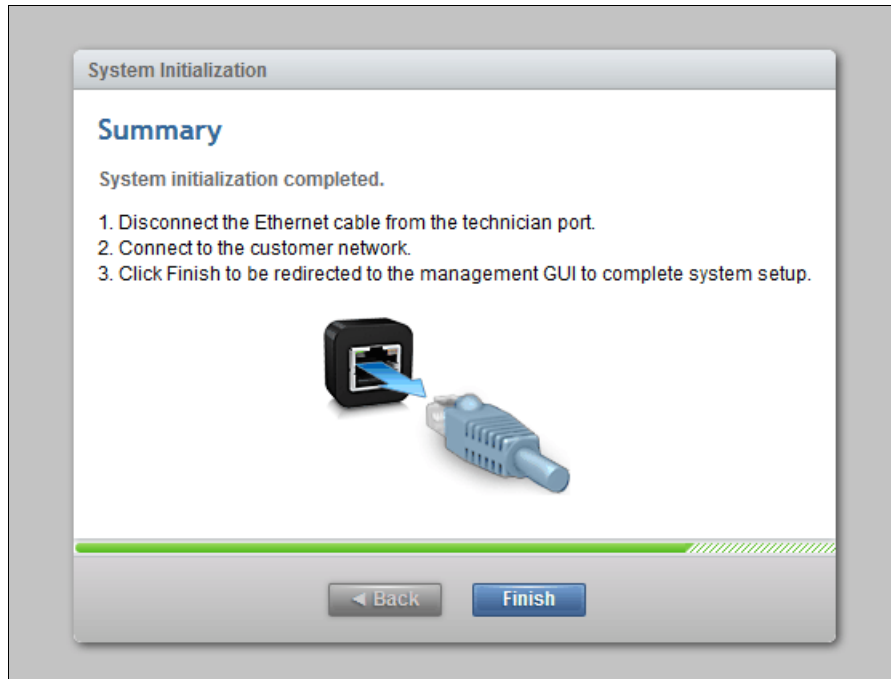


Figure 2-21 System initialization: Completion summary

2.10 Initial configuration

This section describes how to complete the initial configuration, including the following tasks:

- ▶ System components verification
- ▶ Email event notifications
- ▶ System name, date, and time settings
- ▶ License functions
- ▶ Initial storage configuration
- ▶ Initial configuration summary

If you completed the initial setup, that wizard automatically redirects you to the IBM Storwize V5000 Gen2 GUI. Otherwise, complete the following steps to complete the initial configuration process:

1. Start the service configuration wizard by using a web browser on a workstation and point it to the system management IP address that was defined in Figure 2-18 on page 59.
2. Enter a new secure password twice for the superuser user (see Figure 2-22). Click **Log in**.

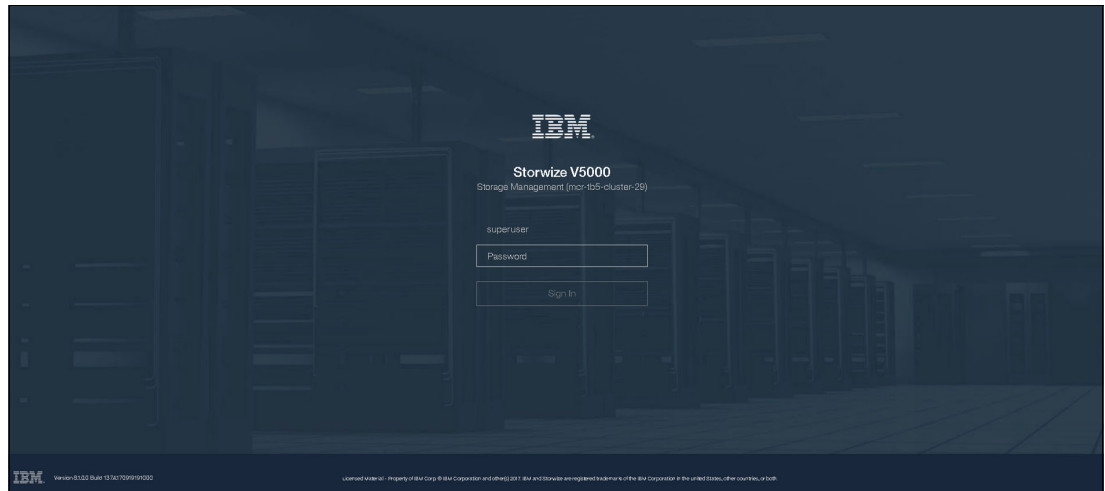


Figure 2-22 Setup wizard: Password prompt

3. Verify the prerequisites in the Welcome window, as shown in Figure 2-23. Click **Next**.

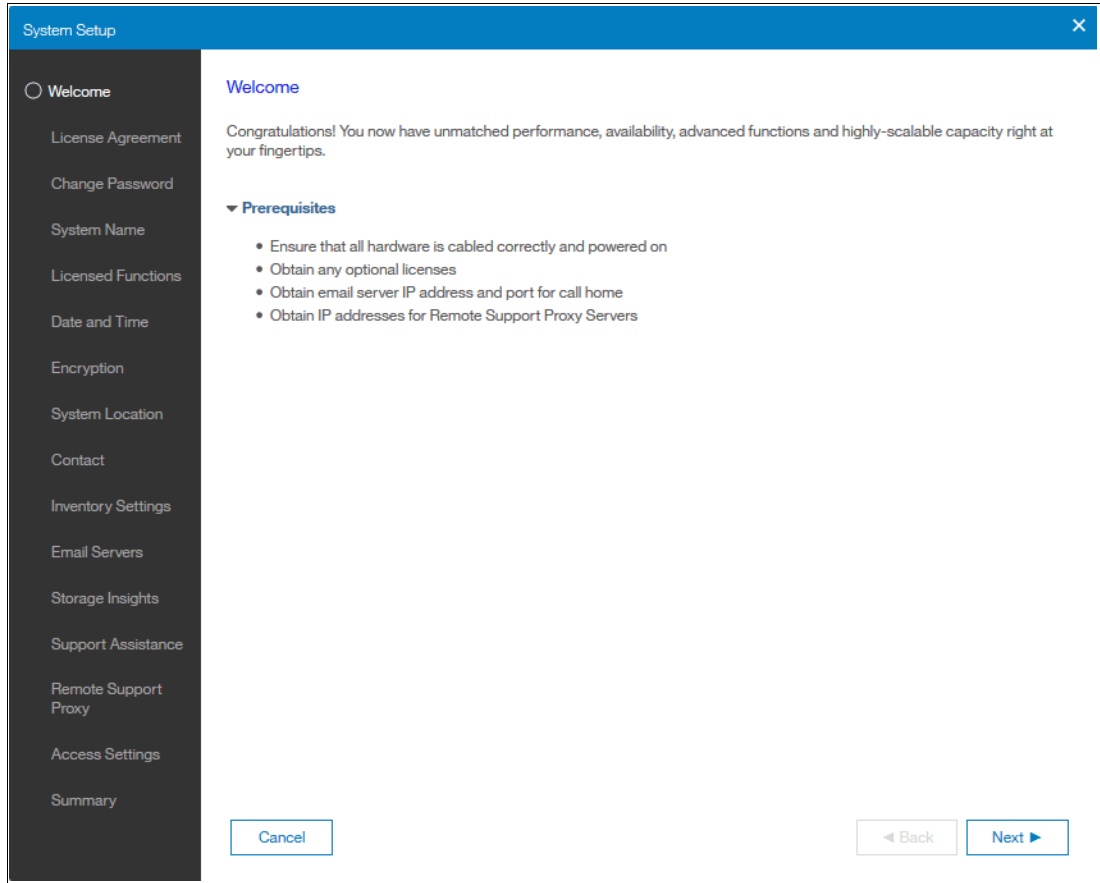


Figure 2-23 Setup wizard: Welcome

4. Accept the license agreement after reading it carefully, as shown in Figure 2-24. Click **Next**.

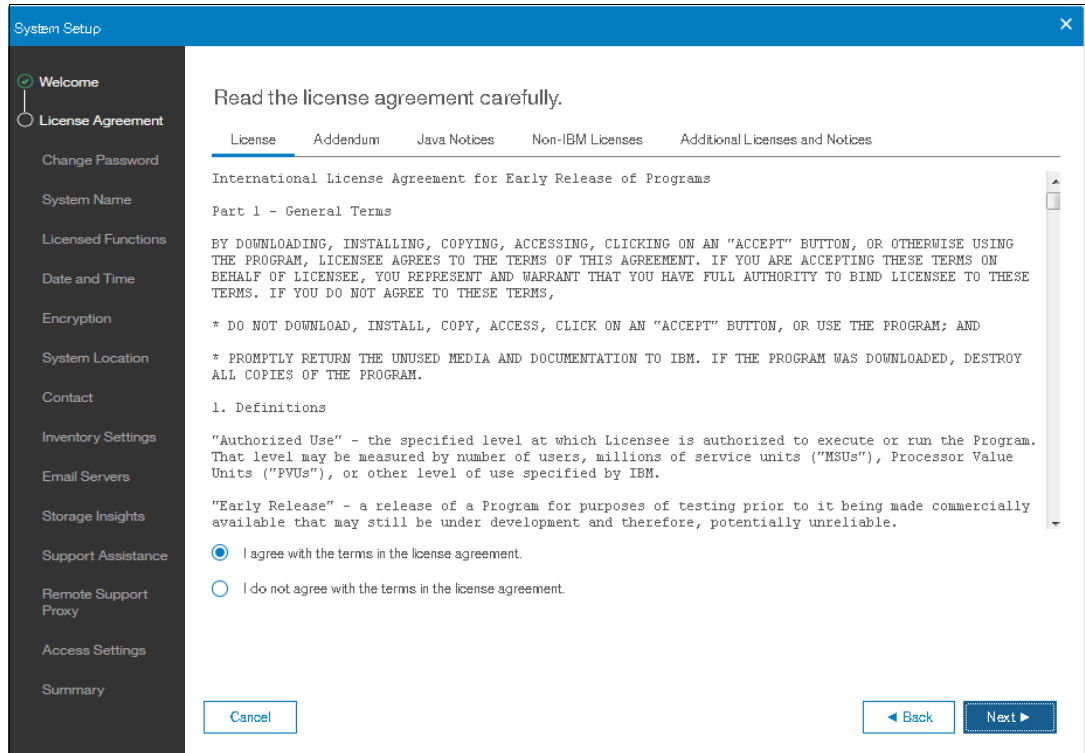


Figure 2-24 Setup wizard: License agreement

5. Change the password for superuser from the default, as shown in Figure 2-25. Then, click **Apply and Next**.

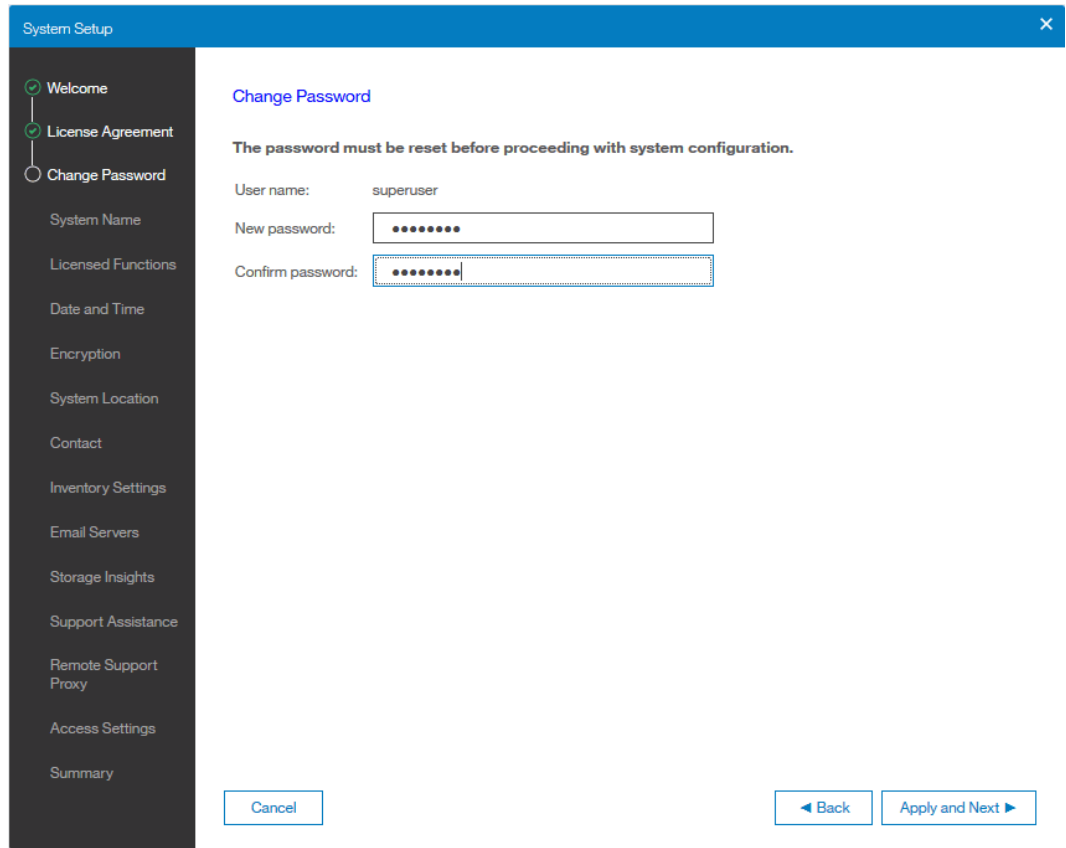


Figure 2-25 Setup wizard: Change password

You see message: “The password was successfully changed”, as shown in Figure 2-26.

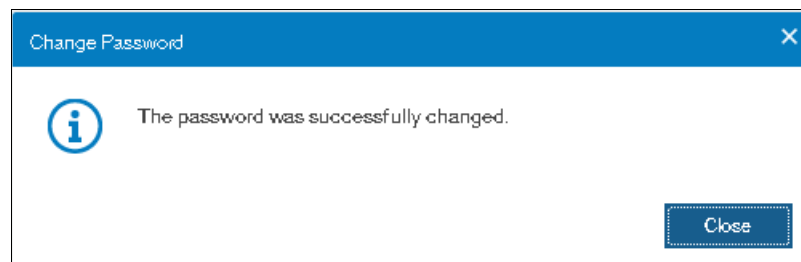


Figure 2-26 Setup wizard: Password changed

6. In the System Name window, enter the system name and click **Apply and Next**, as shown in Figure 2-27.

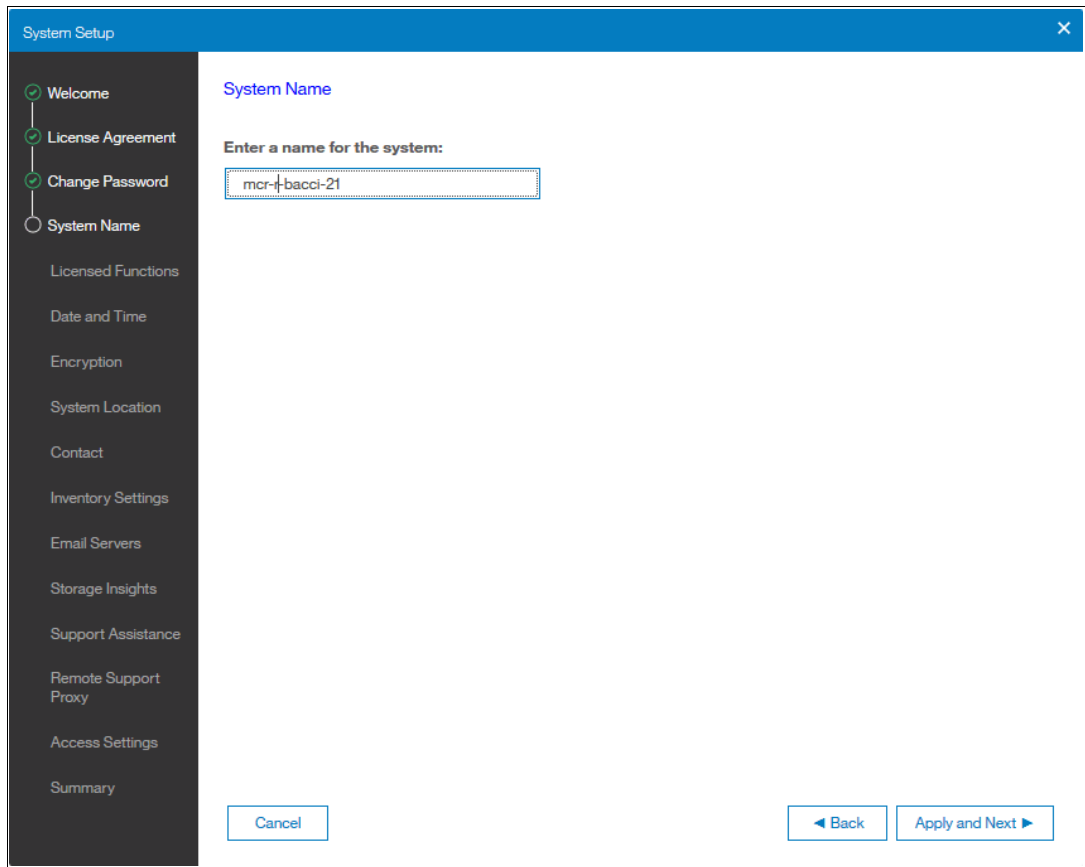


Figure 2-27 Setup wizard: System Name

Note: Use the `chsystem` command to modify the attributes of the clustered system. This command can be used any time after a system is created.

7. In the next window, the IBM Storwize V5000 Gen2 GUI provides help and guidance about additional licenses that are required for certain system functions. A license must be purchased for each enclosure that is attached to, or externally managed by, the IBM Storwize V5000 Gen2. For each of the functions, enter the number of enclosures, as shown in Figure 2-28. Then, click **Apply and Next**.

The screenshot shows the 'System Setup' window with a sidebar on the left containing a list of setup steps: Welcome, License Agreement, Change Password, System Name, Licensed Functions (selected), Date and Time, Encryption, System Location, Contact, Inventory Settings, Email Servers, Storage Insights, Support Assistance, Remote Support Proxy, Access Settings, and Summary. The main area is titled 'Licensed Functions' and contains a sub-header 'Licensed Functions' and a paragraph: 'Additional licenses are required to use certain system functions. For auditing purposes, retain the license agreement for proof of compliance.' Below this are five rows of input fields, each labeled with a function name and 'Number of enclosures': External Virtualization, FlashCopy, Remote Mirroring, Easy Tier, and Compression. Each field contains the number '10'. At the bottom of the window, there are three buttons: 'Need Help' (with a question mark icon), 'Cancel', and 'Apply and Next' (with a right-pointing arrow). The 'Apply and Next' button is highlighted in blue.

Figure 2-28 Setup wizard: Licensed Functions

The following actions are required for each of the licensed functions:

- FlashCopy: Enter the number of enclosures that are licensed to use FlashCopy function.
- Remote copy: Enter the number of Remote Mirroring licenses. This license setting enables the use of Metro Mirror and Global Mirror functions. This value must be equal to the number of enclosures that are licensed for external virtualization, plus the number of attached internal enclosures.
- Easy Tier: Enter the number of enclosures that are licensed to use Easy Tier function.
- External Virtualization: Enter the number of external enclosures that you are virtualizing. An external virtualization license is required for each physical enclosure that is attached to your system.
- Real-time Compression (RtC): Enter the number of enclosures that are licensed to use RtC.

Encryption license: The encryption feature that is available on the Storwize V5020 and V5030 systems uses a special licensing system that differs from the licensing system for the other features. Encryption requires a license key that can be activated in step 10.

- Two options are available for configuring the date and time. Select your preferred method and enter the date and time manually or specify a network address for a Network Time Protocol (NTP) server. After this selection, the Apply and Next option becomes active, as shown in Figure 2-29. Click **Apply and Next**.

The screenshot shows the 'System Setup' wizard window. The left sidebar contains a list of steps: Welcome, License Agreement, Change Password, System Name, Licensed Functions, Date and Time (selected), Encryption, System Location, Contact, Inventory Settings, Email Servers, Storage Insights, Support Assistance, Remote Support Proxy, Access Settings, and Summary. The main content area is titled 'Date and Time' and includes the instruction: 'Select time and date settings. You can enter these settings manually or specify a Network Time Protocol (NTP) server to synchronize time on the system.' There are two radio buttons: 'Manually' (unselected) and 'NTP Server' (selected). Below the radio buttons, there is an 'IP address:' field with the value '132.163.97.1' and a 'Time Zone:' dropdown menu showing '(GMT) Dublin, Edinburgh, London, Lisbon'. At the bottom, there are three buttons: 'Cancel', 'Back', and 'Apply and Next'.

Figure 2-29 Setup wizard: Date and Time

- If you purchased an Encryption License for a Storwize V5020 or Storwize V5030 system, select **Yes**, as shown in Figure 2-30. One license is required for each control enclosure. Therefore, in a Storwize V5030 configuration with two I/O groups, two license keys are required.

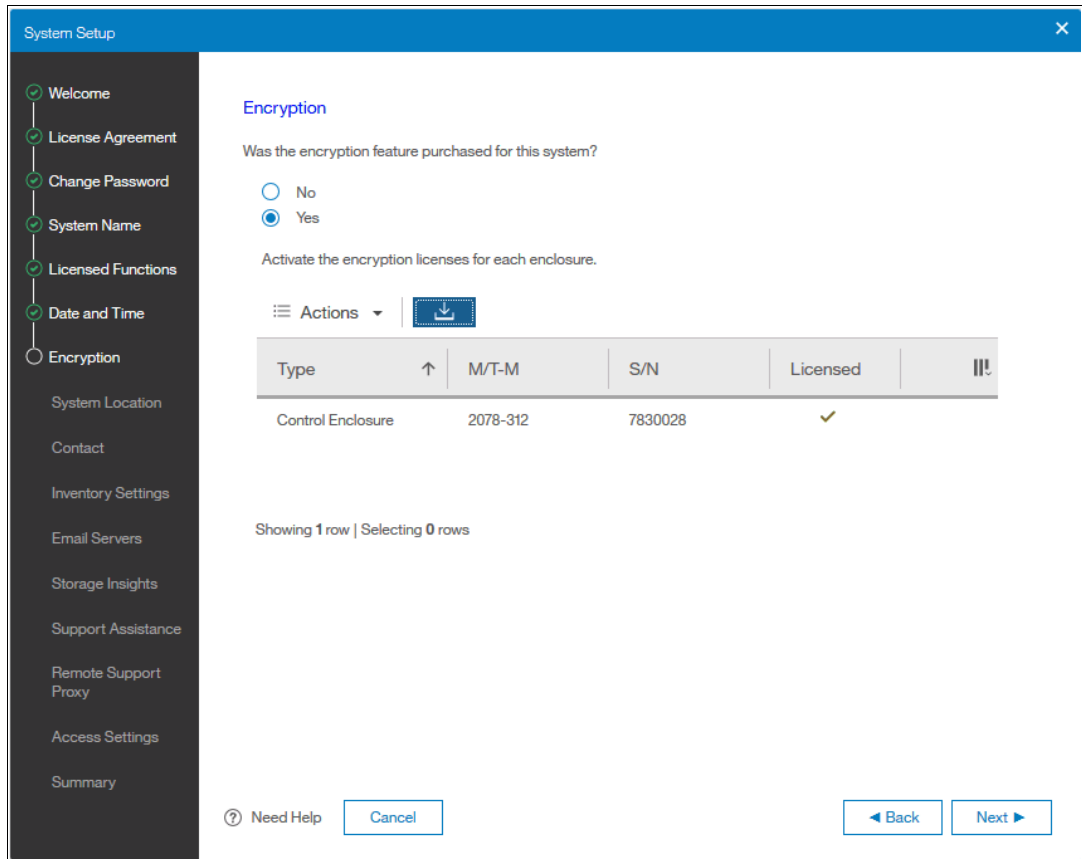


Figure 2-30 Setup wizard: Encryption feature

- The easiest way to activate the encryption license is to highlight each enclosure that you want to activate the license for and choose **Actions** → **Activate License Automatically** and enter the authorization code that came with the purchase agreement for encryption. This action retrieves and applies a license key from `ibm.com`, as shown in Figure 2-31.

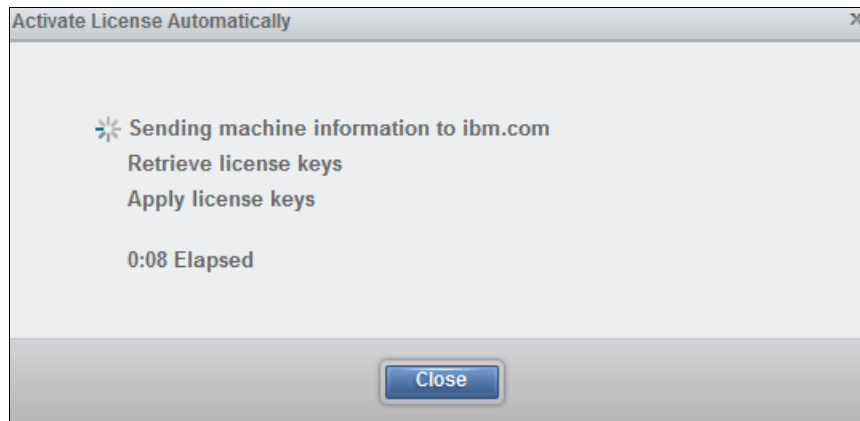
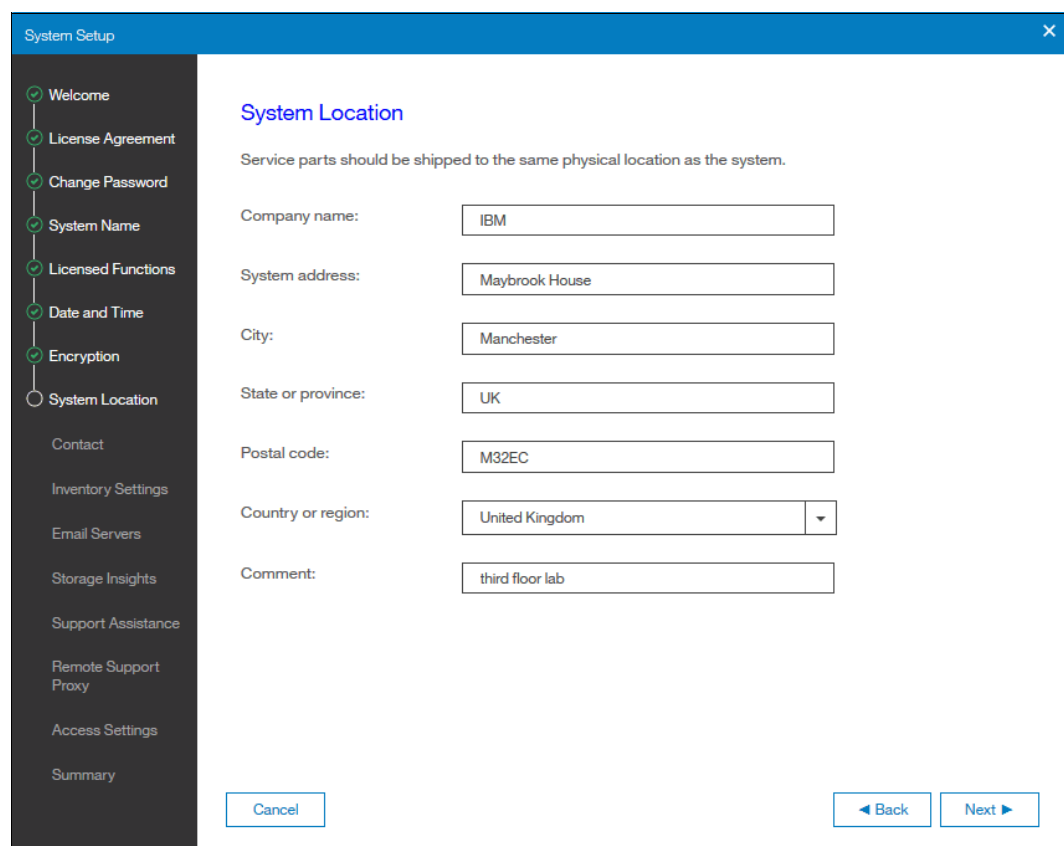


Figure 2-31 Setup wizard: Encryption license activation

11. If automatic activation cannot be performed (for example, if the Storwize V5000 Gen2 system is behind a firewall that prevents it from accessing the internet), choose **Actions** → **Activate License Manually**. Complete the following steps:
- Go to [this website](#).
 - Select **Storwize**. Enter the machine type (2077 or 2078), serial number, and machine signature of the system. You can obtain this information by clicking **Need Help**.
 - Enter the authorization codes that were sent with your purchase agreement for the encryption function.
 - Copy or download the key and paste it into the management GUI to activate the license.
12. When all licenses are active, click **Next** to set up the system location, as shown in Figure 2-32.



The screenshot shows a 'System Setup' window with a sidebar on the left and a main content area on the right. The sidebar contains a list of steps: Welcome, License Agreement, Change Password, System Name, Licensed Functions, Date and Time, Encryption, System Location (selected), Contact, Inventory Settings, Email Servers, Storage Insights, Support Assistance, Remote Support Proxy, Access Settings, and Summary. The main content area is titled 'System Location' and includes a note: 'Service parts should be shipped to the same physical location as the system.' Below this are several input fields: 'Company name' (IBM), 'System address' (Maybrook House), 'City' (Manchester), 'State or province' (UK), 'Postal code' (M32EC), 'Country or region' (United Kingdom), and 'Comment' (third floor lab). At the bottom, there are 'Cancel', 'Back', and 'Next' buttons.

Figure 2-32 Setup wizard: system location

13. After entering the system location, click **Next** to set up the contact person for the system, as shown in Figure 2-33. Then, click **Apply and Next**.

The screenshot shows a 'System Setup' window with a sidebar on the left and a main content area on the right. The sidebar contains a list of steps: Welcome, License Agreement, Change Password, System Name, Licensed Functions, Date and Time, Encryption, System Location, Contact, Inventory Settings, Email Servers, Storage Insights, Support Assistance, Remote Support Proxy, Access Settings, and Summary. The 'Contact' step is currently selected. The main content area is titled 'Contact' and includes the text: 'The support center contacts this person to resolve issues on the system.' Below this text are four input fields: 'Name' (containing 'R.Bacci'), 'Email' (containing 'bacci@noway.com'), 'Phone (primary)' (containing '333-444-555'), and 'Phone (alternate)' (which is empty). At the bottom of the window, there are three buttons: 'Cancel', 'Back', and 'Apply and Next'.

Figure 2-33 Setup wizard: contact person

14. If you plan to enable Call Home, you also can add Inventory and Configuration details. To add this information, you must enable one or both functions, as shown in Figure 2-34.

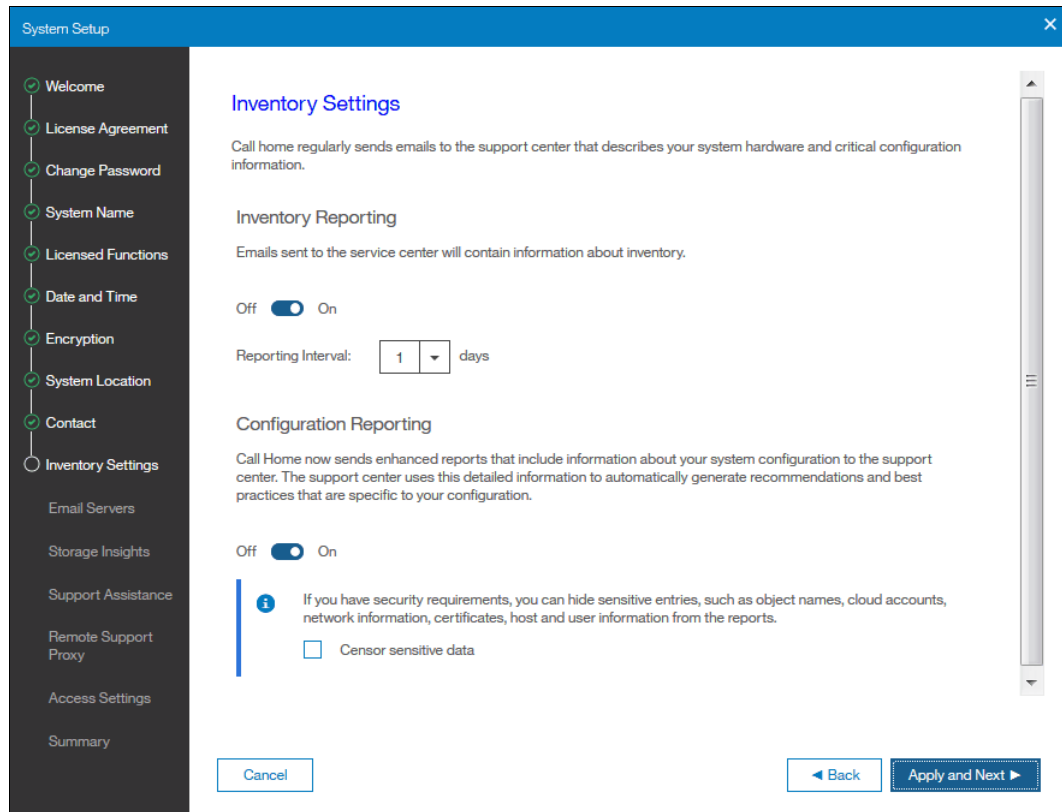


Figure 2-34 Setup wizard: Inventory Settings

15. Click **Apply and Next**.

16. You can configure your system to send email reports to IBM if an issue is detected that requires hardware replacement. This function is called *Call Home*. When this email is received, IBM automatically opens a problem report and contacts you to verify whether replacements parts are required.

Call Home: When Call Home is configured, the IBM Storwize V5000 Gen2 automatically creates a Support Contact with one of the following email addresses, depending on the country or region of installation:

- ▶ US, Canada, Latin America, and Caribbean Islands: callhome1@de.ibm.com
- ▶ All other countries or regions: callhome0@de.ibm.com

The IBM Storwize V5000 Gen2 can use Simple Network Management Protocol (SNMP) traps, syslog messages, and Call Home email to notify you and the IBM Support Center when significant events are detected. Any combination of these notification methods can be used simultaneously.

To set up Call Home, you need the location details of the IBM Storwize V5000 Gen2, Storage Administrator details, and at least one valid SMTP server IP address, as shown in Figure 2-35.

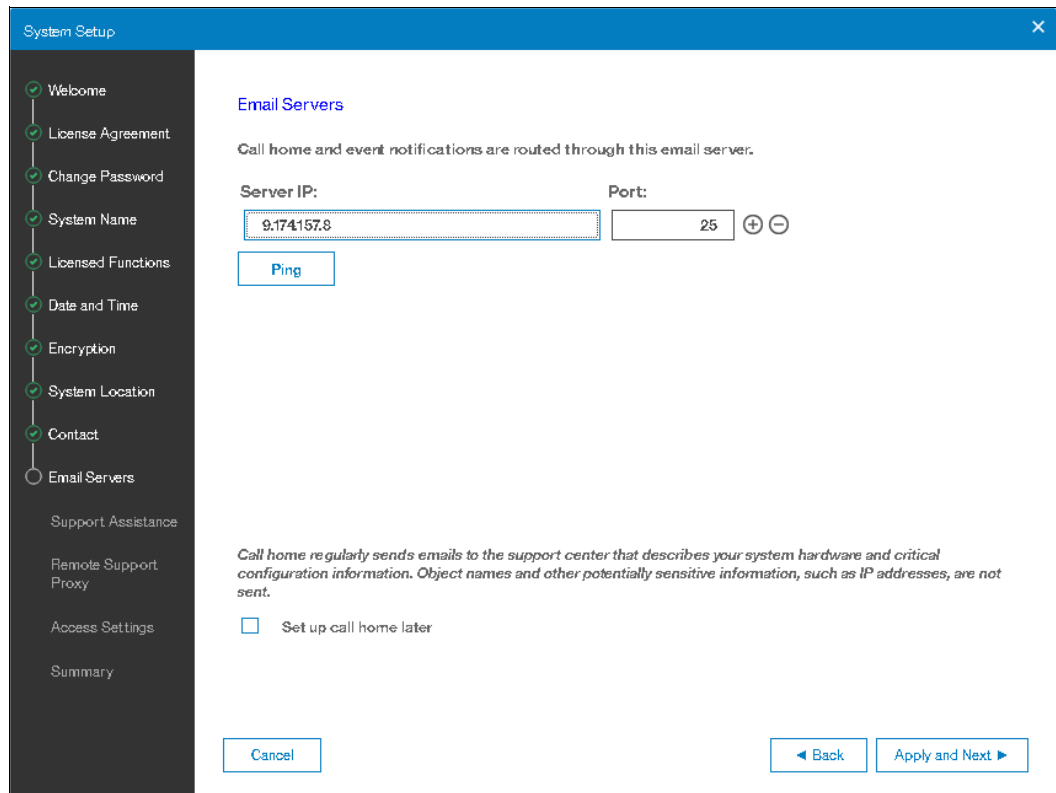


Figure 2-35 Setup wizard: Email server details

Note: If you do not want to configure Call Home now, you can defer it by selecting the option: **Set up call home later** in the GUI and return to it later by clicking **Settings** → **Notifications**.

If your system is under warranty or you purchased a hardware maintenance agreement, we advise you to configure Call Home to enable proactive support of the IBM Storwize V5000 Gen2.

To enter more than one email server, click the plus sign (+) icon, as shown in Figure 2-35. Then, click **Apply and Next** to commit.

17. A new offering is IBM Storage Insights. IBM can gather log packages remotely and provide customers with a unified dashboard that shows the health, capacity, and performance of their IBM block storage systems. If you want to use this free offer, enter your information, as shown in Figure 2-36.

System Setup

Welcome
License Agreement
Change Password
System Name
Licensed Functions
Date and Time
Encryption
System Location
Contact
Inventory Settings
Email Servers
Storage Insights
Support Assistance
Remote Support Proxy
Access Settings
Summary

You're eligible for a new offering called IBM Storage Insights. With Storage Insights, IBM can gather log packages remotely and provide customers with a unified dashboard that shows the health, capacity, and performance of their IBM block storage systems. **It's easy to get started, and it's FREE, so why wait?**

To get started, enter your IBM ID:

IBM ID:

[Don't have an IBM ID? Sign up here.](#)

The following fields were prefilled with the contact information from Call Home. Verify that the contact information can be used for Storage Insights:

First Name:

Last Name:

Company:

Email:

i Why should I use Storage Insights?

Let's face it. Storage performance can be tough to maintain and troubleshoot. Costs skyrocket for every minute you can't access data. Storage Insights monitors performance for easy collaboration with consultants and experts to resolve issues faster. Best of all, it's free and you will get all the credit. Just register your system to start.

[Storage Insights Fact Sheet](#)

I'm not interested in Storage Insights.

[Cancel](#) [Back](#) [Next](#)

Figure 2-36 Setup wizard IBM Storage Insights®

18. The next window is used for setting up support assistance, if wanted, as shown in Figure 2-37.

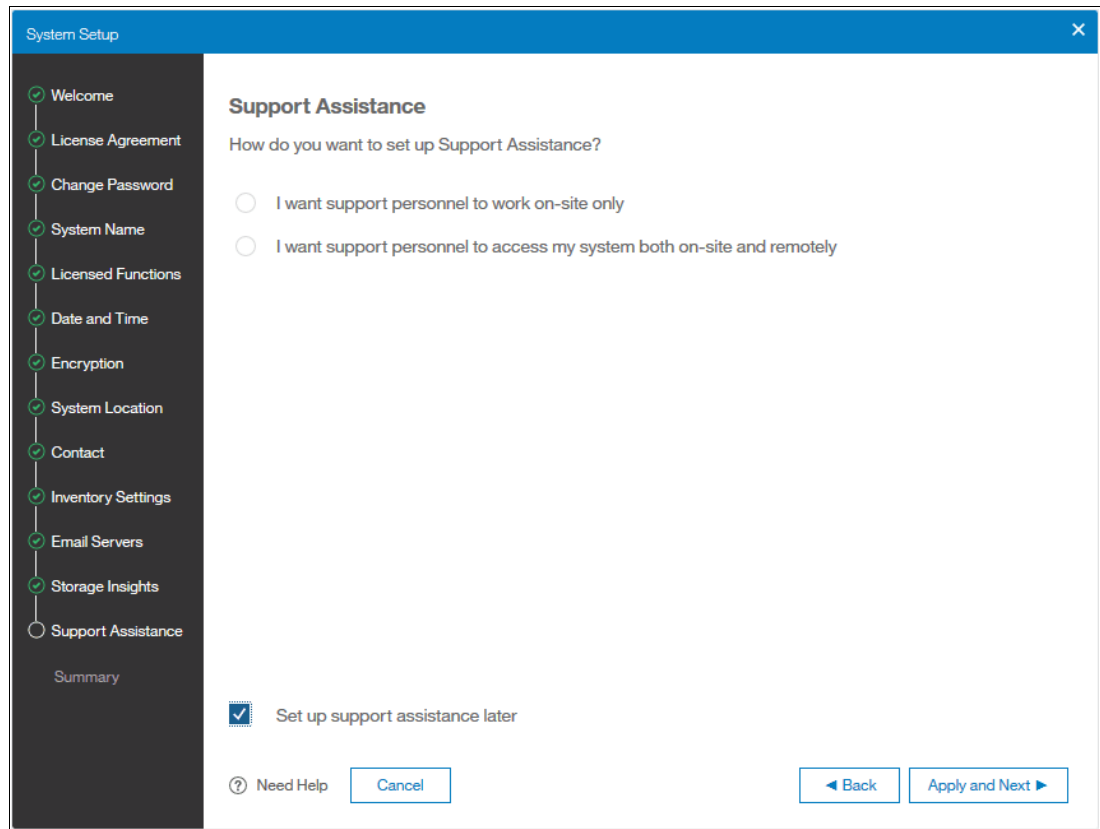


Figure 2-37 Initial setup: Support Assistance

In our setup, we chose to set up the support assistance later because it is described extensively in Chapter 12, “RAS, monitoring, and troubleshooting” on page 623.

19. After clicking **Apply and Next**, the Summary window for the contact details, system location, email server, Call Home, and email notification options opens, as shown in Figure 2-38.

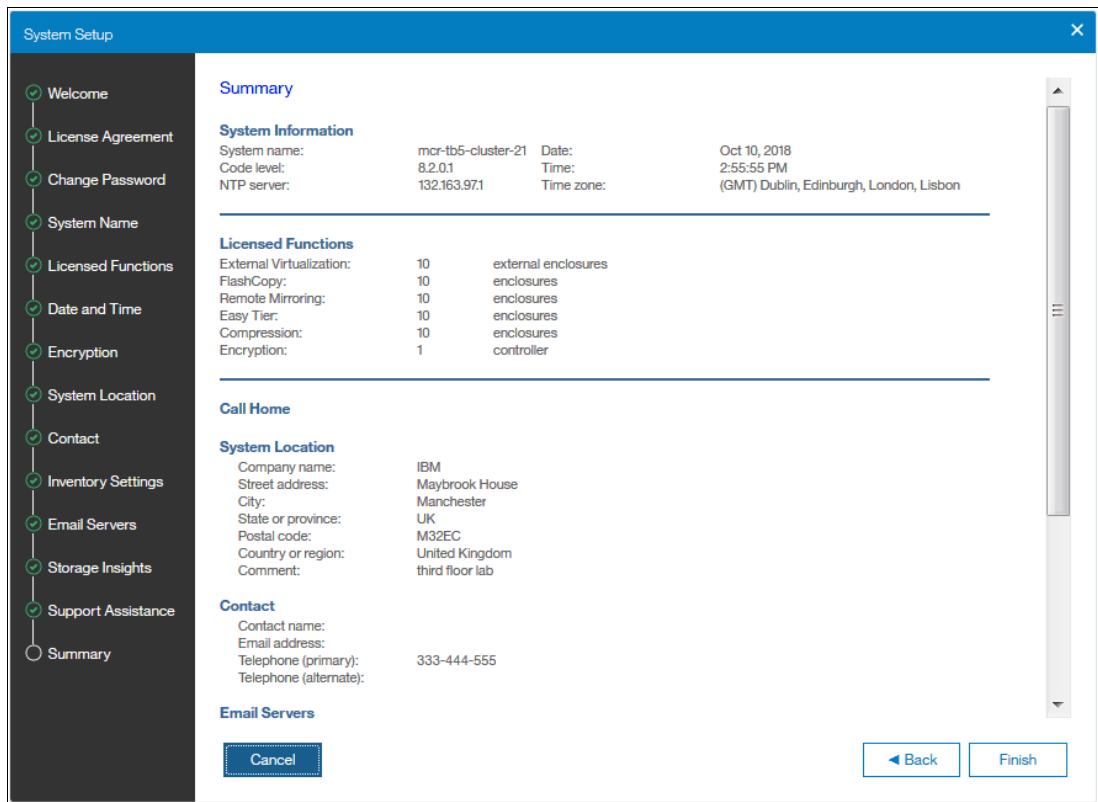


Figure 2-38 Setup wizard: Summary

20. Click **Finish**, and the web browser is redirected to the landing page of management GUI as shown in Figure 2-39.

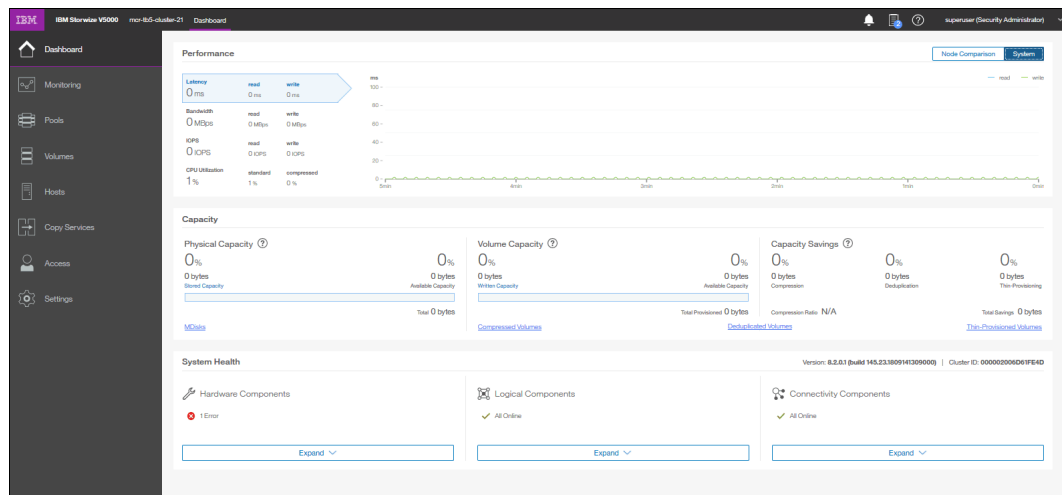


Figure 2-39 Landing page of management GUI

2.10.1 Adding enclosures after the initial configuration

When the initial installation of the IBM Storwize V5000 Gen2 is complete, all expansion enclosures and control enclosures that were purchased at that time must be installed as part of the initial configuration. This process enables the system to make the best use of the enclosures and drives that are available.

Adding a control enclosure

If you are expanding the IBM Storwize V5000 Gen2 after the initial installation by adding a second I/O group (a second control enclosure), you must install it in the rack and connect it to the SAN. Ensure that you rezone your Fibre Channel switches so that the new control enclosure and the existing control enclosure are connected. For more information about zoning the node canisters, see 2.2, “SAN configuration planning” on page 44.

Note: Adding a second I/O group (over the second controller enclosure) is supported on the IBM Storwize V5000 Gen2 model V5030 only.

After the hardware is installed, cabled, zoned, and powered on, a second control enclosure is visible from the IBM Storwize V5000 Gen2 GUI, as shown in Figure 2-40.

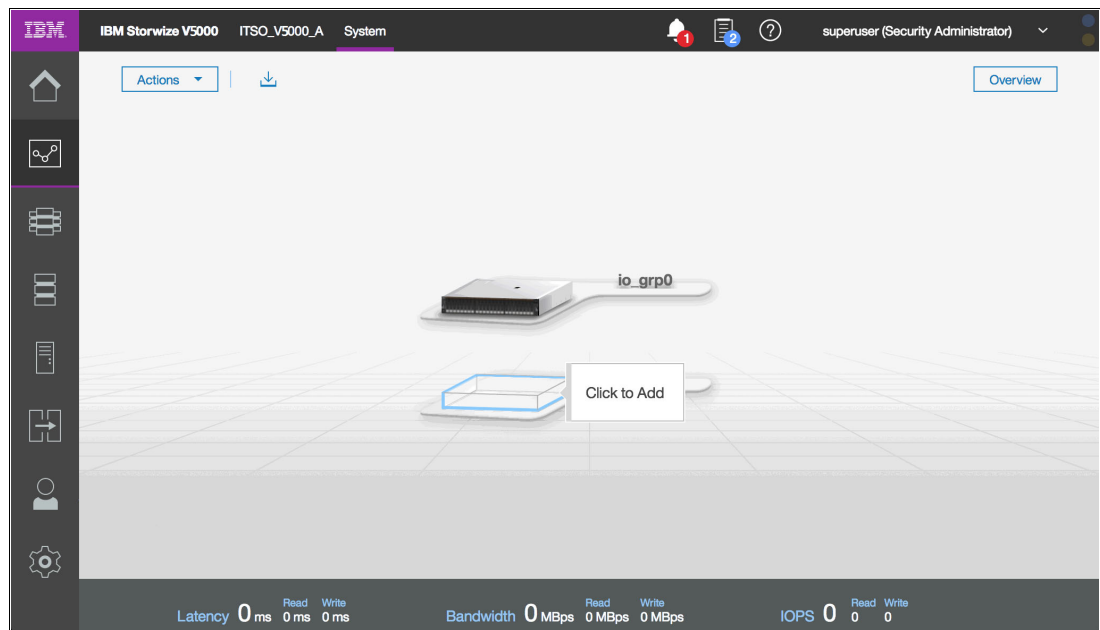


Figure 2-40 Second control enclosure

Complete the following steps to use the management GUI to configure the new enclosure:

1. In the main window, click **Actions** in the upper-left corner and select **Add Enclosures**. Alternatively, you can click the available control enclosure, as shown in Figure 2-41.

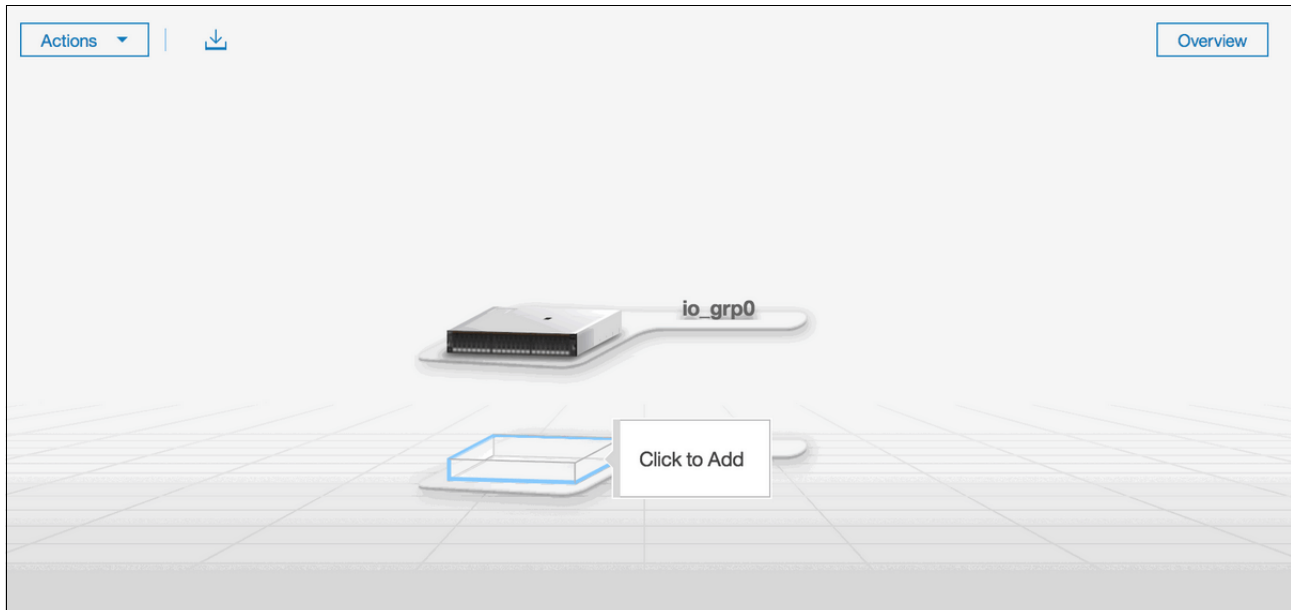


Figure 2-41 Option to add a control enclosure

If the control enclosure is configured correctly, the new control enclosure is identified in the next window, as shown in Figure 2-42.

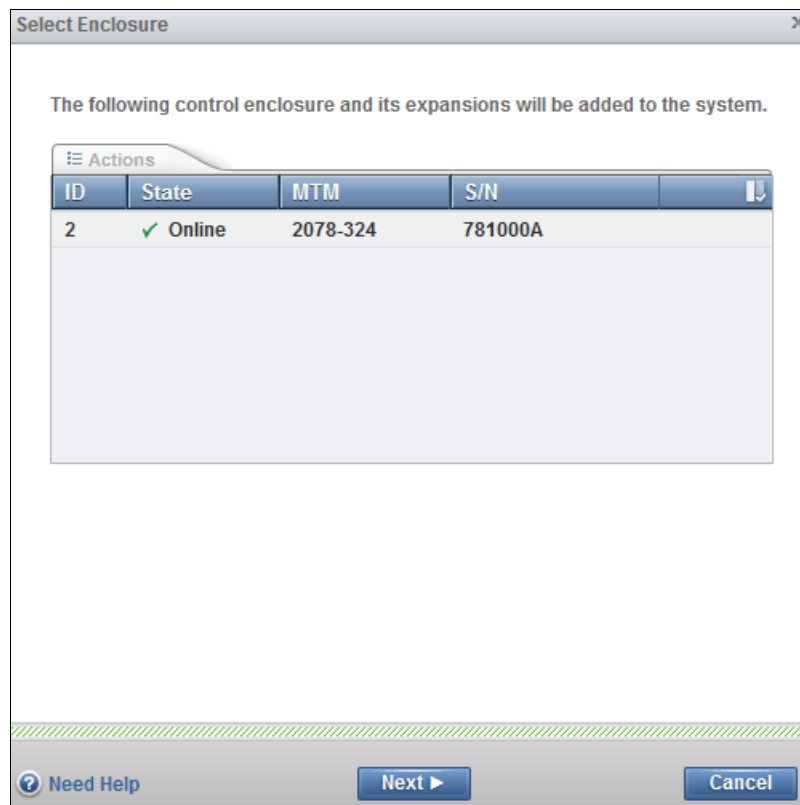


Figure 2-42 New control enclosure identification

2. Select the control enclosure and click **Actions** → **Identify** to turn on the identify LEDs of the new enclosure, if required. Otherwise, click **Next**.
3. The new control enclosure is added to the system, as shown in Figure 2-43. Click **Finish** to complete the operation.

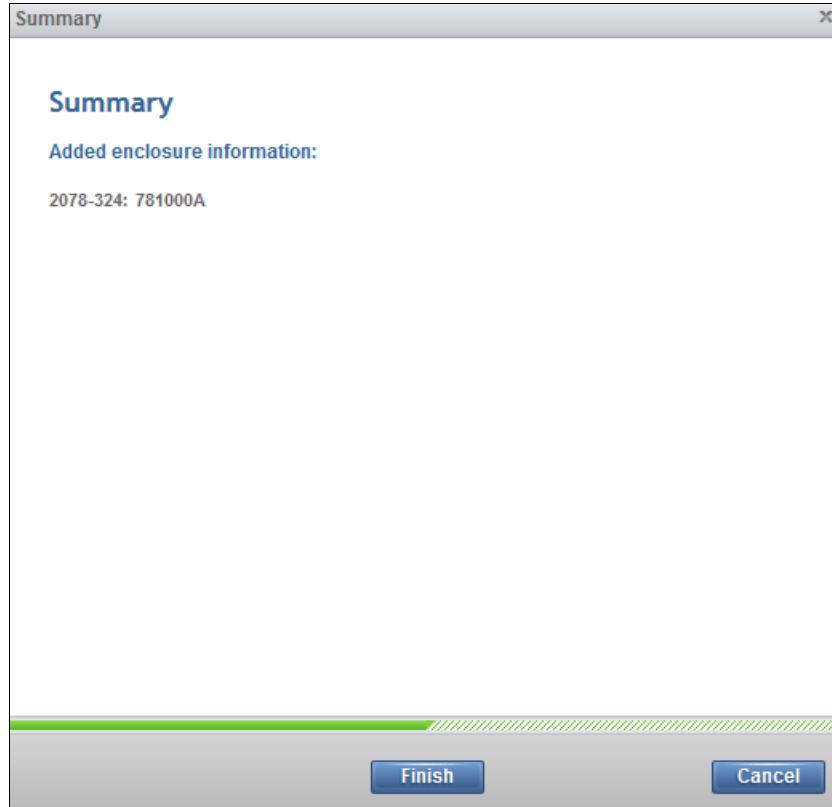


Figure 2-43 Added enclosure summary

When the new enclosure is added, the storage that is provided by the internal drives is available to use, as shown in Figure 2-44.

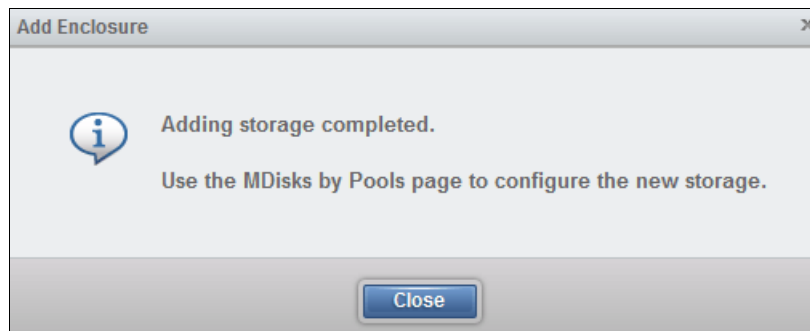


Figure 2-44 Adding storage completed

After the wizard adds the new control enclosure, the IBM Storwize V5000 Gen2 shows the management GUI that contains two I/O groups, as shown in Figure 2-45.

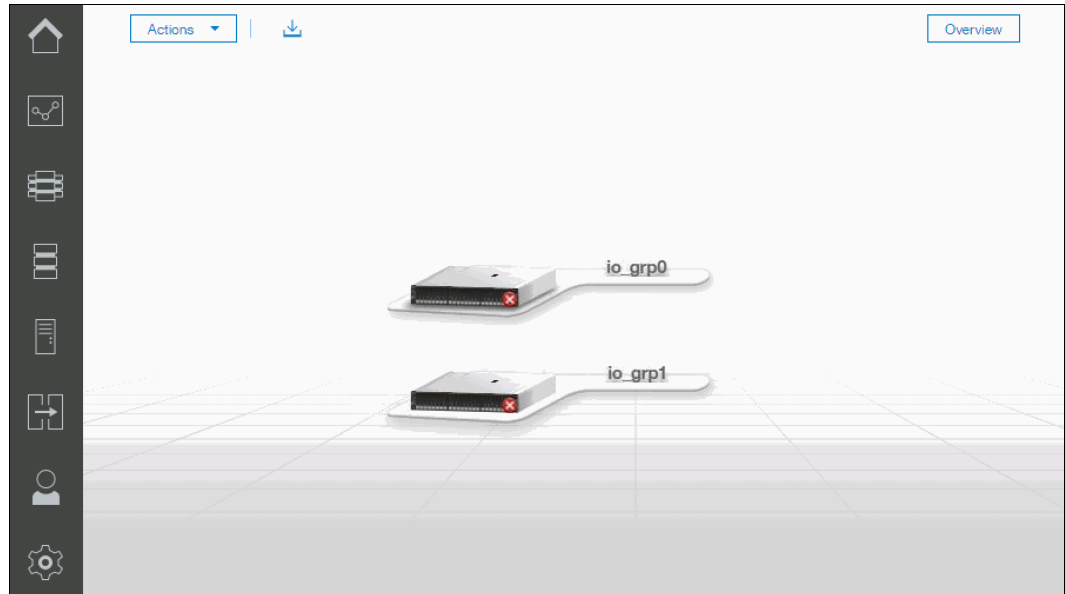


Figure 2-45 IBM Storwize V5000 Gen2 GUI with two I/O groups

Adding an expansion enclosure

Complete the following steps to add an expansion controller:

1. To add an expansion enclosure, change to the Monitoring tab and select **System**. If no new hardware is shown, check your cabling to ensure that the new expansion enclosure is connected correctly and refresh the window.

In the main window, click **Actions** in the upper-left corner and select **Add Enclosures**. Alternatively, you can click the available expansion enclosure, as shown in Figure 2-46.

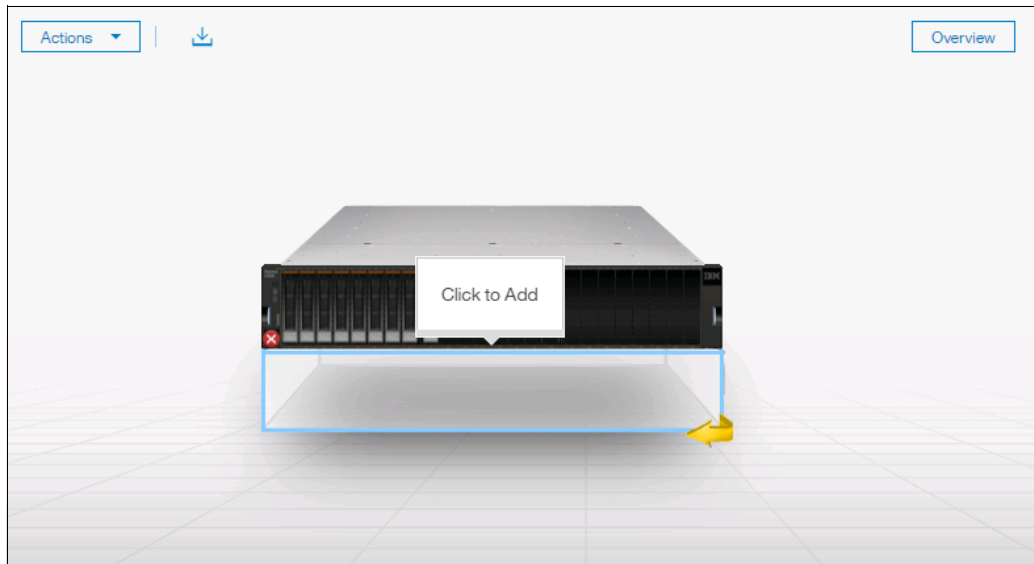


Figure 2-46 Adding an expansion enclosure

2. If the enclosure is cabled correctly, the wizard identifies the candidate expansion enclosure. Select the expansion enclosure and click **Next**, as shown in Figure 2-47.

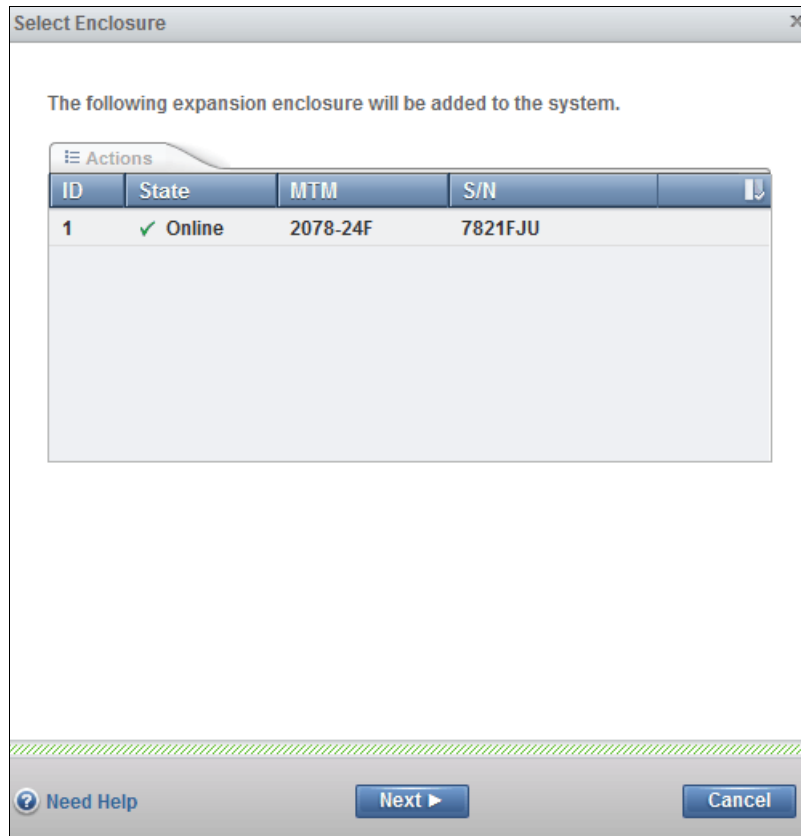


Figure 2-47 Expansion enclosure cable check

3. Select the expansion enclosure and click **Actions** → **Identify** to turn on the identify LEDs of the new enclosure, if required. Otherwise, click **Next**.

4. The new expansion enclosure is added to the system, as shown in Figure 2-48. Click **Finish** to complete the operation.

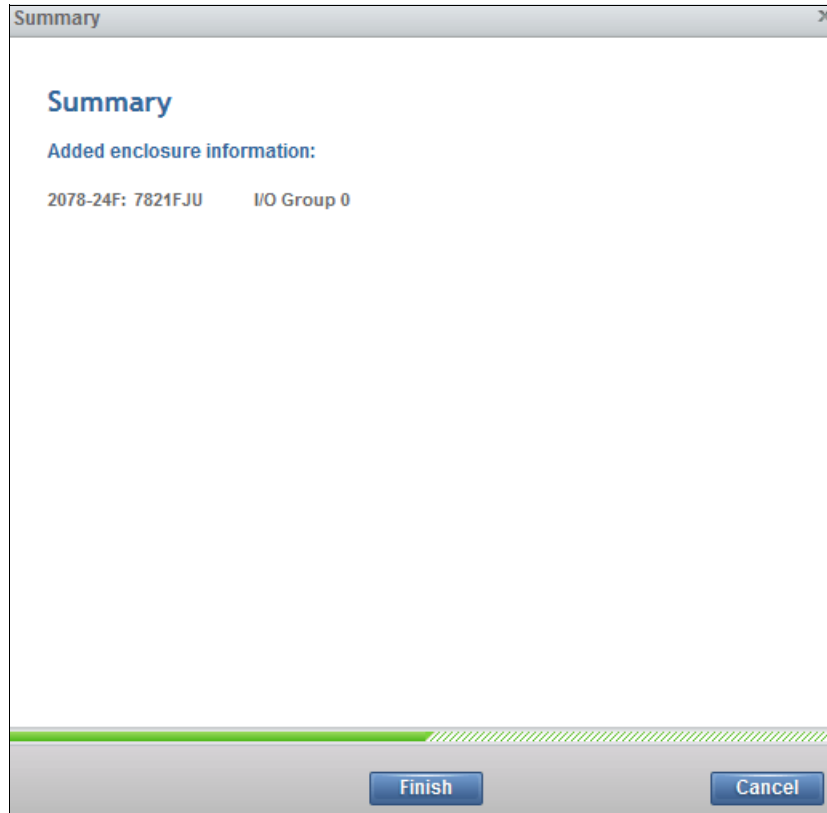


Figure 2-48 Added enclosure summary

After the expansion enclosure is added, the IBM Storwize V5000 Gen2 shows the management GUI that contains two enclosures, as shown in Figure 2-49.



Figure 2-49 IBM Storwize V5000 Gen2 GUI with two enclosures in a single I/O group

2.10.2 Service Assistant Tool

The IBM Storwize V5000 Gen2, as a single I/O group, is configured initially with three IP addresses: one service IP address for each node canister, and a management IP address, which is set when the cluster is started.

The management IP and service IP addresses can be changed within the GUI, as described in Chapter 3, “Graphical user interface overview” on page 85.

IBM Service Assistant (SA) Tool is a web-based GUI that is used to service individual node canisters, primarily when a node has a fault and it is in a service state. A node cannot be active as part of a clustered system while the node is in a service state. The SA Tool is available even when the management GUI is not accessible. The following information and tasks are included:

- ▶ Status information about the connections and the node canister
- ▶ Basic configuration information, such as configuring IP addresses
- ▶ Service tasks, such as restarting the Common Information Model object manager (CIMOM) and updating the worldwide node name (WWNN)
- ▶ Details about node error codes and hints about how to fix the node error

Important: The SA Tool can be accessed by using the superuser account only. You must access SA Tool under the direction of IBM Support only.

The Service Assistance GUI is available by using a service assistant IP address on each node. The SA GUI is accessed through the cluster IP addresses by appending service to the cluster management URL.

If the system is down, the only other method of communicating with the node canisters is through the SA IP address directly. Each node can have a single SA IP address on Ethernet port 1. We advise that these IP addresses are configured on all of the Storwize V5000 Gen2 node canisters.

The default IP address of canister 1 is 192.168.70.121 with a subnet mask of 255.255.255.0.

The default IP address of canister 2 is 192.168.70.122 with a subnet mask of 255.255.255.0.

To open the SA GUI, enter one of the following URLs in any web browser:

- ▶ `http(s)://cluster IP address of your cluster/service`
- ▶ `http(s)://service IP address of a node/service`

The following examples open the SA GUI:

- ▶ Management address:
`http://1.2.3.4/service`
- ▶ SA access address:
`http://1.2.3.5/service`

When you access SA by using the `<cluster address>/service`, the configuration node canister SA GUI login window opens, as shown in Figure 2-50 on page 83.



Figure 2-50 Service Assistant Tool login

The SA interface can view status and run service actions on other nodes and the node where the user is connected.

After you are logged in, you see the Service Assistant Tool Home window, as shown in Figure 2-51.

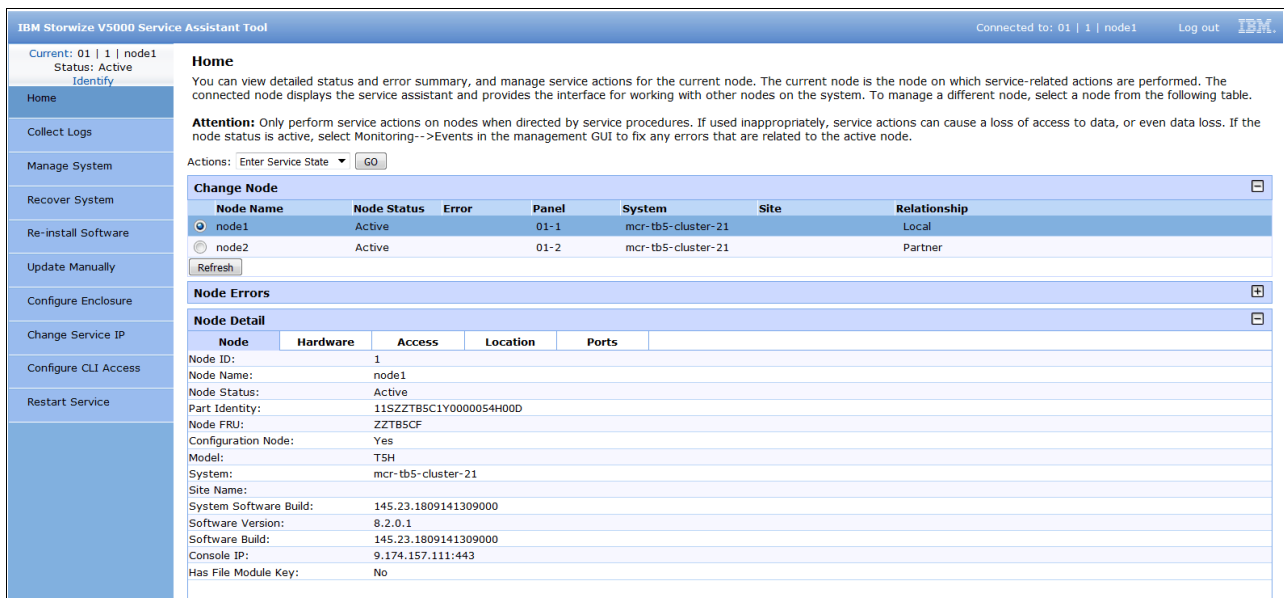


Figure 2-51 Service Assistant Tool Home window

The current canister node is displayed in the upper-left corner of the GUI. As shown in Figure 2-51, the current canister node is node 1.

To change the canister, select the relevant node in the Change Node section of the window. You see that the details in the upper-left corner change to reflect the new canister.

The SA GUI provides access to service procedures and shows the status of the node canisters. We advise that you perform these procedures only if you are directed to use them by IBM Support.

For more information about how to use the SA Tool, see [this website](#).



Graphical user interface overview

This chapter provides an overview of the graphical user interface (GUI) of IBM Spectrum Virtualize on the IBM Storwize V5000 Gen2 and shows you how to use the navigation tools.

This chapter includes the following topics:

- ▶ 3.1, “Overview of IBM Spectrum Virtualize management software” on page 86
- ▶ 3.2, “Monitoring menu” on page 95
- ▶ 3.3, “Pools menu” on page 107
- ▶ 3.4, “Volumes menu” on page 123
- ▶ 3.5, “Hosts menu” on page 128
- ▶ 3.6, “Copy services” on page 134
- ▶ 3.7, “Access menu” on page 143
- ▶ 3.8, “Settings menu” on page 147

3.1 Overview of IBM Spectrum Virtualize management software

A GUI can simplify storage management and provide a fast and more efficient management tool. IBM Spectrum Virtualize V8.1 GUI features significant changes from previous versions, such as the icons, color palette, and object locations. However, usability is a priority, as in all IBM Spectrum products, and usability is maintained in the GUI.

JavaScript: You must enable JavaScript in your browser. For Mozilla Firefox, JavaScript is enabled by default and requires no other configuration. For more information about configuring your web browser, see [this website](#).

3.1.1 Accessing the storage management software

To access the Storwize V5000 Gen2, complete the following steps:

1. To log on to the management software, enter the IP address that was set during the initial setup process into the address line of your web browser. You can connect from any workstation that can communicate with the system. The login window opens (see Figure 3-1 on page 87).

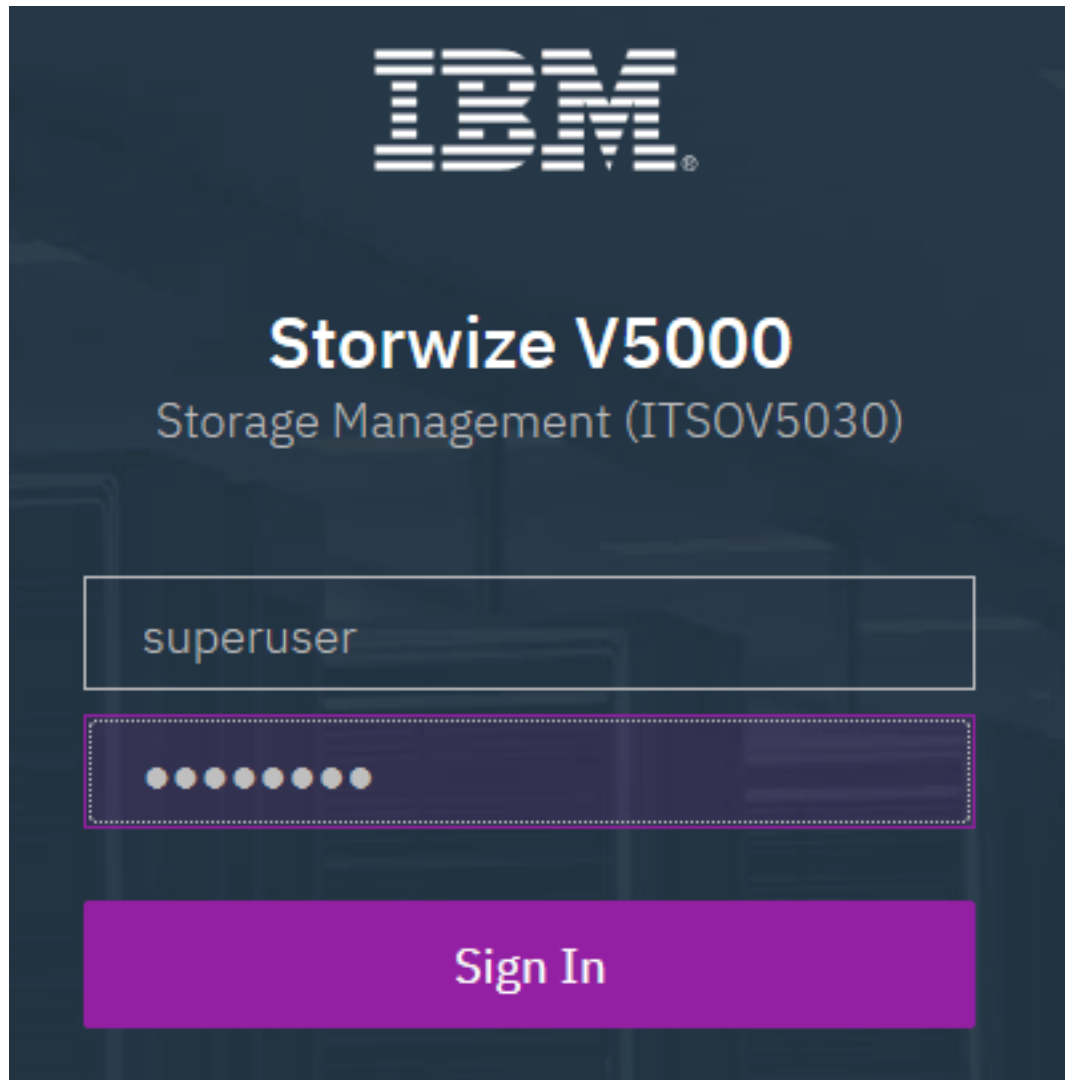


Figure 3-1 Login window

We suggest that each user who operates IBM Spectrum Virtualize has an account that is not shared with someone else. The default user accounts need to be unavailable for remote access, or the passwords need to be changed from the default password and known only to the system owner or kept secured for emergency purposes only.

This approach helps to identify the personnel who are working on the device and to track all of the important changes in the systems. The *Superuser* account must be used for initial configuration only.

- After a successful login, the IBM Spectrum Virtualize System pane displays the Dashboard with all relevant details of your system, as shown in Figure 3-2.

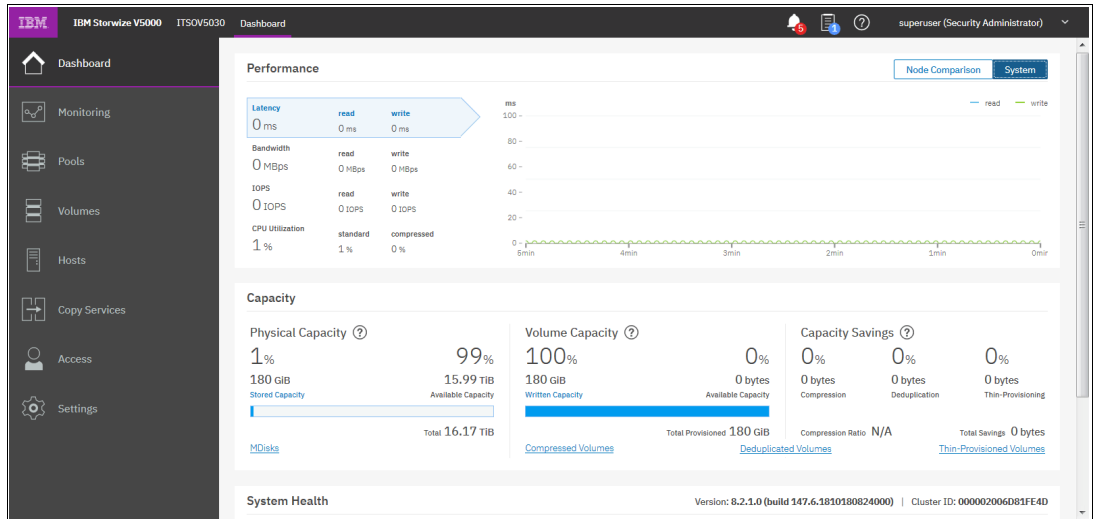


Figure 3-2 A first view

The IBM Spectrum Virtualize System pane is an important user interface. Throughout this chapter, we refer to it as the *IBM Spectrum Virtualize System pane* or *System pane*. In the remaining chapters, we do not explain how to access it each time.

3.1.2 System pane layout

The System pane has four main sections for navigating through the management tool:

- Top

The top menu shows the *navigation path* so that the user knows the exact path.

The Actions option can be used at any time to add more enclosures, modify hardware, or rename the system.

The top menu also has a *system overview* (upper right corner) so that you can see global system topology.

- Left

The *system menu* (or set of *function icons*) is the main object that the users access. By using the system menu, the user can change or view any setting or parameter in the system.

- Center

Within the *system view* object, users can view or change parameters that mainly relate to the hardware of global system settings.

- Bottom

The *informational pane* consists of running tasks, capacity information, basic performance indicator, system health status, and status alerts.

Figure 3-3 shows these main areas.

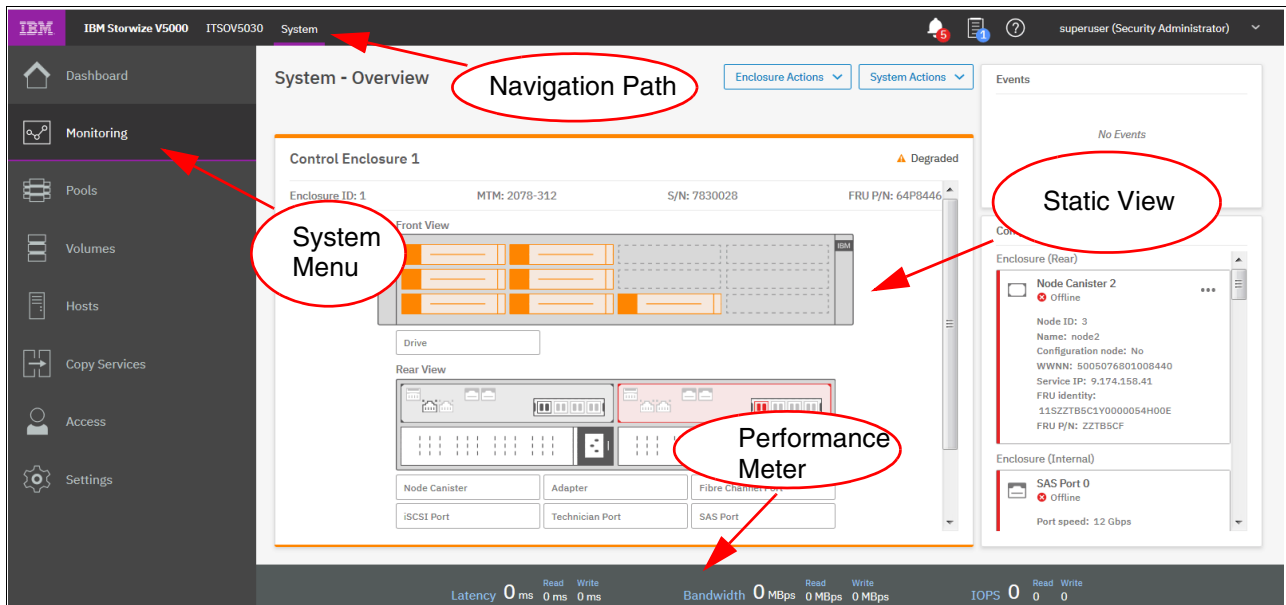


Figure 3-3 Main areas

The following main areas are available:

- ▶ The left side of the window shows eight *function icons*, which are collectively referred to as a *dynamic menu*. The dynamic menu (see Figure 3-4) includes the following function icons:
 - Dashboard
 - Monitoring menu
 - Pools menu
 - Volumes menu
 - Hosts menu
 - Copy Services menu
 - Access menu
 - Settings menu

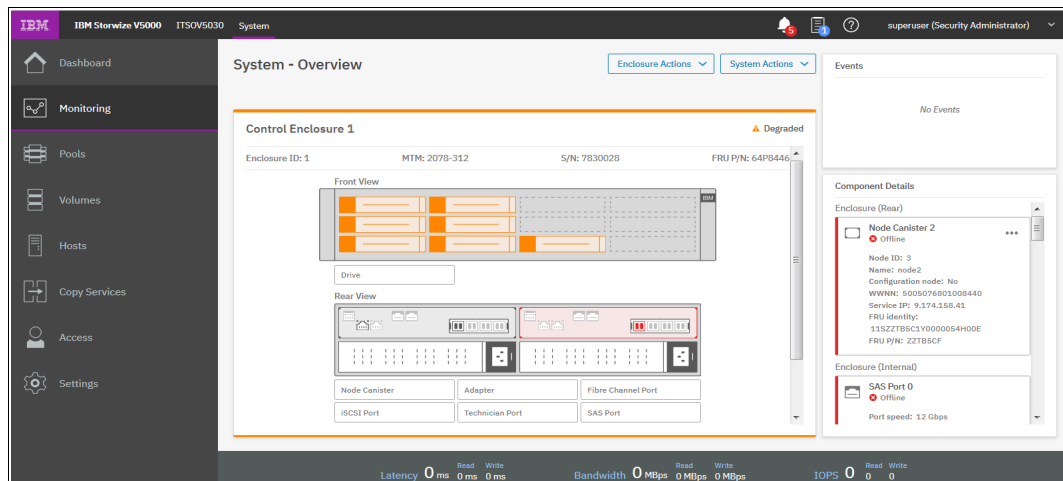


Figure 3-4 Component Model

- ▶ The middle of the window shows a component model of the configuration. Clicking over each component highlights that part and provides a pop-up menu with a description that identifies important parameters and functions of this element.
- ▶ The bottom of the window shows the performance indicator. It gives you information about how your machine is performing right now. This information covers only the external tasks for attached hosts. Figure 3-5 shows the performance indicator.

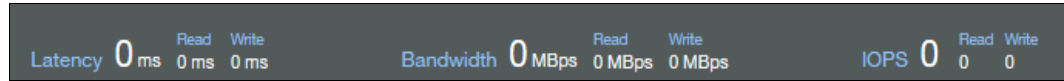


Figure 3-5 Performance indicator

Review the System statistic page to see more information about how internal performance, as shown in Figure 3-6.

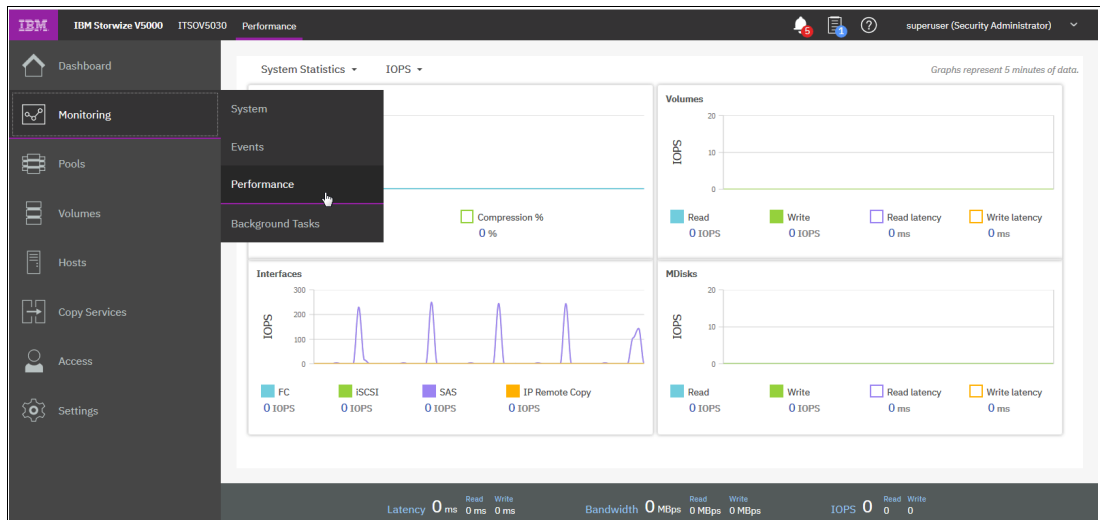


Figure 3-6 System statistic page

- ▶ In the right upper area, you get more information about the health of your system. Clicking the **Event** button shows the information, as shown in Figure 3-7.

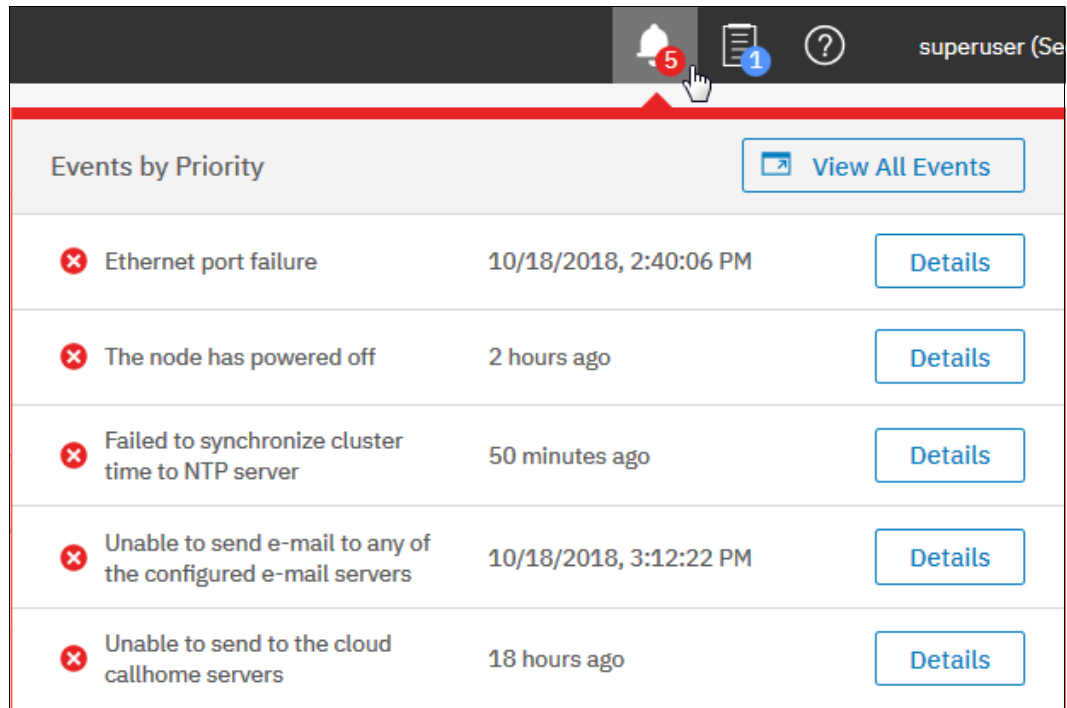


Figure 3-7 Event button

Alternatively, you can click the **Suggested tasks** button, as shown in Figure 3-8.

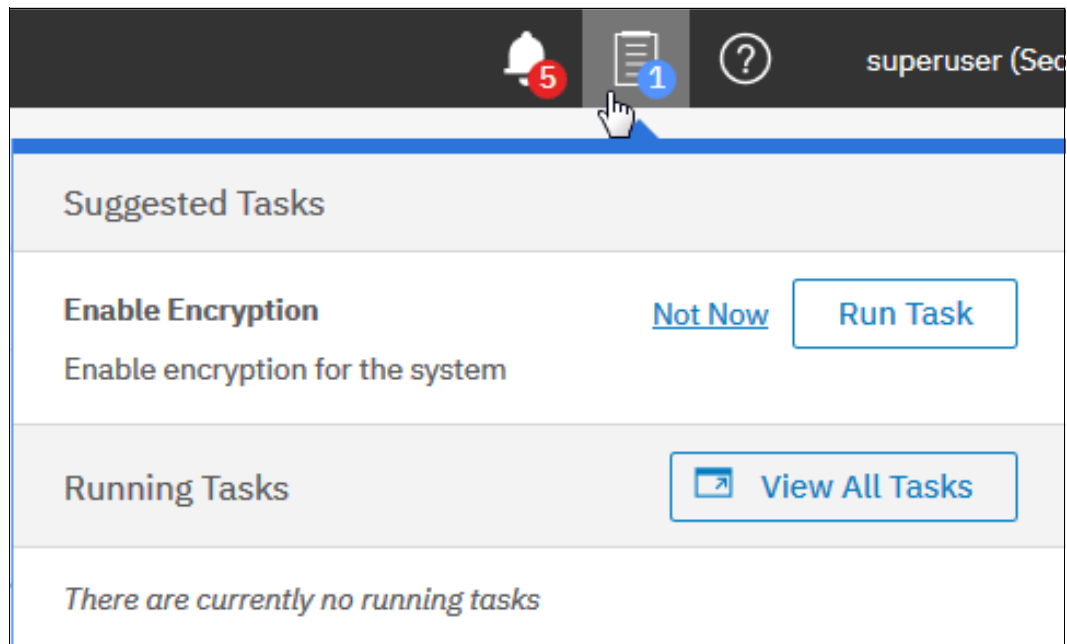


Figure 3-8 Suggested Tasks

A help menu is shown in Figure 3-9.

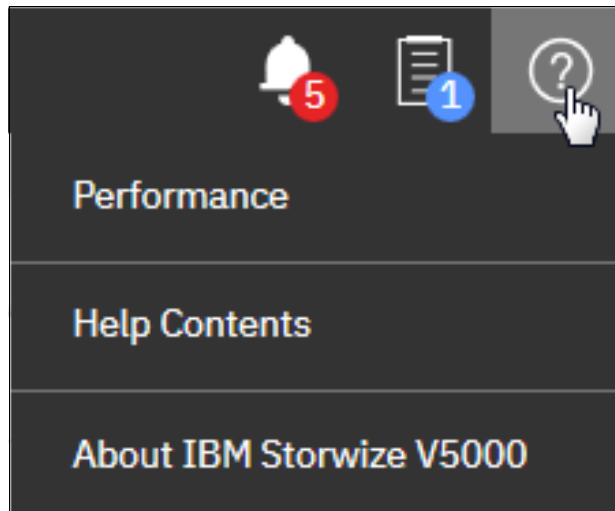


Figure 3-9 Help Menu

Clicking any of these options provides more detailed information about the configuration, situation, or status of the IBM Spectrum Virtualize solution. Click any of these function icons to expand them and minimize them, as required. In an error or warning situation, those indicators are extended by the status alerts icon in the upper-right corner, as shown in Figure 3-7 on page 91.

3.1.3 Navigation

Navigating in the management tool is simple. You can click one of the eight function icons to display a submenu of available options. You also can select specific options. Figure 3-10 shows how to access the **Pools** option, for example.

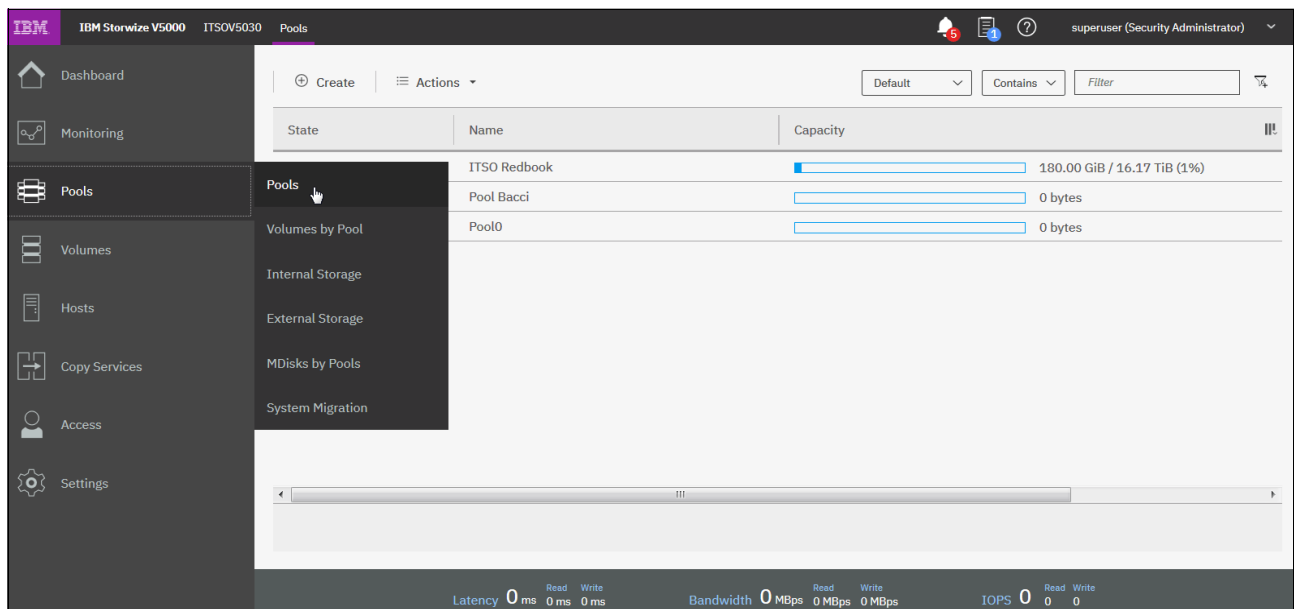


Figure 3-10 Navigating by using the menu options

3.1.4 Multiple selection

With the improved management tool, you can select multiple items by pressing Shift+Ctrl. To select multiple items in a display, click the first item, press and hold Shift, and then, click the last item that you require in the list. All rows between those two items are selected and highlighted in light blue (see Figure 3-11).

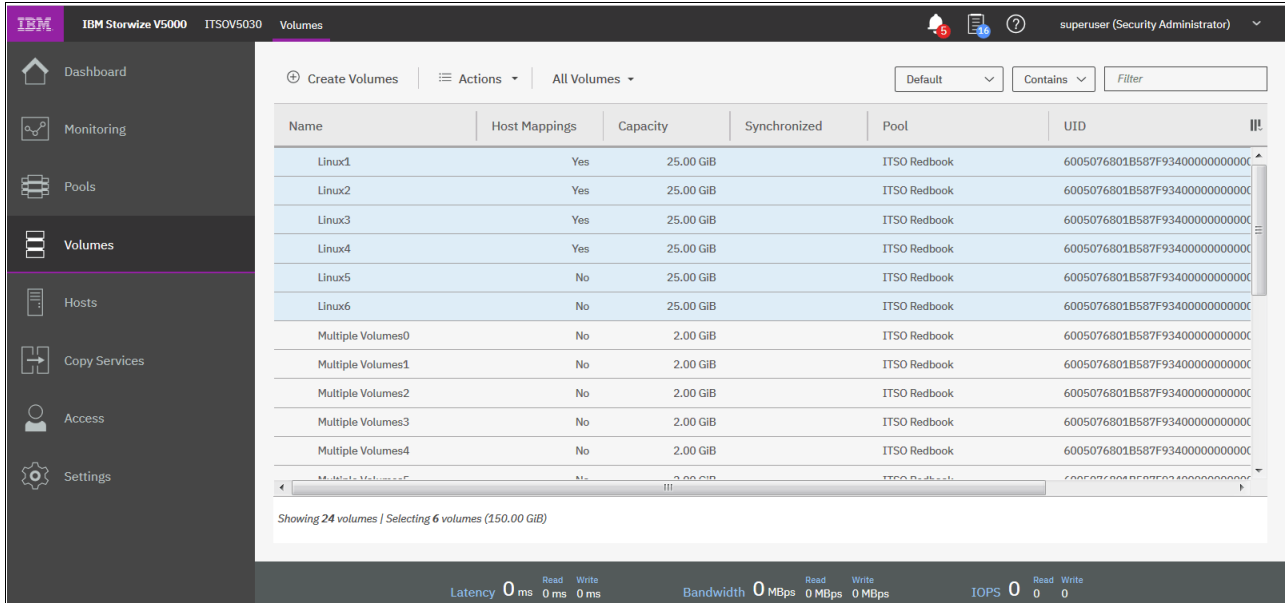


Figure 3-11 Multiple selections by using the Shift key

Similarly, if you want to select multiple items that are not in sequential order, click the first item, press and hold Ctrl, and then, click the other items that you need (see Figure 3-12).

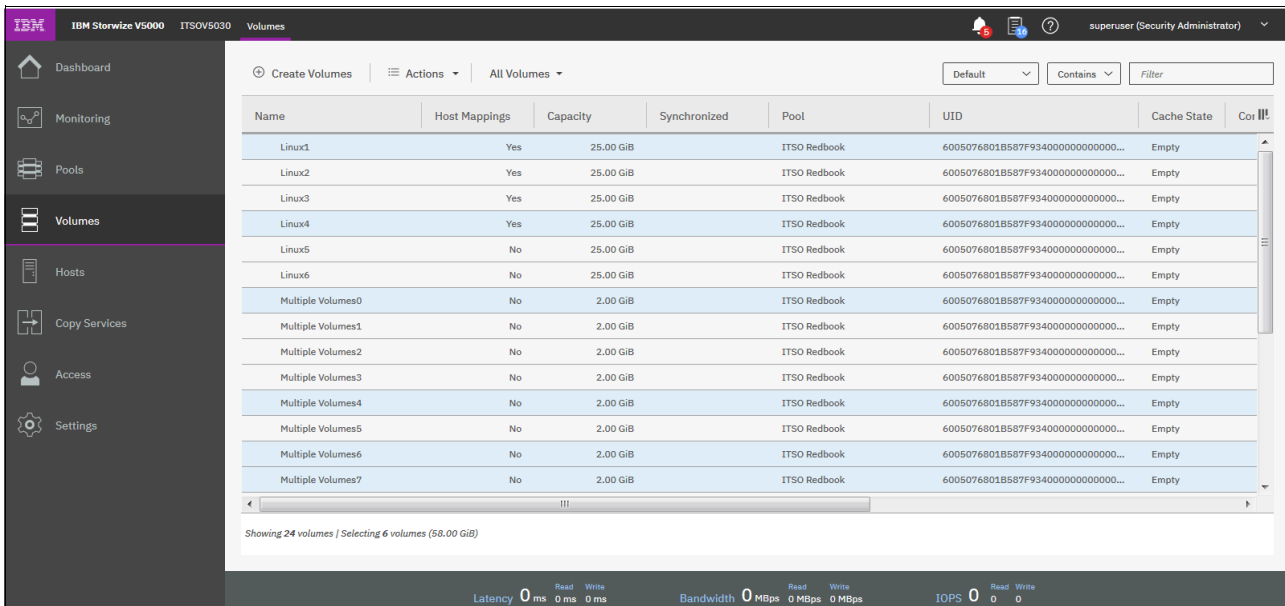


Figure 3-12 Multiple selections by using the Ctrl key

Another option for selecting volumes is selecting by mask. To select all volumes that have VMware in their name, enter VMware into the Filter field and press Enter. All volumes with VMware in their name are displayed. After you filter the volumes, you can easily select all of the displayed volumes or a subset by using the Ctrl key or Shift key technique (see Figure 3-13).

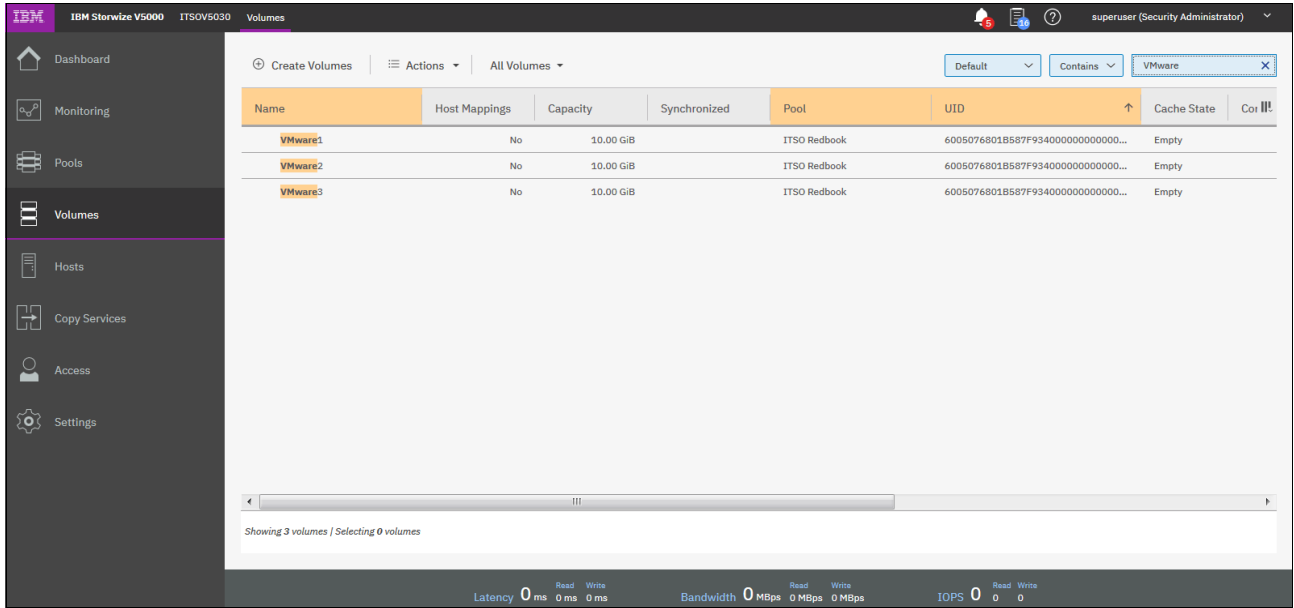


Figure 3-13 Filtering volumes

3.1.5 Status indicators area

The status indicators area at the bottom of the System pane (see Figure 3-14) shows a high-level status of the IBM Storwize V5000 storage system. Information that is shown in this area covers only the performance of the attached Host Systems.

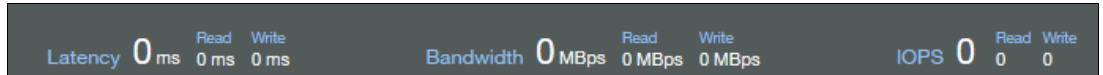


Figure 3-14 Status indicators

Help function

Another useful interface feature is the integrated Help function. You can access help for certain fields and objects by hovering over the question mark (?) icon (see Figure 3-15) that is next to the field. Panel-specific help is available by clicking **Need Help** or by clicking the **Help** link in the upper-right corner of the GUI.

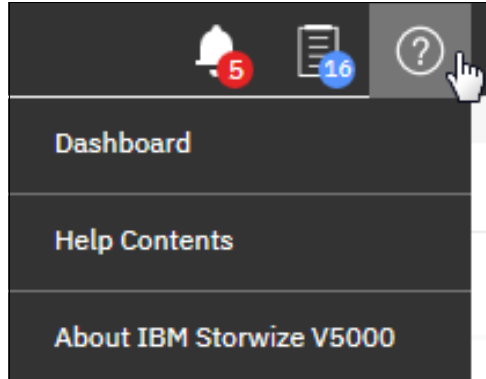


Figure 3-15 Access to panel-specific help

3.2 Monitoring menu

Point to the Monitoring function icon to open the Monitoring menu (see Figure 3-16 on page 97). The Monitoring menu offers the following navigation directions:

System From the System window, you can view details about control and expansion enclosures and various hardware components of the system. The system uses base-2 (binary numeral) as capacity indicators for volumes, drives, and other system objects. The management GUI and the command-line interface (CLI) use different abbreviations to indicate capacity, but the value for these capacity indicators is the same. For more information, see 3.2.1, “System overview” on page 97.

Events The Events window displays two types of events: messages and alerts. It also indicates the cause of any log entry. You can use this window to filter messages that are related to events that occurred in the system. Some alerts have a four-digit error code and a fix procedure that helps you fix the problem.

Other alerts also require action but do not have a fix procedure. Messages are fixed when you acknowledge reading them and mark them as fixed. Each event has a time stamp that indicates when the action occurred or the command was submitted on the system.

When logs are displayed in the command-line interface, the time stamps for the logs in the CLI are the system time. However, when logs are displayed in the management GUI, the time stamps are translated to the local time where the web browser is running.

For more information, see 3.2.3, “Events option” on page 104.

Performance Use real-time statistics to monitor CPU utilization, volume, interface, and MDisk bandwidth of your system and nodes. Each graph represents 5 minutes of collected statistics and provides a means of assessing the overall performance of your system.

You can use system statistics to monitor the bandwidth of all the volumes, interfaces, and MDisks that are being used on your system.

You can also monitor the overall CPU utilization for the system. These statistics summarize the overall performance health of the system and can be used to monitor trends in bandwidth and CPU utilization. You can monitor changes to stable values or differences between related statistics, such as the latency between volumes and MDisks. These differences then can be further evaluated by using performance diagnostic tools.

Additionally, with system-level statistics, you can quickly view bandwidth of volumes, interfaces, and MDisks. Each of these graphs displays the current bandwidth in megabytes per second (Mbps), and a view of bandwidth over time. Each data point can be accessed to determine its individual bandwidth use and to evaluate whether a specific data point might represent performance impacts.

For example, you can monitor the interfaces, such as for Fibre Channel or SAS interfaces, to determine whether the host data-transfer rate is different from the expected rate. You can also select node-level statistics, which can help you determine the performance impact of a specific node. As with system statistics, node statistics help you to evaluate whether the node is operating within normal performance metrics.

For more information, see 3.2.4, “Performance pane” on page 105.

Background Tasks

Use the Background Tasks page to view and manage current tasks that are running on the system. The Background Tasks page displays all long-running tasks that are in progress on the system. Tasks, such as volume synchronization, array initialization, and volume formatting, can take some time to complete. The Background Tasks page displays these tasks and their progress.

After the task completes, the task is automatically deleted from the display. If a task fails with an error, select **Monitoring** → **Events** to determine the problem.

For more information, see 3.2.5, “Background Task” on page 107.

The option that was known as System Details is integrated into the device overview on the general System pane, which is available after login or when you click **System** from the Monitoring menu. For more information, see 3.2.2, “System details” on page 101.

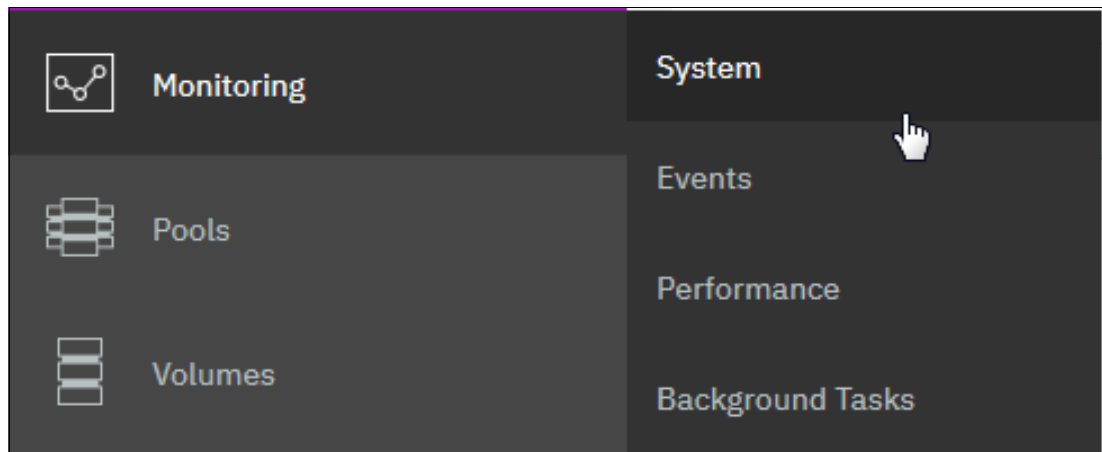


Figure 3-16 Accessing the Monitoring menu

In the following sections, we describe each option on the Monitoring menu.

3.2.1 System overview

The System option on the Monitoring menu provides a general overview about your Storwize V5000 system, including the depiction of all devices in a rack that are directly connected to it (see Figure 3-17).

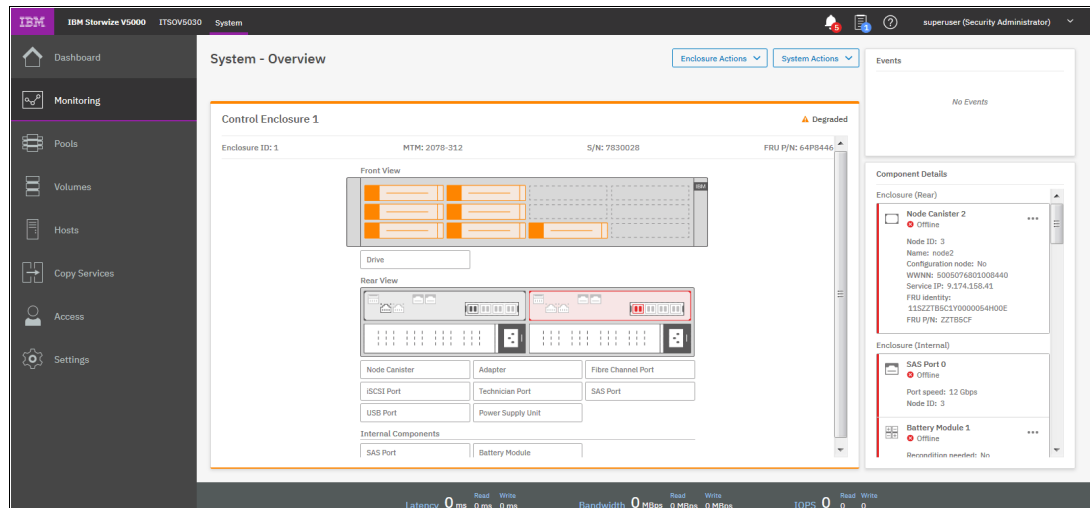


Figure 3-17 System overview

When you hover over a specific component in an enclosure, a pop-up window shows the details of disk drives in the unit. The details of Drive 1 in an enclosure are shown in Figure 3-18.

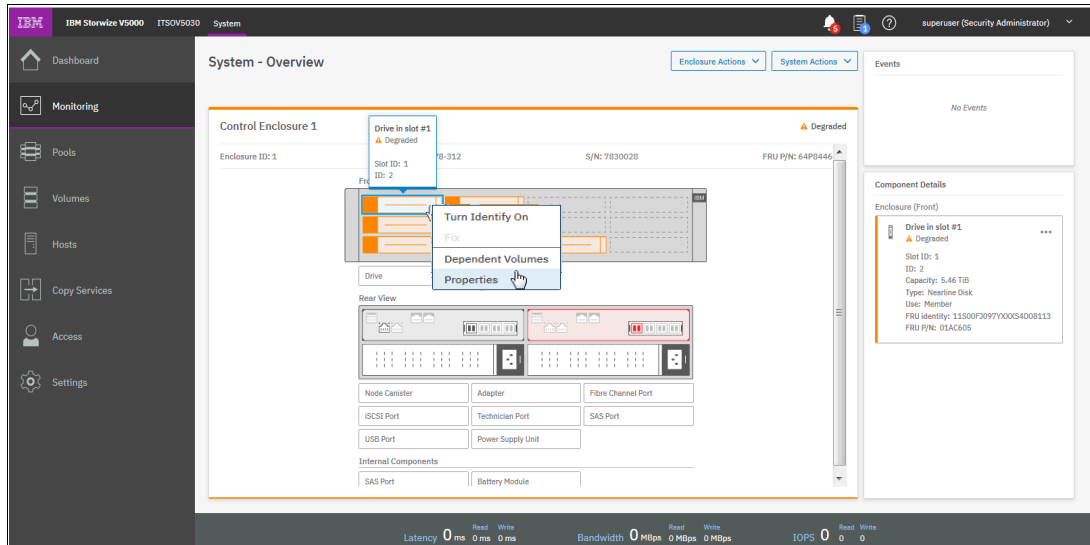


Figure 3-18 Component view

By right-clicking and selecting **Properties**, you see detailed technical parameters, such as capacity, interface speed, rotation speed, and the drive status (online or offline). Click **View more details** in the properties frame, as shown in Figure 3-19.

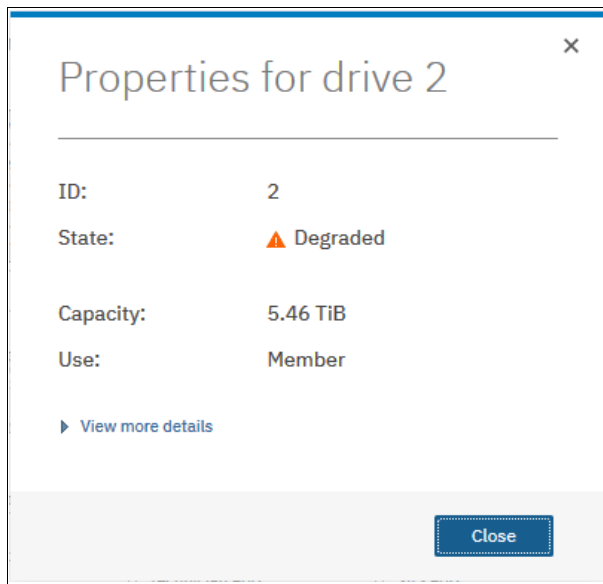


Figure 3-19 Properties and View more details option

A detailed object properties view is shown in Figure 3-20.

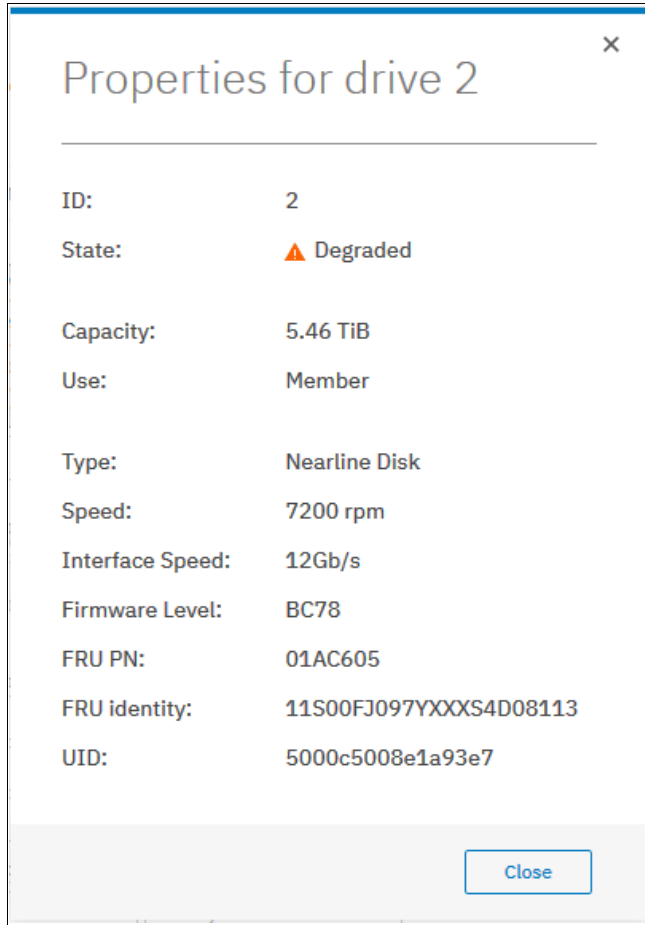


Figure 3-20 Component details

In an environment with multiple Storwize V5000 systems, you can easily navigate the onsite personnel or technician to the correct device by enabling the identification LED on the front pane.

First, right-click the enclosure or drive that you want to identify. Then, click **Identify** in the pop-up window that is shown in Figure 3-21 on page 100 and wait for the confirmation from the technician that the device in the data center was identified correctly. You can also see on the System Overview window that the LED is flashing.

After the confirmation, click **Turn LED Off** (see Figure 3-21).

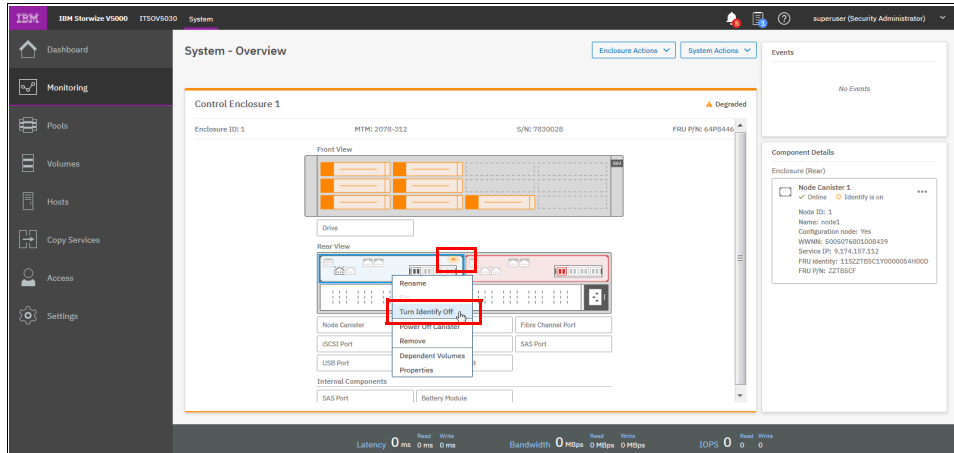


Figure 3-21 Using the identification LED

Alternatively, you can use the IBM Spectrum Virtualize command-line interface (CLI) to obtain the same results. Enter the following sequence of commands:

- ▶ `svctask chenclosure -identify yes 1` (or just `chenclosure -identify yes 1`)
- ▶ `svctask chenclosure -identify no 1` (or just `chenclosure -identify no 1`)

You can use the same CLI to obtain results for a specific controller or drive.

The front and rear view of the system is shown in the system pane view. If malfunctions exist, you can get more information on the right side by clicking the failing part, which is displayed in red, as shown in Figure 3-22.

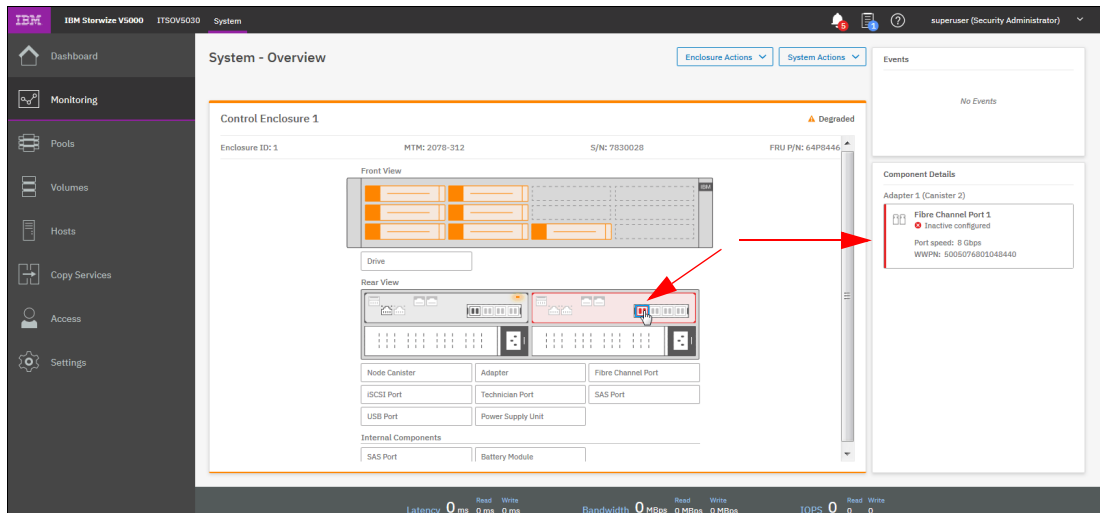


Figure 3-22 Backside of the enclosure

3.2.2 System details

You find the System Details option under **Enclosure Actions** (see Figure 3-23).

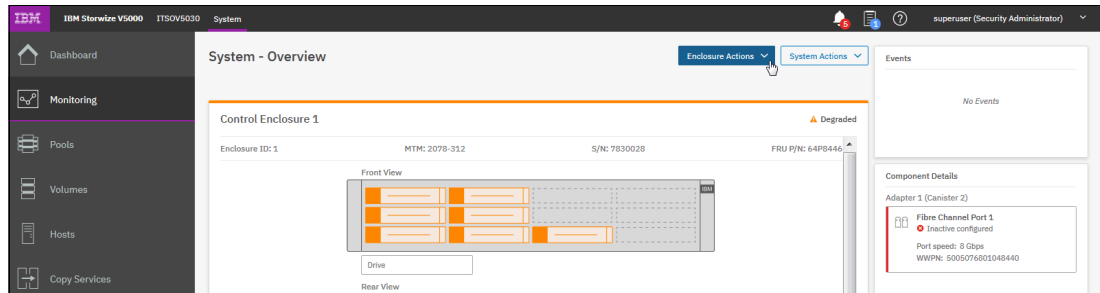


Figure 3-23 Enclosure Actions option

The Enclosure Actions option provides the extended level of parameters and technical details that relate to the system, including the integration of each element with an overall system configuration (see Figure 3-24).

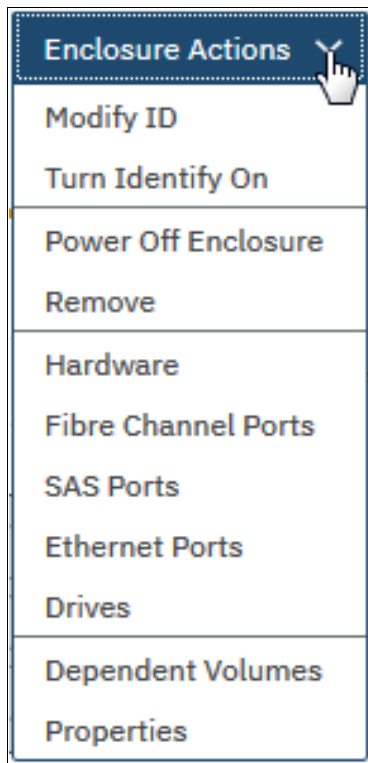


Figure 3-24 Enclosure Actions

By using this menu, you can also power off the machine.

Warning: Powering the machine ON remotely is not possible.

Remove the node or enclosure from the system, or list all volumes that are associated with the system (Dependent Volumes) by selecting **Properties**. The output is shown in Figure 3-25.

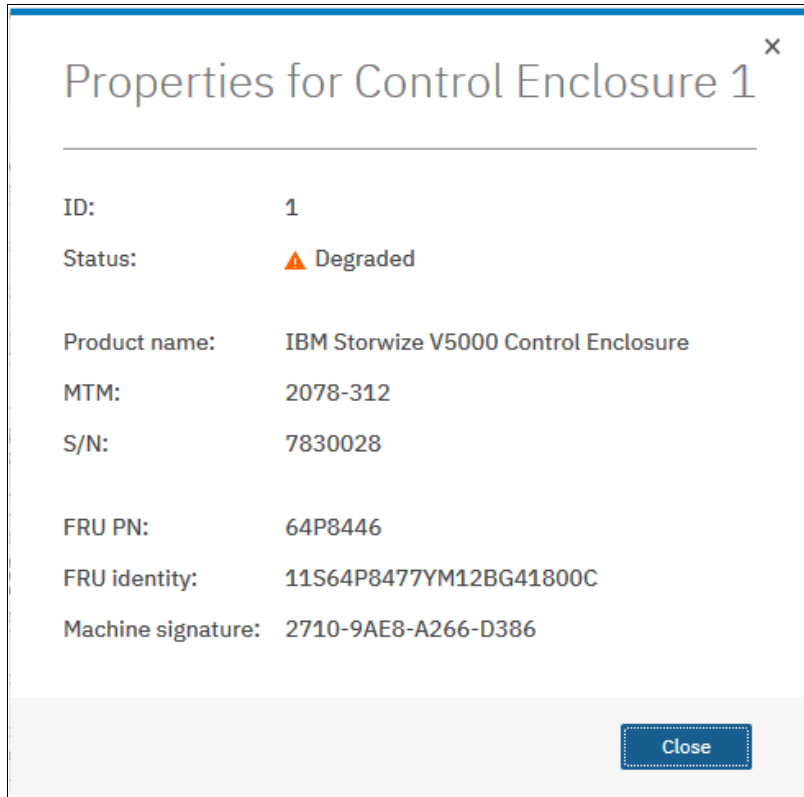


Figure 3-25 Enclosure technical details under the Properties option

Additionally, from the Enclosure Actions pane, you can get an overview (View) of the hardware, available ports, and status for Fibre Channel (FC) and serial-attached SCSI (SAS) ports and Drives (see Figure 3-26).

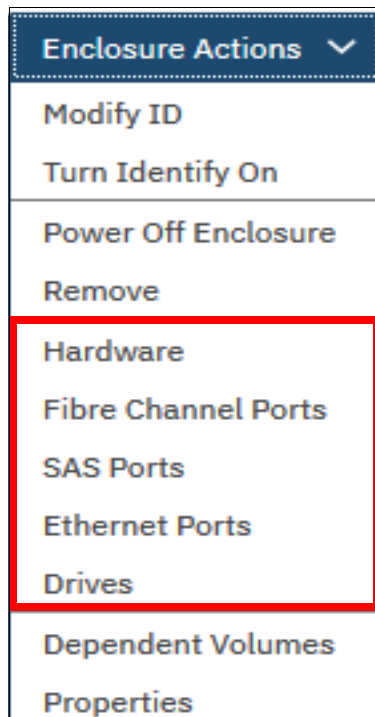


Figure 3-26 Canister details and vital product data

By selecting, for example, **Fibre Channel Ports**, you can see the list and status of available FC ports with their speed and worldwide port names (WWPNs), as shown in Figure 3-27.

The screenshot shows a window titled "Fibre Channel Ports for Control Enclosure 1". At the top, there are controls for "Actions", "Default", "Contains", and a "Filter" input field. Below this is a table with the following columns: "Port ...", "Owning Node", "Virtualized", "State", "Speed", and "WWPN". The table contains 8 rows of data. Below the table is an information message: "You can change the WWPN notation from the actions menu". A "Close" button is located at the bottom right of the window.

Port ...	Owning Node	Virtualized	State	Speed	WWPN
1	node1	No	Active	8Gb	5005076801048439
1	node1	Yes	Active	8Gb	5005076801748439
1	node2	No	Inactive configured	8Gb	5005076801048440
1	node2	Yes	Inactive configured	8Gb	5005076801748440
2	node1	No	Inactive unconfigured	N/A	5005076801088439
2	node1	Yes	Inactive unconfigured	N/A	5005076801788439
2	node2	No	Inactive unconfigured	N/A	5005076801088440
2	node2	Yes	Inactive unconfigured	N/A	5005076801788440

You can change the WWPN notation from the actions menu

Figure 3-27 Status of FC ports in the control enclosure

3.2.3 Events option

The Events option, which is selected from the Monitoring menu (see Figure 3-16 on page 97), tracks all informational, warning, and error messages that occur in the system. You can apply various filters to sort them or export them to an external CSV file. A CSV file can be created from the information that is included in the Events list.

Figure 3-28 on page 105 shows the display after you click **Events** from the Monitoring menu.

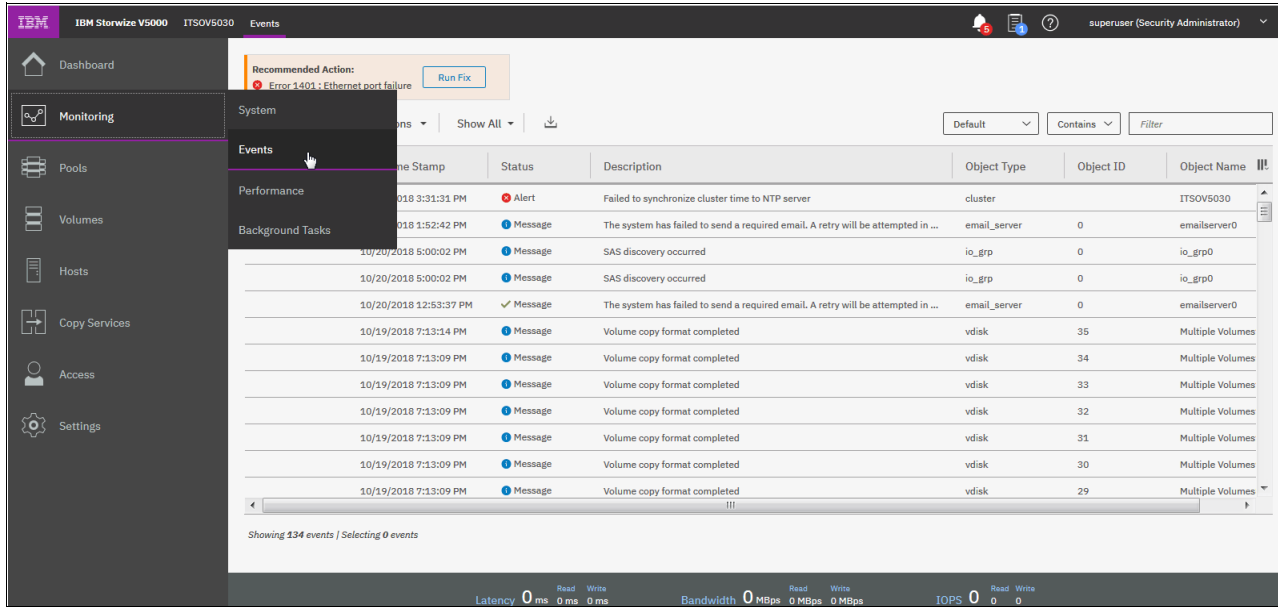


Figure 3-28 Events log

For more information about how to work with the events log and run various fix procedures by using the Events option, see Chapter 12, “RAS, monitoring, and troubleshooting” on page 623.

3.2.4 Performance pane

The Performance pane reports the general system statistics that relate to processor (CPU) use, host and internal interfaces, volumes, and MDisks. The Performance pane might be useful when you compare the performance of each node in the system if problems exist after a node failover occurs (see Figure 3-29).

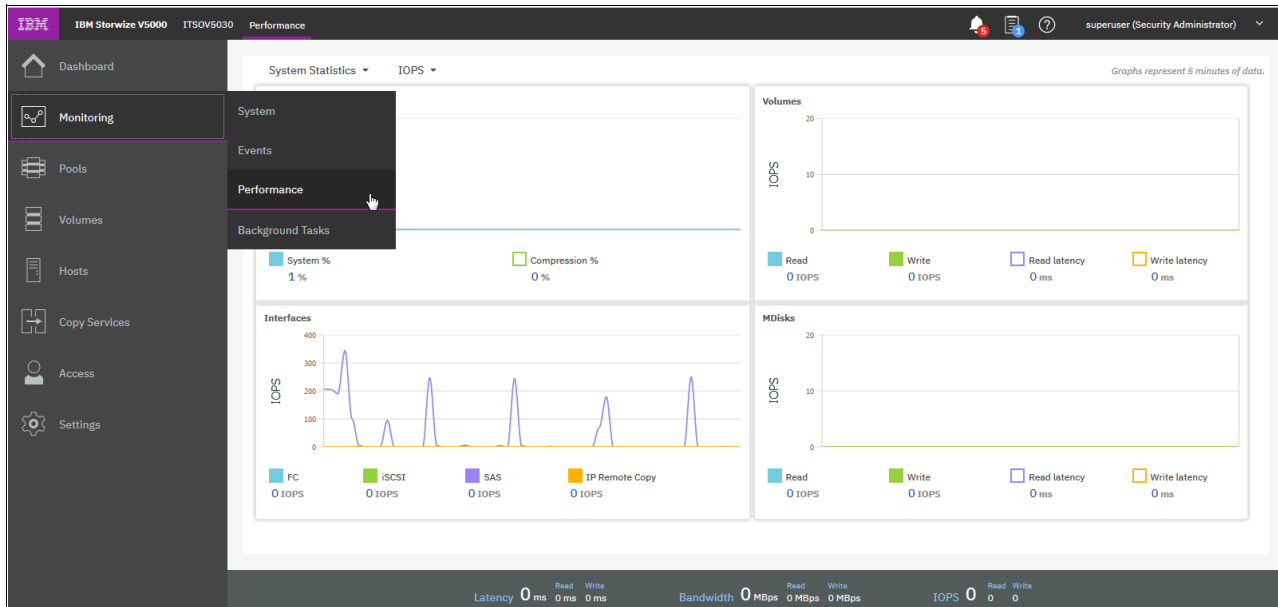


Figure 3-29 Performance statistics of the IBM Storwize V5000 system

You can switch between MBps or IOPS (see Figure 3-30, or even navigate to the statistics at the node level.

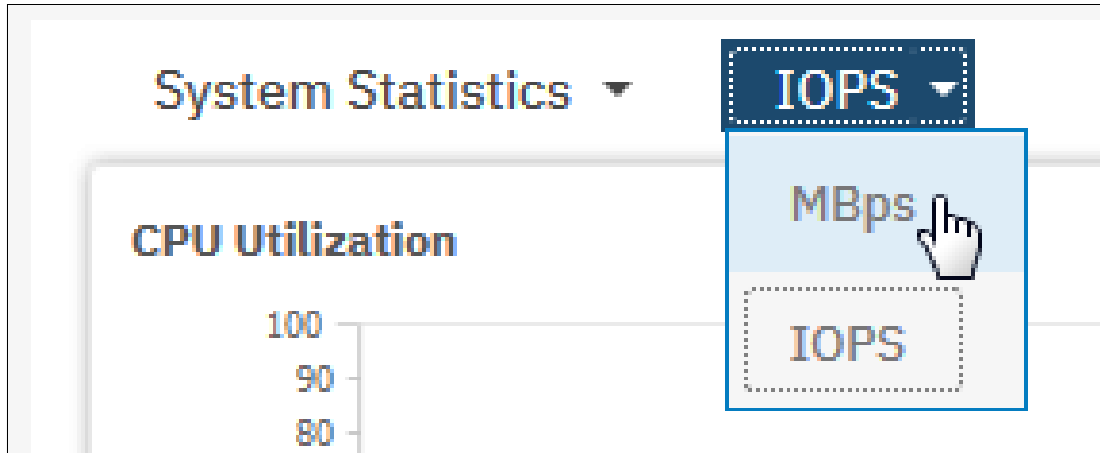


Figure 3-30 Switching between MBps or IOPS

The performance statistics in the GUI shows, by default, the last 5 minutes of data. To see the details of each sample, click the graph and select the time stamp, as shown in Figure 3-31.

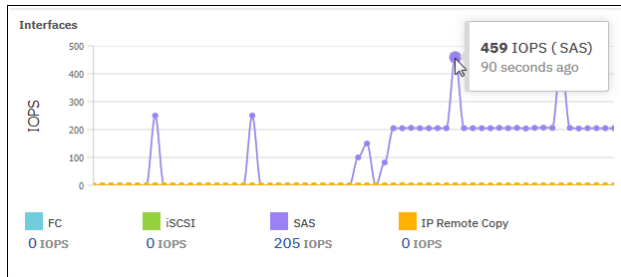


Figure 3-31 Sample details

The previous charts represent 5 minutes of the data stream. For in-depth storage monitoring and performance statistics of your IBM Spectrum Virtualize System with historical data, use the IBM Spectrum Control and the IBM SmartCloud® Virtual Storage Center.

You can also obtain a no-charge unsupported version of the Quick performance overview (**qperf**) for the IBM SAN Volume Controller and Storwize systems from [this website](#).

3.2.5 Background Task

Use the Background Tasks pane to view and manage current tasks that are running on the system (see Figure 3-32 on page 107).

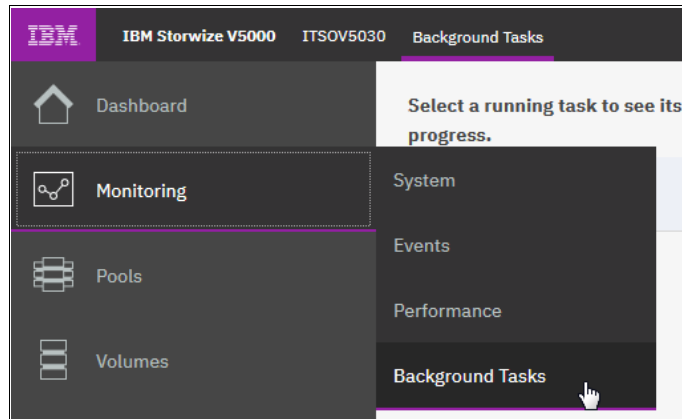


Figure 3-32 Selecting Background Tasks

The Background Tasks pane displays all long-running tasks that are in progress on the system. Tasks, such as volume synchronization, array initialization, and volume formatting, can take some time to complete. The Background Tasks page displays these tasks and their progress. After the task completes, the task is automatically deleted from the display. If a task fails with an error, select **Monitoring** → **Events** to determine the problem. Figure 3-33 shows an example of a FlashCopy Operation.

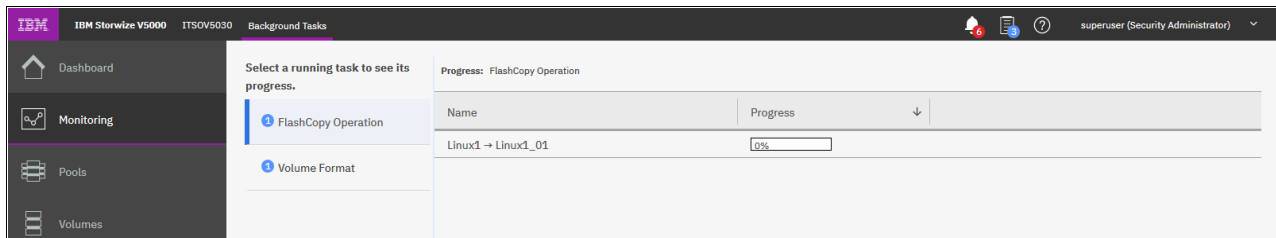


Figure 3-33 Background Task pane

3.3 Pools menu

A *pool* or storage pool is a collection of MDisks that jointly contain all of the data for a specified set of volumes. All MDisks in a pool are split into *extents* of the same size. The size of the extent defines the maximum usable capacity of an IBM Storwize Virtual System. An IBM Storwize Virtual System can address 2^{22} extents. This enables you to address a maximum of 32 PB with 8192 MB extents. Table 3-1 lists the extent sizes.

Table 3-1 Extents

Extent size (MB)	Maximum non thin-provisioned volume capacity in GB	Maximum thin-provisioned volume capacity in GB	Maximum compressed volume size	Maximum MDisk capacity in GB	Total storage capacity manageable per system*
16	2,048 (2 TB)	2,000	2 TB	2,048 (2 TB)	64 TB

Extent size (MB)	Maximum non thin-provisioned volume capacity in GB	Maximum thin-provisioned volume capacity in GB	Maximum compressed volume size	Maximum MDisk capacity in GB	Total storage capacity manageable per system*
32	4,096 (4 TB)	4,000	4 TB	4,096 (4 TB)	128 TB
64	8,192 (8 TB)	8,000	8 TB	8,192 (8 TB)	256 TB
128	16,384 (16 TB)	16,000	16 TB	16,384 (16 TB)	512 TB
256	32,768 (32 TB)	32,000	32 TB	32,768 (32 TB)	1 PB
512	65,536 (64 TB)	65,000	64 TB	65,536 (64 TB)	2 PB
1024	131,072 (128 TB)	130,000	96 TB**	131,072 (128 TB)	4 PB
2048	262,144 (256 TB)	260,000	96 TB**	262,144 (256 TB)	8 PB
4096	262,144 (256 TB)	262,144	96 TB**	524,288 (512 TB)	16 PB
8192	262,144 (256 TB)	262,144	96 TB**	1,048,576 (1024 TB)	32 PB

* The total capacity values assume that all of the storage pools in the system use the same extent size.

** Refer to the following shaded box.

Size limits: If you are using compressed volumes in standard pools, these volumes have the following size limits. If a new or existing compressed volume in a standard pool approaches the maximum size, the system issues an alert. Compressed volumes in data reduction pools do not monitor size of the volumes:

► 96 TB

Maximum virtual size of a new, individual compressed volume. You cannot create a compressed volume that exceeds this size. In addition, you cannot increase the size of a compressed volume beyond this value. If one or more compressed volumes in a system exceed this limit, you receive an alert. To reduce the risk of losing or corrupting data, you must take action soon to remove data from the compressed volume.

► 120 TB

Maximum virtual size of a compressed volume in a system. If any compressed volumes in the system approach or exceed this value, the system issues an alert.

Immediate action is required to remove all data from the compressed volume and prevent the loss of data.

► 128 TB

Maximum physical size of a compressed volume.

For information about how to move data off a compressed volume in a standard pool, view the topic on the [IBM Support website](#) for your product. Search for your product; then, select the Flashes, alerts, and bulletins link under Documents on the support page for your product.

Volumes are created from the extents that are available in the pool. You can add MDisks to a storage pool at any time to increase the number of extents that are available for new volume copies or to expand existing volume copies. Click the **Pools** icon to display the Pools menu options (see Figure 3-34).

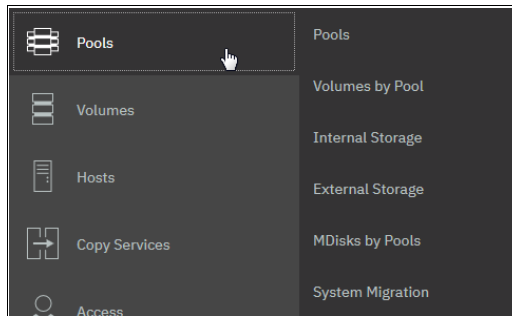


Figure 3-34 Navigating to the Pools menu

The Pools menu includes the following options:

- Pools** Shows a list of pools that are available within the system. It provides an option to create or modify pools and child pools and adding storage.
- Volumes by Pool** Applies the high-level filter, which lists all defined volumes per pool. It also provides a capacity overview, which is valid for a specific, selected pool only. This view is useful when you plan a migration of a volume to another pool so that you have a common overview of all pools and their associated volumes. Unused volumes are not listed.
- Internal Storage** Provides an overview of all disk drives that are installed in the Storwize V5000 system, including its enclosures. You can filter based on disk type and capacity and also see unused volumes that are not assigned to any pool.
- External Storage** Shows all pools and their volumes that are created from the systems that connect to the Storwize V5000 externally and that are integrated in the system repository. It does not show any internal pools or volumes. This type of storage is also called *external virtualization*.
- MDisks by Pools** Provides the list of all managed disks (MDisks) that are internally or externally connected and associated with one of the defined pools. It also lists all unassigned MDisks separately.
- System Migration** Offers the migration wizard to import data from image-mode MDisks to a specific pool. It is useful when you migrate data nondisruptively to the hosts from old external storage to the Storwize V5000.

3.3.1 Pools view

If you plan to add storage to a pool, use the main Pools view. Right-click a pool (or create a pool in advance and add the storage) and select **Add Storage**, as shown in Figure 3-35.

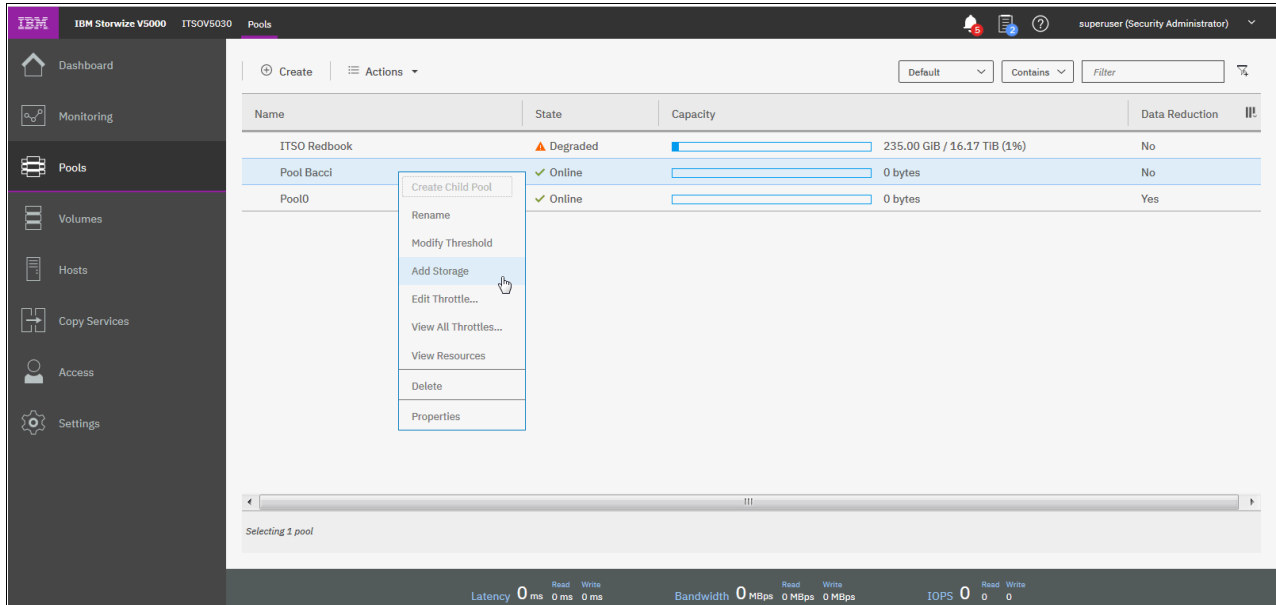


Figure 3-35 Add Storage option

In the next window, you can select **Internal** or **Internal custom**. Selecting Internal custom gives you the ability to select the RAID type, Drive class, and use of a spare drive (see Figure 3-36).

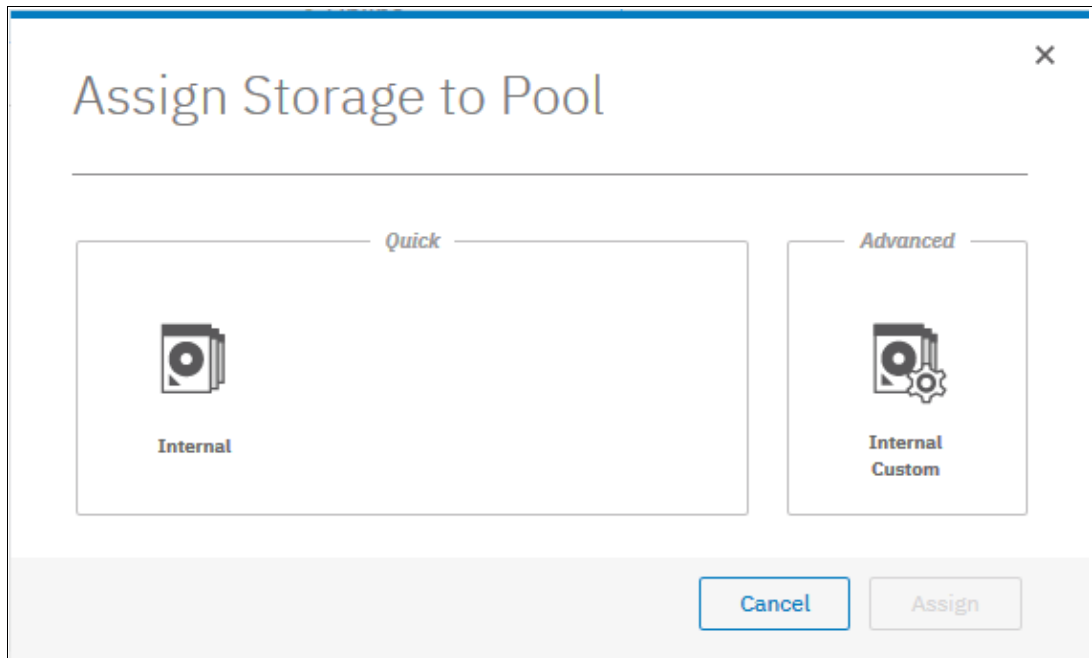


Figure 3-36 Internal storage selection

Figure 3-37 shows the window when you select **Internal**. Choose the default (Quick) settings to create an MDisk for the pool.

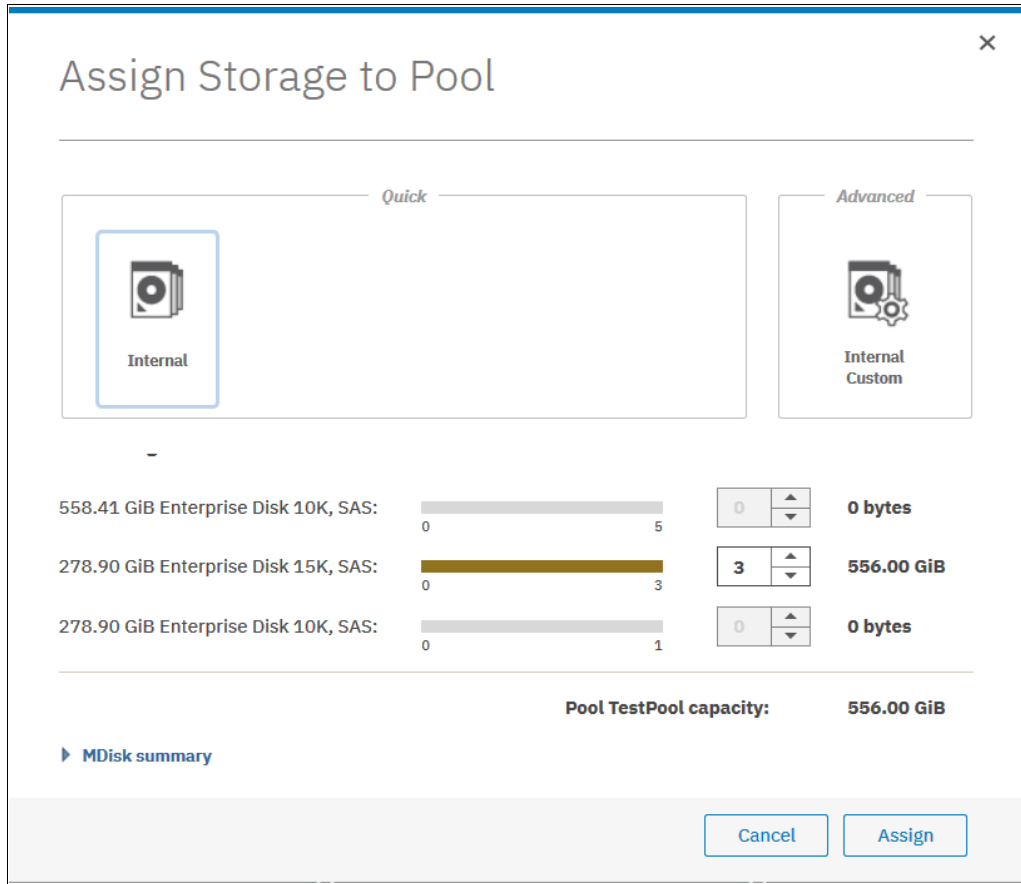


Figure 3-37 Internal storage selection

Figure 3-38 shows you the window if you selected the **Internal Custom storage** window.

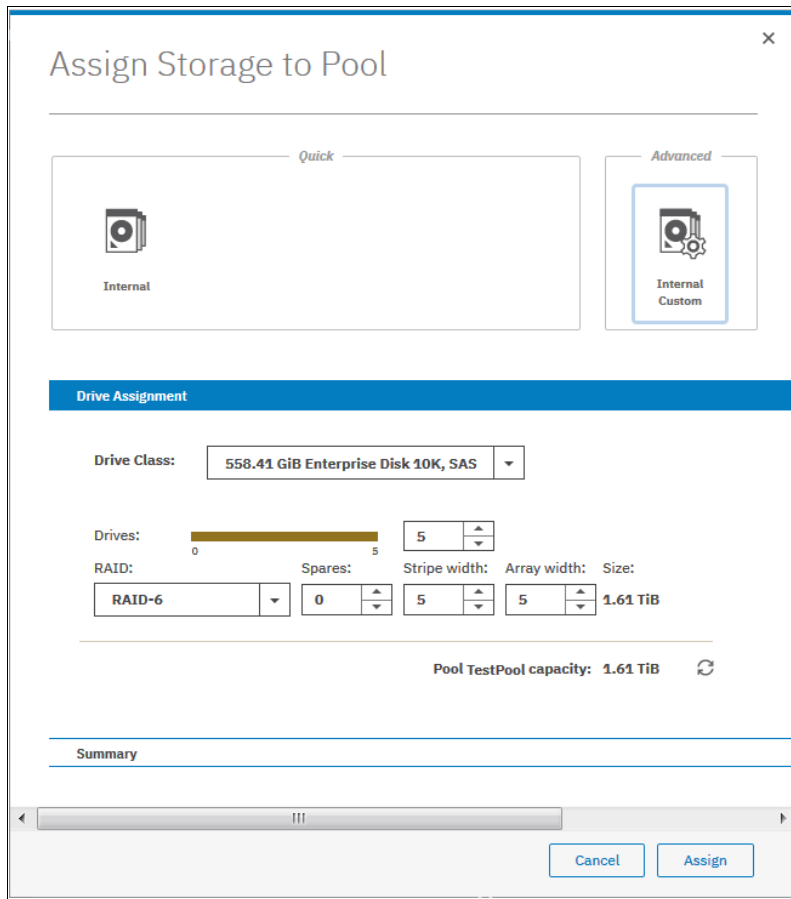


Figure 3-38 Internal Custom storage selection

You can choose from the available drive classes (depending on the installed drives) and RAID sets.

How to rename the pool is shown in Figure 3-39.

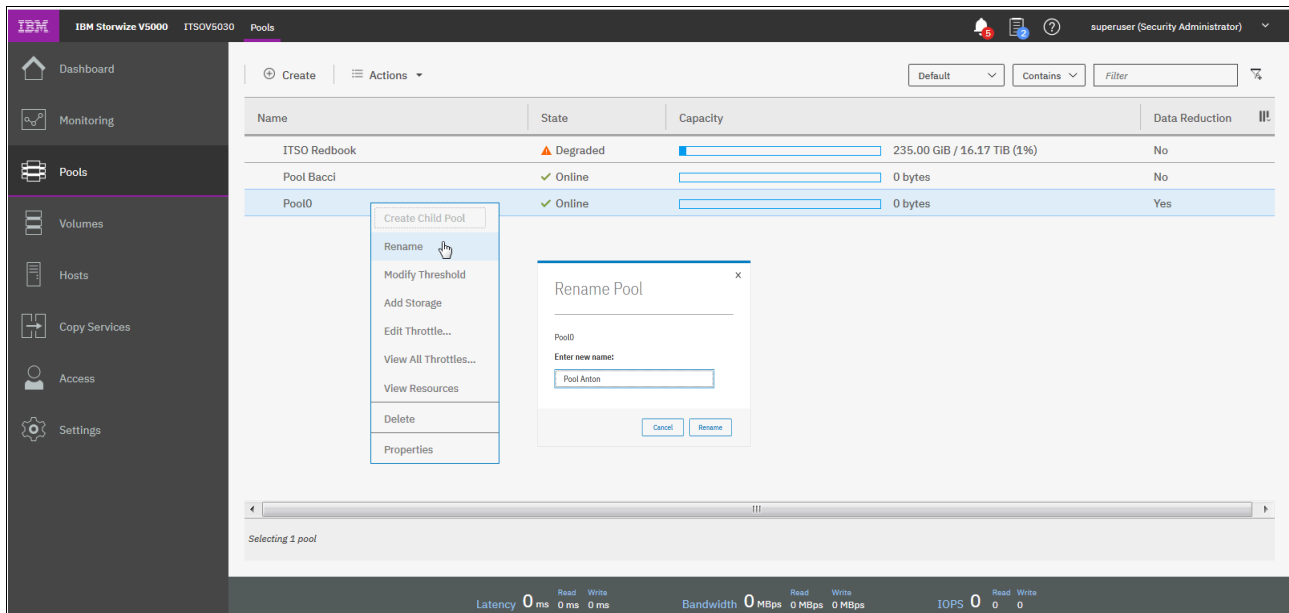


Figure 3-39 Renaming the pool

Throttles for pools

Throttling is a mechanism to control the amount of resources that are used when the system is processing I/Os on a specific pool. If a throttle is defined, the system either processes the I/O, or delays the processing of the I/O to free resources for more critical I/O.

The system also supports throttles to delay processing of I/O operations for pools. If storage systems provide storage to various applications, production pools with more critical I/O can be competing with pools that have lower priority operations.

For example, pools that are used for backup or archive operations can have I/O intensive workloads, which might take bandwidth from production pools. Pool throttles can be used to limit I/Os for these types of pools so that I/O operations for production pools are not affected.

Figure 3-40 shows an example of pool throttling.

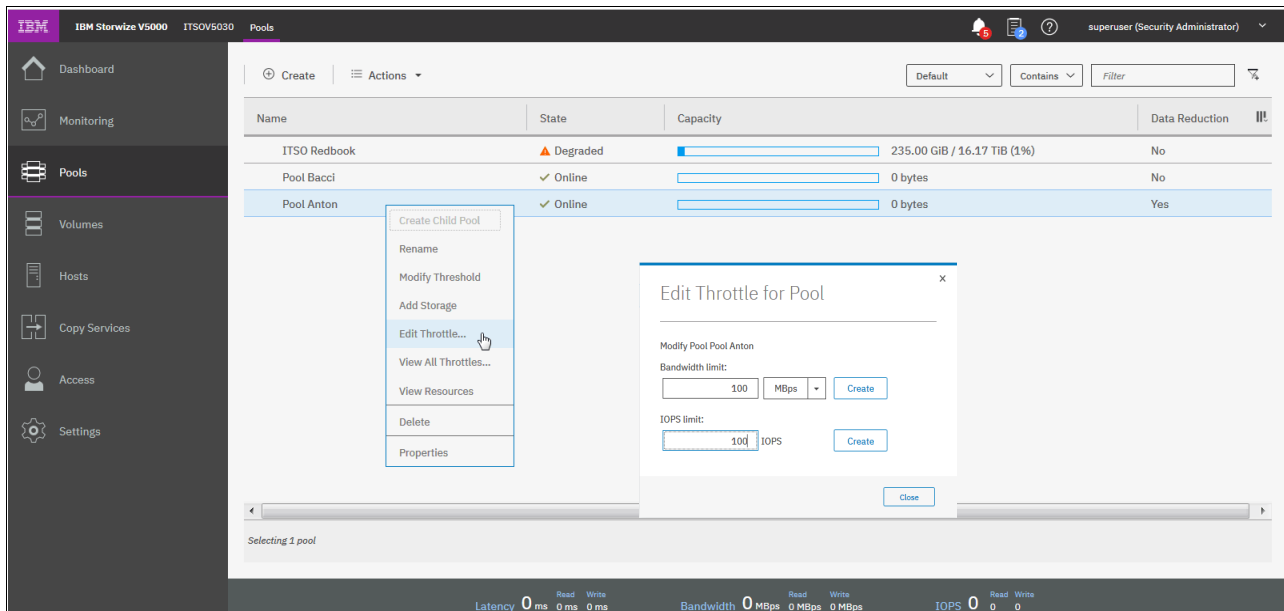


Figure 3-40 Pool throttling

Throttle limits are shown in Figure 3-41.

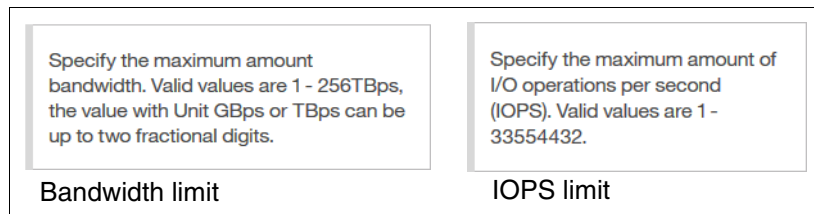


Figure 3-41 Throttling limits

3.3.2 Child pools

Before V7.4, the disk space of a storage pool was provided from MDisks, so the capacity of a storage pool depended on the MDisks' capacity. Creating or splitting a storage pool is impossible. A user cannot create a storage pool and specify the capacity that they want.

A *child pool* is a new logical object that is created from a physical storage pool. A child pool provides most of the functions that pools offer (for example, volume creation), but the user can specify the capacity of the child pool at creation. Administrators can use child pools to control capacity allocation for volumes that are used for specific purposes.

Several administration tasks benefit from being able to define and work with a part of a pool. For example, the system supports VMware vSphere Virtual Volumes (sometimes referred to as *VVols*) that are used in VMware vCenter and VASA applications. Before a child pool can be used for Virtual Volumes for these applications, the system must be enabled for Virtual Volumes.

A child pool is an object that is similar to a storage pool, and can be used interchangeably with a storage pool. A child pool supports volume copy and migration. However, child pools feature the following limitations and restrictions:

- ▶ The maximum capacity cannot exceed the parent pool's size.
- ▶ The capacity can be allocated at creation (thick) or flexible (thin).
- ▶ You must always specify the parent storage pool. The child pool does not own any MDisks.
- ▶ Child pools can be created by using the GUI.
- ▶ The maximum number of child pools is 1023.
- ▶ You are restricted to migrating image-mode volumes to a child pool.
- ▶ Volume extents cannot be migrated out of the child pool.
- ▶ You cannot shrink capacity smaller than the real capacity.

Note: Child pools are not supported in a Data Reduction Pool.

You can view the list of child pools from the Pools menu option by clicking the Plus sign (+) of a parent pool, as shown in Figure 3-42.

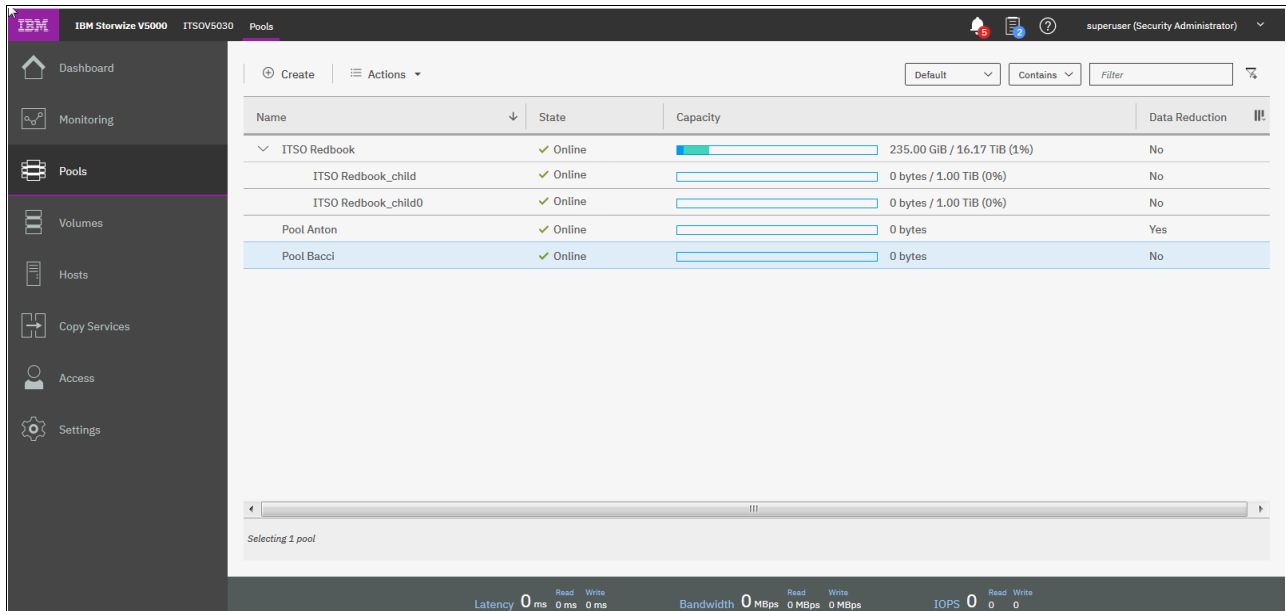


Figure 3-42 Working with child pools

3.3.3 Volumes by pool

The Volumes by Pool menu option lists all defined volumes, which are sorted by their pool assignment (see Figure 3-43 on page 116). Unassigned volumes are not visible in this window. By using this window, you can, for example, create volumes or maps, or unmap volumes to and from hosts, migrate volumes to another pool, and rename, shrink, or expand volumes.

In addition, you can choose a different icon (see Figure 3-43) that represents this pool.

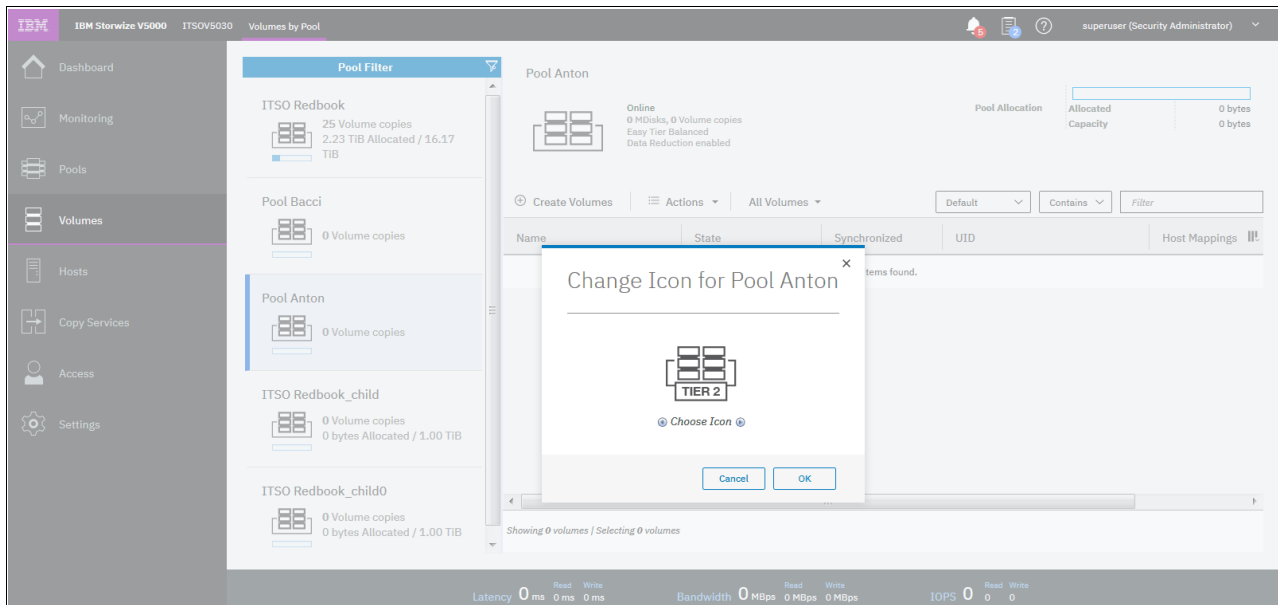


Figure 3-43 Volume by Pool option and changing the icon

To change the icon, click the pen icon, as shown in Figure 3-44.

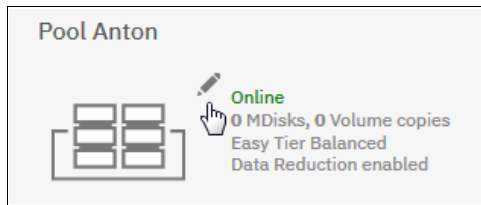


Figure 3-44 Pen sign to change icon

When the pools are defined and the volumes are assigned, the pool shows one of the following operational states:

- Online** The storage pool is online and available. All of the MDisks in the storage pool are available.
- Degraded path** One or more nodes in the clustered system cannot access all of the MDisks in the pool. A degraded path state is most likely the result of the incorrect configuration of either the storage system or the FC fabric. However, hardware failures in the storage system, FC fabric, or a node can also be a contributing factor to this state.
- Degraded ports** One or more 1220 errors were logged against the MDisks in the storage pool. The 1220 error indicates that the remote FC port was excluded from the MDisk. This error might cause reduced performance on the storage system and usually indicates a hardware problem with the storage system.

To fix this problem, you must resolve any hardware problems on the storage system and fix the 1220 errors in the Event log. To resolve these errors in the log, select **Monitoring** → **Events** → **Recommended Actions** → **Run Fix** in the management GUI.

This action displays a list of unfixed errors that are in the Event log. For these unfixed errors, select the error name to begin a guided maintenance procedure to resolve the errors. Errors are listed in descending order with the highest priority error listed first. Resolve the highest priority errors first.

Offline

The storage pool is offline and unavailable. No nodes in the system can access the MDisks. The most likely cause is that one or more MDisks are offline or excluded.

Important: In this view, volumes from child pools are shown the same way as volumes from standard pools. The relationships between the child and parent pools are not visible.

3.3.4 Internal storage

Click the **Internal Storage** option in the Pools menu to open a window that is similar to the window that is shown in Figure 3-45. From this window, you can allocate Redundant Array of Independent Disks (RAID) arrays of internal disk drives into storage pools. This window also offers the option to display internal drives, based on their capacity and speed.

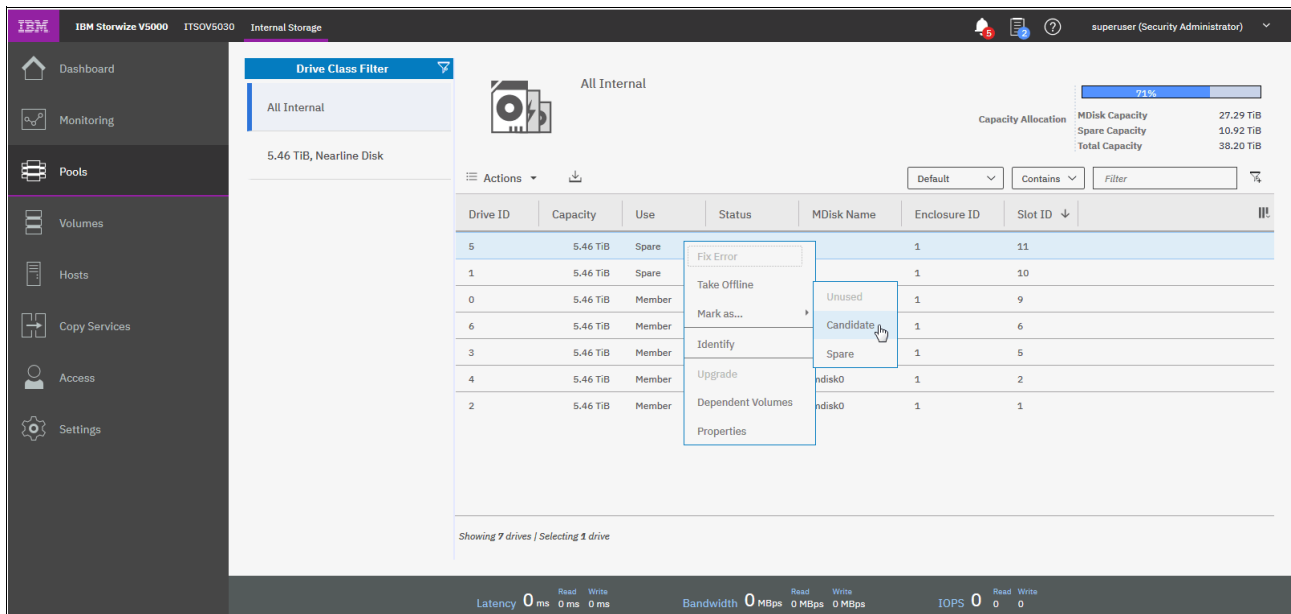


Figure 3-45 Internal storage window

Click **Actions** in the table header, or right-click a specific drive to take the drive offline, show its properties, update firmware of a single drive (or all under Actions), or mark the drive as Unused, Candidate, or Spare.

3.3.5 External storage

Before you can add external storage, your system must be in the Replication Layer, which can be done by using the CLI, as shown in the following command:

```
chsystem -layer replication -cacheprefetch on/off
```

If you must switch it back to storage layer, use following command:

```
chsystem -layer storage -cacheprefetch on/off
```

If you search for external storage and you are not in the replication layer, you receive the warning in the GUI that is shown in Figure 3-46.

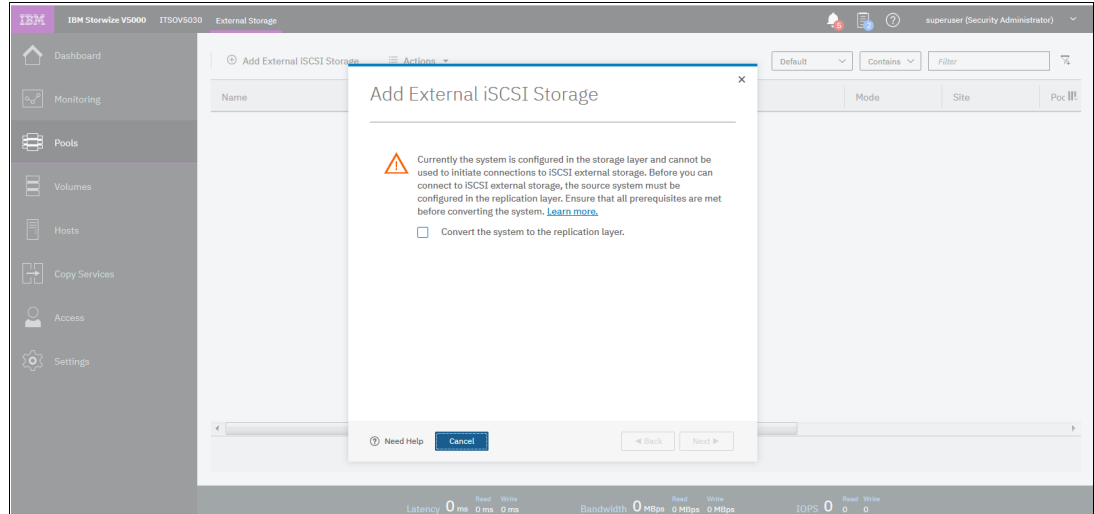


Figure 3-46 Convert System to replication layer

Clicking the **External Storage** option opens the window that is shown in Figure 3-47.

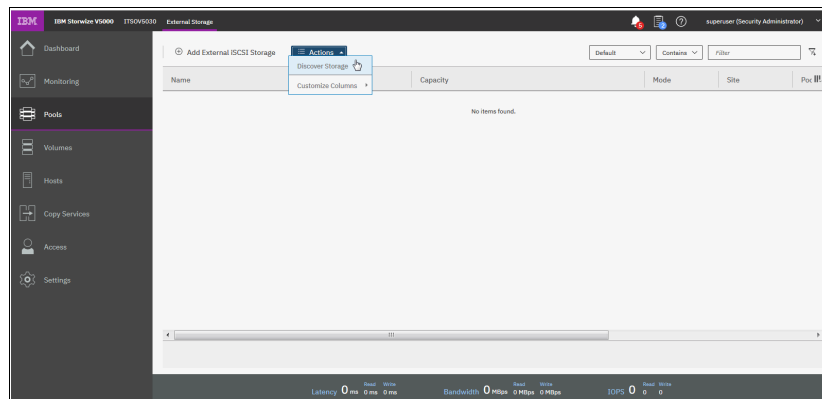


Figure 3-47 Detecting external storage systems

This window provides a list of all externally connected storage area network (SAN) and iSCSI-attached disk systems to the Storwize V5000. The system also supports iSCSI connections to systems that are used as external storage.

Unlike Fibre Channel connections, you must manually configure the connections between the source system and these target external storage systems. Direct attachment between the system and external storage systems is not supported and requires Ethernet switches between the system and the external storage.

To avoid a single point of failure, a dual switch configuration is suggested. For full redundancy, a minimum of two paths between each initiator node and target node must be configured with each path on a separate switch. In addition, extra paths can be configured to increase throughput if both initiator and target nodes support more ports. The system supports a maximum of four paths per node.

When the new external storage system is zoned correctly to the Storwize V5000, run the Discover storage procedure from the Actions menu in the table header or by right-clicking any of the MDisks in the list (see Figure 3-47 on page 118).

A new storage controller (external storage system) is listed automatically when the SAN zoning is configured, but typically without detecting disk drives (see Figure 3-48).

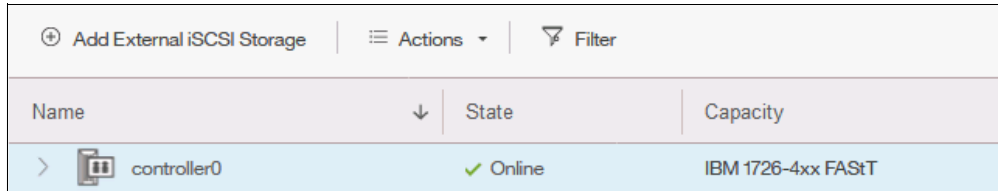


Figure 3-48 Automatically detected new external storage system

By right-clicking a newly detected storage system, you can rename the controller’s default name (see Figure 3-49), in our case, controller0, to reflect the real type of the storage device.

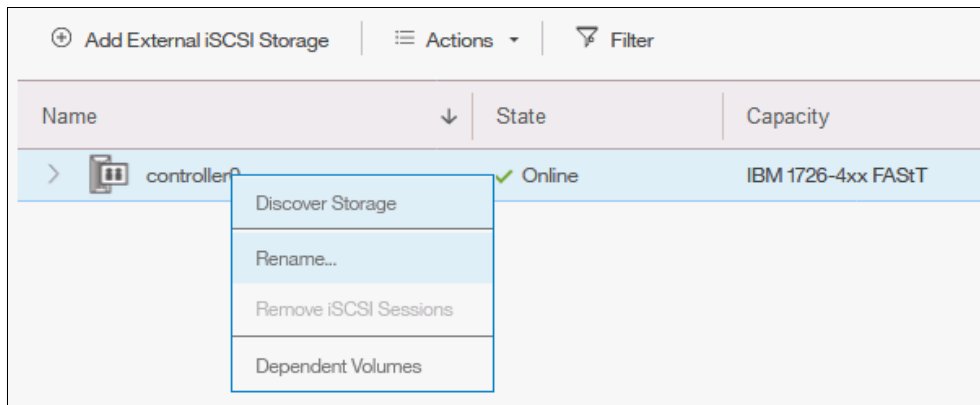


Figure 3-49 Select Rename of the detected system

We suggest that you use a simple naming convention, which in our case is DS3524 (see Figure 3-50).

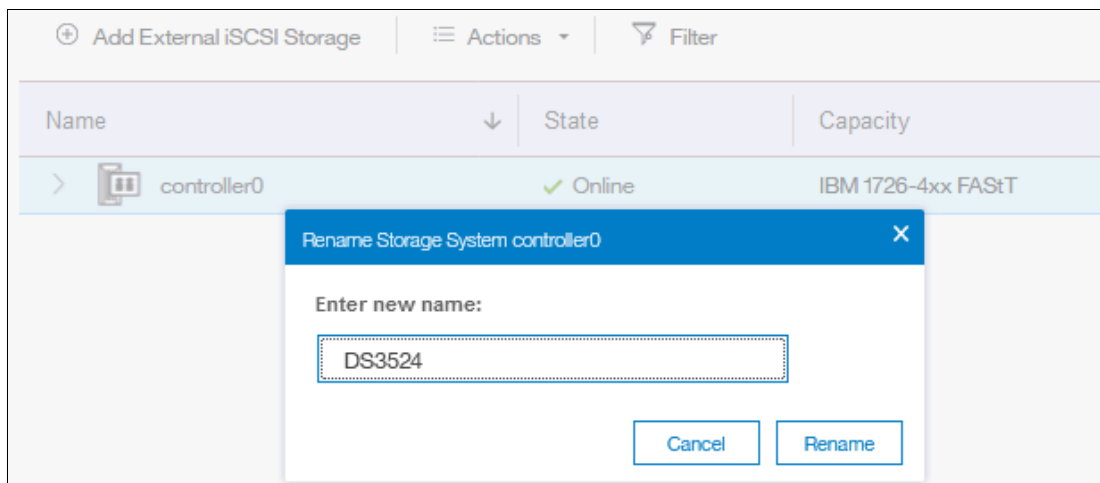


Figure 3-50 Renaming the detected Controller

After the new external storage system is renamed, detect all disks that are configured on that external storage; in our case, IBM DS3524. You can also discover storage from the CLI by running the `svctask detectmdisk` or `detectmdisk` command.

Figure 3-51 shows more information about the detected managed disks.


⊕ Add External iSCSI Storage		☰ Actions ▾	🔍 Filter			
Name	↓	State	Capacity	Mode	Storage System	
∨  DS3524		✓ Online	IBM 1726-4xx FAStT			
mdisk0		✓ Online	64.00 GiB	Unmanaged	DS3524	
mdisk1		✓ Online	64.00 GiB	Unmanaged	DS3524	
mdisk2		✓ Online	64.00 GiB	Unmanaged	DS3524	
mdisk3		✓ Online	64.00 GiB	Unmanaged	DS3524	
mdisk4		✓ Online	10.00 GiB	Unmanaged	DS3524	
mdisk5		✓ Online	32.00 GiB	Unmanaged	DS3524	
mdisk6		✓ Online	32.00 GiB	Unmanaged	DS3524	
mdisk7		✓ Online	32.00 GiB	Unmanaged	DS3524	
mdisk8		✓ Online	32.00 GiB	Unmanaged	DS3524	

Figure 3-51 Newly discovered managed disks

All newly discovered disks are always interpreted in an *unmanaged* mode. You must assign them to the specific pool to be able to operate them.

Important: The MDisks are not physical disk drives; instead, they are storage arrays that are configured on external systems.

If you add a managed disk that contains data to a managed disk group, you lose the data that it contains. The *image* mode is the only mode that preserves its data.

3.3.6 MDisks by pools

This option on the Pools menu provides the list of all managed disks and arrays of disks, internally or externally connected, and associated with one of the defined pools. It also lists all unassigned MDisks (which are provided only by external storage systems) separately, as shown in Figure 3-52.

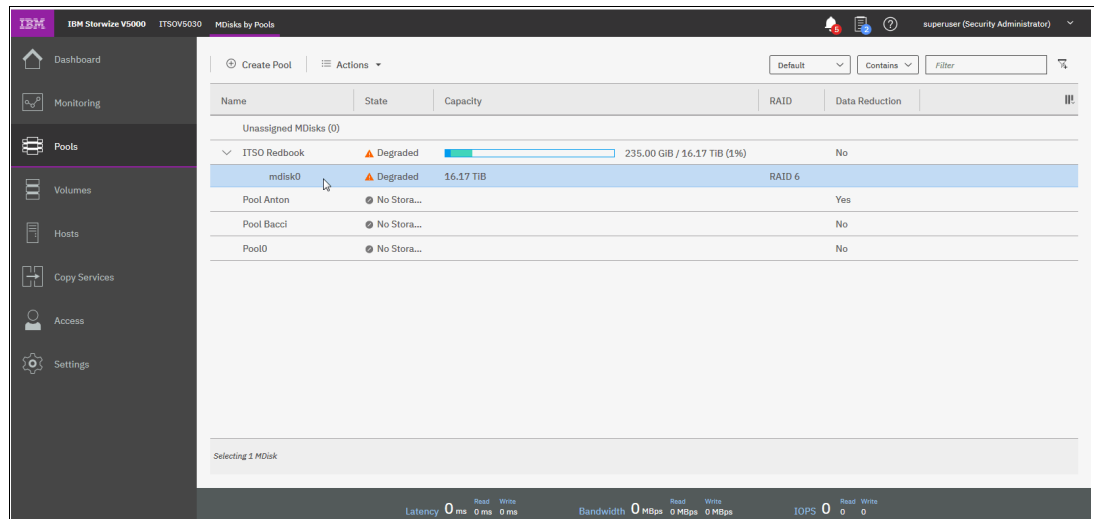


Figure 3-52 List of managed disks that are sorted within pools

All disks that are not yet assigned to any pool are listed in the Unassigned MDisks section. This section is always at the top of the list, even if you sort the list by pool name (click the **Name** header of the table). Right-click a specific disk to open a window where you can assign selected unmanaged disks to the pool.

From the same pane, you can define a new storage pool by clicking **Create Pool** in the upper-left corner of the table (see Figure 3-53).

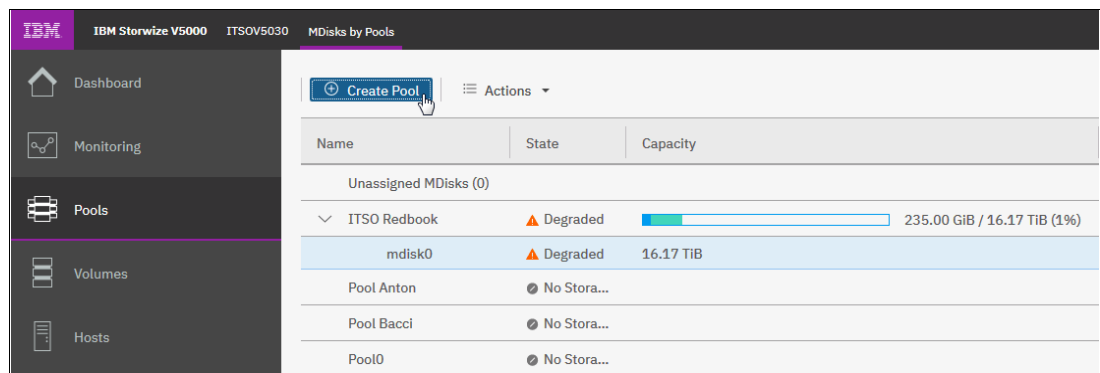


Figure 3-53 Create Pool option

The wizard window opens and you must specify the pool parameters, such as Pool Name, Extent Size, and Warning Threshold. You can directly select Unmanaged MDisks that you want to include in the pool, or skip this task and add MDisks later.

Note: All sort functions in the header of the table apply to MDisks *within* pools. You cannot sort volumes based on specific criteria *across* all pools.

3.3.7 System migration

Migrating data from older storage systems to the Storwize V5000 storage system enables applications to benefit from the new features, such as IBM Easy Tier, Space Efficient volumes, an intuitive management GUI, and advanced storage replication functions that better support applications.

To migrate data, use the IBM Spectrum Virtualize storage migration wizard to guide you through the procedure. This wizard is available by selecting **Pools** → **System Migration**, as shown in Figure 3-54.

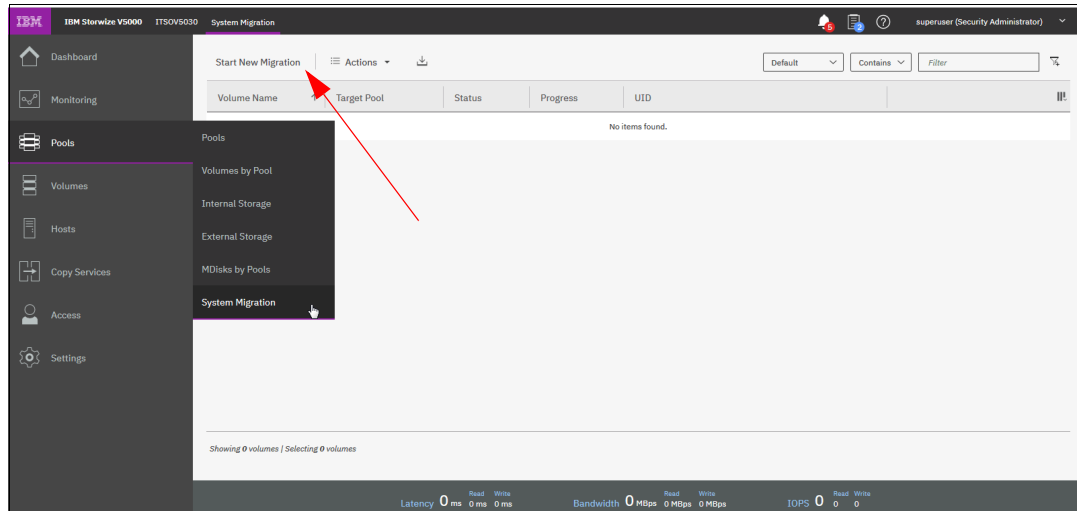


Figure 3-54 System migration

Migrating external volumes to the IBM Storwize V5000 system is one of the key benefits and features of external storage virtualization that are provided by this product. For more information about the migration process, see Chapter 7, “Storage migration” on page 357.

Administrators can migrate data from the external storage system to the system that uses either iSCSI connections, serial-attached SCSI connections, and Fibre Channel or Fibre Channel over Ethernet connections. To use Fibre Channel connections, the system must have the optional Fibre Channel host interface adapter installed.

Note: Before migrating storage, ensure that all host operations are stopped, all the appropriate changes are made to the environment based on the connection type, and the storage that is being migrated is configured to use the device.

At any time, you can pause the running migration processes or create one. No license for External Virtualization is required to migrate from old storage to your new IBM Storwize V5000.

3.4 Volumes menu

A *volume* is a logical disk that the system presents to the attached host. Application servers access volumes, not MDisks or drives. Volumes can be automatically expanded, mirrored, or pre-allocated. Volumes include the following characteristics and differences:

Basic	A basic (<i>fully allocated</i>) volume is the traditional data store method when any host input/output (I/O) is destaged to the drives. Even zeros are destaged. All of the zeros that exist are written.
Mirrored	By using volume mirroring, a volume can have two physical copies. Each volume copy can belong to a different pool, and each copy has the same virtual capacity as the volume. In the management GUI, an asterisk (*) indicates the primary copy of the mirrored volume. The primary copy indicates the preferred volume for read requests.
HyperSwap	HyperSwap volumes create copies on separate sites for systems that are configured with HyperSwap topology. Data that is written to a HyperSwap volume is automatically sent to both copies so that either site can provide access to the volume if the other site becomes unavailable. HyperSwap volumes are supported on Storwize systems (for example, Storwize V5030 or Storwize V5030F systems) that contain more than one I/O group.
Custom	Custom volumes create volumes that are based on user-defined customization rather than taking the standard default settings for each of the options under quick volume creation.
Fully allocated	A fully allocated volume contains virtual capacity and real capacity, which are set when you create the volume.
Thin-provisioned	<p>When you create a volume, you can designate it as thin-provisioned. A thin-provisioned volume has a virtual capacity and a real capacity. <i>Virtual capacity</i> is the volume storage capacity that is available to a host. <i>Real capacity</i> is the storage capacity that is allocated to a volume copy from a storage pool.</p> <p>In a fully allocated volume, the virtual capacity and real capacity are the same. In a thin-provisioned volume, the virtual capacity can be much larger than the real capacity.</p>
Deduplicated	Deduplication can be configured with thin-provisioned and compressed volumes in data reduction pools for added capacity savings. Deduplication is a type of data reduction that eliminates duplicate copies of data.
Compressed	In this special type of volume, data is compressed and thin-provisioned at the same time. Any compressed volume is a thin-provisioned volume by default, and no option is available to change this characteristic. Data within the compressed volume is compressed as it is written to disk. This design saves more space on the storage drive so that you can store more data within the same storage system.
Change volumes	Change volumes are used in Global Mirror relationships where cycling mode is set to Multiple. Change volumes can also be used between HyperSwap volume copies, and other relationship types, to automatically maintain a consistent image of a secondary volume when a relationship is being resynchronized.

Change volumes create periodic point-in-time-copies of the source volumes and replicate them to the secondary site. The use of change volumes lowers bandwidth requirements by addressing only the average throughput and not the peak.

Virtual

The system supports VMware vSphere Virtual Volumes, sometimes referred to as VVols, which allow VMware vCenter to automate the management of system objects like volumes and pools.

Protected

To prevent active volumes or host mappings from being deleted inadvertently, the system supports a global setting that prevents these objects from being deleted if the system detects recent I/O activity.

Important: Compression + Deduplication is available in the IBM Storwize V5030 only and requires 64 GB of RAM. To use these functions, you must obtain the IBM Real-time Compression license.

To keep a volume accessible even when an MDisk on which it depends is unavailable, a mirrored copy can be added to a selected volume. Any volume (generic, thin-provisioned, or compressed) can be mirrored with a mirror from any type, even the same one. Therefore, a volume can be thin-provisioned with compressed copy or compressed with compressed copy. Each volume can have a maximum of two copies.

Each volume copy is created from a set of extents in a storage pool. By using volume mirroring, a volume can have two physical copies. Each volume copy can belong to a different storage pool, and each copy has the same virtual capacity as the volume. In the management GUI, an asterisk (*) indicates the primary copy of the mirrored volume. The primary copy indicates the preferred volume for read requests.

Select the **Volumes** function icon to display the Volumes menu options (see Figure 3-55).

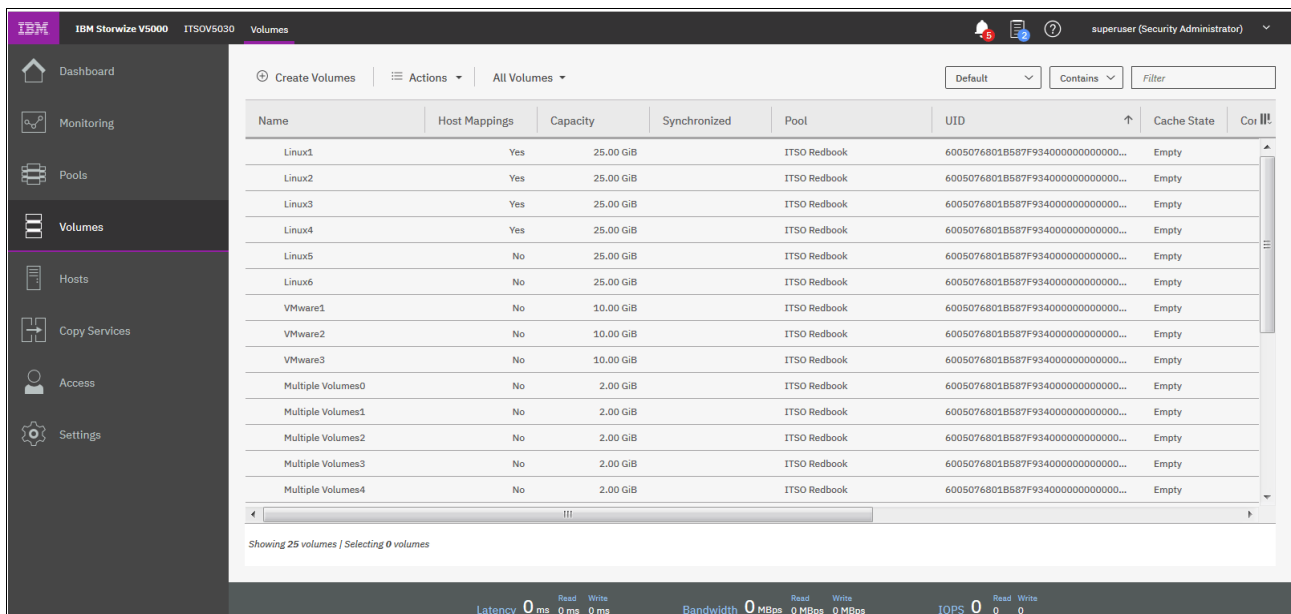


Figure 3-55 Volumes menu

3.4.1 All volumes

Select **Volumes** (see Figure 3-55) to see a list of all defined volumes, which is alphabetically sorted by the volume name (by default). At any time, you can change the sort options by clicking a specific header in the table. You can directly configure a new volume by clicking **Create Volumes**, as shown in Figure 3-56.



Figure 3-56 Create a volume

The wizard opens and the list of volume options is displayed (see Figure 3-57).

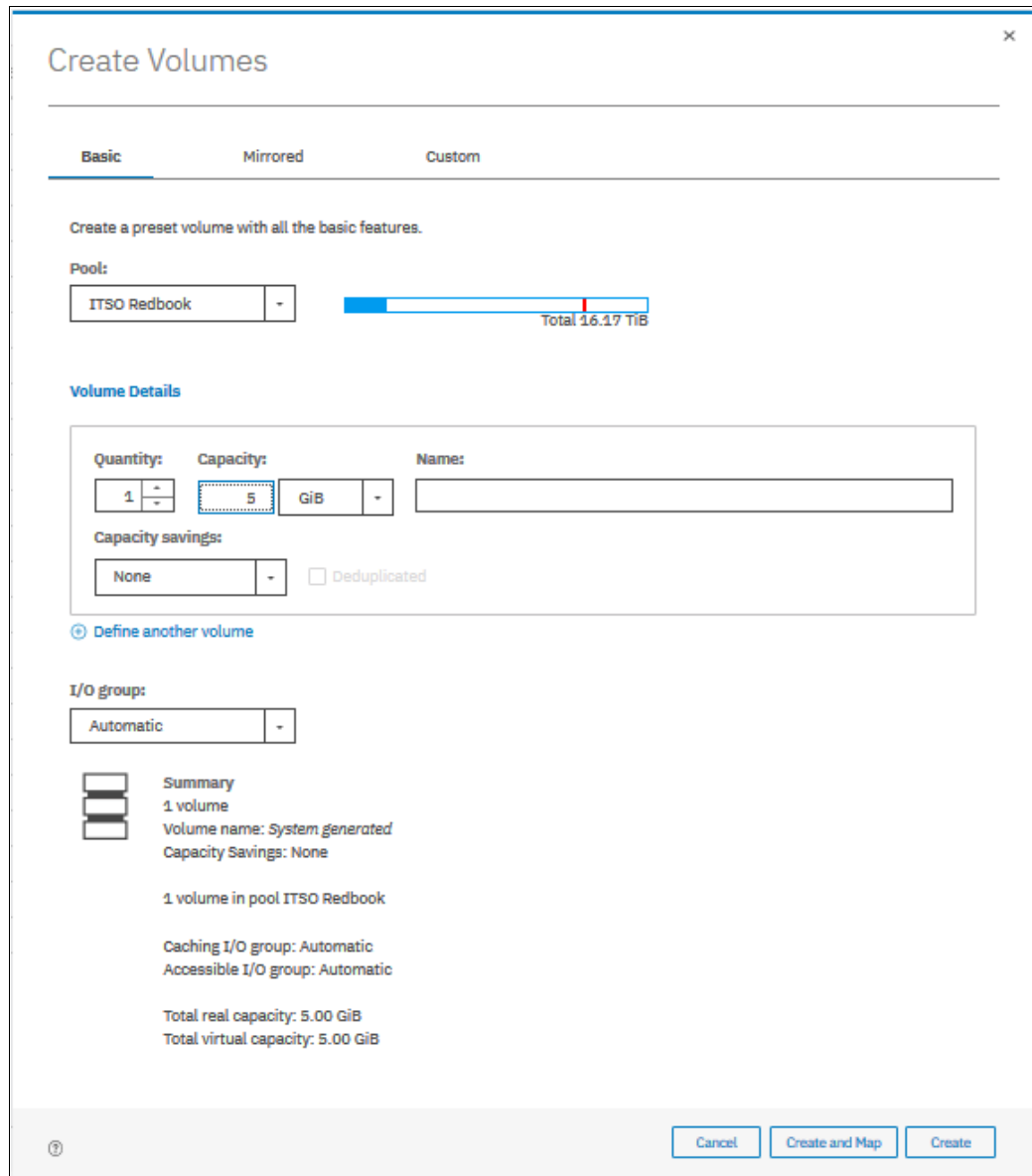


Figure 3-57 Create Volumes wizard

For more information about each type of volume and the procedures for how to effectively create these volumes, see Chapter 6, “Volume configuration” on page 309.

In addition to the volume creation, the following direct volume functions are available:

- ▶ Mapping and unmapping volumes to hosts
- ▶ Renaming, shrinking, or expanding existing volumes
- ▶ Modify Mirror Synchronization Rate
- ▶ Space savings → Estimate compression savings
- ▶ Migrating to a different pool
- ▶ Defining a volume copy

3.4.2 Volumes by pool

This menu is identical to the one that is described in 3.3.3, “Volumes by pool” on page 115.

3.4.3 Volumes by host

Click **Volumes by Host** to open the window that is shown in Figure 3-59. This window shows the volumes that are mapped to a certain host. You can perform the same actions with volumes as in all previous views by clicking **Actions** or by using the menu that opens after you right-click a specific volume (see Figure 3-59).

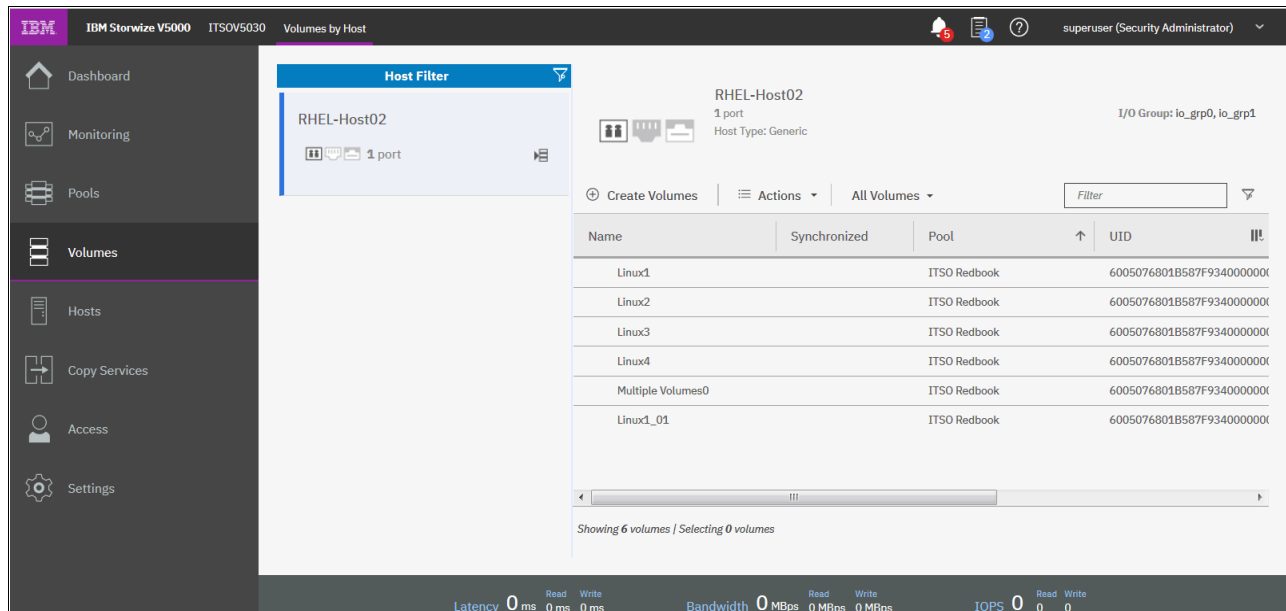


Figure 3-59 Listing volumes by host

3.5 Hosts menu

In a SAN environment, a host system is a computer that is connected to the system through one of the following components:

- ▶ Fibre Channel interface
- ▶ Serial-attached SCSI (SAS) connections
- ▶ IP network

To use Fibre Channel or Fibre Channel over Ethernet connections to a SAN, an optional host interface adapter must be installed. You can use several tools to manage hosts, including the management GUI, CLI, and specialized utilities for working with host bus adapters (HBAs).

To work with hosts in the management GUI, select **Hosts**. When you click the **Host function** icon, the Hosts menu opens, which provides the following options (see Figure 3-60 on page 129):

- ▶ Hosts
- ▶ Host Clusters
- ▶ Ports by Host
- ▶ Host Mappings
- ▶ Volumes by Host

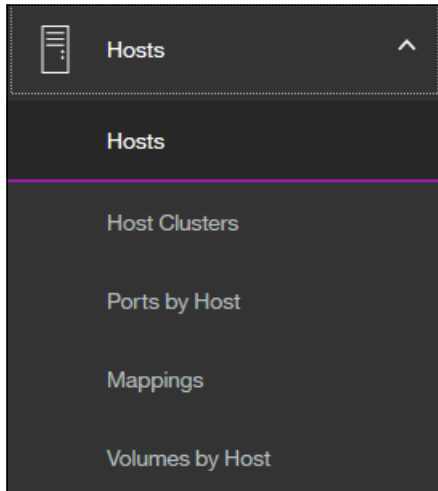


Figure 3-60 Hosts menu

3.5.1 Hosts

This option provides an overview about all hosts that are connected (zoned) to the system, which is detected and configured to be ready for storage allocation. This overview shows the following information about the hosts:

- ▶ The name of the host as defined in IBM Spectrum Virtualize
- ▶ The type of the host
- ▶ Its access status
- ▶ The number of ports that is used for host mapping
- ▶ Whether host mapping is active

From the same pane, you can create, rename, or delete a host, or modify a host mapping. The output of the menu selection is shown in Figure 3-61.

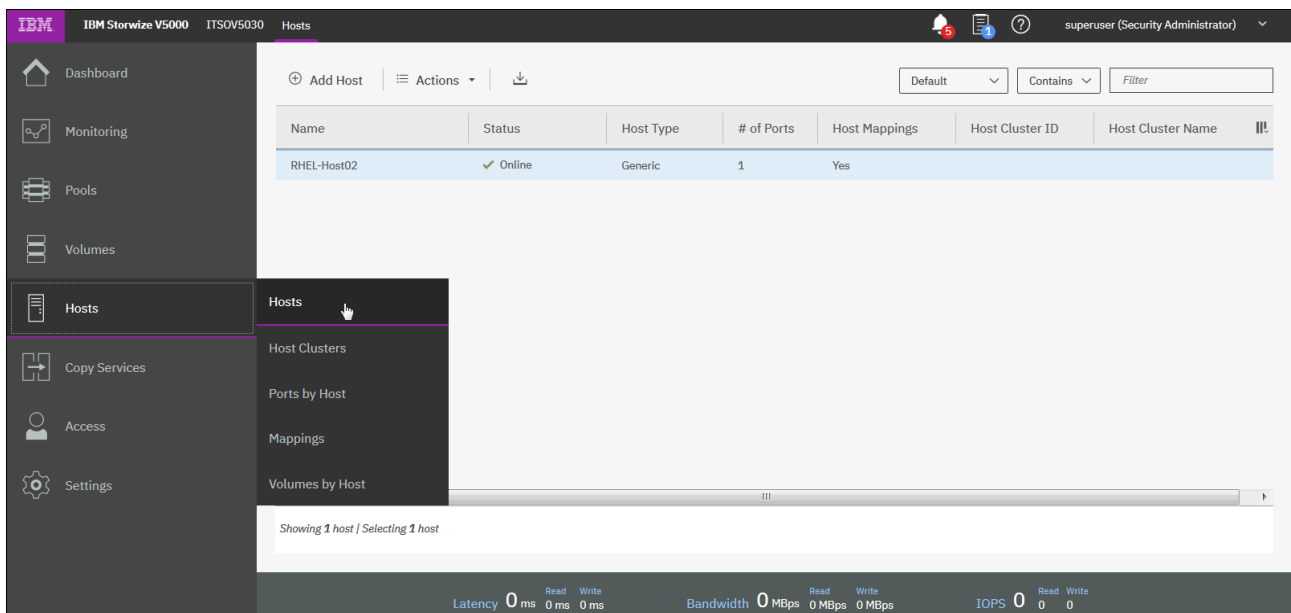


Figure 3-61 Overview of configured hosts

For example, when you click **Add Host** in a pane header, a wizard opens in which you define a Fibre Channel host or an iSCSI host (see Figure 3-62).

Add Host

Required Fields

Name:

Host connections:

Host port (WWPN):

Optional Fields

Host type:

I/O groups:

Host cluster:

Figure 3-62 Add Host wizard

To rename multiple hosts in a single step, mark all hosts that you want by using the Ctrl or Shift key, right-click, and then from the opened menu, select **Rename**. The window that is shown in Figure 3-63 opens.

Rename Hosts

**New Name*

RHEL-Host02

Lenovo

Linux

Figure 3-63 Renaming multiple hosts

Many of the actions that are described are available from different menus. For example, you can select **Volumes** and its option **Volumes by Hosts**, where you can also rename hosts. This flexibility is one of the advantages of the enhanced, redesigned IBM Spectrum Virtualize management GUI.

3.5.2 Host clusters

A host cluster is a group of logical host objects that can be managed together. For example, you can create a volume mapping that is shared by every host in the host cluster (see Figure 3-64).

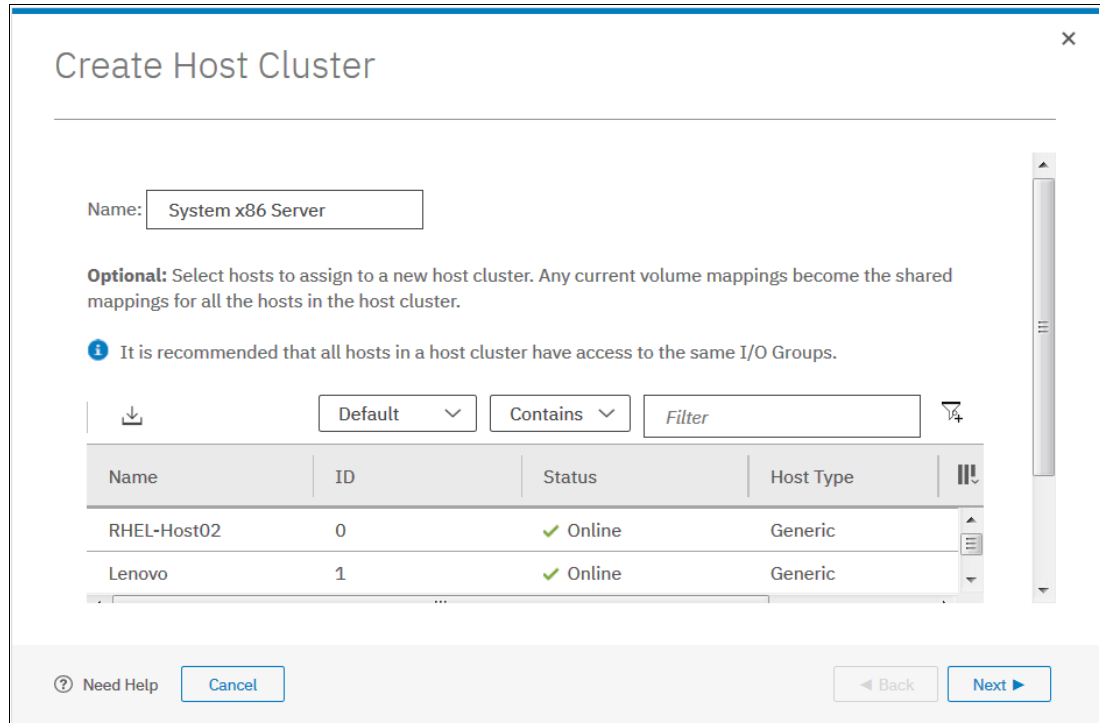


Figure 3-64 Create Host Cluster window

The systems use internal protocols to manage access to the volumes and ensure consistency of the data. Host objects that represent hosts can be grouped in a host cluster and share access to volumes. New volumes can also be mapped to a host cluster, which simultaneously maps that volume to all hosts that are defined in the host cluster. Each host cluster is identified by a unique name and ID, the names of the individual host objects within the cluster, and the status of the cluster.

A host cluster can contain up to 128 hosts. However, a host can be a member of only one host cluster. The management GUI displays the status of each host cluster.

A host cluster can have one of the following statuses:

- ▶ Online: All hosts in the host cluster are online.
- ▶ Host degraded: All hosts in the host cluster are online or degraded.
- ▶ Host cluster degraded: At least one host is offline and at least one host is online or degraded.
- ▶ Offline: All hosts in the host cluster are offline (or the host cluster does not contain any hosts).

By default, hosts within a host cluster inherit all shared volume mappings from that host cluster, as though those volumes were mapped to each host in the host cluster individually. Hosts in a host cluster can also have their own private volume mappings that are not shared with other hosts in the host cluster. With shared mapping, volumes are mapped on a host cluster basis. The volumes are shared by all of the hosts in the host cluster, if there are no Small Computer System Interface (SCSI) LUN conflicts among the hosts.

Volumes that contain data that is needed by other hosts are examples of a shared mapping. If a SCSI LUN conflict occurs, a shared mapping is not created. SCSI LUN conflicts can occur if multiple volumes are mapped with the same SCSI LUN ID or if same volume is mapped to multiple SCSI LUN IDs. The system does not allow a volume to be mapped more than once to the same host. With private mapping, individual volumes are directly mapped to individual hosts.

These volumes are not shared with any other hosts in the host cluster. A host can maintain the private mapping of some volumes and share other volumes with hosts in the host cluster. The SAN boot volume for a host typically is a private mapping.

3.5.3 Ports by host

Click **Ports by Hosts** to open the pane that is shown in Figure 3-65. This pane lists the Fibre Channel and iSCSI ports that are assigned to a particular host.

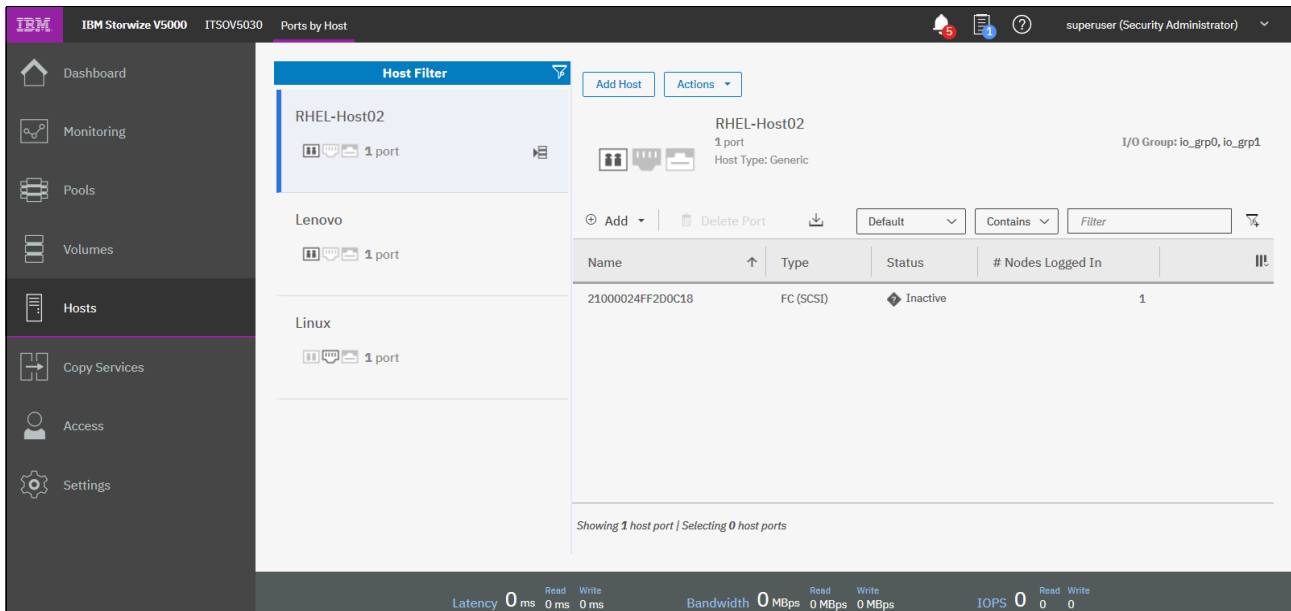


Figure 3-65 Ports by Host window

This overview shows hosts with active, inactive, or degraded ports. You can delete or add a port, or modify its characteristics. Also, you can create a host or rename the existing pane in this pane.

To perform any of the tasks that are shown in Figure 3-66, click **Actions** and select a menu item.

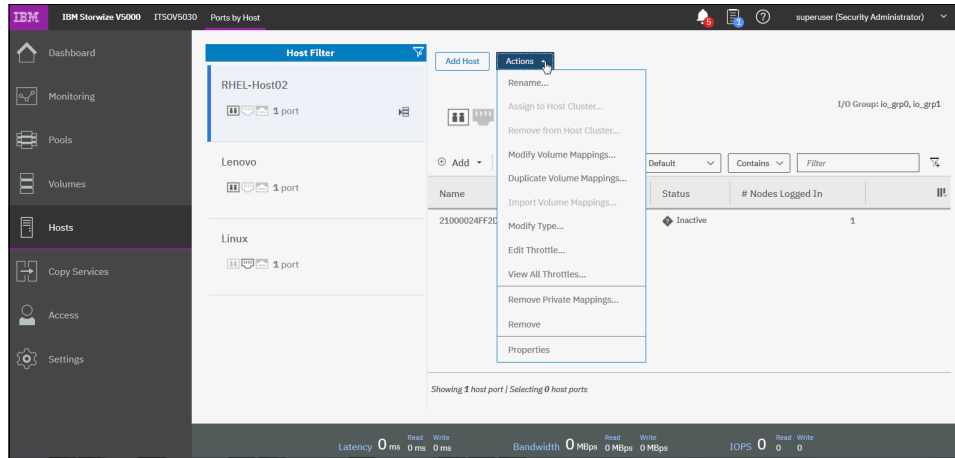


Figure 3-66 Available port actions

To delete multiple ports, select them by using the Ctrl key or Shift key and click **Delete**.

3.5.4 Host mappings

Click **Host Mappings** to open the window that is shown in Figure 3-67. It lists the host name, SCSI identifier, volume name, and volume identifier for all mapped volumes.

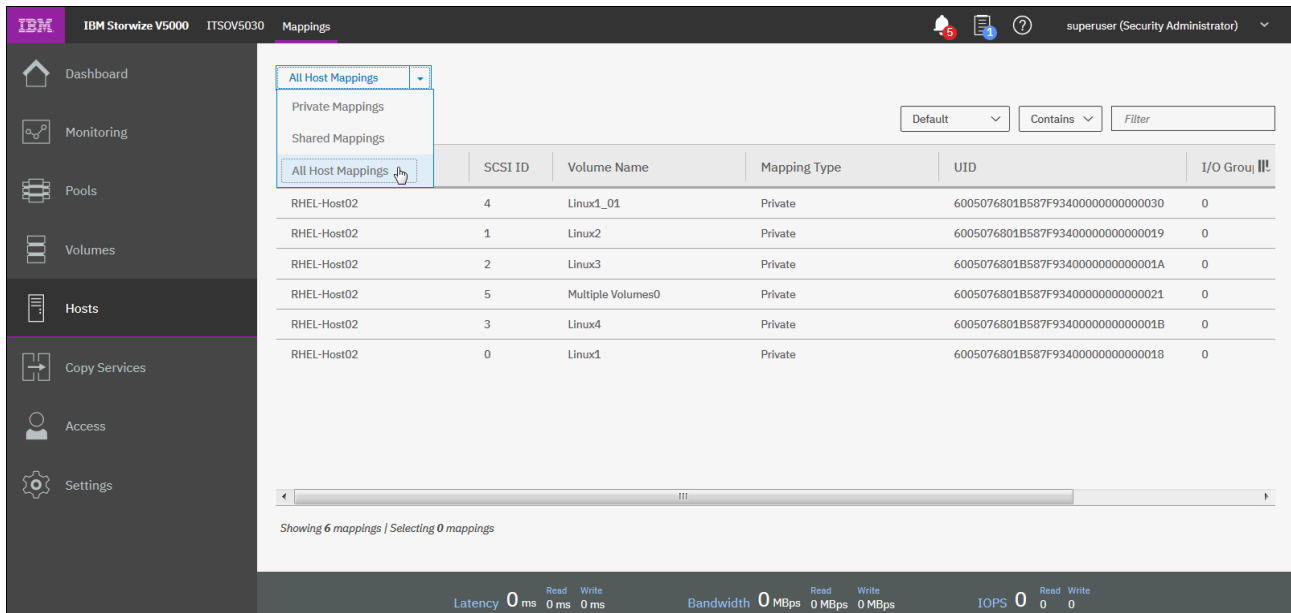


Figure 3-67 Host mappings

From this window, you can view the host properties. You also can obtain the list of mapped volumes or work with port definitions. Right-click the specific host and select **Properties (Host)** from the opened menu. A window similar to the window that is shown in Figure 3-68 on page 134 opens.

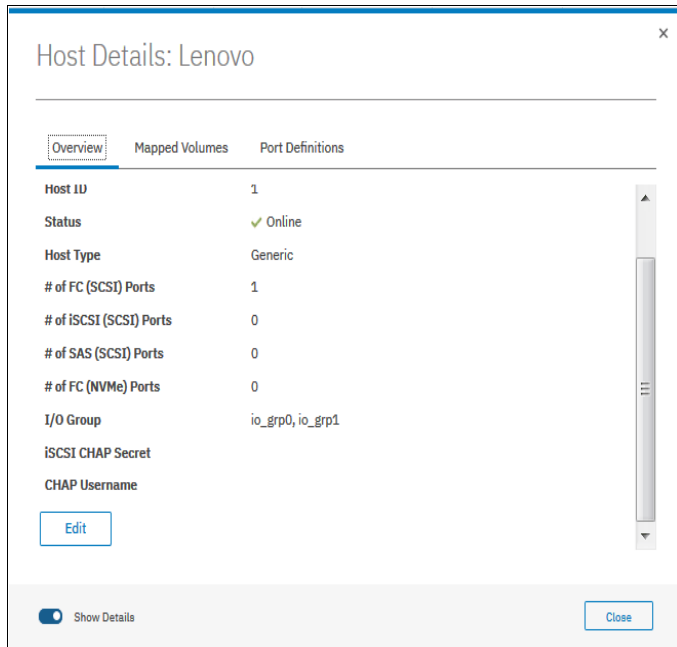


Figure 3-68 Host properties

With enabled details, you can modify host name, host type, I/O group assignment, or iSCSI Challenge Handshake Authentication Protocol (CHAP) Secret by clicking **Edit** and then, **Save** (see Figure 3-68).

3.5.5 Volumes by host

This option is identical to the option that is available in the dynamic menu Volumes. For more information, see 3.4.3, "Volumes by host" on page 128.

3.6 Copy services

The IBM Spectrum Virtualize copy services are part of the IBM Replication Family Services, which are available in all Storwize family products. It consists of the following functions:

- ▶ FlashCopy
- ▶ Metro Mirror and Global Mirror
- ▶ Global Mirror with Changed Volumes
- ▶ Volume Mirroring function (Volume Copy)
- ▶ HyperSwap volume mirroring

Figure 3-69 shows the Copy Services menu functions.

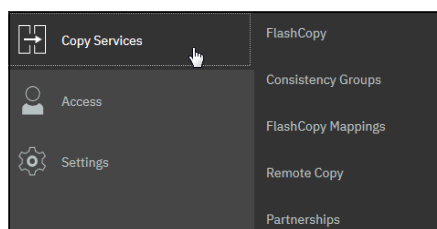


Figure 3-69 Copy Services menu

In this section, we briefly describe how to navigate in the Copy Services menu.

3.6.1 IBM FlashCopy

IBM FlashCopy is a function that you use to create a point-in-time copy of one of your IBM Spectrum Virtualize volumes. This function might be helpful when you back up data or test applications. These copies can be cascaded one on another, read from, written to, and even reversed.

FlashCopy snapshots can conserve storage, if needed, by being space-efficient copies (rather than full copies) that record only items that changed from the originals. Select **FlashCopy** from the dynamic menu to open a pane similar to the pane that is shown in Figure 3-70.

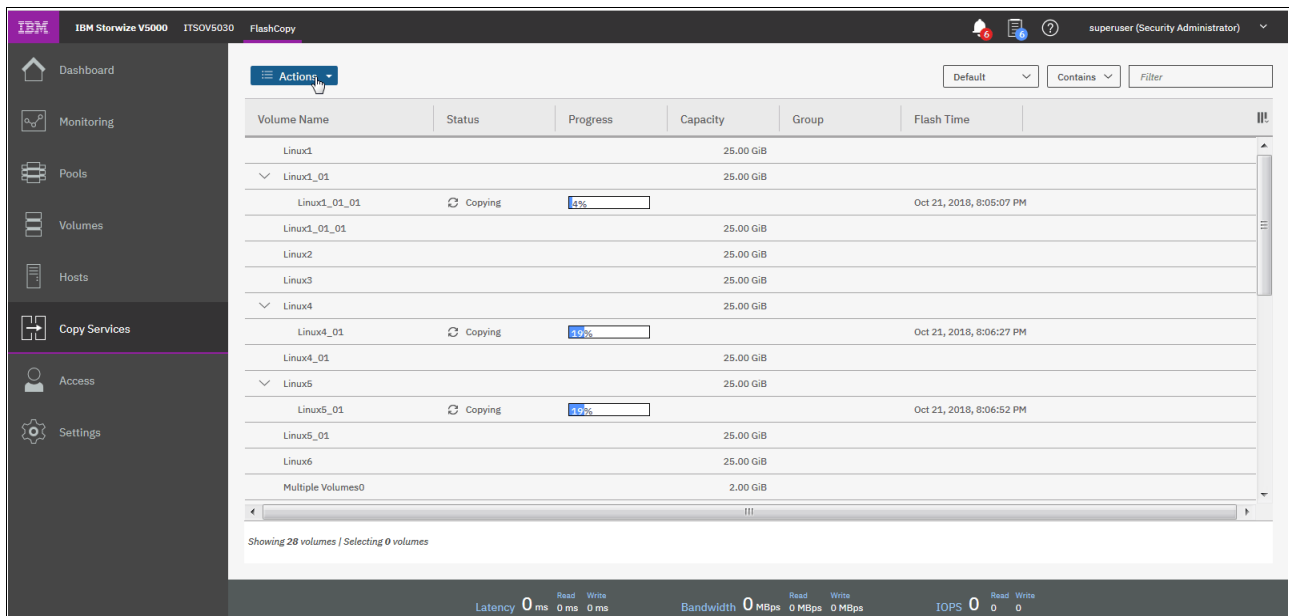


Figure 3-70 FlashCopy operations

If you must create a FlashCopy of another volume, right-click the volume and the list of available functions is displayed. You can perform several tasks, such as start a new snapshot, or clone or back up a volume.

Clicking the volume name opens the window that is shown in Figure 3-71. You can click the tabs at the top of the window to display more information, such as the hosts that the volume or FlashCopy volume is mapped to and its dependent MDisks.

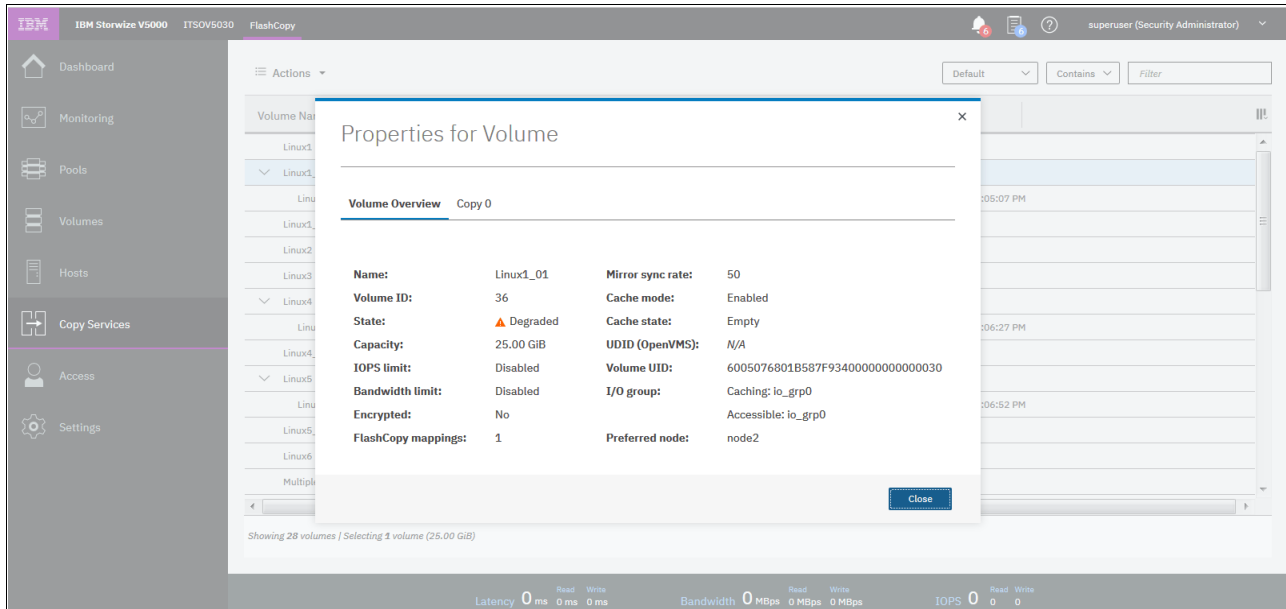


Figure 3-71 FlashCopy volume details

3.6.2 Consistency groups

FlashCopy *consistency groups* can be used to create a consistent point-in-time copy across multiple volumes, and even across multiple managed storage systems, which manages the consistency of dependent writes.

Click **Consistency Group** to open the window that is shown in Figure 3-72. FlashCopy relationships can be placed into a consistency group. You can also use start and stop commands against the FlashCopy consistency group from this window by right-clicking the relationship.

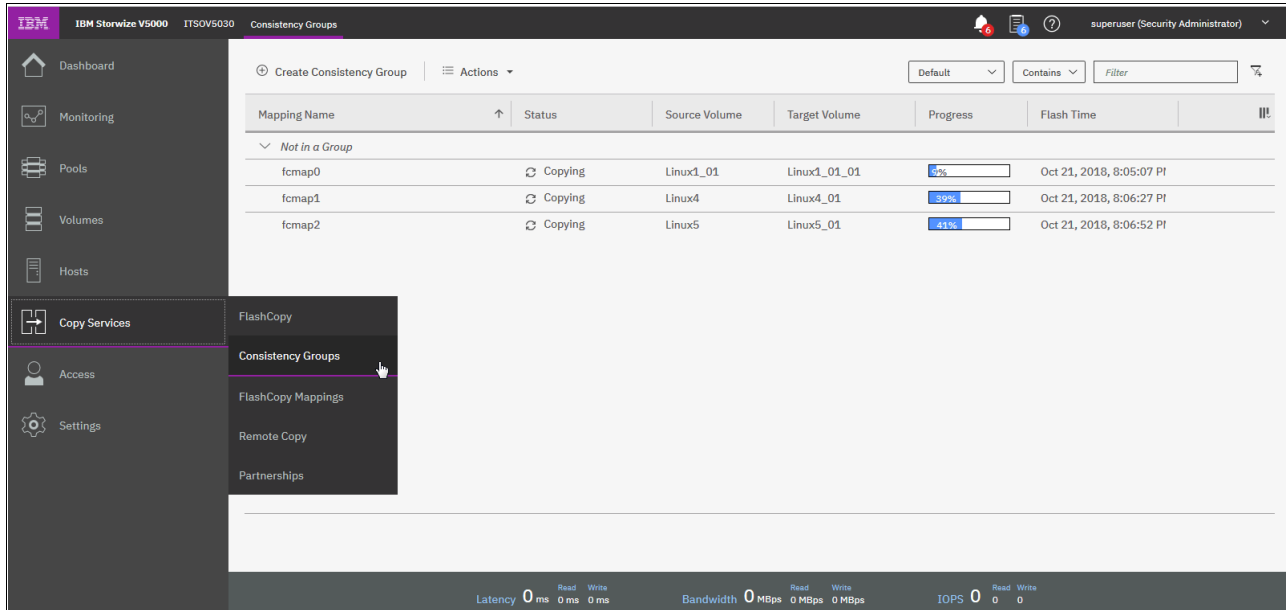


Figure 3-72 FlashCopy Consistency Groups window

When any FlashCopy consistency group is available (empty or with existing relationships), you can move a relationship to that group. Right-click a relationship and select **Move to Consistency Group**, as shown in Figure 3-73. Copying must be finished or idle.

Other actions on the same menu include Remove from Consistency Group, Start (resume) or Stop that FlashCopy operation, Rename Mapping (rename a target volume or FlashCopy mapping), and Delete Mapping.

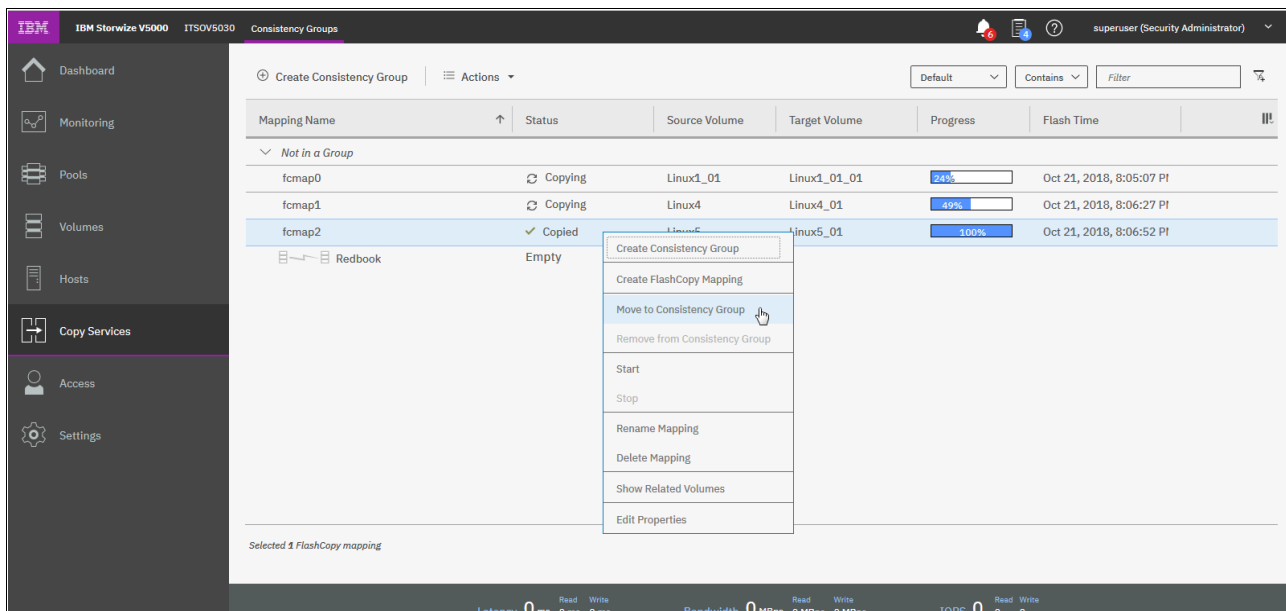


Figure 3-73 Moving a relationship to the consistency group

From the menu, select the appropriate group (in our case, the only one available) and confirm the selection (see Figure 3-74).

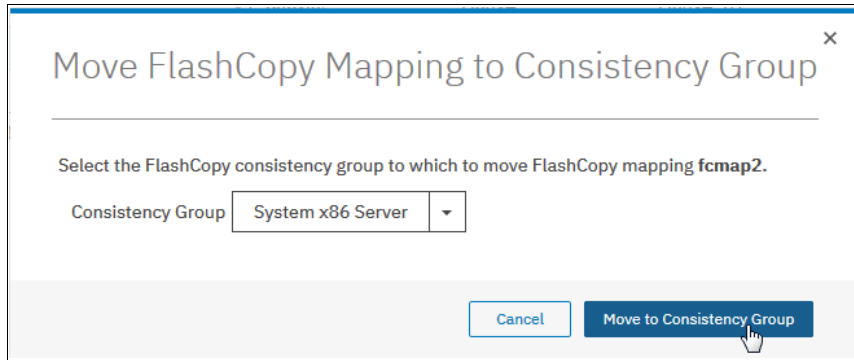


Figure 3-74 Assigning the consistency group

The result of the operation is similar to the result that is shown in Figure 3-75.

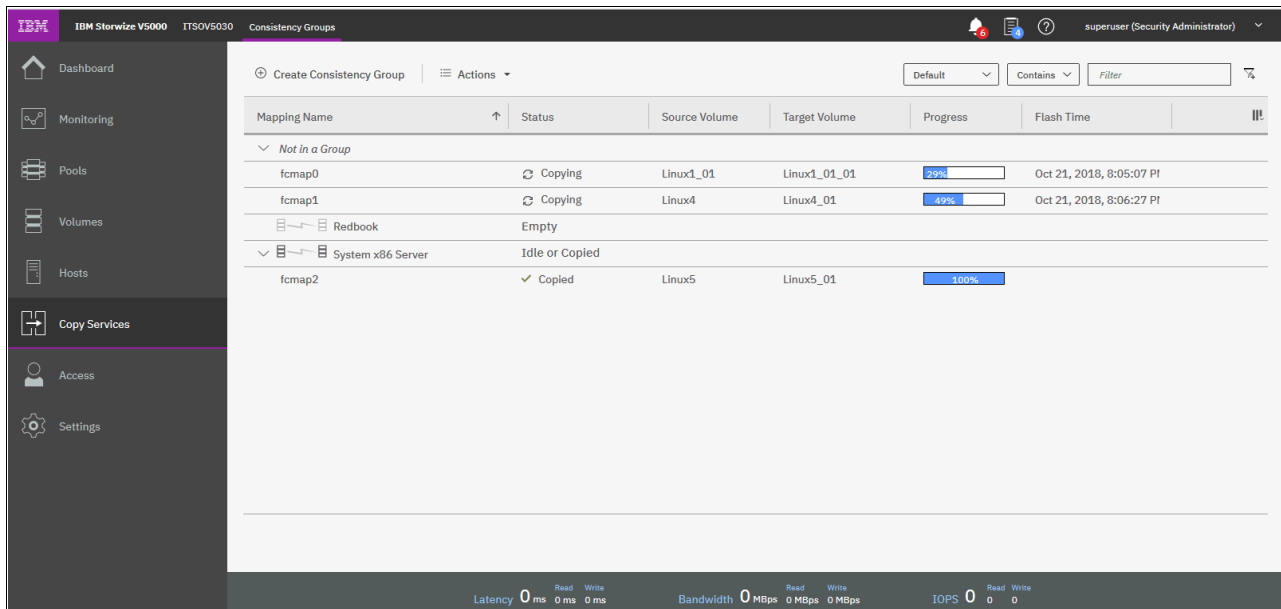


Figure 3-75 Consistency groups

3.6.3 FlashCopy mappings

To create a FlashCopy mapping, click **Create FlashCopy Mapping** (as shown in Figure 3-76) to start a wizard. This wizard maps a source volume to a target volume to prepare for a subsequent copy. This mapping persists until it is deleted. The mapping specifies the source and destination volumes. The destination must be identical in size to the source or the mapping fails.

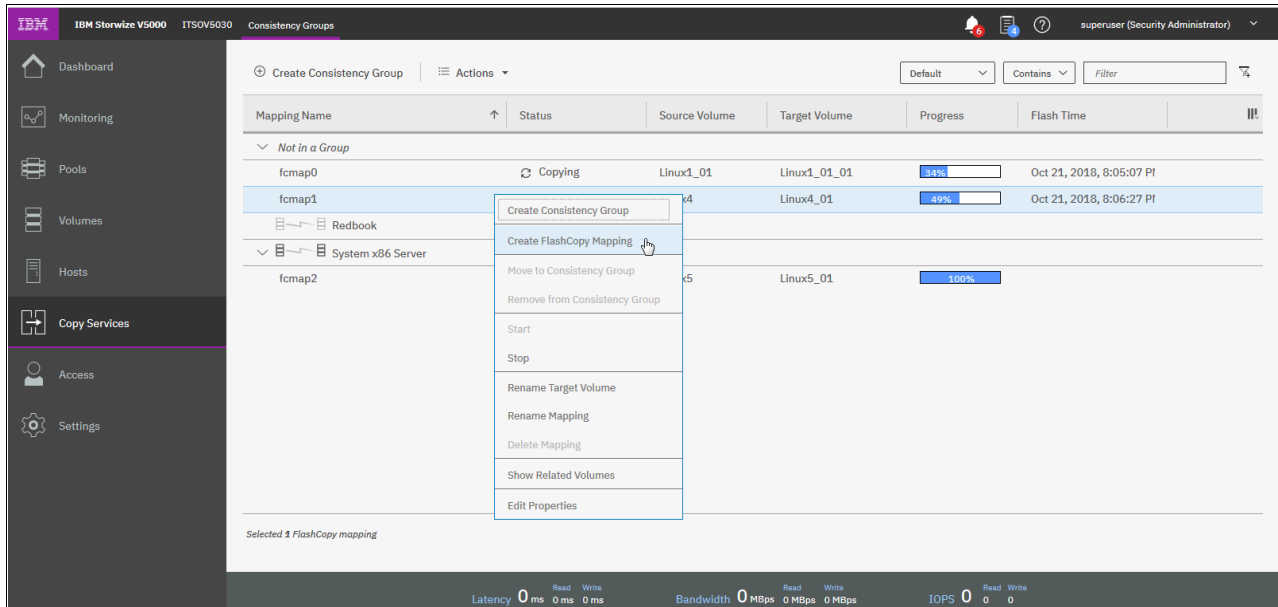


Figure 3-76 FlashCopy mappings

In a single mapping, the source and destination cannot be on the same volume. A mapping is triggered at the point in time when the copy is required. The mapping can optionally be given a name and assigned to a consistency group. These groups of mappings can be triggered at the same time. This process enables multiple volumes to be copied at the same time, which creates a consistent copy of multiple disks. A consistent copy of multiple disks is required for database products in which the database and log files are on separate disks.

If a consistency group (ID or Name) is not specified, the mapping is assigned to the default *group 0*, which is a special group that cannot be started as a whole. Mappings in this group can be started only on an individual basis.

An example of the wizard for FlashCopy mapping creation is shown in Figure 3-77. Select source and target volumes from the wizard.

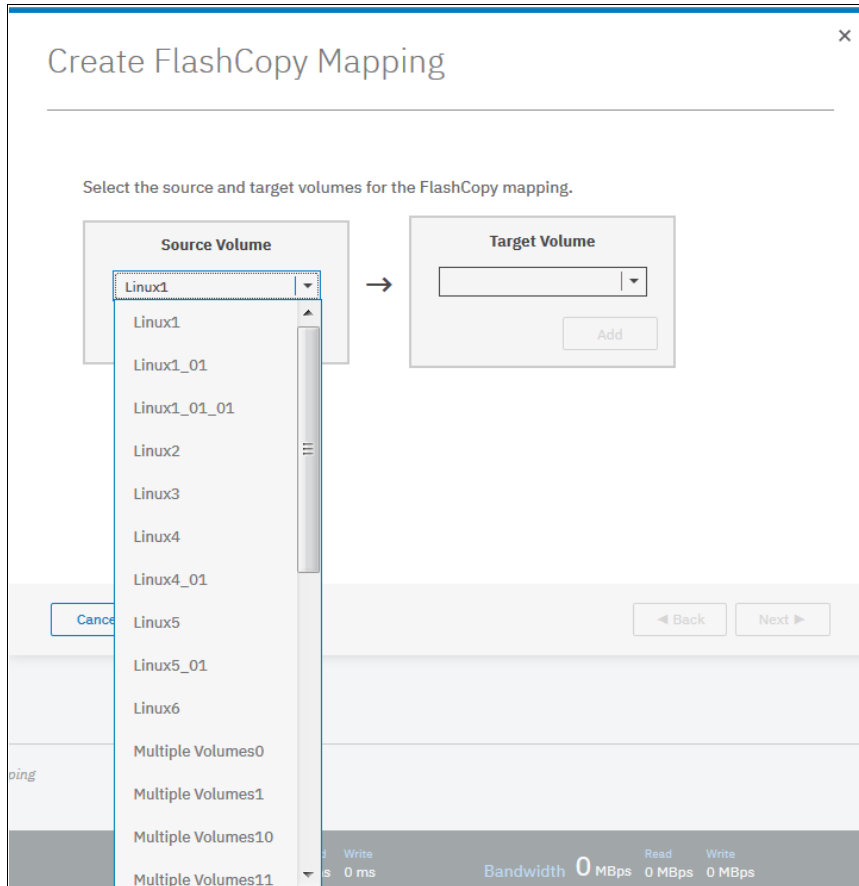


Figure 3-77 Selecting volumes for FlashCopy mappings

You can select the Snapshot (copy-on-write), Clone (replica of the volume without effect on original one), or Backup (data recovery) type of relationship. When selected, you can specify whether you also want to add the mapping to the consistency group.

3.6.4 Remote copy

Click **Remote Copy** to open the window that is shown in Figure 3-78. This window shows the existing remote copy relationships, and you can set up and modify consistency groups. From this window, you can also start and stop relationships, add relationships to a consistency group, and switch the direction of the mirror.

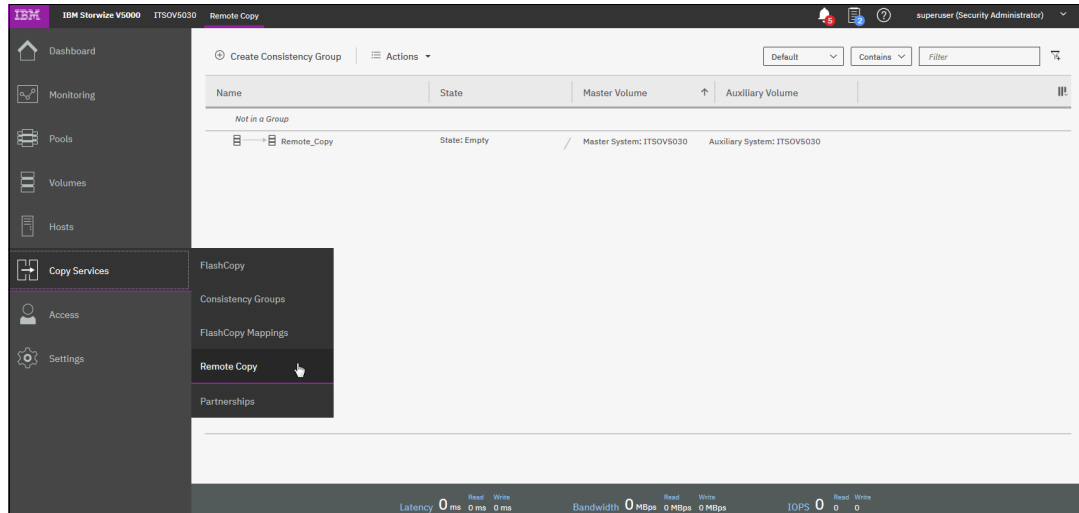


Figure 3-78 Remote Copy window

The menu provides the options to create Metro Mirror, Global Mirror, or Global Mirror with Changed Volumes:

Metro Mirror

This option makes *synchronous* copies. The original write operations are not considered complete until the write operation to the destination disk is confirmed. The distance between your two sites is determined by how much latency your applications can handle.

Global Mirror

This option makes *asynchronous* copies of your disk. The write is considered complete after it is complete at the local disk. It does not wait for the write to be confirmed at the remote cluster as Metro Mirror does. This method greatly reduces the latency that is experienced by your applications if the other cluster is far away.

However, it also means that during a failure, the data on the remote copy might not contain the most recent changes that were committed to the local disk.

Global Mirror with Changed Volumes

This option is best described as “Continuous Remote FlashCopy.” If you use this feature, IBM Spectrum Virtualize essentially takes a periodic FlashCopy of a disk and writes it to your remote destination. This feature completely isolates the local copy from wide area network (WAN) issues and from sudden spikes in workload that might occur.

The drawback is that your remote copy might lag behind the original by a significant amount of data, depending on how you set up the cycle time.

3.6.5 Partnerships

Click **Partnerships** to open the window that is shown in Figure 3-79. You can use this window to set up a new partnership, or delete a partnership for remote mirroring with another Storwize V5000 system.

To create a partnership, click **Create Partnership**. A new window opens. When you select the partnership type (for example, Fibre Channel), the window expands to a more detailed view, as shown in Figure 3-79.

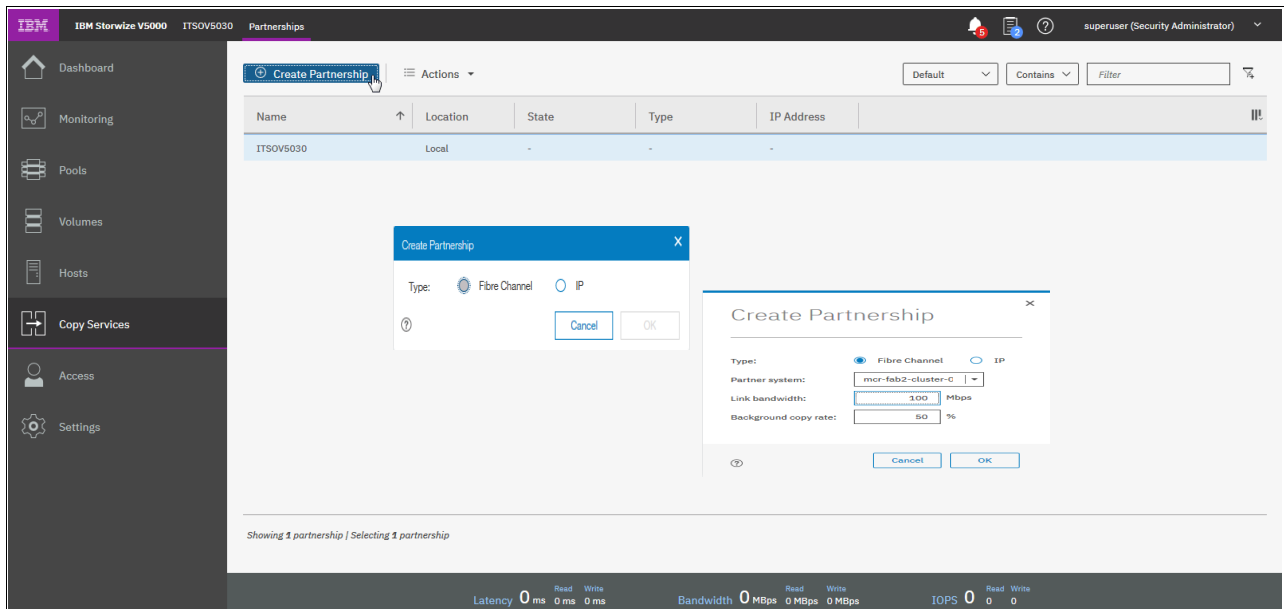


Figure 3-79 Creating a partnership

Clicking a partnership opens a window, as shown in Figure 3-80. From this window, you can also set the background copy rate. This rate specifies the bandwidth, in Mbps, that is used by the background copy process between the clusters (see Figure 3-80). In our case, we configured the partnership only on one side.

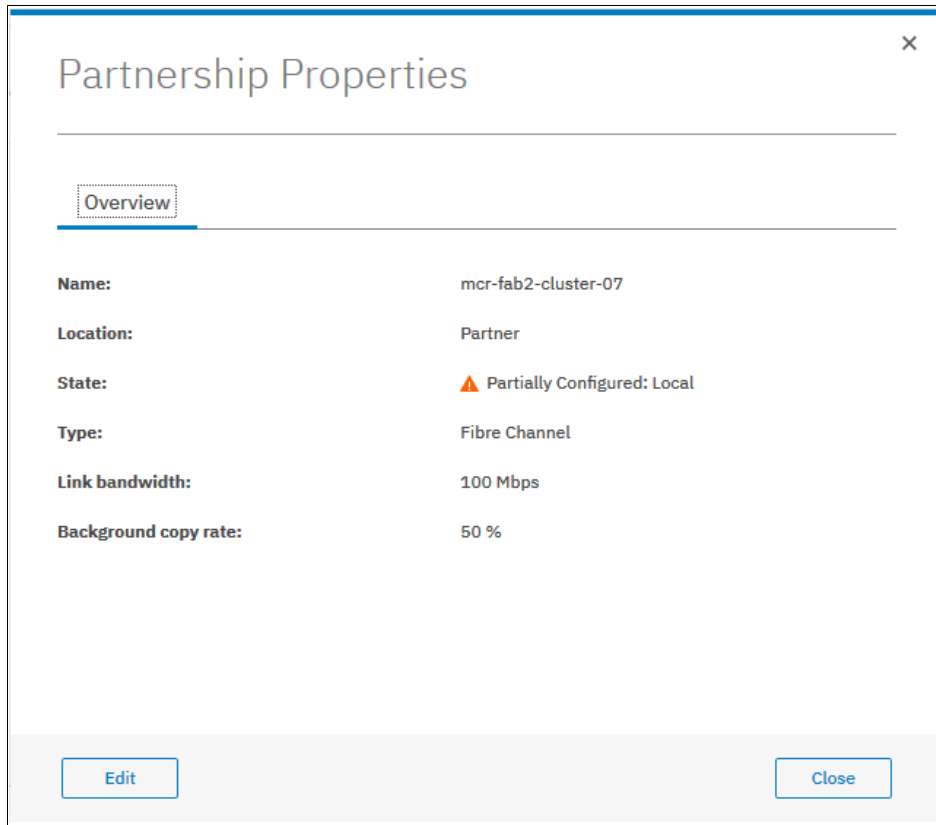


Figure 3-80 Partnership properties window

You can see it in the State row. It shows *Partially Configured: Local*, which is an indication that the configuration was only configured on one side. If you see this message, go to the second system and configure the Create Partnership settings there as well.

3.7 Access menu

The Access menu includes the following options (see Figure 3-81):

- ▶ Users (for user management)
- ▶ Audit Log

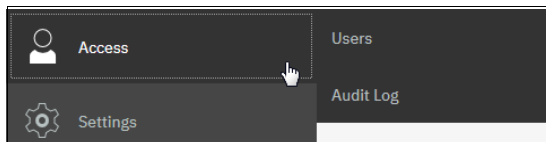


Figure 3-81 Access menu

3.7.1 Users

Figure 3-82 shows the Users window. You can create and delete new users, change and remove passwords, and add and remove Secure Shell (SSH) keys.

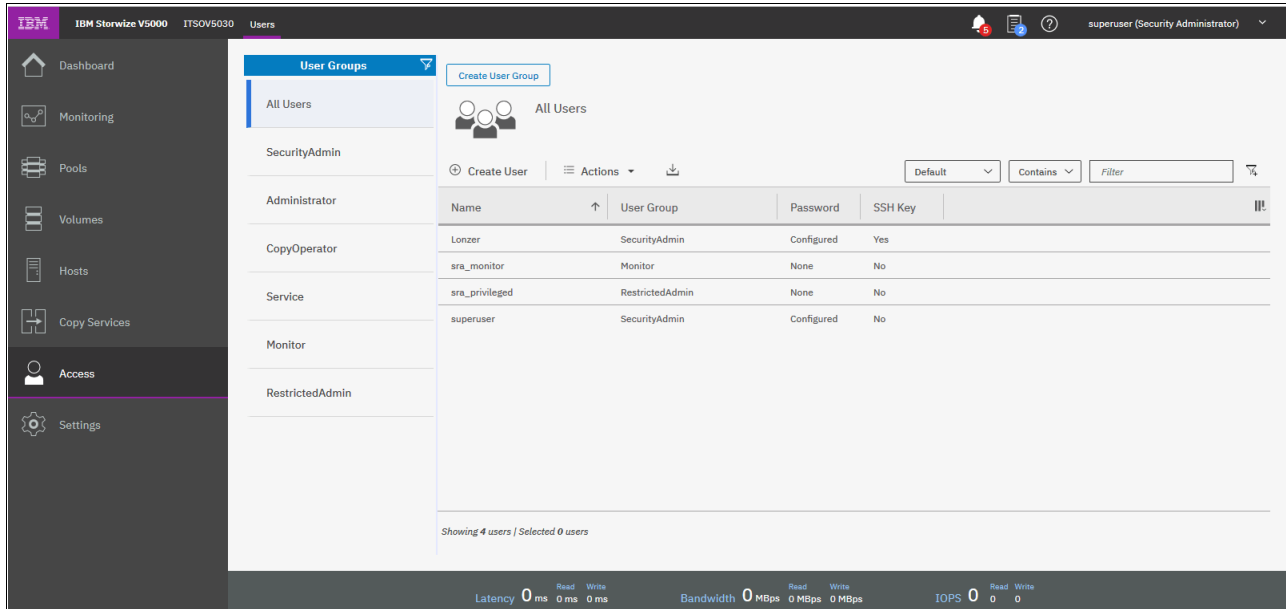


Figure 3-82 Users window

Click **Create User** to open the pane that is shown in Figure 3-83. Use this pane to specify the name and password of the user, and load the SSH key (if the SSH key was generated). An SSH key is not required for CLI access, and you can choose to use SSH or a password for CLI authentication.

Create User [X]

Name
Administrator

Authentication Mode
 Local Remote
User Group: SecurityAdmin

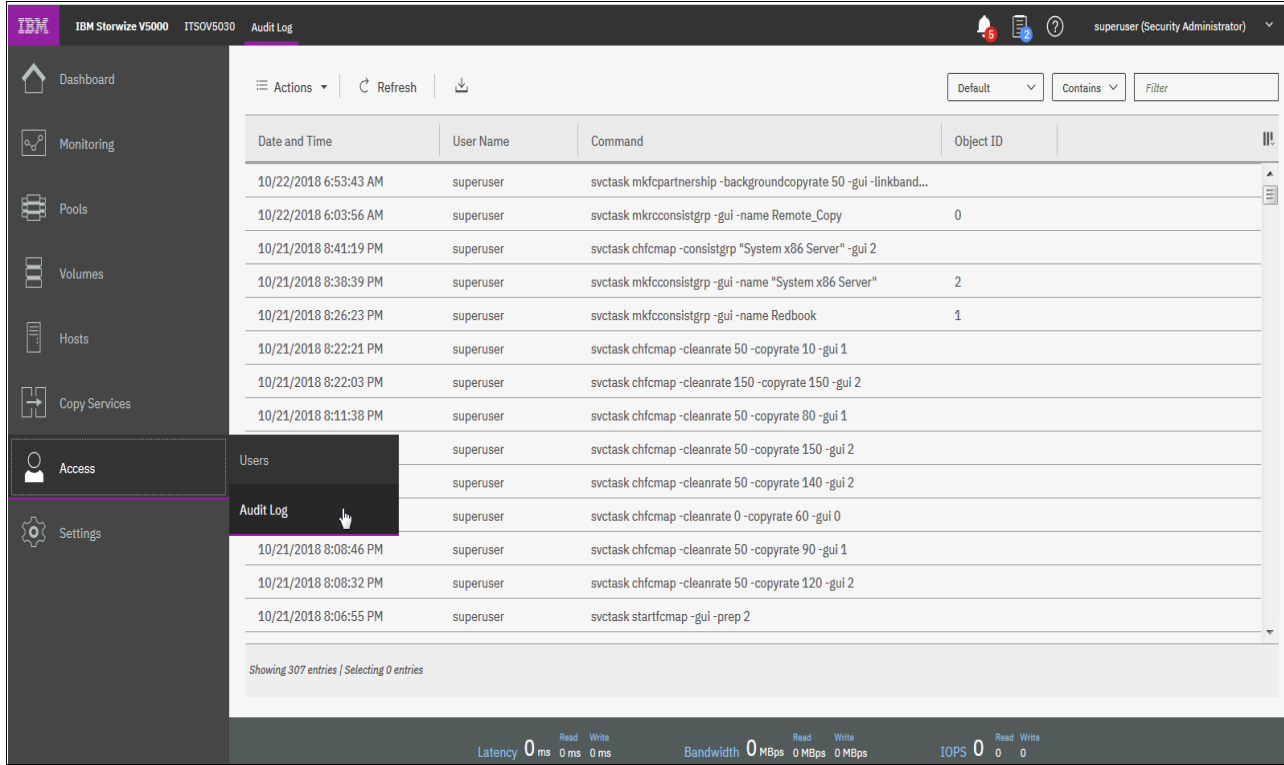
Local Credentials
Users must have a password, an SSH public key, or both.
Password: [Masked] Verify password: [Masked]
SSH Public Key: [Browse...] No file selected.

Cancel Create

Figure 3-83 Creating a user

3.7.2 Audit Log option

Click **Audit Log** to open the window that is shown in Figure 3-84. The cluster maintains an audit log of successfully run commands and displays the users that performed particular actions at certain times.

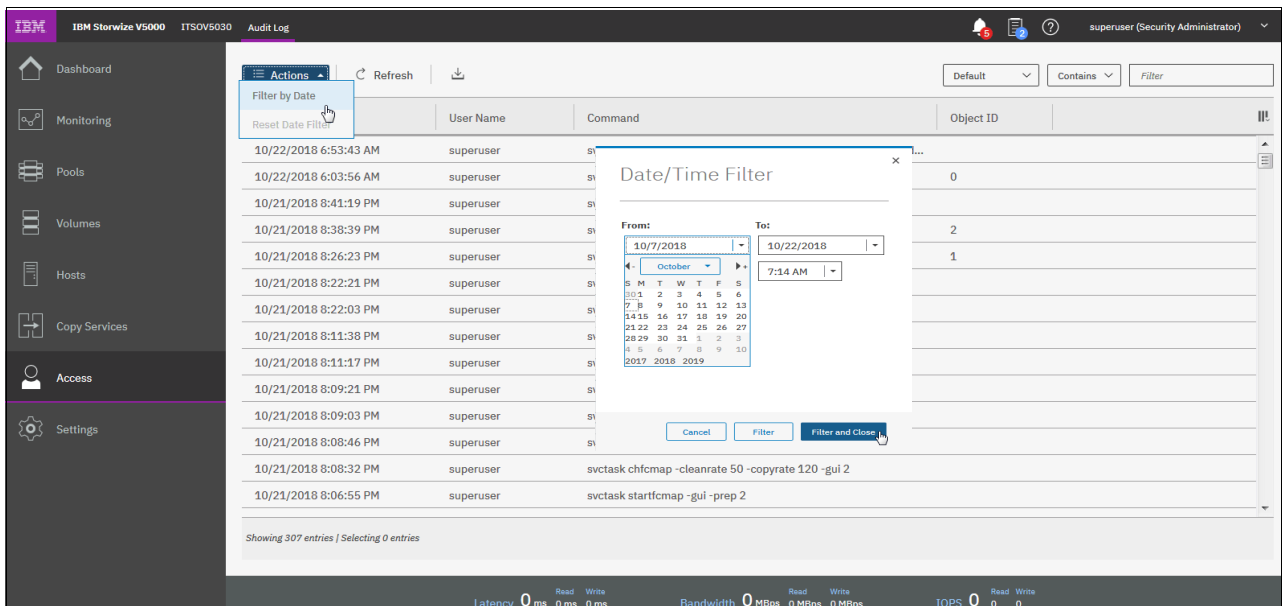


The screenshot shows the IBM Storwize V5000 Audit Log interface. The left sidebar contains navigation options: Dashboard, Monitoring, Pools, Volumes, Hosts, Copy Services, Access, and Settings. The main area displays a table of audit log entries. The table has columns for Date and Time, User Name, Command, and Object ID. The entries show various commands executed by the superuser, such as svctask mkfcpartnership, svctask mkrcconsistgrp, svctask chfcmcp, and svctask startfcmcp. The interface also includes a top navigation bar with the IBM logo, system information (IBM Storwize V5000, ITSOV5030), and user information (superuser (Security Administrator)). A bottom status bar shows performance metrics like Latency, Bandwidth, and IOPS.

Date and Time	User Name	Command	Object ID
10/22/2018 6:53:43 AM	superuser	svctask mkfcpartnership -backgroundcopyrate 50 -gui -linkband...	
10/22/2018 6:03:56 AM	superuser	svctask mkrcconsistgrp -gui -name Remote_Copy	0
10/21/2018 8:41:19 PM	superuser	svctask chfcmcp -consistgrp "System x86 Server" -gui 2	2
10/21/2018 8:38:39 PM	superuser	svctask mkfconsistgrp -gui -name "System x86 Server"	2
10/21/2018 8:26:23 PM	superuser	svctask mkfconsistgrp -gui -name Redbook	1
10/21/2018 8:22:21 PM	superuser	svctask chfcmcp -cleanrate 50 -copyrate 10 -gui 1	
10/21/2018 8:22:03 PM	superuser	svctask chfcmcp -cleanrate 150 -copyrate 150 -gui 2	
10/21/2018 8:11:38 PM	superuser	svctask chfcmcp -cleanrate 50 -copyrate 80 -gui 1	
	superuser	svctask chfcmcp -cleanrate 50 -copyrate 150 -gui 2	
	superuser	svctask chfcmcp -cleanrate 50 -copyrate 140 -gui 2	
	superuser	svctask chfcmcp -cleanrate 0 -copyrate 60 -gui 0	
10/21/2018 8:08:46 PM	superuser	svctask chfcmcp -cleanrate 50 -copyrate 90 -gui 1	
10/21/2018 8:08:32 PM	superuser	svctask chfcmcp -cleanrate 50 -copyrate 120 -gui 2	
10/21/2018 8:06:55 PM	superuser	svctask startfcmcp -gui -prep 2	

Figure 3-84 Audit Log entries

You can filter audit log records by date or within a specific time frame (see Figure 3-85).



The screenshot shows the IBM Storwize V5000 Audit Log interface with a date/time filter dialog box open. The dialog box is titled "Date/Time Filter" and has fields for "From" and "To" dates and times. The "From" date is set to 10/7/2018 and the "To" date is set to 10/22/2018. The time is set to 7:14 AM. The dialog box also includes a calendar view for the month of October 2018. The background shows the same audit log table as in Figure 3-84, but with a "Filter by Date" dropdown menu open in the top left corner of the table area. The interface also includes a top navigation bar and a bottom status bar.

Figure 3-85 Filtering the records

The following commands are *not* recorded in the audit log:

- ▶ All commands that failed
- ▶ `dumpconfig`
- ▶ `cpdumps`
- ▶ `cleardumps`
- ▶ `finderr`
- ▶ `dumperrlog`
- ▶ `dumpinternallog`
- ▶ `svcservicetask dumperrlog`
- ▶ `svcservicetask finderr`

3.8 Settings menu

The Settings menu provides various configurable options to adjust your system parameters according to your needs (see Figure 3-86):

- ▶ Notifications
- ▶ Network
- ▶ Security (remote authentication)
- ▶ System
- ▶ Support
- ▶ GUI Preferences

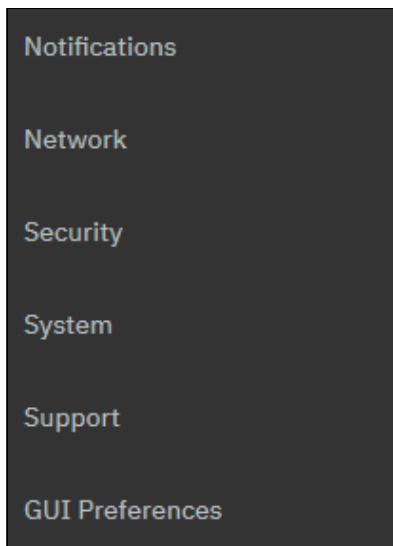


Figure 3-86 Settings menu

3.8.1 Notifications

It is important 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 monitoring for new events that use the management GUI, select the type of event that you want to be notified about; for example, restrict notifications to events that require immediate action.

You can use email, Simple Network Management Protocol (SNMP), or syslog types of notifications (see Figure 3-87).

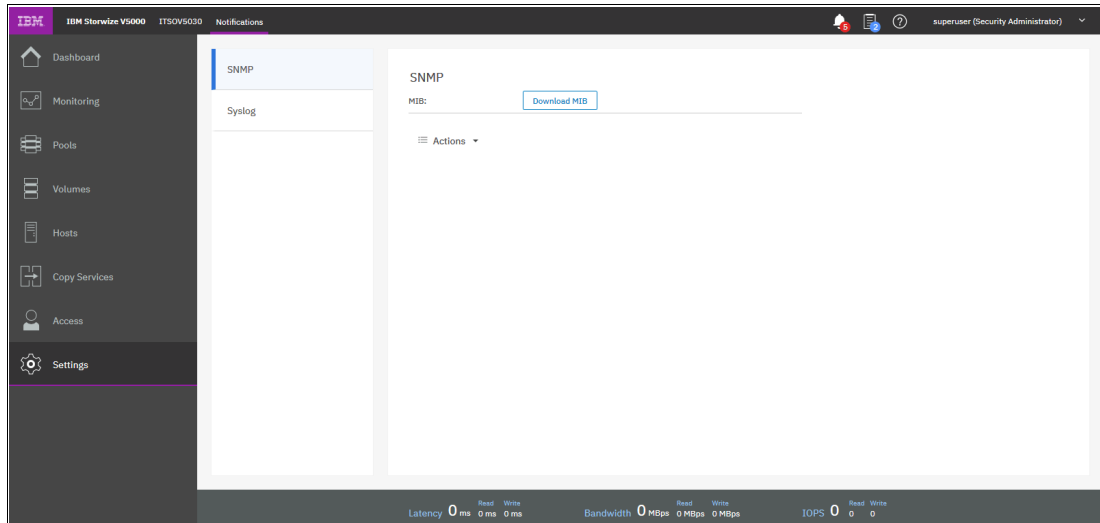


Figure 3-87 SNMP and Syslog settings

If your system is within warranty, or if you use a hardware maintenance agreement, configure your Storwize V5000 system to send email events to IBM directly if an issue that requires hardware replacement is detected. This mechanism is called *Call Home*. For more information, see 3.8.5, “Call Home notifications” on page 156.

3.8.2 Network

Click **Network** to open the window that is shown in Figure 3-88. You can update the network configuration, set up iSCSI definitions, and view information about the Fibre Channel connections.

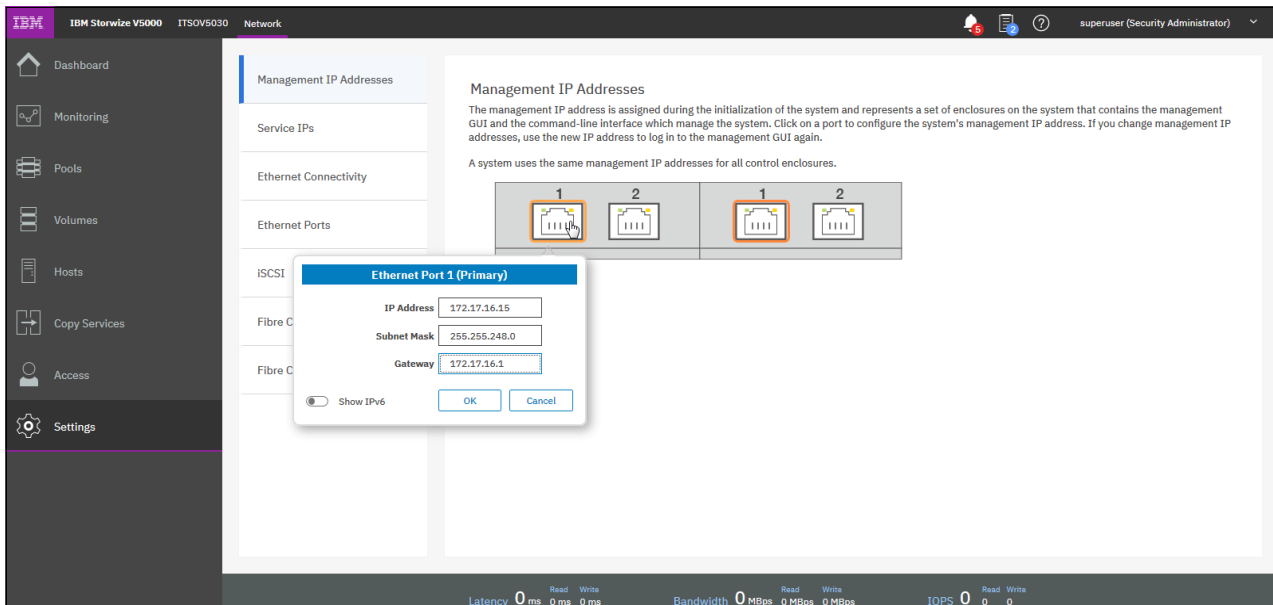


Figure 3-88 Network window

When you click **Fibre Channel Connectivity** (see Figure 3-89), useful information is displayed. In this example, we click **All nodes, storage systems, and hosts** from the menu and then, select **Show Results** to display the details. Other options that are available from the menu include displaying Fibre Channel details for a host, clusters, nodes, or storage systems.

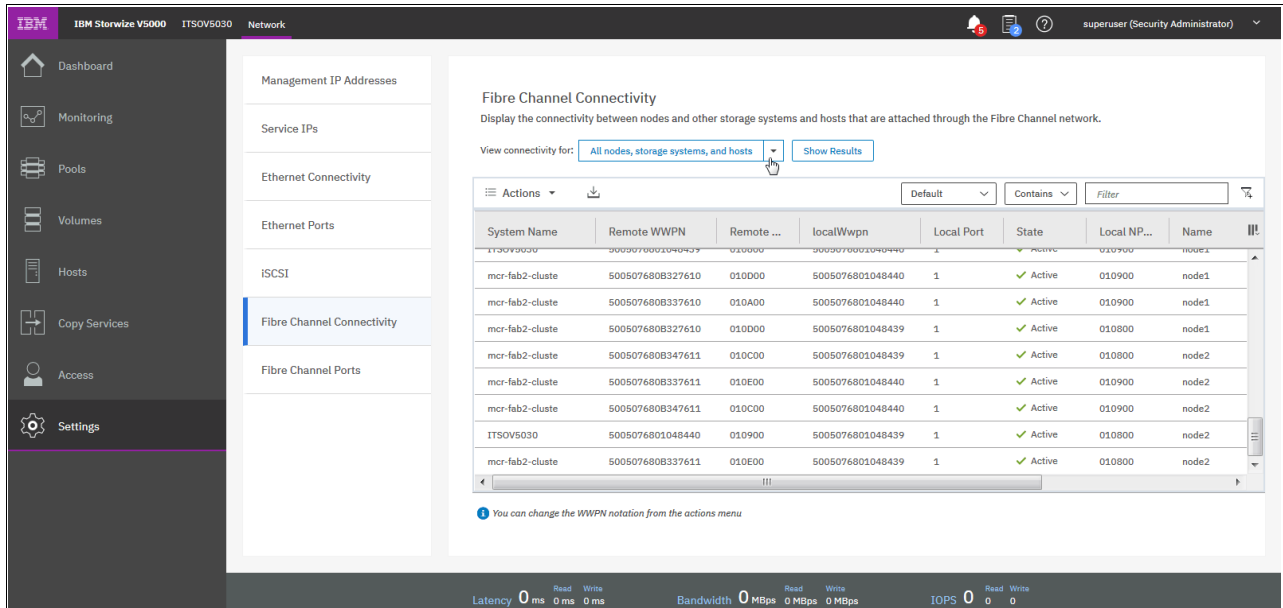


Figure 3-89 Fibre Channel connectivity

3.8.3 Security features

The different security features are described in this section.

Remote authentication

With security and its directory services, the user can remotely authenticate to the IBM Spectrum Virtualize without the need for a local account. Therefore, when you log on, you authenticate with your domain user ID and password rather than a locally created user ID and password.

Remote authentication includes the following benefits:

- ▶ You do not need to configure every user on every IBM Spectrum Virtualize. If multiple machines are in your environment, you can set up authentication more efficiently.
- ▶ When commands are run on the IBM Spectrum Virtualize, the audit log shows the domain user name that issued that command, rather than a local user name, or worse, just “superuser”. (In this case, determining who mapped a volume, acted as the superuser, and so on, might be difficult.)
- ▶ You have central control over access. If someone leaves the company, you remove access at the domain controller, which means that orphan accounts do not remain on your storage equipment.

The access pane to configure remote authentication is shown in Figure 3-90.

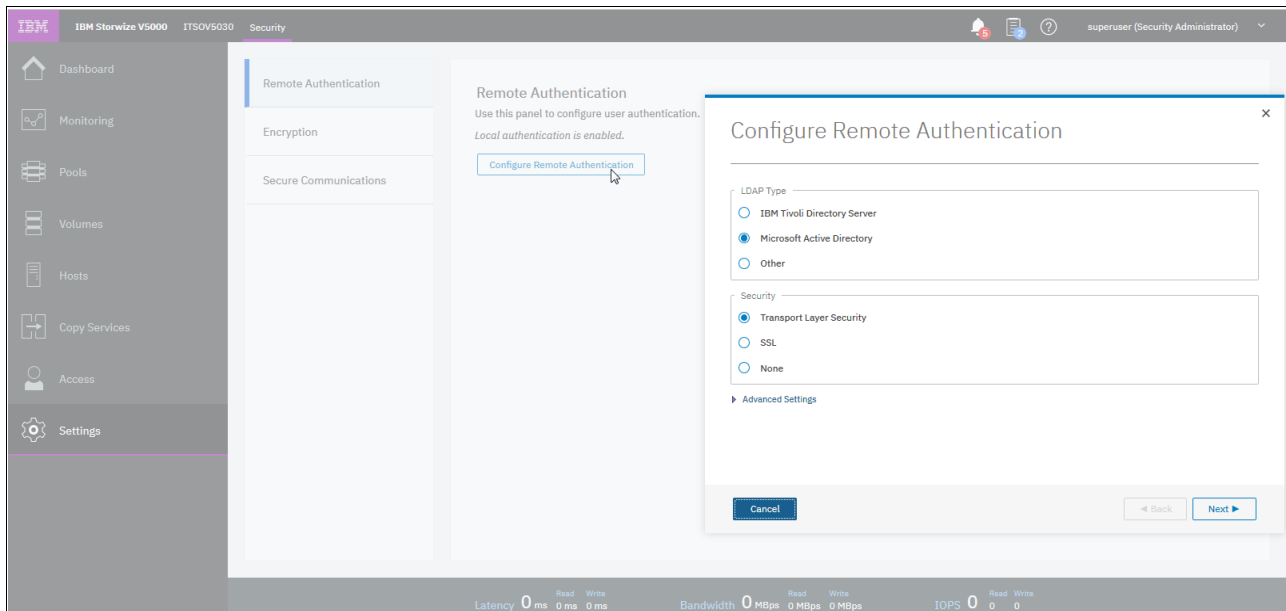


Figure 3-90 Configuring remote authentication

For more information about how to configure remote logon, see the following resources:

- ▶ [IBM Community blog](#)
- ▶ [IBM Knowledge Center](#)

Encryption

In the window that is shown in Figure 3-91, you can enable or disable the encryption function on an IBM Storwize V5030. The window shows that no USB drives that contain encryption keys were detected. These are no longer needed if you have an external Key Management Server.

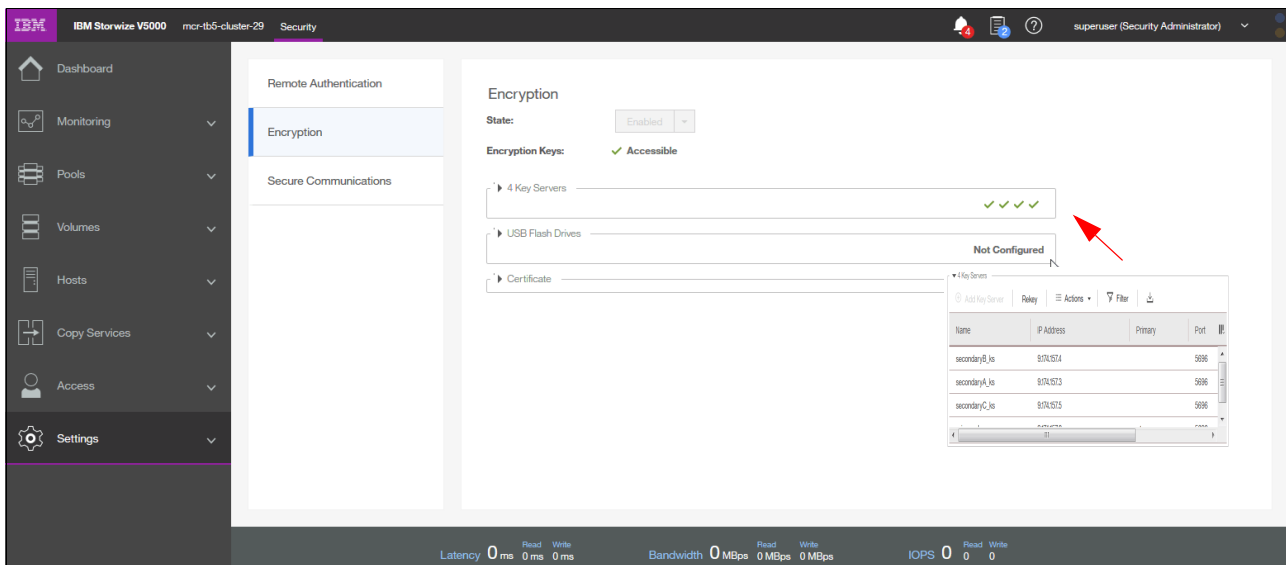


Figure 3-91 Encryption window

Figure 3-91 on page 150 also shows four available external Key Management Server. If no external Server is available, you need the USB keys to encrypt or decrypt your data on start.

External Encryption Management server is implemented in the IBM Spectrum Virtualize code V8.1 onwards.

Secure communications

Use the Secure Communications page to enable and manage secure connections. During system setup, an initial certificate is created to use for secure connections between web browsers. Based on the security requirements for your system, you can create a self-signed certificate or install a signed certificate that is created by a third-party certificate authority.

Self-signed certificates are generated automatically by the system and encrypt communications between the browser and the system. Self-signed certificates can generate web browser security warnings, and they might not comply with organizational security guidelines.

Signed certificates are created by a third-party certificate authority. These certificate authorities ensure that certificates include the required security level for an organization based on purchase agreements. Signed certificates usually have higher security controls for the encryption of data and do not cause browser security warnings.

To use a signed certificate, first generate and download a request for a certificate that is based on the values that are specified on the Secure Communication page. Submit this request to the certificate authority to receive a signed certificate and then install it by using the Secure Communication page.

Before you create a request for either type of certificate, ensure that your current browser does not restrict the type of keys that are used for certificates. Certain browsers limit the use of specific key types for security and compatibility.

The details about the security certificates is shown in Figure 3-92.

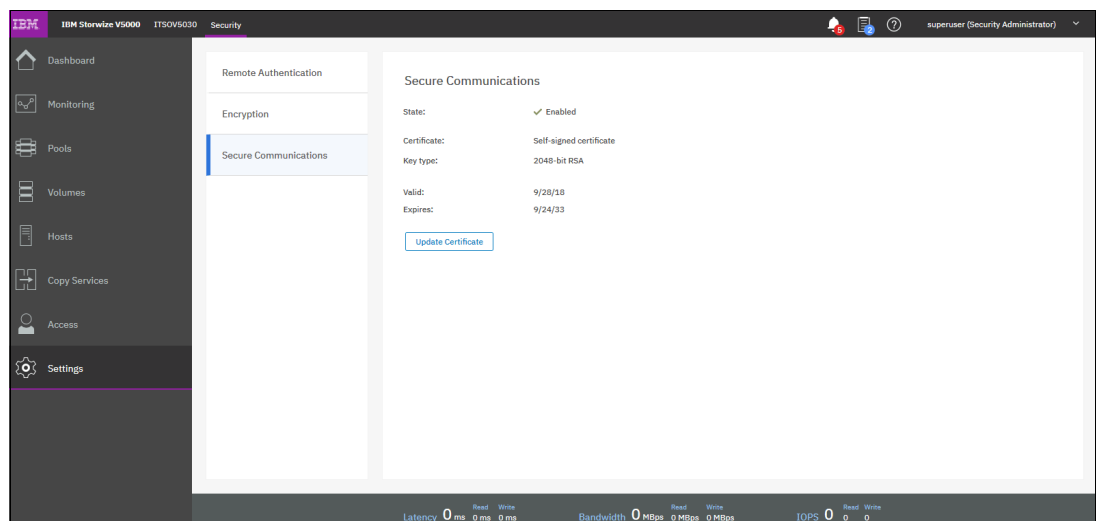


Figure 3-92 Secure communications

If you want to update or change the certificate, click **Update Certificate**.

The Update Certificate window opens, as shown in Figure 3-93.

Update Certificate

Certificate type: Self-signed certificate
 Signed certificate

Key type: 2048-bit RSA

Validity days: 5,475

Country: GB

State:

City: Manchester

Organization: IBM

Organization unit:

Common name: mcr-tb5-cluster-21.stglab.manchester.uk.il

Email address: test.us@ibm.com

Cancel Update

Figure 3-93 Update Certificate window

3.8.4 System menu

The System menu provides the following options:

- ▶ Set the system date and time
- ▶ Manage licenses
- ▶ Upgrade System
- ▶ Virtual Volumes (VVols)
- ▶ IP Quorum
- ▶ I/O Groups
- ▶ DNS

The Date and Time window opens (see Figure 3-94) when you select **Date and Time** from the System menu. You can add a Network Time Protocol (NTP) server, if available.

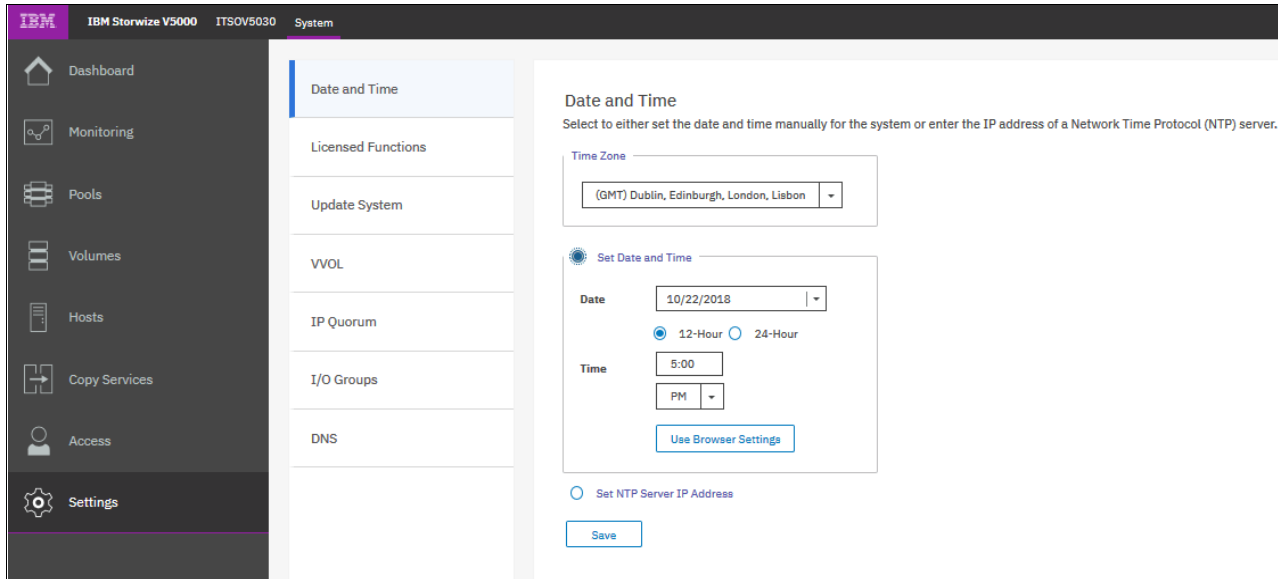


Figure 3-94 Date and Time window

You can also update the license information for specific features, as shown in Figure 3-95.

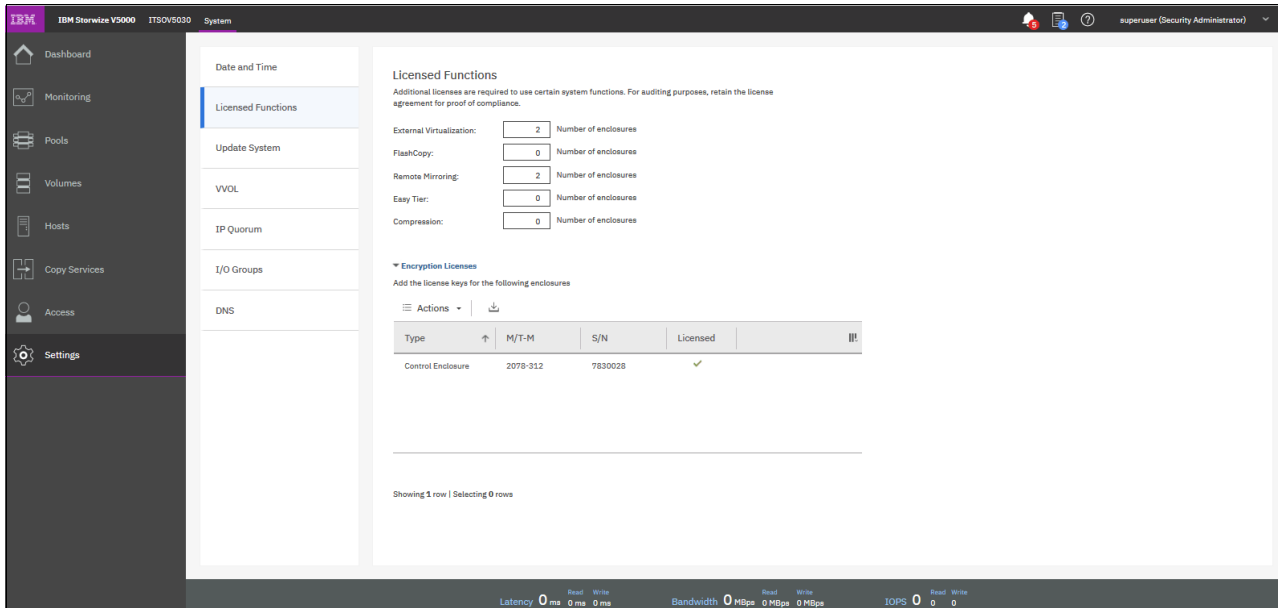


Figure 3-95 Licensing options

To upgrade your IBM Spectrum Virtualize, use the procedure that is described in Chapter 12, “RAS, monitoring, and troubleshooting” on page 623.

Virtual Volume (VVol) is a tape volume that is in a tape volume cache of a virtual tape server (VTS). VVol is a new feature that was introduced in IBM Spectrum Virtualize 7.6. By using this new functionality, users can create volumes on IBM Spectrum Virtualize directly from a VMware vCenter server.

On the VVOL page, you can enable or disable the functionality, as shown in Figure 3-96. Before you can enable VVol, you must set up an NTP server. See the Date and Time settings to set up the NTP server.

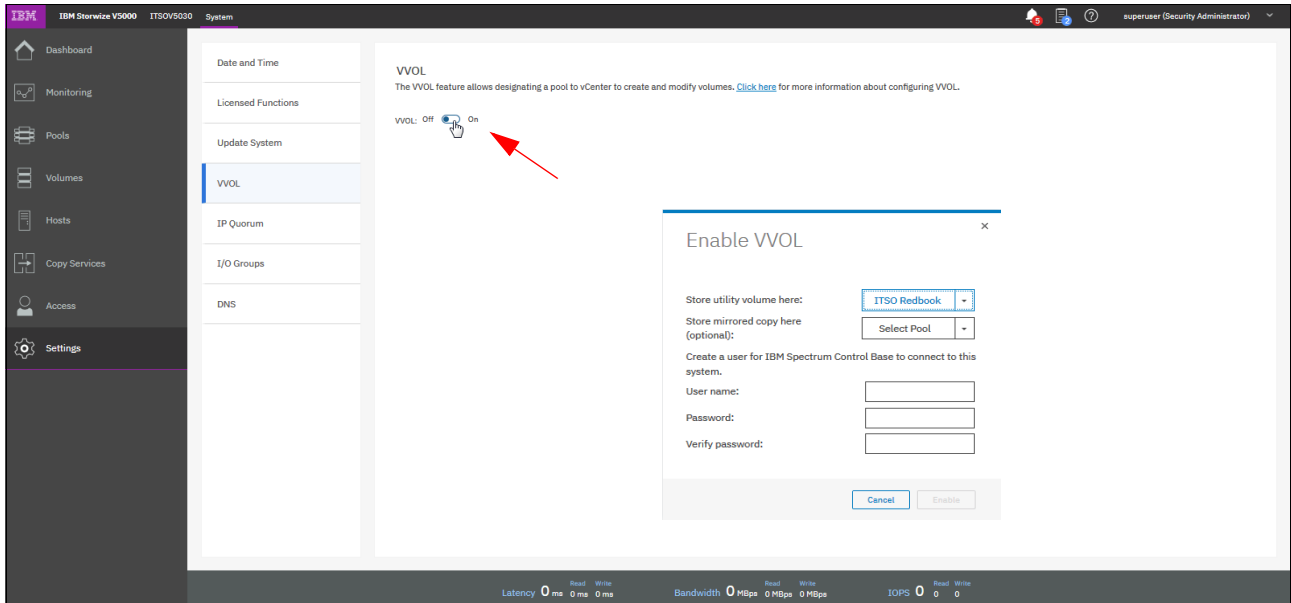


Figure 3-96 Activating VVol

In some HyperSwap configurations, IP quorum applications can be used at the third site as an alternative to third-site quorum disks. No Fibre Channel connectivity at the third site is required to use an IP quorum application as the quorum device. The IP quorum application is a Java application that runs on a host at the third site. The IP network is used for communication between the IP quorum application and node canisters in the system.

If you have a third-site quorum disk, you must remove the third site before you use an IP quorum application. The round-trip time limitations from 80 micro seconds for an IP quorum still exist. Figure 3-97 shows where you can download the Java application.

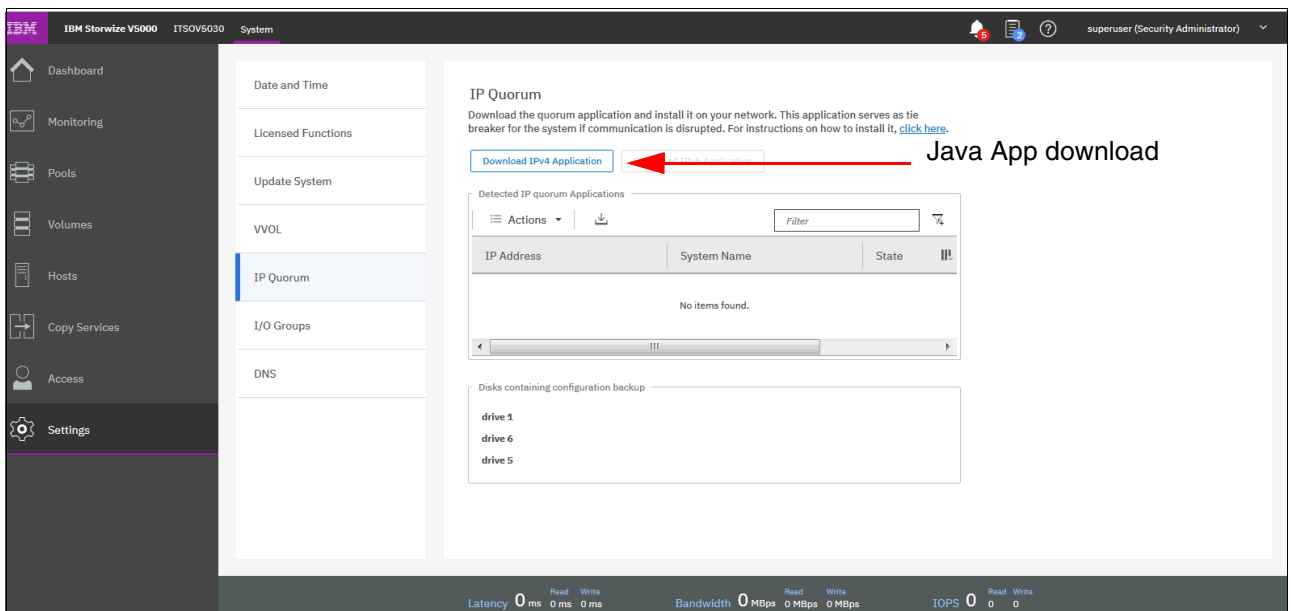


Figure 3-97 IP Quorum

For ports within an I/O group, you can enable virtualization of Fibre Channel ports that are used for host I/O operations. With N_Port ID virtualization (NPIV), the Fibre Channel port consists of a physical port and a virtual port.

When port virtualization is enabled, ports do not come up until they are ready to handle I/O, which improves host behavior around node unpendes. In addition, path failures that are caused by an offline node are masked from hosts. The target port mode on the I/O group indicates the current state of port virtualization.

Figure 3-98 shows the window with the I/O Groups.

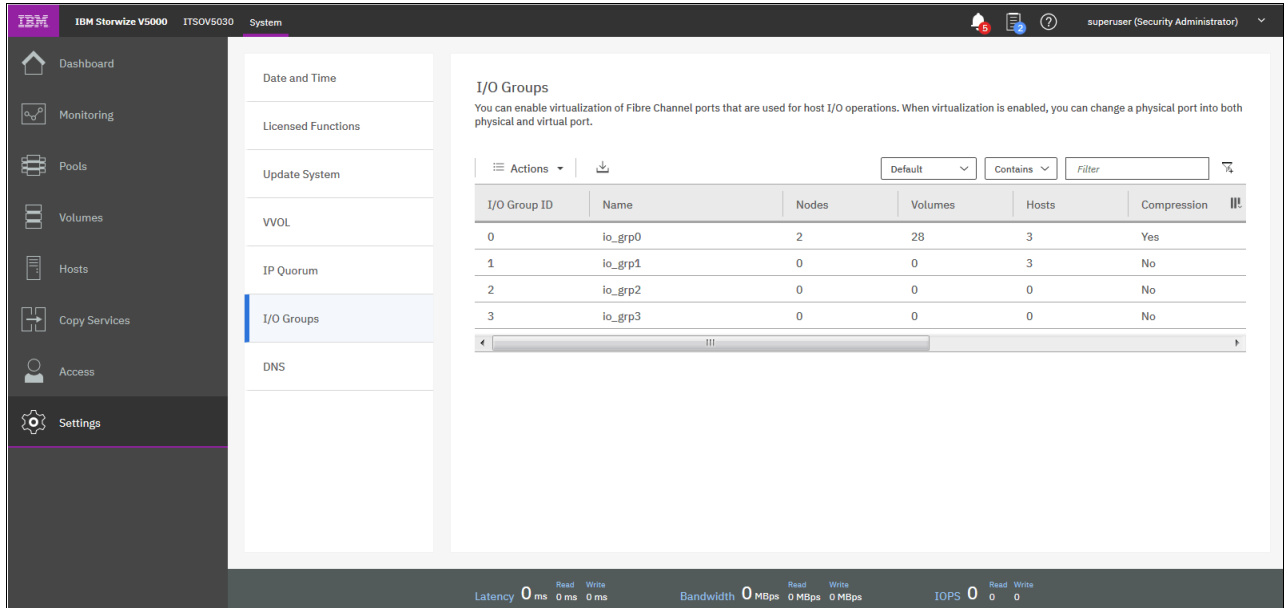


Figure 3-98 I/O Groups

Domain Name System (DNS) translates IP address to host names. You can create, delete, or change domain name servers, which manage names of resources that are on external networks.

You can have up to two DNS servers that are configured on the system. To configure DNS for the system, enter a valid IP address and name for each server. Both IPv4 and IPv6 address formats are supported.

Figure 3-99 shows the DNS setup Window.

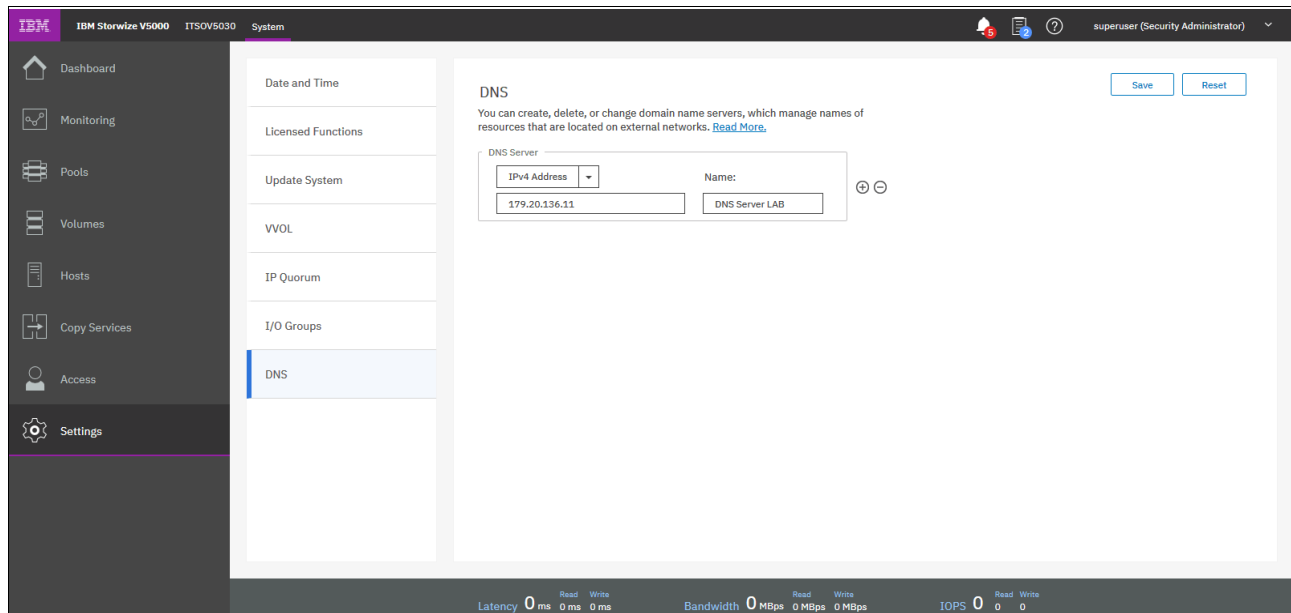


Figure 3-99 DNS

3.8.5 Call Home notifications

When an event is received, the Storwize 5000 Gen2 can automatically open a problem report. If appropriate, IBM contacts you to verify whether replacement parts are required. The configuration window for Call Home notifications is shown in Figure 3-100.

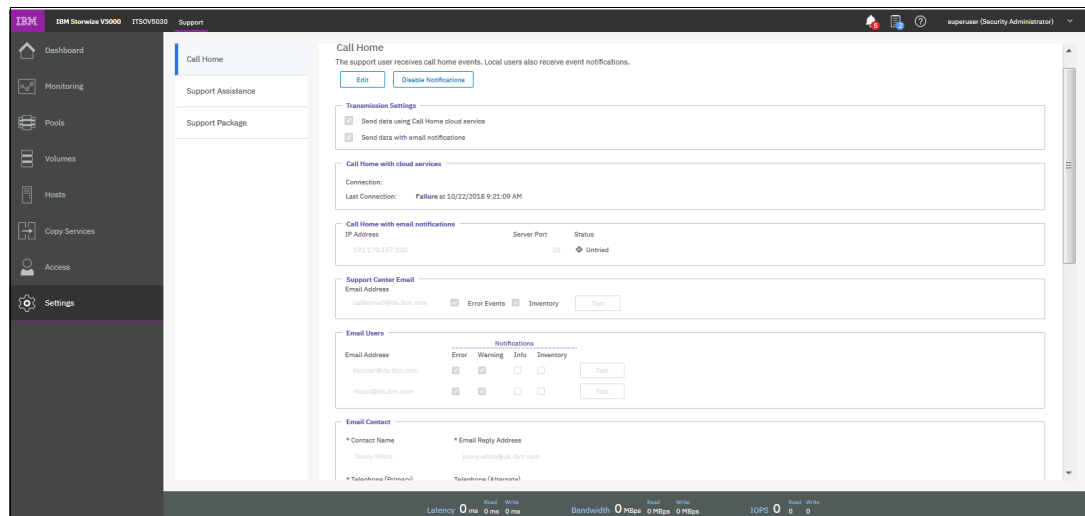


Figure 3-100 Call Home notifications

For more information about enabling Call Home notifications, see Chapter 12, “RAS, monitoring, and troubleshooting” on page 623.

Support assistance enables support personnel to access the system to complete troubleshooting and maintenance tasks. You can configure local support assistance, where support personnel visit your site to fix problems with the system, or remote support assistance.

Local and remote support assistance uses secure connections to protect data exchange between the support center and system. To enable Support assistance, you must enable an email Server. More access controls can be added by the system administrator. Figure 3-101 shows the Support assistance window.

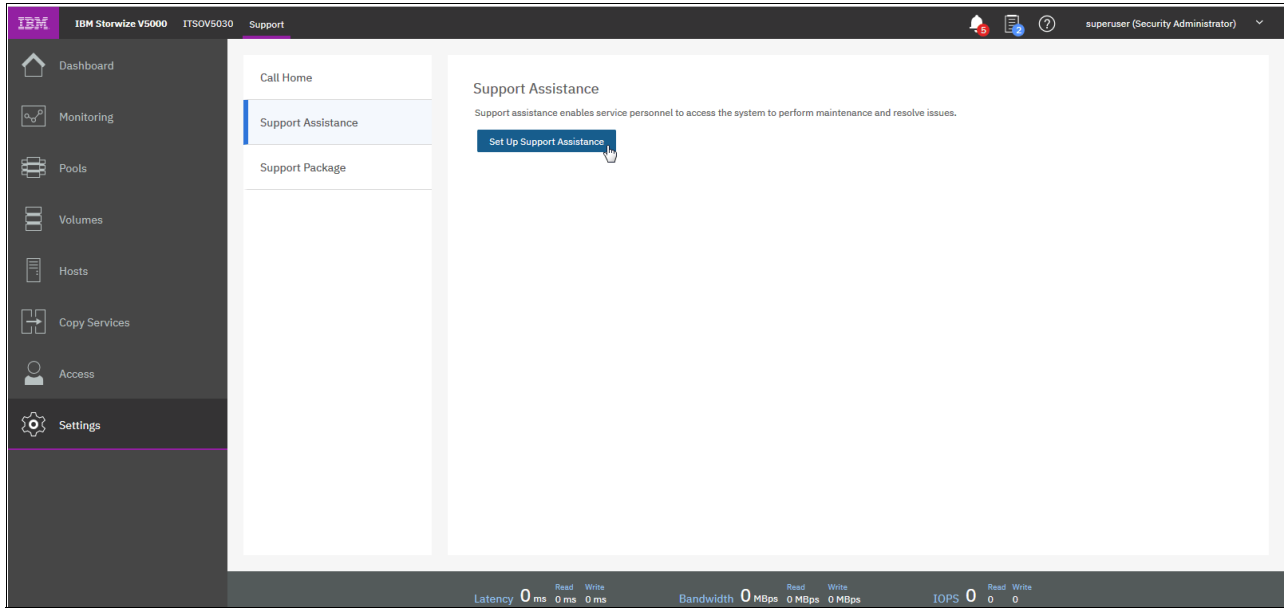


Figure 3-101 Support assistance

Also, if support assistance is configured on your system, you can automatically or manually upload new support packages to the support center to help analyze and resolve errors on the system. You can select individual logs to download to review or send directly to the support center for analysis. Figure 3-102 shows how to upload or download support logs.

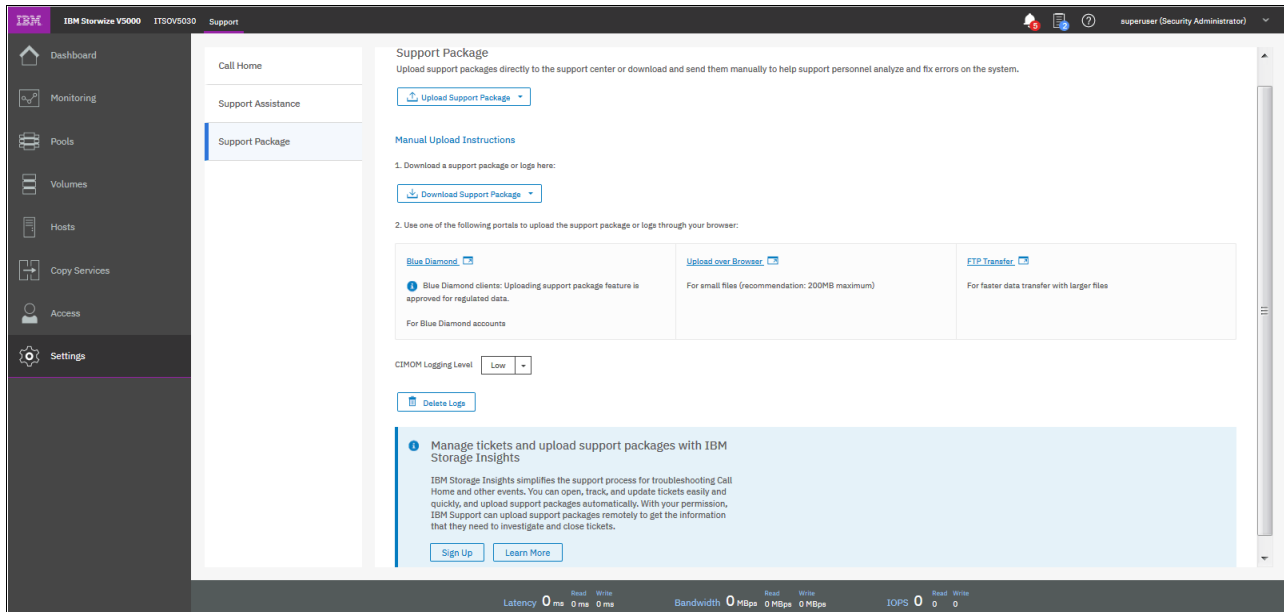


Figure 3-102 Up- or downloading support packages

For more information, see Chapter 12, “RAS, monitoring, and troubleshooting” on page 623.

3.8.6 GUI preferences menu

By using this menu, you can configure the appearance and behavior of the GUI. Click **GUI Preferences** in the Settings option of the Dynamic menu. To display the login message, select **Enable**. You can create a customized login message, as shown in Figure 3-103.

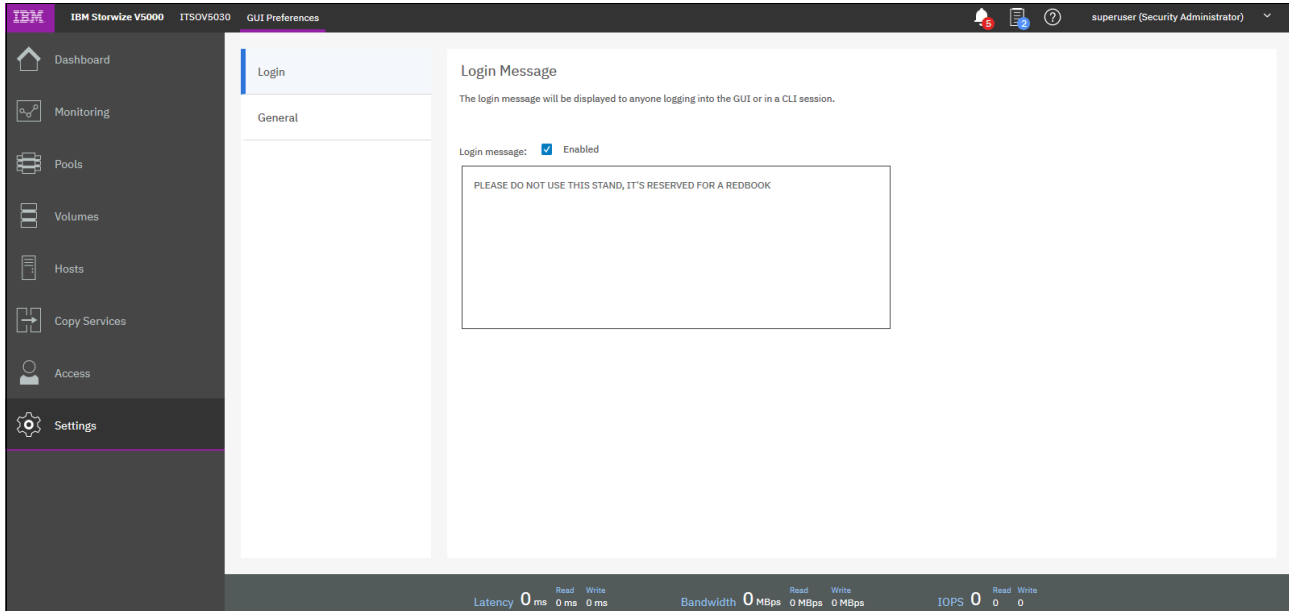


Figure 3-103 GUI preferences

Select **General** to adjust the browser settings, as shown in Figure 3-104.

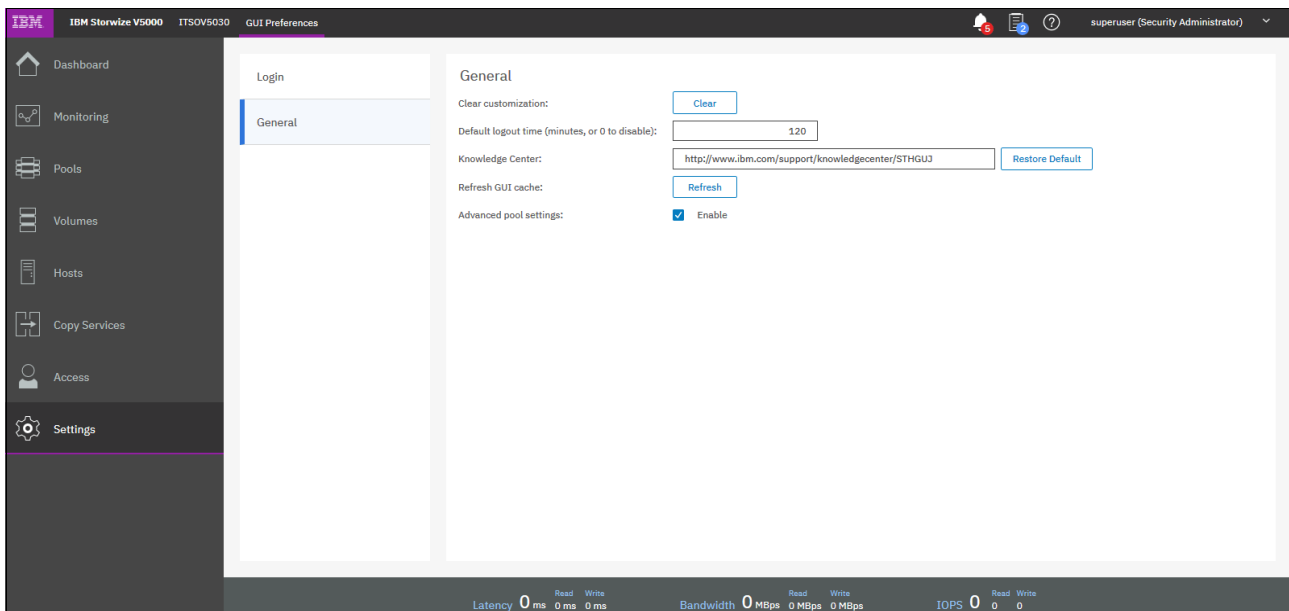


Figure 3-104 General settings



Storage pools

This chapter describes how the IBM Storwize V5000 Gen2 manages physical storage resources. All storage resources that are under IBM Storwize V5000 Gen2 control are managed by using *storage pools*. Storage pools dynamically allocate resources, maximize productivity, and reduce costs.

Internal and external managed disks (MDisks), advanced internal storage, and storage pool management are described in this chapter. For more information about external storage controllers, see Chapter 11, “External storage virtualization” on page 607.

Storage pools can be configured through the Initial Setup wizard when the system is first installed, as described in Chapter 2, “Initial configuration” on page 39. They can also be configured after the initial setup through the management GUI, which provides a set of presets to help you configuring different Redundant Array of Independent Disks (RAID) types.

The recommended configuration presets configure all drives into RAID arrays based on drive class and protect them with the correct number of spare drives. Alternatively, you can configure the storage to your own requirements. Selections include the drive class, the number of drives to configure, whether to configure spare drives, and optimization for performance or capacity.

This chapter includes the following topics:

- ▶ 4.1, “Working with internal drives” on page 160
- ▶ 4.2, “Working with storage pools” on page 172
- ▶ 4.3, “Working with managed disks” on page 188
- ▶ 4.4, “Working with external storage controllers” on page 215

4.1 Working with internal drives

This section describes how to configure the internal storage disk drives by using different RAID levels and optimization strategies. For more information about RAID settings, see section 4.3.2, “RAID configuration” on page 195.

The IBM Storwize V5000 Gen2 storage system provides an Internal Storage window for managing all internal drives. The Internal Storage window can be accessed by opening the System window, clicking the **Pools** option and then, clicking **Internal Storage**, as shown in Figure 4-1.

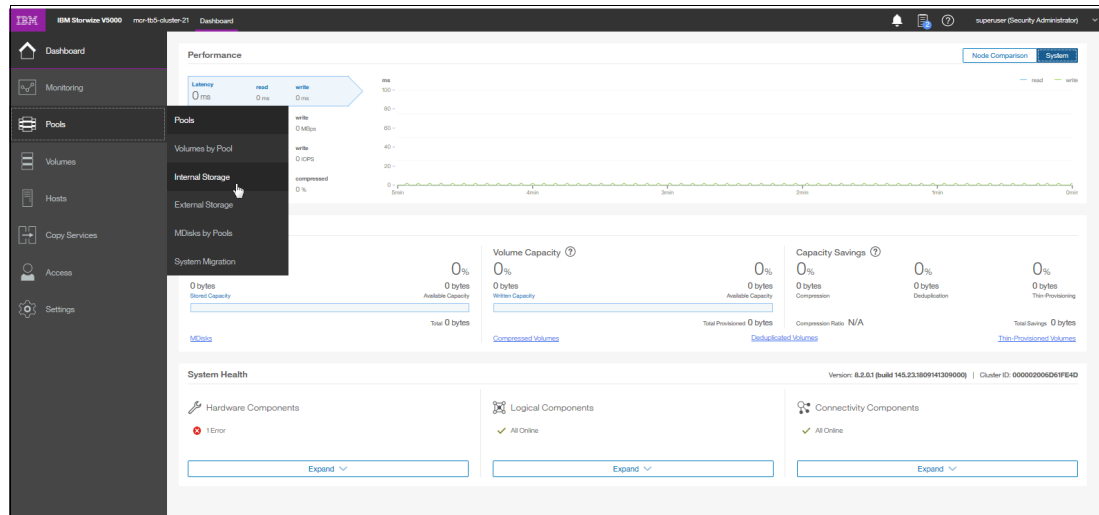


Figure 4-1 Path to Internal Storage window

4.1.1 Internal Storage window

The Internal Storage window (as shown in Figure 4-2 on page 161) provides an overview of the internal drives that are installed in the IBM Storwize V5000 Gen2 storage system. Selecting **All Internal** under the Drive Class Filter shows all the drives that are installed in the managed system, including attached expansion enclosures. Alternatively, you can filter the drives by their type or class. For example, you can choose to show only Enterprise drive class (serial-attached Small Computer System Interface [SCSI] or [SAS]), Nearline SAS, or Flash drives.

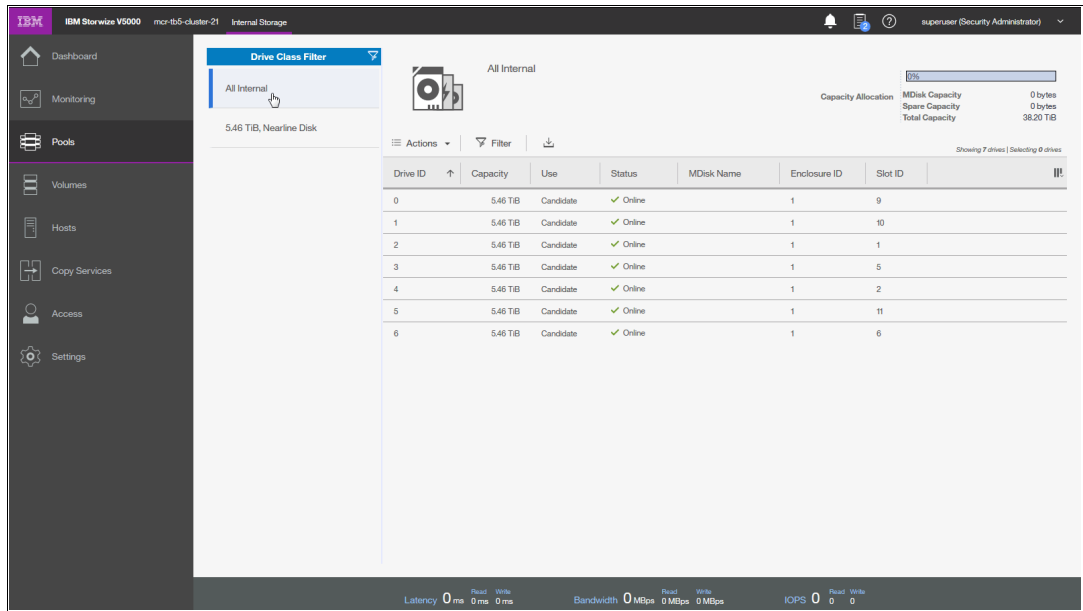


Figure 4-2 Internal Storage window

The right side of the Internal Storage window lists the selected type of internal disk drives. By default, the following information is listed:

- ▶ Logical drive ID
- ▶ Drive capacity
- ▶ Current type of use (unused, candidate, member, spare, or failed)
- ▶ Status (online, offline, and degraded)
- ▶ MDisk name that the drive is a member of
- ▶ Enclosure ID that the drive is installed in
- ▶ Slot ID of the enclosure in which the drive is installed

The default sort order is by enclosure ID. This default setting can be changed to any other column by left-clicking the column header. To toggle between ascending and descending sort order, left-click the column header again. By hovering over the header names, such as Drive ID, a brief description of the items within that column is displayed.

More columns can be included by right-clicking the gray header bar of the table, which opens the selection window, as shown in Figure 4-3. To restore the default column options, select **Restore Default View**.

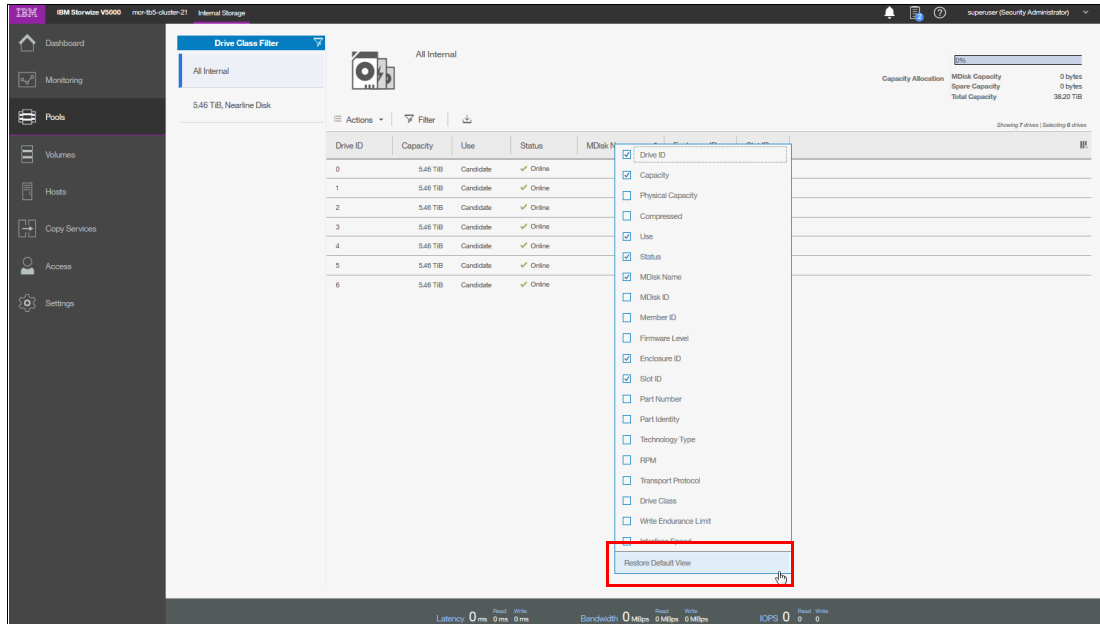


Figure 4-3 More column options for Internal Storage window

The overall internal storage capacity allocation indicator is shown in the upper-right corner. The *Total Capacity* shows the overall capacity of the internal storage that is installed in the IBM Storwize V5000 Gen2 storage system. The *MDisk Capacity* shows the internal storage capacity that is assigned to the MDisks. The *Spare Capacity* shows the internal storage capacity that is used for hot spare disks.

The percentage bar that is shown in Figure 4-4 indicates how much capacity is allocated.

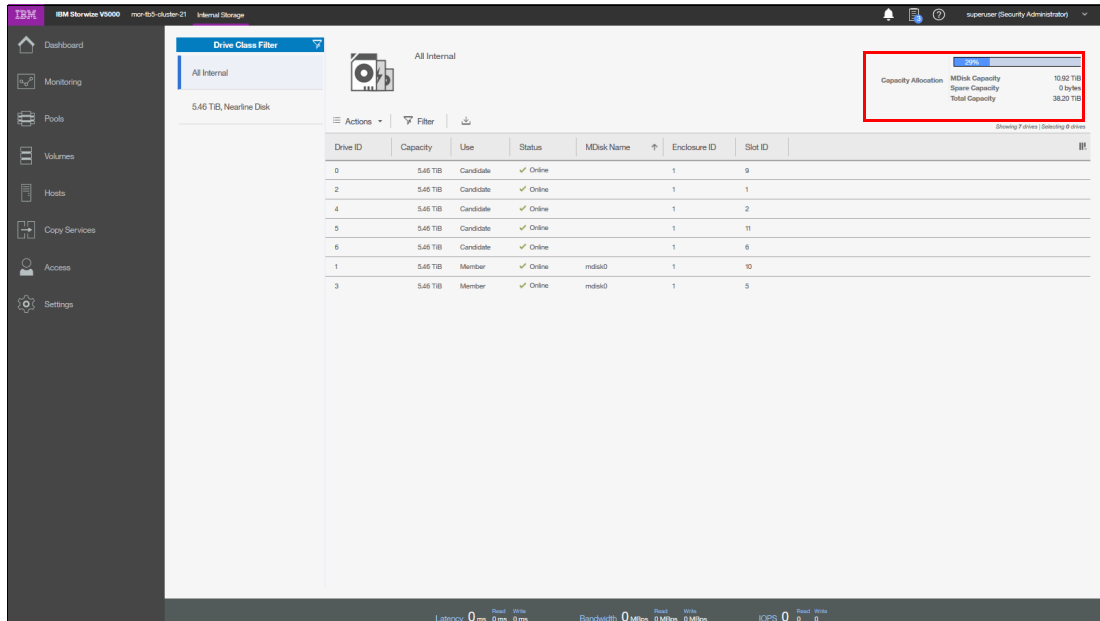


Figure 4-4 Internal Storage allocation indicator

4.1.2 Actions on internal drives

You can perform several actions by right-clicking the internal drives or clicking the **Actions** drop-down menu, as shown in Figure 4-5. If you click **Actions** without selecting any drive, the only options that are available are Upgrade All + Customize Columns.

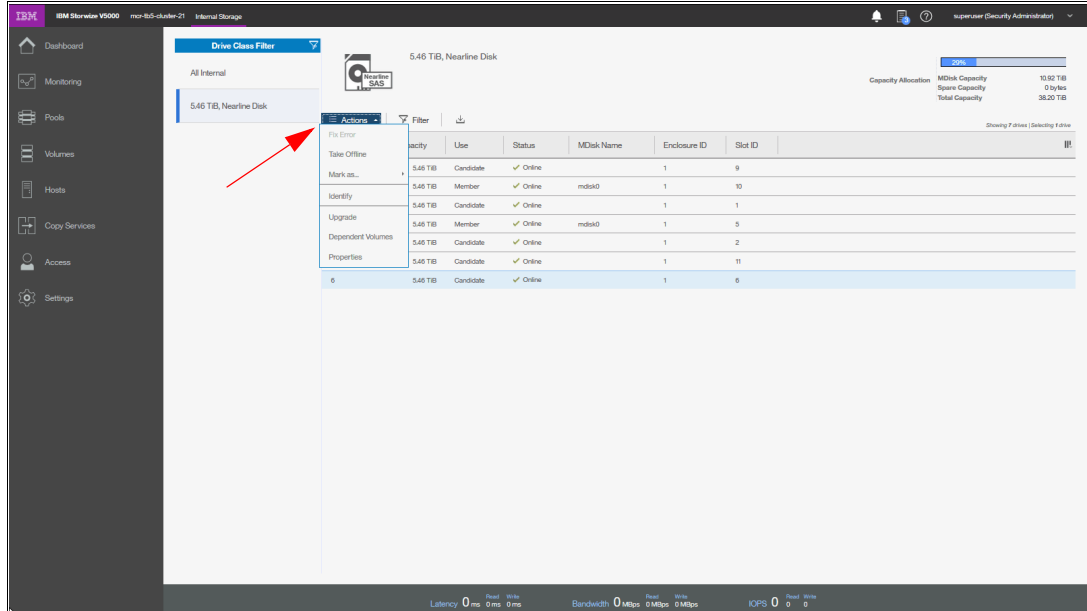


Figure 4-5 Internal drive actions menu

Depending on the status of the selected drive, several actions are available. These actions are described next.

Take Offline

The internal drives can be taken offline if a problem on the drive is identified. A confirmation window opens, as shown in Figure 4-6. The default selection is: only take a drive offline if a spare drive is available, which is strongly recommended and avoids redundancy loss in the *array*. Click **OK** to take the drive offline.

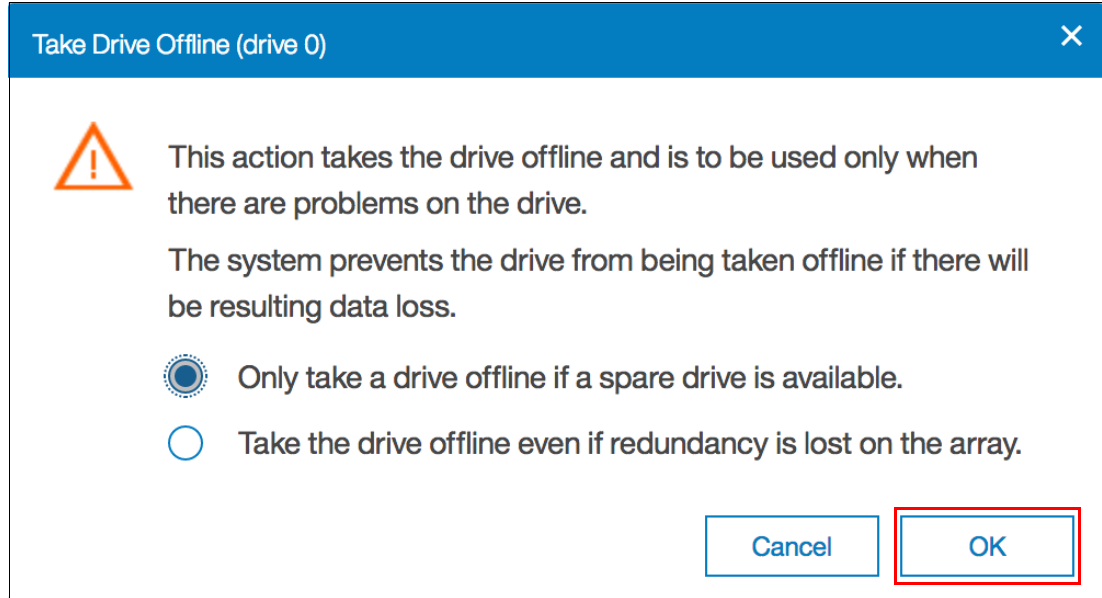


Figure 4-6 Warning before taking offline an internal drive

If the drive fails (as shown in Figure 4-7), the MDisk (from which the failed drive is a member) remains online and a hot spare is automatically reassigned.

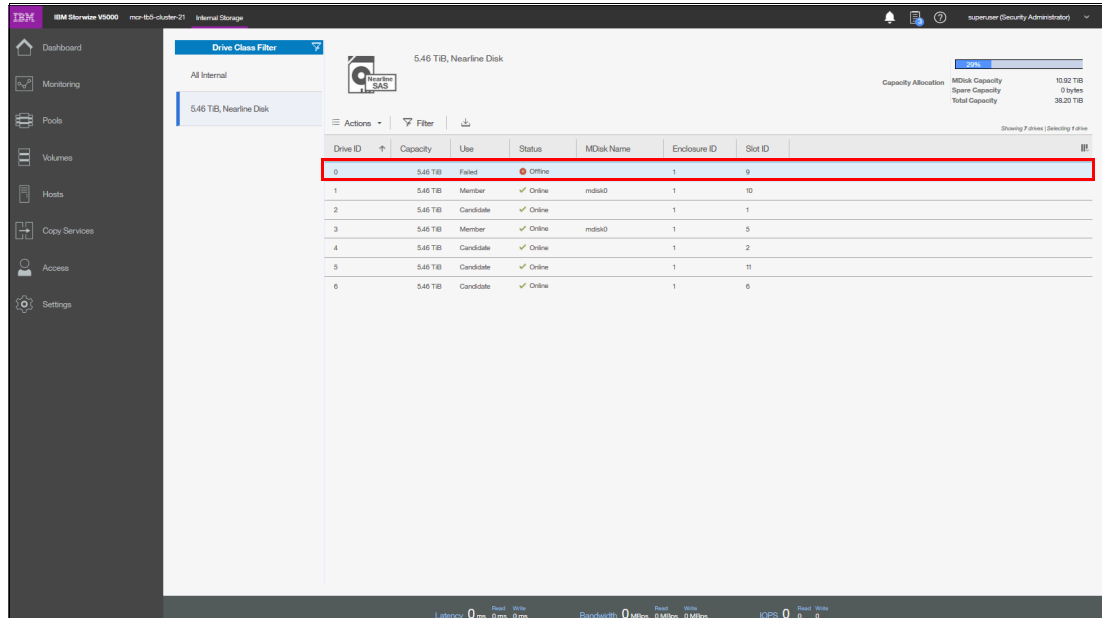


Figure 4-7 Internal drive taken offline

If sufficient spare drives are not available and a drive must be taken offline, the second option for no redundancy (Take the drive offline even if redundancy is lost on the array) must be selected. This option results in a degraded storage pool because of the degraded MDisk, as shown in Figure 4-8.

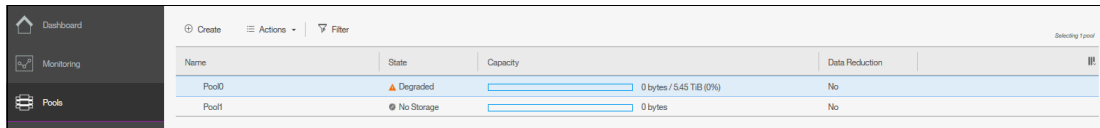


Figure 4-8 Degraded MDisk

The IBM Storwize V5000 Gen2 storage system prevents the drive from being taken offline if it can result in data loss. A drive cannot be taken offline (as shown in Figure 4-9) if no suitable spare drives are available and based on the RAID level of the MDisk, no sufficient redundancy is available. Click **Close** to return to the Internal Storage window.

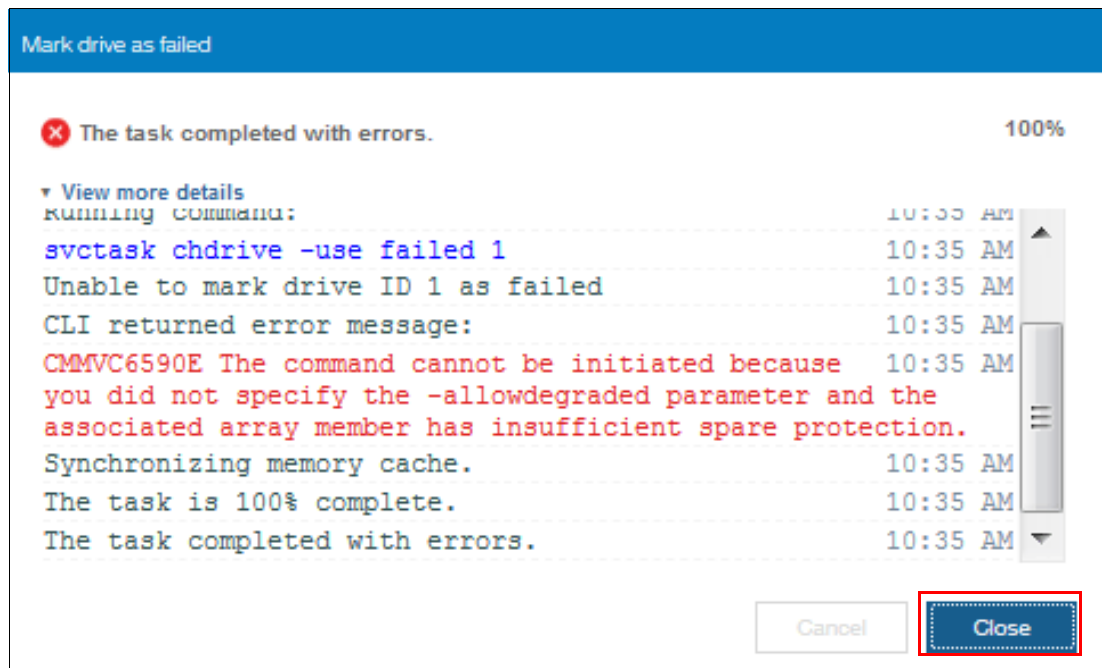


Figure 4-9 Internal drive offline not allowed because of insufficient redundancy

Example 4-1 shows how to use the **chdrive** command-line interface (CLI) command that is used to set the drive to failed.

Example 4-1 Use of the chdrive command to set the drive to failed

```

chdrive -use failed driveID
chdrive -use failed -allowdegraded driveID
  
```

Mark as option

The internal drives in the IBM Storwize V5000 Gen2 storage system can be assigned to the following usage roles by right-clicking the drives and selecting the **Mark as** option, as shown in Figure 4-10:

- ▶ **Unused:** The drive is not in use, and it cannot be used as a spare.
- ▶ **Candidate:** The drive is available for use in an array.
- ▶ **Spare:** The drive can be used as a hot spare, if required.

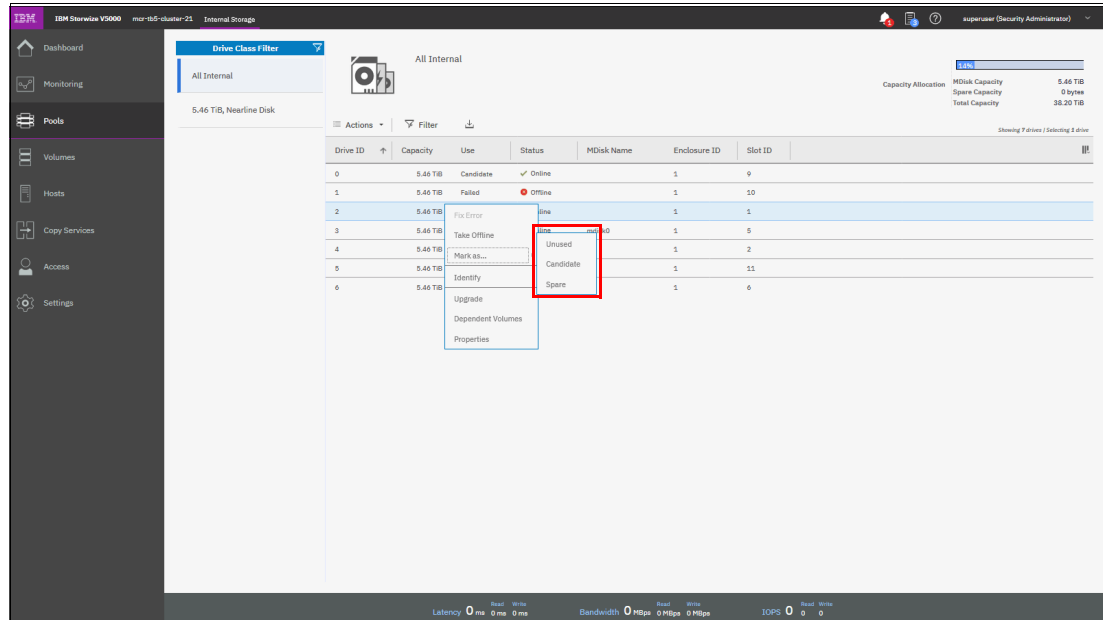


Figure 4-10 Selecting the internal drive “Mark as” action

Defining a new role to a drive depends on the current drive usage role. These dependencies are shown in Figure 4-11.

		To				
		Unused	Candidate	Failed	Member	Spare
From	Unused	allowed	allowed			not allowed
	Candidate	allowed	allowed	no option		allowed
	Failed	allowed	allowed			not allowed
	Member					
	Spare	not allowed	allowed	no option		allowed

Figure 4-11 Internal drive usage role table

Identify

Use the Identify action to turn on the LED light so that you can easily identify a drive that must be replaced or that you want to physically troubleshoot. The window that is shown in Figure 4-12 appears when the LED is on. Click **Turn Identify Off** when you are finished to turn off the drive LED and return to the Internal Storage window.

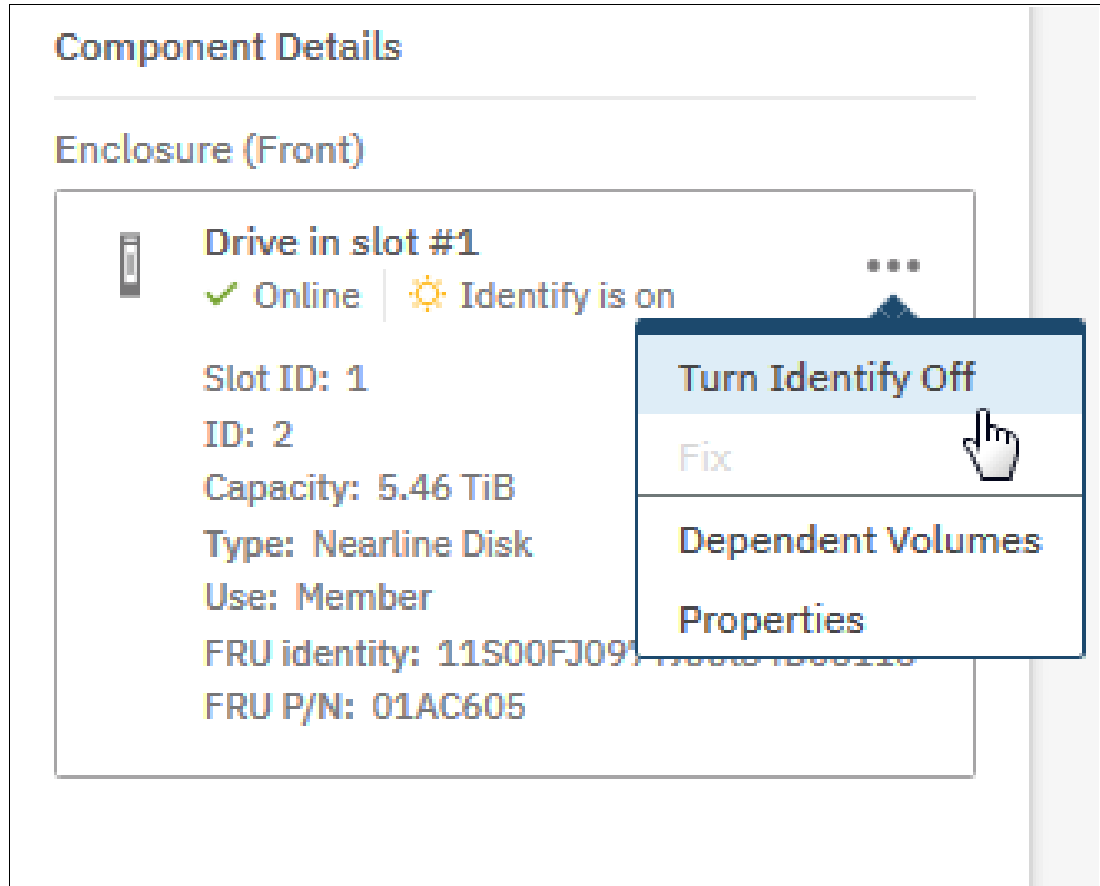


Figure 4-12 Internal drive identification

Example 4-2 shows how to use the `chenclosureslot` command to turn on and turn off the drive LED.

Example 4-2 Use of the `chenclosureslot` command to turn on and turn off the drive LED

```
chenclosureslot -identify yes/no -slot slot enclosureID
```

Upgrade

From this option, you can easily upgrade the drive firmware. You can use the GUI to upgrade individual drives or upgrade all drives for which updates are available. For more information about upgrading drive firmware, see Chapter 12, “RAS, monitoring, and troubleshooting” on page 623 and [this website](#).

Dependent Volumes

Clicking **Dependent Volumes** shows the volumes that depend on the selected drive. Volumes depend on a drive only when their underlying MDisks are in a degraded or inaccessible state and when the removal of more hardware causes the volume to go offline. This condition is true for any RAID 0 MDisk because it has no redundancy, or if the associated MDisk is already degraded.

Use the Dependent Volumes option before you perform any drive maintenance to determine which volumes are affected.

Important: A lack of listed dependent volumes does not imply that no volumes are using the drive.

Figure 4-13 shows an example if no dependent volumes are detected for a specific drive. If a dependent volume is identified, it is listed within this window. When volumes are listed as dependent, you can also check volume saving and throttle by selecting the volume and clicking **Actions**. Click **Close** to return to the Internal Storage window.

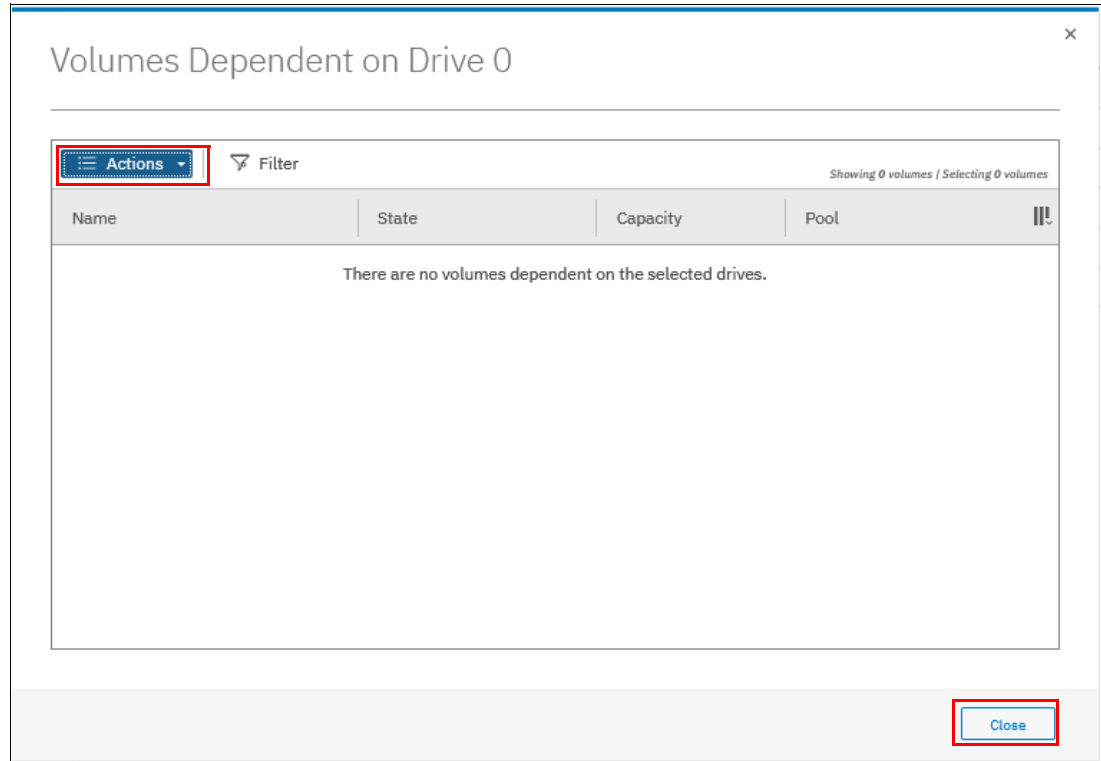


Figure 4-13 Internal drive with no dependent volumes

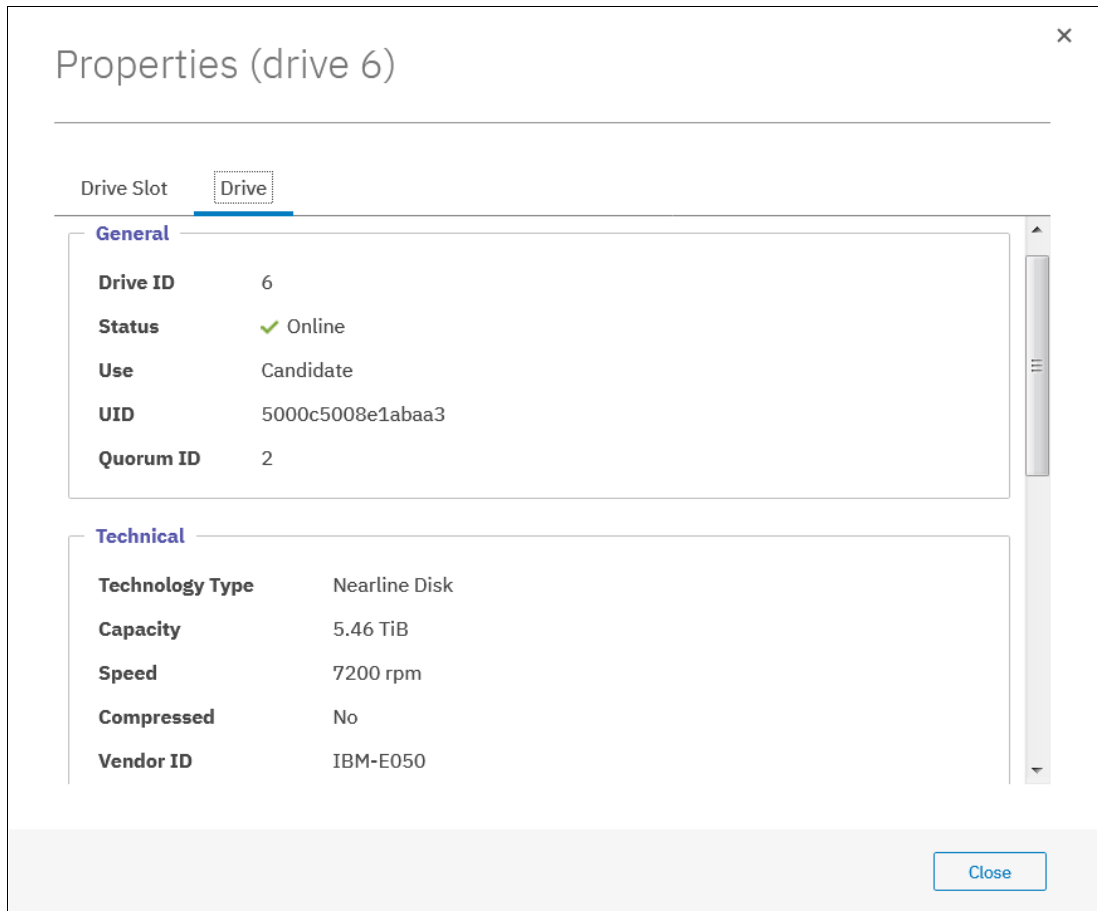
Example 4-3 shows how to view dependent volumes for a specific drive by using the CLI.

Example 4-3 Command to view dependent virtual disks (VDisks) for a specific drive

```
lsdependentvdisks -drive driveID
```

Properties

Clicking **Properties** in the Actions menu or double-clicking the drive provides the vital product data (VPD) and the configuration information, as shown in Figure 4-14. The detailed view opens.



The screenshot shows a window titled "Properties (drive 6)" with a close button in the top right corner. Below the title bar, there are two tabs: "Drive Slot" and "Drive", with "Drive" selected. The main content area is divided into two sections: "General" and "Technical".

General	
Drive ID	6
Status	✓ Online
Use	Candidate
UID	5000c5008e1abaa3
Quorum ID	2

Technical	
Technology Type	Nearline Disk
Capacity	5.46 TiB
Speed	7200 rpm
Compressed	No
Vendor ID	IBM-E050

A "Close" button is located at the bottom right of the window.

Figure 4-14 Detailed internal drive properties

A tab for the Drive Slot is available in the Properties window (as shown in Figure 4-15) to obtain specific information about the slot of the selected drive. Click **Close** to return to the Internal Storage window.

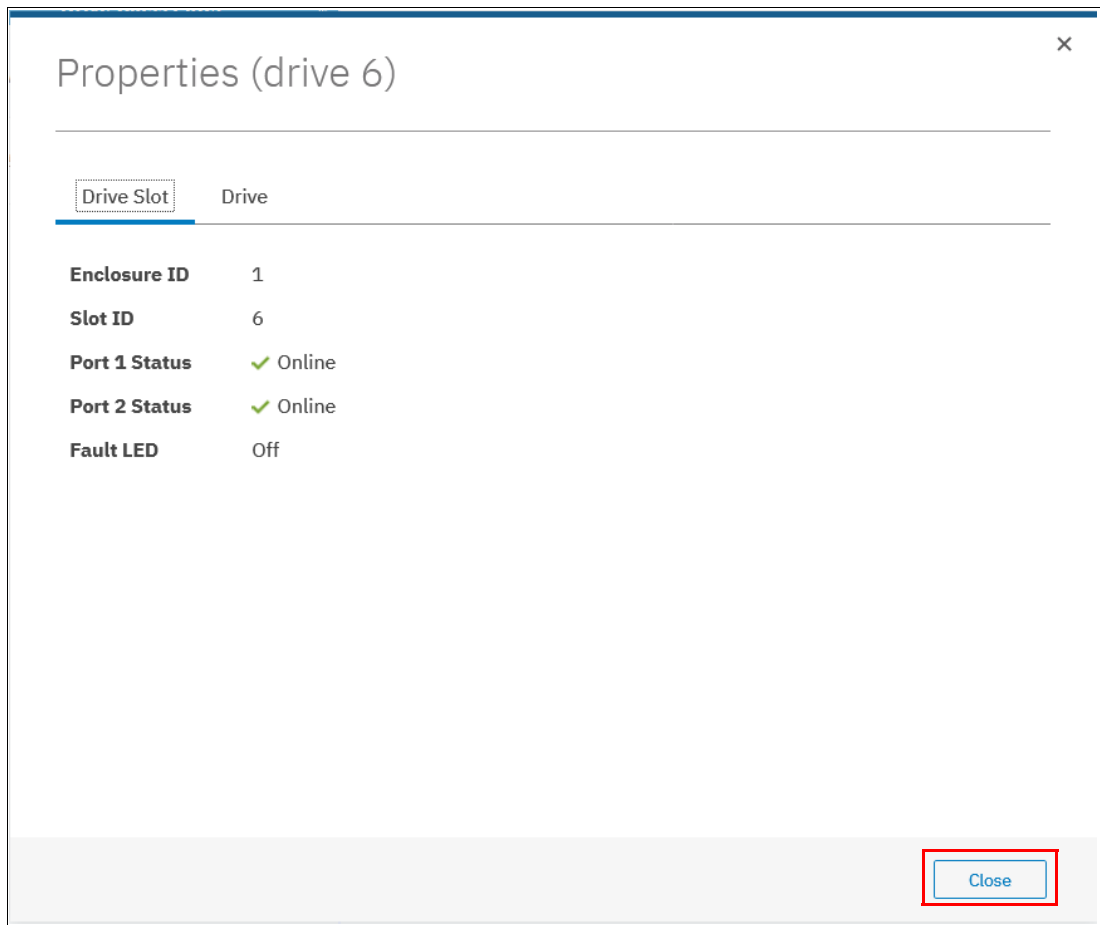


Figure 4-15 Internal drive properties slot

Example 4-4 shows how to use the **lsdrive** command to display the configuration information and drive VPD.

Example 4-4 Use of the lsdrive command to display configuration information and drive VPD

```
IBM_Storage:ITS0 V5000:superuser>lsdrive 1
id 5
status online
error_sequence_number
use candidate
UID 5000c5008e1abaa3
tech_type Nearline Disk
capacity 5.46 TiB
block_size 512
vendor_id IBM-E050
product_id ST600NM0014
FRU_part_number 01AC605
FRU_identity 11S00FJ097YXXS4D080QZ
RPM 150007200
firmware_level BC78
FPGA_level
```



```

mdisk_id
mdisk_name
member_id
enclosure_id 1
slot_id 6
node_id
node_name
quorum_id
port_1_status online
port_2_status online
interface_speed 12Gb
protection_enabled yes
auto_manage inactive
drive_class_id 145
IBM_Storwize:ITS0 V5000:superuser>

```

Customize Columns option

Click **Customize Columns** in the Actions menu to add or remove several columns that are available in the Internal Storage window.

To restore the default column options, select **Restore Default View**, as shown in Figure 4-16.

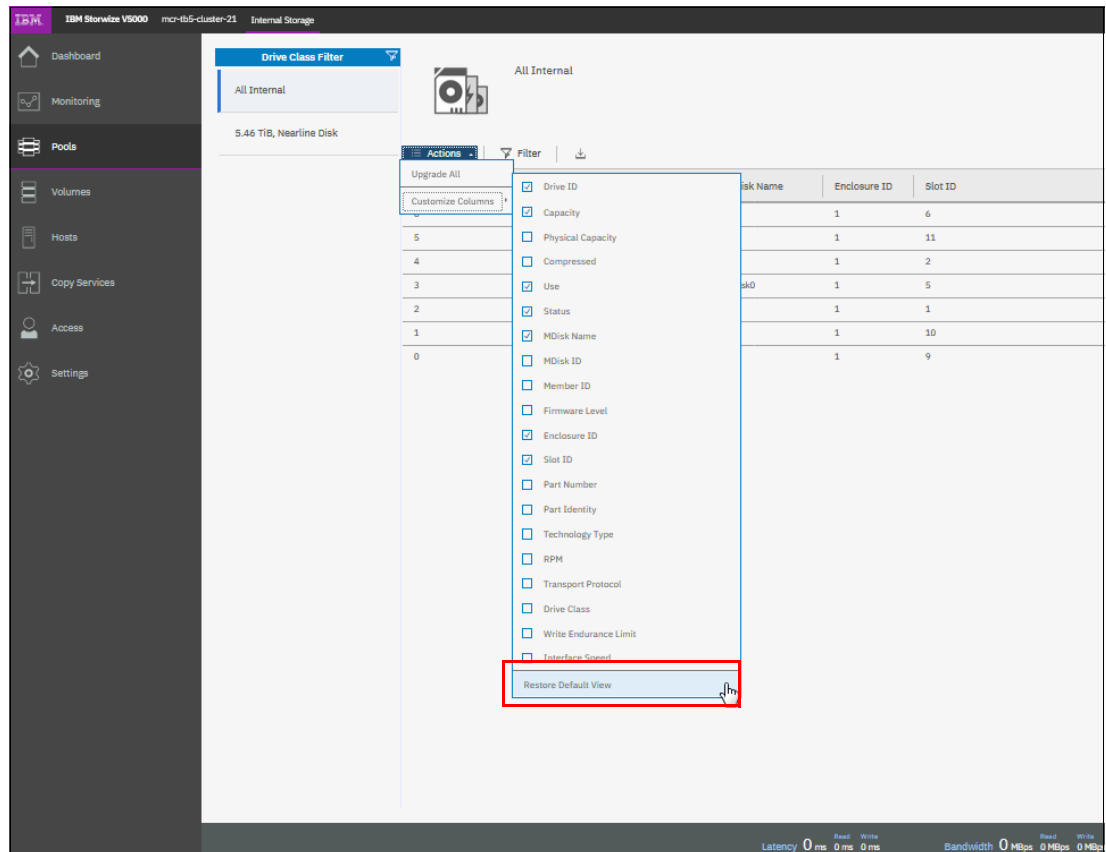


Figure 4-16 Customizing columns on the Internal Storage window

4.2 Working with storage pools

Storage pools (or *pools*) act as containers for MDisks and provision the capacity to volumes. MDisks can be provisioned through internal or external storage. MDisks created from internal storage are created as RAID arrays.

IBM Storwize V5000 Gen2 organizes storage into pools to ease storage management and make it more efficient. All MDisks in a pool are split into extents of the same size and volumes are created from these available extents. The extent size is a property of the storage pool. When an MDisk is added to a pool, the size of the extents that composes it is based on the attribute of the pool to which the MDisk was added.

Storage pools can be further divided into subcontainers named as *child pools*. Child pools inherit the properties of the parent pool and can also be used to provision volumes.

Storage pools are managed by using the Pools window or the MDisks by Pool window. Both windows allow you to run the same actions; however, actions on child pools can be performed only by using the Pools window. To access the Pools window, click **Pools** → **Pools**, as shown in Figure 4-17.

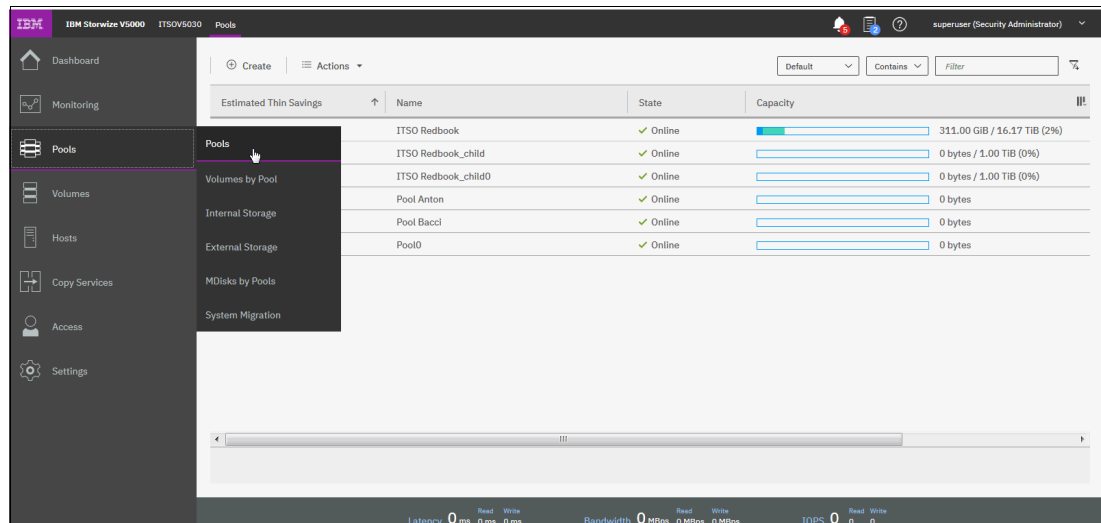


Figure 4-17 Pools window

All available storage pools in the system are listed in the window. If a storage pool has child pools, you can toggle the sign to the left of the storage pool icon to show or hide the child pools.

4.2.1 Creating storage pools

If you are installing a new IBM Storwize V5000 Gen2, no pools are created when you first log in; therefore, the system automatically suggests a pool creation, which leads directly to the Create Pool window. You can access the Pools window in the future through the Pools menu, as shown in Figure 4-17.

To create a storage pool, you can use one of the following methods:

- Browse to **Pools** → **Pools** and click **Create**, as shown in Figure 4-18.

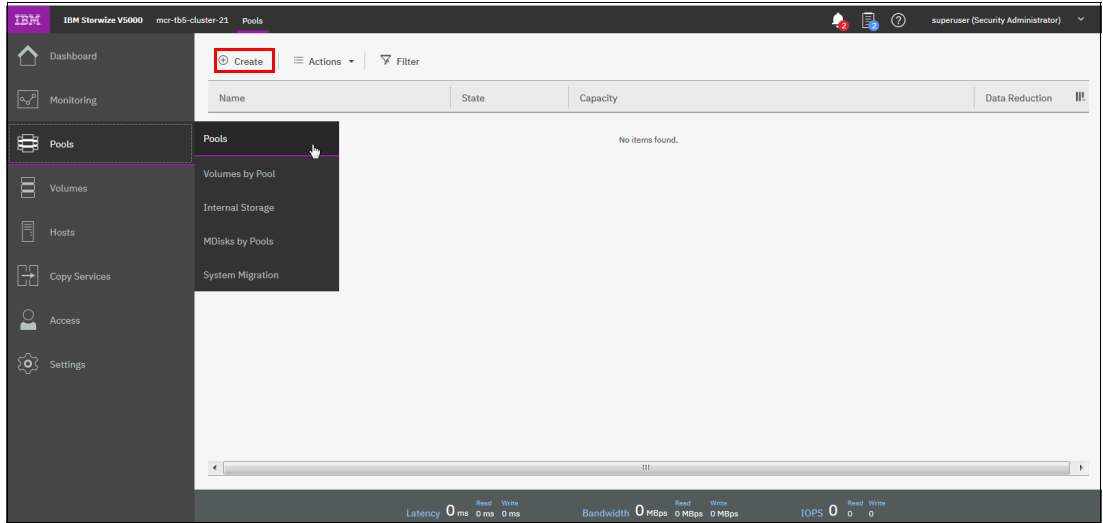


Figure 4-18 **Create** button on Pools window

- Browse to **Pools** → **MDisks by Pools** and click **Create Pool**, as shown in Figure 4-19.

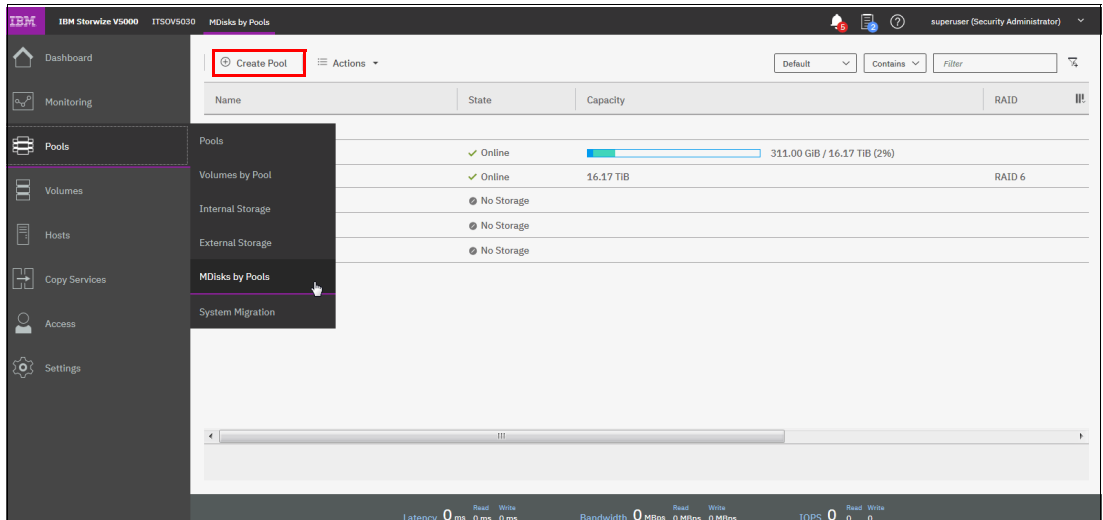


Figure 4-19 **Create Pool** button on MDisks by Pools window

Both of these alternatives open the dialog box that is shown in Figure 4-20.

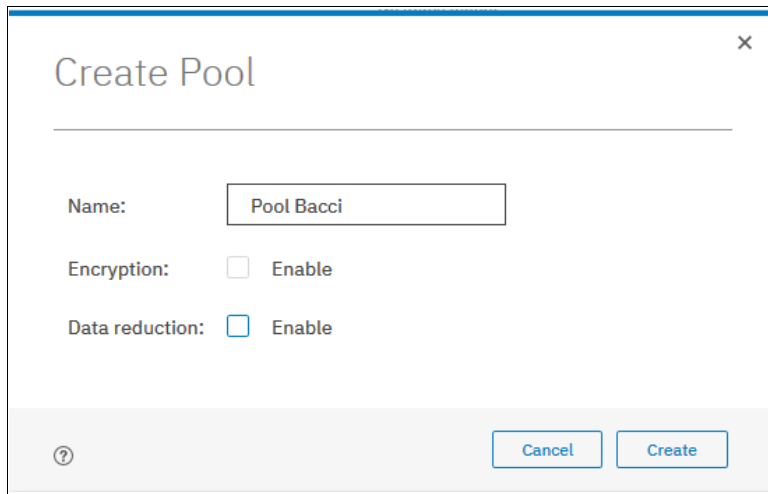


Figure 4-20 Create Pool dialog box

Encryption: Consider the following points:

- ▶ If encryption is enabled, you also can select whether the storage pool is encrypted. The encryption setting of a storage pool is selected at creation time and cannot be changed later. By default, if encryption is enabled, encryption is selected.
- ▶ Select the Data Reduction option to reclaim capacity from this pool for other uses if hosts that use the pool support Unmap. If the physical capacity usage of a data reduction pool exceeds more than 85%, I/O performance can be affected. The system needs 15% of physical capacity available in data reduction pools to ensure that capacity reclamation can be performed efficiently. For more information about Data Reduction Pools, see Chapter 9, “Advanced features for storage efficiency” on page 435.

If advanced pool settings are enabled, you also can select an extent size at the time of the pool creation, as shown in Figure 4-21 on page 175.

Note: Every storage pool that is created by using the GUI has a default extent size of 1 GB. The size of the extent is selected at creation time and cannot be changed later. If you want to specify a different extent size at time of the pool creation, select it in Advanced Pool settings mode. If you select a different Extent size, the GUI shows you also the maximum addressable capacity with each Extent size.

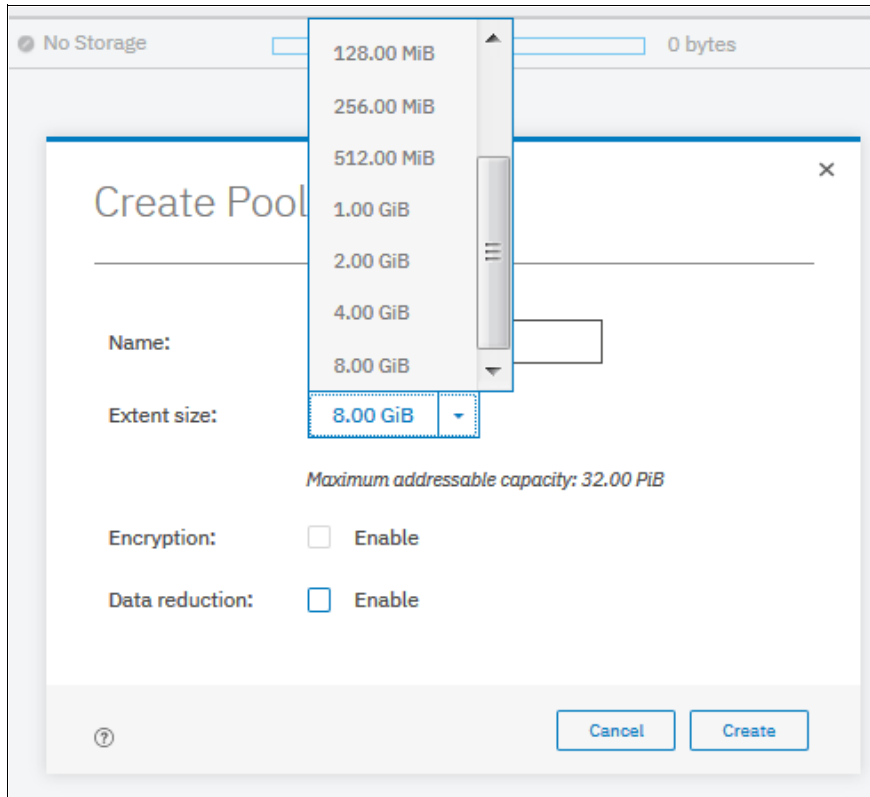


Figure 4-21 Creating pool with advanced pool settings enabled

To enable the advanced setting, click **Settings** → **GUI Preferences** → **General** → **Advanced pool setting** and click **Enable** (see Figure 4-22).

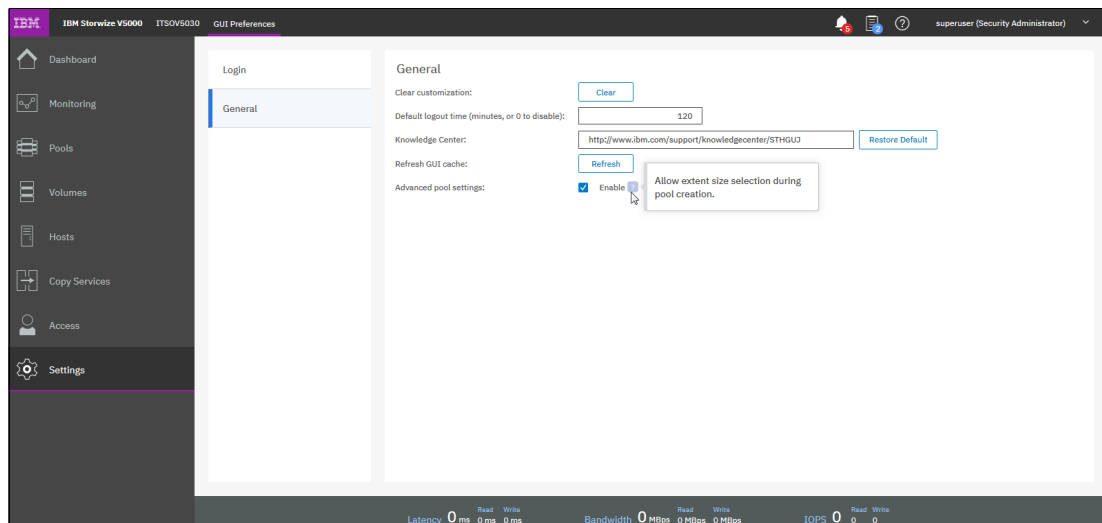


Figure 4-22 Enable the Advanced Pool settings

In the Create Pool dialog box, enter the pool name and click **Create**. The new pool is created and is included in the list of storage pools with zero bytes, as shown in Figure 4-23.

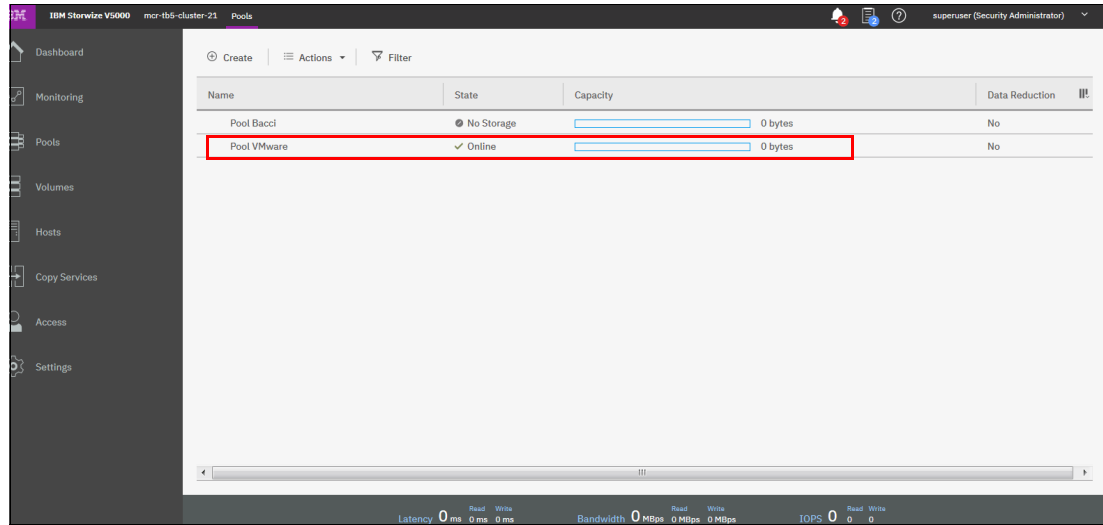


Figure 4-23 New pool with zero bytes included in the list

4.2.2 Actions on storage pools

Several actions can be performed on storage pools, which can be accessed through the Pools window or the MDisks by Pools window. To select an action, select the storage pool and click **Actions**. Alternatively, right-click the storage pool.

Figure 4-24 shows the list of available actions for storage pools that are accessed by using the Pools window.

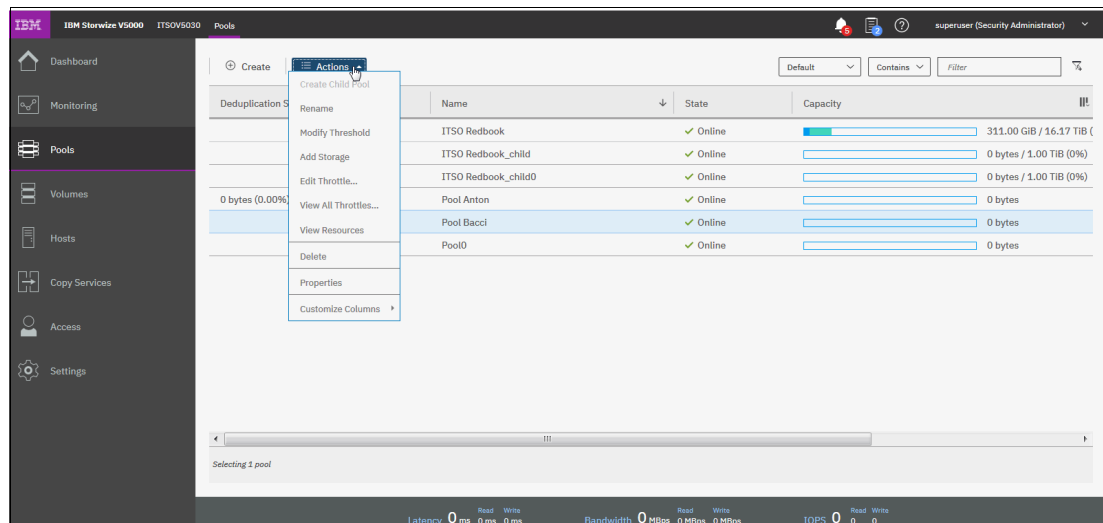


Figure 4-24 Actions list for storage pools

Create child pool option

Selecting **Create Child Pool** starts the wizard to create a child storage pool. For more information about child storage pools and this wizard, see 4.2.3, “Child storage pools” on page 184.

Rename option

Selecting **Rename** at anytime allows you to modify the name of a storage pool, as shown in Figure 4-25. Enter the new name and click **Rename**.

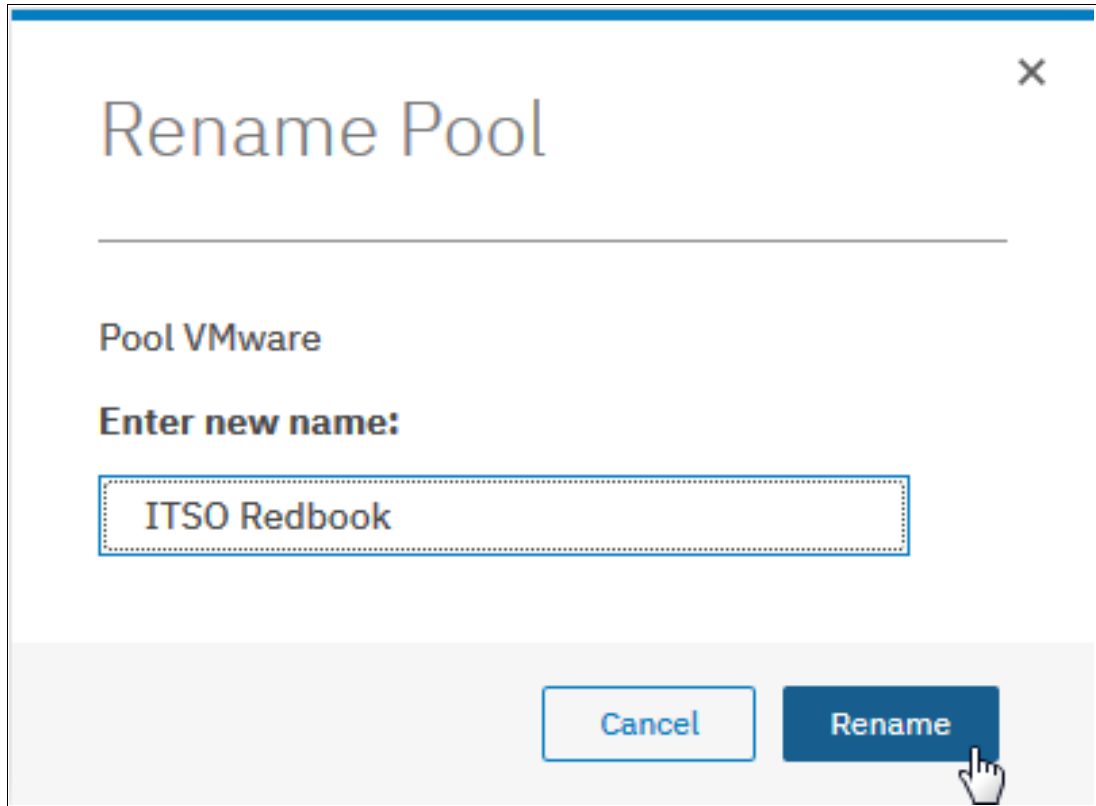


Figure 4-25 Renaming pools

Modify threshold

The storage pool threshold refers to the percentage of storage capacity that must be in use for a warning event to be generated. The threshold is especially useful when thin-provisioned volumes are used that are configured to expand automatically.

The threshold can be modified by selecting **Modify Threshold** and entering the new value, as shown in Figure 4-26. The default threshold is 80%. Warnings can be disabled by setting the threshold to 0%.

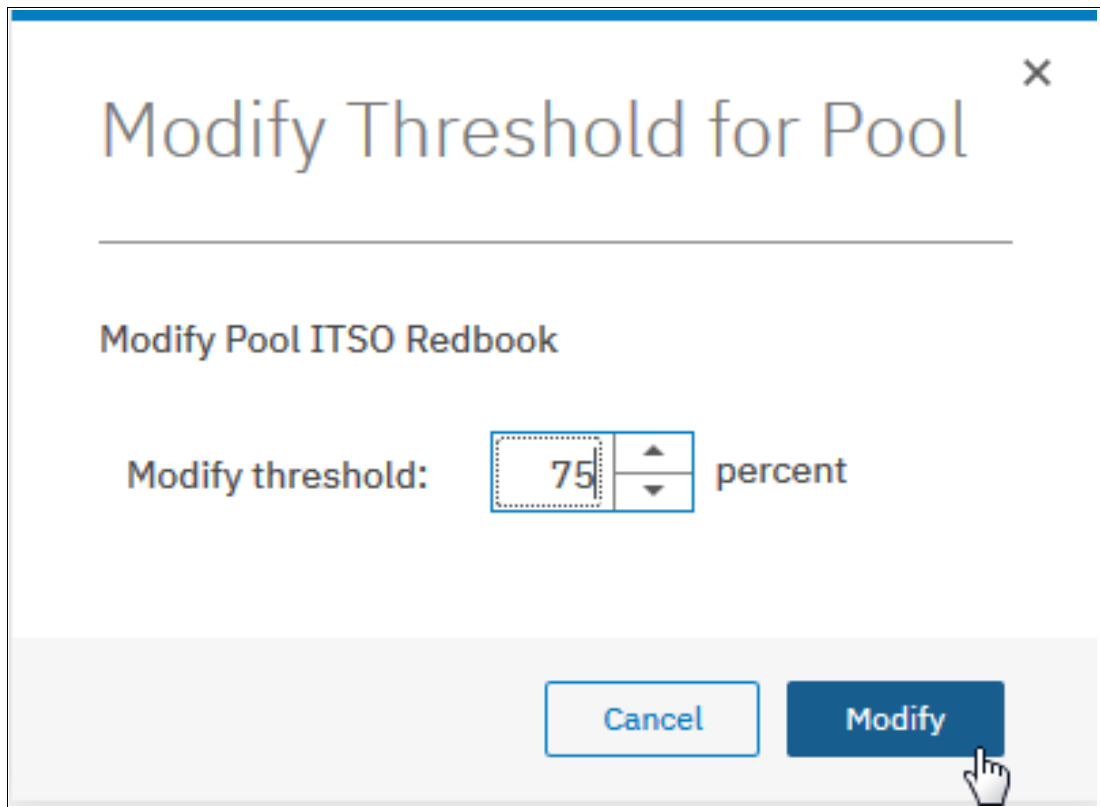


Figure 4-26 Modifying pool threshold

Add storage option

Selecting **Add Storage** starts the wizard to assign storage to the pool. For more information about this wizard, see 4.3.1, “Assigning managed disks to storage pools” on page 189.

Edit Throttle

You can create, modify, and remove throttles for pools by using the management GUI or the command-line interface. *Throttling* is a mechanism to control the amount of resources that are used when the system is processing I/Os on a specific pool. If a throttle is defined, the system processes the I/O or delays the processing of the I/O to free resources for more critical I/O.

The following parameters can be defined by using the Edit Throttle option:

- ▶ Bandwidth limit defines the maximum amount of bandwidth the pool can process before the system delays I/O processing for this pool.
- ▶ IOPS limit defines the maximum I/O operations per second that the pool can process before the system delays I/O processing for this pool.

If the pool does not have throttle settings configured, selecting **Edit Throttle** displays a dialog box with blank fields, as shown in Figure 4-27. Define the limits and click **Create**.

×

Edit Throttle for Pool

Modify Pool ITSO Redbook

Bandwidth limit:

Not enabled MBps ▼ Create

IOPS limit:

Not enabled IOPS Create

Close

Figure 4-27 Edit throttle initial configuration

For a pool that includes defined throttle settings, selecting **Edit Throttle** displays a different dialog box in which the current bandwidth and IOPS limits are displayed, as shown in Figure 4-28. You can change or remove the current bandwidth and IOPS limits by modifying the values and clicking **Save** or clicking **Remove** to disable a limitation.

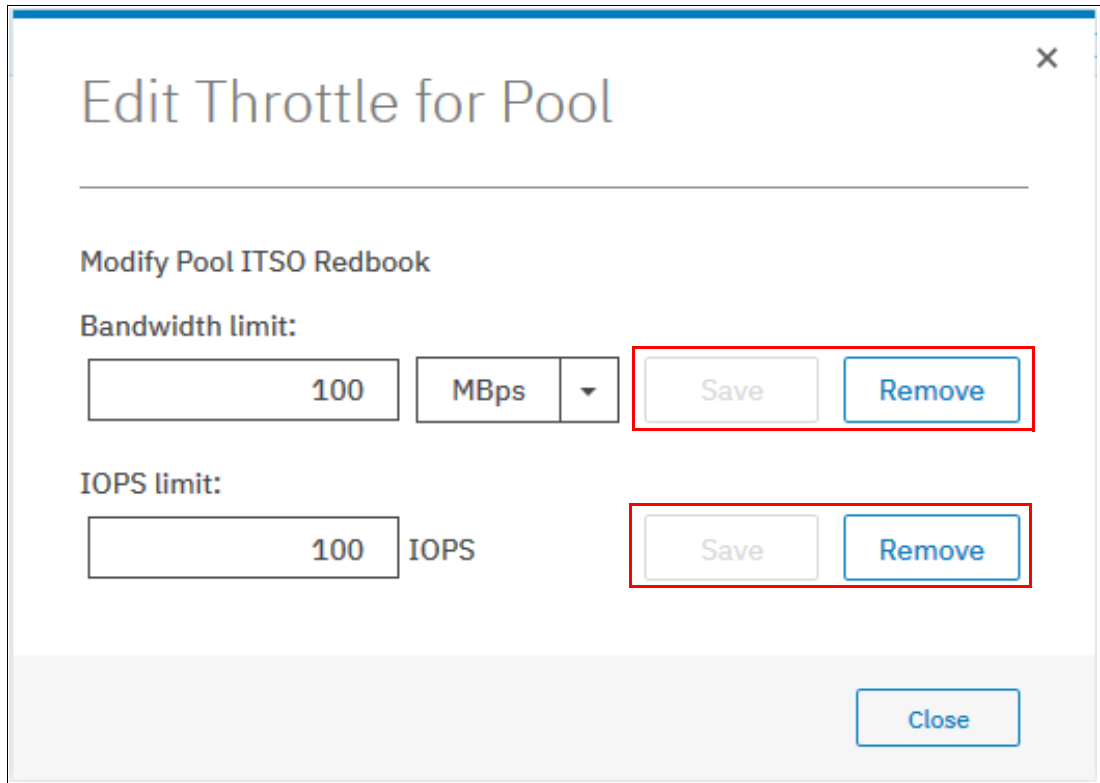
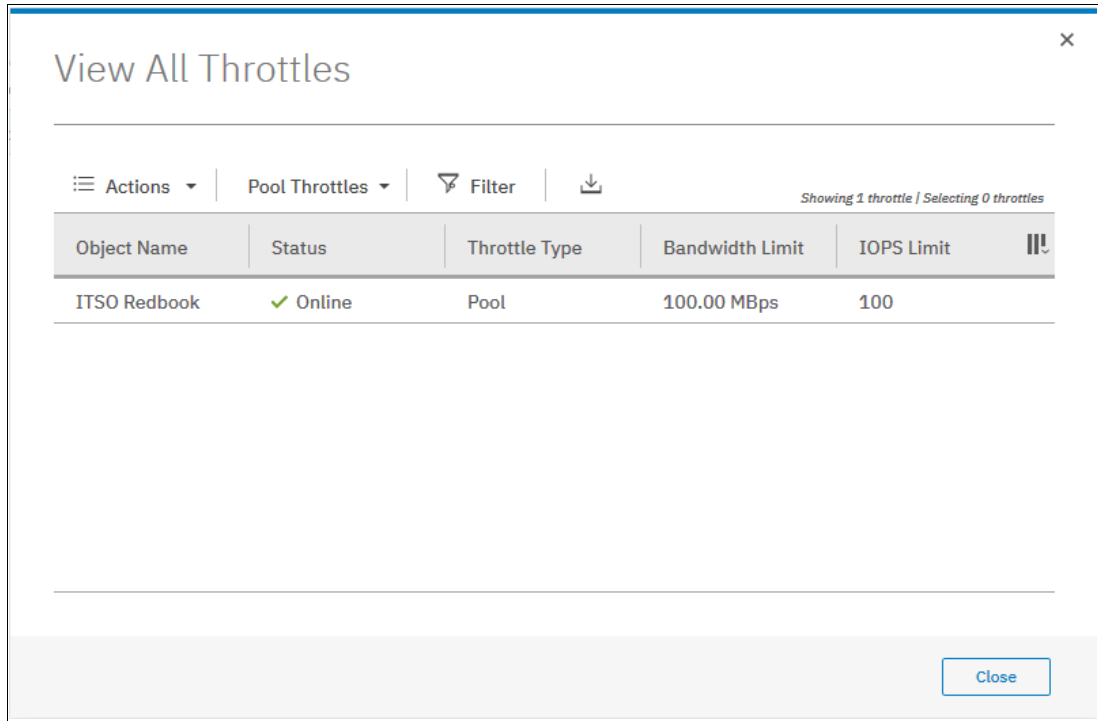


Figure 4-28 Editing throttles

View all Throttles option

Selecting **View All Throttles** opens a window (see Figure 4-29) that displays the current throttle information, which includes the limits that were applied for bandwidth and IOPS.



The screenshot shows a window titled "View All Throttles" with a close button (X) in the top right corner. Below the title bar is a navigation area with "Actions", "Pool Throttles", "Filter", and a download icon. A status indicator shows "Showing 1 throttle | Selecting 0 throttles". The main content is a table with the following data:

Object Name	Status	Throttle Type	Bandwidth Limit	IOPS Limit	
ITSO Redbook	✓ Online	Pool	100.00 MBps	100	

A "Close" button is located in the bottom right corner of the window.

Figure 4-29 View All Throttles window

By default, when the View All Throttles window is opened through the Pools window, it displays throttle information that is related to pools, but through the same window you can select different objects, as shown in Figure 4-30. Selecting a different category displays the throttle information for that specific selection.

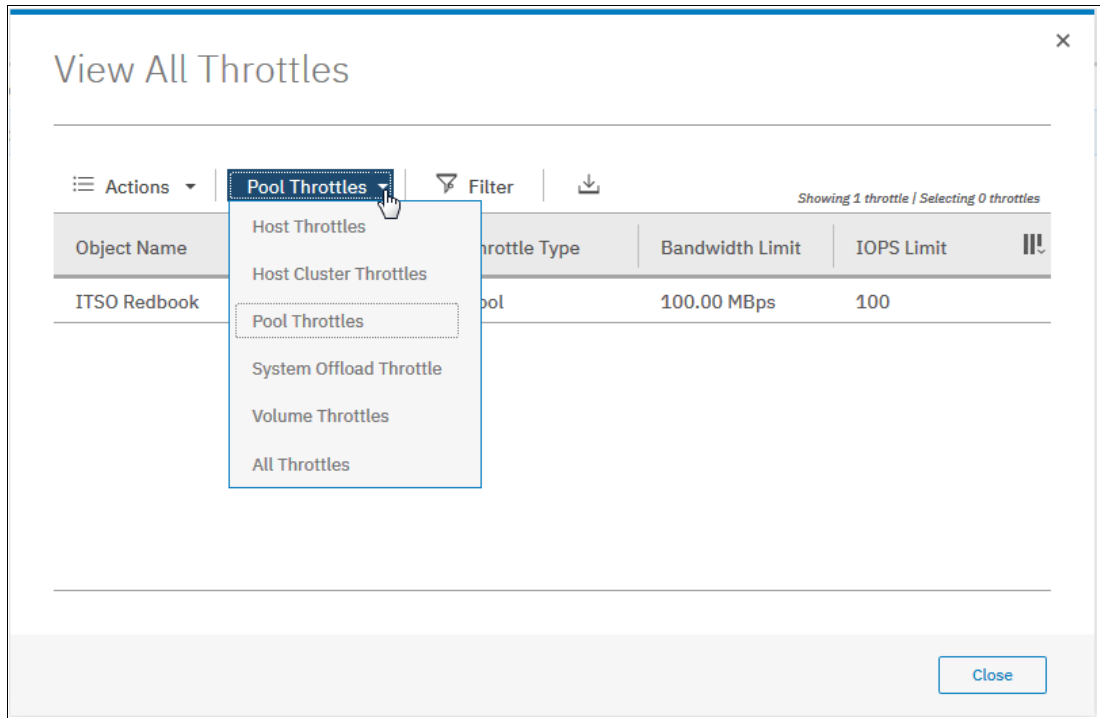


Figure 4-30 Selecting specific throttle information

Delete

Pools can be deleted only by using the GUI if no volumes are assigned to the pool. If the pool has any volumes within it, the option is not available. Selecting **Delete** immediately deletes the pool without other confirmation.

By using the CLI, you can delete a pool and all of its contents by using the **-force** parameter. However, all volumes and host mappings are deleted, and you cannot recover them.

Important: After you delete the pool through the CLI, all data that is stored in the pool is lost except for the image mode MDisks. The image mode MDisk volume definition is deleted, but the data on the imported MDisk remains untouched.

After deleting a pool, all of the managed or image mode MDisks in the pool return to the unmanaged status.

Properties

Selecting **Properties** displays information about the storage pool, as shown in Figure 4-31.

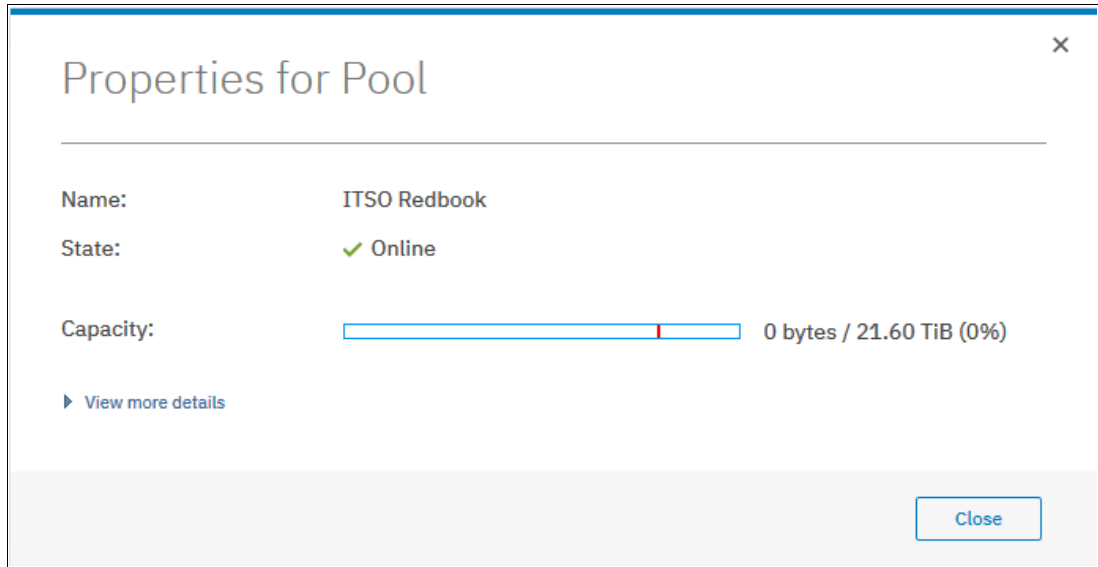


Figure 4-31 Storage pool properties

More information is available by clicking **View more details** and by hovering over the elements on the window, as shown in Figure 4-32. Click **Close** to return to the Pools window.

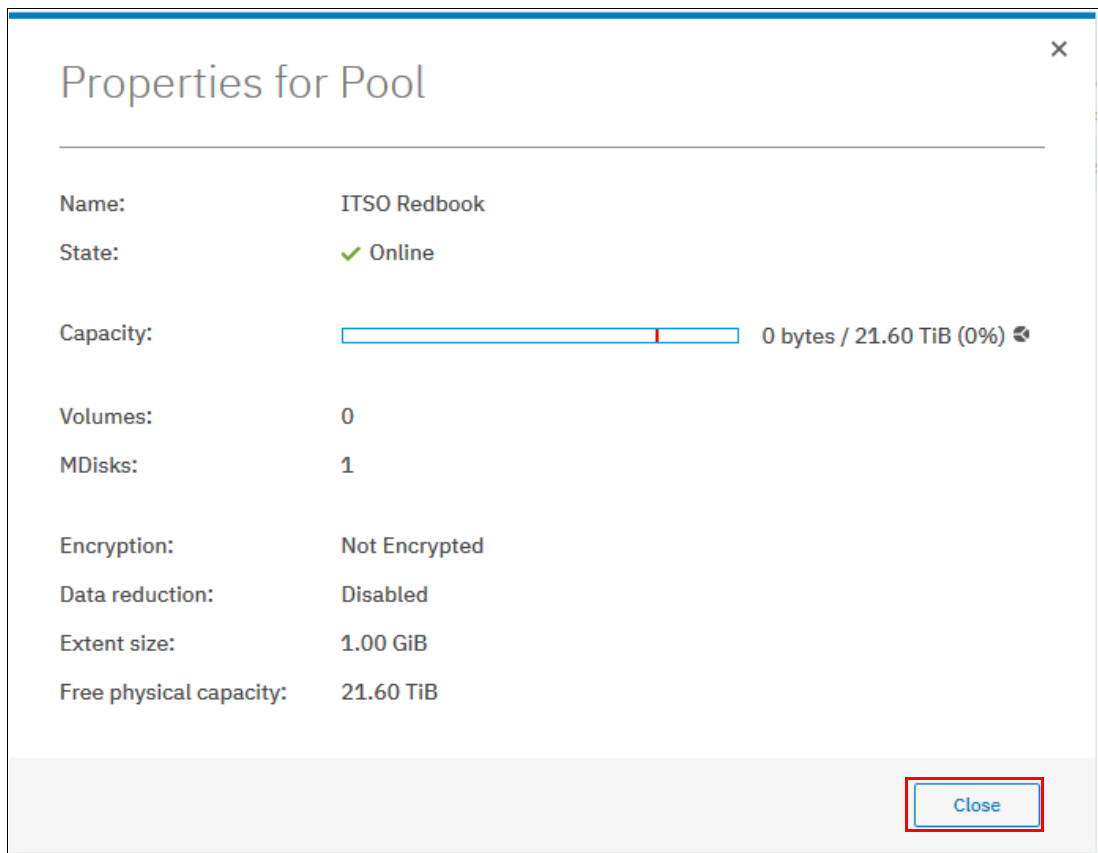


Figure 4-32 More details for storage pool properties

Customize columns

Selecting **Customize Columns** in the Actions menu allows you to include more information fields in the Pools window, as shown in Figure 4-33.

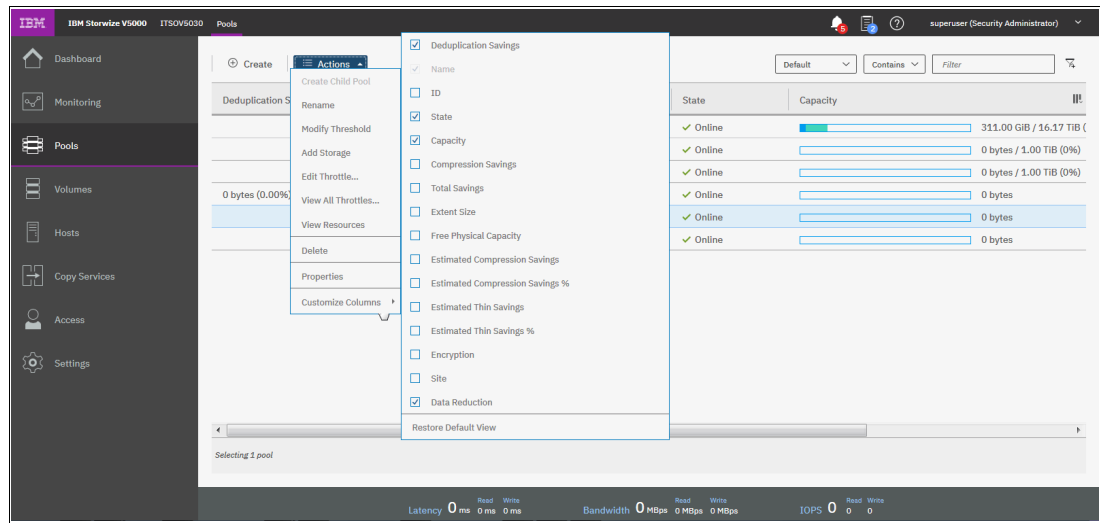


Figure 4-33 Customizing columns in the Pools window

4.2.3 Child storage pools

A *child storage pool* is a storage pool that is created within a storage pool. The storage pool in which the child storage pool is created is called *parent storage pool*.

Unlike a parent pool, a child pool does not contain MDisks; its capacity is provided exclusively by the parent pool in the form of extents. The capacity of a child pool is set at creation time, but can be nondisruptively modified later. The capacity must be a multiple of the parent pool extent size and must be smaller than the free capacity of the parent pool.

Child pools are useful when the capacity allocated to a specific set of volumes must be controlled.

Child pools inherit most properties from their parent pools and cannot be changed. The inherited properties include the following examples:

- ▶ Extent size
- ▶ Easy Tier setting
- ▶ Encryption setting (but only if the parent pool is encrypted)

Creating a child pool

To create a child pool, you can browse to **Pools** → **Pools** → **Actions** or **Pools** → **MDisks by Pools** → **Actions** and select **Create Child Pool**. Alternatively, you can right-click the parent pool, as shown in Figure 4-34.

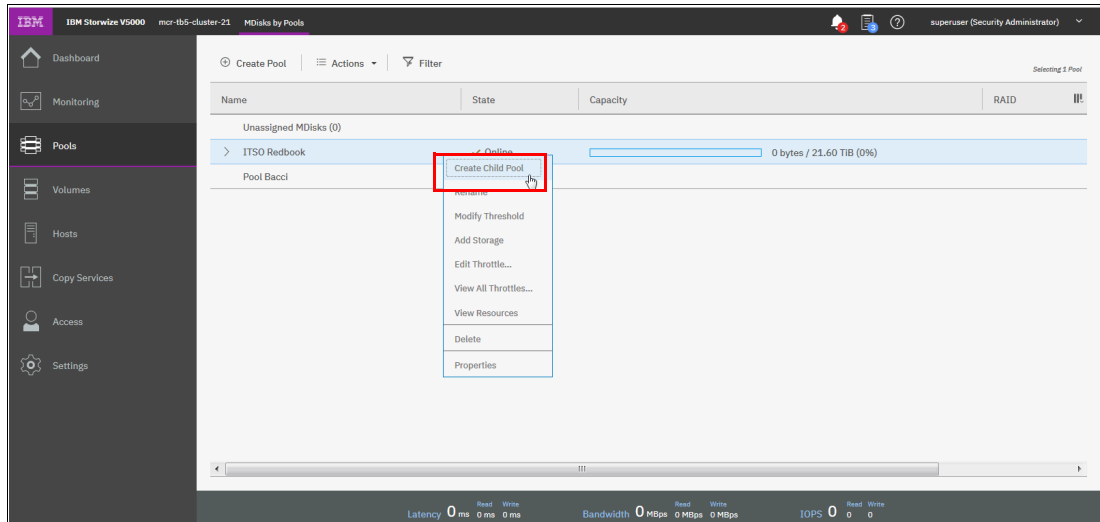


Figure 4-34 Selecting child menu creation

Enter the name and the capacity of the child pool and click **Create**, as shown in Figure 4-35.

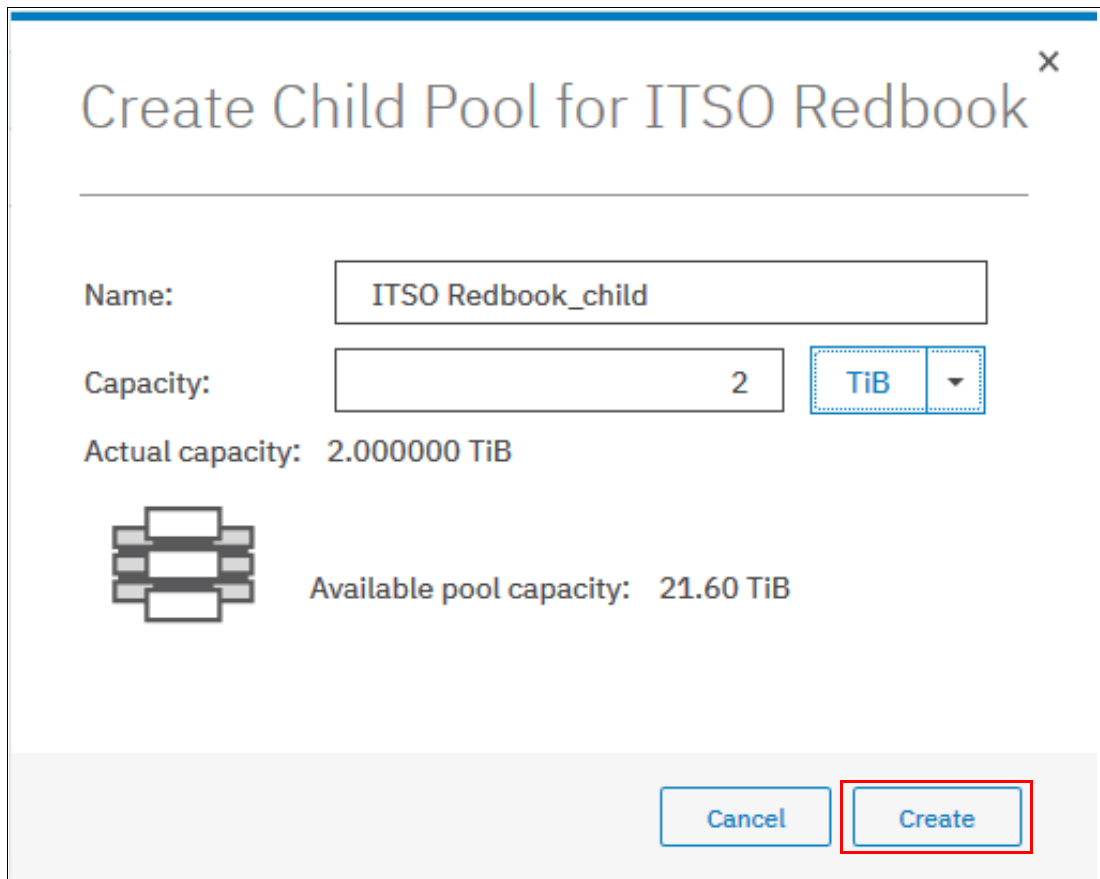


Figure 4-35 Create Child Pool window

Note: You cannot create an encrypted child pool from an unencrypted parent pool if the parent pool contains any unencrypted arrays or an MDisk that is not self-encrypting and there are nodes in the system that do not support software encryption (for example, nodes that do not have encryption license enabled).

An encrypted child pool that is created from an unencrypted parent pool reports as unencrypted if the parent pool contains any unencrypted arrays. Remove these arrays to ensure that the child pool is fully encrypted.

After the child pool is created, it is listed in the Pools window under its parent pool, as shown in Figure 4-36. Toggle the arrow sign in the left of the storage pool name to show or hide the child pools.

Creating a child pool within a child pool is not possible.

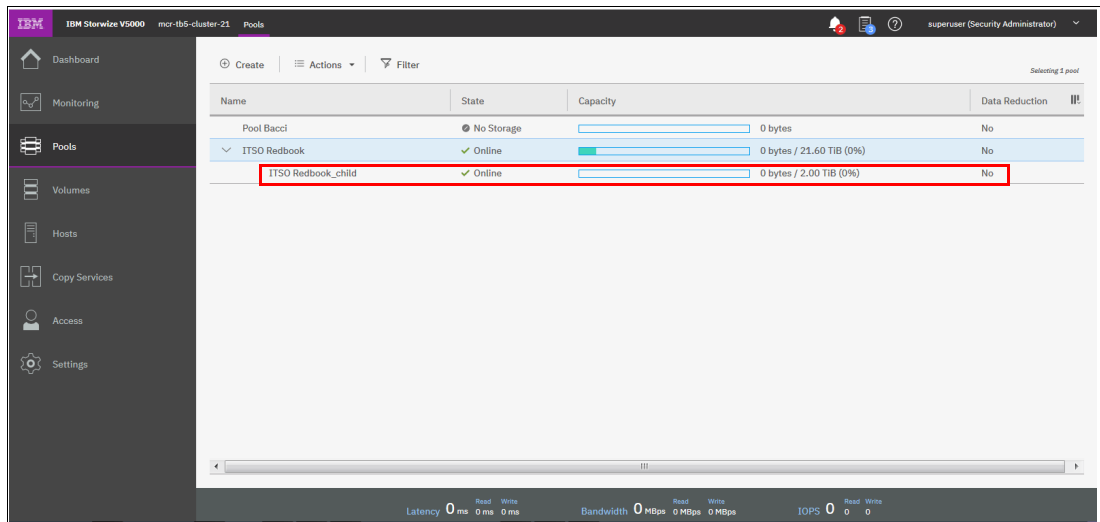


Figure 4-36 Child pool list

Actions on child storage pools

All actions that are supported for parent storage pools are supported for child storage pools, except for Add Storage. Child pools also support the Resize action.

To select an action, right-click the child storage pool, as shown in Figure 4-37. Alternatively, select the storage pool and click **Actions**.

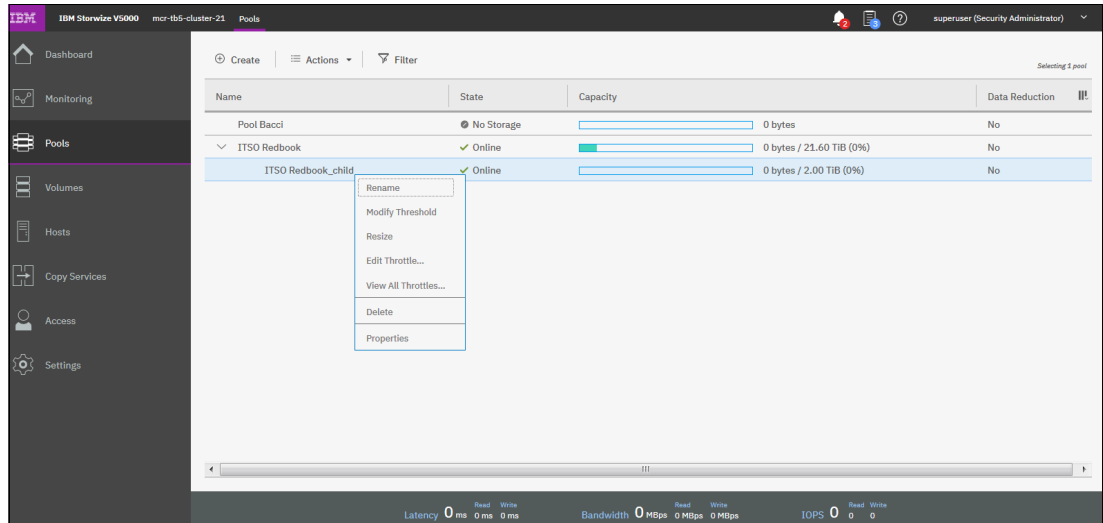


Figure 4-37 Child pools list of actions

Resize

Selecting **Resize** allows you to increase or decrease the capacity of the child storage pool, as shown in Figure 4-38 on page 188. Enter the new pool capacity and click **Resize**.

Note: You cannot shrink a child pool below its real capacity. Therefore, the new size of a child pool must be larger than the capacity that is used by its volumes.

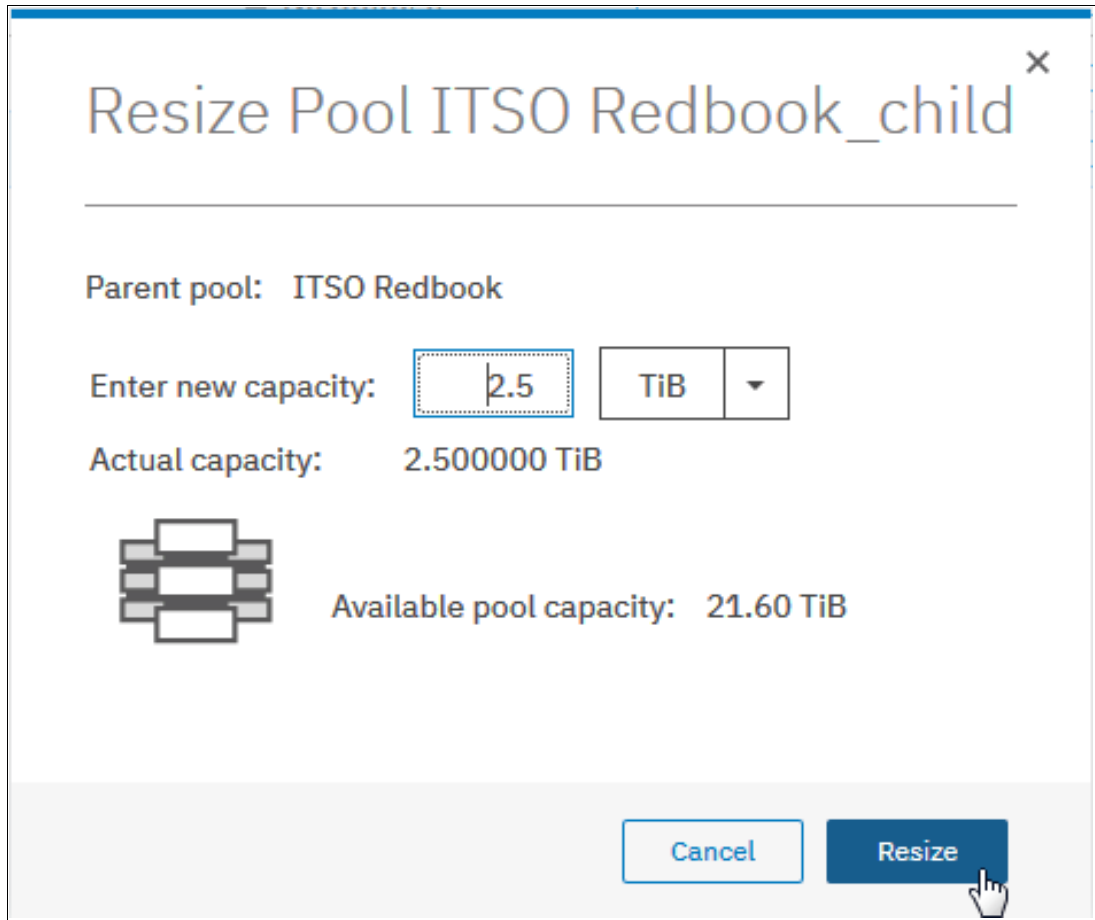


Figure 4-38 Resizing child pools

Delete

Deleting a child pool is a task that is similar to deleting a parent pool. As with a parent pool, the Delete action is disabled if the child pool contains volumes. After deleting a child pool, the extents that were being occupied return to the parent pool as free capacity.

Note: A volume in a child pool can be migrated only to another child pool within the same parent pool or to its own parent pool. In any other case, use volume mirroring instead. During migration from a child pool to its parent pool, or vice versa, there is no real data move. There is only a reassignment of extents between the pools.

4.3 Working with managed disks

A storage pool is created as an empty container, with no storage assigned to it. Storage is then added in the form of MDisks. An MDisk can be an array from internal storage or an LU from an external storage system. The same storage pool can include internal and external MDisks.

Arrays are created from internal storage by using RAID technology to provide redundancy and increased performance. The system supports two types of RAID: traditional RAID and distributed RAID.

Arrays are assigned to storage pools at creation time and cannot be moved between storage pools. It is not possible to have an array that does not belong to any storage pool.

External MDisks can have one of the following modes:

► Unmanaged

External MDisks are discovered by the system as *unmanaged MDisks*. An unmanaged MDisk is not a member of any storage pool, is not associated with any volumes, and has no metadata stored on it. The system does not write to an MDisk that is in unmanaged mode, except when it attempts to change the mode of the MDisk to one of the other modes.

► Managed

When unmanaged MDisks are added to storage pools, they become managed. Managed mode MDisks are always members of a storage pool and provide extents to the storage pool. This mode is the most common and normal mode for an MDisk.

► Image

Image mode provides a direct block-for-block translation from the MDisk to a volume. This mode is provided to satisfy the following major usage scenarios:

- Virtualization of external LUs that contain data that is not written through the IBM Storwize V5000 Gen2.
- Exporting MDisks from the IBM Storwize V5000 Gen2 after volume migrations to image mode MDisks.

MDisks are managed through the MDisks by Pools window. To access the MDisks by Pools window, click **Pools** → **MDisks by Pools**, as shown in Figure 4-39.

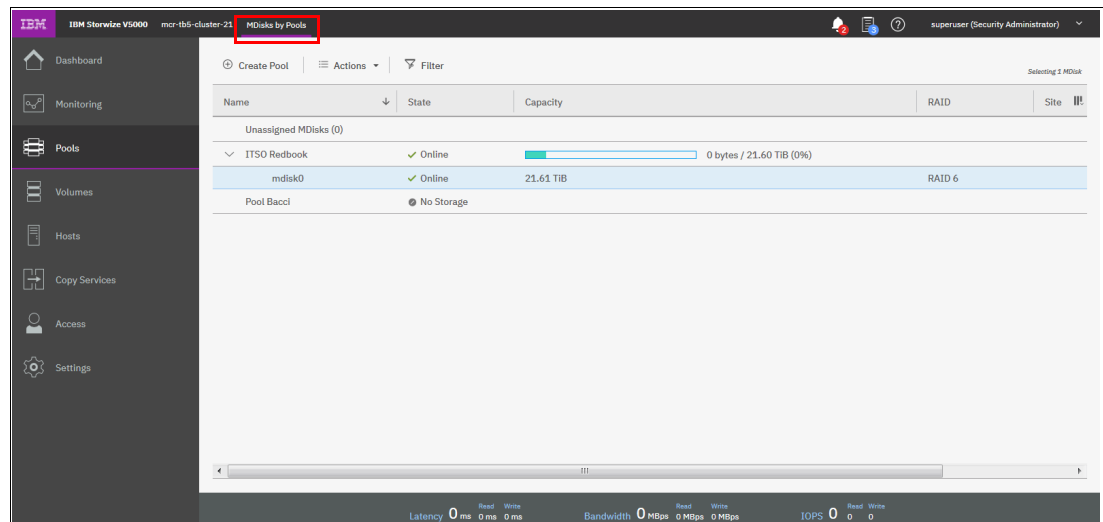


Figure 4-39 MDisks by Pools window

The window lists all the MDisks available in the system under the storage pool to which they belong.

4.3.1 Assigning managed disks to storage pools

MDisks can be assigned to a storage pool at any time to increase the number of extents that are available in the pool. The system automatically balances volume extents between the MDisks to provide the best performance to the volumes.

Arrays are created and assigned to a storage pool at the same time. The following options are available to add storage:

- Option 1: To assign MDisks to a storage pool, browse to **Pools** → **MDisks by Pools**.

Right-click the pool and select **Add Storage**, as shown in Figure 4-40. Alternatively, select a pool and click **Actions**.

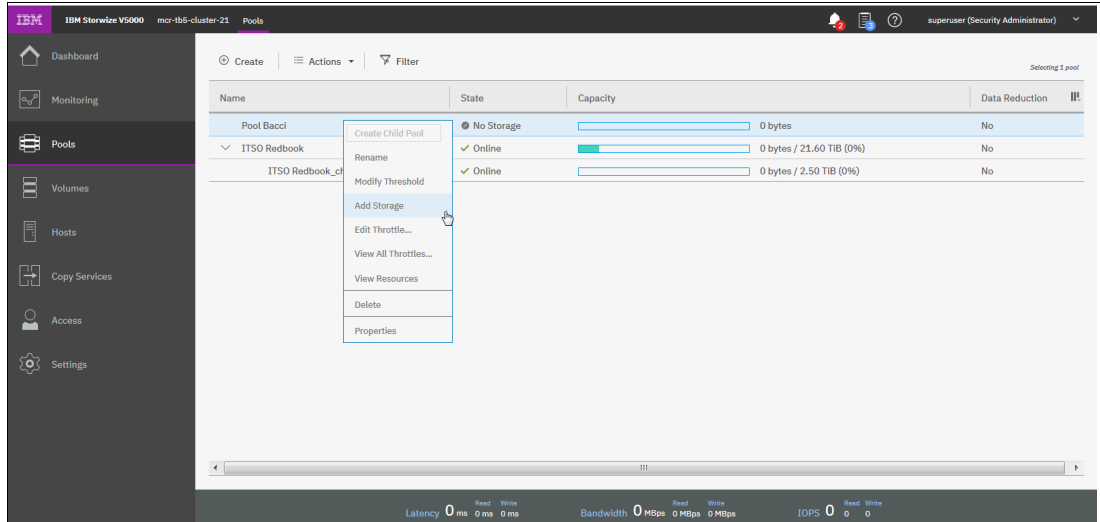


Figure 4-40 Add storage: option 1

- Option 2: Select **Assign** under a specific drive class or external storage controller, as shown in Figure 4-40.

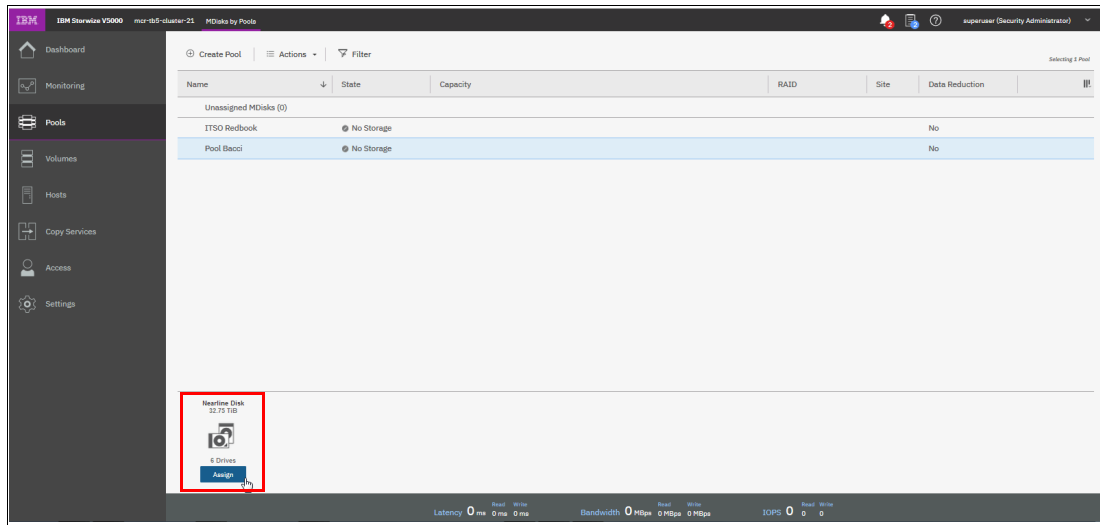


Figure 4-41 Add storage: option 2

The configuration wizard starts, as shown in Figure 4-42.

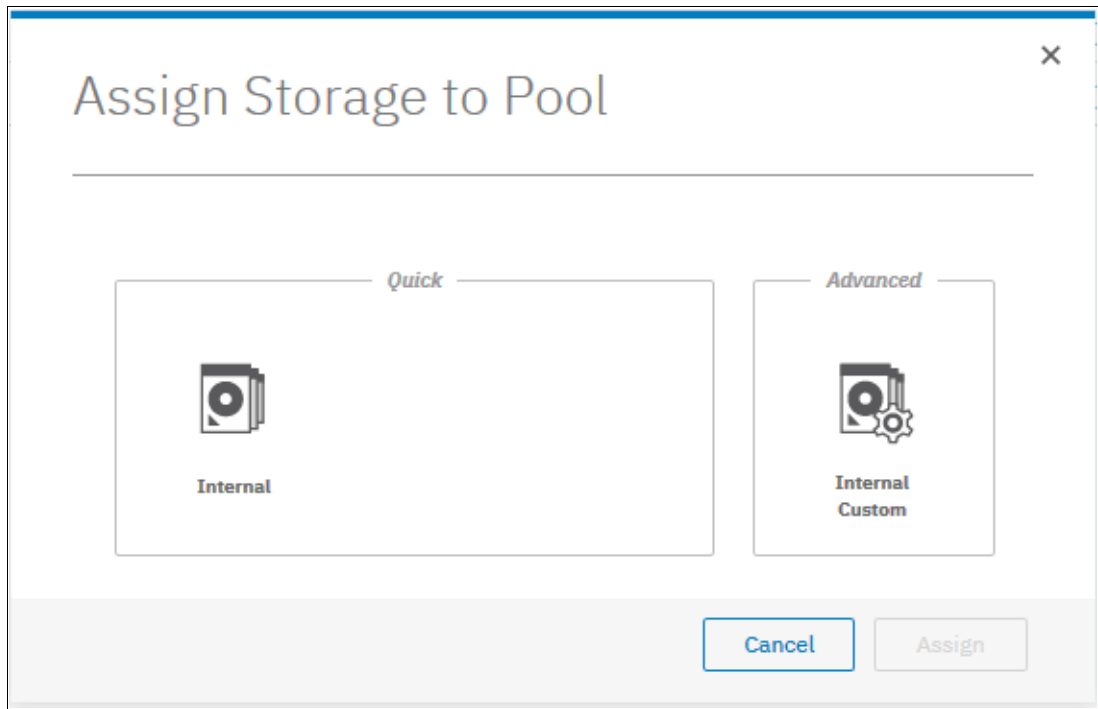


Figure 4-42 Assigning storage to storage pool

Quick internal configuration

Selecting **Internal** suggests a configuration for internal drives that is based on RAID configuration presets, and considers drive class and the number of drives available. It automatically defaults parameters, such as stripe width, number of spares (for traditional RAID), number of rebuild areas (for distributed RAID), and number of drives of each class. The number of drives is the only value that can be adjusted.

Figure 4-43 shows an example of a quick configuration.

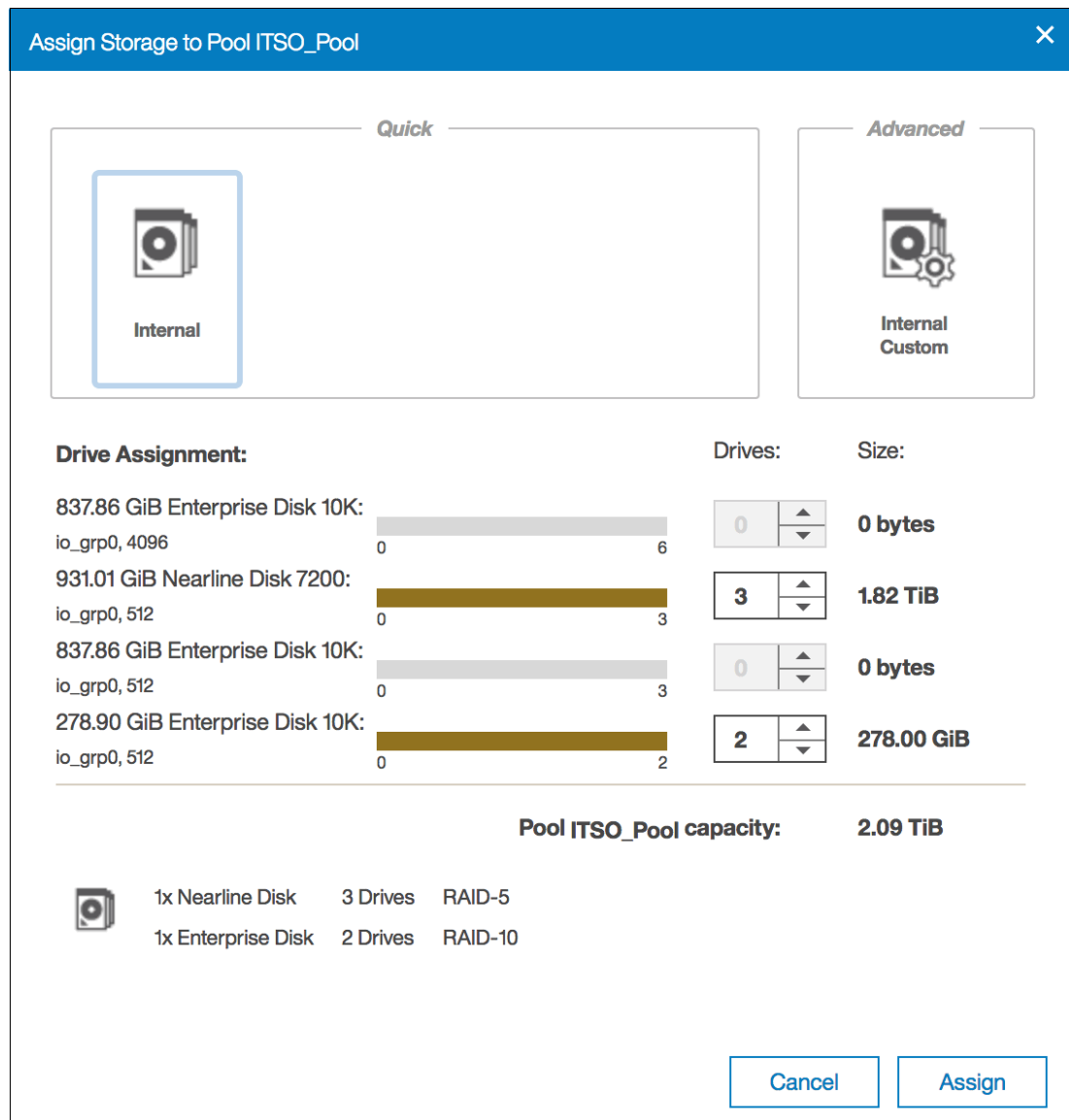


Figure 4-43 Quick configuration wizard

This configuration combines two drive classes that belong to two different tiers of storage (Nearline and Enterprise). This is the default option and takes advantage of the Easy Tier functionality. However, this option can be adjusted by setting the number of drives of different classes to zero, as shown in Figure 4-44.

Note: If any drive class is not compatible with the drives being assigned, that drive class cannot be selected.

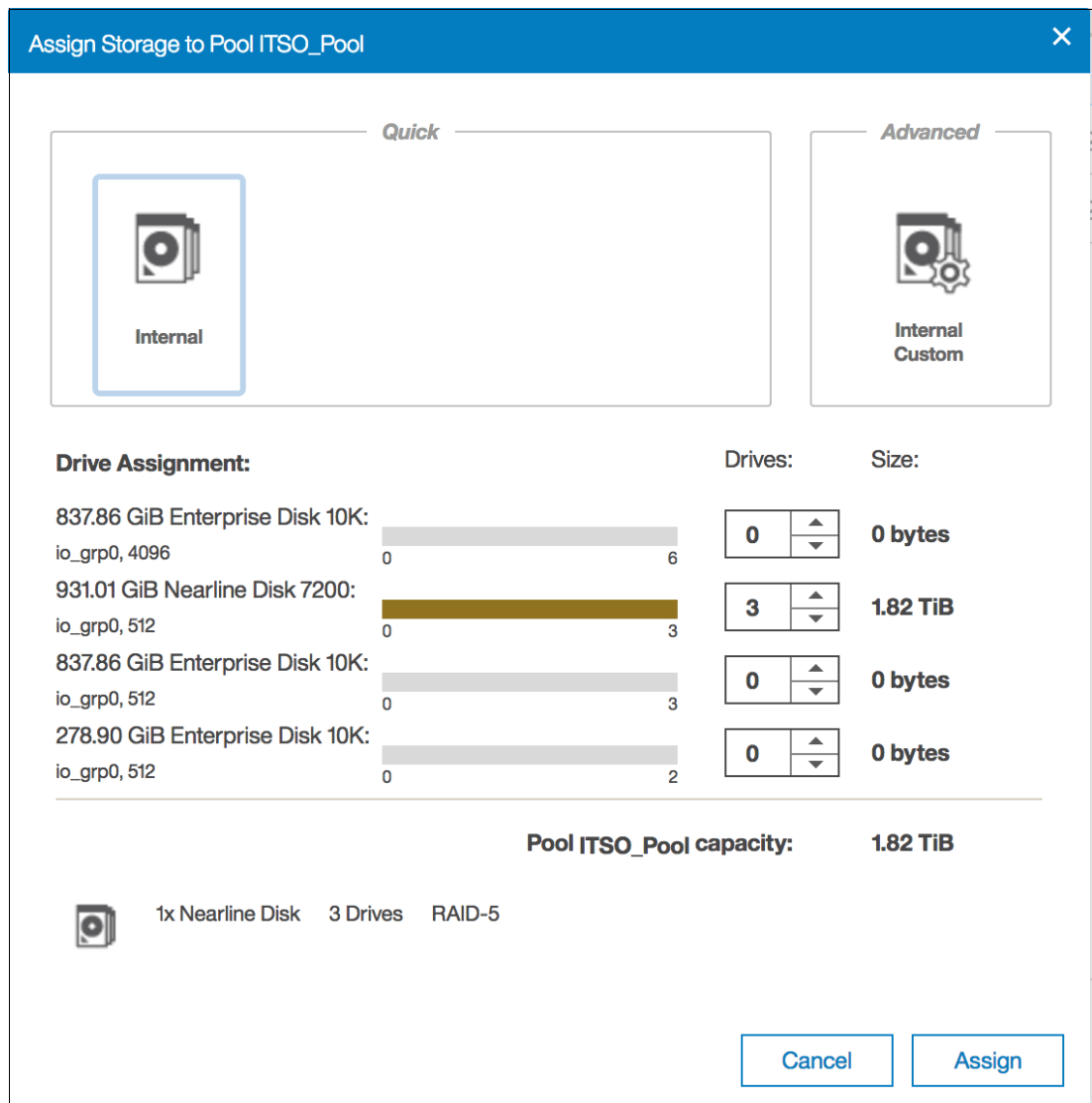


Figure 4-44 Quick configuration wizard with a zeroed storage class

If you are adding storage to a pool with storage already assigned, the existing storage is also considered, with some properties being inherited from existing arrays for a specific drive class. Drive classes that are incompatible with the classes that are in the pool also are disabled.

When you are satisfied with the presented configuration, click **Assign**, as shown in Figure 4-45. The array MDisks are then created and initialized in the background.

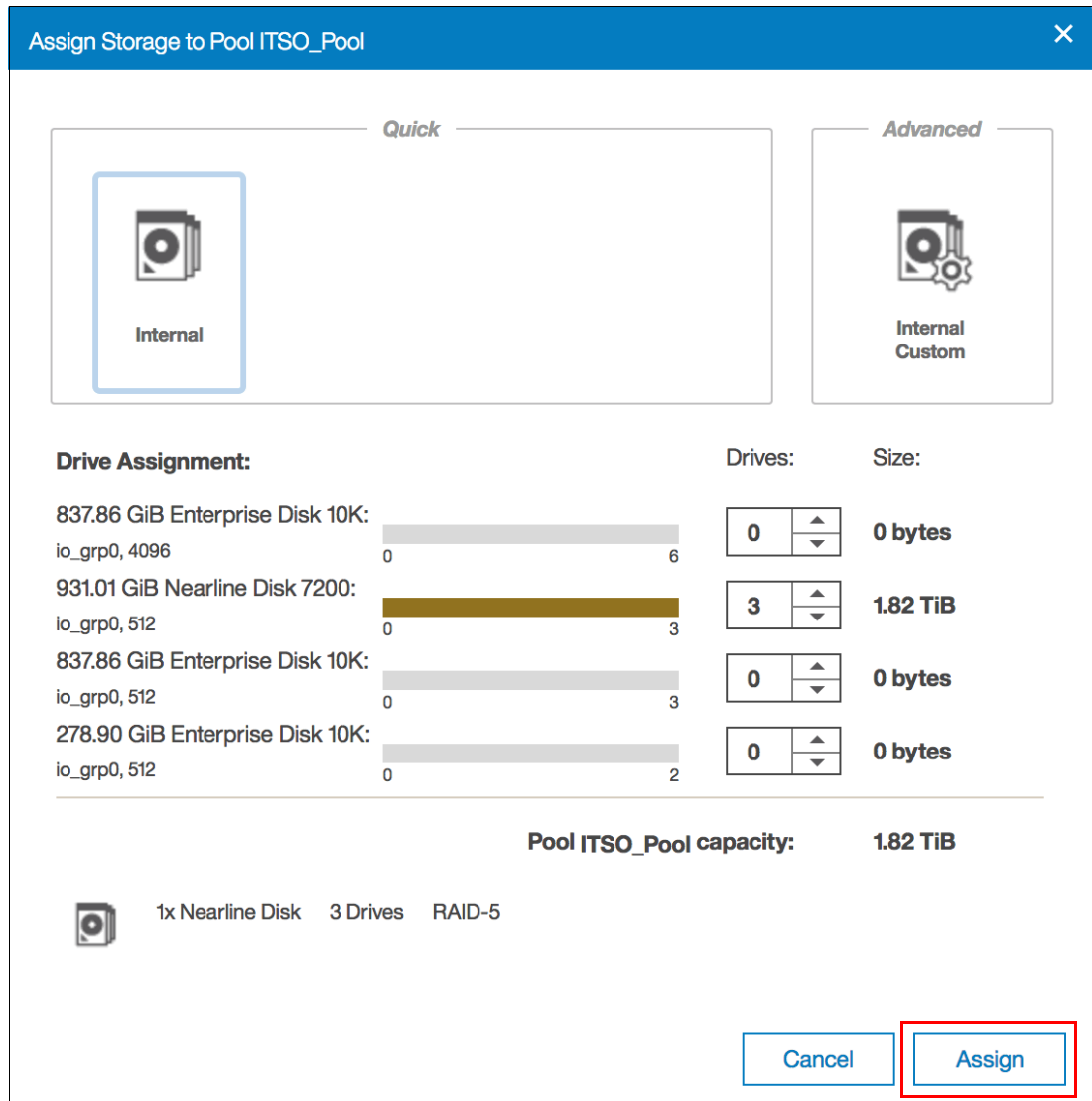


Figure 4-45 Clicking assign on quick configuration wizard

Advanced internal custom configuration

Selecting **Internal Custom** allows the user to customize the configuration for internal drives.

Tip: It is advised to use the advanced configuration only when the quick configuration suggested does not fit your business requirements.

The following values can be customized:

- ▶ RAID level
- ▶ Number of spares
- ▶ Array width
- ▶ Stripe width
- ▶ Number of drives of each class

Figure 4-46 shows an example with six drives that are ready to be configured as RAID 6. Click **Summary** to see the list of MDisk arrays to be created. To return to the default settings, select the refresh icon next to the pool capacity. Click **Assign** to create and assign the arrays.

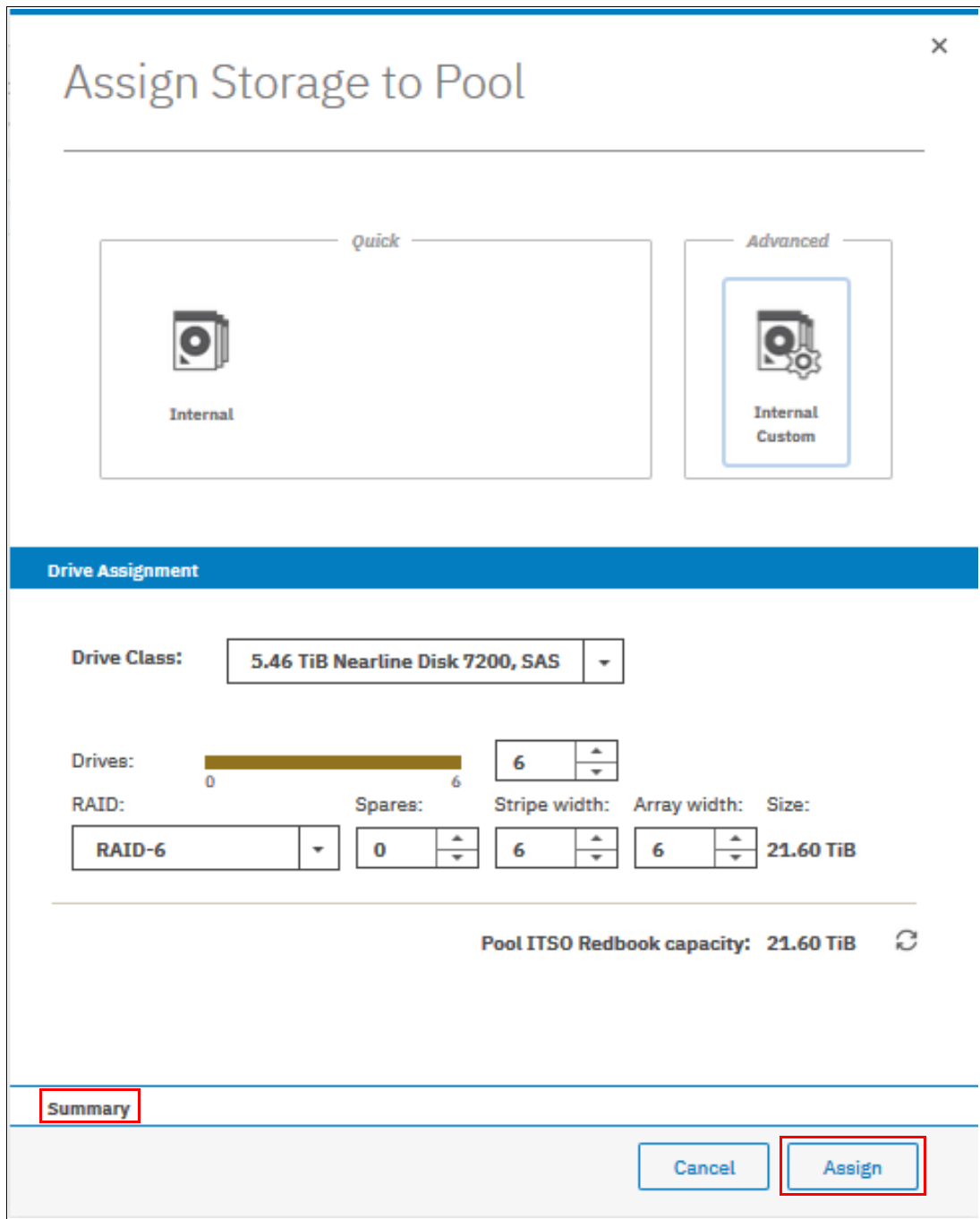


Figure 4-46 Advanced internal custom configuration

4.3.2 RAID configuration

In this section, we describe the Redundant Array of Independent Disks (RAID) configuration and technology.

Introduction to RAID technology

RAID provides two key design goals:

- ▶ Increased data reliability
- ▶ Increased input/output (I/O) performance

When multiple physical disks are set up to use the RAID technology, they are in a *RAID array*. The IBM Storwize V5000 Gen2 provides the following traditional RAID levels:

- ▶ RAID 0
- ▶ RAID 1
- ▶ RAID 5
- ▶ RAID 6
- ▶ RAID 10

RAID technology can provide better performance for data access, high availability for the data, or a combination. RAID levels define a trade-off between high availability, performance, and cost.

The RAID concept must be extended to *disk rebuild time* because of increasing physical disk capacity.

In a disk failure, traditional RAID writes the data to a single spare drive. With increasing capacity, the rebuild time is also increased and the probability of a second failure during the rebuild process becomes more likely. In addition, the spares are idle when they are not being used, which wastes resources.

Distributed RAID (DRAID) addresses these points and is available for the IBM Storwize V5000 Gen2 in two types:

- ▶ Distributed RAID 5 (DRAID 5)
- ▶ Distributed RAID 6 (DRAID 6)

Distributed RAID reduces the recovery time and the probability of a second failure during rebuild. As with traditional RAID, a distributed RAID 5 array can lose one physical drive and survive. If another drive fails in the same array before the bad drive is recovered, the MDisk and the storage pool go offline as they are supposed to. So, distributed RAID does not change the general RAID behavior.

Note: Although Traditional RAID is still supported and is the default choice in the GUI, the suggestion is to use DRAID 6 whenever possible. It depends on the number of disk drives, type, and size.

4.3.3 Distributed RAID

In distributed RAID, all drives are active, which improves performance. Spare capacity is used instead of the idle spare drives from traditional RAID. Because no drives are spare, all drives contribute to performance. The spare capacity is rotated across the disk drives so the write rebuild load is distributed across multiple drives and the bottleneck of one drive is removed.

Figure 4-47 on page 197 shows an example of a distributed RAID with 10 disks. The physical disk drives are divided into multiple packs. The reserved spare capacity (which is marked in yellow) is equivalent to two spare drives, but the capacity is distributed across all of the physical disk drives.

The data is distributed across a single row. For simplification, not all packs are shown in Figure 4-47.

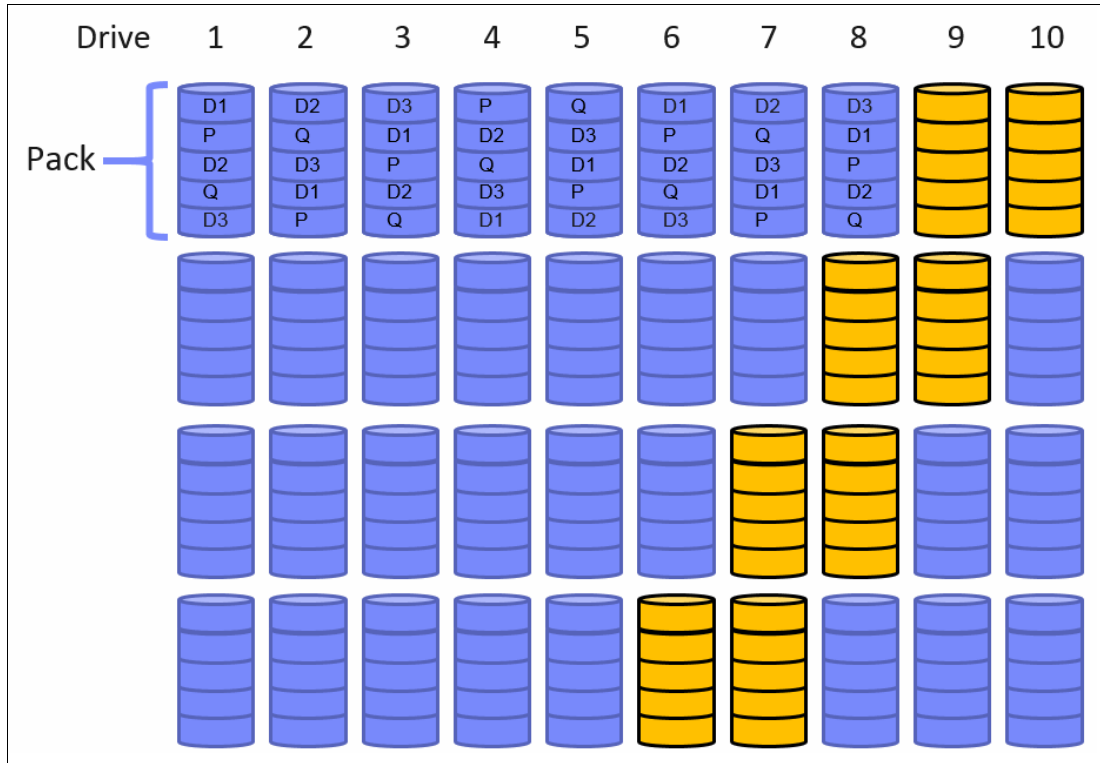


Figure 4-47 Distributed RAID 6

Figure 4-48 on page 198 shows a single drive failure in the distributed RAID 6 (DRAID 6) environment. Physical disk 3 failed and the RAID 6 algorithm is using the spare capacity for a single spare drive in each pack for rebuild (which is marked in green). All disk drives are involved in the rebuild process, which significantly reduces the rebuild time.

For simplification, not all packs are shown in Figure 4-48.

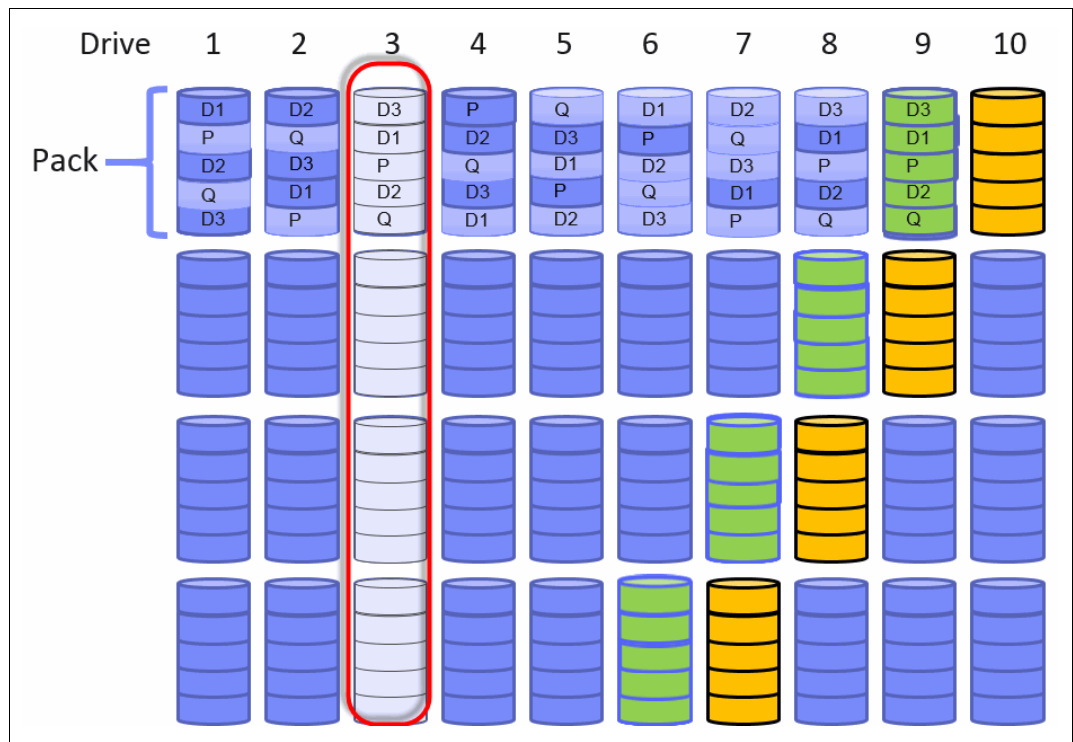


Figure 4-48 Single drive failure with DRAID 6

The use of multiple drives improves the rebuild process, which is up to 10 times faster than traditional RAID. This speed is even more important when you use large drives.

The conversion from traditional RAID to distributed RAID is possible by using volume mirroring or volume migration. Mixing traditional RAID and distributed RAID in the same storage pool is also possible.

Example

The same number of disks can be configured by using traditional or distributed RAID. In our example, we use six disk drives and assign those disks as RAID 6 to a single pool.

Figure 4-49 shows the setup for a traditional RAID 6 environment. The pool consists of one MDisk, with five disk drives. The spare drive is not listed in this summary.

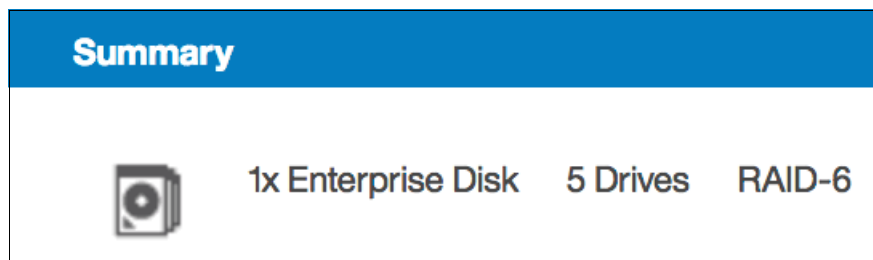


Figure 4-49 Array configuration for a traditional RAID 6 with six disks

Figure 4-50 shows the setup for a distributed RAID 6 environment. The pool consists of a single MDisk with six disk drives. The spare drives are included in this summary.

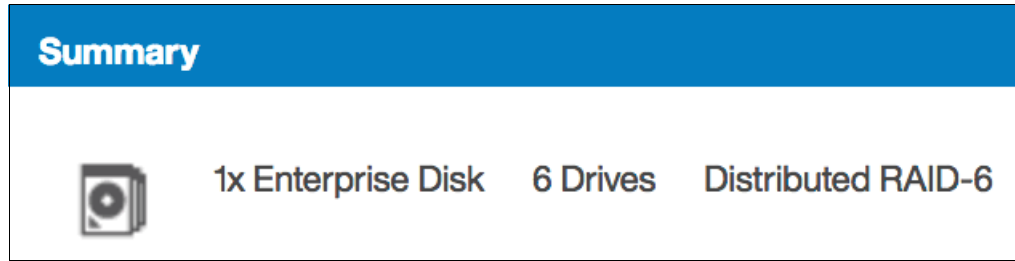


Figure 4-50 Array configuration for a distributed RAID 6 with six disks

4.3.4 RAID configuration presets

RAID configuration presets are used to configure internal drives. They are based on the advised values for the RAID level and drive class. Each preset has a specific goal for the number of drives per array and the number of spare drives to maintain redundancy.

For the best performance with Flash Drives, which are known as solid-state drives (SSDs), arrays with the same number of drives are recommended, which is the same design for traditional RAID arrays.

Table 4-1 lists the presets that are used for Flash drives for the IBM Storwize V5000 Gen2 storage system.

Table 4-1 Flash RAID presets

Preset	Purpose	RAID level	Drives per array goal	Drive count (min - max)	Spare drive goal
Flash RAID 5	Protects against a single drive failure. Data and one stripe of parity are striped across all array members.	5	8	3 - 16	1
Flash Distributed RAID 5	Protects against a single drive failure. Data and one stripe of parity are striped across all array members.	5	8	3 - 16	1
Flash RAID 6	Protects against two drive failures. Data and two stripes of parity are striped across all array members.	6	12	5 - 16	1
Flash Distributed RAID 6	Protects against two drive failures. Data and two stripes of parity are striped across all array members.	6	12	5 - 16	1
Flash RAID 10	Protects against at least one drive failure. All data is mirrored on two array members.	10	8	4 - 16 (even number of drives)	1

Preset	Purpose	RAID level	Drives per array goal	Drive count (min - max)	Spare drive goal
Flash RAID 1	Protects against at least one drive failure. All data is mirrored on two array members.	1	2	2	1
Flash RAID 0	Provides no protection against drive failures.	0	8	1 - 8	0
Flash Easy Tier	Mirrors data to protect against drive failure. The mirrored pairs are spread between storage pools to use for the Easy Tier function.	10	2	4 - 16 (even number of drives)	1

Flash RAID instances: In all Flash RAID instances, drives in the array are balanced across enclosure chains, if possible.

Table 4-2 describes the RAID presets that are used for Enterprise SAS and Nearline SAS drives for the IBM Storwize V5000 Gen2 storage system.

Table 4-2 Hard disk drive (HDD) RAID presets

Preset	Purpose	RAID level	Drives per array goal	Drive count (min - max)	Spare goal depending on drives used	Chain balance
Basic RAID 5	Protects against a single drive failure. Data and one stripe of parity are striped across all array members.	5	8	3 - 16	1	All drives in the array are from the same chain wherever possible.
Distributed RAID 5	Protects against a single drive failure. Data and one stripe of parity are striped across all array members.	5	48 - 60	4 - 128	1: 0 - 36 2: 37 - 72 3: 73 - 100 4: 101 - 128	All drives in the array are from the same chain wherever possible.
Basic RAID 6	Protects against two drive failures. Data and two stripes of parity are striped across all array members.	6	12	5 - 16	1	All drives in the array are from the same chain wherever possible.
Distributed RAID 6	Protects against two drive failures. Data and two stripes of parity are striped across all array members.	6	48 - 60	6 - 128	1: 0 - 36 2: 37 - 72 3: 73 - 100 4: 101 - 128	All drives in the array are from the same chain wherever possible.

Preset	Purpose	RAID level	Drives per array goal	Drive count (min - max)	Spare goal depending on drives used	Chain balance
Basic RAID 10	Protects against at least one drive failure. All data is mirrored on two array members.	10	8	4 - 16 (must be an even number of drives)	1	All drives in the array are from the same chain wherever possible.
Balanced RAID 10	Protects against at least one drive or enclosure failure. All data is mirrored on two array members. The mirrors are balanced across the two enclosure chains.	10	8	4 - 16 (even)	1	Exactly half of the drives are from each chain.
RAID 0	Provides no protection against drive failures.	0	8	1 - 8	0	All drives in the array are from the same chain wherever possible.

Actions on arrays

MDisks that are created from internal storage are RAID arrays and support specific actions that are not supported on external MDisks. Some actions that are supported on traditional RAID arrays are not supported on distributed RAID arrays and vice versa.

To choose an action, select the array and click **Actions**, as shown in Figure 4-51. Alternatively, right-click the array.

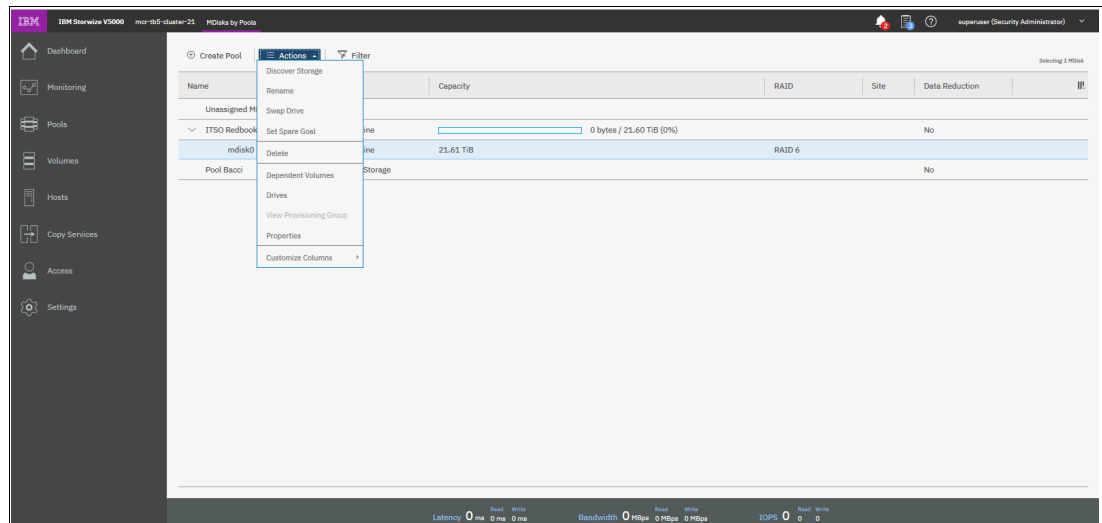


Figure 4-51 Available actions on arrays

Swap drive

Selecting **Swap Drive** allows the user to replace a drive in an array with another drive. The other drive must have the use of Candidate or Spare. This action can be used to replace a drive that is expected to fail soon.

Figure 4-52 shows the dialog box that opens. Select the member drive to be replaced and the replacement drive.

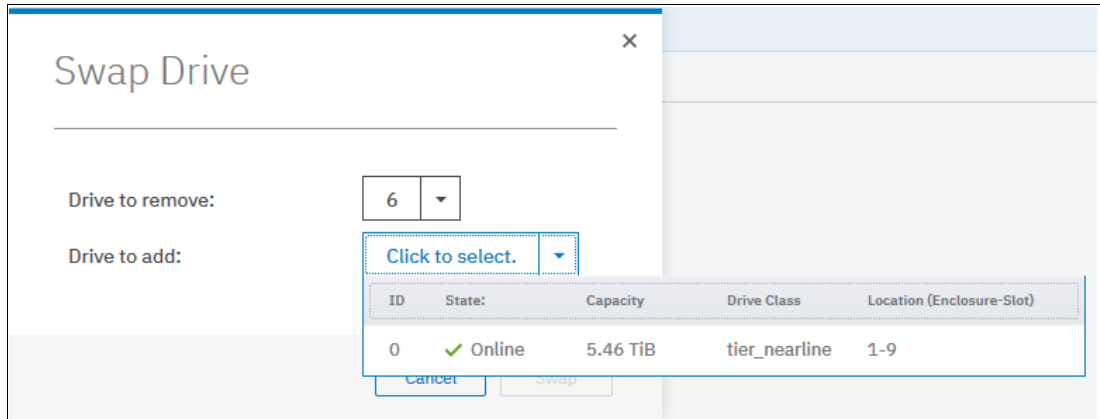


Figure 4-52 Swap Drive window

After defining the disk to be removed and the disk to be added, click **Swap**, as shown in Figure 4-53.

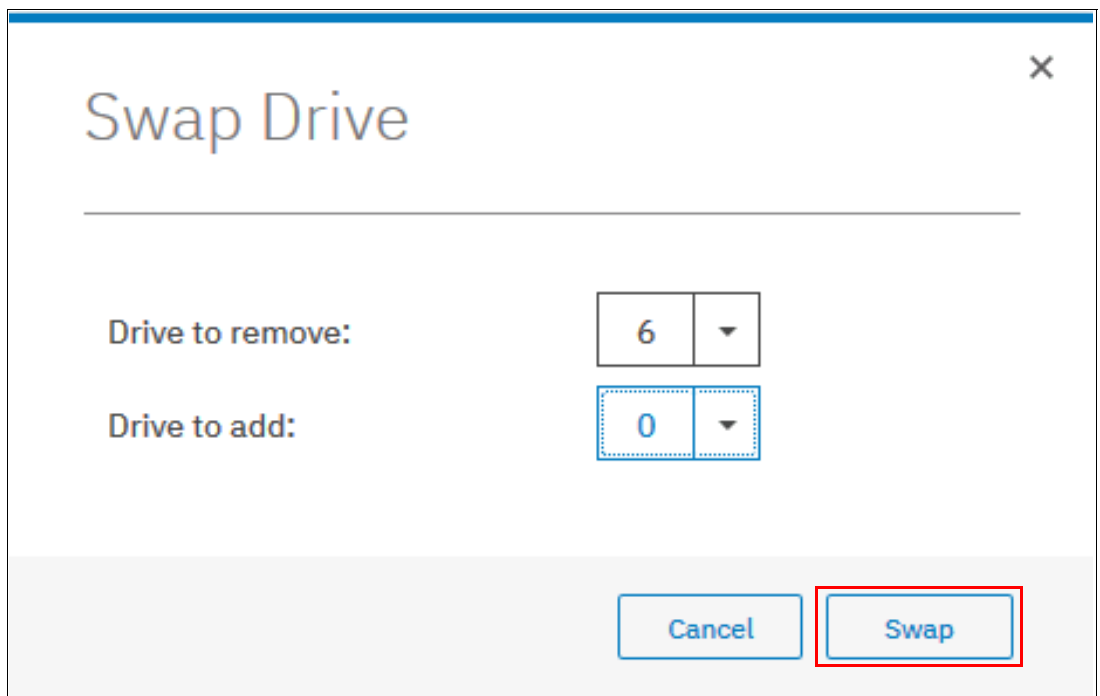


Figure 4-53 Swap button on Swap Drive window

Set Spare Goal

This action is available for traditional RAID arrays only. Selecting **Set Spare Goal** allows you to set the number of spare drives that is required to protect the array from drive failures.

If the number of spare drives available does not meet the configured goal, an error is logged in the Event log. This error can be fixed by adding drives of a compatible drive class as spares.

Figure 4-54 shows the dialog box that opens when you select **Set Spare Goal**. Define the number of required spares and click **Save**.

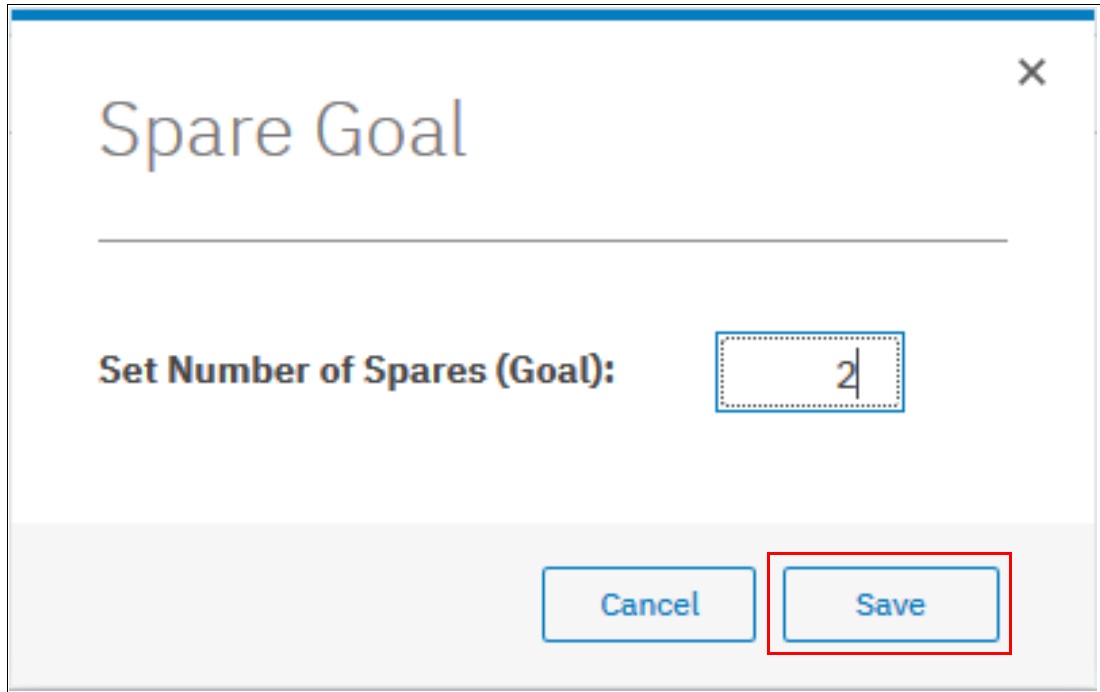


Figure 4-54 Spare Goal window

Set rebuild areas goal

This action is available for distributed RAID arrays only. Selecting **Set Rebuild Areas Goal** enables you to set the number of rebuild areas that is required to protect the array from drive failures.

If the number of rebuild areas available does not meet the configured goal, an error is logged in the Event log. This error can be fixed by replacing the failed drives in the array with new drives of a compatible drive class.

Figure 4-55 shows the dialog box that opens when you select **Set Rebuild Areas Goal**. Define the representative number of required spares that composes the rebuild area and click **Save**.

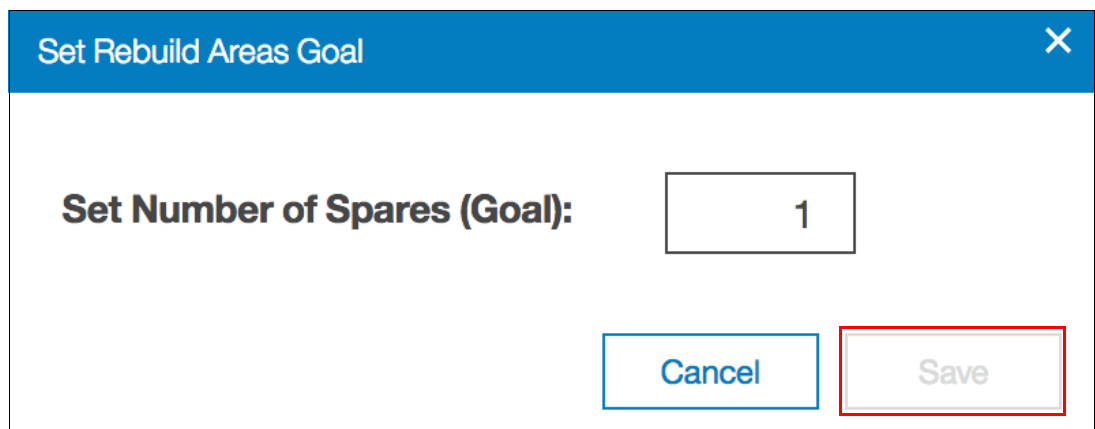


Figure 4-55 Rebuild Areas Goal window

Delete

Selecting **Delete** removes the array from the storage pool and deletes it.

Remember: An array does not exist outside of a storage pool. Therefore, an array cannot be removed from the pool without being deleted.

If there are no volumes that use extents from the array, the deletion command runs immediately without more confirmation. If there are volumes that use extents from the array, you are prompted to confirm the action, as shown in Figure 4-56. Click **Yes** to migrate the volumes or **No** to cancel the deletion process.

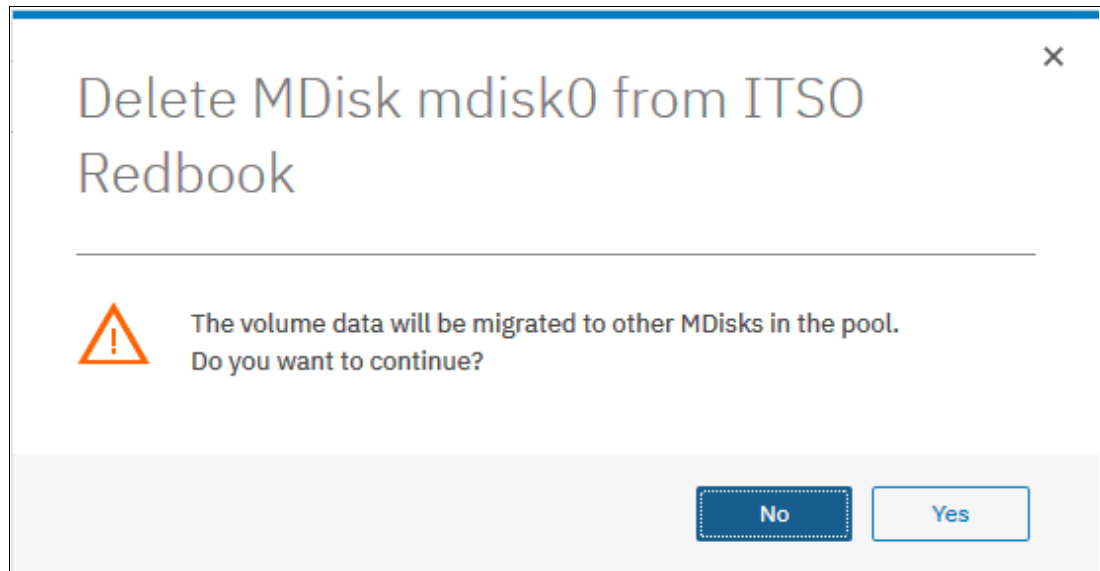


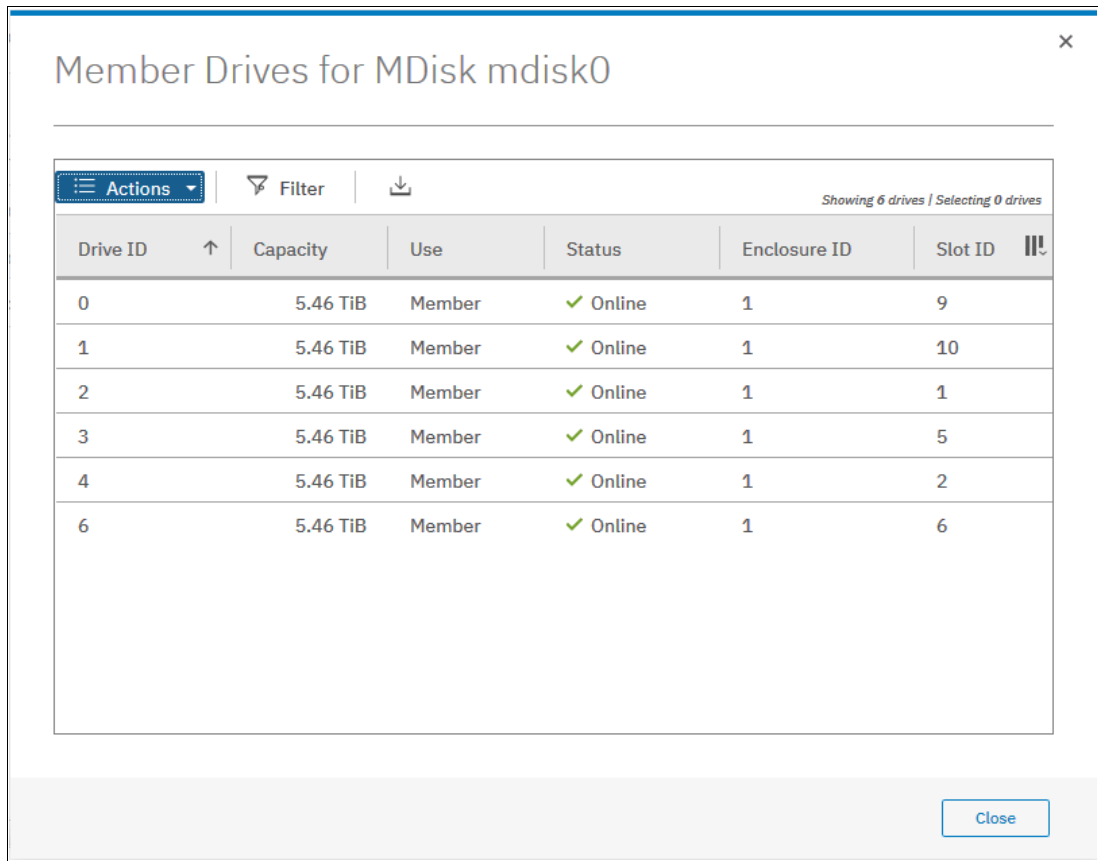
Figure 4-56 MDisk deletion confirmation window

Confirming the action starts the migration of the volumes to extents from other MDisks that remain in the pool; after the action completes, the array is removed from the storage pool and deleted.

Note: Ensure that you have enough available capacity remaining in the storage pool to allocate the data being migrated from the removed array; otherwise, the command fails.

Drives

Selecting **Drives** shows information about the drives that are included in the array, as shown in Figure 4-57.



The screenshot shows a window titled "Member Drives for MDisk mdisk0" with a close button in the top right corner. Below the title bar is a toolbar with "Actions", "Filter", and a download icon. A status bar indicates "Showing 6 drives | Selecting 0 drives". The main content is a table with the following columns: Drive ID, Capacity, Use, Status, Enclosure ID, and Slot ID. The table contains six rows of data, all showing 5.46 TiB capacity, Member use, and Online status.

Drive ID	Capacity	Use	Status	Enclosure ID	Slot ID
0	5.46 TiB	Member	✓ Online	1	9
1	5.46 TiB	Member	✓ Online	1	10
2	5.46 TiB	Member	✓ Online	1	1
3	5.46 TiB	Member	✓ Online	1	5
4	5.46 TiB	Member	✓ Online	1	2
6	5.46 TiB	Member	✓ Online	1	6

A "Close" button is located at the bottom right of the window.

Figure 4-57 Window showing the drives that are members of an MDisk

4.3.5 Actions on external MDisks

External MDisks support specific actions that are not supported on arrays. Some actions are supported on unmanaged external MDisks only and some are supported on managed external MDisks only.

To choose an action, right-click the external MDisk, as shown in Figure 4-58. Alternatively, select the external MDisk and click **Actions**.

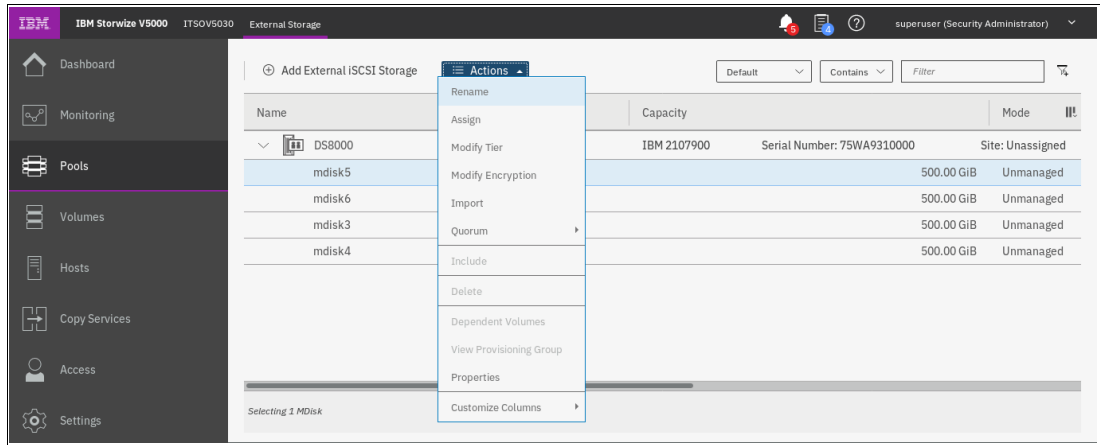


Figure 4-58 Available actions for external MDisk

Assign

This action is available for unmanaged external MDisk only. Selecting **Assign** opens the dialog box that is shown in Figure 4-59. This action acts on the selected MDisk or MDisk only.

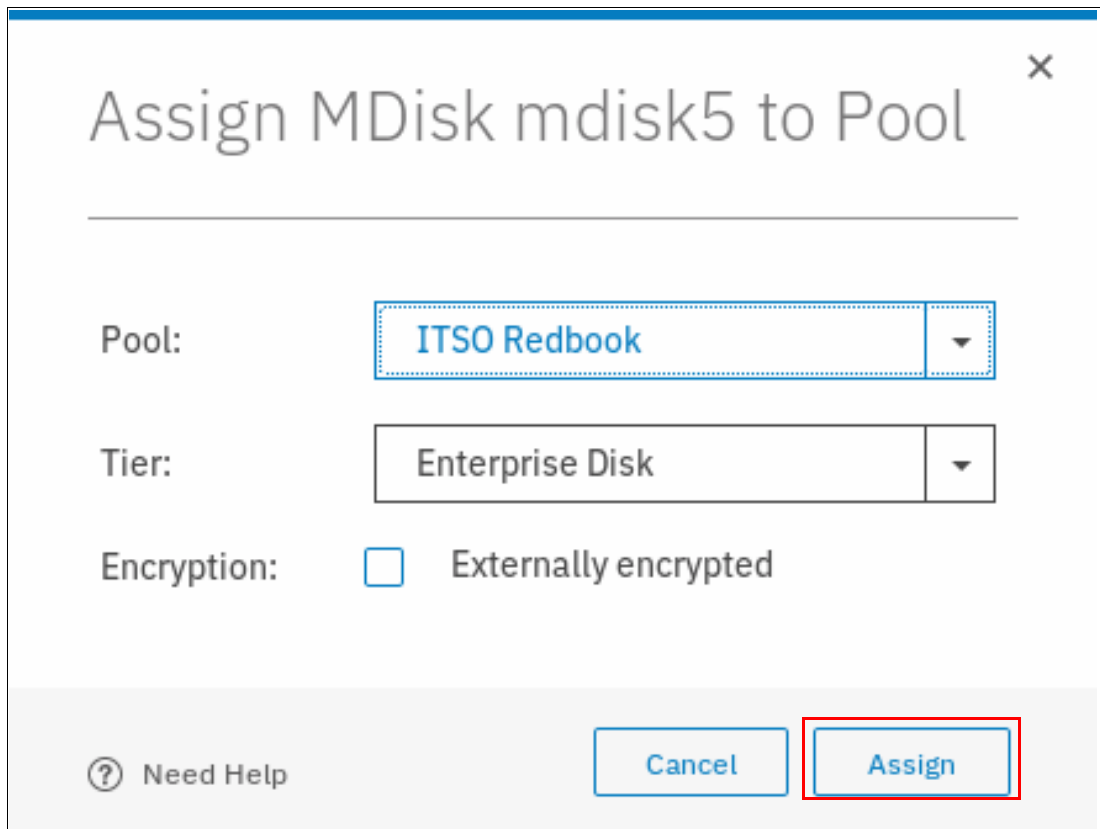


Figure 4-59 Assigning external MDisk to a pool

Important: If you must preserve data on an unmanaged MDisk, do *not* assign it to a storage pool because this action deletes the data on the MDisk. Use the **Import** option instead.

Modify Tier

Selecting **Modify Tier** allows the user to modify the tier to which the external MDisk is assigned, as shown in Figure 4-60. This setting is adjustable because the system cannot detect the tiers that are associated with external storage automatically. The Enterprise Disk (Tier 2) option is selected by default.

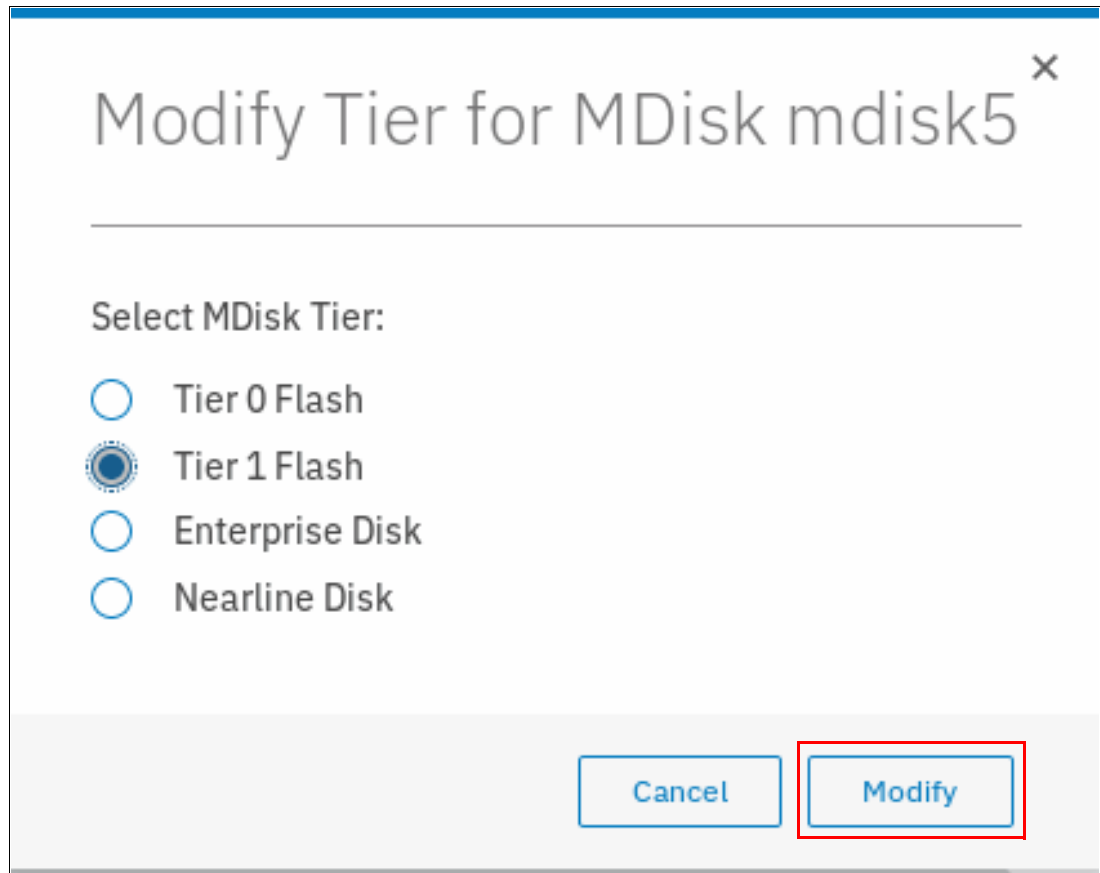


Figure 4-60 Modifying external MDisk tier

Selecting **Modify Encryption** allows the user to modify the encryption setting for the MDisk, as shown in Figure 4-61 on page 208. This option is available only when encryption is enabled.

For example, if the external MDisk is encrypted by the external storage system, change the encryption state of the MDisk to **Externally encrypted**. This process stops the system from encrypting the MDisk again if the MDisk is part of an encrypted storage pool.

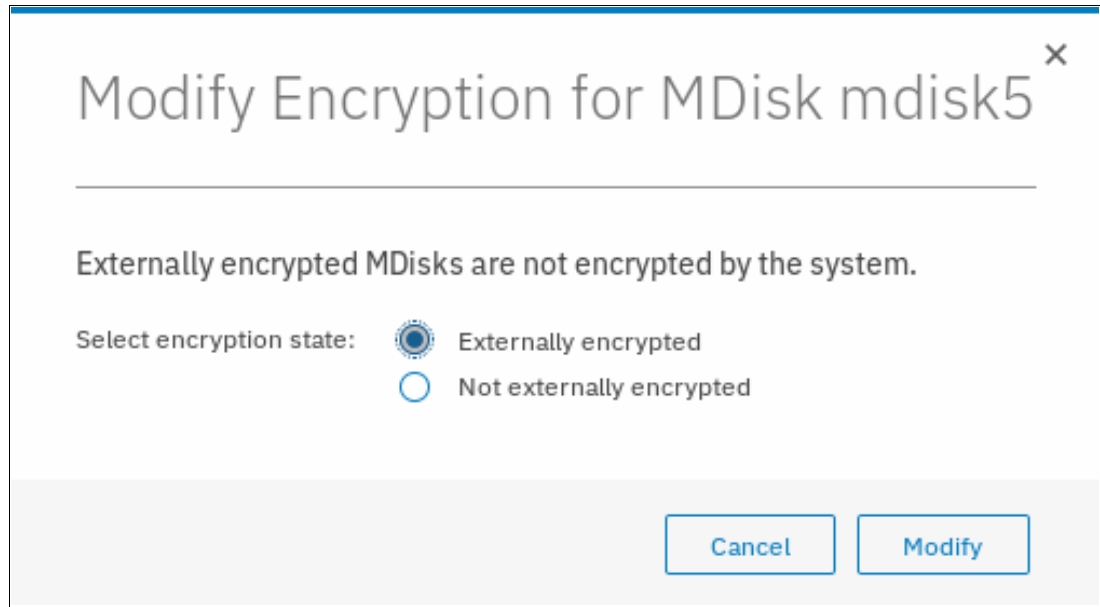


Figure 4-61 Modifying external MDisk encryption

Import

This action is available only for unmanaged MDisks. Importing an unmanaged MDisk allows the user to preserve the data on the MDisk by migrating the data to a new volume or keeping the data on the external system.

Selecting **Import** allows you to choose one of the following migration methods:

- ▶ The Import to temporary pool as image-mode volume option does not migrate data from the source MDisk. It creates an *image-mode volume* that has a direct block-for-block translation of the MDisk. The data is preserved on the external storage system, but it is also accessible from the IBM Storwize V5000 Gen2 system.

If this method is selected, the image-mode volume is created in a temporary migration pool and presented through the IBM Storwize V5000 Gen2. Choose the extent size of the temporary pool and click **Import**, as shown in Figure 4-62 on page 209.

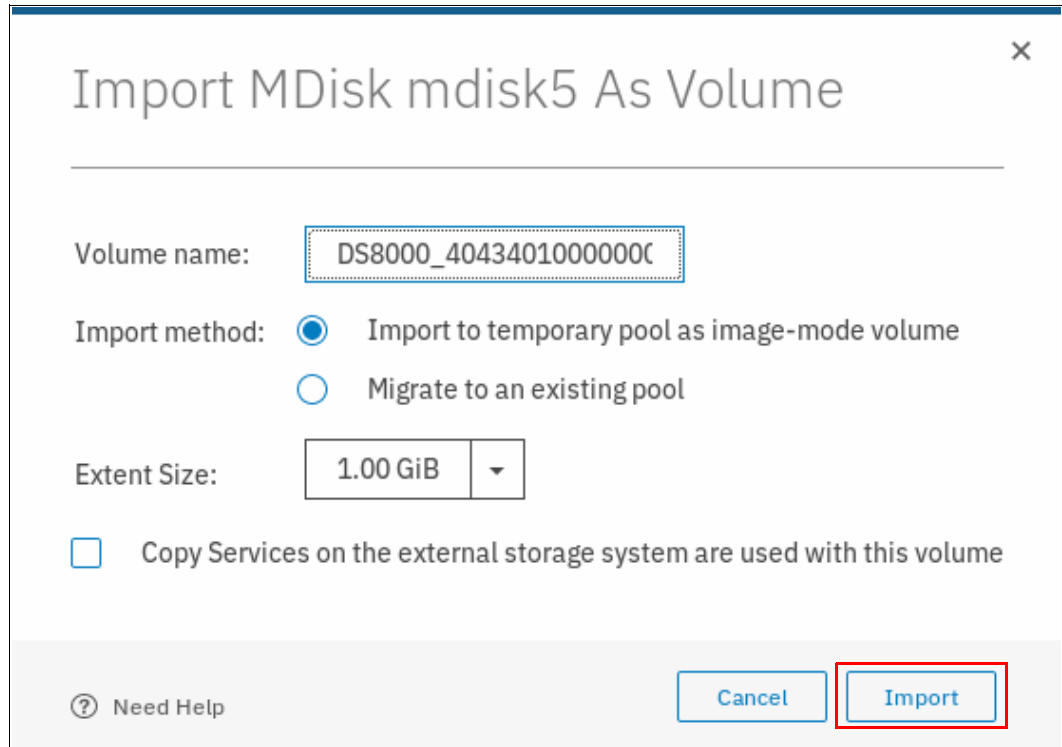


Figure 4-62 Importing an external MDisk as an image-mode volume

The MDisk is imported and listed as an image-mode MDisk in the temporary migration pool, as shown in Figure 4-63.

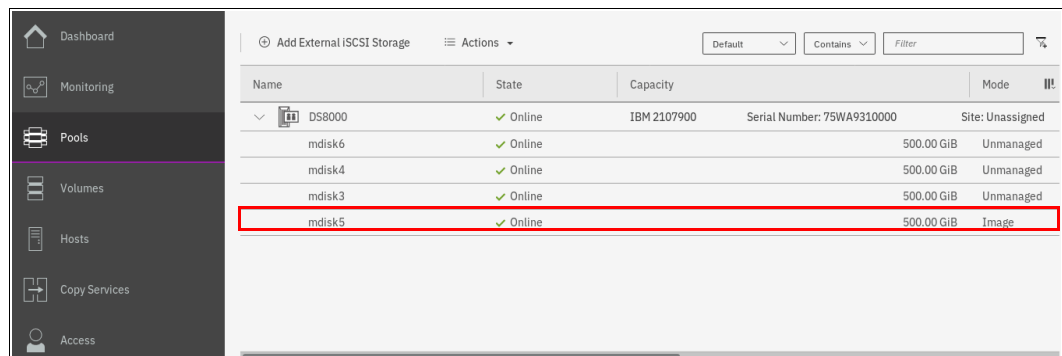


Figure 4-63 Image-mode MDisk

A corresponding image-mode volume is now available in the same migration pool, as shown in Figure 4-64.

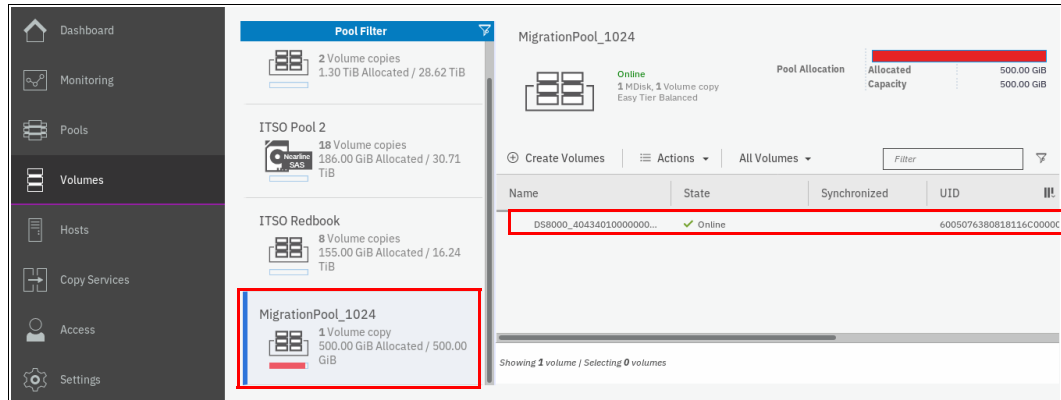


Figure 4-64 Corresponding image-mode volume

The image-mode volume can then be mapped to the original host mode. The data is still physically present on the physical disk of the original external storage controller system and no automatic migration process is running. If needed, the image-mode volume can be migrated manually to another storage pool by using volume migration or volume mirroring later.

- The Migrate to an existing pool option starts by creating an image-mode volume as the first method. However, it then migrates the data from the image-mode volume onto another volume in the selected storage pool. After the migration process completes, the image-mode volume and temporary migration pool are deleted.

If this method is selected, choose the storage pool to hold the new volume and click **Import**, as shown in Figure 4-65. Only pools with sufficient free extent capacity are listed.

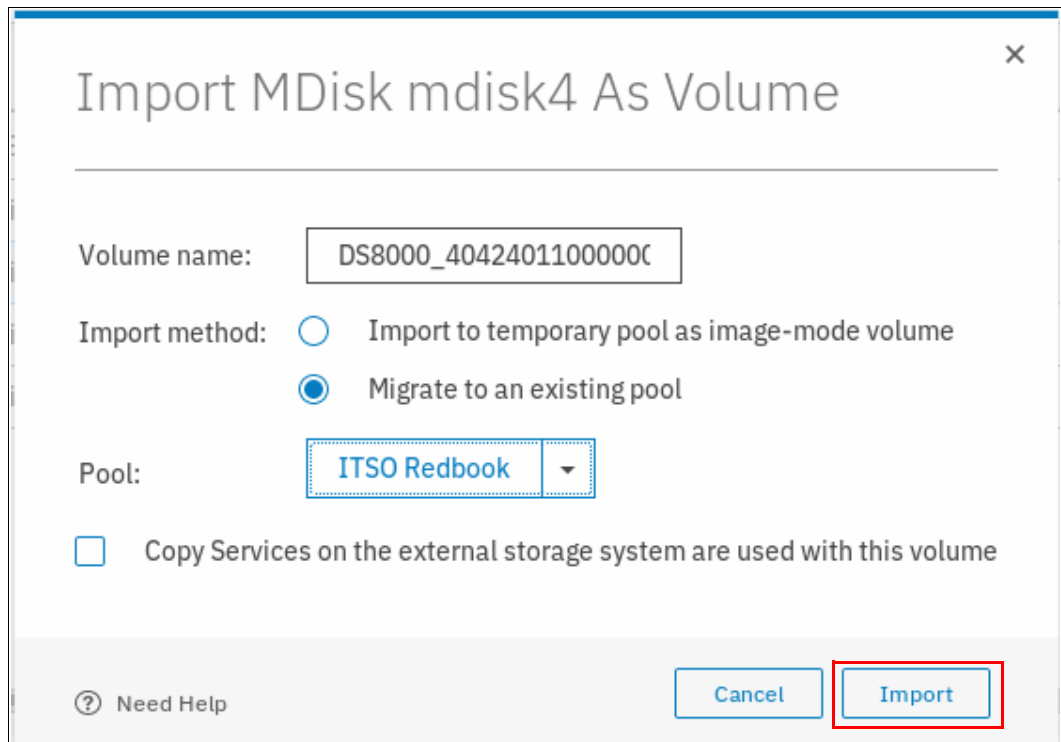


Figure 4-65 Importing an external MDisk to an existing pool

The data migration begins automatically after the MDisk is imported successfully as an image-mode volume. You can check the migration progress by browsing to **Pools** → **System Migration**, as shown in Figure 4-66.

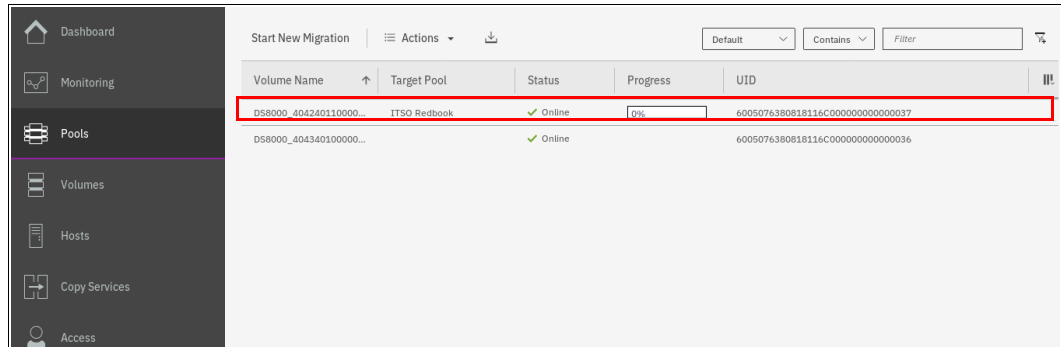


Figure 4-66 Importing external MDisk progress

After the migration completes, the volume is available in the chosen destination pool, as shown in Figure 4-67. This volume is no longer an image-mode volume; it is a normal striped volume.

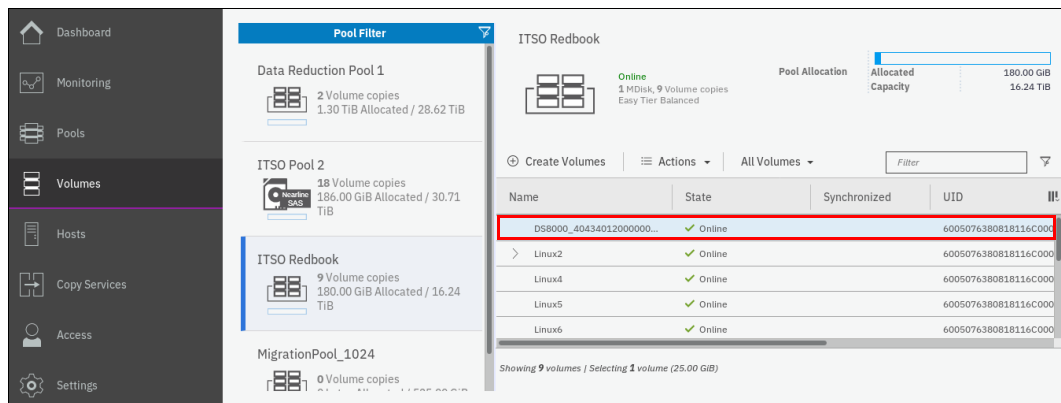


Figure 4-67 Striped volume after migration

At this point, all data is migrated from the source MDisk and the MDisk is no longer in image mode, as shown in Figure 4-68. The MDisk can be removed from the temporary pool and used as a regular MDisk to host volumes.

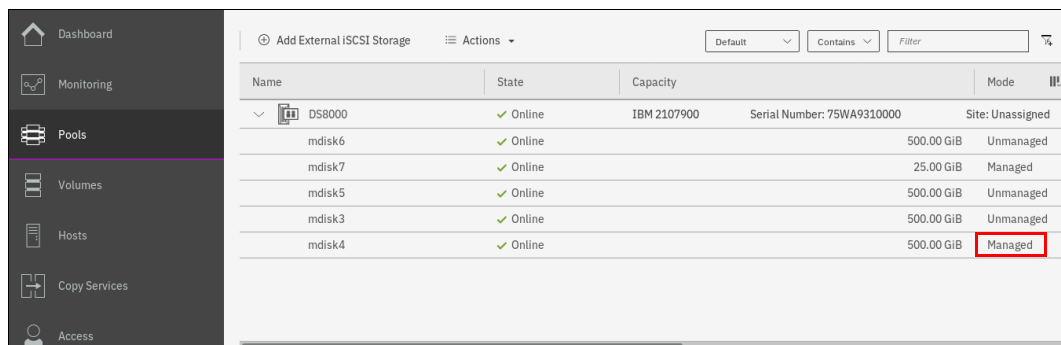


Figure 4-68 Volume shows in Managed mode

Alternatively, import and migration of external MDisks to another pool can be done by selecting **Pools** → **System Migration**.

For more information about migration and the system migration wizard, see Chapter 7, “Storage migration” on page 357.

Include

The system can exclude an MDisk with multiple I/O failures or persistent connection errors from its storage pool to ensure that these errors do not interfere with data access. If an MDisk was automatically excluded, run the fix procedures to resolve any connection and I/O failure errors. Drives that are used by the excluded MDisk with multiple errors might require replacing or reseating.

After the problems are fixed, select **Include** to add the excluded MDisk back into the storage pool.

Remove

In some cases, you might want to remove external MDisks from storage pools to reorganize your storage allocation. Selecting **Remove** removes the MDisk from the storage pool. After the MDisk is removed, it returns to unmanaged. If no volumes are in the storage pool to which this MDisk is allocated, the command runs immediately without any confirmation. If volumes are in the pool, you are prompted to confirm the action, as shown in Figure 4-69. Click **Yes** to migrate the volumes or **No** to cancel the deletion process.

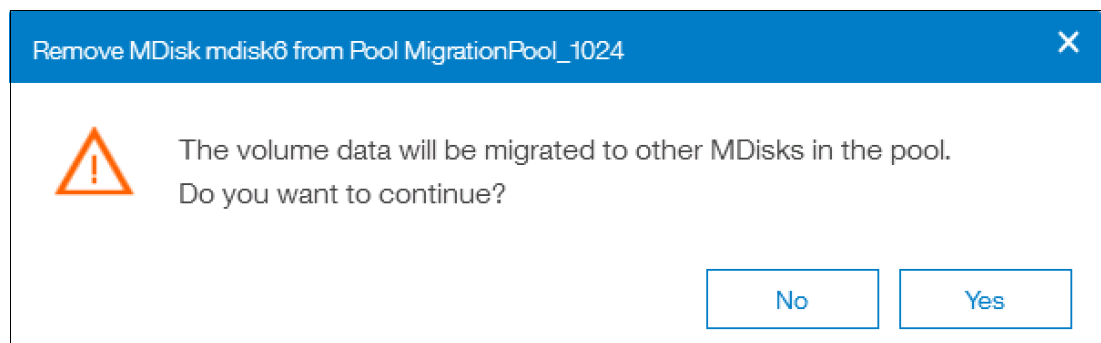


Figure 4-69 Removing an external MDisk

Confirming the action starts the migration of the volumes to extents from other MDisks that remain in the pool. When the action completes, the MDisk is removed from the storage pool and returns to unmanaged.

Important: Ensure that you have enough available capacity remaining in the storage pool to allocate the data being migrated from the removed MDisk or else the command fails.

The MDisk that is removed must remain accessible to the system while all data is copied to other MDisks in the same storage pool. If the MDisk is unmapped before the migration finishes, all volumes in the storage pool go offline and remain in this state until the removed MDisk is connected again.

4.3.6 More actions on MDisks

In this section, we describe several other actions that supported on arrays and external MDisks.

Discover storage

The Discover storage option in the upper left of the MDisks by Pools window is useful if external storage controllers are in your environment. (For more information, see Chapter 11, “External storage virtualization” on page 607). The Discover storage action starts a rescan of the Fibre Channel network. It discovers any new MDisks that were mapped to the IBM Storwize V5000 Gen2 storage system and rebalances MDisk access across the available controller device ports.

This action also detects any loss of controller port availability and updates the IBM Storwize V5000 Gen2 configuration to reflect any changes.

When external storage controllers are added to the IBM Storwize V5000 Gen2 environment, the IBM Storwize V5000 Gen2 automatically discovers the controllers. The logical unit numbers (LUNs) that are presented by those controllers are listed as unmanaged MDisks.

However, if you attached new storage and the IBM Storwize V5000 Gen2 did not detect it, you might need to use the Discover storage option before the system detects the new LUNs. If the configuration of the external controllers is modified later, the IBM Storwize V5000 Gen2 might be unaware of these configuration changes. Use Detect MDisk to rescan the Fibre Channel network and update the list of unmanaged MDisks.

Figure 4-70 shows the Discover storage option.

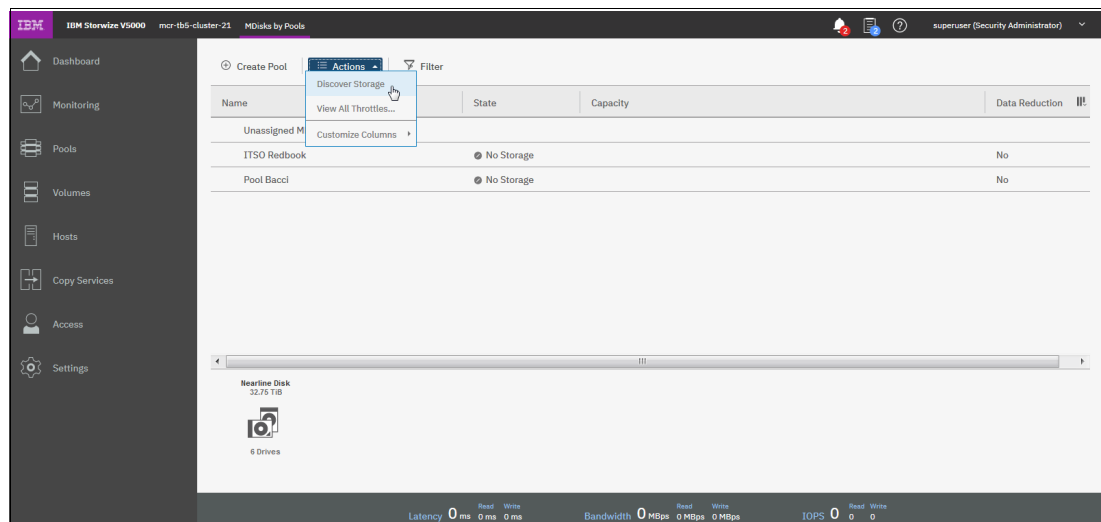


Figure 4-70 Discover storage action

Note: The Discover storage action is asynchronous. Although the task appears to be finished, it might still be running in the background.

Rename

MDisks can be renamed by selecting the MDisk and clicking **Rename** from the Actions menu. Enter the new name of your MDisk (as shown in Figure 4-71) and click **Rename**.

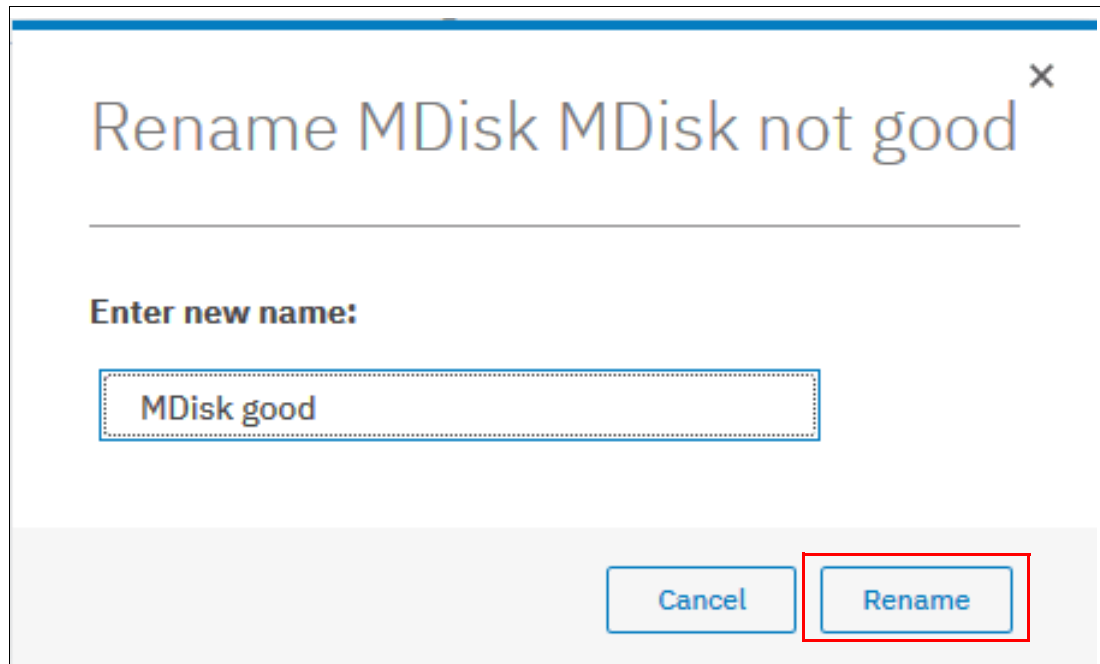


Figure 4-71 Rename MDisk

Show Dependent Volumes

Figure 4-72 shows the volumes that depend on an MDisk. The volumes can be displayed by selecting the MDisk and clicking **Show Dependent Volumes** from the **Actions** menu. The volumes are listed with general information.

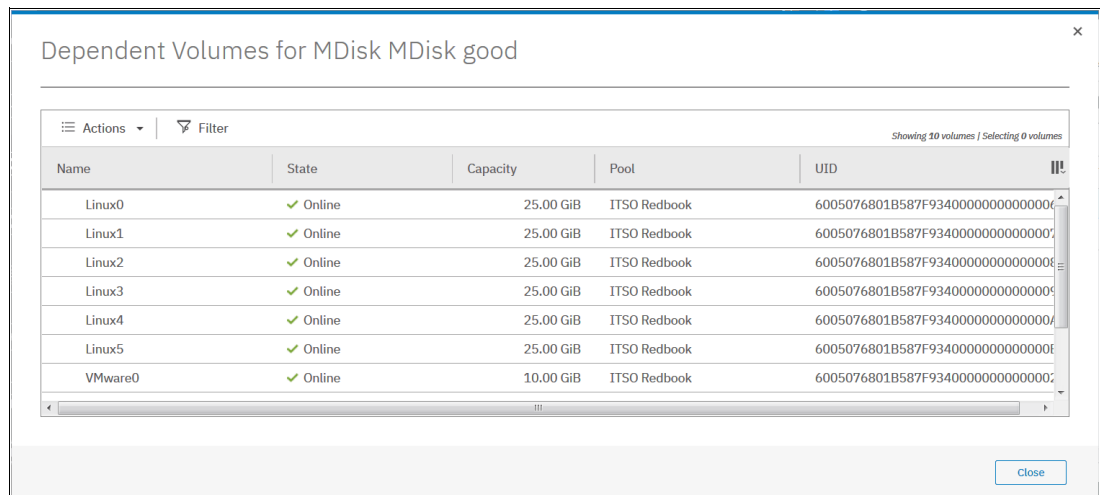


Figure 4-72 Show MDisk dependent volumes

Properties

The Properties action for an MDisk shows the information that you need to identify it. In the MDisks by Pools window, select the MDisk and click **Properties** from the **Actions** menu. Alternatively, right-click the MDisk and select **Properties**. For more information about the selected MDisk, click **Copy 0**, as shown in Figure 4-73.

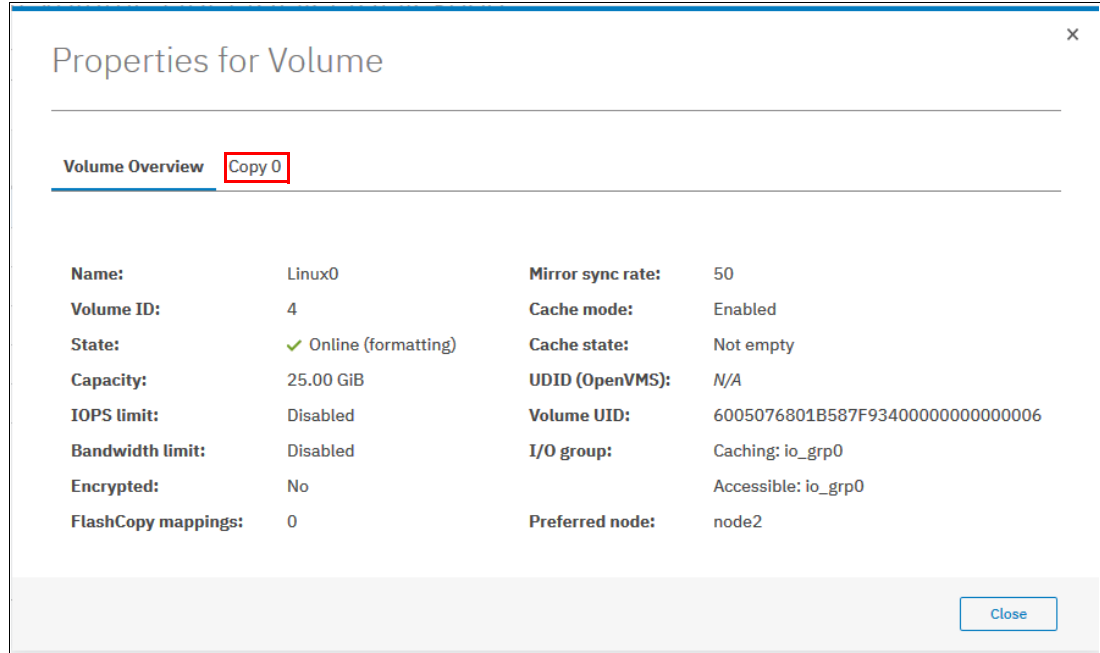


Figure 4-73 MDisk properties

4.4 Working with external storage controllers

After the internal storage configuration is complete, you can find the MDisks that were created by using the internal drives in the MDisks by Pools window. When you use this window, you can manage all MDisks that are made up of internal and external storage.

LUNs that are presented by external storage systems to IBM Storwize V5000 Gen2 are discovered as unmanaged MDisks. Initially, the MDisk is not a member of any storage pools, which means that it is not used by the IBM Storwize V5000 Gen2 storage system.

For more information about external storage, see Chapter 11, “External storage virtualization” on page 607.



Host configuration

This chapter describes how to use the IBM Storwize V5000 graphical user interface (GUI) to create hosts and how to prepare a host to access the volumes that are created. (Volume creation is described in Chapter 6, “Volume configuration” on page 309.)

The chapter includes the following topics:

- ▶ 5.1, “Host attachment overview” on page 218
- ▶ 5.2, “Planning for direct-attached hosts” on page 218
- ▶ 5.3, “Preparing the host operating system” on page 219
- ▶ 5.4, “N-Port ID Virtualization support” on page 246
- ▶ 5.5, “Creating hosts by using the GUI” on page 255
- ▶ 5.6, “Host Clusters” on page 280
- ▶ 5.7, “Proactive Host Failover” on page 306

5.1 Host attachment overview

A host system is an open-systems computer that is connected to the switch through a Fibre Channel (FC), direct-attached, serial-attached SCSI (SAS) connection or an internet Small Computer System Interface (iSCSI).

The IBM Storwize V5000 supports the following host attachment protocols:

- ▶ 16 Gb FC *or* 10 Gb iSCSI/FC over Ethernet (FCoE) as an optional host interface
- ▶ 12 Gb SAS (standard host interface)
- ▶ 1 Gb or 10 Gb iSCSI (standard host interface, depending on the model)

In this chapter, we assume that your hosts are connected to your FC, SAS, or Internet Protocol (IP) network and you completed the steps that are described in Chapter 2, “Initial configuration” on page 39. Follow basic zoning recommendations to ensure that each host has at least two network adapters, that each adapter is on a separate network (or at minimum in a separate zone), and that each adapter is connected to both canisters. This setup ensures four paths for failover and failback.

For SAS attachment, ensure that each host has at least one SAS host bus adapter (HBA) connection to each IBM Storwize V5000 canister. For more information about configuring SAS attached hosts, see 2.4, “SAS direct-attach planning” on page 48.

Before you map the newly created volumes on the host of your choice, preparation goes a long way toward ease of use and reliability. Several steps are required on a host in preparation for mapping new Storwize V5000 volumes to the host. Use the IBM System Storage Interoperation Center (SSIC) to check the code levels that are supported to attach your host to your storage. The SSIC is a web tool that checks the interoperation of host, storage, switches, and multipathing drivers. The SSIC is available at [this website](#).

The complete support matrix is listed in the *IBM Storwize V5000 Supported Hardware List, Device Driver, Firmware, and Recommended Software Levels* document.

This chapter focuses on Windows and VMware. If you want to attach any other hosts (for example, IBM AIX®, Linux, or an Apple system), the required information is available at [this website](#).

5.2 Planning for direct-attached hosts

Starting with V7.5, we supported direct-attached FC hosts. A direct-attached configuration dedicates the entire port bandwidth for use for a specific connection. Planning must account for the volume of expected traffic before you decide how many ports are used in a direct-attached configuration. The Storwize V5000 system offers multiple options for you to decide how to create a direct-attached configuration.

Because of the bandwidth requirements when you use a direct-attached configuration, it is important to determine the volume of expected traffic when you decide the number of required ports. For example, if a Storwize V5000 to Storwize V5000 direct-attached link is configured between nodes, the link might carry intra-node traffic, such as FlashCopy data. Therefore, enough Storwize V5000 to Storwize V5000 bandwidth must be available so that it can carry all possible intra-node traffic without any bandwidth bottleneck.

The following guidelines are provided for direct-attached configurations.

5.2.1 FC direct attachment to host systems

The Storwize V5000 supports direct attachment connectivity between its FC ports and host ports. Host systems can connect to 16 Gb FC ports on the Storwize V5000. No special configuration is required for host systems that use this configuration.

5.2.2 FC direct attachment between nodes in a Storwize V5000 system

Direct connection of the Storwize V5000 FC ports without the use of an FC switch is supported. Such connections between the Storwize V5000 nodes might be useful in small configurations where no FC switch exists. It can also be used to connect nodes in the same input/output (I/O) group to provide a dedicated connection for mirroring the fast write cache data. Ensure that sufficient bandwidth is provisioned between nodes to accommodate all of the intra-node traffic.

The Storwize V5000 Gen2 supports 16 Gb Fibre Channel ports and FC direct attachment on all 16 Gb ports.

Note: Be careful about the maximum length of the FC links in this configuration.

5.3 Preparing the host operating system

The following steps are required to prepare a host to connect to the Storwize V5000:

1. Ensure that the latest supported system updates are applied to your host operating system.
2. Ensure that the HBAs are physically installed in the host.
3. Ensure that the latest supported HBA firmware and driver levels are installed in your host.
4. Configure the HBA parameters. Although settings are provided for each host operating system in the following sections, review the Storwize V5000 IBM Knowledge Center to obtain the latest supported settings.
5. Configure the host I/O parameters, such as the disk I/O timeout value.
6. Install and configure multipath software.
7. Determine the host worldwide port names (WWPNs).
8. Connect the HBA ports to switches by using the correct cables, or directly attach to the ports on the IBM Storwize V5000.
9. Configure the switches, if applicable.
10. Configure SAN Boot (optional).

5.3.1 Windows 2008 R2 and 2012 R2: Preparing for FC attachment

Complete the following steps to prepare a Windows 2008 R2 or Windows 2012 R2 host to connect to an IBM Storwize V5000 by using FC:

1. Ensure that the current operating system service pack and test fixes are applied to your server.
2. Use the current firmware and driver levels on your host system.

3. Install an HBA or HBAs on the Windows server by using the current basic input/output system (BIOS) and drivers.
4. Connect the FC host adapter ports to the switches, or use direct connections.
5. Configure the switches (zoning).
6. Configure the HBA for hosts that run Windows.
7. Set the Windows timeout value.
8. Install the multipath module.

Downloading and installing the supported drivers and firmware

Install a supported HBA driver for your configuration. Use the Windows Device Manager or vendor tools, such as QLogic Converged Console (QCC) or HBAnyware (Emulex), to install the driver.

Brocade adapter software is now maintained by QLogic, so check the QLogic web pages to obtain the correct support for your Brocade adapters. Also, check and update the BIOS (firmware) level of the HBA by using the manufacturer's provided tools. Check the readme file to see whether any Windows registry parameters must be set for the HBA driver.

For more information about current supported levels, see [this website](#).

Configuring Brocade HBAs for Windows

This section applies to Windows hosts with installed Brocade HBAs. After you install the device driver and firmware, you must configure the HBAs. To perform this task, use the Brocade Host Connectivity Manager (HCM) software or restart into the HBA BIOS, load the adapter defaults, and set the following values:

- ▶ Host Adapter BIOS: Disabled (unless the host is configured for storage area network (SAN) Boot)
- ▶ Queue depth: 4

Configuring QLogic HBAs for Windows

This section applies to Windows hosts with installed QLogic HBAs.

After you install the device driver and firmware, you must configure the HBAs. To perform this task, use the QLogic QConverge Console (QCC) command-line interface (CLI) software or restart into the HBA BIOS, load the adapter defaults, and set the following values:

- ▶ Host Adapter BIOS: Disabled (unless the host is configured for SAN Boot)
- ▶ Adapter Hard Loop ID: Disabled
- ▶ Connection Options: 1 (only point to point)
- ▶ Logical unit numbers (LUNs) Per Target: 0
- ▶ Port Down Retry Count: 15

For more information about the QCC Control Center Software, see [this website](#).

Configuring Emulex HBAs for Windows

This section applies to Windows hosts with installed Emulex HBAs.

After you install the device driver and firmware, you must configure the HBAs. To perform this task, use the Emulex HBAnyware software or restart into the HBA BIOS, load the defaults, and set topology to 1 (10F_Port Fabric).

Setting the Windows timeout value

For Windows hosts, the disk I/O timeout value must be set to 60 seconds. To verify this setting, complete the following steps:

1. Click **Start** → **Run**. Alternatively, open a Power Shell window.
2. In the dialog box or Power Shell window, enter `regedit` and press Enter.
3. In the registry editor, locate the `HKEY_LOCAL_MACHINE\System\CurrentControlSet\Services\disk\TimeOutValue` key.
4. Confirm that the value for the key is 60 (decimal value), and, if not, change the value to 60 (see Figure 5-1).

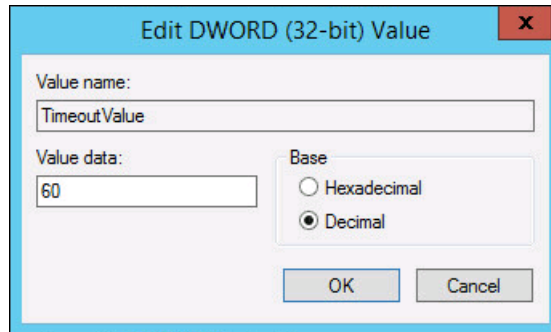


Figure 5-1 Windows timeout value

Installing the multipathing software

Microsoft Multipath Input/Output (MPIO) solutions work with device-specific modules (DSMs) that are written by vendors. However, the MPIO driver package does not, by itself, form a complete solution. This joint solution enables the storage vendors to design device-specific solutions that are tightly integrated with the Windows operating system. MPIO is not included with the Windows operating system. Storage vendors must pack the MPIO drivers with their own DSM.

IBM Subsystem Device Driver DSM (SDDDSM) is the IBM multipath I/O solution that is based on Microsoft MPIO technology. It is a device-specific module that supports IBM storage devices on Windows hosts. The intent of MPIO is to provide better integration of a multipath storage solution with the operating system. It also supports the use of multipath in the SAN infrastructure during the startup process for SAN Boot hosts.

To ensure correct multipathing with the Storwize V5000, SDDDSM must be installed on Windows hosts. To install SDDDSM, complete the following steps:

1. Go to the following SDDDSM download matrix to determine the correct level of SDDDSM to install for Windows 2008 R2 or Windows 2012 R2, and then download the package from this IBM Support [web page](#).
2. Extract the package to your hard disk drive, and run `setup.exe` to install SDDDSM. A command prompt window opens (see Figure 5-2). Confirm the installation by entering Yes.

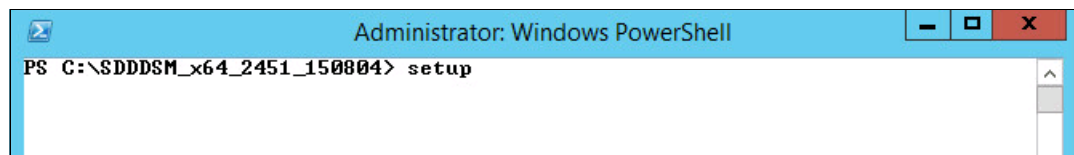
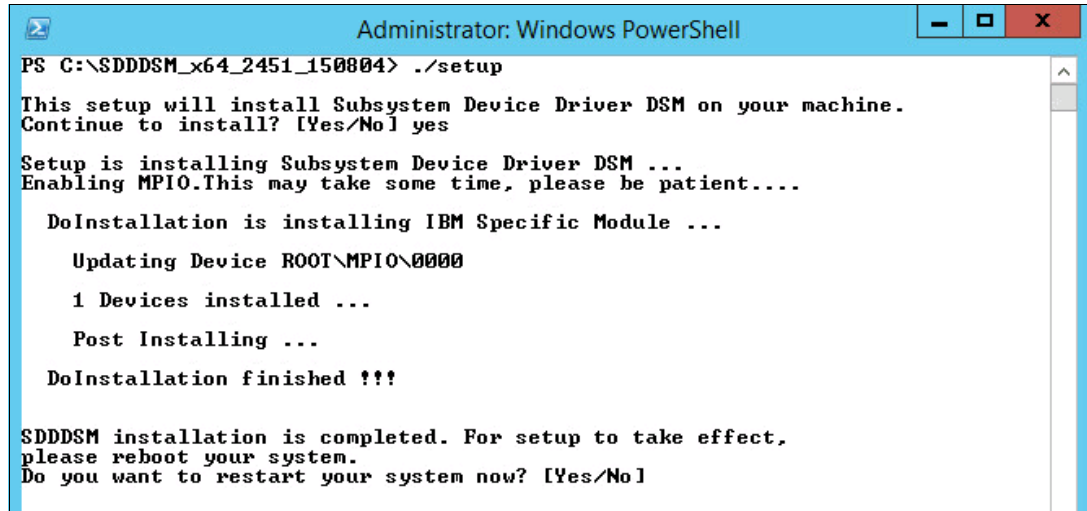


Figure 5-2 SDDDSM setup

During the setup, SDDDSM also determines whether an older version is installed and prompts you to upgrade to the current version.

3. After the setup completes, you are prompted to restart the system. Confirm this restart by entering Yes and pressing Enter (see Figure 5-3).



```
Administrator: Windows PowerShell
PS C:\SDDDSM_x64_2451_150804> ./setup
This setup will install Subsystem Device Driver DSM on your machine.
Continue to install? [Yes/No] yes
Setup is installing Subsystem Device Driver DSM ...
Enabling MPIO. This may take some time, please be patient....
DoInstallation is installing IBM Specific Module ...
Updating Device ROOT\MPIO\0000
1 Devices installed ...
Post Installing ...
DoInstallation finished !!!
SDDDSM installation is completed. For setup to take effect,
please reboot your system.
Do you want to restart your system now? [Yes/No]
```

Figure 5-3 Answer Yes to restart the host

You successfully installed IBM SDDDSM. To check whether IBM SDDDSM is installed correctly, see the next section and “Windows 2008 R2” on page 222.

Windows 2008 R2

To check the installed driver version, complete the following steps:

1. Select **Start** → **All Programs** → **Subsystem Device Driver DSM** → **Subsystem Device Driver DSM**.
2. A command prompt opens. Run **datapath query version** to determine the version that is installed (see Example 5-1) for this Windows 2008 R2 host.

Example 5-1 The datapath query version command

```
C:\Program Files\IBM\SDDDSM>datapath query version
IBM SDDDSM Version 2.4.5.1
Microsoft MPIO Version 6.1.7601.17514
```

3. The worldwide port names (WWPNs) of the FC HBA are required to correctly zone switches and configure host attachment on the IBM Storwize V5000. You can obtain the WWPNs by using vendor tools, HBA BIOS, native Windows command line, or SDDDSM. This command can be used to determine the worldwide port names (WWPNs) of the host. Run **datapath query wwpn** (see Example 5-2) and note the WWPNs of your host because you need them later.

Example 5-2 The datapath query wwpn command

```
C:\Program Files\IBM\SDDDSM>datapath query wwpn
Adapter Name      PortWWN
Scsi Port3:      21000024FF35B960
Scsi Port4:      21000024FF25B961
```

For more information about SDDDSM, see *Multipath Subsystem Device Driver User's Guide*, GC52-1309. The guide is available at this IBM Support [web page](#).

Windows 2012 R2

To check the installed driver version, complete the following steps:

1. Select the **Windows Start** icon in the lower-left corner, as shown in Figure 5-4.



Figure 5-4 Windows 2012 R2 start

2. Expand the view by clicking the down arrow that is shown in Figure 5-5.

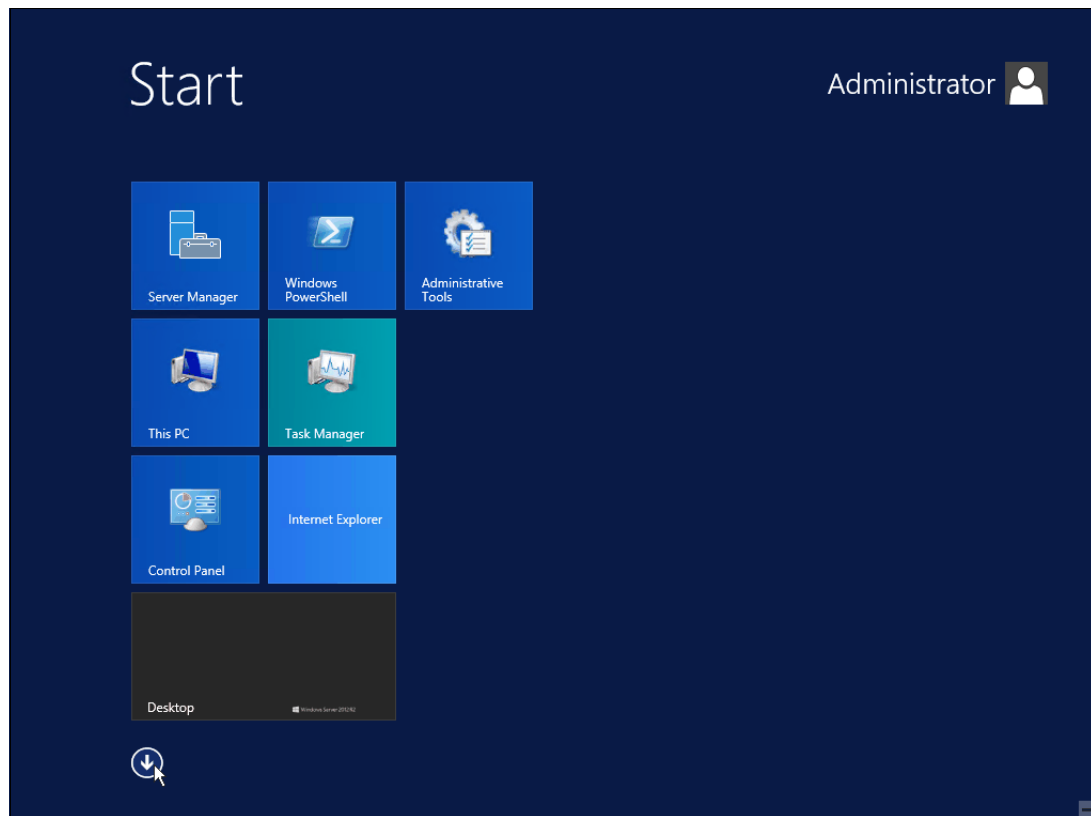


Figure 5-5 Expand view to see all programs that are installed

3. Search for the section *Subsystem Device Driver DSM* (see Figure 5-6).

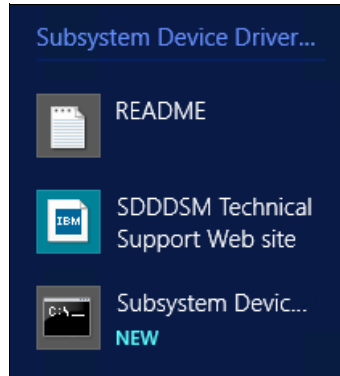


Figure 5-6 Subsystem Device Driver DSM in the all programs menu

4. Click **Subsystem Device Driver DSM** (see Figure 5-7).

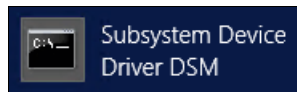


Figure 5-7 Subsystem Device Driver DSM

5. A command prompt opens. Run **datapath query version** to determine the version that is installed (see Figure 5-8).

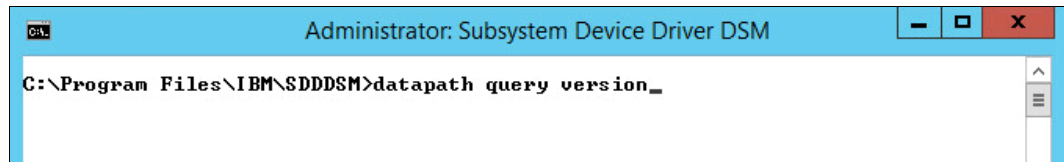


Figure 5-8 Datapath query version

The Windows 2012 R2 host is shown in Example 5-3.

Example 5-3 The datapath query version command

```
C:\Program Files\IBM\SDDDSM>datapath query version
IBM SDDDSM Version 2.4.5.1
Microsoft MPIIO Version 6.3.9600.16384
C:\ProgramFiles\IBM\SDDDSM>
```

6. The WWPNs of the FC HBA are required to correctly zone switches and configure host attachment on the IBM Storwize V5000. You can obtain the WWPNs by using vendor tools, HBA BIOS, native Windows command line, or SDDDSM. This command can be used to determine the WWPNs of the host. Run **datapath query wwpn** (see Example 5-4) and document the WWPNs of your host because you need them later.

Example 5-4 The datapath query wwpn command

```
C:\Program Files\IBM\SDDDSM>datapath query wwpn
Adapter Name      PortWWN
Scsi Port 7      100000051EC76B89
Scsi Port 7      100000051EC76B8A
```

For more information about SDDDSM, see *Multipath Subsystem Device Driver User's Guide*, GC52-1309, which is available at [this website](#).

The Windows host was prepared to connect to the Storwize V5000, and you know the WWPNs of the host. The next step is to configure a host object for the WWPNs by using the Storwize V5000 GUI. For more information about this process, see 5.5.1, “Creating FC hosts” on page 258.

SAN Boot hosts are beyond the intended scope of this book. For more information, search for “[hSAN Boot](#)” in IBM Knowledge Center.

Windows 2003: The examples focus on Windows 2008 R2 and Windows 2012, but the procedure for Windows 2003 is similar. If you use Windows 2003, do not forget to install Microsoft hotfix 912944. If you do not install it before you perform this procedure, preferred pathing is not available.

5.3.2 Windows 2008 R2 and Windows 2012 R2: Preparing for iSCSI attachment

This section describes iSCSI attachment.

Installing and updating supported HBAs

Install a supported HBA model with the latest supported firmware and drivers for your configuration. The latest supported HBAs and levels for Windows 2008 R2 and 2012 R2 are available at [this website](#).

Install the driver by using Windows Device Manager or vendor tools. Also, check and update the firmware level of the HBA by using the manufacturer's provided tools. Always check the readme file to see whether any Windows registry parameters must be set for the HBA driver.

Important: For converged network adapters (CNAs), which can support FC and iSCSI, it is important to ensure that the Ethernet networking driver is installed in addition to the FCoE driver. You are required to install the Ethernet networking driver and the FCoE driver before you configure iSCSI.

If you use a hardware iSCSI HBA, refer to the manufacturer's documentation and the Storwize V5000 IBM Knowledge Center for more information about the hardware and host operating system configuration. The following section describes how to configure iSCSI by using the software initiator.

In Windows 2008 R2 and 2012, the Microsoft iSCSI software initiator is preinstalled.

Complete the following steps:

1. Enter `iscsi` in the search field of the Windows 2008 R2 Start window (see Figure 5-9) and click **iSCSI Initiator**.

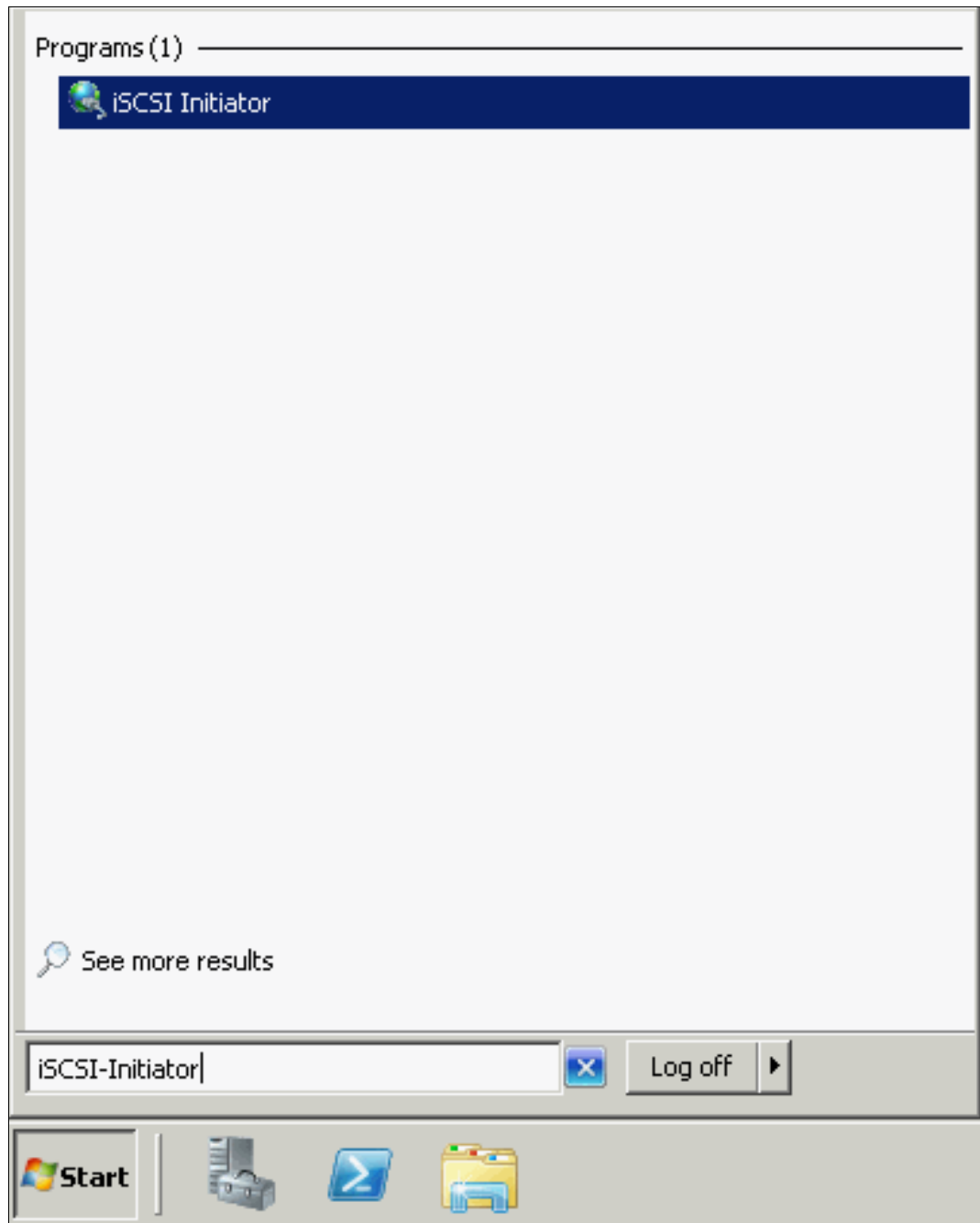


Figure 5-9 Windows 2008 R2 iSCSI Initiator

2. For Windows 2012 R2, go to the all programs menu and enter iSCSI in the search field at the top of the window (see Figure 5-10).



Figure 5-10 iSCSI Initiator Windows 2012 R2

3. Confirm the automatic start of the iSCSI service (see Figure 5-11).

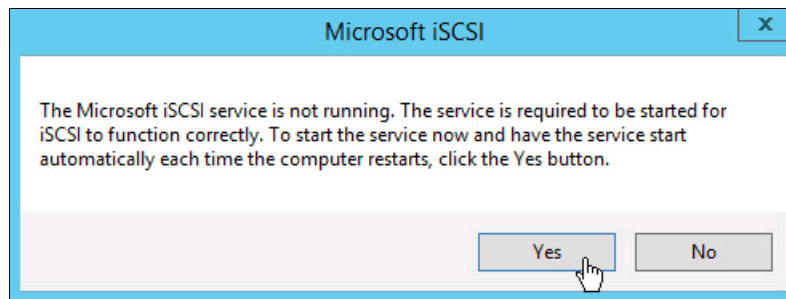


Figure 5-11 Automatic start of the iSCSI service

- The iSCSI Initiator Properties window opens. Select the **Configuration** tab (see Figure 5-12). Write down the initiator name of your Windows host because you use it later.

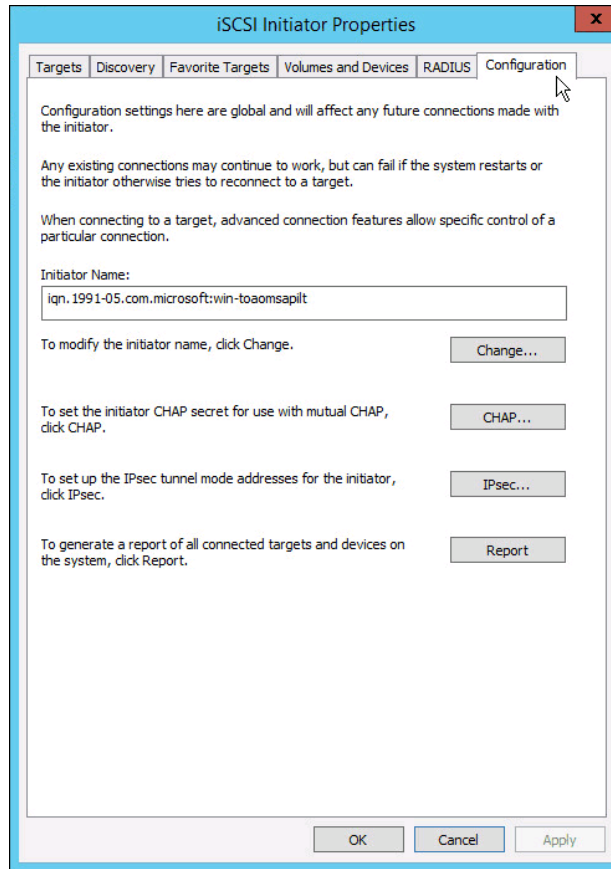


Figure 5-12 iSCSI Initiator Properties window

- You can change the initiator name, or enable advanced authentication, but these actions are out of the scope of our basic setup. By default, iSCSI authentication is not enabled. For more information, see this IBM Knowledge Center for the Storwize V5000 [this web page](#).
- From the Configuration tab, you can change the initiator name, enable CHAP authentication, and more. However, these tasks are beyond the scope of our basic setup. CHAP authentication is disabled, by default. For more information, see [Microsoft iSCSI Initiator Step-by-Step Guide](#).

Configuring Ethernet ports

We suggest that you use separate dedicated ports for host management and iSCSI. In this case, we must configure IPs on the iSCSI ports in the same subnet and virtual LAN (VLAN) as the external storage to which we want to attach.

To configure Ethernet port IPs on Windows 2008 R2 and 2012 R2, complete the following steps:

1. Go to **Control Panel** → **Network and Internet** → **Network and Sharing Center**. The window that is shown in Figure 5-13 opens.

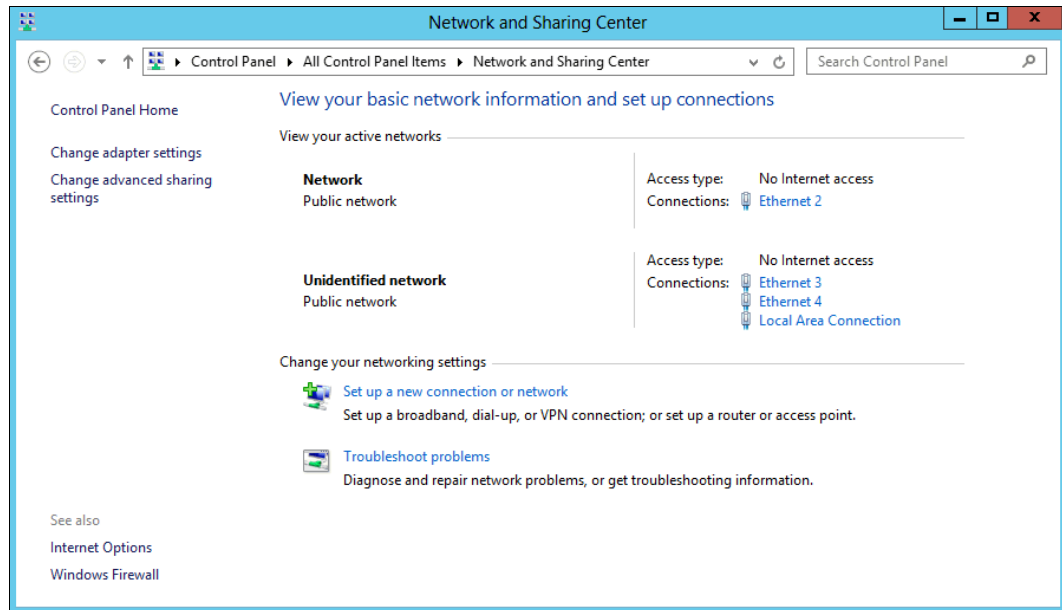


Figure 5-13 Network and Sharing Center in Windows 2012 R2

In this case, two networks are visible to the system. We use the first network to connect to the server. It consists of a single dedicated Ethernet port for management. The second network is our iSCSI network. It consists of two dedicated Ethernet ports for iSCSI. We suggest that you use at least two dedicated ports for failover purposes.

- To configure an IP address, click one of the iSCSI Ethernet connections (in this case, Ethernet 3 or Ethernet 4). Figure 5-14 shows the window that displays the Ethernet status.

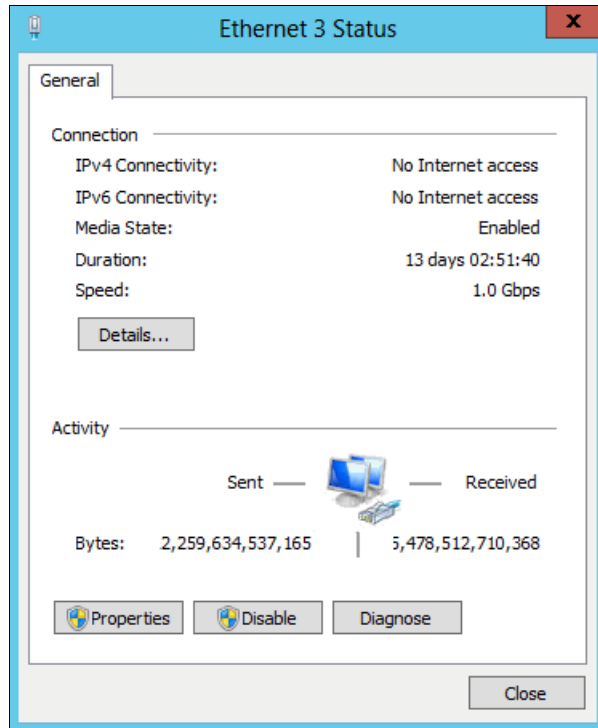


Figure 5-14 Ethernet status

- To configure the IP address, click **Properties** (see Figure 5-15).

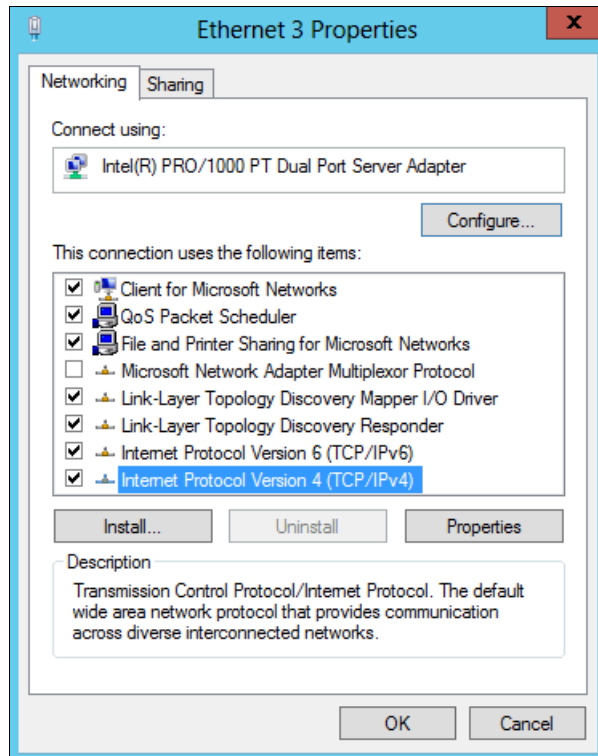


Figure 5-15 Ethernet properties

- If you use IPv6, select **Internet Protocol Version 6 (TCP/IPv6)** and click **Properties**. Otherwise, select **Internet Protocol Version 4 (TCP/IPv4)** and click **Properties** to configure an IPv4 address.
- For IPv4, the window that is shown in Figure 5-16 opens. To manually set the IP, select **Use the following address** and enter an IP address, subnet mask, and gateway. Set the DNS server address, if required. Click **OK** to confirm.

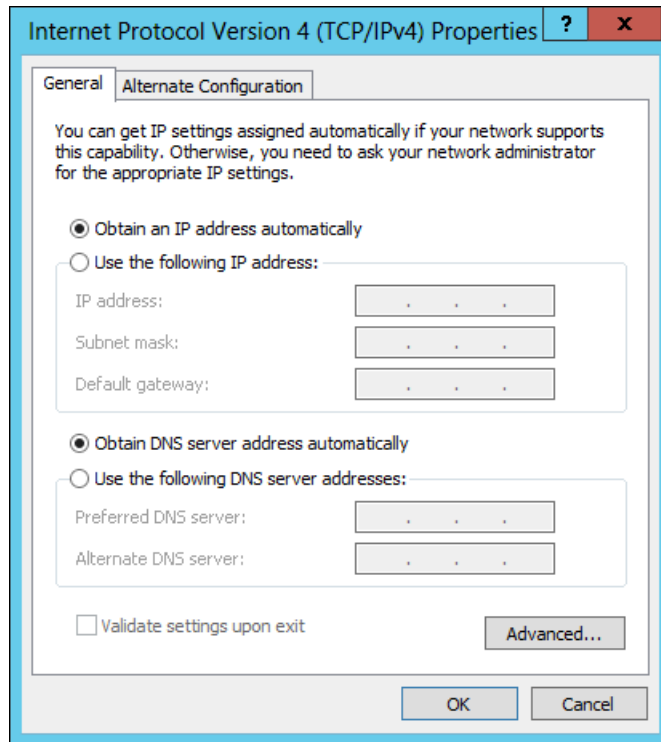


Figure 5-16 Configuring an IPv4 address in Windows 2012 R2

- Repeat these steps for each port that you want to configure for iSCSI attachment.

The Ethernet ports are now prepared for iSCSI attachment.

Setting the Windows registry keys

Complete the following steps to modify the system registry so that your iSCSI operations are more reliable:

- In the search field of the Windows Start window, enter `regedit` and click **regedit.exe**.
- In the registry editor, locate the following key:

```
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Class\{4D36E97B-E325-11CE-BFC1-08002BE10318}\<bus ID>\Parameters\LinkDownTime
```

Confirm that the value for the `LinkDownTime` key is 120 (decimal value) and, if not, change the value to 120.

- In the registry editor, locate the following key:

```
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Class\{4D36E97B-E325-11CE-BFC1-08002BE10318}\<bus ID>\Parameters\MaxRequestHoldTime
```

Confirm that the value for the `MaxRequestHoldTime` key is 120 (decimal value) and, if not, change the value to 120.

4. In the registry editor, locate the following key:

```
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Class\{4D36E97B-E325-11CE-BFC1-08002BE10318}\<bus ID>\Parameters\MaxPendingRequests
```

Confirm that the value for the MaxPendingRequests key is 2048 (decimal value) and, if not, change the value to 2048.

5. In the registry editor, locate the following key:

```
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\Disk\TimeOutValue
```

Confirm that the value for the TimeOutValue key is 60 (decimal value) and, if not, change the value to 60.

6. Restart your host for these changes to take effect.

Multipath support for iSCSI on Windows

For multipathing with iSCSI, we must enable MPIO. For more information about enabling MPIO, see “Installing the multipathing software” on page 221.

Important: IBM Subsystem Device Driver DSM (SDDDSM) is *not* supported for iSCSI attachment. Do not follow the steps to install SDDDSM that you follow to install FC or SAS.

These basic steps are to prepare a Windows 2008 R2 or Windows 2012 R2 host for iSCSI attachment. For more information about configuring the IBM Storwize V5000 for iSCSI connections, see 5.5.3, “Creating iSCSI hosts” on page 270.

5.3.3 Windows 2012 R2: Preparing for SAS attachment

This procedure is described in the following sections.

Installing and updating supported HBAs

Install a supported SAS HBA with the latest supported firmware and drivers for your configuration. A list of the latest supported HBAs and levels for Windows 2008 R2, Windows 2012 R2, and other operating systems is available at this IBM System Storage Interoperation Center (SSIC) [web page](#).

Install the driver by using Windows Device Manager or vendor tools. Also, check and update the firmware level of the HBA by using the manufacturer’s provided tools. Always check the readme file to see whether any Windows registry parameters must be set for the HBA driver.

Determining host WWPNS

The WWPNS of the SAS HBA are required to configure host attachment on the IBM Storwize V5000.

You can obtain the host WWPNS by using vendor tools or the HBA BIOS. However, the easiest way is to connect the SAS cables to the ports on the IBM Storwize V5000, log on to the Storwize CLI through Secure Shell (SSH), and run `svcinfo lsssasportcandidate`, as shown in Example 5-5.

Example 5-5 Finding host WWPNS

```
IBM_Storwize:ITS0_V5000:superuser>svcinfo lsssasportcandidate
sas_WWPN
```

Configuring SAS HBAs on Windows

We suggest that the following settings are used:

- ▶ I/O Timeout for Block Devices: 10
- ▶ I/O Timeout for Sequential Devices: 10
- ▶ I/O Timeout for Other Devices: 10
- ▶ LUNs to Scan for Block Devices: All
- ▶ LUNs to Scan for Sequential Devices: All
- ▶ LUNs to Scan for Other Devices: All

Multipath support for SAS on Windows

For multipathing with SAS, we must enable MPIO and install IBM Subsystem Device Driver DSM (SDDDSM). For more information, see “Installing the multipathing software” on page 221.

We described the basic steps to prepare a Windows 2008 R2 and 2012 R2 host for SAS attachment. For more information about configuring SAS attachment on the IBM Storwize V5000 side, see 5.5.5, “Creating SAS hosts” on page 278.

5.3.4 VMware ESXi: Preparing for Fibre Channel attachment

Complete the following steps to prepare a VMware ESXi host to connect to an IBM Storwize V5000 by using Fibre Channel:

1. Install the HBA or HBAs on the ESXi server.
2. Ensure that the current firmware levels are applied on your host system.
3. Update and configure the HBA for hosts that are running ESXi.
4. Connect the FC host adapter ports to the switches.
5. Configure the switches (zoning).
6. Install the VMware ESXi Hypervisor and load more drivers, if required.

Downloading and installing the supported firmware

Install the current firmware levels to your host server. For the HBAs, check the following [SSIC web page](#).

Download the current supported HBA firmware for your configuration and apply it to your system. Certain HBAs and especially the new converged network adapters (CNAs) require another driver to be loaded into ESXi. Check the VMware Compatibility Guide at [this website](#) to see whether any requirements exist for your configuration.

Configuring QLogic HBAs for VMware ESXi

This section applies to ESXi hosts with installed QLogic HBAs. After you install the firmware, you must configure the HBAs. To perform this task, use the QCC software or HBA BIOS, load the adapter defaults, and set the following values:

- ▶ Host adapter settings:
 - Host Adapter BIOS: Disabled (unless the host is configured for SAN Boot)
 - Frame size: 2048
 - Loop Reset Delay: 5 (minimum)
 - Adapter Hard Loop ID: Disabled
 - Hard Loop ID: 0

- Spinup Delay: Disabled
 - Connection Options 1: Point to point only
 - Fibre Channel Tape Support: Disabled
 - Data Rate: 2
- ▶ Advanced adapter settings:
- Execution throttle: 100
 - LUNs per Target: 0
 - Enable LIP Reset: No
 - Enable LIP Full Login: Yes
 - Enable Target Reset: Yes
 - Login Retry Count: 8
 - Link Down Timeout: 10
 - Command Timeout: 20
 - Extended event logging: Disabled (enable it for debugging only)
 - RIO Operation Mode: 0
 - Interrupt Delay Timer: 0

The QCC management software delivers a unified web-based single-pane-of-glass management console across the QLogic family of storage and networking adapters. A graphical user interface (GUI) or command-line interface (CLI) is available. A VMware vCenter plug-in is also available. The QCC for Windows is available at [this web page](#).

Configuring Emulex HBAs for VMware ESXi

This section applies to ESXi hosts with installed Emulex HBAs. After you install the firmware, load the default settings of all of your adapters that are installed on the host system, and ensure that the Adapter BIOS is disabled, unless you use SAN Boot.

VMware ESXi installation

To install VMware ESXi, complete the following steps:

1. Install your VMware ESXi server and load any other drivers and patches, if required. If you are not familiar with the procedure, see the installation guide that is available at [this web page](#).
2. After you complete your ESXi installation, connect to your ESXi server by using the vSphere web client and go to the Configuration tab.

- Click **Storage Adapters**, and scroll down to your FC HBAs (see Figure 5-17). Document the WWPNs of the installed adapters for later use.

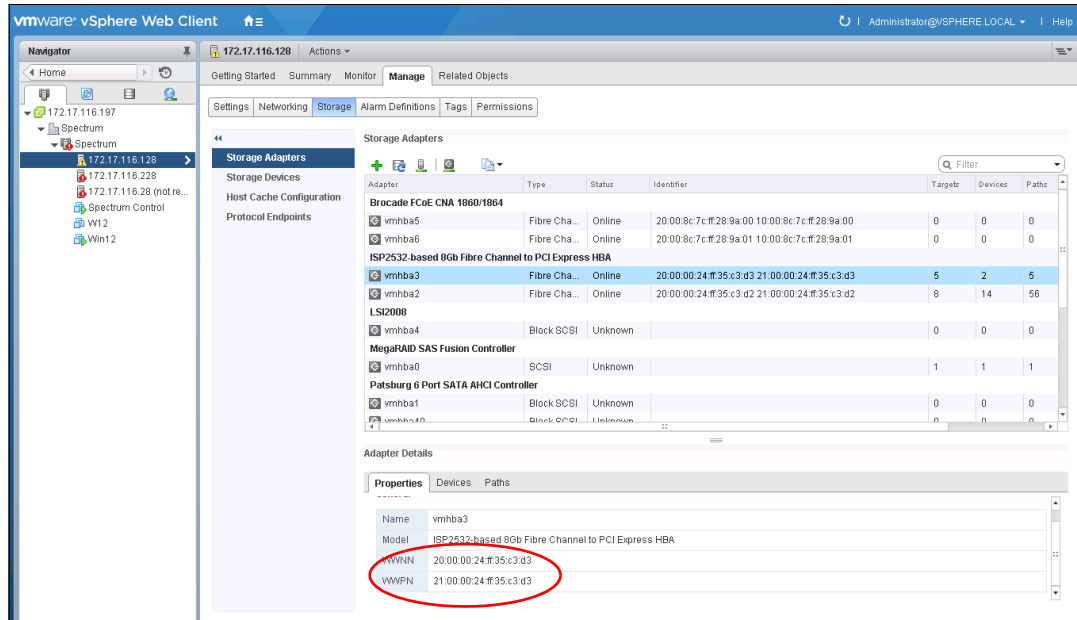


Figure 5-17 Show WWPNS in VMware ESXi

VMware ESXi multipathing

The ESXi server has its own multipathing software. You do not need to install a multipathing driver on the ESXi server or the guest operating systems. The ESXi multipathing policy supports the following operating modes:

- ▶ Round Robin
- ▶ Fixed
- ▶ Most Recently Used (MRU)

The IBM Storwize V5000 is an active/active storage device. The suggested multipathing policy is *Round Robin*. Round Robin performs static load balancing for I/O. If you do not want the I/O balanced over all available paths, the *Fixed* policy is supported also. This policy setting can be selected for every volume.

Set this policy after you attach the Storwize V5000 LUNs to the ESXi host. For more information, see Chapter 6, “Volume configuration” on page 309. If you use an older version of VMware ESX (up to version 3.5), *Fixed* is the suggested policy setting.

MRU selects the first working path, which is discovered at system start time. If this path becomes unavailable, the ESXi/ESX host switches to an alternative path and continues to use the new path while it is available. This policy is the default policy for LUNs that are presented from an Active/Passive array. ESXi/ESX does not return to the previous path if, or when, the previous path returns. It remains on the working path until it fails for any reason.

Determining host WWPNS

The WWPNS of the FC HBA are required to correctly zone switches and configure host attachment on the IBM Storwize V5000. On VMware ESXi, you can obtain these WWPNS through the VMware vSphere Client.

Note: Beginning with VMware ESXi version 5.5, certain new features can be accessed through the vSphere Web Client only. However, we do not demonstrate any of these features. All of the following examples continue to focus on the use of the desktop client.

Connect to the ESXi server (or VMware vCenter) by using the VMware vSphere Client and browse to the Configuration tab. Click **Storage Adapters** to see the HBA WWPNs, as shown in Figure 5-18.

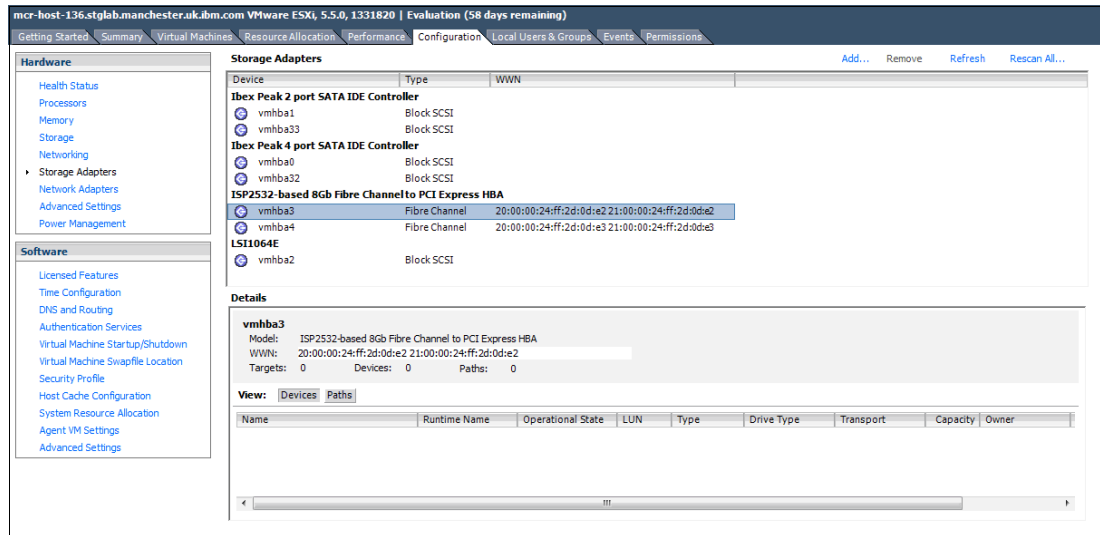


Figure 5-18 FC WWPNs in VMware vSphere Client

After all of these steps are completed, the ESXi host is prepared to connect to the Storwize V5000. For more information about creating the ESXi FC host in the Storwize V5000 GUI, see 5.5.1, “Creating FC hosts” on page 258.

5.3.5 VMware ESXi: Preparing for iSCSI attachment

This section describes how to enable iSCSI on VMware ESXi hosts. We focus on vSphere because the complete iSCSI stack was rewritten in this level. This level offers improved performance and supports useful features, such as jumbo frames and Transmission Control Protocol (TCP) Segmentation Offload. We focus on the basic ESXi iSCSI setup.

For more information, see the VMware vSphere Documentation Center, which is available at [this web page](#).

For more information, see [VMware Compatibility Guide](#).

Important: For converged network adapters (CNAs) that support both FC and iSCSI, it is important to ensure that the Ethernet networking driver is installed in addition to the FCoE driver. The Ethernet networking driver and the FCoE driver are required for the configuration of iSCSI.

For more information about the hardware and host operating configuration, see the manufacturer's documentation and the Storwize V5000 IBM Knowledge Center if you use a hardware iSCSI HBA. The following section describes how to configure iSCSI by using the software initiator.

Complete the following steps to prepare a VMware ESXi host to connect to a Storwize V5000 by using iSCSI:

1. Ensure that the current firmware levels are applied on your host system.
2. Install VMware ESXi and load more drivers, if required.
3. Connect the ESXi server to your network. You must use separate network interfaces for iSCSI traffic.
4. Configure your network to fulfill your security and performance requirements.

The iSCSI initiator is installed by default on your ESXi server, but you must enable it. To enable it, complete the following steps:

1. Connect to your ESXi server by using the vSphere Client. Go to **Manage**, and select **Networking** (see Figure 5-19).

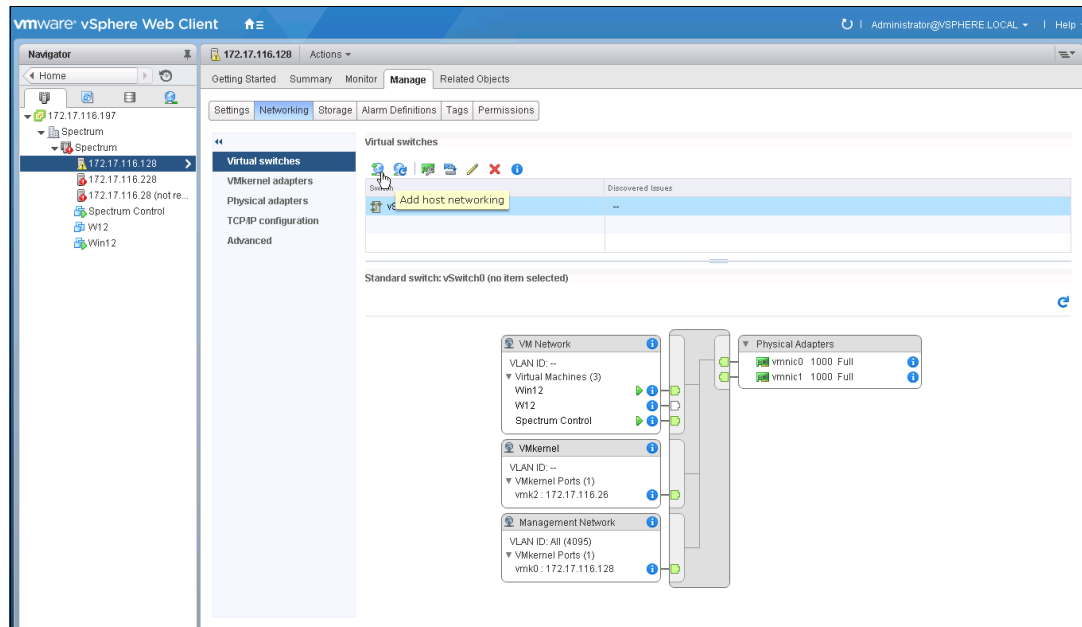


Figure 5-19 Select VMware networking

2. Click **Add Networking** to start the Add Networking wizard (see Figure 5-20). Select **VMkernel Network Adapter** and then, click **Next**.

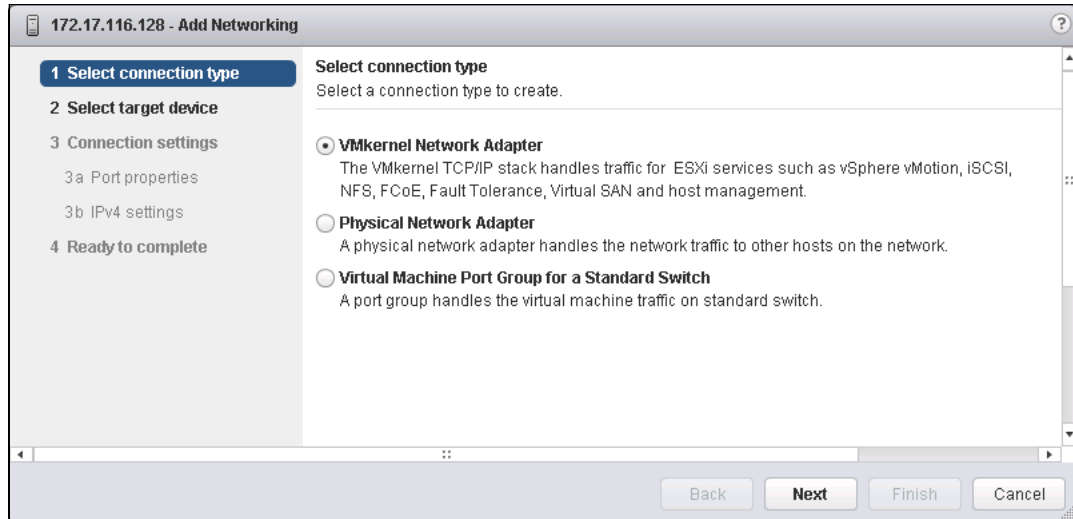


Figure 5-20 VMware: Add Networking wizard

3. Click **Select target device**, as shown in Figure 5-21.

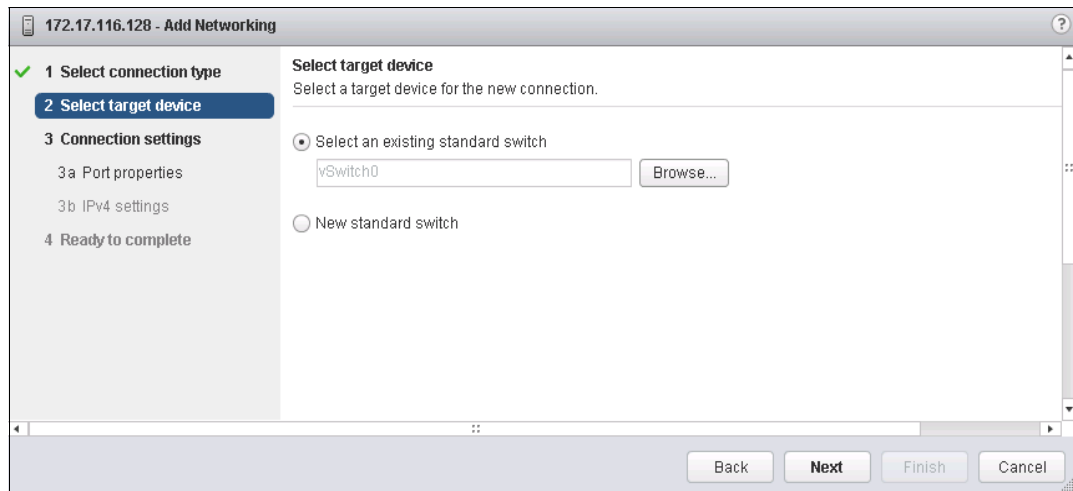


Figure 5-21 Select target device

4. Select one or more network interfaces that you want to use for iSCSI traffic and click **Next** (see Figure 5-22).

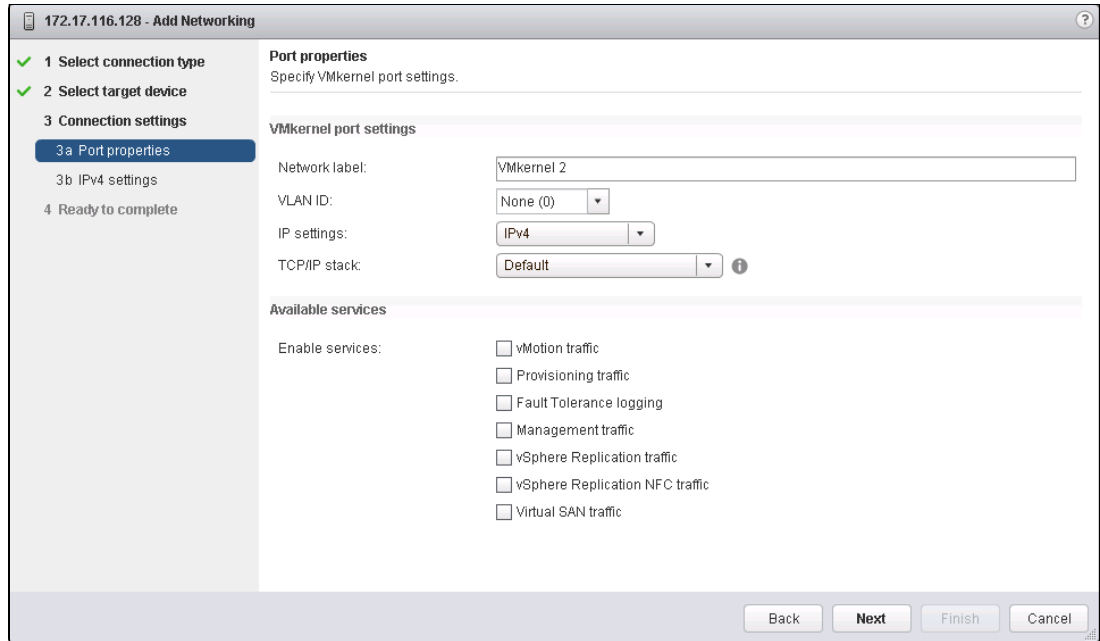


Figure 5-22 Select an iSCSI interface

5. Enter a meaningful network label and click **Next** (see Figure 5-23).

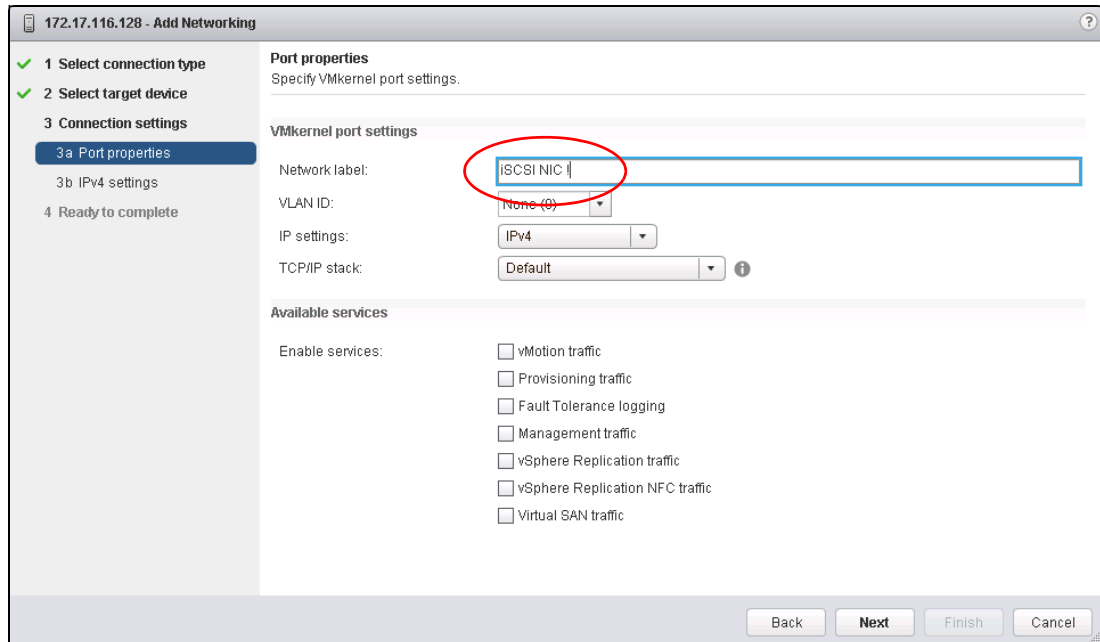


Figure 5-23 Enter a network label

6. Enter an IP address for your iSCSI network. Use a dedicated network for iSCSI traffic (see Figure 5-24).

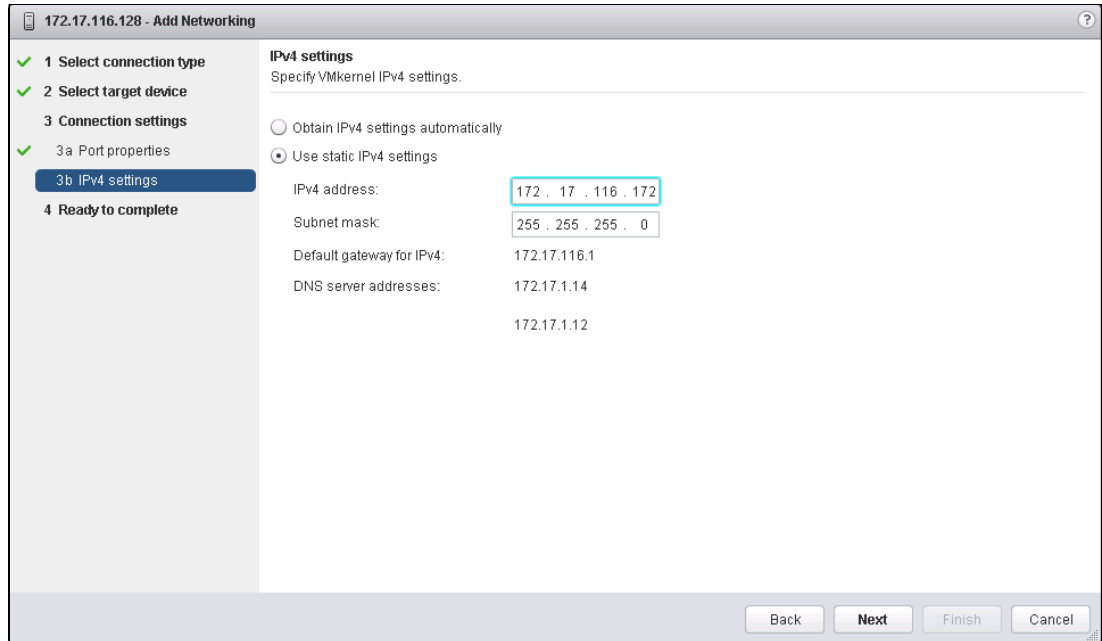


Figure 5-24 Enter an iSCSI network IP

7. Click **Next**, as shown in Figure 5-25.

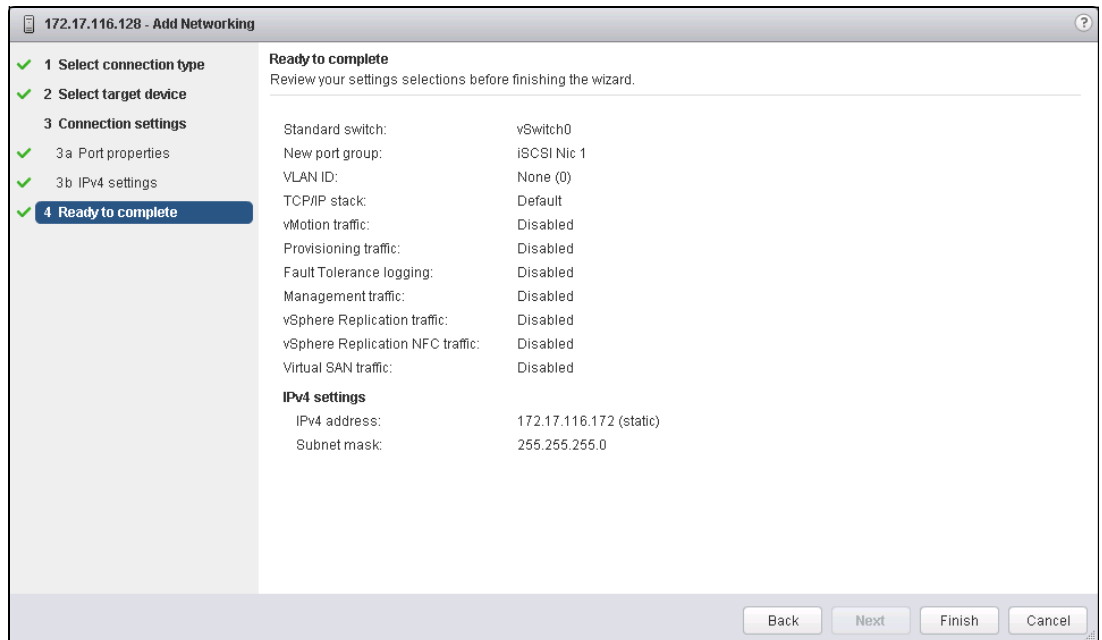


Figure 5-25 Ready to complete

8. Click **Finish** to complete the setup.

9. Check whether an iSCSI software adapter is available. Select **Storage Adapters** on the Manage tab (see Figure 5-26).

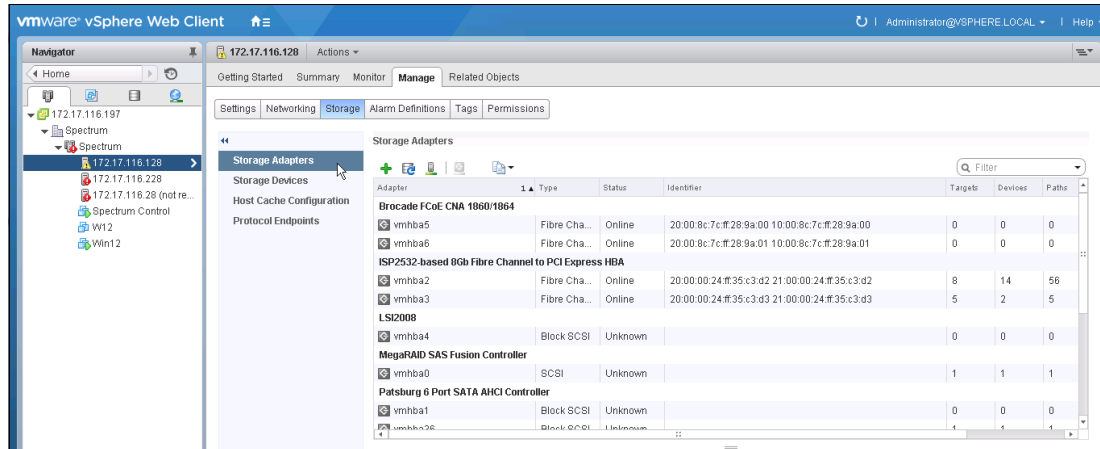


Figure 5-26 Select new iSCSI software adapter

10. Click the plus sign (+) to add an iSCSI software adapter (see Figure 5-27).

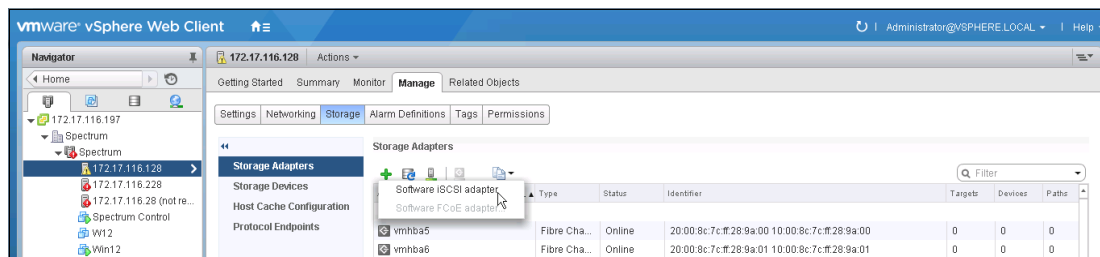


Figure 5-27 Add an iSCSI software adapter

The Add Software iSCSI Adapter window opens (see Figure 5-28).

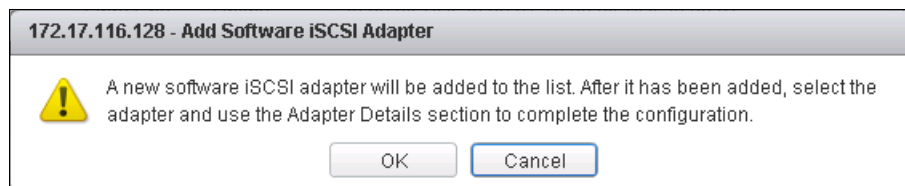


Figure 5-28 Add Software iSCSI Adapter window

11. Click **OK**. A message displays that prompts you to configure the adapter after it is added.

A new iSCSI adapter is added to the Storage Adapters window (see Figure 5-29).

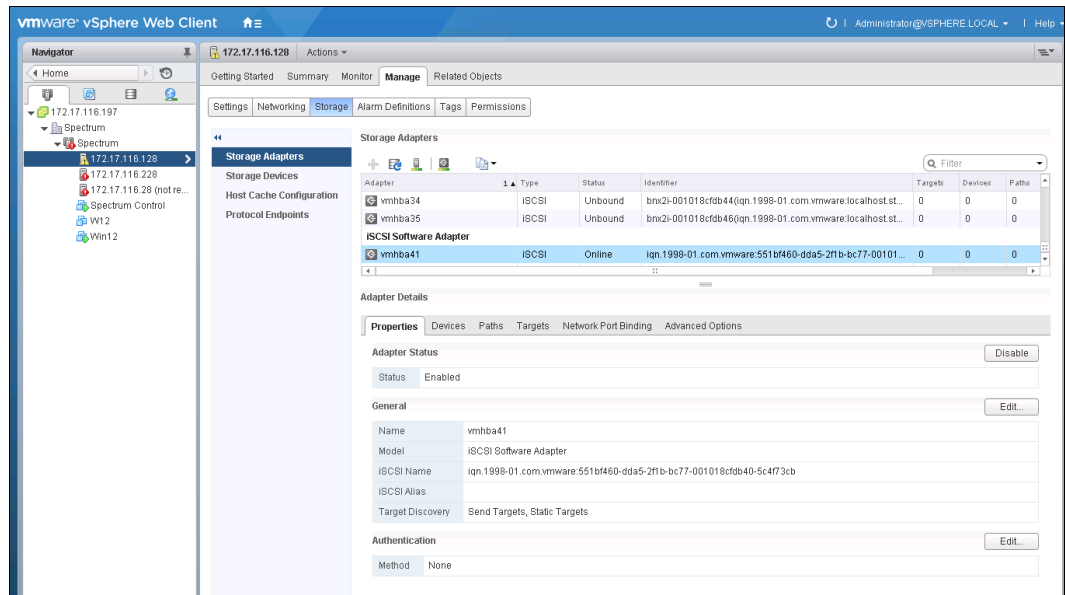


Figure 5-29 New iSCSI Software Adapter

12. Select **Storage Adapters** and scroll to the iSCSI Software Adapter (see Figure 5-30). Highlight it and you see the Adapter Details in the lower part of the window.

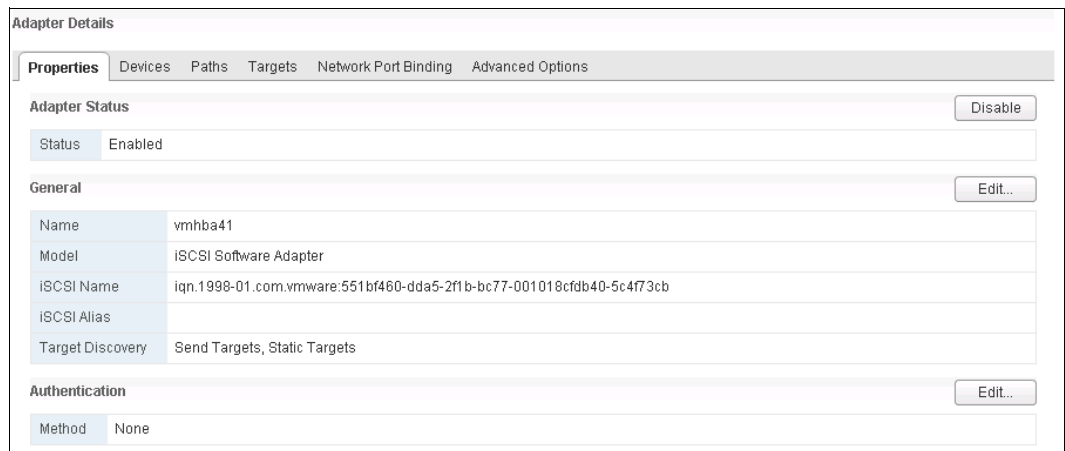


Figure 5-30 iSCSI Software Adapter

13. The iSCSI Software Adapter Properties window opens. Figure 5-31 shows that the initiator is enabled by default. To change this setting, click **Disable**.



Figure 5-31 iSCSI Software Adapter properties

14. The VMware ESX iSCSI initiator is successfully enabled (see Figure 5-32). Document the initiator name for later use.

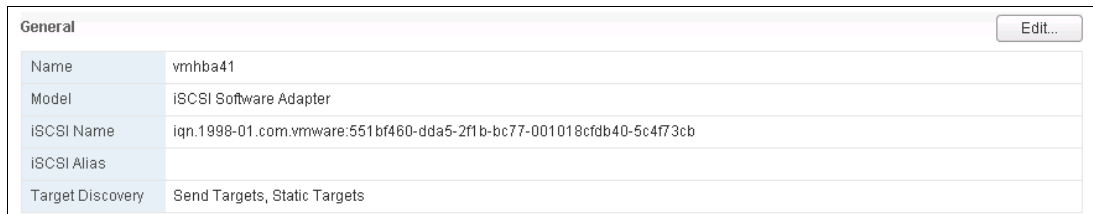


Figure 5-32 Enabled VMware iSCSI initiator

Multipath support for iSCSI on VMware ESXi

As explained in 5.3.4, “VMware ESXi: Preparing for Fibre Channel attachment” on page 233, the ESXi server uses its own multipathing software.

For iSCSI, extra configuration is required in the VMkernel port properties to enable path failover. Each VMkernel port must map to one physical adapter port, which is not the default setting. Complete the following steps:

1. Browse to the **Configuration** tab and select **Networking**. Click **Properties** next to the vSwitch that you configured for iSCSI to open the window that is shown in Figure 5-33.

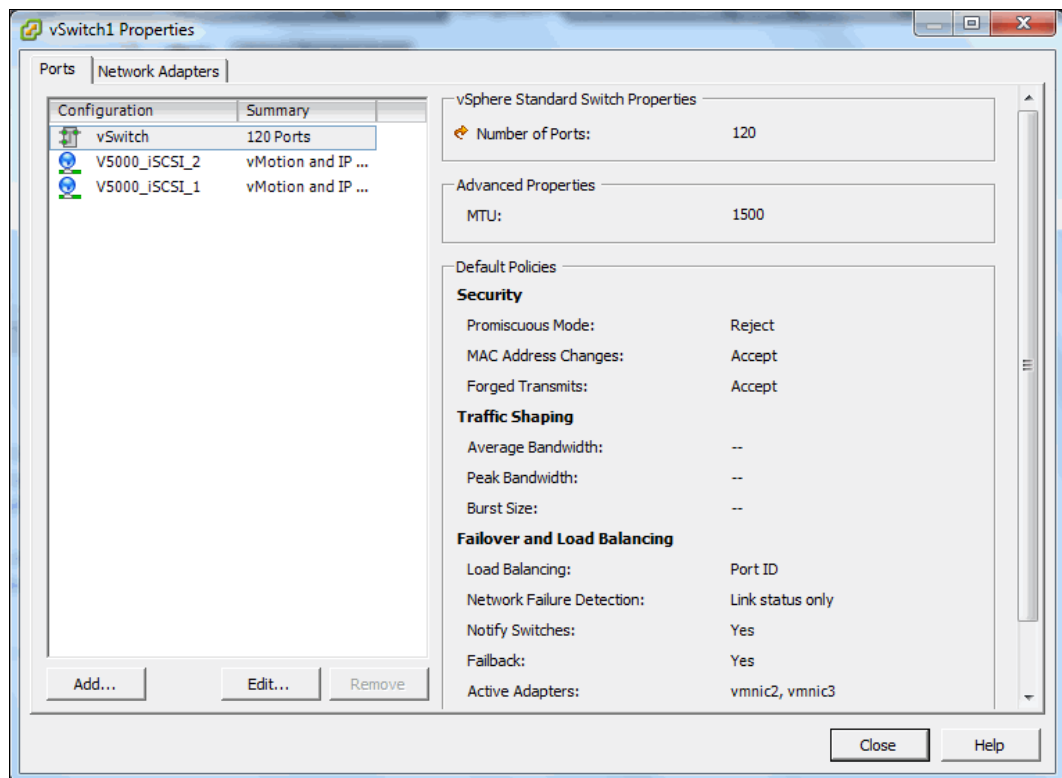


Figure 5-33 View the vSwitch properties with listed VMkernel ports

2. Select one of the VMkernel ports and click **Edit**. The window that is shown in Figure 5-34 opens.

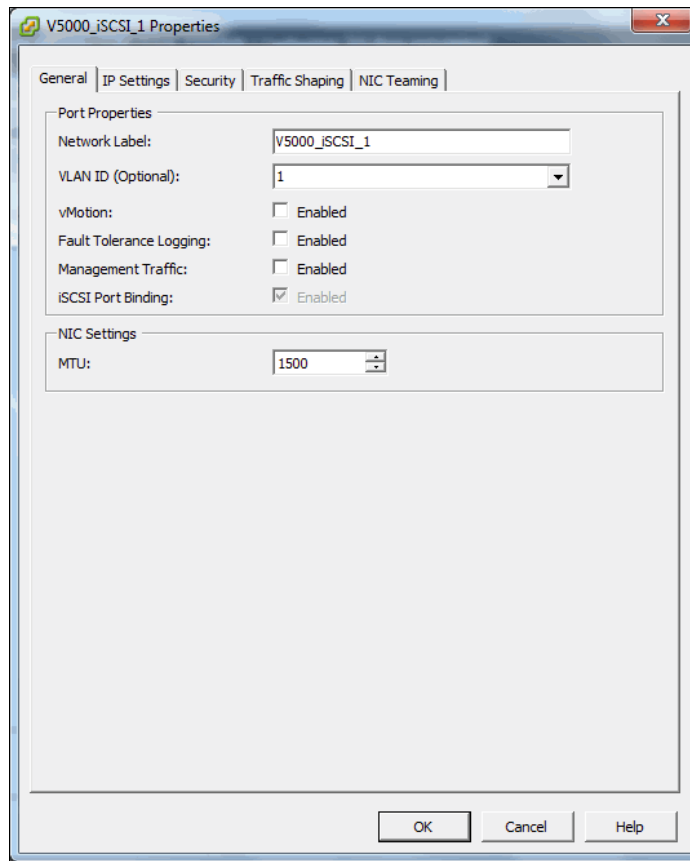


Figure 5-34 Editing a VMkernel port

3. Click the **NIC Teaming** tab. Select **Override switch failover order** and ensure that each port is tied to one physical adapter port, as shown in Figure 5-35.

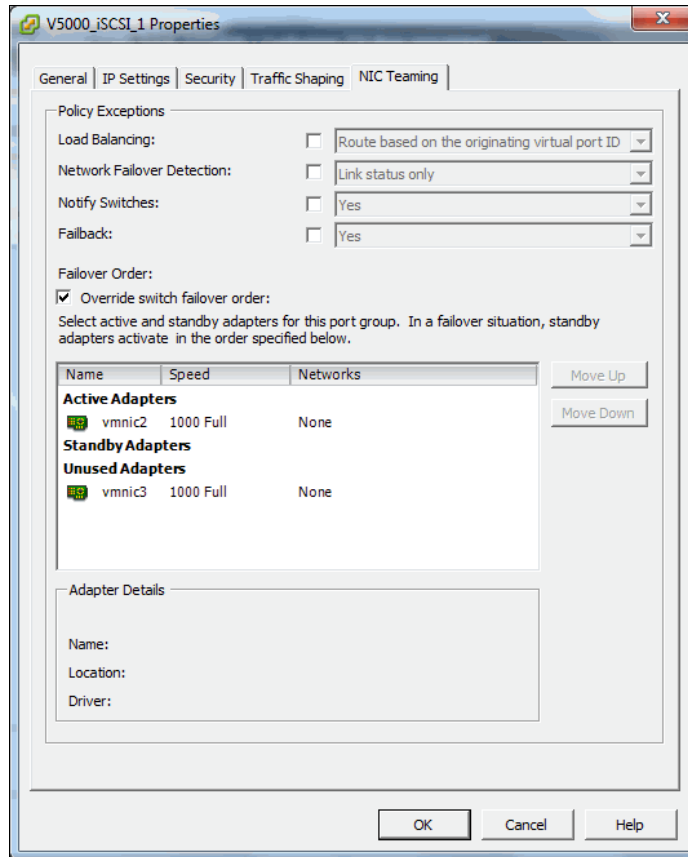


Figure 5-35 Configuring a VMkernel port to bind to a single physical adapter port

These basic steps are required to prepare a VMware ESXi host for iSCSI attachment. For information about configuring iSCSI attachment on the IBM Storwize V5000, see 5.5.3, “Creating iSCSI hosts” on page 270.

For more information about configuring iSCSI attachment on the VMware ESXi side, [this white paper](#) that was published by VMware is a useful resource.

5.3.6 VMware ESXi: Preparing for SAS attachment

This procedure is described in the following sections.

Installing and updating supported HBAs

Install a supported HBA with the latest supported firmware and drivers for your configuration. A list of the latest supported HBAs and levels for VMware ESXi is available at [this website](#).

Install the driver by using VMware vSphere Client, the ESXi CLI, or vendor tools. Also, check and update the firmware level of the HBA by using the manufacturer’s provided tools. Always check the readme file to see whether more configuration is required for the HBA driver.

For more information, see the [VMware Compatibility Guide web page](#).

Configuring SAS HBAs on VMware ESXi

In this example, we used an LSI 9207-8e card and did not need to configure HBA parameters beyond the default settings. We advise that you check the parameters through the HBA BIOS or vendor tools to confirm that they are suitable for your requirements.

Multipath support for SAS on VMware ESXi

As with FC, we can use native ESXi multipathing for SAS attachment on VMware ESXi 5.5. For more information, see 5.3.4, “VMware ESXi: Preparing for Fibre Channel attachment” on page 233.

Determining host WWPNS

The worldwide port names (WWPNs) of the SAS HBA are required to configure host attachment on the IBM Storwize V5000.

The host WWPNs are not directly available through VMware vSphere Client. However, you can obtain them by using vendor tools or the HBA BIOS. The method that is described in 5.3.3, “Windows 2012 R2: Preparing for SAS attachment” on page 232 also works.

These basic steps are required to prepare a VMware ESXi host for SAS attachment. For information about configuring SAS attachment on the IBM Storwize V5000 side, see 5.5.5, “Creating SAS hosts” on page 278.

For more information and guidance about attaching storage with VMware ESXi, [this document](#) that was published by VMware is a useful resource.

5.4 N-Port ID Virtualization support

The usage model for all IBM Spectrum Virtualize products is based on two-way active/active node models. That is, a pair of distinct control modules that share active/active access for a specific volume. These nodes each have their own Fibre Channel WWNN, and thus all ports that are presented from each node have a set of WWPNs that are presented to the fabric.

Traditionally, if one node fails or is removed for any reason, the paths that are presented for volumes from that node go offline. It is up to the native OS multipathing software to failover from using both sets of WWPN to just those that remain online. While this process is exactly what multipathing software is designed to do, occasionally it can be problematic, particularly if paths are not seen as coming back online for some reason.

Starting with IBM Spectrum Virtualize V7.7.0, the IBM Spectrum Virtualize system can be enabled into N-Port ID Virtualization (NPIV) mode. When NPIV mode is enabled on the IBM Spectrum Virtualize system, ports do not come up until they are ready to service I/O, which improves host behavior around node unpendes. In addition, path failures that are caused by an offline node are masked from host multipathing.

When NPIV is enabled on IBM Spectrum Virtualize system nodes, each physical WWPN reports up to three virtual WWPNs, as listed in Table 5-1.

Table 5-1 Spectrum Virtualize NPIV ports

NPIV port	Description
Primary NPIV Port	This is the WWPN that communicates with backend storage, and might be used for node to node traffic (local or remote).

NPIV port	Description
Primary Host Attach Port	This is the WWPN that communicates with hosts. It is a target port only, and this is the primary port, so it represents this local node's WWNN.
Failover Host Attach Port	This is a standby WWPN that communicates with hosts and is only brought online on this node if the partner node in this I/O Group goes offline. This is the same as the Primary Host Attach WWPN on the partner node.

Figure 5-36 shows the three WWPNs that are associated with a SAN Volume Controller port when NPIV is enabled.

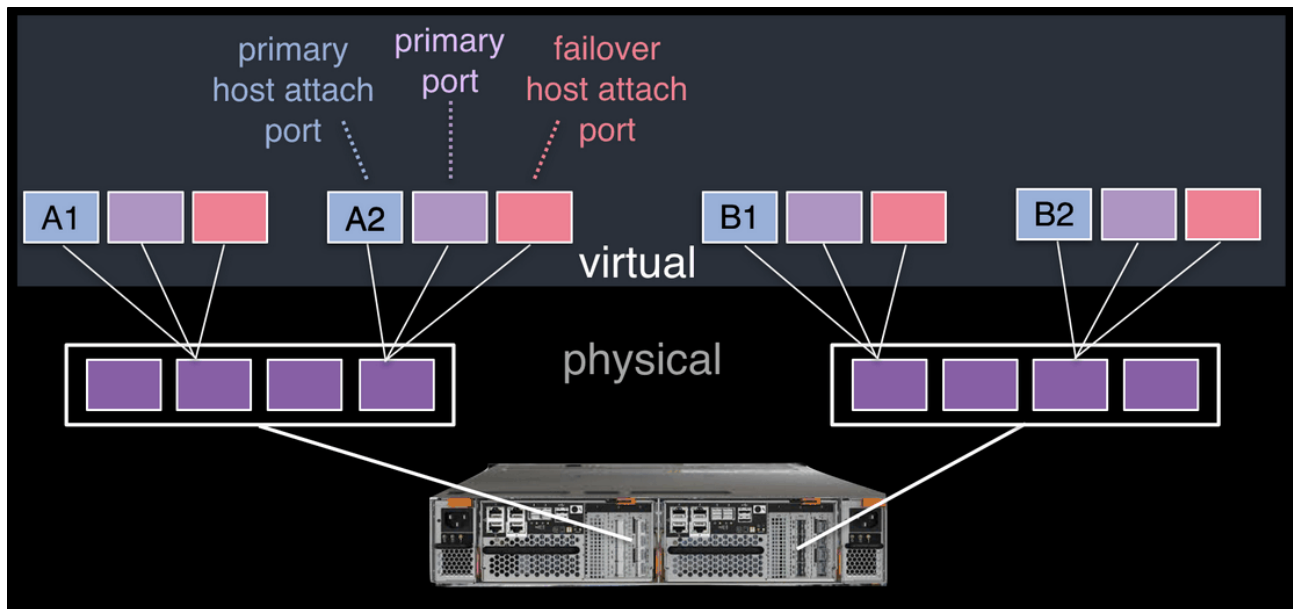


Figure 5-36 Allocation of NPIV virtual WWPN ports per physical port

The failover host attach port (shown in pink in Figure 5-36) is not active at this time. Figure 5-37 on page 248 shows what occurs when the second node fails. Subsequent to the node failure, the failover host attach ports on the remaining node are active and take on the WWPN of the failed node's primary host attach port.

Note: Figure 5-37 on page 248 shows only two ports per node in detail, but the same applies for all physical ports.

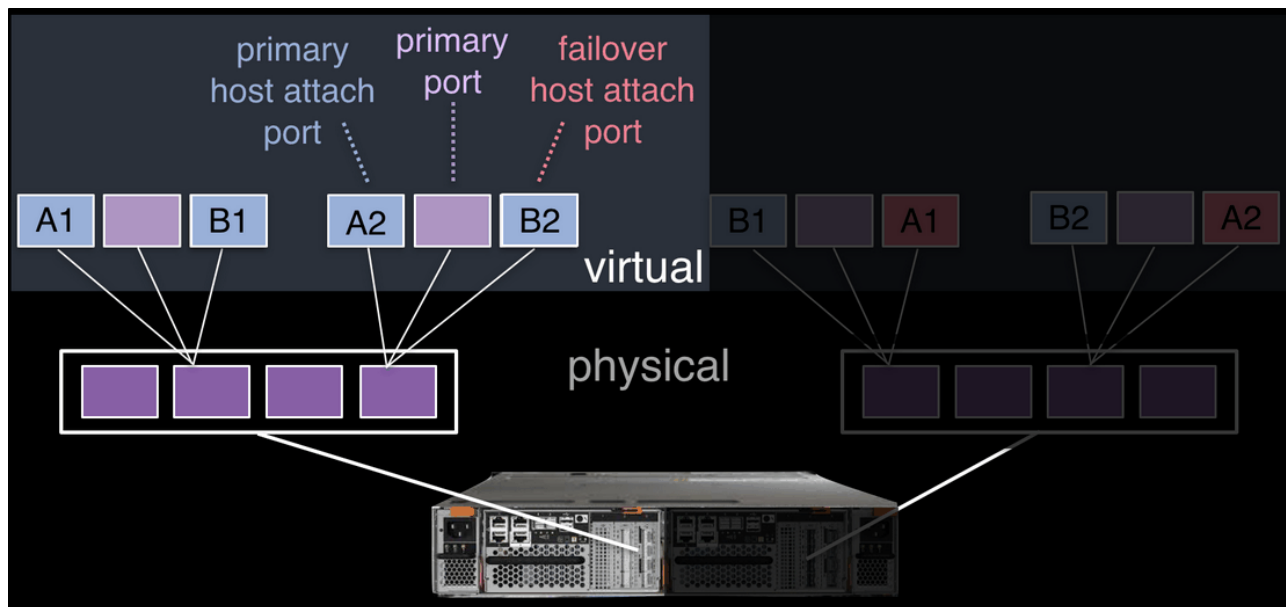


Figure 5-37 Allocation of NPIV virtual WWPN ports per physical port after a node failure

With V7.7.0 onwards, this process occurs automatically when NPIV is enabled at a system level in Spectrum Virtualize. At this time, the failover happens only automatically between the two nodes in an I/O group.

There is a transitional mode for compatibility with an earlier version during the transition period as well.

NPIV is enabled by default in a new IBM Spectrum Virtualize V8.2.1.0.

Note: NPIV is supported for FC protocol only. It is *not* supported for FCoE protocol or iSCSI.

5.4.1 NPIV prerequisites

Consider the following points regarding NPIV enablement:

- ▶ For NPIV enablement, the IBM Spectrum Virtualize system must be at V7.7.0 or later.
- ▶ A V7.7.0 or later system with NPIV enabled as backend storage for a system that is earlier than version 7.7.0 is not supported.
- ▶ Both nodes in an IO group should have identical hardware to allow failover to work as expected.
- ▶ FC switches must permit each physically connected IBM Spectrum Virtualize system port the ability to create two extra NPIV ports.
- ▶ Check this configuration and restriction [web page](#) to ensure that your operating environment is supported for NPIV.

5.4.2 Enabling NPIV on a new system

For V7.7.0 and later, an IBM Spectrum Virtualize system should have NPIV enabled by default. For any case where it is not enabled by default and NPIV is wanted, NPIV can be enabled on a new IBM Spectrum Virtualize system by completing the following steps:

1. Run the `lsiogrp` command to list the I/O groups present in a system, as shown in Example 5-6.

Example 5-6 Listing the I/O groups in a system

```
IBM_Storwize:ITSOV5030:superuser>lsiogrp
id name          node_count vdisk_count host_count site_id site_name
0  io_grp0        2          16          2          0
1  io_grp1        0          0           2          0
2  io_grp2        0          0           0          0
3  io_grp3        0          0           0          0
4  recovery_io_grp 0          0           0          0
IBM_Storwize:ITSOV5030:superuser>
```

2. Run the `lsiogrp` command to view the status of NPIV, as shown in Example 5-7.

Example 5-7 Checking NPIV mode with the `fctargetportmode` field

```
IBM_Storwize:ITSOV5030:superuser>lsiogrp 0
id 0
name io_grp0
node_count 2
vdisk_count 16
host_count 2
flash_copy_total_memory 20.0MB
flash_copy_free_memory 20.0MB
remote_copy_total_memory 20.0MB
remote_copy_free_memory 20.0MB
mirroring_total_memory 20.0MB
mirroring_free_memory 19.9MB
raid_total_memory 80.0MB
raid_free_memory 75.2MB
maintenance no
compression_active no
accessible_vdisk_count 16
compression_supported yes
max_enclosures 21
encryption_supported yes
flash_copy_maximum_memory 2048.0MB
site_id
site_name
fctargetportmode enabled
compression_total_memory 0.0MB
deduplication_supported yes
deduplication_active no
nqn
IBM_Storwize:ITSOV5030:superuser>
```

NPIV is enabled if the resulting output is `fctargetportmode enabled`, as shown in Example 5-7.

- List the virtual WWPNs by using the `lstorageportfc` command, as shown in Example 5-8.

Example 5-8 Listing the virtual WWPNs

```

IBM_Storwize:ITSOV5030:superuser>lstorageportfc
id WWPN           WNNN           port_id owning_node_id current_node_id nportid host_io_permitted
virtualized protocol
1 500507680D0496E7 500507680D0096E7 1 1 1 011200 no no
scsi
2 500507680D7496E7 500507680D0096E7 1 1 1 011201 yes yes
scsi
4 500507680D0896E7 500507680D0096E7 2 1 1 011200 no no
scsi
5 500507680D7896E7 500507680D0096E7 2 1 1 011201 yes yes
scsi
(content removed for brevity)
49 500507680D0496E6 500507680D0096E6 1 2 2 011300 no no
scsi
50 500507680D7496E6 500507680D0096E6 1 2 2 011301 yes yes
scsi
52 500507680D0896E6 500507680D0096E6 2 2 2 011300 no no
scsi
53 500507680D7896E6 500507680D0096E6 2 2 2 011301 yes yes
scsi
(content removed for brevity)
IBM_Storwize:ITSOV5030:superuser>

```

At this point, you can configure zones for hosts by using the primary host attach ports (virtual WWPNs) of the IBM Spectrum Virtualize ports, as shown in **bold** in the output of Example 5-8.

- If the status of `fctargetportmode` is disabled, run the `chiogrp` command to get into transitional mode for NPIV, as shown in Example 5-9.

Example 5-9 Change the NPIV mode to transitional

```

IBM_Storwize:ITSOV5030:superuser>chiogrp -fctargetportmode transitional 0

```

- Verify the transitional mode by using the `lsiogrp` command, as shown in Example 5-10.

Example 5-10 NPIV transitional mode

```

IBM_Storwize:ITSOV5030:superuser>lsiogrp 0
id 0
name io_grp0
node_count 2
vdisk_count 16
host_count 2
flash_copy_total_memory 20.0MB
flash_copy_free_memory 20.0MB
remote_copy_total_memory 20.0MB
remote_copy_free_memory 20.0MB
mirroring_total_memory 20.0MB
mirroring_free_memory 19.9MB
raid_total_memory 80.0MB
raid_free_memory 75.2MB
maintenance no
compression_active no
accessible_vdisk_count 16

```

```

compression_supported yes
max_enclosures 21
encryption_supported yes
flash_copy_maximum_memory 2048.0MB
site_id
site_name
fctargetportmode transitional
compression_total_memory 0.0MB
deduplication_supported yes
deduplication_active no
nqn
IBM_Storwize:ITSOV5030:superuser>

```

In transitional mode, host I/O is permitted on primary ports and primary host attach ports (virtual WWPN), as shown in Example 5-11 under the host_io_permitted column.

Example 5-11 WWPNs in transitional mode

```

IBM_Storwize:ITSOV5030:superuser>ls targetportfc
id WWPN          WNNN          port_id owning_node_id current_node_id nportid host_io_permitted
virtualized protocol
1 500507680D0496E7 500507680D0096E7 1 1 1 011200 yes no
scsi
2 500507680D7496E7 500507680D0096E7 1 1 1 011201 yes yes
scsi
4 500507680D0896E7 500507680D0096E7 2 1 1 011200 yes no
scsi
5 500507680D7896E7 500507680D0096E7 2 1 1 011201 yes yes
scsi
(content removed for brevity)
49 500507680D0496E6 500507680D0096E6 1 2 2 011300 yes no
scsi
50 500507680D7496E6 500507680D0096E6 1 2 2 011301 yes yes
scsi
52 500507680D0896E6 500507680D0096E6 2 2 2 011300 yes no
scsi
53 500507680D7896E6 500507680D0096E6 2 2 2 011301 yes yes
scsi
(content removed for brevity)
IBM_Storwize:ITSOV5030:superuser>

```

6. Enable NPIV by changing the mode from transitional to enabled, as shown in Example 5-12.

Example 5-12 Enabling NPIV

```

IBM_Storwize:ITSOV5030:superuser>chiogrp -fctargetportmode enabled 0

```

NPIV enablement can be verified by checking the fctargetportmode field, as shown in Example 5-13.

Example 5-13 NPIV enablement verification

```

IBM_Storwize:ITSOV5030:superuser>lsiogrp 0
id 0
name io_grp0
node_count 2
vdisk_count 16

```

```

host_count 2
flash_copy_total_memory 20.0MB
flash_copy_free_memory 20.0MB
remote_copy_total_memory 20.0MB
remote_copy_free_memory 20.0MB
mirroring_total_memory 20.0MB
mirroring_free_memory 19.9MB
raid_total_memory 80.0MB
raid_free_memory 75.2MB
maintenance no
compression_active no
accessible_vdisk_count 16
compression_supported yes
max_enclosures 21
encryption_supported yes
flash_copy_maximum_memory 2048.0MB
site_id
site_name
fctargetportmode enabled
compression_total_memory 0.0MB
deduplication_supported yes
deduplication_active no
nqn
IBM_Storwize:ITS0V5030:superuser>

```

At this point, you can configure zones for hosts by using the primary host attach ports (virtual WWPNs) of the IBM Spectrum Virtualize ports, as shown in **bold** in the output of Example 5-8 on page 250.

5.4.3 Enabling NPIV on an existing system

When IBM Spectrum Virtualize systems that are running code before 7.7.1 are upgraded to version 7.7.1 or higher, the NPIV feature is not turned on by default because it might require changes to host side zoning. Enabling NPIV on a system requires that you complete the following steps after meeting the prerequisites:

1. Audit your SAN fabric layout and zoning rules because NPIV has stricter requirements. Ensure that equivalent ports are on the same fabric and in the same zone. For more information, see the topic about zoning considerations for N_PortID Virtualization in [IBM Knowledge Center](#).
2. Check the path count between your hosts and the IBM Spectrum Virtualize system to ensure that the number of paths is half of the usual supported maximum. For more information, see the topic about zoning considerations for N_Port ID Virtualization in [IBM Knowledge Center](#).
3. Check whether your system has targetportmode set to disabled, as shown in Example 5-14.

Example 5-14 fctargetportmode disabled in iogrp

```

IBM_Storwize:ITS0V5030:superuser>lsiogrp 0
id 0
name io_grp0
node_count 2
vdisk_count 16
host_count 0

```

```

flash_copy_total_memory 20.0MB
flash_copy_free_memory 20.0MB
remote_copy_total_memory 20.0MB
remote_copy_free_memory 20.0MB
mirroring_total_memory 20.0MB
mirroring_free_memory 19.9MB
raid_total_memory 80.0MB
raid_free_memory 75.2MB
maintenance no
compression_active no
accessible_vdisk_count 16
compression_supported yes
max_enclosures 21
encryption_supported yes
flash_copy_maximum_memory 2048.0MB
site_id
site_name
fctargetportmode disabled
compression_total_memory 0.0MB
deduplication_supported yes
deduplication_active no
nqn
IBM_Storwize:ITSOV5030:superuser>

```

4. Run the `lstorageportfc` command to note the primary host attach WWPNs (virtual WWPNs), as shown in **bold** in Example 5-15.

Example 5-15 Using the `lstorageportfc` command to get primary host WWPNs (virtual WWPNs)

```

IBM_Storwize:ITSOV5030:superuser>lstorageportfc
id WWPN                WWNN                port_id owning_node_id current_node_id nportid host_io_permitted
virtualized protocol
1 500507680D0496E7 500507680D0096E7 1      1                1                011200 yes             no
scsi
2 500507680D7496E7 500507680D0096E7 1      1                000000 no              yes
scsi
4 500507680D0896E7 500507680D0096E7 2      1                1                011200 yes             no
scsi
5 500507680D7896E7 500507680D0096E7 2      1                000000 no              yes
scsi
(content removed for brevity)
49 500507680D0496E6 500507680D0096E6 1      2                2                011300 yes             no
scsi
50 500507680D7496E6 500507680D0096E6 1      2                000000 no              yes
scsi
52 500507680D0896E6 500507680D0096E6 2      2                2                011300 yes             no
scsi
53 500507680D7896E6 500507680D0096E6 2      2                000000 no              yes
scsi
(content removed for brevity)
IBM_Storwize:ITSOV5030:superuser>

```

5. Include the primary host attach ports (virtual WWPNs) to your host zones.

6. Enable transitional mode for NPIV on IBM Spectrum Virtualize system (see Example 5-16).

Example 5-16 NPIV in transitional mode

```
IBM_Storwize:ITSOV5030:superuser>chiogrp -fctargetportmode transitional 0
```

7. Ensure that the primary host attach WWPNs (virtual WWPNs) now allows host traffic, as shown in **bold** in Example 5-17.

Example 5-17 Host attach WWPNs (virtual WWPNs) permitting host traffic

```
IBM_Storwize:ITSOV5030:superuser>lsfabric
id WWPN                WWNN                port_id owning_node_id current_node_id nportid host_io_permitted
virtualized protocol
1 500507680D0496E7 500507680D0096E7 1      1                1                011200 yes          no
scsi
2 500507680D7496E7 500507680D0096E7 1      1                1                011201 yes        yes
scsi
4 500507680D0896E7 500507680D0096E7 2      1                1                011200 yes          no
scsi
5 500507680D7896E7 500507680D0096E7 2      1                1                011201 yes        yes
scsi
(content removed for brevity)
49 500507680D0496E6 500507680D0096E6 1      2                2                011300 yes          no
scsi
50 500507680D7496E6 500507680D0096E6 1      2                2                011301 yes        yes
scsi
52 500507680D0896E6 500507680D0096E6 2      2                2                011300 yes          no
scsi
53 500507680D7896E6 500507680D0096E6 2      2                2                011301 yes        yes
scsi
(content removed for brevity)
IBM_Storwize:ITSOV5030:superuser>
```

8. Ensure that the hosts are using the NPIV ports for host I/O.

Remember: You can verify that you are logged in to the hosts by running the `lsfabric -host host_id_or_name` command. If I/O activity is occurring, each host has at least one line in the command output that corresponds to a host port and shows active in the activity field.

Consider the following points:

- ▶ Hosts where no I/O was issued in the past 5 minutes do not show active for any login.
- ▶ Hosts that do not adhere to preferred paths might still be processing I/O to primary ports.

In most of host operating systems, rescanning the SAN might be required on some hosts to recognize more paths that now are provided via primary host attach ports (virtual WWPNs).

9. After a minimum of 15 minutes passes since entering transitional mode, change the system to enabled mode by entering the command that is shown in Example 5-18.

Example 5-18 Enabling the NPIV

```
IBM_Storwize:ITS0V5030:superuser>chiogrp -fctargetportmode enabled 0
```

Now NPIV is enabled on the IBM Spectrum Virtualize system, and hosts should also be using the virtualized WWPNs for I/O. At this point, the host zones can be amended appropriately to use primary host attach port WWPNs (virtual WWPNs) only.

Note: If hosts are still configured to use non-virtual ports on Storwize V5000, the system prevents you from changing from transitional mode to enabled mode, as shown in Example 5-19.

Example 5-19 System prevents setting fctargetportmode as enabled

```
IBM_Storwize:ITS0V5030:superuser>chiogrp -fctargetportmode enabled 0
CMMVC8019E Task could interrupt IO and force flag not set.
IBM_Storwize:ITS0V5030:superuser>
```

5.5 Creating hosts by using the GUI

This section describes how to create FC, iSCSI, and SAS hosts by using the IBM Storwize V5000 GUI. We assume that the hosts are prepared for attachment, as described in 5.3, “Preparing the host operating system” on page 219, and that you know the host WWPNs and their iSCSI initiator names.

Considerations when you configure hosts in the IBM Storwize V5000

When you create a host object in the IBM Storwize V5000, it is important to verify the configuration limits and restrictions, which are available at [this website](#).

Complete the following steps:

1. Open the Hosts configuration window by clicking **Hosts** on the host window, as shown in Figure 5-38.

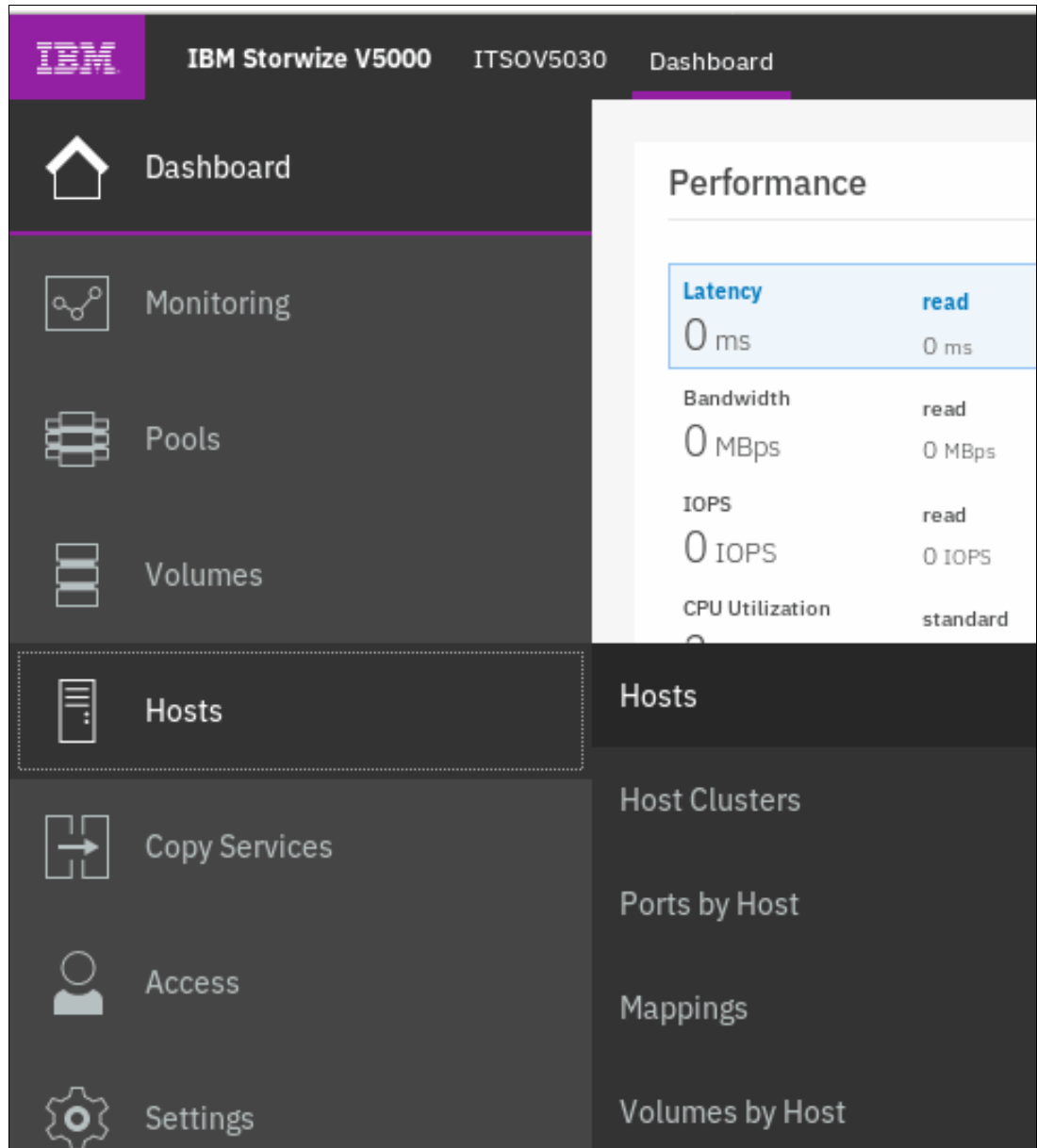


Figure 5-38 Open the Hosts window

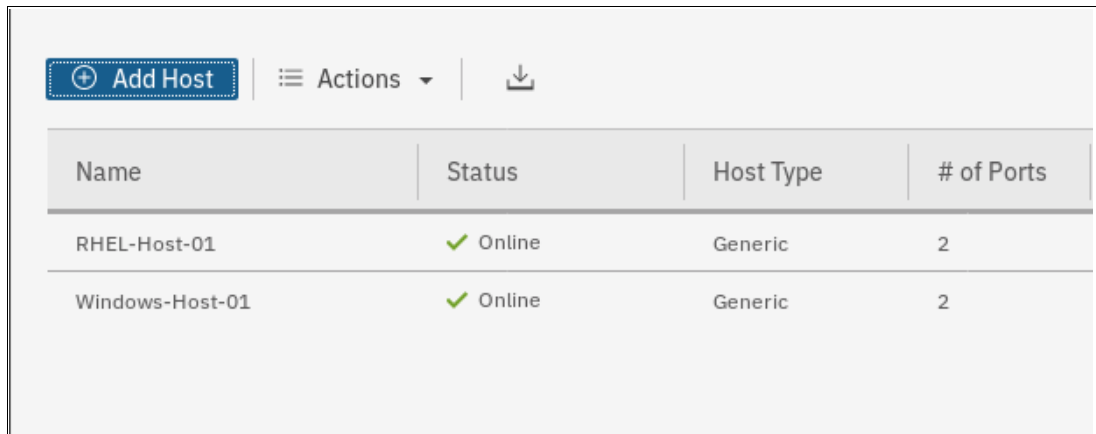
2. A list of the existing hosts is displayed, as shown in Figure 5-39.

The screenshot shows a table with a header row and one data row. Above the table are controls for adding a host, actions, and filters. The table has columns for Name, Status, Host Type, # of Ports, Host Mappings, Host Cluster ID, and Host Cluster Name. The data row shows a host named 'RHEL-Host01' with a status of 'Online', a generic host type, 2 ports, and host mappings.

Name	Status	Host Type	# of Ports	Host Mappings	Host Cluster ID	Host Cluster Name
RHEL-Host01	Online	Generic	2	Yes		

Figure 5-39 Existing host list

3. To create a host, click **Add Host** to start the wizard (see Figure 5-40).



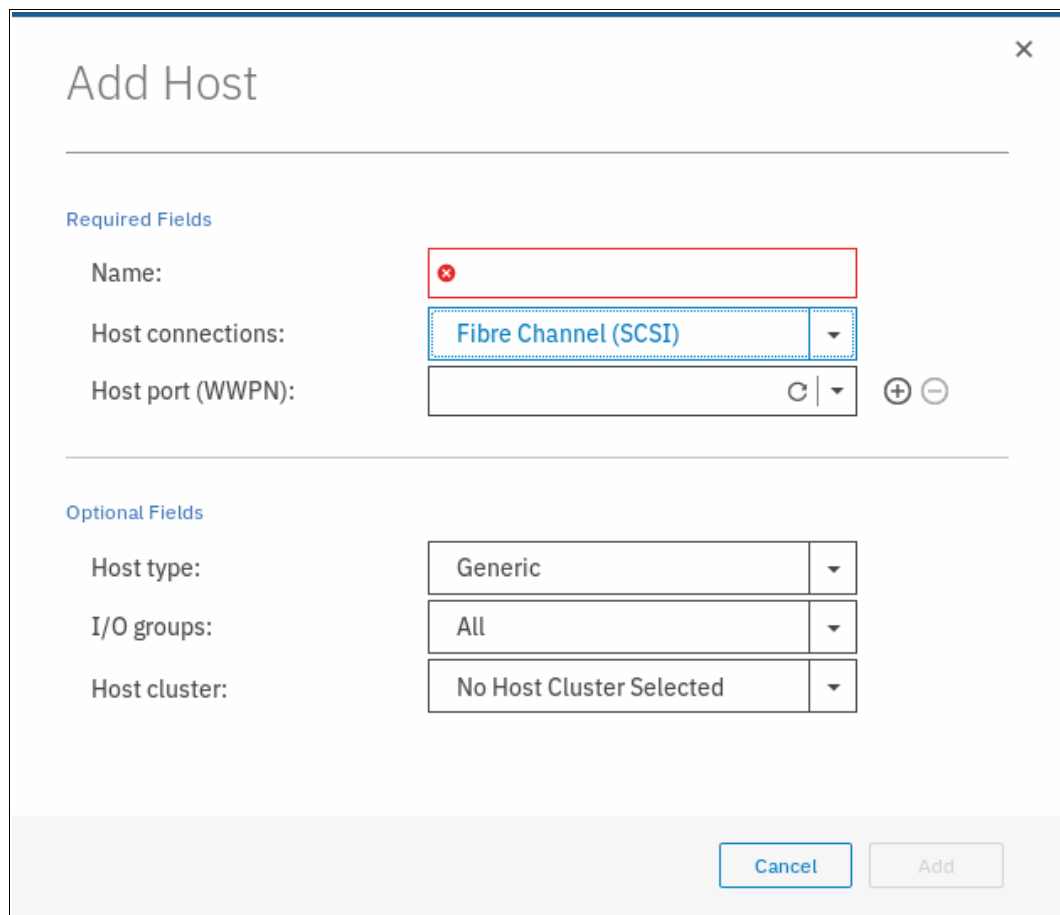
The screenshot shows a user interface with a blue 'Add Host' button on the left. To its right is a 'Actions' dropdown menu and a download icon. Below this is a table with four columns: Name, Status, Host Type, and # of Ports. The table contains two rows of data.

Name	Status	Host Type	# of Ports
RHEL-Host-01	✓ Online	Generic	2
Windows-Host-01	✓ Online	Generic	2

Figure 5-40 Add Host

4. If you want to create an FC host, see 5.5.1, “Creating FC hosts” on page 258. To create iSCSI hosts, see 5.5.3, “Creating iSCSI hosts” on page 270. To create SAS hosts, see 5.5.5, “Creating SAS hosts” on page 278.

5. After you click **Add Host**, the host selection menu opens, as shown in Figure 5-41.



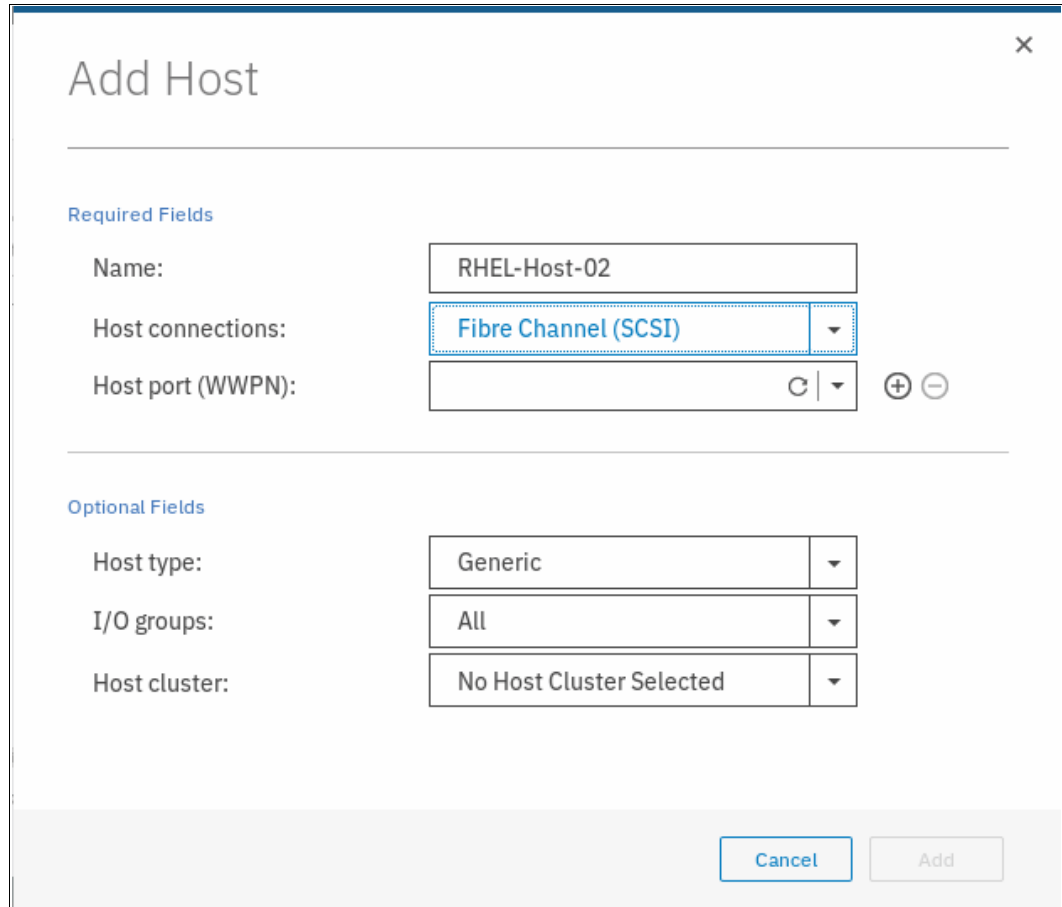
The screenshot shows a 'Add Host' wizard window. It has a title bar with a close button. The main content is divided into 'Required Fields' and 'Optional Fields'. In the 'Required Fields' section, the 'Name' field is empty and has a red error icon. The 'Host connections' dropdown is set to 'Fibre Channel (SCSI)'. The 'Host port (WWPN)' field is empty and has a refresh icon and a plus/minus icon. In the 'Optional Fields' section, the 'Host type' dropdown is set to 'Generic', 'I/O groups' is set to 'All', and 'Host cluster' is set to 'No Host Cluster Selected'. At the bottom right, there are 'Cancel' and 'Add' buttons.

Figure 5-41 Add Host window

5.5.1 Creating FC hosts

To create FC hosts, complete the following steps:

1. Click **Fibre Channel** and enter a host name (see Figure 5-42).



The screenshot shows a dialog box titled "Add Host" with a close button (X) in the top right corner. The dialog is divided into two sections: "Required Fields" and "Optional Fields".

Required Fields:

- Name:** A text input field containing "RHEL-Host-02".
- Host connections:** A dropdown menu with "Fibre Channel (SCSI)" selected.
- Host port (WWPN):** A text input field with a refresh icon and a dropdown arrow, followed by plus (+) and minus (-) icons.

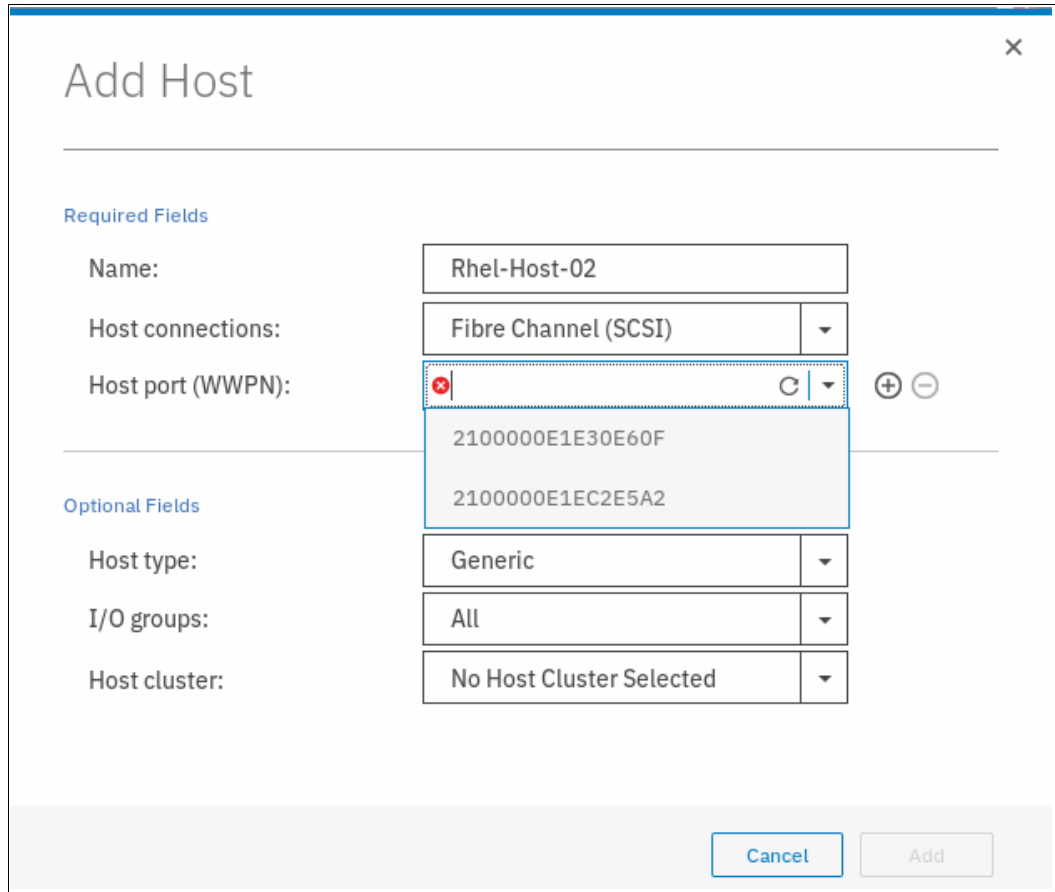
Optional Fields:

- Host type:** A dropdown menu with "Generic" selected.
- I/O groups:** A dropdown menu with "All" selected.
- Host cluster:** A dropdown menu with "No Host Cluster Selected" selected.

At the bottom right of the dialog, there are two buttons: "Cancel" and "Add".

Figure 5-42 Creating an FC host

2. Enter a host name and click the **Host port** drop-down list to get a list of all known WWPNs (see Figure 5-43).



The screenshot shows a window titled "Add Host" with a close button (X) in the top right corner. The window is divided into two sections: "Required Fields" and "Optional Fields".

Required Fields:

- Name:** A text input field containing "Rhel-Host-02".
- Host connections:** A dropdown menu showing "Fibre Channel (SCSI)".
- Host port (WWPN):** A dropdown menu that is open, showing a list of WWPNs: "2100000E1E30E60F" and "2100000E1EC2E5A2". The dropdown has a refresh icon and a close icon (X) in the top left corner. To the right of the dropdown are plus (+) and minus (-) icons.

Optional Fields:

- Host type:** A dropdown menu showing "Generic".
- I/O groups:** A dropdown menu showing "All".
- Host cluster:** A dropdown menu showing "No Host Cluster Selected".

At the bottom right of the window, there are two buttons: "Cancel" and "Add".

Figure 5-43 Available WWPNs

The IBM Storwize V5000 has the host port WWPNs available if you prepared the hosts as described in 5.3, "Preparing the host operating system" on page 219. If they do not appear in the list, scan for new disks in your operating system and click **Rescan** in the configuration wizard. If they still do not appear, check your SAN zoning, correct it, and repeat the scanning.

Note: You can enter WWPNs manually. However, if these WWPNs are not visible to the IBM Storwize V5000, the host object appears as offline and it is unusable for I/O operations until the ports are visible.

3. Select the WWPN for your host (see Figure 5-44).

The screenshot shows a dialog box titled "Add Host" with a close button (X) in the top right corner. The dialog is divided into two sections: "Required Fields" and "Optional Fields".

Required Fields:

- Name:** RHEL-Host-02
- Host connections:** Fibre Channel (SCSI) (dropdown menu)
- Host port (WWPN):** 2100000E1E0405FE (text input field with a refresh icon and a dropdown arrow). To the right of the input field are plus (+) and minus (-) icons.

Optional Fields:

- Host type:** Generic (dropdown menu)
- I/O groups:** All (dropdown menu)
- Host cluster:** No Host Cluster Selected (dropdown menu)

At the bottom right of the dialog, there are two buttons: "Cancel" and "Add".

Figure 5-44 Add a port to a list

4. If you want to add ports to your Host, click the plus sign (+).

5. Add all ports that belong to the host (see Figure 5-45).

Add Host

Name: RHEL-Host-02

Host connections: Fibre Channel (SCSI)

Host port (WWPN): 2100000E1E30E60F

Host port (WWPN): 2100000E1EC2E5A2

Optional Fields

Host type: Generic

I/O groups: All

Host cluster: No Host Cluster Selected

Cancel Add

Figure 5-45 Add all WWPNs

Creating offline hosts: If you want to create hosts that are offline or not connected, you can enter the WWPNs manually. Enter them into the Host port (WWPN) field and add them to the list (see Figure 5-46).

The screenshot shows a dialog box titled "Add Host" with a close button (X) in the top right corner. The dialog is divided into two sections: "Required Fields" and "Optional Fields".

Required Fields:

- Name:** A text input field containing "ESX-01".
- Host connections:** A dropdown menu set to "Fibre Channel (SCSI)".
- Host port (WWPN):** A list of WWPNs. The first entry is "2000000000000000" with plus and minus icons. The second entry is "20000000000000001" with a refresh icon and plus/minus icons. This second entry is highlighted with a blue dotted border.

Optional Fields:

- Host type:** A dropdown menu set to "Generic".
- I/O groups:** A dropdown menu set to "All".

At the bottom right of the dialog, there are two buttons: "Cancel" and "Add".

Figure 5-46 Manually added WWPN

6. If you are creating a Hewlett-Packard UNIX (HP-UX) or Target Port Group Support (TPGS) host, select **Advanced**, and more options appear (see Figure 5-47). Select your host type.

The screenshot shows a window titled "Add Host" with a close button (X) in the top right corner. The window is divided into two main sections. The top section contains the following fields:

- Name:** A text input field containing "ESX-01".
- Host connections:** A dropdown menu showing "Fibre Channel (SCSI)".
- Host port (WWPN):** A text input field containing "200000000000000000". To its right are plus (+) and minus (-) icons.
- Host port (WWPN):** A second text input field containing "200000000000000001". To its right are a refresh icon and plus (+) and minus (-) icons.

The bottom section is titled "Optional Fields" and contains:

- Host type:** A dropdown menu currently showing "Generic". A list is open below it, showing the following options: "Generic", "HP/UX", "OpenVMS", "TPGS", and "VVOL".
- I/O groups:** A text input field.
- Host cluster:** A text input field.

An "Add" button is located at the bottom right of the window.

Figure 5-47 Add Host: Advanced Settings

7. You can set the **I/O Groups** that your host can access. The host objects must belong to the same I/O groups as the volumes that you want to map. Otherwise, these volumes are not visible to the host (see Figure 5-48).

The screenshot shows the 'Add Host' dialog box with the following configuration:

- Name: ESX-01
- Host connections: Fibre Channel (SCSI)
- Host port (WWPN): 2000000000000000
- Host type: Generic
- I/O groups: All (dropdown menu is open showing checked options: All, io_grp0, io_grp1)
- Host cluster: (empty)

Buttons: Cancel, Add

Figure 5-48 Setting I/O groups

Note: IBM Storwize V5000 supports a maximum of two control enclosures for each system, and each control enclosure can have two nodes. The two nodes are arranged as two I/O groups per cluster. Because of the host object limit per I/O group, it is best to create hosts that use single I/O groups for maximum host connectivity.

8. Select the wanted host cluster (if any), as shown in Figure 5-49.

Figure 5-49 Selecting Host cluster

9. Click **Add Host**, and a task completion window is displayed. Click **Close**.

10. On the Hosts window, the host that was created appears, as shown in Figure 5-50.

Name	Status	Host Type	# of Ports	Host Mappings	Host Cluster ID	Host Cluster Name
RHEL-Host-01	Online	Generic	2	No	1	ITSO-Host-Cluster-01
RHEL-Host-02	Online	Generic	2	No	1	ITSO-Host-Cluster-01
RHEL-Host-03	Offline	Generic	2	No		
RHEL-Host-04	Offline	Generic	2	No		
RHEL-Host-05	Offline	Generic	2	No		
Windows-Host-01	Online	Generic	2	No		

Showing 6 hosts | Selecting 1 host

Figure 5-50 Defined hosts

11. Repeat steps 1 - 10 for all of your FC hosts. After you add all of the FC hosts, create volumes and map them to the created hosts. For more information, see Chapter 6, “Volume configuration” on page 309.

5.5.2 Configuring the IBM Storwize V5000 for FC connectivity

You can configure the FC ports on the IBM Storwize V5000 for use for certain connections only. This capability is referred to as *port masking*. In a system with multiple I/O groups and remote partnerships, port masking is a useful tool for ensuring peak performance and availability.

The following options are available per port:

- ▶ Any: Allow local and remote communication between nodes.
- ▶ Local: Allow only local node communication.
- ▶ Remote: Allow only remote node communication.
- ▶ None: Do not allow any node communication.

In all cases, host I/O is still permitted, so the None option can be used to exclusively reserve a port for host I/O and backend if IBM Storwize V5030 also is used for external backend storages.

A limit of 16 logins exists per node from another node before an error is logged. A combination of port masking and SAN zoning can help you manage logins and provide optimum I/O performance with local, remote, and host traffic.

To configure FC ports, complete the following steps:

1. Go to **Settings** → **Network**, as shown in Figure 5-51.

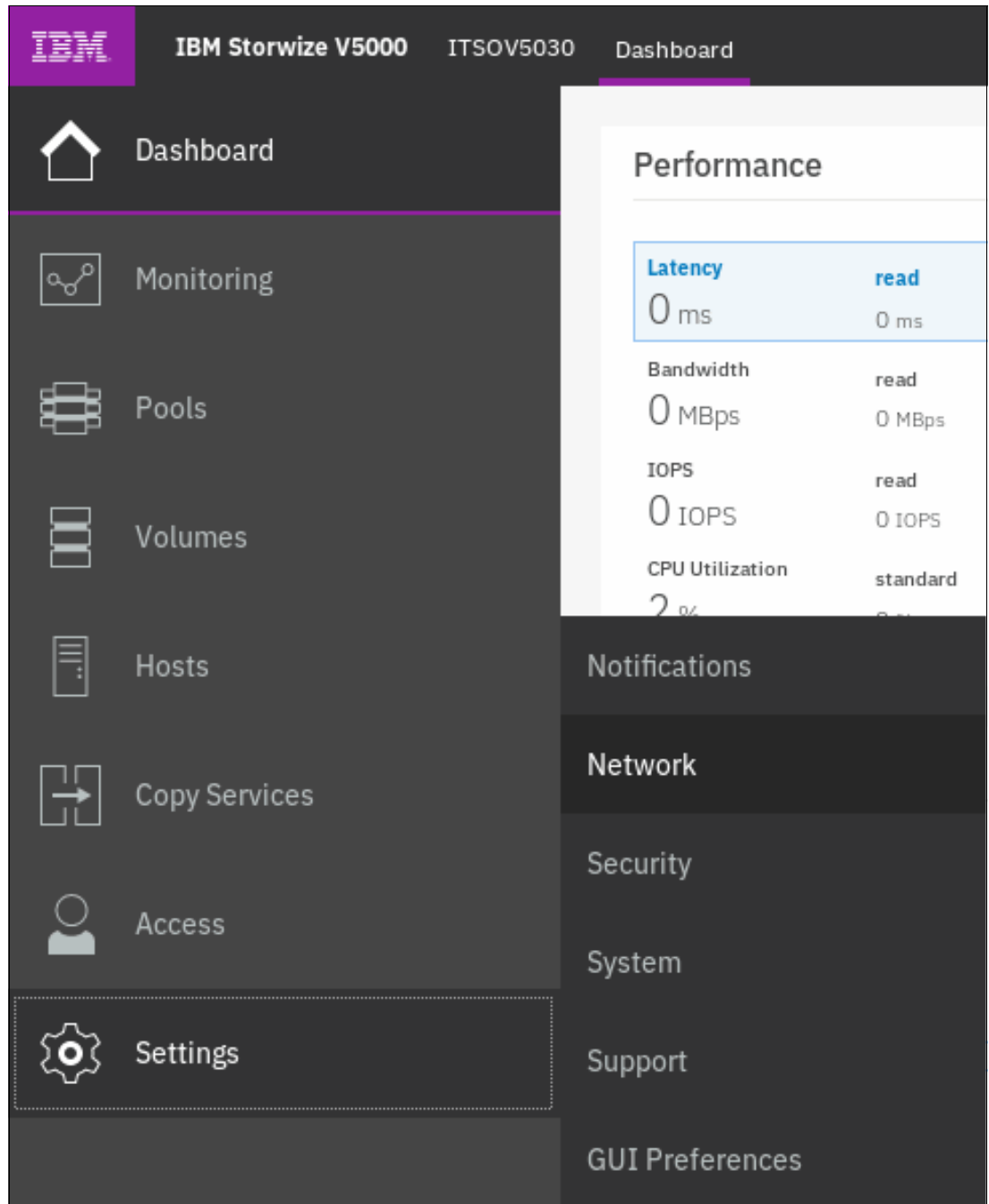


Figure 5-51 Opening the network settings view

2. Select **Fibre Channel Ports** and the Fibre Channel Ports configuration view displays, as shown in Figure 5-52.

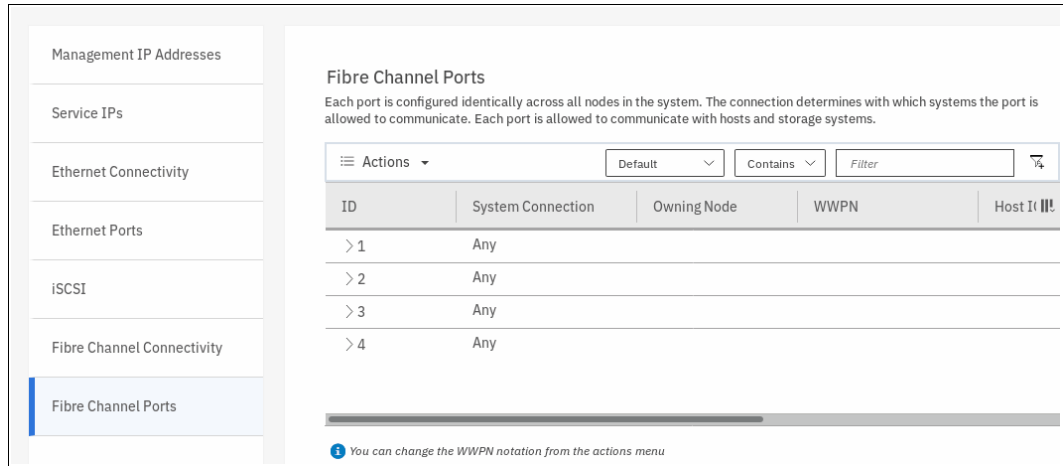


Figure 5-52 The Fibre Channel Ports view

3. To configure a port, select the port, and from the **Actions** menu, select **Modify Connection**. The window that is shown in Figure 5-53 opens.

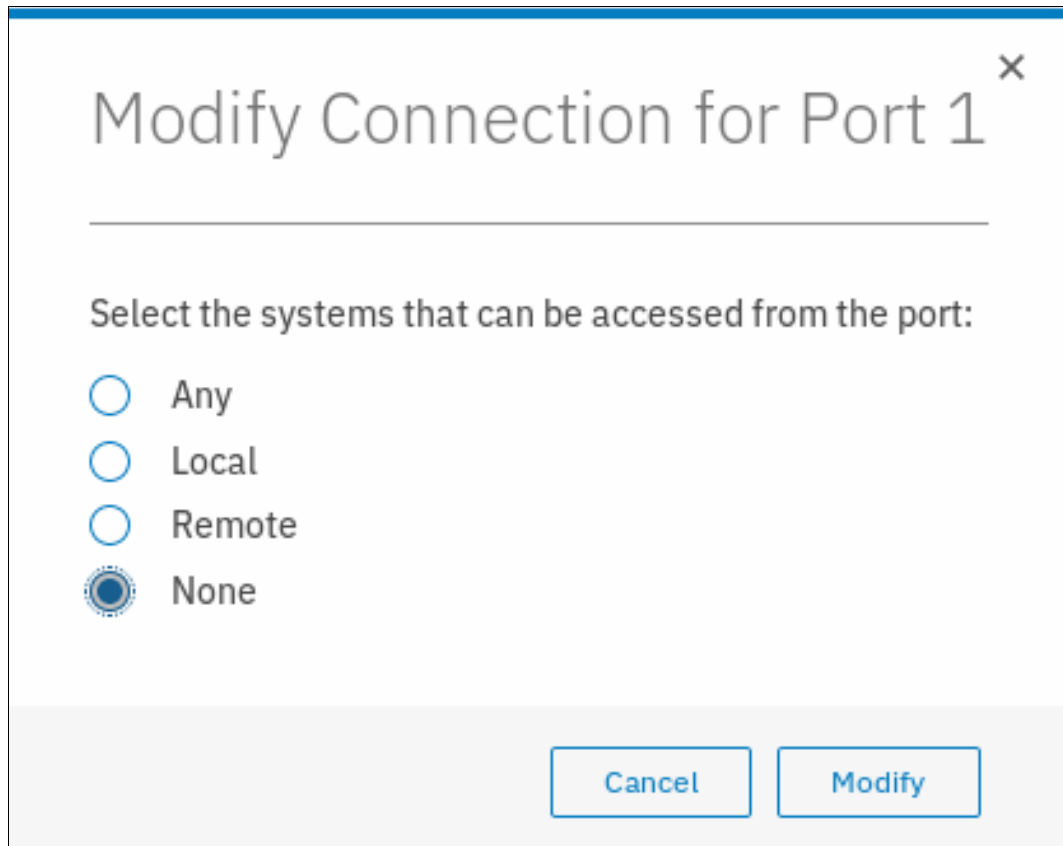


Figure 5-53 Modifying the connection for Ports 3 and 4

In this example, we selected **Local** for ports 1 and 2 to allow node to node traffic, and **None** for ports 3 and 4 to restrict traffic on these ports to host I/O only. Click **Modify** to confirm the selection.

Note: This action configures the selected ports for *all* nodes. You cannot configure FC ports on a per node basis. You can also select multiple ports to modify the connections at once. You also cannot select multiple ports at the same time and set different connections for them. In our example, we first selected ports 1 and 2, modified them for Local; then, we selected ports 3 and 4 and modified them to None.

4. You can view connections between nodes, storage systems, and hosts by selecting **Fibre Channel Connectivity** while you are in the Network settings view. Choose the connections that you want to view and click **Show Results**, as shown in Figure 5-54.

Fibre Channel Connectivity
 Display the connectivity between nodes and other storage systems and hosts that are attached through the Fibre Channel network.

View connectivity for: All nodes, storage systems, and hosts [Show Results](#)

Name	System Name	Remote WWPN	Remote ...	localWwpn	
DS8000		500507630A00C575	011600	500507680D0496E6	1
DS8000		500507630A00C575	011600	500507680D0496E7	1
RHEL-Host-01		210000E1E09E3E9	010100	500507680D0496E7	1

You can change the WWPN notation from the actions menu

Figure 5-54 Viewing FC connections between nodes, storage systems, and hosts

5.5.3 Creating iSCSI hosts

To create iSCSI hosts, complete the following steps:

1. On Hosts window, click **Add Host**. Then, click **iSCSI** and the iSCSI configuration wizard opens (see Figure 5-55).

The screenshot shows a window titled "Add Host" with a close button (X) in the top right corner. The window is divided into two sections: "Required Fields" and "Optional Fields".

Required Fields:

- Name:** A text input field containing "VMware-Host-01".
- Host connections:** A dropdown menu with "iSCSI (SCSI)" selected.
- Host IQN:** An empty text input field with a red border and a red "x" icon, indicating an error.

Optional Fields:

- CHAP authentication:** A checkbox that is unchecked.
- CHAP secret:** A text input field with a placeholder "Enter 1 to 79 characters".
- CHAP username:** A text input field with a placeholder "Enter 1 to 31 characters".
- Host type:** A dropdown menu with "Generic" selected.

At the bottom right of the window, there are two buttons: "Cancel" and "Add".

Figure 5-55 Add an iSCSI host

2. Enter a descriptive host name (the iSCSI initiator name in the name field). Enter the iSCSI Qualified Name (IQN) information, as shown in Figure 5-56 on page 271.
If you want to add several initiator names to one host, repeat this step by clicking the plus sign (+).

3. If you are connecting an HP-UX or TPGS host, select the **Host type** and select the correct host type (see Figure 5-56). Click **Add**.

The screenshot shows a dialog box titled "Add Host" with a close button (X) in the top right corner. The dialog is divided into sections. The first section contains "Host connections:" with a dropdown menu set to "iSCSI (SCSI)" and "Host IQN:" with a text input field containing "iqn-1998-01.com.vmware.iscsi:itsc" and expand/collapse icons (+/-). Below this is a section titled "Optional Fields" in blue. It includes "CHAP authentication:" with an unchecked checkbox, "CHAP secret:" with a text input field containing "Enter 1 to 79 characters", "CHAP username:" with a text input field containing "Enter 1 to 31 characters", "Host type:" with a dropdown menu set to "Generic", and "I/O groups:" with a dropdown menu set to "All". At the bottom right of the dialog are "Cancel" and "Add" buttons.

Figure 5-56 Create an iSCSI host: Optional settings

4. You can set the I/O groups that your host can access.

Important: The host objects must belong to the same I/O groups as the volumes that you want to map. Otherwise, the host cannot see these volumes. For more information, see Chapter 6, "Volume configuration" on page 309.

The IBM Storwize V5000 supports a maximum of four nodes per system. These nodes are arranged as two I/O groups per cluster. Because of the host object limit per I/O group, it is best to create hosts that use single I/O groups for maximum host connectivity.

5. A task completion window is displayed. Click **Close**.

6. A task completion window is displayed. Click **Close**.
7. Repeat these steps for every iSCSI host that you want to create. Figure 5-57 shows all of the hosts that were created in the system. The host names highlighted were created in this section to illustrate the process.

Name	Status	Host Type	# of Ports	Host Mappings	Host Cluster ID	Host Cluster Name
RHEL-Host-01	Online	Generic	2	No	1	ITSO-Host-Cluster-01
RHEL-Host-02	Online	Generic	2	No	1	ITSO-Host-Cluster-01
RHEL-Host-03	Offline	Generic	2	No		
RHEL-Host-04	Offline	Generic	2	No		
RHEL-Host-05	Offline	Generic	2	No		
VMware-Host-01	Offline	Generic	1	No		
VMware-Host-02	Online	Generic	2	No		

Showing 7 hosts | Selecting 2 hosts

Figure 5-57 All hosts

Note: iSCSI hosts might show a Degraded status until the volumes are mapped. This limitation relates to the implementation of iSCSI in the IBM Storwize V5000. This status is not necessarily a problem with network connectivity or the host configuration.

The iSCSI host is now configured on the IBM Storwize V5000. To provide connectivity, the iSCSI Ethernet ports must also be configured.

5.5.4 Configuring the IBM Storwize V5000 for iSCSI host connectivity

Complete the following steps to enable iSCSI connectivity:

1. Switch to the configuration Overview window and select **Network** (see Figure 5-58).

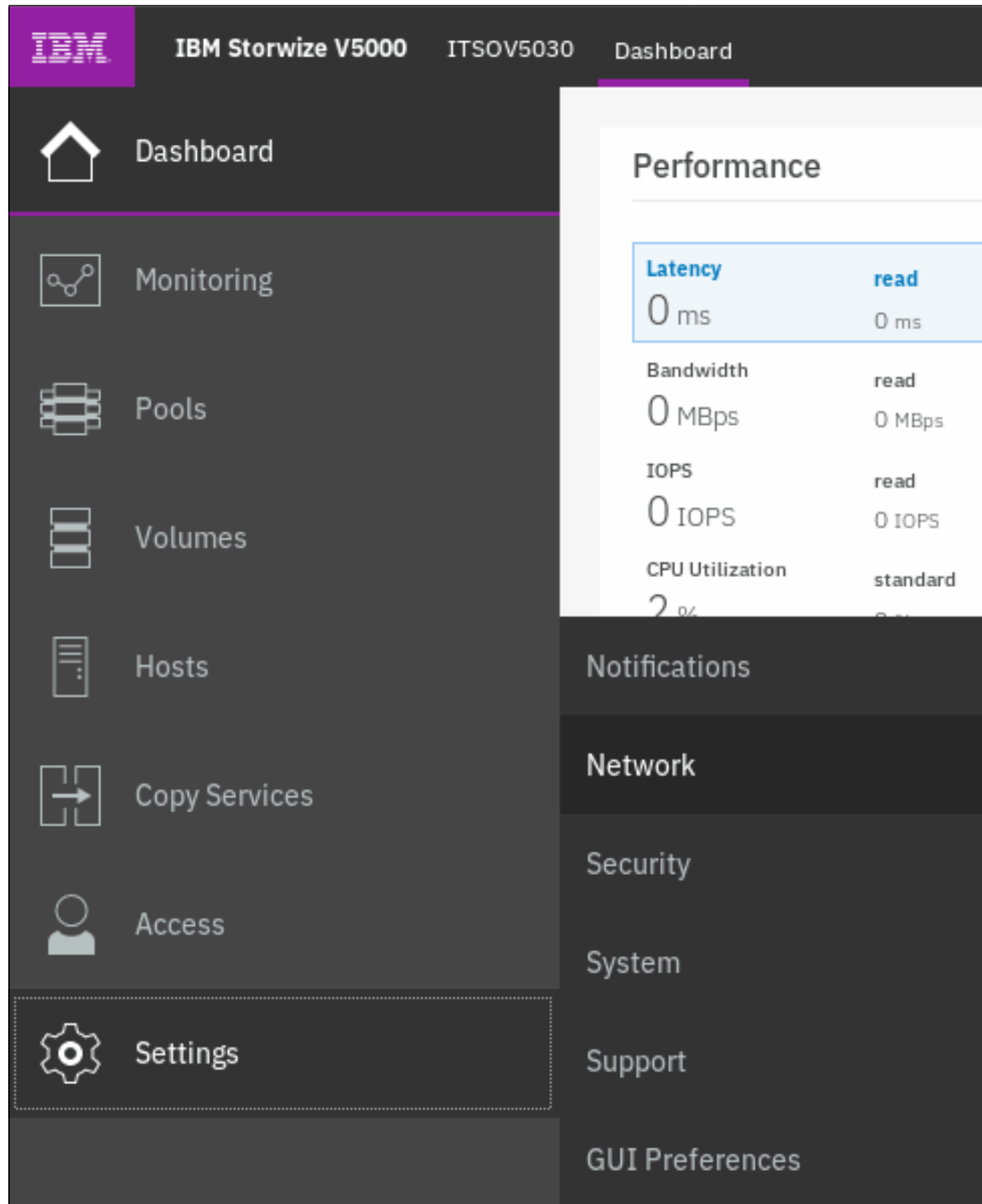


Figure 5-58 Configuration: Network

2. Select **iSCSI** and the iSCSI Configuration window opens (see Figure 5-59).

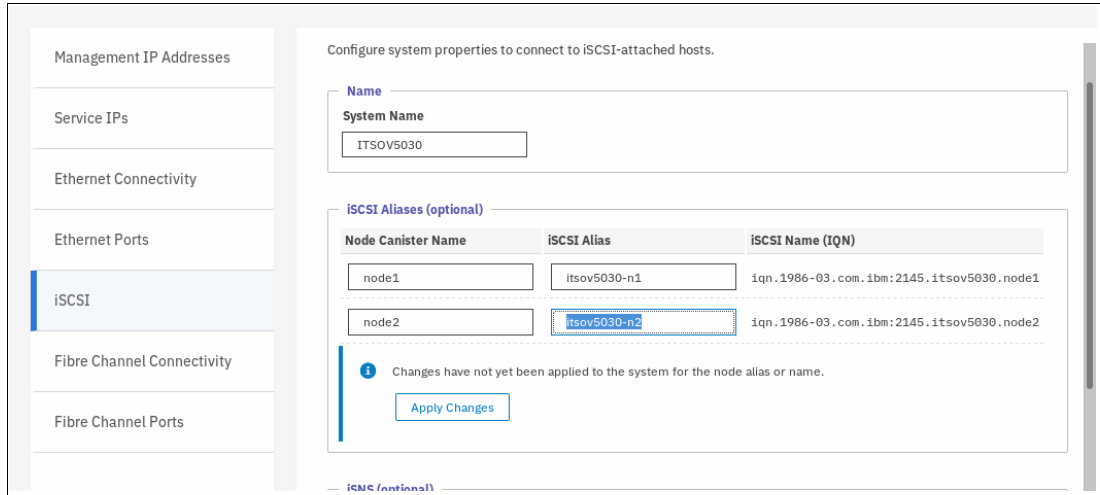


Figure 5-59 iSCSI Configuration window

3. The system waits until you apply the changes that you made. Click **Apply Changes**. A task completion window is displayed. Click **Close**.
4. The Configuration window (see Figure 5-59) shows an overview of all of the iSCSI settings for the IBM Storwize V5000. You can configure the iSCSI alias, internet Storage Name Service (iSNS) addresses, and Challenge Handshake Authentication Protocol (CHAP).
5. Click **Ethernet Ports** to enter the iSCSI IP address (see Figure 5-60). Select the wanted port to configure. Then, select the **Actions** menu, and select **Modify IP Settings**. Repeat this step for each port that you want to use for iSCSI traffic.

Note: We advise that VLANs are used for iSCSI to separate iSCSI traffic from management traffic.

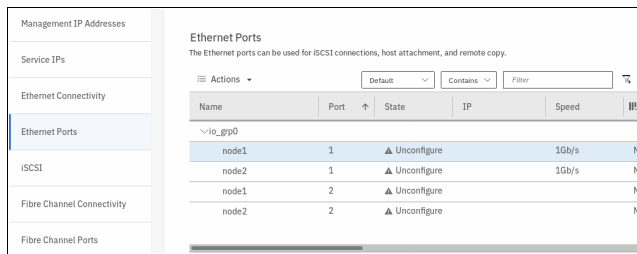


Figure 5-60 Select port to configure iSCSI traffic

- After you enter the IP address, subnet mask, and gateway, click **Modify** to enable the configuration, as shown in Figure 5-61.

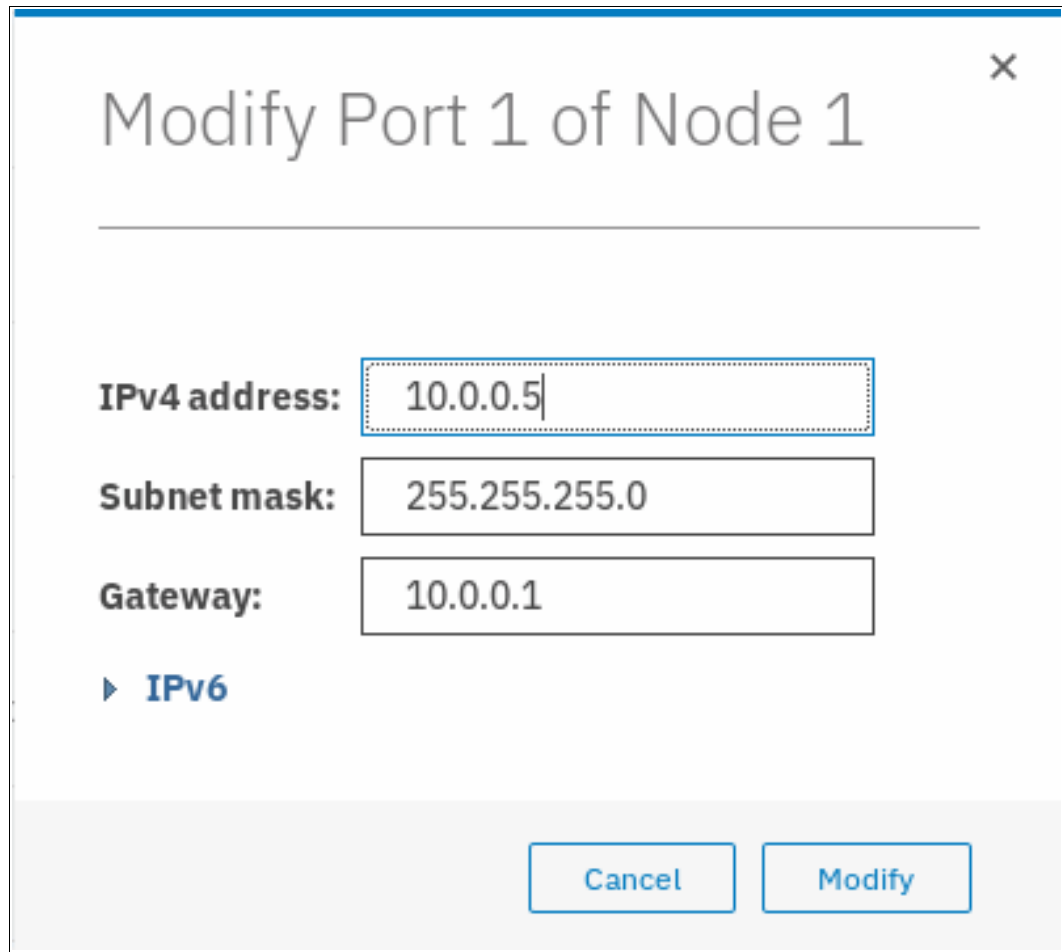


Figure 5-61 Configure iSCSI IP

- After the changes are successfully applied, click **Close**.
- Under Actions, you can select whether IPV4, IPV6, or both types of hosts are enabled for iSCSI traffic (see Figure 5-62).

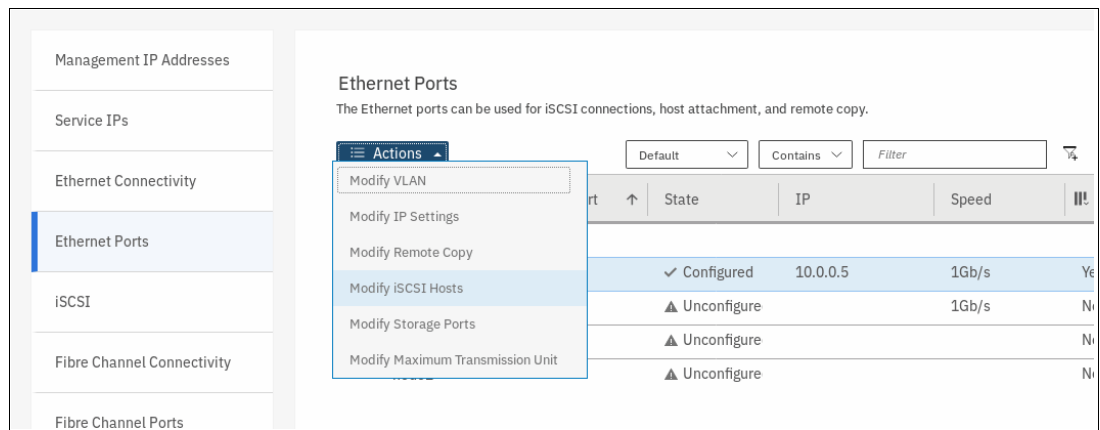


Figure 5-62 Actions menu to modify iSCSI hosts

By default, only IPV4 iSCSI hosts are enabled (see Figure 5-63).

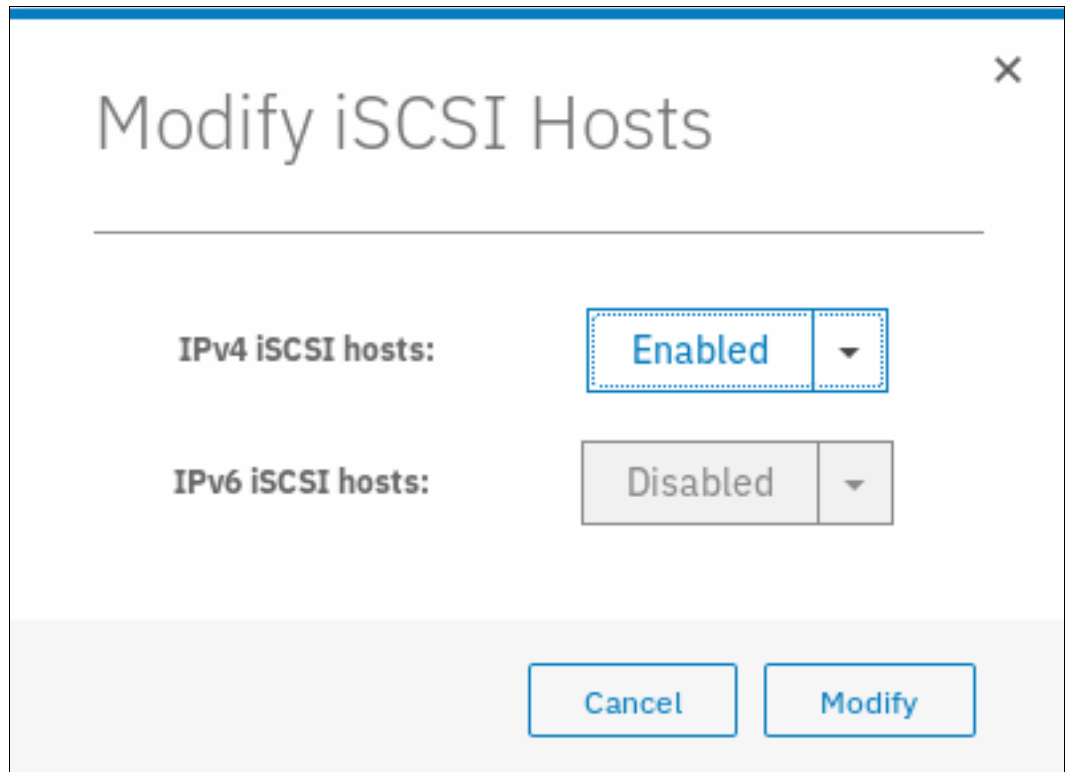


Figure 5-63 Enabled iSCSI hosts

9. To modify virtual LAN (VLAN) settings, select the port again, click **Actions** and then, click **Modify VLAN**. The window that is shown in Figure 5-64 on page 277 opens. Select **Enable** to turn on VLAN tagging and set the tag in the field that is provided. Because the failover port must belong to the same VLAN, leave **Apply change to the failover port too** as selected. Click **Modify** to confirm.

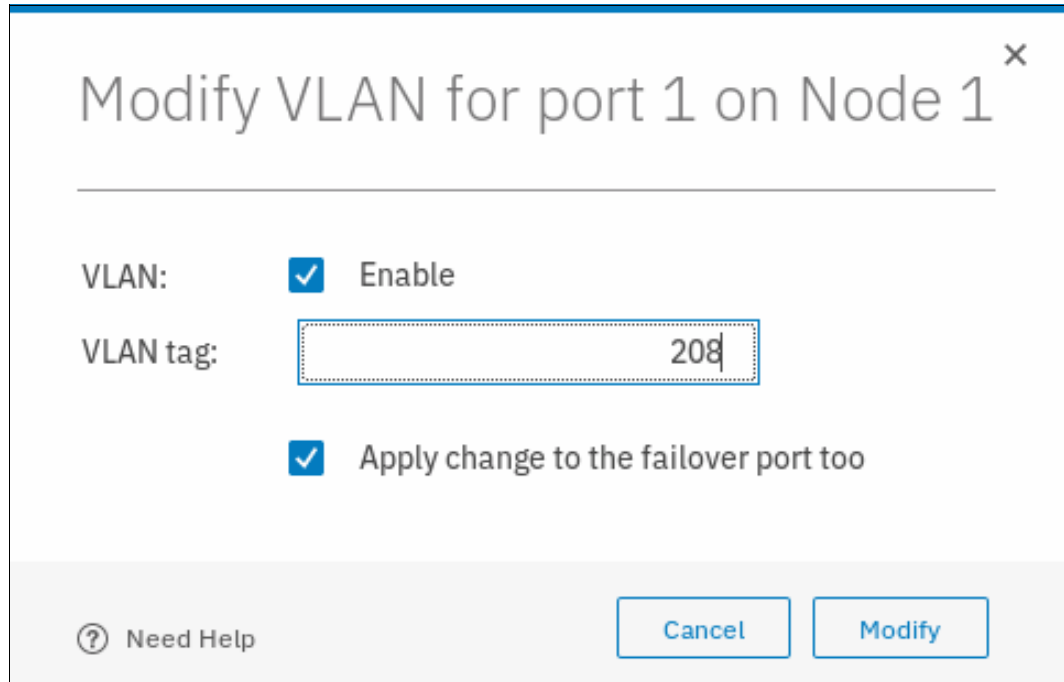


Figure 5-64 Modifying VLAN settings for Port 2 of Node 1

10. Repeat the previous steps for all ports that must be configured.
11. You can also configure iSCSI aliases, an iSNS name, and CHAP authentication. These options are in the iSCSI Configuration view. To access this view, click **iSCSI** in the Network settings view, as shown in Figure 5-65.

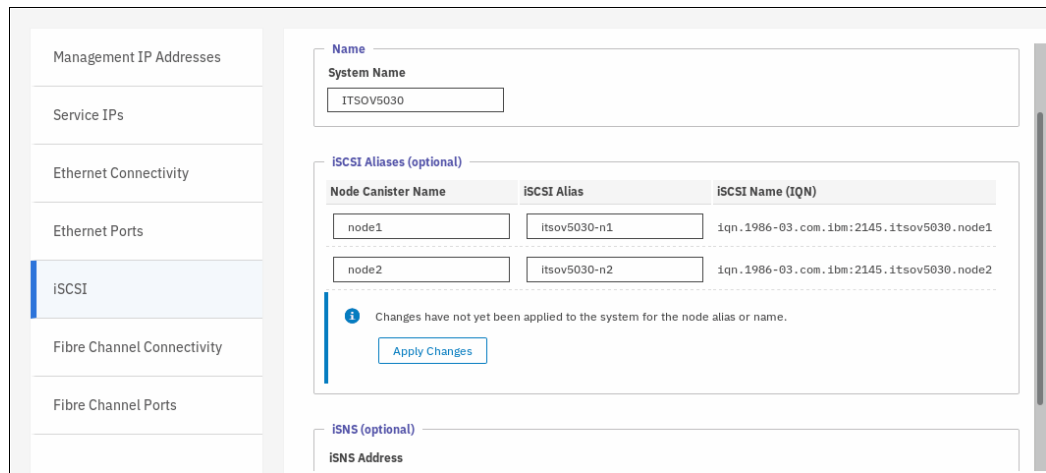


Figure 5-65 Advanced iSCSI configuration

The IBM Storwize V5000 is now configured and ready for iSCSI use. Document the initiator names of your storage canisters (see Figure 5-59 on page 274) because you need them later.

For more information about creating volumes and mapping them to a host, see Chapter 6, “Volume configuration” on page 309.

5.5.5 Creating SAS hosts

To create an SAS host, complete the following steps:

1. From the main window, click **Hosts** → **Hosts** → **Add Host**. The host configuration wizard opens, as shown in Figure 5-66.

Add Host

Required Fields

Name:

Host connections: iSCSI SAS

iSCSI host IQN:

Optional Fields

CHAP authentication:

CHAP secret:

Host type: ▼

I/O groups: ▼

Host cluster: ▼

Figure 5-66 The host configuration wizard

2. Enter a descriptive host name and click **SAS** under the **Host Connections** option, as shown in Figure 5-67.

The screenshot shows the 'Add Host' dialog box with the following fields:

- Required Fields:**
 - Name: ITSO_SAS_HOST
 - Host connections: iSCSI SAS
 - Host port (WWPN): No candidate ports found
- Optional Fields:**
 - Host type: Generic
 - I/O groups: All
 - Host cluster: No Host Cluster Selected

Buttons: Cancel, Add

Figure 5-67 SAS WWPNS that are visible to the system

3. Click the **Host port (WWPN)** drop-down menu and select the wanted WWPNS that belong to the host, as shown in Figure 5-68.

The screenshot shows the 'Add Host' dialog box with the following fields:

- Required Fields:**
 - Name: ITSO_SAS_HOST
 - Host connections: iSCSI SAS
 - Host port (WWPN): 1234567887654321, 8765432112345678
- Optional Fields:**
 - Host type: Generic
 - I/O groups: All
 - Host cluster: No Host Cluster Selected

Buttons: Cancel, Add

Figure 5-68 SAS WWPNS

- If you prepared a SAS host (as described in 5.3, “Preparing the host operating system” on page 219), the WWPNs that you recorded in this section appear. If they do not appear in the list, verify that you completed all of the required steps and check your cabling. Then, click **Rescan** in the configuration wizard. Ensure that the ends of the SAS cables are aligned correctly.

Note: You can enter WWPNs manually. However, if these WWPNs are not visible to the IBM Storwize V5000, the host object appears as offline and it is unusable for I/O operations until the ports are visible.

- Under the **Optional Fields** section, you can set host type, the I/O groups that your host can access, and the host cluster the host belongs if it is defined.

Important: Host objects must belong to the same I/O groups as the volumes that you want to map. Otherwise, the volumes are not visible to the host.

The IBM Storwize V5000 supports a maximum of four nodes per system. These nodes are arranged as two I/O groups per cluster. Because of the host object limit per I/O group, for maximum host connectivity, it is best to create hosts that use single I/O groups.

You can choose the host type. If you use HP/UX, OpenVMS, or TPGS, you must configure the host. Otherwise, the default option (Generic) is acceptable.

- Click **Add Host** and a task completion window opens.
- Click **Close** to return to the host view, which now lists your newly created host object, as shown in Figure 5-69.

Name	Status	Host Type	# of Ports	Host Mappings	Host Cluster ID	Host Cluster Name
ossw003_host	✓ Online	Generic	2	Yes		
ITSO_SAS_HOST	● Offline	Generic	2	No		

Figure 5-69 Hosts view lists the newly created host object

- Repeat these steps for all of your SAS hosts.

After all of the host objects are created, see Chapter 6, “Volume configuration” on page 309 to create volumes and map them to the hosts.

5.6 Host Clusters

IBM Spectrum Virtualize software supports host clusters starting with version 7.7.1. A *host cluster* is a group of logical host objects that can be managed together. For example, you can create a volume mapping that is shared by every host in the host cluster. The systems use internal protocols to manage access to the volumes and ensure consistency of the data. Host objects that represent hosts can be grouped in a host cluster and share access to volumes.

Volume mappings can either be shared or private:

- Shared mappings are volume mappings that are shared among all the hosts that are in a host cluster. When a host cluster is created, any common volume mappings become shared among all the hosts within the host cluster.

- ▶ Private mappings are mappings that are associated with an individual host. If a mapping is not common, it remains a private mapping for that host only.

A host cluster allows a user to create a group of hosts to form a cluster, which is treated as one single entity, thus allowing multiple hosts to have access to the same set of volumes.

Volumes that are mapped to that host cluster are assigned to all members of the host cluster with the same SCSI ID.

By defining a host cluster, the user can map one or more volumes to the host cluster object. As a result, the volume (or set of volumes) in turn are assigned to an individual host object that is part of the host cluster and each of the volumes are mapped with the same SCSI ID to all the hosts that are part of the host cluster with one command.

A host cluster is made up of individual hosts, and volumes can also be assigned to individual hosts that make up the cluster. Although a host is part of host cluster, volumes can still be assigned to a particular host in a non-shared manner. A policy can be devised which might pre-assign a standard set of SCSI IDs for volumes to be assigned to the host cluster, and another set of SCSI IDs to be used for individual assignments to hosts.

Note: For example, SCSIs ID 0 - 100 can be used for individual host assignment, and SCSI IDs higher than 100 can be used for host cluster. By employing such a policy, wanted volumes are not shared, and others can be shared. For example, the boot volume of each host can be kept private, and data and application volumes can be shared.

A typical use case is to define a host cluster that contains all of the WWPNs that belong to the hosts that are participating in a host operating system-based cluster, such as IBM PowerHA® and Microsoft Cluster Server (MSCS).

This section describes the following host cluster operations by using the GUI:

- ▶ Creating a host cluster
- ▶ Adding a member to the host cluster
- ▶ Listing host cluster members
- ▶ Assigning a volume to the host cluster
- ▶ Unmapping a volume from the host cluster
- ▶ Removing a host cluster member
- ▶ Removing the host cluster

Notes: From IBM Spectrum Virtualize V8.1.0 onwards, various Host Cluster-related operations can also be done by using the GUI in addition to CLI.

Host clusters enable you to create individual hosts and add them to a host cluster. Care must be taken to ensure that no loss of access occurs when transitioning to host clusters.

5.6.1 Creating a host cluster

To create a host cluster by using the GUI, complete the following steps:

1. Click **Hosts** → **Hosts** from the main window, as shown in Figure 5-70.

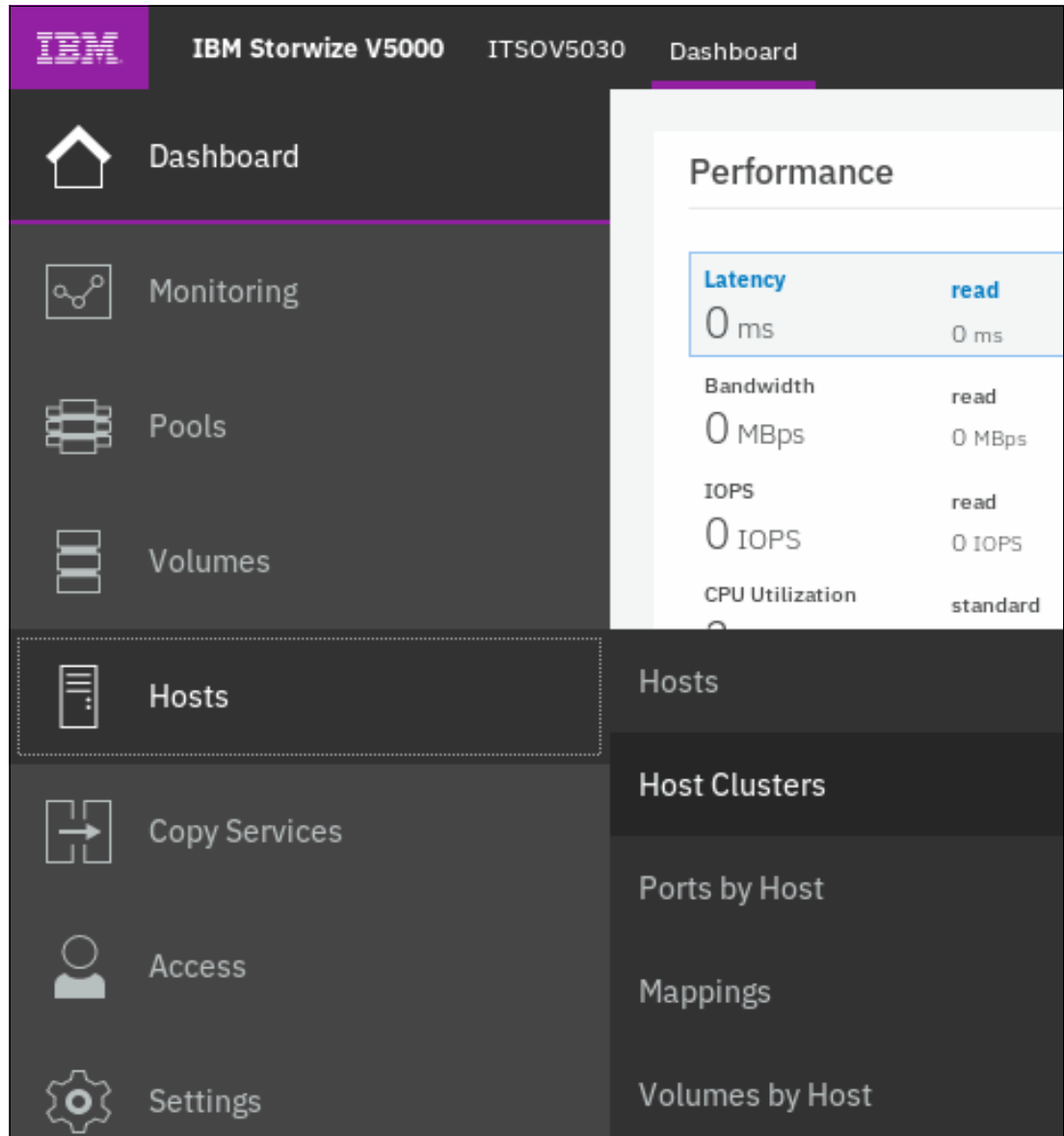


Figure 5-70 Host Clusters option

2. Click **Create Host Cluster**, as shown in Figure 5-71.

The screenshot shows the 'Create Host Cluster' dialog box. At the top left is a 'Create Host Cluster' button. To its right are 'Actions' and a download icon. Further right are dropdown menus for 'Default' and 'Contains', and a 'Filter' input field. Below this is a table with the following data:

ID	Name	Status	Host Count	Mappings Count	
0	ITSO-Host-Cluster-02	Offline	0	0	
2	ITSO-Host-Cluster-03	Offline	0	0	

Figure 5-71 Create Host Cluster option

3. Enter the name of the Host Cluster that you want to create, as shown in Figure 5-72.

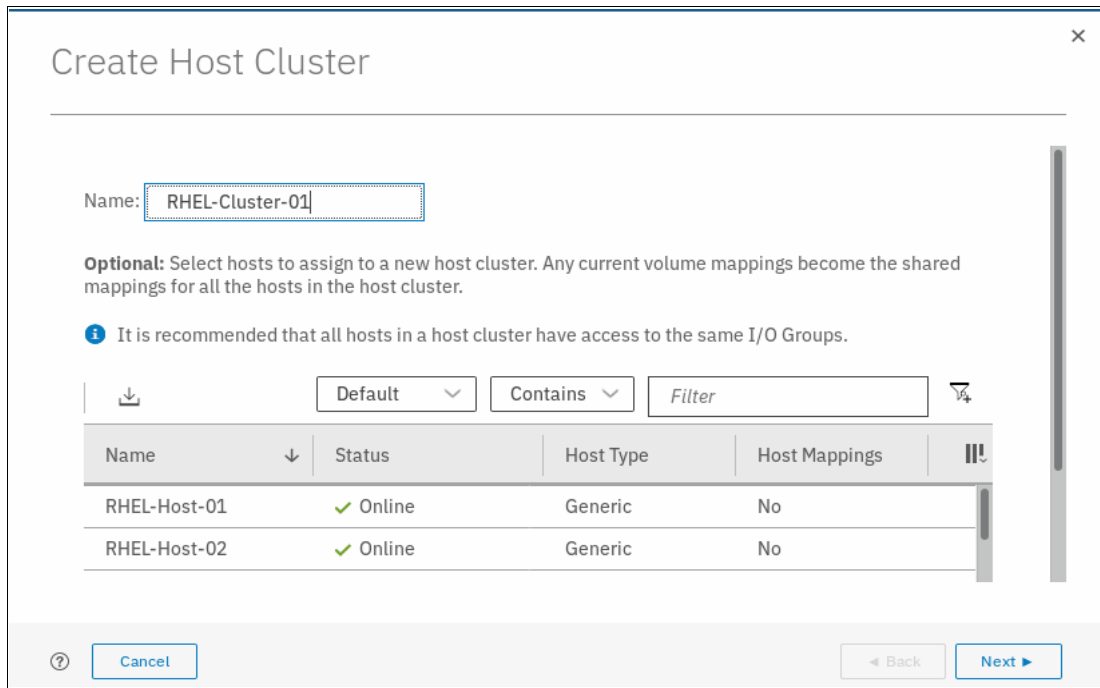


Figure 5-72 Defining Host Cluster

4. Select the hosts that you want to be part of the cluster, as shown in Figure 5-73.

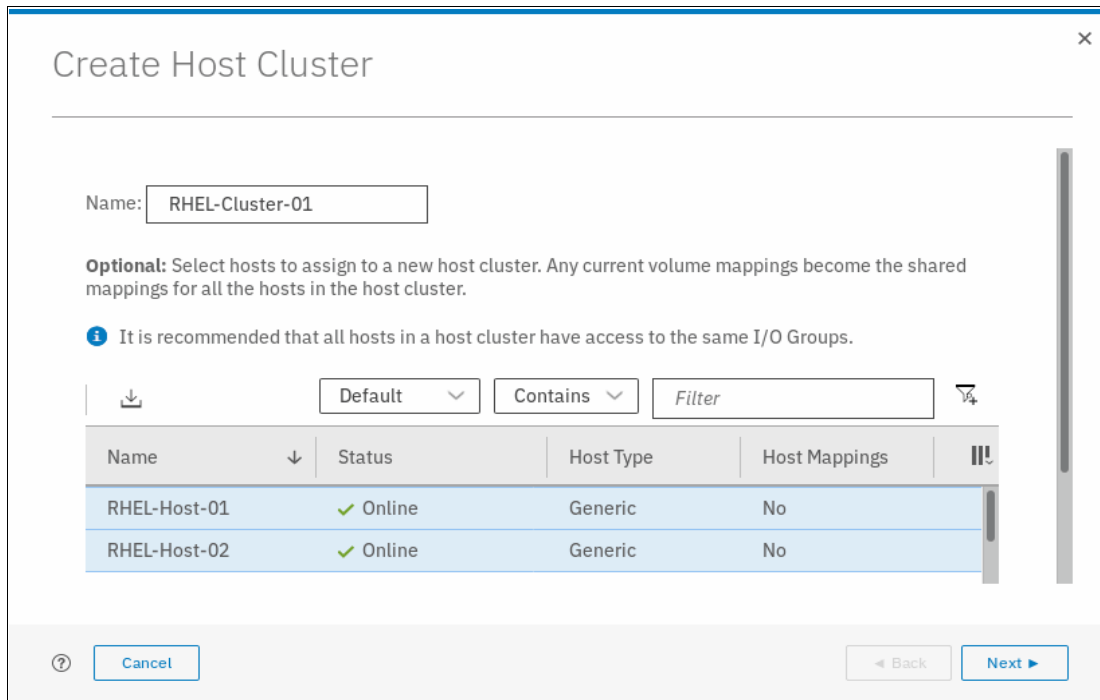


Figure 5-73 Selecting hosts for cluster

5. Click **Next** and a summary is displayed, as shown in Figure 5-74.

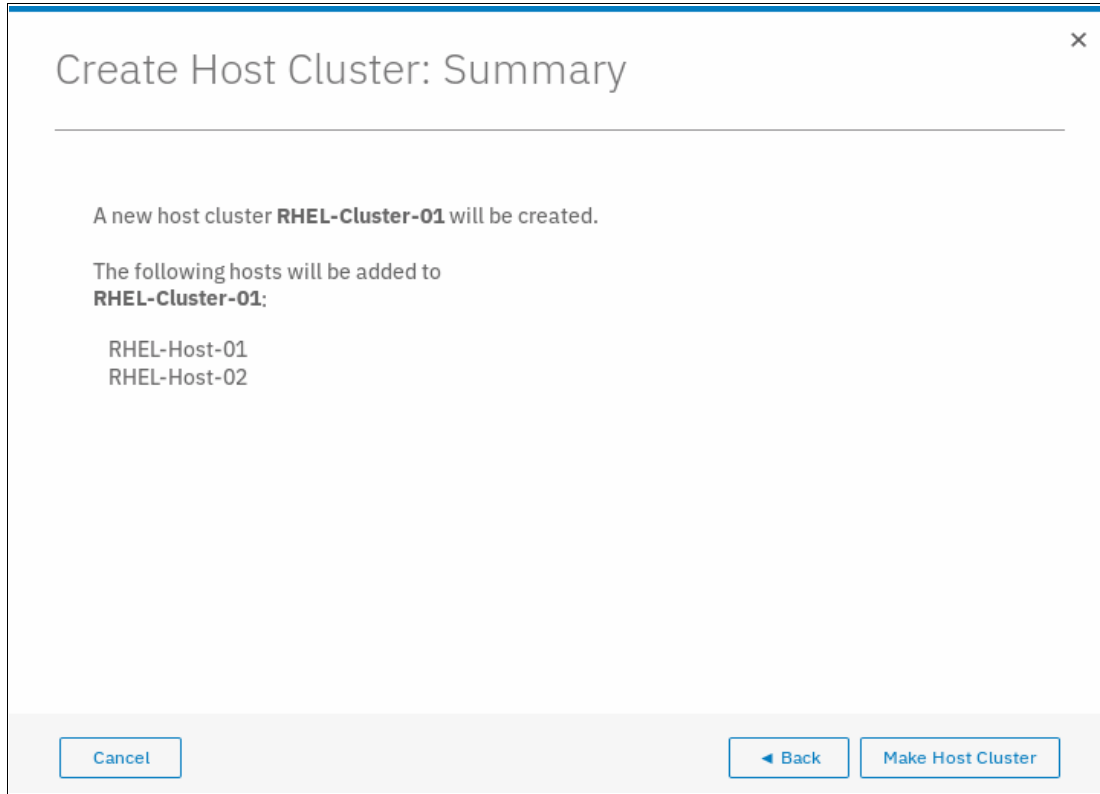


Figure 5-74 Create Host Cluster: Summary

6. Click **Make Host Cluster**, and a task completion window is displayed. Then, click **Close**. The host cluster that you created is available in the list of clusters, as shown in Figure 5-75.

ID	Name	Status	Host Count	Mappings Count
0	ITSO-Host-Cluster-02	Offline	0	0
1	RHEL-Cluster-01	Online	2	0
2	ITSO-Host-Cluster-03	Offline	0	0

Showing 3 host clusters | Selecting 0 host clusters

Figure 5-75 Host Clusters list

5.6.2 Adding a member to a host cluster

To add a member to a host cluster, complete the following steps:

1. From the **Host Clusters** window, select the wanted host cluster to which you want to add members and click **Actions**, as shown in Figure 5-76.

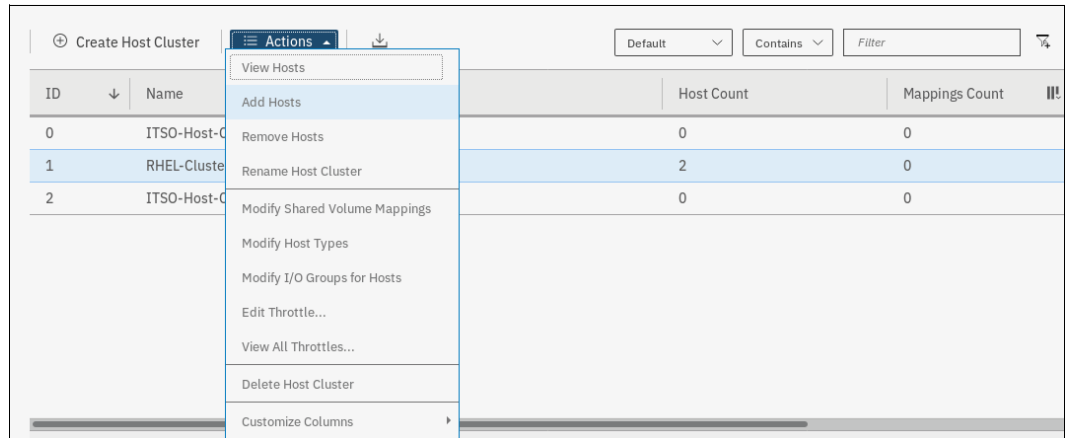


Figure 5-76 Selecting the Host Cluster

2. Click **Add Host** and a selection window opens, as shown in Figure 5-77.

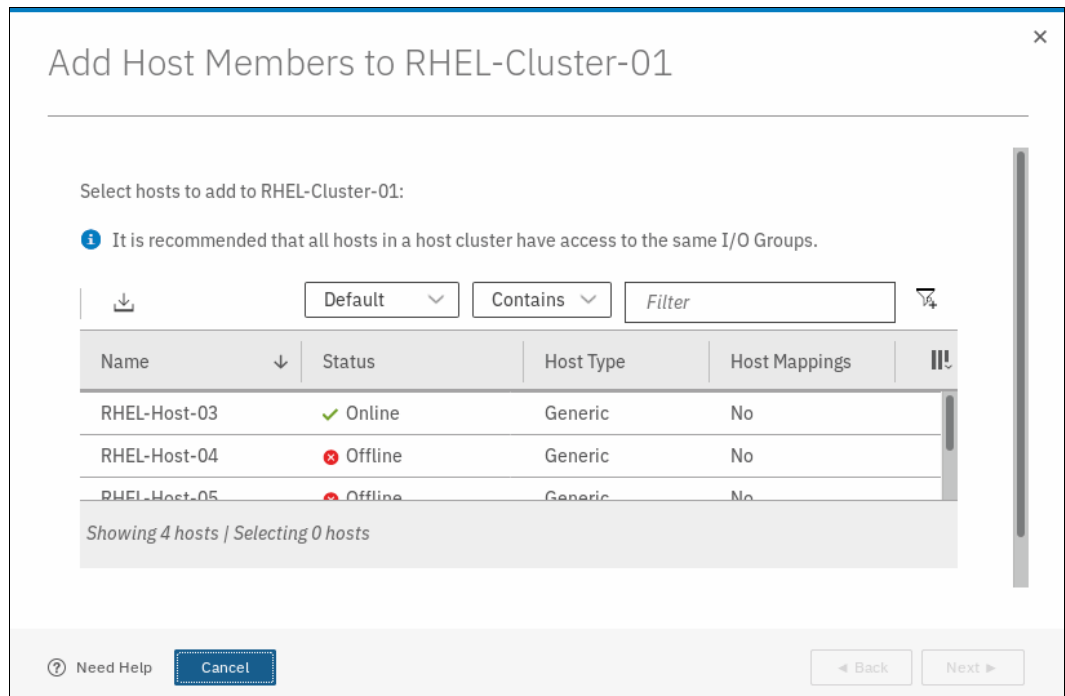


Figure 5-77 Host selection window

3. Select the wanted hosts, as shown in Figure 5-78.

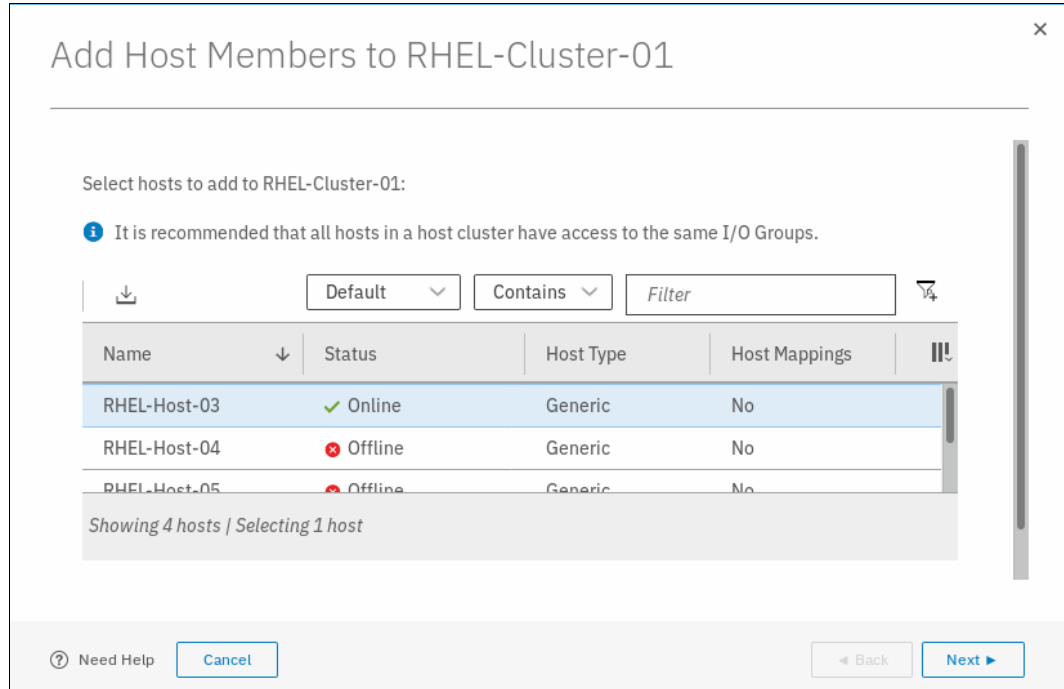


Figure 5-78 Selecting wanted hosts

4. Click **Next** and a summary of the hosts to be added is shown (see Figure 5-79).

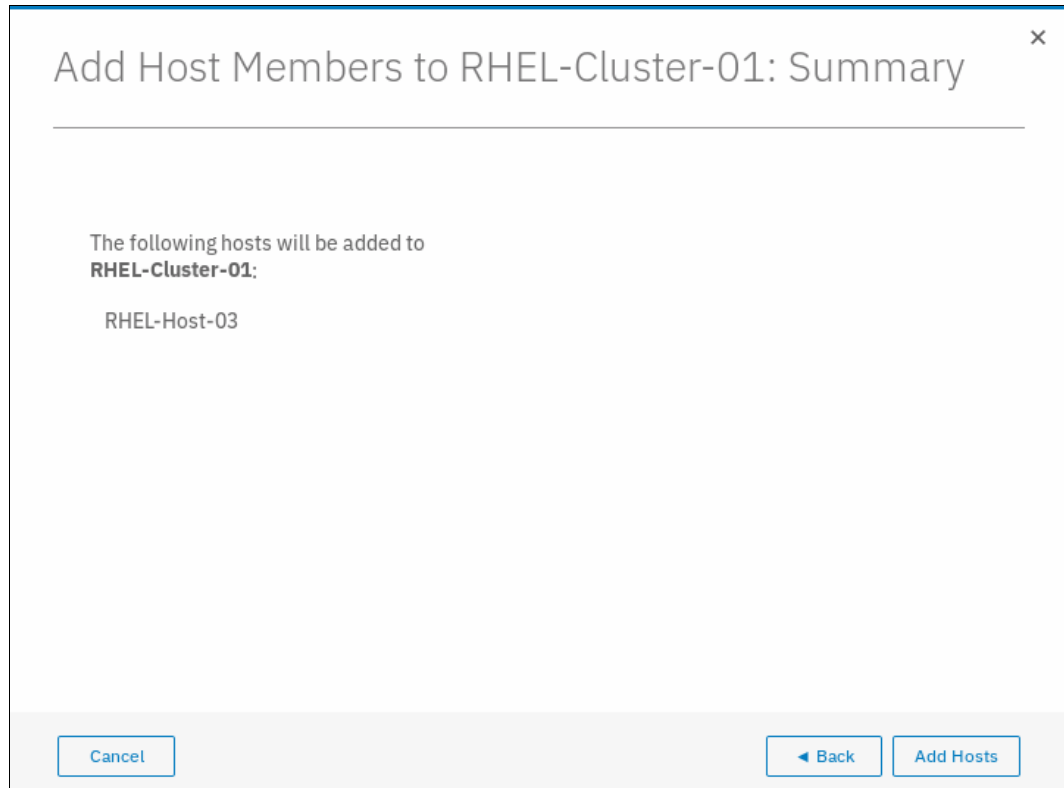


Figure 5-79 Summary of hosts

- Click **Add Hosts** and hosts are added to the host cluster definition. A task completion window opens. Click **Close**.

On the host cluster list, you can see the number of hosts that is contained in the cluster, as shown in Figure 5-80.

ID	Name	Status	Host Count	Mappings Count
0	ITSO-Host-Cluster-02	Offline	0	0
1	RHEL-Cluster-01	Online	3	0
2	ITSO-Host-Cluster-03	Offline	0	0

Figure 5-80 Host count on cluster

5.6.3 Listing a host cluster member

To list members of an existing host cluster, complete the following steps:

- From the **Host Clusters** window, select the host cluster for which you want to view members. Then, click **Actions**, as shown in Figure 5-81.

ID	Name	Status	Host Count	Mappings Count
0	ITSO-Host-C	Offline	0	0
1	RHEL-Cluste	Online	3	0
2	ITSO-Host-C	Offline	0	0

Figure 5-81 Host Clusters list

2. Click **View Hosts** to list the host cluster members, as shown in Figure 5-82.

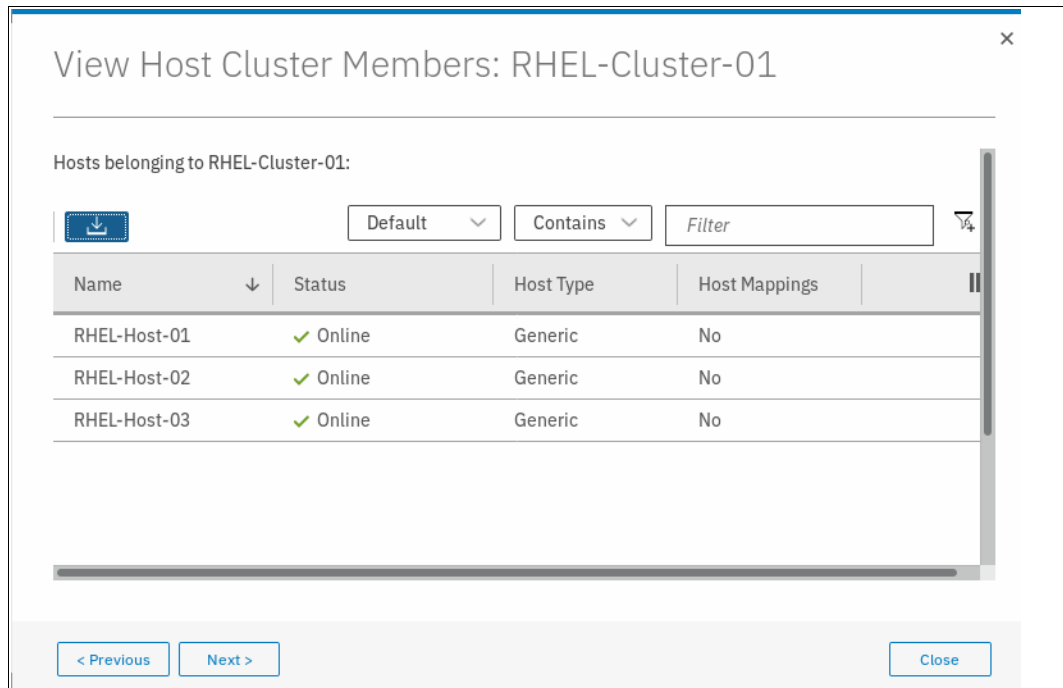


Figure 5-82 Listing host cluster members

Note: You can navigate among the host clusters in your system by using the Previous and Next buttons in this window.

3. Click **Close** and you are returned to the Host Clusters window, as shown in Figure 5-83.

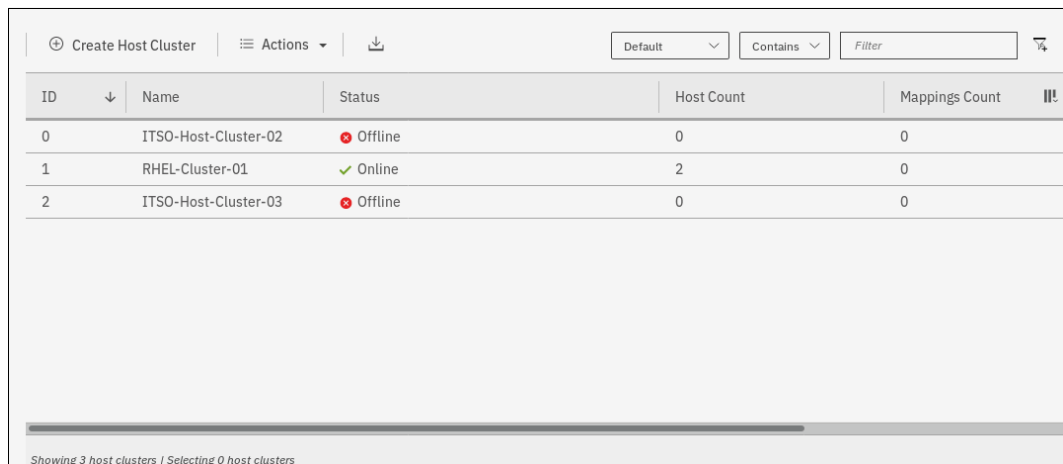
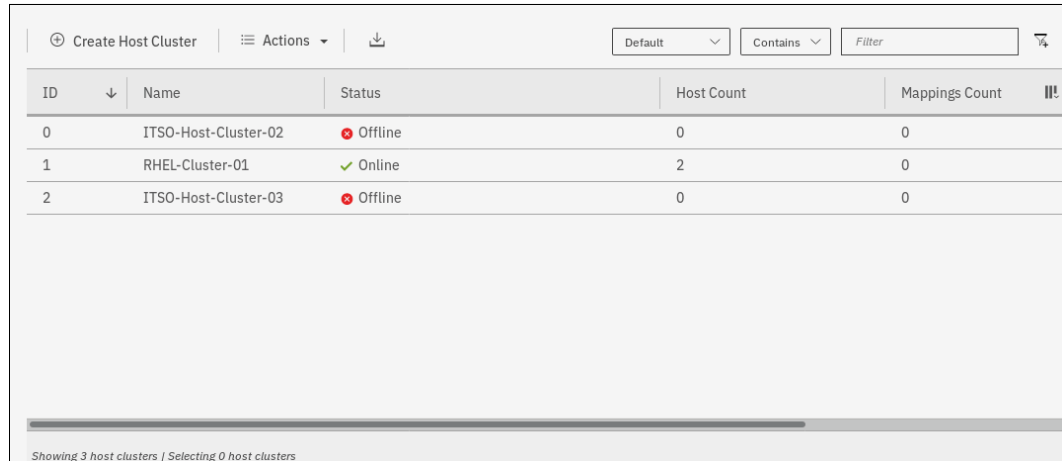


Figure 5-83 Host Clusters list

5.6.4 Assigning a volume to a Host Cluster

To assign a volume to a host cluster, complete the following steps:

1. From the **Host Clusters** window, select the wanted host cluster, as shown in Figure 5-84.

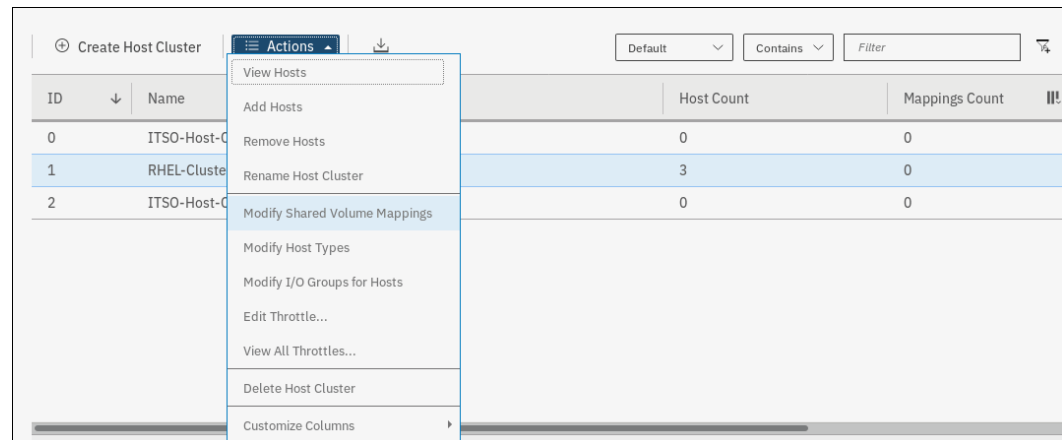


ID	Name	Status	Host Count	Mappings Count
0	ITSO-Host-Cluster-02	Offline	0	0
1	RHEL-Cluster-01	Online	2	0
2	ITSO-Host-Cluster-03	Offline	0	0

Showing 3 host clusters / Selecting 0 host clusters

Figure 5-84 Host Clusters list

2. Click **Actions**, and then, select **Modify Shared Volume Mappings**, as shown in Figure 5-85.



ID	Name	Host Count	Mappings Count
0	ITSO-Host-C	0	0
1	RHEL-Cluste	3	0
2	ITSO-Host-C	0	0

- View Hosts
- Add Hosts
- Remove Hosts
- Rename Host Cluster
- Modify Shared Volume Mappings
- Modify Host Types
- Modify I/O Groups for Hosts
- Edit Throttle...
- View All Throttles...
- Delete Host Cluster
- Customize Columns

Figure 5-85 Modify Shared Volume Mapping for Host Cluster

A window that lists the volumes that are mapped to the selected host cluster opens, as shown in Figure 5-86.

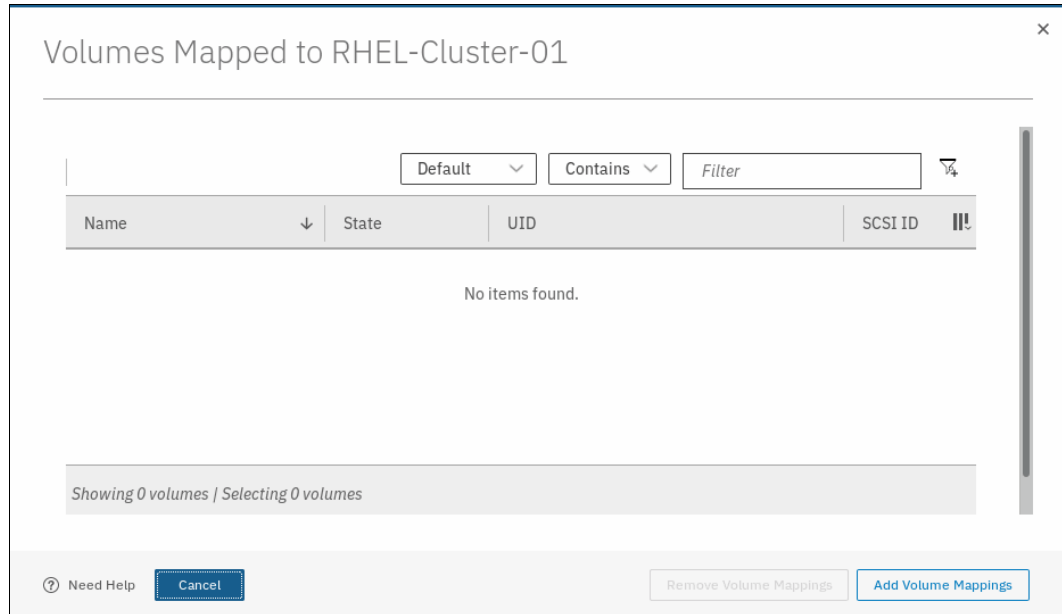


Figure 5-86 Volumes mapped to host cluster

Note: The list is empty if no volumes are mapped to the selected host cluster.

3. Click **Add Volume Mappings**. A window that lists the volumes that you can select to assign to a host cluster opens, as shown in Figure 5-87.

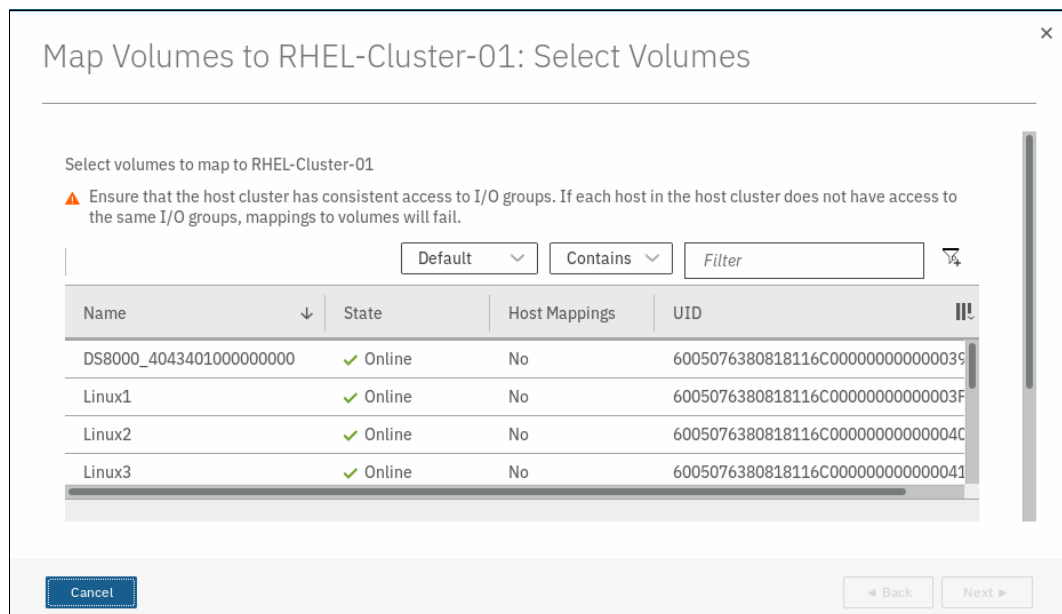


Figure 5-87 Volume list

4. Select the list of volumes to be mapped to the host cluster, as shown in Figure 5-88.

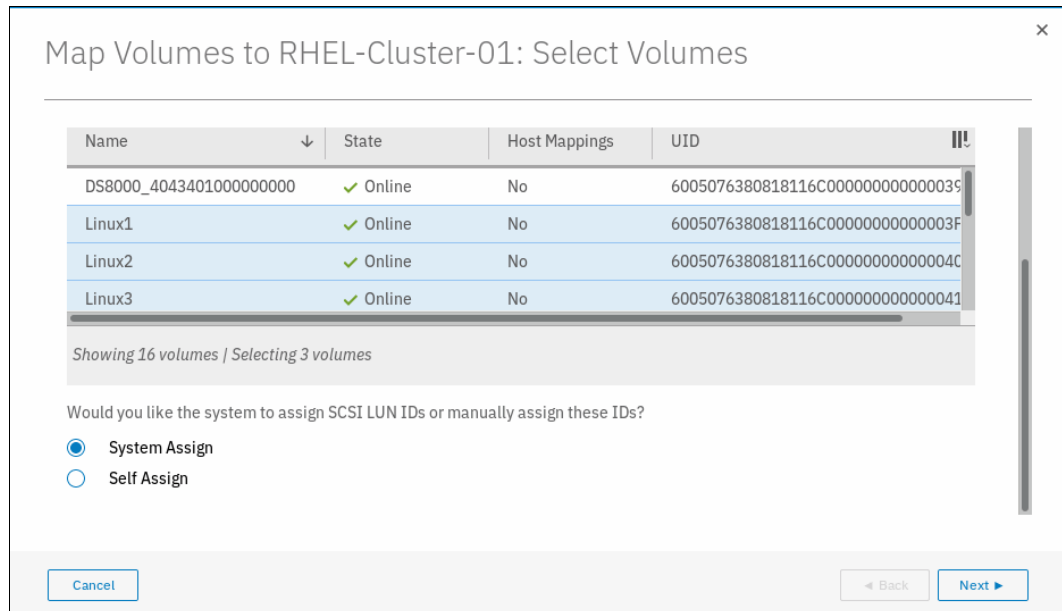


Figure 5-88 Selecting the list of volumes

5. You can choose that the SCSI LUN IDs be assigned automatically by the system or manually assign them. In this example, we chose the system-assigned SCSI LUN IDs.
6. Click **Next**. A summary window that includes the list of volumes to be mapped opens, as shown in Figure 5-89.

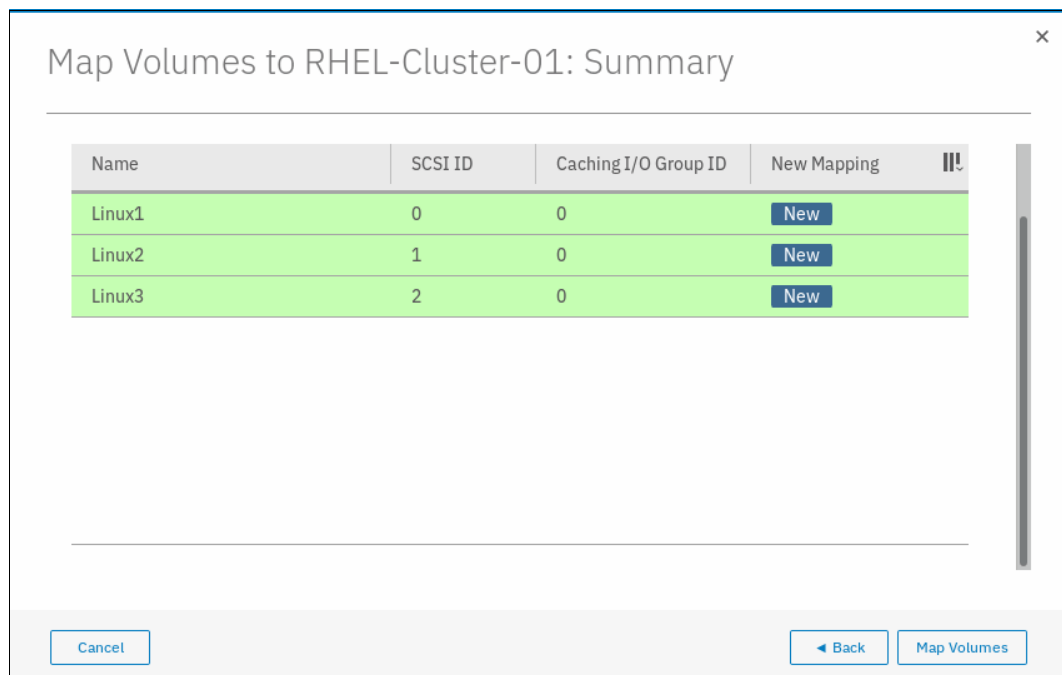


Figure 5-89 Summary of volumes to be mapped to host cluster

7. Click **Map Volumes** and a task completion window opens. Click **Close**.

5.6.5 Removing volume mapping from a host cluster

To remove a volume mapping from a host cluster, complete the following steps:

1. From the **Host Clusters** window, select the wanted host cluster, as shown in Figure 5-90.

ID	Name	Status	Host Count	Mappings Count
0	ITSO-Host-Cluster-02	Offline	0	0
1	RHEL-Cluster-01	Online	2	0
2	ITSO-Host-Cluster-03	Offline	0	0

Showing 3 host clusters | Selecting 0 host clusters

Figure 5-90 Host Clusters

2. Right-click the wanted host cluster and select **Modify Shared Volume Mappings**, as shown in Figure 5-91.

ID	Name	Status	Host Count	Mappings Count
0	ITSO-Host-C	Offline	0	0
1	RHEL-Cluste	Online	3	0
2	ITSO-Host-C	Offline	0	0

- View Hosts
- Add Hosts
- Remove Hosts
- Rename Host Cluster
- Modify Shared Volume Mappings
- Modify Host Types
- Modify I/O Groups for Hosts
- Edit Throttle...
- View All Throttles...
- Delete Host Cluster
- Customize Columns

Figure 5-91 Modify Shared Volume Mapping for Host Cluster

3. A window showing a list of volumes that are mapped to the host cluster opens, as shown in Figure 5-92.

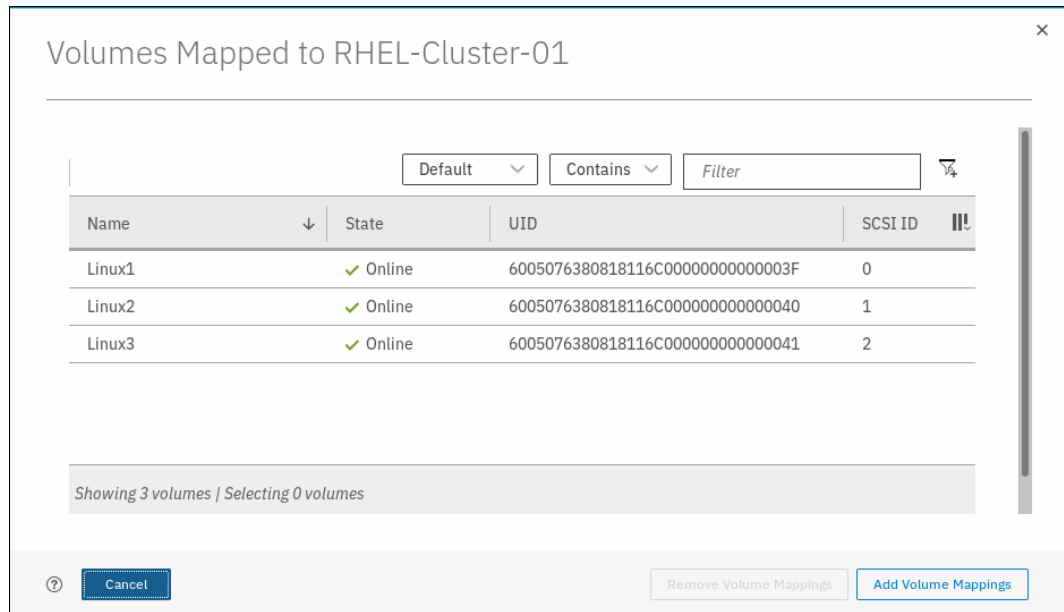


Figure 5-92 List of volumes mapped to a host cluster

4. Select the wanted volume to be unmapped, as shown in Figure 5-93.

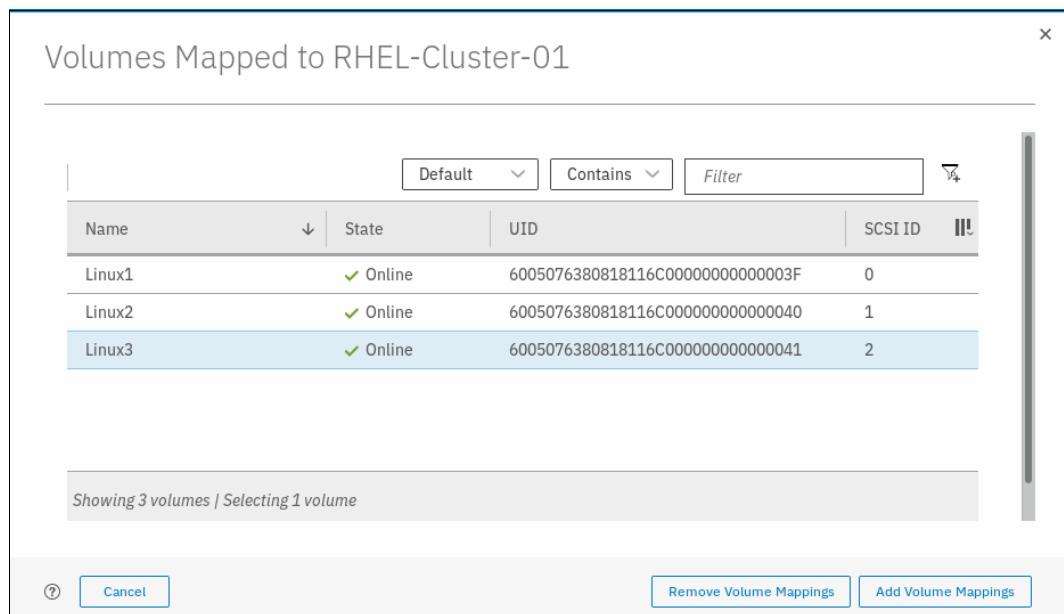


Figure 5-93 Selecting the volume to be unmapped

5. Click **Remove Volume Mappings**. A summary window in which the hosts are listed opens, as shown in Figure 5-94.

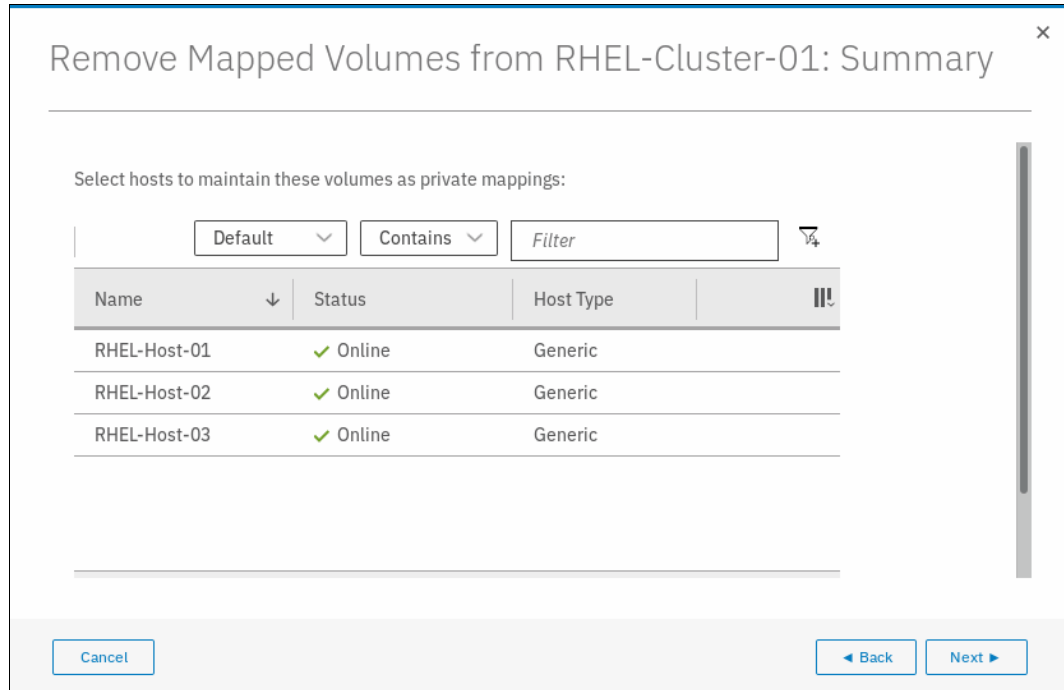


Figure 5-94 Host cluster member list

Note: At this point, you can select any host from the list to keep the private mapping between the selected host and the volume.

6. Click **Next**. A window opens, as shown in Figure 5-95.

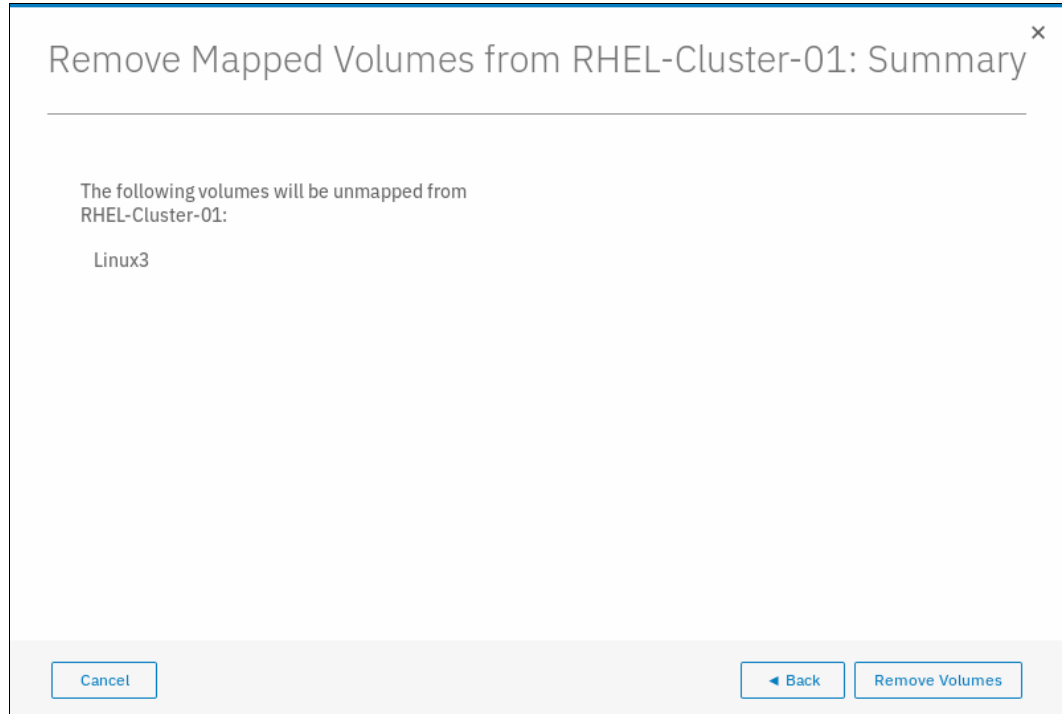


Figure 5-95 Volume unmap from host cluster

7. Click **Remove Volumes**. A task completion window opens. Click **Close**.

5.6.6 Removing a host cluster member

To remove a a host from the host cluster, complete the following steps:

1. From the **Host Clusters** window, select the wanted host cluster, as shown in Figure 5-96.

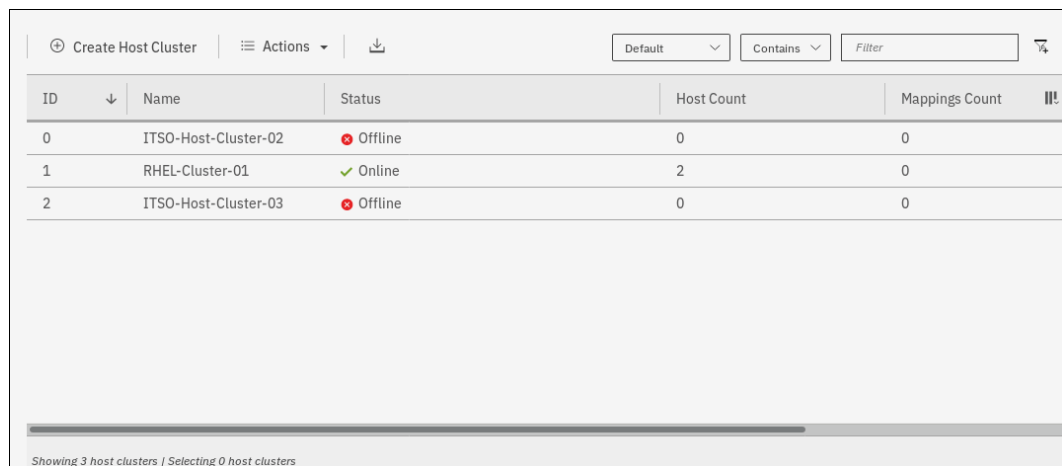


Figure 5-96 Host Clusters

2. Click **Actions** → **Remove Hosts**, as shown in Figure 5-97.

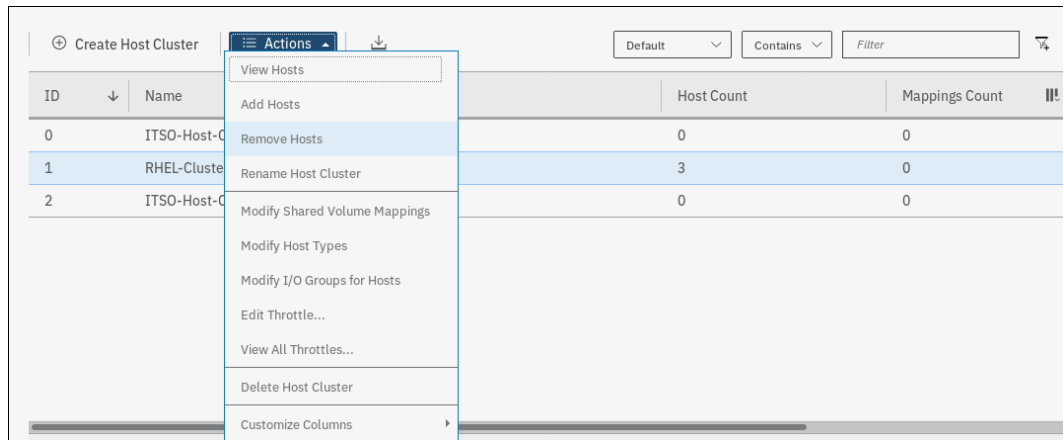


Figure 5-97 Removing hosts from host cluster

3. A window opens in which the hosts that are members of the host cluster are listed, as shown in Figure 5-98.

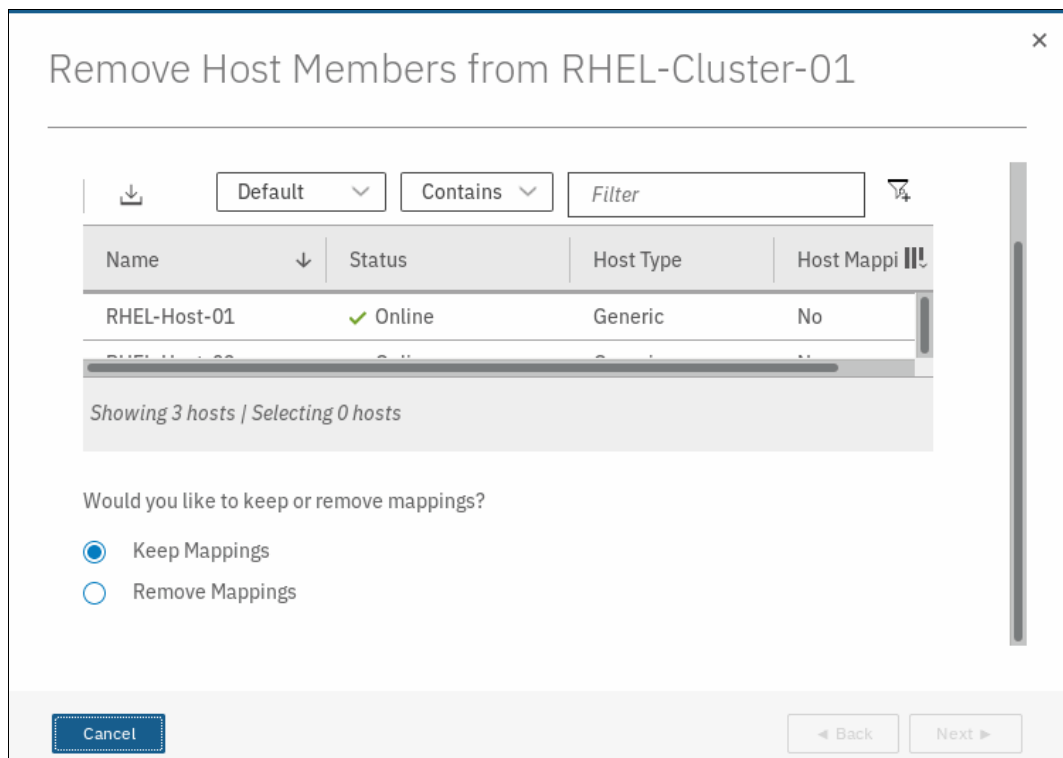


Figure 5-98 Host Cluster member selection

4. Select the host member that must be removed, as shown in Figure 5-99.

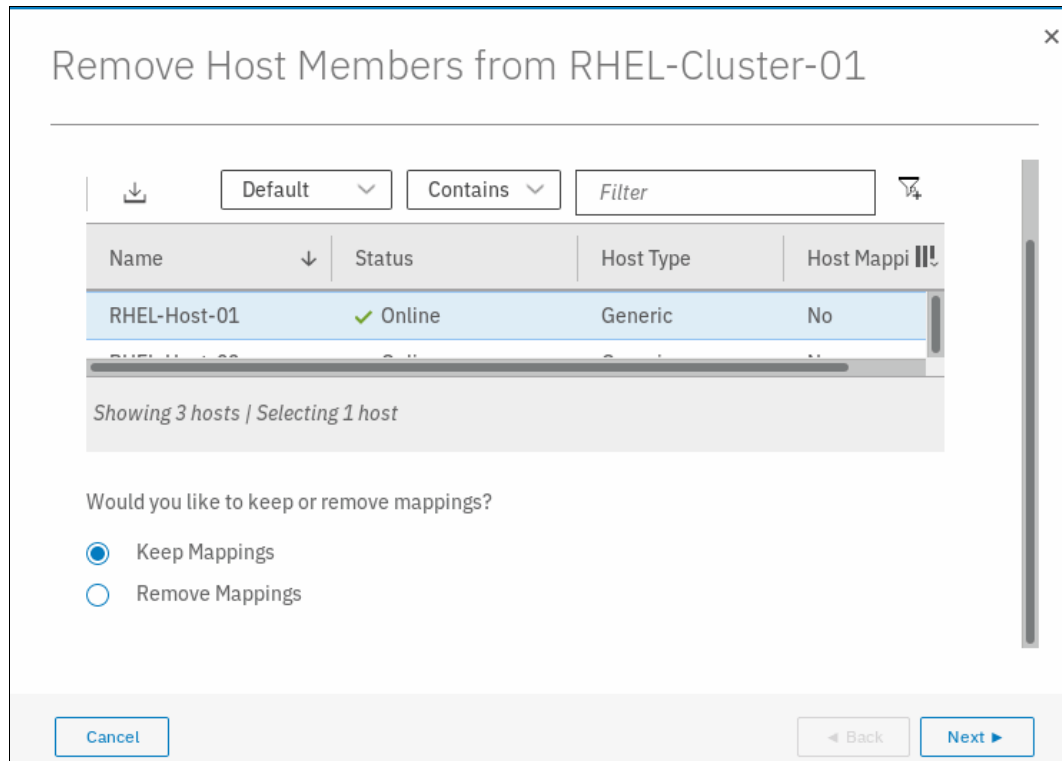


Figure 5-99 Host member selected

5. Select the option to indicate whether you want to keep the mappings after the host member is removed from the host cluster or to remove those mappings. In our example, we chose **Remove Mappings**, as shown in Figure 5-100.

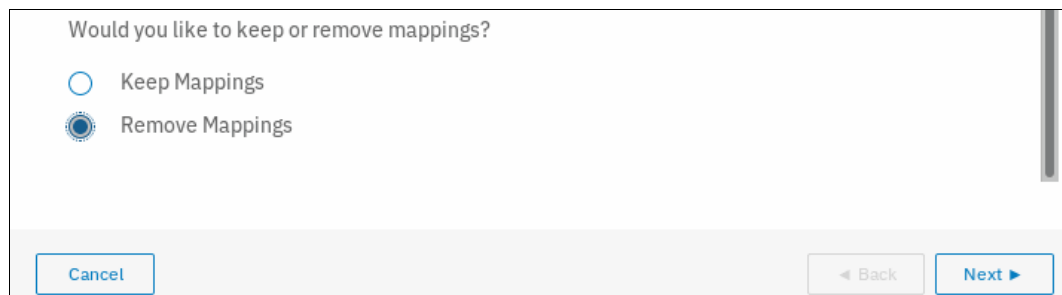


Figure 5-100 Mapping selection during removal of host member from host cluster

Note: Select **Keep Mappings** to retain all the shared mappings in the host cluster as private mappings for the selected hosts. Select **Remove Mappings** to remove all the shared mappings if the hosts that are being removed no longer require access to these volumes.

6. Click **Next**. A window in which the removal of the selected host member from host cluster is confirmed opens, as shown in Figure 5-101.

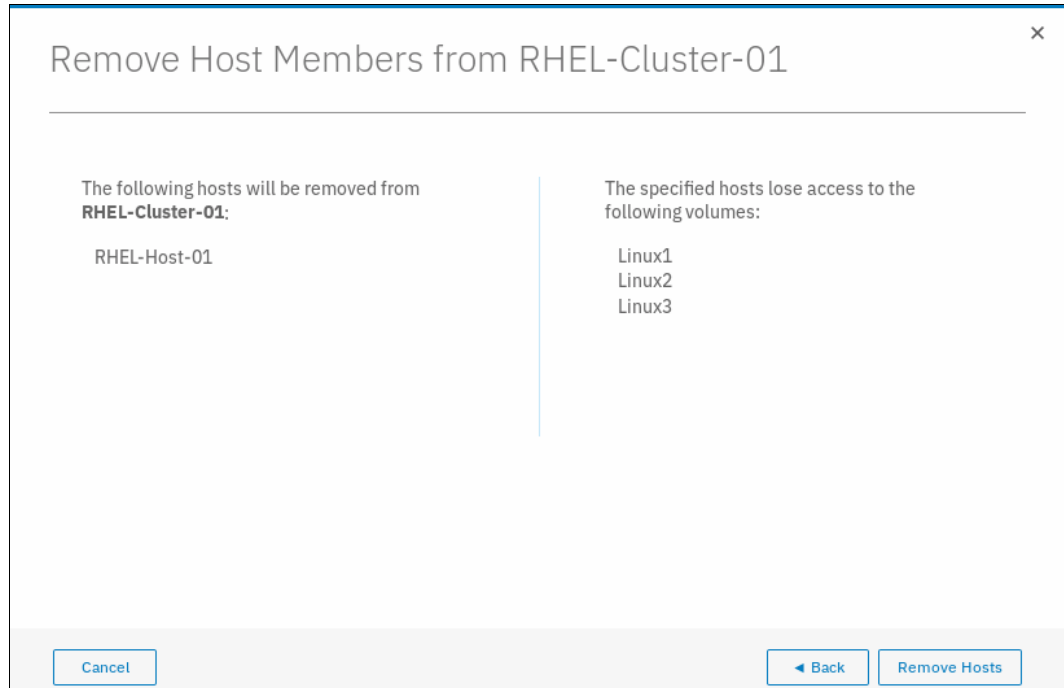


Figure 5-101 Confirmation window

7. Click **Remove Hosts**. A task completion window opens. Click **Close**.

5.6.7 Removing a host cluster

To remove a host cluster, complete the following steps:

1. From the **Host Clusters** window, select the wanted host cluster, as shown in Figure 5-102.

ID	Name	Status	Host Count	Mappings Count
0	ITSO-Host-Cluster-02	Offline	0	0
1	RHEL-Cluster-01	Online	2	0
2	ITSO-Host-Cluster-03	Offline	0	0

Figure 5-102 Host Clusters

2. Right-click the wanted host cluster and select **Delete Host Cluster**, as shown in Figure 5-103.

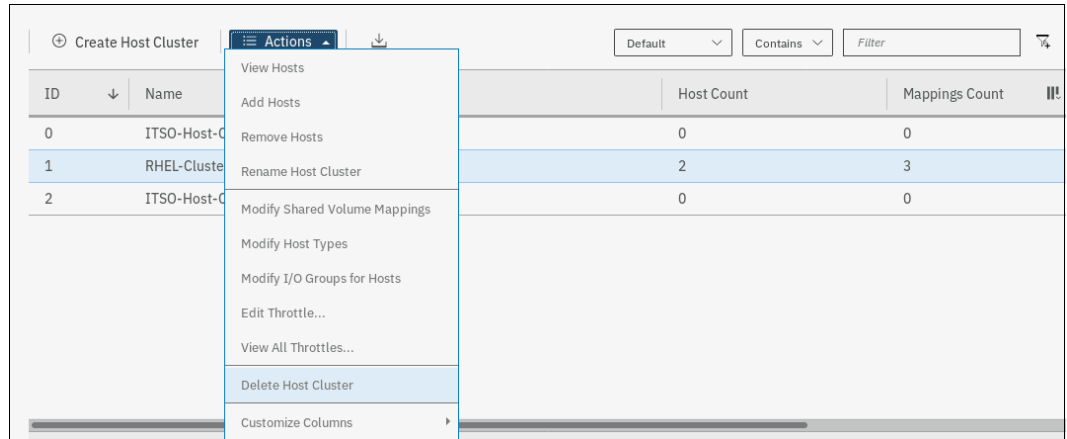


Figure 5-103 Delete Host Cluster selection

3. A window opens in which you are prompted to confirm the deletion of the host cluster object along with your selection for mappings, as shown in Figure 5-104.

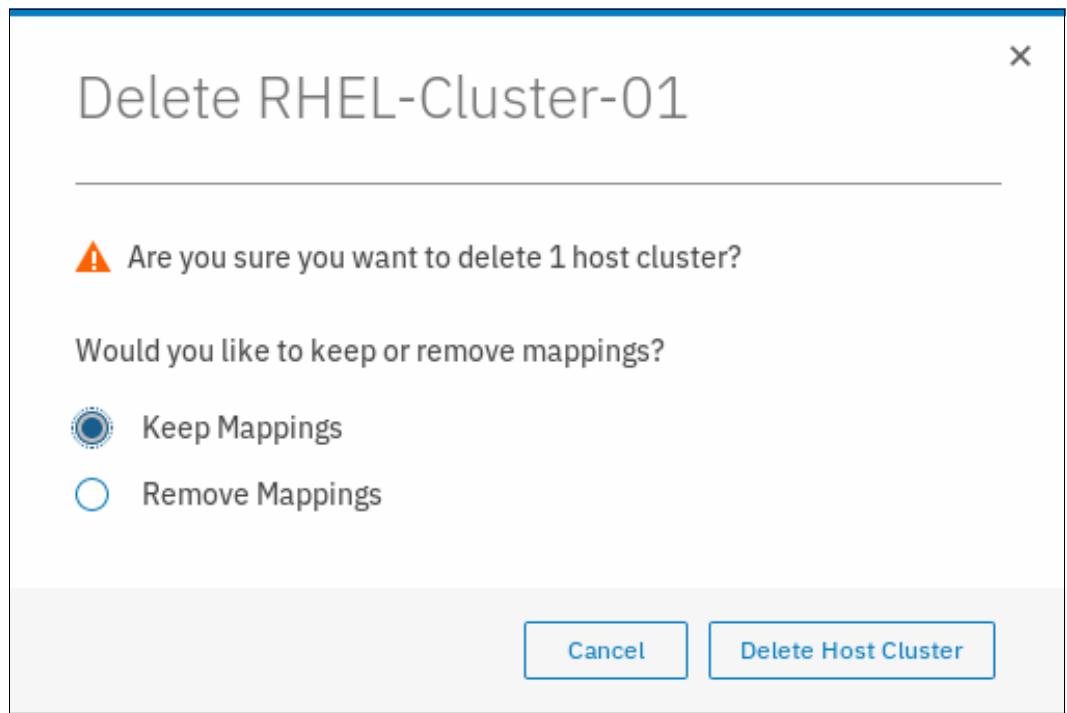


Figure 5-104 Confirm host cluster object deletion

4. Select the wanted mappings option. In our example, we chose **Remove Mappings**, as shown in Figure 5-105.

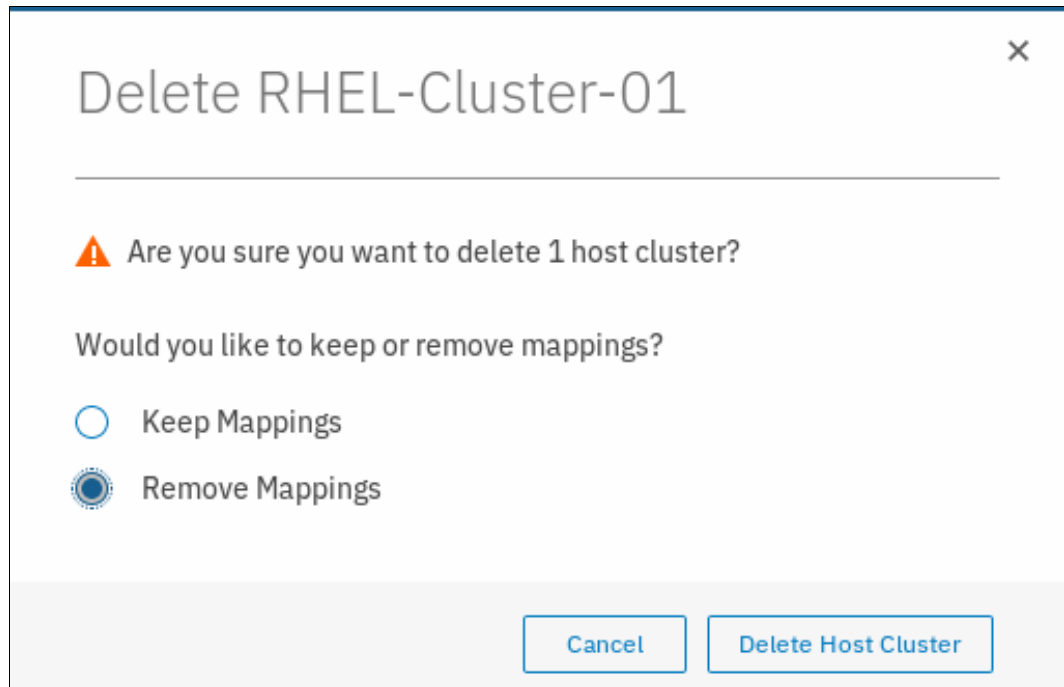


Figure 5-105 Remove mappings after host cluster deletion

Warning: Selecting **Remove Mappings** removes all of the shared mappings if the hosts that are being removed no longer require access to these volumes. Therefore, use this option with caution; otherwise, the servers that are part of the host cluster lose access to all shared volumes.

5. Click **Delete Host Cluster**. A task completion window opens. Click **Close**.

5.6.8 I/O throttling for hosts and host clusters

You can set a limit on the number of I/O operations that are accepted by the storage system. The limit is known as the throttling rate and is set in terms of I/O operations per second (IOPS) or bandwidth. I/O throttling is a way to achieve quality of service (QoS). I/O throttling is a mechanism to limit the volume of I/O that is processed by the storage system at various levels, which results in better distribution of storage system resources. I/O throttling is also referred to as *I/O governing*.

In IBM Spectrum Virtualize, no I/O throttling rate is set by default. However, I/O throttling can be set at the following levels:

- ▶ Host
- ▶ Host clusters
- ▶ Volume
- ▶ MDisk group

When I/O throttling is set, the I/O rate is limited by queuing I/Os if it exceeds preset limits. I/O throttling does not guarantee minimum performance. Internal I/Os, such as FlashCopy, Metro Mirror, and intra-cluster traffic, are not throttled.

I/O throttling can be beneficial in the following scenarios:

- ▶ An aggressive host that is hogging bandwidth of the Spectrum Virtualize system can be limited by a throttle. For example, allow restricted I/Os from a data mining server instead of an application server.
- ▶ Restrict a group of hosts by their throttles. For example, department A gets more bandwidth than department B.
- ▶ Each volume can have a throttle defined. For example, a backup volume can have less bandwidth than a production volume).

In this section, we describe the process of setting I/O throttle on defined hosts and host clusters.

Setting I/O throttle for host

To set I/O throttle on a defined host, complete the following steps:

1. Select the **Hosts** option under the **Hosts** section on the main window, as shown in Figure 5-106 on page 302.

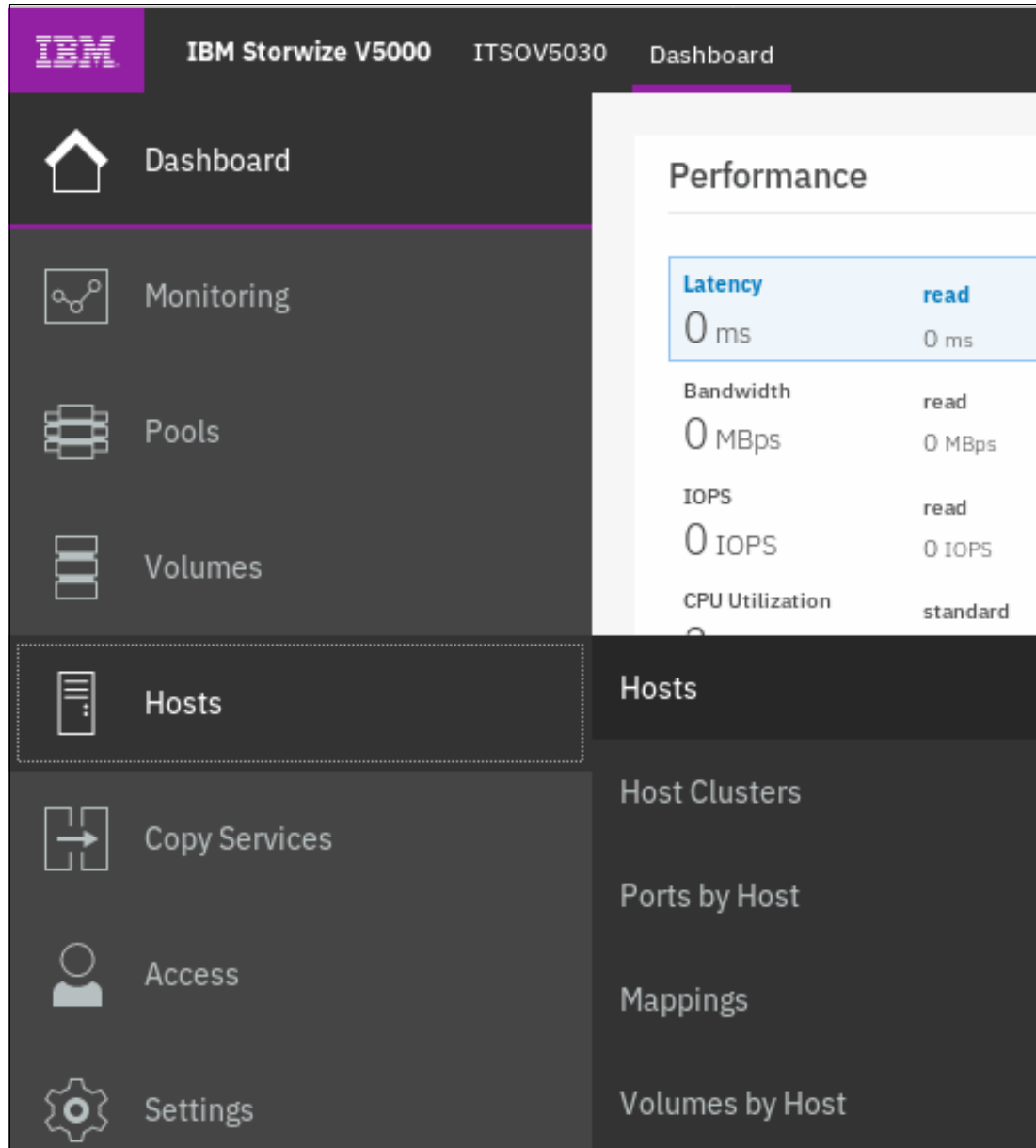


Figure 5-106 Hosts window

2. Select the wanted host. Click **Actions** → **Edit Throttle**, as shown in Figure 5-107.

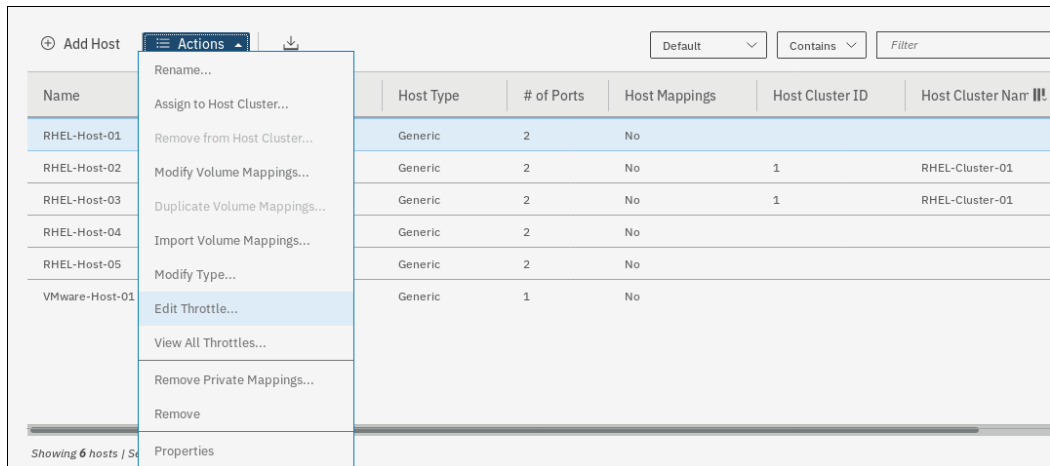


Figure 5-107 Selecting Edit Throttle option

3. Enter the wanted type of I/O throttle (IOPS or Bandwidth). In our example, we set up an IOPS throttle by entering the **IOPS limit**, as shown in Figure 5-108.

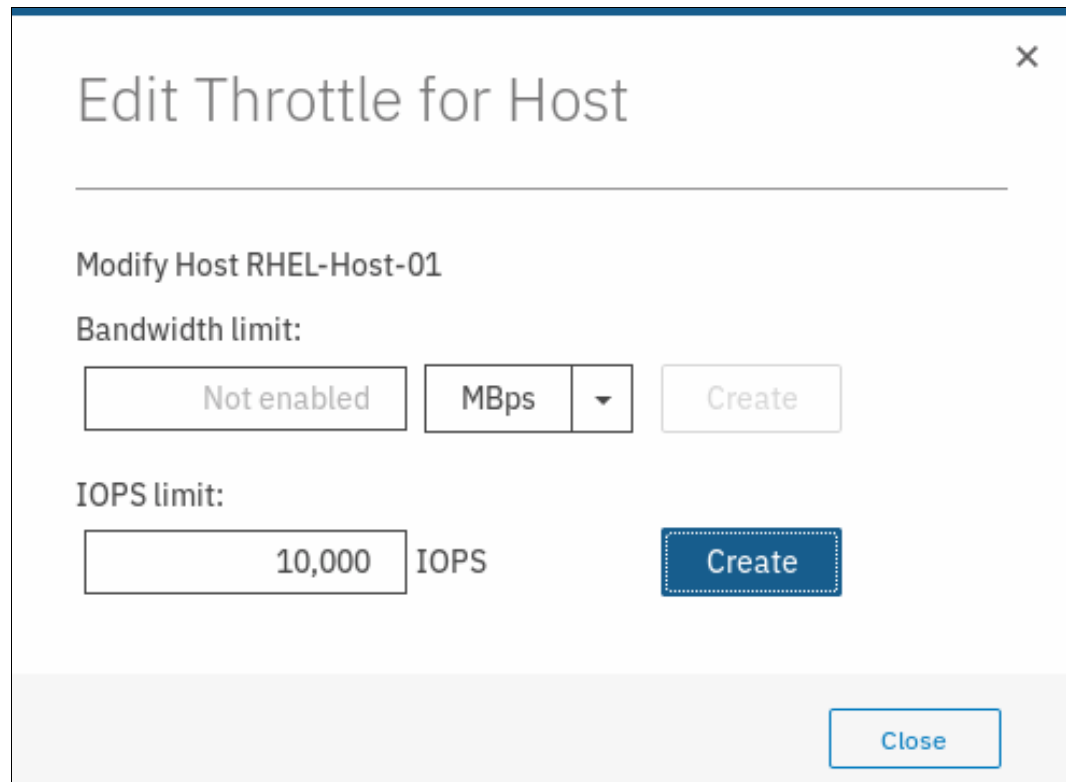


Figure 5-108 Setting IOPS throttle

Note: While defining a throttle for a host, you can define a throttle in terms of IOPS or bandwidth, but not both at the same time. If you want to have host throttle defined for both IOPS and bandwidth, you must define them one after the other.

4. Click **Create**. A task completion window opens. Click **Close**.

Setting I/O throttle for Host Cluster

To set I/O throttle on a defined Host Cluster, complete the following steps:

1. Select the **Hosts** option under the **Host Clusters** section on the main window, as shown in Figure 5-109.

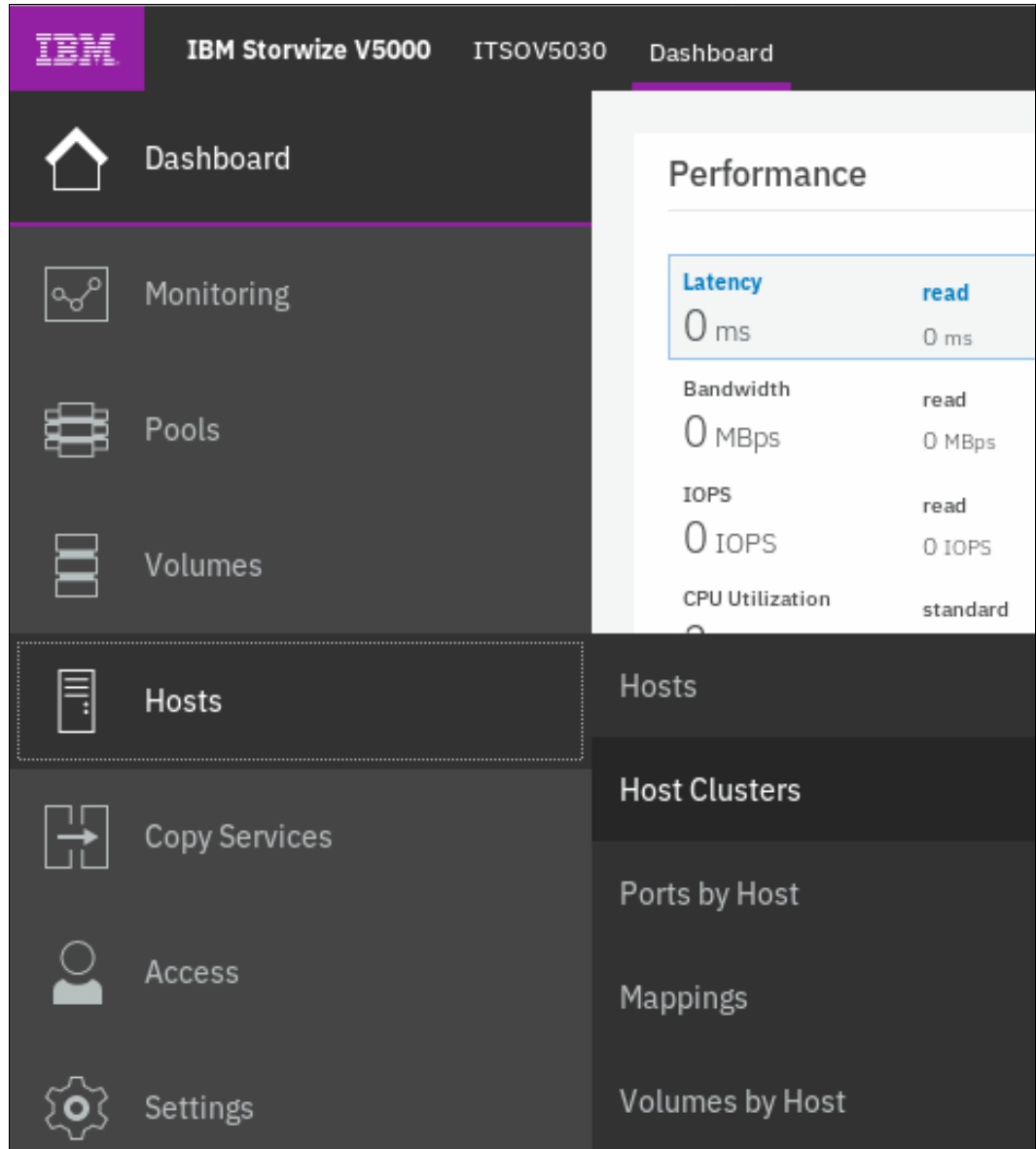


Figure 5-109 Host Clusters

2. Select the wanted host cluster. Click **Actions** → **Edit Throttle**, as shown in Figure 5-110.

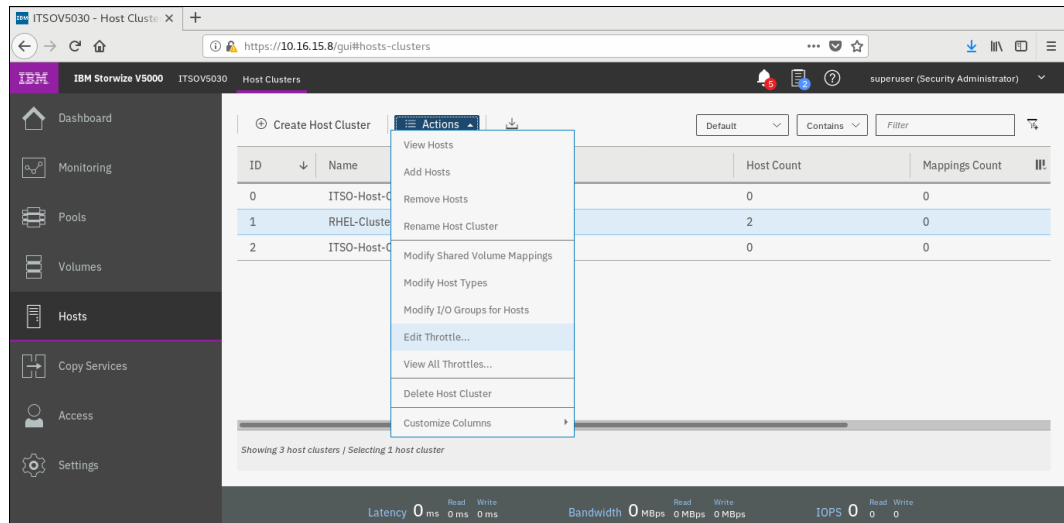


Figure 5-110 Selecting Edit Throttle option

3. Enter the wanted type of I/O throttle (IOPS or Bandwidth). In our example, we set up Bandwidth throttle by entering the **Bandwidth limit**, as shown in Figure 5-111.

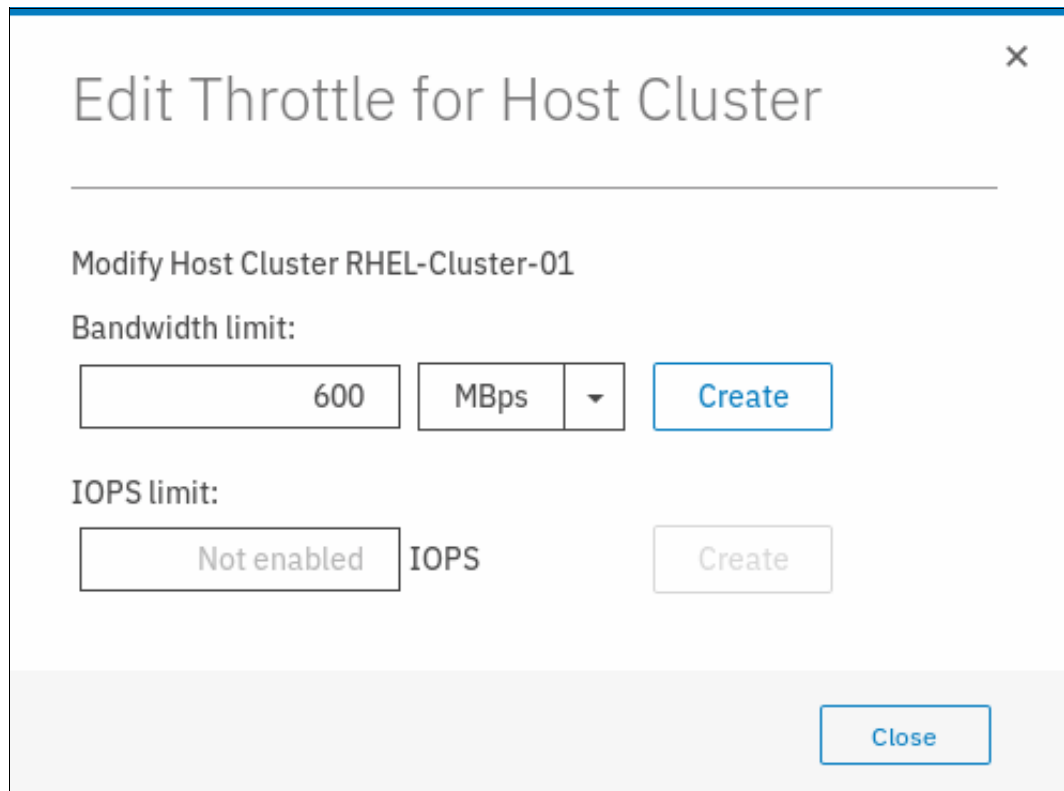


Figure 5-111 Setting bandwidth throttle

4. Click **Create**. A task completion window opens. Click **Close**.

Consider the following points when the I/O throttle is defined on host or host clusters:

- ▶ I/O throttle cannot be defined for host if it is a part of host cluster that has an I/O throttle defined at the host cluster level.
- ▶ If the host cluster does not have an I/O throttle defined, its member hosts can have their individual I/O throttles defined.
- ▶ The mdiskgrp (storage pool) throttles for child pool and parent pool work independently.
- ▶ If a volume has multiple copies, throttling is done for the mdiskgrp (storage pool) that is serving the primary copy. The throttling is not applicable for the secondary pool for mirrored volumes and stretched cluster implementations.
- ▶ A host cannot be added to a host cluster if both have their individual throttles defined.
- ▶ A seeding host that is used for creating a host cluster cannot have a host throttle defined for it.

5.7 Proactive Host Failover

During planned maintenance procedures of IBM Storwize V5000 Gen2, such as firmware upgrade or node canister replacement, the host multipathing driver might experience issues recovering paths to volumes. Most of the issues seem to stem from the fact that the path to a volume was removed without warning to the host and the host must detect that the path is now offline, time out, and redrive all I/Os that were in flight.

As a secondary issue, hosts often attempt to restart sending I/O to the preferred node canister when the ports come online, but the node canister might still be unpending. During this time, the node canister queues all incoming I/O until it has its configuration data and then volumes start coming online.

From a host perspective, the host completes the following tasks:

- ▶ Gracefully fail over I/O to different node canister if the node canister it is using is about to pend to avoid expensive error recovery as a result of losing active paths that can affect business applications.
- ▶ Send I/O to one node canister (even if the preferred node canister is unavailable) because round-robinning across multiple node canisters affects application I/O performance.

To minimize these issues, the Proactive Host Failover feature was added to Spectrum Virtualize since V7.8.1. Because of the Proactive Host Failover feature, the host multipath driver receives notification for node canister removal or node canister reboot during the planned maintenance procedures of IBM Storwize V5000 Gen2.

Because of the notification that is received, the host uses the partner node canister for I/O. Therefore, the I/O does not need to be timed-out and retried.

Note: Proactive Host Failover is an internal feature of IBM Spectrum Virtualize software. No changes are made to the CLI or GUI for this feature.

Consider the following points regarding Proactive Host Failover:

- ▶ When an IBM Spectrum Virtualize system knows a node canister is about to go down, it raises unit attentions to try to trigger host path failovers to surviving node canisters.
- ▶ Requires the host to be tracking the preferred paths; usually requires ALUA support.
- ▶ Delays the failback when a node canister is online until the node canister confirmed it is ready to service I/O.
- ▶ Works with and without NPIV enabled and for all connection protocols that IBM Spectrum Virtualize system supports.
- ▶ Adjusts preferred paths when node canisters are unavailable so that the IBM Spectrum Virtualize system always is presenting a set of preferred paths.



Volume configuration

A volume is a logical disk that is provisioned out of a storage pool and recognized by a host with a unique identifier (UID) field and a parameter list.

The first part of this chapter provides a brief overview of IBM Spectrum Virtualize volumes, the classes of volumes available, and the topologies with which they are associated. It also provides an overview of advanced customization available.

The second part describes how to create volumes by using the GUI's Quick and Advanced volume creation menus, and shows you how to map these volumes to defined hosts.

The third part provides an introduction to the new volume manipulation commands, which are designed to facilitate the creation and administration of volumes used for IBM HyperSwap topology.

Note: For more information about advanced host and volume administration, such as volume migration and creating volume copies, see Chapter 10, “Copy Services” on page 465.

This chapter includes the following topics:

- ▶ 6.1, “Introduction to volumes” on page 310
- ▶ 6.2, “Create Volumes menu” on page 321
- ▶ 6.3, “Creating volumes by using the Volume Creation” on page 325
- ▶ 6.4, “Mapping a volume to the host” on page 329
- ▶ 6.5, “Creating Custom volumes” on page 331
- ▶ 6.6, “HyperSwap and the mkvolume command” on page 338
- ▶ 6.7, “Mapping Volumes to Host after volume creation” on page 342
- ▶ 6.8, “Migrating a volume to another storage pool” on page 347
- ▶ 6.9, “Migrating volumes using the volume copy feature” on page 350
- ▶ 6.10, “I/O throttling” on page 353

6.1 Introduction to volumes

A volume is a logical disk that the system presents to attached hosts. For an IBM Spectrum Virtualize system, the volume that is presented is from a virtual disk (VDisk).

A volume is a discrete area of usable storage that is virtualized, using IBM Spectrum Virtualize code, from storage area network (SAN) storage that is managed by the IBM Spectrum Virtualize cluster. The term *virtual* is used because the presented volume does not necessarily exist on a single physical entity.

Volumes have the following characteristics or attributes:

- ▶ Volumes can be created and deleted.
- ▶ Volumes can be resized (expanded or shrunk).
- ▶ Volume extents can be migrated at run time to another MDisk or storage pool.
- ▶ Volumes can be created as fully allocated or thin-provisioned. A conversion from a fully allocated to a thin-provisioned volume and vice versa can be done at run time.
- ▶ Volumes can be stored in multiple storage pools (mirrored) to make them resistant to disk subsystem failures or to improve the read performance.
- ▶ Volumes can be mirrored synchronously or asynchronously for longer distances. An IBM Spectrum Virtualize system can run active volume mirrors to a maximum of three other IBM Spectrum Virtualize systems, but not from the same volume.
- ▶ Volumes can be copied by using FlashCopy. Multiple snapshots and quick restore from snapshots (reverse FlashCopy) are supported.
- ▶ Volumes can be compressed.
- ▶ Volumes can be virtual. The system supports VMware vSphere Virtual Volumes, sometimes referred to as VVols, which allow VMware vCenter to manage system objects, such as volumes and pools. The system administrator can create these objects and assign ownership to VMware administrators to simplify management of these objects.

Note: A managed disk (MDisk) is a logical unit of physical storage. MDisks are Redundant Arrays of Independent Disks (RAID) from internal storage, or external physical disks that are presented as a single logical disk on the SAN. Each MDisk is divided into several extents, which are numbered (from 0) sequentially from the start to the end of the MDisk. The extent size is a property of the storage pools that the MDisks are added to.

MDisks are not visible to host systems.

Volumes have two major modes: Managed mode and image mode. Managed mode volumes have two policies: The sequential policy and the striped policy. Policies define how the extents of a volume are allocated from a storage pool.

The *type* attribute of a volume defines the allocation of extents that make up the volume copy:

- ▶ A striped volume contains a volume copy that has one extent allocated in turn from each MDisk that is in the storage pool. This option is the default, but you can also supply a list of MDisks to use as the stripe set, as shown in Figure 6-1.

Attention: By default, striped volume copies are striped across all MDisks in the storage pool. If some of the MDisks are smaller than others, the extents on the smaller MDisks are used up before the larger MDisks run out of extents. Manually specifying the stripe set in this case might result in the volume copy not being created.

If you are unsure if sufficient free space exists to create a striped volume copy, select one of the following options:

- ▶ Check the free space on each MDisk in the storage pool by using the `lsfreeextents` command.
- ▶ Allow the system automatically create the volume copy by not supplying a specific stripe set.

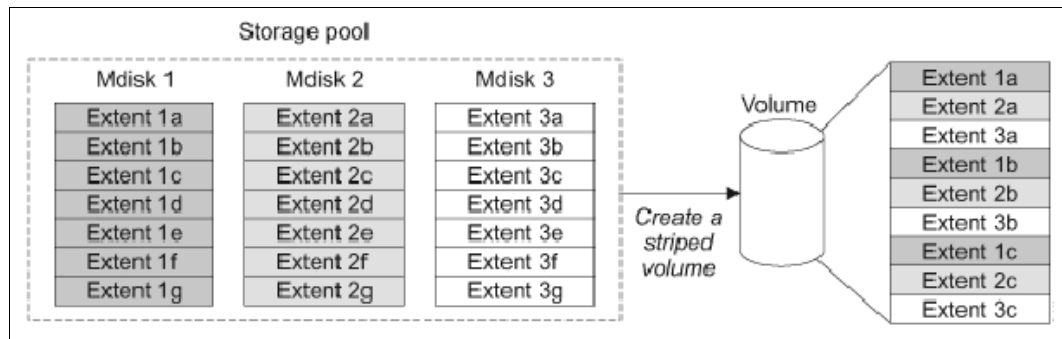


Figure 6-1 Striped extent allocation

- ▶ A *sequential* volume contains a volume copy that has extents that are allocated sequentially on one MDisk.
- ▶ *Image-mode* volumes are a special type of volume that has a direct relationship with one MDisk.

6.1.1 Image mode volumes

Image mode volumes are used to migrate LUNs that were previously mapped directly to host servers over to the control of the IBM Spectrum Virtualize system. Image mode provides a one-to-one mapping between the logical block addresses (LBAs) between a volume and an MDisk. Image mode volumes have a minimum size of one block (512 bytes) and always occupy at least one extent.

An image mode MDisk is mapped to one, and only one, image mode volume.

The volume capacity that is specified must be equal to the size of the image mode MDisk. When you create an image mode volume, the specified MDisk must be in unmanaged mode and must not be a member of a storage pool. The MDisk is made a member of the specified storage pool (Storage Pool_IMG_XXX) as a result of creating the image mode volume.

IBM Spectrum Virtualize also supports the reverse process, in which a managed mode volume can be migrated to an image mode volume. If a volume is migrated to another MDisk, it is represented as being in managed mode during the migration. It also is only represented as an image mode volume after it reaches the state where it is a straight-through mapping.

An image mode MDisk is associated with exactly one volume. If the (image mode) MDisk is not a multiple of the MDisk Group's extent size, the last extent is partial (not filled). An image mode volume is a pass-through one-to-one map of its MDisk. It cannot be a quorum disk and it does not have any metadata extents that are assigned to it from the IBM Spectrum Virtualize system. Managed or image mode MDisks are always members of a storage pool.

It is a preferred practice to put image mode MDisks in a dedicated storage pool and use a special name for it (for example, Storage Pool_IMG_XXX). The extent size that is chosen for this specific storage pool must be the same as the extent size into which you plan to migrate the data. All of the IBM Spectrum Virtualize copy services functions can be applied to image mode disks (see Figure 6-2).

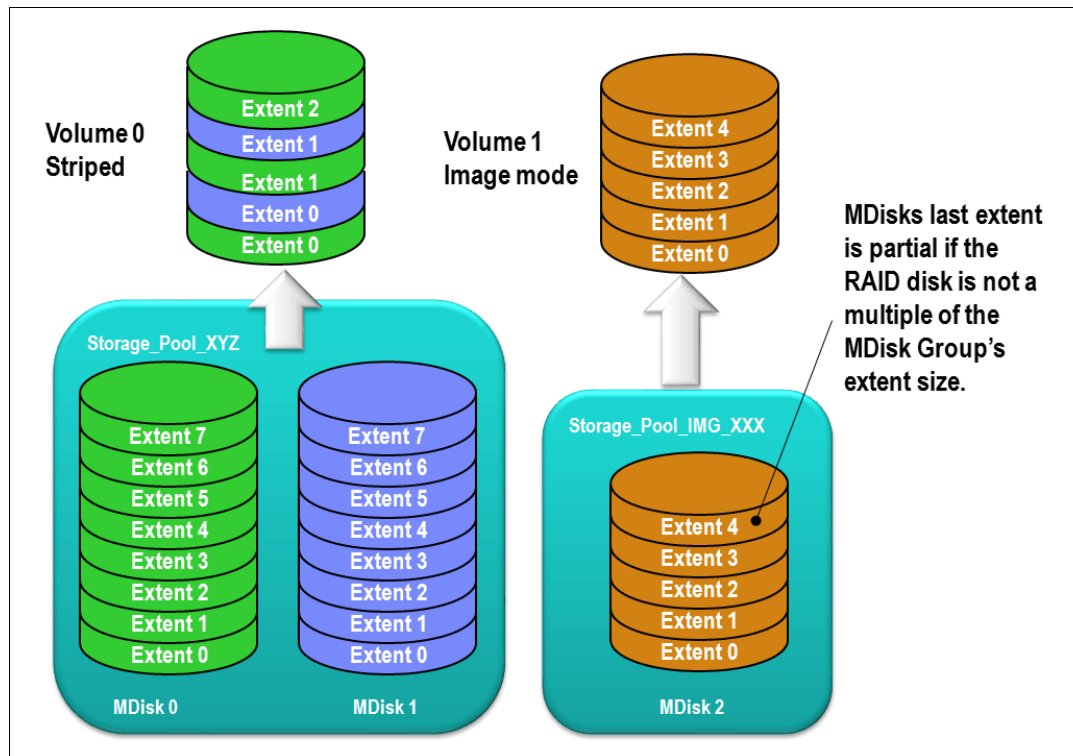


Figure 6-2 Image mode volume versus striped volume

6.1.2 Managed mode volumes

Volumes that are operating in managed mode provide a full set of virtualization functions. Within a storage pool, the IBM Spectrum Virtualize supports an arbitrary relationship between extents on (managed mode) volumes and extents on MDisks. Each volume extent maps to exactly one MDisk extent.

Figure 6-3 shows this mapping. It also shows a volume that consists of several extents that are shown as V0 - V7. Each of these extents is mapped to an extent on one of the MDisks: A, B, or C. The mapping table stores the details of this indirection.

Several of the MDisk extents are unused. No volume extent maps to them. These unused extents are available for use in creating volumes, migration, expansion, and so on.

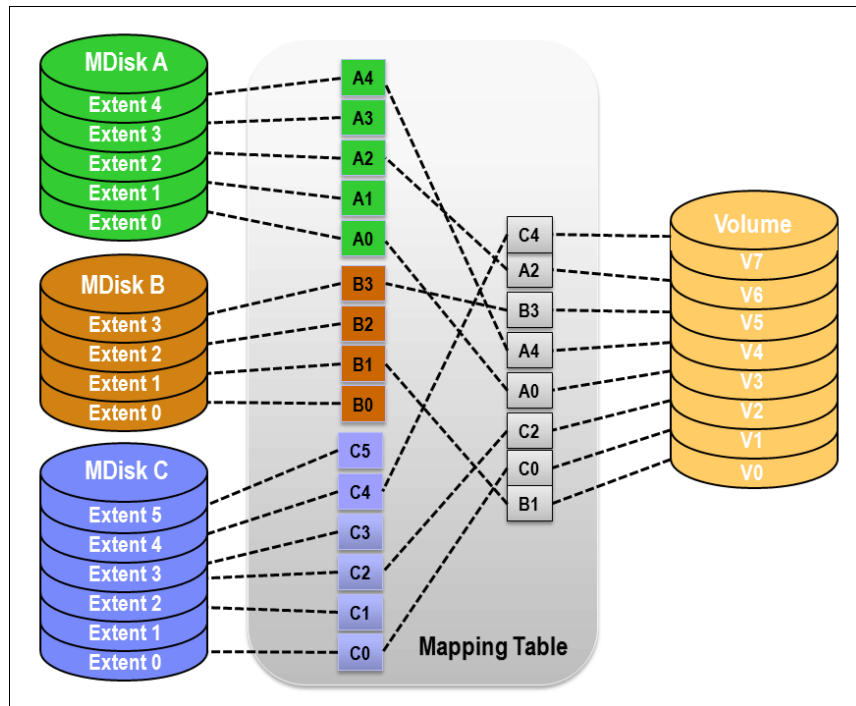


Figure 6-3 Simple view of block virtualization

The allocation of a specific number of extents from a specific set of MDisks is performed by the following algorithm:

- ▶ If the set of MDisks from which to allocate extents contains more than one MDisk, extents are allocated from MDisks in a round-robin fashion.
- ▶ If an MDisk has no free extents when its turn arrives, its turn is missed and the round-robin moves to the next MDisk in the set that has a free extent.

When a volume is created, the first MDisk from which to allocate an extent is chosen in a pseudo-random way rather than by choosing the next disk in a round-robin fashion. The pseudo-random algorithm avoids the situation where the *striping effect* that is inherent in a round-robin algorithm that places the first extent for many volumes on the same MDisk.

Placing the first extent of several volumes on the same MDisk can lead to poor performance for workloads that place a large I/O load on the first extent of each volume, or that create multiple sequential streams.

6.1.3 Cache mode for volumes

It is also possible to define the cache characteristics of a volume. Under normal conditions, a volume's read and write data is held in the cache of its preferred node, with a mirrored copy of write data that is held in the partner node of the same I/O Group. However, it is possible to create a volume with cache disabled. This setting means that the I/Os are passed directly through to the back-end storage controller rather than being held in the node's cache.

Having cache-disabled volumes makes it possible to use the native copy services in the underlying RAID array controller for MDisks (LUNs) that are used as the IBM Spectrum Virtualize image mode volumes. The use of IBM Spectrum Virtualize Copy Services rather than the underlying disk controller copy services gives better results.

Cache characteristics of a volume can have any of the following settings:

► `readwrite`

All read and write I/O operations that are performed by the volume are stored in cache. This mode is the default cache mode for all volumes.

► `readonly`

All read I/O operations that are performed by the volume are stored in cache.

► `disabled`

All read and write I/O operations that are performed by the volume are not stored in cache. Under normal conditions, a volume's read and write data is held in the cache of its preferred node, with a mirrored copy of write data that is held in the partner node of the same I/O Group. With cache disabled volume, the I/Os are passed directly through to the back-end storage controller rather than being held in the node's cache.

Note: Having cache-disabled volumes makes it possible to use the native copy services in the underlying RAID array controller for MDisks (LUNs) that are used as IBM Spectrum Virtualize image mode volumes. Contact IBM Support before turning off the cache for volumes in production environment to avoid any performance degradation.

6.1.4 Mirrored volumes

The mirrored volume feature provides a simple RAID 1 function, so a volume has two physical copies of its data. This approach enables the volume to remain online and accessible, even if one of the MDisks sustains a failure that causes it to become inaccessible.

The two copies of the volume often are allocated from separate storage pools or by using image-mode copies. The volume can participate in FlashCopy and remote copy relationships. It is serviced by an I/O Group, and has a preferred node.

Each copy is not a separate object and cannot be created or manipulated, except in the context of the volume. Copies are identified through the configuration interface with a copy ID of their parent volume. This copy ID can be 0 or 1.

This feature provides a point-in-time copy function that is achieved by “splitting” a copy from the volume. However, the mirrored volume feature does not address other forms of mirroring that are based on remote copy, which is sometimes called *IBM HyperSwap*, that mirrors volumes across I/O Groups or clustered systems. It is also not intended to manage mirroring or remote copy functions in back-end controllers.

Figure 6-4 provides an overview of volume mirroring.

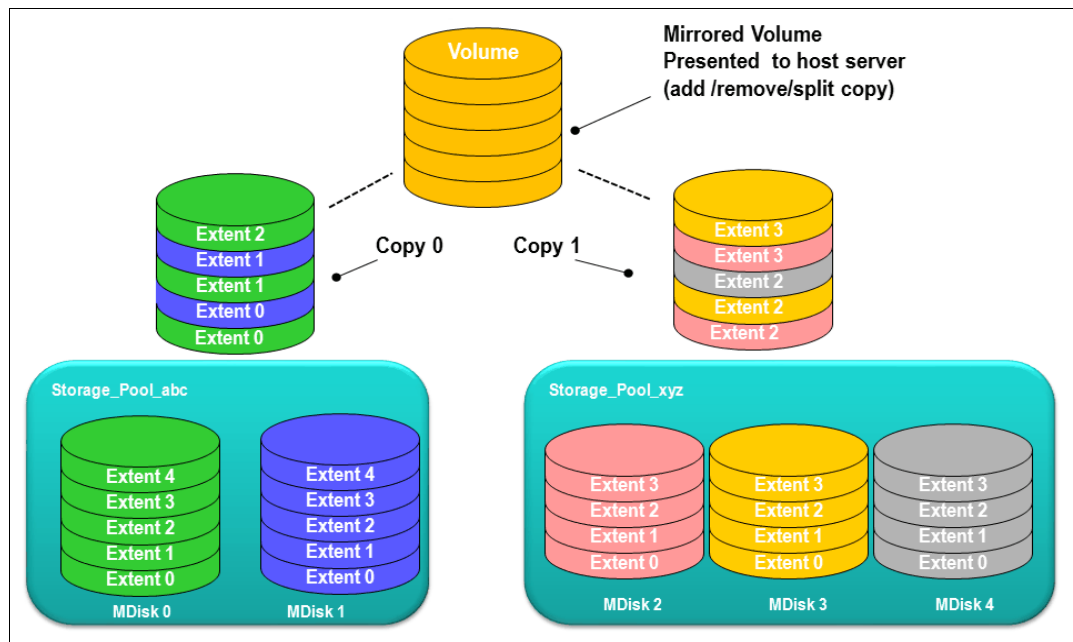


Figure 6-4 Volume mirroring overview

A second copy can be added to a volume with a single copy or removed from a volume with two copies. Checks prevent the accidental removal of the only remaining copy of a volume. A newly created, unformatted volume with two copies initially has the two copies in an out-of-synchronization state. The primary copy is defined as “fresh” and the secondary copy is defined as “stale.”

The synchronization process updates the secondary copy until it is fully synchronized. This update is done at the default *synchronization rate* or at a rate that is defined when the volume is created or modified. The synchronization status for mirrored volumes is recorded on the quorum disk.

If a two-copy mirrored volume is created with the **format** parameter, both copies are formatted in parallel, but the volume remains online while copies are being formatted.

If mirrored volumes are expanded or shrunk, all of their copies are also expanded or shrunk.

If it is known that MDisk space (which is used for creating copies) is already formatted or if the user does not require read stability, a `no_synchronization` option can be selected that declares the copies as synchronized (even when they are not).

To minimize the time that is required to resynchronize a copy that is out of sync, only the 256 kibibyte (KiB) grains that were written to since the synchronization was lost are copied. This approach is known as an *incremental synchronization*. Only the changed grains must be copied to restore synchronization.

Important: An unmirrored volume can be migrated from one location to another by adding a second copy to the wanted destination, waiting for the two copies to synchronize, and then removing the original copy 0. This operation can be stopped at any time. The two copies can be in separate storage pools with different extent sizes.

When there are two copies of a volume, one copy is known as the *primary copy*. If the primary is available and synchronized, reads from the volume are directed to it. The user can select the primary when the volume is created or can change it later.

Placing the primary copy on a high-performance controller maximizes the read performance of the volume.

Write I/O operations data flow with a mirrored volume

For write I/O operations to a mirrored volume, the IBM Spectrum Virtualize preferred node definition, with the multipathing driver on the host, is used to determine the preferred path. The host routes the I/Os through the preferred path, and the corresponding node is responsible for further destaging written data from cache to both volume copies. Figure 6-5 shows the data flow for write I/O processing when volume mirroring is used.

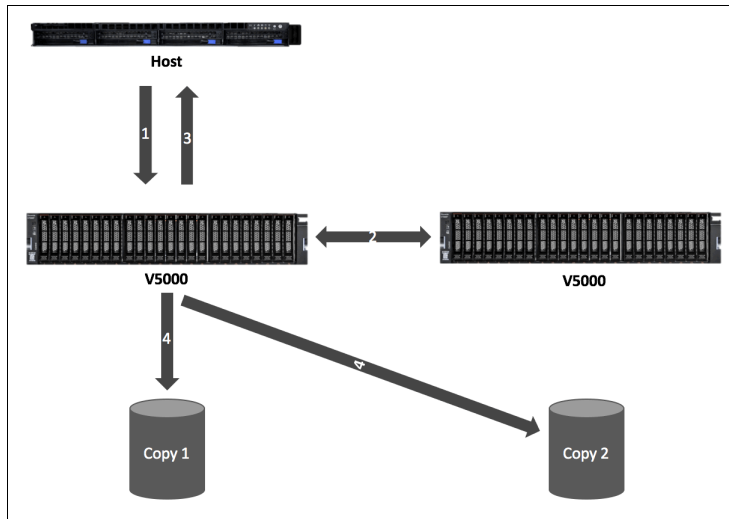


Figure 6-5 Data flow for write I/O processing in a mirrored volume

As shown in Figure 6-5, all the writes are sent by the host to the preferred node for each volume (1). Then, the data is mirrored to the cache of the partner node in the I/O Group (2), and acknowledgment of the write operation is sent to the host (3). The preferred node then destages the written data to the two volume copies (4).

A volume with copies can be checked to see whether all of the copies are identical or consistent. If a medium error is encountered while it is reading from one copy, it is repaired by using data from the other copy. This consistency check is performed asynchronously with host I/O.

Important: Mirrored volumes can be taken offline if no quorum disk is available. This behavior occurs because the synchronization status for mirrored volumes is recorded on the quorum disk.

Mirrored volumes use bitmap space at a rate of 1 bit per 256 KiB grain, which translates to 1 MiB of bitmap space supporting 2 TiB of mirrored volumes. The default allocation of bitmap space is 20 MiB, which supports 40 TiB of mirrored volumes. If all 512 MiB of variable bitmap space is allocated to mirrored volumes, 1 PiB of mirrored volumes can be supported.

Table 6-1 lists the bitmap space default configuration.

Table 6-1 *Bitmap space default configuration*

Copy service	Minimum allocated bitmap space	Default allocated bitmap space	Maximum allocated bitmap space	Minimum ^a functionality when using the default values
Remote copy ^b	0	20 MiB	512 MiB	40 TiB of remote mirroring volume capacity
FlashCopy ^c	0	20 MiB	2 GiB	<ul style="list-style-type: none"> ▶ 10 TiB of FlashCopy source volume capacity ▶ 5 TiB of incremental FlashCopy source volume capacity
Volume mirroring	0	20 MiB	512 MiB	40 TiB of mirrored volumes
RAID	0	40 MiB	512 MiB	<ul style="list-style-type: none"> ▶ 80 TiB array capacity using RAID 0, 1, or 10 ▶ 80 TiB array capacity in three-disk RAID 5 array ▶ Slightly less than 120 TiB array capacity in five-disk RAID 6 array

a. The actual amount of functionality might increase based on settings, such as grain size and strip size. RAID is subject to a 15% margin or error.

b. Remote copy includes Metro Mirror, Global Mirror, and active-active relationships.

c. FlashCopy includes the FlashCopy function, Global Mirror with change volumes, and active-active relationships.

The sum of all bitmap memory allocation for all functions except FlashCopy must not exceed 552 MiB.

6.1.5 Thin-provisioned volumes

Volumes can be configured to be thin-provisioned or fully allocated. A *thin-provisioned* volume behaves as though application reads and writes were fully allocated. When a thin-provisioned volume is created, the user specifies two capacities:

- ▶ The real physical capacity that is allocated to the volume from the storage pool
- ▶ Its virtual capacity that is available to the host

In a *fully allocated* volume, these two values are the same.

Therefore, the real capacity determines the quantity of MDisk extents that is initially allocated to the volume. The *virtual capacity* is the capacity of the volume that is reported to all other IBM Spectrum Virtualize components (for example, FlashCopy, cache, and remote copy), and to the host servers.

The *real capacity* is used to store the user data and the metadata for the thin-provisioned volume. The real capacity can be specified as an absolute value, or as a percentage of the virtual capacity.

Thin-provisioned volumes can be used as volumes that are assigned to the host, by FlashCopy to implement thin-provisioned FlashCopy targets, and with the mirrored volumes feature.

When a thin-provisioned volume is initially created, a small amount of the real capacity is used for initial metadata. I/Os are written to grains of the thin volume that were not previously written, which causes grains of the real capacity to be used to store metadata and the actual user data. I/Os are written to grains that were previously written, which updates the grain where data was previously written.

The grain size is defined when the volume is created. The grain size can be 32 KiB, 64 KiB, 128 KiB, or 256 KiB. The default grain size is 256 KiB, which is the preferred option. If you select 32 KiB for the grain size, the volume size cannot exceed 260 TiB. The grain size cannot be changed after the thin-provisioned volume is created. Generally, smaller grain sizes save space, but they require more metadata access, which can adversely affect performance.

When using thin-provisioned volume as a FlashCopy source or target volume, use 256 KiB to maximize performance. When using thin-provisioned volume as a FlashCopy source or target volume, specify the same grain size for the volume and for the FlashCopy function.

Figure 6-6 shows the thin-provisioning concept.

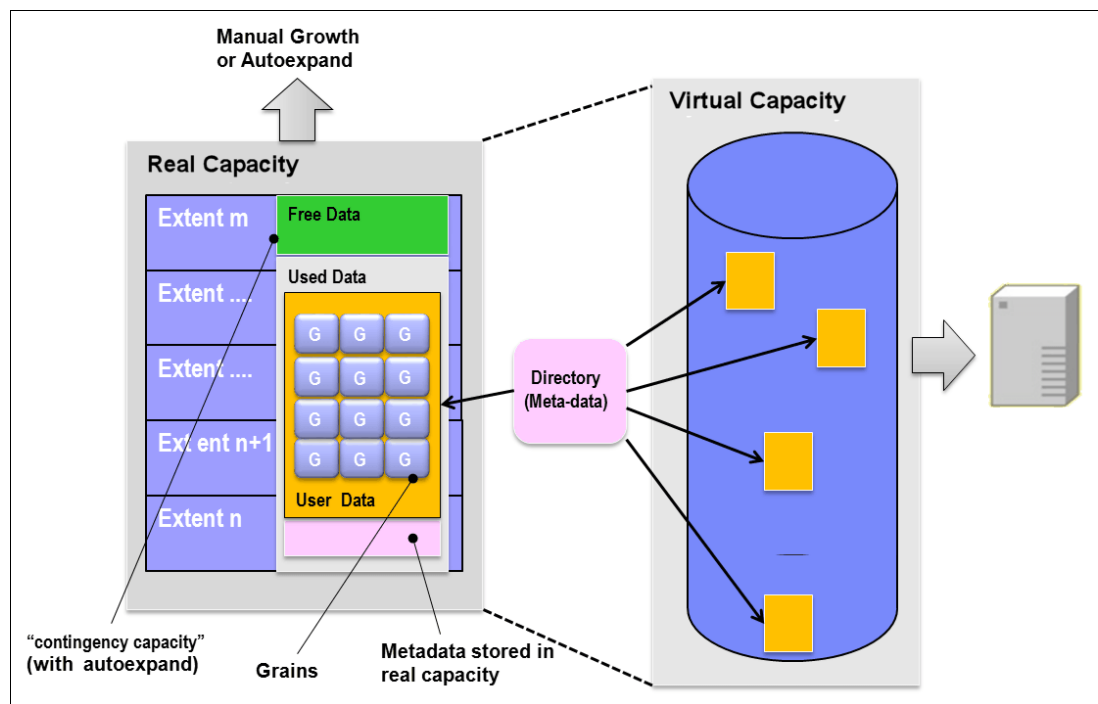


Figure 6-6 Conceptual diagram of thin-provisioned volume

Thin-provisioned volumes store user data and metadata. Each grain of data requires metadata to be stored. Therefore, the I/O rates that are obtained from thin-provisioned volumes are less than the I/O rates that are obtained from fully allocated volumes.

The metadata storage used is never greater than 0.1% of the user data. The resource usage is independent of the virtual capacity of the volume. If you are using the thin-provisioned volume directly with a host system, use a small grain size.

Thin-provisioned volume format: Thin-provisioned volumes do not need formatting. A read I/O that requests data from deallocated data space returns zeros. When a write I/O causes space to be allocated, the grain is “zeroed” before use.

The real capacity of a thin volume can be changed if the volume is not in image mode. Increasing the real capacity enables a larger amount of data and metadata to be stored on the volume. Thin-provisioned volumes use the real capacity that is provided in ascending order as new data is written to the volume. If the user initially assigns too much real capacity to the volume, the real capacity can be reduced to free storage for other uses.

A thin-provisioned volume can be configured to *autoexpand*. This feature causes the IBM Spectrum Virtualize to automatically add a fixed amount of real capacity to the thin volume as required. Therefore, autoexpand attempts to maintain a fixed amount of unused real capacity for the volume, which is known as the *contingency capacity*.

The contingency capacity is initially set to the real capacity that is assigned when the volume is created. If the user modifies the real capacity, the contingency capacity is reset to be the difference between the used capacity and real capacity.

A volume that is created without the autoexpand feature, and therefore has a zero contingency capacity, goes offline when the real capacity is used and it must expand.

Autoexpand does not cause the real capacity to grow much beyond the virtual capacity. The real capacity can be manually expanded to more than the maximum that is required by the current virtual capacity, and the contingency capacity is recalculated.

To support the auto expansion of thin-provisioned volumes, the storage pools from which they are allocated have a configurable capacity warning. When the used capacity of the pool exceeds the warning capacity, a warning event is logged. For example, if a warning of 80% is specified, the event is logged when 20% of the free capacity remains.

A thin-provisioned volume can be converted nondisruptively to a fully allocated volume (or vice versa) by using the volume mirroring function. For example, the system allows a user to add a thin-provisioned copy to a fully allocated primary volume, and then remove the fully allocated copy from the volume after they are synchronized.

The fully allocated-to-thin-provisioned migration procedure uses a zero-detection algorithm so that grains that contain all zeros do not cause any real capacity to be used.

6.1.6 Compressed volumes

This is a custom type of volume where data is compressed as it is written to disk, which saves more space. Compression is a separately orderable license that is set on a per enclosure basis. One license is required for each control or expansion enclosure and each enclosure in any external storage systems that use virtualization. Contact your IBM representative to discuss the licensing options available.

Note: For IBM Storwize V5000 Gen2, only Storwize V5030 and Storwize V5030F models support compression.

For more information about Data Reduction Pools, see Chapter 9, “Advanced features for storage efficiency” on page 435.

6.1.7 Volumes for various topologies

A *Basic* volume is the simplest form of volume. It consists of a single volume copy, which is made up of extents that are *striped* across all MDisks in a storage pool. It services I/O by using *readwrite* cache and is classified as *fully allocated* (reported real capacity and virtual capacity are equal). You can create other forms of volumes, depending on the type of topology that is configured on your system:

- ▶ With standard topology, which is a single-site configuration, you can create a basic volume or a mirrored volume.

By using volume mirroring, a volume can have two physical copies. Each volume copy can belong to a different pool, and each copy has the same virtual capacity as the volume. In the management GUI, an asterisk indicates the primary copy of the mirrored volume. The primary copy indicates the preferred volume for read requests.

- ▶ With HyperSwap topology, which is a three-site HA configuration, you can create a basic volume or a HyperSwap volume.

HyperSwap volumes create copies on separate sites for systems that are configured with HyperSwap topology. Data that is written to a HyperSwap volume is automatically sent to both copies so that either site can provide access to the volume if the other site becomes unavailable.

Note: For IBM Storwize V5000 Gen2, the HyperSwap topology is supported on Storwize V5030 and Storwize V5030F systems only.

- ▶ The IBM Spectrum Virtualize V7.6.0 release also introduces *Virtual Volumes (VVols)*. These volumes are available in a system configuration that supports VMware vSphere Virtual Volumes. These volumes allow VMware vCenter to manage system objects, such as volumes and pools. The Spectrum Virtualize system administrators can create volume objects of this class, and assign ownership to VMware administrators to simplify management.

For more information about configuring vVol with IBM Spectrum Virtualize, see *Configuring VMware Virtual Volumes for Systems Powered by IBM Spectrum Virtualize*, SG24-8328.

Note: From V7.4.0 onwards, it is possible to prevent accidental deletion of volumes, if they recently performed any I/O operations. This feature is called *Volume protection*, and it prevents active volumes or host mappings from being deleted inadvertently. This is done by using a global system setting. For more information, see [IBM Knowledge Center](#).

6.2 Create Volumes menu

The GUI is the simplest means of volume creation, and presents different options in the Create Volumes menu depending on the topology of the system.

To start the process of creating a volume, complete the following steps:

1. Click the **Volumes** on the main window, as shown in Figure 6-7.

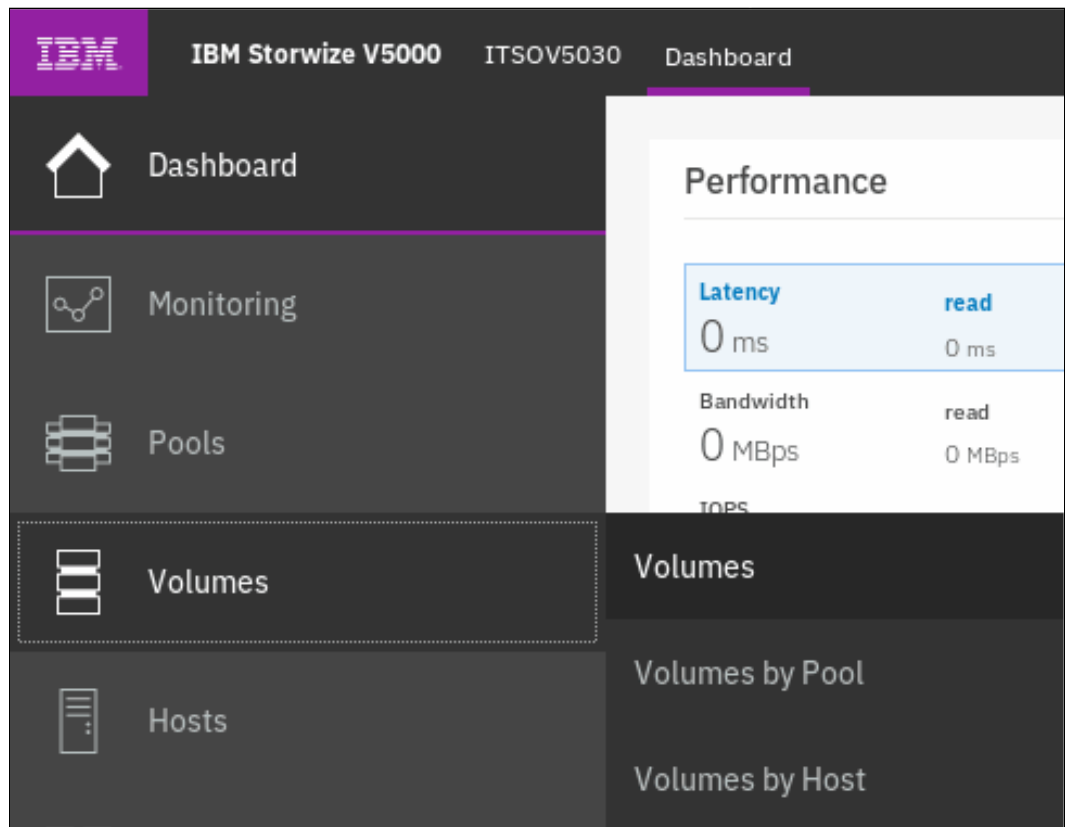


Figure 6-7 Volumes

2. Click **Volumes** in the submenu option. Then, in the pane on the right, you see a list of available volumes, as shown in Figure 6-8.

Name	State	Synchronized	Pool	UID
CompressedVolume-01	Online		ITSO Pool 2	6005076380818116C000000000000015
CompressedVolume-02	Online		ITSO Pool 2	6005076380818116C000000000000016
CompressedVolume-03	Online		ITSO Pool 2	6005076380818116C000000000000017
CompressedVolume-04	Online		ITSO Pool 2	6005076380818116C000000000000018
CompressedVolume-05	Online		ITSO Pool 2	6005076380818116C000000000000019
Linux1	Online		ITSO Redbook	6005076380818116C00000000000000C
Linux2	Online		ITSO Redbook	6005076380818116C00000000000000D
Linux3	Online		ITSO Redbook	6005076380818116C00000000000000E
Linux4	Online		ITSO Redbook	6005076380818116C00000000000000F
Linux5	Online		ITSO Redbook	6005076380818116C000000000000010

Showing 24 volumes | Selecting 0 volumes

Figure 6-8 Existing volumes

3. Click **Create Volumes**, as shown in Figure 6-9.

Name	State	Synchronized	Pool	UID
CompressedVolume-01	Online		ITSO Pool 2	6005076380818116C000000000000015
CompressedVolume-02	Online		ITSO Pool 2	6005076380818116C000000000000016
CompressedVolume-03	Online		ITSO Pool 2	6005076380818116C000000000000017
CompressedVolume-04	Online		ITSO Pool 2	6005076380818116C000000000000018
CompressedVolume-05	Online		ITSO Pool 2	6005076380818116C000000000000019
DS8000_40424010000000...	Online		ITSO Redbook	6005076380818116C00000000000003A
DS8000_40434010000000...	Online		MigrationPool_1024	6005076380818116C000000000000039
DS8000_40434012000000...	Online		ITSO Redbook	6005076380818116C000000000000038
Linux1	Online		ITSO Pool 2	6005076380818116C00000000000000C
Linux2	Online		ITSO Redbook	6005076380818116C00000000000000D

Figure 6-9 Create Volumes

Depending on the topology of the system, clicking **Create Volumes** provides different options.

For standard topology, **Create Volumes** provides a pop-up window with options to create basic, mirrored, or custom volumes, as shown in Figure 6-10.

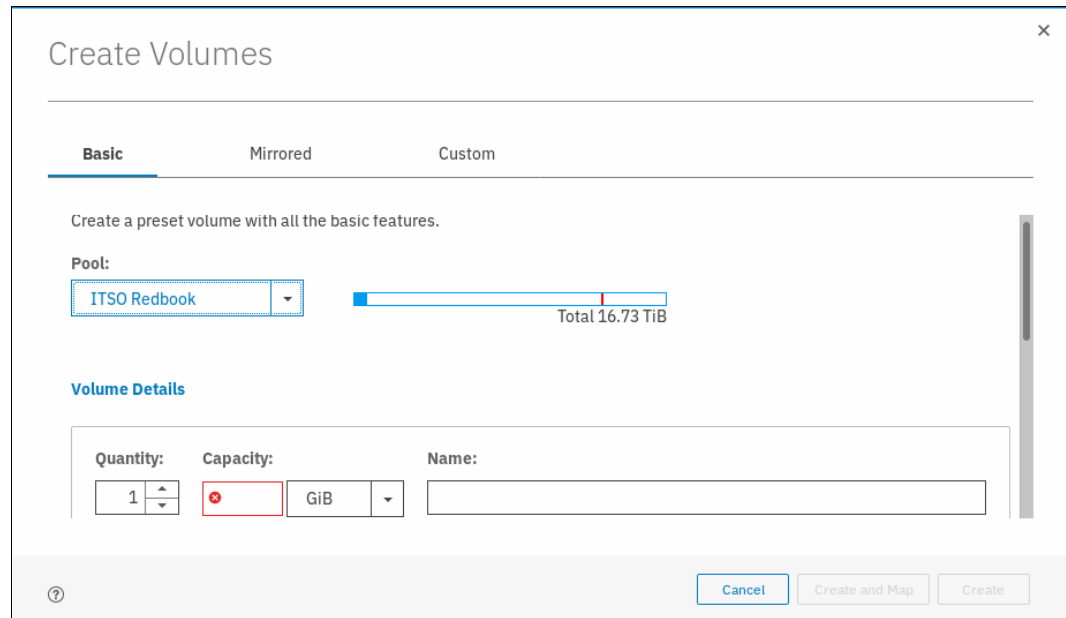


Figure 6-10 Create Volume options for standard topology

For HyperSwap topology, the **Create Volumes** option provides pop-up window with options to create basic, HyperSwap, or custom volumes, as shown in Figure 6-11.

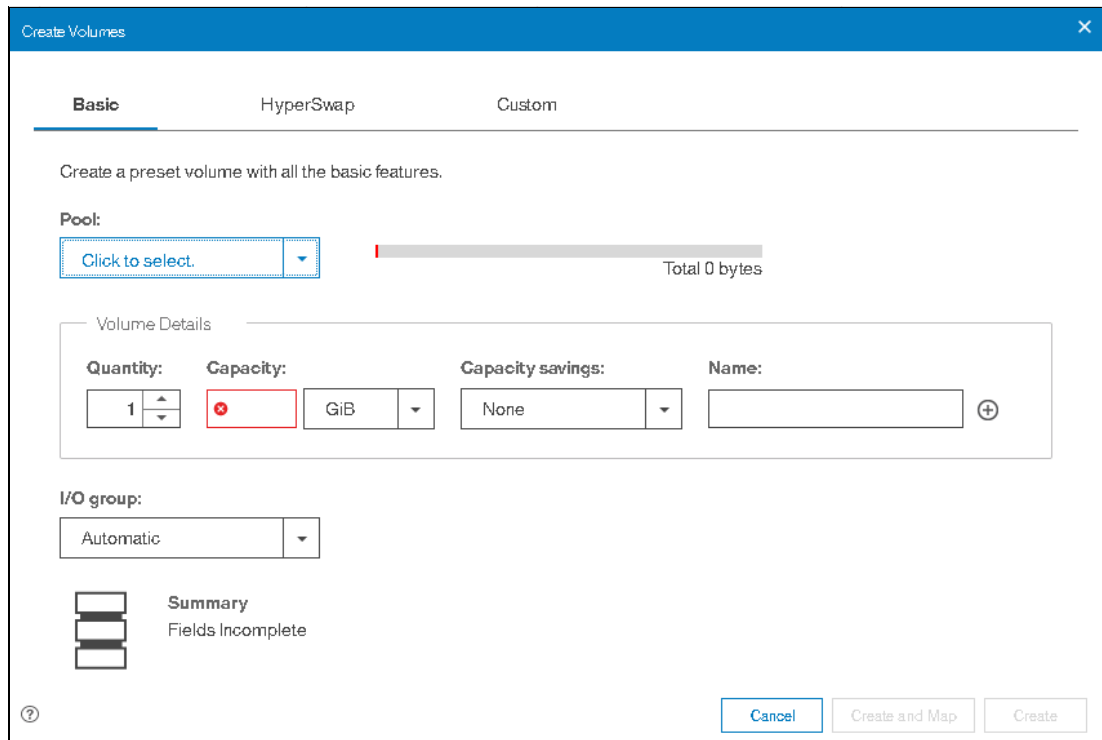


Figure 6-11 Create Volume options for HyperSwap topology

Clicking any of the three choices in the **Create Volumes** window opens a drop-down window in which volume details can be entered. The example that is shown in Figure 6-12 shows a Basic volume to demonstrate this view.

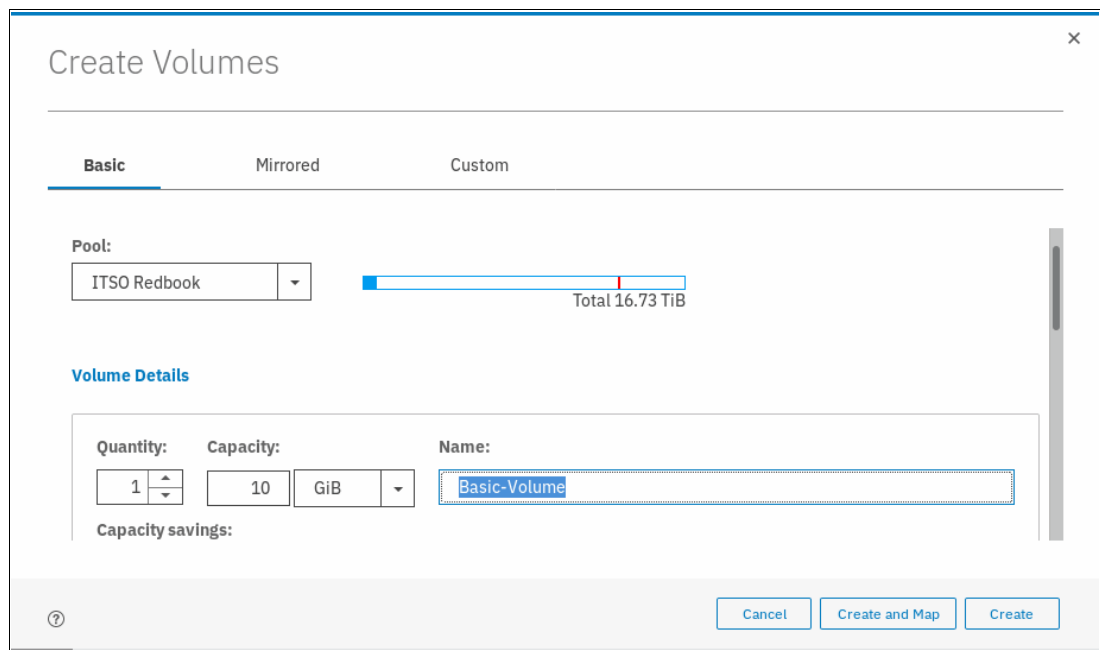


Figure 6-12 Basic Volume

Notes: Consider the following points:

- ▶ A Basic volume is a volume whose data is striped across all available managed disks (MDisks) in one storage pool.
- ▶ A Mirrored volume is a volume with two physical copies, where each volume copy can belong to a different storage pool.
- ▶ In the context of this menu, a Custom volume is a Basic or Mirrored volume with customization from the default parameters.

By using the **Capacity Savings** parameter, Volume Creation also provides the ability to change the default provisioning of a Basic or Mirrored Volume to Thin-provisioned or Compressed.

For more information about volume migration, see 6.8, “Migrating a volume to another storage pool” on page 347.

For more information about creating volume copies, see 6.3.2, “Creating Mirrored volumes by using Volume Creation” on page 327.

6.3 Creating volumes by using the Volume Creation

This section focuses on using the Volume Creation operation to create Basic and Mirrored volumes in a system with standard topology. It also covers creating host-to-volume mapping. Volume Creation is available on four different volume classes:

- ▶ Basic
- ▶ Mirrored
- ▶ Custom
- ▶ HyperSwap

Note: The ability to create HyperSwap volumes by using the GUI simplifies the creation and configuration processes. This simplification is enhanced by the GUI by using the `mkvolume` command.

6.3.1 Creating Basic volumes by using Volume Creation

The most commonly used type of volume is the Basic volume. This type of volume is fully provisioned, with the entire size dedicated to the defined volume. The host and the IBM Spectrum Virtualize system see the fully allocated space.

Create a Basic volume by clicking the **Basic** icon, as shown in Figure 6-13 on page 326. This action opens another input window in which you can define the following information:

- ▶ Pool: The pool in which the volume is created (drop-down).
- ▶ Quantity: The number of volumes to be created (numeric up/down).
- ▶ Capacity: Size of the volume in units (drop-down).
- ▶ Capacity Savings (drop-down):
 - None
 - Thin-provisioned
 - Compressed
- ▶ Name: Name of the volume (cannot start with a number).
- ▶ I/O group

The Basic volume creation process is shown in Figure 6-13.

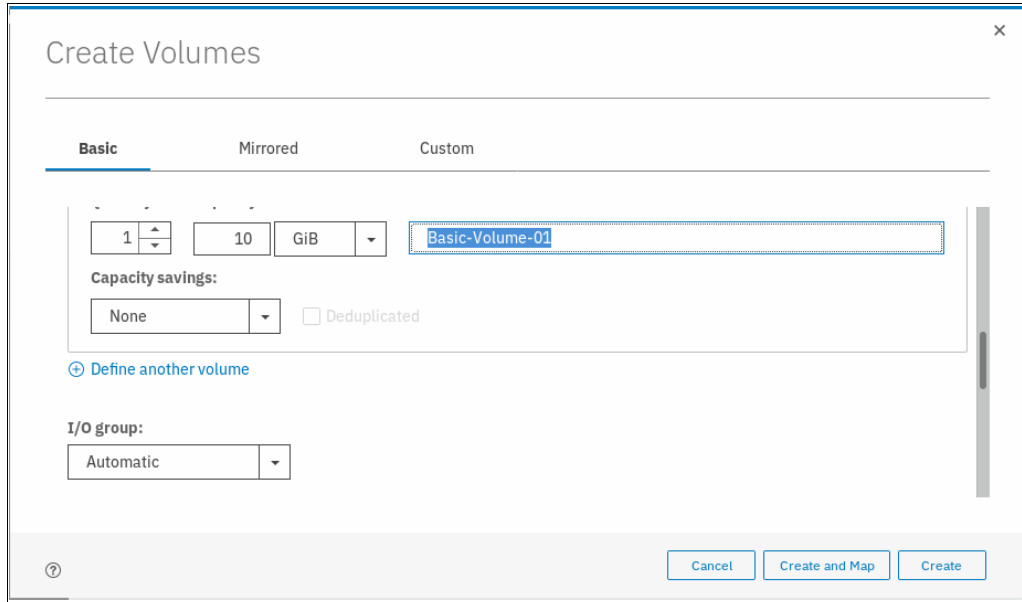


Figure 6-13 Creating Basic volume

An appropriate naming convention is recommended for volumes for easy identification of their association with the host or host cluster. At a minimum, it should contain the name of the pool or some tag that identifies the underlying storage subsystem. It can also contain the host name that the volume is mapped to, or perhaps the content of this volume; for example, name of applications to be installed.

When all of the characteristics of the Basic volume are defined, it can be created by selecting one of the following options:

- ▶ Create
- ▶ Create and Map to Host

Note: The plus sign (+ **Define another volume**) option that is shown in Figure 6-13 can be used to create volumes in the same instance of the volume creation wizard.

In this example, the **Create** option was selected (the volume-to-host mapping can be performed later). At the end of the volume creation process, a task completion window is opened. Success is also indicated by the state of the Basic volume being reported as formatting in the Volumes pane, as shown in Figure 6-14.

Name	State	Synchronized	Pool	UID	
Basic-Volume-01	✓ Online (formatting)		ITSO Redbook	6005076380818116C00000000000003C	!!!

Figure 6-14 Basic Volume Fast-Format

Notes: Consider the following points:

- ▶ Fully allocated volumes are automatically formatted through the quick initialization process after the volume is created. This process makes fully allocated volumes available for use immediately.
- ▶ Quick initialization requires a small amount of I/O to complete, and limits the number of volumes that can be initialized at the same time. Some volume actions, such as moving, expanding, shrinking, or adding a volume copy, are disabled when the specified volume is initializing. Those actions are available after the initialization process completes.
- ▶ The quick initialization process can be disabled in circumstances where it is not necessary. For example, if the volume is the target of a Copy Services function, the Copy Services operation formats the volume. The quick initialization process can also be disabled for performance testing so that the measurements of the raw system capabilities can take place without waiting for the process to complete.

For more information, see [IBM Knowledge Center](#).

6.3.2 Creating Mirrored volumes by using Volume Creation

IBM Spectrum Virtualize offers the capability to mirror volumes, which means a single volume, presented to a host, can have two physical copies. Each volume copy can belong to a different pool, and each copy has the same virtual capacity as the volume. When a server writes to a mirrored volume, the system writes the data to both copies. When a server reads a mirrored volume, the system picks one of the copies to read.

Normally, this copy is the primary copy, as indicated in the management GUI by an asterisk (*). If one of the mirrored volume copies is temporarily unavailable (for example, because the storage system that provides the pool is unavailable), the volume remains accessible to servers. The system remembers which areas of the volume are written and resynchronizes these areas when both copies are available.

The use of mirrored volumes results in the following outcomes:

- ▶ Improves availability of volumes by protecting them from a single storage system failure
- ▶ Provides concurrent maintenance of a storage system that does not natively support concurrent maintenance
- ▶ Provides an alternative method of data migration with better availability characteristics
- ▶ Converts between fully allocated volumes and thin-provisioned volumes

Note: Although volume mirroring is not a true disaster recovery (DR) solution because both copies are accessed by the same node pair and addressable by only a single cluster, it can improve availability.

To create a mirrored volume, complete the following steps:

1. In the Create Volumes window, click **Mirrored**. In the **Mirrored copies** subsection, choose the **Pool** of **Copy1** and **Copy2** by using the drop-down menu. Although the mirrored volume can be created in the same pool, this setup is not typical.
2. In Volume Details section, complete the following fields:
 - Quantity
 - Capacity
 - Capacity savings

- Name

Generally, keep mirrored volumes on a separate set of physical disks (Pools). Leave the I/O group option at its default setting of Automatic (see Figure 6-15).

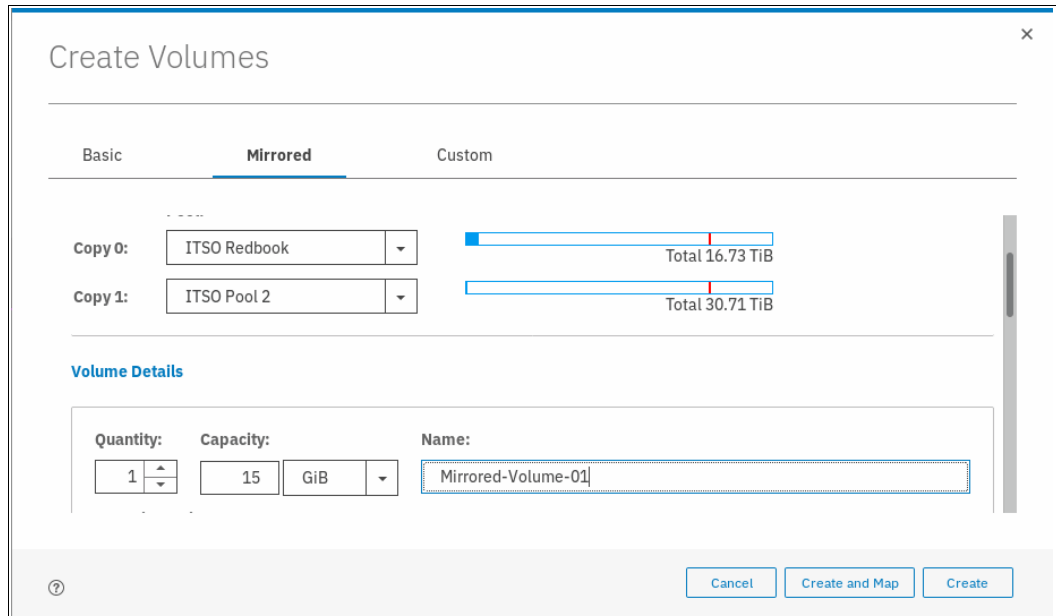


Figure 6-15 Mirrored Volume creation

3. Click **Create** (or **Create and Map to Host**).

A task completion window opens.

Note: When creating a Mirrored volume by using this menu, you are not required to specify the Mirrored Sync rate. It defaults to 2 MBps. Customization of this synchronization rate can be done by using the **Custom** option.

Volume Creation with Capacity Saving options

The Volume Creation operation also provides the ability to alter the provisioning of a Basic or Mirrored volume into Thin-provisioned or Compressed by using the **Capacity Savings** parameter.

Select **Thin-provisioned** or **Compressed** from the drop-down menu, as shown in Figure 6-16.

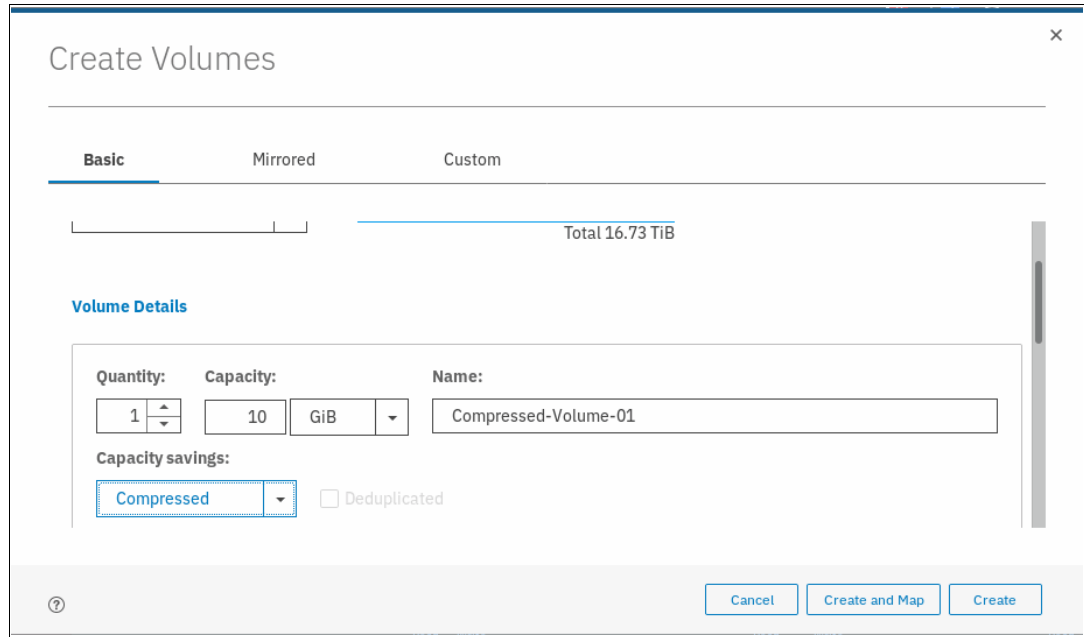


Figure 6-16 Volume Creation with Capacity Saving option set to Compressed

Alternatively, select **Thin-provisioned** from the menu to define a Thin-provisioned volume.

6.4 Mapping a volume to the host

After a volume is created, it can be mapped to a host. Complete the following steps:

1. From the Volumes menu, highlight the volume that you want to create a mapping for and select **Actions** from the menu bar.

Tip: An alternative way of opening the Actions menu is to right-click a volume.

2. From the Actions menu, select **Map to Host or Host Cluster**, as shown in Figure 6-17 on page 330.

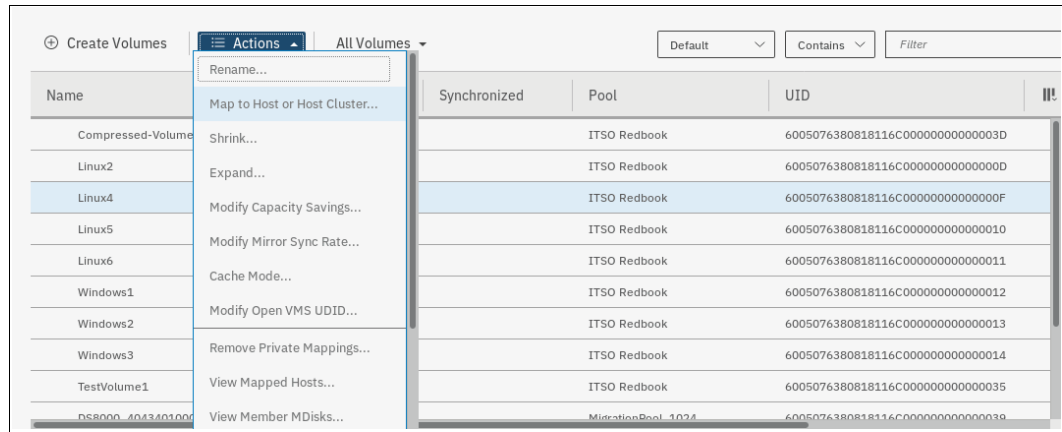


Figure 6-17 Map to Host

3. A Create Mapping window opens. In this window, indicate whether the volume needs to be mapped to a host or host cluster, select the wanted host or host cluster, and whether you want system assigned SCSI ID or self-assigned, as shown in Figure 6-18.

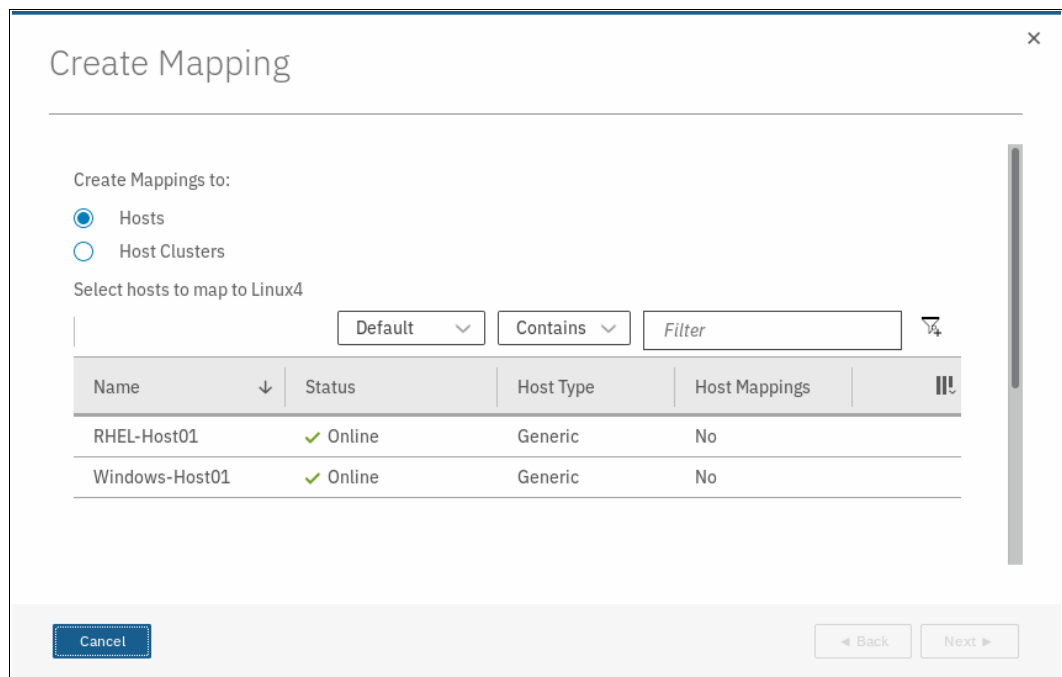


Figure 6-18 Mapping a Volume to Host

4. Click **Next**. A window opens in which mapped volumes to that host are listed along with the new volume to be mapped, as shown in Figure 6-19.

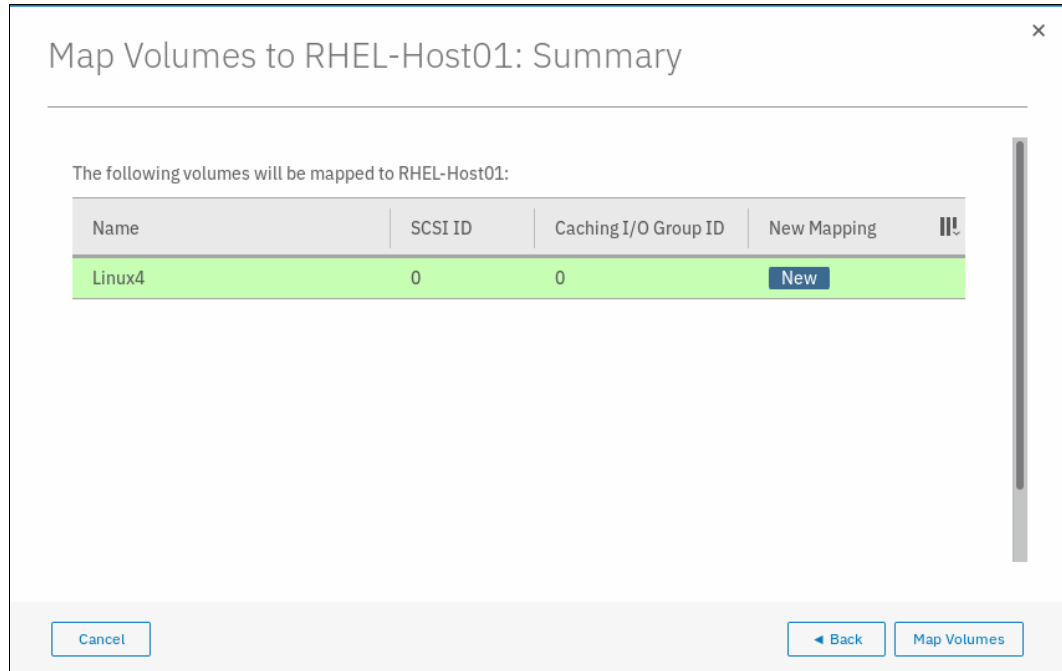


Figure 6-19 Map volume to host

5. Click **Map Volumes** and the Modify Mappings window shows the command details. Then, a Task completed message is displayed.

6.5 Creating Custom volumes

The Create Volumes window enables Custom volume creation. It provides an alternative method of defining Capacity savings options, such as Thin-provisioning and Compression, but also expands on the base level default options for available Basic and Mirrored volumes. A Custom volume can be customized regarding Mirror sync rate, Cache mode, and Fast-Format.

The Custom volume creation operation consists of several options:

- ▶ Volume Location (Mandatory, defines the Pools to be used)
- ▶ Volume Details (Mandatory, defines the Capacity savings option, such as Thin Provisioning or Compression or none)
- ▶ General (for changing default options for Cache mode and Formatting)
- ▶ Summary

Work through these options to customize your Custom volume. Then, commit these changes by using **Create**, as shown in Figure 6-20.

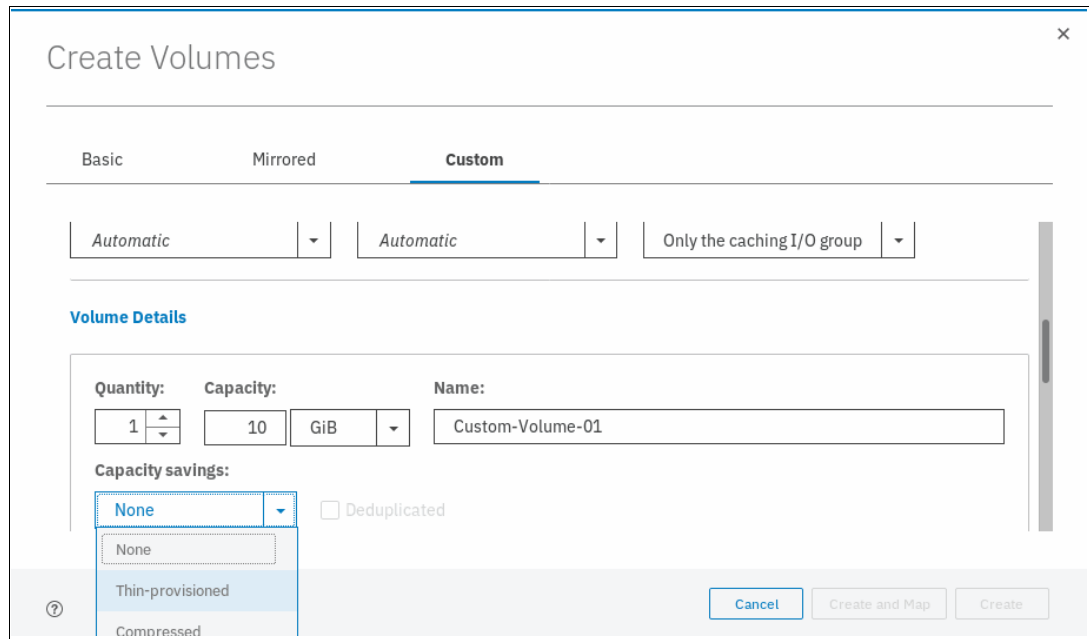


Figure 6-20 Customization submenus

6.5.1 Creating a custom thin-provisioned volume

A thin-provisioned volume can be defined and created by using the **Custom** option. Regarding application reads and writes, thin-provisioned volumes behave as though they were fully allocated.

When creating a thin-provisioned volume, you can specify two capacities:

- ▶ The real physical capacity that is allocated to the volume from the storage pool. The real capacity determines the quantity of extents that are initially allocated to the volume.
- ▶ Its virtual capacity available to the host. The virtual capacity is the capacity of the volume that is reported to all other components (for example, FlashCopy, cache, and remote copy) and to the hosts.

To create a thin-provisioned volume, complete the following steps:

1. From the Create Volumes window, select the **Custom** option. In the **Volume Location** subsection, define the pool in which the volume is created.

- Use the drop-down menu in the **Pool** option to choose the pool. All other options, such as Volume copy type, Caching I/O group, Preferred node, and Accessible I/O groups, can be left with their default options, as shown in Figure 6-21.

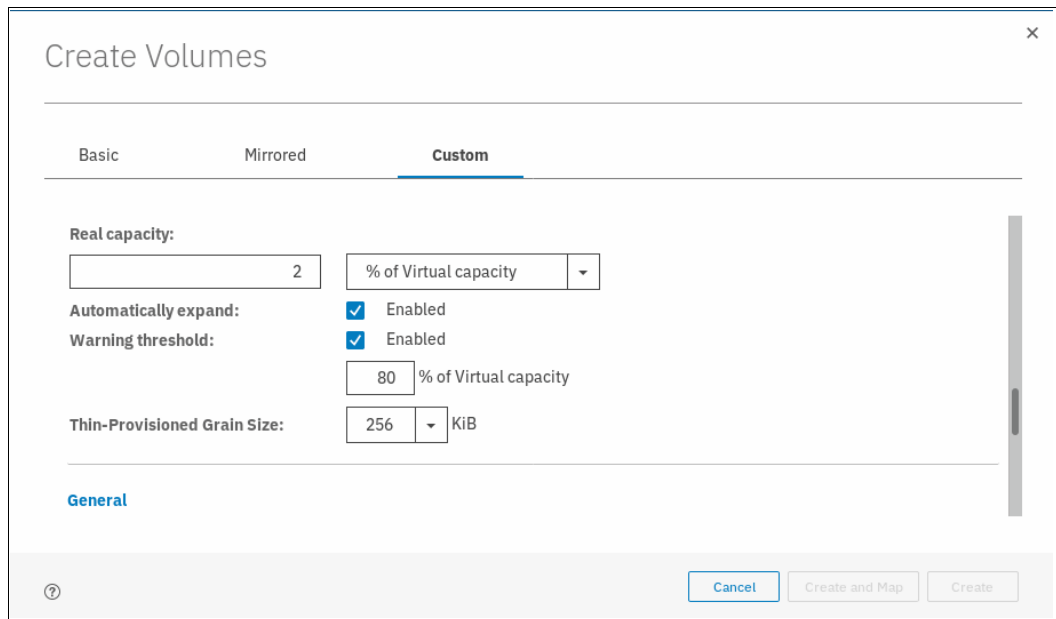


Figure 6-21 Volume Location for thin-provisioned volume

- In the **Volume Details** subsection, you can enter the Quantity, Capacity (virtual), Capacity Savings (choose Thin-provisioned from the drop-down menu), and Name of the volume being created, as shown in Figure 6-22.



Figure 6-22 Volume Details

- In the **Thin Provisioning** subsection, enter the real and virtual capacity, expansion criteria, and grain size, as shown in Figure 6-23.

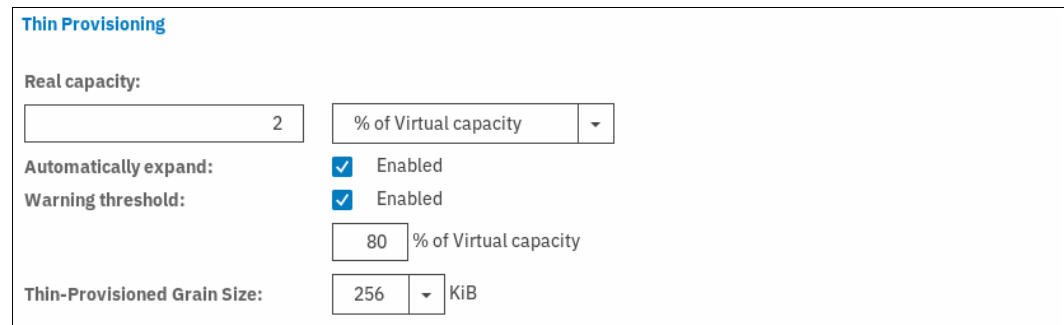


Figure 6-23 Thin Provisioning

The following Thin Provisioning options are available (defaults are in parentheses):

- **Real capacity** (2%): Specify the size of the real capacity space that is used during creation.
- **Automatically Expand** (Enabled): This option enables the automatic expansion of real capacity, if more capacity is to be allocated.
- **Warning threshold** (Enabled): Enter a threshold for receiving capacity alerts.
- **Grain Size** (256 kibibytes (KiB)): Specify the grain size for real capacity. This option describes the size of the chunk of storage to be added to used capacity. For example, when the host writes 1 MB of new data, the capacity is increased by adding four chunks of 256 KiB each.

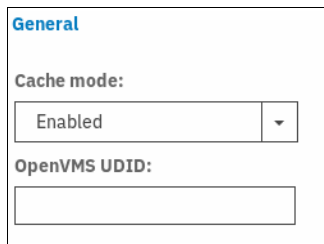
Important: If you do not use the **autoexpand** feature, the volume goes offline after reaching its real capacity.

The default grain size is 256 KiB. The optimum choice of grain size depends on volume use type. For more information, see the “Performance Problem When Using EasyTier With Thin Provisioned Volumes” topic at [this website](#).

Consider the following points:

- ▶ If you are *not* going to use the thin-provisioned volume as a FlashCopy source or target volume, use 256 KiB to maximize performance.
- ▶ If you *are* going to use the thin-provisioned volume as a FlashCopy source or target volume, specify the same grain size for the volume and for the FlashCopy function.

5. In the General subsection, enter the caching mode as Enabled, Read-only, or Disabled, as shown in Figure 6-24. Also, enter unit device identifier (UDID) if this volume is going to be mapped to an OpenVMS host.



The screenshot shows a configuration window titled "General". It contains two fields: "Cache mode:" with a dropdown menu currently set to "Enabled", and "OpenVMS UDID:" with an empty text input field.

Figure 6-24 General details

6. Click **Create** to define the volume. A task completion window opens.

6.5.2 Creating Custom Compressed volumes

The configuration of compressed volumes is similar to thin-provisioned volumes. To create a Compressed volume, complete the following steps:

1. From the **Create Volumes** window, select the **Custom** option. In the **Volume Location** subsection, define the pool in which the volume is created. Use the drop-down menu in the **Pool** option to choose the pool.

All other options, such as **Volume copy type**, **Caching I/O group**, **Preferred node**, and **Accessible I/O groups**, can be left with their default options, as shown in Figure 6-25.

The screenshot shows a 'Create Volumes' dialog box with three tabs: 'Basic', 'Mirrored', and 'Custom'. The 'Custom' tab is selected. Under the heading 'volume Location', there are three sections: 'Volume copy type' with a dropdown menu set to 'None'; 'Pool' with a dropdown menu set to 'ITSO Redbook'; and three dropdown menus for 'Caching I/O group' (set to 'Automatic'), 'Preferred node' (set to 'Automatic'), and 'Accessible I/O groups' (set to 'Only the caching I/O group'). At the bottom right, there are three buttons: 'Cancel', 'Create and Map', and 'Create'. A help icon (?) is located at the bottom left.

Figure 6-25 Defining a volume as compressed using the Capacity savings option

2. In the **Volume Details** subsection, you can enter the Quantity, Capacity (virtual), Capacity Savings (choose Compressed from the drop-down menu), and Name of the volume being created, as shown in Figure 6-26.

The screenshot shows a 'Volume Details' form. It has three columns: 'Quantity' with a spinner box set to '1'; 'Capacity' with a text box '25', a unit dropdown 'GiB', and a small dropdown arrow; and 'Name' with a text box containing 'Compressed-Volume-01'. Below these is a 'Capacity savings' section with a dropdown menu set to 'Compressed' and a checkbox labeled 'Deduplicated' which is unchecked. At the bottom left, there is a blue link with a plus icon that says '+ Define another volume'.

Figure 6-26 Volume Details

3. In the **Compressed** subsection, enter the real capacity in terms of % of virtual capacity or in GiB, expansion criteria, warning threshold, and wanted % of the virtual capacity at which you should receive a warning, as shown in Figure 6-27.

Compressed

Real capacity: % of Virtual capacity ▾

Automatically expand: Enabled

Warning threshold: Enabled

% of Virtual capacity

Figure 6-27 Compressed details

4. In the **General** subsection, enter the Cache mode as Enabled, Read-only, or Disabled, as shown in Figure 6-28. Also, enter unit device identifier (UDID) if this volume is going to be mapped to an OpenVMS host.

General

Cache mode: ▾

OpenVMS UDID:

Figure 6-28 General details

5. Click **Create** to define the volume. A task completion window opens.

6.5.3 Custom Mirrored Volumes

The Custom option in the Create Volumes window is used to customize volume creation. By using this feature, the default options can be overridden and volume creation can be tailored to the specifics of the clients environment.

Modifying the Mirror sync rate

The Mirror sync rate can be changed from the default setting by using the Custom option, subsection Volume Location, of the Create Volumes window.

This option sets the priority of copy synchronization progress, which enables a preferential rate to be set for more important volumes (see Figure 6-29).

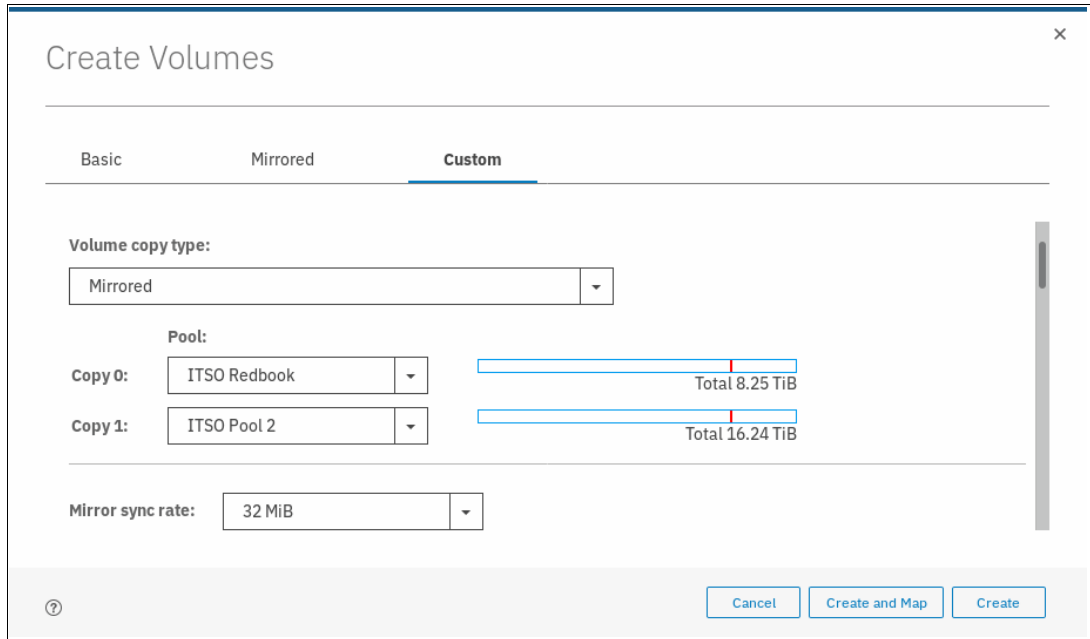


Figure 6-29 Customization of Mirrored sync rate

The progress of formatting and synchronization of a newly created Mirrored Volume can be checked from the Running Tasks menu. In this menu, the progress of all currently running tasks is reported, including Volume Format and Volume Synchronization (see Figure 6-30).

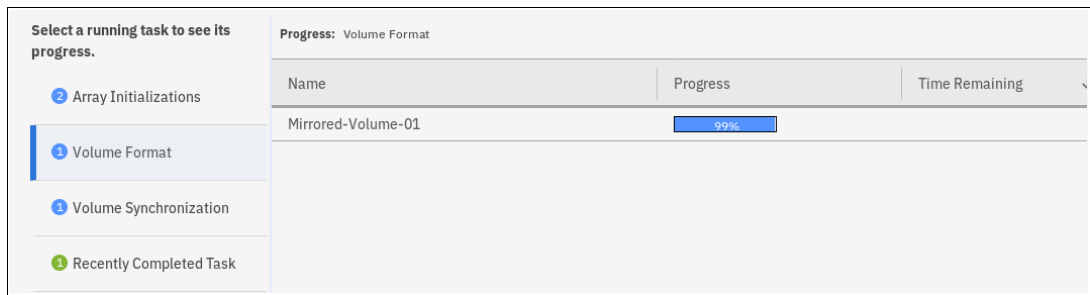


Figure 6-30 Progress of all currently running tasks

Creating a Custom Thin-provisioned Mirrored volume

The Custom option in the Create Volumes window is used to customize volume creation. By using this feature, the default options can be overridden and volume creation can be tailored to the specifics of the clients environment.

The Mirror Sync rate can be changed from the default setting under the Volume Location subsection of the Create Volume window. This option sets the priority of copy synchronization progress, which enables a preferential rate to be set for more important mirrored volumes.

The Summary shows you the capacity information and the allocated space. You can customize the thin-provision settings or the mirror synchronization rate. After you create the volume, the task completion window opens.

The initial synchronization of thin-mirrored volumes is fast when a small amount of real and virtual capacity is used.

6.6 HyperSwap and the mkvolume command

HyperSwap volume configuration is not possible until site awareness are configured.

When the HyperSwap topology is configured, the GUI uses the `mkvolume` command to create volumes instead of the traditional `mkvdisk` command. This section describes the `mkvolume` command that is used in HyperSwap topology. The GUI continues to use the `mkvdisk` command when all other classes of volumes are created.

Note: It is still possible to create HyperSwap volumes as in the V7.5.0 release, as described in the following IBM Support [white paper](#).

For more information, see *IBM Storwize V7000, Spectrum Virtualize, HyperSwap, and VMware Implementation*, [SG24-8317](#).

HyperSwap volumes are a new type of HA volumes that are supported by IBM Spectrum Virtualize. They are built off two IBM Spectrum Virtualize technologies:

- ▶ Metro Mirror
- ▶ Volume Mirroring

These technologies were combined in an active-active configuration that is deployed by using Change Volumes (as used in the Global Mirror with Change Volumes) to create a Single Volume (from a host perspective) in an HA form. Although the volume that is presented is a combination of four “traditional” volumes, it is a single entity from a host (and administrative) perspective, as shown in Figure 6-31.

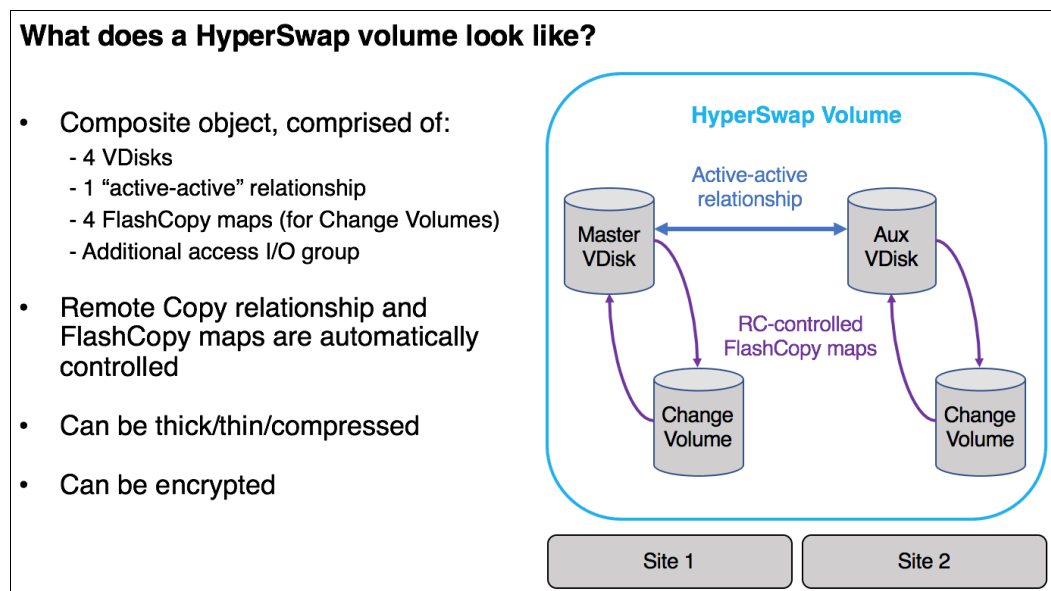


Figure 6-31 What makes up a HyperSwap Volume

The GUI simplifies the complexity of HyperSwap volume creation by presenting only the volume class of HyperSwap as a Volume Creation option after HyperSwap topology is configured.

In the following example, HyperSwap topology is configured and the Volume Creation window is being used to define a HyperSwap Volume, as shown in Figure 6-32.

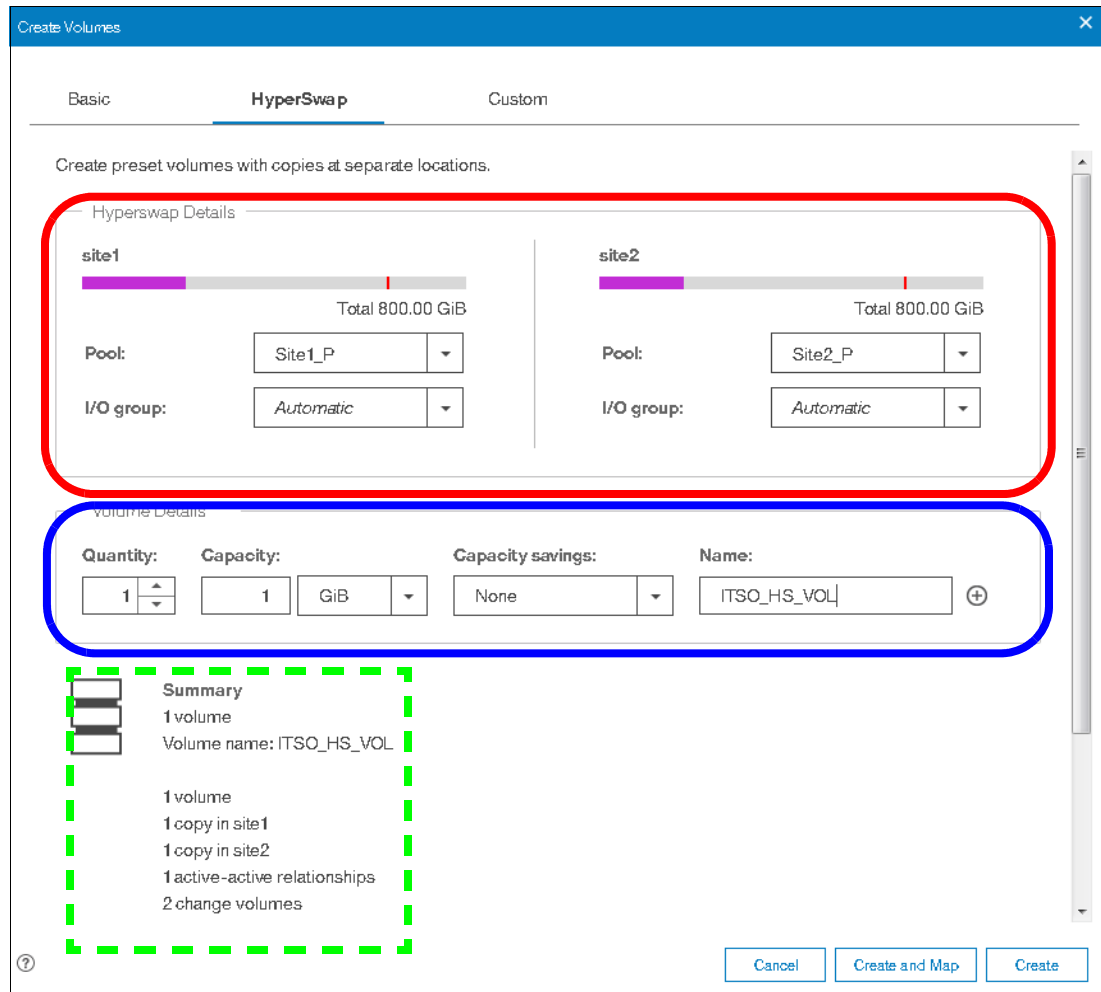


Figure 6-32 HyperSwap Volume creation with Summary of actions

The capacity and name characteristics are defined as for a Basic volume (highlighted in blue in Figure 6-32) and the mirroring characteristics are defined by the Site parameters (highlighted in red in Figure 6-32).

The drop-down menus help during creation, and the Summary (lower left of the creation window) indicates the actions that are carried out when the Create option is selected.

As shown in Figure 6-32, a single volume is created, with volume copies in site1 and site2. This volume is in an active-active (Metro Mirror) relationship with extra resilience provided by two change volumes.

The command that is issued to create this volume is shown in Figure 6-33, and can be summarized as follows:

```
svctask mkvolume -name <name_of_volume> -pool <X:Y> -size <Size_of_volume> -unit <units>
```

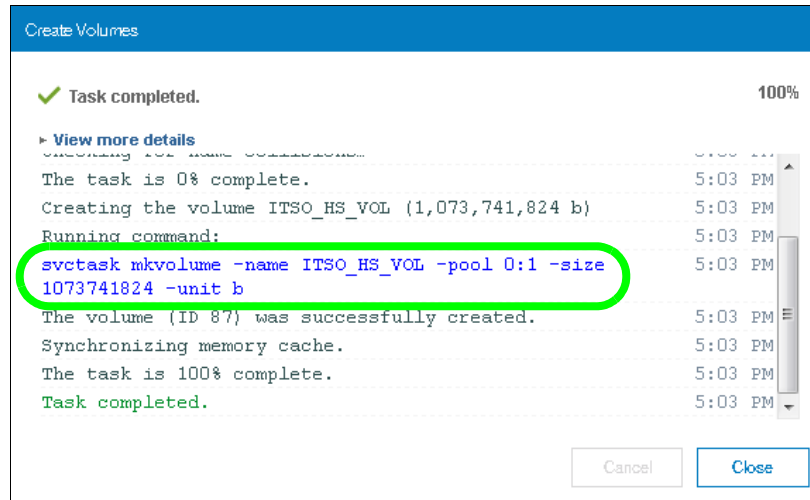


Figure 6-33 Example mkvolume command

6.6.1 Volume manipulation commands

Five CLI commands for administering volumes were released in IBM Spectrum Virtualize V7.6.0. However, the GUI continues to use commands for all volume administration, except for HyperSwap volume creation (`mkvolume`) and deletion (`rmvolume`). The following CLI commands are available for administering volumes:

- ▶ `mkvolume`
- ▶ `mkimagevolume`
- ▶ `addvolumecopy`
- ▶ `rmvolumecopy`
- ▶ `rmvolume`

In addition, the `lsvdisk` command and GUI functionality is available. The `lsvdisk` command now includes `volume_id`, `volume_name`, and `function` fields to easily identify the individual VDisk that make up a HyperSwap volume. These views are “rolled-up” in the GUI to provide views that reflect the client’s view of the HyperSwap volume and its site-dependent copies, as opposed to the “low-level” VDisks and VDisk-change-volumes.

As shown in the Figure 6-34, clicking **Volumes** → **Volumes** shows the HyperSwap Volume ITS0_HS_VOL with an expanded view opened by using the twistie (V) to reveal two volume copies: ITS0_HS_VOL (site1) (Master VDisk) and ITS0_HS_VOL (site2) (Auxiliary VDisk). We do not show the VDisk-Change-Volumes.

Name	State	Synchronized	Pool	UID	Host Mappings	Capacity
ITS0_HS_VOL	Online	No	Multiple	600507630088830788000000000009F	No	1.00 GB
ITS0_HS_VOL (site1)	Online	Yes	Site1_P	600507630088830788000000000009F	No	1.00 GB
ITS0_HS_VOL (site2)	Online	Yes	Site2_P	60050763008883078800000000000A0	No	1.00 GB

Figure 6-34 Hidden Change Volumes

Likewise, the status of the HyperSwap volume is reported at a “parent” level. If one of the copies is synchronized or not or offline, the HyperSwap volume reflects this state, as shown in Figure 6-35.

Name	State	Synchronized	Pool	Volume Group	UID	Host Mappings	Capacity
ITSO_HS_VOL	✓ Online (formatting)		ARCDS8KFCL60		60050780C868245F0000000000000069	No	1.00 GB
Copy 0*	✓ Online (formatting)	Yes	ARCDS8KFCL60		60050780C868245F0000000000000069	No	1.00 GB
Copy 1	✓ Online (formatting)	No	ARCDS8KXK700		60050780C868245F0000000000000069	No	1.00 GB

Figure 6-35 Parent volume reflects state of copy volume

The following HyperSwap-related individual commands are available:

► **mkvolume**

Create an empty volume by using storage from existing storage pools. The type of volume created is determined by the system topology and the number of storage pools specified. Volume is always formatted (zeroed). This command can be used to create the following items:

- Basic volume: Any topology
- Mirrored volume: Standard topology
- HyperSwap volume: HyperSwap topology

► **rmvolume**

Remove a volume. For a HyperSwap volume, this process includes deleting the active-active relationship and the change volumes.

The **-force** parameter with **rmvdisk** is replaced by individual override parameters, which makes it clearer to the user exactly what protection they are bypassing.

► **mkimagevolume**

Create an image mode volume. This command can be used to import a volume, which preserves data. Implemented as a separate command to provide greater differentiation between the action of creating an empty volume and creating a volume by importing data on an MDisk.

► **addvolumecopy**

Add a copy to a volume. The new copy is always synchronized from the copy. For HyperSwap topology systems, this process creates a highly available volume. This command can be used to create the following volume types:

- Mirrored volume: Standard topology
- HyperSwap volume: HyperSwap topology

► **rmvolumecopy**

Remove a copy of a volume. Leaves the volume intact. Converts a Mirrored or HyperSwap volume into a basic volume. For a HyperSwap volume, this process includes deleting the active-active relationship and the change volumes.

This command enables a copy to be identified by its site.

The **-force** parameter with **rmvdiskcopy** is replaced by individual override parameters, making it clearer to the user exactly what protection they are bypassing.

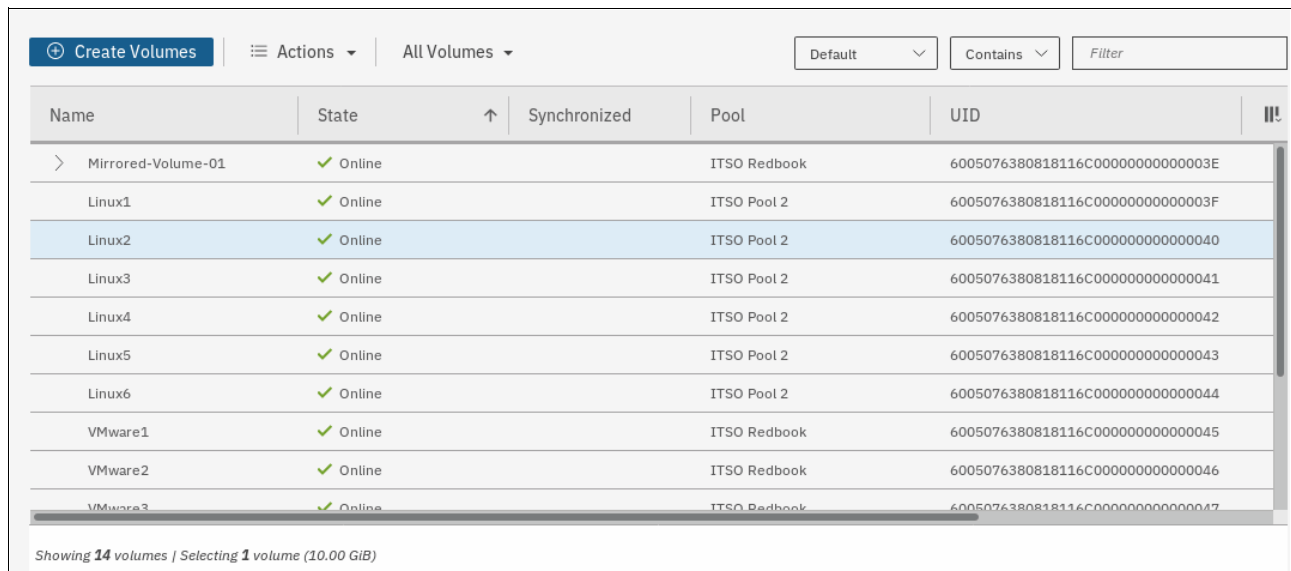
For more information, see [IBM Knowledge Center](#).

6.7 Mapping Volumes to Host after volume creation

Newly created volumes can be mapped to the host at creation time, or later. If the volume was not mapped to a host during creation, follow the steps that are described in 6.7.1, “Mapping newly created volumes to the host using the wizard” on page 342 to map it to a host.

6.7.1 Mapping newly created volumes to the host using the wizard

This section continues the process to map the volume that was created in 6.3, “Creating volumes by using the Volume Creation” on page 325. We assume that you followed that procedure and are on the Volumes pane that shows a list of volumes, as shown in Figure 6-36.



Name	State	Synchronized	Pool	UID
Mirrored-Volume-01	Online		ITSO Redbook	6005076380818116C00000000000003E
Linux1	Online		ITSO Pool 2	6005076380818116C00000000000003F
Linux2	Online		ITSO Pool 2	6005076380818116C000000000000040
Linux3	Online		ITSO Pool 2	6005076380818116C000000000000041
Linux4	Online		ITSO Pool 2	6005076380818116C000000000000042
Linux5	Online		ITSO Pool 2	6005076380818116C000000000000043
Linux6	Online		ITSO Pool 2	6005076380818116C000000000000044
VMware1	Online		ITSO Redbook	6005076380818116C000000000000045
VMware2	Online		ITSO Redbook	6005076380818116C000000000000046
VMware3	Online		ITSO Redbook	6005076380818116C000000000000047

Showing 14 volumes | Selecting 1 volume (10.00 GiB)

Figure 6-36 Volume list

To map a volume, complete the following steps:

1. Right-click the volume name to be mapped and select the **Map to Host or Host Cluster** menu option, as shown in Figure 6-37.

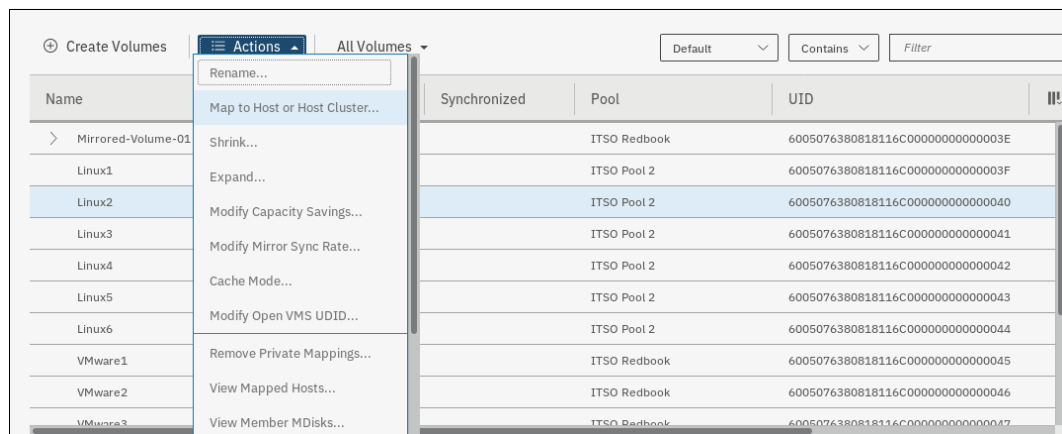


Figure 6-37 Map to host or host cluster

2. Select a host or a host cluster to which the new volume should be attached, as shown in Figure 6-38.

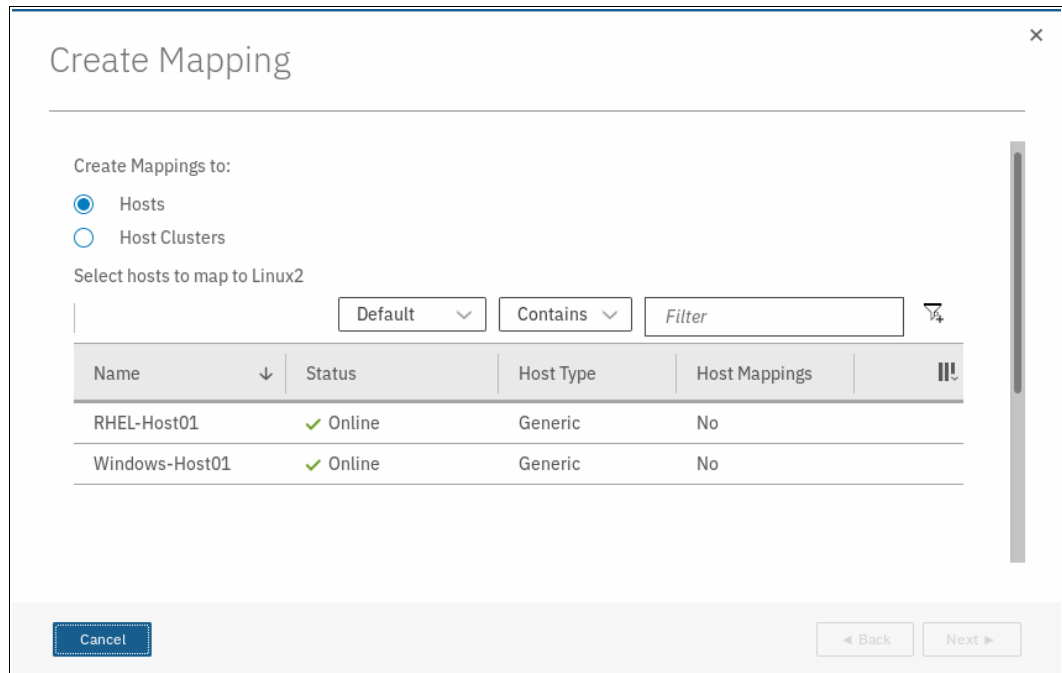


Figure 6-38 Select a host or a host cluster

Note: At this point, you can let the system assign a SCSI ID or choose it to assign it manually by selecting the **Self Assign** option. In this example, we chose to let system assign a SCSI ID.

3. Click **Next**. A summary window opens in which the volume to be mapped along with existing volumes that mapped to the host are listed, as shown in Figure 6-39.

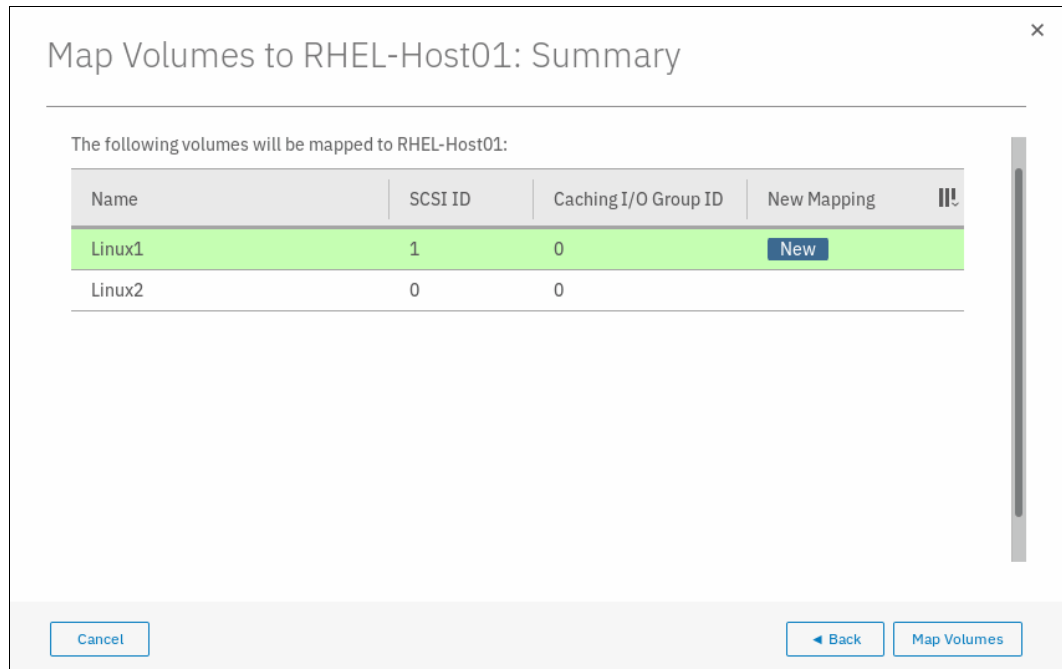


Figure 6-39 Map Volume summary

4. Click **Map Volumes**. A task completion window opens.
5. Click **close**.

6. You can verify the host mapping by clicking **Hosts** from the main window and selecting **Hosts**, as shown in Figure 6-40.

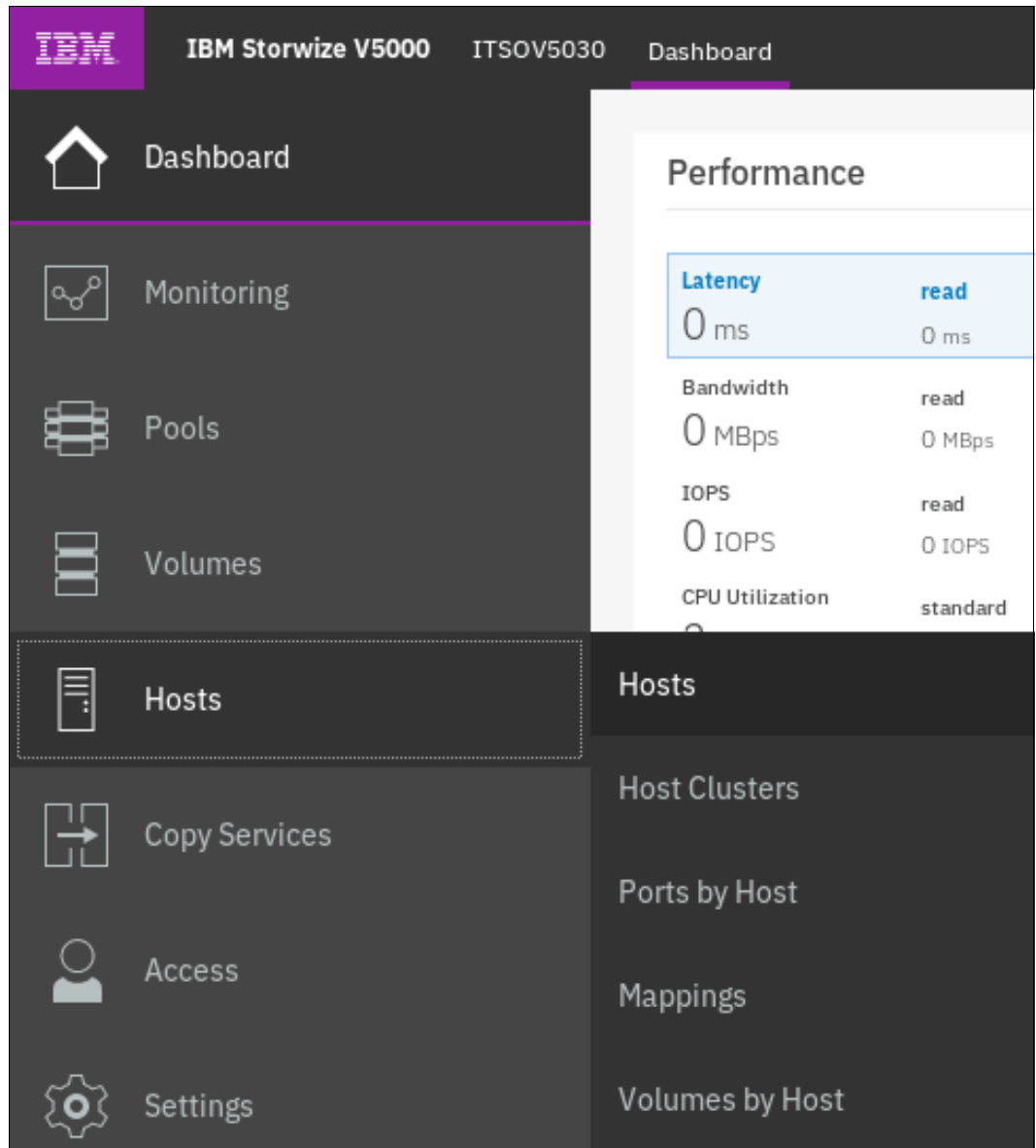


Figure 6-40 Hosts from main window

- Right-click the host to which the volume was mapped and select **Modify Volume Mappings**, as shown in Figure 6-41.

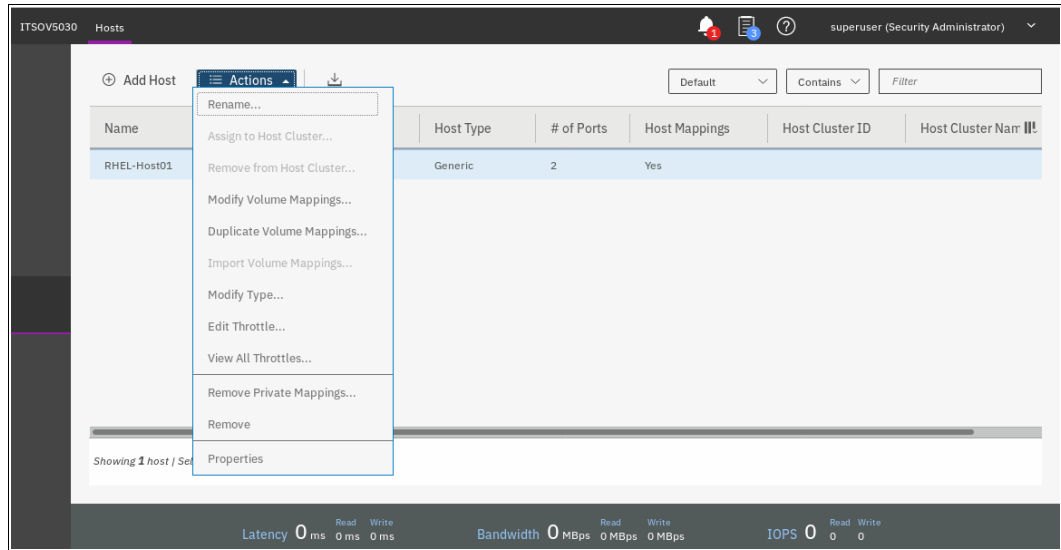


Figure 6-41 Modify host mappings

- A pop-up window opens that shows volumes that are mapped to the selected host, as shown in Figure 6-42.

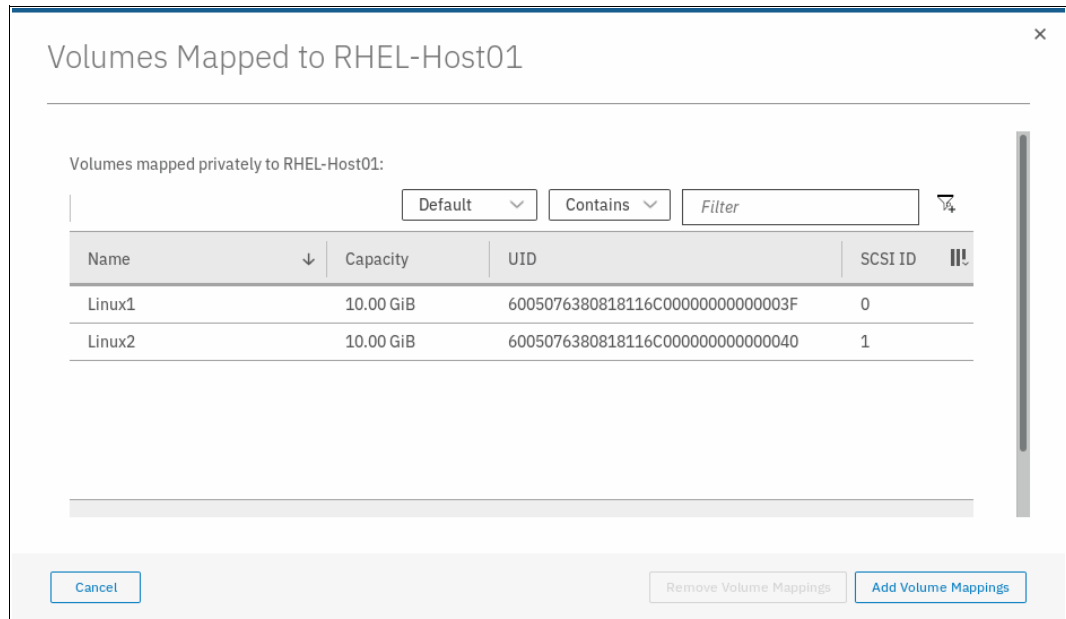


Figure 6-42 Volumes mapped to a host

The host now can access the volumes and store data on them. For information about discovering the volumes on the host and making more host settings, see 6.8, “Migrating a volume to another storage pool” on page 347.

Multiple volumes can also be created in preparation for discovering them later, and customize mappings.

6.8 Migrating a volume to another storage pool

IBM Spectrum Virtualize provides online volume migration while applications are running. Storage pools are managed disk groups, as described in Chapter 4, “Storage pools” on page 159. With volume migration, data can be moved between storage pools, regardless of whether the pool is an internal pool, or a pool on another external storage system. This migration is done without the server and application knowing that it even occurred.

The migration process is a low priority process that does not affect the performance of the IBM Spectrum Virtualize system. However, it moves one extent after another to the new storage pool, so the performance of the volume is affected by the performance of the new storage pool after the migration process.

To migrate a volume to another storage pool, complete the following steps:

1. Click **Volumes** in the main window and select the **Volumes** submenu option, as shown in Figure 6-43.

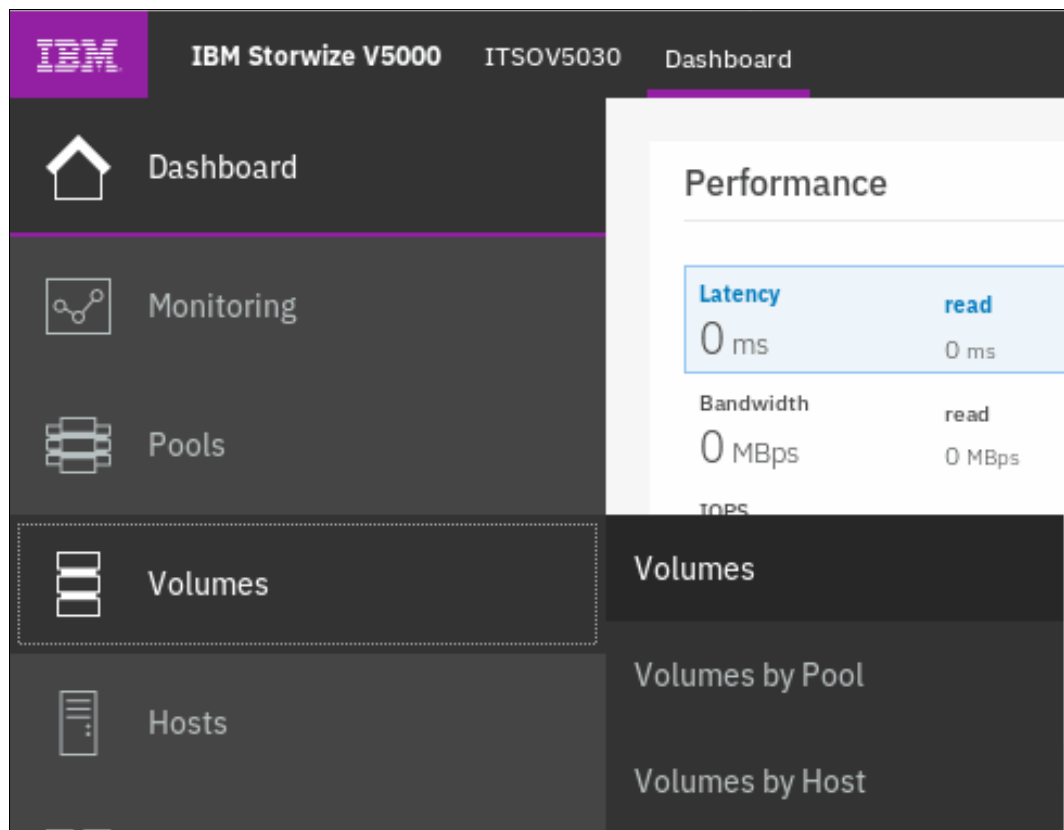


Figure 6-43 Volumes submenu option from main window

2. Right-click the wanted volume and select **Migrate to Another Pool**, as shown in Figure 6-44.

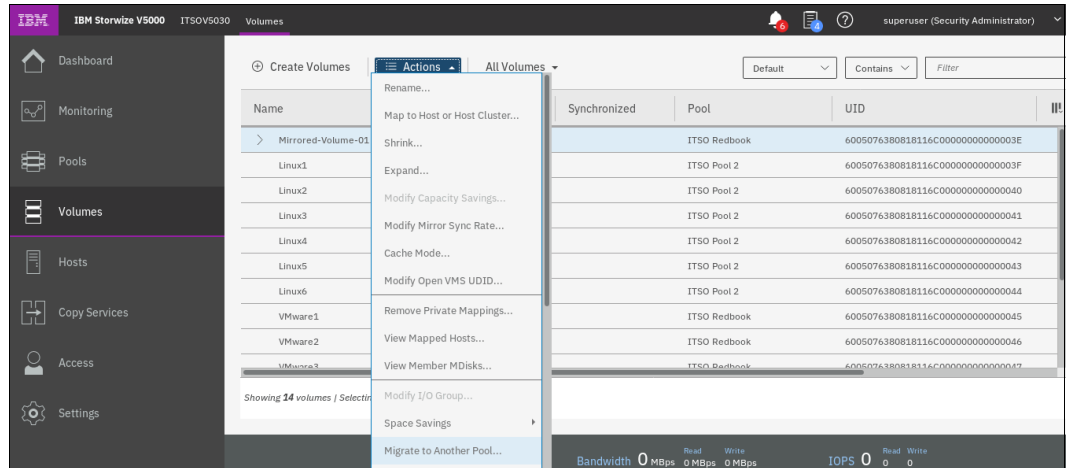


Figure 6-44 Selecting Migrate to Another Pool option

3. The Migrate Volume Copy window opens. If your volume consists of more than one copy, select the copy (from the menu that is shown in Figure 6-45) that you want to migrate to another storage pool. If the selected volume consists of one copy, this selection menu is not available.

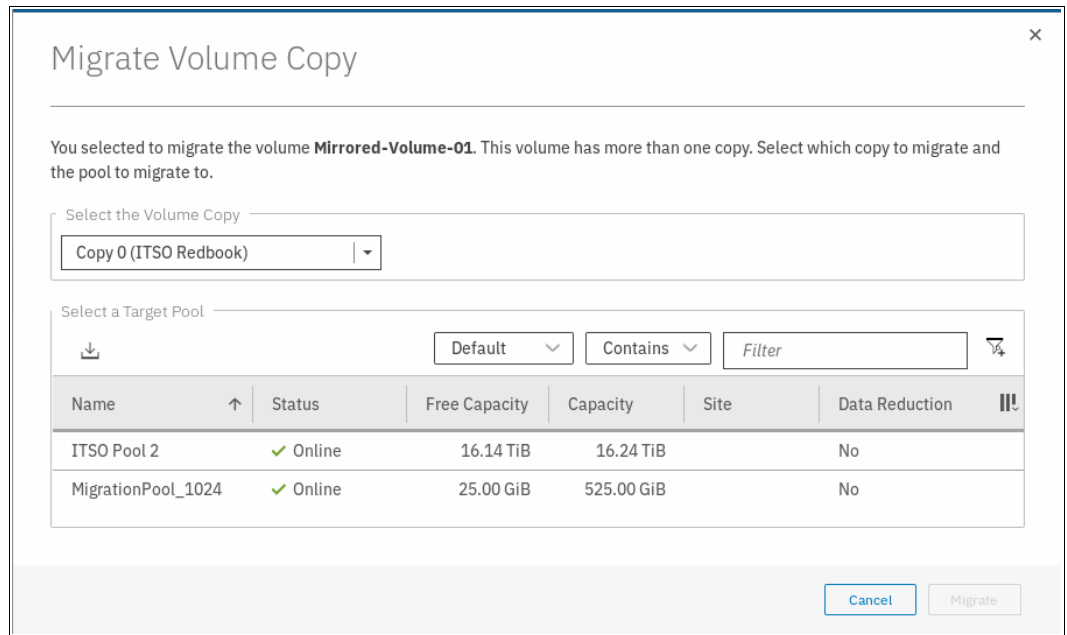


Figure 6-45 Migrate Volume Copy window: Select copy

- Select the new target storage pool, as shown in Figure 6-46.

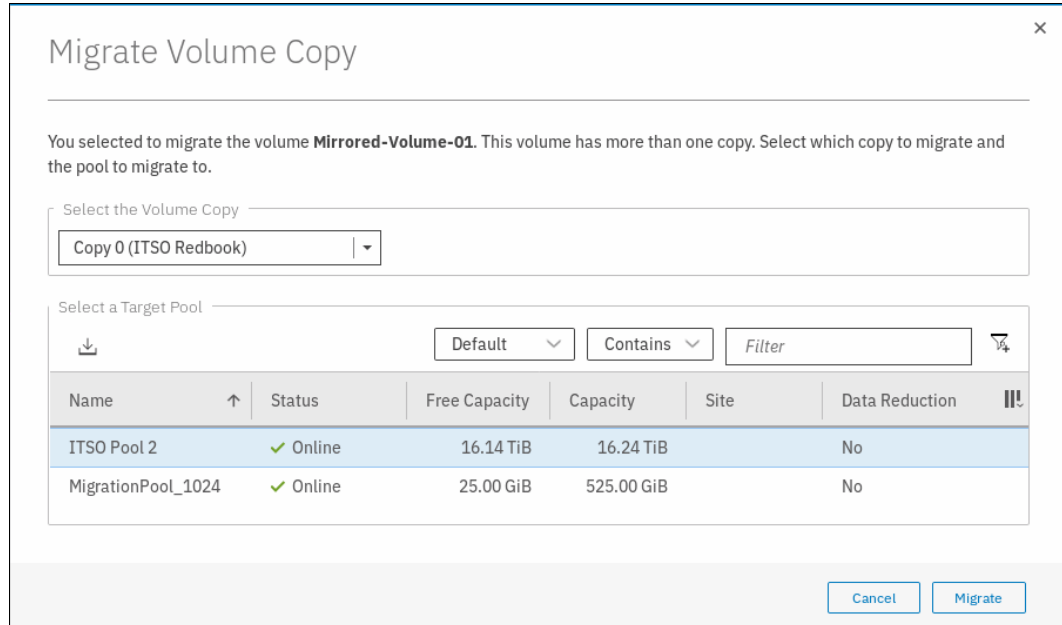


Figure 6-46 Selecting the new pool for volume migration

- Click **Migrate** and the volume copy migration starts. A task completion window opens. Then, click **Close** to return to the Volumes pane.
- Depending on the size of the volume, the migration process takes some time. However, you can monitor the status of the migration in the running tasks bar at the bottom of the GUI, as shown in Figure 6-47.

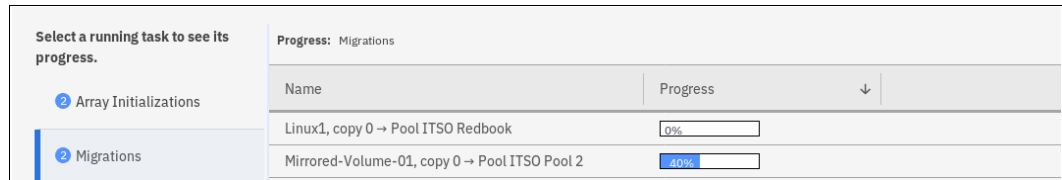


Figure 6-47 Migration progress

After the migration is completed, the volume is shown in the new storage pool. Figure 6-48 shows that it was moved from pool ITSO Redbook to the pool ITSO Pool 2.

Name	State	Synchronized	Pool	UID
<div style="display: flex; align-items: center;"> ∨ Mirrored-Volume-01 </div>	✓ Online		ITSO Pool 2	6005076380818116C00000000000003E
Copy 0*	✓ Online	Yes	ITSO Pool 2	6005076380818116C00000000000003E
Copy 1	✓ Online	Yes	ITSO Pool 2	6005076380818116C00000000000003E
Linux1	✓ Online		ITSO Redbook	6005076380818116C00000000000003F

Figure 6-48 Migration complete

The volume copy is now migrated without any host or application downtime to the new storage pool. It is also possible to migrate both volume copies to other pools online.

Another way to migrate volumes to another pool is by performing the migration by using the volume copies, as described in 6.9, “Migrating volumes using the volume copy feature” on page 350.

Note: Migrating a volume between storage pools with different extent sizes is *not* supported. If you need to migrate a volume to a storage pool with a different extent size, use volume copy features instead.

6.9 Migrating volumes using the volume copy feature

IBM Spectrum Virtualize supports creating, synchronizing, splitting, and deleting volume copies. A combination of these tasks can be used to migrate volumes to other storage pools.

The easiest way to migrate volume copies is to use the migration feature that is described in 6.8, “Migrating a volume to another storage pool” on page 347. If you use this feature, one extent after another is migrated to the new storage pool. However, the use of volume copies provides another possibility to migrate volumes.

To migrate a volume, complete the following steps:

1. Select the volume and select **Add Volume Copy** operation, as shown in Figure 6-49.

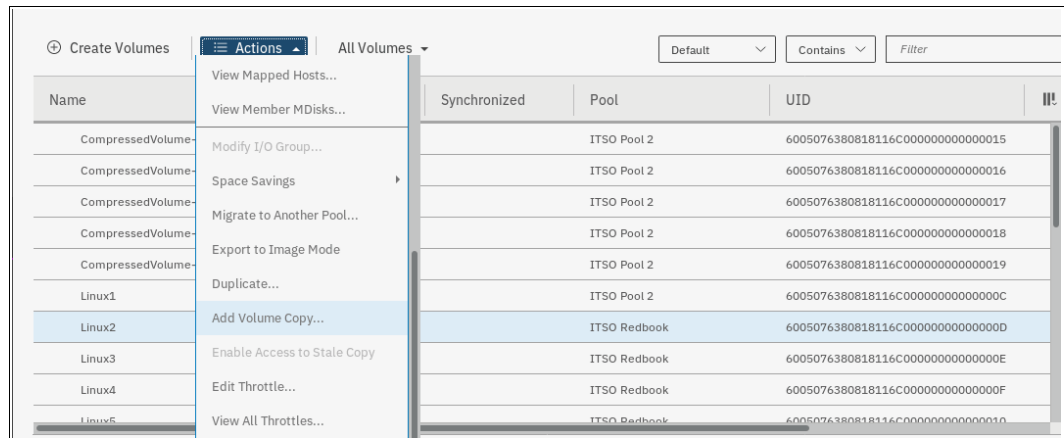


Figure 6-49 Adding the volume copy to another pool

2. Select the wanted pool into which a new copy is to be created, as shown in Figure 6-50.

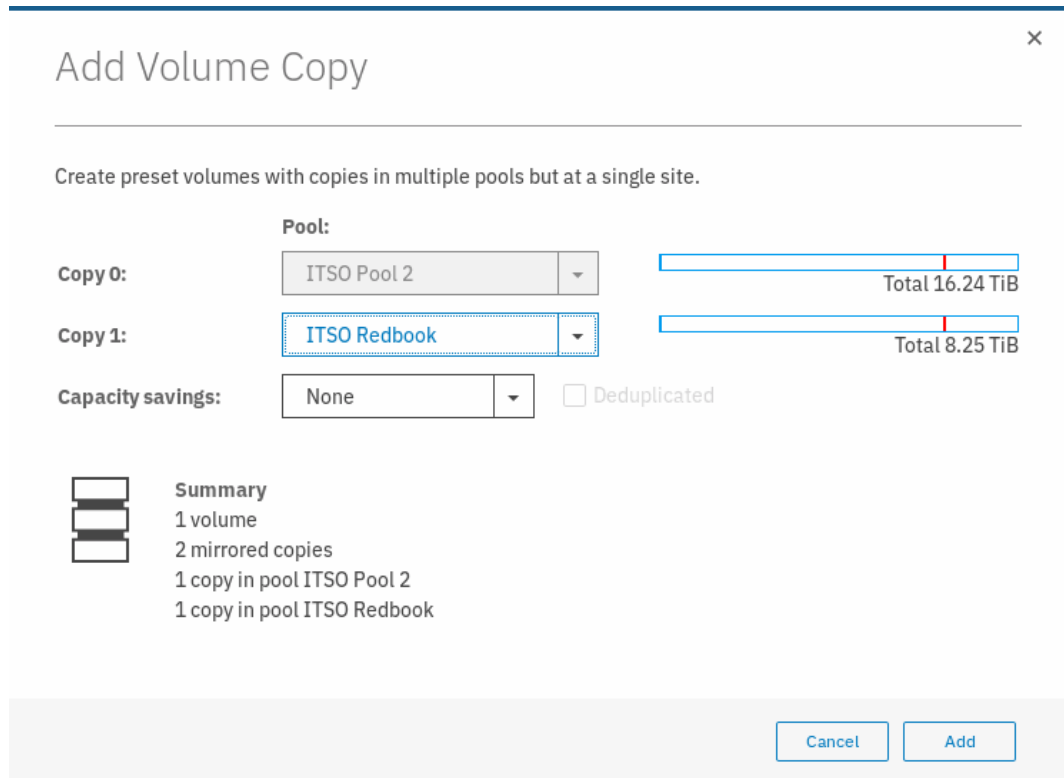


Figure 6-50 Add Volume Copy

3. A task completion window opens. Click **Close**.
4. Wait until the copies are synchronized. Then, change the role of the copies and make the new copy of the primary copy, as shown in Figure 6-51.

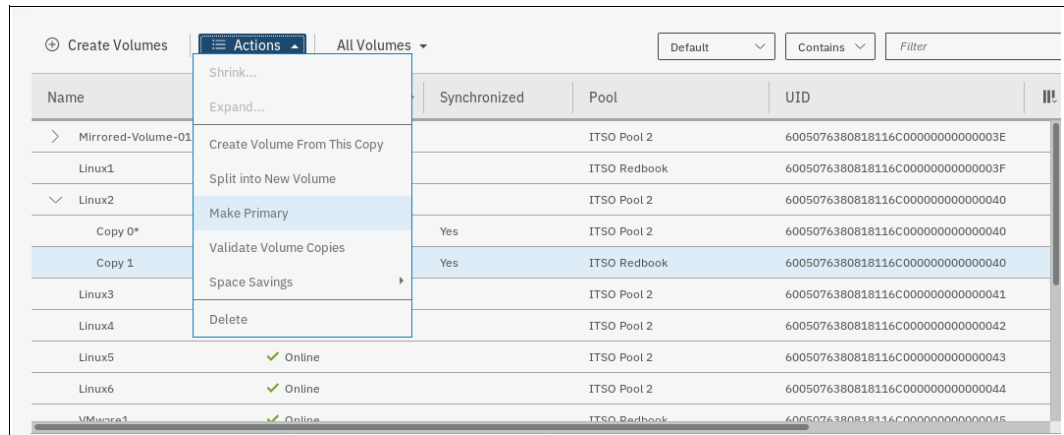


Figure 6-51 Making the new copy in a different storage pool as primary

5. A task completion window opens. Click **Close**.

6. Delete the old copy from the volume, as shown in Figure 6-52.

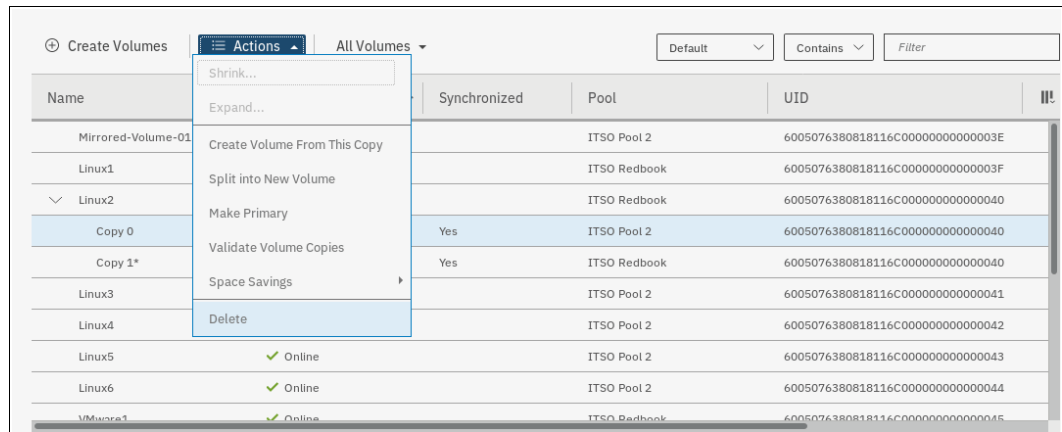


Figure 6-52 Deleting the older copy

7. Ensure that the new copy is in the target storage pool, as shown in Figure 6-53.

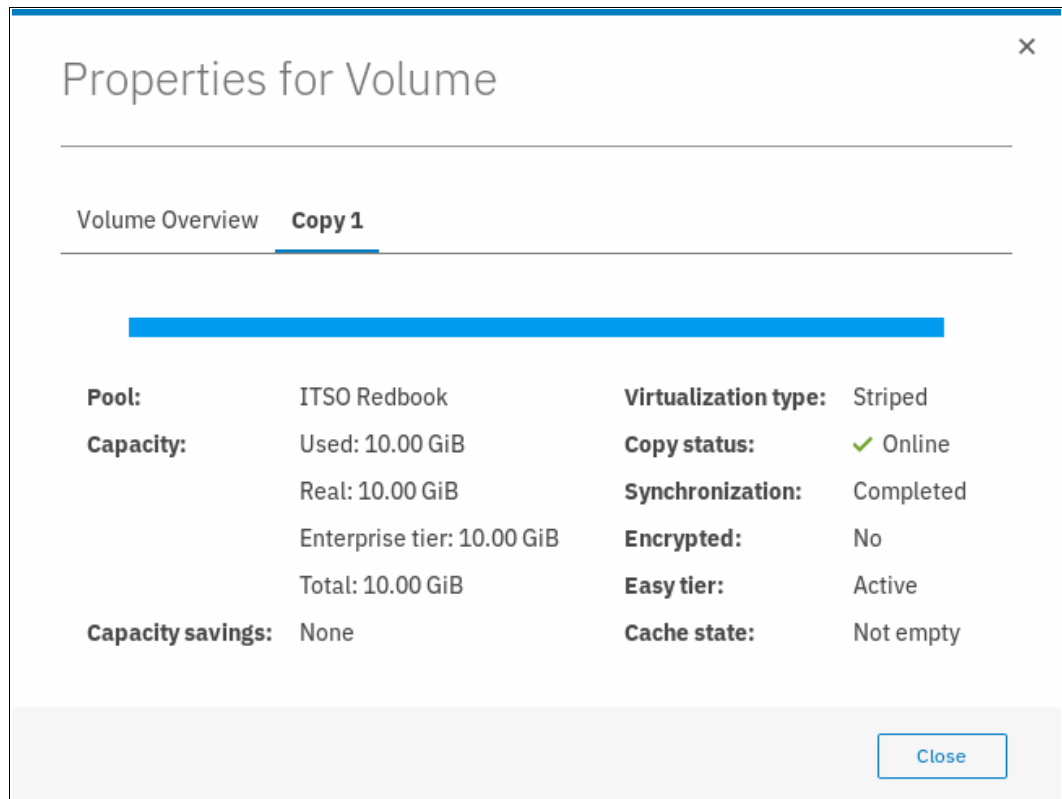


Figure 6-53 Verifying the new copy in the target storage pool

This migration process requires more user interaction, but it offers some benefits; for example, if you migrate a volume from a tier 1 storage pool to a lower performance tier 2 storage pool. In step 1, you create the copy on the tier 2 pool. All reads are still performed in the tier 1 pool to the primary copy. After the synchronization, all writes are destaged to both pools, but the reads are still only done from the primary copy.

Now you can switch the role of the copies online (step 3), and test the performance of the new pool. If you are done testing your lower performance pool, you can split or delete the old copy in tier 1, or switch back to tier 1 in seconds, in case tier 2 pool did not meet your performance requirements.

6.10 I/O throttling

You can set a limit on the number of I/O operations that are accepted for a volume. The limit is set in terms of I/O operations per second (IOPS) or MBps. By default, no I/O throttling rate is set when a volume is created. I/O throttling is also referred to as I/O governing.

Base the choice between I/O and MB as the I/O governing throttle on the disk access profile of the application. Database applications generally issue large amounts of I/O, but they transfer only a relatively small amount of data. In this case, setting an I/O governing throttle that is based on MBps does not achieve much. It is better to use an IOPS as a second throttle.

At the other extreme, a streaming video application generally issues a small amount of I/O, but it transfers large amounts of data. In contrast to the database example, setting an I/O governing throttle that is based on IOPS does not achieve much, so it is better to use an MBps throttle.

An I/O governing rate of 0 does not mean that zero IOPS (or MBps) can be achieved. It means that no throttle is set.

Note: Consider the following points:

- ▶ I/O governing does not affect FlashCopy and data migration I/O rates.
- ▶ I/O governing on a Metro Mirror and Global Mirror secondary volume does not affect the rate of data copy from the primary volume.

6.10.1 Defining throttle on a volume

To define a throttle on a volume, complete the following steps:

1. From the Volumes window, select the wanted volume to throttle, as shown in Figure 6-54.

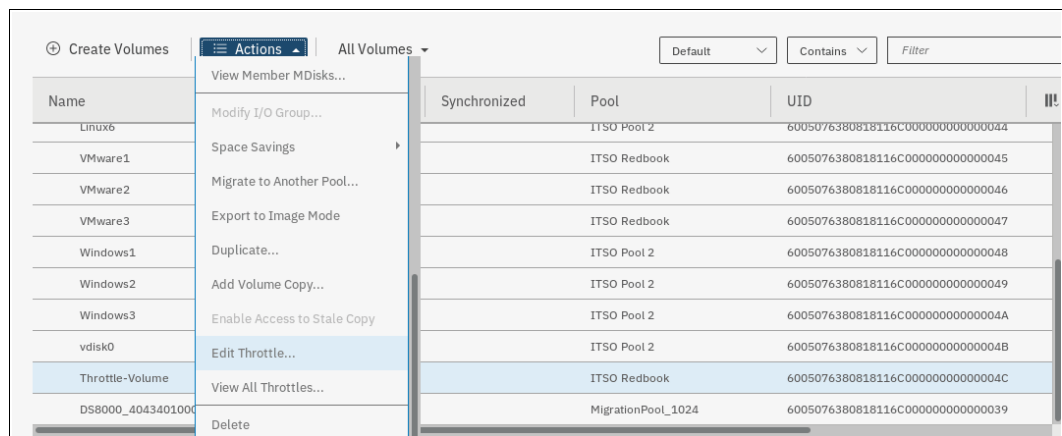


Figure 6-54 Edit Throttle

2. A window opens in which you can set the throttle in terms of IOPS or bandwidth (MBps) or both. In our example, we set the throttle on IOPS, as shown in Figure 6-55.

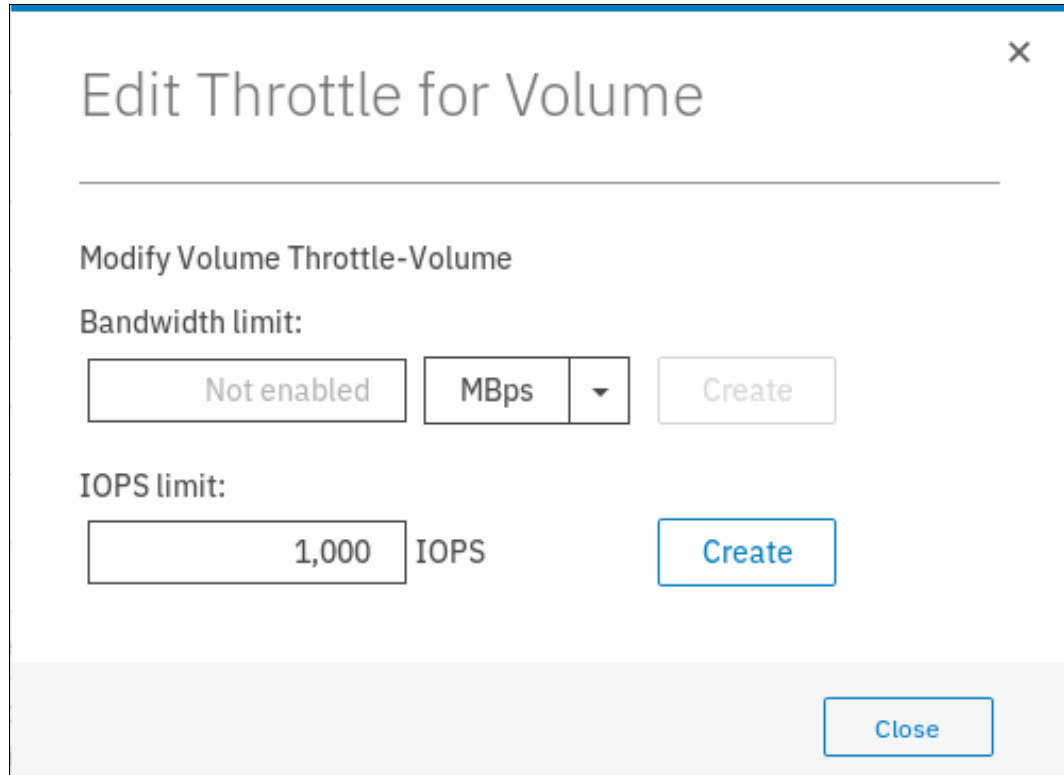


Figure 6-55 IOPS throttle on a volume

3. Click **Create** and a task completion window opens. Click **Close**.

6.10.2 Removing a throttle from a volume

To remove a throttle from a volume, complete the following steps:

1. From the Volumes window, select the wanted volume to remove throttle, as shown in Figure 6-56.

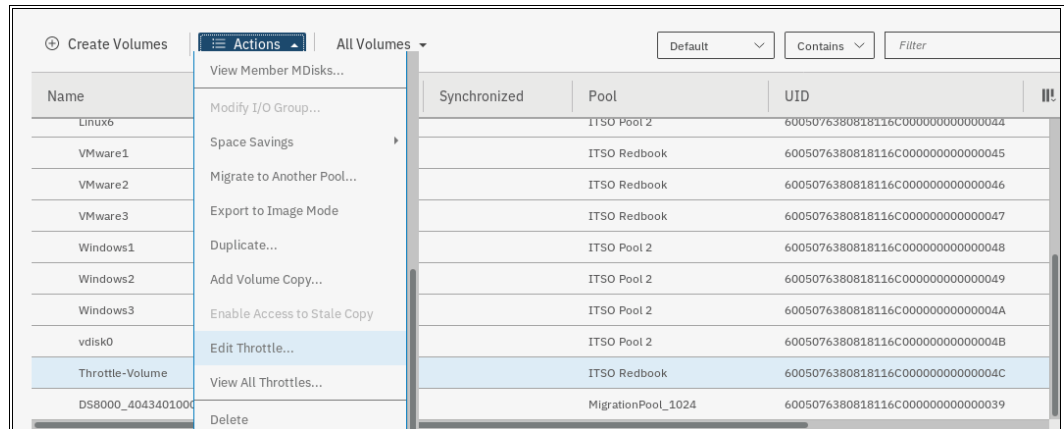


Figure 6-56 Edit Throttle

2. A window opens in which you can remove the throttle that was defined. In our example, we remove the throttle on IOPS by clicking **Remove**, as shown in Figure 6-57.

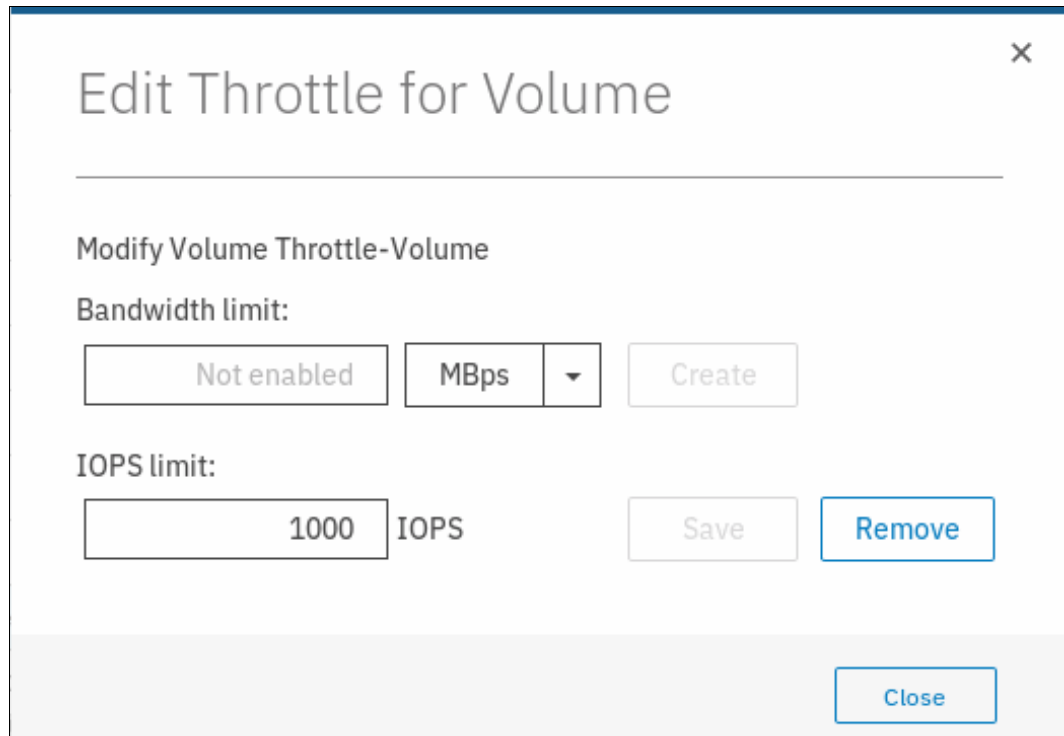


Figure 6-57 Remove throttle

3. A task completion window opens. Click **close**.



Storage migration

This chapter provides a detailed walkthrough of the storage migration wizard. The storage migration wizard migrates data from external storage systems to the internal capacity of the Storwize V5000 Gen2.

The data migration from other storage systems to the Storwize V5000 Gen2 consolidates storage and enables the benefits of the Storwize V5000 Gen2 features across all the volumes, such as the intuitive GUI, internal virtualization, thin provisioning, and FlashCopy.

This chapter includes the following topics:

- ▶ 7.1, “Storage migration wizard overview” on page 358
- ▶ 7.2, “Interoperation and compatibility” on page 358
- ▶ 7.3, “Storage migration wizard” on page 359

7.1 Storage migration wizard overview

The storage migration wizard uses the Volume Mirroring functionality to allow reads and writes during the migration, which eliminates disruption and downtime. After the end of the migration, the external storage can be retired. The storage migration wizard can also be used to migrate data from other Storwize systems, such as the Storwize V5000 Gen1 or Storwize V3700.

There are multiple reasons to use the Storwize V5000 Gen2 migration features:

- ▶ To redistribute workload within a clustered system across the disk subsystem
- ▶ To move workload onto newly installed storage
- ▶ To move workload off old or failing storage, ahead of decommissioning it
- ▶ To migrate data from an older disk subsystem to Storwize V5000 Gen2
- ▶ To migrate data from one disk subsystem to another disk subsystem

Specifically, this chapter provides information about the following topics:

- ▶ Interoperation and compatibility
- ▶ Storage migration wizard

For more information: For more information about the command-line interface setup, see Appendix A, “CLI setup and SAN Boot” on page 797.

For more information about migrating data manually, see Chapter 11, “External storage virtualization” on page 607.

7.2 Interoperation and compatibility

Interoperation is an important consideration when a new storage system is deployed into an environment that includes a storage infrastructure. This section describes how to check that the storage environment, storage system, and IBM Storwize V5000 Gen2 are ready for the data migration process.

To ensure interoperation and compatibility across all the elements that connect to the storage area network (SAN) fabric, check the proposed configuration with the IBM System Storage Interoperation Center (SSIC). SSIC can confirm whether a solution is supported and provide recommendations for hardware and software levels.

SSIC validates the components within a single storage solution. To confirm the interoperation between multiple storage solutions, it is necessary to request separate validation for each of them.

To require an interoperation validation within SSIC website, you must select **IBM System Storage Midrange Disk** in the Storage Family and select **Storwize V5000 Storage Controller Support** in the Storage Model. After defining the physical storage information, you must select the storage version and define the external storage that is to be attached to the Storwize V5000 Gen2.

Figure 7-1 shows the SSIC website window with the necessary values to request an interoperation validation for the Storwize V5000 Gen2.

IBM System Storage Interoperation Center (SSIC)

Start your search with ANY of the below selection boxes. You are NOT required to perform your query from the top down. Please view the details of your selected configuration. This requires clicking the Submit button or exporting your data.

Revise Selected Criteria - click link below to change search query
New Search > Storage Family > Storage Model >

Configuration Results = 2 890 [SSIC Education and Help](#)

Storage Family: IBM System Storage Midrange Disk
Storage Model: Storwize V5000 Storage Controller Support
Storage Version (Export Selected Version): 8 selections
Storage Controller: 402 selections

Submit

Configuration Results = 2 890 [SSIC Education and Help](#)

Figure 7-1 SSIC website window

For more information about SSIC, see this [IBM Support website](#).

If the required configuration is not listed for support in SSIC, contact your IBM marketing representative to ask for a request for price quotation (RPQ) for your specific configuration.

7.3 Storage migration wizard

The Storwize V5000 Gen2 storage migration wizard simplifies the migration task. The wizard features intuitive windows that guide users through the entire process.

7.3.1 External virtualization capability

All of the Storwize V5000 Gen2 models can migrate data from external storage controllers, including migrating from any other previous Storwize system generation. Storwize V5000 Gen2 uses the functionality that is provided by its external virtualization capability to perform the migration. This capability places external Fibre Channel (FC)-connected logical units (LUs) under the control of Storwize V5000 Gen2. After the volumes are virtualized, hosts continue to access them through the IBM Storwize V5000 Gen2, which acts as a proxy.

The difference between Storwize V5000 Gen2 models is that the Storwize V5010 and Storwize V5020 models can perform data migration only; external storage controllers cannot be virtualized on them. Storwize V5030 can be used to migrate data and externally virtualize storage from external storage controllers. For more information about external virtualization, see Chapter 11, “External storage virtualization” on page 607.

7.3.2 Model and adapter card considerations

Storage migration can use the FC SAN or direct-attach serial-attached Small Computer System Interface (SCSI) (SAS) connections to perform the data transfer.

FC migration to a Storwize V5000 Gen2 requires the purchase of a pair of the optional 16 Gb FC adapter cards.

SAS migration on the Storwize V5010 and Storwize V5030 systems requires the purchase of a pair of optional SAS adapter cards. SAS migration on V5020 can be performed without an adapter card by using the onboard SAS host attachment ports.

Table 7-1 lists the requirements for each model.

Table 7-1 Comparison of Storwize V5000 Gen2 models for storage migration

	Storwize V5010	Storwize V5020	Storwize V5030
External virtualization	Not supported	Not supported	Licensed feature
FC SAN migration	With 16 G FC cards	With 16 G FC cards	With 16 G FC cards
SAS device adapter (DA) migration	With 12 G SAS cards	Yes (onboard ports)	With 12 G SAS cards

7.3.3 Overview of the storage migration wizard

Consider the following points regarding the storage migration wizard process:

- ▶ Typically, storage systems segregate storage into many Small Computer System Interface (SCSI) LUs that are presented to hosts through a Fibre Channel SAN. Storage can also be presented through direct SAS attachment to a host. In general, the steps to migrate either one of these storage systems are the same.
- ▶ Input/output (I/O) to the Logical Unit Numbers (LUN) must be stopped and changes must be made to the mapping of the storage system LUs and to the SAN fabric zoning or SAS connectivity so that the original LUs are presented directly to the Storwize V5000 Gen2 and not to the hosts. Storwize V5000 Gen2 discovers the external LUs as unmanaged (not a member of any storage pools) managed disks (MDisks).
- ▶ The unmanaged MDisks are then imported to the Storwize V5000 Gen2 as image mode MDisks and placed in a storage pool. This storage pool is now a logical container for the externally attached LUs.
- ▶ Each volume has a one-to-one mapping with an image mode MDisk. From a data perspective, the image mode volume represents the SAN-attached LUs exactly as they were before the import operation. The image mode volumes are on the same physical drives of the storage system and the data remains unchanged. Storwize V5000 Gen2 presents active images of the SAN-attached LUs and it acts as a proxy.
- ▶ You must remove the storage system multipath device driver from the hosts. Then, configure the hosts for the Storwize V5000 Gen2 attachment. If you are migrating over Fibre Channel, further zoning changes are made for host-to-V5000 Gen2 SAN connections. The Storwize V5000 Gen2 hosts are defined with worldwide port names (WWPNs) and the volumes are mapped to the hosts. After the volumes are mapped, the hosts discover the Storwize V5000 Gen2 volumes through a host rescan or reboot operation.
- ▶ The Storwize V5000 Gen2 volume mirror operations are then initiated. The image mode volumes are mirrored to generic volumes. The mirrors are online migration tasks, which means that a defined host can still access and use the volumes during the mirror synchronization process.
- ▶ After the mirror operations are complete, the volume mirror relationships and the image mode volumes are removed. The other storage system LUs are migrated and the now redundant existing storage can be retired or reused elsewhere.

Important: If you are migrating volumes from another IBM Storwize product, be aware that the target Storwize V5000 Gen2 system must be configured at the replication layer for the source to discover the target system as a host. The default layer setting for the Storwize V5000 Gen2 is storage.

Ensure that the Storwize V5000 Gen2 system is configured as a replication layer system. Enter the following command:

```
chsystem -layer replication
```

If you do not enter this command, you cannot add the target system as a host on the source storage system or see source volumes on the target system.

To change the source Storwize system to the storage layer, enter the following command:

```
chsystem -layer storage
```

For more information about layers and how to change them, see Chapter 10, “Copy Services” on page 465.

7.3.4 Storage migration wizard tasks

The storage migration wizard is designed for an easy migration of data from other storage systems to the internal capacity of the Storwize V5000 Gen2.

This section describes the following storage migration wizard tasks:

- ▶ Avoiding data loss
- ▶ Verifying prerequisites for Fibre Channel connections
- ▶ Verifying prerequisites for SAS connections
- ▶ Verifying prerequisites for iSCSI connections
- ▶ Identifying restrictions and prerequisites for the wizard
- ▶ Preparing the environment for migration
- ▶ Mapping storage
- ▶ Migrating MDisks
- ▶ Configuring hosts
- ▶ Mapping volumes to hosts
- ▶ Selecting a storage pool
- ▶ Finishing the storage migration wizard
- ▶ Finalizing migrated volumes

Avoiding data loss

It is prudent to avoid any potential data loss by creating a backup of all of the data that is stored on the hosts, the storage systems, and the Storwize V5000 Gen2 before you use the storage migration wizard.

Verifying prerequisites for Fibre Channel connections

Cable this system into the SAN of the external storage that you want to migrate. Ensure that your system is cabled and zoned into the same storage area network (SAN) as the external storage system that you are migrating.

If you are using Fibre Channel, connect the Fibre Channel cables to the Fibre Channel ports in both canisters of your system, and then to the Fibre Channel network. If you are using Fibre Channel over Ethernet (FCoE), connect Ethernet cables to the 10 Gbps Ethernet ports.

Ensure that all systems are running a software level that enables them to recognize the other nodes in the cluster. Also, ensure that the systems use Fibre Channel adapters at the same speed. To avoid performance bottlenecks, do not use a combination of 8 Gbps and 16 Gbps links.

Examples of how to connect a Storwize V5000 Gen1 to a Storwize V5030 system are shown in Figure 7-2 and Figure 7-3.

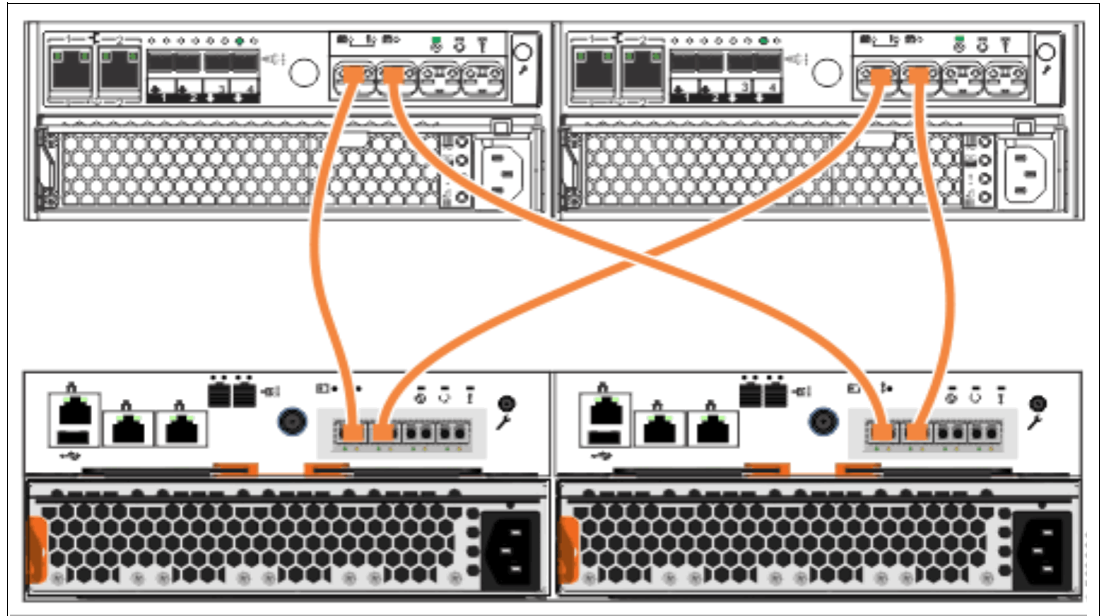


Figure 7-2 Direct Fibre Channel connections between systems

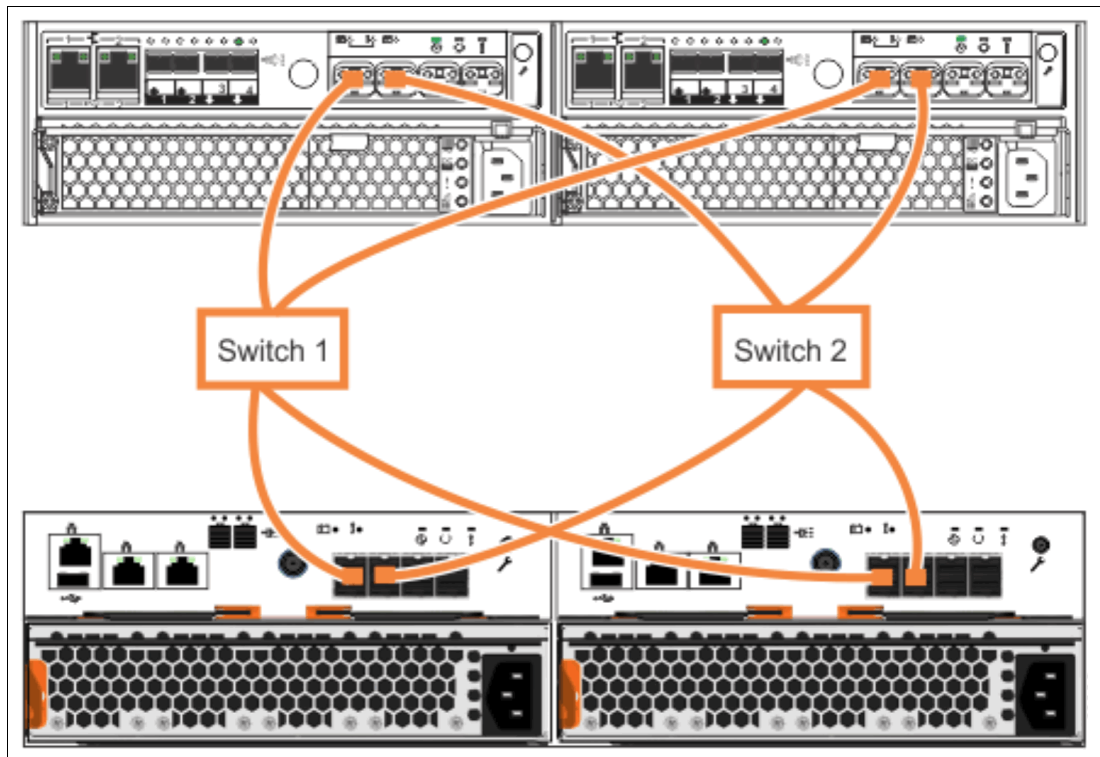


Figure 7-3 Fibre Channel connections using switches between the systems

Verifying prerequisites for SAS connections

For migrating from a Storwize V3500, Storwize V3700, or Storwize V5000 Gen1, ensure that all systems are running a software level that can support SAS migration.

Cable the Storwize V5000 Gen2 directly to the external storage system that you want to migrate. Depending on the Storwize V5000 Gen2 model, the cabling differs slightly. The Storwize V5010 and Storwize V5030 need four SAS cables (two SAS cables per node canister) that are connected to the optional SAS card. The Storwize V5020 needs four SAS cables (two SAS cables per node canister) that are connected to SAS port 2 and SAS port 3.

The Storwize V3500 or Storwize V3700 source systems require two cables per node canister. Each canister must be connected to each Storwize V5000 Gen2 canister. On the V3500 or V3700, you can use SAS ports 1, 2, or 3. Do not use SAS port 4.

Examples of how to connect a Storwize V5000 Gen2 to the Storwize V3500/V3700 are shown in Figure 7-4, Figure 7-5, and Figure 7-6 on page 364.

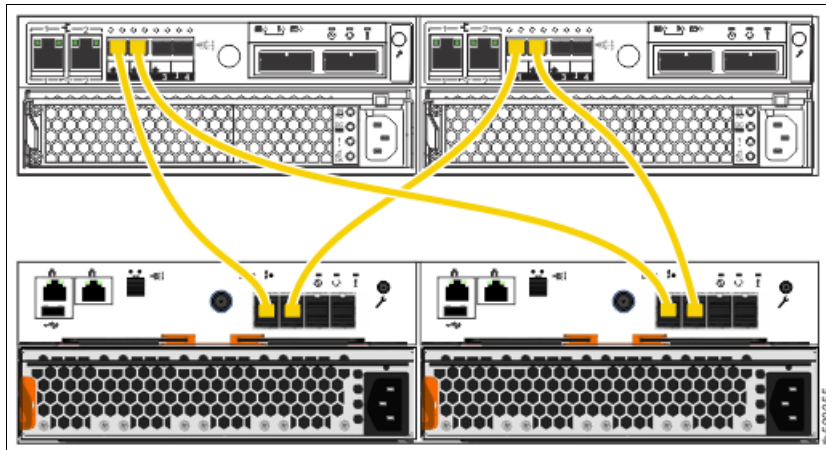


Figure 7-4 Connecting SAS cables from a Storwize V3500 or V3700 to a Storwize V5010 system

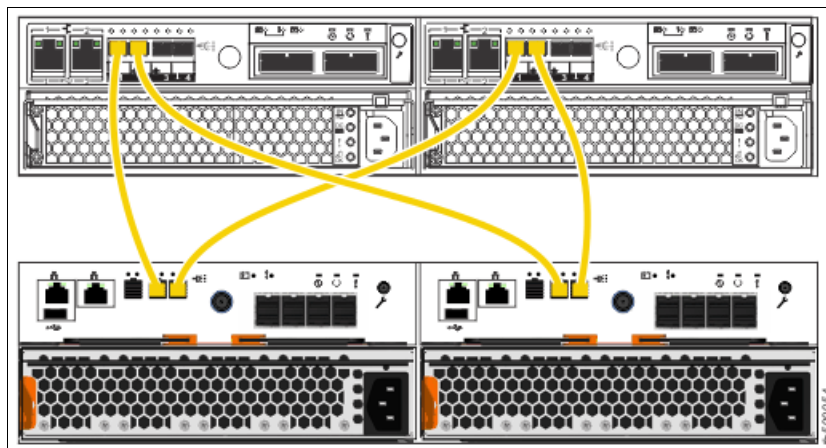


Figure 7-5 Connecting SAS cables from a Storwize V3500 or V3700 to a Storwize V5020 system

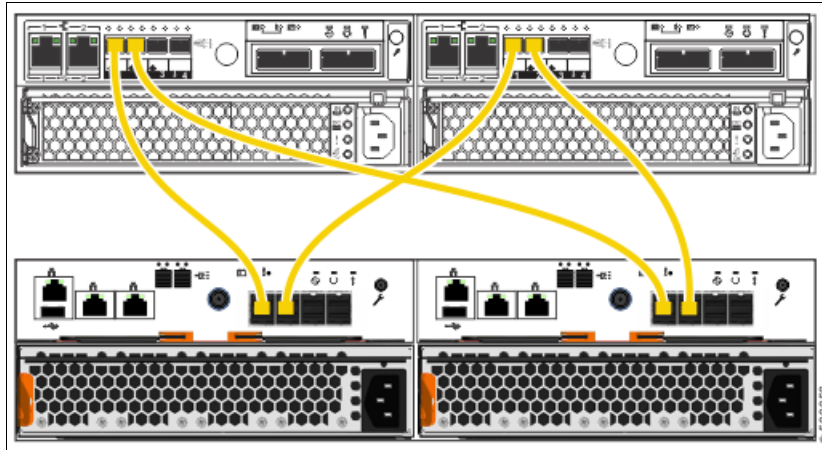


Figure 7-6 Connecting SAS cables from a Storwize V3500 or V3700 to a Storwize V5030 system

Storwize V5000 Gen1 source systems require two cables per node canister. Each canister must be connected to each Storwize V5000 Gen2 canister. On the V5000 Gen1, you must use SAS port 1 or 2. Do *not* use SAS port 3 or 4.

Examples of how to connect a Storwize V5000 Gen2 to a Storwize V5000 Gen1 are shown in Figure 7-7, Figure 7-8 on page 365, and Figure 7-9 on page 365.

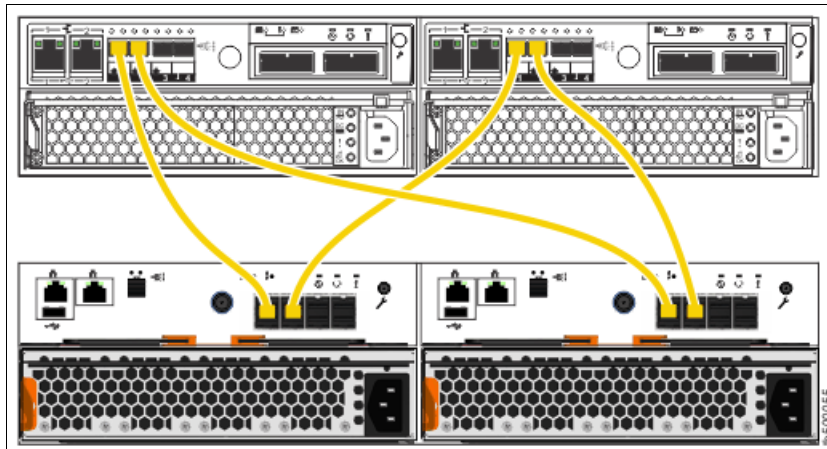


Figure 7-7 Connecting SAS cables from a Storwize V5000 Gen1 system to a Storwize V5010 system

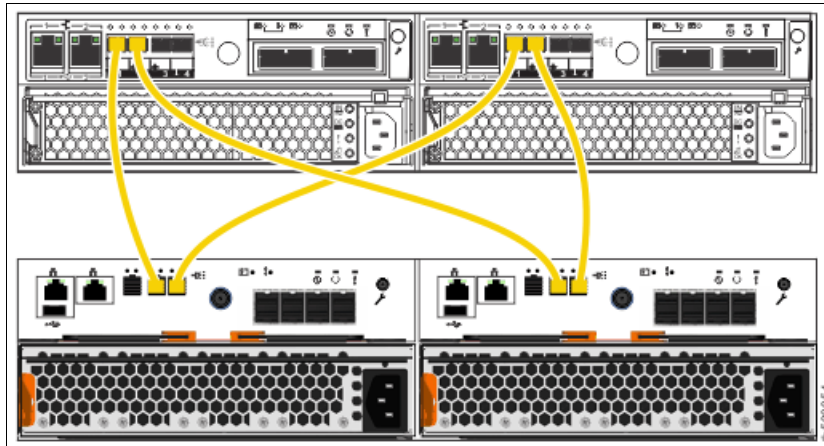


Figure 7-8 Connecting SAS cables from a Storwize V5000 Gen1 system to a Storwize V5020 system

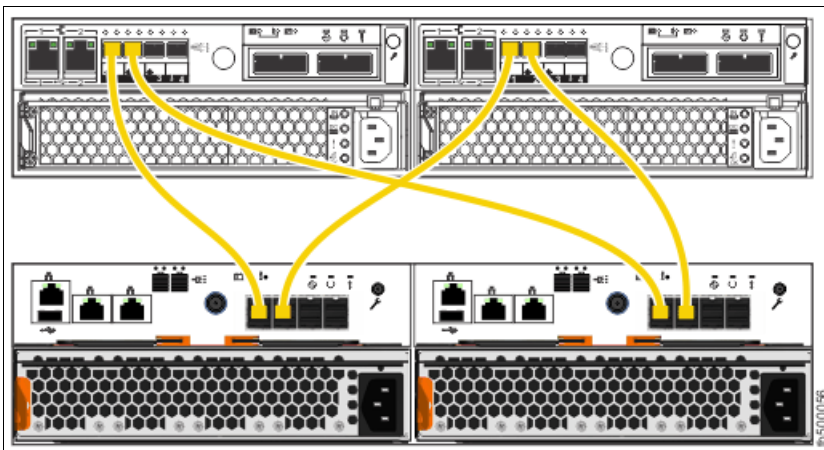


Figure 7-9 Connecting SAS cables from a Storwize V5000 Gen1 system to a Storwize V5030 system

Migration from DS3200 and DS3500: For more information about how to cable and perform migration from a DS3200 or DS3500 by using SAS, see the Storwize V5000 Gen2 [IBM Knowledge Center](#).

Although the Storwize V5000 Gen2 IBM Knowledge Center shows an SAS cabling example related to IBM Storwize V5000 Gen2 V7.8.1, the solution also applies to IBM Storwize V5000 Gen2 V8.2.0.0.

Verifying prerequisites for iSCSI connections

On Storwize V5000 Gen2 systems, you can use iSCSI connections to migrate data from different systems (Storwize V5010, V5020, and V5030) and to virtualize external storage systems (Storwize V5030).

Migration considerations and configurations can vary depending on the type of system to be migrated or virtualized. You can use an iSCSI attachment to migrate data from a Storwize family system to a Storwize V5000 Gen2 system. Storwize V5000 Gen2 does not support iSCSI connections to migrate data from Storwize V3500.

You can use any available Ethernet port to establish iSCSI connectivity between the Storwize V5000 Gen2 system and the backend storage controller.

Note: If you are using onboard Ethernet ports on a Storwize V5010 or Storwize V5020 system, ensure that the onboard Ethernet port 2 on the system is not configured to be used as the technician port.

To avoid performance bottlenecks, the iSCSI initiator and target systems must use Ethernet ports at the same speed. Do not use a combination of Ethernet links that run at different speeds.

For full redundancy and increased throughput, use two or more Ethernet switches. Similarly numbered Ethernet ports on each node of each system must be connected to the same switch. They must also be configured on the same subnet or VLAN.

Figure 7-10 shows iSCSI connections between a Storwize V5030 system (iSCSI initiator) and a Storwize V3700 system (iSCSI target).

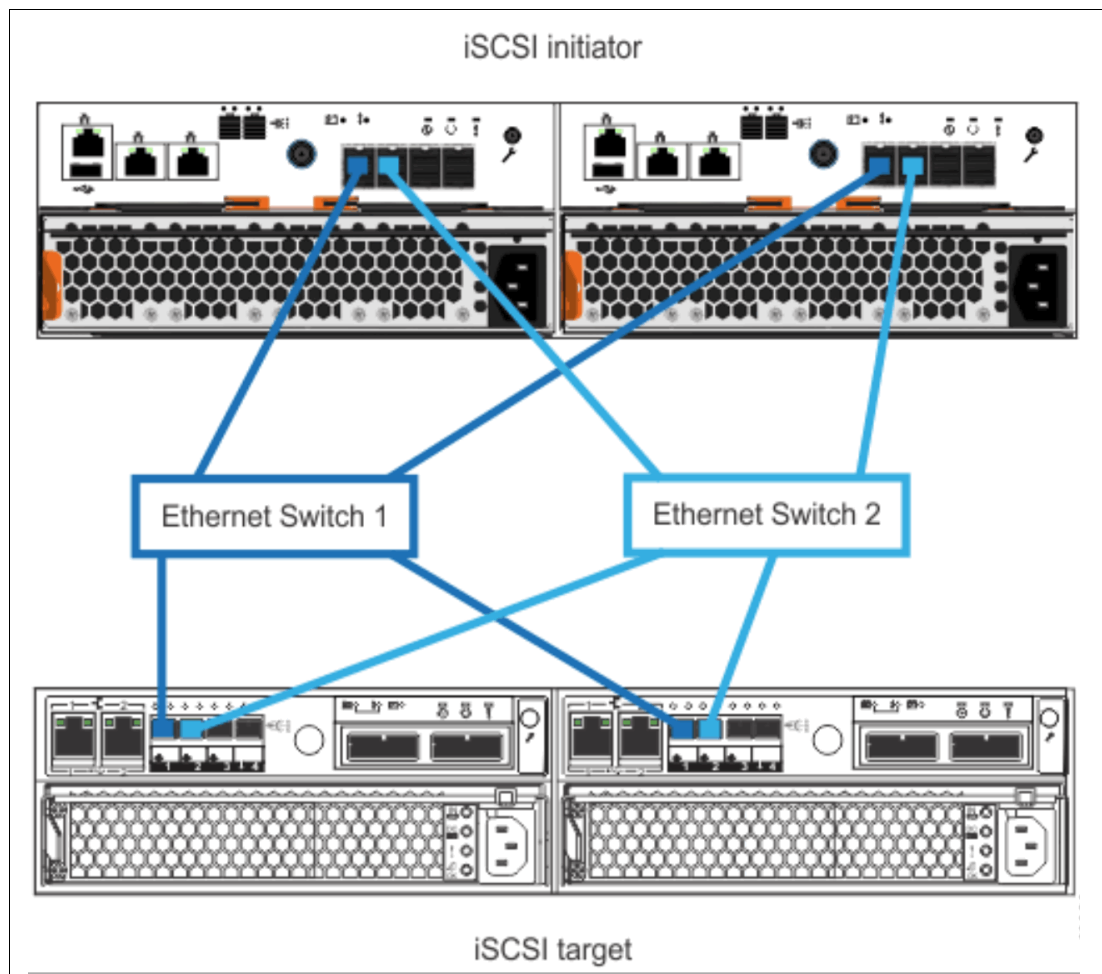


Figure 7-10 iSCSI connections between a Storwize V5030 system and a Storwize V3700 system

Accessing the storage migration wizard

Select **System Migration** in the Pools menu (see Figure 7-11) to open the System Migration window.

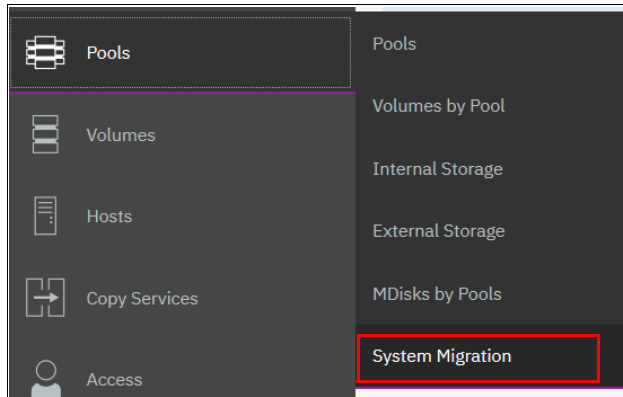


Figure 7-11 Pools menu

The System Migration window provides access to the storage migration wizard and displays the migration progress information. Click **Start New Migration** to begin the storage migration wizard. Figure 7-12 shows the System Migration window.

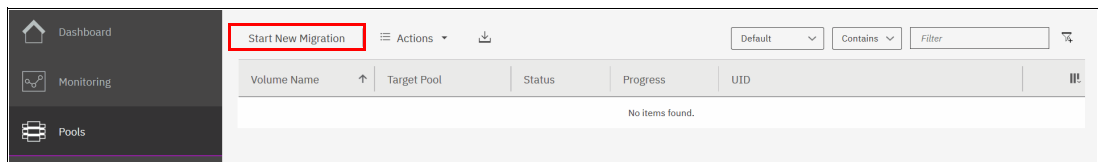


Figure 7-12 System Migration window

Important: Consider the following points:

- ▶ You might receive a warning message as shown in Figure 7-13 that indicates that no externally attached storage controllers were found if you did not configure your zoning correctly (or if the layer was incorrectly set if another Storwize system is attached). Click **Close** and correct the problem before you start the migration wizard again.
- ▶ The subsequent windows in the migration wizard, as shown in Figure 7-15 on page 370, direct you to remove the host zoning to the external storage and create zones between the Storwize V5000 Gen2 and the external storage. However, these steps must be performed *before* you start the wizard.

For more information about the steps you must complete before you start the data migration wizard, see “Preparing the environment for migration” on page 370 and “Mapping storage” on page 371.

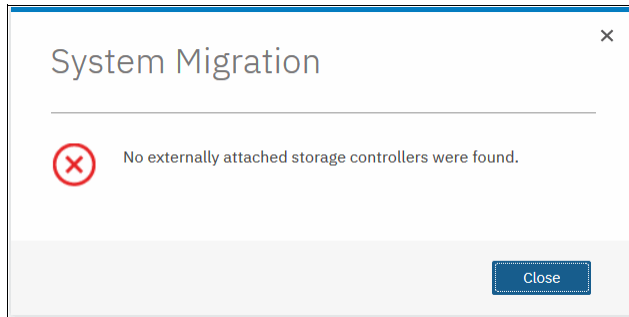


Figure 7-13 Error message that is displayed when no external storage is found

Identifying restrictions and prerequisites for the wizard

This window of the storage migration wizard describes the restrictions and prerequisites for the wizard, as shown in Figure 7-14.

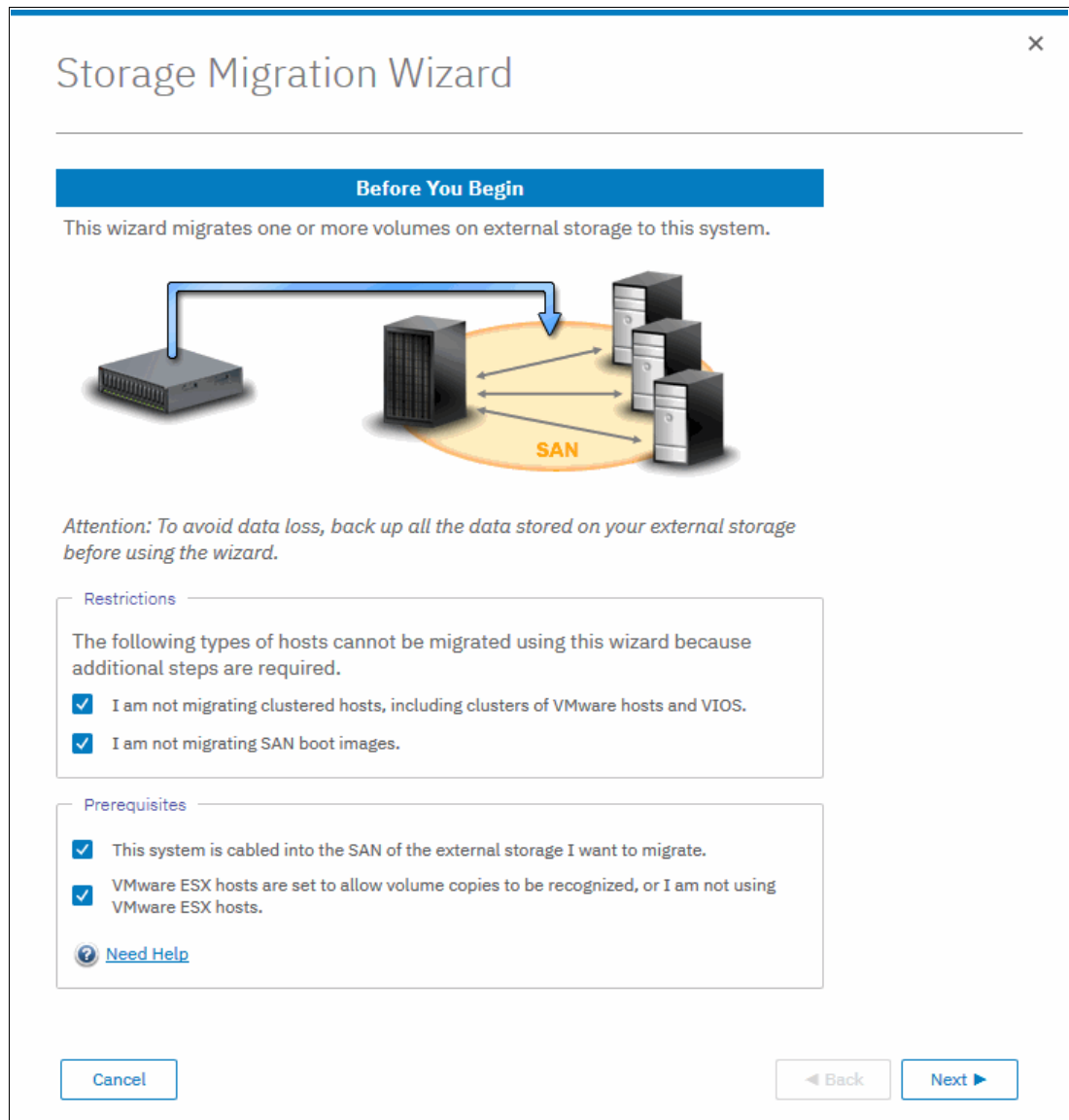


Figure 7-14 Before you begin the storage migration wizard

Restrictions

Confirm that the following conditions are met:

- ▶ You are not using the storage migration wizard to migrate cluster hosts, including clusters of VMware hosts and Virtual I/O Servers (VIOS).
- ▶ You are not using the storage migration wizard to migrate SAN Boot images.

If you identify that any of the restriction options cannot be selected, the migration must be performed outside of this wizard because more steps are required. For more information, see the Storwize V5000 Gen2 [IBM Knowledge Center](#).

The VMware vSphere Storage vMotion feature might be an alternative for migrating VMware clusters. For more information, see [this website](#).

For more information about migrating SAN Boot images, see Appendix A, “CLI setup and SAN Boot” on page 797.

Prerequisites

Confirm that the following prerequisites are met:

- ▶ Ensure that the Storwize V5000 Gen2, existing storage system, hosts, and Fibre Channel ports are physically connected to the SAN fabrics.
- ▶ If VMware ESX hosts are involved in the data migration, ensure that the VMware ESX hosts are set to allow volume copies to be recognized. For more information, see the VMware ESX product documentation at [this website](#).

If all options can be selected, click **Next**. In all other cases, the option is not available and the data must be migrated without the use of this wizard.

Preparing the environment for migration

Figure 7-15 shows the Prepare Environment for Migration window. Carefully follow the instructions that are provided here. When all of the required tasks are complete, click **Next**.

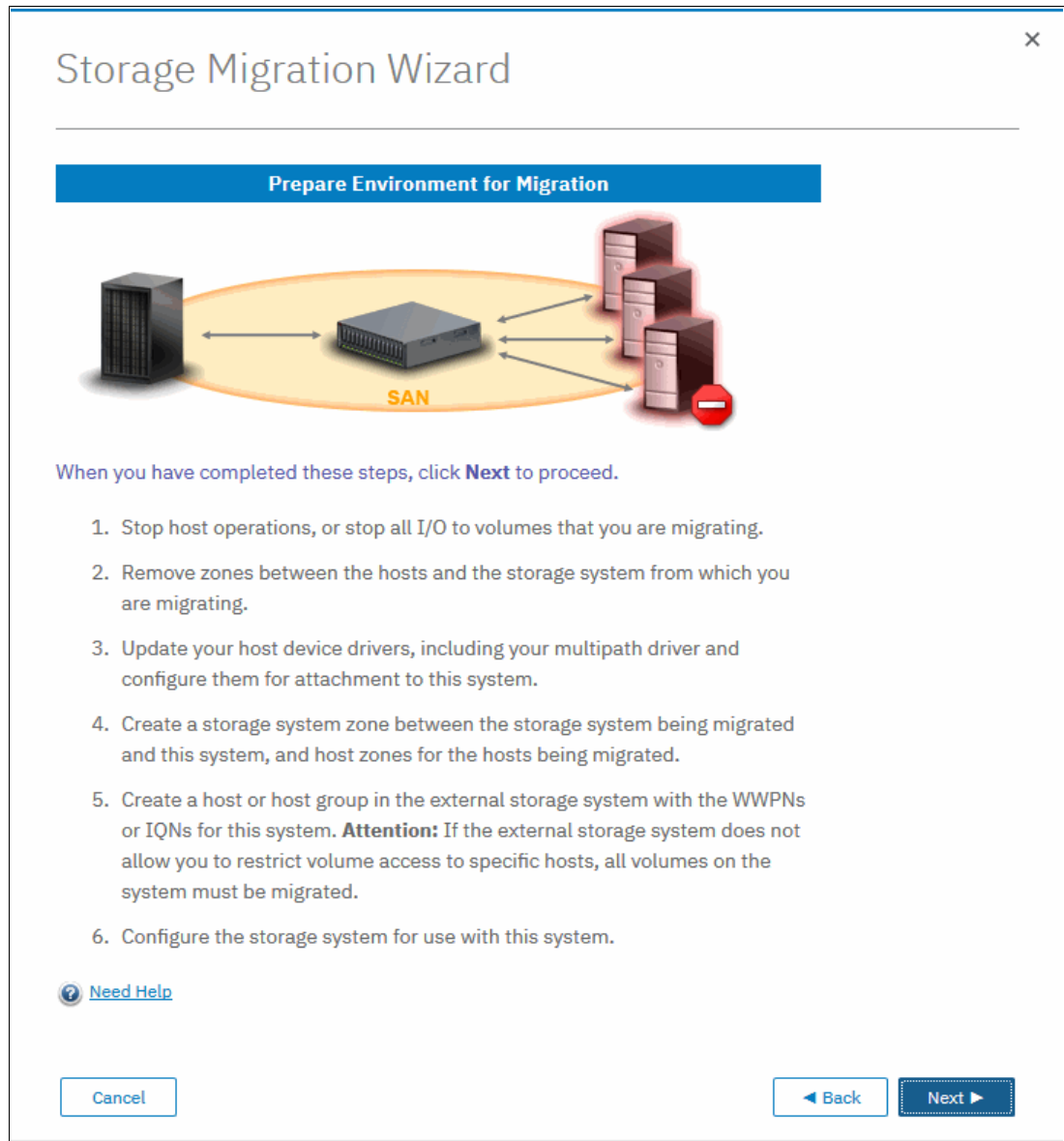


Figure 7-15 Prepare the migration environment

Mapping storage

Follow the instructions that are shown in the Map Storage window that is shown in Figure 7-16 and click **Next**. Record all of the details carefully because the information is used later.

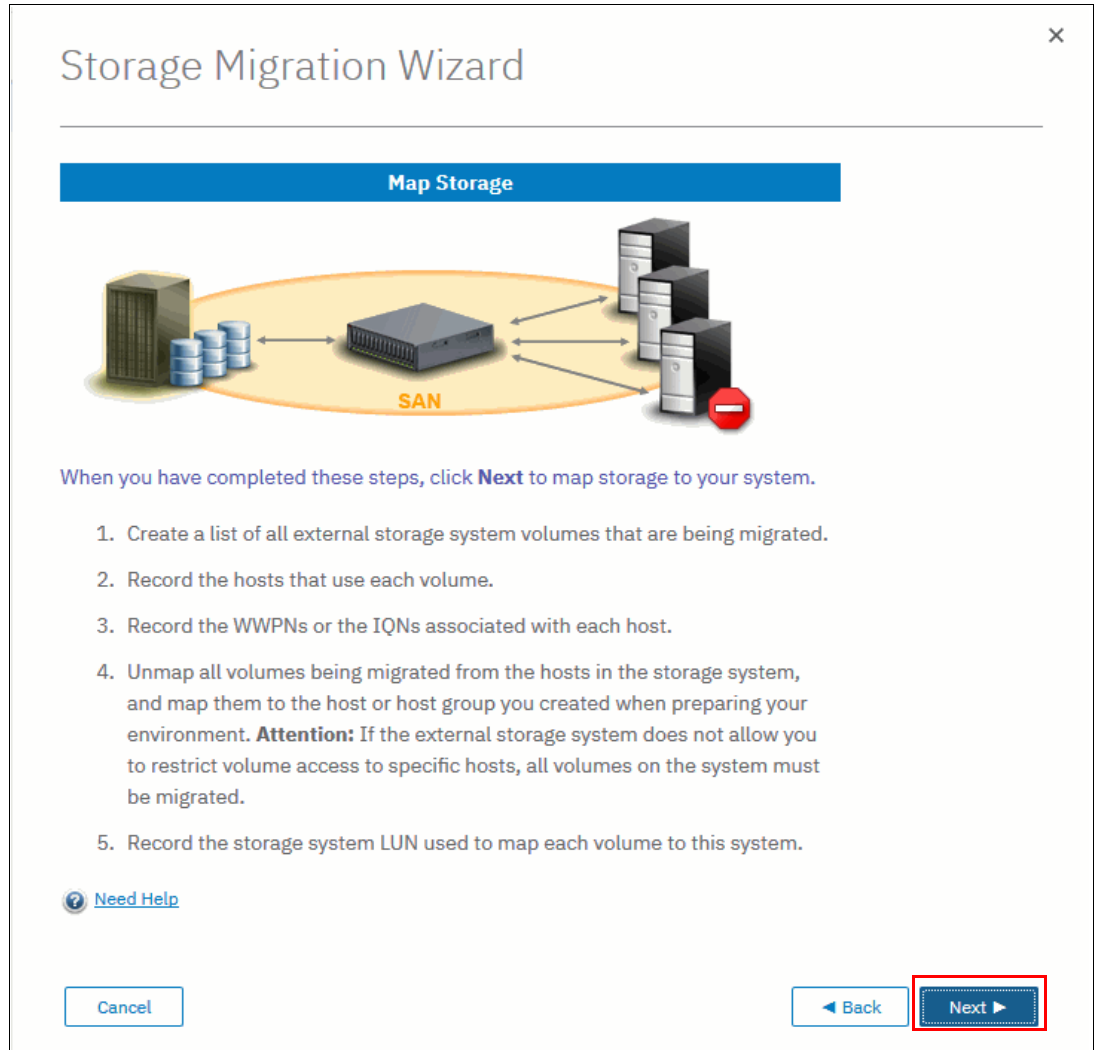


Figure 7-16 Directions to record migration data

Table 7-2 shows an example table for you to capture the information that relates to the external storage system LUs.

Table 7-2 Example table to capture the external LU information

LU name	Controller	Array	SCSI ID	Host name	Capacity
V3700external0	Node2	V3700	0		50 GiB
V3700external1	Node1	V3700	1		50 GiB
V3700external2	Node2	V3700	2		50 GiB
V3700external3	Node1	V3700	3		50 GiB
V3700external4	Node2	V3700	4		50 GiB
V3700external5	Node1	V3700	5		50 GiB

SCSI ID: Record the SCSI ID of the LUs to which the host is originally mapped. Certain operating systems do not support the change of the SCSI ID during the migration.

Table 7-3 shows an example table to capture host information.

Table 7-3 Example table to capture host information

Host name	Adapter/Slot/Port	Worldwide port name (WWPN)	Host bus adapter (HBA) firmware	HBA device driver	Operating system	V5000 multipath software
mcr-host-153	QLE2562/2/1	21000024FF2D076C	2.10	9.1.9.25	Red Hat Enterprise Linux 5 (RHEL5)	Device Mapper
mcr-host-153	QLE2562/2/2	21000024FF2D076D	2.10	9.1.9.25	RHEL5	Device Mapper

After all of the data is collected and the tasks are performed in the Map Storage section, click **Next**. The Storwize V5000 Gen2 runs the discover devices task and sequentially shows the Migrating MDisks window.

Migrating MDisks

Select the MDisks from the storage system to migrate and click **Next**. Figure 7-17 shows the Migrating MDisks window.

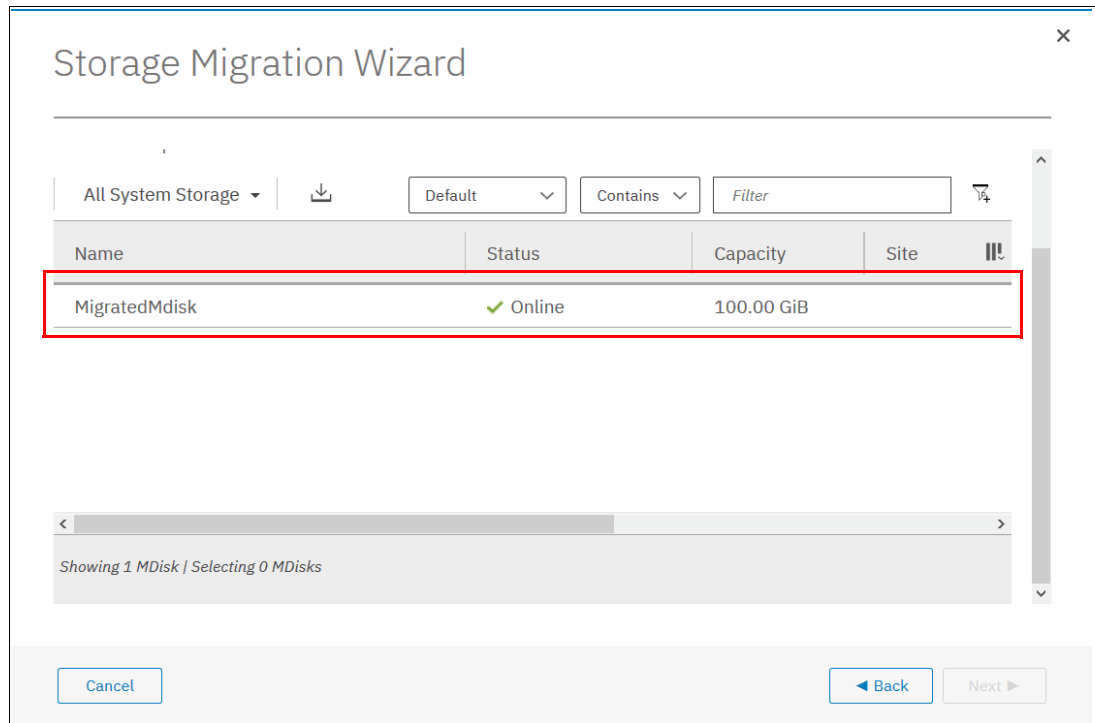


Figure 7-17 Migrating MDisks window

The Storwize V5000 Gen2 runs the Import MDisks task and sequentially shows the Configuring Hosts window.

MDisk selection: Select only the MDisks that are applicable to the current migration plan. After the current migration completes, you can start another migration to migrate any remaining MDisks.

Configuring hosts

Note: This step is optional. You can bypass it by selecting **Next** and moving to “Mapping volumes to hosts” on page 374.

Complete the following steps of the wizard to select or configure new hosts as required. Figure 7-18 shows the Configure Hosts (optional) window. If hosts are defined, they are listed in the window, as shown in Figure 3 on page 374.

1. If no hosts are defined, they can be created by selecting the **Add Host** option.

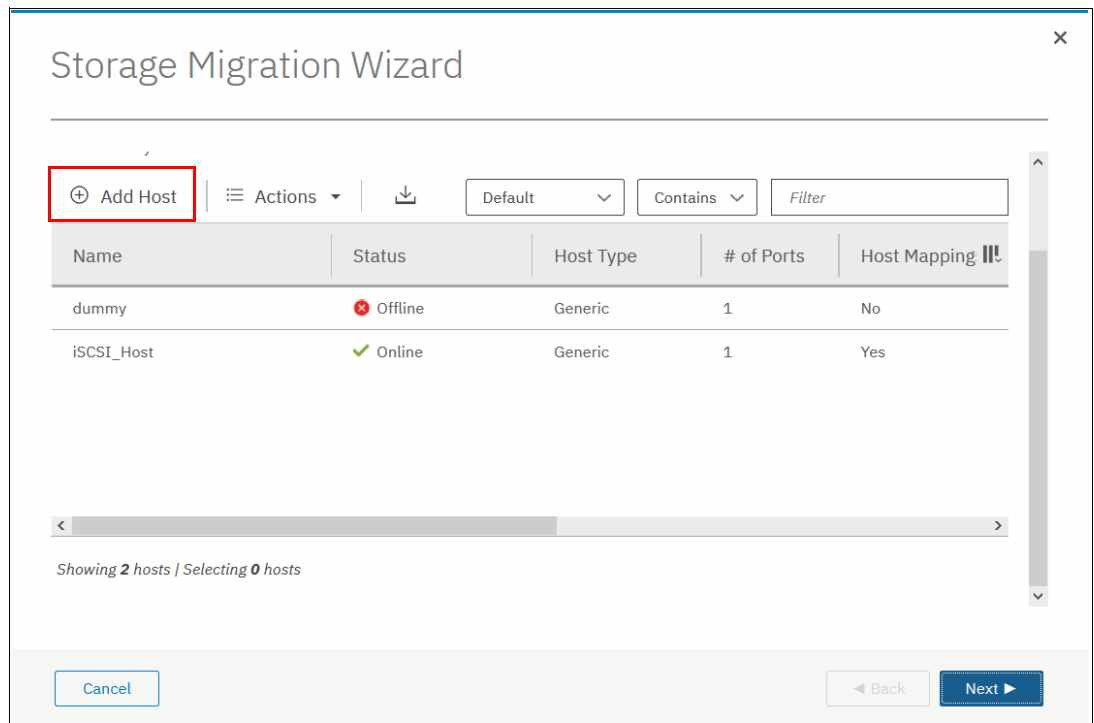


Figure 7-18 Configure Hosts window

2. Select your connection type, name the host, and assign the ports (in this case, Fibre Channel WWPNs). In the advanced settings, assign the I/O group ownership and host type, as shown in Figure 7-19. Click **Add** to complete the task. For more information about I/O group assignment, see Chapter 5, “Host configuration” on page 217.

Add Host

Required Fields

Name: ISCSI HOST

Host connections: iSCSI (SCSI)

Site: None

Host IQN: 24532

Optional Fields

CHAP authentication:

CHAP secret: Enter 1 to 79 characters

CHAP username: Enter 1 to 31 characters

Host type: Generic

I/O groups: All

Host cluster: No Host Cluster Selected

Cancel Add

Figure 7-19 The details to add a host are complete

3. The host is listed in the original Configure Hosts (optional) window. Click **Next** to display the Map Volumes to Host (optional) window.

Mapping volumes to hosts

Note: This step is optional. You can bypass it by selecting **Next** and moving to “Selecting a storage pool” on page 377.

Complete the following steps of the wizard to select volumes that were migrated from the external storage system to the Storwize V5000 Gen2 and map them to hosts:

1. Hold Ctrl and click the volume names to select multiple volumes. Click **Map to Host** to open the Create Mapping window. Figure 7-20 shows the Map Volumes to Hosts (optional) window.

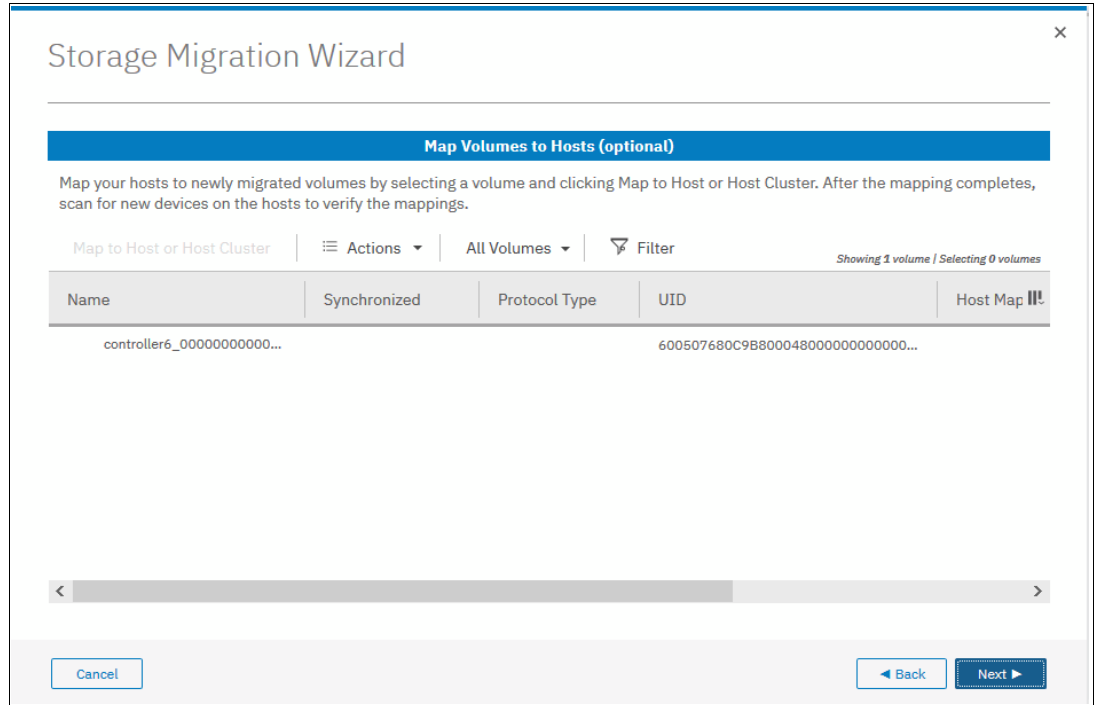


Figure 7-20 Map Volumes to Hosts window

2. The image mode volumes are listed and the names are assigned automatically by the Storwize V5000 Gen2 storage system. These names can be changed to reflect more meaningful names to the user by selecting the volume and clicking **Rename** in the Actions menu.

Names: The names of the image mode volumes must begin with a letter. The name can be a maximum of 63 characters. You can use the following valid characters:

- ▶ Uppercase letters (A - Z)
- ▶ Lowercase letters (a - z)
- ▶ Digits (0 - 9)
- ▶ Underscore (_)
- ▶ Period (.)
- ▶ Hyphen (-)
- ▶ Blank space

The names cannot begin or end with a space.

3. Select from the host list the hosts to which the imported volumes will be mapped, as shown in Figure 7-21. Click **Next**.

Create Mapping

Create Mappings to:

Hosts

Host Clusters

Select hosts to map to controller6_0000000000000002

Filter

Showing 2 hosts | Selecting 1 host

Name	Status	Host Type	Host Mappings	Protocol
dummy	Offline	Generic	No	SCSI
iSCSI_Host	Online	Generic	Yes	SCSI

Would you like the system to assign SCSI LUN IDs or manually assign these IDs?

System Assign

Self Assign

Cancel Back Next

Figure 7-21 Modify host mappings

Note: If you select Host Clusters in the Create Mapping window, you must ensure that the host cluster has consistent access to I/O groups. If each host in the host cluster does not have access to the same I/O groups, mappings to volumes fail.

4. A confirmation window opens with the task summary, as shown in Figure 7-22. Click **Map Volumes** to finish the task and return to the Map Volumes to Hosts (optional) window.

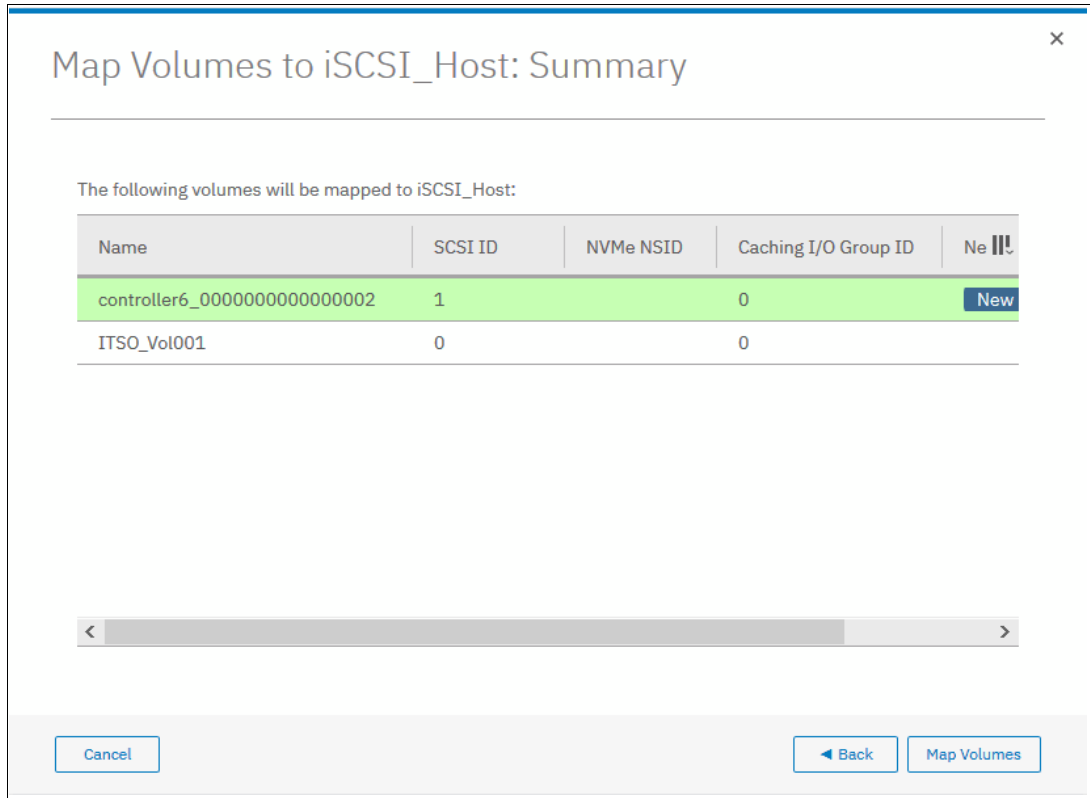


Figure 7-22 Mapping volumes summary

5. Check that in the **Mappings** Column it shows Yes. This means that the volumes were successfully mapped to the hosts. Click **Next**.

Selecting a storage pool

Note: This step is optional. You can bypass it by avoiding a pool selection, clicking **Next**, and moving to “Finishing the storage migration wizard” on page 379.

To continue with the storage migration wizard, select a storage pool to migrate the imported volumes to, as shown in Figure 7-23. Click **Next** to proceed to the last window of the storage migration wizard. The process uses the volume mirroring function that is included within the Storwize V5000 Gen2.

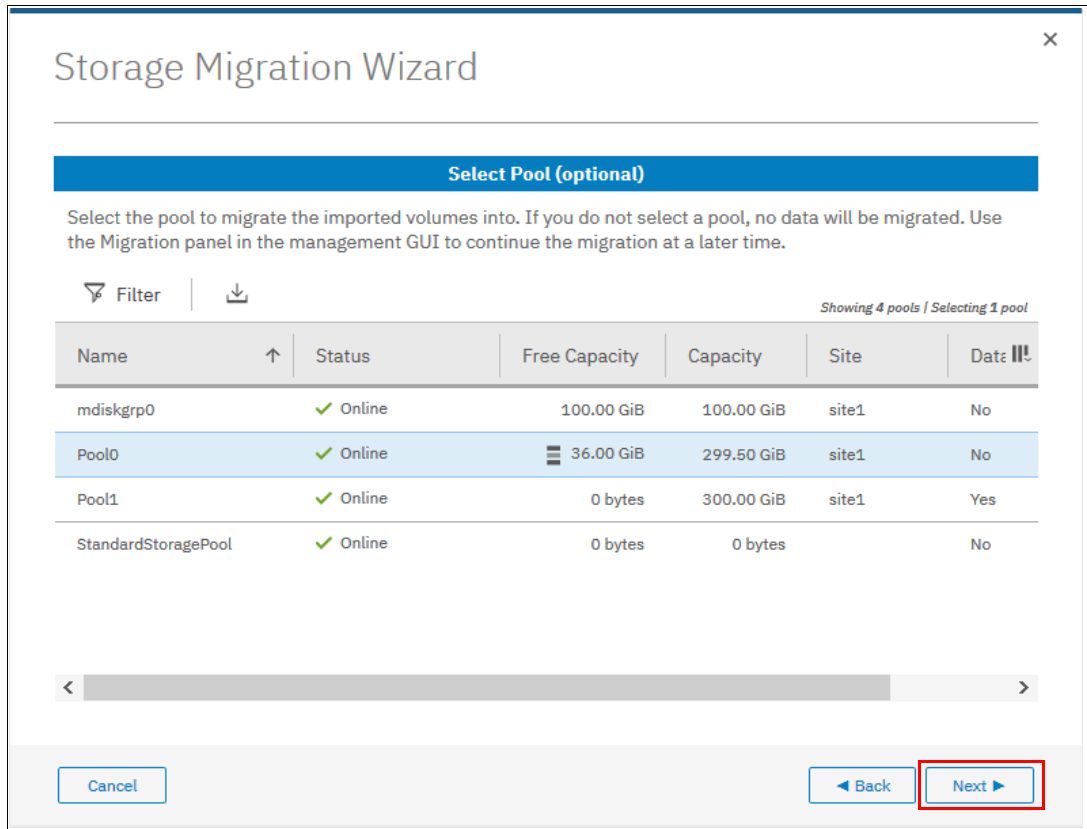


Figure 7-23 Storage pool selection

Finishing the storage migration wizard

To complete the wizard, complete the following steps:

1. Click **Finish** to end the storage migration wizard, as shown in Figure 7-24.

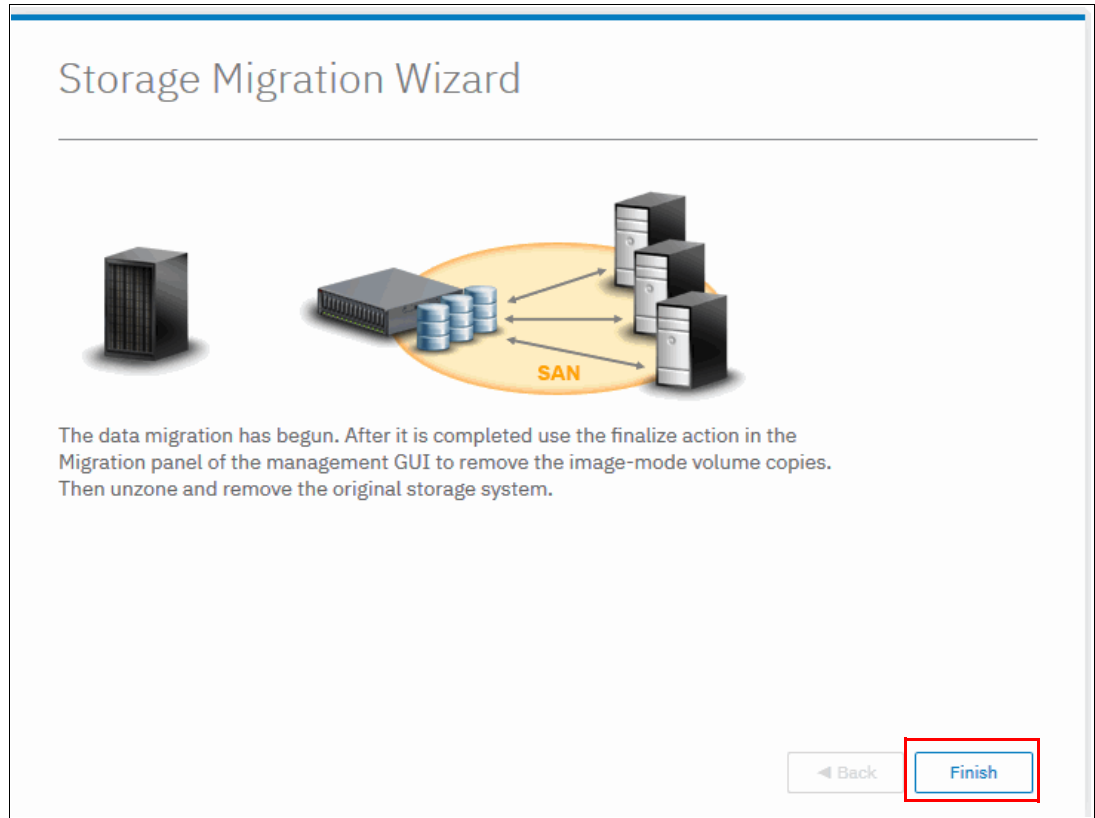


Figure 7-24 Migration wizard complete

2. The end of the storage migration wizard is not the end of the data migration task. It is still in progress. A percentage indicator is displayed in the Storage Migration window, as shown in Figure 7-25.

The screenshot shows a table with columns for Volume Name, Target Pool, Status, Progress, and UID. The Progress column contains a progress bar showing 13% completion. The table has a header row and one data row.

Volume Name	Target Pool	Status	Progress	UID
controller0_0000000000...	Test Pool	✓ Online	13%	600507680C9B80004800000000000034

Figure 7-25 Storage migration progress

3. If you want to check the progress by using the CLI, run the **lsvdisksyncprogress** command because the process is essentially a volume copy, as shown in Figure 7-26.

```
IBM Storwize:ITSOV5030: superuser>lsvdisksyncprogress
vdisk_id vdisk_name                copy_id progress estimated_completion_time
26      controller0 0000000000000000 0 1 _ 7      181016175637
IBM Storwize:ITSOV5030: superuser>
```

Figure 7-26 CLI command showing migration progress

Finalizing migrated volumes

When the migration completes with all of the progress indicators at 100%, complete the following steps:

1. Select all of the volume migrations that you want to finalize by holding down Ctrl and clicking the volumes.
2. Select **Actions** → **Finalize**, as shown in Figure 7-27. Alternatively, right-click the selected volumes and click **Finalize**.

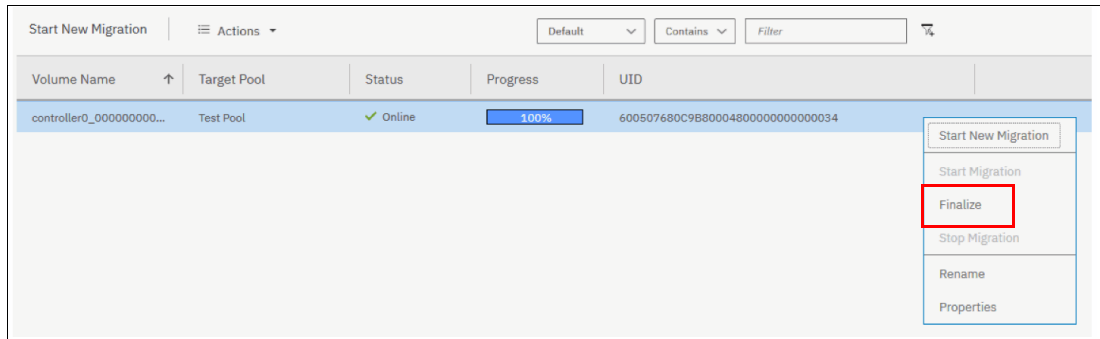


Figure 7-27 Finalize storage migration

3. You are prompted to confirm the number of volume migrations that you want to finalize, as shown in Figure 7-28. Verify that the volume names and the number of migrations are correct and click **OK**.

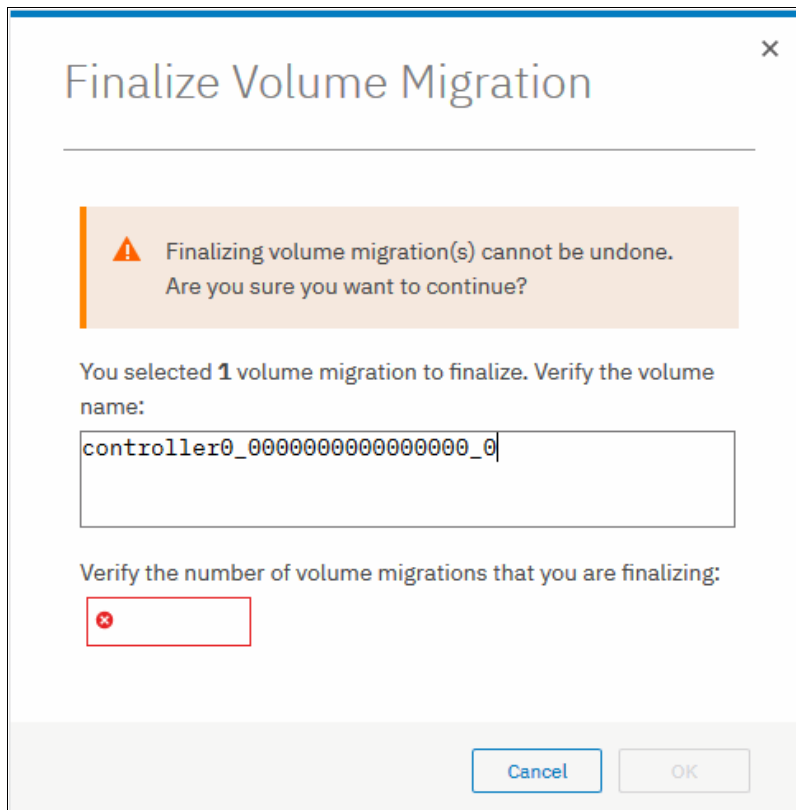


Figure 7-28 Confirm volumes to finalize

When the finalization completes, the data migration to the IBM Storwize V5000 Gen2 is completed. The zoning can be removed and the external storage system can be retired.



Advanced host and volume administration

The IBM Storwize V5000 Gen2 offers many functions for volume and host configuration. The basic host and volume features of the IBM Storwize V5000 are described in Chapter 5, “Host configuration” on page 217, and Chapter 6, “Volume configuration” on page 309. These chapters also describe how to create hosts and volumes and how to map them to a host.

This chapter focuses on advanced host and volume administration and includes the following topics:

- ▶ 8.1, “Advanced host administration” on page 384
- ▶ 8.2, “Adding and deleting host ports” on page 400
- ▶ 8.3, “Advanced volume administration” on page 407
- ▶ 8.4, “Volume properties and volume copy properties” on page 420
- ▶ 8.5, “Advanced volume copy functions” on page 422
- ▶ 8.6, “Volumes by storage pool” on page 430
- ▶ 8.7, “Volumes by host” on page 432

8.1 Advanced host administration

This section describes advanced host administration, including host modification, host mappings, and deleting hosts. Basic host creation and mapping are described in Chapter 5, “Host configuration” on page 217. We assume that hosts are defined and volumes are mapped to them.

The following topics are covered in this section:

- ▶ Modifying hosts, as described in 8.1.1, “Modifying volume mappings” on page 385
- ▶ Ports by host, as described in 8.2, “Adding and deleting host ports” on page 400
- ▶ Host mappings, as described in 8.2, “Adding and deleting host ports” on page 400

The top-level Hosts menu is shown in Figure 8-1.

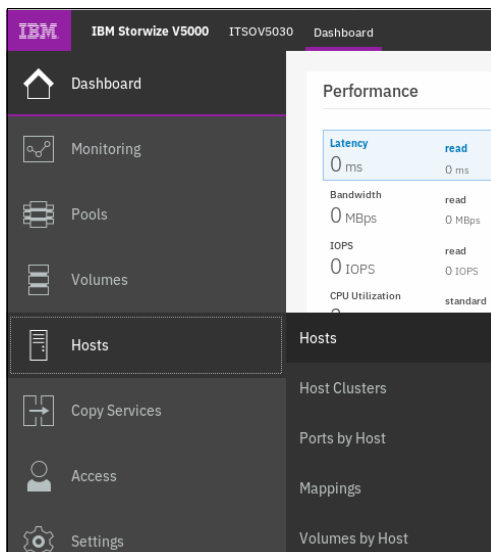


Figure 8-1 Hosts menu

Complete the following steps:

1. Select **Hosts** to open the Hosts window, as shown in Figure 8-2.

The image shows a screenshot of the 'Hosts' window in the management console. At the top, there are controls for 'Add Host', 'Actions', and a download icon. On the right, there are dropdown menus for 'Default', 'Contains', and a 'Filter' input field. Below these is a table with the following columns: Name, Status, Host Type, # of Ports, Host Mappings, Host Cluster ID, and Host Cluster Name. The table contains one row of data.

Name	Status	Host Type	# of Ports	Host Mappings	Host Cluster ID	Host Cluster Name
RHEL-Host01	Online	Generic	2	Yes		

Figure 8-2 Hosts window

2. Select a host and click **Actions** (as shown in Figure 8-3), or right-click the host to show the available actions.

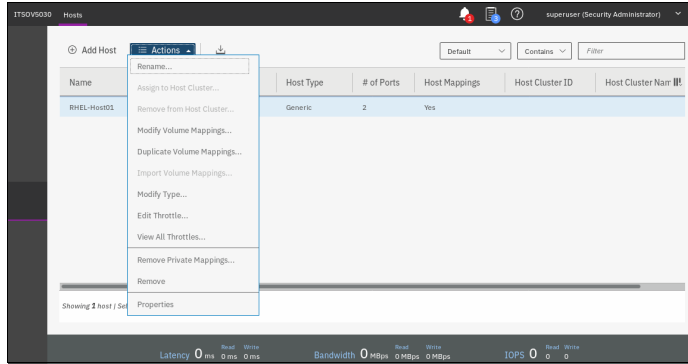


Figure 8-3 Actions menu on the Hosts window

As shown in Figure 8-3, some actions are associated with host mapping. For more information, see 8.1.1, “Modifying volume mappings” on page 385, and 8.1.2, “Unmapping volumes from a host” on page 388.

8.1.1 Modifying volume mappings

Complete the following steps:

1. From the Hosts window, select a host and click **Actions**. Then, select **Modify Volume Mappings** to open the Modify Host Mappings window, as shown in Figure 8-4.

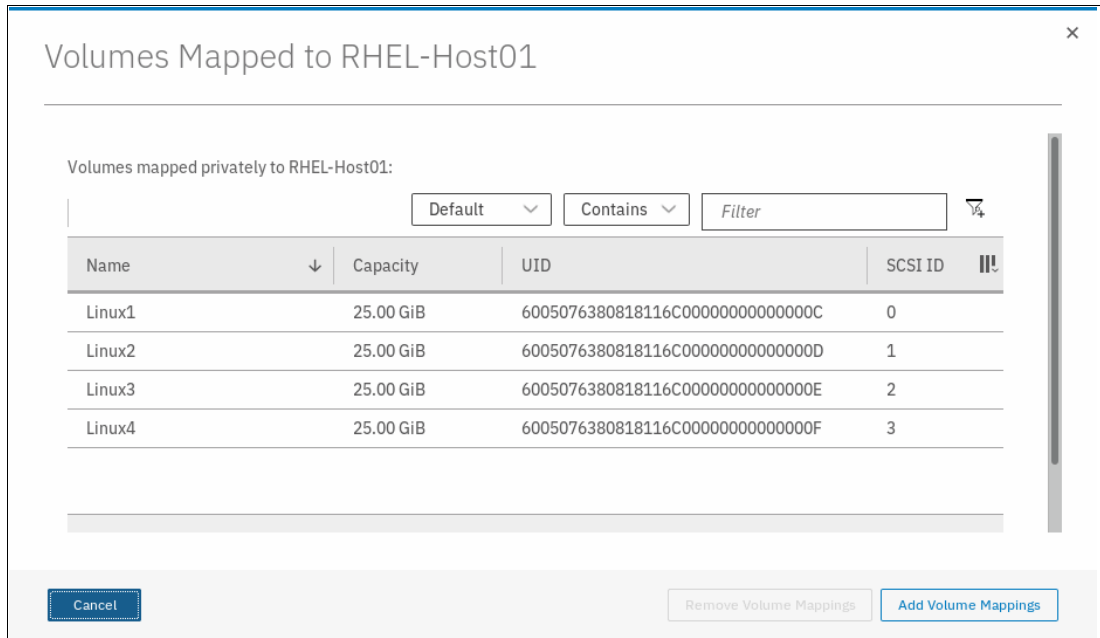


Figure 8-4 Host mappings window

2. Click **Add Volume Mappings**. A window opens in which all of the other volumes that can be mapped to the selected host are listed, as shown in Figure 8-5.

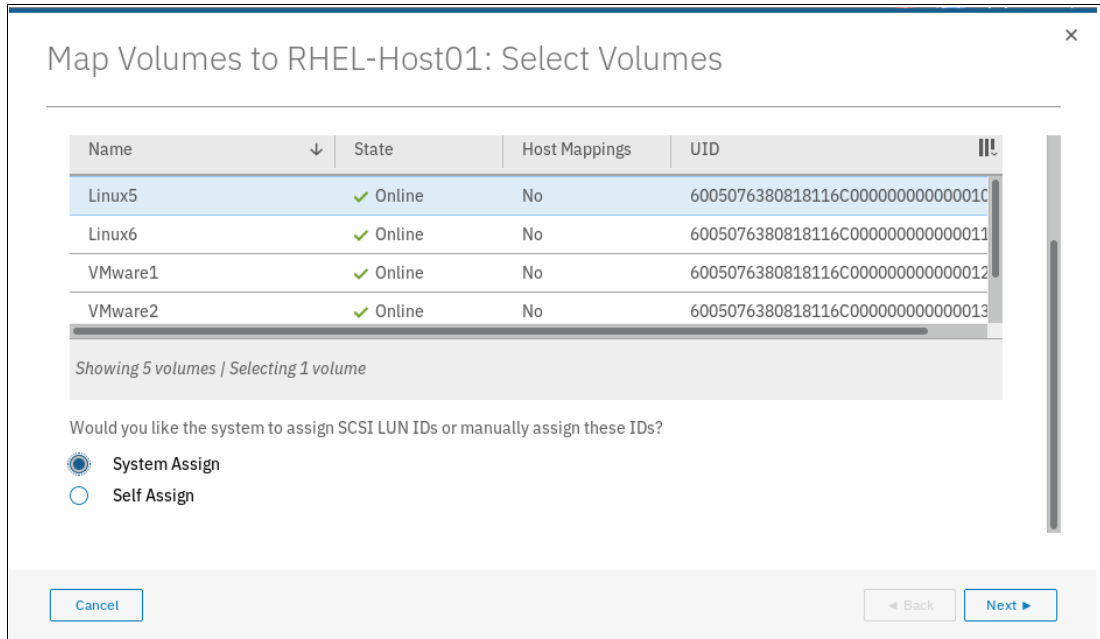


Figure 8-5 Select volumes to map to host

3. Select the volume that you need to map to the host. Also, indicate whether you want to allow the system to assign the SCSI ID or if you want to assign the SCSI ID manually. In this example, Linux2 volume is being mapped with the user supplied SCSI ID, as shown in Figure 8-6.

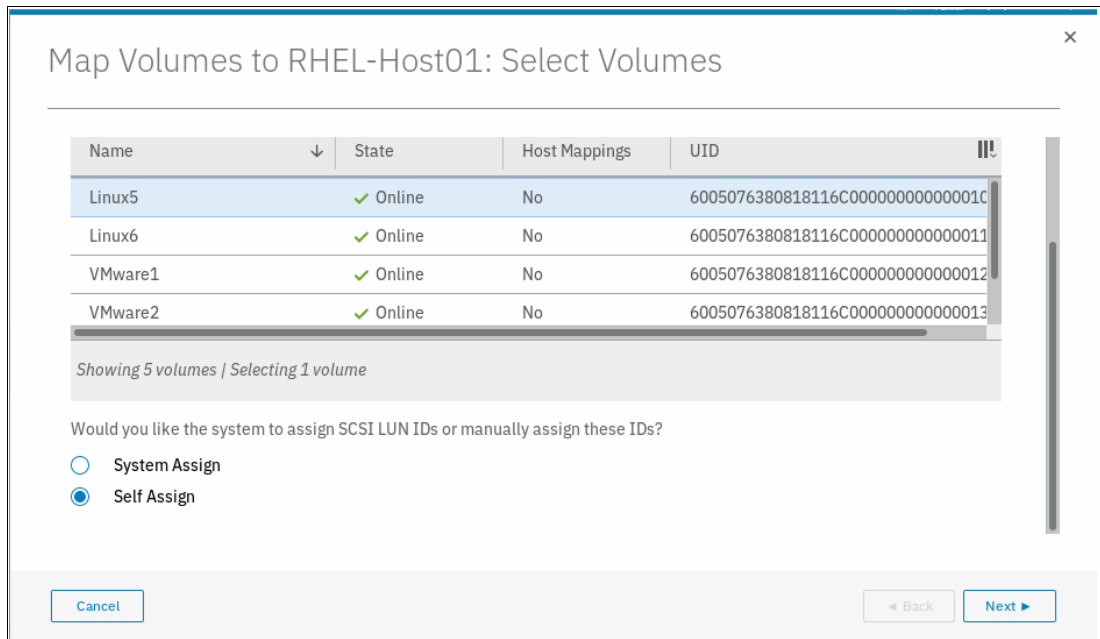


Figure 8-6 Selecting the volume and SCSI ID for mapping to a host

- Click **Next**. As shown in Figure 8-7, a window opens in which the user can provide the SCSI ID to be used for the mapping. The right side of the window also shows the current SCSI IDs that are used for mapping other volumes to the same host.

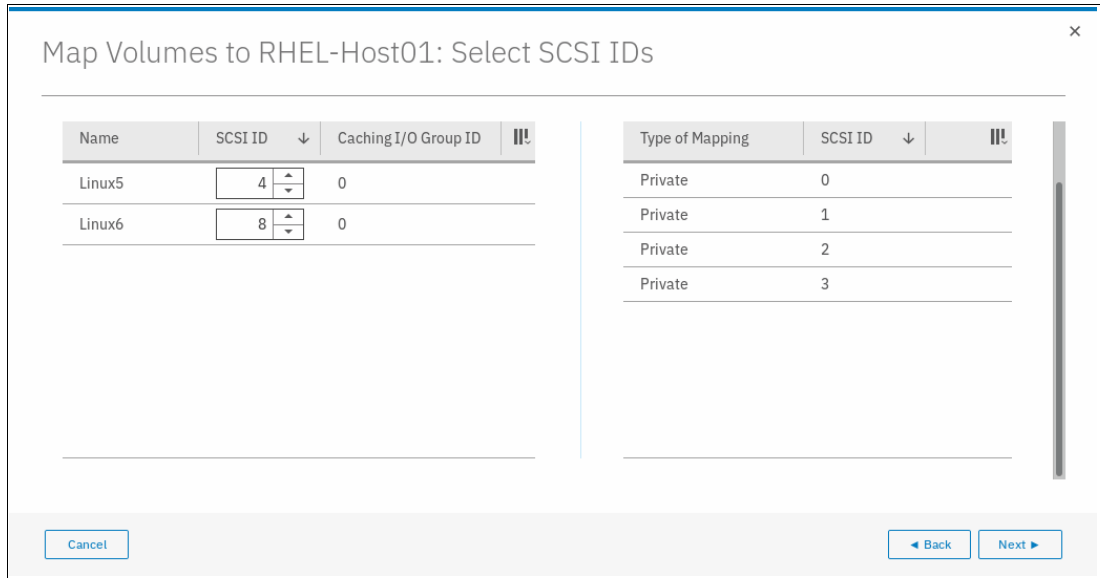


Figure 8-7 Selecting the SCSI ID for the mapping

Important: The IBM Storwize V5000 automatically assigns the lowest available SCSI ID if none is specified. However, you can set a SCSI ID for the volume. The SCSI ID cannot be changed while the volume is assigned to the host.

- Click **Next**. A window opens in which new mapping to the host is shown, along with mapped volumes to that host, as shown in Figure 8-8 on page 388. SCSI IDs for each of the mappings are also shown as part of this process.

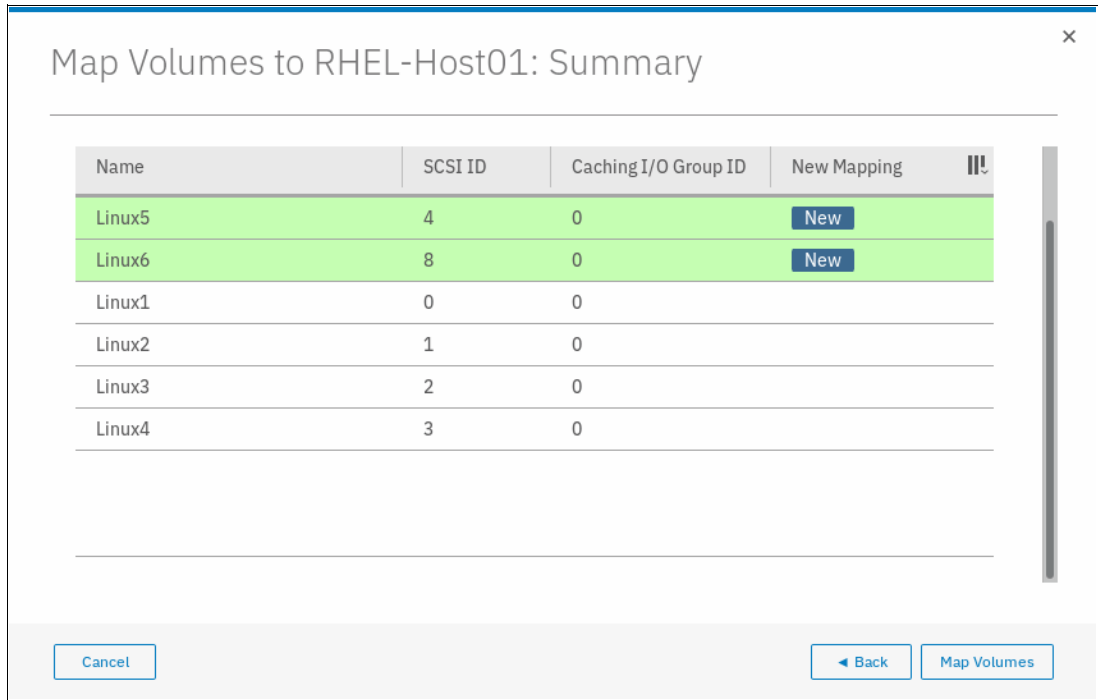


Figure 8-8 Volume map

6. Click **Map Volumes**. A confirmation window opens.

Note: If the host is part of a host cluster, the host cluster SCSI IDs also is shown, which prevents you from picking an existing SCSI ID on the host cluster to add to your private host mapping.

8.1.2 Unmapping volumes from a host

To unmap a volume from a host, complete the following steps:

1. From the main navigation window, click **Hosts**, as shown in Figure 8-9.

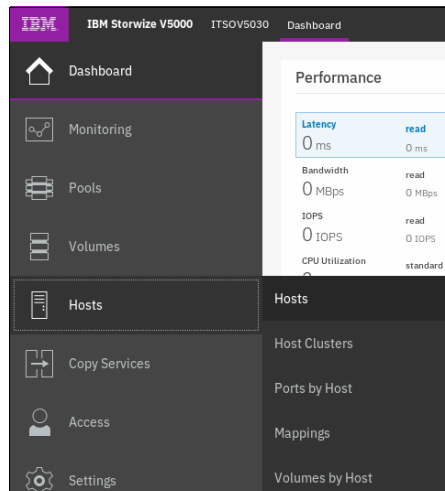


Figure 8-9 Hosts

- From the list, select the host for which you want to unmap a volume, and then, click **Action**, as shown in Figure 8-10.

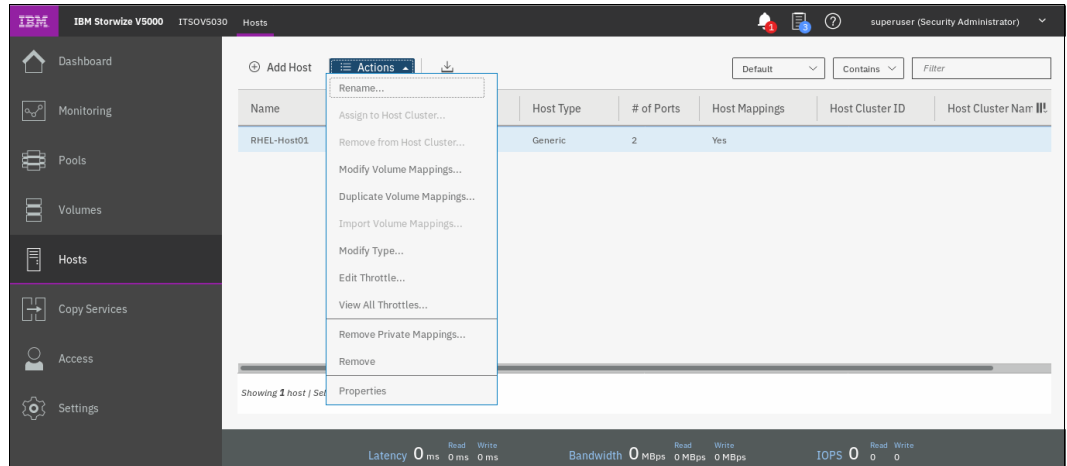


Figure 8-10 Action for the selected host

- Click **Modify Volume Mapping**. A window that includes a list of volumes that are mapped to the selected host opens, as shown in Figure 8-11.

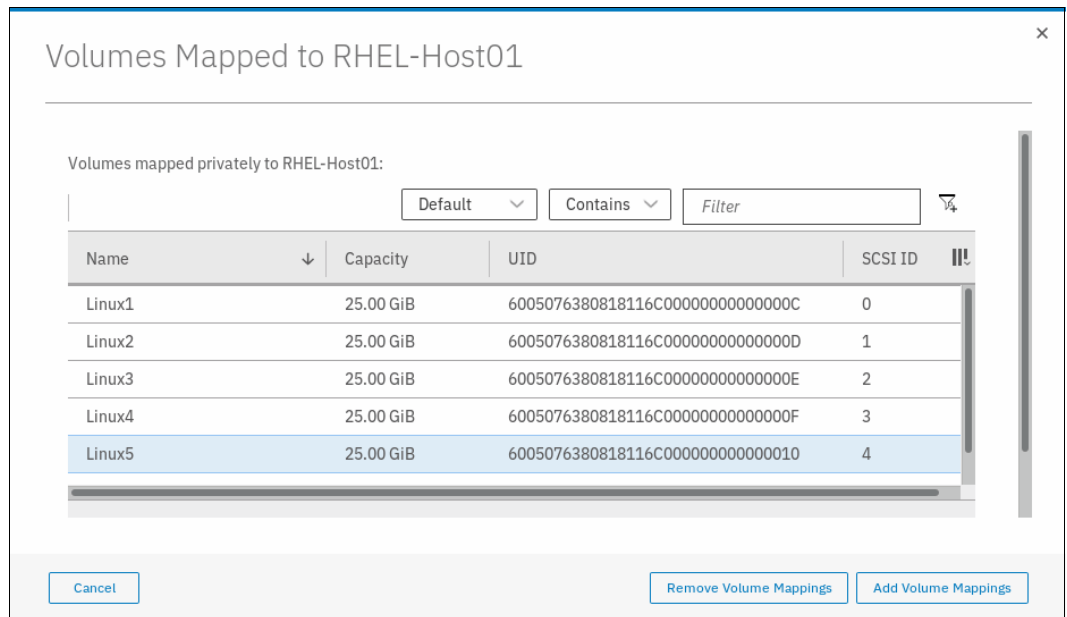


Figure 8-11 List of volumes that are mapped to the host

- Select the volume that you want to unmap and click **Remove Volume Mappings**, as shown in Figure 8-12.

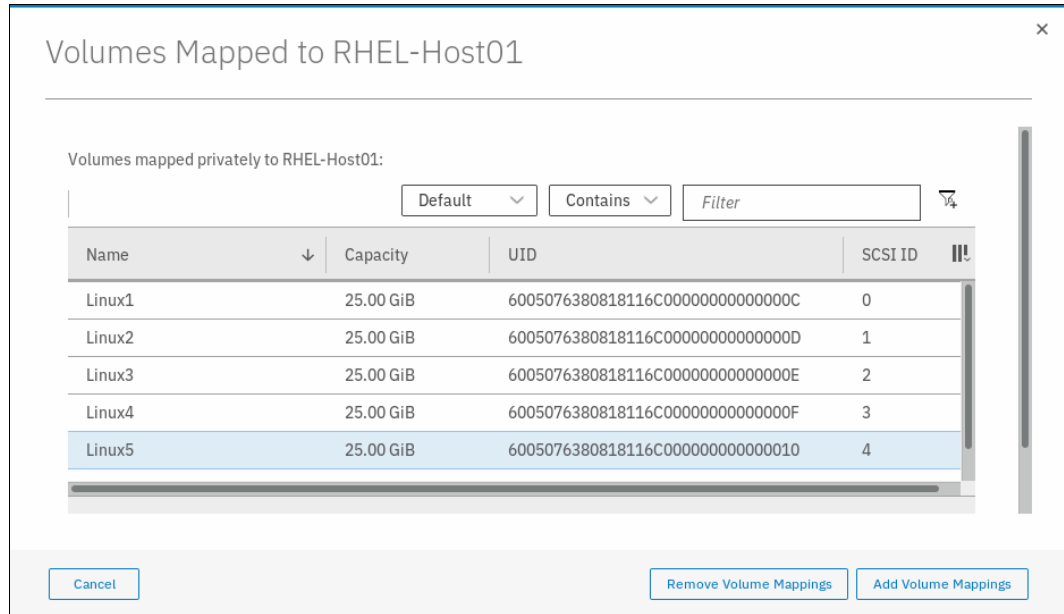


Figure 8-12 Selecting the volume to be unmapped

Note: To unmap multiple volumes, click and hold the Shift key and select each volume in the window. If the volumes you want to unmap are not consecutive, click the Ctrl key and select each volume as wanted.

- A window opens in which the volume that is going to be unmapped is listed, as shown in Figure 8-13.

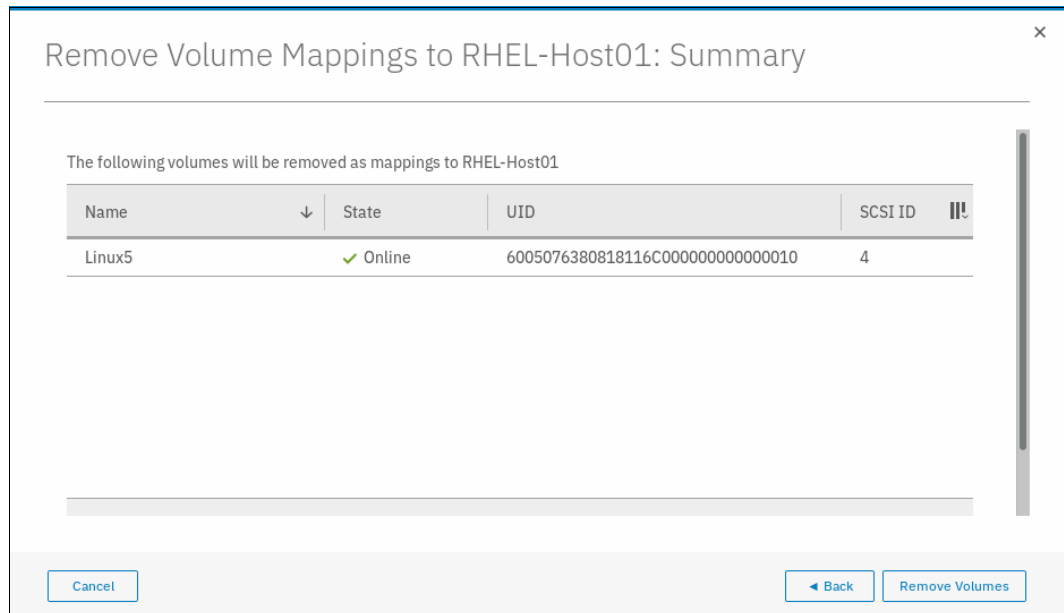


Figure 8-13 Remove volume mapping summary window

6. Click **Remove Volumes** and a task completion confirmation is displayed and your volume is unmapped.

Note: For host clusters, the same procedure can be used by selecting **Host Clusters** in **Hosts** pane, and then selecting **Modify Shared Volume Mappings**.

8.1.3 Renaming a host

To rename a host, complete the following steps:

1. From the main navigation window, select **Hosts**, as shown in Figure 8-14.

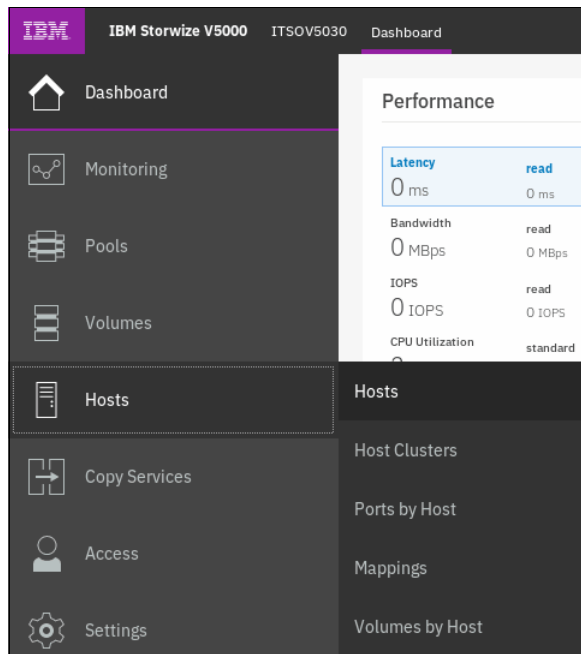


Figure 8-14 Hosts

2. Select the host that must be renamed and click **Action**. Then, click **Rename**, as shown in Figure 8-15.

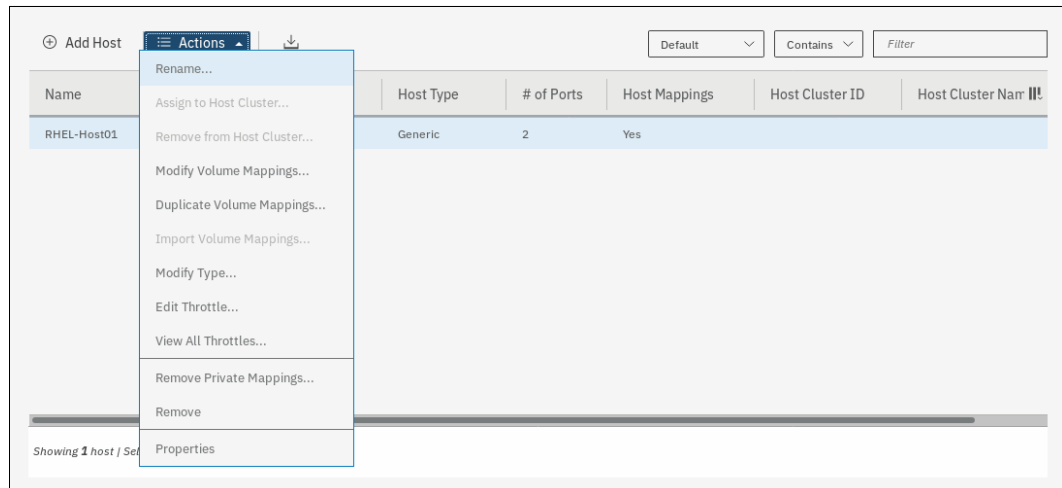


Figure 8-15 Selecting Rename action for host

3. A window opens in which you can enter the new name, as shown in Figure 8-16.

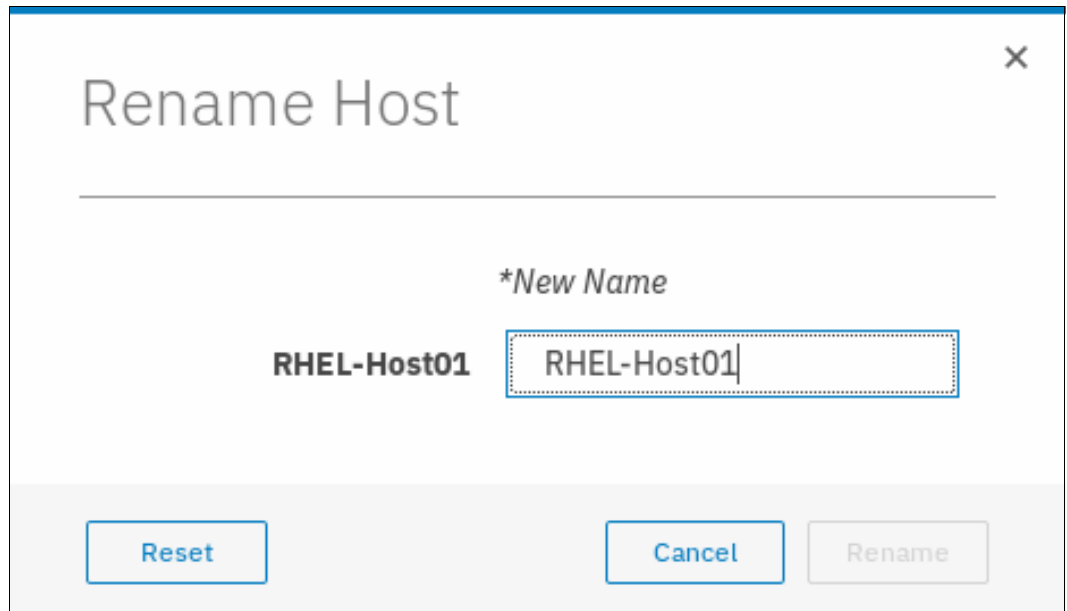


Figure 8-16 Rename a host window

4. Enter the new host name and click **Rename**, as shown in Figure 8-17.

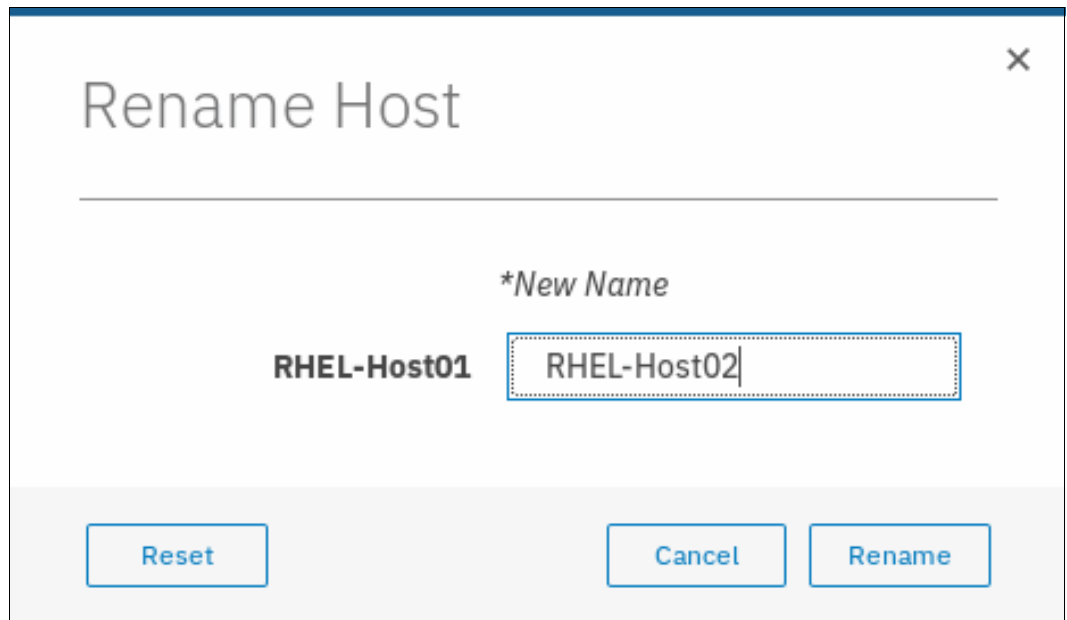


Figure 8-17 Renaming a host

5. The selected host is renamed and a task completion confirmation window is displayed. The host now features the new name.

Note: For host clusters, the same procedure can be used by selecting **Host Clusters** in **Hosts** window and selecting **Rename Host Cluster**.

8.1.4 Removing a host

To remove a host, complete the following steps:

1. From the main navigation window, select **Hosts**, as shown in Figure 8-18.

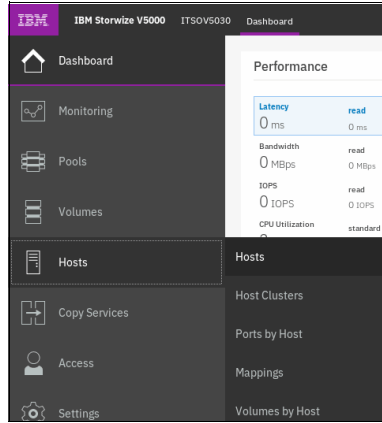


Figure 8-18 Hosts

2. Select the host that needs to be removed and click **Action**, as shown in Figure 8-19.

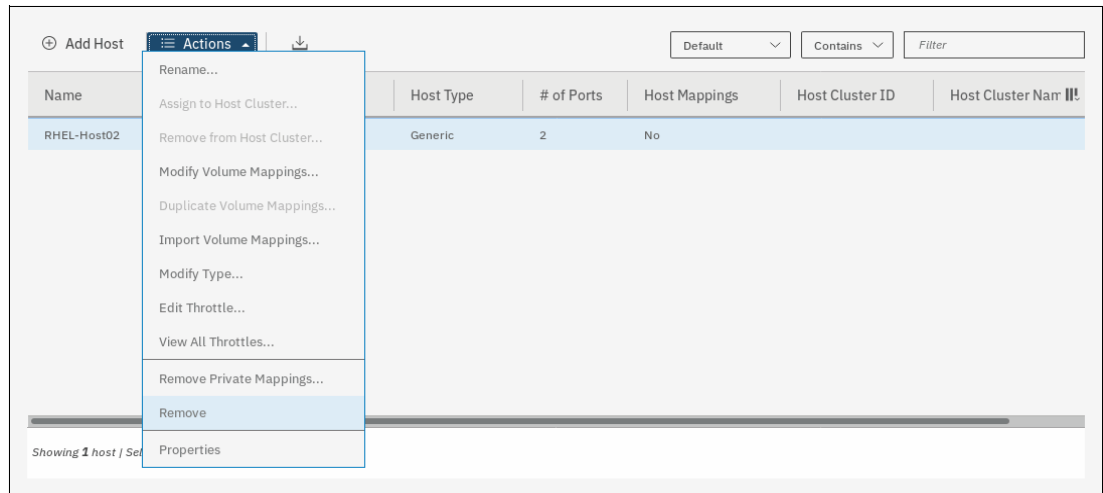


Figure 8-19 Selecting Remove operation

3. Verify the number of hosts you are removing, along with confirmation to remove the host even if the host includes mapped volumes, as shown in Figure 8-20.

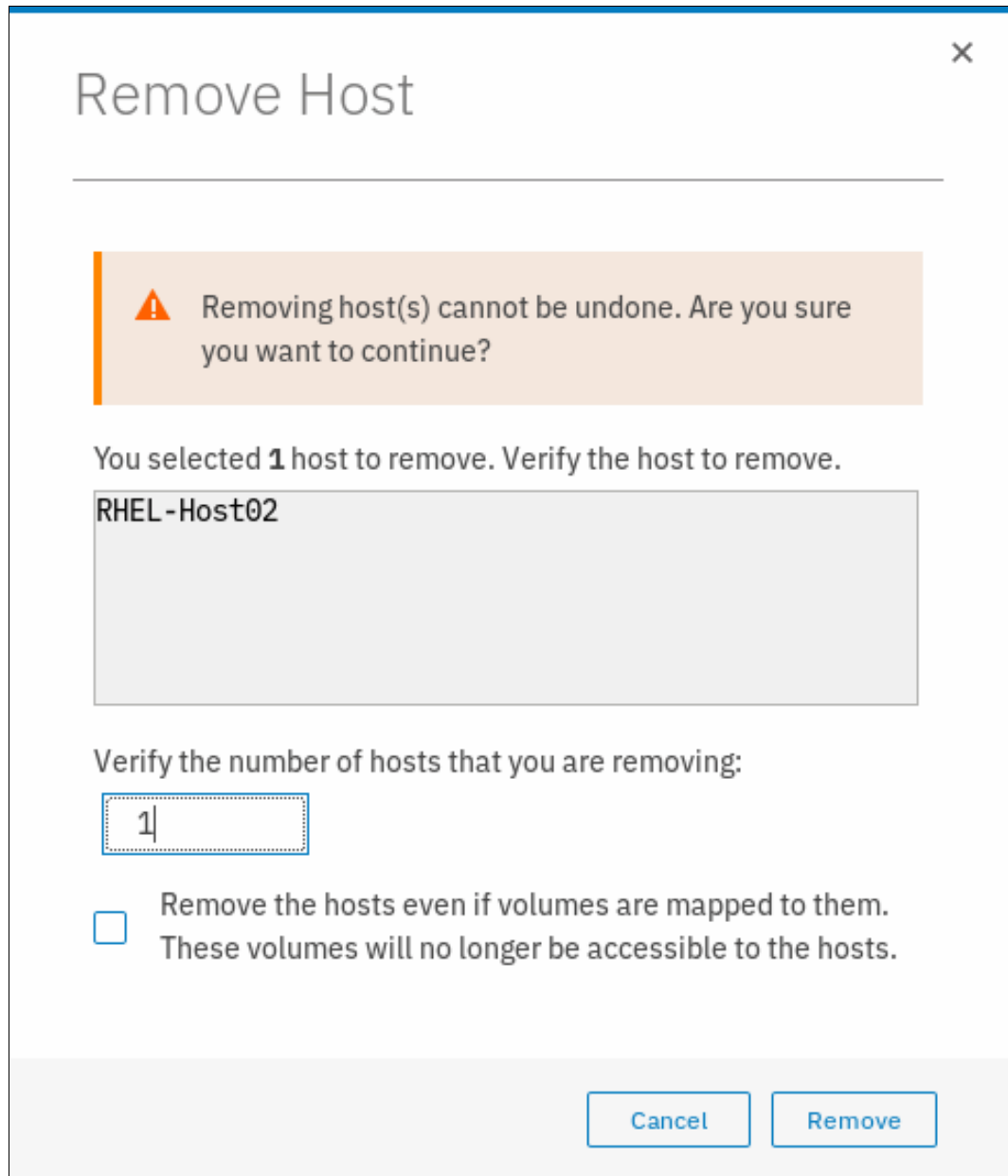


Figure 8-20 Removing a host

Note: If a host includes mapped volumes, then to remove the host, you must select the **Remove the hosts even if volumes are mapped to them. These volumes will no longer be accessible to the hosts** option to force the action.

4. A task completion window is displayed. Click **Close**.

8.1.5 Host properties

This section describes the host properties, which provide the following information:

- ▶ Overview
- ▶ Mapped volumes
- ▶ Port definitions

Overview

To open the Host Details window, complete the following steps:

1. Select the host.
2. From the Actions menu, click **Properties**. You can also highlight the host and right-click to access the **Actions** menu, as shown in Figure 8-21.

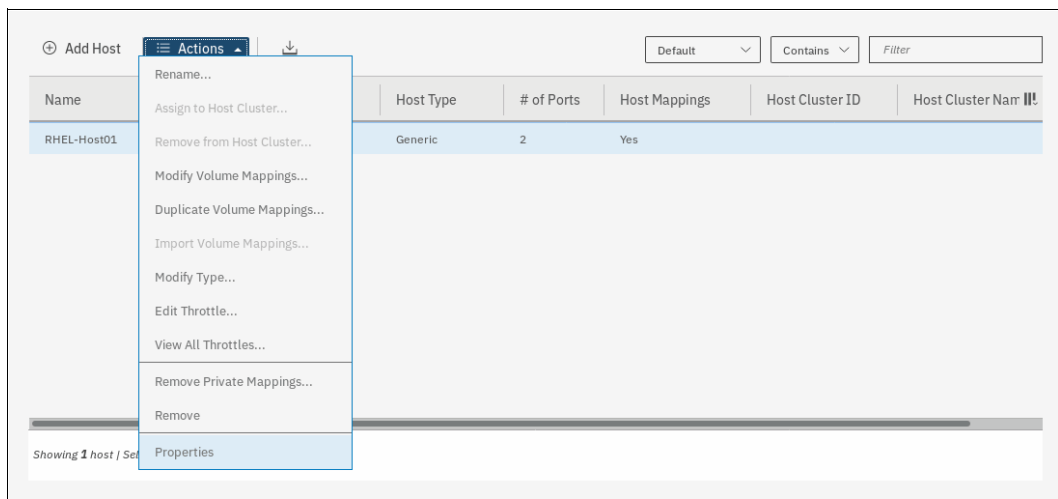


Figure 8-21 Opening host properties

3. Figure 8-22 on page 396 shows the Overview tab of the Host Details window. Select the **Show Details** option in the lower-left corner of the window to see more information about the host.

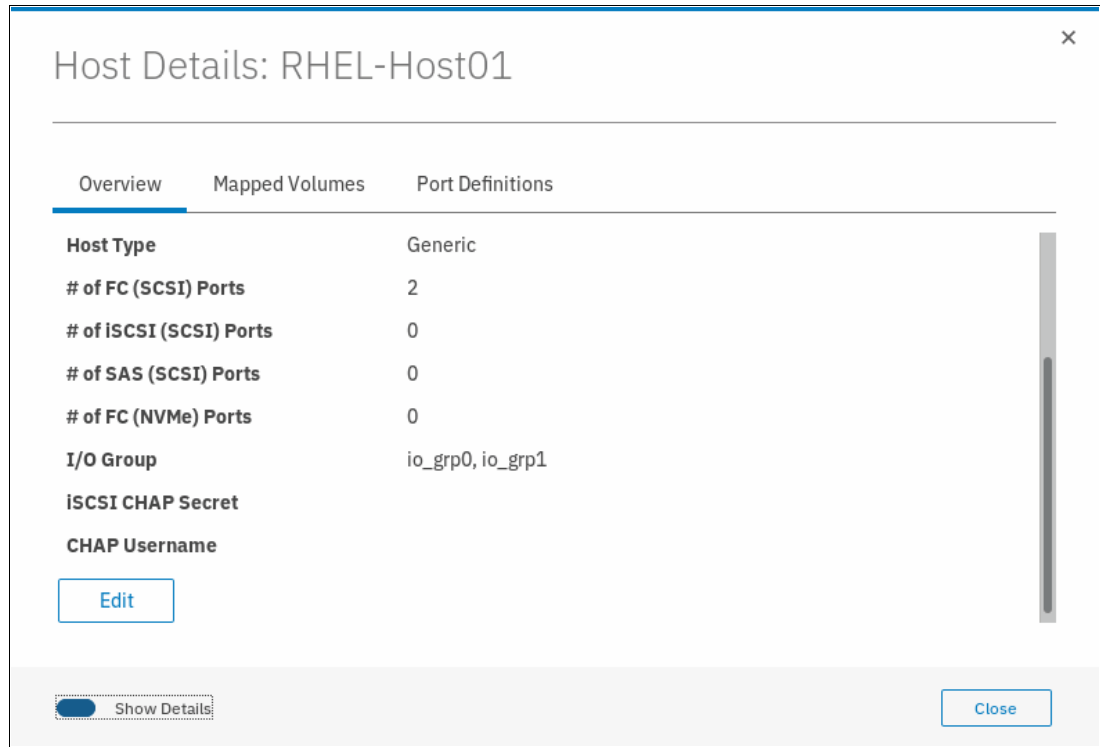


Figure 8-22 Host details

The Overview tab provides the following information:

- Host name: Host object name.
- Host ID: Host object identification number.
- Status: The current host object status. This value can be Online, Offline, or Degraded.
- Host type: The type of host can be Generic, Generic (hidden secondary volumes), HP/UX, OpenVMS, Target Port Group Support (TPGS), and VMware Virtual Volume (VVOL).
- Number of Fibre Channel (FC) ports: The number of host Fibre Channel ports.
- Number of internet SCSI (iSCSI) ports: The number of host iSCSI names or host iSCSI qualified names (IQN) IDs.
- Number of serial-attached SCSI (SAS) ports: The number of host SAS ports.
- Number of Non-Volatile Memory express (NVMe) ports: The number of host NVMe Qualified names (NQM) IDs.

Note: Non-Volatile Memory Express (NVMe) are not supported on Storwize V5000 systems.

- I/O group: The I/O group from which the host can access a volume (or volumes).
- iSCSI Challenge Handshake Authentication Protocol (CHAP) secret: The CHAP information if it exists or if it is configured.
- CHAP Username: The CHAP user name the host uses to authenticate to the system.

4. To change the host properties, click **Edit**. Several fields can be edited, as shown in Figure 8-23.

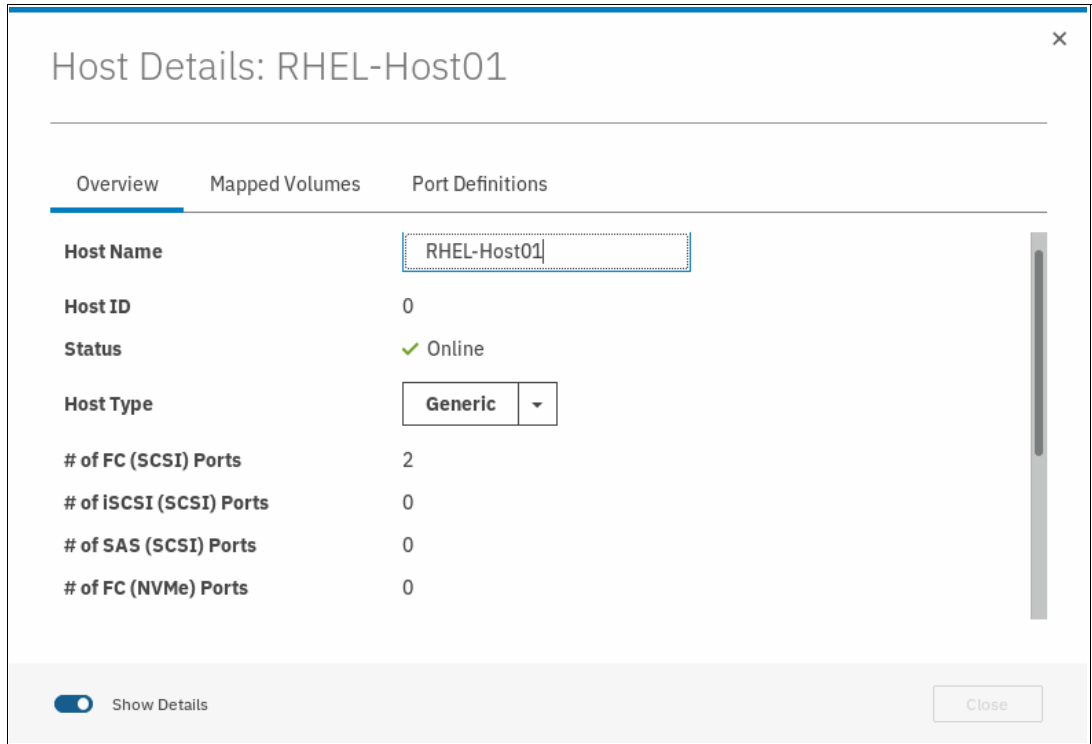


Figure 8-23 Host properties: Editing the host information

5. For the host type, choose one of the following values:
- Generic
 - Generic (hidden secondary volumes)
 - HP/UX, OpenVMS
 - TPGS
 - VVOL
6. After you change any host information, click **Save** to apply your changes.

Mapped Volumes

Figure 8-24 shows the Mapped Volumes tab, which provides an overview of the volumes that are mapped to the host. This tab provides the following information:

- ▶ SCSI ID
- ▶ Volume Name
- ▶ Unique identifier (UID)
- ▶ Caching I/O group ID

SCSI ID	Name	UID	Caching I/...
0	Linux1	6005076380818116C0000000000000...	0
1	Linux3	6005076380818116C0000000000000...	0
2	Linux4	6005076380818116C0000000000000...	0

Figure 8-24 Host Details: Mapped Volumes tab

Port Definitions

Figure 8-25 on page 399 shows the Port Definitions tab, which shows the configured host ports and their status. This tab provides the following information:

- ▶ Name: The worldwide port names (WWPNs) (for SAS and FC hosts) or iSCSI Qualified Name (IQN) for iSCSI hosts
- ▶ Type: Port type
- ▶ Status: Current port status
- ▶ Number of nodes that are logged in: Lists the number of IBM Storwize V5000 node canisters to which each port (initiator port) is logged in

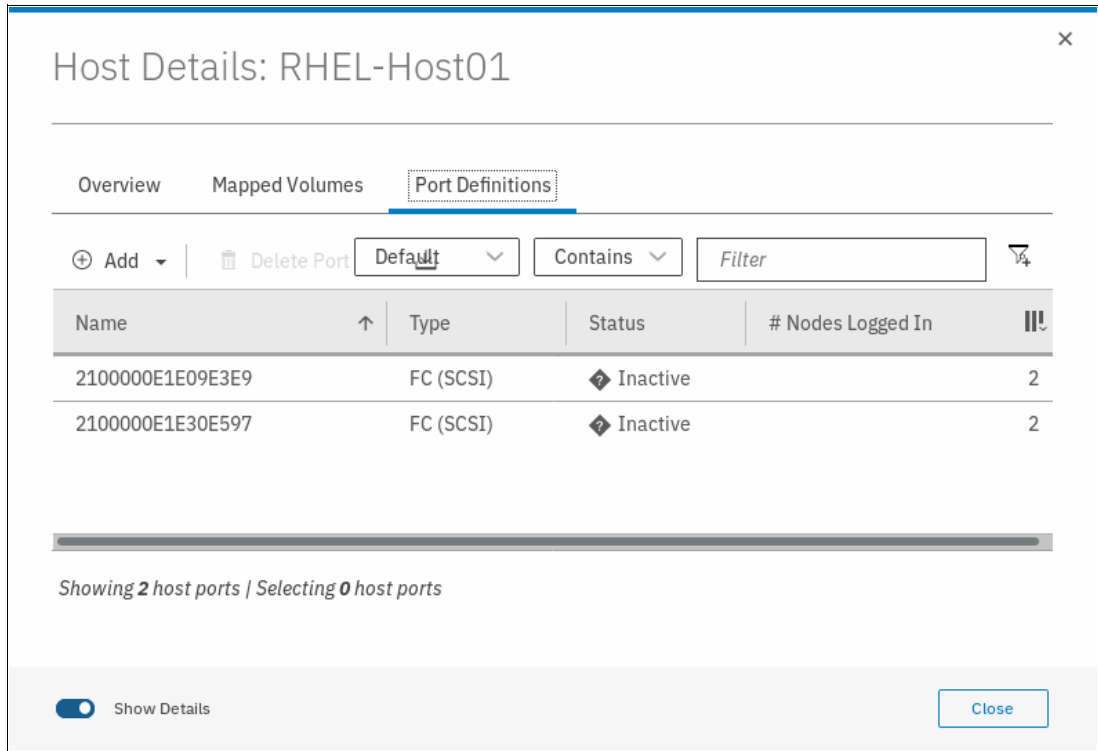


Figure 8-25 Host port definitions

Click **Close** to close the Host Details window.

8.2 Adding and deleting host ports

In this section, we describe how to add and delete ports to and from a host definition. The examples that are shown here are for Fibre Channel ports, but the steps are applicable for SAS and iSCSI ports as well.

8.2.1 Adding host port

To add or delete host ports, go to the **Port Definitions** tab (as described in “Port Definitions” on page 398) and complete the following the steps:

1. Click **Add**, as shown in Figure 8-26.

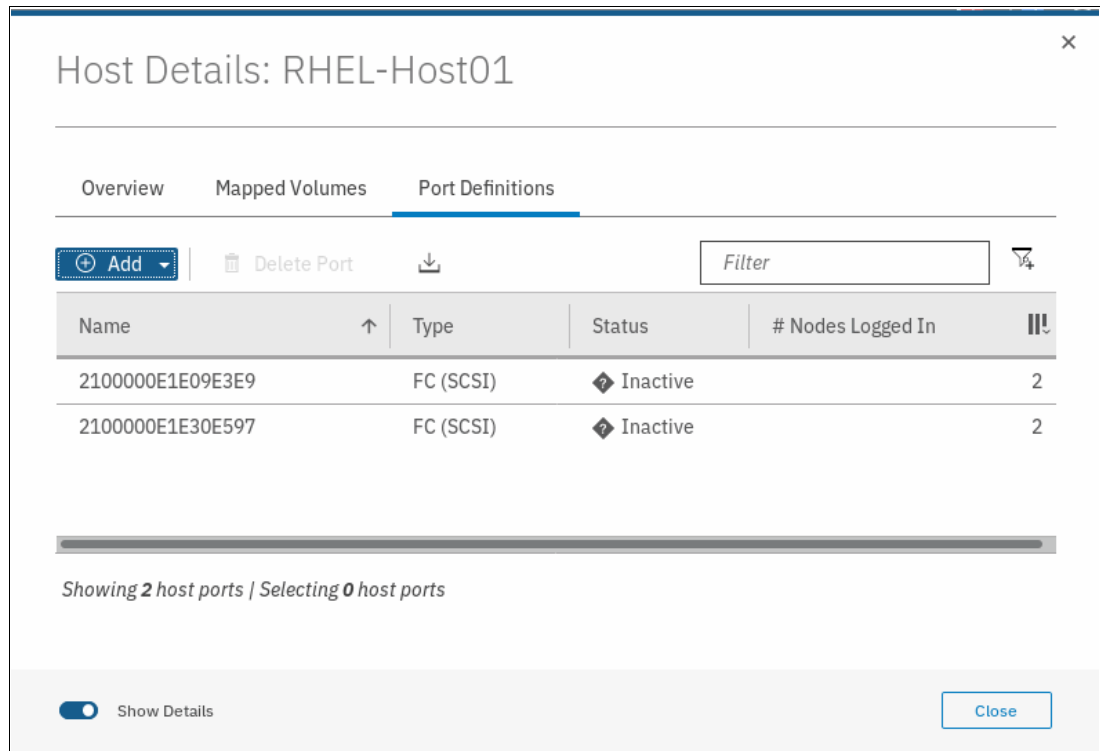


Figure 8-26 Add host port

2. Select the type of port that you want to add. In this example, we chose Fibre Channel port. Then, you are directed to a drop-down list that includes the available connected WWPN. Select the wanted WWPN to be added to the host, as shown in Figure 8-27.

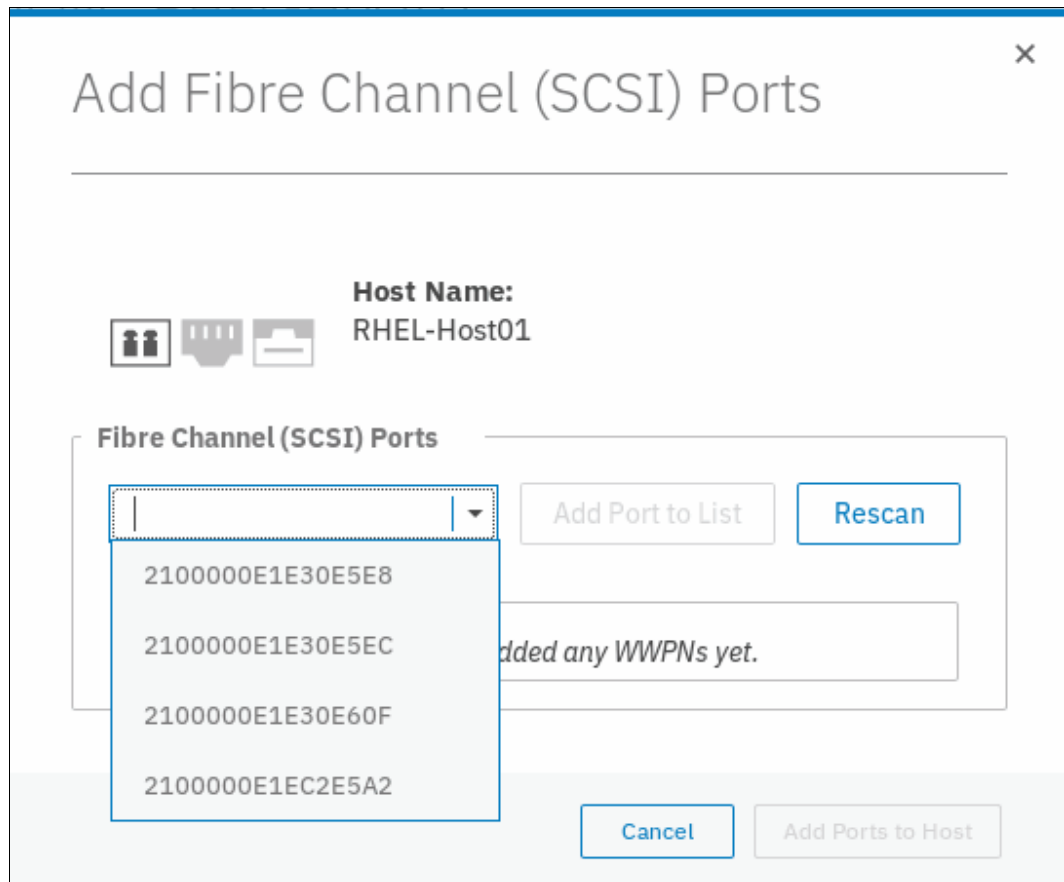


Figure 8-27 Select WWPNs to be added

Note: If the WWPN does not show in the drop-down list, click **Rescan** and try again. If the port still does not appear, review the zoning.

3. Select the wanted WWPN and click **Add Port to List**, as shown in Figure 8-28.

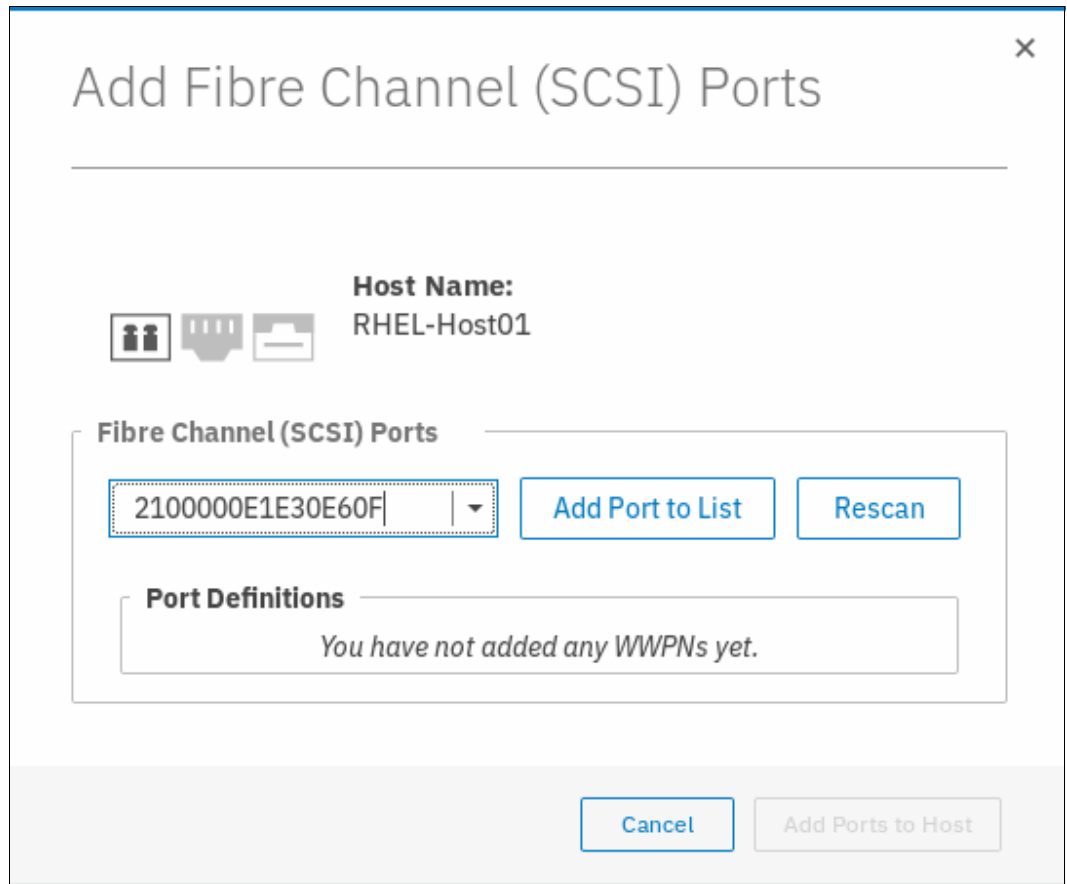


Figure 8-28 Add Port to List

4. The selected port is shown under **Port Definitions**, as shown in Figure 8-29.

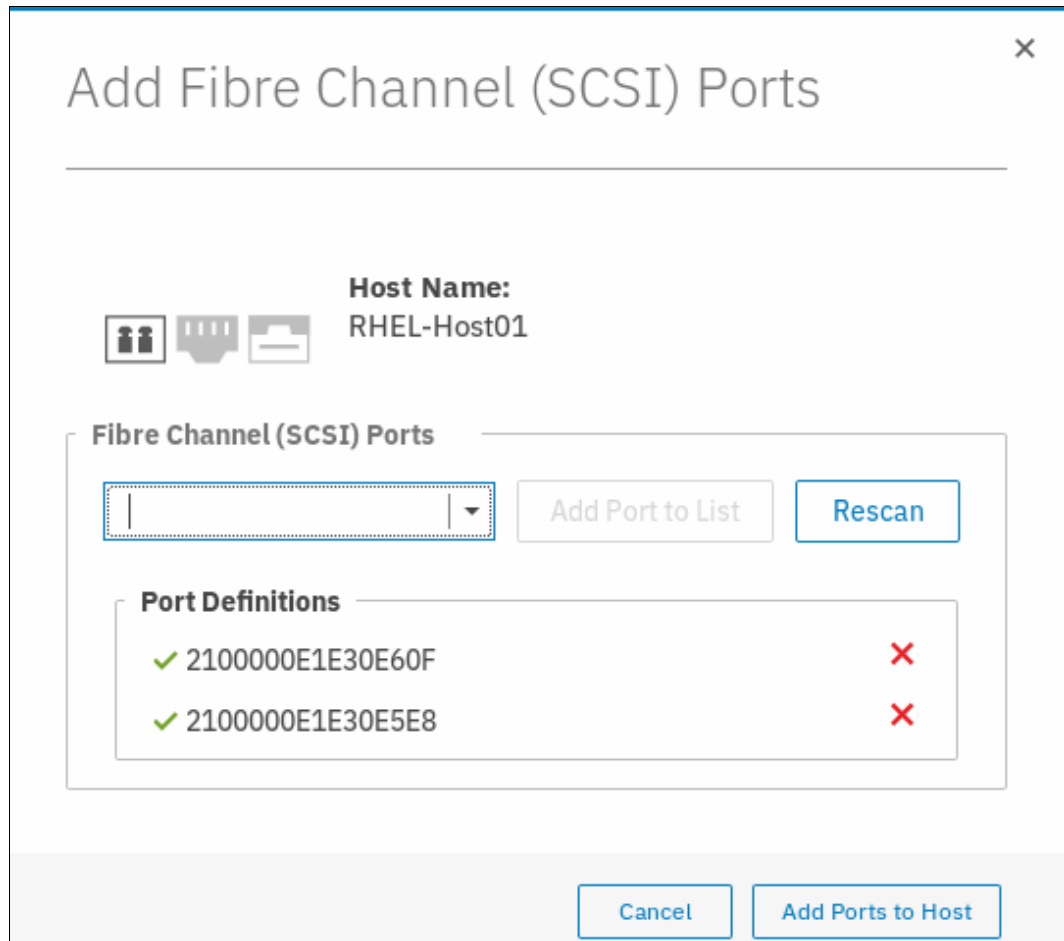


Figure 8-29 Port Definitions

Note: If the selected port is not the port you wanted, click the red X to the right to delete it from the selection.

5. Click **Add Ports to Host**. A task completion window is displayed.

The Host Details window now shows the ports that are defined for the host, including the recently added one, as shown in Figure 8-30.

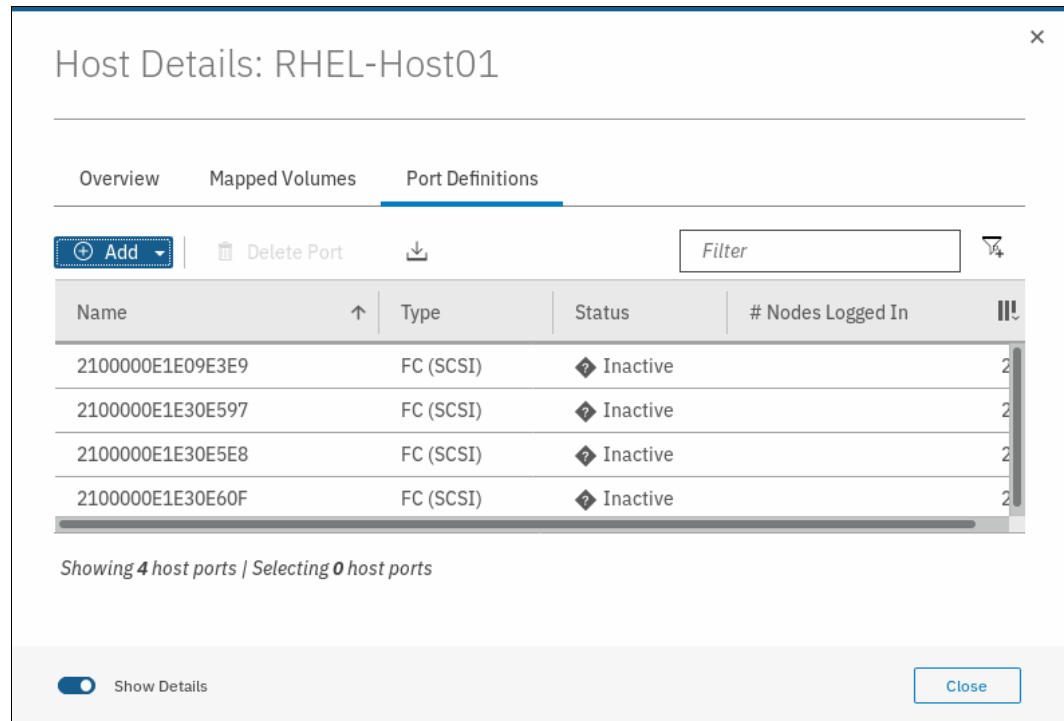


Figure 8-30 Port Definitions after adding a port

8.2.2 Deleting a host port

To add or delete host ports, go to the **Port Definitions** tab (as described in “Port Definitions” on page 398) and complete the following steps:

1. Select the port that you want to delete, as shown in Figure 8-31 on page 405.

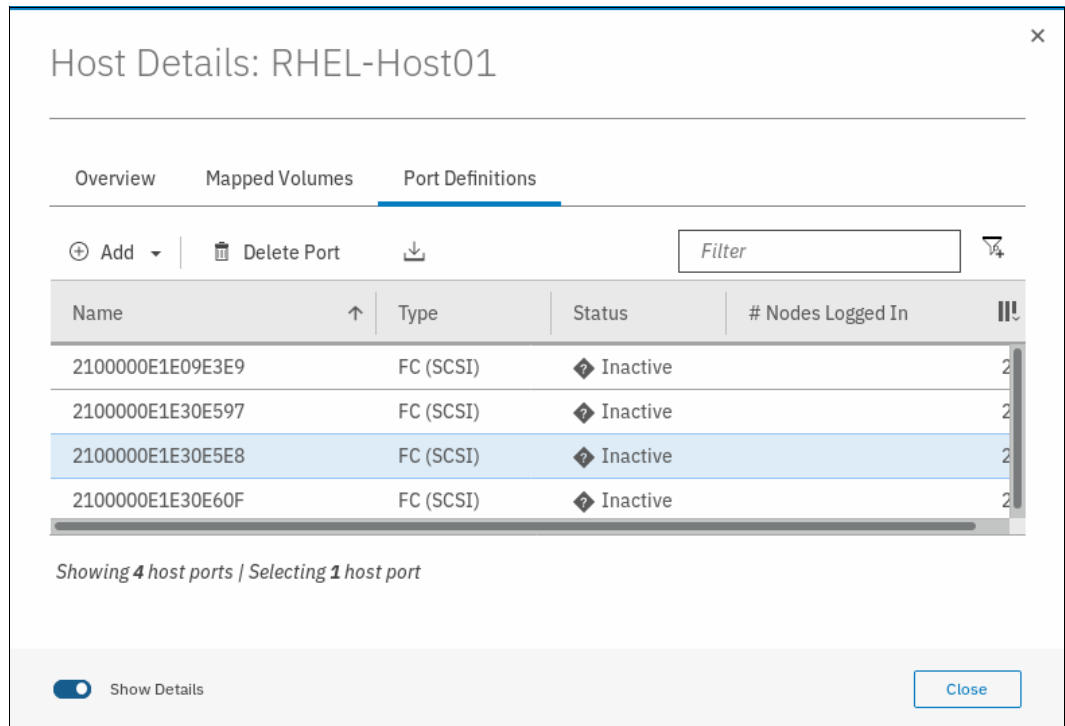


Figure 8-31 Select port to delete

2. Select **Delete Port**, as shown in Figure 8-32.

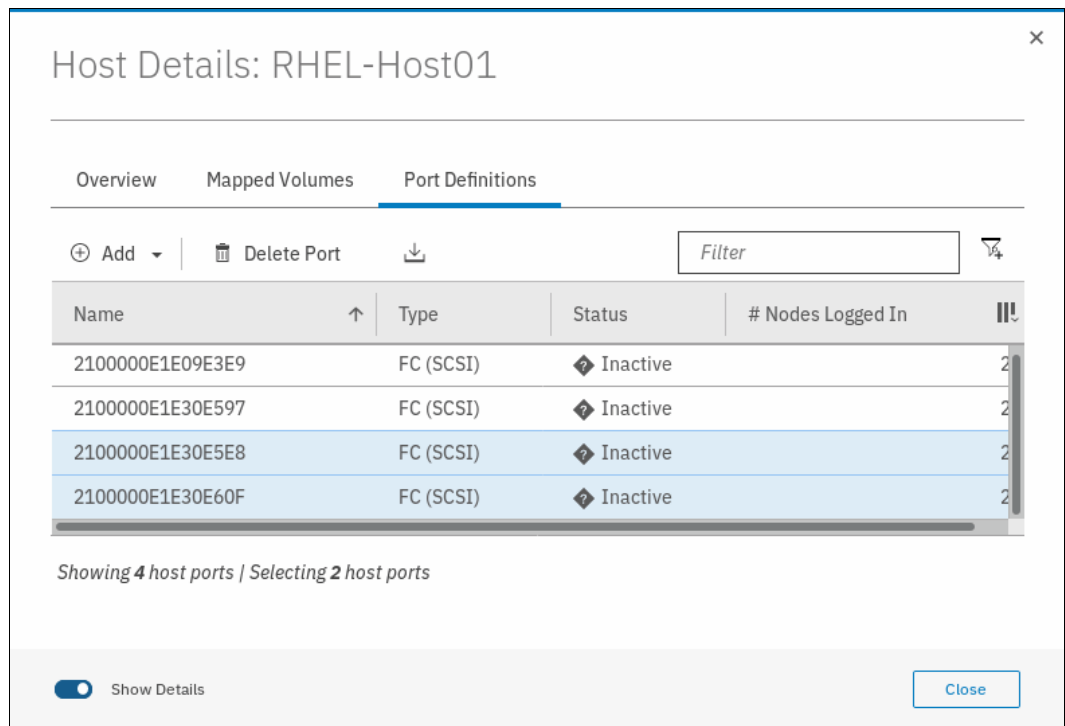


Figure 8-32 Delete Port operation

3. Verify the number of ports and the port to be deleted, as shown in Figure 8-33.

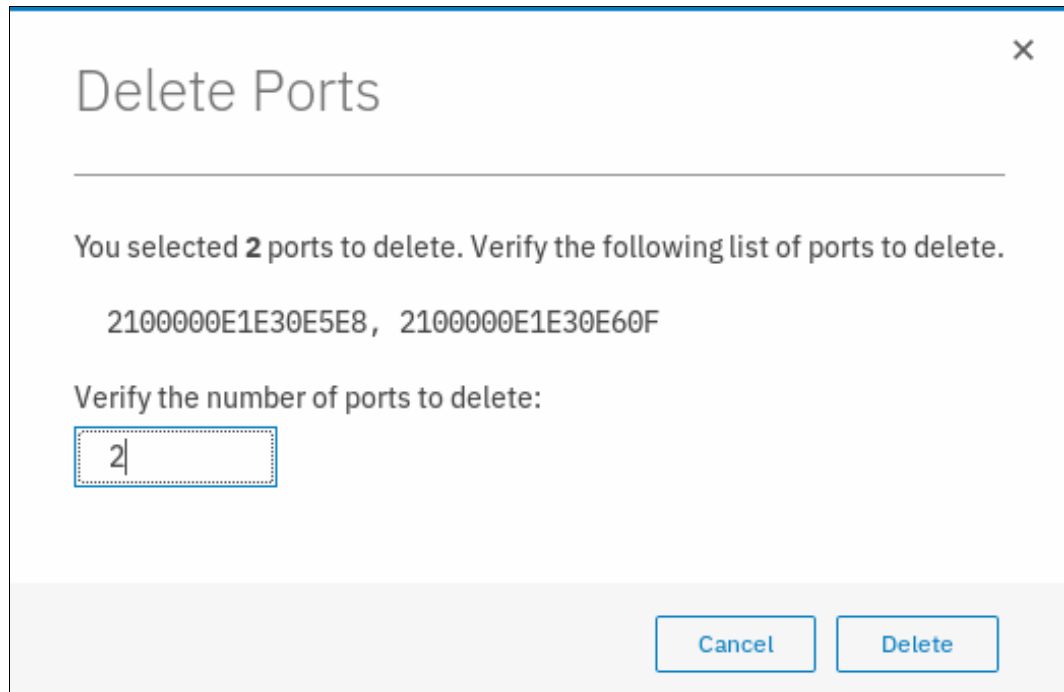


Figure 8-33 Delete port

4. Click **Delete**. A window opens that indicates that the port deletion Task is completed. Then, a window opens in which the current host ports are listed, as shown in Figure 8-34.

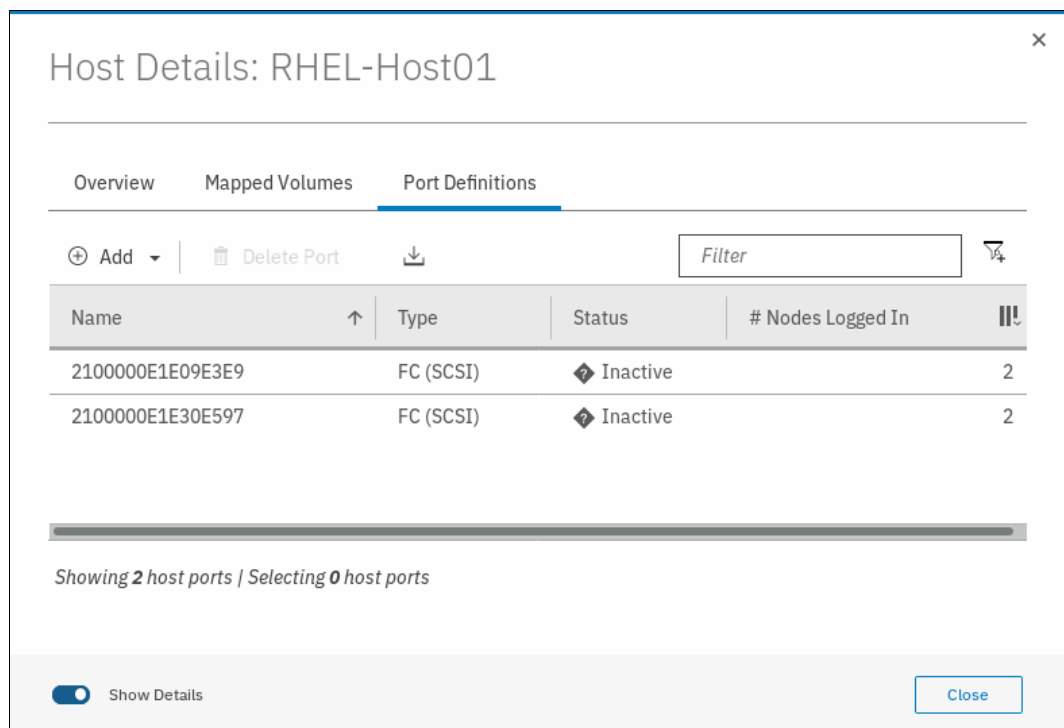


Figure 8-34 Host ports after deletion of a port

8.3 Advanced volume administration

This section describes volume administration tasks, such as volume modification and the creation of volume copies. We assume that you completed Chapter 6, “Volume configuration” on page 309, and that you are familiar with volume creation and generic, thin-provisioned, mirrored, thin-mirrored, and compressed volumes.

Figure 8-35 shows the following advanced feature administration options, which are available in the Volumes menu:

- ▶ Volumes
- ▶ Volumes by Pool
- ▶ Volumes by Host

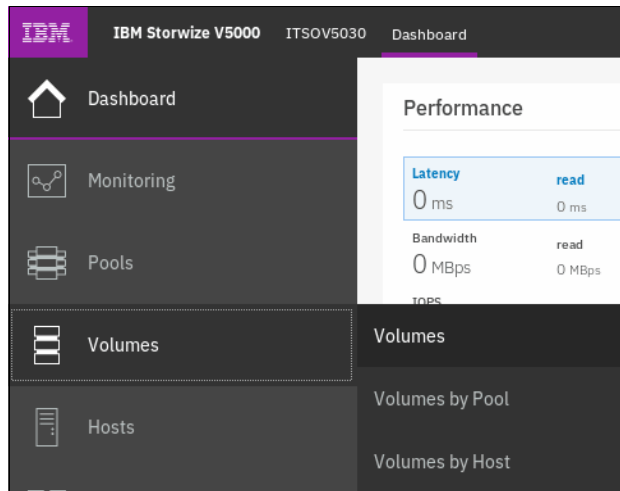


Figure 8-35 Volumes menu

8.3.1 Advanced volume functions

To perform advanced volume functions, complete the following steps:

1. Click **Volumes** and the Volumes window opens, as shown in Figure 8-36.

The screenshot shows the 'Volumes' window in the management console. At the top, there are controls for 'Create Volumes', 'Actions', and 'All Volumes'. There are also dropdown menus for 'Default' and 'Contains', and a 'Filter' input field. Below these controls is a table with the following columns: Name, State, Synchronized, Pool, and UID. The table contains 10 rows of data, all with a 'State' of 'Online'. The first five rows are 'CompressedVolume-01' through 'CompressedVolume-05', all associated with 'ITSO Pool 2'. The last five rows are 'Linux1' through 'Linux5', all associated with 'ITSO Redbook'. At the bottom of the window, it says 'Showing 24 volumes | Selecting 0 volumes'.

Figure 8-36 Volumes window

This window lists all configured volumes on the system and provides the following information:

- Name: Shows the name of the volume. If a twistie sign (>) appears before the name, two copies of this volume exist. Click the twistie sign (>) to expand the view and list the volume copies, as shown in Figure 8-36 on page 407.
- State: Provides the status information about the volume, which can be online, offline, or degraded.
- Synchronized: For mirrored volumes, whether the copies are synchronized.
- Pool: Shows in which storage pool the volume is stored. The primary copy, which is marked with an asterisk (*) is shown unless you expand the volume copies.
- UID: The volume unique identifier.
- Host mappings: Shows whether a volume has host mapping: Yes when host mapping exists and No when no hosting mappings exist.
- Capacity: The disk capacity that is presented to the host. If a thin provisioned or compressed volume exists, an icon is shown next to the capacity. Therefore, the listed capacity is the virtual capacity, which might be larger than the real capacity on the system.

Tip: Right-click anywhere in the blue title bar to customize the volume attributes that are displayed. You might want to add useful information, such as the caching I/O group and the real capacity.

2. To create a volume, click **Create Volumes** and complete the steps that are described in Chapter 6, “Volume configuration” on page 309.
3. Right-clicking or selecting a volume and opening the Actions menu shows the available actions for a volume, as shown in Figure 8-37.

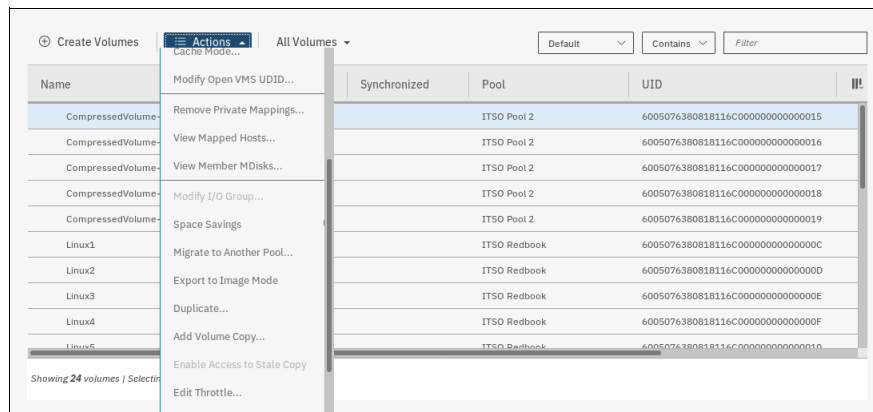


Figure 8-37 Actions menu for a volume

Depending on the Storwize V5000 model capabilities and volume types, the following volume actions are available:

- ▶ Rename (8.3.5, “Renaming a volume” on page 411)
- ▶ Map to Host or Host Cluster (8.3.2, “Mapping a volume to a host” on page 409)
- ▶ Shrink (8.3.6, “Shrinking a volume” on page 412)
- ▶ Expand (8.3.7, “Expanding a volume” on page 413)
- ▶ Modify Capacity Savings (choose between none, Thin Provisioning, and Compression)

- ▶ **Modify Mirror Synch Rate** (Set the synchronization rate value. For more information, see 8.4, “Volume properties and volume copy properties” on page 420.)
- ▶ **Cache Mode** (Choose between Enabled, Read Only, and Disabled.)
- ▶ **Modify Open VMS unit device identifier (UDID)**
- ▶ **Remove Private Mappings** (For more information, see 8.3.3, “Unmapping volumes from private hosts” on page 410)
- ▶ **View Mapped Hosts** (For more information, see 8.3.4, “Viewing which host is mapped to a volume” on page 410)
- ▶ **View Member MDisks**
- ▶ **Modify I/O group** (only applicable to multiple I/O group systems)
- ▶ **Space Savings** (only for compressed volumes)
- ▶ **Migrate to Another Pool** (for more information, see 8.3.8, “Migrating a volume to another storage pool” on page 414)
- ▶ **Export to Image Mode** (for more information, see 8.3.9, “Exporting to an image mode volume” on page 414)
- ▶ **Duplicate** (for more information, see 8.3.11, “Duplicating a volume” on page 417)
- ▶ **Add Volume Copy** (for more information, see 8.3.12, “Adding a volume copy” on page 418)
- ▶ **Enable access to stale copy** (available for IBM HyperSwap volumes if the copy is not up-to-date and inaccessible, but contains consistent data from an earlier time)
- ▶ **Edit Throttle**
- ▶ **View All Throttles**
- ▶ **Delete** (for more information, see 8.3.10, “Deleting a volume” on page 416)
- ▶ **Volume Copy Actions** (for more information, see 8.5, “Advanced volume copy functions” on page 422)
- ▶ **Modify Properties**
- ▶ **Properties** (for more information, see 8.4, “Volume properties and volume copy properties” on page 420)

Other actions are available for copies of volumes. For more information, see 8.5, “Advanced volume copy functions” on page 422.

8.3.2 Mapping a volume to a host

To map a volume to a host, see 6.7, “Mapping Volumes to Host after volume creation” on page 342.

8.3.3 Unmapping volumes from private hosts

To remove all host mappings from a volume, complete the following steps:

1. Select **Remove Private Mappings** from the Actions menu. This action removes all host mappings, which means that no hosts can access this volume.
2. Confirm the number of mappings to remove, and click **Unmap**, as shown in Figure 8-38.

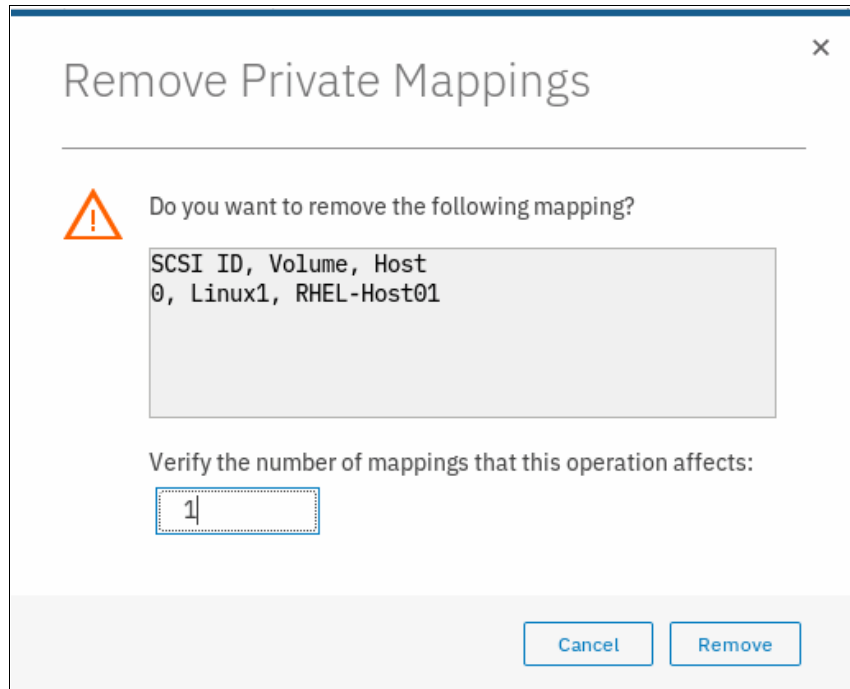


Figure 8-38 Unmapping from host or hosts

3. After the task completes, click **Close** to return to the Volumes window.

Important: Ensure that the required procedures are run on the host OS *before* you run the unmapping procedure.

8.3.4 Viewing which host is mapped to a volume

To determine which host mappings are configured, complete the following steps:

1. Highlight a volume and select **View Mapped Host** from the Actions menu. The Host Maps tab of the Volume Details window opens, as shown in Figure 8-39 on page 411. In this example, host RHEL-Host01 is mapped to the Linux1 volume.

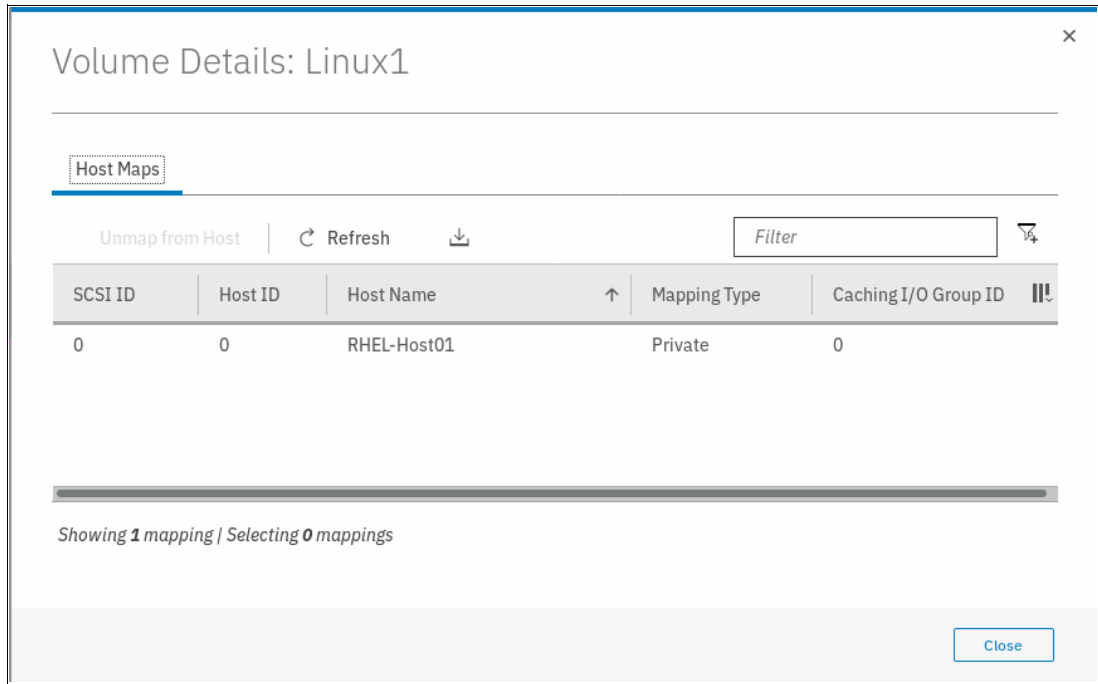


Figure 8-39 Volume to host mapping

2. To remove a mapping, highlight the host and click **Unmap from Host**. If several hosts are mapped to this volume, only the selected host is removed.

8.3.5 Renaming a volume

To rename a volume, complete the following steps:

1. Select **Rename** from the Actions menu. The Rename Volume window opens.
2. Enter the new name, as shown in Figure 8-40.



Figure 8-40 Renaming a volume

3. Click **Reset** to reset the name field to the original name of the volume.
4. Click **Rename** to apply the changes.
5. Click **Close** to close the window.

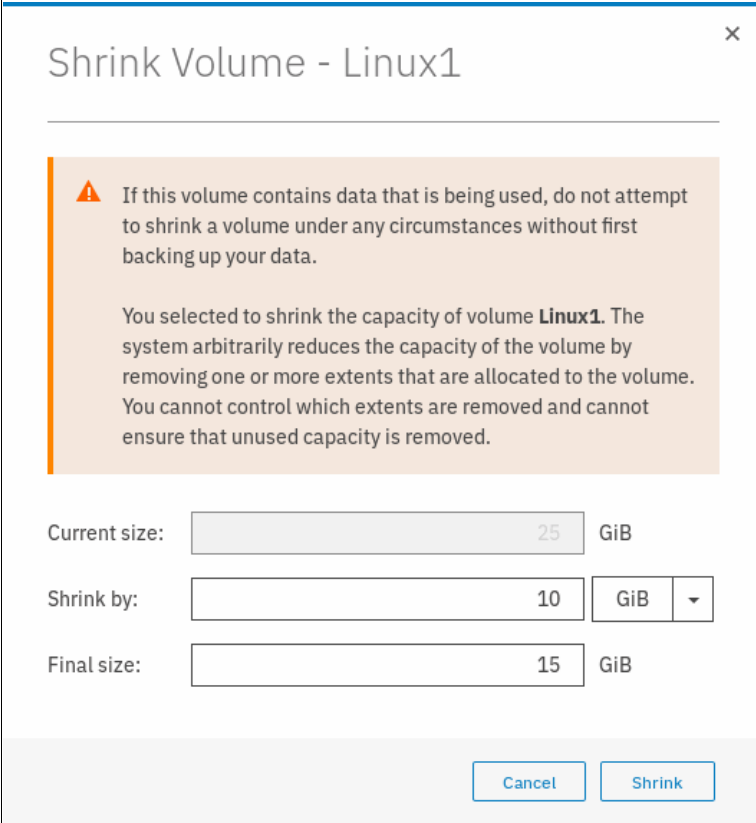
8.3.6 Shrinking a volume

IBM Storwize V5000 can shrink volumes. To shrink a volume, complete the following steps:

1. Select **Shrink** from the Actions menu.
2. Enter the new volume size or the value by which to shrink the volume, as shown in Figure 8-41.

Important: If the volume contains data that is used, do not attempt to shrink a volume under any circumstances without first backing up your data.

The system arbitrarily reduces the capacity of the volume by removing one or more extents that are allocated to the volume. You cannot control which extents are removed or ensure that unused capacity is removed.



Shrink Volume - Linux1

⚠ If this volume contains data that is being used, do not attempt to shrink a volume under any circumstances without first backing up your data.

You selected to shrink the capacity of volume **Linux1**. The system arbitrarily reduces the capacity of the volume by removing one or more extents that are allocated to the volume. You cannot control which extents are removed and cannot ensure that unused capacity is removed.

Current size: GiB

Shrink by: GiB

Final size: GiB

Figure 8-41 Shrink Volume window

3. Click **Shrink** to start the process.
4. Click **Close** when the task completes to return to the Volumes window.
5. Run the required procedures on the host OS after the shrinking process.

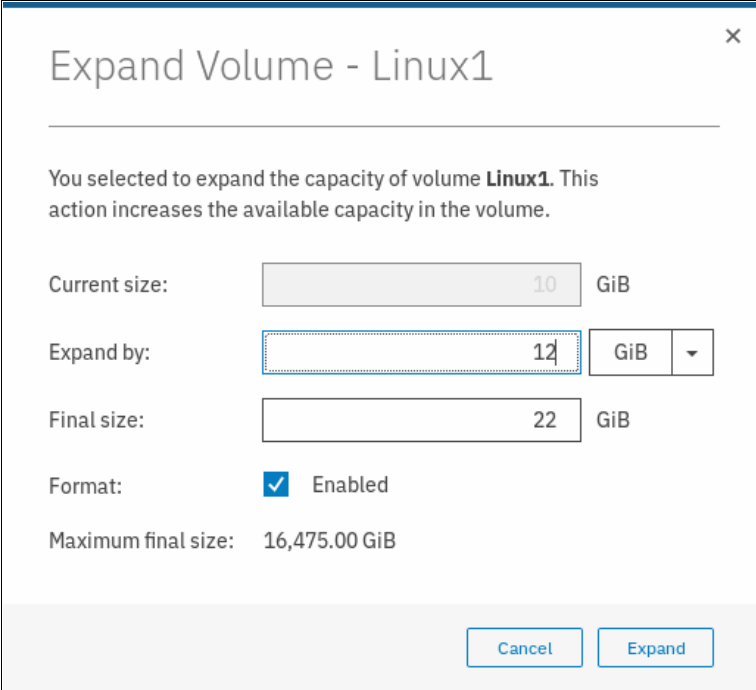
Important: For volumes that contain more than one copy, you might receive a CMMVC6354E error. Check the Running tasks window and wait for the copy to synchronize. If you want the synchronization process to complete more quickly, increase the rate by increasing the Mirror Sync Rate value in the Actions menu. When the copy is synchronized, resubmit the shrink process.

Similar errors might occur if other tasks (for example, volume expand or format operations) are running on the volume. The solution is to wait until these operations finish, then restart the shrink process.

8.3.7 Expanding a volume

The IBM Storwize V5000 can expand volumes. Use this feature only if the host OS supports it. This capability increases the capacity that is allocated to the particular volume by the amount that is specified. To expand a volume, complete the following steps:

1. Select **Expand** from the Actions menu.
2. Enter the new volume size or enter the amount by which the volume needs to expand.
3. Click **Expand**, as shown in Figure 8-42.



Expand Volume - Linux1

You selected to expand the capacity of volume **Linux1**. This action increases the available capacity in the volume.

Current size: 10 GiB

Expand by: 12 GiB

Final size: 22 GiB

Format: Enabled

Maximum final size: 16,475.00 GiB

Cancel Expand

Figure 8-42 Expand Volume window

4. If the task completion window remains open, review the results of the operation and click **Close** to return to the Volumes window.
5. Run the required procedures on the host operating system to use the full available space.

Note: You can expand the capacity of volumes in Metro Mirror and Global Mirror relationships that are in `consistent_synchronized` state if those volumes are using thin-provisioned or compressed copies.

You cannot expand the following types of volumes:

- ▶ Volumes in HyperSwap relationships or in Global Mirror relationships that are operating in cycling mode
- ▶ Volumes in relationships where a change volume is configured
- ▶ Volumes that have a fully allocated copy

For more information about volume expansion, see [this website](#).

8.3.8 Migrating a volume to another storage pool

For more information about migrating a volume to another storage pool, see 6.8, “Migrating a volume to another storage pool” on page 347.

8.3.9 Exporting to an image mode volume

Image mode provides a direct block-for-block translation from a managed disk (MDisk) to a volume with no virtualization. An image mode MDisk is associated with one volume only. This feature can be used to export a volume to a non-virtualized disk and to remove the volume from storage virtualization.

Note: Among the IBM Storwize V5000 family, this feature is available only on the IBM Storwize V5030 storage systems.

To export to an image mode volume, complete the following steps:

1. Select the volume that you want. From the Actions menu, choose **Export to Image Mode**, as shown in Figure 8-43.

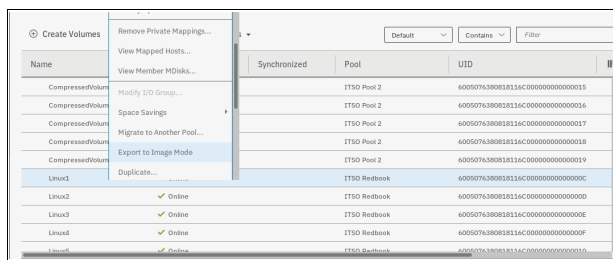


Figure 8-43 Exporting a volume to an image mode

- The Export to Image Mode wizard opens and displays the available MDisks. Select the MDisk to which to export the volume, and click **Next**, as shown in Figure 8-44.

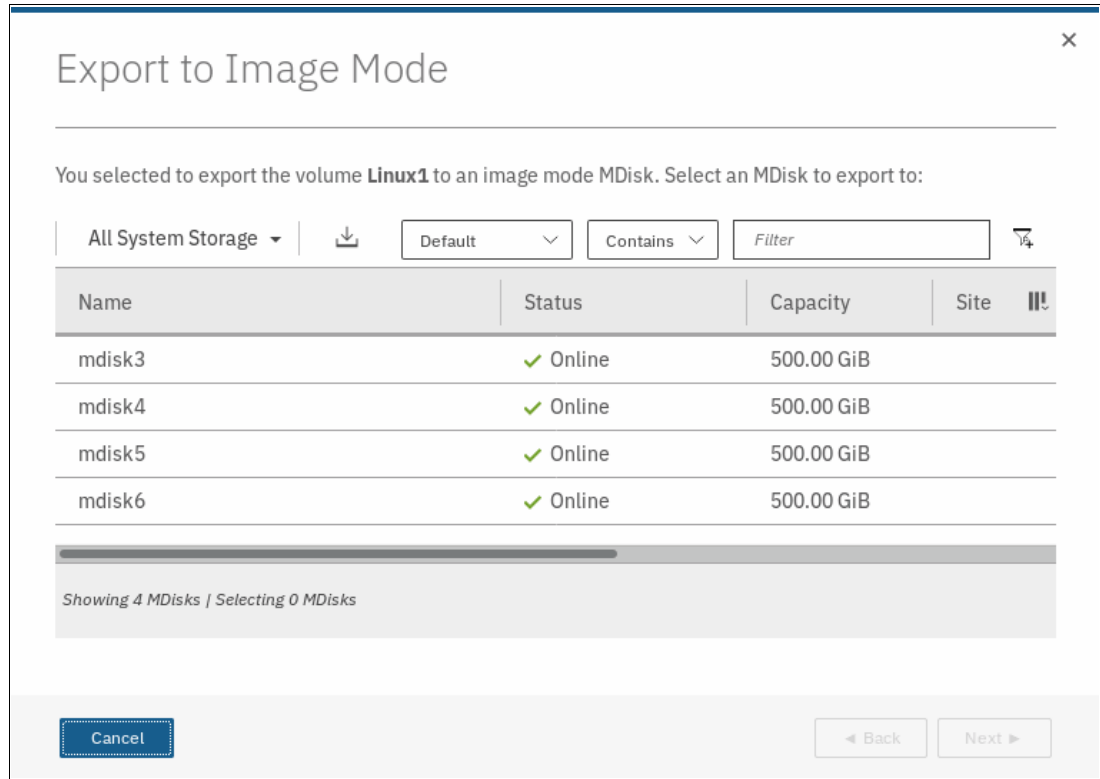


Figure 8-44 Selecting the MDisk to which to export the volume

3. Select a storage pool into which the image-mode volume is placed after the migration completes, as shown in Figure 8-45.

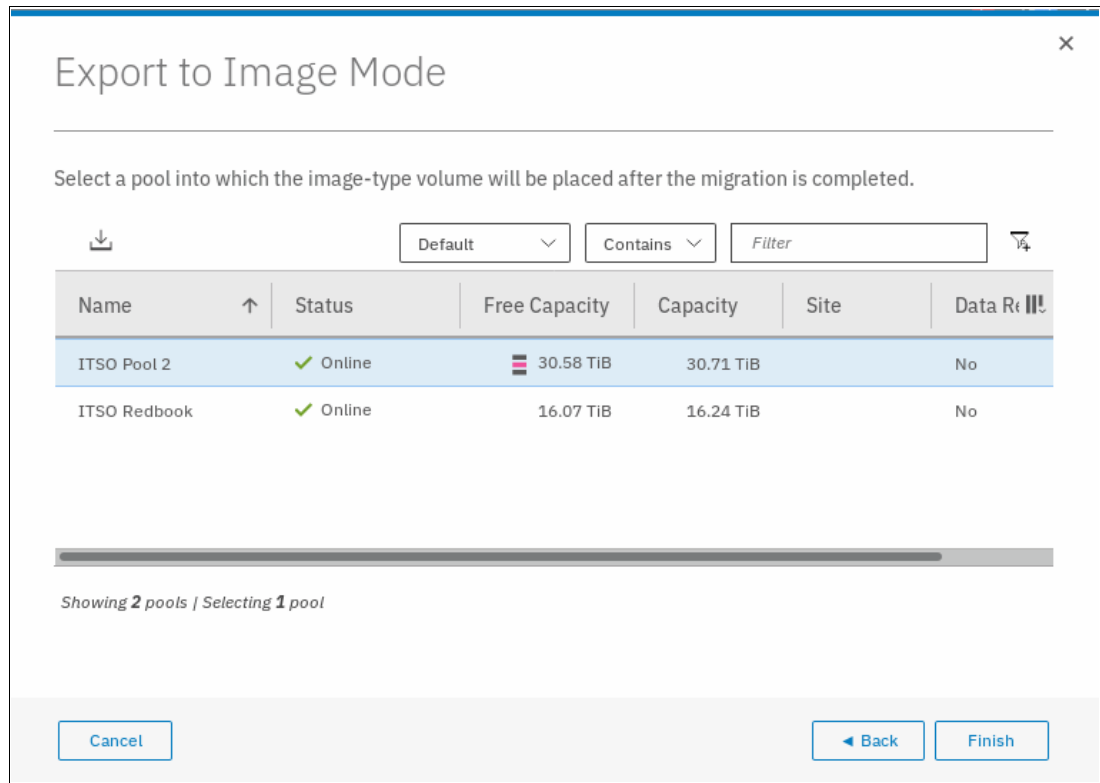


Figure 8-45 Select the storage pool

4. Click **Finish** to start the migration. After the task is complete, check the results and click **Close** to return to the Volumes window.

Important: Use image mode to import or export data into or out of the IBM Storwize V5000. Migrate data from image mode MDisks to other storage pools to benefit from storage virtualization.

For more information about importing volumes from external storage, see Chapter 7, “Storage migration” on page 357, and Chapter 4, “Storage pools” on page 159.

8.3.10 Deleting a volume

To delete a volume, complete the following steps:

1. Select **Delete** from the Actions menu.
2. Confirm the number of volumes and select the option if you want to force the deletion. Figure 8-46 on page 417 shows the Delete Volume window.

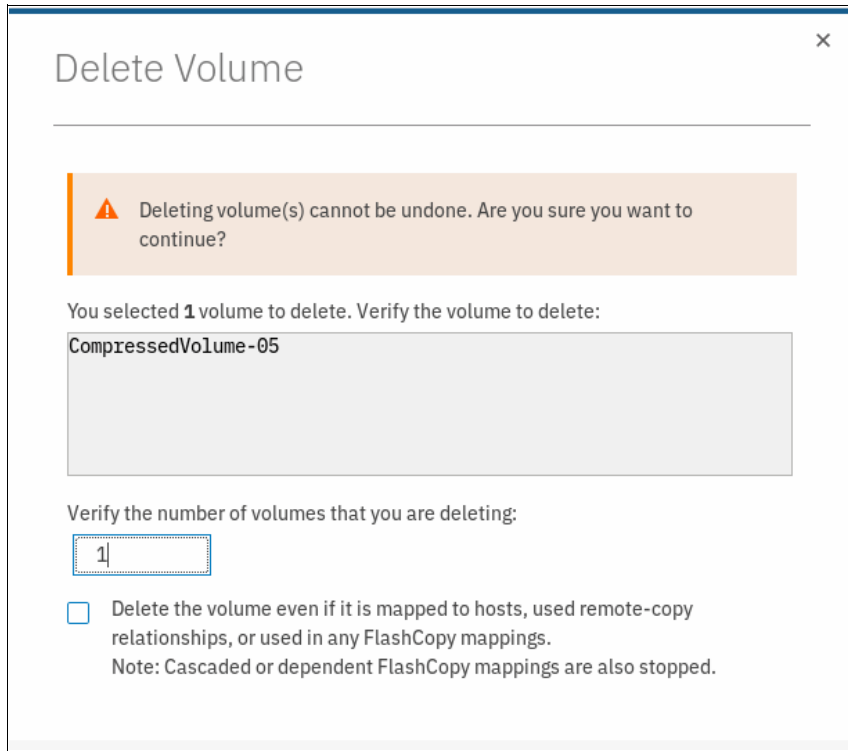


Figure 8-46 Delete Volume window

3. Click **Delete** to remove the selected volume or volumes from the system. After the task completes, click **Close** to return to the Volumes window.

Important: You must force the deletion if the volume has host mappings or if the volume is used in FlashCopy mappings. To be cautious, always ensure that the volume has no association before you delete it.

8.3.11 Duplicating a volume

You can create a volume by using the same presets and parameters as an existing volume. The following parameters are shown:

- ▶ Volume preset (generic, thin-provision, and compressed)
- ▶ Volume size
- ▶ Storage pool
- ▶ Access and caching I/O group
- ▶ Caching mode
- ▶ Easy Tier status
- ▶ Virtualization type

Important: Duplicating a volume does not duplicate the volume data. The duplicating task creates a volume with the same preset and volume parameters as the source volume. Duplicating mirrored and image-mode volumes is not supported.

To duplicate a volume, complete the following steps:

1. Select **Duplicate** from the Actions menu (see Figure 8-47).

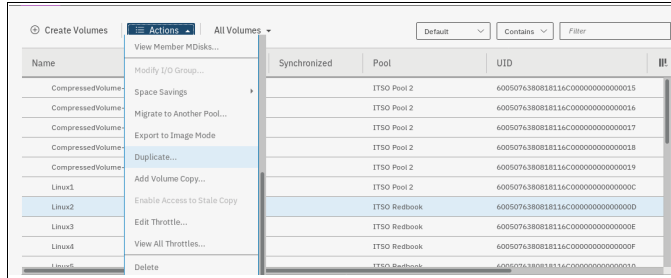


Figure 8-47 Duplicate volume option

2. The Duplicate Volume window (see Figure 8-48) can be used to change the name of the new volume. By default, a sequence integer is appended to the name of the volume to be duplicated.

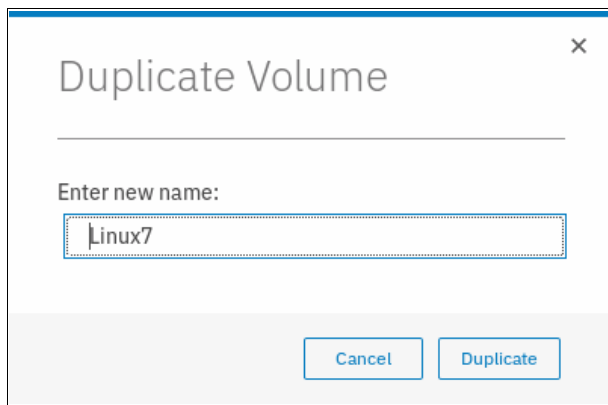


Figure 8-48 Duplicate Volume

3. Click **Duplicate** to start the process. If the task completion window remains open, review the process results and click **Close**.

8.3.12 Adding a volume copy

If a volume consists of only one copy, you can add a mirrored copy of the volume. This copy can be generic or thin-provisioned.

You can also use this method to migrate data across storage pools with different extent sizes.

To add a second copy, complete the following steps:

1. Select the volume and click **Actions** → **Add Volume Copy**, as shown in Figure 8-49.

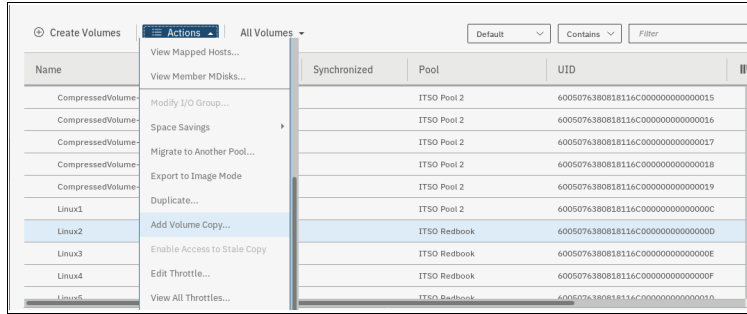


Figure 8-49 Add Volume Copy option

2. Select the storage pool in which to create the copy. Select capacity savings, between None, Thin-provisioned, or Compressed. Click **Add**, as shown in Figure 8-50.

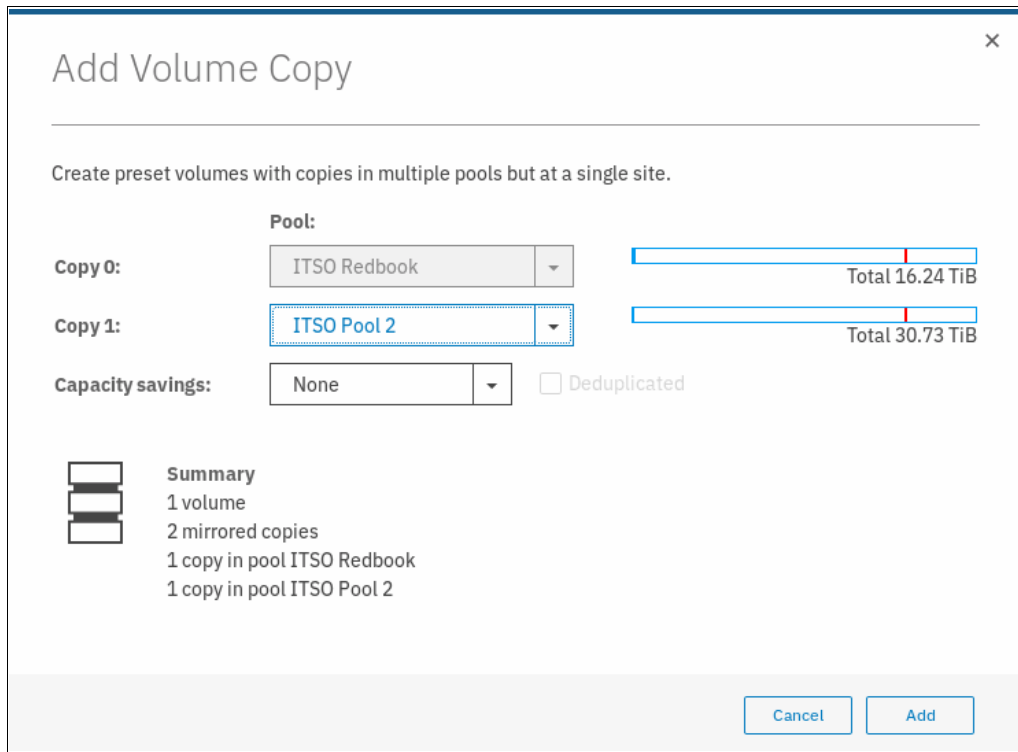


Figure 8-50 Add Volume Copy: Select a storage pool

3. The copy is created after you click **Add** and data starts to synchronize as a background task. If the task completion window remains open, review the results and click **Close**.

The volume that is named Linux2 now has two volume copies, which are stored in separate storage pools, as shown in Figure 8-51.

Name	State	Synchronized	Pool	UID	
CompressedVolume-05	✓ Online		ITSO Pool 2	6005076380818116C000000000000019	
Linux1	✓ Online		ITSO Pool 2	6005076380818116C00000000000000C	
Linux2	✓ Online		ITSO Redbook	6005076380818116C00000000000000D	
Copy 0*	✓ Online	Yes	ITSO Redbook	6005076380818116C00000000000000D	
Copy 1	✓ Online	No	ITSO Pool 2	6005076380818116C00000000000000D	
Linux3	✓ Online		ITSO Redbook	6005076380818116C00000000000000E	

Figure 8-51 Volume copies

Note: If your volume is in a standard storage pool and you want to add a volume copy in a Data Reduction Pool, the only available capacity saving options are Thin-provisioned and None. If your volume is in a Data Reduction Pool and you want to add a volume copy in a standard storage pool, all capacity saving options are available.

8.4 Volume properties and volume copy properties

This section provides an overview of the IBM Storwize V5000 volumes.

To open the advanced view of a volume, complete the following steps:

1. Select **Properties** from the Actions menu. The full list of volume properties is shown (see Figure 8-52).

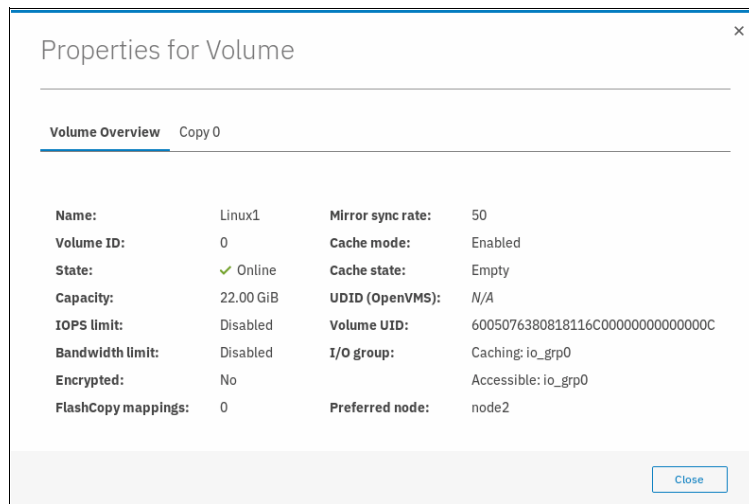


Figure 8-52 Volume details overview

The following details are available on Volume Overview tab:

- Name: The name of the volume.
- ID: The SCSI ID of the volume.
- State: Status information about the volume, which can be online, offline, or degraded.

- Capacity: The capacity of the volume. If the volume is thin-provisioned or compressed, this number is the virtual capacity.
 - IOPS Limit: The maximum rate at which the volume processes the I/O operations in IOPS, in case you set I/O throttling for the volume. Disabled means there is no IOPS I/O throttling enabled for the volume.
 - Bandwidth Limit: The maximum rate at which the volume processes the I/O operations in MBps, in case you set I/O throttling for the volume. Disabled means there is no bandwidth I/O throttling enabled for the volume.
 - Encrypted: Whether the volume is encrypted.
 - FlashCopy Mappings: The number of existing FlashCopy relationships. For more information, see Chapter 10, “Copy Services” on page 465.
 - Mirror sync rate: The ratio that determines how quickly the copies of the volume are synchronized. It also affects the format process speed. By default, this rate is set to 50; 0 means no synchronization occurs.
 - Cache mode: Whether the cache is enabled or disabled for this volume.
 - Cache state: Whether open I/O requests are in the cache that is not destaged to the disks.
 - UDID (OpenVMS): The unit device identifiers (UDIDs) that are used by OpenVMS hosts to access the volume.
 - Volume UID: The volume unique identifier.
 - I/O Group: The volume caching I/O group and the I/O groups that the host can use to access the volume.
 - Preferred node: The ID of the preferred node for the volume.
2. To open the advanced view of a volume copy, select the main volume properties. Click the wanted copy tab, as shown in Figure 8-53.

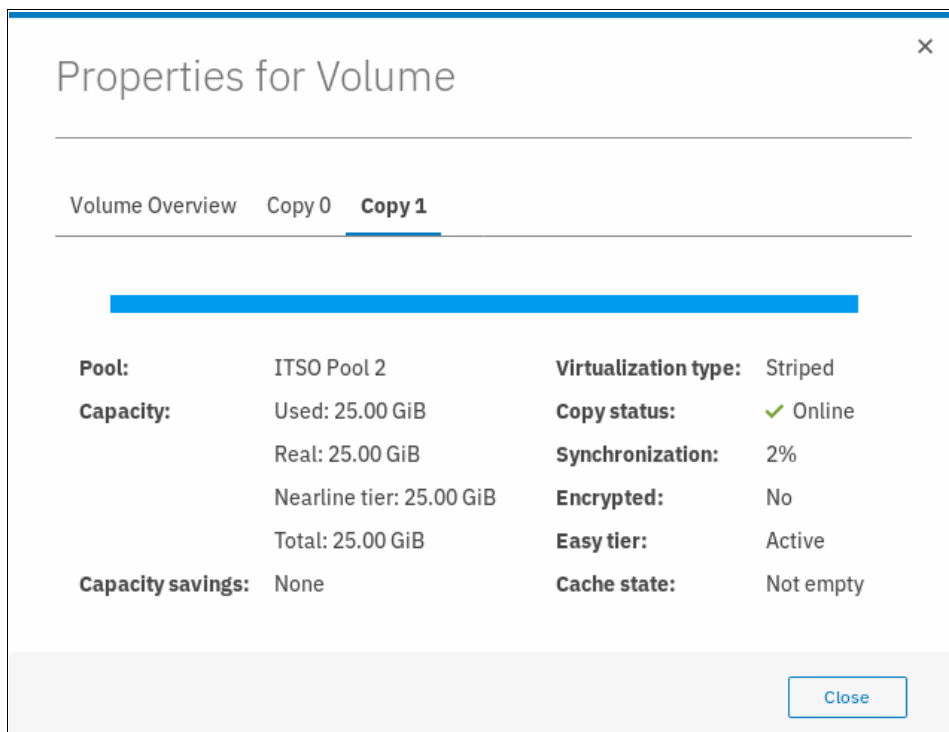


Figure 8-53 Volume copy properties

The following I volume copy information is available:

- Pool: The storage pool where the volume is located.
- Capacity: The capacity of the volume and the capacity on each tier. If the volume is thin-provisioned or compressed, it also shows the real and the used capacity.
- Capacity Savings: The capacity savings type of the volume, which can be Thin-provisioned, Compressed, or None.
- Virtualization type: The virtualization type of the volume. The value can be striped, seq, or image.
- Copy status: The status of the volume copy.
- Synchronization: The status of the copy synchronization process.
- Encrypted: Whether the volume copy is encrypted.
- Easy tier: The state of the Easy Tier.
- Cache state: Whether open I/O requests are in the cache that is not destaged to the disks.

3. To change the volume copy properties for the thin-provisioned or compressed volume copy type, select a volume copy from the Volumes window, and click **Actions** → **Modify Properties**. Use the Modify Properties window (see Figure 8-54) to customize the thin-provisioning values: Warning Threshold and Enable Autoexpand.

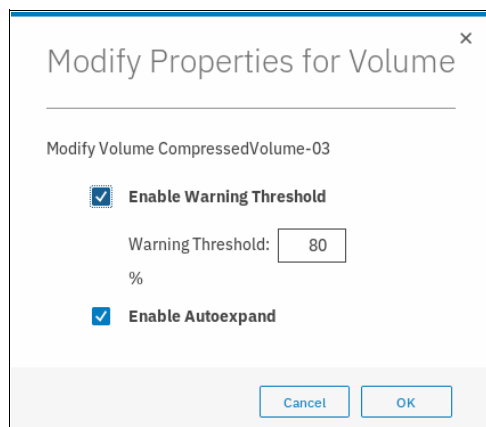


Figure 8-54 Modify volume copy properties

4. Modify the values if needed and click **OK**. After the task completes, check the results of the operation and click **Close** to return to the Volumes window.

8.5 Advanced volume copy functions

In 8.3.1, “Advanced volume functions” on page 407, we described all of the available actions at a volume level and how to create a second volume copy. In this section, we focus on volumes that consist of two volume copies and how to apply the concept of two copies for business continuity and data migration.

To access advanced volume copy functions, complete the following steps:

1. Select the wanted volume for multiple copies.
2. Click the twistie (>) to show all the copies.

3. Select a volume copy.
4. Open the Actions menu to display the following volume copy actions (see Figure 8-55).

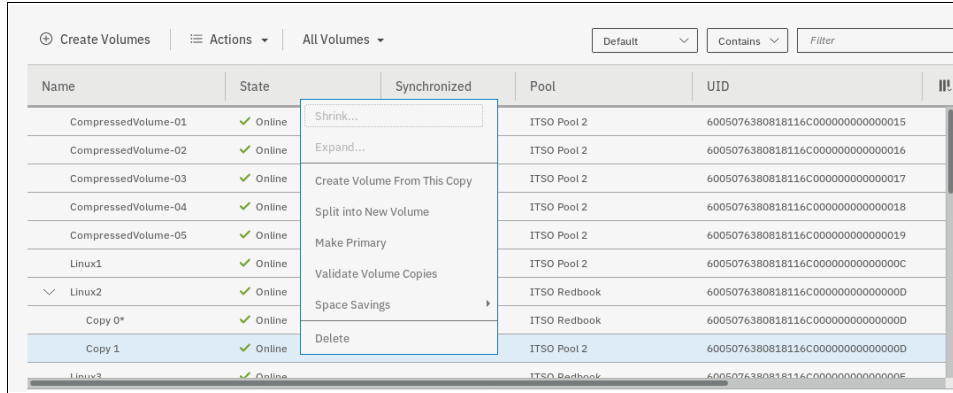


Figure 8-55 Volume Copy Actions menu

The following functions are available on the Actions menu:

- ▶ Create a volume from this copy
- ▶ Split into a new volume (for more information, see 8.5.2, “Splitting into a new volume” on page 424)
- ▶ Make primary (for more information, see 8.5.1, “Volume copy: Make Primary” on page 423)
- ▶ Validate volume copies (for more information, see 8.5.3, “Validate Volume Copies option” on page 426)
- ▶ Space Savings
- ▶ Delete (for more information, see 8.5.4, “Delete volume copy option” on page 428)

8.5.1 Volume copy: Make Primary

When you review the volume copies, you can see that one of the copies shows an asterisk (*) next to its name, as shown in Figure 8-56.

Name	State	Synchronized	Pool	UID	!!!
CompressedVolume-05	✓ Online		ITSO Pool 2	6005076380818116C000000000000019	
Linux1	✓ Online		ITSO Pool 2	6005076380818116C00000000000000C	
Linux2	✓ Online		ITSO Redbook	6005076380818116C00000000000000D	
Copy 0*	✓ Online	Yes	ITSO Redbook	6005076380818116C00000000000000D	
Copy 1	✓ Online	Yes	ITSO Pool 2	6005076380818116C00000000000000D	
Linux3	✓ Online		ITSO Redbook	6005076380818116C00000000000000E	

Figure 8-56 Volume copy names

Each volume has a primary and a secondary copy, and the asterisk indicates the primary copy. The two copies are always synchronized, which means that all writes are destaged to both copies, but all reads are always performed from the primary copy. The maximum configurable number of copies per volume is two. The roles of the copies can be changed.

To change roles of the copies, complete the following steps:

1. Select the secondary copy. Click **Actions** → **Make Primary**. Usually, it is a preferred practice to place the volume copies on storage pools with similar performance because the write performance is constrained if one copy is placed on a lower-performance pool.

Figure 8-57 shows the secondary copy Actions menu.

Name	State	Synchronized	Pool	UID
CompressedVolume-01	✓ Online		ITSO Pool 2	6005076380818116C000000000000015
CompressedVolume-02	✓ Online		ITSO Pool 2	6005076380818116C000000000000016
CompressedVolume-03	✓ Online		ITSO Pool 2	6005076380818116C000000000000017
CompressedVolume-04	✓ Online		ITSO Pool 2	6005076380818116C000000000000018
CompressedVolume-05	✓ Online		ITSO Pool 2	6005076380818116C000000000000019
Linux1	✓ Online		ITSO Pool 2	6005076380818116C00000000000000C
Linux2	✓ Online		ITSO Redbook	6005076380818116C00000000000000D
Copy 0*	✓ Online		ITSO Redbook	6005076380818116C00000000000000D
Copy 1	✓ Online		ITSO Pool 2	6005076380818116C00000000000000D
Linux 3	✓ Online		ITSO Redbook	6005076380818116C00000000000000E

Figure 8-57 Make primary option

2. If you require high read performance, you can place the primary copy in a solid-state drive (SSD) pool or an externally virtualized Flash System and then place the secondary copy in a normal disk storage pool. This action maximizes the read performance of the volume and guarantees that a synchronized second copy is in your less expensive disk pool.

You also can migrate online copies between storage pools. For more information about how to select the copy that you want to migrate, see 8.3.8, “Migrating a volume to another storage pool” on page 414.

3. Click **Make Primary** and the role of the Copy 1 is changed to Primary, as shown in Figure 8-58.

Name	State	Synchronized	Pool	UID
CompressedVolume-05	✓ Online		ITSO Pool 2	6005076380818116C000000000000019
Linux1	✓ Online		ITSO Pool 2	6005076380818116C00000000000000C
Linux2	✓ Online		ITSO Pool 2	6005076380818116C00000000000000D
Copy 0	✓ Online	Yes	ITSO Redbook	6005076380818116C00000000000000D
Copy 1*	✓ Online	Yes	ITSO Pool 2	6005076380818116C00000000000000D
Linux3	✓ Online		ITSO Redbook	6005076380818116C00000000000000E

Figure 8-58 Change the primary volume copy

4. If the task completion window remains open, check the process output and click **Close**.

The volume copy feature is also a powerful option for migrating volumes, as described in 8.5.5, “Migrating volumes using the volume copy features” on page 429.

8.5.2 Splitting into a new volume

If the two-volume copies are synchronized, you can split one of the copies to a new volume and map this volume to another host. From a storage perspective, this procedure can be performed online, which means that you can split one copy from the volume and create a copy from the remaining volume without affecting the host. However, if you want to use the split copy for testing or backup, you must ensure that the data inside the volume is consistent. Therefore, the data must be flushed to storage to make the copies consistent.

For more information about flushing the data, see your operating system documentation. The easiest way to flush the data is to shut down the hosts or application before a copy is split.

In our example, volume Linux2 has two copies: Copy 0 is primary and Copy 1 is secondary.

To split a copy, complete the following steps:

1. Click **Split into New Volume** (see Figure 8-59) on any copy and the remaining secondary copy automatically becomes the primary for the source volume.

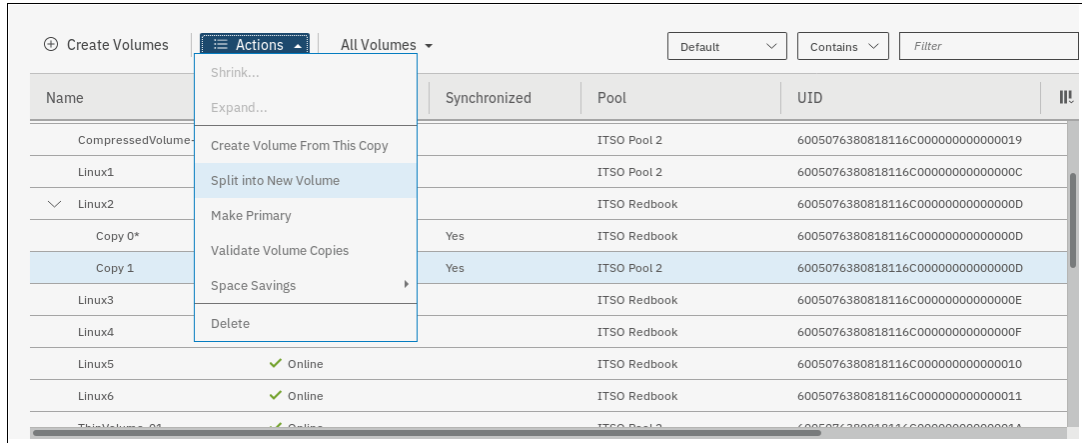


Figure 8-59 Split into New Volume option

Figure 8-60 shows the Split Volume Copy window to specify a name for the new volume.

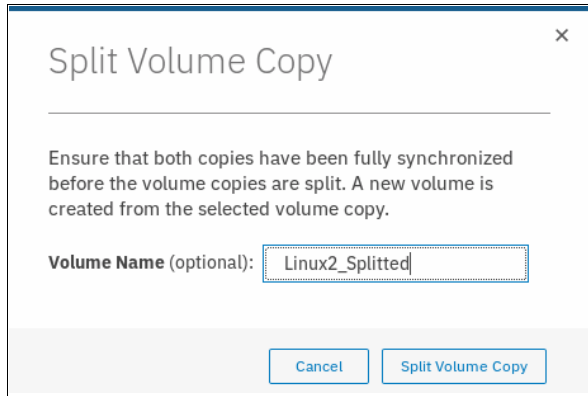


Figure 8-60 Split Volume Copy window

2. If the task completion window remains open after the task completes, review the results and click **Close** to return to the Volumes window.

As shown in Figure 8-61, the copy appears as a new volume that is named Linux2_Splitted (as specified during the split process). The new volume can be mapped to a host.

Name	State	Synchronized	Pool	UID
Linux2	✓ Online		ITSO Redbook	6005076380818116C00000000000000D
Linux2_Splitted	✓ Online		ITSO Pool 2	6005076380818116C00000000000002B

Figure 8-61 Volumes: New volume from the split copy operation

Important: If you receive error message code CMMVC6357E while you are splitting a volume copy, click the **Running Tasks** icon to view the synchronization status. Then, wait for the copy to synchronize and repeat the splitting process.

8.5.3 Validate Volume Copies option

You can check whether the volume copies are identical and process the differences between them if they are not.

To validate the copies of a mirrored volume, complete the following steps:

1. From the Actions menu, select **Validate Volume Copies**, as shown in Figure 8-62.

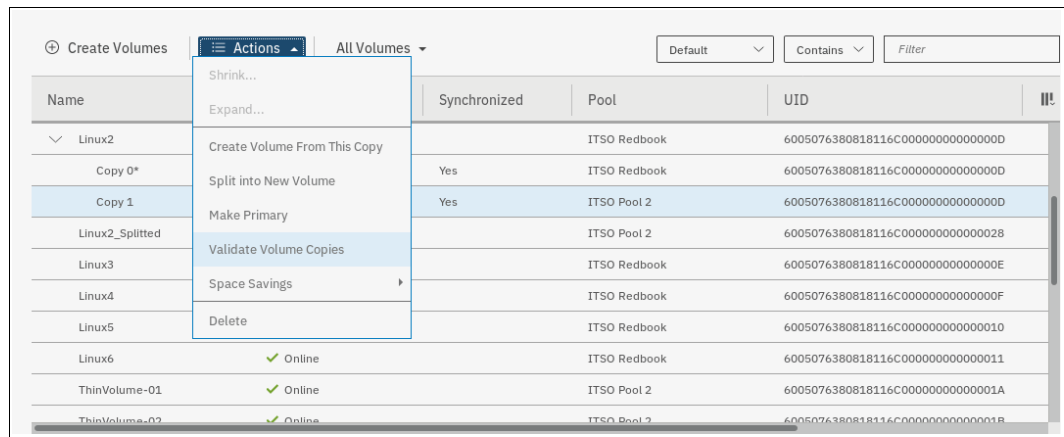


Figure 8-62 Actions menu: Validate Volume Copies

The Validate Volume Copies window opens, as shown in Figure 8-63.

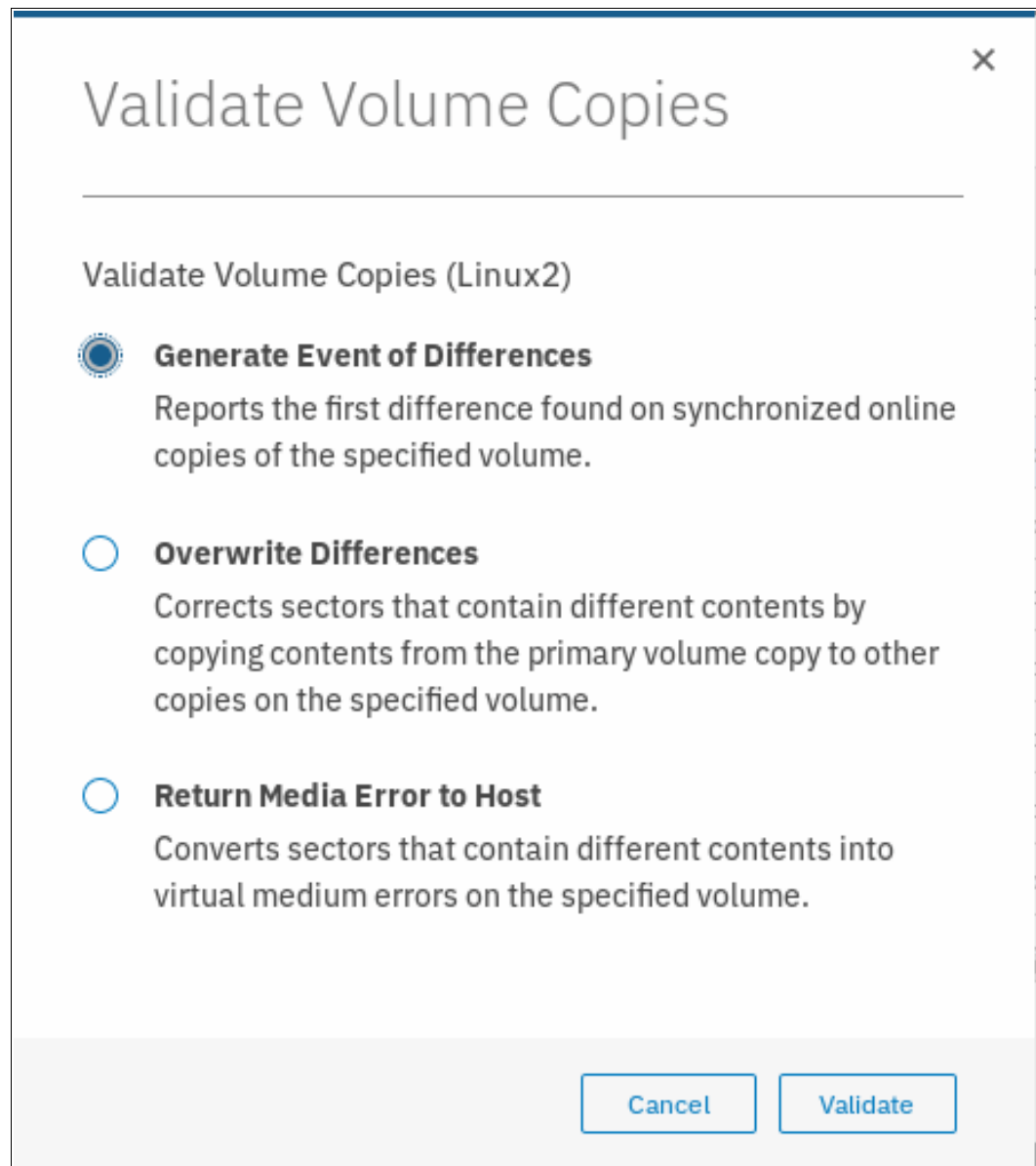


Figure 8-63 Validate Volume Copies window

The following options are available:

- **Generate Event of Differences**
Use this option if you want to verify that the mirrored volume copies are identical. If any difference is identified, the command stops and logs an error that includes the logical block address (LBA) and the length of the first difference. You can use this option, starting at a different LBA each time, to count the number of differences on a volume.
- **Overwrite Differences**
Use this option to overwrite contents from the primary volume copy to the other volume copy. The command corrects any differing sectors by copying the sectors from the primary copy to the copies, which are compared. Upon completion, the command process logs an event, which indicates the number of differences that were corrected.

Use this option if you are sure that the primary volume copy data is correct or that your host applications can handle incorrect data.

– Return Media Error to Host

Use this option to convert sectors on all volume copies, which contain different contents, into virtual medium errors. Upon completion, the command logs an event, which indicates the number of differences that were found, the number that were converted into medium errors, and the number that were not converted. Use this option if you are unsure what the correct data is and you do not want an incorrect version of the data to be used.

2. Select which action to perform and click **Validate** to start the task.
3. The volume is now checked. If the task window remains open, review the task results and click **Close**.

The validation process runs as a background process and takes time, depending on the volume size. You can check the status in the Background Tasks window, as shown in Figure 8-64.

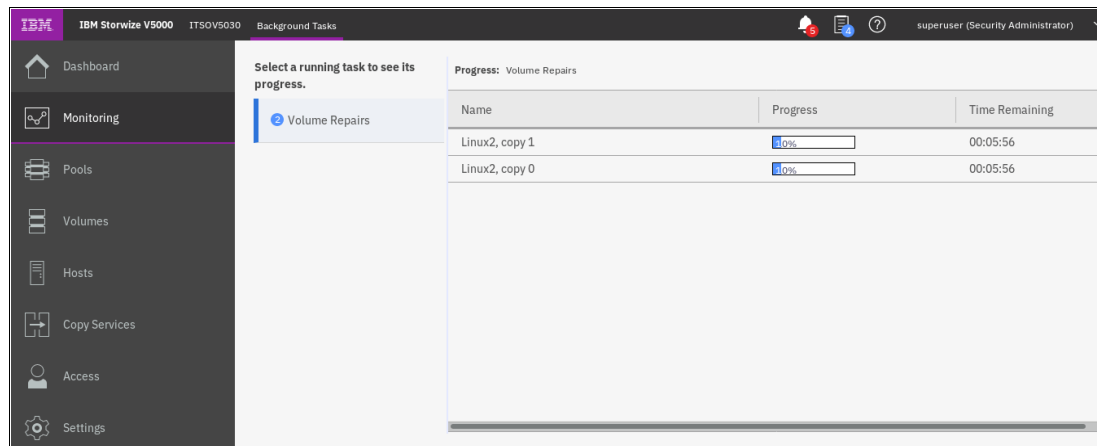


Figure 8-64 Validate Volume Copies: Running Tasks

8.5.4 Delete volume copy option

To delete a volume copy, complete the following steps:

1. Click **Delete** (as shown in Figure 8-65) to delete a volume copy.

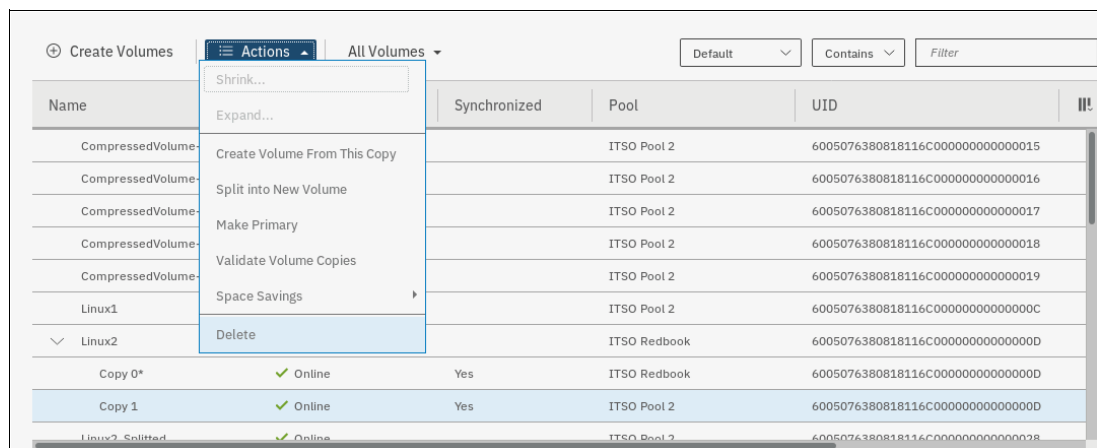


Figure 8-65 Actions menu: Delete a volume copy

2. Confirm the deletion process by clicking **Yes**. Figure 8-66 shows the copy deletion warning window.

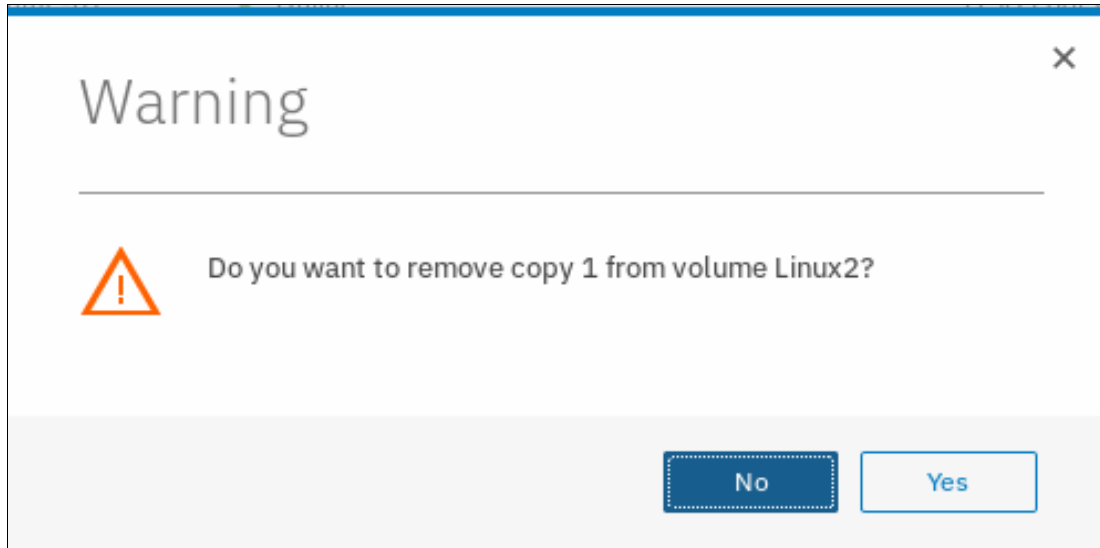


Figure 8-66 Delete a copy confirmation

3. If the task completion window remains open after the copy is deleted, review the results of the operation and click **Close** to return to the Volumes window.

The copy is deleted, but the volume stays online by using the remaining copy (see Figure 8-67).

Name	State	Synchronized	Pool	UID	
Linux2	✓ Online		ITSO Redbook	6005076380818116C000000000000000	!!!

Figure 8-67 Volume still online using the remaining copy

8.5.5 Migrating volumes using the volume copy features

In the previous sections, we showed how to create, synchronize, split, and delete volume copies. A combination of these tasks can be used to migrate volumes to other storage pools.

The easiest way to migrate volume copies is to use the migration feature that is described in 8.3.8, “Migrating a volume to another storage pool” on page 414. By using this feature, one extent after another is migrated to the new storage pool. However, the use of volume copies provides another way to migrate volumes if the storage pool extent sizes differ.

To migrate a volume, complete the following steps:

1. Create a second copy of your volume in the target storage pool. For more information, see 8.3.12, “Adding a volume copy” on page 418.
2. Wait until the copies are synchronized.
3. Change the role of the copies and make the new copy the primary copy. For more information, see 8.5, “Advanced volume copy functions” on page 422.
4. Split or delete the old copy from the volume. For more information, see 8.5.2, “Splitting into a new volume” on page 424 or 8.5.4, “Delete volume copy option” on page 428.

This migration process requires more user interaction with the IBM Storwize V5000 GUI, but it offers benefits. For example, we look at migrating a volume from a tier 1 storage pool to a lower-performance tier 2 storage pool.

In step 1, you create the copy on the tier 2 pool, while all reads are still performed in the tier 1 pool to the primary copy. After the synchronization, all writes are destaged to both pools, but the reads are still only from the primary copy.

Because the copies are fully synchronized, you can switch their roles online (step 3), and analyze the performance of the new pool. After you test your lower performance pool, you can split or delete the old copy in tier 1 or switch back to tier 1 in seconds if the tier 2 storage pool did not meet your requirements.

8.6 Volumes by storage pool

To see the layout of volumes within pools, complete the following steps:

1. Click **Volumes by Pool**, as shown in Figure 8-68.

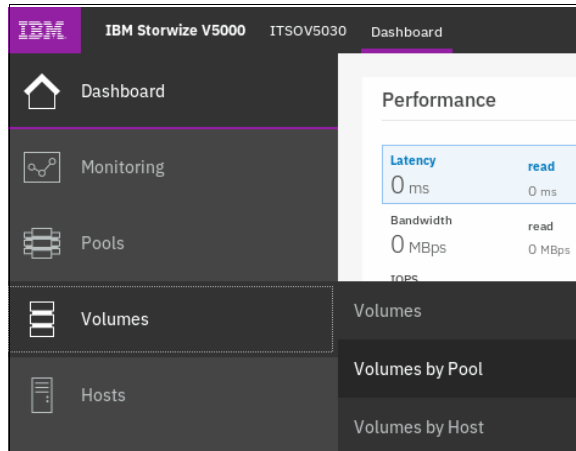


Figure 8-68 Volumes by Pool menu

The Volumes by Pool window opens, as shown in Figure 8-69.

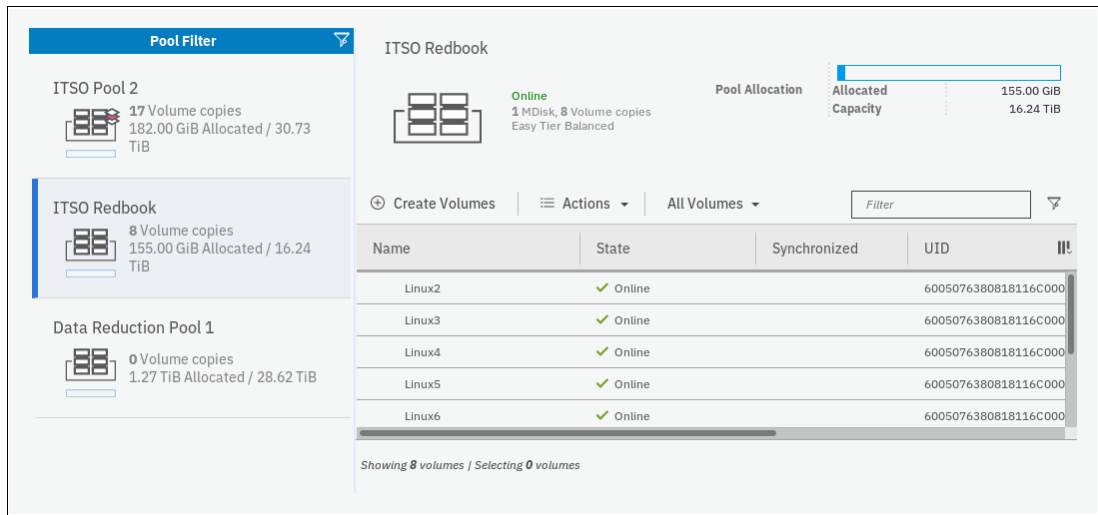


Figure 8-69 Volumes by Pool window

The left pane is called the *pool filter*. The storage pools are displayed in the pool filter. For more information about storage pools, see Chapter 4, “Storage pools” on page 159.

In the upper right, you see information about the pool that you selected in the pool filter. The following information is also shown:

- Pool icon: Because storage pools can have different characteristics, you can change the storage pool icon by clicking it. For more information, see 4.2, “Working with storage pools” on page 172.
- Pool name: The name that was entered when the storage pool was created. Click it to change the name, if needed.
- Pool details: Shows you the information about the storage pools, such as the status, number of managed disks, and Easy Tier status.
- Volume allocation: Shows you the amount of capacity that is allocated to volumes from this storage pool.

The lower-right section lists all volumes with at least one copy in the selected storage pool. The following information is provided:

- Name: Shows the name of the volume.
- State: Shows the status of the volume.
- Synchronized: Shows whether the volume copies are synchronized.
- UID: Shows the volume unique identifier (UID).
- Host mappings: Shows whether host mappings exist.
- Capacity: Shows the capacity that is presented to hosts.

2. You can create volumes from this window. Click **Create Volumes** to open the Volume Creation window. For more information, see Chapter 6, “Volume configuration” on page 309.

Selecting a volume and opening the Actions menu or right-clicking the volume shows the same options, as described in 8.3, “Advanced volume administration” on page 407.

8.7 Volumes by host

To see an overview of the volumes that a host can access, complete the following steps:

1. Click **Volumes by Host**, as shown in Figure 8-70.

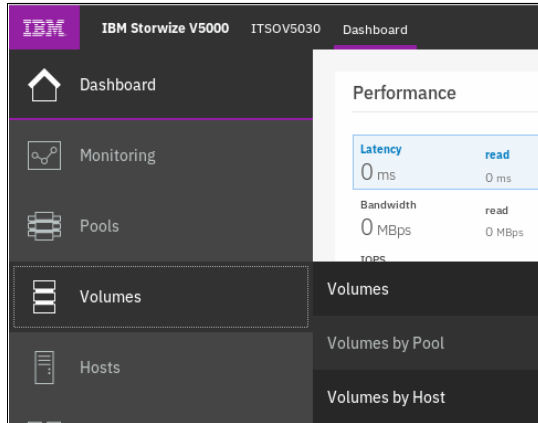


Figure 8-70 Volumes by Host option

2. The Volumes by Host window opens, as shown in Figure 8-71.

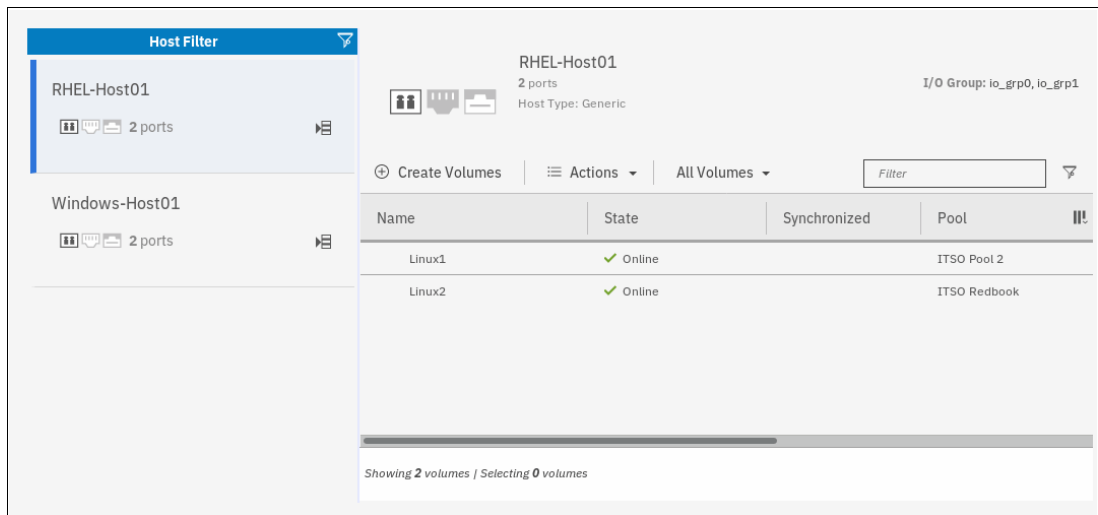


Figure 8-71 Volumes by Host window

3. The *host filter* is in the left pane of the view. Selecting a host shows its properties in the right pane, such as the host name, number of ports, host type, and the I/O group to which it has access.

The right pane, next to the host name, shows icons for Fibre Channel, iSCSI, and SAS connectivity. Depending on the type of host connectivity, the respective icon is highlighted and the other icons are disabled.

The volumes that are mapped to this host are listed in the table in the lower-right part of the window.

4. You can create a volume from this window. Click **Create Volumes** to open the same wizard as described in Chapter 6, “Volume configuration” on page 309.

Selecting a volume and opening the Actions menu or right-clicking the volume shows the same options, as described in 8.3, “Advanced volume administration” on page 407.



Advanced features for storage efficiency

IBM Spectrum Virtualize running inside the IBM Storwize V5000 offers several functions for storage optimization and efficiency.

This chapter introduces the basic concepts of those functions. It also provides a short technical overview and implementation recommendations.

For more information about planning and configuration of storage efficiency features, see the following publications. Some of these publications are dedicated to SAN Volume Controller and Storwize V7000, but are still applicable to Storwize V5000 because it uses the same principles:

- ▶ *IBM System Storage SAN Volume Controller and Storwize V7000 Best Practices and Performance Guidelines*, SG24-7521
- ▶ *Introduction and Implementation of Data Reduction Pools and Deduplication*, SG24-8430
- ▶ *IBM Real-time Compression in IBM SAN Volume Controller and IBM Storwize V7000*, REDP-4859
- ▶ *Implementing IBM Real-time Compression in SAN Volume Controller and IBM Storwize V7000*, TIPS1083
- ▶ *Implementing IBM Easy Tier with IBM Real-time Compression*, TIPS1072

This chapter includes the following topics:

- ▶ 9.1, “Easy Tier” on page 436
- ▶ 9.2, “Thin provisioned volumes” on page 448
- ▶ 9.3, “Unmap” on page 450
- ▶ 9.4, “Data Reduction Pools” on page 452
- ▶ 9.5, “Compression with standard pools” on page 460
- ▶ 9.6, “Saving estimation for compression and deduplication” on page 462

9.1 Easy Tier

IBM Spectrum Virtualize includes the IBM System Storage Easy Tier function. It enables automated subvolume data placement throughout different storage tiers. It also automatically moves extents within the same storage tier to intelligently align the system with current workload requirements, and to optimize the usage of Flash drives or flash arrays.

Many applications exhibit a significant skew in the distribution of I/O workload. A small fraction of the storage is responsible for a disproportionately large fraction of the total I/O workload of an environment.

Easy Tier acts to identify this skew and to automatically place data to take advantage of it. By moving the “hottest” data onto the fastest tier of storage, the workload on the remainder of the storage is significantly reduced. By servicing most of the application workload from the fastest storage, Easy Tier acts to accelerate application performance and increase overall server utilization. This can reduce costs in servers and application licenses.

Note: Easy Tier is a licensed function. To run Easy Tier, you must have the appropriate number of licenses that are installed on Storwize V5000.

9.1.1 Easy Tier concepts

Easy Tier is a performance optimization function that automatically migrates (or moves) extents that belong to a volume between different storage tiers, based on their I/O load. Movement of the extents is online and unnoticed from the host perspective.

As a result of extent movement, the volume no longer has all its data in one tier, but rather in two or three tiers. Each tier provides optimal performance for the extent, as shown in Figure 9-1.



Figure 9-1 Easy Tier

Easy Tier monitors the I/O activity and latency of the extents on all Easy Tier enabled storage pools. Based on the performance log, it creates an extent migration plan and *promotes* (moves) high activity or hot extents to a higher disk tier within the same storage pool. It also *demotes* extents whose activity dropped off, or cooled, by moving them from a higher disk tier MDisk back to a lower tier MDisk.

If a pool contains one type of MDisk, Easy Tier goes into balancing mode. With it, it moves extents from busy MDisks to less busy MDisks of the same tier.

Tiers of storage

The MDisks (external LUs or array type) that are presented to the Storwize V5000 are likely to have different performance attributes because of the type of disk or RAID array on which they are located.

Depending on performance, the system divides available storage into the following tiers:

- ▶ Tier 0 flash

Tier 0 flash drives are high-performance flash drives that use enterprise flash technology.

- ▶ Tier 1 flash

Tier 1 flash drives represent the Read-Intensive (RI) flash drive technology. Tier 1 flash drives are lower-cost flash drives that typically offer capacities larger than enterprise class flash, but lower performance and write endurance characteristics.

- ▶ Enterprise tier

Enterprise tier exists when the pool contains MDisks on enterprise-class hard disk drives, which are disk drives that are optimized for performance.

- ▶ Nearline tier

Nearline tier exists when the pool has MDisks on nearline-class disks drives that are optimized for capacity.

For array type MDisks, the system automatically sets its tier because it knows the capabilities of array members, physical drives, or modules. External MDisks needs manual tier assignment when they are added to a storage pool.

Note: The tier of MDisks mapped from certain types of IBM System Storage Enterprise Flash is fixed to tier0_flash, and cannot be changed.

Even though IBM Storwize V5000 can distinguish four tiers, Easy Tier manages only a three tier storage architecture. MDisk tiers are mapped to Easy Tier tiers, depending on the pool configuration.

Figure 9-2 shows the possible combinations for the pool configuration with four MDisk tiers.

	EasyTier Tier (by configuration)													
	T0	T0+T1	T0+T1+T2	T0+T1+T2+T3	T0+T2	T0+T2+T3	T0+T3	T1	T1+T2	T1+T2+T3	T1+T3	T2	T2+T3	T3
T0 (Tier0 Flash)	1	1	1	1	1	1	1							
T1 (Tier1 Flash)		2	2	2				2	2	1	2			
T2 (Tier2 HDD)			3	2	2	2			3	2		2	2	
T3 (Tier3 NearLi ne)				3		3	2			3	3		3	3

Figure 9-2 Tier combinations

The table columns that are shown in Figure 9-2 represent all the possible pool configurations; the rows report in which Easy Tier tier each MDisk tier is mapped. For example, consider a pool with all the possible tiers configured that corresponds with the T0+T1+T2+T3 configuration in the table. With this configuration, the T1 and T2 are mapped to the same tier. If there is no Tier 0 flash in a storage pool, Tier 1 flash is used as the highest performance tier.

For more information about planning and configuration considerations or best practices, see *IBM System Storage SAN Volume Controller and Storwize V7000 Best Practices and Performance Guidelines*, SG24-7521.

Easy Tier actions

The Easy Tier function continuously monitors volumes for host I/O activity. It collects performance statistics for each extent, and derives averages for a rolling 24-hour period of I/O activity. Random and sequential I/O rate and bandwidth for reads and writes are collected, and I/O response time.

Different types of analytics are used to decide whether extent data migration is required. Once per day Easy Tier analyzes the statistics to work out which data should be sent to a higher performing tier or might be sent to a tier with lower performance.

It analyzes the statistics four times per day to identify if any data must be rebalanced between managed disks in the same tier. Easy Tier check the statistics once every 5 minutes to identify if any of the managed disks is overloaded.

All of these analysis phases generate a list of migrations that should be executed. The system then spends as much time as needed running the migration plan.

The migration plan can consist of the following actions on volume extents (as shown in Figure 9-3):

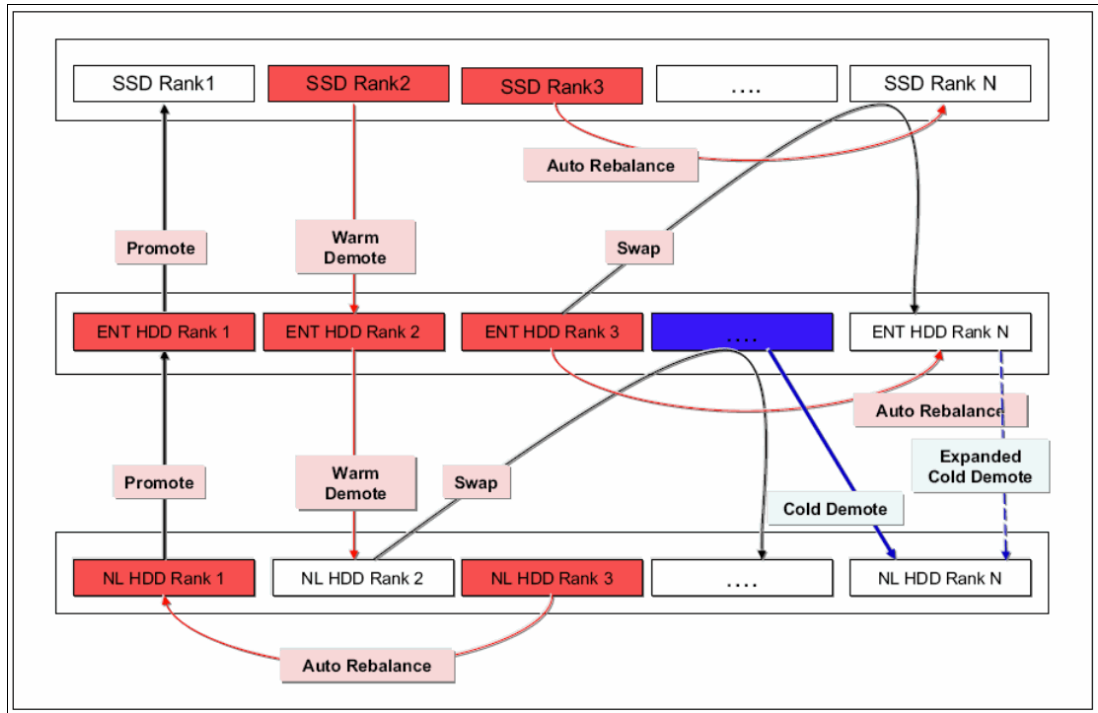


Figure 9-3 Actions on extents

- ▶ **Promote**
Moves the hotter extents to the MDisks (ranks) of the higher-performance tier with available capacity. Promote occurs within adjacent tiers.
- ▶ **Demote**
Demotes colder extents from the higher tier to the lower tier. Demote occurs within adjacent tiers.
- ▶ **RB-Move (Auto Rebalance)**
Auto-rebalance, which is also known as *intra-tier rebalancing*, is a capability of Easy Tier that automatically rebalances the workload across all MDisks of a storage tier within a managed extent pool. It also automatically populates new ranks that were added to the pool.
- ▶ **Warm demote**
Easy Tier continuously ensures that the higher performance tier does not suffer from saturation or overload conditions that might affect the overall performance in the pool. This action is triggered when bandwidth or IOPS exceeds a predefined threshold of an MDisk and causes the movement of selected extents from the higher-performance tier to the lower-performance tier to prevent MDisk overload.
- ▶ **Cold demote**
Easy Tier automatically locates and demotes inactive (or cold) extents that are on a higher performance tier to its adjacent lower-cost tier. In that way, it automatically frees extents on the higher storage tier before the extents on the lower tier become hot, and then helps the system to be more responsive to new hot data.
Cold demote occurs between tiers 2 and 3 only.

- ▶ Expanded cold demote

Demotes appropriate sequential workloads to the lowest tier to better use nearline tier bandwidth.

- ▶ Swap

A swap moves a “hot” extent from a lower performance disk tier to a higher disk tier while simultaneously moving a “cold” extent from the higher disk tier to a lower performance disk tier.

Extent migration occurs at a maximum rate of 12 GB every 5 minutes for the entire system. It prioritizes actions as listed here:

- ▶ Promote and rebalance get equal priority.
- ▶ Demote is guaranteed 1 GB every 5 minutes, and then gets whatever is left.

Note: Extent promotion and demotion occur between adjacent tiers only. In a three-tier storage pool, Easy Tier does not move extents from a flash tier directly to nearline tier or vice versa without moving to the enterprise tier first.

The Easy Tier overload protection is designed to avoid overloading any type of MDisk with too much work. To achieve this, Easy Tier must have an indication of the maximum capability of a managed disk.

For an array that is made of locally attached drives, the system can calculate the performance of the MDisk because it is pre-programmed with performance characteristics for different drives. For a SAN-attached managed disk, the system cannot calculate the performance capabilities, so the system has a number of pre-defined levels that can be configured manually for each MDisk. This is called the Easy Tier load parameter (low, medium, high, and very_high).

If you analyze the statistics and find that the system does not appear to be sending enough IOPS to your external SSD MDisk, you can increase the load parameter.

Easy Tier operating modes

Easy Tier includes the following main operating modes:

- ▶ Off

When off, no statistics are recorded and no cross-tier extent migration occurs. Also, with Easy Tier turned off, no storage pool balancing across MDisks in the same tier is performed, even in single tier pools.

- ▶ Evaluation or measurement only

When in this mode, Easy Tier collects only usage statistics for each extent in a storage pool, if it is enabled on both the volume and the pool. No extents are moved. This collection is typically done for a single-tier pool that contains only HDDs so that the benefits of adding Flash drives to the pool can be evaluated before any major hardware acquisition.

- ▶ Automatic data placement/Storage pool balancing

In this mode, usage statistics and extent information are collected. Extent migration is performed between tiers (if there is more than one pool in a tier). Also, auto-balance between MDisks of each tier is performed.

Note: The auto-balance process automatically balances data when new MDisks are added into a pool. However, it does not migrate extents from existing MDisks to achieve even extent distribution among all old and new MDisks in the storage pool. The Easy Tier migration plan is based on performance, and not on the capacity of the underlying MDisks or on the number of extents on them.

Implementation considerations

Consider the following implementation and operational rules when you use the IBM System Storage Easy Tier function on the Storwize V5000:

- ▶ Volumes that are added to storage pools use extents from the “middle” tier of three-tier model, if available. Easy Tier then collects usage statistics to determine which extents to move to “faster” T0 or “slower” T2 tiers. If there are no free extents in T1, extents from the other tiers are used.
- ▶ When an MDisk with allocated extents is deleted from a storage pool, extents in use are migrated to MDisks in the same tier as the MDisk that is being removed, if possible. If insufficient extents exist in that tier, extents from the other tier are used.
- ▶ Easy Tier monitors extent I/O activity of each copy of a mirrored volume. Easy Tier works with each copy independently of the other copy.

Note: Volume mirroring can have different workload characteristics on each copy of the data because reads are normally directed to the primary copy and writes occur to both copies. Therefore, the number of extents that Easy Tier migrates between the tiers might differ for each copy.

- ▶ For compressed volumes on standard pools, only reads are analyzed by Easy Tier.
- ▶ Easy Tier automatic data placement is not supported on image mode or sequential volumes. However, it supports evaluation mode for such volumes. I/O monitoring is supported and statistics are accumulated.
- ▶ When a volume is migrated out of a storage pool that is managed with Easy Tier, Easy Tier automatic data placement mode is no longer active on that volume. Automatic data placement is also turned off while a volume is being migrated, even when it is between pools that both have Easy Tier automatic data placement enabled. Automatic data placement for the volume is reenabled when the migration is complete.

When the system migrates a volume from one storage pool to another, it attempts to migrate each extent to an extent in the new storage pool from the same tier as the original extent, if possible.

- ▶ When Easy Tier automatic data placement is enabled for a volume, you cannot use the `svctask migrateexts` CLI command on that volume.

9.1.2 Implementing and tuning Easy Tier

The Easy Tier function is enabled by default if the license is set up. It starts monitoring I/O activity immediately after storage pool and volumes are created, and starts extent migration when the necessary I/O statistics are collected.

A few parameters can be adjusted. Also, Easy Tier can be switched off on selected volumes in storage pools.

MDisk settings

The tier for internal (array) MDisks is detected automatically and depends on the type of drives that are its members. No adjustments are needed.

For an external MDisk, the tier is assigned when it is added to a storage pool. To assign the MDisk, navigate to **Pools** → **External Storage**, select the MDisk (or MDisks) to add, and click **Assign**.

Note: The tier of MDisks mapped from certain types of IBM System Storage Enterprise Flash is fixed to tier0_flash, and cannot be changed.

You can choose the target storage pool and storage tier that is assigned, as shown in Figure 9-4.

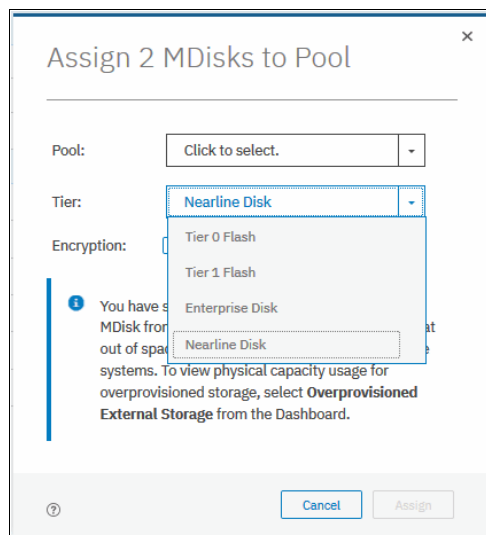


Figure 9-4 Choosing tier when assigning MDisks

To change the storage tier for an MDisk that is assigned, click **Pools** → **External Storage**. Then, right-click one or more selected MDisks and choose **Modify Tier**, as shown in Figure 9-5.

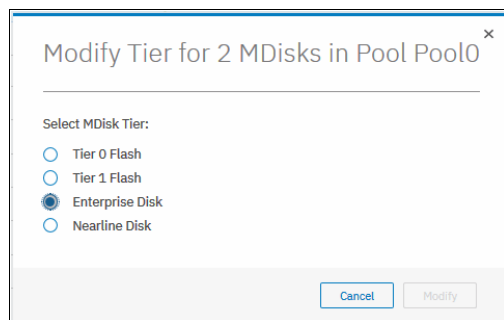


Figure 9-5 Changing MDisk tier

Note: Assigning a tier that does not match to a physical back end storage type to an external MDisk is not supported by IBM and might lead to unpredictable consequences.

To determine which tier is assigned to an MDisk, click **Pools** → **External Storage**. Then, click **Actions** → **Customize columns** and select **Tier**. This includes the current tier setting into a list of MDisk parameters that are shown in the **External Storage** pane. You can also find this information in MDisk properties. To show this information, right-click MDisk, select **Properties**, and expand “View more details” section, as shown in Figure 9-6.

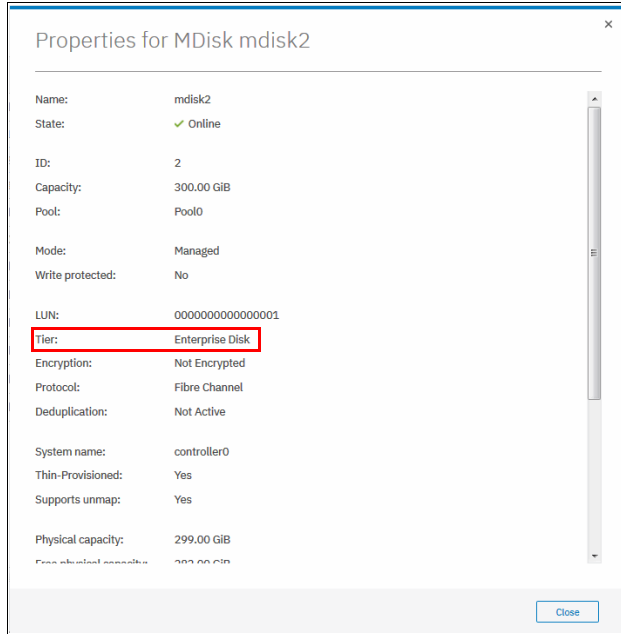


Figure 9-6 MDisk properties

To list MDisk parameters with the CLI, use the `lsmdisk` command. The current tier for each MDisk is shown. To change the external MDisk tier, use `chmdisk` with the `-tier` parameter as shown in Example 9-1.

Example 9-1 Listing and changing tiers for MDisks (partially shown)

```
IBM_Storwize:ITS0V5K:superuser>lsmdisk
id name status mode mdisk_grp_id ... tier encrypt
1 mdisk1 online unmanaged ... tier0_flash no
2 mdisk2 online managed 0 ... tier_enterprise no
3 mdisk3 online managed 0 ... tier_enterprise no
<...>
IBM_Storwize:ITS0V5K:superuser>chmdisk -tier tier1_flash mdisk2
IBM_Storwize:ITS0V5K:superuser>
```

For an external MDisk, the system cannot calculate its exact performance capabilities, so it has a number of predefined levels. In rare cases, statistics analysis might show that Easy Tier is overusing or under-utilizing an MDisk. If so, levels can be adjusted, which can be done only by using the CLI. Use `chmdisk` with `-easytierload` parameter. To reset Easy Tier load to system-default for chosen MDisk, use `-easytier default`, as shown in Figure 9-2 on page 444.

Note: Adjust Easy Tier load settings only if instructed to do so by IBM Technical Support or your solution architect.

To list the current Easy Tier load setting of an MDisk, use `lsmdisk` with MDisk name or ID as a parameter.

Example 9-2 Changing Easy Tier load

```
IBM_Storwize:ITS0V5K:superuser>chmdisk -easytierload default mdisk2
IBM_Storwize:ITS0V5K:superuser>
IBM_Storwize:ITS0V5K:superuser>lsmdisk mdisk2 | grep tier
tier tier_enterprise
easy_tier_load high
IBM_Storwize:ITS0V5K:superuser>
```

Storage pool settings

When a storage pool (standard pool or Data Reduction Pool) is created, Easy Tier is switched on by default. The system automatically enables Easy Tier functions when the storage pool contains an MDisk from more than one tier. It also enables automatic rebalancing when the storage pool contains an MDisk from only one tier.

You can disable Easy Tier or switch it to measure-only mode when creating a pool or any moment later. This is not possible with the GUI; only with the system CLI.

To check current Easy Tier function state on a pool, click **Pools** → **Pools**, right-click the selected pool, choose **Properties** and expand the **View more details** section, as shown in Figure 9-7 on page 445.

Easy Tier status can be:

- ▶ **active**
Indicates that a pool is being managed by Easy Tier, and extent migrations between tiers can be performed. Performance-based pool balancing is also enabled.
This state is expected for a pool with two or more tiers of storage.
- ▶ **balanced**
Indicates that a pool is being managed by Easy Tier to provide performance-based pool balancing.
This state is the expected for a pool with a single tier of storage.
- ▶ **inactive**
Indicates that Easy Tier is inactive (switched off).
- ▶ **measured**
Shows that Easy Tier statistics are being collected, but no extent movement can be performed.

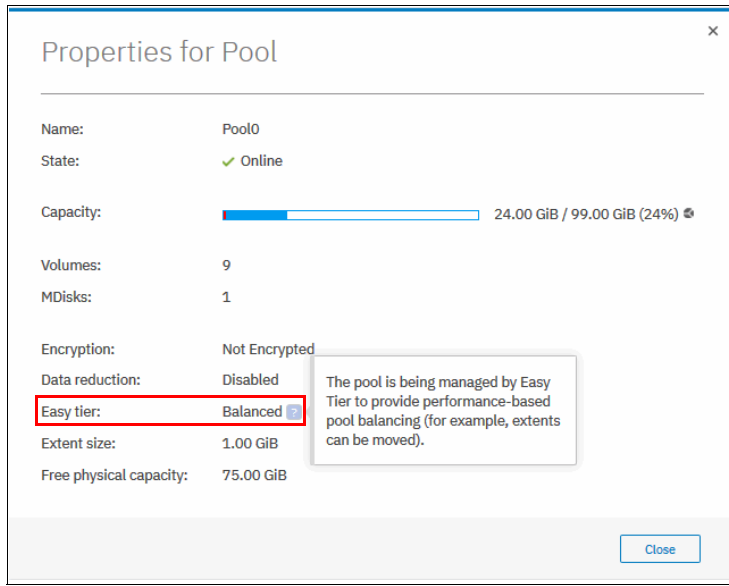


Figure 9-7 Pool properties

To find the status of the Easy Tier function on the pools with the CLI, use the `lsmdiskgrp` command without any parameters. To switch Easy Tier off or back on, use the `chmdiskgrp`, as shown in Example 9-3. By running `lsmdiskgrp` with pool name/ID as a parameter, you can also determine how much storage of each tier is available within the pool.

Example 9-3 Listing and changing Easy Tier status on pools

```
IBM_Storwize:ITS0V5K:superuser>lsmdiskgrp
id name status mdisk_count ... easy_tier easy_tier_status
0 Pool0 online 1 ... auto balanced
2 Pool1 online 3 ... auto balanced
IBM_Storwize:ITS0V5K:superuser>chmdiskgrp -easytier measure Pool0
IBM_Storwize:ITS0V5K:superuser>chmdiskgrp -easytier auto Pool0
IBM_Storwize:ITS0V5K:superuser>
```

Volume settings

By default, each striped-type volume allows Easy Tier to manage its extents. If you must fix the volume extent location (for example, to prevent extent demotes and to keep the volume in the higher-performing tier), you can turn off Easy Tier management for a particular volume copy.

Note: Thin-provisioned and compressed volumes in a data reduction pool cannot have Easy Tier switched off. It is possible to switch off Easy Tier only at a pool level.

This can be done by using only the CLI. Use the `lsvdisk` command to check and the `chvdisk` command to modify Easy Tier function status on a volume copy, as shown in Example 9-4.

Example 9-4 Checking and modifying Easy Tier settings on a volume

```
IBM_Storwize:ITS0-V7k:superuser>lsvdisk vdisk0 |grep easy_tier
easy_tier on
easy_tier_status balanced
IBM_Storwize:ITS0-V7k:superuser>chvdisk -easytier off vdisk0
```

IBM_Storwize:ITS0V5K:superuser>

System-wide settings

The system-wide setting Easy Tier acceleration is disabled by default. Turning it on makes Easy Tier move extents up to four times faster than the default setting. In accelerate mode, Easy Tier can move up to 48 GiB per 5 minutes; whereas in normal mode it moves up to 12 GiB. The following are the two most probable use cases for acceleration:

- ▶ When adding capacity to the pool, accelerating Easy Tier can quickly spread volumes onto the new MDisks.
- ▶ Migrating the volumes between the storage pools when the target storage pool has more tiers than the source storage pool, so Easy Tier can quickly promote or demote extents in the target pool.

Note: Enabling Easy Tier acceleration is advised only during periods of low system activity only after migrations or storage reconfiguration occurs. It is recommended to keep Easy Tier acceleration mode off during normal system operation.

This setting can be changed online, but only by using the CLI. To turn on or off Easy Tier acceleration mode, use the `chsystem` command and the `lssystem` command to check its state, as shown in Example 9-5.

Example 9-5 The chsystem command

```
IBM_Storwize:ITS0V5K:superuser>lssystem |grep easy_tier
easy_tier_acceleration off
IBM_Storwize:ITS0V5K:superuser>chsystem -easytieracceleration on
IBM_Storwize:ITS0V5K:superuser>
```

9.1.3 Monitoring Easy Tier activity

When Easy Tier is active, it constantly monitors and records I/O activity and collects extent heat data. Heat data files are produced approximately once a day and summarizes the activity per volume since the prior heat data file was produced.

The IBM Storage Tier Advisor Tool is a Windows console application that can analyze heat data files that are produced by Easy Tier. It also can produce a graphical display of the amount of “hot” data per volume and provide predictions about how more solid-state drive (T0) capacity, Enterprise Drive (T1), and Nearline Drive (T2) can benefit the system and storage pool performance.

For more information about the IBM Storage Tier Advisor Tool, see this [IBM Support web page](#).

You can download the IBM Storage Tier Advisor Tool and install it on your Windows computer. The tool is included as an ISO file, which must be extracted to a temporary location.

The tool installer is in `temporary_location\IMAGES\STAT\Disk1\InstData\NoVM\`. Storage Tier Advisor Tool by default is installed in the `C:\Program Files\IBM\STAT\` directory.

On IBM Storwize V5000, the heat data files are found in the `/dumps/easytier` directory on the configuration node, and are named `dpa_heat.node_panel_name.time_stamp.data`. Any heat data file is erased when it older than 7 days.

Heat files must be offloaded and IBM Storage Tier Advisor Tool started from a Windows command prompt console with the file specified as a parameter, as shown in Example 9-6.

Example 9-6 Running STAT in Windows command prompt

```
C:\Program Files (x86)\IBM\STAT>stat dpa_heat.7830038-1.181107.110150.data
```

The IBM Storage Tier Advisor Tool creates a set of HTML and .csv files that can be used for Easy Tier analysis.

To download a heat data file, open **Settings** → **Support** → **Support Package** → **Download Support Package** → **Download Existing Package**, as shown in Figure 9-8.

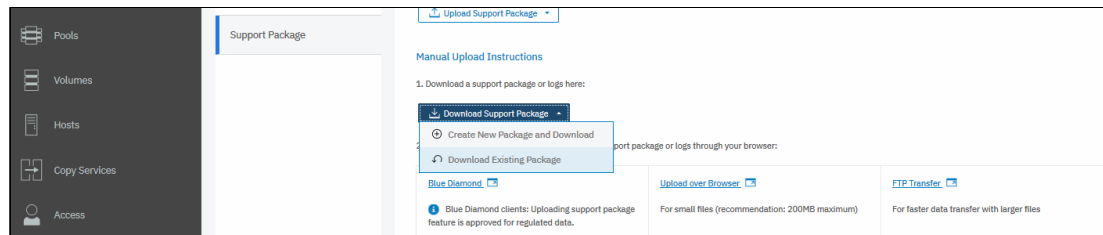


Figure 9-8 Download Easy Tier heat file: Download Support Package

A download window opens that shows all files in /dumps and its subfolders on a current configuration node. You can filter the list by using the `easytier` keyword, select the `dpa_heat` files that will be analyzed, and click **Download**, as shown in Figure 9-9. Save them in a convenient location (for example, to a subfolder that holds the IBM Storage Tier Advisor Tool executable file).

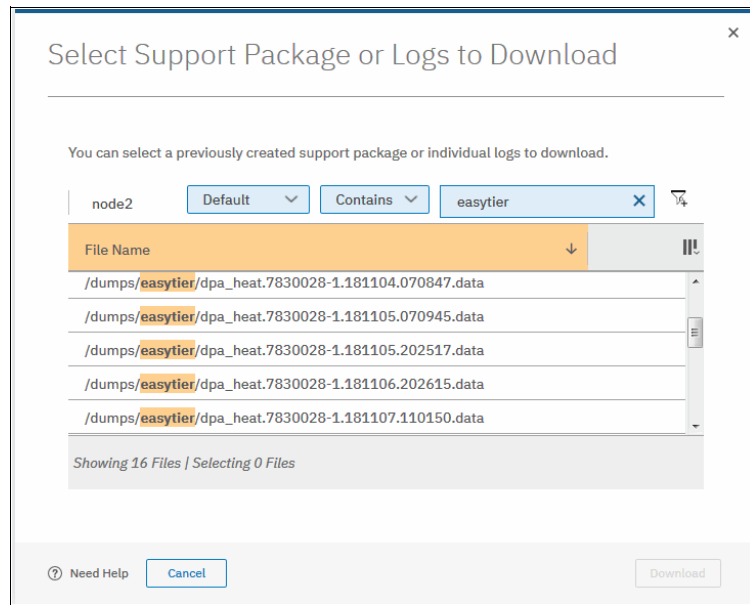


Figure 9-9 Downloading Easy Tier heat data file: dpa_heat files

You can also specify the output directory if you want. IBM Storage Tier Advisor Tool creates a set of HTML files, and the user can then open the `index.html` file in a browser to view the results. Also, the following .csv files are created and placed in the `Data_files` directory:

- ▶ `<panel_name>_data_movement.csv`
- ▶ `<panel_name>_skew_curve.csv`

► <panel_name>_workload_ctg.csv

These files can be used as input data for other utilities.

For more information about how to interpret IBM Storage Tier Advisor Tool tool output and CSV files analysis, see *IBM System Storage SAN Volume Controller and Storwize V7000 Best Practices and Performance Guidelines*, SG24-7521.

9.2 Thin provisioned volumes

In a shared storage environment, thin provisioning is a method for optimizing the use of available storage. It relies on the allocation of blocks of data on demand versus the traditional method of allocating all of the blocks up front. This method eliminates almost all white space, which helps avoid the poor usage rates (often as low as 10%) that occur in the traditional storage allocation method where large pools of storage capacity are allocated to individual servers but remain unused (not written to).

Thin provisioning presents more storage space to the hosts or servers that are connected to the storage system than is available on the storage system.

9.2.1 Concepts

Volumes can be configured as thin-provisioned or fully allocated. Both can be configured in standard pools and data reduction pools.

In IBM Storwize V5000, each volume has virtual capacity and real capacity parameters. *Virtual capacity* is the volume storage capacity that is available to a host Operating System (OS) and is used by it to create a filesystem. *Real capacity* is the storage capacity that is allocated to a volume from a pool. It shows the amount of space that is used on a physical storage.

Fully allocated volumes are created with the same amount of real capacity and virtual capacity. This type uses no storage efficiency features.

A thin-provisioned volume presents a different capacity to mapped hosts than the capacity that the volume uses in the storage pool. Therefore, real and virtual capacities might not be equal.

The virtual capacity of a thin-provisioned volume is typically significantly larger than its real capacity. As more information is written by the host to the volume, more of the real capacity is used. The system identifies read operations to unwritten parts of the virtual capacity and returns zeros to the server without using any real capacity.

The autoexpand feature prevents a thin-provisioned volume from using up its capacity and going offline. As a thin-provisioned volume uses capacity, the autoexpand feature maintains a fixed amount of unused real capacity, which is called the *contingency capacity*. For thin-provisioned volumes in standard pools, the autoexpand feature can be turned on and off. For thin-provisioned volumes in data reduction pools, the autoexpand feature is always enabled.

A thin-provisioned volume can be converted non-disruptively to a fully allocated volume, or vice versa, by using the volume mirroring function. For example, you can add a thin-provisioned copy to a fully allocated primary volume and then remove the fully allocated copy from the volume after they are synchronized.

The fully allocated to thin-provisioned migration procedure uses a zero-detection algorithm so that grains that contain all zeros do not cause any real capacity to be used. Usually, if the Storwize V5000 is to detect zeros on the volume, you must use software on the host side to write zeros to all unused space on the disk or file system.

9.2.2 Implementation

For more information about creating thin-provisioned volumes, see Chapter 6, “Volume configuration” on page 309.

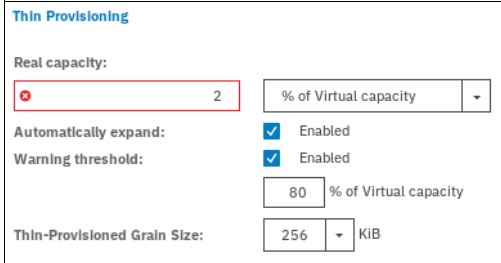
In a standard pool, the system uses the real capacity to store data that is written to the volume and metadata that describes the thin-provisioned configuration of the volume. Metadata uses volume real capacity and usually needs less than 0.5% of virtual capacity to store its data.

This means that if your host used 100% of virtual capacity, some extra space is required on your storage pool to store thin provisioning metadata. In a worst case scenario, the real size of a thin provisioned volume can be 100.5% of its virtual capacity.

In a data reduction pool, metadata for a thin provisioned volume is stored separately from user data, and does not count in the volumes real capacity.

Volume parameters

When creating a thin provisioned volume in Custom mode, some of its parameters can be modified, as shown in Figure 9-10.



The screenshot shows a configuration window titled "Thin Provisioning". It contains several settings:

- Real capacity:** A text input field containing the value "2", with a red box around it. To its right is a dropdown menu labeled "% of Virtual capacity".
- Automatically expand:** A checkbox that is checked, with the label "Enabled" next to it.
- Warning threshold:** A checkbox that is checked, with the label "Enabled" next to it. Below this is a text input field containing "80" and a dropdown menu labeled "% of Virtual capacity".
- Thin-Provisioned Grain Size:** A text input field containing "256" and a dropdown menu labeled "KiB".

Figure 9-10 Thin VDisk parameters

In a DRP pool, thin provisioned volume fine-tuning is not required. Real capacity (rsize) value is ignored, and Grain Size is fixed to 8 KB.

When a thin-provisioned volume is created in a standard pool, Real capacity (rsize) defines both initial volume real capacity and the amount of contingency capacity, which is used by autoexpand.

Write I/Os to the grains of the thin volume in a standard pool that were not previously written to causes grains of the real capacity to be used to store metadata and user data. Write I/Os to the grains (that were previously written to) updates the grain where data was written. The grain is defined when the volume is created, and can be 32 KiB, 64 KiB, 128 KiB, or 256 KiB.

Smaller granularities can save more space, but they have larger metadata directories. When you use thin-provisioning with FlashCopy, specify the same grain size for the thin-provisioned volume and FlashCopy.

Host considerations

Do not use defragmentation applications on thin-provisioned volumes. The defragmentation process can write data to different areas of a volume, which can cause a thin-provisioned volume to grow up to its virtual size.

9.3 Unmap

Spectrum Virtualize systems running V8.1.0 and later support the SCSI Unmap command. This enables hosts to notify the storage controller of capacity that is no longer required, which might improve capacity savings.

9.3.1 SCSI unmap command

Unmap is a set of SCSI primitives that allow hosts to indicate to a SCSI target that space allocated to a range of blocks on a target storage volume is no longer required. This command allows the storage controller to take measures and optimize the system so that the space can be reused for other purposes. The most common use case, for example, is a host application, such as VMware freeing storage within a file system. The storage controller can then optimize the space, such as reorganizing the data on the volume so that space is better used.

When a host allocates storage, the data is placed in a volume. To free the allocated space back to the storage pools, human intervention is needed on the storage controller. The SCSI Unmap feature is used to allow host operating systems to unprovision storage on the storage controller, which means that the resources can automatically be freed up in the storage pools and used for other purposes.

A SCSI unmappable volume is a volume that can have storage unprovision and space reclamation being triggered by the host operating system. IBM Storwize V5000 can pass the SCSI unmap command through to back-end storage controllers that support the function.

9.3.2 Back-end SCSI Unmap

The system can generate and send SCSI Unmap commands to specific backend storage controllers.

This occurs when volumes are deleted, extents are migrated, or an Unmap command is received from the host. SCSI Unmap commands are sent only if Storwize V5030 can detect that the back end controller supports it.

This helps prevent a thin-provisioning storage controller from running out of free capacity for write I/O requests. This means when you are using supported thin provisioned back-end storage, SCSI Unmap should normally be left enabled.

This feature is turned on by default, and it is recommended to keep back-end unmap enabled.

To verify that sending unmap commands to back end is enabled, use the CLI command `lssystem`, as shown in Example 9-7.

Example 9-7 Verifying back end unmap support status

```
IBM_Storwize:ITS0V5K:superuser>lssystem | grep backend_unmap
backend_unmap on
```

9.3.3 Host SCSI Unmap

The Spectrum Virtualize system can advertise support for SCSI Unmap to hosts. With it, some host types (for example, Windows, Linux, or VMware) will then change their behavior when creating a new filesystem on a volume, issuing SCSI Unmap commands to the whole capacity of the volume. This causes the system to overwrite the whole capacity with zero-data, and the format completes when all of these writes complete. Some host types run a background process (for example `fstrim` on Linux), which periodically issues SCSI Unmap commands for regions of a filesystem that are no longer required.

Host Unmap commands can increase the free capacity reported by the Data Reduction Pool, when received by thin-provisioned or compressed volumes. This does not apply to standard storage pools. Also, IBM Storwize V5000 returns end SCSI Unmap commands to controllers that support them if host unmaps for corresponding blocks are received.

Host SCSI Unmap commands drive more I/O workload to back-end storage. In some circumstances (for example, volumes that are on a heavily loaded nearline SAS array), this can cause an increase in response times on volumes using the same storage. Also, host formatting time is likely to increase, compared to a system that does not advertise support the SCSI Unmap command.

If you are using Data Reduction Pools, it is recommended to turn on SCSI Unmap support.

If only standard pools are being used, you might consider keeping host Unmap support switched off because you do not see any improvement in the system's behavior.

To check and modify current setting for host SCSI Unmap support, use the `lssystem` and `chsystem` CLI commands, as shown in Example 9-8.

Example 9-8 Turning host unmap support on

```
IBM_Storwize:ITS0V5K:superuser>lssystem | grep host_unmap
host_unmap off
IBM_Storwize:ITS0V5K:superuser>chsystem -hostunmap on
IBM_Storwize:ITS0V5K:superuser>
```

Note: You can switch host Unmap support on and off non-disruptively on the system side. However, hosts might need to rediscover storage, or (in the worst case) be restarted.

9.3.4 Offload IO throttle

Throttles are a mechanism to control the amount of resources that are used when the system is processing I/Os on supported objects. If a throttle limit is defined, the system processes the I/O for that object, or delays the processing of the I/O to free resources for more critical I/O operations.

SCSI offload commands, such as Unmap and Xcopy, are used by hosts to format new filesystems, or copy volumes without the host needing to read and then write data.

Some host types might request large amounts of I/O on the storage system by issuing Write Same/Unmap commands. If the underlying storage cannot handle the amount of I/O that is generated, performance of volumes can be affected.

Spectrum Virtualize offload throttling limits the concurrent I/O that can be generated by such commands, which can prevent the MDisk overloading. This limits the rate at which host features, such as VMware VMotion, can copy data.

Note: For systems that are managing any nearline storage, it might be recommended to set the offload throttle to 100 MBps.

To implement offload throttle, you can use the `mkthrottle` command with `-type offload` parameter or GUI, as shown in Figure 9-11.

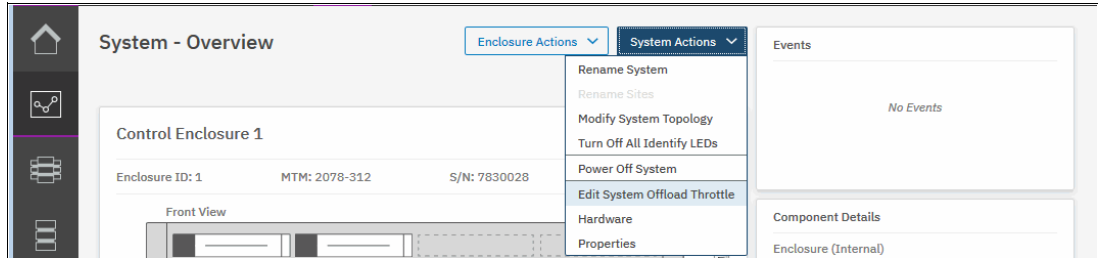


Figure 9-11 Setting offload throttle

9.4 Data Reduction Pools

Data Reduction Pools (DRP) increase infrastructure capacity usage by using new efficiency functions and reducing storage costs. The pools enable you to automatically de-allocate and reclaim capacity of thin-provisioned volumes that contain deleted data and for the first time, enable this reclaimed capacity to be reused by other volumes. Data reduction pools support volume compression. Also, they support deduplication that can be configured with thin-provisioned and compressed volumes.

Note: This book provides only an overview of DRP aspects. For more information, see *Introduction and Implementation of Data Reduction Pools and Deduplication*, SG24-8430.

9.4.1 Introduction to DRP

At its core, a data reduction pool uses a Log Structured Array (LSA) to allocate capacity. 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 allows you to allocate data sequentially and provide a directory that provides a lookup to match LBA with physical address within the array. Therefore, the volume 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.

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 requests because of 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, directory update, and data write.

DRP technology allows you to create five types of volumes:

- ▶ Fully allocated

This type provides no storage efficiency.

- ▶ Thin-provisioned
This type provides some storage efficiency but no compression or deduplication. Volume capacity is allocated on demand as storage is first written to.
- ▶ Thin and Compressed
In addition to on-demand space allocation, data is compressed before being written to storage.
- ▶ Thin and Deduplicated
In addition to on-demand space allocation, duplicates of data blocks are detected and are replaced with references to the first copy.
- ▶ Thin, Compressed, and Deduplicated
This type provides maximum storage efficiency and capacity savings by combining both methods.

Note: Consider the following points:

- ▶ V8.1.2 or higher is required for Data Reduction Pools.
- ▶ Nodes must have 32 GB of memory each to support compression and deduplication.
- ▶ Compression and deduplication is not supported on Storwize V5010, V5020, and Storwize V5030 with 16 GB onboard memory.

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

9.4.2 Data Reduction Pools benefits

DRPs are a new type of storage pool that implement 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 a 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, lower power consumption, is cheaper to cool, and has higher density. However, the cost of Flash storage is still higher than disk storage.

With 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 because of deduplication. The deduplication process identifies unique data patterns and stores the 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 might 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 because of compression. While 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 is trying 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 de-duplicated and compressed, data is de-duplicated first, and then compressed. Therefore, deduplication references are created on the compressed data stored on the physical domain.

DRPs offer a new implementation of data compression that is integrated into the I/O stack. 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, which results in more performance gains.

The third benefit of DRPs comes in the form of performance improvements because of Easy Tier. The metadata of DRPs does not fit in RAM; therefore, it is stored on disk on metadata volumes that are separate from data volumes.

The metadata volumes of DRPs are small and frequently accessed. They are good candidates for promotion through Easy Tier. In contrast, data volumes are large; however, because the metadata is stored separately, Easy Tier can accurately identify frequently used data. Performance gains are expected because Easy Tier promotes metadata to the fastest available storage tier.

DRPs support end-to-end SCSI 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), which results in the freeing of all the capacity allocated within that unmap. Similarly, deleting a volume at the DRP level frees all the capacity back to the pool.

When a DRP 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 DRP to maximize space within the pool.

9.4.3 Implementing DRP with Compression and Deduplication

The implementation process for DRP pools is similar to standard pools, but has its own specifics.

Creating pools and volumes

To create a DRP, select the **Data Reduction Enable** option in the **Create Pool** window, which is available by clicking **Pools** → **Pools**. For more information about how to create storage pool and populate it with MDisks, see Chapter 4, “Storage pools” on page 159.

There is a maximum number of four data reduction pools in a system. When this limit is reached, creating any other DRPs is not possible.

Note: The best practice for Storwize V5030 is to create one DRP only.

A DRP makes use of metadata. Even when there are no volumes in the pool, some of the space of the pool is used to store the metadata. Regardless of the type of volumes the pool contains, metadata is always stored separate from customer data.

Metadata capacity depends on the total capacity of a storage pool and on a pool extent size. This should be taken into account when planning capacity.

Note: If your DRP has a total capacity below 50 TB, you might need to decrease the extent size from the default for DRP (which is 4 GB) to 1 GB for optimal space savings.

The main purpose of DRPs is to be a fast and efficient container for volumes with data reduction capability. DRPs can also contain fully allocated volumes.

To create a volume on a DRP, browse to **Volumes** → **Volumes**, and click **Create Volumes**.

Figure 9-12 shows the Create Volumes window. Under the Capacity Savings menu, you can select None, Thin Provisioned and Compressed. If Compressed or Thin Provisioned are selected, the Deduplicated option also becomes available and can be selected.

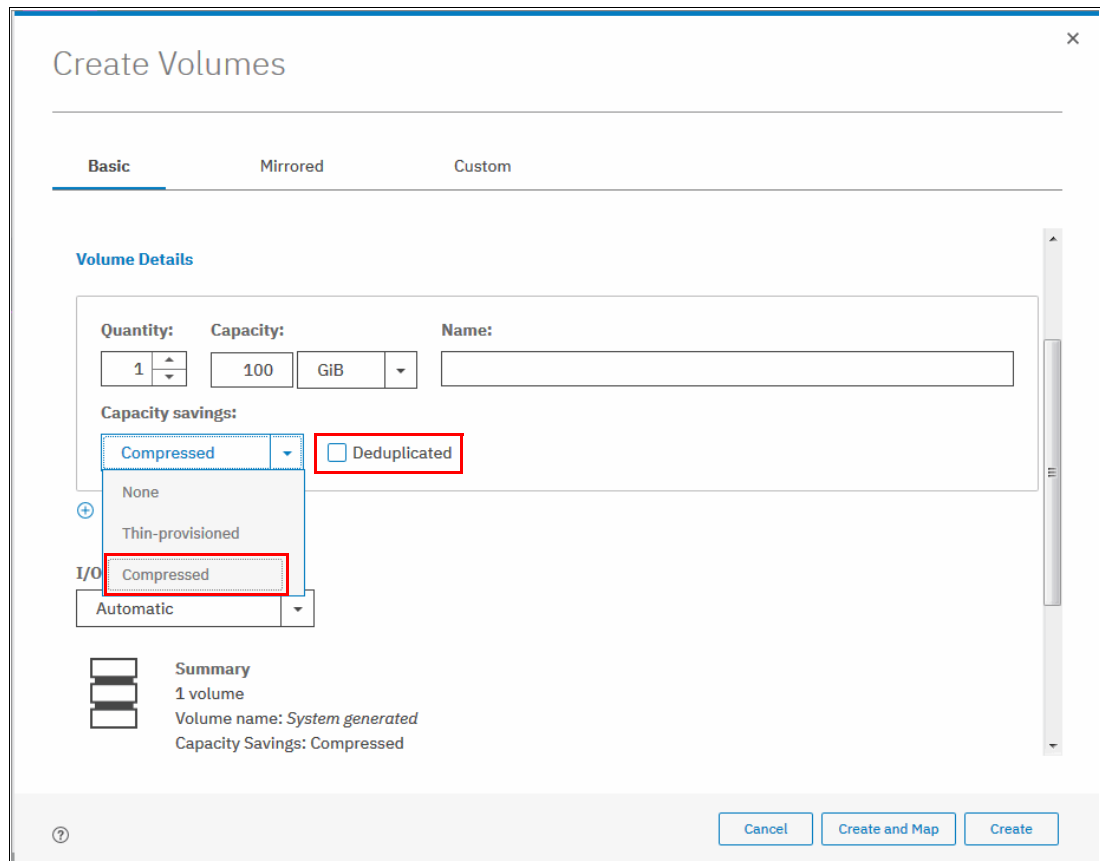


Figure 9-12 Create compressed volume

Capacity monitoring

Figure 9-13 on page 456 shows the Volumes window with a list of volumes, which are created in the DRP pool. Only Capacity (which is virtual capacity, available to a host) is shown. 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 DRPs, be aware that there is no volume-copy level reporting on used capacity.

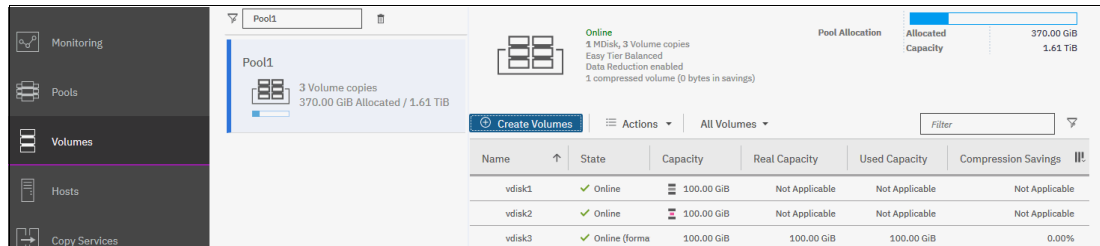


Figure 9-13 Volumes in DRP pool

The only parameter (except virtual capacity) that is available on a volume level to monitor thin, compressed, or deduplicated volume is `used_capacity_before_reduction`. It indicates the total amount of data written to a thin-provisioned or compressed volume copy in a data reduction storage pool before data reduction occurs. This field is blank for fully allocated volume copies and volume copies not in a DRP. Capacity that is assigned to the volume in the tier is also not reported.

To find this value, use the `lsvdisk` command with volume name or ID as a parameter, as shown in Example 9-9. It shows a non-compressed thin-provisioned volume with virtual size 100 GiB on a DRP, which was mounted to a host, and had a 72 GiB file created on it.

Example 9-9 Volume in a DRP capacity monitoring

```
IBM_Storwize:ITS0V5K:superuser>lsvdisk vdisk1
id 15
name vdisk1
<...>
capacity 100.00GB
<...>
used_capacity
real_capacity
free_capacity

tier tier0_flash
tier_capacity 0.00MB
tier tier1_flash
tier_capacity 0.00MB
tier tier_enterprise
tier_capacity 0.00MB
tier tier_nearline
tier_capacity 0.00MB
compressed_copy no
uncompressed_used_capacity
deduplicated_copy no
used_capacity_before_reduction 71.09GB
```

Capacity and space saving reporting is available in the storage pool views and Out of space warning thresholds are configured at the storage pool level. You can check savings with the GUI Pools view, GUI Dashboard, or by using the CLI command `lsmdiskgrp` with DRP ID or name as a parameter, as shown in Example 9-10.

Example 9-10 Pool savings monitoring

```
IBM_Storwize:ITS0V5K:superuser>lsmdiskgrp 3
id 3
name Pool1
```

```

<...>
capacity 8.64TB
free_capacity 7.53TB
virtual_capacity 1.79TB
used_capacity 882.38GB
real_capacity 883.95GB
overallocation 20
<...>
tier tier0_flash
tier_mdisk_count 0
tier_capacity 0.00MB
tier_free_capacity 0.00MB
tier tier1_flash
tier_mdisk_count 1
tier_capacity 8.64TB
tier_free_capacity 7.78TB
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
<...>
data_reduction yes
used_capacity_before_reduction 1.30TB
used_capacity_after_reduction 793.38GB
overhead_capacity 89.00GB
deduplication_capacity_saving 10.80GB
reclaimable_capacity 0.00MB
physical_capacity 8.64TB
physical_free_capacity 7.78TB
shared_resources no

```

The output reports real capacity used on each of the storage tiers. Deduplication savings are also shown. That compression-related values show 0MB because they belong to RtC (standard pool) compression.

For more information about every reported value, see IBM Knowledge Center for the [lsmdiskgrp](#) command.

Migrating to and from DRP

Although data can be migrated regardless of the nature of the new volumes, the type of these volumes determines the migration strategy used.

The reason for the following strategies is compression. Compressed volumes in a standard pool cannot coexist with compressed or deduplicated volumes in a DRP:

- Migration strategy for non-compressed source volumes on a standard pool to a DRP

If you need to migrate fully allocated or thin-provisioned volumes from a standard pool to fully allocated or thin-provisioned (with or without compression and deduplication) volume in a DRP pool, you can do this by creating a second volume copy.

To create a second copy, right-click the source volume and choose **Add Volume Copy**, as shown in Figure 9-14. Choose target DRP for the second copy and **Capacity savings** type: None, Thin-provisioned, or Compressed.

You can choose **Compressed** and select **Deduplicated** only if there are no RtC compressed volumes in standard pools of your I/O group.

After you click **Add**, synchronization starts. The time synchronization takes to complete depends on the size of the volume and system performance. You can increase the synchronization rate by right-clicking the volume and selecting **Modify Mirror Sync Rate**.

When copies are synchronized, Yes is displayed for both copies in the Synchronized column in the Volumes pane.

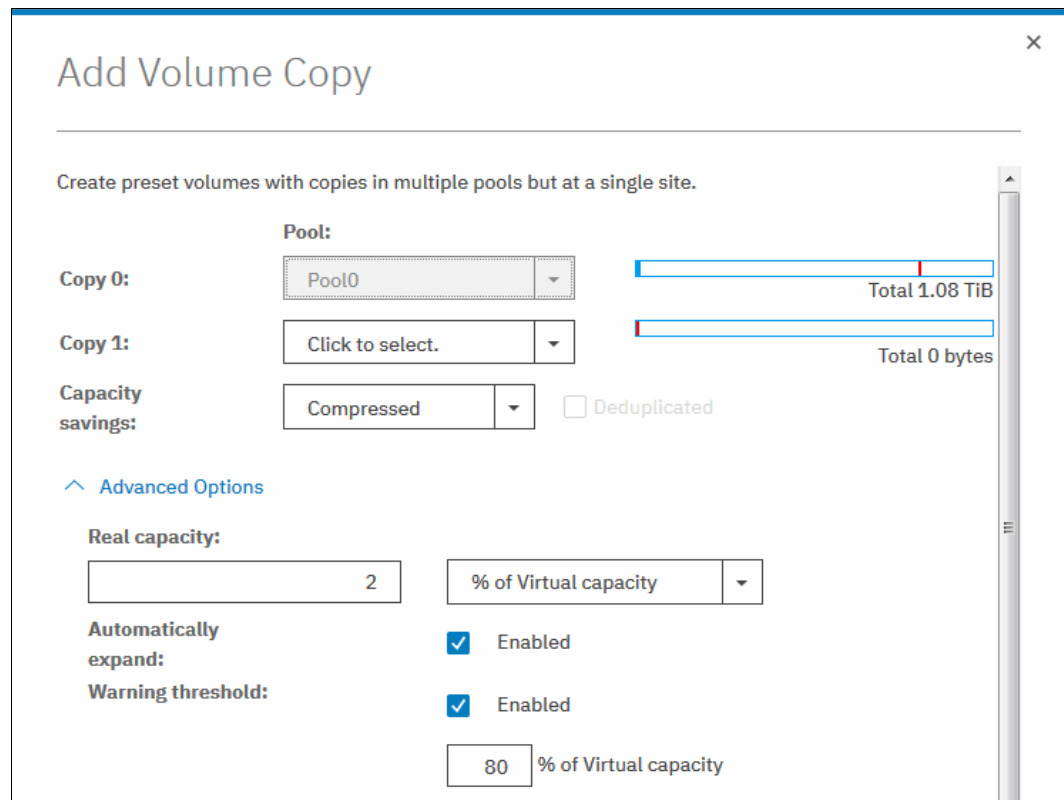


Figure 9-14 Add volume copy window

You can also track synchronization process with Running tasks pane, as shown in Figure 9-15. After it reaches 100% and copies are in-sync, you can complete migration by deleting the source copy in a standard pool.

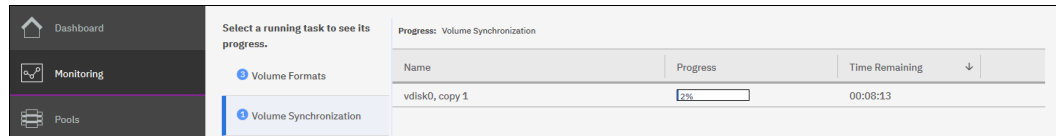


Figure 9-15 Synchronization progress

- ▶ Migration strategy for volumes on DRP to non-compressed volumes in a standard pool

This process is the same as the migration strategy for non-compressed source volumes on a standard pool to a DRP. The difference is that the target pool for a second copy is standard.
- ▶ Migration strategy for a compressed volume on a standard pool to thin and deduplicated or compressed and deduplicated volume on a DRP

Deduplicated or compressed volumes on DRP cannot coexist in the same I/O group with compressed volumes in standard pools. To migrate a compressed volume from a standard pool to a thin-provisioned deduplicated or compressed deduplicated volume, a two-step procedure is required:

- a. For a source compressed volume, create a second uncompressed, non-deduplicated copy in a DRP, wait for synchronization to complete, and delete the source copy.

Complete the step for all compressed volumes in all standard pools in an I/O group, and verify that no other compressed volumes are left in the standard pools.
- b. For each non-compressed volume in a DRP, create a second copy on the same DRP, but switch capacity savings to **Compressed** and select **Deduplicated**, if wanted. Wait for synchronization, and delete the source copy to complete migration.

Alternatively, right-click a non-compressed volume in a DRP, click **Modify Capacity Savings**, and choose **Compressed** and check **Deduplicated**, if needed. A second copy is created and the source copy is deleted automatically after synchronization.

Garbage collection and volume deletion

DRP includes built-in services to enable garbage collection of unused blocks. Garbage Collection is a DRP process that reduces the amount of data that is stored on external storage systems and internal drives by reclaiming previously used storage resources that are no longer needed by host systems.

When a DRP is created, the system monitors the pool for reclaimable capacity from host unmap operations. When space is freed from a host operating system, it is a process called *unmapping*. By using this process, hosts indicate that the allocated capacity is no longer required on a target volume. The freed space can be collected and reused on the system, without the reallocation of capacity on the storage.

Volume delete is a special case of Unmap that zeros an entire volume. There is a background Unmap process that sequentially walks the volume capacity, emulating large Unmap requests. The purpose is to notify the garbage collection technology to free up the physical space.

Because this process can require some time, the volume deletion process in DRPs is asynchronous from the command. After you delete a volume by using the GUI or CLI, it goes into *deleting* state, which can be noticed with the `lsvdisk` CLI command. After the Unmap process completes, the volume disappears.

Both host Unmaps and volume deletions increase unused space capacity. It is displayed by the system as `reclaimable_capacity`, which is shown with the `lsmdiskgrp` command, as shown in Example 9-10 on page 456. Unused capacity is freed after it is reclaimed by garbage collection.

9.5 Compression with standard pools

Random Access Compression Engine (RACE) technology was first introduced in the IBM Real-time Compression Appliances. It is integrated into the IBM Storwize V5000 software stack as the IBM Real-time Compression (RtC) solution.

RACE or RtC is used for compression of the volume copies, allocated from standard pools. DRPs use a different IBM Storwize V5000 compression function.

For more information about RtC compression, see *IBM Real-time Compression in IBM SAN Volume Controller and IBM Storwize V7000*, REDP-4859.

Note: RACE compression is supported on Storwize V5030 only.

9.5.1 Real-time Compression concepts

At a high level, the IBM RACE component compresses data that is written into the storage system dynamically. This compression occurs transparently, so Fibre Channel and iSCSI connected hosts are not aware of the compression.

RACE is an inline compression technology, which means that each host write is compressed as it passes through IBM Spectrum Virtualize to the disks. This technology has a clear benefit over other compression technologies that are post-processing based.

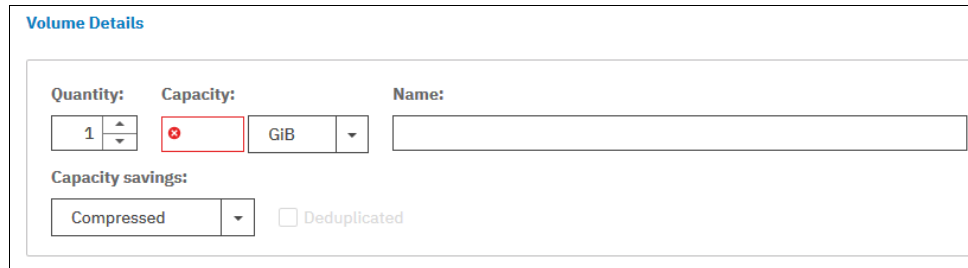
These technologies do not provide immediate capacity savings. Therefore, they are not a good fit for primary storage workloads, such as databases and active data set applications.

RACE is based on the Lempel-Ziv lossless data compression algorithm and operates by using a real-time method. When a host sends a write request, the request is acknowledged by the write cache of the system, and then staged to the storage pool. As part of its staging, the write request passes through the compression engine and is then stored in compressed format onto the storage pool. Therefore, writes are acknowledged immediately after they are received by the write cache with compression occurring as part of the staging to internal or external physical storage.

Capacity is saved when the data is written by the host because the host writes are smaller when they are written to the storage pool. IBM RtC is a self-tuning solution, adapting to the workload that runs on the system at any particular moment.

9.5.2 Implementing RtC compression

To create a compressed volume, choose **Capacity Savings - Compressed** in the Create Volumes window, as shown in Figure 9-16. For more information about creating volumes, see Chapter 6, “Volume configuration” on page 309.



The screenshot shows the 'Volume Details' window. It has three input fields: 'Quantity' with a spinner set to 1, 'Capacity' with a dropdown set to GiB and a red error icon, and 'Name' with an empty text box. Below these is the 'Capacity savings:' section, which has a dropdown menu set to 'Compressed' and an unchecked checkbox for 'Deduplicated'.

Figure 9-16 Creating compressed VDisk

In addition to compressing data in real time, you can also compress data sets (convert volume to compressed). To do so, you must change the capacity savings settings of the volume by right-clicking it and selecting **Modify Capacity Settings**. In the menu, select **Compression** as the Capacity Savings option, as shown in Figure 9-17.

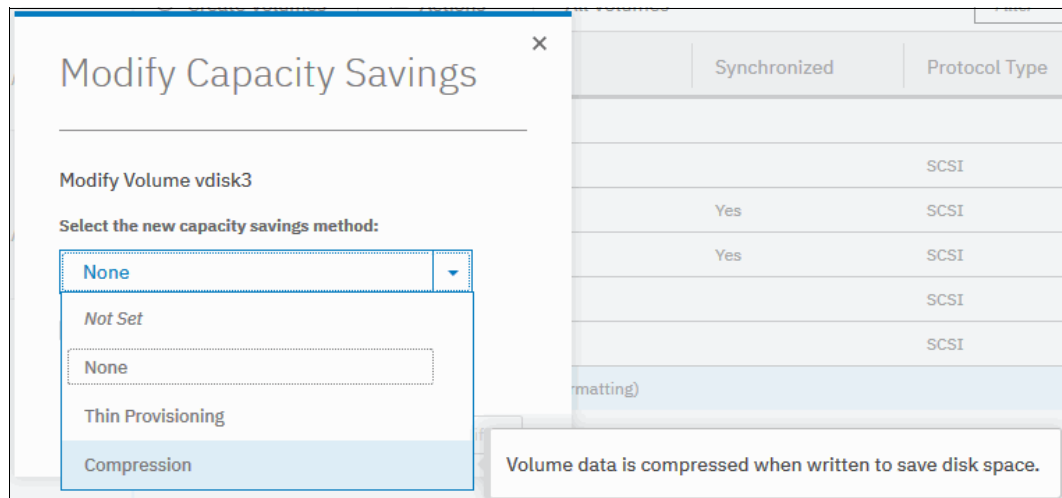


Figure 9-17 Selecting capacity setting

After the copies are fully synchronized, the original volume copy is deleted automatically.

As a result, you compressed data on the volume. This process is nondisruptive, so the data remains online and accessible by applications and users.

This capability enables clients to regain space from the storage pool, which can then be reused for other applications.

With the virtualization of external storage systems, the ability to compress that is stored data significantly enhances and accelerates the benefit to users. This capability enables them to see a tremendous return on their Storwize V5000 investment.

On the initial purchase of an Storwize V5000 with Real-time Compression, clients can defer their purchase of new storage. When new storage is needed, IT purchases a lower amount of the required storage before compression.

For more information about volume migration for compressed volumes on standard pools to DRPs, see “Migrating to and from DRP” on page 457.

9.6 Saving estimation for compression and deduplication

This section provides information about the specific tools that are used for sizing the environment for compression and deduplication.

9.6.1 Evaluate compression savings by using IBM Comprestimator

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 of showing the benefits of using compression. The utility performs read only operations and therefore has no effect on the data that is being stored on device.

If the compression savings prove to be beneficial in your environment, volume mirroring can be used to convert volumes to compressed volumes in the data reduction pools.

To analyze all the volumes that are currently on the system, run the CLI command `analyzevdiskbysystem`.

This command analyzes all the current volumes that are created on the system. Volumes that are created during or after the analysis are not included and can be analyzed individually. Progress for analyzing of all the volumes on system depends on the number of volumes that are being analyzed and results can be expected at about a minute per volume. For example, if a system has 50 volumes, compression savings analysis takes approximately 50 minutes.

You can run analysis on a single volume by specifying its name or ID as a parameter for the `analyzevdisk` CLI command.

To check the progress of the analysis run the `lsvdiskanalysisprogress` command. This command displays the total number of volumes on the system, total number of volumes that are remaining to be analyzed, and estimated time of completion.

The command `lsvdiskanalysis` is used to display information for thin provisioning and compression estimation analysis report for all volumes.

You can also use the GUI to run analysis. Browse to the Volumes pane, right-click any volume and select **Space Savings** → **Estimate Compression Savings**, as shown in Figure 9-18.

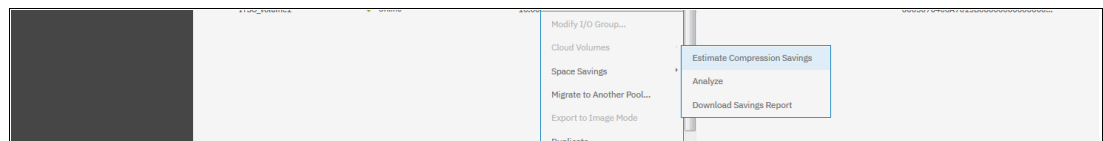


Figure 9-18 Comprestimator submenu

This action shows you the results of the latest estimation cycle and the date when it was completed. If there has been no analysis done yet, the system suggests running it.

To run or rerun on a single volume, select **Analyze** from the **Space Savings** submenu.

If you are using a version of IBM Spectrum Virtualize that is before V7.6 or if you want to estimate the compression savings of another IBM or non-IBM storage system, the separate IBM Comprestimator Utility can be installed on a host that is connected to the device that needs to be analyzed. For more information and the latest version of this utility, see this [IBM Support web page](#).

Consider the following recommended best practices for using the Comprestimator utility:

- ▶ Run the utility before implementing an IBM Spectrum Virtualize solution and before implementing DRPs.
- ▶ Download the latest version of the Comprestimator utility if you are not using one that is included in your IBM Spectrum Virtualize solution.
- ▶ Use the Comprestimator to analyze volumes that contain as much active data as possible rather than volumes that are nearly empty or newly created. This practice ensures 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.

9.6.2 Evaluating compression and deduplication

To help with the profiling and analysis of user workloads that must be migrated to the new system, IBM provides a highly accurate data reduction estimation tool that supports deduplication and compression. The tool operates by scanning target workloads on any older array (from IBM or third party) and then merging all scan results to provide an integrated system level data reduction estimate.

The Data Reduction Estimator Tool (DRET) utility uses advanced mathematical and statistical algorithms to perform an analysis with very low memory footprint. The utility runs on a host that can access the devices to be analyzed. It performs only read operations, so it has no effect on the data stored on the device.

The following sections provide information about installing DRET on a host and using it to analyze devices on it. Depending on the environment configuration, in many cases DRET is used on more than one host to analyze more data types.

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 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, which makes 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 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 deleted, but might still have traces on the device.

For more information and the latest version of this utility, see [this web page](#).



Copy Services

This chapter describes the Copy Services functions that are provided by the IBM Storwize V5000 Gen2 storage system, including FlashCopy, Remote Copy, and HyperSwap. Copy services functions are useful for making data copies for backup, application test, recovery, and so on. The IBM Storwize V5000 Gen2 system makes it easy to apply these functions to your environment through its intuitive graphical user interface (GUI).

This chapter includes the following topics:

- ▶ 10.1, “IBM FlashCopy” on page 466
- ▶ 10.2, “FlashCopy functional overview” on page 471
- ▶ 10.3, “Implementing FlashCopy” on page 472
- ▶ 10.4, “Managing FlashCopy by using the GUI” on page 493
- ▶ 10.5, “Volume mirroring and migration options” on page 521
- ▶ 10.6, “Native IP replication” on page 522
- ▶ 10.7, “Remote Copy services” on page 531
- ▶ 10.8, “Consistency protection for Remote and Global mirror” on page 558
- ▶ 10.9, “Remote Copy commands” on page 559
- ▶ 10.10, “Managing Remote Copy using the GUI” on page 567
- ▶ 10.11, “Troubleshooting remote copy” on page 590
- ▶ 10.12, “HyperSwap” on page 593

10.1 IBM FlashCopy

By using the IBM FlashCopy function of the IBM Spectrum Virtualize, you can perform a *point-in-time copy* of one or more volumes. This section describes the inner workings of FlashCopy and provides details about its configuration and use.

You can use FlashCopy to help you solve critical and challenging business needs that require duplication of data of your source volume. Volumes can remain online and active while you create consistent copies of the data sets. Because the copy is performed at the block level, it operates below the host operating system and its cache. Therefore, the copy is not apparent to the host.

Important: Because FlashCopy operates at the block level below the host operating system and cache, those levels do need to be flushed for consistent FlashCopies.

While the FlashCopy operation is performed, the source volume is briefly halted to initialize the FlashCopy bitmap, and then input/output (I/O) can resume. Although several FlashCopy options require the data to be copied from the source to the target in the background, which can take time to complete, the resulting data on the target volume is presented so that the copy appears to complete immediately.

This process is performed by using a bitmap (or bit array), which tracks changes to the data after the FlashCopy is started, and an indirection layer, which enables data to be read from the source volume transparently.

10.1.1 Business requirements for FlashCopy

FlashCopy can fulfil various business needs, it provides features that covers and satisfies many types of environments, database and applications. Common use cases for FlashCopy include, but are not limited to, the following examples:

- ▶ Rapidly creating consistent backups of dynamically changing data
- ▶ Rapidly creating consistent copies of production data to facilitate data movement or migration between hosts
- ▶ Rapidly creating copies of production data sets for application development and testing
- ▶ Rapidly creating copies of production data sets for auditing purposes and data mining
- ▶ Rapidly creating copies of production data sets for quality assurance

Regardless of your business needs, FlashCopy within the IBM Spectrum Virtualize is flexible and offers a broad feature set, which makes it applicable to many scenarios.

10.1.2 Backup improvements with FlashCopy

FlashCopy does not reduce the time that it takes to perform a backup to traditional backup infrastructure. However, it can be used to minimize and, under certain conditions, eliminate application downtime that is associated with performing backups. FlashCopy can also transfer the resource usage of performing intensive backups from production systems.

After the FlashCopy is performed, the resulting image of the data can be backed up to tape, as though it were the source system. After the copy to tape is completed, the image data is redundant and the target volumes can be discarded. For time-limited applications, such as these examples, “no copy” or incremental FlashCopy is used most often. The use of these methods puts less load on your infrastructure.

When FlashCopy is used for backup purposes, the target data usually is managed as read-only at the operating system level. This approach provides extra security by ensuring that your target data was not modified and remains true to the source.

10.1.3 Restore with FlashCopy

FlashCopy can perform a restore from any existing FlashCopy mapping. Therefore, you can restore (or copy) from the target to the source of your regular FlashCopy relationships. When restoring data from FlashCopy, this method can be qualified as reversing the direction of the FlashCopy mappings.

This capability has the following benefits:

- ▶ There is no need to worry about pairing mistakes; you trigger a restore.
- ▶ The process appears instantaneous.
- ▶ You can maintain a pristine image of your data while you are restoring what was the primary data.

This approach can be used for various applications, such as recovering your production database application after an errant batch process that caused extensive damage.

Preferred practices: Although restoring from a FlashCopy is quicker than a traditional tape media restore, you must not use restoring from a FlashCopy as a substitute for good archiving practices. Instead, keep one to several iterations of your FlashCopies so that you can near-instantly recover your data from the most recent history, and keep your long-term archive as appropriate for your business.

In addition to the restore option, which copies the original blocks from the target volume to modified blocks on the source volume, the target can be used to perform a restore of individual files. To do that, you make the target available on a host. We suggest that you do not make the target available to the source host, because seeing duplicates of disks causes problems for most host operating systems. Copy the files to the source using normal host data copy methods for your environment.

10.1.4 Moving and migrating data with FlashCopy

FlashCopy can be used to facilitate the migration of data between hosts while minimizing downtime for applications. By using FlashCopy, application data can be copied from source volumes to target volumes while applications remain online. After the volumes are fully copied and synchronized, the application can be brought down and then immediately brought back up on the new server that is accessing the new FlashCopy target volumes.

This method differs from the other migration methods, which are described later in this chapter. Common uses for this capability are host and back-end storage hardware refreshes.

10.1.5 Application testing with FlashCopy

It is often important to test a new version of an application or operating system that is using actual production data. This testing ensures the highest quality possible for your environment. FlashCopy makes this type of testing easy to accomplish without putting the production data at risk or requiring downtime to create a constant copy.

You create a FlashCopy of your source and use that for your testing. This copy is a duplicate of your production data down to the block level so that even physical disk identifiers are copied. Therefore, it is impossible for your applications to tell the difference.

10.1.6 Host and application considerations to ensure FlashCopy integrity

Because FlashCopy is at the block level, it is necessary to understand the interaction between your application and the host operating system. From a logical standpoint, it is easiest to think of these objects as “layers” that sit on top of one another. The application is the topmost layer, and beneath it is the operating system layer.

Both of these layers have various levels and methods of caching data to provide better speed. Because the FlashCopy sit below these layers, they are unaware of the cache at the application or operating system layers.

To ensure the integrity of the copy that is made, it is necessary to flush the host operating system and application cache for any outstanding reads or writes before the FlashCopy operation is performed. Failing to flush the host operating system and application cache produces what is referred to as a *crash consistent* copy.

The resulting copy requires the same type of recovery procedure, such as log replay and file system checks, that is required following a host crash. FlashCopies that are crash consistent often can be used following file system and application recovery procedures.

Various operating systems and applications provide facilities to stop I/O operations and ensure that all data is flushed from host cache. If these facilities are available, they can be used to prepare for a FlashCopy operation. When this type of facility is unavailable, the host cache must be flushed manually by quiescing the application and unmounting the file system or drives.

Preferred practice: From a practical standpoint, when you have an application that is backed by a database and you want to make a FlashCopy of that application’s data, it is sufficient in most cases to use the write-suspend method that is available in most modern databases because the database maintains strict control over I/O.

This method is the opposite of flushing data from the application and the backing database, which is always the suggested method because it is safer. However, this method can be used when facilities do not exist or your environment includes time sensitivity.

10.1.7 FlashCopy attributes

The FlashCopy function in IBM Spectrum Virtualize features the following attributes:

- ▶ The target is the time-zero copy of the source, which is known as *FlashCopy mapping targets*.
- ▶ FlashCopy produces a copy of the source volume, including any metadata that was written by the host operating system, logical volume manager, and applications.

- ▶ The source volume and target volume are available (almost) immediately following the FlashCopy operation.
- ▶ The source and target volumes must be the same “virtual” size.
- ▶ The source and target volumes must be on the same IBM Storwize system.
- ▶ The source and target volumes do not need to be in the same I/O Group or storage pool.
- ▶ The storage pool extent sizes can differ between the source and target.
- ▶ The source volumes can have up to 256 target volumes (Multiple Target FlashCopy).
- ▶ The target volumes can be the source volumes for other FlashCopy relationships (*cascaded FlashCopy*).
- ▶ Consistency groups are supported to enable FlashCopy across multiple volumes at the same time.
- ▶ Up to 255 FlashCopy consistency groups are supported per system.
- ▶ Up to 512 FlashCopy mappings can be placed in one consistency group.
- ▶ The target volume can be updated independently of the source volume.
- ▶ Bitmaps that are governing I/O redirection (I/O indirection layer) are maintained in both node canisters of the IBM Storwize I/O Group to prevent a single point of failure.
- ▶ FlashCopy mapping and Consistency Groups can be automatically withdrawn after the completion of the background copy.
- ▶ Thin-provisioned FlashCopy (or Snapshot in the graphical user interface [GUI]) use disk space only when updates are made to the source or target data, and not for the entire capacity of a volume copy.
- ▶ FlashCopy licensing is based on the virtual capacity of the source volumes.
- ▶ Incremental FlashCopy copies all of the data when you first start FlashCopy and then only the changes when you stop and start FlashCopy mapping again. Incremental FlashCopy can substantially reduce the time that is required to re-create an independent image.
- ▶ Reverse FlashCopy enables FlashCopy targets to become restore points for the source without breaking the FlashCopy relationship, and without having to wait for the original copy operation to complete.
- ▶ The maximum number of supported FlashCopy mappings is 4096 per clustered system.
- ▶ The size of the source and target volumes cannot be altered (increased or decreased) while a FlashCopy mapping is defined.

10.1.8 Reverse FlashCopy

Reverse FlashCopy enables FlashCopy targets to become restore points for the source without breaking the FlashCopy relationship, and without having to wait for the original copy operation to complete. It supports multiple targets (up to 256) and therefore, multiple rollback points.

A key advantage of the IBM Spectrum Virtualize Multiple Target Reverse FlashCopy function is that the reverse FlashCopy does not destroy the original target, which enables processes that are using the target, such as a tape backup, to continue uninterrupted.

IBM Spectrum Virtualize also provides the ability to create an optional copy of the source volume to be made before the reverse copy operation starts. This ability to restore back to the original source data can be useful for diagnostic purposes.

Complete the following steps to restore from an on-disk backup:

1. (Optional) Create a target volume (volume Z) and use FlashCopy to copy the production volume (volume X) onto the new target for later problem analysis.
2. Create a FlashCopy map with the backup to be restored (volume Y) or (volume W) as the source volume and volume X as the target volume, if this map does not exist.
3. Start the FlashCopy map (volume Y → volume X) with the `-restore` option to copy the backup data onto the production disk. If the `-restore` option is specified and no FlashCopy mapping exists, the command is ignored, which preserves your data integrity.

The production disk is instantly available with the backup data. Figure 10-1 shows an example of Reverse FlashCopy.

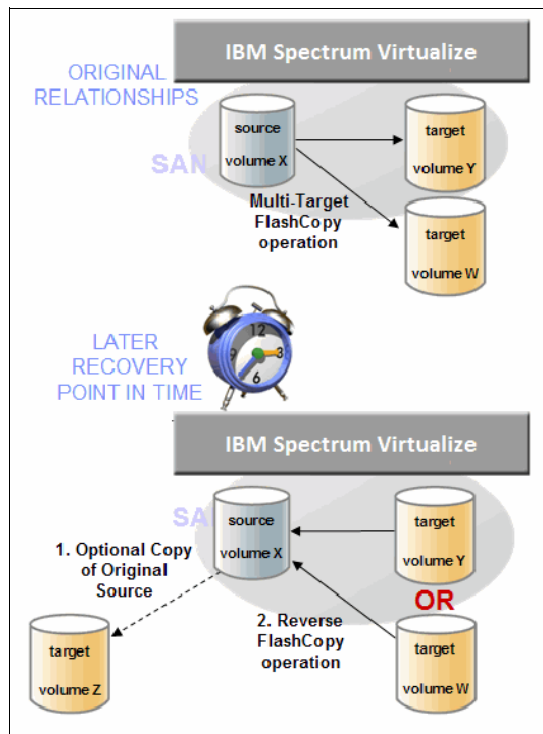


Figure 10-1 Reverse FlashCopy

Regardless of whether the initial FlashCopy map (volume X → volume Y) is incremental, the Reverse FlashCopy operation copies the modified data only.

Consistency Groups are reversed by creating a set of new reverse FlashCopy maps and adding them to a new reverse Consistency Group. Consistency Groups cannot contain more than one FlashCopy map with the same target volume.

10.1.9 IBM Spectrum Protect Snapshot

The management of many large FlashCopy relationships and Consistency Groups is a complex task without a form of automation for assistance. IBM Spectrum Protect™ Snapshot (formerly *IBM Tivoli® FlashCopy Manager*) provides fast application-aware backups and restores using advanced point-in-time image technologies in the IBM Spectrum Virtualize.

In addition, it allows you to manage frequent, near-instant, nondisruptive, application-aware backups and restores using integrated application and VMware snapshot technologies. IBM Spectrum Protect Snapshot can be widely used in IBM and non-IBM storage systems.

Note: For more information about how IBM Spectrum Protect Snapshot can help your business, see [this web page](#).

10.2 FlashCopy functional overview

FlashCopy works by defining a FlashCopy mapping that consists of one source volume with one target volume. Multiple FlashCopy mappings can be defined, and point-in-time consistency can be maintained across multiple individual mappings by using Consistency Groups. For more information, see “Consistency Group with Multiple Target FlashCopy” on page 475.

Before you start a FlashCopy (regardless of the type and options specified), you must issue a **prestartfcmap** or **prestartfcconsistgrp** command, which puts the cache into write-through mode and provides a flushing of the I/O currently bound for your volume. After FlashCopy is started, an effective copy of a source volume to a target volume is created.

The content of the source volume is presented immediately on the target volume and the original content of the target volume is lost. This FlashCopy operation is also referred to as a *time-zero copy* (T0).

Tip: Rather than using **prestartfcmap** or **prestartfcconsistgrp**, you can also use the **-prep** parameter in the **startfcmap** or **startfcconsistgrp** command to prepare and start FlashCopy in one step.

The source and target volumes are available for use immediately after the FlashCopy operation. The FlashCopy operation creates a bitmap that is referenced and maintained to direct I/O requests within the source and target relationship. This bitmap is updated to reflect the active block locations as data is copied in the background from the source to the target, and updates are made to the source.

For more information about background copy, see 10.3.5, “Grains and the FlashCopy bitmap” on page 477.

Figure 10-2 shows the redirection of the host I/O toward the source volume and the target volume.

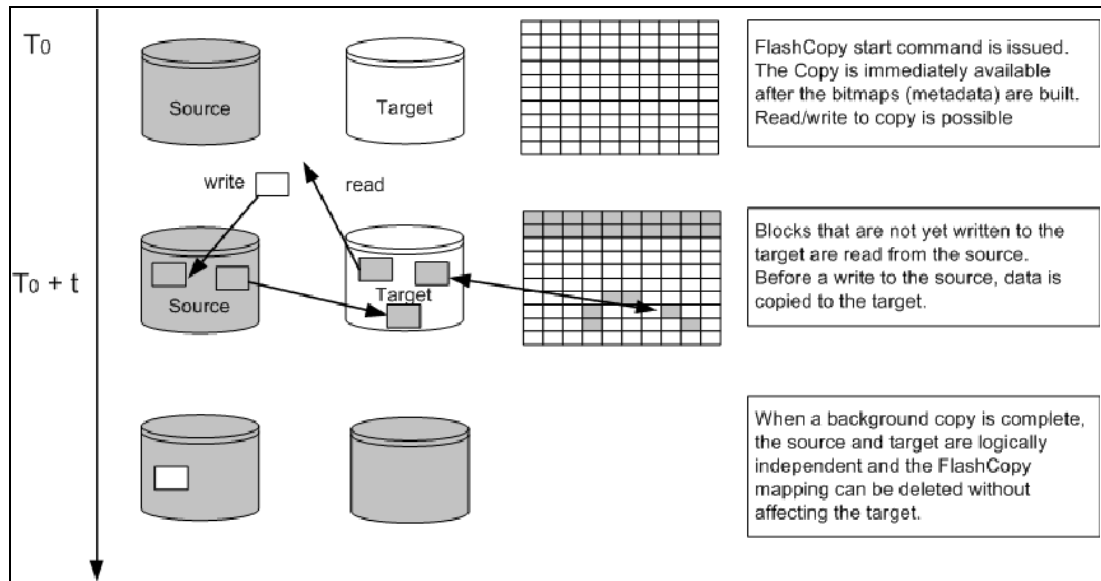


Figure 10-2 Redirection of host I/O

10.3 Implementing FlashCopy

This section describes how FlashCopy is implemented in the IBM Spectrum Virtualize running on IBM Storwize.

10.3.1 FlashCopy mappings

FlashCopy occurs between a source volume and a target volume. The source and target volumes must be the same size. The minimum granularity that IBM Spectrum Virtualize supports for FlashCopy is an entire volume. It is not possible to use FlashCopy to copy only part of a volume.

Important: As with any point-in-time copy technology, you are bound by operating system and application requirements for interdependent data and the restriction to an entire volume.

The source and target volumes must belong to the same IBM Storwize V5000 system Gen2 system, but they do not have to be in the same I/O Group or storage pool. FlashCopy associates a source volume to a target volume through FlashCopy mapping.

To become members of a FlashCopy mapping, source and target volumes must be the same size. Volumes that are members of a FlashCopy mapping cannot have their size increased or decreased while they are members of the FlashCopy mapping.

A *FlashCopy mapping* is the act of creating a relationship between a source volume and a target volume. FlashCopy mappings can be stand-alone or a member of a Consistency Group. You can perform the actions of preparing, starting, or stopping FlashCopy on a stand-alone mapping or a Consistency Group.

Figure 10-3 shows the concept of FlashCopy mapping.

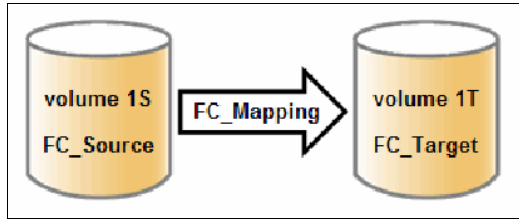


Figure 10-3 FlashCopy mapping

10.3.2 Multiple Target FlashCopy

The IBM Storwize V5000 Gen2 system supports up to 256 target volumes from a single source volume. Each copy is managed by a unique mapping. Figure 10-4 shows the Multiple Target FlashCopy implementation.

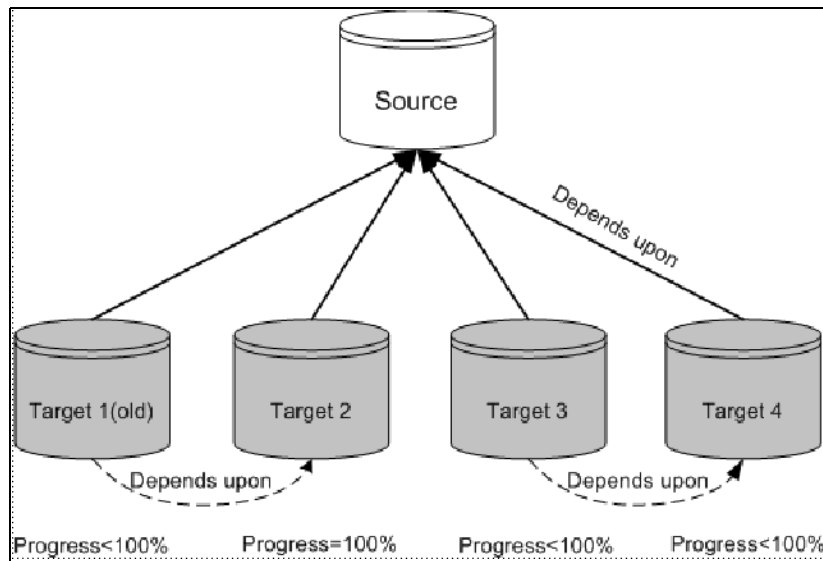


Figure 10-4 Multiple Target FlashCopy implementation

Figure 10-4 also shows four targets and mappings that are taken from a single source, along with their interdependencies. In this example, Target 1 is the oldest (as measured from the time that it was started) through to Target 4, which is the newest. The ordering is important because of how the data is copied when multiple target volumes are defined and because of the dependency chain that results.

A write to the source volume does not cause its data to be copied to all of the targets. Instead, it is copied to the newest target volume only (Target 4 in Figure 10-4). The older targets refer to new targets first before referring to the source.

From the point of view of an intermediate target disk (not the oldest or the newest), it treats the set of newer target volumes and the true source volume as a type of composite source. It treats all older volumes as a kind of target (and behaves like a source to them).

If the mapping for an intermediate target volume shows 100% progress, its target volume contains a complete set of data. In this case, mappings treat the set of newer target volumes (up to and including the 100% progress target) as a form of composite source.

A dependency relationship exists between a particular target and all newer targets (up to and including a target that shows 100% progress) that share the source until all data is copied to this target and all older targets.

For more information about Multiple Target FlashCopy, see 10.3.6, “Interaction and dependency between multiple target FlashCopy mappings” on page 478.

10.3.3 Consistency Groups

Consistency Groups address the requirement to preserve point-in-time data consistency across multiple volumes for applications that include related data that spans multiple volumes. For these volumes, Consistency Groups maintain the integrity of the FlashCopy by ensuring that “dependent writes” are run in the application’s intended sequence.

When Consistency Groups are used, the FlashCopy commands are issued to the FlashCopy Consistency Group, which performs the operation on all FlashCopy mappings that are contained within the Consistency Group at the same time.

Figure 10-5 shows a Consistency Group that includes two FlashCopy mappings.

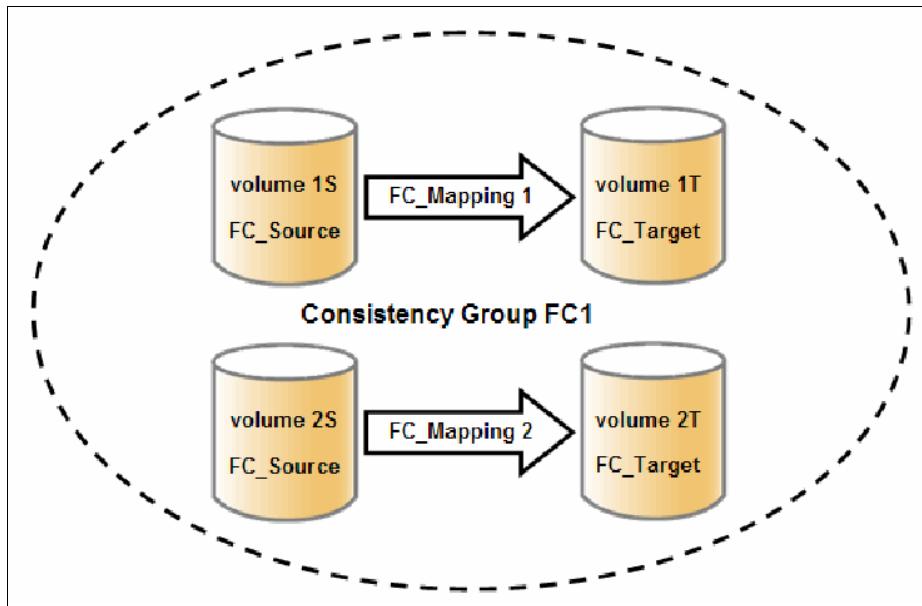


Figure 10-5 FlashCopy Consistency Group

Important: After an individual FlashCopy mapping is added to a Consistency Group, it can be managed as part of the group only. Operations, such as prepare, start, and stop, are no longer allowed on the individual mapping.

Dependent writes

To show why it is crucial to use Consistency Groups when a data set spans multiple volumes, consider the following typical sequence of writes for a database update transaction:

1. A write is run to update the database log, which indicates that a database update is about to be performed.
2. A second write is run to perform the actual update to the database.

3. A third write is run to update the database log, which indicates that the database update completed successfully.

The database ensures the correct ordering of these writes by waiting for each step to complete before the next step is started. However, if the database log (updates 1 and 3) and the database (update 2) are on separate volumes, it is possible for the FlashCopy of the database volume to occur before the FlashCopy of the database log. This sequence can result in the target volumes seeing writes 1 and 3 but not 2 because the FlashCopy of the database volume occurred before the write was completed.

In this case, if the database was restarted by using the backup that was made from the FlashCopy target volumes, the database log indicates that the transaction completed successfully. In fact, it did not complete successfully because the FlashCopy of the volume with the database file was started (the bitmap was created) before the write completed to the volume. Therefore, the transaction is lost and the integrity of the database is in question.

To overcome the issue of dependent writes across volumes and to create a consistent image of the client data, a FlashCopy operation must be performed on multiple volumes as an atomic operation. To accomplish this method, the IBM Spectrum Virtualize supports the concept of *Consistency Groups*.

FlashCopy commands can then be issued to the FlashCopy Consistency Group and, therefore, simultaneously for all of the FlashCopy mappings that are defined in the Consistency Group.

For example, when a FlashCopy **start** command is issued to the Consistency Group, all of the FlashCopy mappings in the Consistency Group are started at the same time. This simultaneous start results in a point-in-time copy that is consistent across all of the FlashCopy mappings that are contained in the Consistency Group.

Consistency Group with Multiple Target FlashCopy

A Consistency Group aggregates FlashCopy mappings, not volumes. Therefore, where a source volume has multiple FlashCopy mappings, they can be in the same or separate Consistency Groups.

If a particular volume is the source volume for multiple FlashCopy mappings, you might want to create separate Consistency Groups to separate each mapping of the same source volume. Regardless of whether the source volume with multiple target volumes is in the same consistency group or in separate consistency groups, the resulting FlashCopy produces multiple identical copies of the source data.

Maximum configurations

Table 10-1 lists the FlashCopy properties and maximum configurations.

Table 10-1 FlashCopy properties and maximum configurations

FlashCopy property	Maximum	Comment
FlashCopy targets per source	256	This maximum is the number of FlashCopy mappings that can exist with the same source volume.
FlashCopy mappings per system	4096	The number of mappings is no longer limited by the number of volumes in the system, so the FlashCopy component limit applies.
FlashCopy Consistency Groups per system	255	This maximum is an arbitrary limit that is policed by the software.

FlashCopy property	Maximum	Comment
FlashCopy volume capacity per I/O Group	4 pebibytes (PiB)	This maximum is a limit on the quantity of FlashCopy mappings that are using bitmap space from this I/O Group. This maximum configuration uses all 4 gibibytes (GiB) of bitmap space for the I/O Group and allows no Metro or Global Mirror bitmap space. The default is 40 tebibytes (TiB).
FlashCopy mappings per Consistency Group	512	This limit is because of the time that is taken to prepare a Consistency Group with many mappings.

10.3.4 FlashCopy indirection layer

The *FlashCopy indirection layer* governs the I/O to the source and target volumes when a FlashCopy mapping is started, which is done by using a FlashCopy bitmap. The purpose of the FlashCopy indirection layer is to enable the source and target volumes for read and write I/O immediately after the FlashCopy is started.

To show how the FlashCopy indirection layer works, we examine what happens when a FlashCopy mapping is prepared and then started.

When a FlashCopy mapping is prepared and started, the following sequence is applied:

1. Flush the write cache to the source volume or volumes that are part of a Consistency Group.
2. Put cache into write-through mode on the source volumes.
3. Discard cache for the target volumes.
4. Establish a sync point on all of the source volumes in the Consistency Group (which creates the FlashCopy bitmap).
5. Ensure that the indirection layer governs all of the I/O to the source volumes and target volumes.
6. Enable cache on the source volumes and target volumes.

FlashCopy provides the semantics of a point-in-time copy by using the indirection layer, which intercepts I/O that is directed at the source or target volumes. The act of starting a FlashCopy mapping causes this indirection layer to become active in the I/O path, which occurs automatically across all FlashCopy mappings in the Consistency Group.

The indirection layer then determines how each I/O is to be routed, based on the following factors:

- ▶ The volume and the logical block address (LBA) to which the I/O is addressed
- ▶ Its direction (read or write)
- ▶ The state of an internal data structure, the FlashCopy bitmap

The indirection layer allows the I/O to go through to the underlying volume, redirects the I/O from the target volume to the source volume, or queues the I/O while it arranges for data to be copied from the source volume to the target volume. To explain in more detail which action is applied for each I/O, we first look at the FlashCopy bitmap.

10.3.5 Grains and the FlashCopy bitmap

When data is copied between volumes, it is copied in units of address space that are known as *grains*. Grains are units of data that are grouped to optimize the use of the bitmap that tracks changes to the data between the source and target volume. You can use 64 kibibytes (KiB) or 256 KiB grain sizes (256 KiB is the default). The FlashCopy bitmap contains 1 bit for each grain, and is used to show whether the source grain was copied to the target. The 64 KiB grain size uses bitmap space at a rate of four times the default 256 KiB size.

The FlashCopy bitmap dictates read and write behavior for the source and target volumes.

Source reads

Reads are performed from the source volume, which is the same as for non-FlashCopy volumes.

Source writes

Writes to the source cause one of the following actions:

- ▶ If the grain was not copied to the target yet, the grain is copied before the actual write is performed to the source. The bitmap is updated to indicate that this grain is already copied to the target.
- ▶ If the grain was copied, the write is performed to the source as usual.

Target reads

Reads are performed from the target if the grain was copied. Otherwise, the read is performed from the source and no copy is performed.

Target writes

Writes to the target cause one of the following actions:

- ▶ If the grain was not copied from the source to the target, the grain is copied from the source to the target before the actual write is performed to the source. The bitmap is updated to indicate that this grain is already copied to the target.
- ▶ If the entire grain is being updated on the target, the target is marked as split with the source (if there is no I/O error during the write) and the write goes directly to the target.
- ▶ If the grain in question was already copied from the source to the target, the write goes directly to the target.

The FlashCopy indirection layer algorithm

Imagine the FlashCopy indirection layer as the I/O traffic director when a FlashCopy mapping is active. The I/O is intercepted and handled according to whether it is directed at the source volume or at the target volume, depending on the nature of the I/O (read or write) and the state of the grain (whether it was copied).

Figure 10-6 shows how the background copy runs while I/Os are handled according to the indirection layer algorithm.

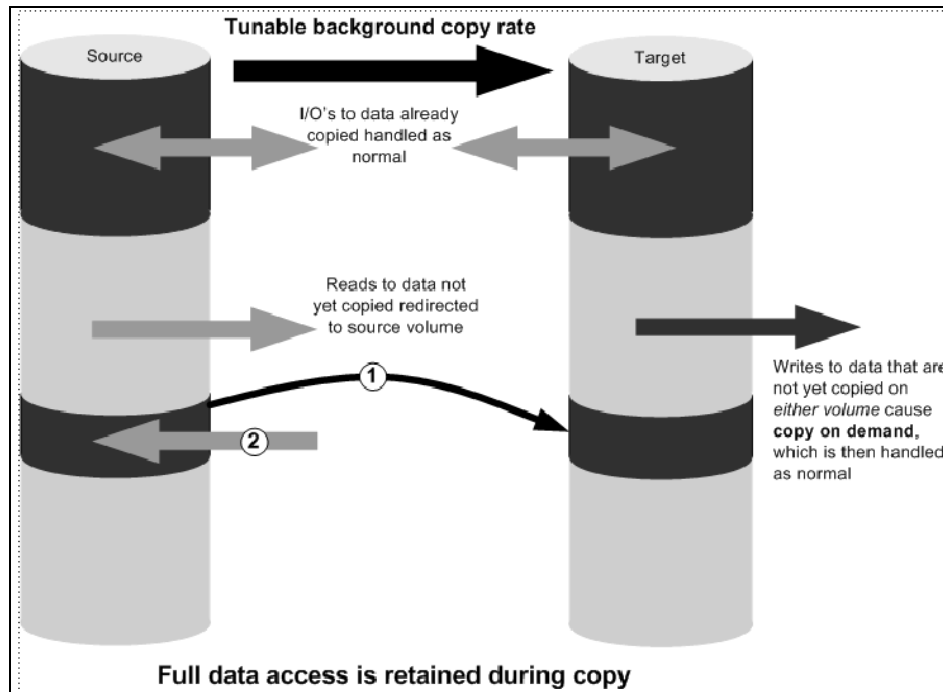


Figure 10-6 I/O processing with FlashCopy

10.3.6 Interaction and dependency between multiple target FlashCopy mappings

Figure 10-7 shows a set of four FlashCopy mappings that share a common source. The FlashCopy mappings target volumes Target 0, Target 1, Target 2, and Target 3.

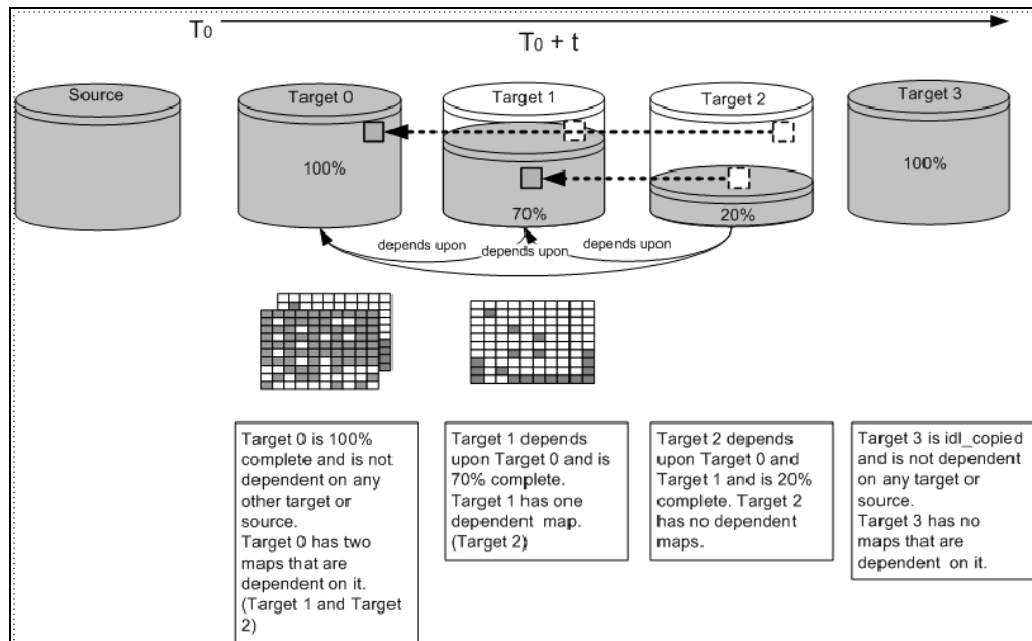


Figure 10-7 Interactions among multiple target FlashCopy mappings

In Figure 10-7 on page 478, Target 0 is not dependent on a source because it completed copying. Target 0 has two dependent mappings (Target 1 and Target 2).

Target 1 depends on Target 0. It remains dependent until all of Target 1 is copied. Target 2 depends on it because Target 2 is 20% copy complete. After all of Target 1 is copied, it can then move to the `idle_copied` state.

Target 2 is dependent upon Target 0 and Target 1 and remains dependent until all of Target 2 is copied. No target depends on Target 2; therefore, when all of the data is copied to Target 2, it can move to the `idle_copied` state.

Target 3 completed copying, so it is not dependent on any other maps.

Target writes with Multiple Target FlashCopy

A write to an intermediate or the newest target volume must consider the state of the grain within its own mapping, and the state of the grain of the next oldest mapping.

If the grain of the next oldest mapping is not yet copied, it must be copied before the write can proceed, to preserve the contents of the next oldest mapping. The data that is written to the next oldest mapping comes from a target or source.

If the grain in the target that is being written is not yet copied, the grain is copied from the oldest copied grain in the mappings that are newer than the target, or from the source if none is copied. After this copy is done, the write can be applied to the target.

Target reads with Multiple Target FlashCopy

If the grain being read is copied from the source to the target, the read returns data from the target that is being read. If the grain is not yet copied, each of the newer mappings is examined in turn, and the read is performed from the first copy that is found. If none is found, the read is performed from the source.

Stopping the copy process

When a **stop** command is issued to a mapping that contains a target that has dependent mappings, the mapping enters the `stopping` state and begins copying all grains that are uniquely held on the target volume of the mapping that is being stopped to the next oldest mapping that is in the `Copying` state. The mapping remains in the `stopping` state until all grains are copied, and then enters the `stopped` state.

Note: The stopping copy process can be ongoing for several mappings that share the source at the same time. At the completion of this process, the mapping automatically makes an asynchronous state transition to the `stopped` state, or the `idle_copied` state if the mapping was in the `copying` state with `progress = 100%`.

For example, if the mapping that is associated with Target 0 was issued a **stopfcmap** or **stopfcconsistgrp** command, Target 0 enters the `stopping` state while a process copies the data of Target 0 to Target 1. After all of the data is copied, Target 0 enters the `stopped` state, and Target 1 is no longer dependent upon Target 0; however, Target 1 remains dependent on Target 2.

10.3.7 Summary of the FlashCopy indirection layer algorithm

Table 10-2 summarizes the indirection layer algorithm.

Table 10-2 Summary table of the FlashCopy indirection layer algorithm

Accessed volume	Was the grain copied?	Host I/O operation	
		Read	Write
Source	No	Read from the source volume.	Copy grain to most recently started target for this source, then write to the source.
	Yes	Read from the source volume.	Write to the source volume.
Target	No	If any newer targets exist for this source in which this grain was copied, read from the oldest of these targets. Otherwise, read from the source.	Hold the write. Check the dependency target volumes to see whether the grain was copied. If the grain is not copied to the next oldest target for this source, copy the grain to the next oldest target. Then, write to the target.
	Yes	Read from the target volume.	Write to the target volume.

10.3.8 Interaction with the cache

Starting with V7.3, the entire cache subsystem was redesigned and changed. Cache was divided into upper and lower cache. Upper cache serves mostly as write cache and hides the write latency from the hosts and application. Lower cache is a read/write cache and optimizes I/O to and from disks. Figure 10-8 shows the new IBM Spectrum Virtualize cache architecture.

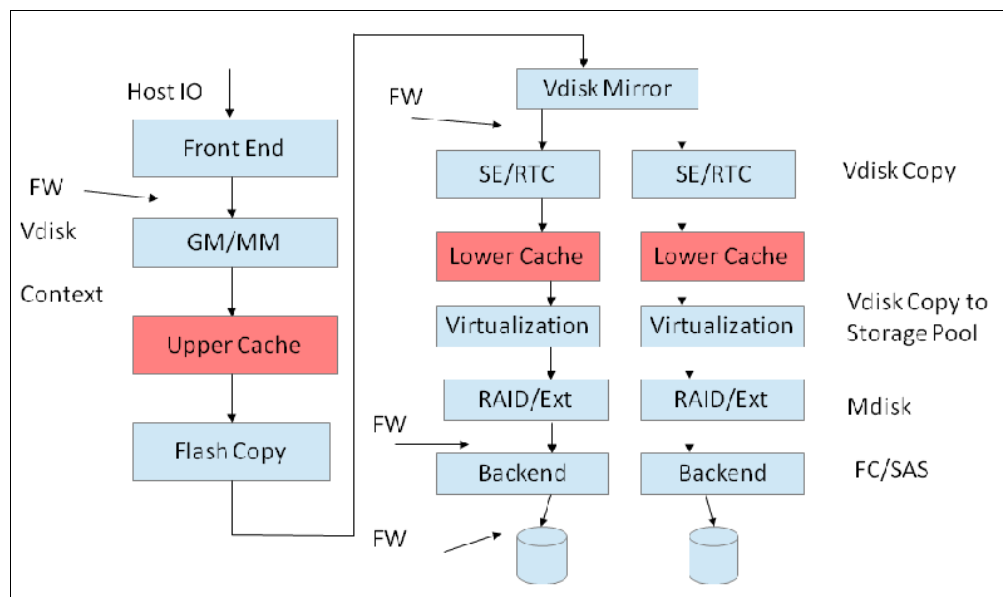


Figure 10-8 New cache architecture

This copy-on-write process introduces significant latency into write operations. To isolate the active application from this extra latency, the FlashCopy indirection layer is placed logically between upper and lower cache. Therefore, the extra latency that is introduced by the copy-on-write process is encountered only by the internal cache operations and not by the application.

Figure 10-9 shows the logical placement of the FlashCopy indirection layer.

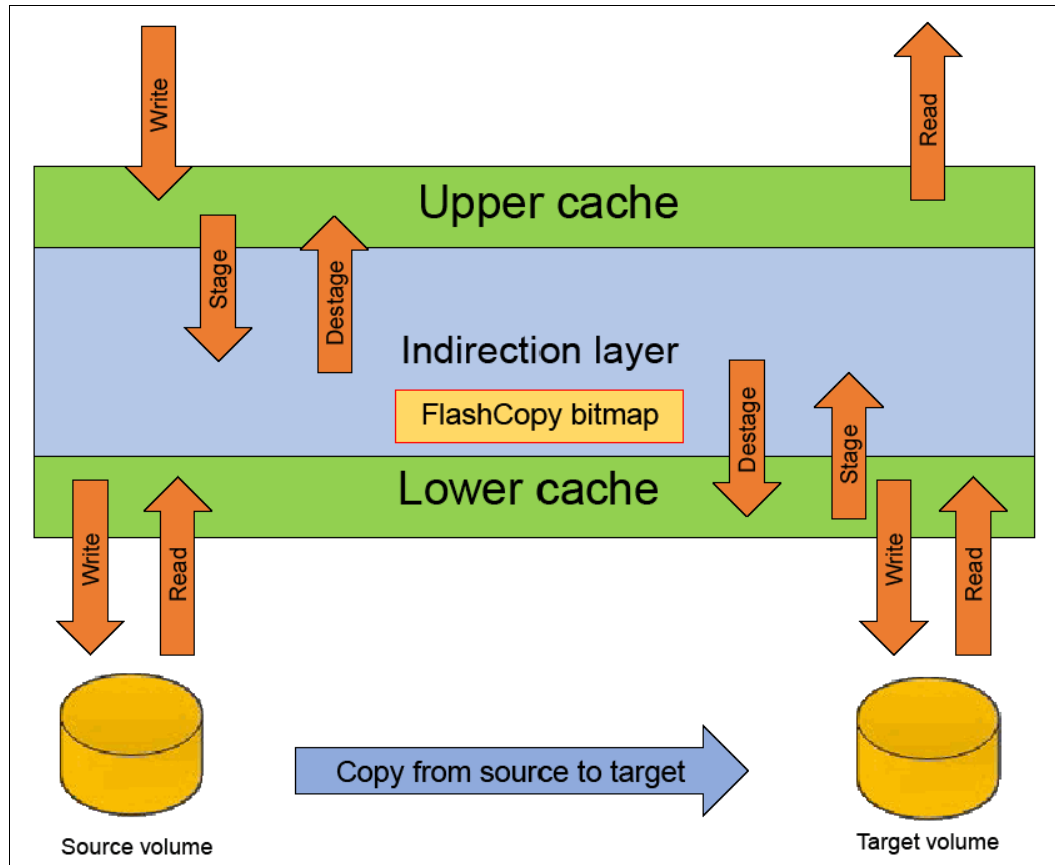


Figure 10-9 Logical placement of the FlashCopy indirection layer

Introduction of the two-level cache provides more performance improvements to the FlashCopy mechanism. Because now the FlashCopy layer is above lower cache in the IBM Spectrum Virtualize software stack, it can benefit from read prefetching and coalescing writes to backend storage. Also, preparing FlashCopy is much faster, because upper cache write data does not have to go directly to backend storage but to lower cache layer.

Also, in the multitarget FlashCopy the target volumes of the same image share cache data. This design is opposite to previous IBM Spectrum Virtualize code versions, where each volume had its own copy of cached data.

10.3.9 FlashCopy and image mode volumes

FlashCopy can be used with image mode volumes. Because the source and target volumes must be the same size, you must create a target volume with the same size as the image mode volume when you are creating a FlashCopy mapping. To accomplish this task, use the `svcinfolsvdisk -bytes volume_name` command. The size in bytes is then used to create the volume that is used in the FlashCopy mapping.

This method provides an exact number of bytes because image mode volumes might not line up one-to-one on other measurement unit boundaries. Example 10-1 lists the size of the ITS0-V5Krs001 volume. The ITS0-V5Krs002-image volume is then created, which specifies the same size.

Example 10-1 Listing the size of a volume in bytes and creating a volume of equal size

```
IBM_Storwize:ITS0V5030:superuser>lsvdisk -bytes ITS0-V5Krs001
id 40
name ITS0-V5Krs001
IO_group_id 0
IO_group_name io_grp0
status online
mdisk_grp_id 1
mdisk_grp_name ITS0 Redbook
capacity 1073741824
type striped
...

IBM_Storwize:ITS0V5030:superuser>mkvdisk -mdiskgrp 1 -iogrp 0 -size 1073741824
-unit b -vtype striped -name ITS0-V5Krs002-image
Virtual Disk, id [41], successfully created
IBM_Storwize:ITS0V5030:superuser>

IBM_Storwize:ITS0V5030:superuser>lsvdisk -delim " "
40 ITS0-V5Krs001 0 io_grp0 online 1 ITS0 Redbook 1.00GB striped
6005076801B587F93400000000000034 0 1 not_empty 0 no 0 1 ITS0 Redbook no no 40
ITS0-V5Krs001
41 ITS0-V5Krs002-image 0 io_grp0 online 1 ITS0 Redbook 1.00GB striped
6005076801B587F93400000000000035 0 1 not_empty 0 no 0 1 ITS0 Redbook yes no 41
ITS0-V5Krs002-image
IBM_Storwize:ITS0V5030:superuser>
```

Tip: Alternatively, you can use the `expandvdisksize` and `shrinkvdisksize` volume commands to modify the size of the volume. These actions must be performed before a mapping is created. You can use an image mode volume as a FlashCopy source volume or target volume.

10.3.10 FlashCopy mapping events

In this section, we describe the events that modify the states of a FlashCopy. We also describe the mapping events that are listed in Table 10-3 on page 483.

Overview of a FlashCopy sequence of events: The following tasks show the FlashCopy sequence:

1. Associate the source data set with a target location (one or more source and target volumes).
2. Create a FlashCopy mapping for each source volume to the corresponding target volume. The target volume must be equal in size to the source volume.
3. Discontinue access to the target (application dependent).
4. Prepare (pre-trigger) the FlashCopy:
 - a. Flush the cache for the source.
 - b. Discard the cache for the target.
5. Start (trigger) the FlashCopy:
 - a. Pause I/O (briefly) on the source.
 - b. Resume I/O on the source.
 - c. Start I/O on the target.

Table 10-3 Mapping events

Mapping event	Description
Create	<p>A FlashCopy mapping is created between the specified source volume and the specified target volume. The operation fails if any one of the following conditions is true:</p> <ul style="list-style-type: none"> ▶ The source volume is a member of 256 FlashCopy mappings. ▶ The node has insufficient bitmap memory. ▶ The source and target volumes are different sizes.
Prepare	<p>The prestartfcmap or prestartfcconsistgrp command is directed to a Consistency Group for FlashCopy mappings that are members of a normal Consistency Group or to the mapping name for FlashCopy mappings that are stand-alone mappings. The prestartfcmap or prestartfcconsistgrp command places the FlashCopy mapping into the Preparing state.</p> <p>The prestartfcmap or prestartfcconsistgrp command can corrupt any data that was on the target volume because cached writes are discarded. Even if the FlashCopy mapping is never started, the data from the target might be changed logically during the act of preparing to start the FlashCopy mapping.</p>
Flush done	<p>The FlashCopy mapping automatically moves from the preparing state to the prepared state after all cached data for the source is flushed and all cached data for the target is no longer valid.</p>

Mapping event	Description
Start	<p>When all of the FlashCopy mappings in a Consistency Group are in the prepared state, the FlashCopy mappings can be started. To preserve the cross-volume Consistency Group, the start of all of the FlashCopy mappings in the Consistency Group must be synchronized correctly concerning I/Os that are directed at the volumes by using the startfcmap or startfcconsistgrp command.</p> <p>The following actions occur during the running of the startfcmap command or the startfcconsistgrp command:</p> <ul style="list-style-type: none"> ▶ New reads and writes to all source volumes in the Consistency Group are paused in the cache layer until all ongoing reads and writes beneath the cache layer are completed. ▶ After all FlashCopy mappings in the Consistency Group are paused, the internal cluster state is set to enable FlashCopy operations. ▶ After the cluster state is set for all FlashCopy mappings in the Consistency Group, read and write operations continue on the source volumes. ▶ The target volumes are brought online. <p>As part of the startfcmap or startfcconsistgrp command, read and write caching is enabled for the source and target volumes.</p>
Modify	<p>The following FlashCopy mapping properties can be modified:</p> <ul style="list-style-type: none"> ▶ FlashCopy mapping name ▶ Clean rate ▶ Consistency group ▶ Copy rate (for background copy or stopping copy priority) ▶ Automatic deletion of the mapping when the background copy is complete
Stop	<p>The following separate mechanisms can be used to stop a FlashCopy mapping:</p> <ul style="list-style-type: none"> ▶ Issue a command ▶ An I/O error occurred
Delete	<p>This command requests that the specified FlashCopy mapping is deleted. If the FlashCopy mapping is in the copying state, the force flag must be used.</p>
Flush failed	<p>If the flush of data from the cache cannot be completed, the FlashCopy mapping enters the stopped state.</p>
Copy complete	<p>After all of the source data is copied to the target and there are no dependent mappings, the state is set to copied. If the option to automatically delete the mapping after the background copy completes is specified, the FlashCopy mapping is deleted automatically. If this option is not specified, the FlashCopy mapping is not deleted automatically and can be reactivated by preparing and starting again.</p>
Bitmap online/offline	<p>The node failed.</p>

10.3.11 FlashCopy mapping states

This section describes the states of a FlashCopy mapping.

Idle_or_copied

The source and target volumes act as independent volumes, even if a mapping exists between the two. Read and write caching is enabled for the source and the target volumes.

If the mapping is incremental and the background copy is complete, the mapping records the differences between the source and target volumes only. If the connection to both nodes in the I/O group that the mapping is assigned to is lost, the source and target volumes are offline.

Copying

The copy is in progress. Read and write caching is enabled on the source and the target volumes.

Prepared

The mapping is ready to start. The target volume is online, but is not accessible. The target volume cannot perform read or write caching. Read and write caching is failed by the Small Computer System Interface (SCSI) front end as a hardware error. If the mapping is incremental and a previous mapping is completed, the mapping records the differences between the source and target volumes only. If the connection to both nodes in the I/O group that the mapping is assigned to is lost, the source and target volumes go offline.

Preparing

The target volume is online, but not accessible. The target volume cannot perform read or write caching. Read and write caching is failed by the SCSI front end as a hardware error. Any changed write data for the source volume is flushed from the cache. Any read or write data for the target volume is discarded from the cache.

If the mapping is incremental and a previous mapping is completed, the mapping records the differences between the source and target volumes only. If the connection to both nodes in the I/O group that the mapping is assigned to is lost, the source and target volumes go offline.

Performing the cache flush that is required as part of the **startfcmap** or **startfcconsistgrp** command causes I/Os to be delayed while they are waiting for the cache flush to complete. To overcome this problem, FlashCopy supports the **prestartfcmap** or **prestartfcconsistgrp** commands, which prepare for a FlashCopy start while still allowing I/Os to continue to the source volume.

In the Preparing state, the FlashCopy mapping is prepared by completing the following steps:

1. Flushing any modified write data that is associated with the source volume from the cache. Read data for the source is left in the cache.
2. Placing the cache for the source volume into write-through mode so that subsequent writes wait until data is written to disk before the **write** command that is received from the host is complete.
3. Discarding any read or write data that is associated with the target volume from the cache.

Stopped

The mapping is stopped because you issued a **stop** command or an I/O error occurred. The target volume is offline and its data is lost. To access the target volume, you must restart or delete the mapping. The source volume is accessible and the read and write cache is enabled. If the mapping is incremental, the mapping is recording write operations to the source volume. If the connection to both nodes in the I/O group that the mapping is assigned to is lost, the source and target volumes go offline.

Stopping

The mapping is copying data to another mapping. If the background copy process is complete, the target volume is online while the stopping copy process completes. If the background copy process is not complete, data is discarded from the target volume cache. The target volume is offline while the stopping copy process runs. The source volume is accessible for I/O operations.

Suspended

The mapping started, but it did not complete. Access to the metadata is lost, which causes the source and target volume to go offline. When access to the metadata is restored, the mapping returns to the copying or stopping state and the source and target volumes return online. The background copy process resumes. Any data that was not flushed and was written to the source or target volume before the suspension is in cache until the mapping leaves the suspended state.

Summary of FlashCopy mapping states

Table 10-4 lists the various FlashCopy mapping states, and the corresponding states of the source and target volumes.

Table 10-4 FlashCopy mapping state summary

State	Source		Target	
	Online/Offline	Cache state	Online/Offline	Cache state
Idling/Copied	Online	Write-back	Online	Write-back
Copying	Online	Write-back	Online	Write-back
Stopped	Online	Write-back	Offline	N/A
Stopping	Online	Write-back	► Online if copy complete ► Offline if copy incomplete	N/A
Suspended	Offline	Write-back	Offline	N/A
Preparing	Online	Write-through	Online but not accessible	N/A
Prepared	Online	Write-through	Online but not accessible	N/A

10.3.12 Thin-provisioned FlashCopy

FlashCopy source and target volumes can be thin-provisioned.

Source or target thin-provisioned

The most common configuration is a fully allocated source and a thin-provisioned target. By using this configuration, the target uses a smaller amount of real storage than the source. With this configuration, use the NOCOPY (background copy rate = 0%) option only. Although the COPY option is supported, this option creates a fully allocated target, which defeats the purpose of thin provisioning.

Source and target thin-provisioned

When the source and target volumes are thin-provisioned, only the data that is allocated to the source is copied to the target. In this configuration, the background copy option has no effect.

Performance: The best performance is obtained when the grain size of the thin-provisioned volume is the same as the grain size of the FlashCopy mapping.

Thin-provisioned incremental FlashCopy

The implementation of thin-provisioned volumes does not preclude the use of incremental FlashCopy on the same volumes. It does not make sense to have a fully allocated source volume and then use incremental FlashCopy (which is always a full copy the first time) to copy this fully allocated source volume to a thin-provisioned target volume. However, this action is not prohibited.

Consider the following optional configurations:

- ▶ A thin-provisioned source volume can be copied incrementally by using FlashCopy to a thin-provisioned target volume. Whenever the FlashCopy is performed, only data that was modified is recopied to the target. If space is allocated on the target because of I/O to the target volume, this space is not reclaimed with subsequent FlashCopy operations.
- ▶ A fully allocated source volume can be copied incrementally by using FlashCopy to another fully allocated volume at the same time as it is being copied to multiple thin-provisioned targets (taken at separate points in time). By using this combination, a single full backup can be kept for recovery purposes, and the backup workload is separated from the production workload. At the same time, older thin-provisioned backups can be retained.

10.3.13 Background copy

With FlashCopy background copy enabled, the source volume data is copied to the corresponding target volume. With the FlashCopy background copy disabled, only data that changed on the source volume is copied to the target volume.

The benefit of using a FlashCopy mapping with background copy enabled is that the target volume becomes a real clone (independent from the source volume) of the FlashCopy mapping source volume after the copy is complete. When the background copy function is not performed, the target volume remains a valid copy of the source data only while the FlashCopy mapping remains in place.

The *background copy rate* is a property of a FlashCopy mapping that is defined as a value 0 - 100. The background copy rate can be defined and changed dynamically for individual FlashCopy mappings. A value of 0 disables the background copy.

Table 10-5 on page 488 shows the relationship of the background copy rate value to the attempted number of grains to be copied per second.

Table 10-5 Background copy rate

Value	Data copied per second	Grains per second (256 KB grain)	Grains per second (64 KB grain)
01 - 10	128 KiB	0.5	2
11 - 20	256 KiB	1	4
21 - 30	512 KiB	2	8
31 - 40	1 mebibyte (MiB)	4	16
41 - 50	2 MiB	8	32
51 - 60	4 MiB	16	64
61 - 70	8 MiB	32	128
71 - 80	16 MiB	64	256
81 - 90	32 MiB	128	512
91 - 100	64 MiB	256	1,024
101 - 110	128 MiB	512	2,048
111 - 120	256 MiB	1,024	4,096
121 - 130	512 MiB	2,048	8,192
131 - 140	1 GiB	4,096	16,384
141 - 150	2 GiB	8,192	32,768

The *grains per second* numbers represent the maximum number of grains that the IBM Storwize V5000 copies per second, assuming that the bandwidth to the managed disks (MDisks) can accommodate this rate.

If the IBM Storwize V5000 Gen2 system cannot achieve these copy rates because of insufficient width from the nodes to the MDisks, the background copy I/O contends for resources on an equal basis with the I/O that is arriving from the hosts. Background copy I/O and I/O that is arriving from the hosts tend to see an increase in latency and a consequential reduction in throughput.

Background copy and foreground I/O continue to make progress, and do not stop, hang, or cause the node to fail. The background copy is performed by both nodes of the I/O Group in which the source volume is found.

10.3.14 Serialization of I/O by FlashCopy

In general, the FlashCopy function in the IBM Spectrum Virtualize introduces no explicit serialization into the I/O path. Therefore, many concurrent I/Os are allowed to the source and target volumes.

However, there is a lock for each grain. The lock can be in shared or exclusive mode. For multiple targets, a common lock is shared, and the mappings are derived from a particular source volume. The lock is used in the following modes under the following conditions:

- The lock is held in shared mode during a read from the target volume, which touches a grain that was not copied from the source.

- ▶ The lock is held in exclusive mode while a grain is being copied from the source to the target.

If the lock is held in shared mode and another process wants to use the lock in shared mode, this request is granted unless a process is already waiting to use the lock in exclusive mode.

If the lock is held in shared mode and it is requested to be exclusive, the requesting process must wait until all holders of the shared lock free it.

Similarly, if the lock is held in exclusive mode, a process that is wanting to use the lock in shared or exclusive mode must wait for it to be freed.

10.3.15 Event handling

When a FlashCopy mapping is not copying or stopping, the FlashCopy function does not affect the handling or reporting of events for error conditions that are encountered in the I/O path. Event handling and reporting are affected only by FlashCopy when a FlashCopy mapping is copying or stopping; that is, actively moving data.

We describe these scenarios next.

Node failure

Normally, two copies of the FlashCopy bitmap are maintained. One copy of the FlashCopy bitmap is on each of the two nodes that make up the I/O Group of the source volume. When a node fails, one copy of the bitmap for all FlashCopy mappings whose source volume is a member of the failing node's I/O Group becomes inaccessible.

FlashCopy continues with a single copy of the FlashCopy bitmap that is stored as non-volatile in the remaining node in the source I/O Group. The system metadata is updated to indicate that the missing node no longer holds a current bitmap. When the failing node recovers or a replacement node is added to the I/O Group, the bitmap redundancy is restored.

Path failure (Path Offline state)

In a fully functioning system, all of the nodes have a software representation of every volume in the system within their application hierarchy.

Because the storage area network (SAN) that links IBM Storwize V5000 Gen2 system node canisters to each other and to the MDisks is made up of many independent links, it is possible for a subset of the nodes to be temporarily isolated from several of the MDisks. When this situation happens, the managed disks are said to be *path offline* on certain nodes.

Other nodes: Other nodes might see the managed disks as online because their connection to the managed disks is still functioning.

Path Offline for the source volume

If a FlashCopy mapping is in the copying state and the source volume goes path offline, this path offline state is propagated to all target volumes up to, but not including, the target volume for the newest mapping that is 100% copied but remains in the copying state. If no mappings are 100% copied, all of the target volumes are taken offline. Path offline is a state that exists on a per-node basis. Other nodes might not be affected. If the source volume comes online, the target and source volumes are brought back online.

Path Offline for the target volume

If a target volume goes path offline but the source volume is still online, and if there are any dependent mappings, those target volumes also go path offline. The source volume remains online.

10.3.16 Asynchronous notifications

FlashCopy raises informational event log entries for certain mapping and Consistency Group state transitions. These state transitions occur as a result of configuration events that complete asynchronously. The informational events can be used to generate Simple Network Management Protocol (SNMP) traps to notify the user.

Other configuration events complete synchronously, and no informational events are logged as a result of the following events:

▶ **PREPARE_COMPLETED**

This state transition is logged when the FlashCopy mapping or Consistency Group enters the prepared state as a result of a user request to prepare. The user can now start (or stop) the mapping or Consistency Group.

▶ **COPY_COMPLETED**

This state transition is logged when the FlashCopy mapping or Consistency Group enters the `idle_or_copied` state when it was in the copying or stopping state. This state transition indicates that the target disk now contains a complete copy and no longer depends on the source.

▶ **STOP_COMPLETED**

This state transition is logged when the FlashCopy mapping or Consistency Group enters the stopped state as a result of a user request to stop. It is logged after the automatic copy process completes. This state transition includes mappings where no copying needed to be performed. This state transition differs from the event that is logged when a mapping or group enters the stopped state as a result of an I/O error.

10.3.17 Interoperation with Metro Mirror and Global Mirror

A volume can be part of any copy relationship (FlashCopy, Metro Mirror, or Remote Mirror). Therefore, FlashCopy can work with Metro Mirror and Global Mirror to provide better protection of the data.

For example, Metro Mirror copy can be performed to duplicate data from Site_A to Site_B and then, perform a daily FlashCopy to back up the data to another location.

Note: If a volume is set to Transparent Cloud Tiering function, it cannot be part of FlashCopy, Metro Mirror, or Remote Mirror.

Table 10-6 on page 491 lists the supported combinations of FlashCopy and remote copy. In the table, *remote copy* refers to Metro Mirror and Global Mirror.

Table 10-6 FlashCopy and remote copy interaction

Component	Remote copy primary site	Remote copy secondary site
FlashCopy Source	Supported	Supported latency: When the FlashCopy relationship is in the preparing and prepared states, the cache at the remote copy secondary site operates in write-through mode. This process adds latency to the latent remote copy relationship.
FlashCopy Target	This is a supported combination and has the following restrictions: <ul style="list-style-type: none"> ▶ Issuing a stop -force might cause the remote copy relationship to be fully resynchronized. ▶ Code level must be 6.2.x or later. ▶ I/O Group must be the same. 	This is a supported combination with the major restriction that the FlashCopy mapping cannot be copying, stopping, or suspended. Otherwise, the restrictions are the same as at the remote copy primary site.

10.3.18 FlashCopy presets

The IBM Spectrum Virtualize GUI interface provides three FlashCopy presets (Snapshot, Clone, and Backup). Although these presets meet most FlashCopy requirements, they do not support all possible FlashCopy options. If more specialized options are required that are not supported by the presets, the options must be performed by using CLI commands.

This section describes the preset options and their use cases.

Snapshot

This preset creates a copy-on-write point-in-time copy. The snapshot is not intended to be an independent copy. Instead, the copy is used to maintain a view of the production data at the time that the snapshot is created. Therefore, the snapshot holds only the data from regions of the production volume that changed since the snapshot was created. Because the snapshot preset uses thin provisioning, only the capacity that is required for the changes is used.

Snapshot uses the following preset parameters:

- ▶ Background copy: None
- ▶ Incremental: No
- ▶ Delete after completion: No
- ▶ Cleaning rate: No
- ▶ Primary copy source pool: Target pool

Use case

The user wants to produce a copy of a volume without affecting the availability of the volume. The user does not anticipate many changes to be made to the source or target volume; a significant proportion of the volumes remains unchanged.

By ensuring that only changes require a copy of data to be made, the total amount of disk space that is required for the copy is reduced. Therefore, many snapshot copies can be used in the environment.

Snapshots are useful for providing protection against corruption or similar issues with the validity of the data, but they do not provide protection from physical controller failures. Snapshots can also provide a vehicle for performing repeatable testing (including “what-if” modeling that is based on production data) without requiring a full copy of the data to be provisioned.

Clone

The clone preset creates a replica of the volume, which can be changed without affecting the original volume. After the copy completes, the mapping that was created by the preset is automatically deleted.

Clone uses the following preset parameters:

- ▶ Background copy rate: 50
- ▶ Incremental: No
- ▶ Delete after completion: Yes
- ▶ Cleaning rate: 50
- ▶ Primary copy source pool: Target pool

Use case

Users want a copy of the volume that they can modify without affecting the original volume. After the clone is established, there is no expectation that it is refreshed or that there is any further need to reference the original production data again. If the source is thin-provisioned, the target is thin-provisioned for the auto-create target.

Backup

The backup preset creates a point-in-time replica of the production data. After the copy completes, the backup view can be refreshed from the production data, with minimal copying of data from the production volume to the backup volume.

Backup uses the following preset parameters:

- ▶ Background Copy rate: 50
- ▶ Incremental: Yes
- ▶ Delete after completion: No
- ▶ Cleaning rate: 50
- ▶ Primary copy source pool: Target pool

Use case

The user wants to create a copy of the volume that can be used as a backup if the source becomes unavailable, as in the case of loss of the underlying physical controller. The user plans to periodically update the secondary copy, and does not want to suffer from the resource demands of creating a copy each time (and incremental FlashCopy times are faster than full copy, which helps to reduce the window where the new backup is not yet fully effective). If the source is thin-provisioned, the target is also thin-provisioned in this option for the auto-create target.

Another use case, which is not supported by the name, is to create and maintain (periodically refresh) an independent image that can be subjected to intensive I/O (for example, data mining) without affecting the source volume’s performance.

10.4 Managing FlashCopy by using the GUI

It is often easier to work with the FlashCopy function from the GUI if you have a reasonable number of host mappings. However, in enterprise data centers with many host mappings, we suggest that you use the CLI to run your FlashCopy commands.

This section describes the tasks that you can perform at a FlashCopy level using the IBM Spectrum Virtualize GUI.

The following methods can be used to visualize and manage your FlashCopy:

- Open the GUI and move the mouse pointer over Copy Services menu and click FlashCopy, as shown in Figure 10-10.

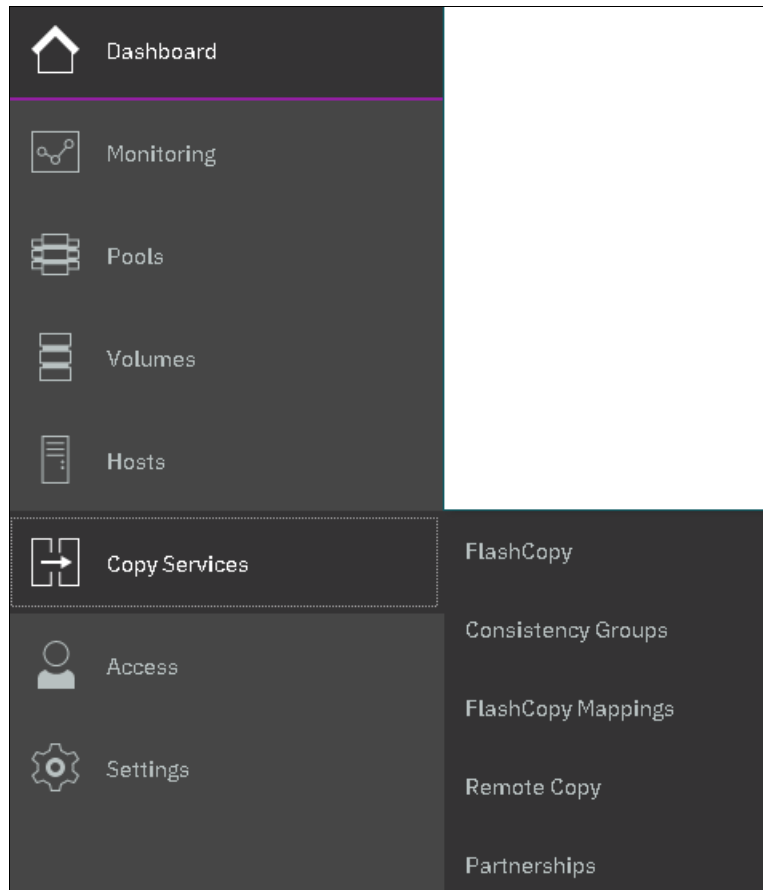


Figure 10-10 FlashCopy panel

- From the Copy Services option on the main panel, use the Consistency Groups option, as shown in Figure 10-11. A *Consistency Group* is a container for mappings. You can add many mappings to a Consistency Group.



Figure 10-11 Consistency Groups panel

- From the Copy Services option on the main panel, use the FlashCopy Mappings panel, as shown in Figure 10-12. A *FlashCopy mapping* defines the relationship between a source volume and a target volume.

Mapping Name	Status	Source Volume	Target Volume	Progress	Group	Flash Time
fomap0	✓ Copied	Linux01_01	Linux01_01_01	100%		Oct 22, 2018, 12:05:07 AM
fomap2	✓ Copied	Linux6	Linux01_01	100%	System x86 Server	

Figure 10-12 FlashCopy Mappings panel

10.4.1 Creating a FlashCopy mapping

In this section, FlashCopy mappings are created for volumes.

Complete the following steps:

1. From the main panel, move the mouse pointer over Copy Services click **FlashCopy**. The FlashCopy panel opens, as shown in Figure 10-13.

Volume Name	Status	Progress	Capacity	Group	Flash Time
ITSO-V8Kcs001			1.00 GiB		
ITSO-V8Kcs002-image			1.00 GiB		
Linux01			25.00 GiB		
Linux01_01			25.00 GiB		
Linux01_01_01	✓ Copied	100%			Oct 22, 2018, 12:05:07 AM
Linux01_01_01			25.00 GiB		
Linux02			25.00 GiB		

Figure 10-13 FlashCopy panel

2. Select the volume for which you want to create the FlashCopy relationship (see Figure 10-14).

Multiple FlashCopy mappings: To create multiple FlashCopy mappings at once, select multiple volumes by holding down Ctrl and clicking the entries that you want.

Volume Name	Status	Progress	Capacity	Group	Flash Time
ITSO-V8Kcs001			1.00 GiB		
ITSO-V8Kcs002-image			1.00 GiB		
Linux01			25.00 GiB		
Linux01_01			25.00 GiB		
Linux01_01_01			25.00 GiB		
Linux02			25.00 GiB		

Figure 10-14 FlashCopy mapping: Select the volume (or volumes)

Depending on whether you created the target volumes for your FlashCopy mappings or you want the system to create the target volumes for you, the following options are available:

- If you created the target volumes, see “Using existing target volumes” on page 495.
- If you want the system to create the target volumes for you, see “Creating target volumes” on page 498.

Using existing target volumes

Complete the following steps to use existing target volumes for the FlashCopy mappings:

1. Select the source volume that you want to use. Then, click **Actions** → **Advanced FlashCopy** → **Use Existing Target Volumes**, as shown in Figure 10-15.

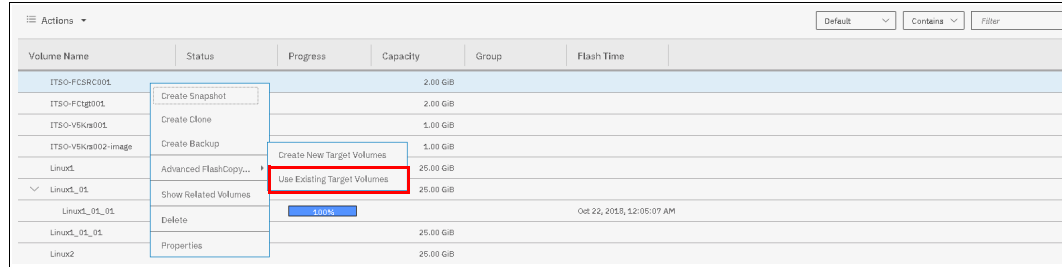


Figure 10-15 Using existing target volumes

2. The Create FlashCopy Mapping window opens (see Figure 10-16). In this window, select the target volume and then, click **Add**.

Important: The source volume and the target volume must be of equal size. Therefore, only targets of the same size are shown in the list for a source volume.

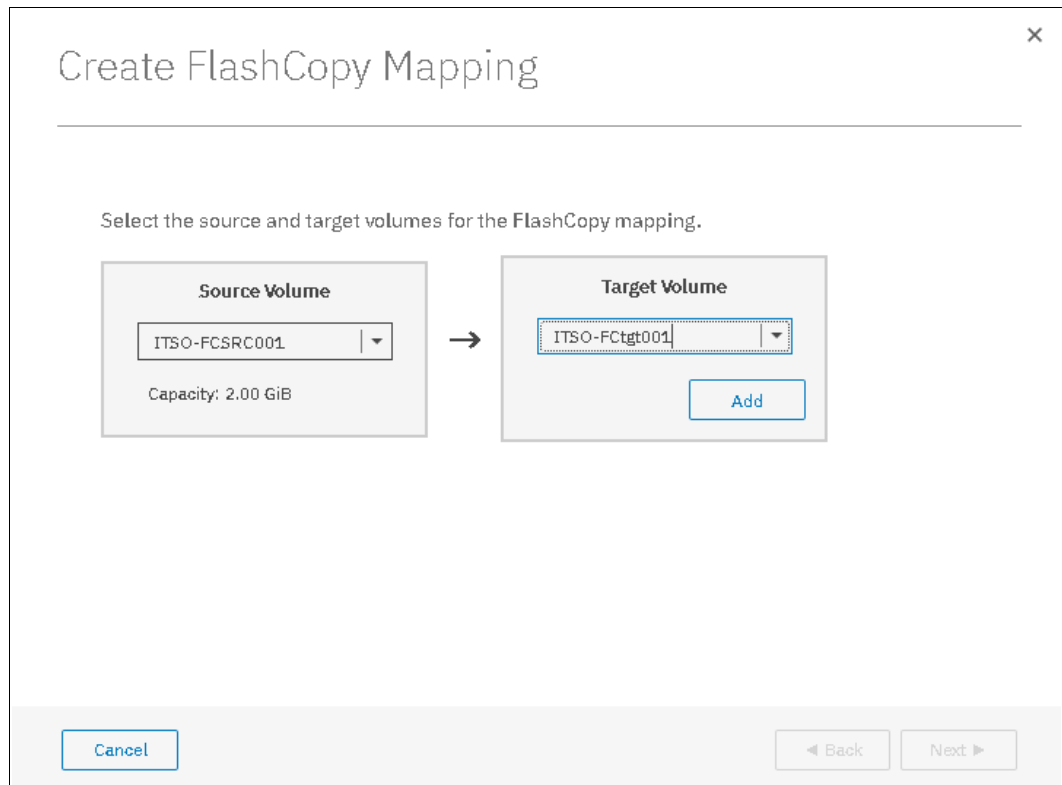


Figure 10-16 Create a FlashCopy Mapping by using an existing target volume

To remove a relationship that was created, click **X**, as shown in Figure 10-17 on page 496.

3. Click **Next** after you create all of the relationships that you need, as shown in Figure 10-17.

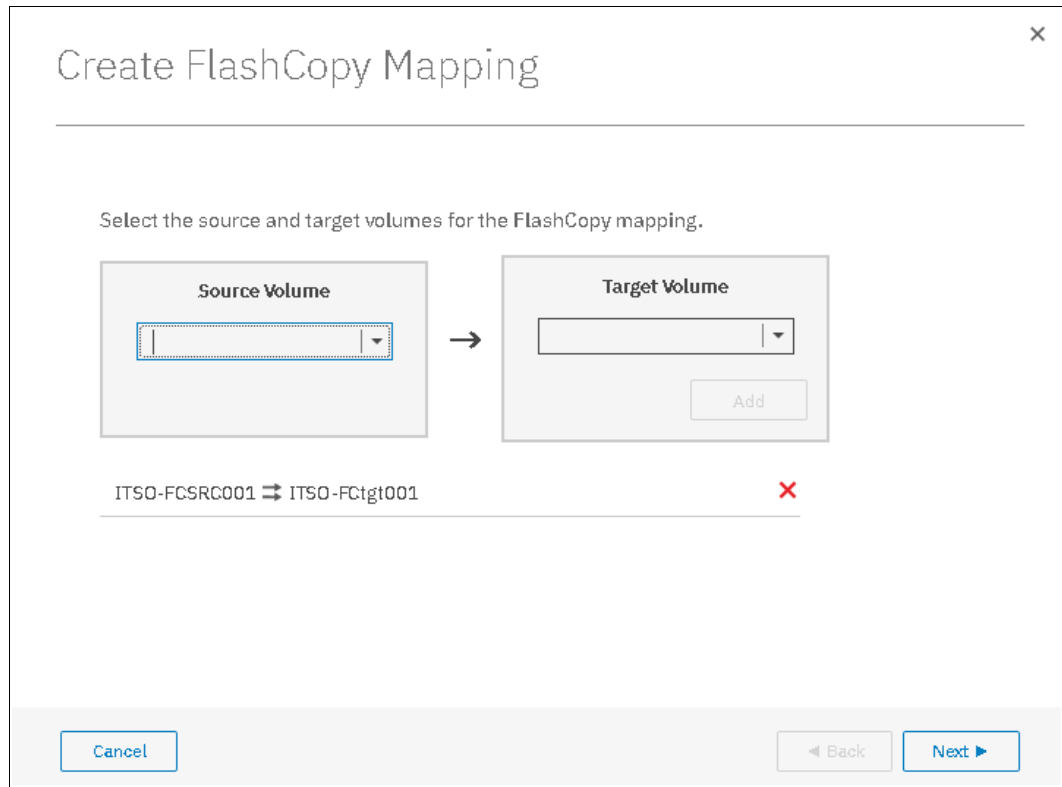


Figure 10-17 Create FlashCopy Mapping window

4. In the next window, select one FlashCopy preset. The GUI provides the following presets to simplify common FlashCopy operations.
 - Snapshot: Creates a copy-on-write point-in-time copy.
 - Clone: Creates a replica of the source volume on a target volume. The copy can be changed without affecting the original volume.
 - Backup: Creates a FlashCopy mapping that can be used to recover data or objects if the system experiences data loss. These backups can be copied multiple times from source and target volumes.

5. For each preset, you can customize various advanced options. You can access these settings by clicking the preset. The preset options are shown in Figure 10-18.

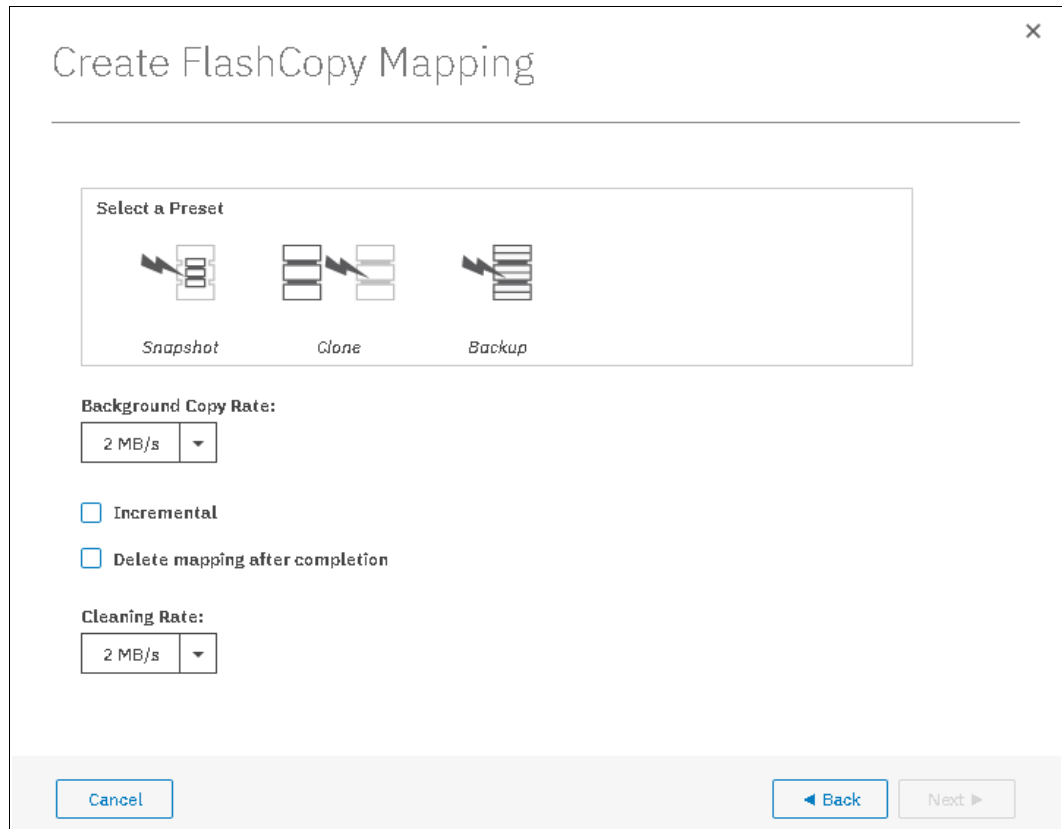


Figure 10-18 Create FlashCopy Mapping Presets

If you prefer not to customize these settings, go directly to step 7.

You can customize the following advanced setting options, as shown in Figure 10-18.

6. After you complete your modifications, click **Next**.
7. Choose whether to add the mappings to a Consistency Group, as shown in Figure 10-19 on page 498. Click **Finish**.

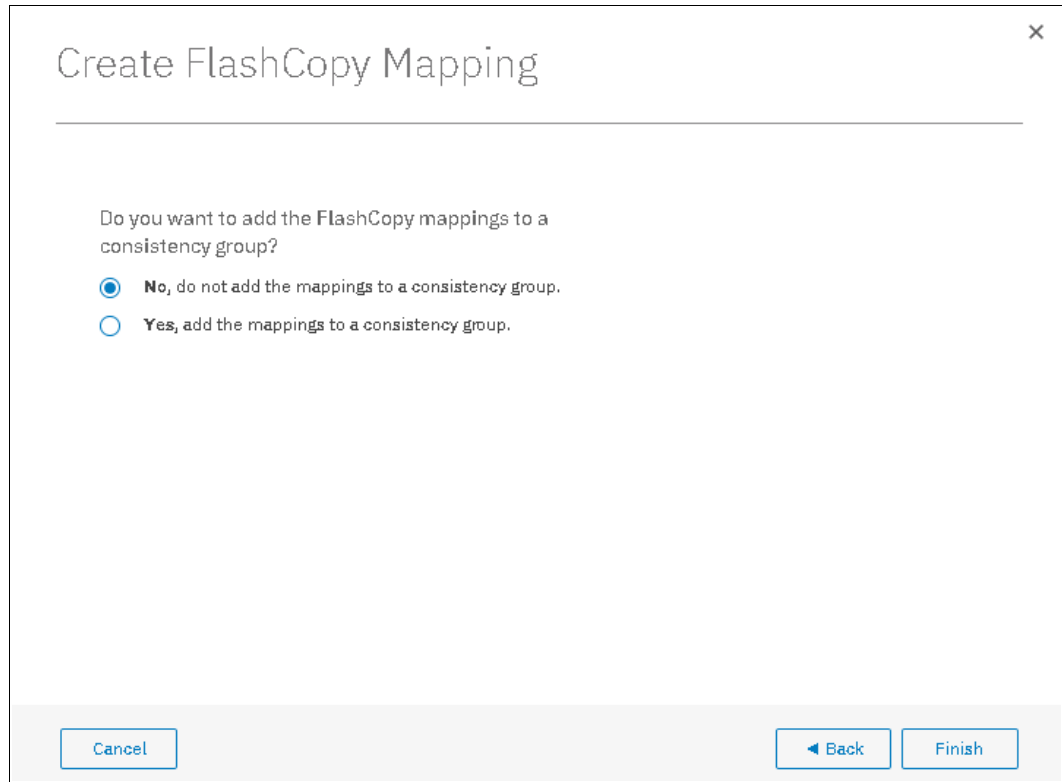


Figure 10-19 Add the mappings to a Consistency Group

8. Check the result of this FlashCopy mapping. From the main panel, click **Copy Services** → **FlashCopy Mappings**, as shown in Figure 10-20.

Mapping Name	Status	Source Volume	Target Volume	Progress	Group	Flash Time
fcmapp0	Copied	Linux01_01	Linux01_01_01	100%		Oct 22, 2015, 12:05:07 AM
fcmapp1	Idle	ITSO-FCSRC001	ITSO-FCtgt001	0%		
fcmapp2	Copied	Linux05	Linux05_01	100%	System x86 Server	

Figure 10-20 FlashCopy mappings

9. For each FlashCopy mapping relationship that was created, a mapping name is automatically generated that starts with fcmappX, where X is the next available number. If needed, you can rename these mappings, as described in 10.4.11, “Renaming FlashCopy mapping” on page 515.

Creating target volumes

Complete the following steps to create target volumes for FlashCopy mapping:

1. Select the source volume by left-click; then, click **Actions** → **Advanced FlashCopy** → **Create New Target Volumes**, as shown in Figure 10-21 on page 499.

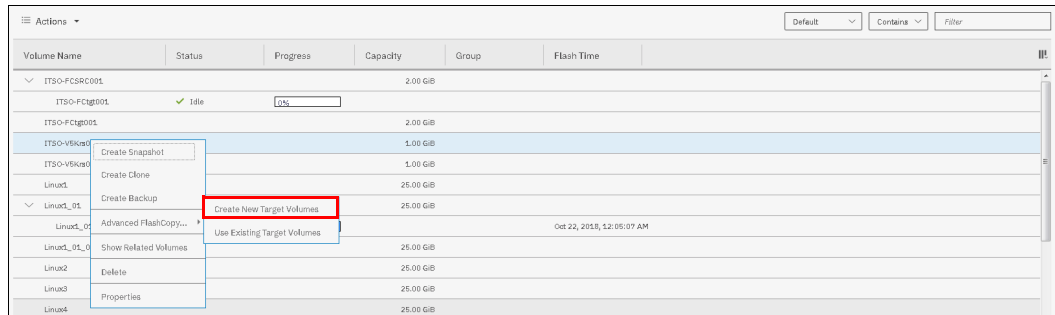


Figure 10-21 Create new target volumes

Target volume naming: If the target volume does not exist, the target volume is created. The target volume name is based on its source volume and a generated number at the end; for example, source_volume_name_XX, where XX is a number that was generated dynamically.

- In the Create FlashCopy Mapping window (see Figure 10-22), you must select one FlashCopy preset where advanced options can be configured as shown.

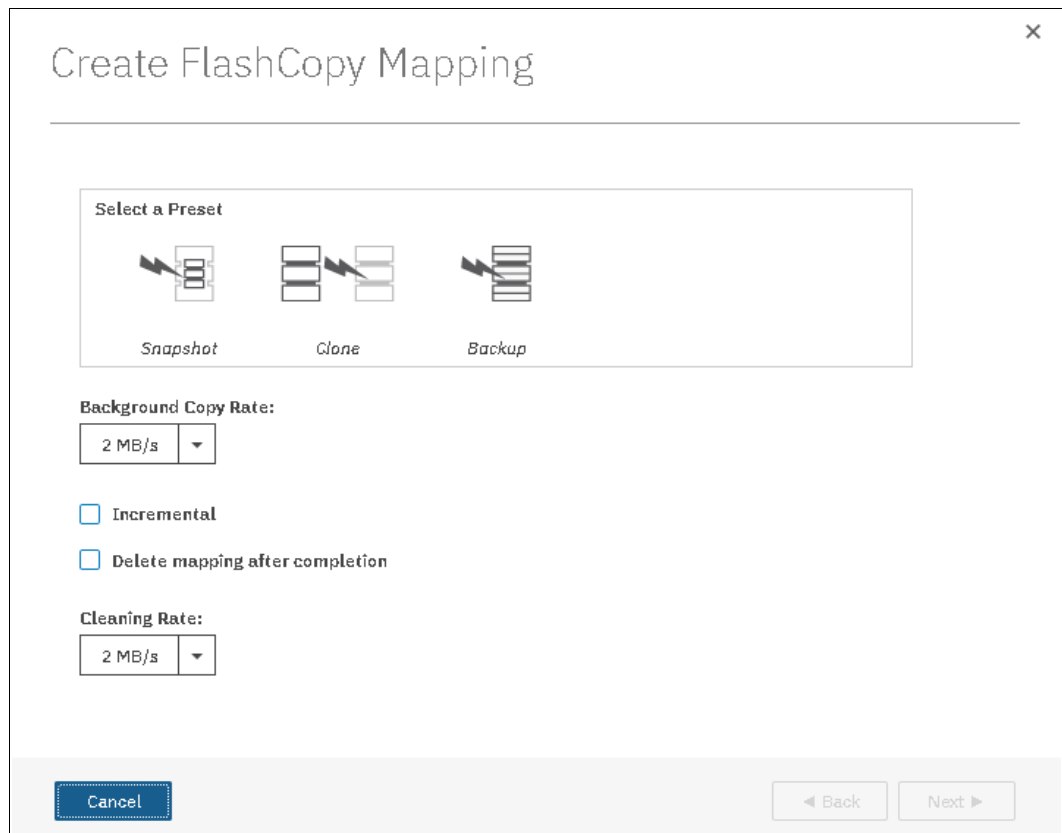


Figure 10-22 Create FlashCopy Mapping window

3. You can choose whether to add this FlashCopy mapping to a Consistency Group as shown in Figure 10-23. Click **Next**.

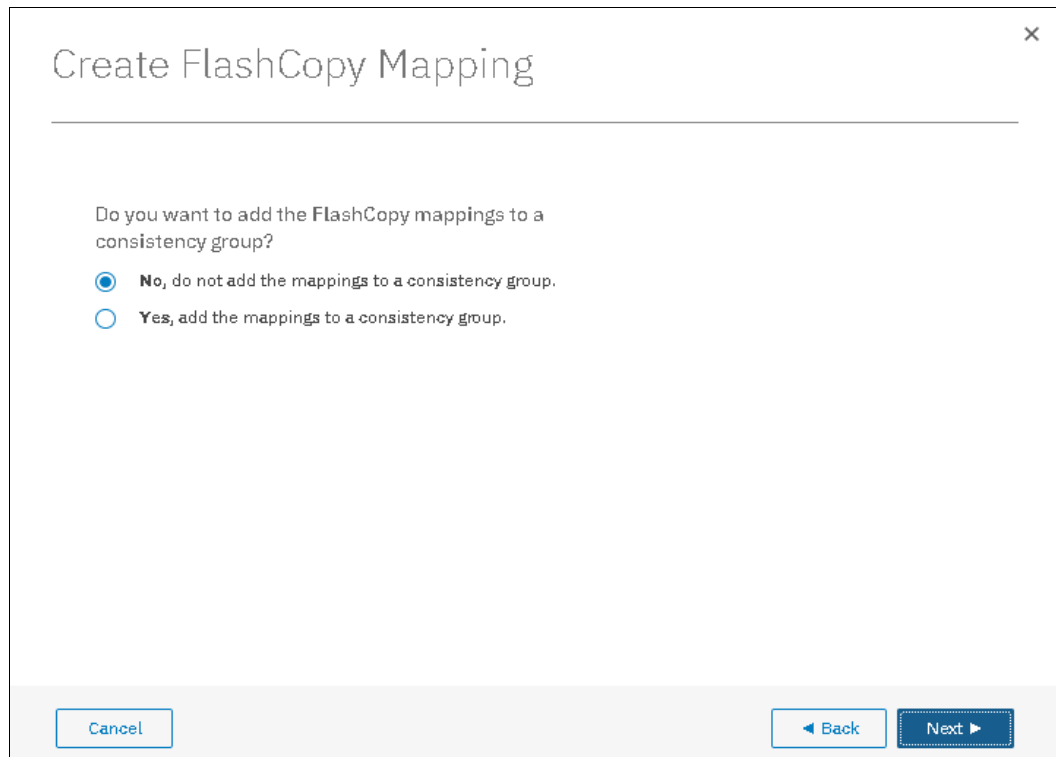


Figure 10-23 Creating FlashCopy mapping

4. In the next window, select the pool that is used to create the target, as shown in Figure 10-24 on page 501. Click **Next**.

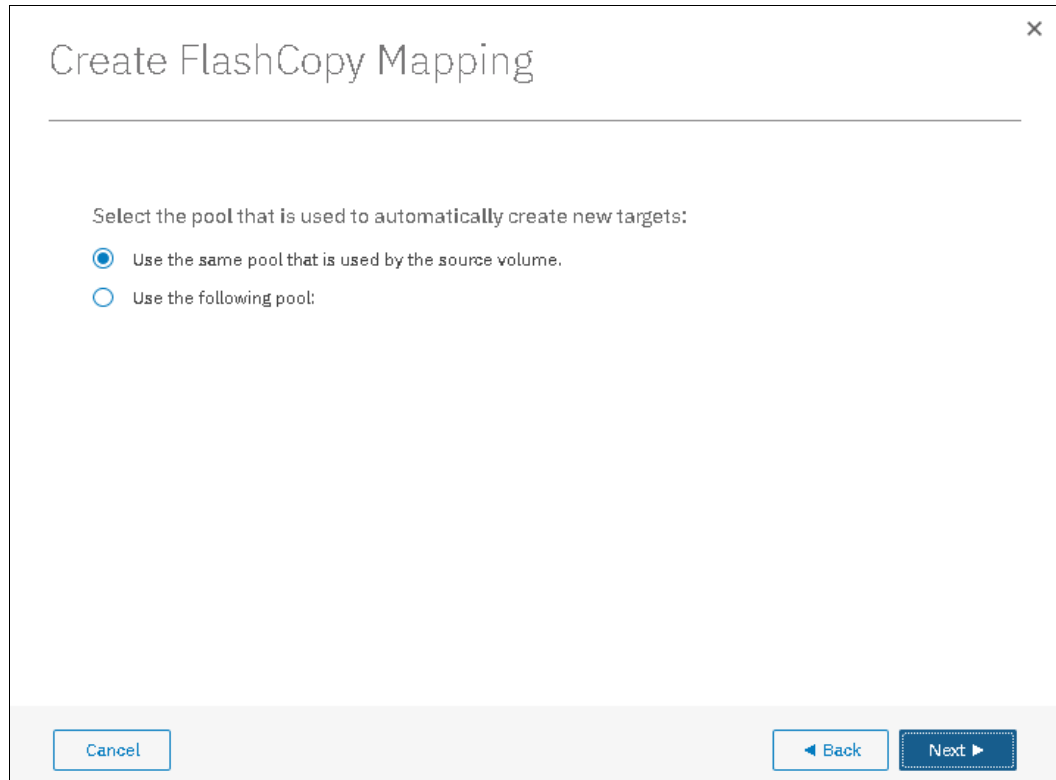


Figure 10-24 Select pool to create new target

5. In the next window, select the volume type, as shown in Figure 10-25.

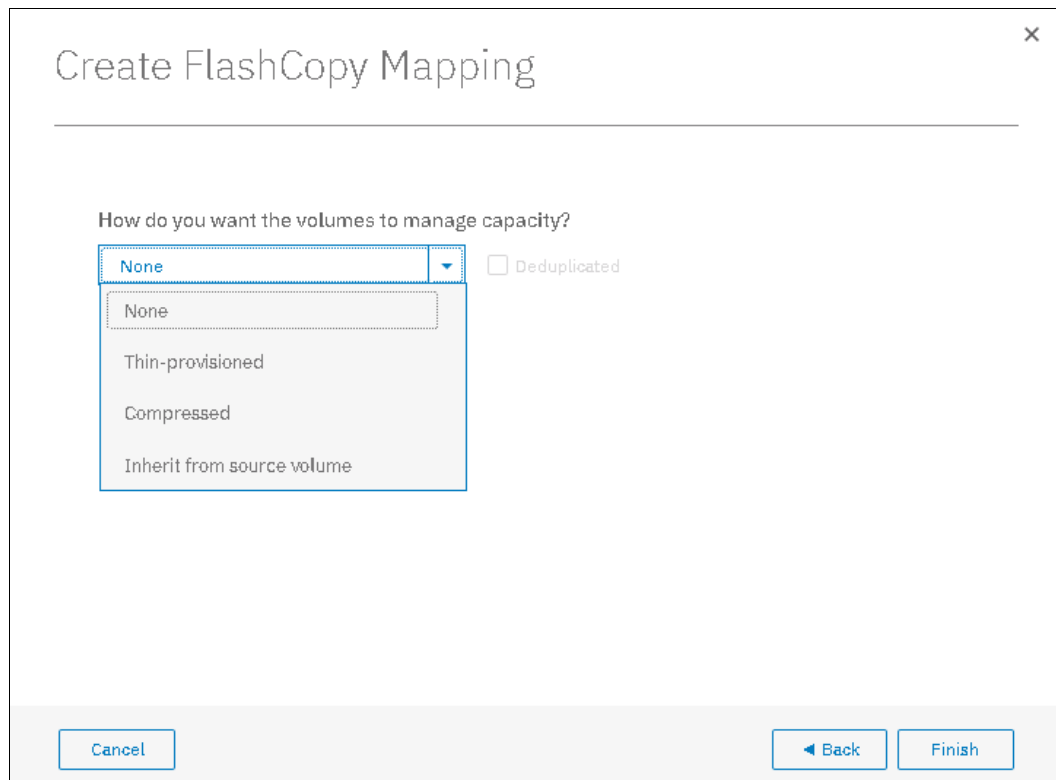


Figure 10-25 Select volume type

The volumes can be managed as None, Thin-provisioned, Compressed or Inherit from source volume.

6. Click **Finish** to run the new FlashCopy mapping creation.
7. Check the result of this FlashCopy mapping, as shown in Figure 10-26. For each FlashCopy mapping relationship that is created, a mapping name is automatically generated that starts with `fcmapiX` where `X` is the next available number.

Volume Name	Status	Progress	Capacity	Group	Flash Time
ITSO-FCsrc001	Idle	100%	2.00 GiB		
ITSO-FCtgt001			2.00 GiB		
ITSO-VSkra001			1.00 GiB		
ITSO-VSkra001_01			1.00 GiB		
ITSO-VSkra002-image			1.00 GiB		

Figure 10-26 FlashCopy mappings

The FlashCopy mapping is ready for use.

Tip: You can start FlashCopy from the GUI. However, the use of the GUI might be impractical if you plan to handle many FlashCopy mappings or Consistency Groups periodically. In these cases, creating a script by using the CLI might be more convenient.

10.4.2 Single-click snapshot

The *snapshot* creates a point-in-time backup of production data. The snapshot is not intended to be an independent copy. Instead, it is used to maintain a view of the production data at the time that the snapshot is created. Therefore, the snapshot holds only the data from regions of the production volume that changed since the snapshot was created. Because the snapshot preset uses thin provisioning, only the capacity that is required for the changes is used.

Snapshot uses the following preset parameters:

- ▶ Background copy: No
- ▶ Incremental: No
- ▶ Delete after completion: No
- ▶ Cleaning rate: No
- ▶ Primary copy source pool: Target pool

To create and start a snapshot, complete the following steps:

1. From the main panel, click **Copy Services** → **FlashCopy**.
2. Select the volume that you want to create a snapshot of and right-click **Create Snapshot**, as shown in Figure 10-27.

Volume Name	Status	Progress	Capacity	Group	Flash Time
ITSO-FCsrc001	Idle	100%	2.00 GiB		
ITSO-FCtgt001			2.00 GiB		
ITSO-S3C-R801			2.00 GiB		
ITSO-VSkra001			1.00 GiB		
ITSO-VSkra001			1.00 GiB		
ITSO-VSkra002			1.00 GiB		
Linux2			25.00 GiB		
Linux2_01			25.00 GiB		
Linux2_01_01			25.00 GiB		Oct 22, 2018, 12:05:07 AM
Linux2_01_01			25.00 GiB		
Linux2			25.00 GiB		

Figure 10-27 Create FlashCopy Snapshot

A volume is created as a target volume for this snapshot in the same pool as the source volume. The FlashCopy mapping is created and started.

3. You can check the FlashCopy progress in the Progress column Status area, as shown in Figure 10-28.

Volume Name	Status	Progress	Capacity	Group	Flash Time
ITSO-FCSRC001			2.00 GiB		
ITSO-FCTg001	Idle	0%			
ITSO-FCTg001			2.00 GiB		
ITSO-S3C-RS01			2.00 GiB		
ITSO-S3C-RS01_01	Copying	0%			Oct 24, 2015, 4:06:51 PM
ITSO-S3C-RS01_01			2.00 GiB		

Figure 10-28 Snapshot created and started

10.4.3 Single-click clone

The *clone preset* creates an exact replica of the volume, which can be changed without affecting the original volume. After the copy completes, the mapping that was created by the preset is automatically deleted.

The clone preset uses the following parameters:

- ▶ Background copy rate: 50
- ▶ Incremental: No
- ▶ Delete after completion: Yes
- ▶ Cleaning rate: 50
- ▶ Primary copy source pool: Target pool

To create and start a clone, complete the following steps:

1. From the main panel, click **Copy Services** → **FlashCopy**.
2. Select the volume that you want to clone, **Right-Click** → **Create Clone**, as shown in Figure 10-29.

Volume Name	Status	Progress	Capacity	Group	Flash Time
ITSO-FCSRC001			2.00 GiB		
ITSO-FCTg001	Idle	0%			
ITSO-FCTg001			2.00 GiB		
ITSO-S3C-RS01			2.00 GiB		
ITSO-S3C-RS01_01	Copying	0%			Oct 24, 2015, 4:06:51 PM
ITSO-S3C-RS01_01			2.00 GiB		
ITSO-S3C-RS02CL			2.00 GiB		
ITSO-V8Kra001			1.00 GiB		
ITSO-V8Kra001_01			1.00 GiB		
ITSO-V8Kra002-image			1.00 GiB		
Linux1			25.00 GiB		
Linux1_01			25.00 GiB		
Linux1_01_01		100%			Oct 22, 2015, 12:05:07 AM
Linux1_01_01			25.00 GiB		
Linux2			25.00 GiB		

Figure 10-29 Create Clone option

3. A volume is created as a target volume for this clone in the same pool as the source volume. The FlashCopy mapping is created and started. You can check the FlashCopy progress in the FlashCopy mappings option or by clicking **Monitoring** → **Background Tasks** from the main panel.

After the FlashCopy clone is created, the mapping is removed and the new cloned volume becomes available, as shown in Figure 10-30.

Volume Name	Status	Progress	Capacity	Group	Flash Time
ITSO-S3C-RS03CL			1.00 MiB		
ITSO-S3C-RS03CL_01			1.00 MiB		

Figure 10-30 Clone created and FlashCopy relationship removed

10.4.4 Single-click backup

The backup creates a point-in-time replica of the production data. After the copy completes, the backup view can be refreshed from the production data, with minimal data copied from the production volume to the backup volume. The backup preset uses the following parameters:

- ▶ Background Copy rate: 50
- ▶ Incremental: Yes
- ▶ Delete after completion: No
- ▶ Cleaning rate: 50
- ▶ Primary copy source pool: Target pool

To create and start a backup, complete the following steps:

1. From the main panel, click **Copy Services** → **FlashCopy**.
2. Select the volume that you want to back up, and **Right-click** → **Create Backup**, as shown in Figure 10-31.

Volume Name	Status	Progress	Capacity	Group	Flash Time
ITSO-FC3RC001			2.00 GiB		
ITSO-FCag001	Idle	0%			
ITSO-FCag001			2.00 GiB		
ITSO-S3C-004			1.00 MiB		
ITSO-S3C-RS01			2.00 GiB		
ITSO-S3C-RS01_01		0%			Oct 24, 2018, 4:06:51 PM
ITSO-S3C-RS02CL			2.00 GiB		
ITSO-S3C-RS02CL			2.00 GiB		
ITSO-S3C-RS02		45%			Oct 24, 2018, 4:11:06 PM
ITSO-S3C-RS02CL			2.00 GiB		
ITSO-S3C-RS03CL			1.00 MiB		
ITSO-S3C-RS03CL			1.00 MiB		

Figure 10-31 Create Backup option

A volume is created as a target volume for this backup in the same pool as the source volume. The FlashCopy mapping is created and started.

3. You can check the FlashCopy progress in the Progress column, as shown in Figure 10-32, or in the Running Tasks from the main panel.

Volume Name	Status	Progress	Capacity	Group	Flash Time
ITSO-S3C-004			1.00 MiB		
ITSO-S3C-004_01	Copied	100%			Oct 24, 2018, 4:21:29 PM
ITSO-S3C-004_01			1.00 MiB		

Figure 10-32 Backup created and started

10.4.5 Creating a FlashCopy Consistency Group

To create a FlashCopy Consistency Group in the GUI, complete the following steps:

1. From the main panel, click **Copy Services** → **Consistency Group**. The Consistency Group panel is shown in Figure 10-33.

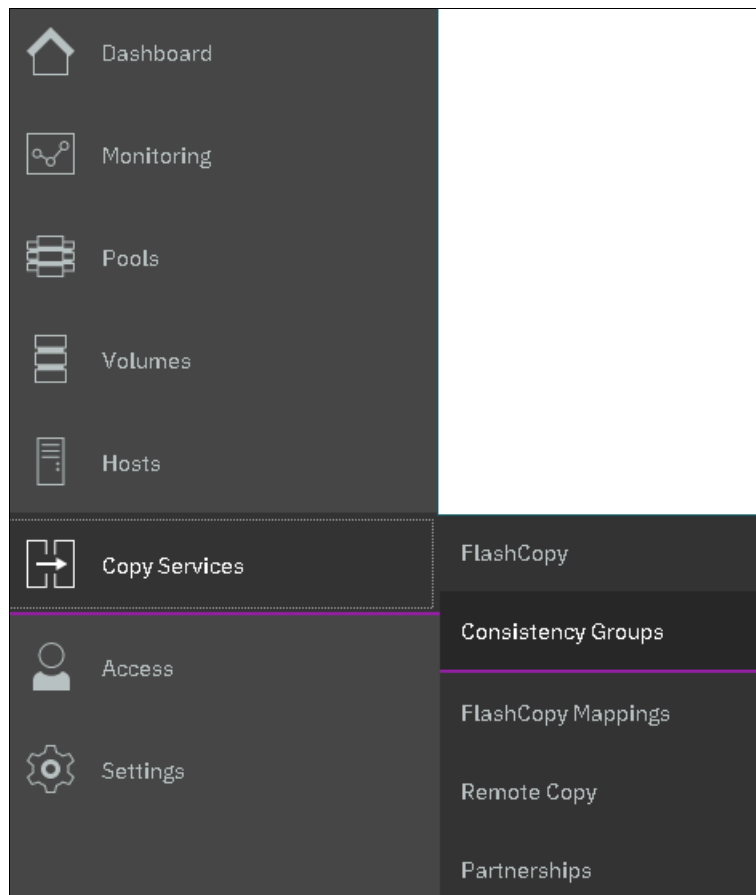


Figure 10-33 FlashCopy Consistency Group

2. Click **Create Consistency Group** and enter the FlashCopy Consistency Group name that you want to use. Then, click **Create** (see Figure 10-34).

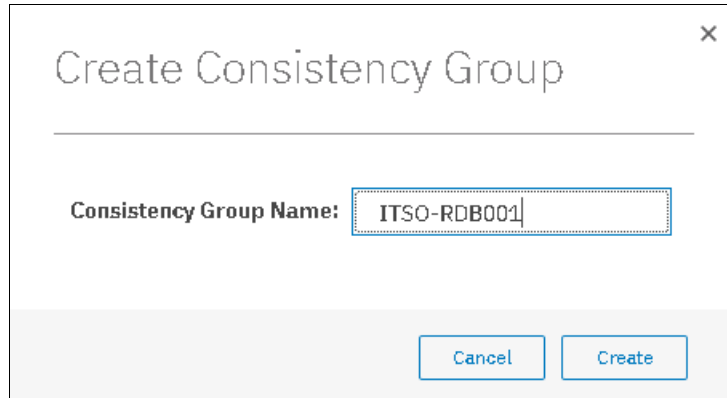


Figure 10-34 Create Consistency Group window

Consistency Group name: You can use the letters A - Z and a - z, the numbers 0 - 9, and the underscore (_) character. The volume name can be 1 - 63 characters.

10.4.6 Creating FlashCopy mappings in a Consistency Group

This section describes how to create FlashCopy mappings for volumes and their related targets. The source and target volumes were created before this operation.

Complete the following steps:

1. From the main panel, click **Copy Services** → **Consistency Group**. The Consistency Groups panel opens, as shown in Figure 10-35.



Figure 10-35 Consistency Group panel

2. In this window, click **Actions** → **Create FlashCopy Mappings**, as shown in Figure 10-36.



Figure 10-36 Create FlashCopy mappings

3. The Create FlashCopy Mapping window opens, as shown in Figure 10-37. In this window, you must create the relationships between the source volumes and the target volumes.

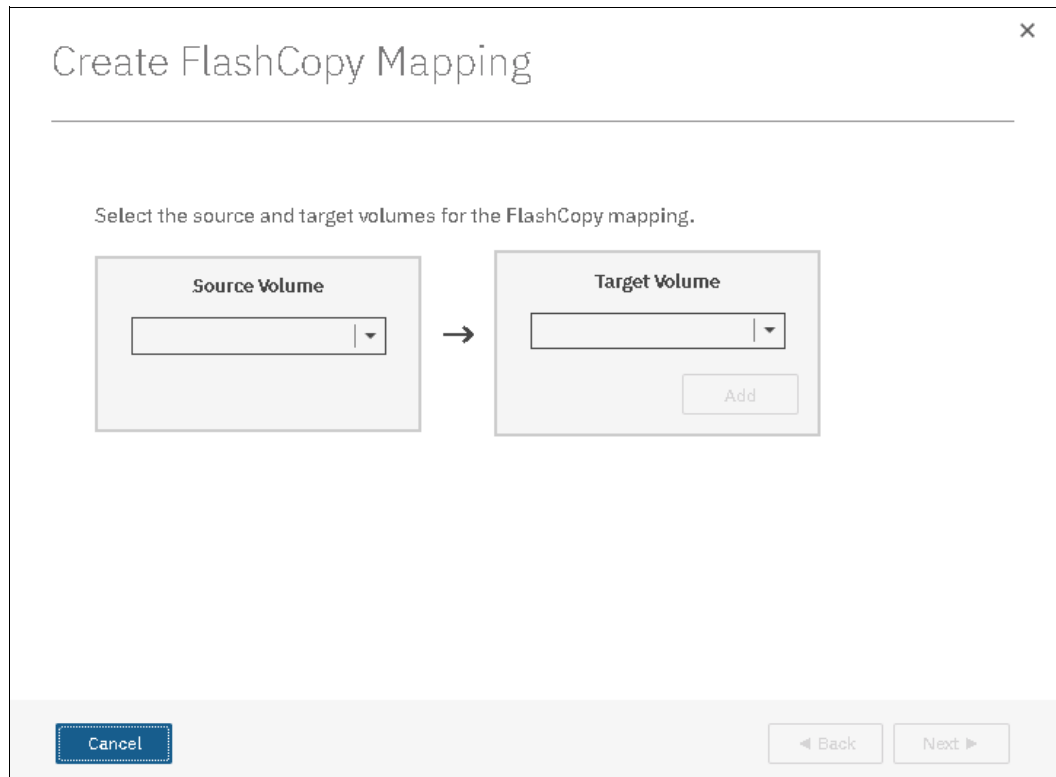


Figure 10-37 FlashCopy mapping creation

4. Select a volume in the Source Volume column by using the drop-down list. Then, select a volume in the Target Volume column by using the drop-down list. Click **Add**, as shown in Figure 10-38. Repeat this step to create other relationships.

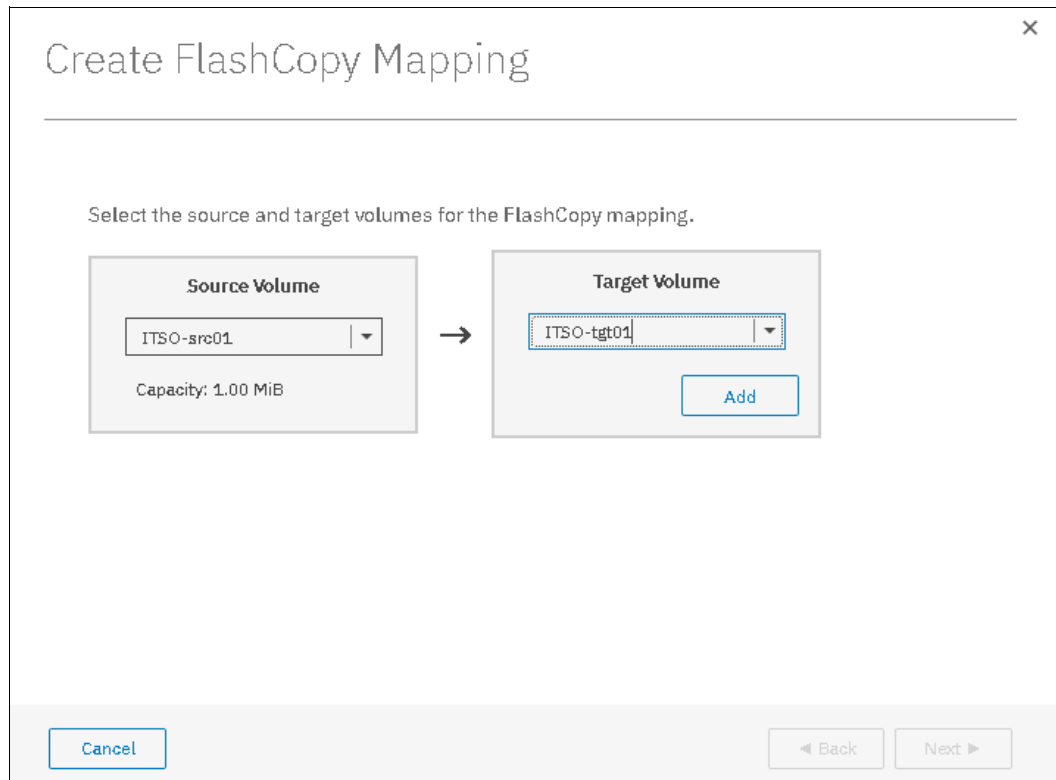
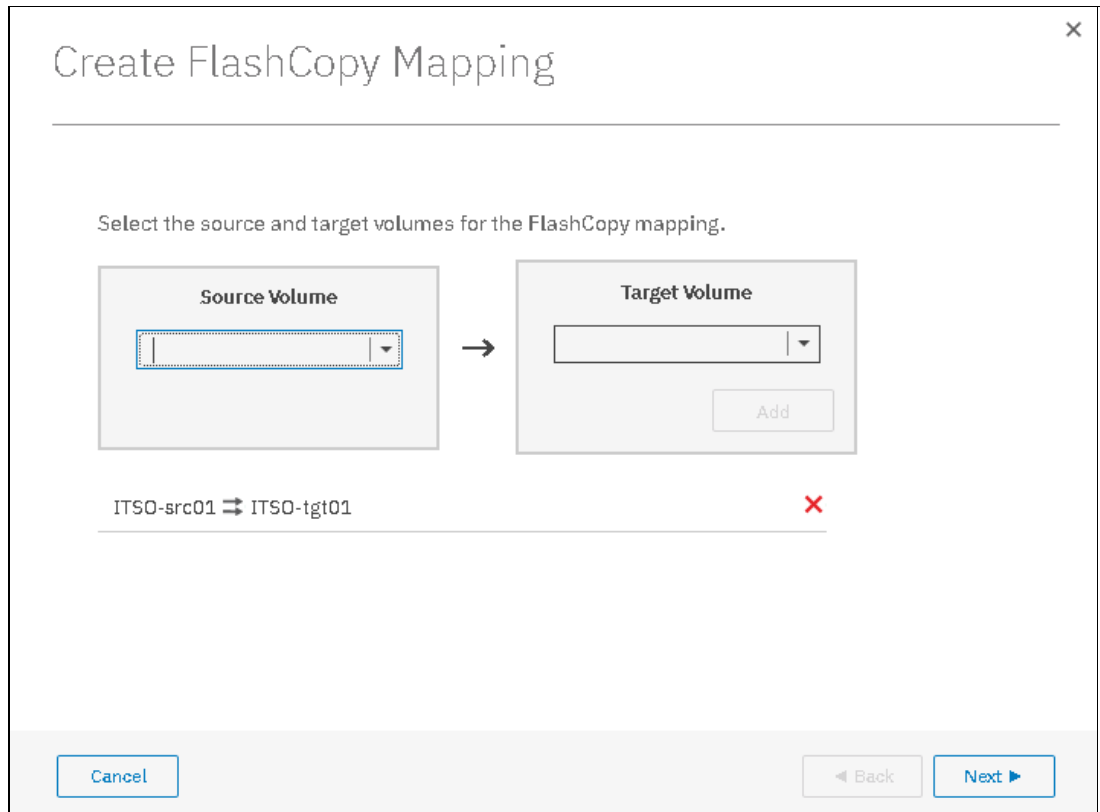


Figure 10-38 Source and target volumes

5. To remove a relationship that was created, click **X**.

6. Click **Next** after all of the relationships that you want to create are shown (see Figure 10-39).



7. In the next window, you must select one FlashCopy preset along with their customization options. The GUI provides the following presets to simplify common FlashCopy operations, as shown in Figure 10-40 on page 510:
- Snapshot: Creates a copy-on-write point-in-time copy.
 - Clone: Creates a replica of the source volume on a target volume. The copy can be changed without affecting the original volume.
 - Backup: Creates a FlashCopy mapping that can be used to recover data or objects if the system experiences data loss. These backups can be copied multiple times from the source and target volumes.

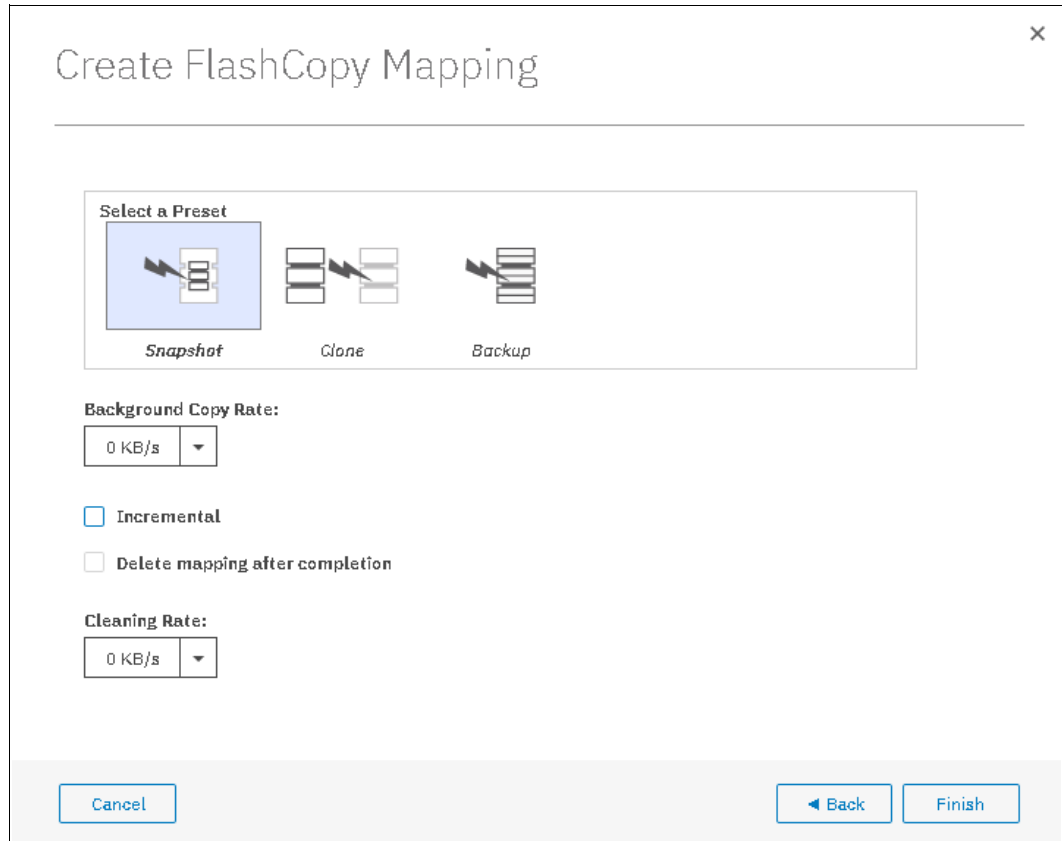


Figure 10-40 Create FlashCopy Mapping window

Select the preset. Based on that selection, you can customize options that are based on the selected preset. Click **Finish**.

8. Check the result of this FlashCopy mapping in the Consistency Groups window, as shown in Figure 10-41.

For each FlashCopy mapping relationship that you created, a mapping name is automatically generated that starts with `fmapX` where `X` is an available number.



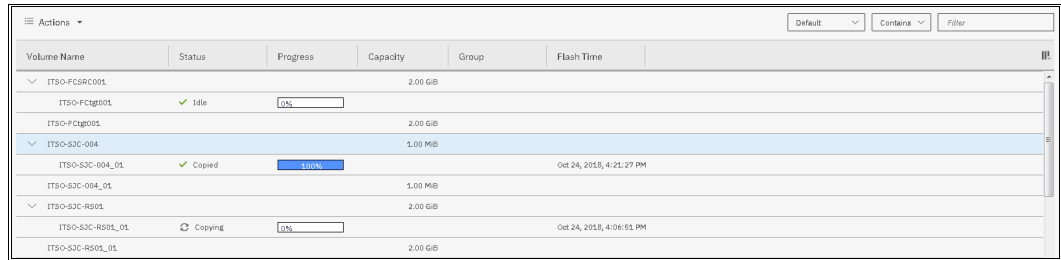
Figure 10-41 Create FlashCopy mappings result

Tip: You can start FlashCopy from the IBM Spectrum Virtualize GUI. However, if you plan to handle many FlashCopy mappings or Consistency Groups periodically, or at varying times, creating a script by using the operating system shell CLI might be more convenient.

10.4.7 Showing related volumes

Complete the following steps to show related volumes for a specific FlashCopy mapping:

1. From main panel, click **Copy Services** → **FlashCopy**, as shown in Figure 10-42.



Volume Name	Status	Progress	Capacity	Group	Flash Time
ITSO-FCSRC001			2.00 GiB		
ITSO-FCtg001	✓ Idle	0%			
ITSO-FCtg001			2.00 GiB		
ITSO-SJC-004			1.00 MiB		
ITSO-SJC-004_01	✓ Copied	100%	1.00 MiB		Oct 24, 2018, 4:21:27 PM
ITSO-SJC-004_01			1.00 MiB		
ITSO-SJC-RS01			2.00 GiB		
ITSO-SJC-RS01_01	⊞ Copying	0%	2.00 GiB		Oct 24, 2018, 4:06:51 PM
ITSO-SJC-RS01_01			2.00 GiB		

Figure 10-42 Volumes

2. Select the volume or the FlashCopy mapping to show the related volumes, as shown in Figure 10-43.

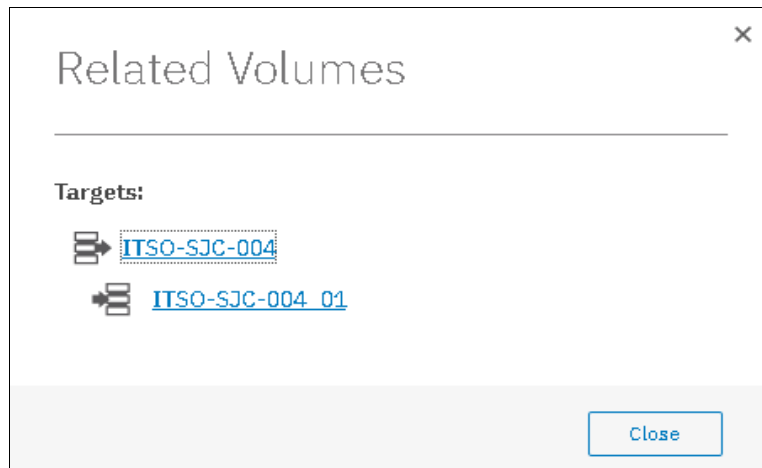


Figure 10-43 Show Related Volumes

10.4.8 Moving a FlashCopy mapping to a Consistency Group

Complete the following steps to move a FlashCopy mapping to the Consistency Group:

1. From the main panel, click **Copy Services** → **FlashCopy** or **Consistency Groups**, or **FlashCopy Mappings**.
2. Select the FlashCopy mapping that you want to move to a Consistency Group or the FlashCopy mapping for which you want to change the Consistency Group.
3. Click the FlashCopy mapping relationship to move to a consistency group, as shown in Figure 10-44 on page 512.

Tip: You can also right-click a FlashCopy mapping and select **Move to Consistency Group**.

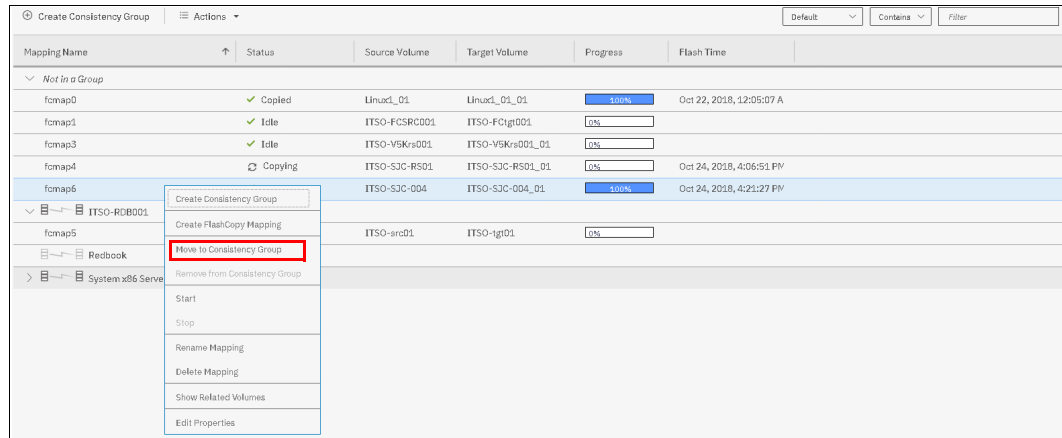


Figure 10-44 Move to Consistency Group action

- In the Move FlashCopy Mapping to Consistency Group window, select the Consistency Group for this FlashCopy mapping by using the drop-down list (see Figure 10-45).

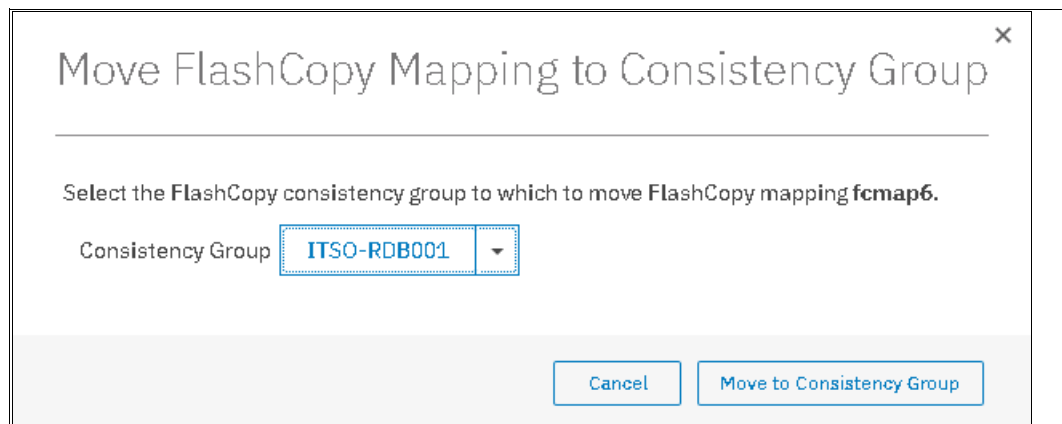


Figure 10-45 Move FlashCopy mapping to Consistency Group window

- Click **Move to Consistency Group** to confirm your changes.

10.4.9 Removing a FlashCopy mapping from a Consistency Group

Complete the following steps to remove a FlashCopy mapping from a Consistency Group:

- From the main panel, click **Copy Services** → **FlashCopy** or **Consistency Groups**, or **FlashCopy Mappings**.
- Select the FlashCopy mapping that you want to remove from a Consistency Group.
- Right-click the FlashCopy relationship mapping to be removed, as shown in Figure 10-46 on page 513.

Tip: You can also right-click a FlashCopy mapping and select **Remove from Consistency Group**.

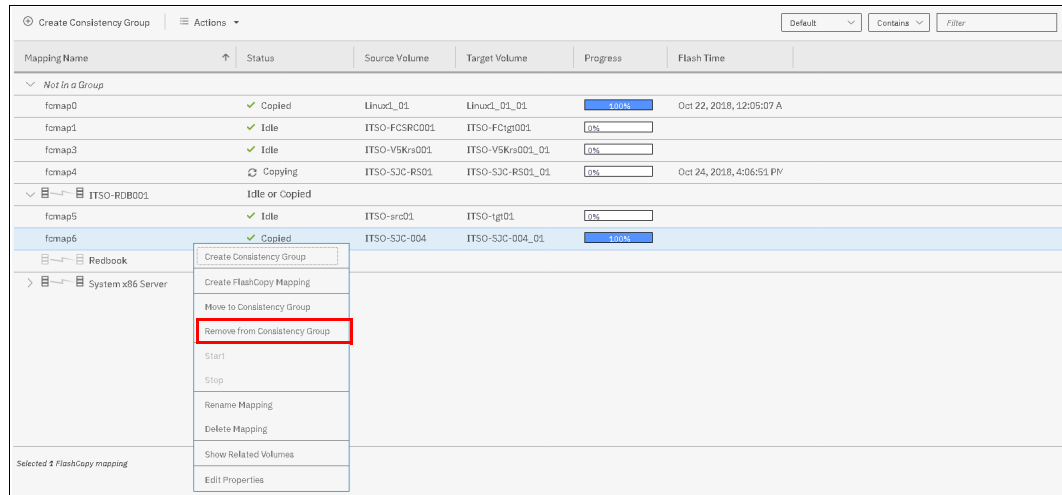


Figure 10-46 Remove from Consistency Group action

- In the Remove FlashCopy Mapping from Consistency Group window, click **Remove**, as shown in Figure 10-47.

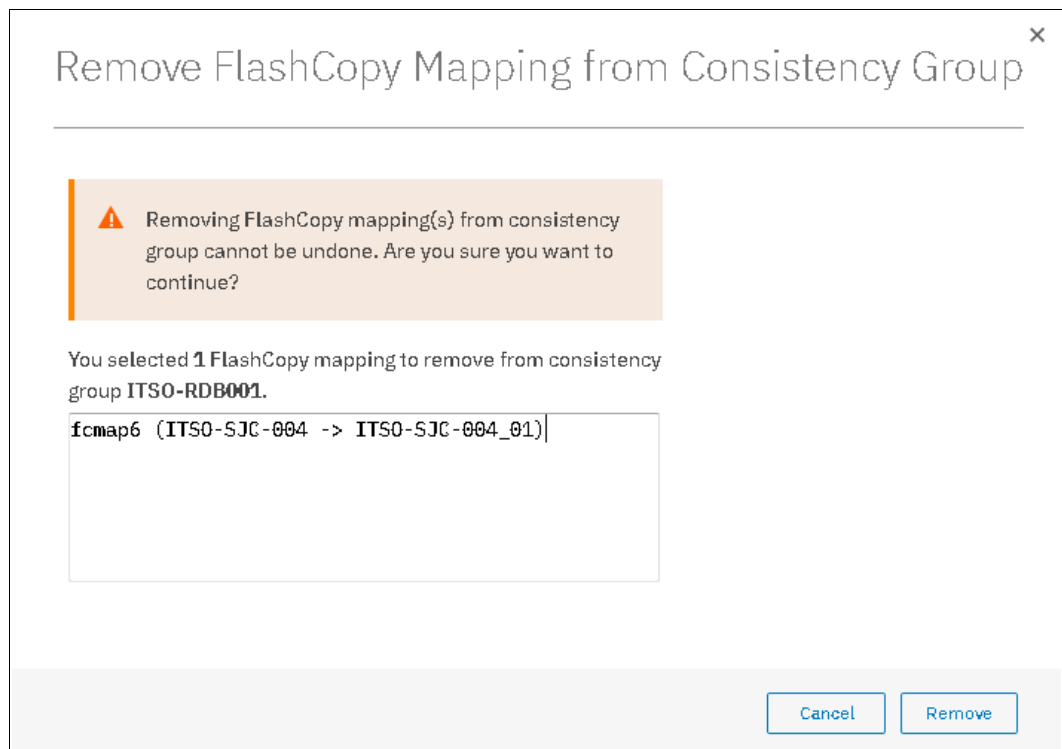


Figure 10-47 Remove FlashCopy Mapping from Consistency Group

10.4.10 Modifying a FlashCopy mapping

Complete the following steps to modify a FlashCopy mapping:

- From the main panel, click **Copy Services** → **FlashCopy** or **Consistency Groups**, or **FlashCopy Mappings**.
- Select the FlashCopy mapping that you want to modify, and right-click **Edit Properties**, as shown in Figure 10-48 on page 514.

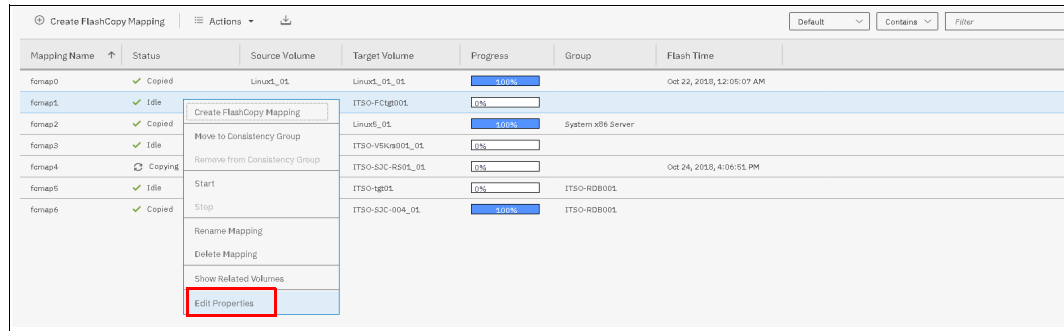


Figure 10-48 Edit Properties

3. In the Edit FlashCopy Mapping window, you can modify the Background Copy Rate and cleaning Rate from the drop-down, as shown in Figure 10-49.

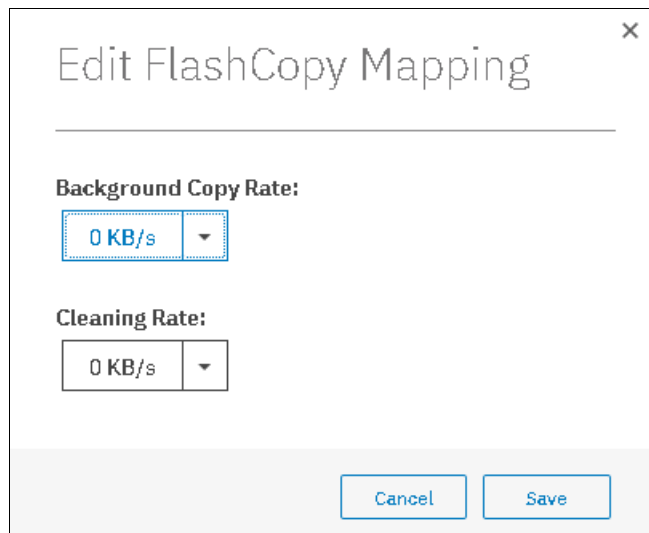


Figure 10-49 Modifying Background Copy Rate

Background Copy Rate: This option determines the priority that is given to the copy process. A faster rate increases the priority of the process, which might affect the performance of other operations.

For FlashCopy background copy rates, starting from V7.8.1, IBM Spectrum Virtualize software allows the background copy rate up to 2 GBps.

Cleaning Rate: This option minimizes the amount of time that a mapping is in the stopping state. If the mapping is not complete, the target volume is offline while the mapping is stopping.

For FlashCopy background cleaning rates, starting from V7.8.1, IBM Spectrum Virtualize software allows the background cleaning rate up to 2 GBps.

4. Click **Save** to confirm your changes.

10.4.11 Renaming FlashCopy mapping

Complete the following steps to rename a FlashCopy mapping:

1. From the main panel, click **Copy Services** → **FlashCopy** or **Consistency Groups**, or **FlashCopy Mappings**.
2. Select the FlashCopy mapping that you want to rename, right-click **Rename Mapping**, as shown in Figure 10-50.

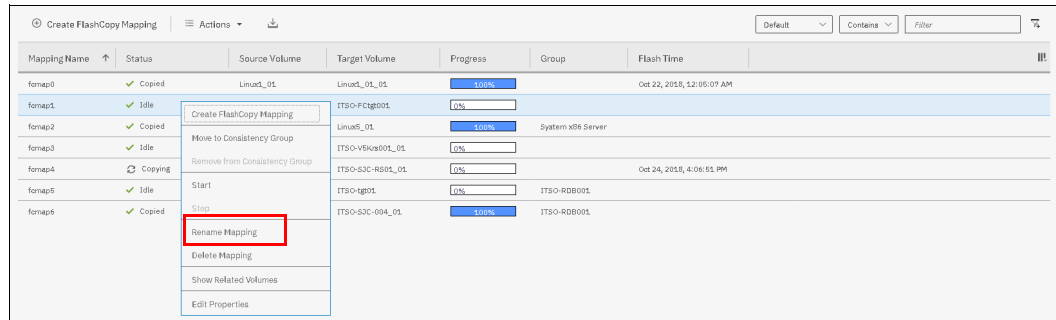


Figure 10-50 Rename Mapping action

Tip: You can also right-click a FlashCopy mapping and select **Rename Mapping**.

3. In the Rename FlashCopy Mapping window, enter the new name that you want to assign to the FlashCopy mapping and click **Rename**, as shown in Figure 10-51.

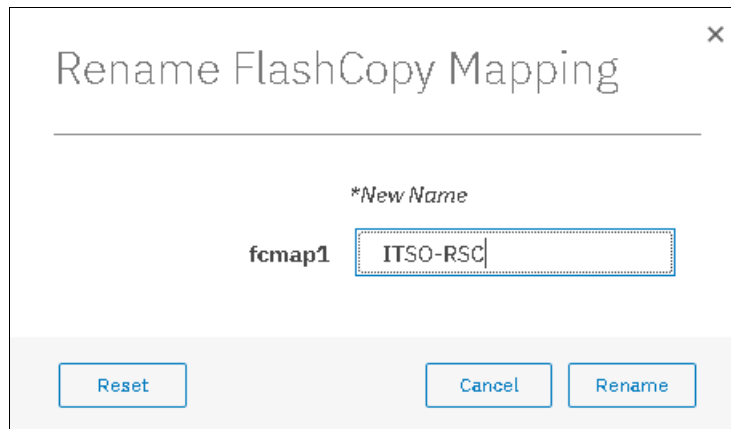


Figure 10-51 Rename Mapping

FlashCopy mapping name: You can use the letters A - Z and a - z, the numbers 0 - 9, and the underscore (_) character. The FlashCopy mapping name can be 1 - 63 characters.

10.4.12 Renaming a Consistency Group

To rename a Consistency Group, complete the following steps:

1. From the main panel, click **Copy Services** → **Consistency Groups**.

2. Select the Consistency Group to rename. Then, select right-click **Rename**, as shown in Figure 10-52.

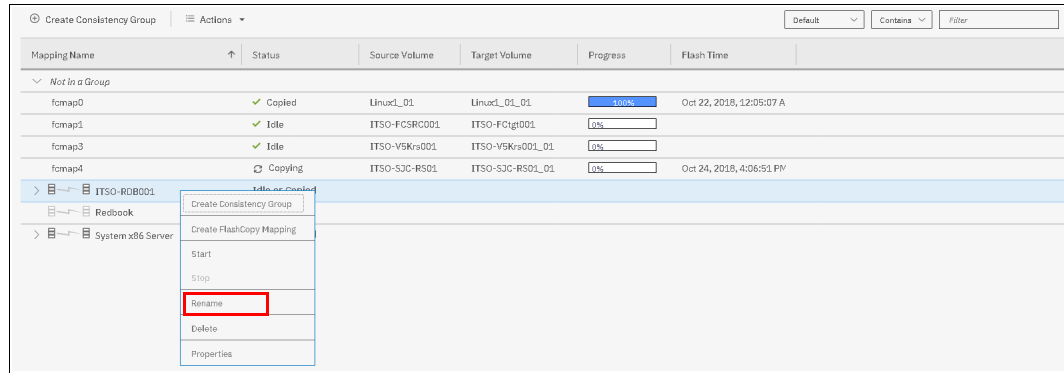


Figure 10-52 Renaming a Consistency Group

3. Enter the new name that you want to assign to the Consistency Group and click **Rename**, as shown in Figure 10-53.

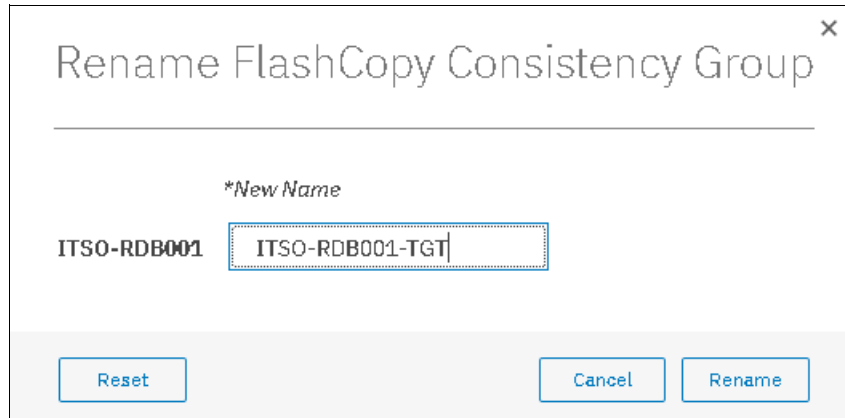


Figure 10-53 Changing the name for a Consistency Group

Consistency Group name: The name can consist of the letters A - Z and a - z, the numbers 0 - 9, the dash (-), and the underscore (_) character. The name can be 1 - 63 characters. However, the name cannot start with a number, dash, or underscore.

10.4.13 Deleting FlashCopy mapping

Complete the following steps to delete a FlashCopy mapping:

1. From the main panel, click **Copy Services** → **FlashCopy** or **Consistency Groups**, or **FlashCopy Mappings**.
2. Select the FlashCopy mapping that you want to delete, and right-click, as shown in Figure 10-54.

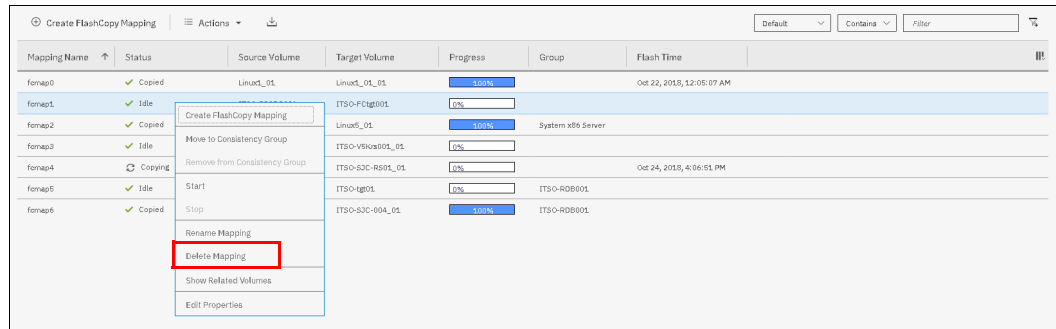


Figure 10-54 Mapping delete

3. The Delete FlashCopy Mapping window opens, as shown in Figure 10-55 on page 518. In the “Verify the number of FlashCopy mappings that you are deleting” field, you must enter the number of volumes that you want to remove. This verification was added to help avoid deleting the wrong mappings.

If you still have target volumes that are inconsistent with the source volumes and you want to delete these FlashCopy mappings, select **Delete the FlashCopy mapping even when the data on the target volume is inconsistent, or if the target volume has other dependencies**.

4. Click **Delete**, as shown in Figure 10-55 on page 518.

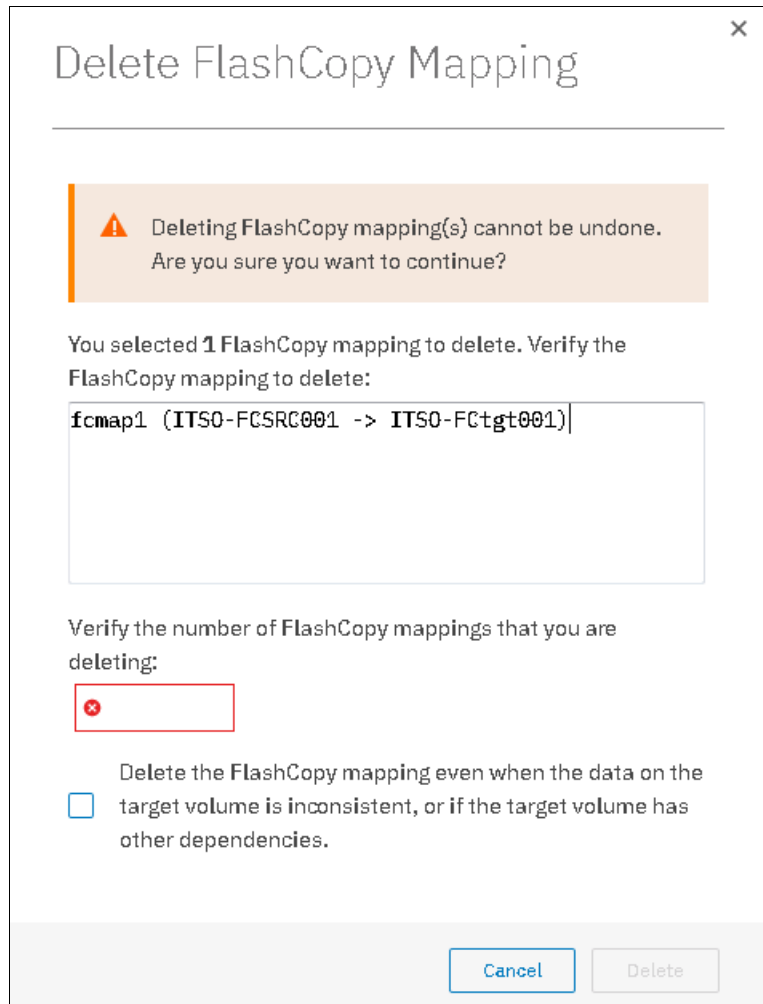


Figure 10-55 Delete FlashCopy Mapping

10.4.14 Deleting FlashCopy Consistency Group

Important: Deleting a Consistency Group does not delete the FlashCopy mappings.

Complete the following steps to delete a FlashCopy Consistency Group:

1. From the main panel, click **Copy Services** → **Consistency Groups**.
2. Select the FlashCopy Consistency Group that you want to delete, right-click **Delete**, as shown in Figure 10-56 on page 519.

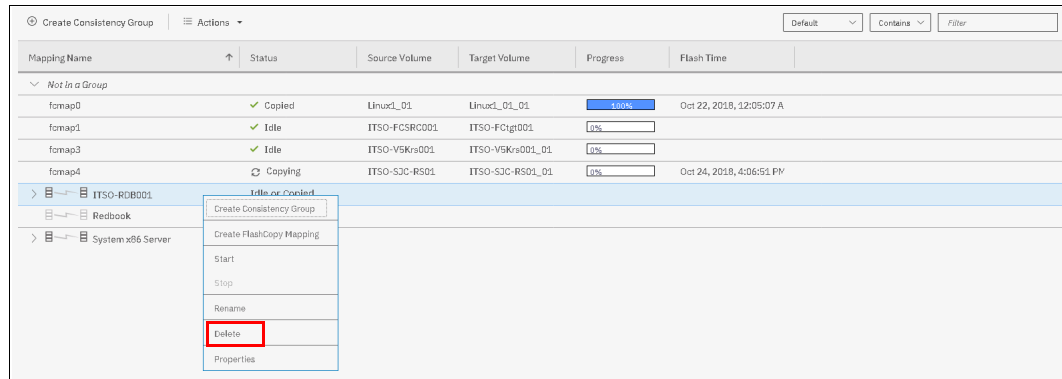


Figure 10-56 Delete Consistency Group

3. The Warning window opens, as shown in Figure 10-57. Click **Yes**.

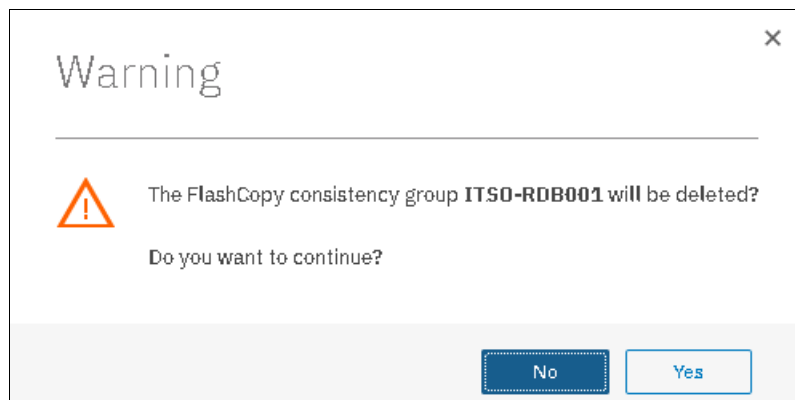


Figure 10-57 Warning window

10.4.15 Starting FlashCopy process

When the FlashCopy mapping is created, the copy process can be started. Only mappings that are not members of a Consistency Group can be started individually. Complete the following steps:

1. From the main panel, click **Copy Services** → **FlashCopy Mappings**.
2. Select the FlashCopy mapping that you want to start, and right-click **Start** (as shown in Figure 10-58 on page 520) to start the FlashCopy process.

Tip: You can also right-click a FlashCopy mapping and select **Start**.

Mapping Name	Status	Source Volume	Target Volume	Progress	Group	Flash Time
fomap0	✓ Copied	Linux_01	Linux_01_01	100%		Oct 22, 2018, 12:05:07 AM
fomap1	✓ Idle	ITSO-FC8RC001	ITSO-FC8R001	0%		
fomap2	✓ Copied		x8_01	100%	System x86 Server	
fomap3	✓ Idle		>V8K01_01	0%		
fomap4	⊗ Copying		>S3C-R801_01	0%		Oct 24, 2018, 4:06:51 PM
fomap5	✓ Idle		>tgr01	0%	ITSO-RDB001	
fomap6	✓ Copied		>S3C-004_01	100%	ITSO-RDB001	

Figure 10-58 Start the FlashCopy process action

- You can check the FlashCopy progress in the Progress column of the table or in the Running Tasks status area. After the task completes, the FlashCopy mapping status is in a Copied state, as shown in Figure 10-59.

Mapping Name	Status	Source Volume	Target Volume	Progress	Group	Flash Time
fomap0	✓ Copied	Linux_01	Linux_01_01	100%		Oct 22, 2018, 12:05:07 AM
fomap1	⊗ Copying	ITSO-FC8RC001	ITSO-FC8R001	0%		Oct 24, 2018, 6:58:53 PM

Figure 10-59 Checking the FlashCopy progress

10.4.16 Stopping FlashCopy process

When a FlashCopy copy process is stopped, the target volume becomes invalid and it is set offline by the system. The FlashCopy mapping copy must be retriggered to bring the target volume online again.

Important: Stop a FlashCopy copy process only when the data on the target volume is not useful and can be discarded, or if you want to modify the FlashCopy mapping. When a FlashCopy mapping is stopped, the target volume becomes invalid and it is set offline by the system.

Complete the following steps to stop a FlashCopy copy process:

- From the main panel, click **Copy Services** → **FlashCopy Mappings**.
- Select the FlashCopy mapping that you want to stop, and right-click **Stop** (as shown in Figure 10-60).

Mapping Name	Status	Source Volume	Target Volume	Progress	Group	Flash Time
fomap0	✓ Copied	Linux_01	Linux_01_01	100%		Oct 22, 2018, 12:05:07 AM
fomap1	⊗ Copying	ITSO-FC8RC001	ITSO-FC8R001	0%		Oct 24, 2018, 6:58:53 PM
fomap2	✓ Copied		x8_01	100%	System x86 Server	
fomap3	✓ Idle		>V8K01_01	0%		
fomap4	⊗ Copying		>R801_01	0%		Oct 24, 2018, 4:06:51 PM
fomap5	✓ Idle		>tgr01	0%	ITSO-RDB001	
fomap6	✓ Copied		>004_01	100%	ITSO-RDB001	

Figure 10-60 Stopping the FlashCopy copy process

The FlashCopy Mapping status changes to Stopped, as shown in Figure 10-61.

Mapping Name	Status	Source Volume	Target Volume	Progress	Group	Flash Time
femapi0	Copied	Linux01_01	Linux01_01	100%		Oct 22, 2015, 12:05:07 AM
femapi1	Stopped	ITSO-FCSRC001	ITSO-FCUG001	0%		

Figure 10-61 FlashCopy Mapping status

10.5 Volume mirroring and migration options

Volume mirroring is a simple RAID 1-type function that enables a volume to remain online, even when the storage pool that is backing it becomes inaccessible. Volume mirroring is designed to protect the volume from storage infrastructure failures by seamless mirroring between storage pools.

Volume mirroring is provided by a specific volume mirroring function in the I/O stack, and it cannot be manipulated like a FlashCopy or other types of copy volumes. However, this feature provides migration functionality, which can be obtained by splitting the mirrored copy from the source, or by using the *migrate to* function. Volume mirroring cannot control backend storage mirroring or replication.

With volume mirroring, host I/O completes when both copies are written, and this feature is enhanced with a tunable latency tolerance. This tolerance provides an option to give preference to losing the redundancy between the two copies. This tunable timeout value is Latency or Redundancy.

The Latency tuning option, which is set with `chvdisk -mirrorwritepriority latency`, is the default. It prioritizes host I/O latency, which yields a preference to host I/O over availability. However, you might need to give preference to redundancy in your environment when availability is more important than I/O response time. Use the `chvdisk -mirrorwritepriority redundancy` command to set the redundancy option.

Regardless of which option you choose, volume mirroring can provide extra protection for your environment.

Migration offers the following options:

- ▶ **Export to Image mode.** By using this option, you can move storage from managed mode to image mode, which is useful if you are using the IBM Storwize as a migration device. For example, vendor A's product cannot communicate with vendor B's product, but you must migrate data from vendor A to vendor B. By using Export to image mode, you can migrate data by using Copy Services functions and then return control to the native array while maintaining access to the hosts.
- ▶ **Import to Image mode.** By using this option, you can import a storage MDisk or logical unit number (LUN) with its data from an external storage system without putting metadata on it so that the data remains intact. After you import it, all copy services functions can be used to migrate the storage to other locations while the data remains accessible to your hosts.
- ▶ **Volume migration by using volume mirroring and then by using Split into New Volume.** By using this option, you can use the available RAID 1 functionality. You create two copies of data that initially has a set relationship (one volume with two copies, one primary and one secondary) but then break the relationship (two volumes, both primary and no relationship between them) to make them independent copies of data.

You can use this option to migrate data between storage pools and devices. You might use this option if you want to move volumes to multiple storage pools. Each volume can have two copies at a time, so you can add only one copy to the original volume, and then you must split those copies to create another copy of the volume.

- ▶ Volume migration by using move to another pool. By using this option, you can move any volume between storage pools without any interruption to the host access. This option is a quicker version of the Volume Mirroring and Split into New Volume option. You might use this option if you want to move volumes in a single step, or you do not have a volume mirror copy already.

Migration: While these migration methods do not disrupt access, you must take a brief outage to install the host drivers for your IBM Storwize V5000 Gen2 system if you do not already have them installed.

With volume mirroring, you can move data to different MDisk within the same storage pool or move data between different storage pools. The use of volume mirroring over volume migration is beneficial because with volume mirroring, storage pools do not need to have the same extent size as is the case with volume migration.

Note: Volume mirroring does not create a second volume before you split copies. Volume mirroring adds a second copy of the data under the same volume, so you end up having one volume presented to the host with two copies of data connected to this volume. Only splitting copies creates another volume, and then both volumes have only one copy of the data.

Starting with V7.3 and the introduction of the new cache architecture, mirrored volume performance is significantly improved. Now, lower cache is beneath the volume mirroring layer, which means that both copies have their own cache.

This approach helps in cases of having copies of different types; for example, generic and compressed, because now both copies use its independent cache and performs its own read prefetch. Destaging of the cache can now be done independently for each copy, so one copy does not affect performance of a second copy.

Also, because the Storwize destage algorithm is MDisk aware, it can tune or adapt the destaging process (depending on MDisk type and usage) for each copy independently.

10.6 Native IP replication

Before we describe Remote Copy features that benefit from the use of multiple IBM Storwize V5000 Gen2 systems, it is important to describe the partnership option introduced with V7.2 native IP replication.

For more information about implementing native IP replication and how to use it, see *IBM SAN Volume Controller and Storwize Family Native IP Replication*, REDP-5103.

10.6.1 Native IP replication technology

Remote Mirroring over IP communication is supported on the IBM SAN Volume Controller and Storwize Family systems by using Ethernet communication links. The IBM Spectrum Virtualize Software IP replication uses innovative *Bridgeworks SANSlide* technology to optimize network bandwidth and utilization. This new function enables the use of a lower-speed and lower-cost networking infrastructure for data replication.

Bridgeworks' SANSlide technology, which is integrated into the IBM Spectrum Virtualize Software, uses artificial intelligence to improve network bandwidth use and adapt to changing workload and network conditions.

This technology can improve remote mirroring network bandwidth usage up to three times, which can enable clients to deploy a less costly network infrastructure, or speed up remote replication cycles to enhance disaster recovery effectiveness.

With an Ethernet network data flow, the data transfer can slow down over time. This condition occurs because of the latency that is caused by waiting for the acknowledgment of each set of packets that are sent. The next packet set cannot be sent until the previous packet is acknowledged, as shown in Figure 10-62.

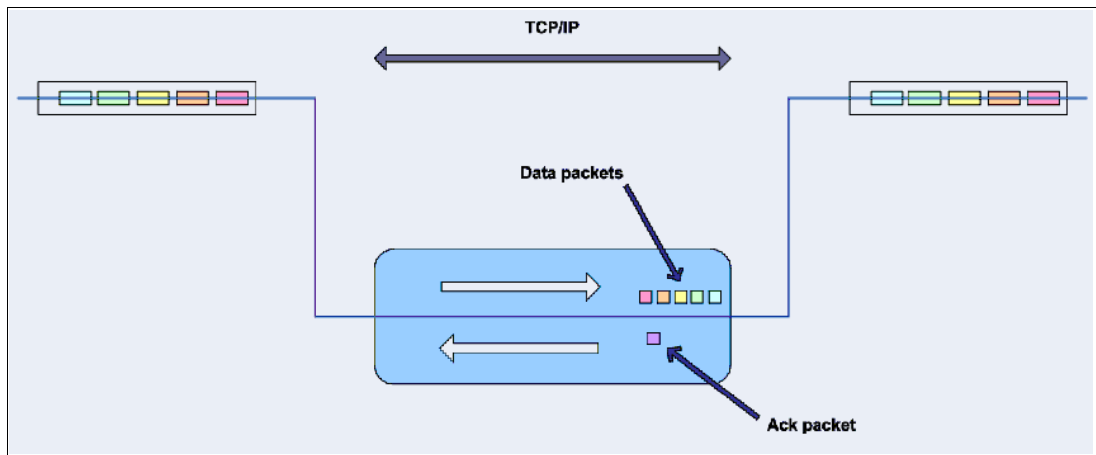


Figure 10-62 Typical Ethernet network data flow

However, by using the embedded IP replication, this behavior can be eliminated with the enhanced parallelism of the data flow by using multiple virtual connections (VC) that share IP links and addresses. The artificial intelligence engine can dynamically adjust the number of VCs, receive window size, and packet size as appropriate to maintain optimum performance. While the engine is waiting for one VC's ACK, it sends more packets across other VCs. If packets are lost from any VC, data is automatically retransmitted, as shown in Figure 10-63.

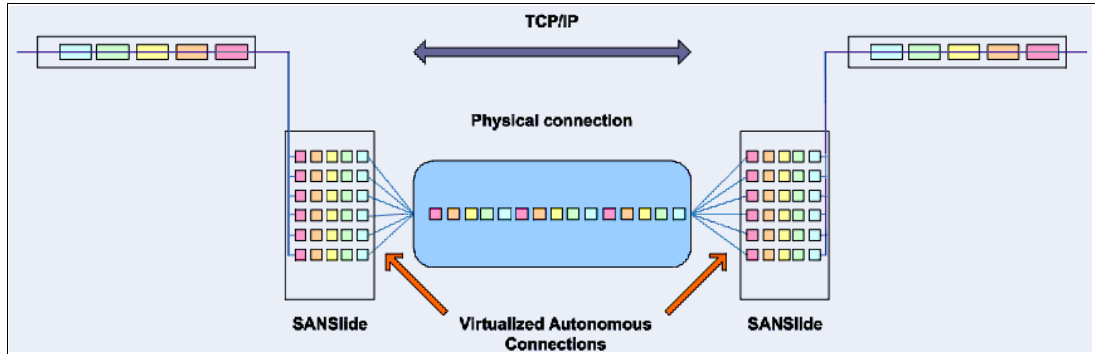


Figure 10-63 Optimized network data flow by using Bridgeworks SANSlide technology

For more information about this technology, see *IBM Storwize V7000 and SANSlide Implementation*, REDP-5023.

With native IP partnership, the following Copy Services features are supported:

- ▶ Metro Mirror (MM)

Referred to as *synchronous replication*, MM provides a consistent copy of a source virtual disk on a target virtual disk. Data is written to the target virtual disk synchronously after it is written to the source virtual disk so that the copy is continuously updated.

- ▶ Global Mirror (GM) and GM with Change Volumes

Referred to as *asynchronous replication*, GM provides a consistent copy of a source virtual disk on a target virtual disk. Data is written to the target virtual disk asynchronously so that the copy is continuously updated. However, the copy might not contain the last few updates if a disaster recovery (DR) operation is performed. An added extension to GM is GM with Change Volumes. GM with Change Volumes is the preferred method for use with native IP replication.

10.6.2 IBM Storwize System Layers

An IBM Storwize family system can be in one of two layers: the *replication* layer or the *storage* layer. The system layer affects how the system interacts with IBM Storwize V5000 Gen2 systems and IBM SAN Volume Controller systems. The IBM SAN Volume Controller is always set to replication layer and this parameter is unchangeable.

In the storage layer, an Storwize family system has the following characteristics and requirements:

- ▶ The system can perform MM and GM replication with other storage-layer systems.
- ▶ The system can provide external storage for replication-layer systems or IBM SAN Volume Controller.
- ▶ The system cannot use a storage-layer system as external storage.

In the replication layer, an IBM SAN Volume Controller or an IBM Storwize family system has the following characteristics and requirements:

- ▶ The system can perform MM and GM replication with other replication-layer systems or IBM SAN Volume Controller.
- ▶ The system cannot provide external storage for a replication-layer system or an IBM SAN Volume Controller.
- ▶ The system can use a storage-layer system as external storage.

A Storwize family system is in the storage layer by default, but the layer can be changed. For example, you might want to change a Storwize V5000 to a replication layer to complete Global Mirror or Metro Mirror replication with an IBM SAN Volume Controller system.

To change the storage layer of a Storwize system with internal storage before you add a second system into the SAN zone, you do not need to stop I/O operations. However, if the system has Fibre Channel (FC) connections to another Storwize family or IBM SAN Volume Controller system in the SAN fabric, I/O operations must be stopped temporarily. In this scenario, the FC ports must be disabled (for example, by unplugging all the FC ports, changing zoning, and disabling switch ports) before you change the system layer. Then, you must re-enable the FC ports.

Note: Before you change the layer of a Storwize family system, the following conditions must be met:

- ▶ No host object can be configured with worldwide port names (WWPNs) from a Storwize family system.
- ▶ No system partnerships can be defined.
- ▶ No Storwize family system can be visible on the SAN fabric.

In your IBM Storwize system, use the `lssystem` command to check the current system layer, as shown in Example 10-2.

Example 10-2 Output from lssystem command showing the system layer

```
IBM_Storwize:ITS0V5030:superuser>lssystem
id 000002006D81FE4D
name ITS0V5030
...
easy_tier_acceleration off
has_nas_key no
layer storage
...

```

Note: Consider the following rules for creating remote partnerships between the IBM SAN Volume Controller and Storwize Family systems:

- ▶ An IBM SAN Volume Controller is always in the replication layer.
- ▶ By default, the IBM Storwize systems are in the storage layer but can be changed to the replication layer.
- ▶ A system can form partnerships only with systems in the same layer.
- ▶ Starting in software V6.4, an IBM SAN Volume Controller or Storwize system in the replication layer can virtualize an IBM Storwize in the storage layer.

10.6.3 IP partnership limitations

The following prerequisites and assumptions must be considered before IP partnership between two IBM Spectrum Virtualize systems can be established:

- ▶ The IBM SAN Volume Controller or IBM Storwize systems are successfully installed with V7.2 or later code levels.
- ▶ The systems must have the necessary licenses that enable remote copy partnerships to be configured between two systems. No separate license is required to enable IP partnership.
- ▶ The storage SANs are configured correctly and the correct infrastructure to support the Spectrum Virtualize systems in remote copy partnerships over IP links is in place.
- ▶ The two systems must be able to ping each other and perform the discovery.
- ▶ The maximum number of partnerships between the local and remote systems, including both IP and FC partnerships, is limited to the current maximum that is supported, which is three partnerships (four systems total).
- ▶ Only a single partnership over IP is supported.
- ▶ A system can have simultaneous partnerships over FC and IP, but with separate systems. The FC zones between two systems must be removed before an IP partnership is configured.
- ▶ IP partnerships are supported on both 10 gigabits per second (Gbps) links and 1 Gbps links. However, the intermix of both on a single link is not supported.
- ▶ The maximum supported round-trip time is 80 milliseconds (ms) for 1 Gbps links.
- ▶ The maximum supported round-trip time is 10 ms for 10 Gbps links.
- ▶ The minimum supported link bandwidth is 10 Mbps.
- ▶ The inter-cluster heartbeat traffic uses 1 Mbps per link.
- ▶ Only nodes from two I/O Groups can have ports that are configured for an IP partnership.
- ▶ Migrations of remote copy relationships directly from FC-based partnerships to IP partnerships are not supported.
- ▶ IP partnerships between the two systems can be over IPv4 or IPv6 only, but not both.
- ▶ Virtual LAN (VLAN) tagging of the IP addresses that are configured for remote copy is supported starting with V7.4.0.
- ▶ Management IP and Internet SCSI (iSCSI) IP on the same port can be in a different network starting with V7.4.0.
- ▶ An added layer of security is provided by using Challenge Handshake Authentication Protocol (CHAP) authentication.
- ▶ Transmission Control Protocol (TCP) ports 3260 and 3265 are used for IP partnership communications. Therefore, these ports must be open in firewalls between the systems.
- ▶ Only a single Remote Copy (RC) data session per physical link can be established. It is intended that only one connection (for sending/receiving RC data) is made for each independent physical link between the systems.

Note: A physical link is the physical IP link between the two sites, A (local) and B (remote). Multiple IP addresses on local system A can be connected (by Ethernet switches) to this physical link. Similarly, multiple IP addresses on remote system B can be connected (by Ethernet switches) to the same physical link. At any time, only a single IP address on cluster A can form an RC data session with an IP address on cluster B.

- ▶ The maximum throughput is restricted based on the use of 1 Gbps or 10 Gbps Ethernet ports, and varies based on distance (for example, round-trip latency) and quality of communication link (for example, packet loss):
 - One 1 Gbps port might transfer up to 110 megabytes per second (MBps) unidirectional, 190 MBps bidirectional
 - Two 1 Gbps ports might transfer up to 220 MBps unidirectional, 325 MBps bidirectional
 - One 10 Gbps port might transfer up to 240 MBps unidirectional, 350 MBps bidirectional
 - Two 10 Gbps port might transfer up to 440 MBps unidirectional, 600 MBps bidirectional

Note: The Bandwidth setting definition when the IP partnerships are created changed. Previously, the bandwidth setting defaulted to 50 MB, and was the maximum transfer rate from the primary site to the secondary site for initial sync/resyncs of volumes.

The Link Bandwidth setting is now configured by using megabits (Mb) not MB. You set the Link Bandwidth setting to a value that the communication link can sustain, or to what is allocated for replication. The Background Copy Rate setting is now a percentage of the Link Bandwidth. The Background Copy Rate setting determines the available bandwidth for the initial sync and resyncs or for GM with Change Volumes.

Note: For more information about IP replication requirements and limitations and supported configurations, see [IBM Knowledge Center for IBM Storwize V5000](#).

10.6.4 VLAN support

Starting with V7.4.0, VLAN tagging is supported for both iSCSI host attachment and IP replication. Hosts and remote-copy operations can connect to the system through Ethernet ports. Each traffic type has different bandwidth requirements, which can interfere with each other if they share the same IP connections. VLAN tagging creates two separate connections on the same IP network for different types of traffic. The system supports VLAN configuration on both IPv4 and IPv6 connections.

When the VLAN ID is configured for the IP addresses that are used for iSCSI host attach or IP replication, the appropriate VLAN settings on the Ethernet network and servers must be configured correctly to not experience connectivity issues. After the VLANs are configured, changes to the VLAN settings disrupts iSCSI and IP replication traffic to and from the partnerships.

During the VLAN configuration for each IP address, the VLAN settings for the local and failover ports on two nodes of an I/O Group can differ. To avoid any service disruption, switches must be configured so the failover VLANs are configured on the local switch ports and the failover of IP addresses from a failing node to a surviving node succeeds.

If failover VLANs are not configured on the local switch ports, there are no paths to IBM Storwize V5000 Gen2 system node canisters during a node canister failure and the replication fails.

Consider the following requirements and procedures when implementing VLAN tagging:

- ▶ VLAN tagging is supported for IP partnership traffic between two systems.
- ▶ VLAN provides network traffic separation at the layer 2 level for Ethernet transport.
- ▶ VLAN tagging by default is disabled for any IP address of a node port. You can use the CLI or GUI to optionally set the VLAN ID for port IPs on both systems in the IP partnership.
- ▶ When a VLAN ID is configured for the port IP addresses that are used in remote copy port groups, appropriate VLAN settings on the Ethernet network must also be properly configured to prevent connectivity issues.

Setting VLAN tags for a port is disruptive. Therefore, VLAN tagging requires that you stop the partnership first before you configure VLAN tags. Then, restart again when the configuration is complete.

Note: For more information about configuring VLAN for IP replication, see [IBM Knowledge Center for IBM Storwize V5000](#).

10.6.5 IP partnership and terminology

The IP partnership terminology and abbreviations that are used are listed in Table 10-7.

Table 10-7 Terminology for IP partnership

IP partnership terminology	Description
Remote copy group or Remote copy port group	<p>The following numbers group a set of IP addresses that are connected to the same physical link. Therefore, only IP addresses that are part of the same remote copy group can form remote copy connections with the partner system:</p> <ul style="list-style-type: none"> ▶ 0 – Ports that are not configured for remote copy ▶ 1 – Ports that belong to remote copy port group 1 ▶ 2 – Ports that belong to remote copy port group 2 <p>Each IP address can be shared for iSCSI host attach and remote copy functionality. Therefore, appropriate settings must be applied to each IP address.</p>
IP partnership	Two systems that are partnered to perform remote copy over native IP links.
FC partnership	Two systems that are partnered to perform remote copy over native Fibre Channel links.
Failover	Failure of a node within an I/O group causes the volume access to go through the surviving node. The IP addresses fail over to the surviving node in the I/O group. When the configuration node of the system fails, management IPs also fail over to an alternative node.
Failback	When the failed node rejoins the system, all failed over IP addresses are failed back from the surviving node to the rejoined node, and virtual disk access is restored through this node.

IP partnership terminology	Description
linkbandwidthmbits	Aggregate bandwidth of all physical links between two sites in Mbps.
IP partnership or partnership over native IP links	These terms are used to describe the IP partnership feature.
Discovery	<p>Process by which two IBM Spectrum Virtualize systems exchange information about their IP address configuration. For IP-based partnerships, only IP addresses configured for Remote Copy are discovered.</p> <p>For example, the first Discovery takes place when the user is running the <code>mkippartnership</code> CLI command. Subsequent Discoveries can take place as a result of user activities (configuration changes) or as a result of hardware failures (for example, node failure, ports failure, and so on).</p>

10.6.6 States of IP partnership

The different partnership states in IP partnership are listed in Table 10-8.

Table 10-8 States of IP partnership

State	Systems connected	Support for active remote copy I/O	Comments
Partially_Configured_Local	No	No	This state indicates that the initial discovery is complete.
Fully_Configured	Yes	Yes	Discovery successfully completed between two systems, and the two systems can establish remote copy relationships.
Fully_Configured_Stopped	Yes	Yes	The partnership is stopped on the system.
Fully_Configured_Remote_Stopped	Yes	No	The partnership is stopped on the remote system.
Not_Present	Yes	No	The two systems cannot communicate with each other. This state is also seen when data paths between the two systems are not established.
Fully_Configured_Exceeded	Yes	No	There are too many systems in the network, and the partnership from the local system to remote system is disabled.
Fully_Configured_Excluded	No	No	The connection is excluded because of too many problems, or either system cannot support the I/O work load for the Metro Mirror and Global Mirror relationships.

The following steps must be completed to establish two systems in the IP partnerships:

1. The administrator configures the CHAP secret on both the systems. This step is *not* mandatory, and users can choose to not configure the CHAP secret.
2. The administrator configures the system IP addresses on both local and remote systems so that they can discover each other over the network.

3. If you want to use VLANs, configure your LAN switches and Ethernet ports to use VLAN tagging (for more information about VLAN tagging, see 10.6.4, “VLAN support” on page 527).
4. The administrator configures the systems ports on each node in both of the systems by using the GUI (or the `cfgport ip` CLI command), and completes the following steps:
 - a. Configure the IP addresses for remote copy data.
 - b. Add the IP addresses in the respective remote copy port group.
 - c. Define whether the host access on these ports over iSCSI is allowed.
5. The administrator establishes the partnership with the remote system from the local system where the partnership state then changes to the `Partially_Configured_Local` state.
6. The administrator establishes the partnership from the remote system with the local system, and if successful, the partnership state then changes to the `Fully_Configured` state, which implies that the partnerships over the IP network were successfully established. The partnership state momentarily remains in the `Not_Present` state before moving to the `Fully_Configured` state.
7. The administrator creates MM, GM, and GM with Change Volume relationships.

Partnership consideration: When the partnership is created, no master or auxiliary status is defined or implied. The partnership is equal. The concepts of *master or auxiliary* and *primary or secondary* apply to volume relationships only, not to system partnerships.

10.6.7 Remote copy groups

This section describes remote copy groups and different ways to configure the links between the two remote systems. The two IBM Spectrum Virtualize systems can be connected to each other over one link or, at most, two links. To address the requirement to enable the systems to know about the physical links between the two sites, the concept of remote copy port groups was introduced.

Remote copy port group ID is a numerical tag that is associated with an IP port of IBM Storwize V5000 Gen2 system to indicate to which physical IP link it is connected. Multiple nodes can be connected to the same physical long-distance link, and must therefore share the same remote copy port group ID.

In scenarios where there are two physical links between the local and remote clusters, two remote copy port group IDs must be used to designate which IP addresses are connected to which physical link. This configuration must be done by the system administrator by using the GUI or the `cfgport ip` CLI command.

Remember: IP ports on both partners must be configured with identical remote copy port group IDs for the partnership to be established correctly.

The IBM Storwize V5000 Gen2 system IP addresses that are connected to the same physical link are designated with identical remote copy port groups and is supported three remote copy groups: 0, 1, and 2. The IP addresses are, by default, in remote copy port group 0.

Ports in port group 0 are not considered for creating remote copy data paths between two systems. For partnerships to be established over IP links directly, IP ports must be configured in remote copy group 1 if a single inter-site link exists, or in remote copy groups 1 and 2 if two inter-site links exist. You can assign one IPv4 address and one IPv6 address to each Ethernet

port on the system platforms. Each of these IP addresses can be shared between iSCSI host attach and the IP partnership. The user must configure the required IP address (IPv4 or IPv6) on an Ethernet port with a remote copy port group.

The administrator might want to use IPv6 addresses for remote copy operations and use IPv4 addresses on that same port for iSCSI host attach. This configuration also implies that for two systems to establish an IP partnership, both systems must have IPv6 addresses that are configured.

Administrators can choose to dedicate an Ethernet port for IP partnership only. In that case, host access must be explicitly disabled for that IP address and any other IP address that is configured on that Ethernet port.

Note: To establish an IP partnership, each IBM Storwize node must have only a single remote copy port group that is configured, 1 or 2. The remaining IP addresses must be in remote copy port group 0.

10.7 Remote Copy services

This section describes the Remote Copy services, which are a synchronous remote copy called *Metro Mirror (MM)*, asynchronous remote copy called *Global Mirror (GM)*, and Global Mirror with Change Volumes. Remote Copy in the IBM Storwize V5000 Gen2 system is similar to Remote Copy in the IBM System Storage DS8000 family at a functional level, but the implementation differs.

The IBM Storwize V5000 Gen2 system provides a single point of control when remote copy is enabled in your network (regardless of the disk subsystems that are used) if those disk subsystems are supported by the IBM Storwize V5000 Gen2 system.

The general application of remote copy services is to maintain two real-time synchronized copies of a volume. Often, two copies are geographically dispersed between two IBM Storwize V5000 Gen2 systems, although it is possible to use MM or GM within a single system (within an I/O Group). If the master copy fails, you can enable an auxiliary copy for I/O operation.

Tips: Intracluster MM/GM uses more resources within the system when compared to an intercluster MM/GM relationship, where resource allocation is shared between the systems. Use intercluster MM/GM when possible. For mirroring volumes in the same system, it is better to use Volume Mirroring or the FlashCopy feature.

A typical application of this function is to set up a dual-site solution that uses two IBM Storwize V5000 Gen2 systems. The first site is considered the *primary site* or *production site*, and the second site is considered the *backup site* or *failover site*, which is activated when a failure at the first site is detected.

10.7.1 Multiple IBM Storwize V5000 system mirroring

Each IBM Storwize V5000 Gen2 system can maintain up to three partner system relationships, which enables as many as four systems to be directly associated with each other. This system partnership capability enables the implementation of disaster recovery (DR) solutions.

Note: For more information about restrictions and limitations of native IP replication, see 10.6.3, “IP partnership limitations” on page 526.

Figure 10-64 shows an example of a multiple system mirroring configuration.

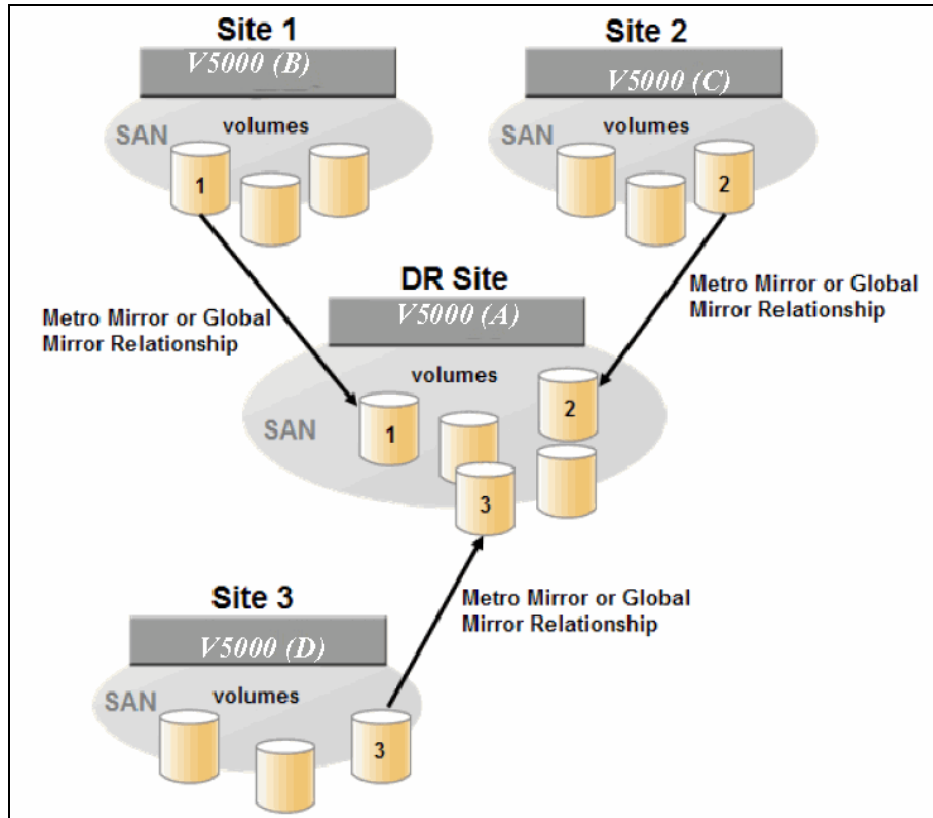


Figure 10-64 Multiple system mirroring configuration example

Supported multiple system mirroring topologies

Multiple system mirroring supports various partnership topologies, as shown in the example in Figure 10-65. This example is a star topology (A →B, A →C, and A →D).

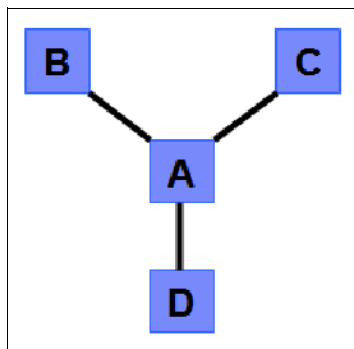


Figure 10-65 Star topology

Figure 10-65 shows four systems in a star topology, with System A at the center. System A can be a central DR site for the three other locations.

By using a star topology, you can migrate applications by using a process, such as the one described in the following example:

1. Suspend application at A.
2. Remove the $A \rightarrow B$ relationship.
3. Create the $A \rightarrow C$ relationship (or the $B \rightarrow C$ relationship).
4. Synchronize to system C, and ensure that $A \rightarrow C$ is established:
 - $A \rightarrow B$, $A \rightarrow C$, $A \rightarrow D$, $B \rightarrow C$, $B \rightarrow D$, and $C \rightarrow D$
 - $A \rightarrow B$, $A \rightarrow C$, and $B \rightarrow C$

Figure 10-66 shows an example of a triangle topology ($A \rightarrow B$, $A \rightarrow C$, and $B \rightarrow C$).

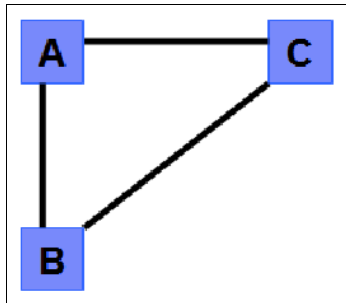


Figure 10-66 Triangle topology

Figure 10-67 shows an example of an IBM Storwize V5000 Gen2 system fully connected topology ($A \rightarrow B$, $A \rightarrow C$, $A \rightarrow D$, $B \rightarrow D$, and $C \rightarrow D$).

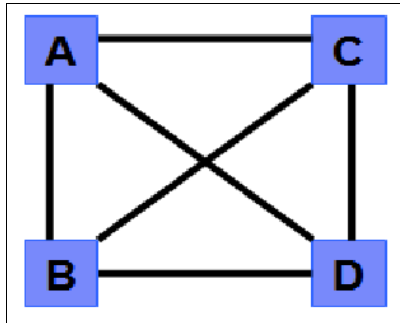


Figure 10-67 Fully connected topology

Figure 10-67 is a fully connected mesh in which every system has a partnership to each of the three other systems. This topology enables volumes to be replicated between any pair of systems, for example $A \rightarrow B$, $A \rightarrow C$, and $B \rightarrow C$.

Figure 10-68 shows a daisy-chain topology.

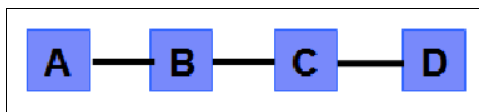


Figure 10-68 Daisy-chain topology

Although systems can have up to three partnerships, volumes can be part of only one remote copy relationship; for example, $A \rightarrow B$.

System partnership intermix: All of the preceding topologies are valid for the intermix of the IBM SAN Volume Controller with the Storwize V5000 Gen2 system if the Storwize V5000 Gen2 system is set to the replication layer and running IBM Spectrum Virtualize code 6.3.0 or later.

10.7.2 Importance of write ordering

Many applications that use block storage have a requirement to survive failures, such as loss of power or a software crash, and to not lose data that existed before the failure. Because many applications must perform large numbers of update operations in parallel, maintaining write ordering is key to ensuring the correct operation of applications after a disruption.

An application that performs a high volume of database updates is designed with the concept of dependent writes. With dependent writes, it is important to ensure that an earlier write completed before a later write is started. Reversing or performing the order of writes differently than the application intended can undermine the application's algorithms and can lead to problems, such as detected or undetected data corruption.

The IBM Spectrum Virtualize Metro Mirror and Global Mirror implementation operates in a manner that is designed to always keep a consistent image at the secondary site. The Global Mirror implementation uses complex algorithms that operate to identify sets of data and number those sets of data in sequence. The data is then applied at the secondary site in the defined sequence.

Operating in this manner ensures that if the relationship is in a `Consistent_Synchronized` state, the Global Mirror target data is at least crash consistent, and supports quick recovery through application crash recovery facilities.

Remote Copy Consistency Groups

A Remote Copy Consistency Group can contain an arbitrary number of relationships up to the maximum number of MM/GM relationships that is supported by the IBM Spectrum Virtualize system. MM/GM commands can be issued to a Remote Copy Consistency Group.

Therefore, these commands can be issued simultaneously for all MM/GM relationships that are defined within that Consistency Group, or to a single MM/GM relationship that is not part of a Remote Copy Consistency Group. For example, when a `startrcconsistgrp` command is issued to the Consistency Group, all of the MM/GM relationships in the Consistency Group are started at the same time.

Figure 10-69 shows the concept of Metro Mirror Consistency Groups. The same applies to Global Mirror Consistency Groups.

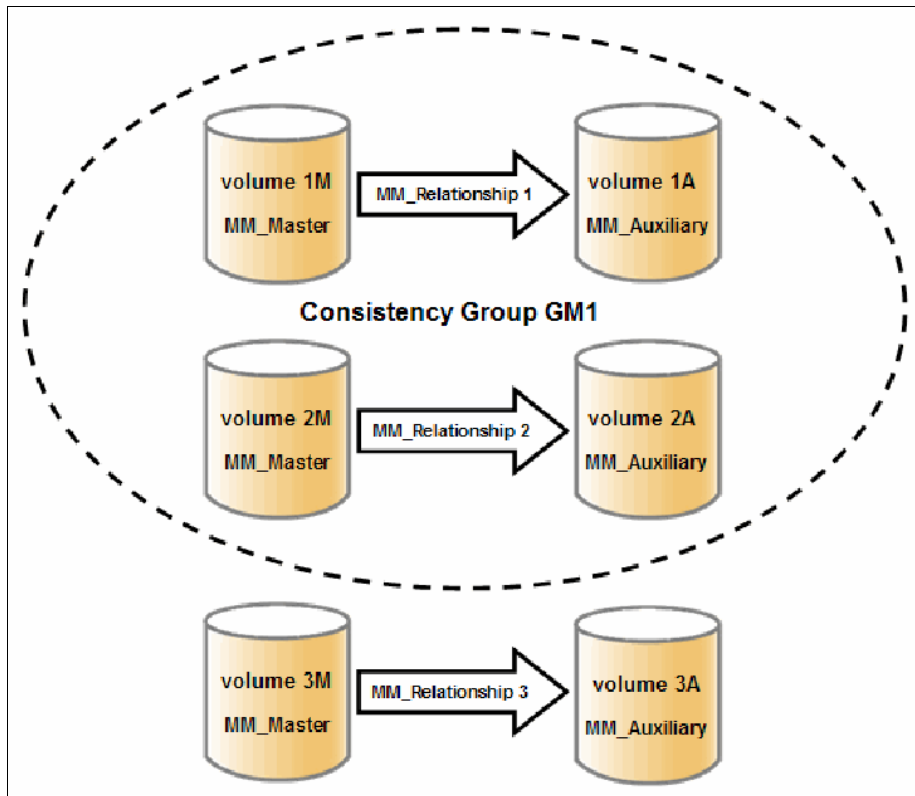


Figure 10-69 Metro Mirror Consistency Group

Because the MM_Relationship 1 and 2 are part of the Consistency Group, they can be handled as one entity. The stand-alone MM_Relationship 3 is handled separately.

Certain uses of MM/GM require the manipulation of more than one relationship. Remote Copy Consistency Groups can group relationships so that they are manipulated in unison.

Consider the following points:

- ▶ MM/GM relationships can be part of a Consistency Group, or they can be stand-alone and, therefore, are handled as single instances.
- ▶ A Consistency Group can contain zero or more relationships. An empty Consistency Group with zero relationships in it has little purpose until it is assigned its first relationship, except that it has a name.
- ▶ All relationships in a Consistency Group must have corresponding master and auxiliary volumes.
- ▶ All relationships in one Consistency Group must be the same type, for example only Metro Mirror or only Global Mirror.

Although Consistency Groups can be used to manipulate sets of relationships that do not need to satisfy these strict rules, this manipulation can lead to unwanted side effects. The rules behind a Consistency Group mean that certain configuration commands are prohibited. These configuration commands are not prohibited if the relationship is not part of a Consistency Group.

For example, consider the case of two applications that are independent, yet they are placed into a single Consistency Group. If an error occurs, synchronization is lost and a background copy process is required to recover synchronization. While this process is progressing, MM/GM rejects attempts to enable access to the auxiliary volumes of either application.

If one application finishes its background copy more quickly than the other application, MM/GM still refuses to grant access to its auxiliary volumes, even though it is safe in this case. The MM/GM policy is to refuse access to the entire Consistency Group if any part of it is inconsistent.

Stand-alone relationships and Consistency Groups share a common configuration and state model. All of the relationships in a non-empty Consistency Group have the same state as the Consistency Group.

10.7.3 Remote copy intercluster communication

In the traditional FC, the intercluster communication between systems in a MM and GM partnership is performed over the SAN. This section describes this communication path.

For more information about intercluster communication between systems in an IP partnership, see 10.6.6, “States of IP partnership” on page 529.

The IBM Storwize V5000 Gen2 system FC ports on each system must communicate with each other to create the partnership. Switch zoning is critical to facilitating intercluster communication.

Intercluster communication channels

When an IBM Spectrum Virtualize system partnership is defined on a pair of systems, the following intercluster communication channels are established:

- ▶ A single control channel, which is used to exchange and coordinate configuration information
- ▶ I/O channels between each of these nodes in the systems

These channels are maintained and updated as nodes and links appear and disappear from the fabric, and are repaired to maintain operation where possible. If communication between the systems is interrupted or lost, an event is logged (and the MM and GM relationships stop).

Alerts: You can configure the system to raise Simple Network Management Protocol (SNMP) traps to the enterprise monitoring system to alert on events that indicate an interruption in internode communication occurred.

Intercluster links

All IBM Storwize V5000 Gen2 node canisters maintain a database of other devices that are visible on the fabric. This database is updated as devices appear and disappear.

Devices that advertise themselves as IBM SAN Volume Controller or Storwize family product nodes are categorized according to the system to which they belong. Nodes that belong to the same system establish communication channels between themselves and begin to exchange messages to implement clustering and the functional protocols of IBM Spectrum Virtualize.

Nodes that are in separate systems do not exchange messages after initial discovery is complete, unless they are configured together to perform a remote copy relationship.

The intercluster link carries control traffic to coordinate activity between two systems. The link is formed between one node in each system. The traffic between the designated nodes is distributed among logins that exist between those nodes.

If the designated node fails (or all of its logins to the remote system fail), a new node is chosen to carry control traffic. This node change causes the I/O to pause, but it does not put the relationships in a `ConsistentStopped` state.

Note: It is advised to use `chsystem` with `-partnerfcportmask` to dedicate several FC ports only to system-to-system traffic to ensure that remote copy is not affected by other traffic, such as host-to-node traffic or node-to-node traffic within the same system.

10.7.4 Metro Mirror overview

Metro Mirror establishes a synchronous relationship between two volumes. The volumes in a Metro Mirror relationship are referred to as the master volume (primary) and the auxiliary volume (secondary). Traditional FC Metro Mirror is primarily used in a metropolitan area or geographical area, up to a maximum distance of 300 km (186.4 miles) to provide synchronous replication of data.

With synchronous copies, host applications write to the master volume, but they do not receive confirmation that the write operation completed until the data is written to the auxiliary volume. This action ensures that both volumes have identical data when the copy completes. After the initial copy completes, the Metro Mirror function always maintains a fully synchronized copy of the source data at the target site.

Metro Mirror has the following characteristics:

- ▶ Zero recovery point objective (RPO)
- ▶ Synchronous
- ▶ Production application performance that is affected by round-trip latency

Increased distance directly affects host I/O performance because the writes are synchronous. Use the requirements for application performance when you are selecting your Metro Mirror auxiliary location.

Consistency Groups can be used to maintain data integrity for dependent writes, which is similar to FlashCopy Consistency Groups (for more information about FlashCopy Consistency Groups, see 10.3, “Implementing FlashCopy” on page 472).

The IBM Spectrum Virtualize provides intracluster and intercluster Metro Mirror.

Intracluster Metro Mirror

Intracluster Metro Mirror performs the intracluster copying of a volume, in which both volumes belong to the same system and I/O Group within the system. Because it is within the same I/O Group, there must be sufficient bitmap space within the I/O Group for both sets of volumes and licensing on the system.

Important: Performing Metro Mirror across I/O Groups within a system is not supported.

Intercluster Metro Mirror

Intercluster Metro Mirror performs intercluster copying of a volume, in which one volume belongs to a system and the other volume belongs to a separate system.

Two IBM Spectrum Virtualize systems must be defined in a partnership, which must be performed on both systems to establish a fully functional Metro Mirror partnership.

By using standard single-mode connections, the supported distance between two systems in a Metro Mirror partnership is 10 km (6.2 miles), although greater distances can be achieved by using extenders. For extended distance solutions, contact your IBM representative.

Limit: When a local fabric and a remote fabric are connected for Metro Mirror purposes, the inter-switch link (ISL) hop count between a local node and a remote node cannot exceed seven.

10.7.5 Synchronous remote copy

Metro Mirror is a fully synchronous remote copy technique that ensures that writes are committed at both the master and auxiliary volumes before write completion is acknowledged to the host, but only if writes to the auxiliary volumes are possible.

Events, such as a loss of connectivity between systems, can cause mirrored writes from the master volume and the auxiliary volume to fail. In that case, Metro Mirror suspends writes to the auxiliary volume and enables I/O to the master volume to continue to avoid affecting the operation of the master volumes.

Figure 10-70 on page 538 shows how a write to the master volume is mirrored to the cache of the auxiliary volume before an acknowledgment of the write is sent back to the host that issued the write. This process ensures that the auxiliary is synchronized in real time if it is needed in a failover situation.

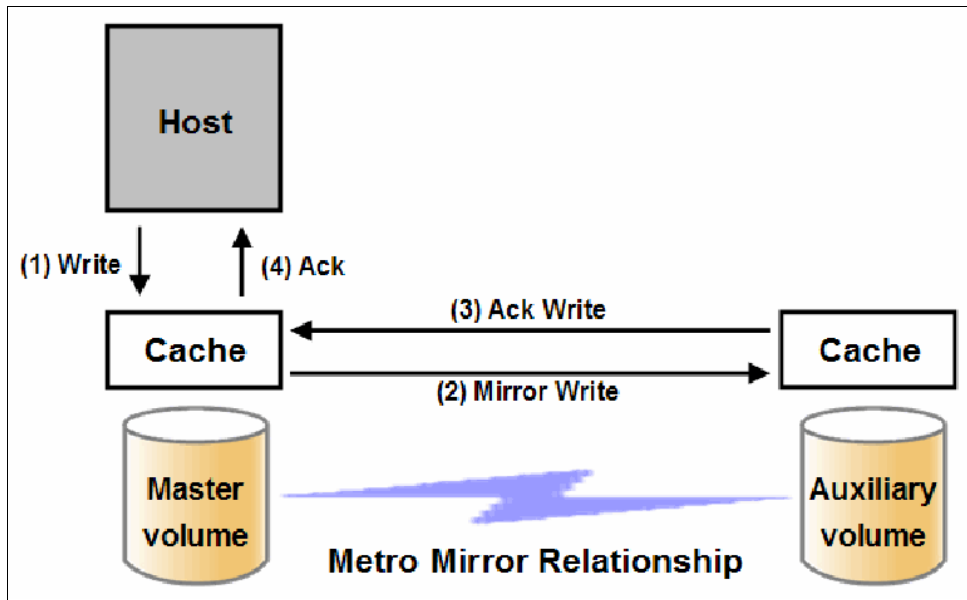


Figure 10-70 Write on volume in Metro Mirror relationship

However, this process also means that the application is exposed to the latency and bandwidth limitations (if any) of the communication link between the master and auxiliary volumes. This process might lead to unacceptable application performance, particularly when placed under peak load. Therefore, the use of traditional FC Metro Mirror has distance limitations that are based on your performance requirements. The IBM Spectrum Virtualize does not support more than 300 km (186.4 miles).

10.7.6 Metro Mirror features

The IBM Spectrum Virtualize Metro Mirror function supports the following features:

- ▶ Synchronous remote copy of volumes that are dispersed over metropolitan distances.
- ▶ The Metro Mirror relationships between volume pairs, with each volume in a pair that is managed by a Storwize system or IBM SAN Volume Controller system (requires V6.3.0 or later).
- ▶ Supports intracluster Metro Mirror where both volumes belong to the same system (and I/O Group).
- ▶ The IBM Spectrum Virtualize supports intercluster Metro Mirror where each volume belongs to a separate system. You can configure a specific system for partnership with another system. All intercluster Metro Mirror processing occurs between two IBM Spectrum Virtualize systems that are configured in a partnership.
- ▶ Intercluster and intracluster Metro Mirror can be used concurrently.
- ▶ The IBM Storwize V5000 Gen2 system does not require that a control network or fabric is installed to manage Metro Mirror. For intercluster Metro Mirror, the system maintains a control link between two systems. This control link is used to control the state and coordinate updates at either end. The control link is implemented on top of the same FC fabric connection that the IBM Storwize V5000 Gen2 system uses for Metro Mirror I/O.
- ▶ The IBM Spectrum Virtualize implements a configuration model that maintains the Metro Mirror configuration and state through major events, such as failover, recovery, and resynchronization, to minimize user configuration action through these events.

The IBM Spectrum Virtualize supports the resynchronization of changed data so that write failures that occur on the master or auxiliary volumes do not require a complete resynchronization of the relationship.

10.7.7 Metro Mirror attributes

The Metro Mirror function in IBM Spectrum Virtualize possesses the following attributes:

- ▶ A partnership is created between two IBM Storwize V5000 Gen 2 systems or an IBM SAN Volume Controller system and IBM Storwize V5000 Gen2 system that are operating in the replication layer (for intercluster Metro Mirror).
- ▶ A Metro Mirror relationship is created between two volumes of the same size.
- ▶ To manage multiple Metro Mirror relationships as one entity, relationships can be made part of a Metro Mirror Consistency Group, which ensures data consistency across multiple Metro Mirror relationships and provides ease of management.
- ▶ When a Metro Mirror relationship is started and when the background copy completes, the relationship becomes consistent and synchronized.
- ▶ After the relationship is synchronized, the auxiliary volume holds a copy of the production data at the primary, which can be used for DR.
- ▶ The auxiliary volume is in read-only mode when relationship is active.

- ▶ To access the auxiliary volume, the Metro Mirror relationship must be stopped with the access option enabled before write I/O is allowed to the auxiliary.
- ▶ The remote host server is mapped to the auxiliary volume, and the disk is available for I/O.

10.7.8 Practical use of Metro Mirror

The master volume is the production volume, and updates to this copy are mirrored in real time to the auxiliary volume. The contents of the auxiliary volume that existed when the relationship was created are deleted.

Switching copy direction: The copy direction for a Metro Mirror relationship can be switched so that the auxiliary volume becomes the master, and the master volume becomes the auxiliary, which is similar to the FlashCopy restore option. However, although the FlashCopy target volume can operate in read/write mode, the target volume of the started remote copy is always in read-only mode.

While the Metro Mirror relationship is active, the auxiliary volume is not accessible for host application write I/O at any time. The IBM Storwize V5000 Gen2 system allows read-only access to the auxiliary volume when it contains a consistent image. IBM Storwize allows boot time operating system discovery to complete without an error, so that any hosts at the secondary site can be ready to start the applications with minimum delay, if required.

For example, many operating systems must read LBA zero to configure a logical unit. Although read access is allowed at the auxiliary in practice, the data on the auxiliary volumes cannot be read by a host because most operating systems write a “dirty bit” to the file system when it is mounted. Because this write operation is not allowed on the auxiliary volume, the volume cannot be mounted.

This access is provided only where consistency can be ensured. However, coherency cannot be maintained between reads that are performed at the auxiliary and later write I/Os that are performed at the master.

To enable access to the auxiliary volume for host operations, you must stop the Metro Mirror relationship by specifying the `-access` parameter. While access to the auxiliary volume for host operations is enabled, the host must be instructed to mount the volume before the application can be started, or instructed to perform a recovery process.

For example, the Metro Mirror requirement to enable the auxiliary copy for access differentiates it from third-party mirroring software on the host, which aims to emulate a single, reliable disk regardless of what system is accessing it. Metro Mirror retains the property that there are two volumes in existence, but it suppresses one volume while the copy is being maintained.

The use of an auxiliary copy demands a conscious policy decision by the administrator that a failover is required. The tasks to be performed on the host that is involved in establishing the operation on the auxiliary copy are substantial. The goal is to make this copy rapid (much faster when compared to recovering from a backup copy) but not seamless.

The failover process can be automated through failover management software. The IBM Storwize V5000 Gen2 system provides SNMP traps and programming (or scripting) for the CLI to enable this automation.

10.7.9 Global Mirror overview

This section describes the Global Mirror copy service, which is an asynchronous remote copy service. This service provides and maintains a consistent mirrored copy of a source volume to a target volume.

Global Mirror establishes a Global Mirror relationship between two volumes. The volumes in a Global Mirror relationship are referred to as the *master* (source) volume and the *auxiliary* (target) volume, which is the same as Metro Mirror. Consistency Groups can be used to maintain data integrity for dependent writes, which is similar to FlashCopy Consistency Groups.

Global Mirror writes data to the auxiliary volume asynchronously, which means that host writes to the master volume provide the host with confirmation that the write is complete before the I/O completes on the auxiliary volume.

Global Mirror has the following characteristics:

- ▶ Near-zero RPO
- ▶ Asynchronous
- ▶ Production application performance that is affected by I/O sequencing preparation time

Intracluster Global Mirror

Although Global Mirror is available for intracluster, it has no functional value for production use. Intracluster Metro Mirror provides the same capability with less processor use. However, leaving this functionality in place simplifies testing and supports client experimentation and testing (for example, to validate server failover on a single test system). As with Intracluster Metro Mirror, you must consider the increase in the license requirement because source and target exist on the same IBM Spectrum Virtualize system.

Intercluster Global Mirror

Intercluster Global Mirror operations require a pair of IBM Spectrum Virtualize systems that are connected by several intercluster links. The two systems must be defined in a partnership to establish a fully functional Global Mirror relationship.

Limit: When a local fabric and a remote fabric are connected for Global Mirror purposes, the ISL hop count between a local node and a remote node must not exceed seven hops.

10.7.10 Asynchronous remote copy

Global Mirror is an asynchronous remote copy technique. In asynchronous remote copy, the write operations are completed on the primary site and the write acknowledgment is sent to the host before it is received at the secondary site. An update of this write operation is sent to the secondary site at a later stage, which provides the capability to perform remote copy over distances that exceed the limitations of synchronous remote copy.

The Global Mirror function provides the same function as Metro Mirror remote copy, but over long-distance links with higher latency without requiring the hosts to wait for the full round-trip delay of the long-distance link.

Figure 10-71 on page 542 shows that a write operation to the master volume is acknowledged back to the host that is issuing the write before the write operation is mirrored to the cache for the auxiliary volume.

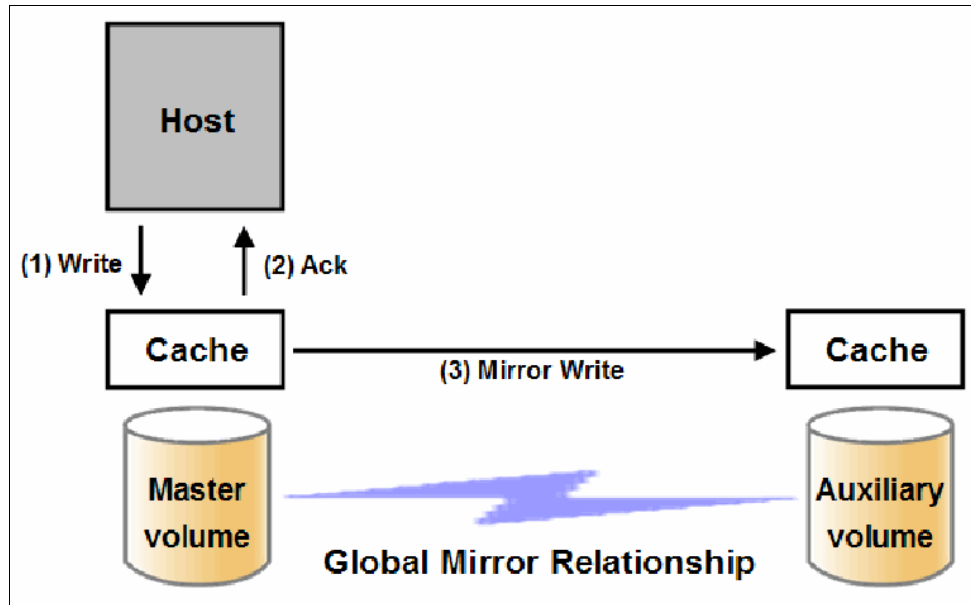


Figure 10-71 Global Mirror write sequence

The Global Mirror algorithms maintain a consistent image on the auxiliary always. They achieve this consistent image by identifying sets of I/Os that are active concurrently at the master, assigning an order to those sets, and applying those sets of I/Os in the assigned order at the secondary. As a result, Global Mirror maintains the features of Write Ordering and Read Stability.

The multiple I/Os within a single set are applied concurrently. The process that marshals the sequential sets of I/Os operates at the secondary system. Therefore, the process is not subject to the latency of the long-distance link. These two elements of the protocol ensure that the throughput of the total system can be grown by increasing system size while maintaining consistency across a growing data set.

Global Mirror write I/O from production system to a secondary system requires serialization and sequence-tagging before being sent across the network to a remote site (to maintain a write-order consistent copy of data).

To avoid affecting the production site, IBM Spectrum Virtualize supports more parallelism in processing and managing Global Mirror writes on the secondary system by using the following methods:

- ▶ Secondary system nodes store replication writes in new redundant non-volatile cache
- ▶ Cache content details are shared between nodes
- ▶ Cache content details are batched together to make node-to-node latency less of an issue
- ▶ Nodes intelligently apply these batches in parallel as soon as possible
- ▶ Nodes internally manage and optimize Global Mirror secondary write I/O processing

In a failover scenario where the secondary site must become the master source of data, certain updates might be missing at the secondary site. Therefore, any applications that use this data must have an external mechanism for recovering the missing updates and reapplying them; for example, a transaction log replay.

Global Mirror is supported over FC, FC over IP (FCIP), FC over Ethernet (FCoE), and native IP connections. The maximum supported round-trip latency between sites depends on the type of partnership between systems, the version of software, and the system hardware that is used.

Figure 10-72 lists the maximum round-trip latency. This restriction applies to all variant of remote mirroring. More configuration requirements and guidelines apply to systems that perform remote mirroring over extended distances, where the round-trip time is > 80 ms.

Software version	System node hardware	Partnership		
		FC	1 Gbps IP	10 Gbps IP
7.3.0 and earlier	All	80 ms	80 ms	10 ms
7.4.0 and later	• Storwize® V5000 Gen2	250 ms		
	All other models	80 ms		

Figure 10-72 Supported Remote mirroring latency

10.7.11 Global Mirror features

IBM Spectrum Virtualize Global Mirror supports the following features:

- ▶ Asynchronous remote copy of volumes that are dispersed over metropolitan-scale distances.
- ▶ The IBM Spectrum Virtualize implements the Global Mirror relationship between a volume pair, with each volume in the pair being managed by an IBM SAN Volume Controller or IBM Storwize system running IBM Spectrum Virtualize.
- ▶ The IBM Storwize V5000 Gen2 supports intracluster Global Mirror where both volumes belong to the same system (and I/O Group).
- ▶ The IBM Storwize V5000 Gen2 intercluster Global Mirror in which each volume belongs to its separate IBM Storwize V5000 Gen2 system. An IBM Storwize V5000 Gen2 system can be configured for partnership with 1 - 3 other systems. For more information about IP partnership restrictions, see 10.6.3, “IP partnership limitations” on page 526.
- ▶ Intercluster and intracluster Global Mirror can be used concurrently, but not for the same volume.
- ▶ The IBM Storwize V5000 Gen2 system does not require a control network or fabric to be installed to manage Global Mirror. For intercluster Global Mirror, the IBM Storwize V5000 Gen2 system maintains a control link between the two systems. This control link is used to control the state and to coordinate the updates at either end. The control link is implemented on top of the same FC fabric connection that the IBM Storwize V5000 Gen2 system uses for Global Mirror I/O.
- ▶ The IBM Storwize V5000 Gen2 system implements a configuration model that maintains the Global Mirror configuration and state through major events, such as failover, recovery, and resynchronization, to minimize user configuration action through these events.
- ▶ The IBM Storwize V5000 Gen2 system implements flexible resynchronization support, enabling it to resynchronize volume pairs that experienced write I/Os to both disks, and to resynchronize only those regions that changed.
- ▶ An optional feature for Global Mirror is a delay simulation to be applied on writes that are sent to auxiliary volumes. It is useful in intracluster scenarios for testing purposes.

Colliding writes

The Global Mirror algorithm requires that only a single write is active on a volume. I/Os that overlap an active I/O are sequential and, this is called *colliding writes*. If a further write is received from a host while the auxiliary write is still active, the new host write is delayed until the auxiliary write is complete. This rule is needed if a series of writes to the auxiliary must be tried again and, is called *reconstruction*. Conceptually, the data for reconstruction comes from the master volume.

If multiple writes are allowed to be applied to the master for a sector, only the most recent write gets the correct data during reconstruction. If reconstruction is interrupted for any reason, the intermediate state of the auxiliary is inconsistent. Applications that deliver such write activity do not achieve the performance that Global Mirror is intended to support. A volume statistic is maintained about the frequency of these collisions.

An attempt is made to allow multiple writes to a single location to be outstanding in the Global Mirror algorithm. There is still a need for master writes to be sequential, and the intermediate states of the master data must be kept in a non-volatile journal while the writes are outstanding to maintain the correct write ordering during reconstruction. Reconstruction must never overwrite data on the auxiliary with an earlier version. The volume statistic that is monitoring colliding writes is now limited to those writes that are not affected by this change.

Figure 10-73 shows a colliding write sequence example.

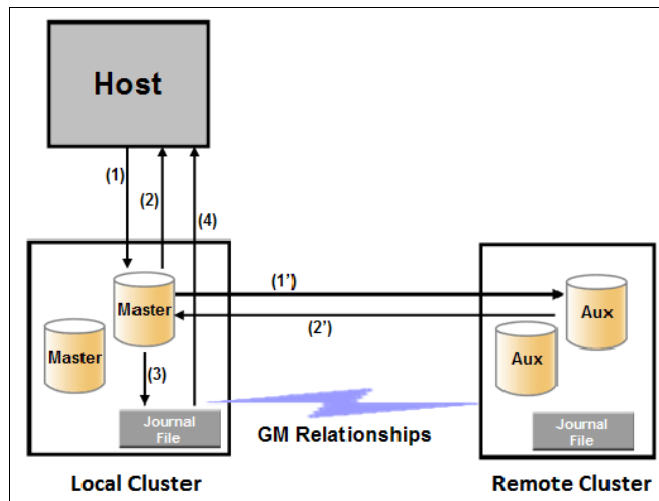


Figure 10-73 Colliding writes example

The following numbers correspond to the numbers that are shown in Figure 10-73:

- ▶ (1) The first write is performed from the host to LBA X.
- ▶ (2) The host is provided acknowledgment that the write completed, even though the mirrored write to the auxiliary volume is not yet complete.
- ▶ (1') and (2') occur asynchronously with the first write.
- ▶ (3) The second write is performed from the host also to LBA X. If this write occurs before (2'), the write is written to the journal file.
- ▶ (4) The host is provided acknowledgment that the second write is complete.

Delay simulation

An optional feature for Global Mirror enables a delay simulation to be applied on writes that are sent to auxiliary volumes. This feature enables you to perform testing that detects colliding writes. Therefore, you can use this feature to test an application before the full deployment of the feature. The feature can be enabled separately for each of the intracluster or intercluster Global Mirrors.

You specify the delay setting by using the `chsystem` command and view the delay by using the `lssystem` command. The `gm_intra_cluster_delay_simulation` field expresses the amount of time that intracluster auxiliary I/Os are delayed. The `gm_inter_cluster_delay_simulation` field expresses the amount of time that intercluster auxiliary I/Os are delayed. A value of zero disables the feature.

Tip: If you are experiencing repeated problems with the delay on your link, make sure that the delay simulator was properly disabled.

10.7.12 Using Change Volumes with Global Mirror

Global Mirror is designed to achieve an RPO as low as possible so that data is as up-to-date as possible. This design places several strict requirements on your infrastructure. In certain situations with low network link quality, congested hosts, or overloaded hosts, you might be affected by multiple 1920 congestion errors.

Congestion errors occur in the following primary situations:

- ▶ Congestion at the source site through the host or network
- ▶ Congestion in the network link or network path
- ▶ Congestion at the target site through the host or network

Global Mirror includes functionality that is designed to address the following conditions, which might negatively affect certain Global Mirror implementations:

- ▶ The estimation of the bandwidth requirements tends to be complex.
- ▶ Ensuring the latency and bandwidth requirements can be met is often difficult.
- ▶ Congested hosts on the source or target site can cause disruption.
- ▶ Congested network links can cause disruption with only intermittent peaks.

To address these issues, *Change Volumes* were added as an option for Global Mirror relationships. Change Volumes use the FlashCopy functionality, but they cannot be manipulated as FlashCopy volumes because they are for a special purpose only. Depending on the cycling mode defined, Change Volumes replicate point-in-time images on a cycling period.

Note: The cycling mode can be either `none` or `multi`. When cycling mode is set to `none`, the Global Mirror behaves identically to Global Mirror without Change Volumes. When cycling mode is set to `multi`, the Global Mirror behaves as described in this section.

The cycling mode can be changed only when the relationship is stopped and in `consistent_stopped` or `inconsistent_stopped` status.

The default cycling period is 300 seconds.

Your change rate needs to include only the condition of the data at the point-in-time that the image was taken, rather than all the updates during the period. The use of this function can provide significant reductions in replication volume.

Global Mirror with Change Volumes has the following characteristics:

- ▶ Larger RPO
- ▶ Point-in-time copies
- ▶ Asynchronous

- Possible system performance resource requirements because point-in-time copies are created locally

Figure 10-74 shows a simple Global Mirror relationship without Change Volumes.

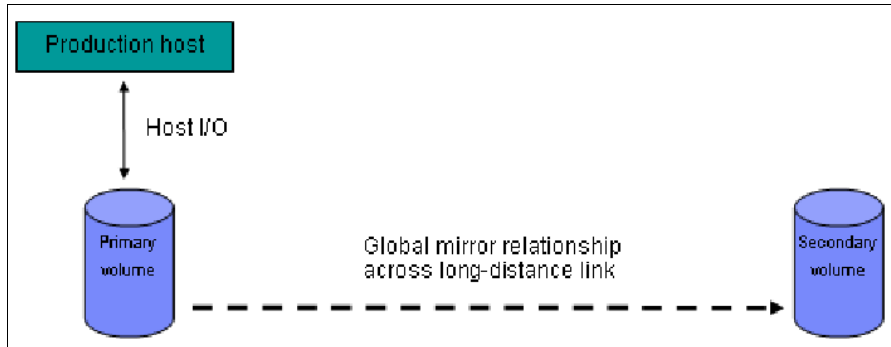


Figure 10-74 Global Mirror without Change Volumes

With Change Volumes, this environment looks as it is shown in Figure 10-75.

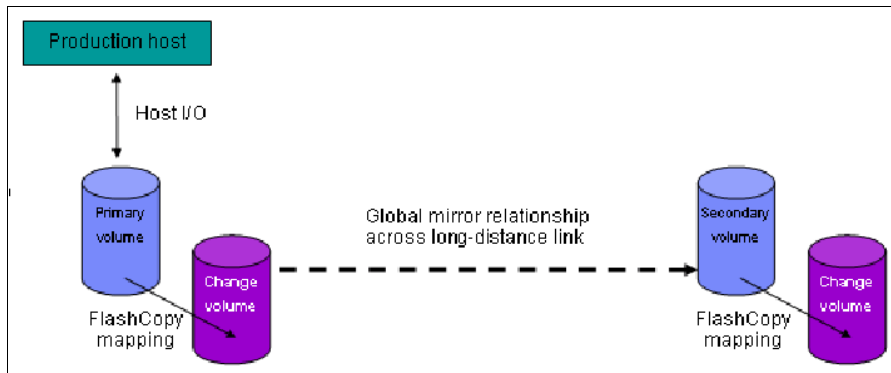


Figure 10-75 Global Mirror with Change Volumes

With Change Volumes, a FlashCopy mapping exists between the primary volume and the primary Change Volume. The mapping is updated on the cycling period (60 seconds to one day). The primary Change Volume is then replicated to the secondary Global Mirror volume at the target site, which is then captured in another Change Volume on the target site. This approach provides an always consistent image at the target site and protects your data from being inconsistent during resynchronization.

Figure 10-76 shows how Change Volumes might save you replication traffic.

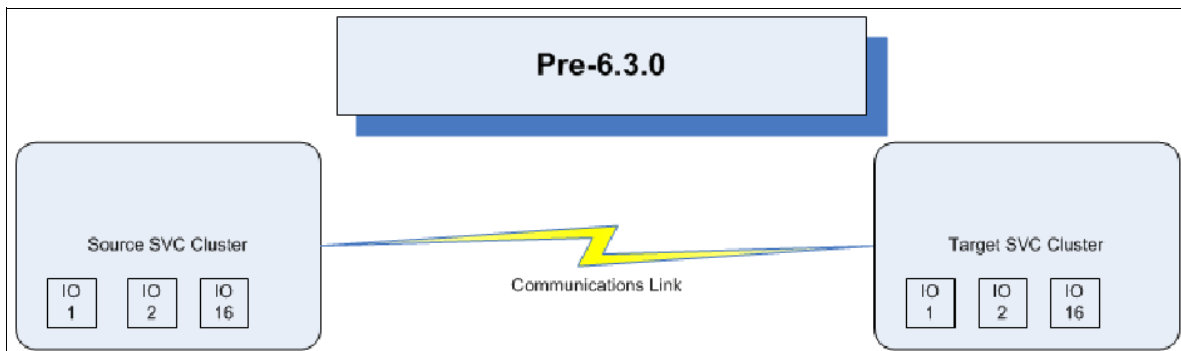


Figure 10-76 Global Mirror I/O replication without Change Volumes

In Figure 10-76 on page 546, you can see several I/Os on the source and the same number on the target, and in the same order. Assuming that this data is the same set of data being updated repeatedly, this approach results in wasted network traffic. The I/O can be completed much more efficiently, as shown in Figure 10-77.

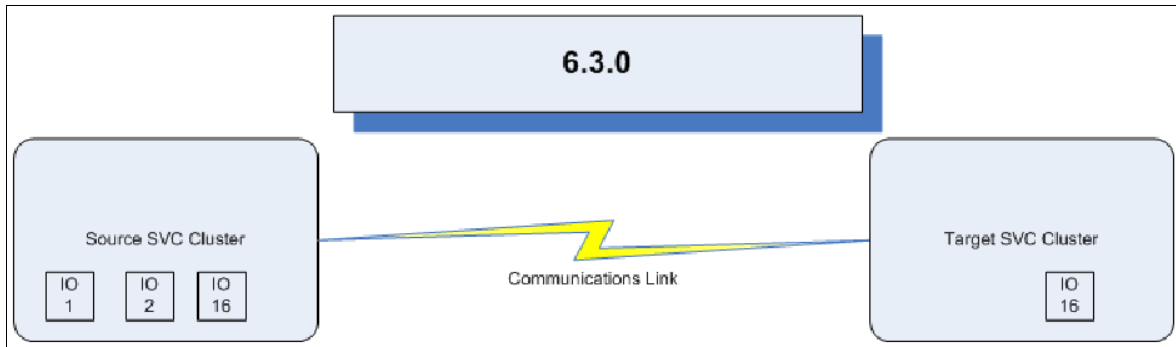


Figure 10-77 Global Mirror I/O with Change Volumes V6.3.0 and beyond

In Figure 10-77, the same data is being updated repeatedly. Therefore, Change Volumes demonstrate significant I/O transmission savings by needing to send I/O number 16 only, which was the last I/O before the cycling period.

You can adjust the cycling period by using the `chrcrelationship -cycleperiodseconds <60 - 86400>` command from the CLI. If a copy does not complete in the cycle period, the next cycle does not start until the prior cycle completes. For this reason, the use of Change Volumes gives you the following possibilities for RPO:

- ▶ If your replication completes in the cycling period, your RPO is twice the cycling period.
- ▶ If your replication does not complete within the cycling period, RPO is twice the completion time. The next cycling period starts immediately after the prior cycling period is finished.

Carefully consider your business requirements versus the performance of Global Mirror with Change Volumes. Global Mirror with Change Volumes increases the intercluster traffic for more frequent cycling periods. Therefore, selecting the shortest cycle periods possible is not always the answer. In most cases, the default must meet requirements and perform well.

Important: When you create your Global Mirror volumes with Change Volumes, make sure that you remember to select the Change Volume on the auxiliary (target) site. Failure to do so leaves you exposed during a resynchronization operation.

10.7.13 Distribution of work among nodes

For the best performance, MM/GM volumes must have their preferred nodes evenly distributed among the nodes of the systems. Each volume within an I/O Group has a preferred node property that can be used to balance the I/O load between nodes in that group. MM/GM also uses this property to route I/O between systems.

If this preferred practice is not maintained (for example, source volumes are assigned to only one node in the I/O group), you can change the preferred node for each volume to distribute volumes evenly between the nodes. You can also change the preferred node for volumes that are in a remote copy relationship without affecting the host I/O to a particular volume.

The remote copy relationship type does not matter. (The remote copy relationship type can be MM, GM, or GM with Change Volumes.) You can change the preferred node to the source and target volumes that are participating in the remote copy relationship.

10.7.14 Background copy performance

The background copy performance is subject to sufficient Redundant Array of Independent Disks (RAID) controller bandwidth. Performance is also subject to other potential bottlenecks, such as the intercluster fabric, and possible contention from host I/O for the IBM SAN Volume Controller or IBM Storwize V5000 Gen2 system bandwidth resources.

Background copy I/O is scheduled to avoid bursts of activity that might have an adverse effect on system behavior. An entire grain of tracks on one volume is processed at around the same time, but not as a single I/O. Double buffering is used to try to use sequential performance within a grain. However, the next grain within the volume might not be scheduled for some time. Multiple grains might be copied simultaneously, and might be enough to satisfy the requested rate, unless the available resources cannot sustain the requested rate.

Global Mirror paces the rate at which background copy is performed by the appropriate relationships. Background copy occurs on relationships that are in the `InconsistentCopying` state with a status of `Online`.

The quota of background copy (configured on the intercluster link) is divided evenly between all nodes that are performing background copy for one of the eligible relationships. This allocation is made irrespective of the number of disks for which the node is responsible. Each node in turn divides its allocation evenly between the multiple relationships that are performing a background copy.

The default value of the background copy is 25 megabytes per second (MBps), per volume.

Important: The background copy value is a system-wide parameter that can be changed dynamically, but only on a per-system basis and not on a per-relationship basis. Therefore, the copy rate of all relationships changes when this value is increased or decreased. In systems with many remote copy relationships, increasing this value might affect overall system or intercluster link performance. The background copy rate can be changed from 1 - 1000 MBps.

10.7.15 Thin-provisioned background copy

Metro Mirror and Global Mirror relationships preserve the space-efficiency of the master. Conceptually, the background copy process detects a deallocated region of the master and sends a special *zero buffer* to the auxiliary.

If the auxiliary volume is thin-provisioned and the region is deallocated, the special buffer prevents a write and, therefore, an allocation. If the auxiliary volume is not thin-provisioned or the region in question is an allocated region of a thin-provisioned volume, a buffer of “real” zeros is synthesized on the auxiliary and written as normal.

10.7.16 Methods of synchronization

This section describes two methods that can be used to establish a synchronized relationship.

Full synchronization after creation

The full synchronization after creation method is the default method. It is the simplest method in that it requires no administrative activity apart from issuing the necessary commands. However, in certain environments, the available bandwidth can make this method unsuitable.

Use the following command sequence for a single relationship:

- ▶ Run `mkrcrelationship` without specifying the `-sync` option.
- ▶ Run `starttrcrelationship` without specifying the `-clean` option.

Synchronized before creation

In this method, the administrator must ensure that the master and auxiliary volumes contain identical data before creating the relationship by using the following technique:

- ▶ Both disks are created with the security delete feature to make all data zero.
- ▶ A complete tape image (or other method of moving data) is copied from one disk to the other disk.

With this technique, do not allow I/O on the master or auxiliary before the relationship is established. Then, the administrator must run the following commands:

- ▶ Run `mkrcrelationship` with the `-sync` flag.
- ▶ Run `starttrcrelationship` without the `-clean` flag.

Important: Failure to perform these steps correctly can cause MM/GM to report the relationship as consistent when it is not, which creates a data loss or data integrity exposure for hosts that are accessing data on the auxiliary volume.

10.7.17 Practical use of Global Mirror

The practical use of Global Mirror is similar to the Metro Mirror that is described in 10.7.8, “Practical use of Metro Mirror” on page 540. The main difference between the two remote copy modes is that Global Mirror and Global Mirror with Change Volumes are mostly used on much larger distances than Metro Mirror.

Weak link quality or insufficient bandwidth between the primary and secondary sites can also be a reason to prefer asynchronous Global Mirror over synchronous Metro Mirror. Otherwise, the use cases for Metro Mirror and Global Mirror are the same.

10.7.18 Valid combinations of FlashCopy, Metro Mirror, and Global Mirror

Table 10-9 lists the combinations of FlashCopy and Metro Mirror or Global Mirror functions that are valid for a single volume.

Table 10-9 Valid combination for a single volume

FlashCopy	Metro Mirror or Global Mirror source	Metro Mirror or Global Mirror target
FlashCopy Source	Supported	Supported
FlashCopy Target	Supported	Not supported

10.7.19 Remote Copy configuration limits

Table 10-10 lists the Metro Mirror and Global Mirror configuration limits.

Table 10-10 Metro Mirror configuration limits

Parameter	Value
Number of Metro Mirror or Global Mirror Consistency Groups per system	256
Number of Metro Mirror or Global Mirror relationships per system	4096
Number of Metro Mirror or Global Mirror relationships per Consistency Group	No limit is imposed beyond the Remote Copy relationships per system limit.
Total volume size per I/O Group	There is a per-I/O Group limit of 1024 terabytes (TB) on the quantity of master and auxiliary volume address spaces that can participate in Metro Mirror and Global Mirror relationships. This maximum configuration uses all 512 MiB of bitmap space for the I/O Group and allows 10 MiB of space for all remaining copy services features.
Total number of Global Mirror with Change Volumes relationships per system	256

For more information about the configuration limits, search for “Configuration Limits and Restrictions for IBM Storwize V5000” at this IBM Knowledge Center [web page](#).

10.7.20 Remote Copy states and events

This section describes the various states of a MM/GM relationship and the conditions that cause them to change. In Figure 10-78, the MM/GM relationship diagram shows an overview of the status that can apply to a MM/GM relationship in a connected state.

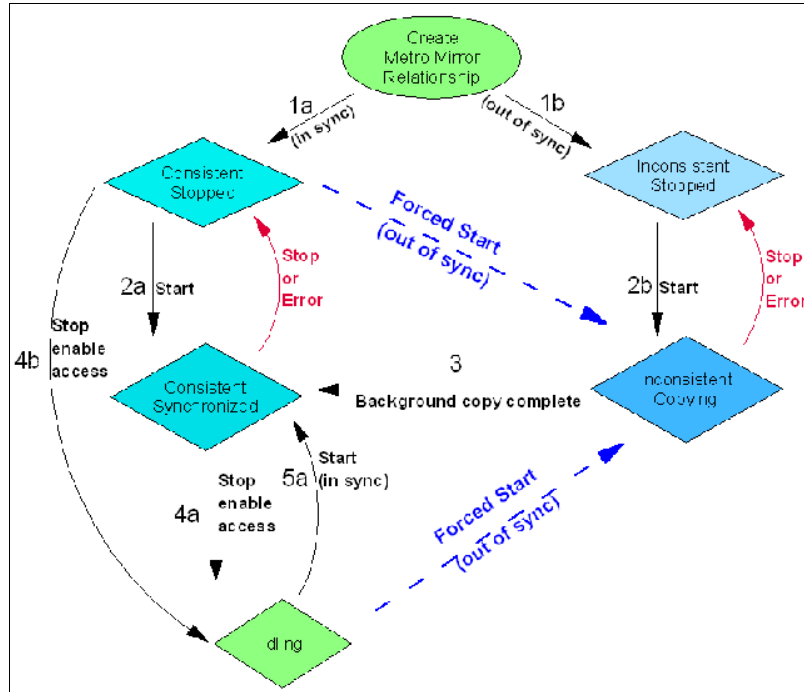


Figure 10-78 Metro Mirror or Global Mirror mapping state diagram

When the MM/GM relationship is created, you can specify whether the auxiliary volume is already in sync with the master volume, and the background copy process is then skipped. This capability is useful when MM/GM relationships are established for volumes that were created with the format option.

The following step identifiers are shown in Figure 10-78:

- ▶ Step 1:
 - a. The MM/GM relationship is created with the **-sync** option, and the MM/GM relationship enters the `ConsistentStopped` state.
 - b. The MM/GM relationship is created without specifying that the master and auxiliary volumes are in sync, and the MM/GM relationship enters the `InconsistentStopped` state.
- ▶ Step 2:
 - a. When a MM/GM relationship is started in the `ConsistentStopped` state, the MM/GM relationship enters the `ConsistentSynchronized` state. Therefore, no updates (write I/O) were performed on the master volume while in the `ConsistentStopped` state. Otherwise, the **-force** option must be specified, and the MM/GM relationship then enters the `InconsistentCopying` state while the background copy is started.
 - b. When a MM/GM relationship is started in the `InconsistentStopped` state, the MM/GM relationship enters the `InconsistentCopying` state while the background copy is started.

- ▶ Step 3:
When the background copy completes, the MM/GM relationship transitions from the `InconsistentCopying` state to the `ConsistentSynchronized` state.
- ▶ Step 4:
 - a. When a MM/GM relationship is stopped in the `ConsistentSynchronized` state, the MM/GM relationship enters the `Idling` state when you specify the `-access` option, which enables write I/O on the auxiliary volume.
 - b. When a MM/GM relationship is stopped in the `ConsistentSynchronized` state without an `-access` parameter, the auxiliary volumes remain read-only and the state of the relationship changes to `ConsistentStopped`.
 - c. To enable write I/O on the auxiliary volume, when the MM/GM relationship is in the `ConsistentStopped` state, issue the `svctask stopprcrelationship` command, which specifies the `-access` option, and the MM/GM relationship enters the `Idling` state.
- ▶ Step 5:
 - a. When a MM/GM relationship is started from the `Idling` state, you must specify the `-primary` argument to set the copy direction. If no write I/O was performed (to the master or auxiliary volume) while in the `Idling` state, the MM/GM relationship enters the `ConsistentSynchronized` state.
 - b. If write I/O was performed to the master or auxiliary volume, the `-force` option must be specified and the MM/GM relationship then enters the `InconsistentCopying` state while the background copy is started. The background process copies only the data that changed on the primary volume while the relationship was stopped.

Stop on Error

When a MM/GM relationship is stopped (intentionally, or because of an error), the state changes. For example, the MM/GM relationships in the `ConsistentSynchronized` state enter the `ConsistentStopped` state, and the MM/GM relationships in the `InconsistentCopying` state enter the `InconsistentStopped` state.

If the connection is broken between the two systems that are in a partnership, all (intercluster) MM/GM relationships enter a `Disconnected` state. For more information, see “Connected versus disconnected” on page 552.

Common states: Stand-alone relationships and Consistency Groups share a common configuration and state model. All MM/GM relationships in a Consistency Group have the same state as the Consistency Group.

State overview

In the following sections, we provide an overview of the various MM/GM states.

Connected versus disconnected

Under certain error scenarios (for example, a power failure at one site that causes one complete system to disappear), communications between two systems in an MM/GM relationship can be lost. Alternatively, the fabric connection between the two systems might fail, which leaves the two systems running, but they cannot communicate with each other.

When the two systems can communicate, the systems and the relationships that spans them are described as *connected*. When they cannot communicate, the systems and the relationships spanning them are described as *disconnected*.

In this state, both systems are left with fragmented relationships and are limited regarding the configuration commands that can be performed. The disconnected relationships are shown as having a changed state. The new states describe what is known about the relationship and the configuration commands that are permitted.

When the systems can communicate again, the relationships are reconnected. MM/GM automatically reconciles the two state fragments, considering any configuration or other event that occurred while the relationship was disconnected. As a result, the relationship can return to the state that it was in when it became disconnected, or it can enter a new state.

Relationships that are configured between volumes in the same IBM Storwize V5000 Gen2 system (intracluster) are never described as being in a disconnected state.

Consistent versus inconsistent

Relationships that contain volumes that are operating as secondaries can be described as being consistent or inconsistent. Consistency Groups that contain relationships can also be described as being consistent or inconsistent. The consistent or inconsistent property describes the relationship of the data on the auxiliary to the data on the master volume. It can be considered a property of the auxiliary volume.

An auxiliary volume is described as *consistent* if it contains data that might be read by a host system from the master if power failed at an imaginary point while I/O was in progress, and power was later restored. This imaginary point is defined as the *recovery point*.

The requirements for consistency are expressed regarding activity at the master up to the recovery point. The auxiliary volume contains the data from all of the writes to the master for which the host received successful completion and that data was not overwritten by a subsequent write (before the recovery point).

Consider writes for which the host did not receive a successful completion (that is, it received bad completion or no completion at all). If the host then performed a read from the master of that data that returned successful completion and no later write was sent (before the recovery point), the auxiliary contains the same data as the data that was returned by the read from the master.

From the point of view of an application, consistency means that an auxiliary volume contains the same data as the master volume at the recovery point (the time at which the imaginary power failure occurred). If an application is designed to handle with an unexpected power failure, this assurance of consistency means that the application can use the auxiliary and begin operation as though it was restarted after the hypothetical power failure. Again, maintaining the application write ordering is the key property of consistency.

If a relationship (or set of relationships) is inconsistent and an attempt is made to start an application by using the data in the secondaries, the following outcomes are possible:

- ▶ The app might decide that the data is corrupted and crash or exit with an event code.
- ▶ The application might fail to detect that the data is corrupted and return erroneous data.
- ▶ The application might work without a problem.

Because of the risk of data corruption (and in particular undetected data corruption), MM/GM strongly enforces the concept of consistency and prohibits access to inconsistent data.

Consistency as a concept can be applied to a single relationship or a set of relationships in a Consistency Group. Write ordering is a concept that an application can maintain across several disks that are accessed through multiple systems. Therefore, consistency must operate across all of those disks.

When you are deciding how to use Consistency Groups, the administrator must consider the scope of an application's data and consider all of the interdependent systems that communicate and exchange information.

If two programs or systems communicate and store details as a result of the information exchanged, either of the following actions might occur:

- ▶ All of the data that is accessed by the group of systems must be placed into a single Consistency Group.
- ▶ The systems must be recovered independently (each within its own Consistency Group). Then, each system must perform recovery with the other applications to become consistent with them.

Consistent versus synchronized

A copy that is consistent and up-to-date is described as *synchronized*. In a synchronized relationship, the master and auxiliary volumes differ only in regions where writes are outstanding from the host.

Consistency does not mean that the data is up-to-date. A copy can be consistent and yet contain data that was frozen at a point in the past. Write I/O might continue to a master, but not be copied to the auxiliary. This state arises when it becomes impossible to keep data up-to-date and maintain consistency. An example is a loss of communication between systems when you are writing to the auxiliary.

When communication is lost for an extended period, MM/GM tracks the changes that occurred on the master, but not the order or the details of such changes (write data). When communication is restored, it is impossible to synchronize the auxiliary without sending write data to the auxiliary out of order. Therefore, consistency is lost.

The following policies can be used to cope with this situation:

- ▶ Make a point-in-time copy of the consistent auxiliary before you allow the auxiliary to become inconsistent. If there is a disaster before consistency is achieved again, the point-in-time copy target provides a consistent (although out-of-date) image.
- ▶ Accept the loss of consistency and the loss of a useful auxiliary while synchronizing the auxiliary.

Detailed states

In the following sections, we describe the states that are portrayed to the user for either Consistency Groups or relationships. We also describe information that is available in each state. The major states are designed to provide guidance about the available configuration commands.

InconsistentStopped

`InconsistentStopped` is a connected state. In this state, the master is accessible for read and write I/O, but the auxiliary is not accessible for read or write I/O. A copy process must be started to make the auxiliary consistent.

This state is entered when the relationship or Consistency Group was `InconsistentCopying` and suffered a persistent error or received a **stop** command that caused the copy process to stop.

A **start** command causes the relationship or Consistency Group to move to the `InconsistentCopying` state. A **stop** command is accepted, but has no effect.

If the relationship or Consistency Group becomes disconnected, the auxiliary side transitions to `InconsistentDisconnected`. The master side transitions to `IdlingDisconnected`.

InconsistentCopying

`InconsistentCopying` is a connected state. In this state, the master is accessible for read and write I/O, but the auxiliary is not accessible for read or write I/O. This state is entered after a **start** command is issued to an `InconsistentStopped` relationship or a Consistency Group.

It is also entered when a forced start is issued to an `Idling` or `ConsistentStopped` relationship or Consistency Group. In this state, a background copy process runs that copies data from the master to the auxiliary volume.

In the absence of errors, an `InconsistentCopying` relationship is active, and the copy progress increases until the copy process completes. In certain error situations, the copy progress might halt or even regress.

A persistent error or **stop** command places the relationship or Consistency Group into an `InconsistentStopped` state. A **start** command is accepted but has no effect.

If the background copy process completes on a stand-alone relationship or on all relationships for a Consistency Group, the relationship or Consistency Group transitions to the `ConsistentSynchronized` state.

If the relationship or Consistency Group becomes disconnected, the auxiliary side transitions to `InconsistentDisconnected`. The master side transitions to `IdlingDisconnected`.

ConsistentStopped

`ConsistentStopped` is a connected state. In this state, the auxiliary contains a consistent image, but it might be out-of-date in relation to the master. This state can arise when a relationship was in a `ConsistentSynchronized` state and experienced an error that forces a Consistency Freeze. It can also arise when a relationship is created with a `CreateConsistentFlag` set to `TRUE`.

Normally, write activity that follows an I/O error causes updates to the master, and the auxiliary is no longer synchronized. In this case, consistency must be given up for a period to reestablish synchronization. You must use a **start** command with the **-force** option to acknowledge this condition, and the relationship or Consistency Group transitions to `InconsistentCopying`. Enter this command only after all outstanding events are repaired.

In the unusual case where the master and the auxiliary are still synchronized (perhaps following a user stop, and no further write I/O was received), a **start** command takes the relationship to `ConsistentSynchronized`. No **-force** option is required. Also, in this case, you can enter a **switch** command that moves the relationship or Consistency Group to `ConsistentSynchronized` and reverses the roles of the master and the auxiliary.

If the relationship or Consistency Group becomes disconnected, the auxiliary transitions to `ConsistentDisconnected`. The master transitions to `IdlingDisconnected`.

An informational status log is generated whenever a relationship or Consistency Group enters the `ConsistentStopped` state with a status of `Online`. You can configure this event to generate an SNMP trap that can be used to trigger automation or manual intervention to issue a **start** command following a loss of synchronization.

ConsistentSynchronized

ConsistentSynchronized is a connected state. In this state, the master volume is accessible for read and write I/O, and the auxiliary volume is accessible for read-only I/O. Writes that are sent to the master volume are also sent to the auxiliary volume. Either successful completion must be received for both writes, the write must be failed to the host, or a state must transition out of the ConsistentSynchronized state before a write is completed to the host.

A **stop** command takes the relationship to the ConsistentStopped state. A **stop** command with the **-access** parameter takes the relationship to the Idling state.

A **switch** command leaves the relationship in the ConsistentSynchronized state, but it reverses the master and auxiliary roles (it switches the direction of replicating data). A **start** command is accepted, but has no effect.

If the relationship or Consistency Group becomes disconnected, the same transitions are made as for ConsistentStopped.

Idling

Idling is a connected state. Both master and auxiliary volumes operate in the master role. Therefore, both master and auxiliary volumes are accessible for write I/O.

In this state, the relationship or Consistency Group accepts a **start** command. MM/GM maintains a record of regions on each disk that received write I/O while they were idling. This record is used to determine what areas must be copied following a **start** command.

The **start** command must specify the new copy direction. A **start** command can cause a loss of consistency if either volume in any relationship received write I/O, which is indicated by the Synchronized status. If the **start** command leads to loss of consistency, you must specify the **-force** parameter.

Following a **start** command, the relationship or Consistency Group transitions to ConsistentSynchronized if there is no loss of consistency, or to InconsistentCopying if there is a loss of consistency.

Also, the relationship or Consistency Group accepts a **-clean** option on the **start** command while in this state. If the relationship or Consistency Group becomes disconnected, both sides change their state to IdlingDisconnected.

IdlingDisconnected

IdlingDisconnected is a disconnected state. The target volumes in this half of the relationship or Consistency Group are all in the master role and accept read or write I/O.

The priority in this state is to recover the link to restore the relationship or consistency.

No configuration activity is possible (except for deletes or stops) until the relationship becomes connected again. At that point, the relationship transitions to a connected state. The exact connected state that is entered depends on the state of the other half of the relationship or Consistency Group, which depends on the following factors:

- ▶ The state when it became disconnected
- ▶ The write activity since it was disconnected
- ▶ The configuration activity since it was disconnected

If both halves are IdlingDisconnected, the relationship becomes Idling when it is reconnected.

While `IdlingDisconnected`, if a write I/O is received that causes the loss of synchronization (synchronized attribute transitions from `true` to `false`) and the relationship was not stopped (either through a user stop or a persistent error), an event is raised to notify you of the condition. This same event also is raised when this condition occurs for the `ConsistentSynchronized` state.

When the relationship or Consistency Group becomes connected again, the relationship becomes `InconsistentCopying` automatically unless either of the following conditions are true:

- ▶ The relationship was `InconsistentStopped` when it became disconnected.
- ▶ The user issued a **stop** command while disconnected.

In either case, the relationship or Consistency Group becomes `InconsistentStopped`.

ConsistentDisconnected

`ConsistentDisconnected` is a disconnected state. The target volumes in this half of the relationship or Consistency Group are all in the auxiliary role, and accept read I/O but *not* write I/O.

This state is entered from `ConsistentSynchronized` or `ConsistentStopped` when the auxiliary side of a relationship becomes disconnected.

In this state, the relationship or Consistency Group displays an attribute of `FreezeTime`, which is the point when Consistency was frozen. When it is entered from `ConsistentStopped`, it retains the time that it had in that state. When it is entered from `ConsistentSynchronized`, the `FreezeTime` shows the last time at which the relationship or Consistency Group was known to be consistent. This time corresponds to the time of the last successful heartbeat to the other system.

A **stop** command with the `-access` flag set to `true` transitions the relationship or Consistency Group to the `IdlingDisconnected` state. This state allows write I/O to be performed to the auxiliary volume and is used as part of a DR scenario.

When the relationship or Consistency Group becomes connected again, the relationship or Consistency Group becomes `ConsistentSynchronized` only if this action does not lead to a loss of consistency. The following conditions must be true:

- ▶ The relationship was `ConsistentSynchronized` when it became disconnected.
- ▶ No writes received successful completion at the master while disconnected.

Otherwise, the relationship becomes `ConsistentStopped`. The `FreezeTime` setting is retained.

Empty

This state applies only to Consistency Groups. It is the state of a Consistency Group that has no relationships and no other state information to show. It is entered when a Consistency Group is first created. It is exited when the first relationship is added to the Consistency Group, at which point the state of the relationship becomes the state of the Consistency Group.

10.8 Consistency protection for Remote and Global mirror

Before V7.8.1, Metro Mirror and regular Global Mirror relationships and consistency groups stop when one of the following events occurs:

- ▶ The link between systems goes down
- ▶ A secondary volume goes offline

Consistency of the secondary volume is lost during resynchronization, so the relationship is automatically stopped and a 1720 error is raised, which requires the user to restart the relationship manually.

Global Mirror with Change Volumes (GMCV) relationships use a secondary change volume to retain a consistent copy during resync and automatically restart when they can.

From V7.8.1, Metro Mirror and regular Global Mirror also behave like GMCV relationship if a secondary change volume is configured which does the following:

- ▶ Makes Metro Mirror and Global Mirror more suited to links with intermittent connectivity and IP replication
- ▶ Stop as before if no secondary change volume configured

The consistency protection mechanism for Metro Mirror and regular Global Mirror uses change volumes and has following characteristics:

- ▶ It is a tweak to Metro Mirror and Global Mirror copy types by using technology that is in Global Mirror with Change Volumes
- ▶ Does not need FlashCopy license
- ▶ Uses two FlashCopy maps per relationship per system (so maximum of 2500 relationships on a 10k volume-capable system)
- ▶ Supported on all systems that can have remote mirroring license
- ▶ Requires both participating systems to be at V7.8.1 or later

Consistency protection for Metro or regular Global Mirror can be enabled by configuring a secondary change volume and no further configuration is needed to enable this behavior. All relationships in a consistency group must be so configured for this behavior to work on any relationship in the consistency group.

Table 10-11 lists the events and the expected behavior when consistency protection mechanism is enabled for metro mirror and regular global mirror.

Table 10-11 Events and expected behavior

Event	Expected behavior for the relationship
Link down or secondary volume offline Relationship is consistent_synchronized (started and in sync)	<ul style="list-style-type: none"> ▶ Retain the secondary consistent copy ▶ Prepare for resynchronization ▶ Go to the consistent_copying state ▶ Automatically resume replication as and when possible ▶ Return to consistent_synchronized when complete

Event	Expected behavior for the relationship
Relationship is restarted Relationship was stopped (consistent_stopped or idling) and the two copies are different	<ul style="list-style-type: none"> ▶ Retain the secondary consistent copy ▶ Prepare for resynchronization ▶ Go to the consistent_copying state ▶ Replicate differences as and when possible ▶ Return to consistent_synchronized when complete

Read/write access can be enabled as normal during resynchronization, which rewinds the secondary to the last consistent image. The hosts that are reading the secondary volume during resynchronization see data from the last consistent image. Also, at the end of the cycle, data might need to be cleaned from the change volume to the next FlashCopy map in the cascade.

Note: A change volume should be created as thin provisioned, but in theory, can grow to 100%.

A change volume must be:

- ▶ Used by the relationship that owns it.
- ▶ In the same I/O group as the associated master or auxiliary volume.
- ▶ The same size as the associated master or auxiliary volume.

A change volume is owned and used by the associated Remote Copy relationship. Therefore, it cannot be:

- ▶ Mapped to a host.
- ▶ Used as source or target of any FlashCopy maps.
- ▶ Part of any other relationship.
- ▶ A filesystem disk

Assigning a change volume to a relationship requires FlashCopy mappings to be created between the master or auxiliary volume and the associated change volume. Therefore, there must be sufficient unallocated FlashCopy memory in the target I/O group or the command fails.

10.9 Remote Copy commands

This section describes commands that must be issued to create and operate Remote Copy services.

10.9.1 Remote Copy process

The MM/GM process includes the following steps:

1. A system partnership is created between two IBM Storwize V5000 Gen2 systems or IBM SAN Volume Controller (for intercluster MM/GM).
2. A MM/GM relationship is created between two volumes of the same size.
3. To manage multiple MM/GM relationships as one entity, the relationships can be made part of a MM/GM Consistency Group to ensure data consistency across multiple MM/GM relationships, or for ease of management.

4. The MM/GM relationship is started and when the background copy completes, the relationship is consistent and synchronized.
5. When synchronized, the auxiliary volume holds a copy of the production data at the master that can be used for disaster recovery.
6. To access the auxiliary volume, the MM/GM relationship must be stopped with the access option enabled before write I/O is submitted to the auxiliary.

Following these commands, the remote host server is mapped to the auxiliary volume and the disk is available for I/O.

Note: For more information about MM/GM commands, see *IBM Spectrum Virtualize and SAN Volume Controller and Storwize family Command-Line Interface User's Guide*, [GC27-2287](#).

The command set for MM/GM contains the following broad groups:

- ▶ Commands to create, delete, and manipulate relationships and Consistency Groups
- ▶ Commands to cause state changes

If a configuration command affects more than one system, MM/GM performs the work to coordinate configuration activity between the systems. Certain configuration commands can be performed only when the systems are connected, and fail with no effect when they are disconnected.

Other configuration commands are permitted, even though the systems are disconnected. The state is reconciled automatically by MM/GM when the systems become connected again.

For any command (with one exception) a single system receives the command from the administrator. This design is significant for defining the context for a CreateRelationship **mkrcrelationship** or CreateConsistencyGroup **mkrcconsistgrp** command, in which case the system that is receiving the command is called the *local system*.

The exception is a command that sets systems into a MM/GM partnership. The **mkfcpartnership** and **mkippartnership** commands must be issued on the local and remote systems.

The commands in this section are described as an abstract command set, and are implemented by either of the following methods:

- ▶ CLI can be used for scripting and automation.
- ▶ GUI can be used for one-off tasks.

10.9.2 Listing available system partners

Use the **lspartnershipcandidate** command to list the systems that are available for setting up a two-system partnership. This command is a prerequisite for creating MM/GM relationships.

Note: This command is not supported on IP partnerships. Use **mkippartnership** for IP connections.

10.9.3 Changing the system parameters

When you want to change system parameters specific to any remote copy or Global Mirror only, use the **chsystem** command. The **chsystem** command features the following parameters for MM/GM:

▶ **-relationshipbandwidthlimit** *cluster_relationship_bandwidth_limit*

This parameter controls the maximum rate at which any one remote copy relationship can synchronize. The default value for the relationship bandwidth limit is 25 MBps, but this value can now be specified 1 - 100,000 MBps. The partnership overall limit is controlled by the **chpartnership -linkbandwidthhmbits** command, and must be set on each involved system.

Important: Do not set this value higher than the default without first establishing that the higher bandwidth can be sustained without affecting the host's performance. The limit must never be higher than the maximum that is supported by the infrastructure connecting the remote sites, regardless of the compression rates that you might achieve.

▶ **-gmlinktolerance** *link_tolerance*

This parameter specifies the maximum period that the system tolerates delay before stopping Global Mirror relationships. Specify values 60 - 86,400 seconds in increments of 10 seconds. The default value is 300. Do not change this value except under the direction of IBM Support.

▶ **-gmmaxhostdelay** *max_host_delay*

This parameter specifies the maximum time delay, in milliseconds, at which the Global Mirror link tolerance timer starts counting down. This threshold value determines the additional effect that Global Mirror operations can add to the response times of the Global Mirror source volumes. You can use this parameter to increase the threshold from the default value of 5 milliseconds.

▶ **-maxreplicationdelay** *max_replication_delay*

This parameter sets a maximum replication delay in seconds. The value must be a number 1 - 360. This feature sets the maximum number of seconds to be tolerated to complete a single I/O. If I/O cannot complete within the **max_replication_delay**, the 1920 event is reported. This is the system-wide setting. When set to 0, the feature is disabled. This applies to Metro Mirror and Global Mirror relationships.

Use the **chsystem** command to adjust these values, as shown in the following example:

```
chsystem -gmlinktolerance 300
```

You can view all of these parameter values by using the **lssystem <system_name>** command.

We focus on the **gmlinktolerance** parameter in particular. If poor response extends past the specified tolerance, a 1920 event is logged and one or more GM relationships automatically stop to protect the application hosts at the primary site. During normal operations, application hosts experience a minimal effect from the response times because the GM feature uses asynchronous replication.

However, if GM operations experience degraded response times from the secondary system for an extended period, I/O operations begin to queue at the primary system. This queue results in an extended response time to application hosts. In this situation, the **gm1inktolerance** feature stops GM relationships, and the application host's response time returns to normal.

After a 1920 event occurs, the GM auxiliary volumes are no longer in the `consistent_synchronized` state until you fix the cause of the event and restart your GM relationships. For this reason, ensure that you monitor the system to track when these 1920 events occur.

You can disable the **gm1inktolerance** feature by setting the **gm1inktolerance** value to zero. However, the **gm1inktolerance** feature cannot protect applications from extended response times if it is disabled. It might be appropriate to disable the **gm1inktolerance** feature under the following circumstances:

- ▶ During SAN maintenance windows in which degraded performance is expected from SAN components, and application hosts can withstand extended response times from GM volumes.
- ▶ During periods when application hosts can tolerate extended response times and it is expected that the **gm1inktolerance** feature might stop the GM relationships. For example, if you test by using an I/O generator that is configured to stress the back-end storage, the **gm1inktolerance** feature might detect the high latency and stop the GM relationships. Disabling the **gm1inktolerance** feature prevents this result at the risk of exposing the test host to extended response times.

A 1920 event indicates that one or more of the SAN components cannot provide the performance that is required by the application hosts. This situation can be temporary (for example, a result of a maintenance activity) or permanent (for example, a result of a hardware failure or an unexpected host I/O workload).

If 1920 events are occurring, it can be necessary to use a performance monitoring and analysis tool, such as the IBM Virtual Storage Center, to help identify and resolve the problem.

10.9.4 System partnership

To create an IBM SAN Volume Controller or an IBM Storwize system partnership, use the **mkfcpartnership** command for traditional Fibre Channel (FC or FCoE) connections or **mkippartnership** for IP-based connections.

The **svctask mkfcpartnership** command

Use the **mkfcpartnership** command to establish a one-way MM/GM partnership between the local system and a remote system. Alternatively, use **mkippartnership** to create IP-based partnership.

To establish a fully functional MM/GM partnership, you must issue this command on both systems. This step is a prerequisite for creating MM/GM relationships between volumes on the IBM Spectrum Virtualize systems.

When the partnership is created, you can specify the bandwidth to be used by the background copy process between the local and remote system. If it is not specified, the bandwidth default is 50 MBps. The bandwidth must be set to a value that is less than or equal to the bandwidth that can be sustained by the intercluster link.

Background copy bandwidth effect on foreground I/O latency

The background copy bandwidth determines the rate at which the background copy is attempted for MM/GM. The background copy bandwidth can affect foreground I/O latency in one of the following ways:

- ▶ The following result can occur if the background copy bandwidth is set too high compared to the MM/GM intercluster link capacity:
 - The background copy I/Os can back up on the MM/GM intercluster link.
 - There is a delay in the synchronous auxiliary writes of foreground I/Os.
 - The foreground I/O latency increases as perceived by applications.
- ▶ If the background copy bandwidth is set too high for the storage at the primary site, background copy read I/Os overload the primary storage and delay foreground I/Os.
- ▶ If the background copy bandwidth is set too high for the storage at the secondary site, background copy writes at the secondary site overload the auxiliary storage and again delay the synchronous secondary writes of foreground I/Os.

To set the background copy bandwidth optimally, ensure that you consider all three resources: primary storage, intercluster link bandwidth, and auxiliary storage. Provision the most restrictive of these three resources between the background copy bandwidth and the peak foreground I/O workload.

Perform this provisioning by calculation or by determining experimentally how much background copy can be allowed before the foreground I/O latency becomes unacceptable. Then, reduce the background copy to accommodate peaks in workload.

The `chpartnership` command

To change the bandwidth that is available for background copy in the system partnership, use the `chpartnership -backgroundcopyrate <percentage_of_link_bandwidth>` command to specify the percentage of whole link capacity to be used by background copy process.

10.9.5 Creating a Metro Mirror/Global Mirror consistency group

Use the `mkrconsistgrp` command to create an empty MM/GM Consistency Group.

The MM/GM consistency group name must be unique. If the consistency group involves two systems, the systems must be in communication throughout the creation process.

The new consistency group does not contain any relationships and is in the Empty state. You can add MM/GM relationships to the group (upon creation or afterward) by using the `chrcrelationship` command.

10.9.6 Creating a Metro Mirror/Global Mirror relationship

Use the `mkrcrelationship` command to create a MM/GM relationship. This relationship persists until it is deleted.

Optional parameter: If you do not use the `-global` optional parameter, a Metro Mirror relationship is created rather than a Global Mirror relationship.

The auxiliary volume must be equal in size to the master volume or the command fails. If both volumes are in the same system, they must be in the same I/O Group. The master and auxiliary volume cannot be in an existing relationship, and they cannot be the target of a FlashCopy mapping. This command returns the new relationship (`relationship_id`) when successful.

When the MM/GM relationship is created, you can add it to a Consistency Group, or it can be a stand-alone MM/GM relationship if no Consistency Group is specified.

The `lsrcrelationshipcandidate` command

Use the `lsrcrelationshipcandidate` command to list the volumes that are eligible to form an MM/GM relationship.

When the command is issued, you can specify the master volume name and auxiliary system to list the candidates that comply with the prerequisites to create a MM/GM relationship. If the command is issued with no parameters, all of the volumes that are not disallowed by another configuration state, such as being a FlashCopy target, are listed.

10.9.7 Changing Metro Mirror/Global Mirror relationship

Use the `chrcrelationship` command to modify the following properties of an MM/GM relationship:

- ▶ Change the name of an MM/GM relationship.
- ▶ Add a relationship to a group.
- ▶ Remove a relationship from a group by using the `-force` flag.

Adding an MM/GM relationship: When an MM/GM relationship is added to a Consistency Group that is not empty, the relationship must have the same state and copy direction as the group to be added to it.

10.9.8 Changing Metro Mirror/Global Mirror consistency group

Use the `chrconsistgrp` command to change the name of an MM/GM Consistency Group.

10.9.9 Starting Metro Mirror/Global Mirror relationship

Use the `startrcrelationship` command to start the copy process of an MM/GM relationship.

When the command is issued, you can set the copy direction if it is undefined, and, optionally, you can mark the auxiliary volume of the relationship as clean. The command fails if it is used as an attempt to start a relationship that is a part of a consistency group.

You can issue this command only to a relationship that is connected. For a relationship that is idling, this command assigns a copy direction (master and auxiliary roles) and begins the copy process. Otherwise, this command restarts a previous copy process that was stopped by a `stop` command or by an I/O error.

If the resumption of the copy process leads to a period when the relationship is inconsistent, you must specify the **-force** parameter when the relationship is restarted. This situation can arise if, for example, the relationship was stopped and then further writes were performed on the original master of the relationship.

The use of the **-force** parameter here is a reminder that the data on the auxiliary becomes inconsistent while resynchronization (background copying) occurs. Therefore, this data is unusable for Disaster Recovery purposes before the background copy completes.

In the `Idling` state, you must specify the master volume to indicate the copy direction. In other connected states, you can provide the **-primary** argument, but it must match the existing setting.

10.9.10 Stopping Metro Mirror/Global Mirror relationship

Use the **stoprcrelationship** command to stop the copy process for a relationship. You can also use this command to enable write access to a consistent auxiliary volume by specifying the **-access** parameter.

This command applies to a stand-alone relationship. It is rejected if it is addressed to a relationship that is part of a Consistency Group. You can issue this command to stop a relationship that is copying from master to auxiliary.

If the relationship is in an inconsistent state, any copy operation stops and does not resume until you issue a **startrcrelationship** command. Write activity is no longer copied from the master to the auxiliary volume. For a relationship in the `ConsistentSynchronized` state, this command causes a Consistency Freeze.

When a relationship is in a consistent state, you can use the **-access** parameter with the **stoprcrelationship** command to enable write access to the auxiliary volume.

10.9.11 Starting Metro Mirror/Global Mirror consistency group

Use the **startrcconsistgrp** command to start an MM/GM consistency group. You can issue this command only to a consistency group that is connected.

If the consistency group is in the `ConsistentStopped` state and not synchronized or is in the `Idling` state, run **startconsistgrp** with **-force** parameter to restart the copy.

For a consistency group that is idling, this command assigns a copy direction (master and auxiliary roles) and begins the copy process. Otherwise, this command restarts a previous copy process that was stopped by a **stop** command or by an I/O error.

10.9.12 Stopping Metro Mirror/Global Mirror consistency group

Use the **stoprcconsistgrp** command to stop the copy process for an MM/GM consistency group. You can also use this command to enable write access to the auxiliary volumes in the group if the group is in a consistent state.

When a consistency group is in a consistent state, to allow write access to the auxiliary volume use the **-access** parameter with the **stoprcconsistgrp**.

If the consistency group is in an inconsistent state, any copy operation stops and does not resume until you issue the **startrcconsistgrp** command. Write activity is no longer copied from the master to the auxiliary volumes that belong to the relationships in the group.

For a consistency group in the ConsistentSynchronized state, this command causes a Consistency Freeze.

10.9.13 Deleting Metro Mirror/Global Mirror relationship

Use the `rmrcrelationship` command to delete the relationship. Deleting a relationship delete only the logical relationship between the two volumes.

If the relationship is disconnected at the time that the command is issued, the relationship is deleted only on the system on which the command is being run. When the systems reconnect, the relationship is automatically deleted on the other system. Alternatively, if the systems are disconnected and you still want to remove the relationship on both systems, you can issue the `rmrcrelationship` command independently on both systems.

A relationship cannot be deleted if it is part of a consistency group. You must first remove the relationship from the consistency group. If you delete an inconsistent relationship, the auxiliary volume becomes accessible even though it is still inconsistent. This situation is the one case in which MM/GM does not inhibit access to inconsistent data.

10.9.14 Deleting Metro Mirror/Global Mirror consistency group

Use the `rmrcconsistgrp` command to delete an MM/GM consistency group. You can issue this command for any consistency group.

If the consistency group is disconnected at the time that the command is issued, the consistency group is deleted only on the system on which the command is being run. When the systems reconnect, the consistency group is automatically deleted on the other system.

Alternatively, if the systems are disconnected and you still want to remove the consistency group on both systems, you can issue the `rmrcconsistgrp` command separately on both of the systems.

If the consistency group is not empty, the relationships within it are removed from the consistency group before the group is deleted. These relationships then become stand-alone relationships. The state of these relationships is not changed by the action of removing them from the consistency group.

10.9.15 Reversing Metro Mirror/Global Mirror relationship

Use the `switchrcrelationship` command to reverse the roles of the master volume and the auxiliary volume when a stand-alone relationship is in a consistent state. When the command is issued, the wanted master must be specified.

10.9.16 Reversing Metro Mirror/Global Mirror consistency group

Use the `switchrcconsistgrp` command to reverse the roles of the master volume and the auxiliary volume when a consistency group is in a consistent state. This change is applied to all of the relationships in the consistency group. When the command is issued, the wanted master must be specified.

Important: Remember that by reversing the roles, your current source volumes become targets, and target volumes become source volumes. Therefore, you lose write access to your current primary volumes.

10.10 Managing Remote Copy using the GUI

In this section, we describe the tasks to perform Remote Copy management by using the GUI.

The following panels are used to visualize and manage your remote copies:

- ▶ The Remote Copy panel, as shown in Figure 10-79.

To access the Remote Copy panel, select Copy Services at main panel and click **Remote Copy**.

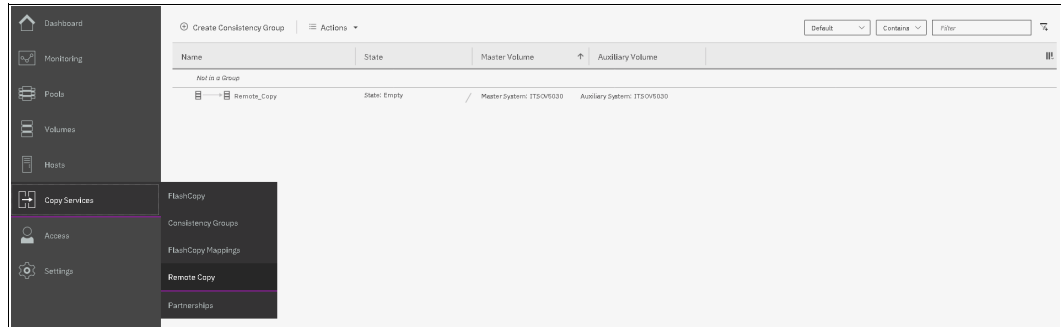


Figure 10-79 Remote Copy panel

- ▶ The Partnerships panel, as shown in Figure 10-80.

To access the Partnerships panel, select Copy Services at main panel and click **Partnerships**.

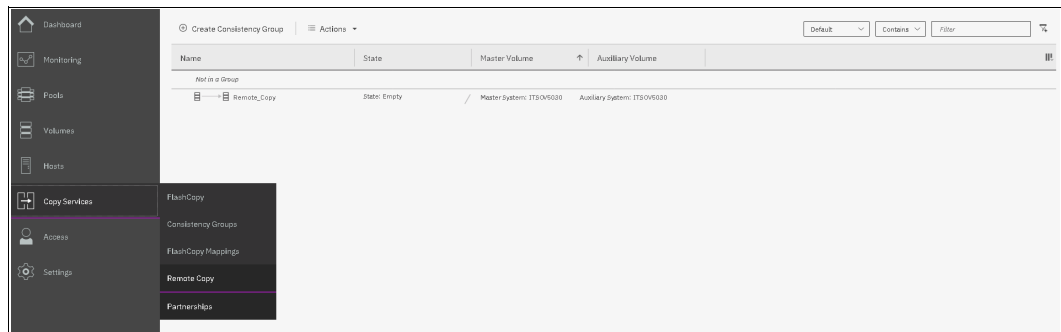


Figure 10-80 Partnerships panel

10.10.1 Creating Fibre Channel partnership

To create an FC partnership between the systems running IBM Spectrum Virtualize, use the GUI and complete the following steps:

1. From the main panel, click **Copy Services** → **Partnerships**.

The Partnership panel opens, as shown in Figure 10-81.



Figure 10-81 Partnership panel

- Click **Create Partnership** to create a partnership with another IBM SAN Volume Controller or IBM Storwize system, as shown in Figure 10-82.



Figure 10-82 Create a partnership

- In the **Create Partnership** window, select Fibre Channel or IP, as shown in Figure 10-83.

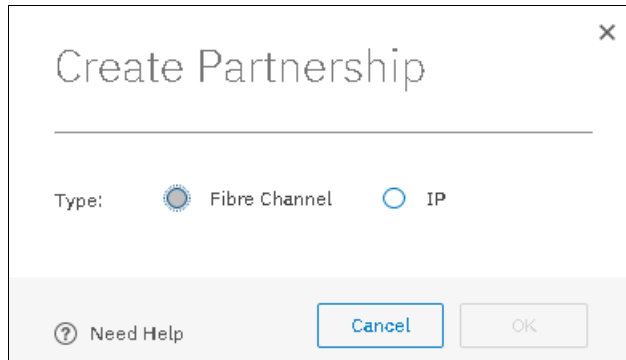


Figure 10-83 Partnership type

- For Fibre Channel partnership, select an available partner system from the drop-down list. If no candidate is available, the following error message is displayed:

This system does not have any candidates.

 - Enter a link bandwidth in megabits per second (Mbps) that is used by the background copy process between the systems in the partnership.
 - Enter the background copy rate.
 - Click **OK** to confirm the partnership relationship creation.

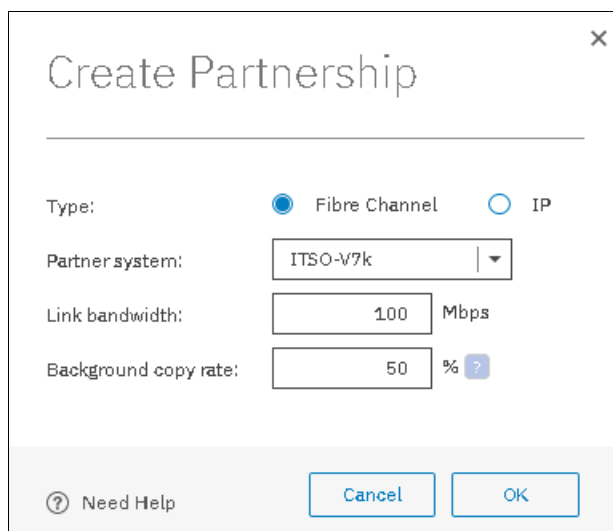


Figure 10-84 Create Fibre Channel partnership

Note: If you choose IP partnership, you must provide the IP address of the partner system and the partner system's CHAP key.

- The partnership is in the Partially Configured state because this work was performed only on one side, as shown in Figure 10-85. To fully configure the partnership between both systems, perform the same steps on the other system.

Name	Location	State	Type	IP Address
ITS0V630	Local	-	-	-
ITS0V76	Remote	Partially Configured	Fibre Channel	-

Figure 10-85 Partnership partially configured

- At the partner system perform the same steps, and as soon as the partnership is configured, it appears as shown in Figure 10-86.

Name	Location	State	Type	IP Address
ITS0V630	Local	-	-	-
ITS0V76	Remote	Fully Configured	Fibre Channel	-

Figure 10-86 Fully configured

10.10.2 Creating stand-alone remote copy relationships

In this section, we create remote copy mappings for volumes with their respective remote targets. The source and target volumes were created before this operation was done on both systems. The target volume must have the same size as the source volume.

Complete the following steps to create stand-alone copy relationships:

- From the main navigation panel, select **Copy Services** → **Remote Copy**.
- Select **Not in a Group**, click **Action** and **Create Relationship**, as shown in Figure 10-87.

Name	State	Master Volume	Auxiliary Volume
Not in a Group			
Remote_Copy	State: Empty	Master System: ITS0V630	Auxiliary System: ITS0V630

Figure 10-87 Creating a new Remote Copy relationship without consistency group

3. In the Create Relationship window, select one of the types of relationships that is to be created (as shown in Figure 10-88 on page 570).

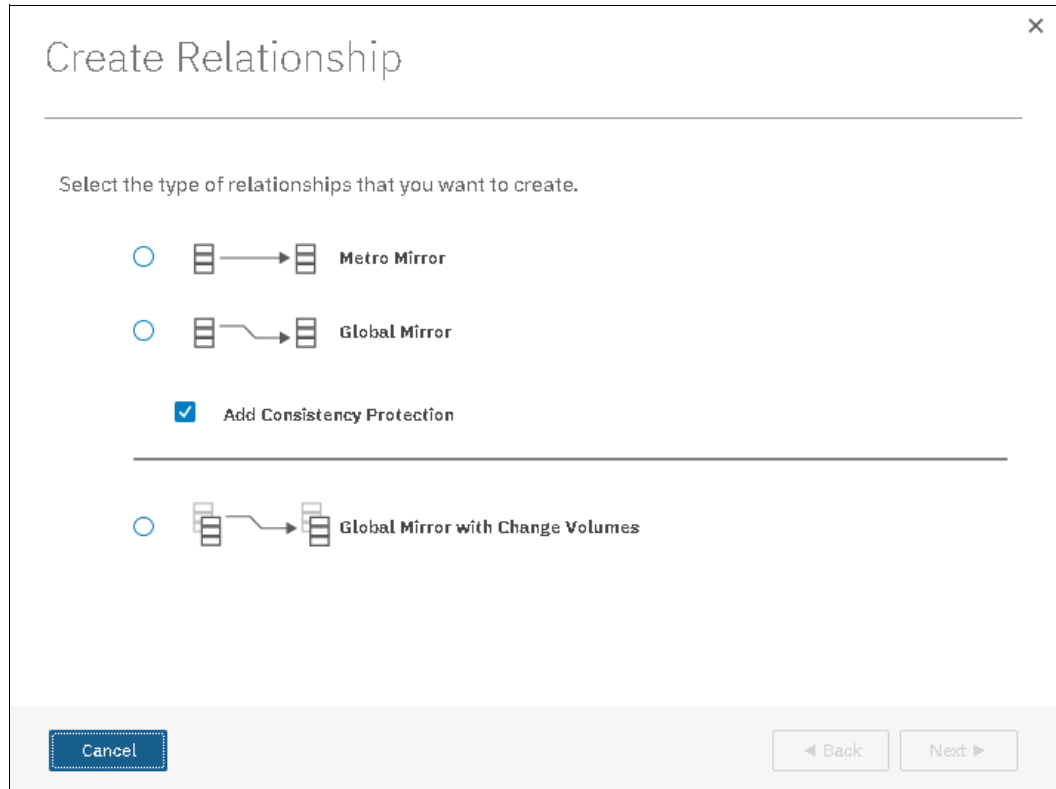
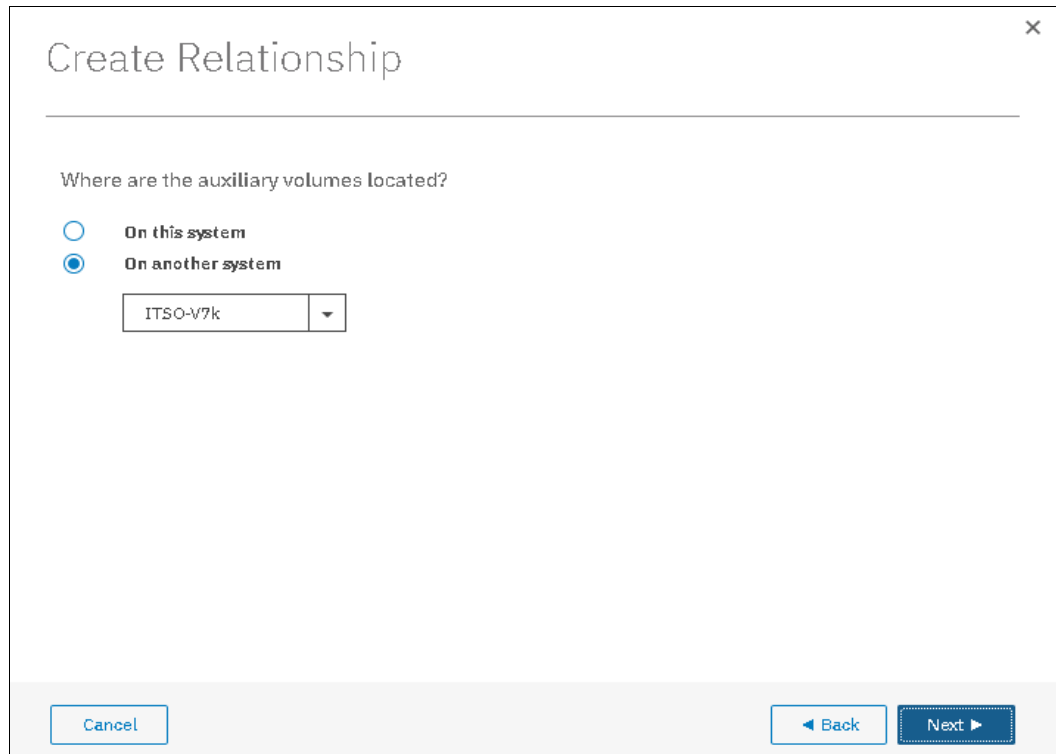


Figure 10-88 Select type of relationship

In this example, the Metro Mirror relationship is selected. Click **Next**.

Note: Starting from V7.8.1, consistency protection by using Change Volume was enabled by default. For more information about consistency protection, see 10.8, “Consistency protection for Remote and Global mirror” on page 558.

4. In the next window, select the location of the auxiliary volumes, as shown in Figure 10-89:



Create Relationship

Where are the auxiliary volumes located?

On this system

On another system

ITSO-V7k

Cancel Back Next

Figure 10-89 Create relationship

- **On this system**, which means that the volumes are local.
- **On another system**, which means that you select the remote system from the drop-down list.

After you make a selection, click **Next**.

5. In the New Relationship window that is shown in Figure 10-90, you can create relationships. Select a master volume in the Master drop-down list. Then, select an auxiliary volume in the Auxiliary drop-down list for this master and click **Add**. If needed, repeat this step to create other relationships.

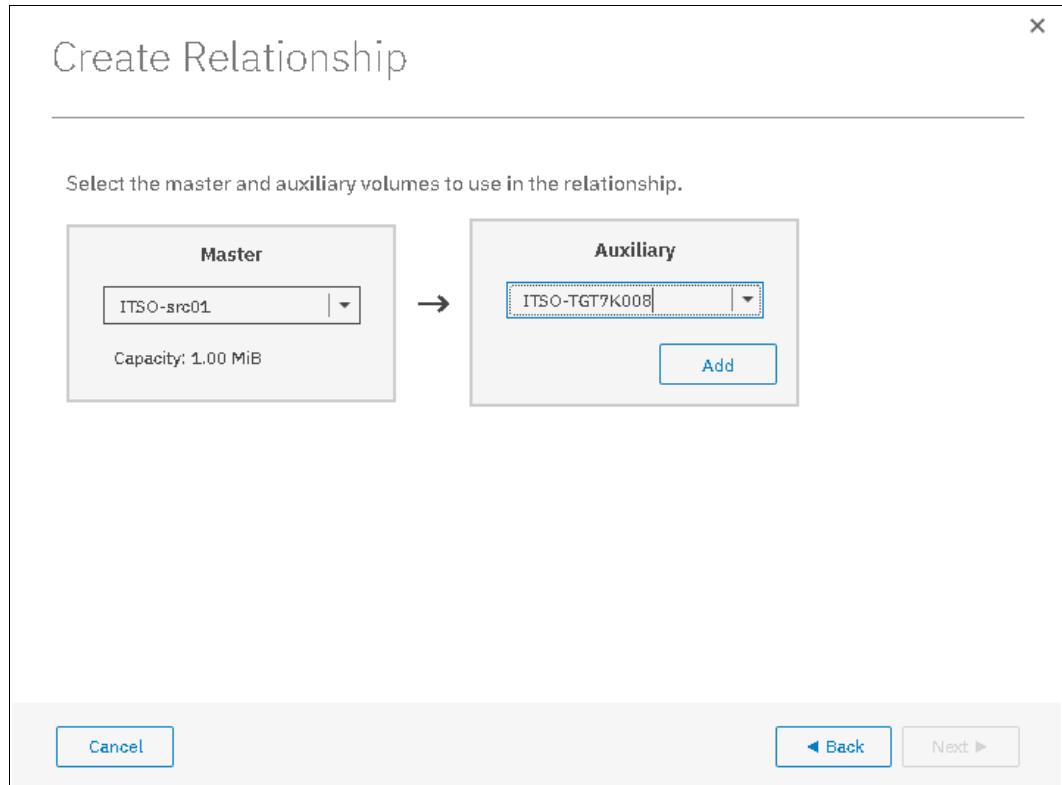


Figure 10-90 Select a volume for mirroring

Important: The master and auxiliary volumes must be of equal size. Therefore, only the targets with the appropriate size are shown in the list for a specific source volume.

6. Because **Add Consistency Protection** was selected, a window opens in which you are prompted if you want to add a change volume, as shown in Figure 10-91.

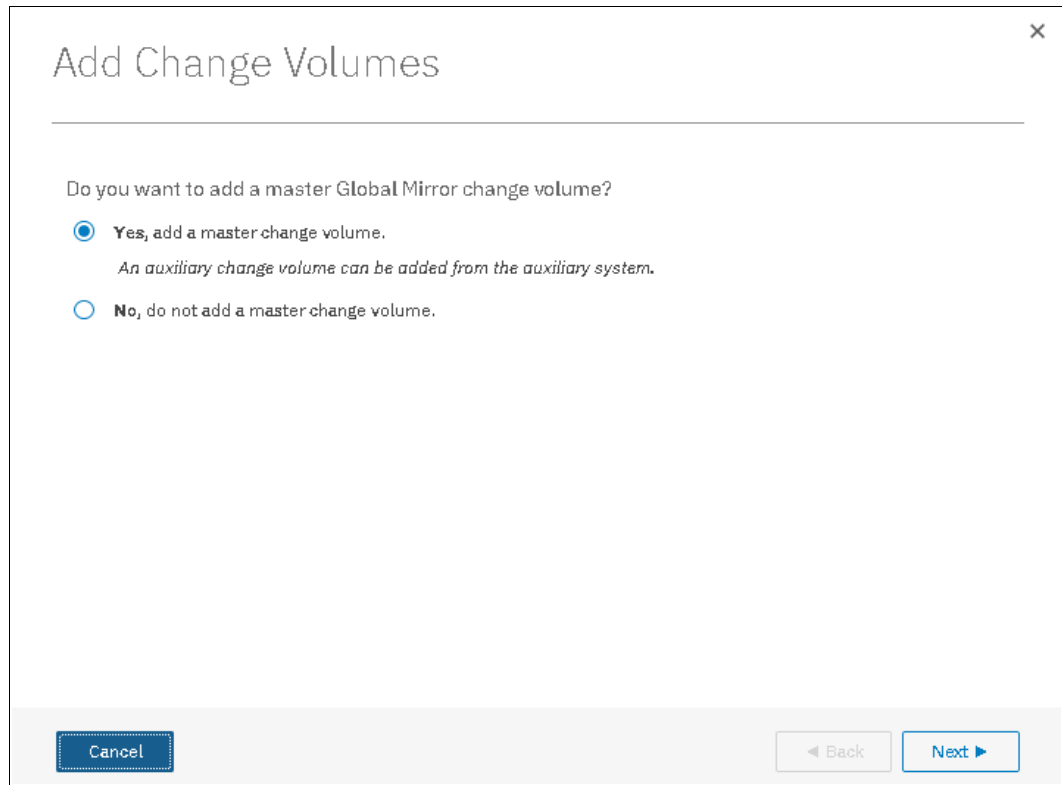


Figure 10-91 Add Change Volume

7. Click **Next**. A window opens in which you are prompted whether you want to add a new change volume or use an existing change volume. In our example, we chose to create a master change volume, as shown in Figure 10-92.

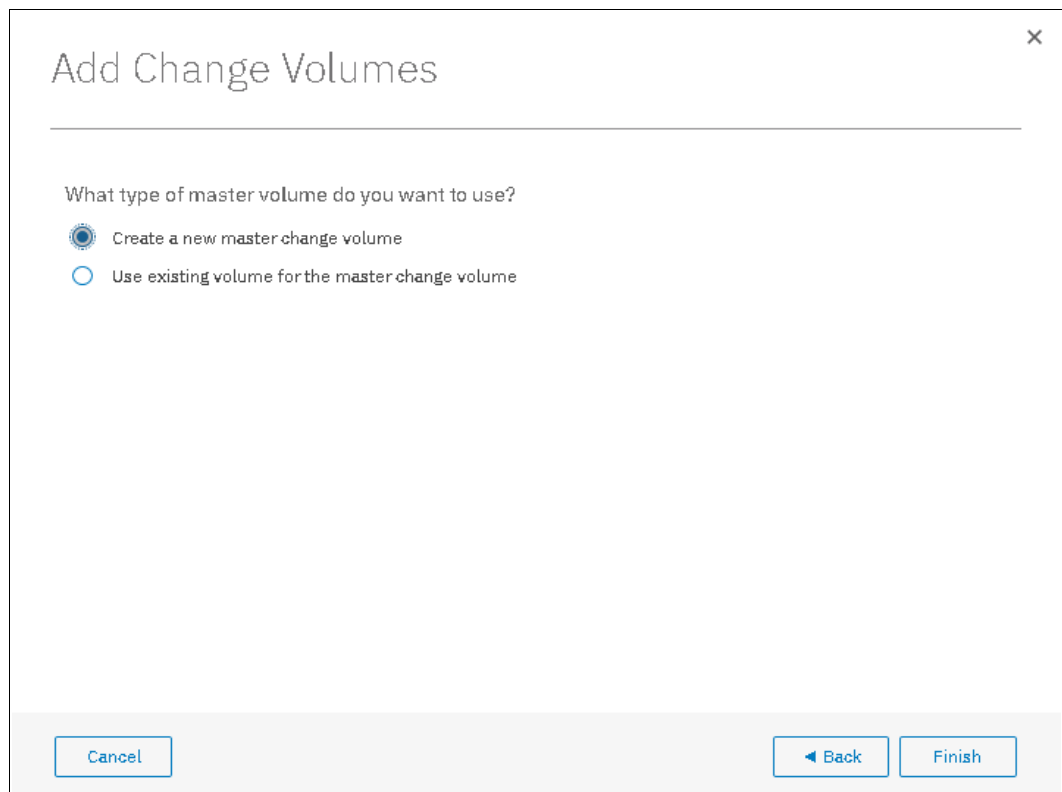


Figure 10-92 Create a master change volume or use an existing volume

8. Click **Finish**. A window opens in which you can remove a relationship that was created by clicking **X**, as shown in Figure 10-93.

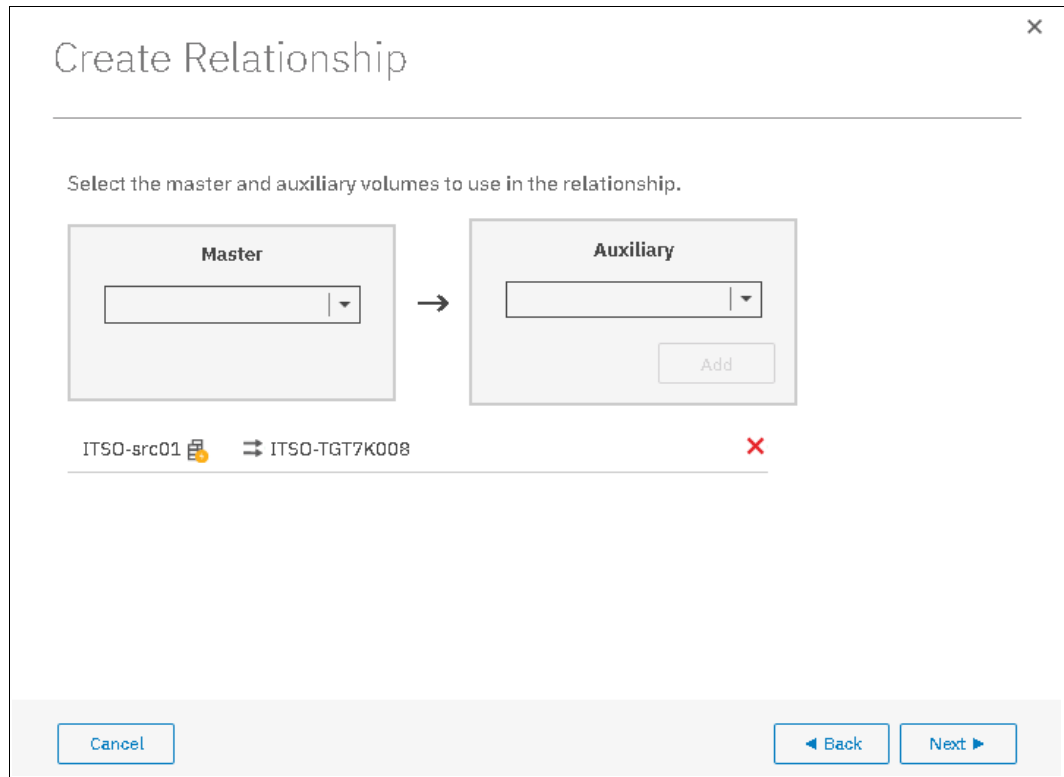
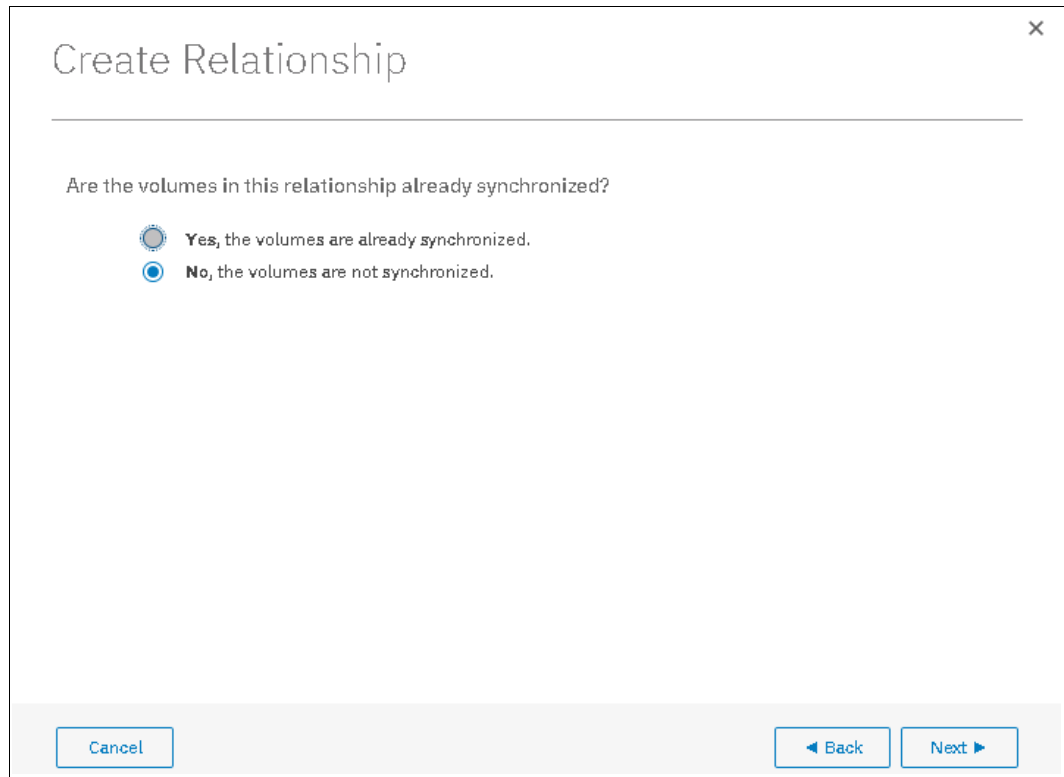


Figure 10-93 Create the relationships between the master and auxiliary volumes

9. After all of the relationships that you want to create are shown, click **Next**.

10. Specify whether the volumes are synchronized, as shown in Figure 10-94. Then, click **Next**.



Create Relationship

Are the volumes in this relationship already synchronized?

Yes, the volumes are already synchronized.

No, the volumes are not synchronized.

Cancel ◀ Back Next ▶

Figure 10-94 Volumes are already synchronized

11. In the next window, select whether you want to start to copy the data and click **Finish**, as shown in Figure 10-95.

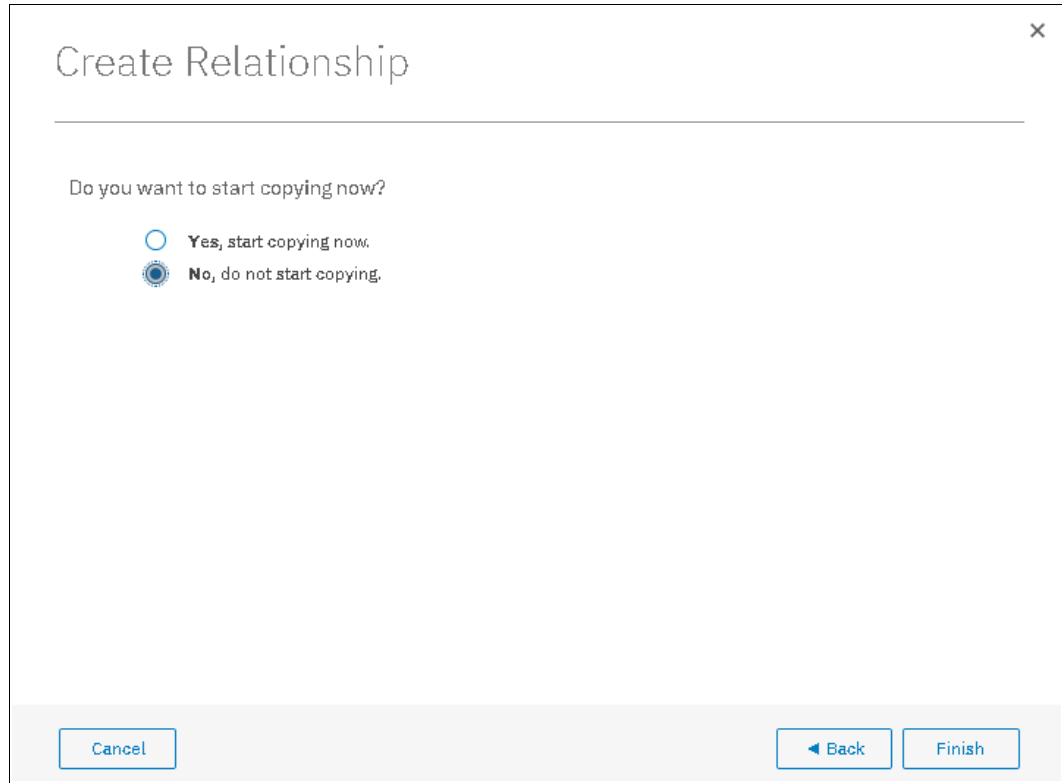


Figure 10-95 Select start copy option

The relationships are visible in the Remote Copy panel, as shown in Figure 10-96. After the copy is finished, the relationship status changes to Consistent synchronized.

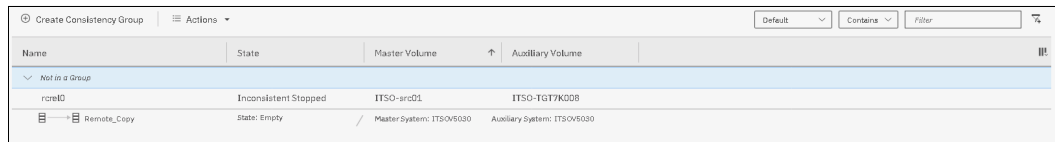


Figure 10-96 Remote Copy relationship created

10.10.3 Creating a Consistency Group

To create a Consistency Group, complete the following steps:

1. From the main navigation panel, select **Copy Services** → **Remote Copy**.
2. Click **Create Consistency Group**, as shown in Figure 10-97.



Figure 10-97 Create Consistency panel

3. Enter a name for the Consistency Group, and then, click **Next**, as shown in Figure 10-98.

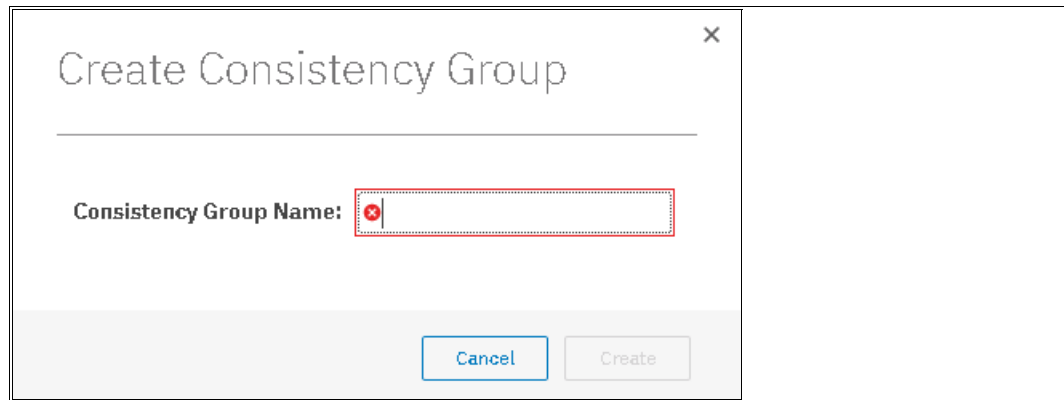


Figure 10-98 Enter a Consistency Group name

The new Consistency Group appears on that Consistency Group panel.

10.10.4 Renaming Consistency Group

To rename a Consistency Group, complete the following steps:

1. From the main navigation panel, select **Copy Services** → **Consistency Group**.
2. In the panel, select the Consistency Group that you want to rename. Then, select **Actions** → **Rename**, as shown in Figure 10-99.

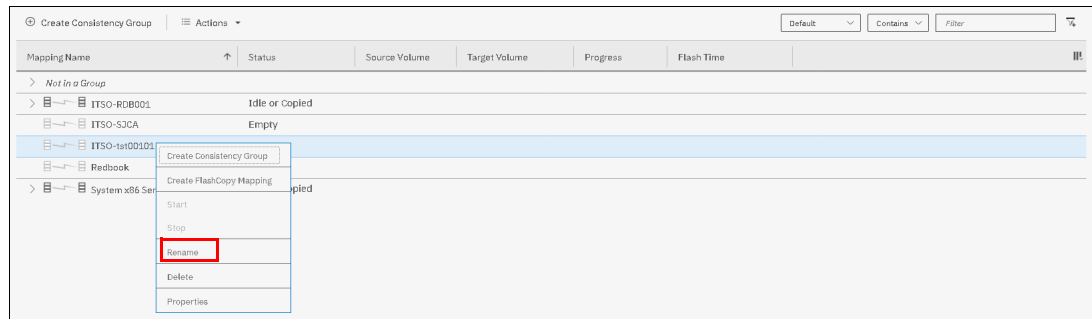


Figure 10-99 Rename Consistency Group

Note: You can also right-click a remote copy consistency group and select **Rename**.

3. Enter the new name and click **Rename**, as shown in Figure 10-100.



Figure 10-100 Changing the name for a Consistency Group

The new Consistency Group name is displayed on the Remote Copy panel.

10.10.5 Renaming remote copy relationship

Complete the following steps to rename a remote copy relationship:

1. From the main navigation panel, select **Copy Services** → **Remote Copy**.
2. In the table, select the remote copy relationship mapping to rename and right-click **Rename**, as shown in Figure 10-101.

Tip: You can also right-click a remote copy relationship and select **Rename**.

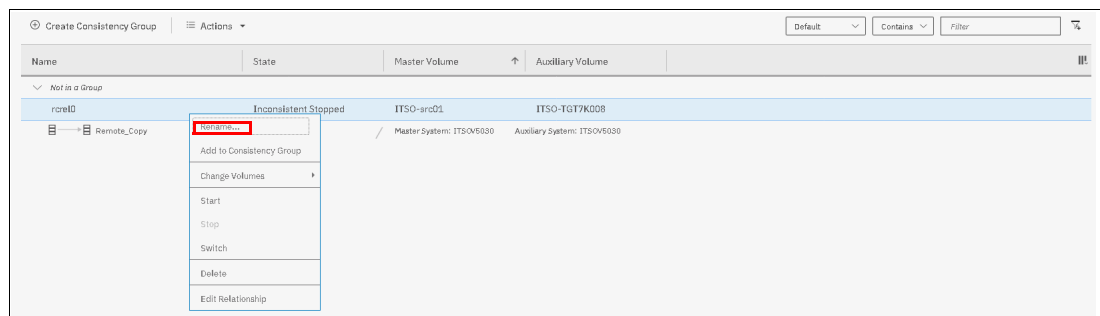


Figure 10-101 Rename remote copy relationship action

3. In the Rename Relationship window, enter the new name and click **Rename**, as shown in Figure 10-102.

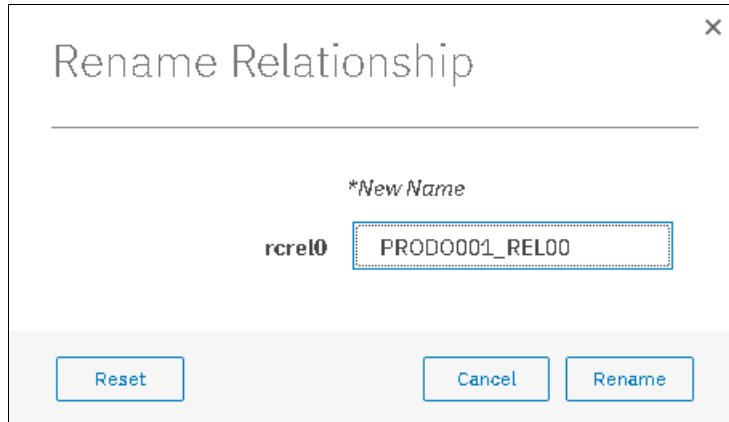


Figure 10-102 Renaming a remote copy relationship

Remote copy relationship name: You can use the letters A - Z and a - z, the numbers 0 - 9, and the underscore (_) character. The remote copy name can be 1 - 15 characters. No blanks are allowed.

10.10.6 Moving stand-alone remote copy relationship to Consistency Group

Complete the following steps to move a remote copy relationship to a Consistency Group:

1. From the main navigation panel, click **Copy Services** → **Remote Copy**.
2. Expand the **Not in a Group** column.
3. Select the relationship to move to the Consistency Group, right-click **Add to Consistency Group**, as shown in Figure 10-103.

Tip: You can also right-click a remote copy relationship and select **Add to Consistency Group**.

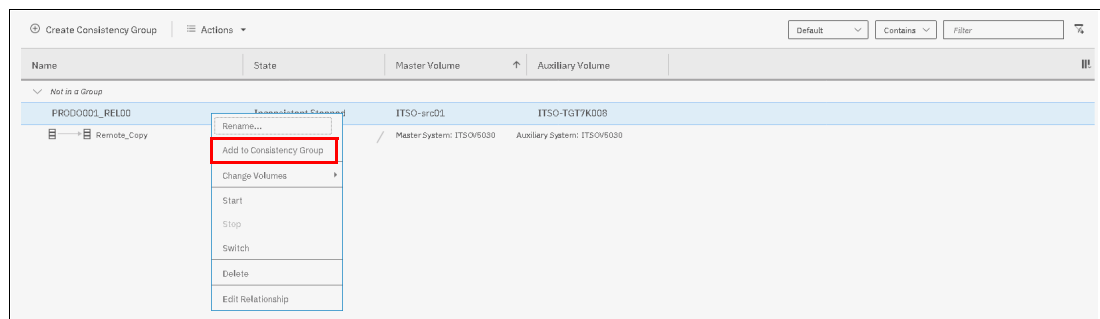


Figure 10-103 Add to Consistency Group

- In the Add Relationship to Consistency Group window, select the Consistency Group by using the drop-down list, as shown in Figure 10-104. Click **Add to Consistency Group** to confirm your changes.

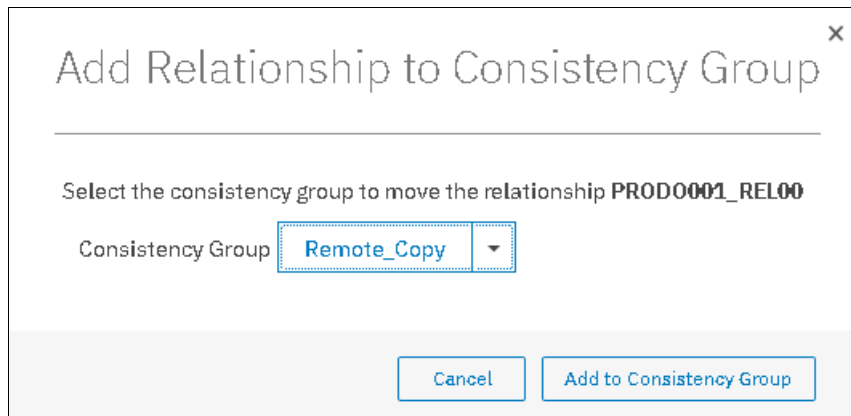


Figure 10-104 Adding a relationship to a Consistency Group

Note: The state of the remote copy consistency group and the recopy copy relationship that is being added must match; otherwise, you cannot add that remote copy relationship into the existing remote copy consistency group.

10.10.7 Removing remote copy relationship from Consistency Group

Complete the following steps to remove a remote copy relationship from a Consistency Group:

- From the main navigation panel, select **Copy Services** → **Remote Copy**.
- Select a Consistency Group.
- Select the remote copy relationship to remove from the Consistency Group and right-click **Remove from Consistency Group**, as shown in Figure 10-105.

Tip: You can also right-click a remote copy relationship and select **Remove from Consistency Group**.

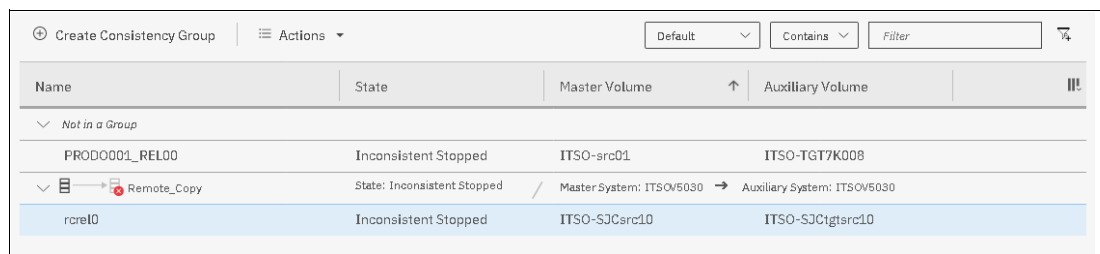


Figure 10-105 Remove from Consistency Group action

- In the Remove Relationship From Consistency Group window, click **Remove**, as shown in Figure 10-106.

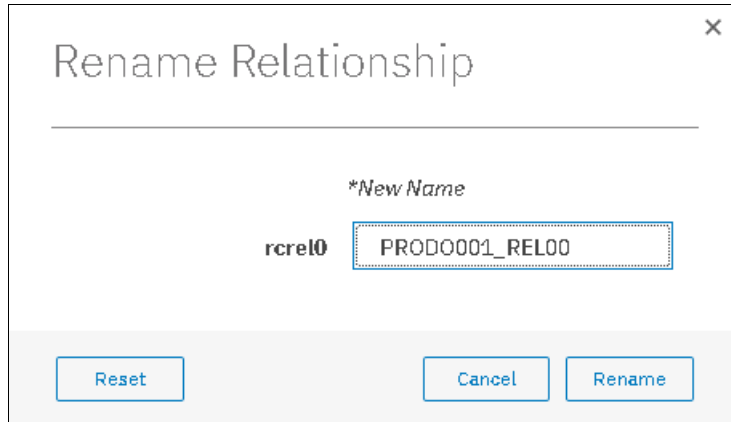


Figure 10-106 Remove a relationship from a Consistency Group

10.10.8 Starting remote copy relationship

When a remote copy relationship is created, the remote copy process can be started. Only relationships that are not members of a Consistency Group, or the only relationship in a Consistency Group, can be started individually.

Complete the following steps to start a remote copy relationship:

- From the main navigation panel, select **Copy Services** → **Remote Copy**.
- Expand the **Not in a Group** column.
- In the table, select the remote copy relationship that you want to start.
- Click **Actions** → **Start** to start the remote copy process, as shown in Figure 10-107.

Tip: You can also right-click a relationship and select **Start** from the list.

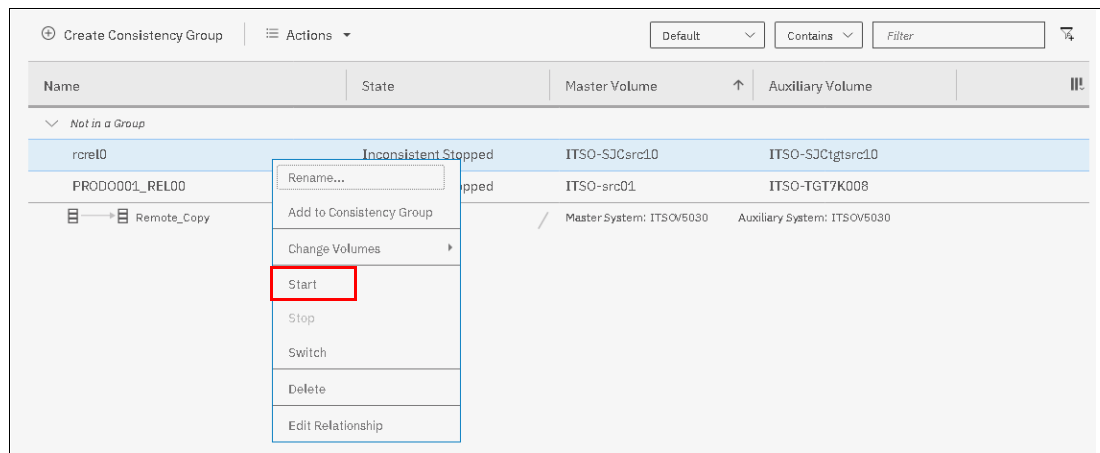


Figure 10-107 Starting the remote copy process

- After the task is complete, the remote copy relationship status has a Consistent Synchronized state, as shown in Figure 10-108.

Name	State	Master Volume	Auxiliary Volume
Not in a Group			
rcrel0	Consistent Synchronized	ITSO-SJCsrc10	ITSO-SJC1gtsrc10
PRODO001_REL00	Inconsistent Stopped	ITSO-src01	ITSO-TGT7K008
Remote_Copy		State: Empty / Master System: ITSOV5030 Auxiliary System: ITSOV5030	

Figure 10-108 Consistent Synchronized remote copy relationship

10.10.9 Starting remote copy Consistency Group

All of the mappings in a Consistency Group are brought to the same state. To start the remote copy Consistency Group, complete the following steps:

- From the main navigation panel, select **Copy Services** → **Remote Copy**.
- Select the Consistency Group that you want to start, as shown in Figure 10-109.

Name	State	Master Volume	Auxiliary Volume
Not in a Group			
PRODO001_REL00	Inconsistent Stopped	ITSO-src01	ITSO-TGT7K008
Remote_Copy		State: Consistent Synchronized / Master System: ITSOV5030 → Auxiliary System: ITSOV5030	
rcrel0	Consistent Synchronized	ITSO-SJCsrc10	ITSO-SJC1gtsrc10

Figure 10-109 Remote Copy Consistency Groups view

- Select the Consistency group and **Right-Click** → **Start** (see Figure 10-110) to start the remote copy Consistency Group.

Name	State	Master Volume	Auxiliary Volume
Not in a Group			
PRODO001_REL00	Inconsistent Stopped	ITSO-src01	ITSO-TGT7K008
Remote_Copy		State: Consistent Synchronized / Master System: ITSOV5030 → Auxiliary System: ITSOV5030	
rcrel0	Inconsistent Stopped	ITSO-SJCsrc10	ITSO-SJC1gtsrc10

- Create Relationship
- Rename
- Start**
- Stop
- Switch
- Edit Consistency Group
- Delete

Figure 10-110 Start Consistency Group

After the task completes, the Consistency Group and all of its relationships becomes in a Consistent Synchronized state.

10.10.10 Switching copy direction

When a remote copy relationship is in the Consistent synchronized state, the copy direction for the relationship can be changed. Only relationships that are not a member of a Consistency Group (or the only relationship in a Consistency Group) can be switched individually. These relationships can be switched from master to auxiliary or from auxiliary to master, depending on the case.

Complete the following steps to switch a remote copy relationship:

1. From the System panel, select **Copy Services** → **Remote Copy**.
2. Expand the **Not in a Group** column.
3. In the table, select the remote copy relationship that you want to switch and right-click **Switch** to start the remote copy process, as shown in Figure 10-111.

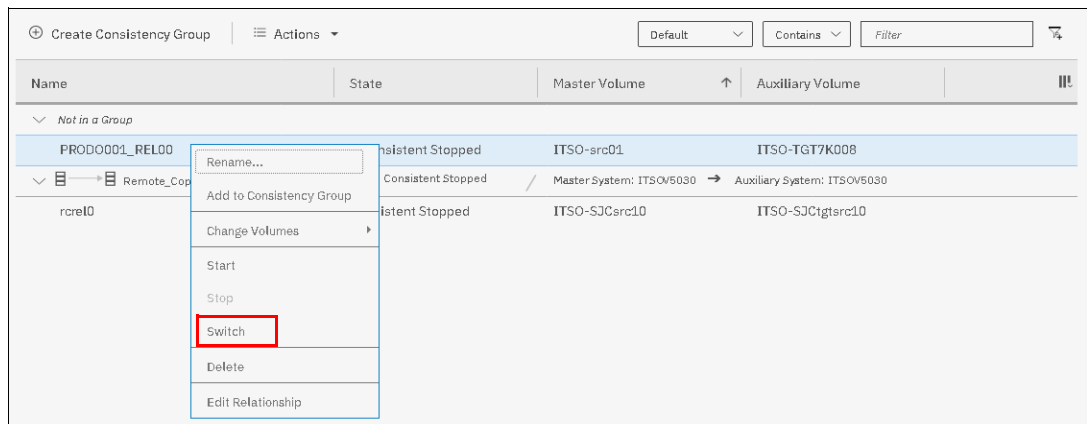


Figure 10-111 Switch remote copy

4. The Warning window that is shown in Figure 10-112 opens. The remote copy is switched from the master volume to the auxiliary volume. Click **Yes**.

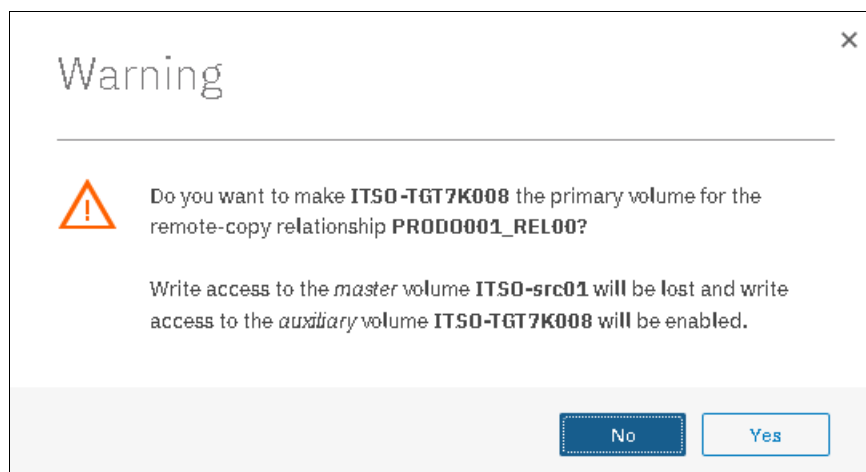


Figure 10-112 Warning window

The copy direction is now switched, as shown in Figure 10-113 with a red circle. The auxiliary volume is now accessible and shown as the primary volume. Also, the auxiliary volume is now synchronized to the master volume.

Name	State	Master Volume	Auxiliary Volume
PRODO001_REL00	Consistent Synchronized	ITSO-src01	ITSO-TGT7K008

Figure 10-113 Checking remote copy synchronization direction

10.10.11 Switching the copy direction for a Consistency Group

When a Consistency Group is in the Consistent Synchronized state, the copy direction for this Consistency Group can be changed.

Important: When the copy direction is switched, it is crucial that no outstanding I/O exists to the volume that changes from primary to secondary because all of the I/O is inhibited to that volume when it becomes the secondary. Therefore, careful planning is required before you switch the copy direction for a Consistency Group.

Complete the following steps to switch a Consistency Group:

1. From the main navigation panel, select **Copy Services** → **Remote Copy**.
2. Select the Consistency Group to switch and right-click **Switch** (as shown in Figure 10-114) to start the remote copy process.

Name	State	Master Volume	Auxiliary Volume
PRODO001_REL00	Consistent Synchronized	ITSO-src01	ITSO-TGT7K008
rcrel0	Consistent Synchronized	ITSO-SJCsrc10	ITSO-SJCTgtsrc10

Figure 10-114 Switching Consistency Group

- The warning window that is shown in Figure 10-115 opens. In the example, the Consistency Group is switched from the master group to the auxiliary group. Click **Yes**.

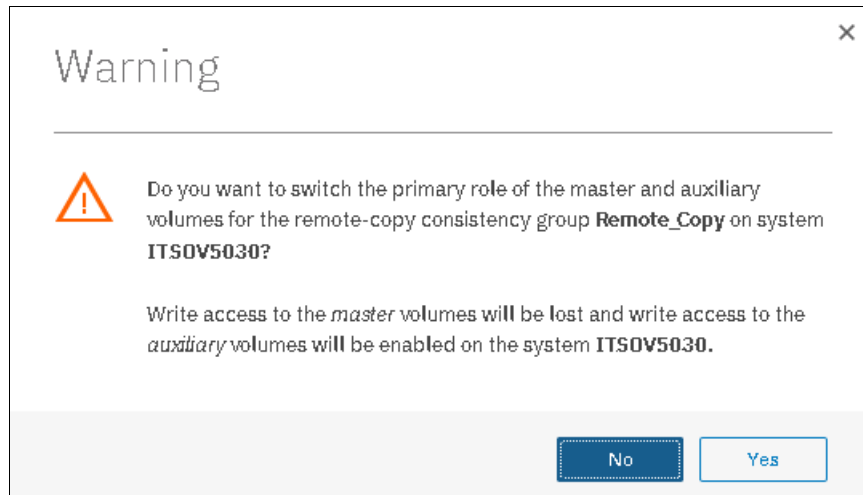


Figure 10-115 Warning window before switching the relationship

The remote copy direction is now switched, as shown in Figure 10-116. The auxiliary volume is now accessible and shown as a primary volume.

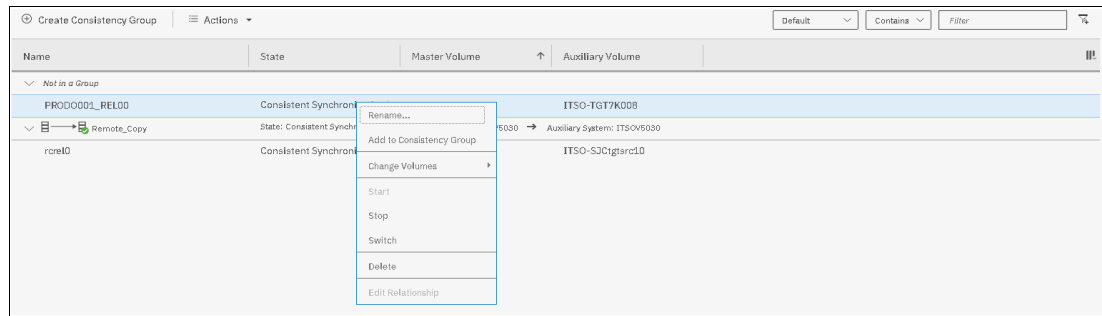


Figure 10-116 Relationship switched

10.10.12 Stopping a remote copy relationship

Complete the following steps to stop a remote copy relationship:

- From the main navigation panel, select **Copy Services** → **Remote Copy**.
- Expand the **Not in a Group** column.
- In the table, select the remote copy relationship to stop and right-click **Stop** (as shown in Figure 10-117 on page 587) to stop the remote copy process.

Tip: You can also right-click a relationship and select **Stop** from the list.

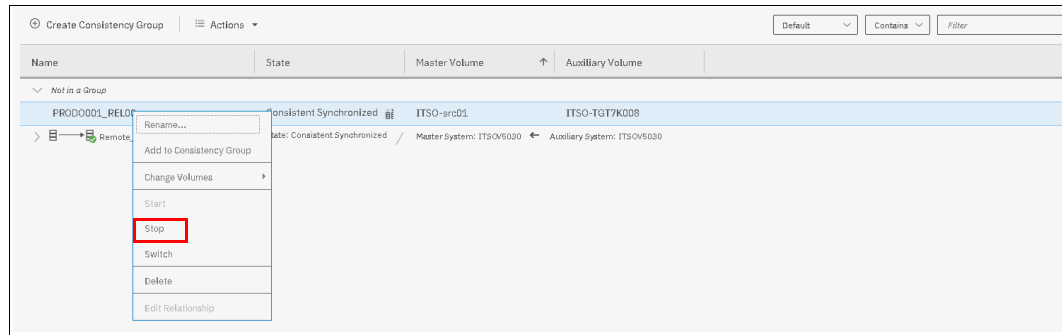


Figure 10-117 Stop remote copy relationship

- The Stop Remote Copy Relationship window opens, as shown in Figure 10-118. To allow secondary read/write access, select **Allow secondary read/write access**. Then, click **Stop Relationship**.

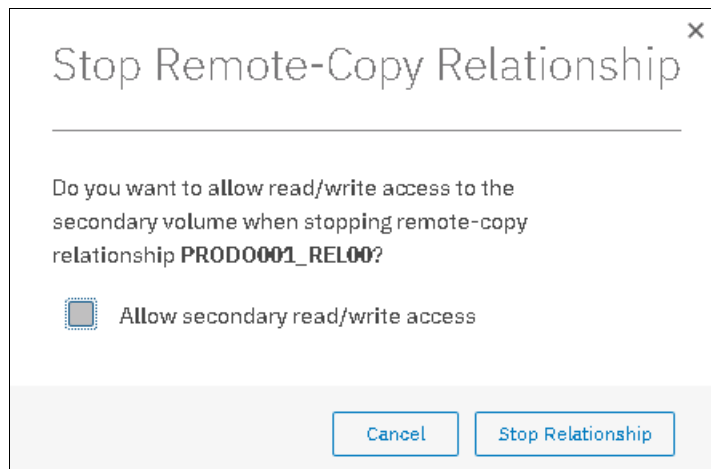


Figure 10-118 Stop Remote Copy Relationship

- The status can be checked through status column, as shown in Figure 10-119.



Figure 10-119 Relationship status

10.10.13 Stopping Consistency Group

To stop the Consistency Group, follow the steps below.

Perform the following steps to stop a Consistency Group:

- From the main navigation panel, select **Copy Services** → **Remote Copy**.
- In the table, select the Consistency Group to stop and right-click **Stop** (as shown in Figure 10-120 on page 588) to stop the remote copy Consistency Group.

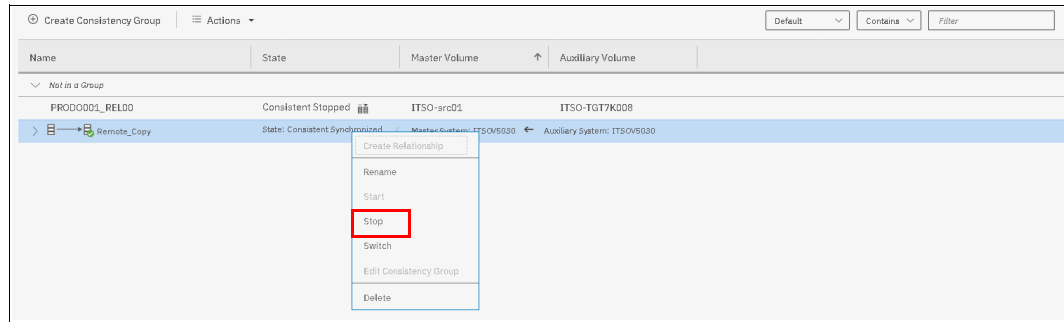


Figure 10-120 Selecting the Stop option

3. The Stop Remote Copy Consistency Group window opens, as shown in Figure 10-121. To allow secondary read/write access, select **Allow secondary read/write access**. Then, click **Stop Consistency Group**.

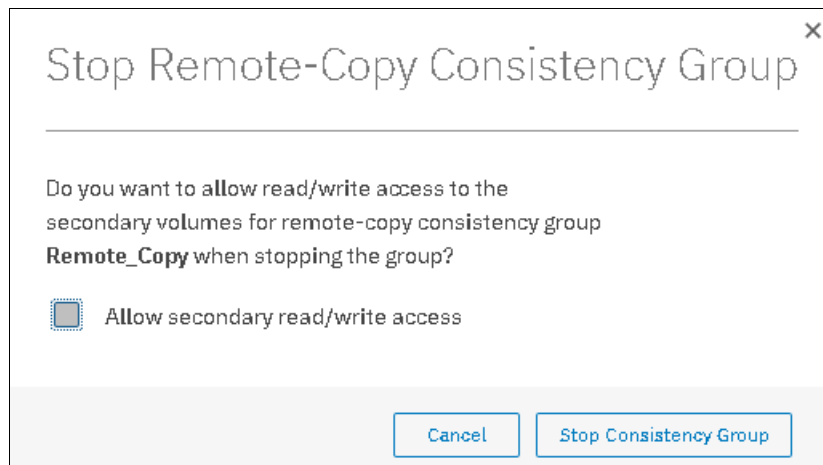


Figure 10-121 Stop Remote Copy Consistency Group window

The new relationship status can be checked, as shown in Figure 10-122. The relationship is now Consistent Stopped.

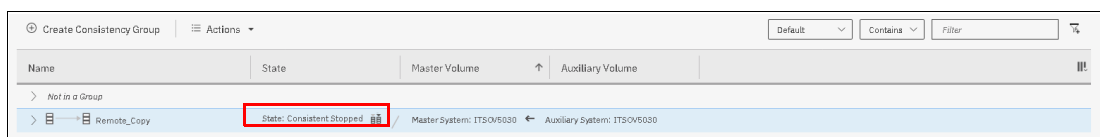


Figure 10-122 Checking remote copy synchronization status

10.10.14 Deleting stand-alone remote copy relationships

Complete the following steps to delete a stand-alone remote copy mapping:

1. From the main navigation panel, select **Copy Services** → **Remote Copy**.
2. In the table, select the remote copy relationship that you want to delete and right-click **Delete**, as shown in Figure 10-123 on page 589.

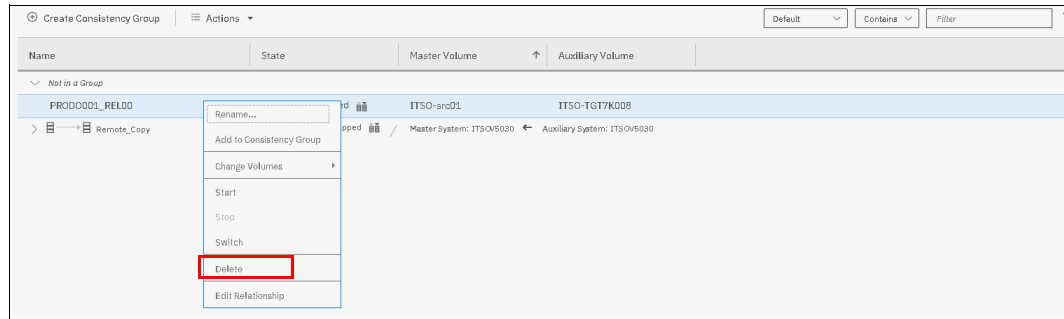


Figure 10-123 Selecting the Delete Relationship option

3. The Delete Relationship window opens (see Figure 10-124). In the “Verify the number of relationships that you are deleting” field, enter the number of volumes that you want to remove. This verification was added to help to avoid deleting the wrong relationships. Click **Delete**.

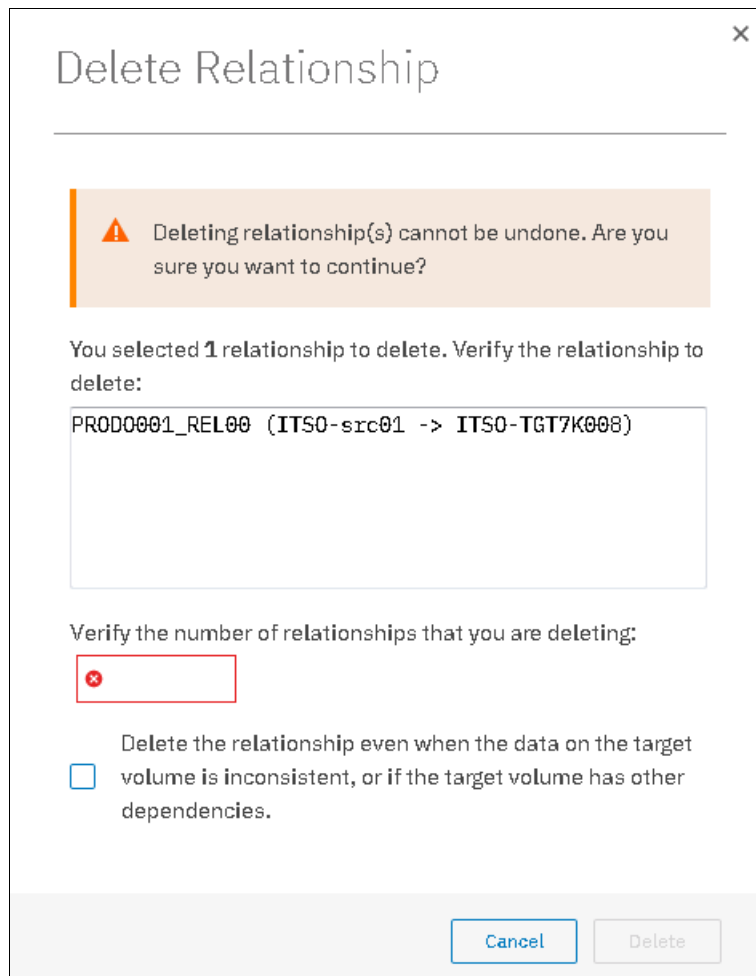


Figure 10-124 Delete remote copy relationship

10.10.15 Deleting Consistency Group

Important: Deleting a Consistency Group does not delete its remote copy mappings.

Complete the following steps to delete a Consistency Group:

1. From the main navigation panel, select **Copy Services** → **Remote Copy**.
2. In the left column, select the Consistency Group to delete and right-click **Delete**, as shown in Figure 10-125.

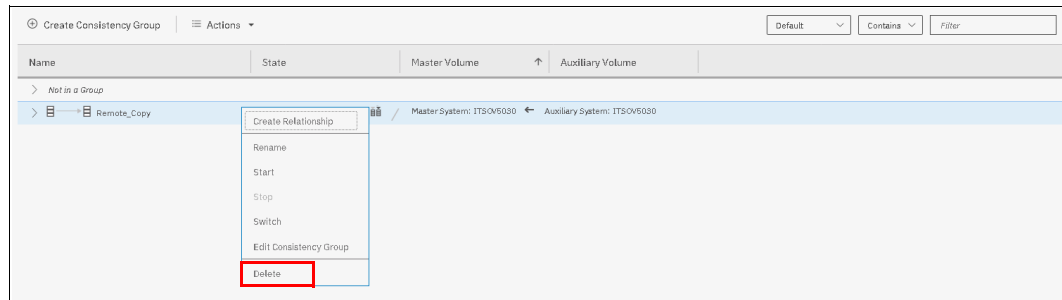


Figure 10-125 Selecting the Delete Consistency Group option

3. The warning window that is shown in Figure 10-126 opens. Click **Yes**.

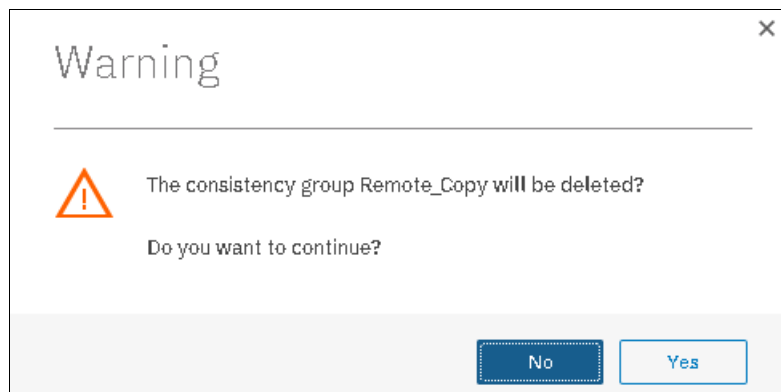


Figure 10-126 Warning Remote Copy deletion

10.11 Troubleshooting remote copy

Remote copy (Metro Mirror and Global Mirror) has two primary error codes that are displayed:

- ▶ A 1920 is a congestion error. This error means that the source, the link between the source and target, or the target cannot keep up with the requested copy rate.
- ▶ A 1720 error is a heartbeat or system partnership communication error. This error often is more serious because failing communication between your system partners involves extended diagnostic time.

10.11.1 1920 error

This error is deliberately generated by the system and is considered as a control mechanism. It occurs after 985004 (Maximum replication delay has been exceeded) events or 985003 (Unable to find path to disk in the remote cluster (system) within the time-out period).

A 1920 error (event ID 050010) can have several triggers, including the following probable causes:

- ▶ Primary IBM Spectrum Virtualize system or SAN fabric problem (10%)
- ▶ Primary IBM Spectrum Virtualize system or SAN fabric configuration (10%)
- ▶ Secondary IBM Spectrum Virtualize system or SAN fabric problem (15%)
- ▶ Secondary IBM Spectrum Virtualize system or SAN fabric configuration (25%)
- ▶ Intercluster link problem (15%)
- ▶ Intercluster link configuration (25%)

In practice, the most often overlooked cause is latency. Global Mirror has a round-trip-time tolerance limit of 80 or 250 milliseconds, depending on the firmware version and the hardware model (see Figure 10-72 on page 543). A message that is sent from your source IBM Spectrum Virtualize system to your target system and the accompanying acknowledgment must have a total time of 80 or 250 milliseconds round trip. That is, it must have up to 40 or 125 milliseconds latency each way.

The primary component of your round-trip time is the physical distance between sites. For every 1000 kilometers (621.4 miles), you observe a 5-millisecond delay each way. This delay does not include the time that is added by equipment in the path. Every device adds a varying amount of time depending on the device, but a good rule is 25 microseconds for pure hardware devices.

For software-based functions (such as compression that is implemented in applications), the added delay tends to be much higher (usually in the millisecond plus range.) Next, we describe an example of a physical delay.

Physical delay example

Company A has a production site that is 1900 kilometers (1180.6 miles) away from its recovery site. The network service provider uses a total of five devices to connect the two sites. In addition to those devices, Company A employs a SAN FC router at each site to provide Fibre Channel over IP (FCIP) to encapsulate the FC traffic between sites.

Now, there are seven devices, and 1900 kilometers (1180.6 miles) of distance delay. All the devices are adding 200 microseconds of delay each way. The distance adds 9.5 milliseconds each way, for a total of 19 milliseconds. Combined with the device latency, the delay is 19.4 milliseconds of physical latency minimum, which is under the 80-millisecond limit of Global Mirror until you realize that this number is the best case number.

The link quality and bandwidth play a large role. Your network provider likely ensures a latency maximum on your network link. Therefore, be sure to stay as far beneath the Global Mirror round-trip-time (RTT) limit as possible. You can easily double or triple the expected physical latency with a lower quality or lower bandwidth network link. Then, you are within the range of exceeding the limit if high I/O occurs that exceeds the existing bandwidth capacity.

When you get a 1920 event, always check the latency first. If the FCIP routing layer is not properly configured, it can introduce latency. If your network provider reports a much lower latency, you might have a problem at your FCIP routing layer. Most FCIP routing devices have built-in tools to enable you to check the RTT.

When you are checking latency, remember that TCP/IP routing devices (including FCIP routers) report RTT using standard 64-byte ping packets.

In Figure 10-127, you can see why the effective transit time must be measured only by using packets that are large enough to hold an FC frame, or 2148 bytes (2112 bytes of payload and 36 bytes of header). Allow estimated resource requirements to be a safe amount, because various switch vendors have optional features that might increase this size. After you verify your latency by using the proper packet size, proceed with normal hardware troubleshooting.

Packet Size	Link Size	Serialization Delay (Time Required to Send Data)	Unit
64	256 Kbps	2.0E+03	microseconds
64	1.5 Mbps	3.4E+02	microseconds
64	100 Mbps	5.1E+00	microseconds
64	155 Mbps	3.3E+00	microseconds
64	622 Mbps	8.2E-01	microseconds
64	1 Gbps	5.1E-04	microseconds
64	10 Gbps	5.1E-05	microseconds
1500	256 Kbps	4.7E+04	microseconds
1500	1.5 Mbps	8.0E+03	microseconds
1500	100 Mbps	1.2E+02	microseconds
1500	155 Mbps	7.7E+01	microseconds
1500	622 Mbps	1.9E+01	microseconds
1500	1 Gbps	1.2E+01	microseconds
1500	10 Gbps	1.2E+00	microseconds
2148	256 Kbps	6.7E+04	microseconds
2148	1.5 Mbps	1.1E+04	microseconds
2148	100 Mbps	1.7E+02	microseconds
2148	155 Mbps	1.1E+02	microseconds
2148	622 Mbps	2.8E+01	microseconds
2148	1 Gbps	1.7E+01	microseconds
2148	10 Gbps	1.7E-03	microseconds

Figure 10-127 Effect of packet size (in bytes) versus the link size

Before proceeding, look at the second largest component of your RTT, which is *serialization delay*. Serialization delay is the amount of time that is required to move a packet of data of a specific size across a network link of a certain bandwidth. The required time to move a specific amount of data decreases as the data transmission rate increases.

Figure 10-127 also shows the orders of magnitude of difference between the link bandwidths. It is easy to see how 1920 errors can arise when your bandwidth is insufficient. Never use a TCP/IP ping to measure RTT for FCIP traffic.

In Figure 10-127, the amount of time in microseconds that is required to transmit a packet across network links of varying bandwidth capacity is compared. The following packet sizes are used:

- ▶ 64 bytes: The size of the common ping packet
- ▶ 1500 bytes: The size of the standard TCP/IP packet
- ▶ 2148 bytes: The size of an FC frame

Finally, your path maximum transmission unit (MTU) affects the delay that is incurred to get a packet from one location to another location. An MTU might cause fragmentation or be too large and cause too many retransmits when a packet is lost.

10.11.2 1720 error

The 1720 error (event ID 050020) is the other problem remote copy might encounter. The amount of bandwidth that is needed for system-to-system communications varies based on the number of nodes. It is important that it is not zero. When a partner on either side stops communication, you see a 1720 appear in your error log. According to the product documentation, there are no likely field-replaceable unit breakages or other causes.

The source of this error is most often a fabric problem or a problem in the network path between your partners. When you receive this error, check your fabric configuration for zoning of more than one host bus adapter (HBA) port for each node per I/O Group if your fabric has more than 64 HBA ports zoned. One port for each node per I/O Group per fabric that is associated with the host is the suggested zoning configuration for fabrics.

For those fabrics with 64 or more host ports, this recommendation becomes a rule. Therefore, you see four paths to each volume discovered on the host because each host needs to have at least two FC ports from separate HBA cards, each in a separate fabric. On each fabric, each host FC port is zoned to two of node ports where each port comes from one node canister. This gives four paths per host volume. More than four paths per volume are supported but not recommended.

Improper zoning can lead to SAN congestion, which can inhibit remote link communication intermittently. Checking the zero buffer credit timer from IBM Virtual Storage Center and comparing against your sample interval reveals potential SAN congestion. If a zero buffer credit timer is above 2% of the total time of the sample interval, it might cause problems.

Next, always ask your network provider to check the status of the link. If the link is acceptable, watch for repeats of this error. It is possible in a normal and functional network setup to have occasional 1720 errors, but multiple occurrences might indicate a larger problem.

If you receive multiple 1720 errors, recheck your network connection and then check the system partnership information to verify its status and settings. Then, proceed to perform diagnostics for every piece of equipment in the path between the two Storwize systems. It often helps to have a diagram that shows the path of your replication from both logical and physical configuration viewpoints.

If your investigations fail to resolve your remote copy problems, contact your IBM Support representative for a more complete analysis.

10.12 HyperSwap

The HyperSwap high availability function allows business continuity in a hardware failure, power failure, connectivity failure, or disasters. HyperSwap is available on the IBM Spectrum Virtualize, IBM Storwize V7000, IBM Storwize V7000 Unified, and IBM Storwize V5000 products.

The HyperSwap feature provides dual-site, active-active access to a volume. HyperSwap functions are available on systems that can support more than one I/O group. This function provides HA solution that are accessible through two sites at up to 300 km (186.4 miles). A fully independent copy of the data is maintained at each site. When data is written by hosts at either site, both copies are synchronously updated before the write operation is completed. The HyperSwap function automatically optimizes itself to minimize data that is transmitted between two sites and to minimize host read and write latency.

If the nodes go offline or the storage at either site goes offline, leaving an online and accessible up-to-date copy, the HyperSwap function can automatically fail over access to the online copy. The HyperSwap function also automatically resynchronizes the two copies when possible.

HyperSwap capability enables each volume to be presented by two I/O groups. The configuration tolerates combinations of node and site failures, by using the same flexible choices of host multipathing driver interoperability that are available for the IBM Storwize. The use of FlashCopy helps maintain a golden image during automatic resynchronization.

Important: Because Remote Mirroring is used to support the HyperSwap capability, Remote Mirroring licensing is a requirement for using HyperSwap.

The HyperSwap function uses a hyperswap topology by spreading the control enclosure of the system across two sites, with storage at a third site that acts as a tie-breaking quorum device. Consider the following points:

- ▶ The HyperSwap topology requires at least one control enclosure in each main data site. Therefore, to get a volume that is resiliently stored on both sites, one control enclosure with the necessary storage capacity is required.
- ▶ The HyperSwap topology uses more system resources to support a fully independent cache on each site, which provides full performance, even if one site is lost.
- ▶ The HyperSwap function can be configured by using the GUI or command-line interface (CLI).
- ▶ The hosts, Storwize V5000 control enclosures, and Storwize V5000 storage enclosures are in one of two failure domains or sites. External virtualized storage capacity can also be used.
- ▶ Volumes are visible as a single object across both sites (the Storwize V5000 control enclosure).

At least two control enclosures are required for HyperSwap. System scalability depends on the hardware details, as listed in Table 10-12.

Table 10-12 HyperSwap support

	V5010	V5020	V5030 /V5030F	V7000	IBM SAN Volume Controller
Maximum number of I/O groups	1	1	2	4	4
Support for HyperSwap	No	No	Yes	Yes	Yes
Support for Stretched Cluster	No	No	No	No	Yes

So, a V5000 HyperSwap cluster is always restricted to a single control enclosure per site. The Storwize V7000 and IBM SAN Volume Controller can provide more scalability and offer more flexibility.

10.12.1 Introduction to HyperSwap volumes

The HyperSwap function is built on the Remote Copy features that include Metro Mirror, Global Mirror, and Global Mirror with Change Volumes.

The HyperSwap function works with the standard multipathing drivers that are available on various host types. No extra host support is required to access the highly available volume. Where multipathing drivers support Asymmetric Logical Unit Access (ALUA), the storage system informs the multipathing driver about the nodes that are in the same site and the nodes that need to be used to minimize I/O latency.

The host and Storwize V5000 site attributes must be configured to enable this optimization and to enable HyperSwap functionality. A three-site setup is required. Two sites are used as the main data center to provide two independent data copies. A quorum disk or an IP-based quorum can be used as a quorum device. However, the quorum device must be placed in a third, independent site.

The quorum disk must be supported as an “extended quorum device”. The connection can be implemented by using Fibre Channel, Fibre Channel through wavelength-division multiplexing (WDM), synchronous digital hierarchy (SDH) and synchronous optical network (SONET), or FCIP. The minimum bandwidth is 2 MBps.

The IP quorum substitutes the active quorum disk’s tiebreaker role. Redundancy can be implemented by using multiple quorum apps, similar to multiple quorum disks. However, only one app is active at a time. The other apps are available if the active quorum device app fails.

Note: For more information about quorum devices, see the [Storwize V5000 IBM Knowledge Center](#).

Because HyperSwap is running as a single cluster that is distributed across two main data centers, one Storwize V5000 control enclosure is required in each site. Both control enclosures must be added to the same cluster. Only the Storwize V5030 supports the clustering of two control enclosures, so two Storwize V5030 control enclosures are required for HyperSwap. Metro Mirror is used to keep both data copies in sync.

The host accesses both I/O groups, as shown in Figure 10-128. The original Metro Mirror target ID is not used for host access. Instead, HyperSwap presents the Metro Mirror source ID for the target volume to the host. From the host perspective, the same volume is available on both I/O groups, although the Storwize V5000 volumes are connected through Metro Mirror.

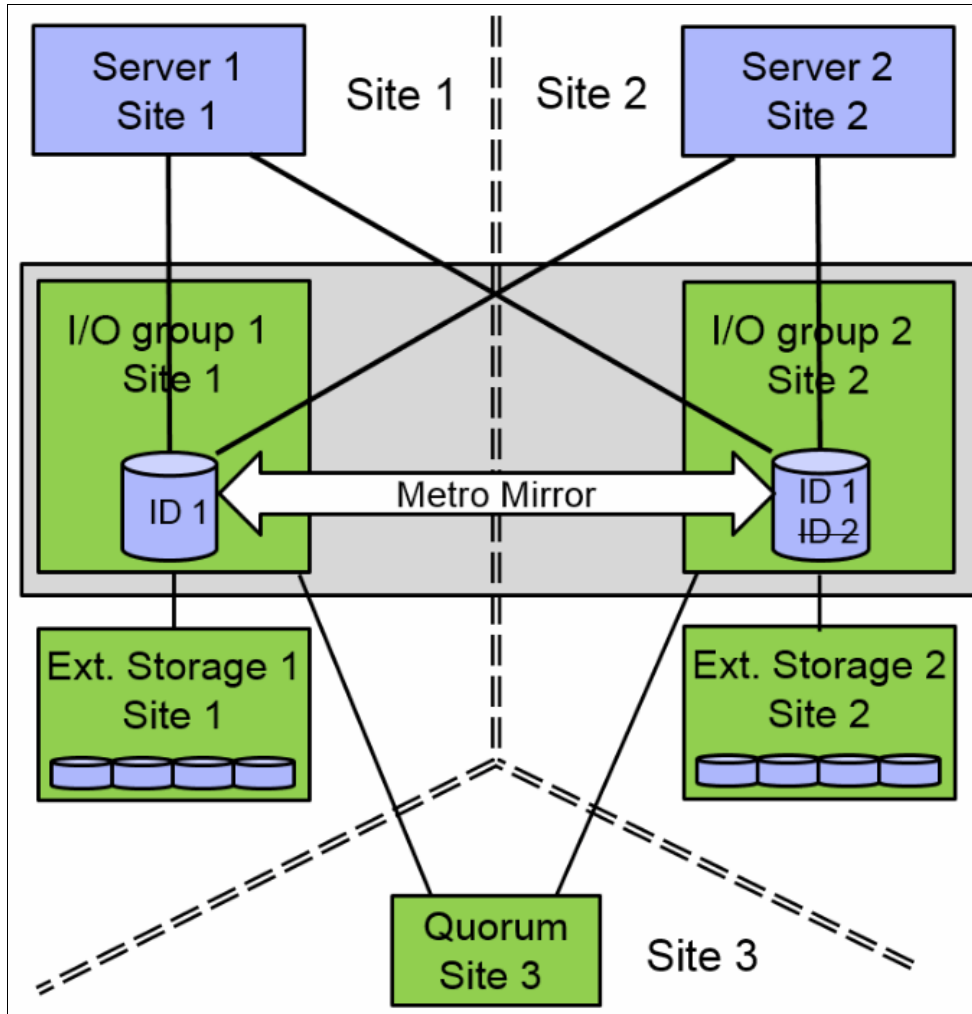


Figure 10-128 HyperSwap

A site attribute must be set for any host, Storwize V5000 storage system, and external virtualized storage system. The host uses the local I/O group (same site attribute) for data access, as shown in Figure 10-129.

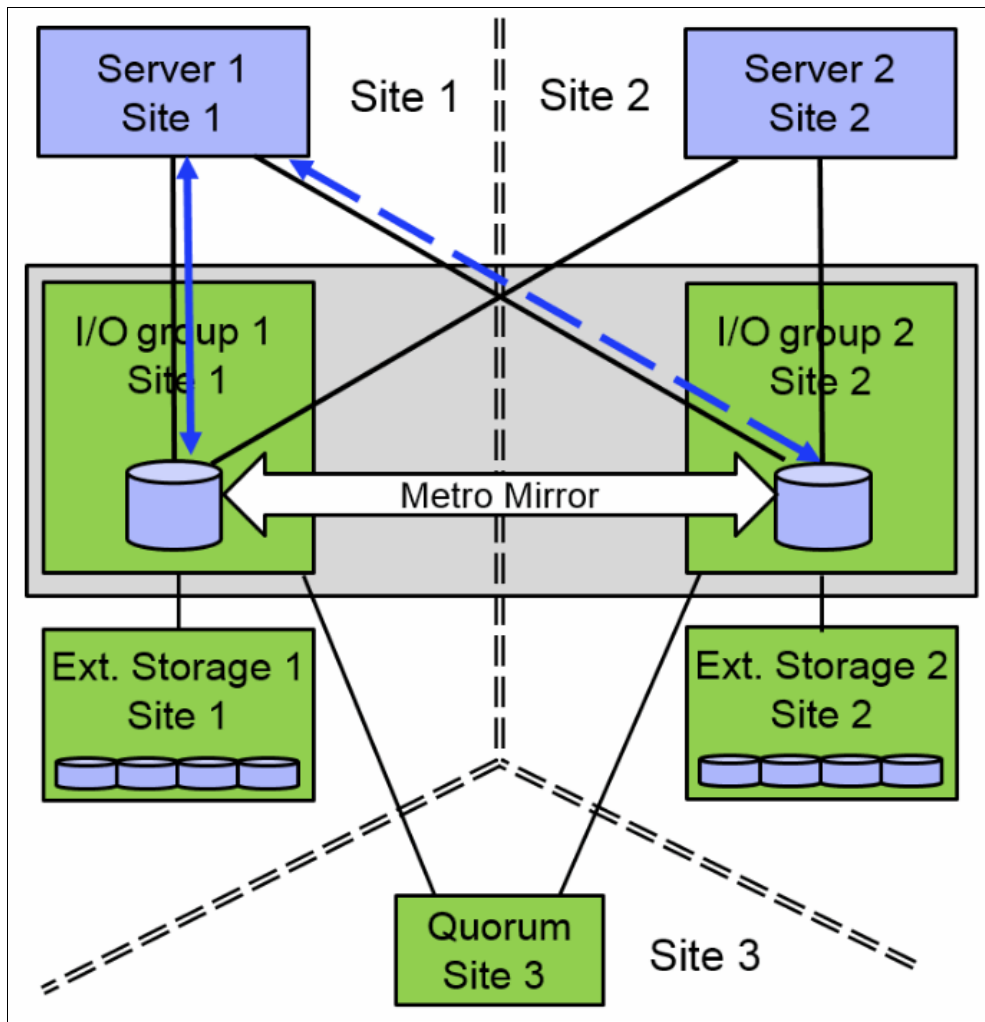


Figure 10-129 Data access

The continuous blue line shows the host default access path to the volume at the same site. The dotted blue line shows the non-preferred access path that is used if the preferred access path is not available. Accessing both I/O groups doubles the number of paths from host to volume. Take note of the limited number of supported paths for your multipath device driver and limit the number of paths to an acceptable level.

Data flow

The host reads and writes data to the local I/O group within the same site. The HyperSwap system sends the data to the remote site by using internal Metro Mirror, as shown in Figure 10-130.

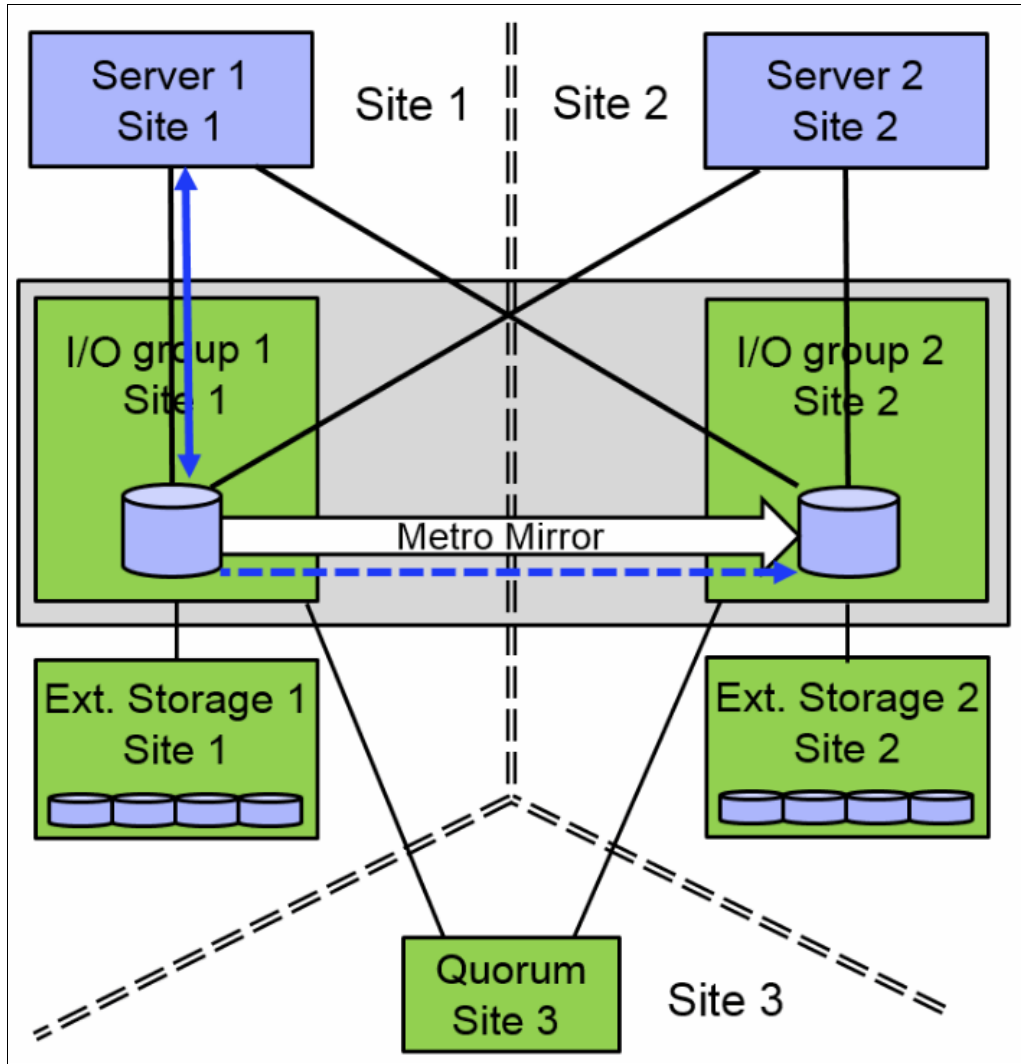


Figure 10-130 Data flow for a single volume

If a host accesses the volume on the Metro Mirror target site, all read and write requests can be forwarded to the I/O group that acts as the Metro Mirror source volume, as shown in Figure 10-131. All host-related traffic must be handled from the remote I/O group, which increases the long-distance data traffic.

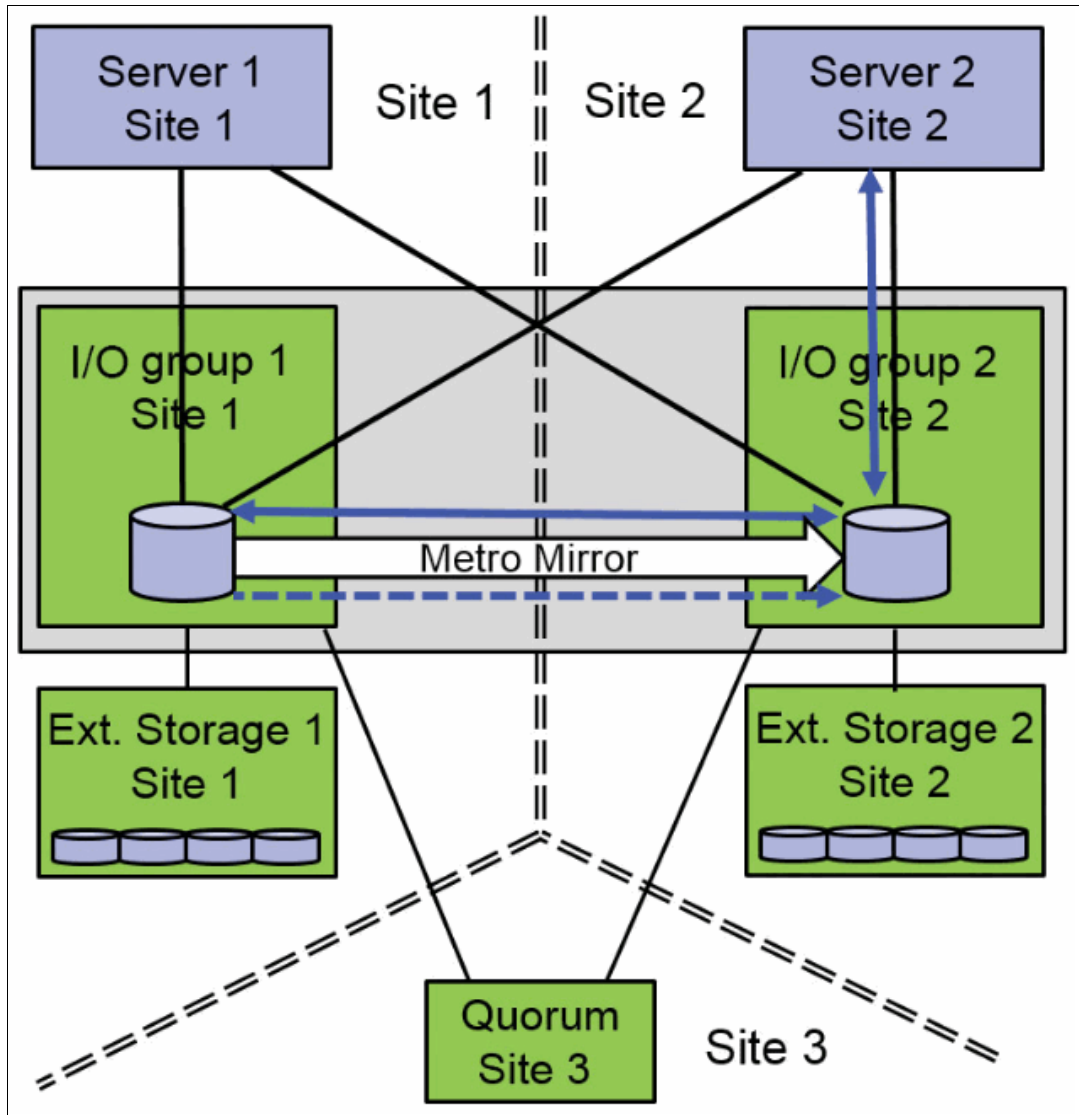


Figure 10-131 Data flow from the “wrong” site

10.12.2 Failure scenarios

If one node fails, the other node in the same I/O group takes over the responsibility for all volumes that are owned by the affected I/O group, as shown in Figure 10-133. The system can deactivate the write cache, which might influence the overall performance. The multipath driver can switch the active paths to the named node. The Metro Mirror relationship continues to operate and provides a synchronous, consistent copy at the remote I/O group.

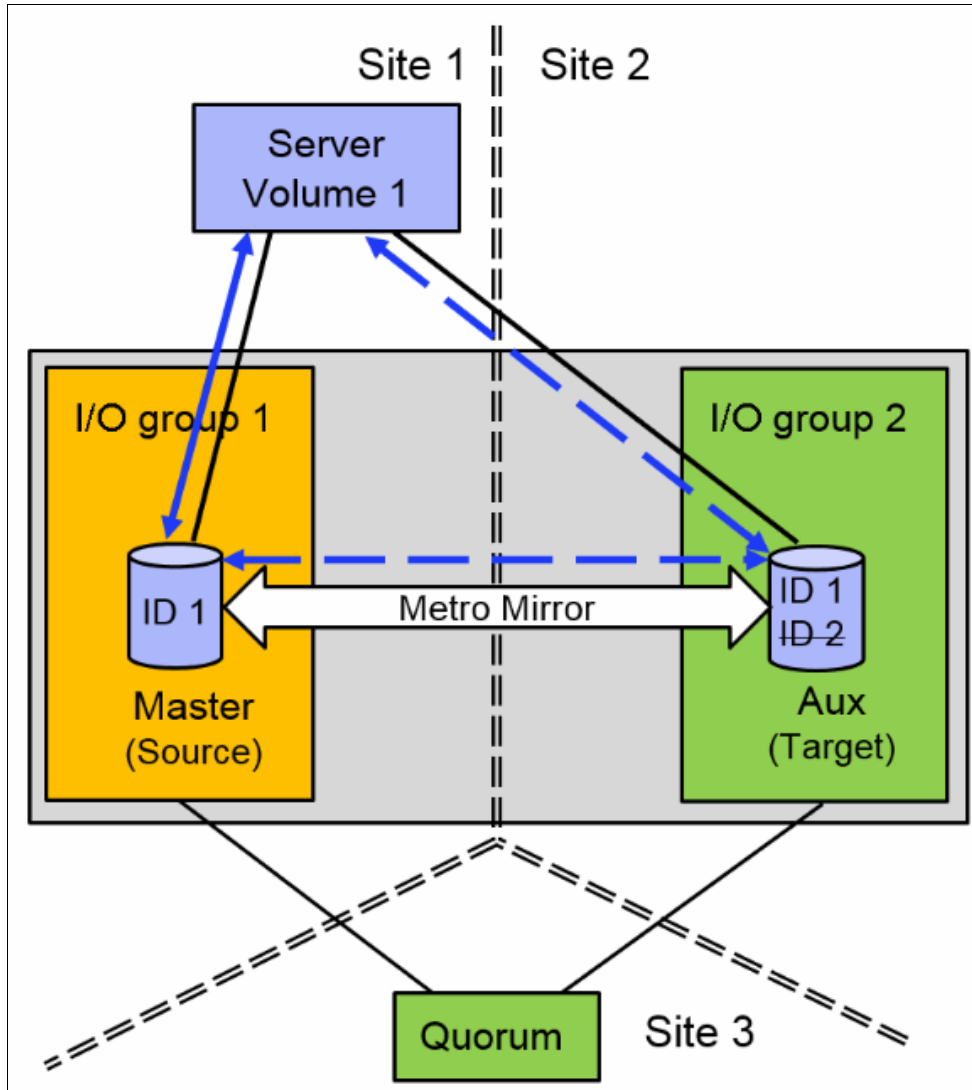


Figure 10-133 Single node failure in an I/O group

If an I/O group fails, the host can use the second I/O group at the remote site, as shown in Figure 10-134. The remote I/O group handles all volume-related traffic, but HyperSwap cannot keep both copies in sync anymore because of an inactive I/O group.

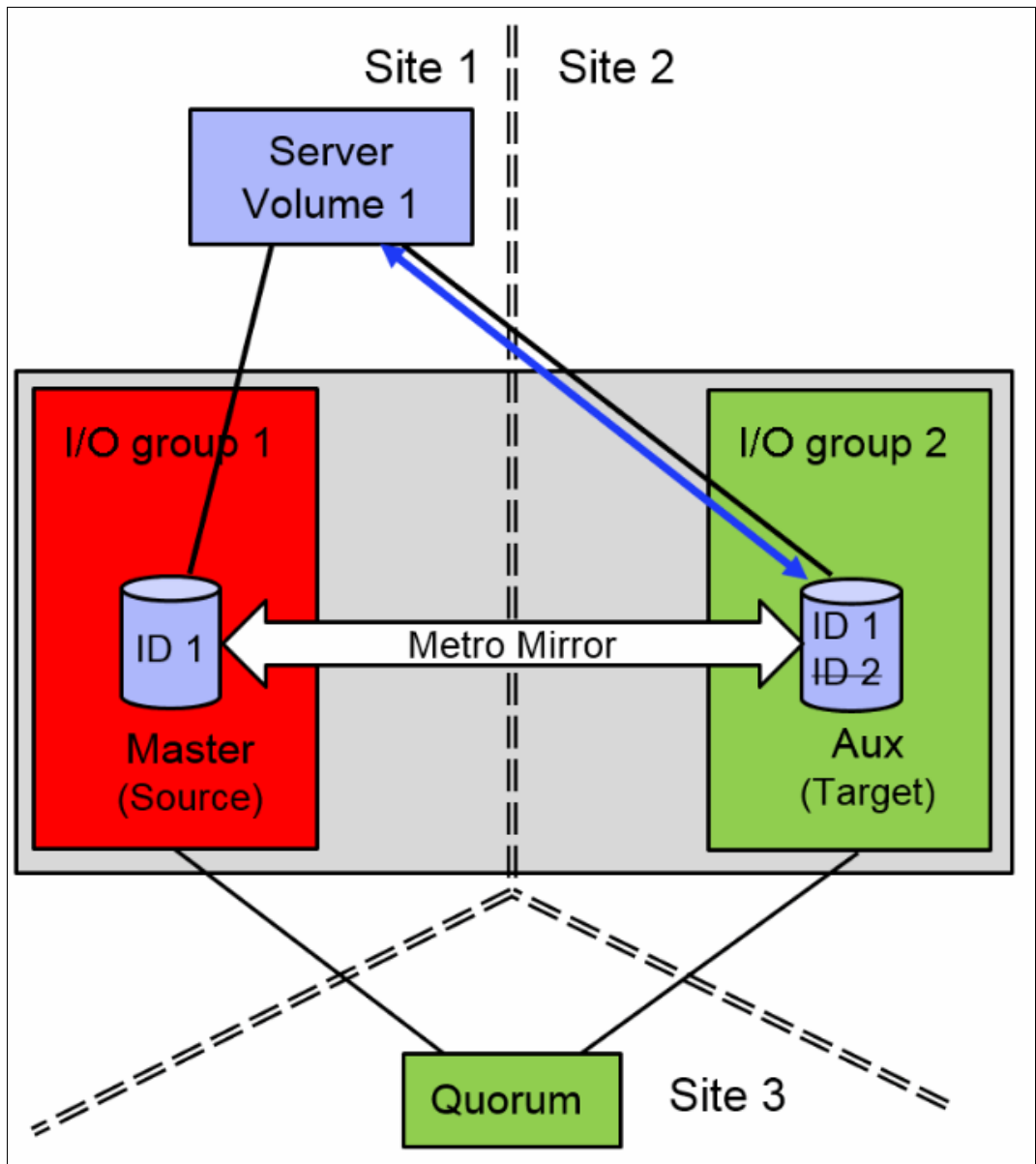


Figure 10-134 I/O group failure

When the failed I/O group is back, the system can automatically resynchronize both copies in the background. Before the resynchronization, the Storwize V5000 can perform a FlashCopy on HyperSwap source and target volumes, as shown in Figure 10-135. Each change volume requires two FlashCopy relationships, one relationship in each direction. So, four FlashCopy relationships are required for each HyperSwap volume.

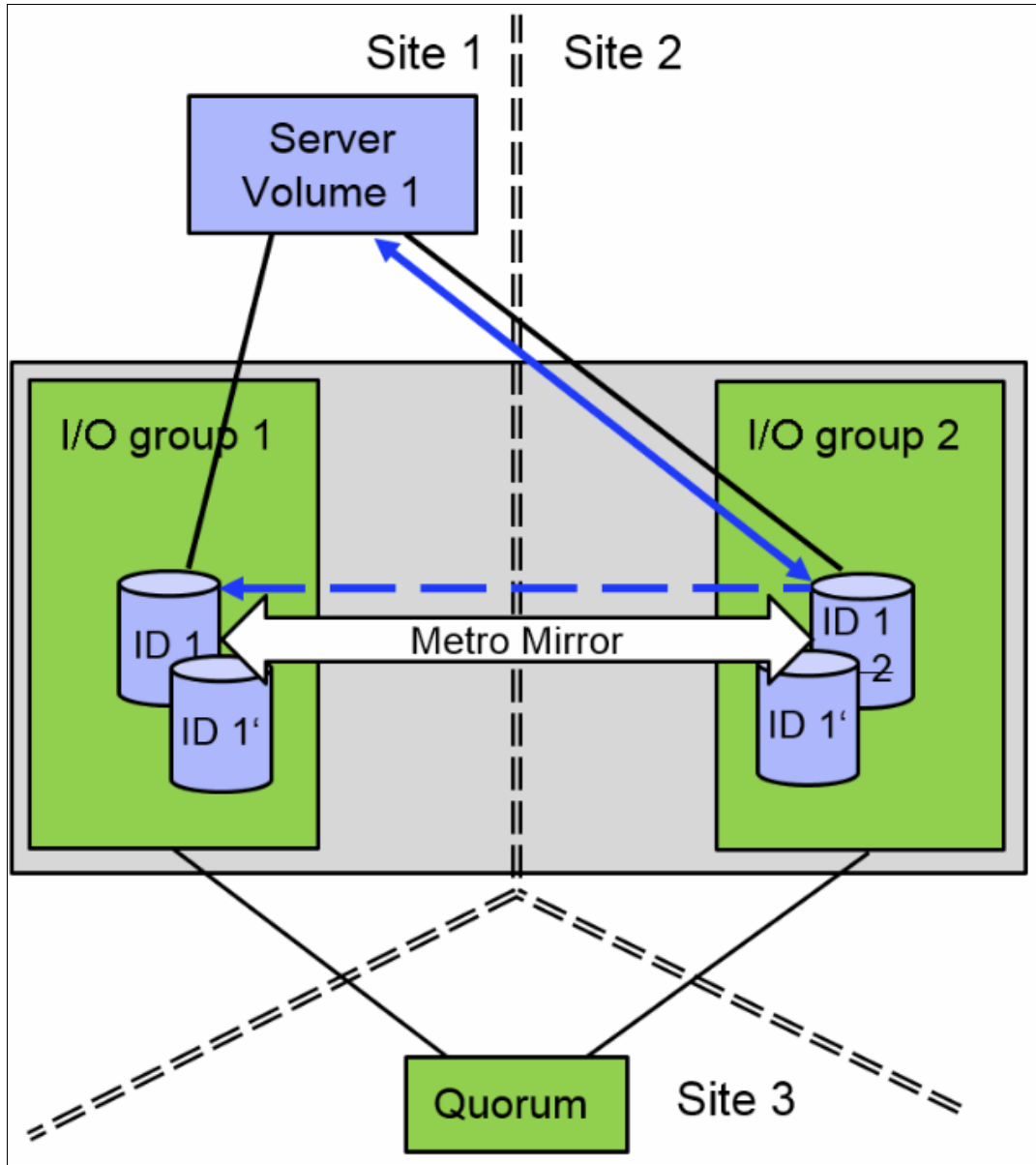


Figure 10-135 Resynchronization

To provide an easy to use GUI, those relationships and FlashCopy volumes are not shown in the GUI. They are only visible and manageable by using the CLI.

After successful resynchronization, the host switches automatically to the I/O group at the same site for the best performance and limited inter-switch link (ISL) usage, as shown in Figure 10-136.

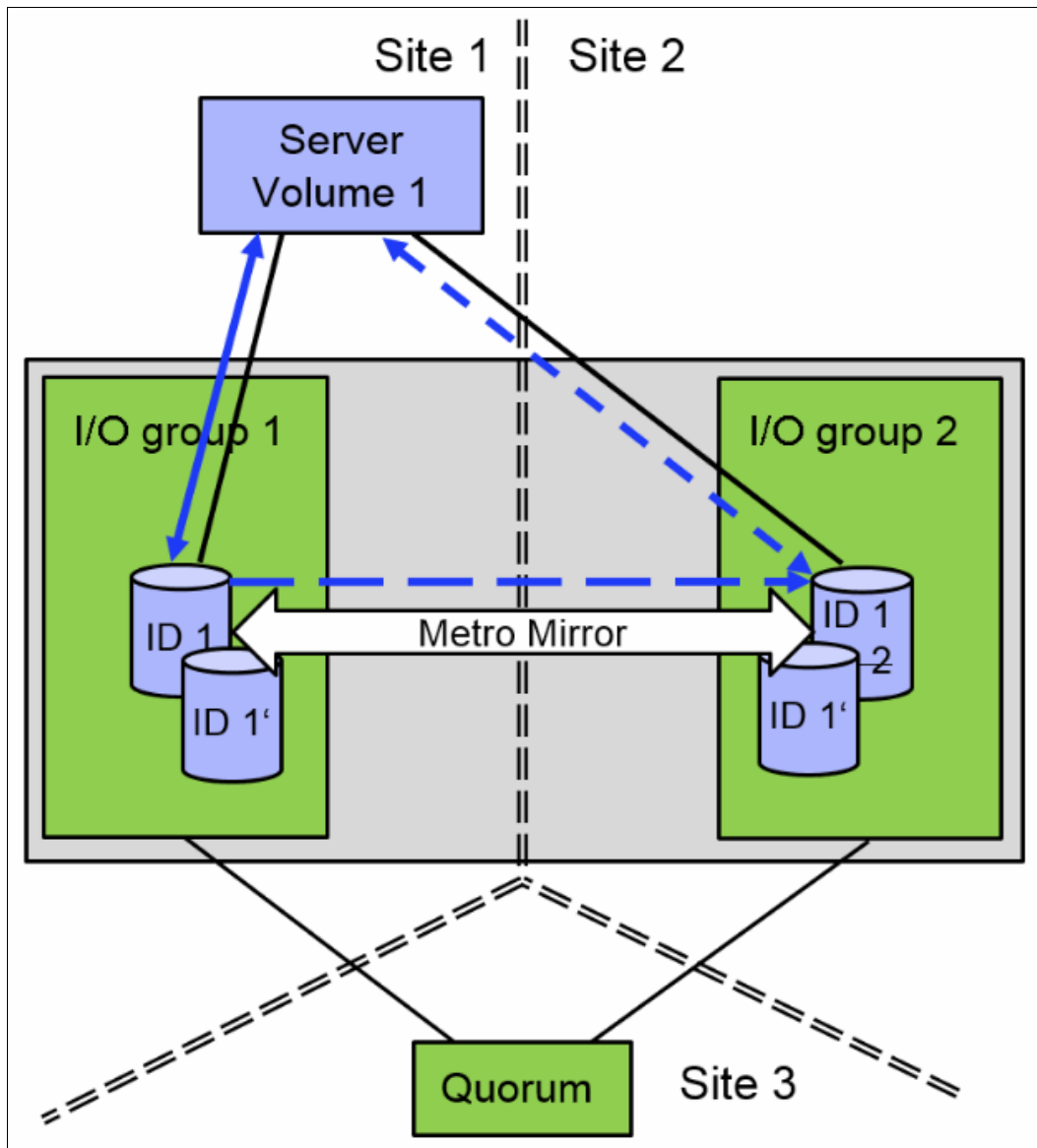


Figure 10-136 Data flow after resynchronization

HyperSwap uses Metro Mirror technology, which enables the usage of Metro Mirror consistency groups that are described in “Remote Copy Consistency Groups” on page 534.

10.12.3 Current HyperSwap limitations

HyperSwap has the following limitations:

- ▶ Cluster internal Metro Mirror is used for replication, so the size of a HyperSwap volume cannot be changed by using `expandvdisksize` and `shrinkvdisksize`.
- ▶ A cascaded Remote Copy is not available. HyperSwap volumes cannot be replicated to a second, independent storage system by using Remote Copy functionality.

The number of HyperSwap volumes supported per IBM Storwize V5000 system (for V5030) is 1024. More FlashCopy requirements can reduce the number of possible HyperSwap volumes.

For more information about the configuration limits search for “Configuration Limits and Restrictions for IBM Storwize V5000” at [IBM Knowledge Center](#).

Consider the following points:

- ▶ Hosts that access a HyperSwap volume through internet Small Computer System Interface (iSCSI) or serial-attached SCSI (SAS) cannot take advantage of the high availability function.
- ▶ FlashCopy usage can be complicated because the Metro Mirror source volume and target volume can switch during daily operation. Because of this possibility, the identification of the copy direction is required for a successful FlashCopy.
- ▶ The Remote Copy relationship must be removed first for a reverse FlashCopy operation. After a reverse FlashCopy, all HyperSwap functions must be implemented manually again (Remote Mirror + FlashCopy relationships).
- ▶ IBM FlashCopy Manager is not supported by HyperSwap volumes.

Note: For more information about HyperSwap, see *IBM Storwize V7000, Spectrum Virtualize, HyperSwap, and VMware Implementation*, [SG24-8317](#).



External storage virtualization

This chapter describes how to incorporate external storage systems, or *storage controllers*, into the virtualized world of the IBM Storwize V5030.

A key feature of the IBM Storwize V5030 is its ability to consolidate disk controllers from various vendors into storage pools. By external disk controllers, the storage administrator can manage and provision storage to applications from a single user interface and use a common set of advanced functions across all of the storage systems under the control of the IBM Storwize V5030.

A distinction must be made between virtualizing external storage and importing data into the IBM Storwize V5030. Virtualizing external storage means the creation of logical units with no data on them and the addition of these logical units to storage pools under the IBM Storwize V5030 control. In this way, the external storage can benefit from the IBM Storwize V5030 features, such as Easy Tier and Copy Services.

When data needs to be put under the control of the IBM Storwize V5030, it must first be imported as an *image mode volume*. It is strongly recommended to copy the data onto internal or external storage that is under the control of the IBM Storwize V5030 instead of allowing the data within an image mode volume, so the data can benefit from the IBM Storwize V5030 features.

Note: External storage virtualization is available on the Storwize V5030 model only. It is not available on the Storwize V5010 or Storwize V5020. However, these models can still import data from external storage systems. For more information about storage migration, see Chapter 7, “Storage migration” on page 357, and 11.2.4, “Importing image mode volumes” on page 615.

This chapter includes the following topics:

- ▶ 11.1, “Planning for external storage virtualization” on page 608
- ▶ 11.2, “Working with external storage” on page 611

11.1 Planning for external storage virtualization

This section describes how to plan for virtualizing external storage with the IBM Storwize V5030. Virtualizing the storage infrastructure with the IBM Storwize V5030 makes your storage environment more flexible, cost-effective, and easy to manage. The combination of the IBM Storwize V5030 and an external storage system allows more storage capacity benefits from the powerful software functions within the IBM Storwize V5030.

The external storage systems that are incorporated into the IBM Storwize V5030 environment can be new or existing systems. Any data on the existing storage systems can be easily migrated to an environment that is managed by the IBM Storwize V5030, as described in Chapter 7, “Storage migration” on page 357.

11.1.1 License for external storage virtualization

From a licensing standpoint, when external storage systems are virtualized by the IBM Storwize V5030, a per-enclosure External Virtualization license is required.

Migration: If the IBM Storwize V5030 is used as a general storage management tool, you must order the correct External Virtualization licenses. The only exception is if you want to migrate data from external storage systems to IBM Storwize V5030 internal storage and then remove the external storage. You can temporarily configure your External Storage license for a 45-day period. For more than a 45-day migration requirement, the correct External Virtualization license must be ordered.

You can configure the IBM Storwize V5030 licenses by clicking the **Settings** icon and then, **System** → **Licensed Functions**. For more information about setting licenses on the IBM Storwize V5030, see Chapter 2, “Initial configuration” on page 39.

For assistance with licensing questions or to purchase any of these licenses, contact your IBM account team or IBM Business Partner.

11.1.2 Configuration planning for external virtualization

External virtualization is supported by using Fibre Channel (FC) or Fibre Channel over Ethernet (FCoE). IBM Storwize V5030 can also virtualize back-end storage systems that are connected over internet Small Computer System Interface (iSCSI).

Note: It is not supported to present single back-end storage LU to IBM Storwize with both FC and iSCSI at the same time. Only one protocol must be used.

Planning for FC attachment

The following prerequisites must be met when you are planning for FC attachment:

- ▶ Base IBM Storwize V5030 system has no FC ports. A Host Interface Card (HIC) needs to be installed into each Storwize V5030 node.

To virtualize a Fibre Channel attached controller, install a pair of the optional 16 Gb Fibre Channel adapter cards (or a pair of the optional 10 GbE adapter cards if you use FCoE).

Any supported controller can be temporarily direct attached for the purposes of migrating the data from that controller onto the Storwize system. For virtualization, external storage controllers are connected through storage area network (SAN) switches.

Note: At the time of this writing, two back-end storage systems were supported as permanently directly attached: IBM DS8000 and IBM FlashSystem 900. For more information, see this [IBM Support web page](#).

- ▶ Ensure that the switches or directors are at the firmware levels that are supported by the IBM Storwize V5030 and that the port login maximums that are listed in the restriction document are not exceeded. For more information, search for “Configuration Limits and Restrictions for IBM Storwize V5000” at [this IBM Knowledge Center web page](#).

The suggested SAN configuration is based in a dual fabric solution. The ports on external storage systems and the IBM Storwize V5030 ports must be evenly split between the two fabrics to provide redundancy if one of the fabrics goes offline.

- ▶ After the IBM Storwize V5030 and external storage systems are connected to the SAN fabrics, zoning on the switches must be configured. In each fabric, create a zone with the four IBM Storwize V5030 worldwide port names (WWPNs), two from each node canister with up to a maximum of eight WWPNs from each external storage system.

Ports: The IBM Storwize V5030 supports a maximum of 16 ports or WWPNs from an externally virtualized storage system.

Figure 11-1 shows an example of how to cable devices to the SAN. Refer to this example as we describe the zoning. For this example, we used an IBM Storwize V3700 as our external storage.

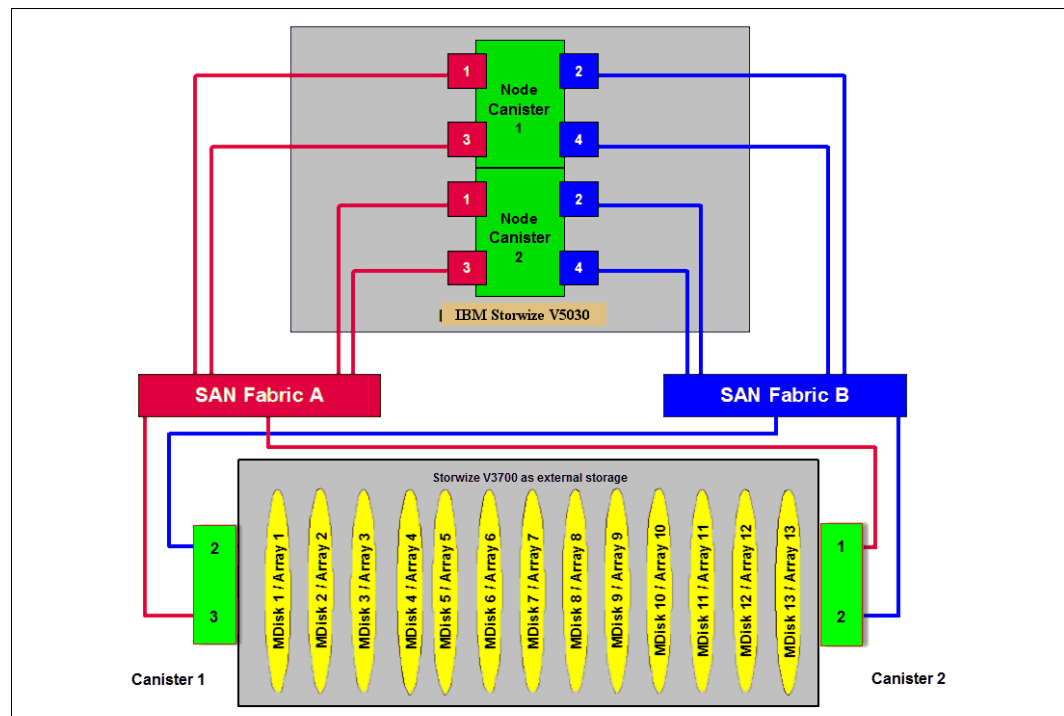


Figure 11-1 SAN cabling and zoning example

- ▶ Create an IBM Storwize V5030/external storage zone for each storage system to be virtualized, as shown in the following examples:
 - Zone the external Storwize V3700 canister 1 port 2 with the Storwize V5030 canister 1 port 2 and canister 2 port 2 in the blue fabric.

- Zone the external Storwize V3700 canister 2 port 2 with the Storwize V5030 canister 1 port 4 and canister 2 port 4 in the blue fabric.
- Zone the external Storwize V3700 canister 1 port 3 with the Storwize V5030 canister 1 port 1 and canister 2 port 1 in the red fabric.
- Zone the external Storwize V3700 canister 2 port 1 with the Storwize V5030 canister 1 port 3 and canister 2 port 3 in the red fabric.

Planning for iSCSI attachment

The iSCSI protocol has many different implementations across products and vendors. Therefore, be aware of some controller-specific considerations when virtualizing storage over iSCSI.

Back-end systems need to be attached to IBM Storwize V5030 through Ethernet switches. Direct connection is not supported for iSCSI attachment.

For security reasons, you might want to physically or logically separate iSCSI traffic that is being sent by the IBM Storwize product to the external storage from other parts of the network. You can do this separation by provisioning a separate subnet or VLAN for traffic between the IBM Storwize V5030 storage system and external storage.

In addition to segmenting the traffic between the system and the external storage away from the rest of the network, you might want to configure the authentication between the two systems. To do this, configure one-way CHAP.

For more information about planning and considerations for your specific controller, refer to [IBM Knowledge Center](#) and *iSCSI Implementation and Best Practices on IBM Storwize Storage Systems*, SG24-8327.

Planning for system layers

If the external controller is a Storwize system, system layers need to be configured. By default, IBM Storwize V5030 is set to *storage layer*. With it, it cannot virtualize another IBM Storwize system, which is configured to the same layer. To be able to do that, IBM Storwize V5030 needs to be set to *replication layer*. Then, it can see external Storwize systems as storage controllers and can virtualize LUs provided by them.

You also must switch to *replication layer* if you are going to virtualize iSCSI backend controllers.

Make sure that the layers are correct before zoning the two systems together.

For more information about layers and how to change them, see [IBM Knowledge Center](#).

11.1.3 External storage configuration planning

Logical units that are created on the external storage system must provide redundancy through various RAID levels, which prevents a single physical disk failure from causing a managed disk (MDisk), storage pool, or associated host volume from getting offline. To minimize the risk of data loss, virtualize storage systems only where logical unit numbers (LUNs) are configured by using a RAID level other than RAID 0 (RAID 1, RAID 10, RAID 0+1, RAID 5, RAID 6, Distributed RAID 5, or Distributed RAID 6):

- ▶ Verify that the storage controllers to be virtualized by the IBM Storwize V5030 meet the configuration restrictions. For more information, search for “Configuration Limits and Restrictions for IBM Storwize V5000” at this [IBM Knowledge Center web page](#).

- ▶ Ensure that the firmware or microcode levels of the storage controllers to be virtualized are supported by the IBM Storwize V5030. For more information, see the [IBM System Storage Interoperation Center \(SSIC\)](#).

The IBM Storwize V5030 must have exclusive access to the LUNs from the external storage system that are presented to it. LUNs cannot be shared between the IBM Storwize V5030 and other storage virtualization platforms or between an IBM Storwize V5030 and hosts. However, different LUNs can be mapped from the same external storage system to an IBM Storwize V5030 and other hosts in the SAN through different storage ports.

- ▶ Ensure that the external storage subsystem LUN masking is configured to map all LUNs to all of the WWPNs (IQNs) in the IBM Storwize V5030 storage system. Ensure that you check the IBM Storwize V5030 [IBM Knowledge Center](#) and review the “Configuring and servicing external storage system” topic before you prepare the external storage systems for discovery from the IBM Storwize V5030 system.

11.1.4 Guidelines for virtualizing external storage

When external storage is virtualized by using the IBM Storwize V5030, the following guidelines must be followed:

- ▶ Avoid splitting arrays into multiple LUNs at the external storage system level. When possible, create a single LUN per array for mapping to the IBM Storwize V5030.
- ▶ Use 6 - 8 disks per RAID group when you create the external LUNs. If you create more than eight disks, a longer rebuild time might result if a single disk fails, which can affect the performance of the LUN and expose it to complete failure if a second disk fails during the rebuild.

Also, the smaller the number of disks, the more likely it is that write operations span an entire stripe (stripe size, which is multiplied by the number of members, minus one). In this case, write performance is improved.

- ▶ Except for Easy Tier, do not mix MDisks that vary in performance or reliability in the same storage pool. Put only MDisks of the same size and performance into the same storage pool. Likewise, group MDisks from different arrays into different pools. For more information about Easy Tier, see Chapter 9, “Advanced features for storage efficiency” on page 435.
- ▶ Do not leave volumes in image mode. Use image mode only to import or export data into or out of the IBM Storwize V5030. Migrate data from image mode volumes and associated MDisks to other storage pools to benefit from storage virtualization and the enhanced benefits of the Storwize V5030.

11.2 Working with external storage

This section describes how to manage external storage by using an IBM Storwize V5030.

The basic concepts of managing an external storage system are the same as the concepts for managing internal storage. The IBM Storwize V5030 discovers LUs from the external storage system as MDisks. These MDisks are added to a storage pool in which volumes are created and mapped to hosts, as needed.

To virtualize iSCSI backend, or if backend storage is another IBM Storwize system, IBM Storwize V5030 needs to be put into *replication layer*. This can be done with the CLI. You also get a warning that your Storwize is in a *storage layer* if you attempt to add iSCSI storage controller with GUI.

Note: Before you change the system layer, the following conditions must be met:

- ▶ No host object can be configured with worldwide port names (WWPNs) from a Storwize family system.
- ▶ No system partnerships can be defined.
- ▶ No Storwize family system can be visible on the SAN fabric.

To switch IBM Storwize V5030 system layer, use the `chsystem` CLI command, as shown in Example 11-1. If run successfully, no output is returned.

Example 11-1 Changing system layer

```
IBM_Storwize:ITS0V5030:superuser>lsystem |grep layer
layer storage
IBM_Storwize:ITS0V5030:superuser>chsystem -layer replication
IBM_Storwize:ITS0V5030:superuser>
```

For more information about layers and how to change them, see [IBM Knowledge Center](#).

To work with external storage controllers and external MDisks, use **Pools** → **External Storage**, as shown in Figure 11-2.

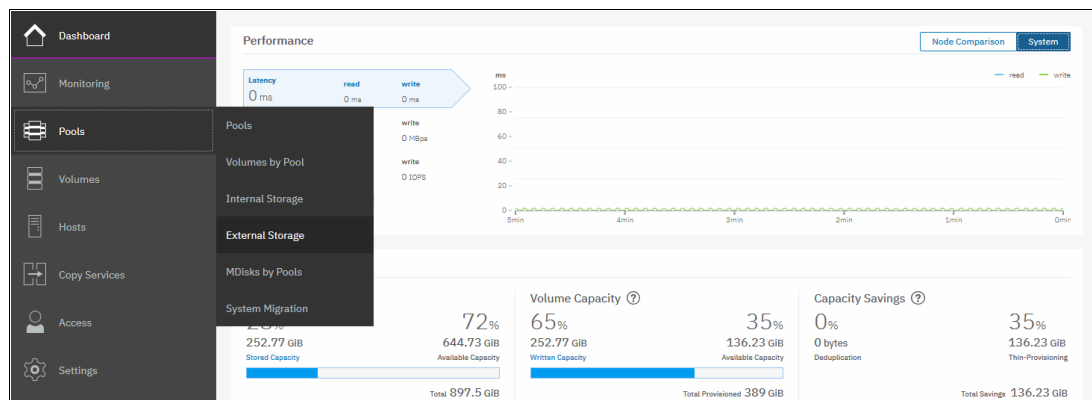


Figure 11-2 Accessing external storage management menu

Note: If the External Storage submenu is not present in Pools menu, you must set up the external virtualization license, as explained in 11.1.1, “License for external storage virtualization” on page 608.

11.2.1 Adding external FC controllers

To add FC attached external storage controller to the IBM Storwize V5030 virtualized environment, zone a minimum of two and a maximum of 16 FC ports from the external storage system with all available FC ports on the IBM Storwize V5030 system. (For more information about zoning, see 11.1.2, “Configuration planning for external virtualization” on page 608.) Because the IBM Storwize V5030 is virtualizing your storage, hosts must be zoned with the Storwize V5030 controller’s WWPNs.

The FC storage controller is automatically detected by IBM Storwize V5030 after it is zoned. To confirm this status, click **Pools** → **External Storage**, as shown in Figure 11-3.

Name	State	Capacity	Mode	Site
controller0	Online	IBM 2145	Unassigned	50
controller1	Online	IBM 2145	Unassigned	50
controller2	Online	IBM 2145	Unassigned	50
controller3	Online	IBM 2145	Unassigned	50

Figure 11-3 List of visible external controllers

Depending on the type of backend system, it might be detected as one or more controller objects. The system name that is configured on your external system is not visible on Storwize V5030. You can distinguish your controllers by WWNN or vendor ID.

To check which external controllers are visible with CLI, use the `lscontroller` command, as shown in Example 11-2.

Example 11-2 `lscontroller` output (some columns are not shown)

```

IBM_Storwize:ITS0V5030:superuser>lscontroller
id_controller_name ctrl_s/n          vendor_id          product_id_low
0 controller0      2076              IBM                2145
1 controller1      2076              IBM                2145
2 controller2      2076              IBM                2145
3 controller3      2076              IBM                2145

```

11.2.2 Adding external iSCSI controllers

You must manually configure iSCSI connections between the IBM Storwize V5030 and the external storage controller. Until then, the controller is not listed in the External Storage pane.

To start virtualizing iSCSI backend controller, you must follow the documentation in [IBM Knowledge Center](#) to perform configuration steps that are specific to your backend storage controller.

For more information about configuring Storwize V5030 to virtualize external storage with iSCSI, see *iSCSI Implementation and Best Practices on IBM Storwize Storage Systems*, SG24-8327.

11.2.3 Working with MDisks

Complete the following steps to prepare LUs on your backend system to be used as IBM Storwize V5030 MDisks:

1. By using the storage partitioning or LUN masking feature of external storage controller, create a group that includes all IBM Storwize V5030 WWPNs (IQNs).
2. Create equal size arrays on the external system by using any RAID level, except zero.
3. Create a single LUN per RAID array.
4. Map the LUNs to all Fibre Channel ports (IQNs) of the IBM Storwize V5030 system by assigning them to the group that was created in step 1.

Verify that the IBM Storwize V5030 discovered the LUNs as unmanaged MDisks. If the external storage does not show up automatically, click **Discover storage** from the Actions menu on the External Storage window, as shown in Figure 11-4.

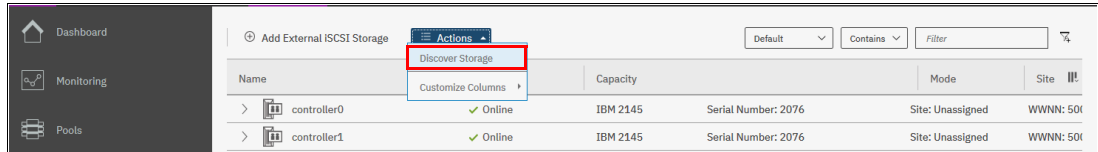


Figure 11-4 Discover storage

You can also start storage discovery by using the CLI command `detectmdisk`. It returns no output. Although it might appear that the `detectmdisk` command completed, some extra time might be required for it to run. It is asynchronous and returns a prompt while the command continues to run in the background.

Add the MDisks to an existing pool or create a pool to include them. If the storage pool does not exist, follow the procedure that is described in Chapter 4, “Storage pools” on page 159.

Figure 11-5 shows how to add selected MDisk to a storage pool. Click **Assign** under the Actions menu or right-click the MDisk and select **Assign** to add MDisk to a storage pool. If the pool does not exist, see Chapter 4, “Storage pools” on page 159 for more information about how to create it.

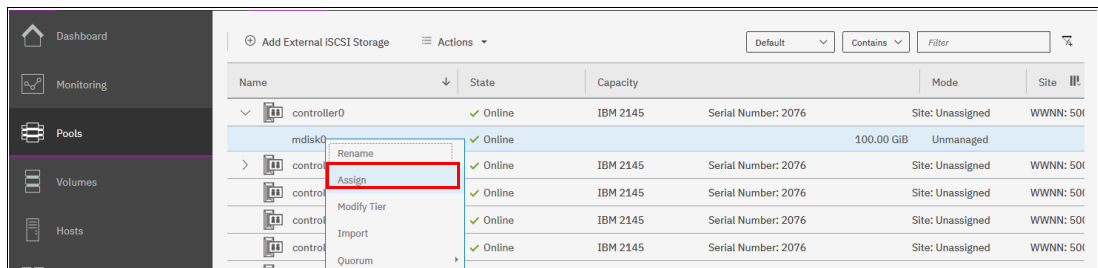


Figure 11-5 Assigning MDisk to a storage pool

After you click **Assign**, a window opens, as shown in Figure 11-6. Select **target pool**, **MDisk storage tier**, and **external encryption** settings.

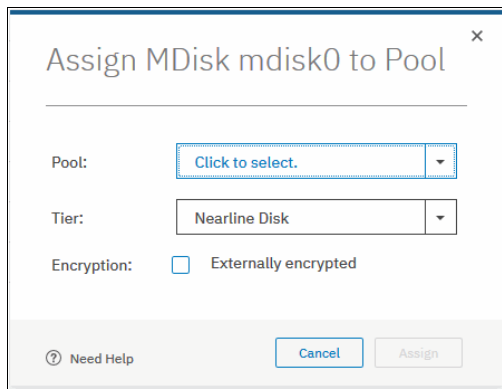


Figure 11-6 Assign MDisk window

When adding MDisks to pools, it is necessary to assign them to the correct storage tiers. It is important to set the tiers correctly if you plan to use the Easy Tier feature because the wrong tier can cause improper EasyTier operations and affect system performance. For more information about storage tiers, see Chapter 9, “Advanced features for storage efficiency” on page 435.

The storage tier setting can be also changed after MDisk is assigned to the pool.

Select the **Externally encrypted** option if backend storage performs data encryption. For more information about IBM Storwize V5030 encryption, see Chapter 13, “Encryption” on page 725.

After the task completes, click **Close**.

Important: If the external storage volumes to virtualize behind the Storwize V5030 contain data and this data must be retained, do *not* use the Assign to pool option to manage the MDisks.

This option destroys the data on the LU.

Instead, use the Import option. For more information, see 11.2.4, “Importing image mode volumes” on page 615.

The external MDisks that are assigned to a pool within IBM Storwize V5030 are displayed by clicking **Pools** → **MDisks by Pools**, as shown in Figure 11-7. Create volumes from the storage pool and map them to hosts, as needed. For more information about how to create and map volumes to hosts, see Chapter 6, “Volume configuration” on page 309.

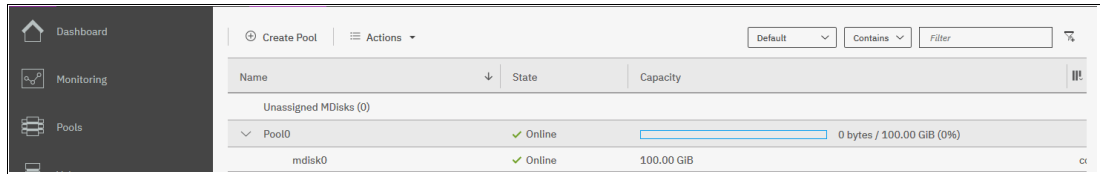


Figure 11-7 MDisks by Pools pane

11.2.4 Importing image mode volumes

If the external storage systems are not new systems and data exists on the LUNs that must be kept after virtualization, you must import the existing LUNs. The process of importing data on external volumes is simplified by using the storage migration wizard, which is described in Chapter 7, “Storage migration” on page 357.

To manually import volumes, they must not be assigned to a storage pool and must remain in Unmanaged state. Hosts that access data from these external storage system LUNs can continue to access data, but they must be rezoned and mapped to the Storwize V5030 to use these external storage system LUNs after they are presented through the IBM Storwize V5030.

Complete the following steps to import an unmanaged MDisk:

1. Select the unmanaged MDisk and click **Import** from the Actions drop-down menu (see Figure 11-8). Multiple MDisks can be selected by using the Ctrl key.

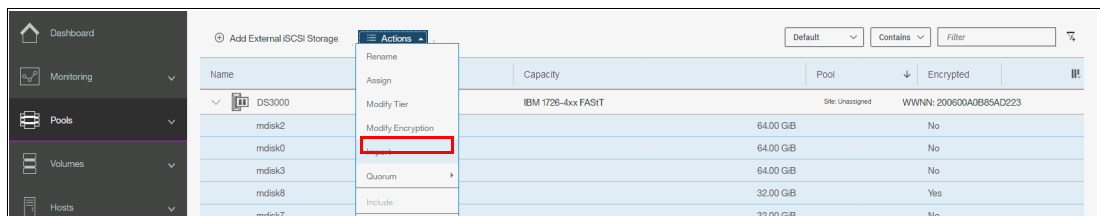


Figure 11-8 Import MDisk option

2. Selecting the Import option opens a new window that requires more volume information, as shown in Figure 11-9.

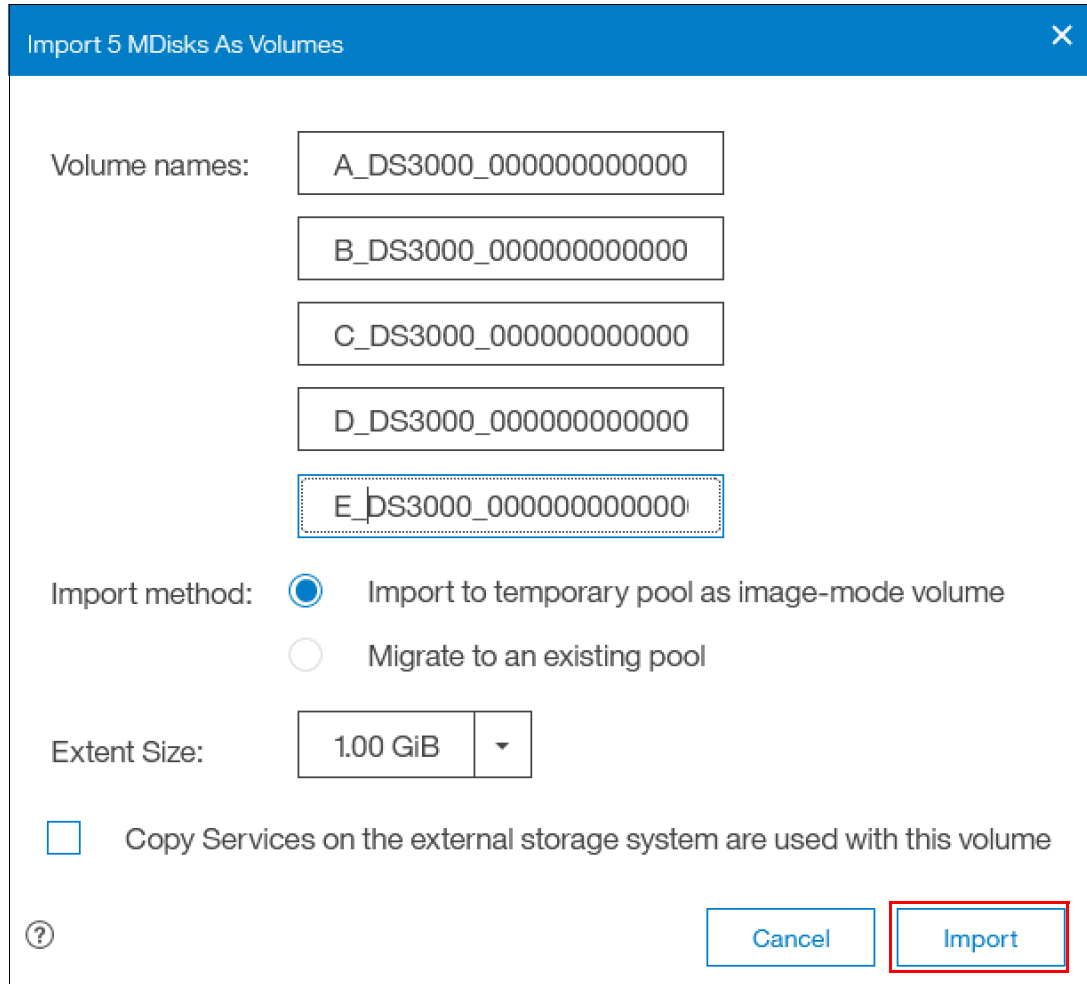


Figure 11-9 Import MDisks as Volumes window

Consider the following points:

- You can change the default volume names to more meaningful names by editing the Volume names text boxes.
- You can choose between importing the volume to a temporary pool as an *image mode volume*, which the IBM Storwize V5030 can create and name for you, or migrating the volume to an existing pool.

An image mode volume has a direct block-for-block translation from the imported MDisk and the external LUN. Therefore, the data is preserved. In this state, the IBM Storwize V5030 is acting as a proxy and the image mode volume is a “pointer” to the external LUN.

Because of the way that virtualization works on the IBM Storwize V5030, the external LUN is presented as an MDisk, but we cannot map an MDisk directly to a host. Therefore, the IBM Spectrum Virtualize software must create the image mode volume to allow hosts to perform the mapping through the Storwize V5030.

3. If you choose a temporary pool, you must first select the extent size for the pool. The default value for extents is 1 GB. If you plan to migrate this volume to another pool later, ensure that the extent size matches the extent size of the prospective target pool. For more information about extent sizes, see Chapter 4, “Storage pools” on page 159.

If an existing storage pool is chosen, the Storwize V5030 can perform a migration task. The external LUN can be imported into a temporary migration pool and a migration task can run in the background to copy data to MDisks that are in the target storage pool. At the end of the migration, the external LUN and its associated MDisk can be in the temporary pool and show as managed, but they can be removed from the Storwize V5030.

Figure 11-10 shows how to select an existing pool for volume migration. The pool must have enough available capacity to store the volumes that are being imported.

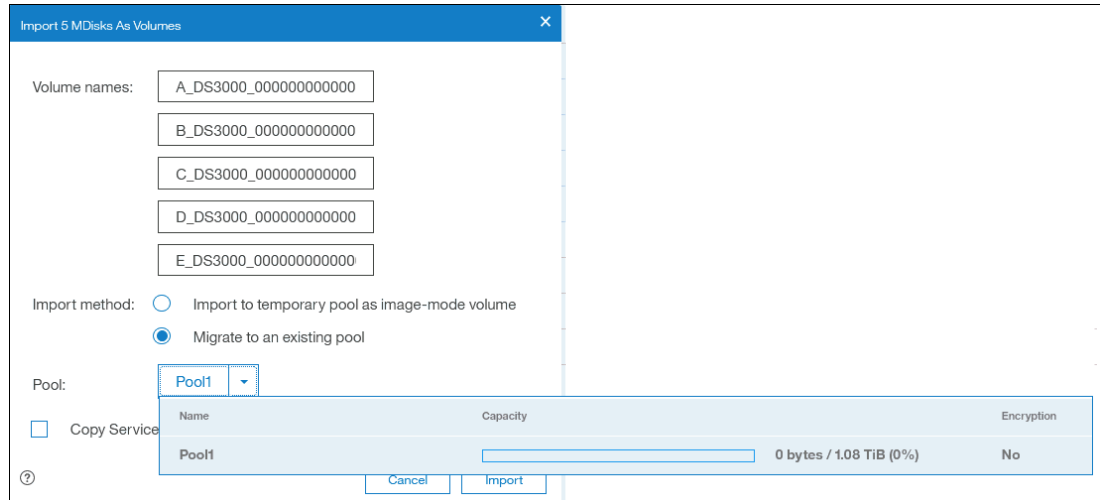


Figure 11-10 Import volumes into an existing pool

4. Select **Copy Services** if copy services (replication functionality) are used on the external storage system that hosts the LUN. Click **Import** to confirm your selections and to start the import process.

Note: Only pools with sufficient capacity are shown because the import of an MDisk to a storage pool can migrate storage. This storage can be migrated only if sufficient capacity exists in the target pool to create a copy of the data on its own MDisks.

A migration task starts and can be tracked through the **System Migration** window within the **Pools** menu, as shown in Figure 11-11. The data migration begins after the MDisk is imported successfully.

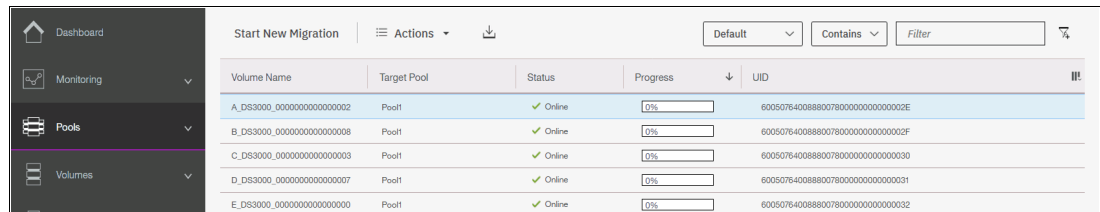


Figure 11-11 Checking the migration status

When the migration completes, the migration status disappears and the volume is displayed in the target pool, as shown in Figure 11-12.

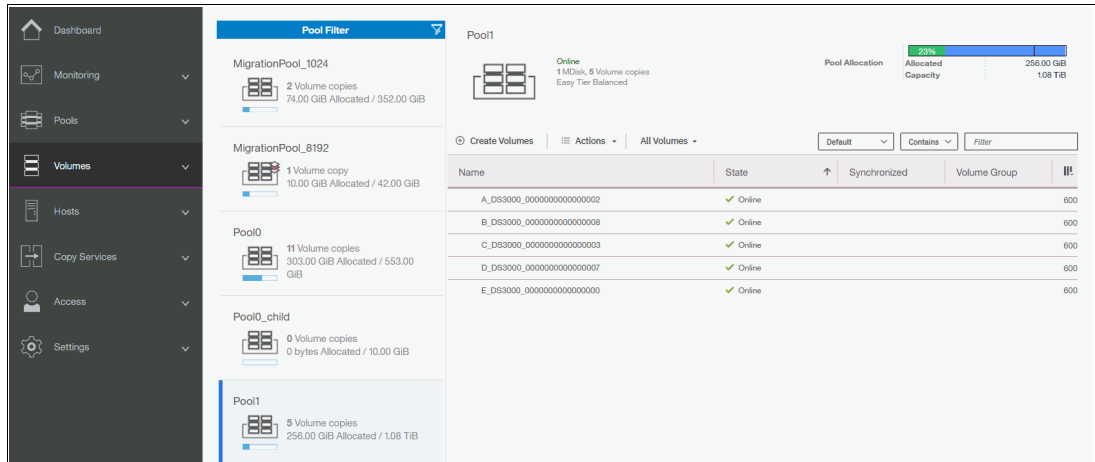


Figure 11-12 Volume is assigned to the target pool

- After the migration completes, the image mode volume is automatically deleted, but the external LUN exists as a managed MDisk in the temporary storage pool. It is unassigned from the pool and listed as an unassigned MDisk. Later, you can retire the external LUN and remove it completely from the Storwize V5030 by unmapping the volume at the external storage and by clicking **Discover Storage** on the Storwize V5030. For more information about removing external storage, see 11.2.6, “Removing external storage” on page 621.

If you choose to import a volume as an image mode volume, the external LUN appears as an MDisk with an associated image mode volume name and can be listed, as shown in Figure 11-13.

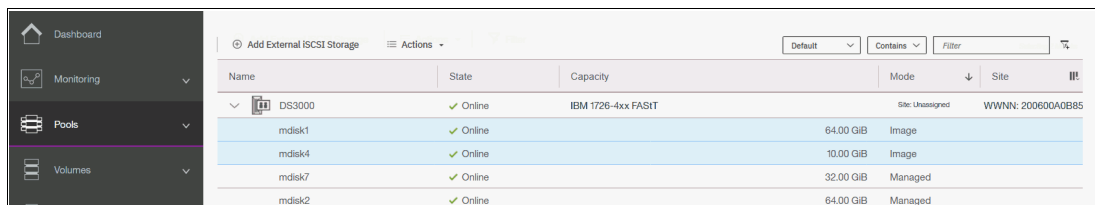


Figure 11-13 Image mode volumes

The volume is also listed in the **System Migration** window because the IBM Storwize V5030 expects you to migrate these volumes later, as shown in Figure 11-14.

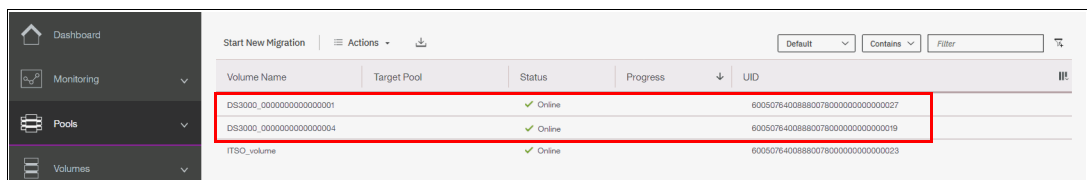


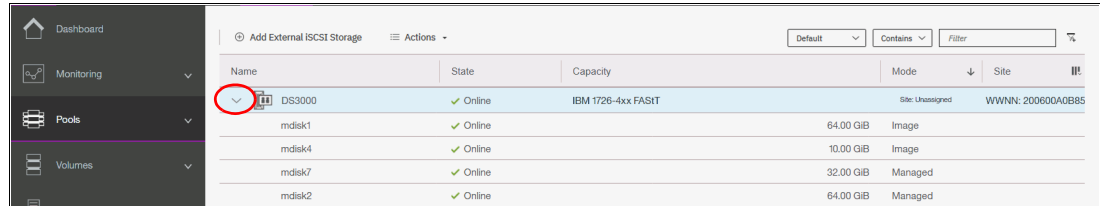
Figure 11-14 Migrations are available for image mode volumes

At the end of this process, the volume can be mapped to a host.

11.2.5 Managing external storage controllers

To perform actions on back-end storage controllers on IBM Storwize V5030, click **Pools** → **External Storage** menu. Complete the following steps:

1. Clicking the arrow sign that precedes each of the external storage controllers (see Figure 11-15) provides a list of the MDisks that are mapped from it.



Name	State	Capacity	Mode	Site
DS3000	Online	IBM 1726-4xx FASIT	Site: Unassigned	WWNN: 200600A0B85
mdisk1	Online		64.00 GiB	Image
mdisk4	Online		10.00 GiB	Image
mdisk7	Online		32.00 GiB	Managed
mdisk2	Online		64.00 GiB	Managed

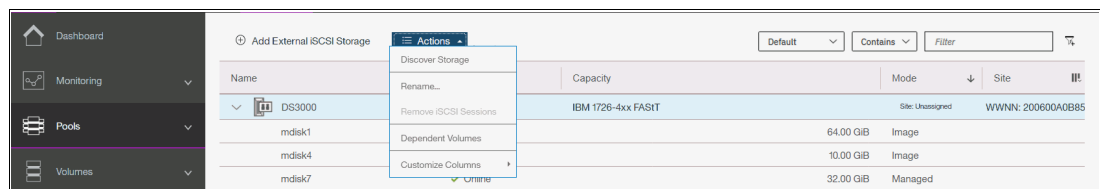
Figure 11-15 External Storage window

To list available controllers, use the `lscontroller` command, as shown in Example 11-2 on page 613. To list the controller's resources with the CLI, run the `lsmdisk` command. If given without parameters, the command displays a list of all MDisks available on the system. Output can be filtered to display MDisk objects that belong only to a single controller, as shown in Example 11-3.

Example 11-3 `lsmdisk` output (some columns are not shown)

```
IBM_Storwize:ITS0V5030:superuser>lsmdisk -filtervalue controller_name=DS3000
id name  status mode  mdisk_grp_id mdisk_grp_name  capacity
1  mdisk1 online image  5           MigrationPool_1024 64.0GB
2  mdisk4 online image  6           MigrationPool_8192 10.0GB
3  mdisk7 online managed 2           Pool1           32.0GB
4  mdisk2 online managed 2           Pool1           64.0GB
<...>
```

2. In the External Storage window, options are available in the Actions menu that can be applied to external storage controllers, as shown in Figure 11-16. Select the external controller and click **Actions** to display the available options. Alternatively, right-click the external controller.



Name	State	Capacity	Mode	Site
DS3000	Online	IBM 1726-4xx FASIT	Site: Unassigned	WWNN: 200600A0B85
mdisk1	Online		64.00 GiB	Image
mdisk4	Online		10.00 GiB	Image
mdisk7	Online		32.00 GiB	Managed

Figure 11-16 External controllers options under Actions menu

3. You can change the name of any external storage system by right-clicking the controller and selecting **Rename**. Alternatively, use the Actions drop-down menu and select **Rename**.

With CLI, a controller can be renamed by using the `chcontroller` command with `-name` and current controller name or ID as parameters.

- Click **Show Dependent Volumes** to display the logical volumes that depend on the status of selected external storage system, as shown in Figure 11-17. If a controller goes offline, the dependent volumes also go offline. Before you take a controller offline for maintenance, you can use the command to ensure that you do not lose access to any volumes.

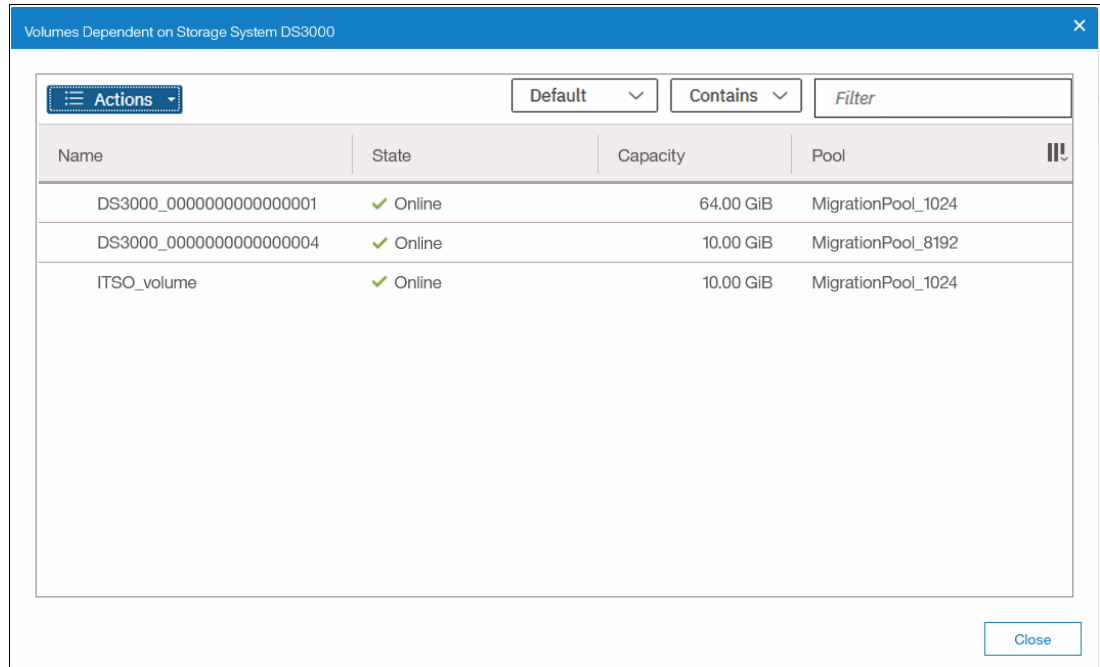


Figure 11-17 Volumes that depend on the external storage

Use the `lscntrollerdependentvdisks` command with the controller name or ID as a parameter to get more information with the CLI.

- You can right-click any volume in the list to see actions that can be performed on it. For more information about the volume actions of the IBM Storwize V5030 storage system, see Chapter 8, “Advanced host and volume administration” on page 383.
- In the **External Storage** window, you can also right-click an **MDisk** (or use the Actions drop-down menu) to display the available options for a selected MDisk, as shown in Figure 11-18.

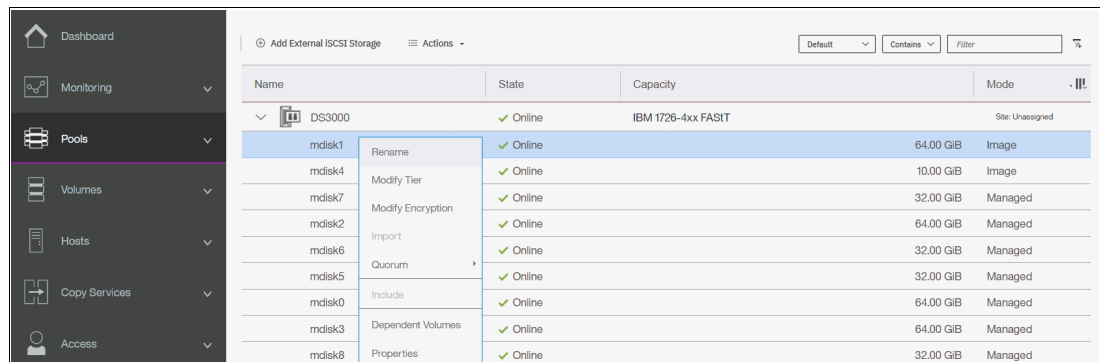


Figure 11-18 MDisk Actions menu in the External Storage window

Run the CLI command `chmdisk` to change MDisk name, tier, or external encryption status, as shown in Example 11-4.

Example 11-4 chmdisk usage

```
IBM_Storwize:ITSOV5030:superuser>chmdisk -name MigratedMdisk mdisk0
IBM_Storwize:ITSOV5030:superuser>chmdisk -tier tier_enterprise mdisk1
IBM_Storwize:ITSOV5030:superuser>chmdisk -encrypt yes mdisk2
```

11.2.6 Removing external storage

If you want to remove the external storage systems from the IBM Storwize V5030 virtualized environment, the following options are available:

- ▶ To remove the external storage systems and discard the data on them, complete the following steps:
 - a. Stop any host I/O on the volumes.
 - b. Remove the volumes from the host file systems, logical volume, or volume group and remove the volumes from the host device inventory.
 - c. Remove the host mapping of volumes and the volumes themselves on the IBM Storwize V5030.
 - d. Remove the storage pools to which the external storage systems belong, or you can keep the storage pool and remove the MDisks of the external storage from the storage pools.
 - e. Unzone and disconnect the external storage systems from the IBM Storwize V5030.
 - f. Click **Actions** → **Discover Storage** on **Pools** → **External Storage** or **Pools** → **MDisks by Pools** pane for the IBM Storwize V5030 to discover the removal of the external storage systems. You can also use the `detecmdisk` CLI command, as described in 11.2.3, “Working with MDisks” on page 613.
- ▶ To remove the external storage systems and keep the volumes and their data on the IBM Storwize V5030, complete the following steps:
 - a. Migrate volumes and their data to other internal or external storage pools that are on the IBM Storwize V5030.
 - b. Remove the storage pools to which the external storage systems belong, or you can keep the storage pools and remove the MDisks of the external storage from the storage pools.
 - c. Unzone and disconnect the external storage systems from the IBM Storwize V5030.
 - d. Click **Actions** → **Discover Storage** on **Pools** → **External Storage** or **Pools** → **MDisks by Pools**, or run the `detecmdisk` CLI command for the IBM Storwize V5030 to rediscover storage.
- ▶ To remove the external storage systems from the IBM Storwize V5030 control and keep the volumes and their data on other external storage systems, complete the following steps:
 - a. Migrate volumes and their data to other internal or external storage pools on the IBM Storwize V5030, as described in Chapter 7, “Storage migration” on page 357.
 - b. Remove the storage pools to which the original external storage systems belong, or you can keep the storage pools and remove the MDisks of that external storage from the storage pools.

- c. Export the volumes that were migrated in step a to image mode with the new MDisks on the target external storage systems. For more information about the restrictions and prerequisites for migration, see Chapter 7, “Storage migration” on page 357.

You must record pre-migration information; for example, the original Small Computer System Interface (SCSI) identifiers (IDs) that the volumes used when they were mapped to hosts. Certain operating systems do not support a change of the SCSI ID during migration. Unzone and disconnect the external storage systems from the IBM Storwize V5030.

- d. Click **Actions** → **Discover Storage** on **Pools** → **External Storage** or **Pools** → **MDisks by Pools**, or run `detecmdisk` CLI command for the IBM Storwize V5030 to rediscover storage.



RAS, monitoring, and troubleshooting

This chapter describes the reliability, availability, and serviceability (RAS) features and ways to monitor and troubleshoot the IBM Storwize V5000 Gen2.

This chapter includes the following topics:

- ▶ 12.1, “Reliability, availability, and serviceability features” on page 624
- ▶ 12.2, “System components” on page 625
- ▶ 12.3, “Configuration backup” on page 642
- ▶ 12.4, “System update” on page 647
- ▶ 12.5, “Monitoring” on page 668
- ▶ 12.6, “Audit log” on page 679
- ▶ 12.7, “Event log” on page 681
- ▶ 12.8, “Support assistance” on page 691
- ▶ 12.9, “Collecting support information” on page 705
- ▶ 12.10, “Powering off the system and shutting down the infrastructure” on page 719

12.1 Reliability, availability, and serviceability features

This section describes the reliability, availability, and serviceability (RAS) features of the IBM Storwize V5000 Gen2, and monitoring and troubleshooting. RAS features are important concepts in the design of the IBM Storwize V5000 Gen2. Hardware and software features, design considerations, and operational guidelines all contribute to make the IBM Storwize V5000 Gen2 reliable.

Fault tolerance and a high level of availability are achieved with the following features:

- ▶ The RAID capabilities of the underlying disk subsystems
- ▶ The software architecture that is used by the IBM Storwize V5000 Gen2 nodes
- ▶ Auto-restart of nodes that are stopped
- ▶ Battery units to provide cache memory protection in a site power failure
- ▶ Host system multipathing and failover support

High levels of serviceability are achieved with the following features:

- ▶ Cluster error logging
- ▶ Asynchronous error notification
- ▶ Dump capabilities to capture software-detected failures
- ▶ Concurrent diagnostic procedures
- ▶ Directed maintenance procedures
- ▶ Concurrent log analysis and memory dump data recovery tools
- ▶ Concurrent maintenance of all of the IBM Storwize V5000 Gen2 components
- ▶ Concurrent upgrade of IBM Storwize V5000 Gen2 software and microcode of drives
- ▶ Concurrent addition or deletion of a node canister in a cluster
- ▶ Software recovery through the Service Assistant Tool
- ▶ Automatic software version correction when a node is replaced
- ▶ Detailed status and error conditions that are displayed through the Service Assistant Tool
- ▶ Error and event notification through Simple Network Management Protocol (SNMP), syslog, and email
- ▶ Access to the Service Assistant Tool through the tech port for network connection problems
- ▶ Remote support personnel is able to access the system to complete troubleshooting and maintenance tasks

At the core of the IBM Storwize V5000 Gen2 is a redundant pair of *node canisters*. The two canisters share the load of transmitting and receiving data between the attached hosts and the disk arrays.

12.2 System components

This section describes each of the components that make up the IBM Storwize V5000 Gen2 system. The components are described in terms of location, function, and serviceability.

12.2.1 Enclosure midplane

The *enclosure midplane* connects the node or expansion canisters to the power supply units and to the drives. The midplane is part of the enclosure midplane assembly, which consists of the midplane and the front section of the enclosure.

During the basic system configuration, vital product data (VPD) is written to the enclosure midplane. On a control enclosure midplane, the VPD contains information, such as worldwide node name (WWNN) 1, WWNN 2, machine type and model, machine part number, and serial number. On an expansion enclosure midplane, the VPD contains information, such as machine type and model, machine part number, and serial number.

The enclosure midplane is initially generic and it is configured as a control enclosure midplane or expansion enclosure midplane only when the VPD is written. After the VPD is written, a control enclosure midplane is no longer interchangeable with an expansion enclosure midplane and vice versa.

Important: The enclosure midplane must be replaced by a trained IBM service support representative (SSR) only.

For more information about the midplane replacement process, see [IBM Storwize V5000 Gen2 Knowledge Center](#).

For a complete overview of maintenance tasks, see [IBM Knowledge Center](#).

12.2.2 Node canisters

Two node canister slots are on the top of the unit. The left slot is canister 1, and the right slot is canister 2.

Figure 12-1 shows the rear view of a fully equipped control enclosure.

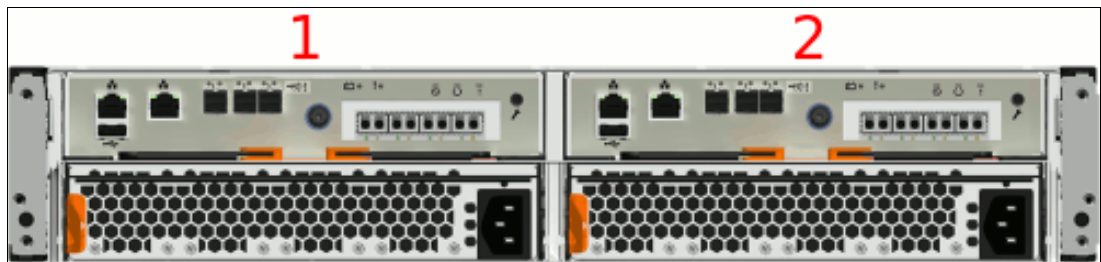


Figure 12-1 Rear view of a control enclosure with two node canisters (the Storwize V5020)

USB ports

Each node canister has one USB port. The location of the port is the same on every model, and no indicators are associated with it.

Figure 12-2 shows the location of the USB port.

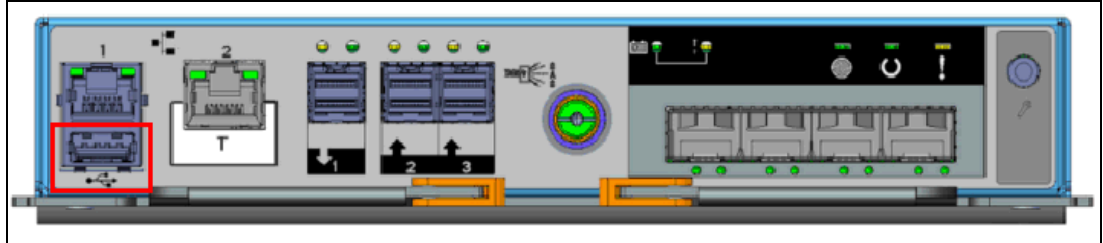


Figure 12-2 Node canister USB port (the Storwize V5020)

The USB flash drive is not required to initialize the system configuration. However, it can be used for other functions. Using the USB flash drive is required in the following situations:

- ▶ When you cannot connect to a node canister in a control enclosure by using the service assistant or the technician port, and you want to see the status of the node or re-enable the technician port.
- ▶ When you do not know, or cannot use, the service IP address for the node canister in the control enclosure and must set the address.
- ▶ When you forget the superuser password and must reset the password.

Ethernet ports

The Storwize V5010 and Storwize V5020 node canisters have two 100/1000 Mbps Ethernet ports. Both ports can be used for management, internet Small Computer System Interface (iSCSI) traffic, and Internet Protocol (IP) replication. Additionally, port 2 can be used as a technician port (the white box with “T” in the center of the box) for system initialization and servicing. After initialization, the technician port is disabled. It is possible to reactivate the technician port later again by using CLI commands.

Figure 12-3 shows the Ethernet ports on the Storwize V5010.

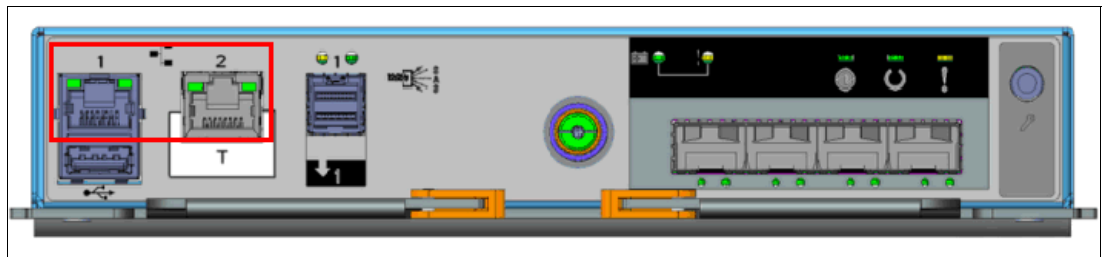


Figure 12-3 Storwize V5010 Ethernet ports

Figure 12-4 shows the Ethernet ports on the Storwize V5020.

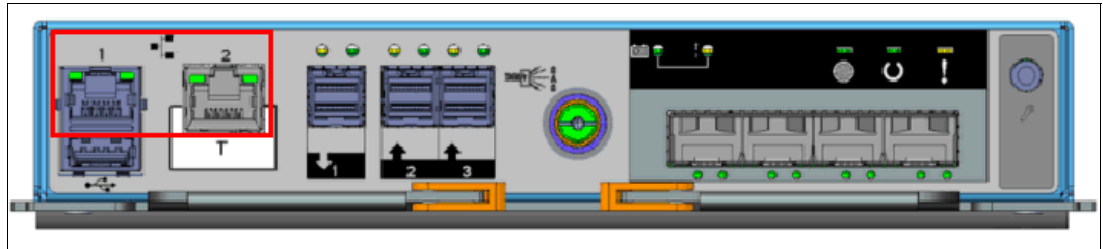


Figure 12-4 Storwize V5020 Ethernet ports

Each Storwize V5030 node canister has two 1/10 Gbps Ethernet ports and one Ethernet technician port. Port 1 and 2 can be used for management, iSCSI traffic, and IP replication. Port T can be used as a technician port for system initialization and service only.

Figure 12-5 shows the Ethernet ports on the Storwize V5030.

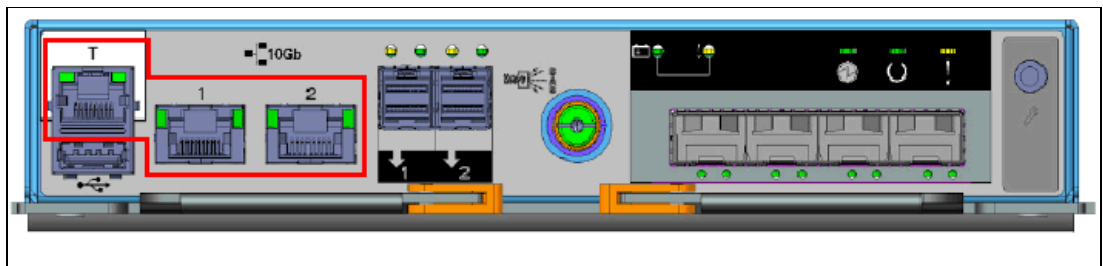


Figure 12-5 Storwize V5030 Ethernet ports

Each port has two LEDs that display the status of its activity. Their meanings are listed in Table 12-1.

Table 12-1 Ethernet port status LEDs

Name and position	Color	State	Meaning
Activity (left)	Green	Flashing	The link is active.
		Off	The link is inactive.
Link speed (right)	Green	Solid	A connection exists to a remote device at 1 Gbps or more.
		Off	No connection exists to a remote device, or the link is connected at less than 1 Gbps.

Serial-attached SCSI ports

Each Storwize V5010 node canister uses one 12 Gbps serial-attached SCSI (SAS) port to connect optional expansion enclosures. This port does not support host attachment.

Figure 12-6 shows the SAS ports on the Storwize V5010.

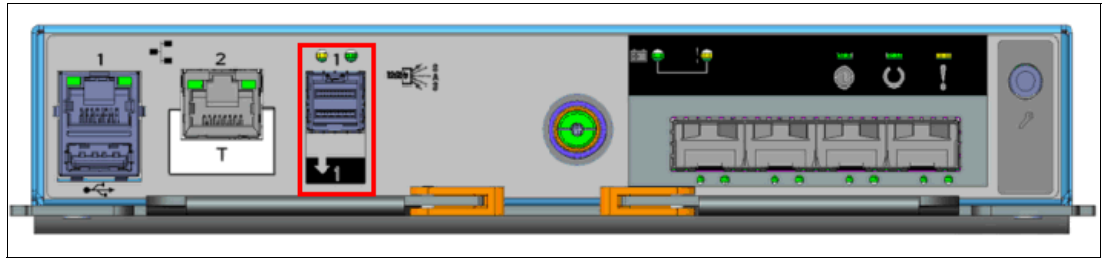


Figure 12-6 Storwize V5010 SAS ports

Each Storwize V5020 node canister has three 12 Gbps SAS ports. Port 1 can be used to connect optional expansion enclosures, and ports 2 and 3 can be used for host attachment.

Figure 12-7 shows the SAS ports on the Storwize V5020.

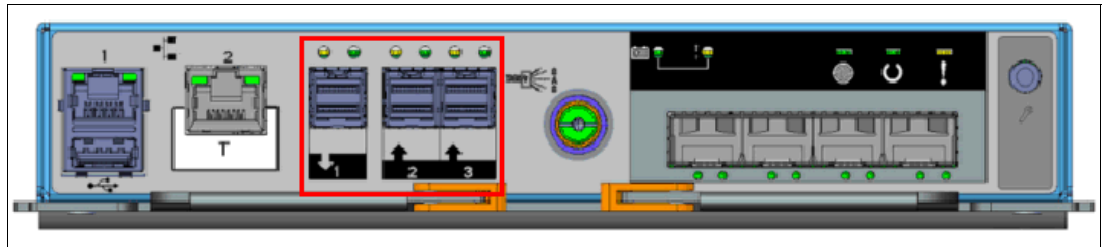


Figure 12-7 Storwize V5020 SAS ports

Each Storwize V5030 node canister has two 12 Gbps SAS ports to connect optional expansion enclosures. This port does not support host attachment.

Figure 12-8 shows the SAS ports on the Storwize V5030.

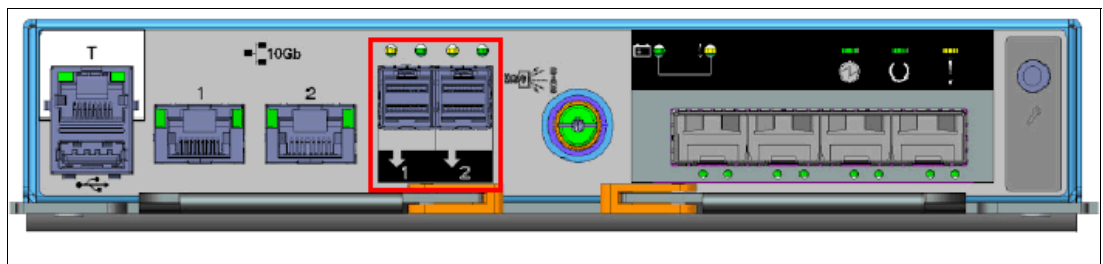


Figure 12-8 Storwize 5030 SAS ports

Each port has two LEDs that display the status of its activity. Their meanings are listed in Table 12-2.

Table 12-2 SAS port status LEDs

Name and position	Color	State	Meaning
Fault (left)	Amber	Solid	One of the following conditions has occurred: <ul style="list-style-type: none"> ▶ One or more, but not all, of the four lanes are up. (If no lanes are up. The activity light is off.) ▶ One or more of the lanes is running at a different speed to the others. ▶ One or more of the up lanes are attached to a different address to the others. ▶ An unsupported device is plugged into this SAS port.
		Off	No fault exists. All four lanes (phys) have a connection.
Link (right)	Green	Solid	A connection exists on at least one lane (phy).
		Off	None of the SAS connections are working.

Battery status

Each node canister houses a battery, the status of which is displayed by two LEDs on the back of the unit, as shown in Figure 12-9.

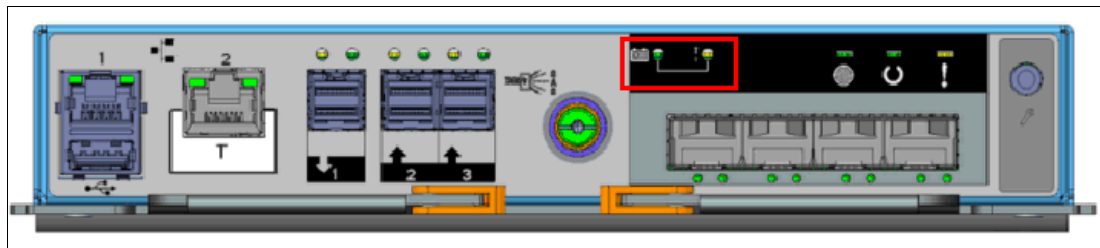


Figure 12-9 Battery status LEDs (the Storwize V5020)

The meaning of each LED is described in Table 12-3.

Table 12-3 Battery status LEDs

Name and position	Color	State and Meaning
Battery status (left)	Green	<ul style="list-style-type: none"> ▶ Fast flash: The battery is charging. It does not have a sufficient charge to perform a “fire hose” dump. Power-on will not be possible. ▶ Flash: The battery has sufficient charge to perform one “fire hose” dump. ▶ On: The battery is fully charged and has sufficient charge to perform two “fire hose” dumps. ▶ Off: The battery is not available for use.
Fault (right)	Amber	<ul style="list-style-type: none"> ▶ Off: No known conditions are preventing normal operation, unless the battery status LED is also on. ▶ On: An active condition or fault might compromise normal operation. ▶ Slow flash: A non-critical fault occurred with the battery.
Battery in use	Green	<ul style="list-style-type: none"> ▶ Off: The battery is not being used to power the canister. ▶ Fast flash: The battery is providing power for a “fire hose” dump.

Canister status

The status of each canister is displayed by three LEDs on the back of the unit, as shown in Figure 12-10.

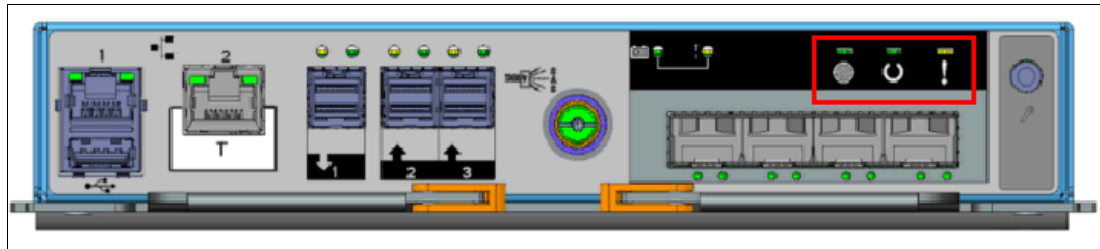


Figure 12-10 Node canister status LEDs (the Storwize V5020)

The meaning of each LED is described in Table 12-4.

Table 12-4 Canister status LEDs

Name and position	Color	State and Meaning
Power (left)	Green	<ul style="list-style-type: none"> ▶ Off: No power is available or power is coming from the battery. ▶ Slow flash: Power is available but the main CPU is not running; the system is in standby mode. ▶ Fast flash: System is in self-test. ▶ On: Power is available and the system code is running.
Status (middle)	Green	<ul style="list-style-type: none"> ▶ Off: Indicates one of the following conditions: <ul style="list-style-type: none"> – No power to the canister – Canister is in standby mode or self-test – Operating system is loading ▶ Flash: The canister is in candidate or service state. It is not performing I/O. It is safe to remove the node. ▶ Flash fast: The canister is performing a fire hose dump. ▶ On: The canister is active, can perform I/O, or starting. The system is part of a cluster.
Canister Fault (right)	Amber	<ul style="list-style-type: none"> ▶ Off: The node is in candidate or active state. Any error that was detected is not severe enough to stop the node participating in a cluster or performing I/O. ▶ Flash: The canister is being identified. A fault condition might exist. ▶ On: The node is in service state or an error exists that might be stopping the system code from starting (node error 550). The node canister cannot become active in the system until the problem is resolved. The problem is not necessarily related to a hardware component.

Replaceable components

The IBM Storwize V5000 Gen2 node canister contains the following field-replaceable (client-replaceable) components:

- ▶ Host Interface Card
- ▶ Memory
- ▶ Battery

Note: The procedure for replacing Host Interface Cards also applies to upgrading controllers.

Figure 12-11 shows the location of these parts within the node canister.

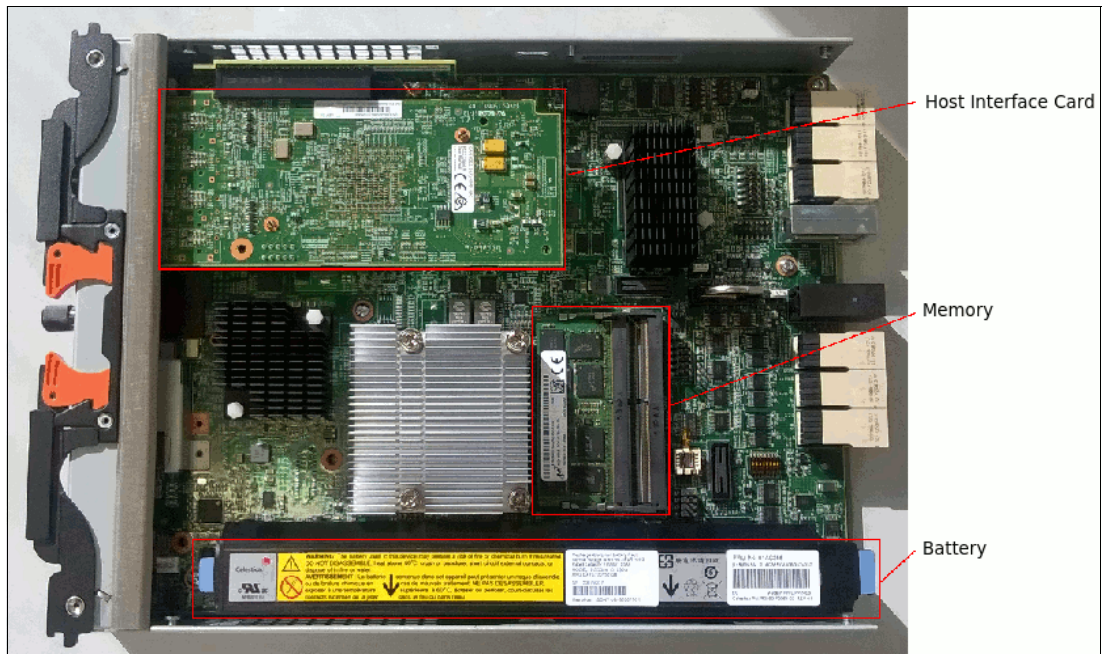


Figure 12-11 Node canister client-replaceable components

Note: Because these components are inside the node canister, their replacement leads to a redundancy loss until the replacement is complete.

Host Interface Card replacement procedure

For information about the Host Interface Card (HIC) replacement process, see [IBM Storwize V5000 Gen2 Knowledge Center](#).

Figure 12-12 shows an HIC replacement.

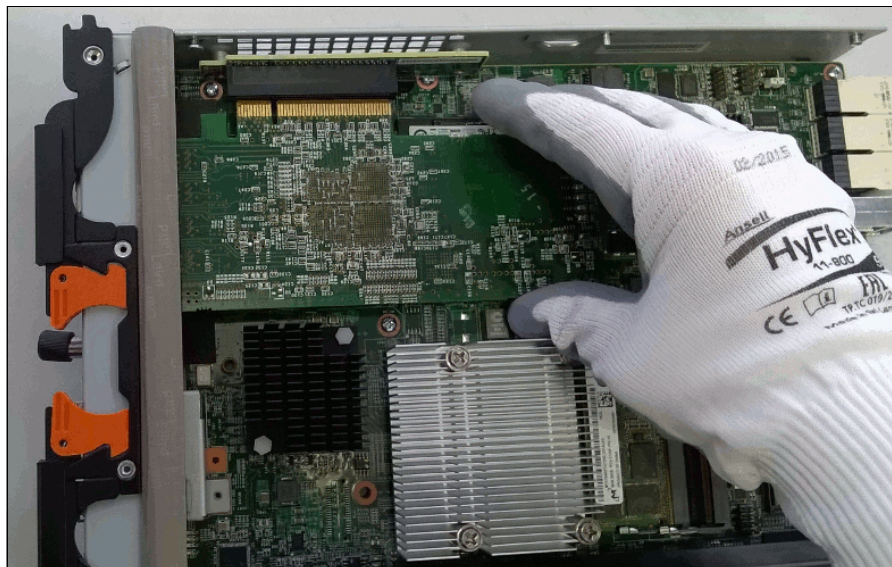


Figure 12-12 HIC replacement

Memory replacement procedure

For more information about the memory replacement process, see [IBM Storwize V5000 Gen2 Knowledge Center](#).

Figure 12-13 shows the location of the memory modules.

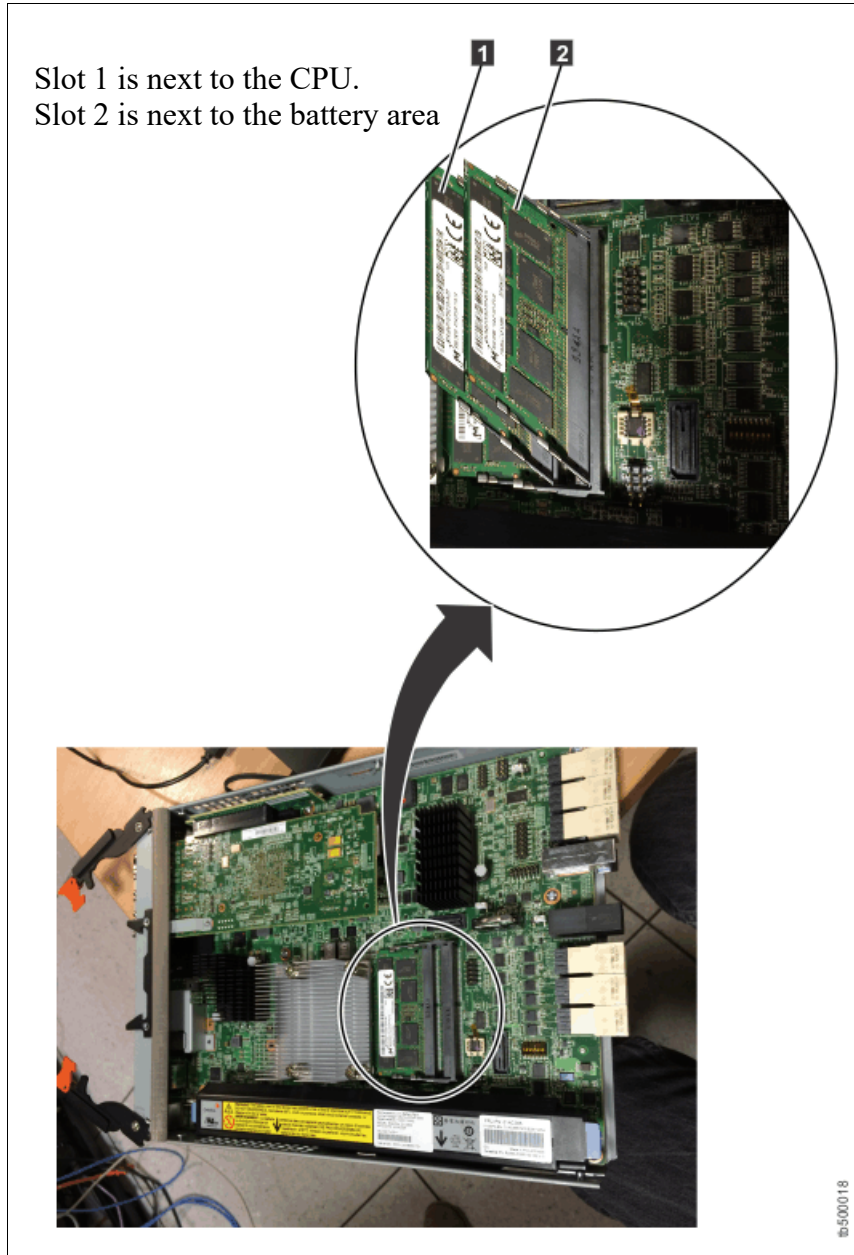


Figure 12-13 Location of memory modules

Figure 12-14 shows a memory replacement.

Note: The memory modules do not stand up. They lie in a cascading fashion.

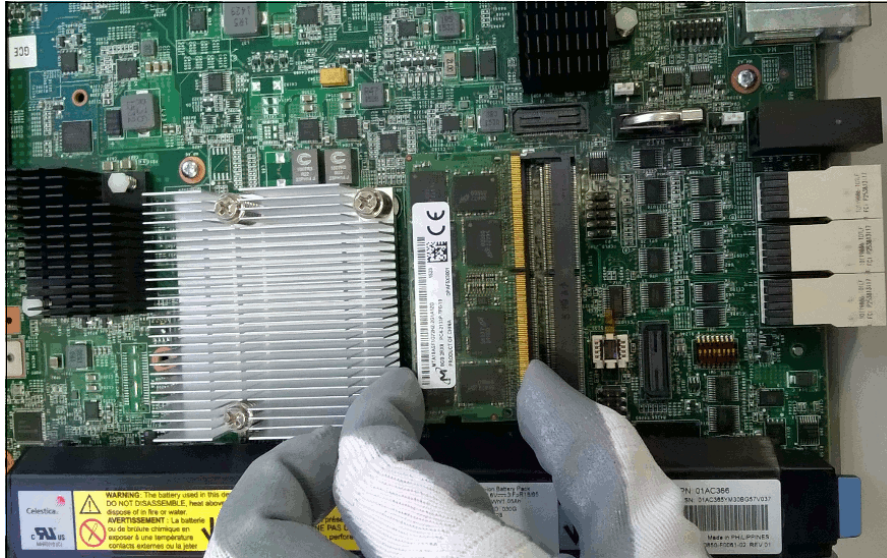


Figure 12-14 Memory replacement

Battery Backup Unit replacement procedure

Attention: The battery is a lithium ion battery. To avoid a possible explosion, do not incinerate the battery. Exchange the battery only with the part that is approved by IBM.

Because the Battery Backup Unit (BBU) replacement leads to a redundancy loss until the replacement is complete, we advise that you replace the BBU only when you are instructed to replace it. We advise you to follow the Directed Maintenance Procedure (DMP).

While you lift and lower the battery during the procedure, grasp the blue handle on each end of the battery and keep the battery parallel to the canister system board, as shown in Figure 12-15.



Figure 12-15 BBU replacement

Important: During the replacement, the battery must be kept parallel to the canister system board while the battery is removed or replaced. Keep equal force, or pressure, on each end.

For more information about the BBU replacement process, see [IBM Knowledge Center](#).

For more information about replacement procedures, see [this website](#).

12.2.3 Expansion canisters

Two expansion canister slots are on the top of the unit. As with the control enclosure, the left slot is canister 1 and the right slot is canister 2.

Figure 12-16 shows the rear view of a fully equipped expansion enclosure.

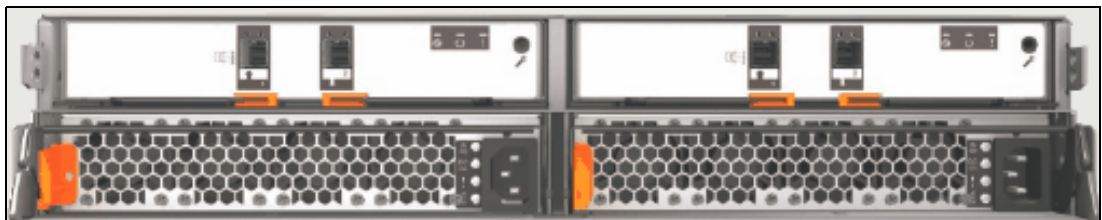


Figure 12-16 Rear view of an expansion enclosure with two expansion canisters

SAS ports

SAS ports are used to connect the expansion canister to the node canister or to an extra expansion canister in the chain. Figure 12-17 shows the SAS ports that are on the expansion canister.

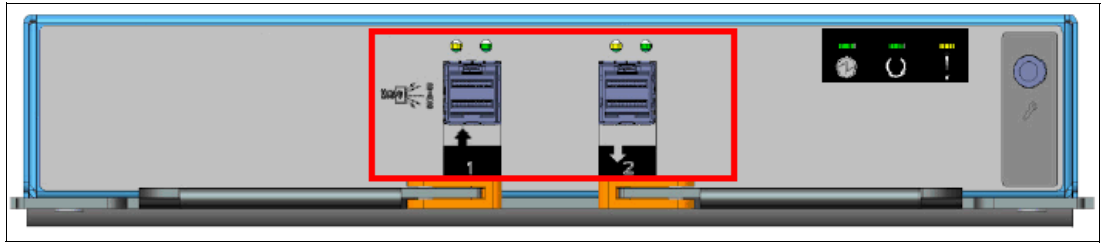


Figure 12-17 Expansion canister SAS ports

Each port has two LEDs that display the status of its activity. Their meanings are shown in Table 12-5.

Table 12-5 SAS port status LEDs

Name and position	Color	State	Meaning
Fault (left)	Amber	Solid	One of the following errors exists: <ul style="list-style-type: none"> ▶ Only 1, 2, or 3 lanes (phys) have a connection. ▶ Not all of the lanes (phys) that have a connection are running at the same speed. ▶ Not all of the lanes (phys) that have a connection are attached to the same address. ▶ An unsupported device is connected to the port.
		Off	No fault exists. All four lanes (phys) have a connection.
Link (right)	Green	Solid	A connection exists on at least one lane (phy).
		Off	No connection exists on any lane (phy).

Canister status

The status of each expansion canister is displayed by three LEDs on the back of the unit, as shown in Figure 12-18.

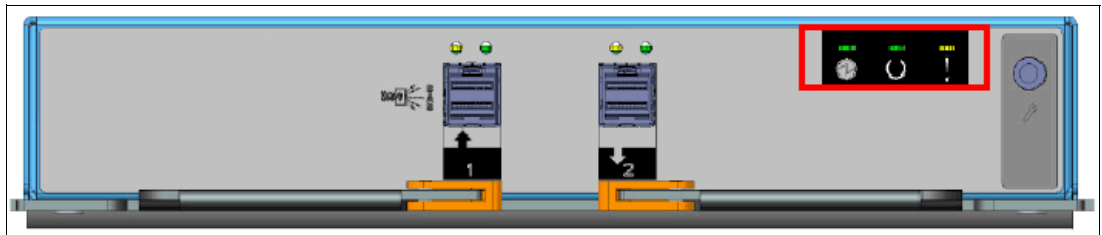


Figure 12-18 Enclosure canister status LEDs

The meaning of each LED is listed in Table 12-6.

Table 12-6 Expansion canister status LEDs

Name and position	Color	State	Meaning
Power (left)	Green	Solid	The canister is receiving power.
		Off	No power is available, or the power is coming from the battery.

Name and position	Color	State	Meaning
Status (middle)	Green	Solid	The canister is running normally.
		Flashing	The canister is unable to read data from the midplane.
		Off	The system is off, in standby, or running a self-test, or the operating system is loading.
Fault (right)	Amber	Solid	A fault requires part replacement, or the canister is still starting.
		Flashing	The canister is being identified. A fault might or might not exist.
		Off	The canister has no faults that require part replacement.

12.2.4 Disk subsystem

This section describes the parts of the IBM Storwize V5000 Gen2 disk subsystem, which is made up of control and expansion enclosures.

The Storwize V5010 and Storwize V5020 can have one control enclosure, also known as an *I/O Group*. The Storwize V5030 can consist of one or two control enclosures (two I/O Groups at most).

Each Storwize V5010 and Storwize V5020 control enclosure can attach up to 10 expansion enclosures. Each Storwize V5030 control enclosure can attach up to 20 expansion enclosures.

SAS cabling

Expansion enclosures are attached to control enclosures and between each other by using SAS cables.

A set of correctly interconnected enclosures is called a *chain*. Each chain is made up of two *strands*. A strand runs through the canisters that are in the same position in each enclosure in the chain. Canister 1 of an enclosure is cabled to canister 1 of the downstream enclosure. Canister 2 of an enclosure is cabled to canister 2 of the downstream enclosure.

Each strand consists of four phys, and each phy operates at 12 Gbps; therefore, a strand has a usable speed of 48 Gbps.

A strand starts with a SAS initiator chip inside an IBM Storwize V5000 Gen2 node canister and progresses through SAS expanders, which connect to the disk drives. Each canister contains an *expander*. Each drive has two ports, each of which is connected to a different expander and strand. This configuration means that both nodes directly access each drive, and no single point of failure exists.

At system initialization, when devices are added to or removed from strands (and at other times), the IBM Storwize V5000 Gen2 software performs a discovery process to update the state of the drive and enclosure objects.

The Storwize V5010 supports one SAS chain for each control enclosure, and up to 10 expansion enclosures can be attached to this chain. The node canister uses SAS port 1 for expansion enclosures.

Always cable from the top down. Cabling from the bottom to the top is not supported.

Figure 12-19 shows the SAS cabling on a Storwize V5010 with three attached expansion enclosures.

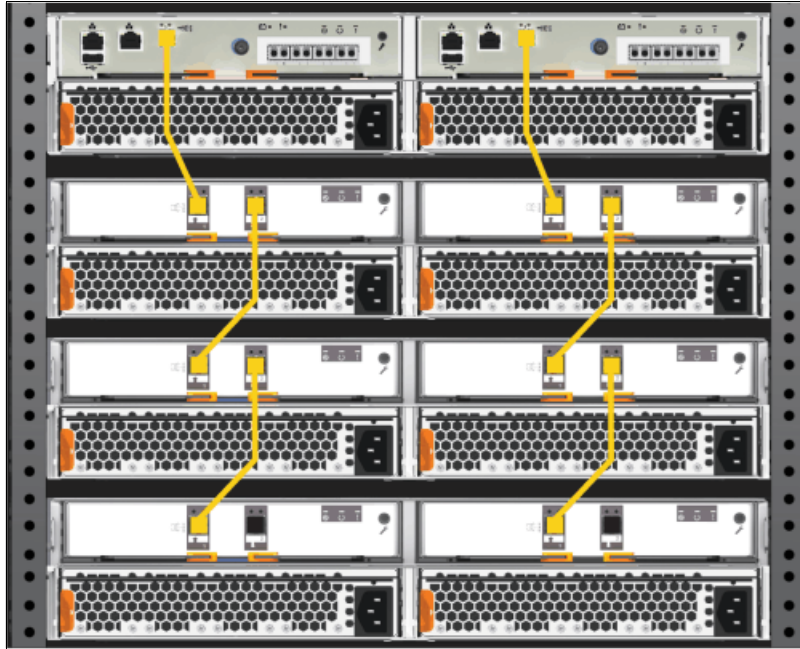


Figure 12-19 SAS expansion cabling on the Storwize V5010

The Storwize V5020 supports one SAS chain for each control enclosure, and up to 10 expansion enclosures can be attached to this chain. The node canister uses SAS port 1 for expansion enclosures.

Always cable from the top down. Cabling from the bottom to the top is not supported.

Figure 12-20 shows the SAS cabling on a Storwize V5020 with three attached expansion enclosures.

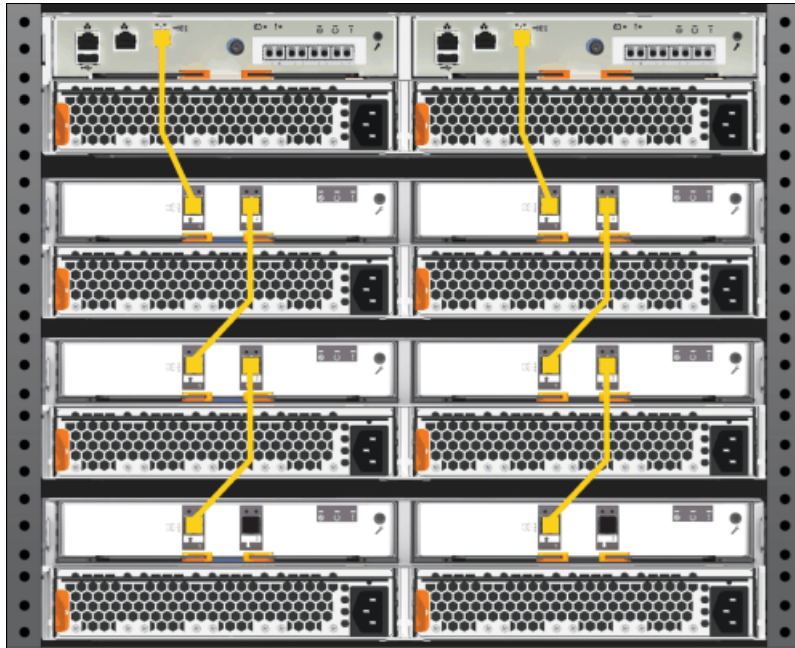


Figure 12-20 SAS expansion cabling on the Storwize V5020

The Storwize V5030 supports two SAS chains for each control enclosure, and up to 10 expansion enclosures can be attached to each chain. The node canister uses SAS port 1 for expansion enclosures.

Always cable from the top down. Cabling from the bottom to the top is not supported.

Figure 12-21 shows the SAS cabling on a Storwize V5030 with six attached expansion enclosures (three enclosures in each chain).

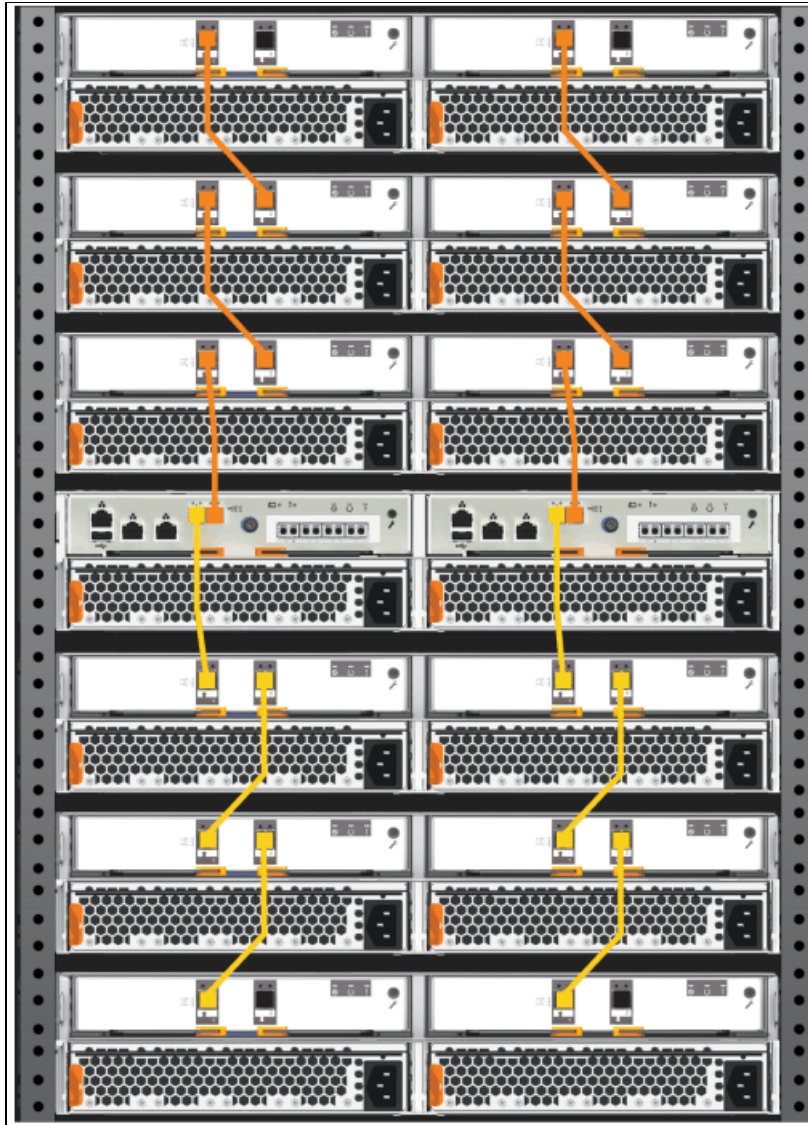


Figure 12-21 SAS expansion cabling on the Storwize V5030

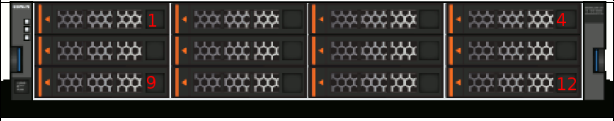

Important: When a SAS cable is inserted, ensure that the connector is oriented correctly by confirming that the following conditions are met:

- ▶ The pull tab must be below the connector.
- ▶ Insert the connector gently until it clicks into place. If you feel resistance, the connector is probably oriented the wrong way. Do not force it.
- ▶ When the connector is inserted correctly, the connector can be removed only by pulling the tab.
- ▶ Cabling is done from the controller view top →down. Top/down button up is *not* supported.

Drive slots

The IBM Storwize V5000 Gen2 has different types of enclosures, depending on the model, warranty, and number of drive slots. Table 12-7 shows the drive slots on each enclosure type.

Table 12-7 Drive slots for each enclosure type

Enclosure type	Drive slots
<ul style="list-style-type: none">▶ Control enclosure 2077/2078-112▶ Control enclosure 2077/2078-212▶ Control enclosure 2077/2078-312▶ Expansion enclosure 2077/2078-12F	12 x 3.5-inch slots  A photograph of a drive enclosure showing 12 drive slots arranged in two rows of six. The slots are numbered 1 through 12 in red. Each slot contains a drive carrier with a starburst pattern.
<ul style="list-style-type: none">▶ Expansion enclosure 2077/2078-92F	92 x 3.5-inch slots (usage of 2.5-inch drives possible with 3.5-inch carriers)
<ul style="list-style-type: none">▶ Control enclosure 2077/2078-124▶ Control enclosure 2077/2078-224▶ Control enclosure 2077/2078-324▶ Control enclosure 2078-U5A▶ Expansion enclosure 2077/2078-24F	24 x 2.5-inch slots  A photograph of a drive enclosure showing 24 drive slots arranged in two rows of 12. The slots are numbered 1 through 24 in red. Each slot contains a drive carrier with a starburst pattern.

Drive replacement procedure

You can reseal or replace a failed drive in a Storwize V5000 Gen2 by removing it from its enclosure and replacing it with the correct new drive without requiring the Directed Maintenance Procedure to supervise the service action.

The system can automatically perform the drive hardware validation tests and promote the drive into the configuration if these tests pass, which automatically configures the inserted drive as a spare. The status of the drive after the promotion can be recorded in the event log as an informational message or an error if a hardware failure occurs during the system action.

For more information about the drive replacement process, see the following IBM Storwize V5000 Gen2 Knowledge Center web pages:

- ▶ [Replacing a 3.5-inch drive assembly](#)
- ▶ [Replacing a 2.5-inch drive assembly](#)

12.2.5 Power supply units

All enclosures require two power supply units (PSUs) for normal operation. A single PSU can power the entire enclosure for redundancy. We advise that you supply AC power to each PSU from different power distribution units (PDUs).

Figure 12-22 shows a fully equipped control enclosure with two supply units. The PSUs are identical between the control and expansion enclosures.

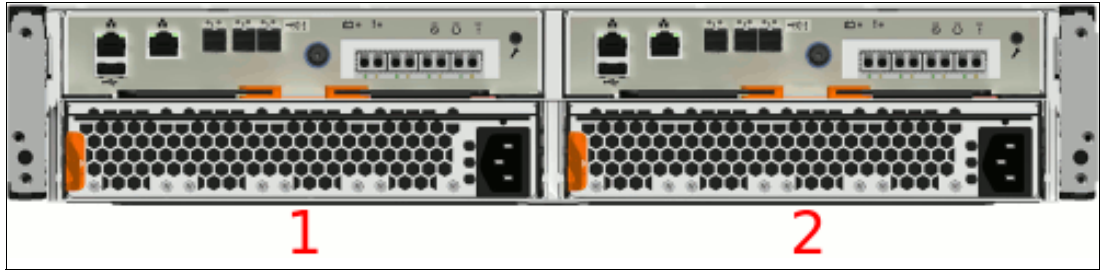


Figure 12-22 Power supply units

The left PSU is numbered 1, and the right PSU is numbered 2.

Power supplies in both control and expansion enclosures are hot-swappable and replaceable without a need to shut down a node or cluster. If the power is interrupted in one node canister for less than 2.5 seconds, the canister cannot perform a fire hose dump (destaging the data from cache to the Flash Memory) and continues operation from battery. If the power is interrupted for more than 2.5 seconds, it performs a firehose dump to save the data.

PSU status

Each PSU has three LEDs that display the status of its activity. The LEDs are the same for the control and expansion units.

Figure 12-23 shows the PSU status LEDs.



Figure 12-23 PSU status LEDs

The meaning of each LED is listed in Table 12-8.

Table 12-8 PSU status LEDs

Name and position	Color	State	Meaning
Input status (top)	Green	Solid	Input power is available.
		Off	No input power is available.
Output status (middle)	Green	Solid	PSU is providing DC output power.
		Off	PSU is not providing DC output power.
Fault (bottom)	Amber	Solid	A fault exists with the PSU.
		Flashing	The PSU is being identified. A fault might exist.
		Off	No fault is detected.

PSU replacement procedure

For more information about the PSU replacement process, see [IBM Storwize V5000 Gen2 Knowledge Center](#).

12.3 Configuration backup

The configuration backup file must be used if a serious failure occurs that requires the system configuration to be restored. The file contains configuration data of arrays, pools, volumes, and so on (but no client data).

The configuration backup file can be downloaded and saved by using the graphical user interface (GUI) or the command-line interface (CLI). The CLI option requires you to log in to the system and download the file by using Secure Copy Protocol (SCP). It is a preferred practice for an automated backup of the configuration.

Important: Save the configuration files of the IBM Storwize V5000 Gen2 regularly. The best approach is to save daily and automate this task. Always perform the additional manual backup before you perform any critical maintenance task, such as an update of the microcode or software version.

The backup file is updated by the cluster every day and stored in the /dumps directory. Even so, it is important to start a manual backup after you change your system configuration.

To successfully perform the configuration backup, follow the prerequisites and requirements:

- ▶ All nodes must be online.
- ▶ No independent operations that change the configuration can be running in parallel.
- ▶ No object name can begin with an underscore.

Important: You can perform an ad hoc backup of the configuration only from the CLI. However, the output of the command can be downloaded from both the CLI and the GUI.

12.3.1 Generating a manual configuration backup by using the CLI

You can use the CLI to trigger a configuration backup manually on an ad hoc basis or by an automatic process regularly. The use of the **svconfig backup** command generates a new backup file. Triggering a backup by using the GUI is not possible, but you can save the output from the GUI.

Example 12-1 shows the output of the **svconfig backup** command.

Example 12-1 Triggering a backup by using the CLI

```
>svconfig backup
.....
.....
.....
CMMVC6155I SVCCONFIG processing completed successfully
```

The **svconfig backup** command creates three files that provide information about the backup process and cluster configuration. These files are created in the /dumps directory on the configuration node and can be retrieved by using SCP. Use the **lsdumps** command to list them, as shown in Example 12-2.

Example 12-2 Listing the backup files by using the CLI

```
>lsdumps
id filename
...
48 svc.config.backup.xml_781000E-1
49 svc.config.backup.sh_781000E-1
50 svc.config.backup.log_781000E-1
...
```

The three files that are created by the backup process are listed in Table 12-9.

Table 12-9 Files that are created by the backup process

File name	Description
svc.config.backup.xml_<serial>	This file contains the cluster configuration data.
svc.config.backup.sh_<serial>	This file contains the names of the commands that were issued to create the backup of the cluster.
svc.config.backup.log_<serial>	This file contains details about the backup, including any error information that might be reported.

12.3.2 Downloading a configuration backup by using the GUI

The IBM Storwize V5000 Gen2 does not offer an option to initiate a backup from the GUI. However, you can download existing daily backups or manual backups that were triggered from the CLI.

To download a configuration backup file by using the GUI, complete the following steps:

1. Browse to **Settings** → **Support** → **Support Package** and select from under **Manual Upload Instructions** → **Download Support Package** (see Figure 12-24).

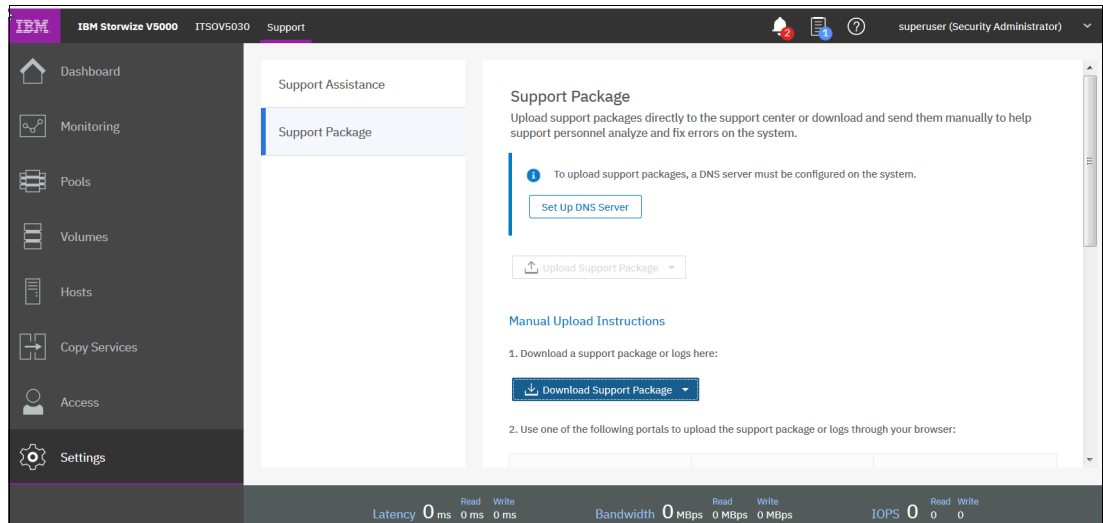


Figure 12-24 Manual Upload Instructions

When you select **Download Support Package**, a window opens (see Figure 12-25).

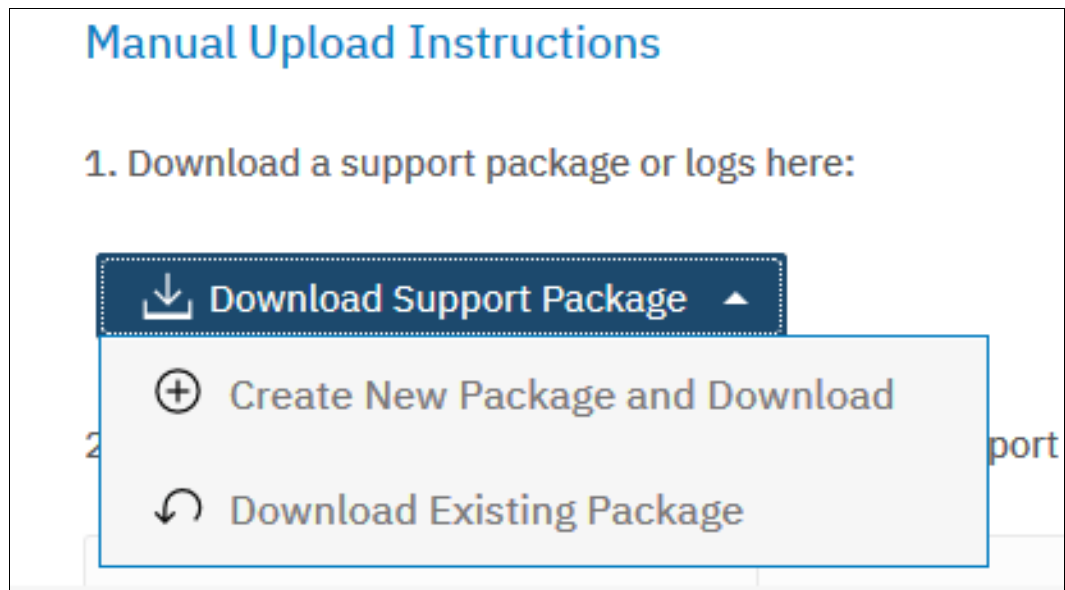


Figure 12-25 Download Support Package

2. Clicking **Download Existing Package** takes you to the next option, where you can select the different kinds of Support packages and all the available log files that are stored on the configuration node, as shown in Figure 12-26.

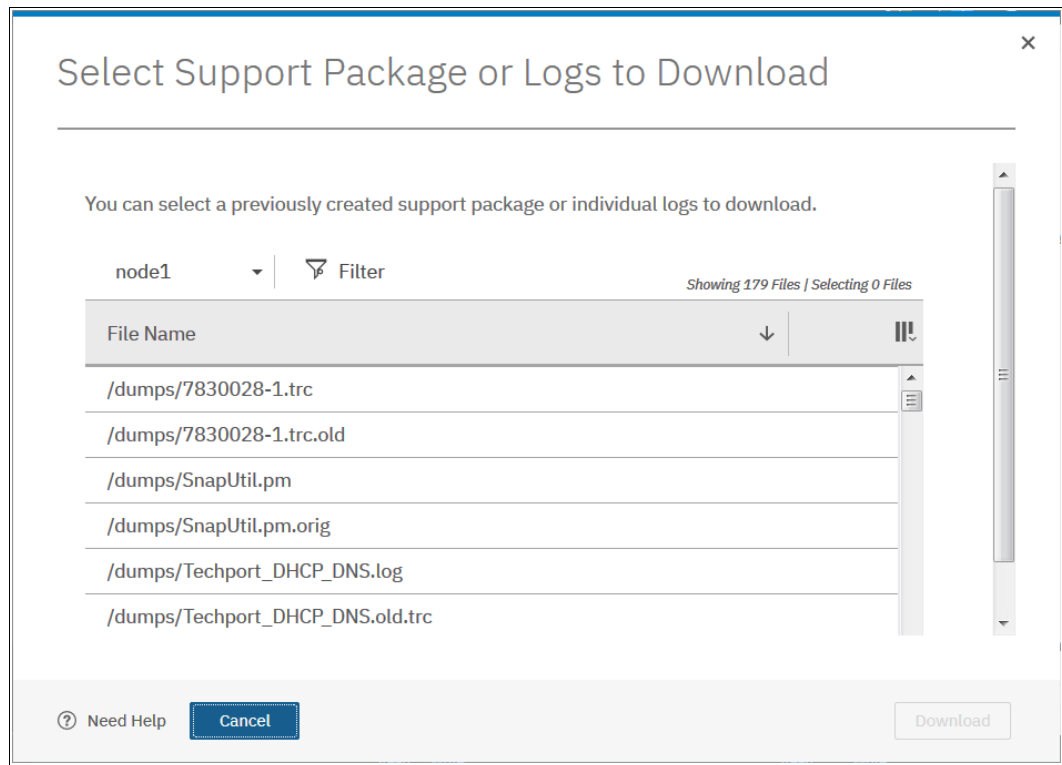


Figure 12-26 Full log listing option

3. Search for the files that are named `svc.config.backup.xml_*`, `svc.config.backup.sh_*`, and `svc.config.backup.log_*`. Select the files, right-click, and select **Download**, as shown in Figure 12-27.

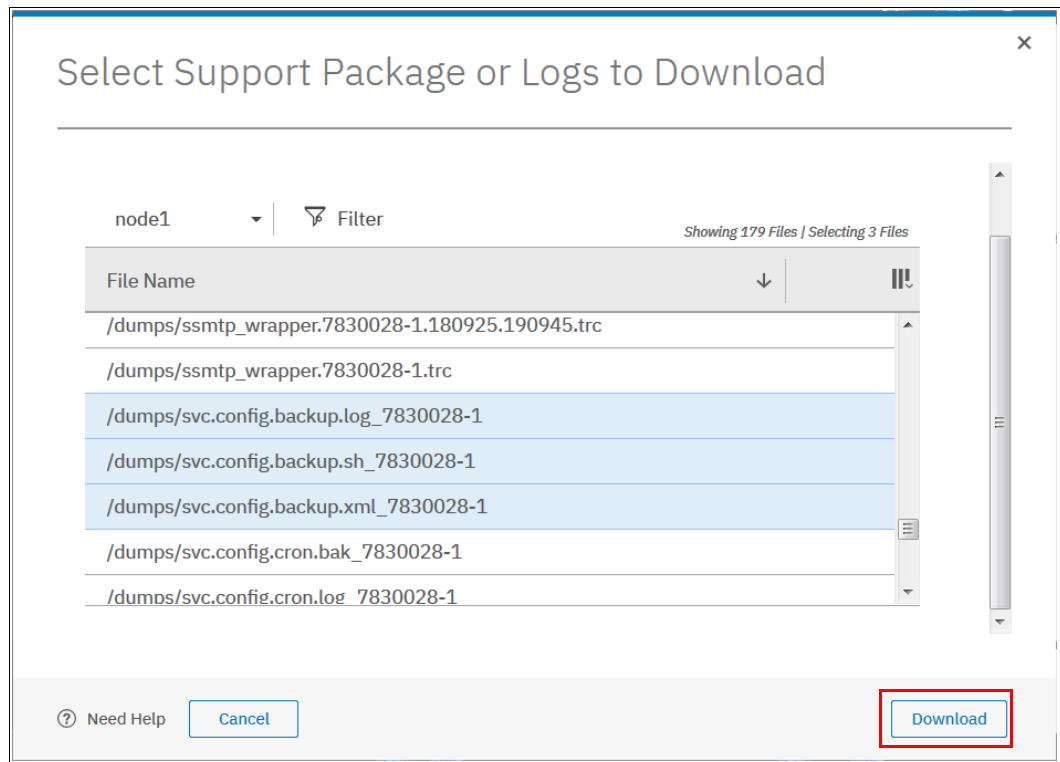


Figure 12-27 Backup files download

4. Although the configuration backup files are updated automatically daily, it might be useful to verify the time stamp of the actual file.

Open the `svc.config.backup.xml_xx` file with a text editor and search for the string `timestamp=`, which is near the top of the file.

Figure 12-28 shows the file and the timestamp information.

Tree View	XSL Output
xml	version="1.0"
xml	
xml	
label	Configuration Back-up
version	821
file_version	147.4.0000000000003
timestamp	2018/09/24 07:40:20 UTC
#comment	cluster_banner section
#comment	cloudaccount section
#comment	cluster section
object	
#comment	clusterip section
object	
object	
#comment	controller section
#comment	drive section
object	

Figure 12-28 Timestamp in the backup xml file

12.4 System update

The system update process involves updating the entire IBM Storwize V5000 Gen2 environment.

The node canister software and the drive firmware are updated separately, so these tasks are separately.

Note: Storwize V5000 Gen1 hardware is not supported by IBM Spectrum Virtualize V8.1 or later. The V7.7.1 and V7.8.1 code streams will continue to be updated with critical fixes for this hardware.

12.4.1 Updating node canister software

For more information about the latest software and to download the software package, see this IBM Support [web page](#).

When you start or restart the GUI, you see a warning that a software update is available, as shown in Figure 12-29.

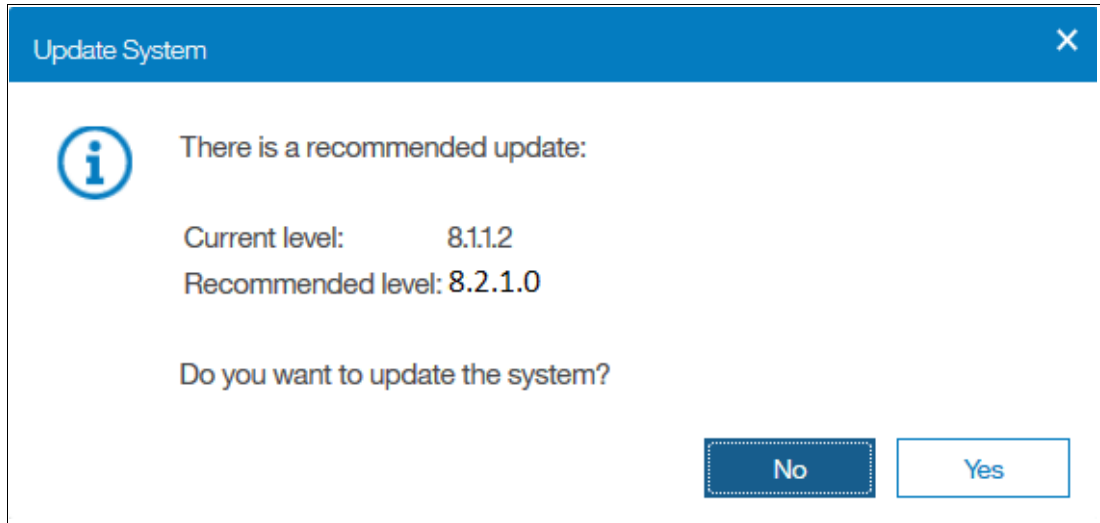


Figure 12-29 New software is available

The GUI also shows whether a software update is available and the latest software level when you navigate to **Settings** → **System** → **Update System**, as shown in Figure 12-30.

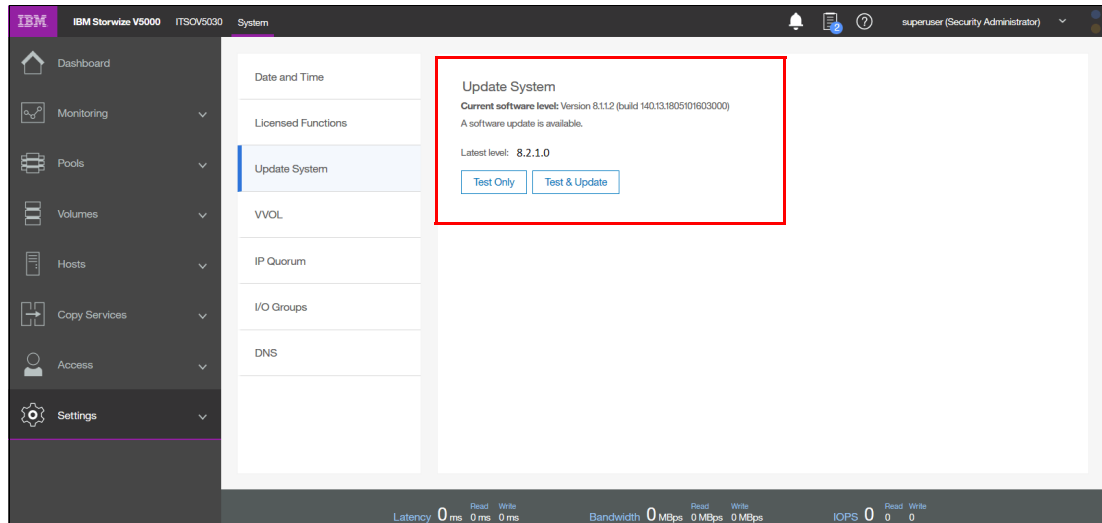


Figure 12-30 Latest software level available

Important: Certain levels of code support updates from only specific previous levels. If you update to more than one level above your current level, you might be required to install an intermediate level. For more information about update compatibility, see this [IBM Support website](#).

Preparing for the update

Allow sufficient time to plan your tasks, review your preparatory update tasks, and complete the update of the IBM Storwize V5000 Gen2 environment. The update procedures can be divided into the following general update tasks, as listed in Table 12-10.

Table 12-10 Software update tasks

Sequence	Upgrade tasks
1	Decide whether you want to update automatically or manually. During an automatic update procedure, the clustered system updates each of the nodes systematically. The automatic method is the preferred procedure for updating software on nodes. However, you can update each node manually.
2	Ensure that Common Information Model (CIM) object manager (CIMOM) clients are working correctly. When necessary, update these clients so that they can support the new version of the IBM Storwize V5000 Gen2 code. Examples can be operating system (OS) versions and options, such as FlashCopy Manager or VMware plug-ins.
3	Ensure that multipathing drivers in the environment are fully redundant. If you experience failover issues with multipathing driver support, resolve these issues before you start normal operations.
4	Update other devices in the IBM Storwize V5000 Gen2 environment. Examples might include updating the hosts and switches to the correct levels.
5	Update your IBM Storwize V5000 Gen2.

Important: Ensure that no unfixed errors are in the log and that the system date and time are correctly set before you start the update.

The amount of time that it takes to perform a node canister update can vary depending on the amount of preparation work that is required and the size of the environment. Generally, to update the node software, allow 20 - 40 minutes for each node canister and a single 30-minute wait when the update is halfway complete.

One node in each I/O group can be upgraded to start, then the system can wait 30 minutes before it upgrades the second node in each I/O group. The 30-minute wait allows the recently updated node canister to come online and be confirmed as operational, and it allows time for the host multipath to recover. So, plan at least two hours for an update.

The software update can be performed concurrently with normal user I/O operations. After the updating node is unavailable, all I/O operations fail to that node and the failed I/O operations are directed to the partner node of the working pair. Applications do not see any I/O failures.

The maximum I/O rate that can be sustained by the system might degrade while the code is uploaded to a node, the update is in progress, the node is rebooted, and the new code is committed because write caching is disabled during the node canister update process.

Important: Ensure that the multipathing drivers are fully redundant with every available path and online. You might see errors that are related to the paths, which can go away (failover), and the error count can increase during the update. When the paths to the nodes return, the nodes fall back to become a fully redundant system.

When new nodes are added to the system, the upgrade package is automatically downloaded to the new nodes from the IBM Storwize V5000 Gen2 system.

Update test utility

The Storwize V5000 Gen2 update test utility checks for known issues that can cause problems during a software update. For more information about and to download the utility, see [this website](#).

The software update test utility can be downloaded in advance of the update process, or it can be downloaded and run directly during the software update, as guided by the update wizard. You can run the utility multiple times on the same system to perform a readiness check-in preparation for a software update.

The installation and use of this utility is non-disruptive, and it does not require a restart of any node. Therefore, host I/O is not interrupted. The utility is installed on the current configuration node only.

System administrators must continue to check whether the version of code that they plan to install is the latest version.

Updating the software automatically by using the GUI

Complete the following steps to automatically update the node canister software by using the GUI:

1. Browse to **Settings** → **System** → **Update System** and select **Test and Update**, as shown in Figure 12-31.

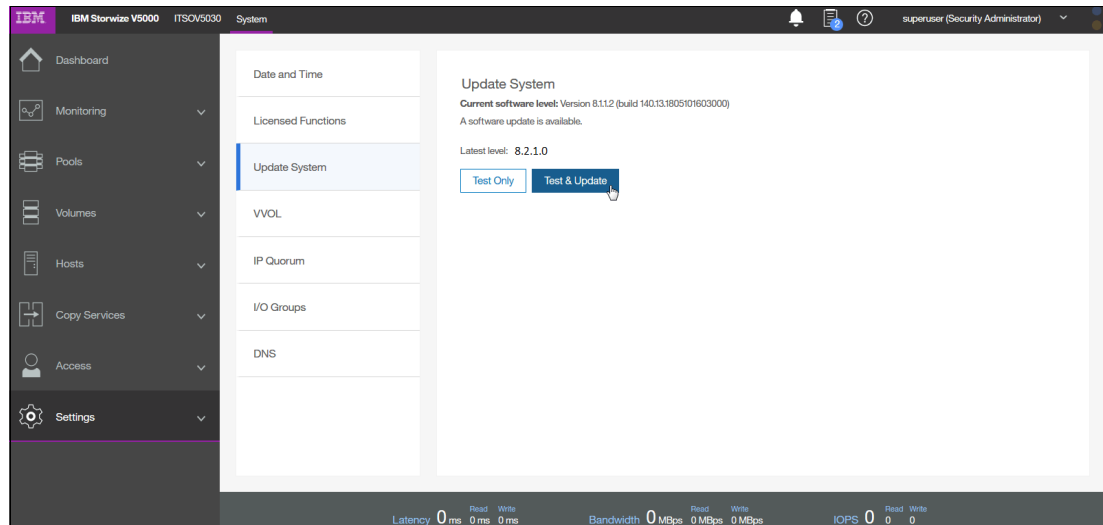


Figure 12-31 Update system window

Alternatively, you can run only the test utility by selecting **Test Only**.

2. Select the test utility and update package files by clicking the folder icons, as shown in Figure 12-32. The code levels are entered automatically.

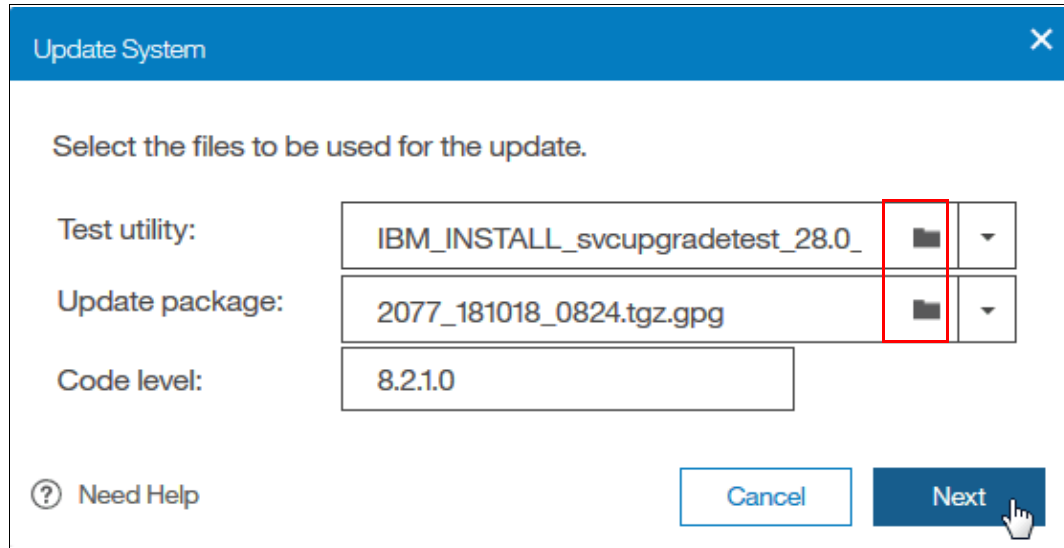


Figure 12-32 File selection

Alternatively, for the **Test Only** option, upload only the test utility and enter the code level manually.

3. Select **Automatic update** and click **Next** to come to the next question regarding paused update, as shown in Figure 12-33. The **Automatic update** option is the default and advised choice.

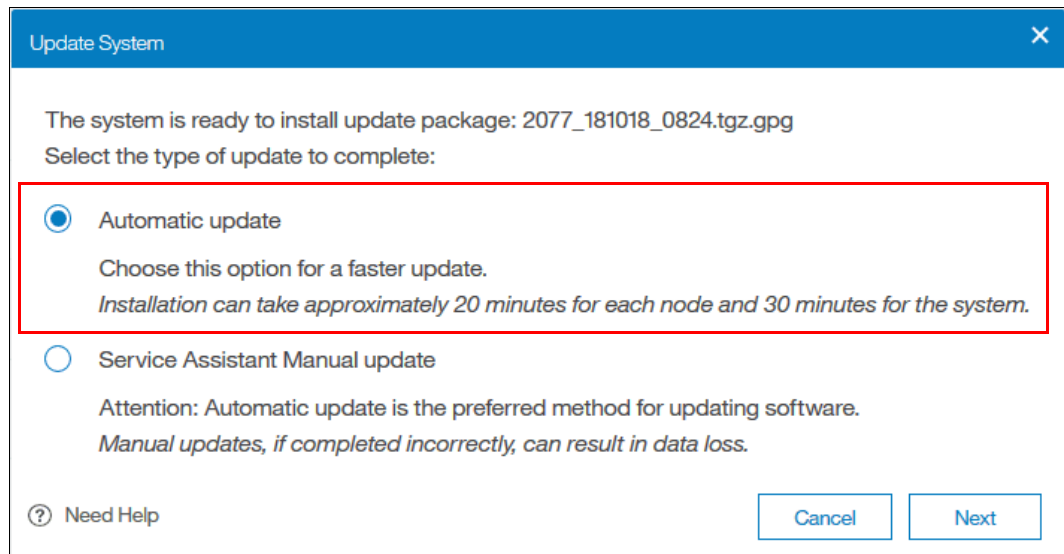


Figure 12-33 Automatic update selection

- As shown in Figure 12-34, you can choose if you want to pause the update or not. Default is **Fully automatic**. Click **Finish** to start the update.

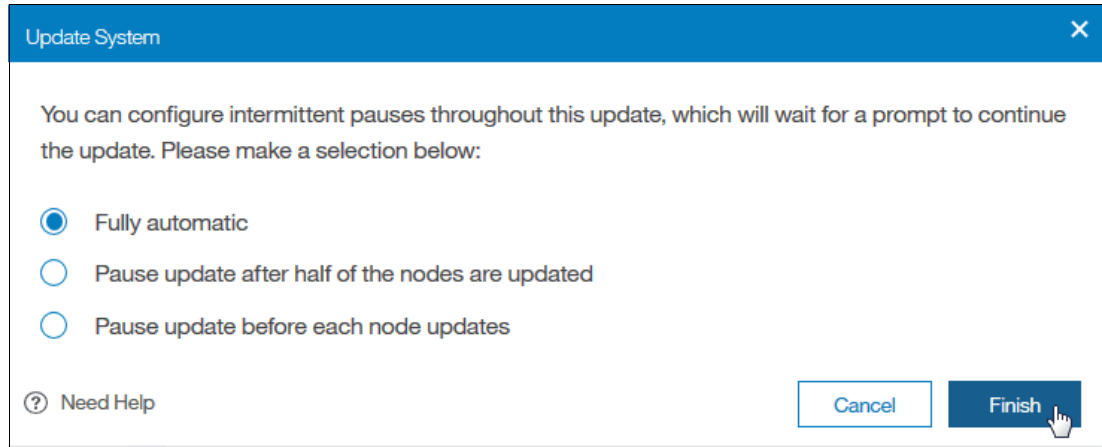


Figure 12-34 Fully automatic

- Wait for the test utility and update package to upload to the system, as shown in Figure 12-35.

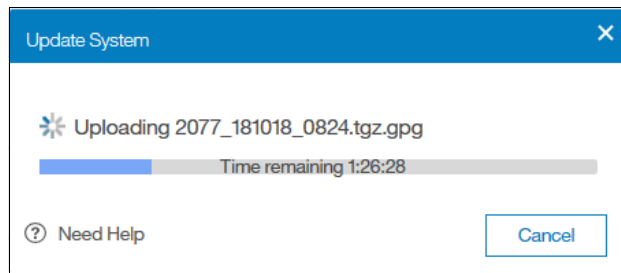


Figure 12-35 File upload

If you click the progress bar, you can toggle between the Time remaining and the Upload speed (see Figure 12-36).

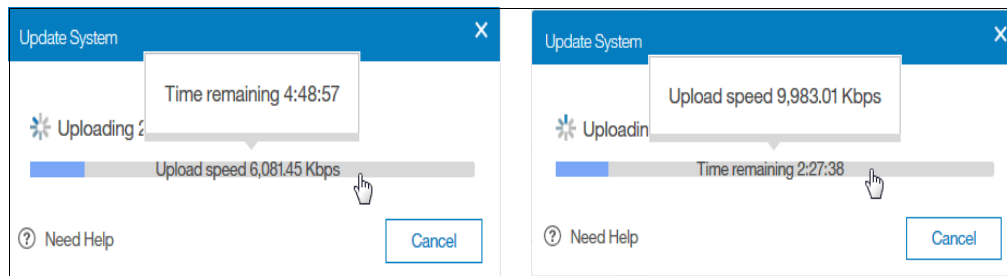


Figure 12-36 File upload

- After the files upload, the test utility is automatically run, as shown in Figure 12-37. The test utility verifies that no issues exist with the current system environment, such as failed components and drive firmware that is not at the latest level.

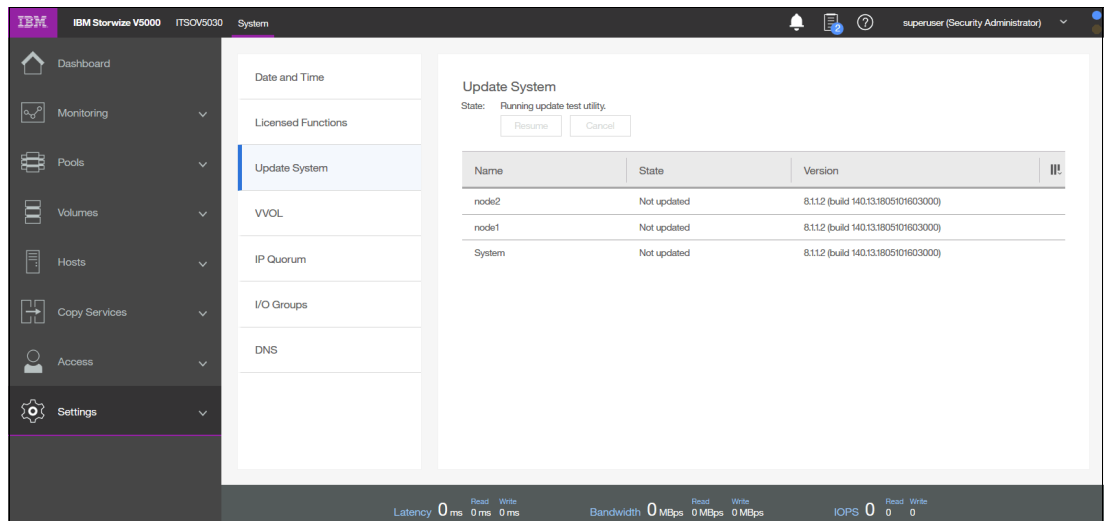


Figure 12-37 State while the test utility runs

- If the test utility discovers any warnings or errors, a window opens to inform the user, as shown in Figure 12-38. Click **Read more** to get more information.

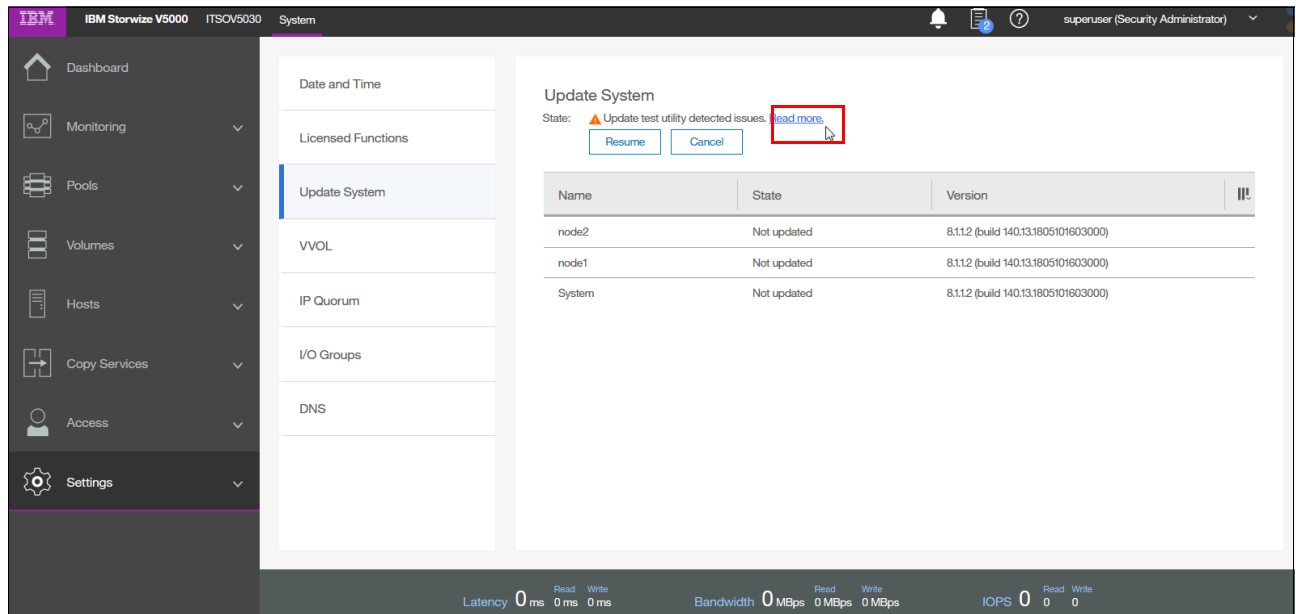


Figure 12-38 Warning about the issues that were detected

Figure 12-39 shows that the test utility identified one warning in this example.

Update Test Utility Results [Close]

***** Warning found *****

This tool has found the internal disks of this system are not running the recommended firmware versions.
Details follow:

Model	Latest FW	Current FW	Drive Info
ST6000NM0014	BC7A	BC78	Drive 0 in slot 1 in enclosure 1
			Drive 1 in slot 2 in enclosure 1
			Drive 2 in slot 11 in enclosure 1
			Drive 3 in slot 9 in enclosure 1
			Drive 4 in slot 5 in enclosure 1
			Drive 5 in slot 10 in enclosure 1
			Drive 6 in slot 6 in enclosure 1

We recommend that you upgrade the drive microcode at an appropriate time. If you believe you are running the latest version of microcode, then check for a later version of this tool. You do not need to upgrade the drive firmware before starting the software upgrade.

***** Error found *****

The system identified that one or more drives in the system are running microcode with a known issue.

The following flashes are appropriate for your drives:

* Please see the following web page for details:
<http://www.ibm.com/support/docview.wss?rs=591&uid=ssg1S1005289>

The following drives are affected by this issue:
0, 1, 2, 3, 4, 5, 6

Results of running svcupgradetest:
=====

The tool has found 1 errors and 1 warnings.

[Need Help](#) [Close](#) [Download Results](#)

Figure 12-39 Test utility results

Warnings do not prevent the software update from continuing, even if the recommended procedure is to fix all issues before you proceed.

8. Close the window and select **Resume** or **Cancel**, as shown in Figure 12-40. Clicking **Resume** continues the software update. Clicking **Cancel** cancels the software update so that the user can correct any issues.

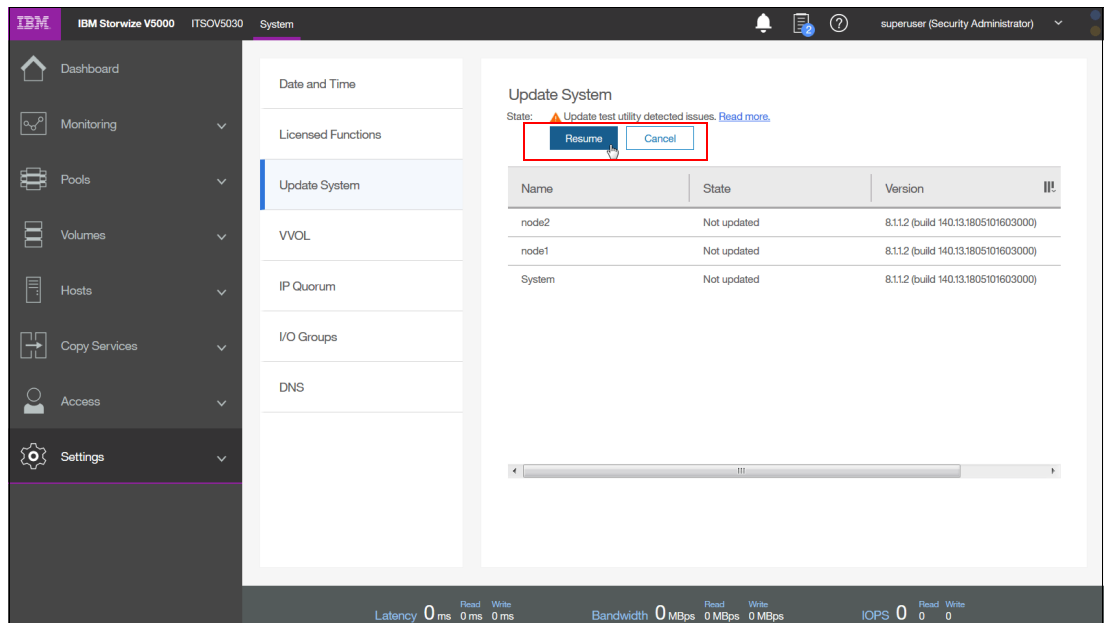


Figure 12-40 State after you run the test utility

9. Selecting **Resume** prompts the user to confirm the action, as shown in Figure 12-41.

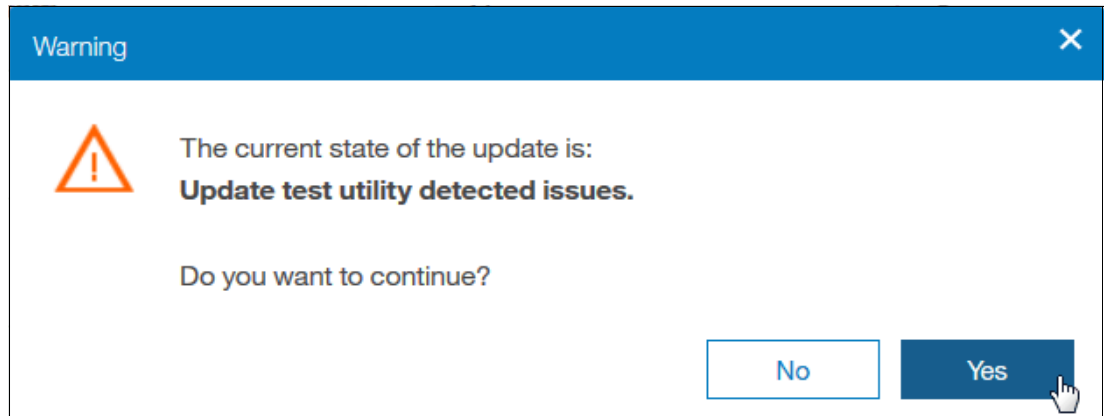


Figure 12-41 Resume confirmation window

10. You are prompted again if you really want to initialize the system (see Figure 12-42). Click **Yes** to proceed.

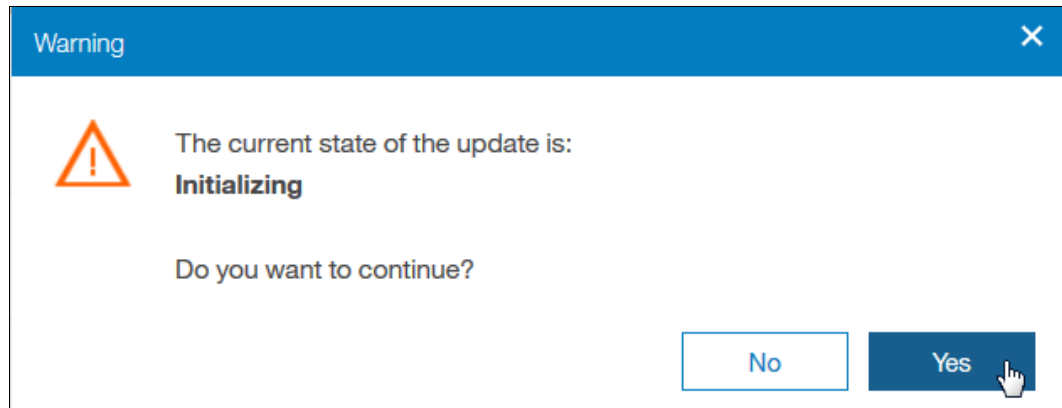


Figure 12-42 Initializing

11. Wait for each node to be updated and rebooted individually until the update process is complete. The GUI displays the overall progress of the update and the current state of each node, as shown in Figure 12-43.

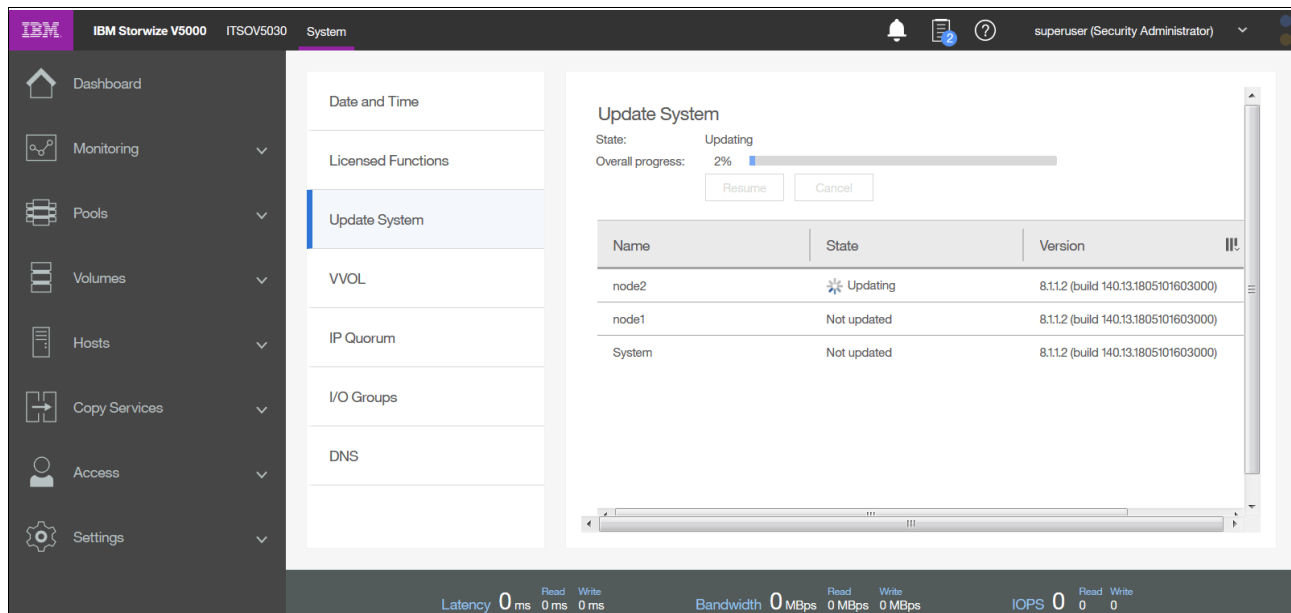


Figure 12-43 Automatic update progress

After the first node completes its upgrade, the system pauses for at least for 30 minutes. You can see in the GUI when the system finalizes the software update (see Figure 12-44).

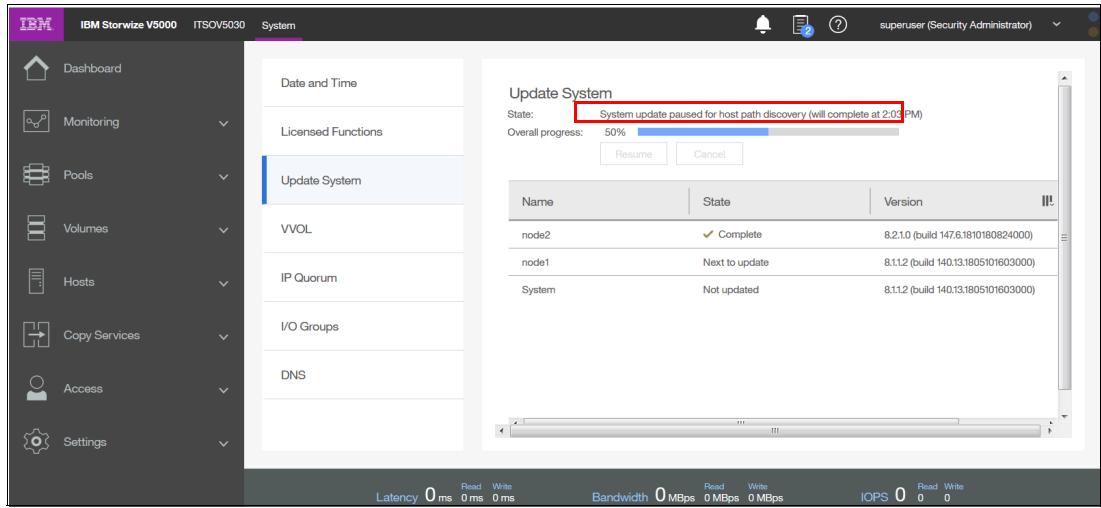


Figure 12-44 First node finished

12. During the update process, a node fails over and you can temporarily lose connection to the GUI. After this situation occurs, a warning is displayed, as shown in Figure 12-45. Select **Yes**.

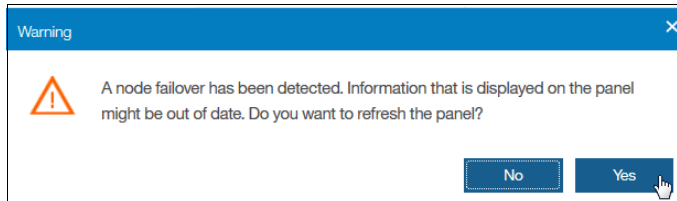


Figure 12-45 Configuration node failover warning

Updating the software manually by using the GUI and Service Assistant Tool

Important: We advise that you update the IBM Storwize V5000 Gen2 automatically by following the update wizard. If a manual update is used, ensure that you do not skip any steps.

Complete the following steps to manually update the software by using the GUI and Service Assistant Tool (SAT):

1. Browse to **Settings** → **System** → **Update System** and select **Test and Update**, as shown in Figure 12-46.

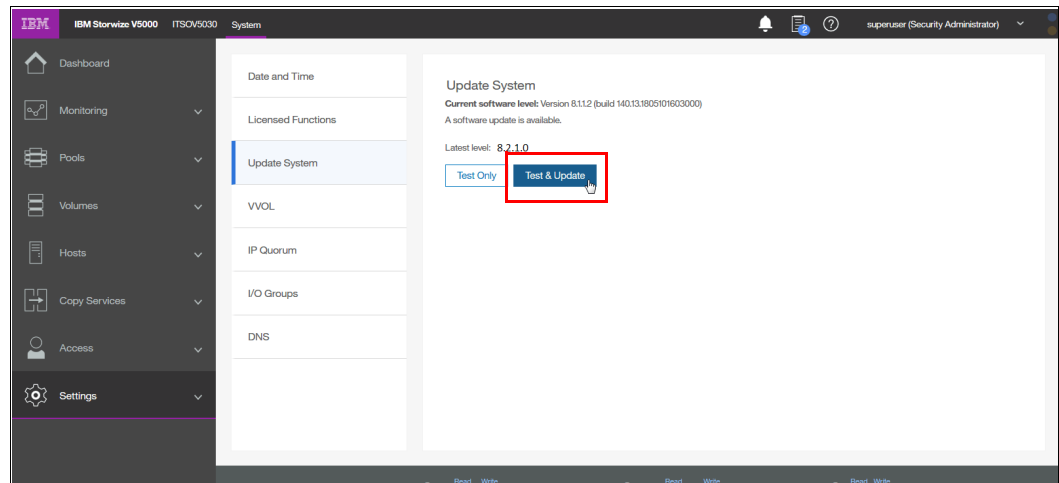


Figure 12-46 Update system window

Alternatively, you can run the test utility by selecting **Test Only**.

2. Select the test utility and update package files by clicking the folder icons, as shown in Figure 12-47. The code levels are entered automatically.

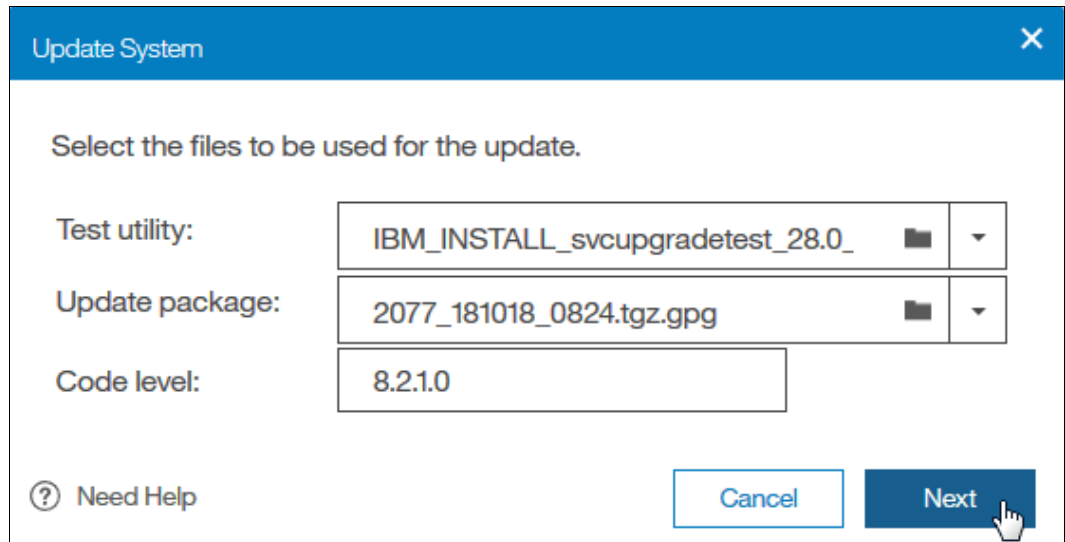


Figure 12-47 File selection

Alternatively, for the **Test Only** option, upload only the test utility and enter the code level manually.

3. Select **Service Assistant Manual update** and click **Finish**, as shown in Figure 12-48.

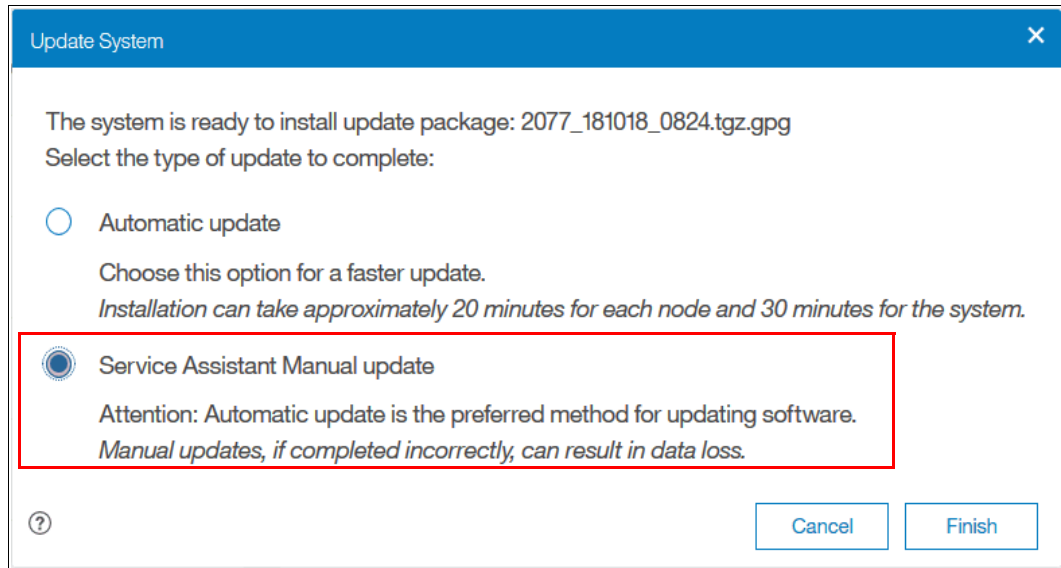


Figure 12-48 Manual update selection

4. Wait for the test utility and update package to upload to the system, as shown in Figure 12-49.

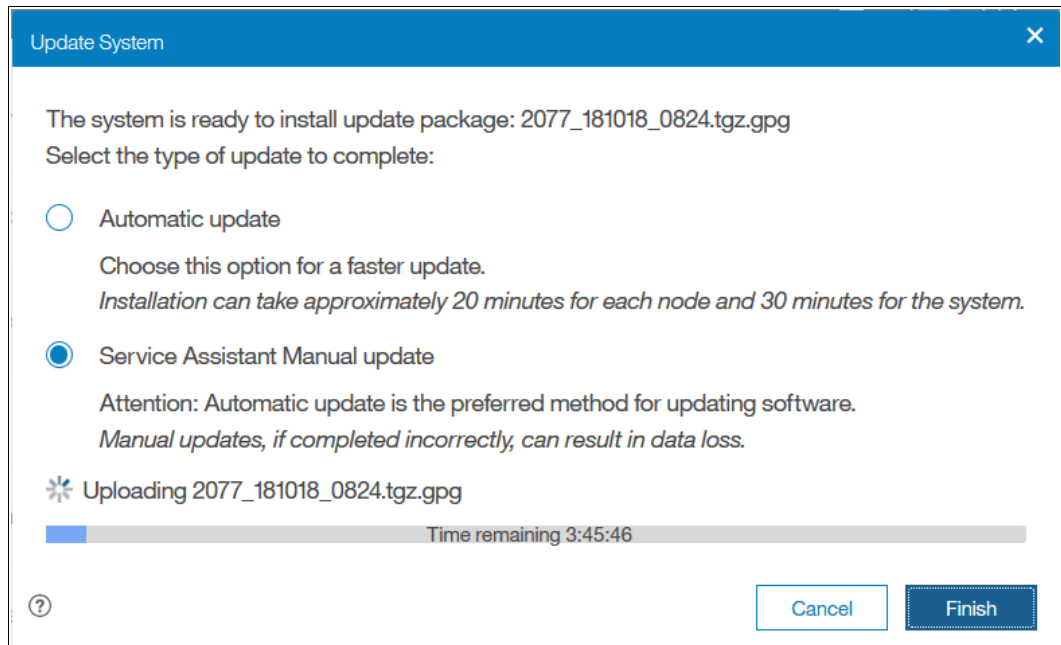


Figure 12-49 File upload

- After the files upload, the test utility is automatically run, as shown in Figure 12-50. The test utility verifies that no issues exist with the current system environment, such as failed components and drive firmware that is not at the latest level.

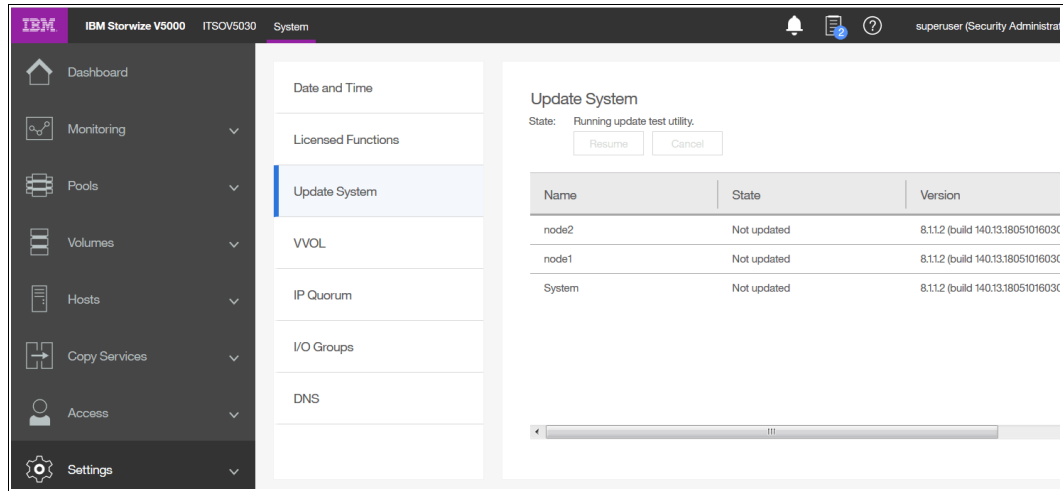


Figure 12-50 State while the test utility run

If the utility identifies no issues, the system is ready for the user to initiate the manual upgrade, as shown in Figure 12-51.

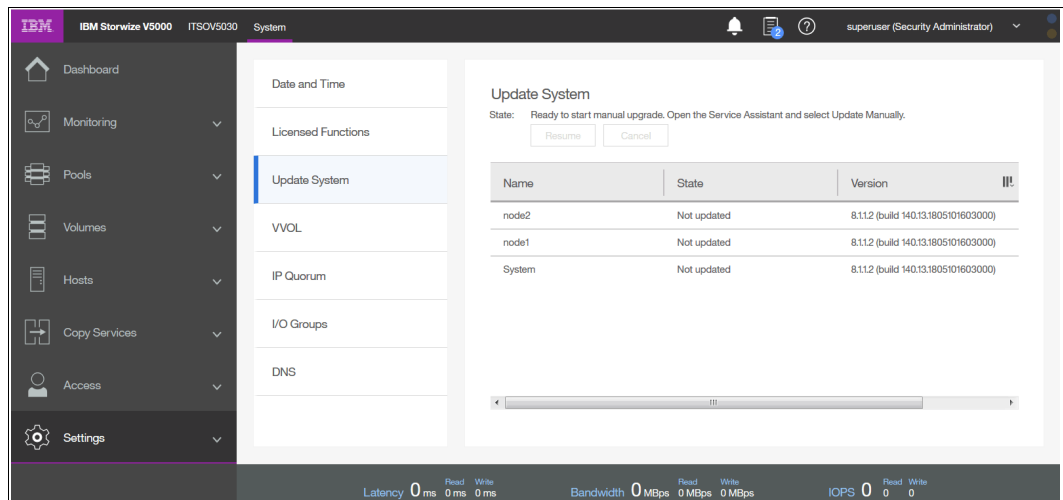


Figure 12-51 State while you wait for the manual upgrade to start

- Choose a node to update. Non-configuration nodes must be updated first. Update the configuration node last. Browse to **Monitoring** → **System** and click the canister. Select **Properties** to confirm the nodes that are the non-configuration nodes, as shown in Figure 12-52.

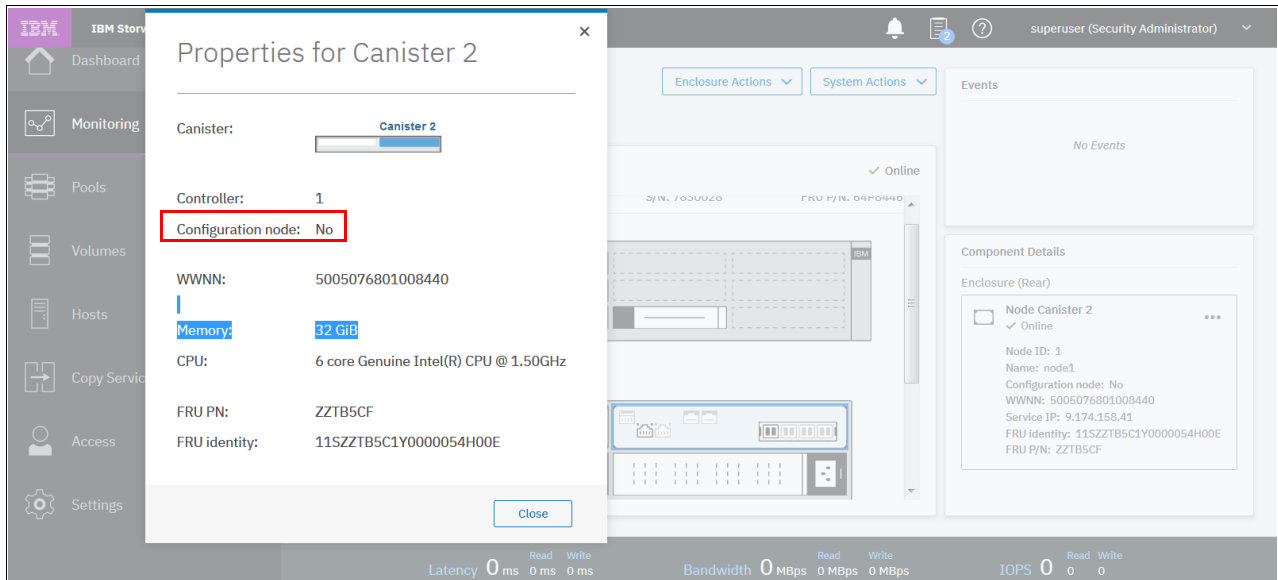


Figure 12-52 Checking the configuration node status

- Right-click the canister that contains the node that you want to update and select **Remove**, as shown in Figure 12-53.

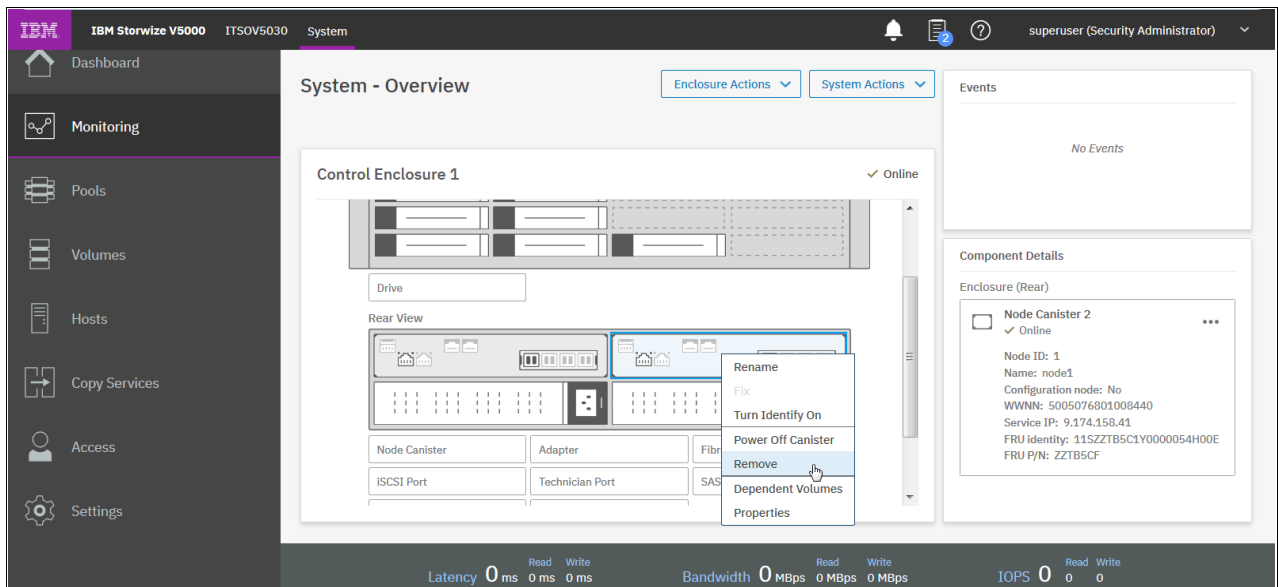


Figure 12-53 Removing a node canister

Important: Ensure that you select the non-configuration nodes first.

- A warning message appears in which you are prompted whether you want to remove the node, as shown in Figure 12-54. Click **Yes**.

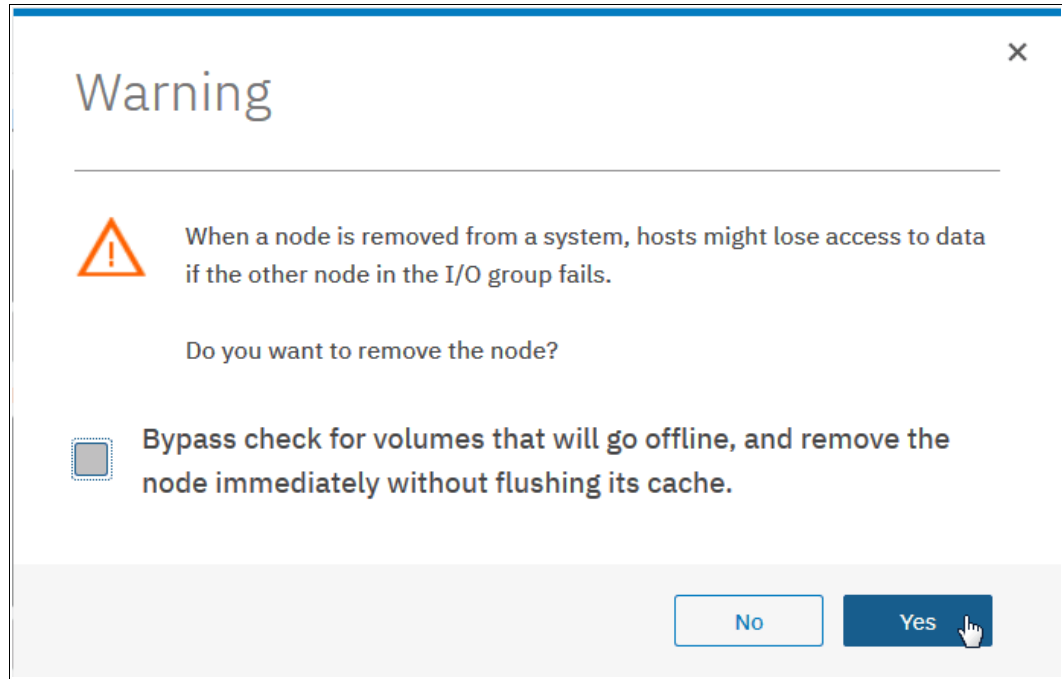


Figure 12-54 Node removal confirmation window

The non-configuration node is removed from the management GUI Update System window and is shown as Unconfigured when you hover over the node after you select **Monitoring** → **System**.

- Open the Service Assistant Tool for the node that you removed. Enter the Service IP Address followed by /service into a browser window. Without /service, the browser opens the associated GUI to this service IP. No HTTP:// or HTTPS:// is needed.

Example: 172.163.18.34/service

- In the Service Assistant Tool, ensure that the node that is ready for update is selected. The node can be in the Service status, display a 690 error, and show no available cluster information, as shown in Figure 12-55.

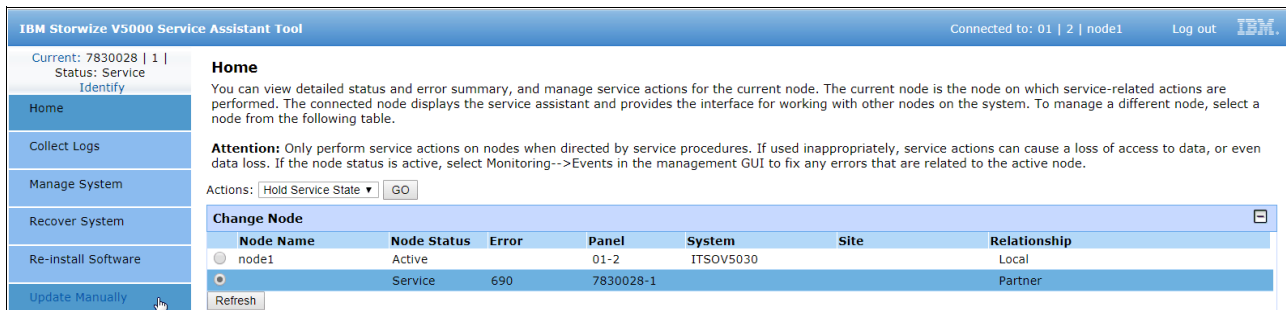


Figure 12-55 Node to update in the Service Assistant Tool

11. In the Service Assistant Tool, select **Update Manually**, and choose the required node canister software upgrade file, as shown in Figure 12-56.

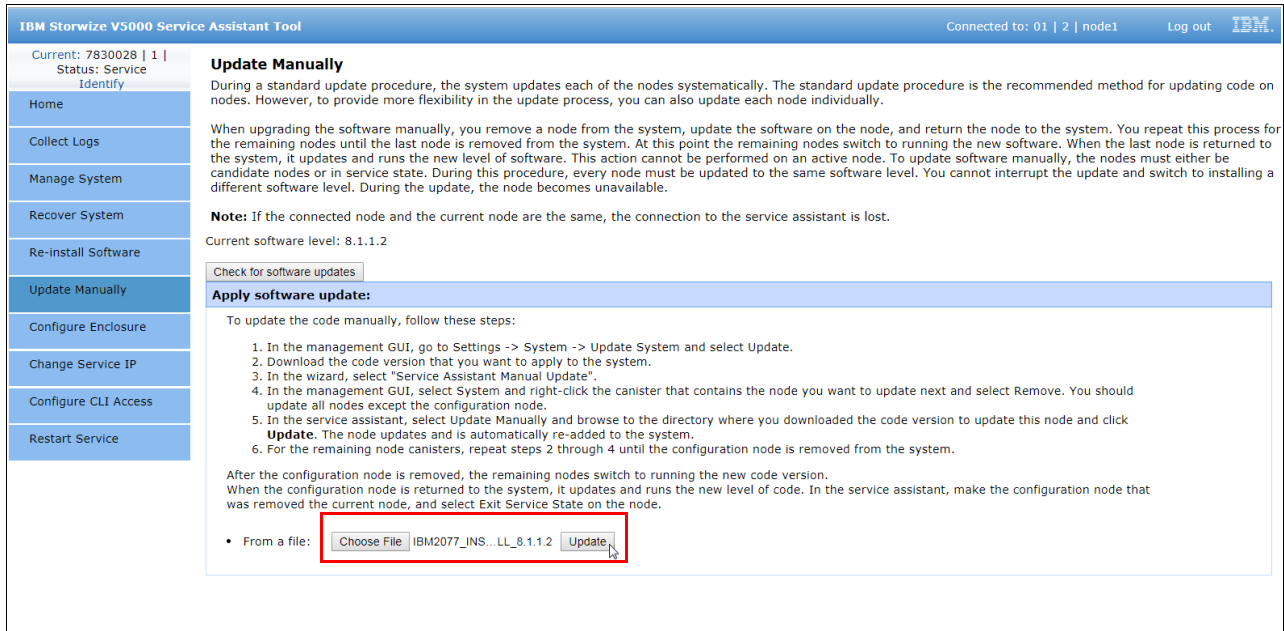


Figure 12-56 Starting the update in the Service Assistant Tool

12. Click **Update** to start the update process on the first node and wait for the node to finish updating.

Non-configuration nodes can be reintroduced automatically into the system after the update finishes. Updating and adding the node again can last 20 - 40 minutes.

The management GUI shows the progress of the update, as shown in Figure 12-57.

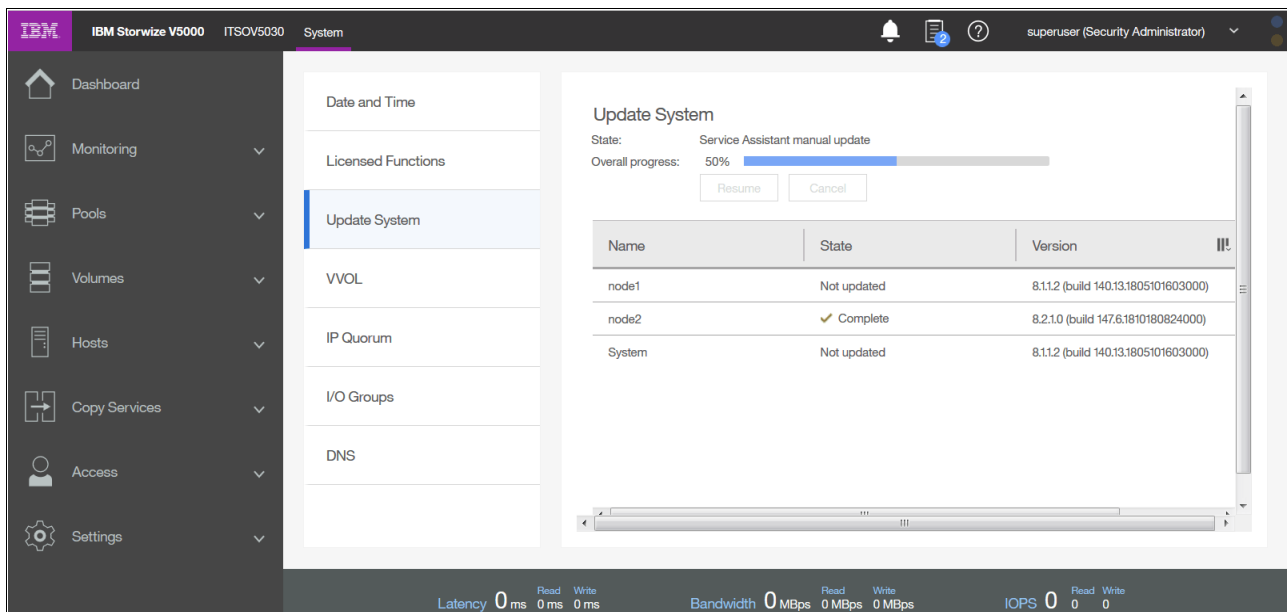


Figure 12-57 Manual update progress

13. Repeat steps 7 - 12 for the remaining nodes, leaving the configuration node until last.

14. After you remove the configuration node from the cluster, you are asked whether you want to refresh the window, as shown in Figure 12-58. Select **Yes**.

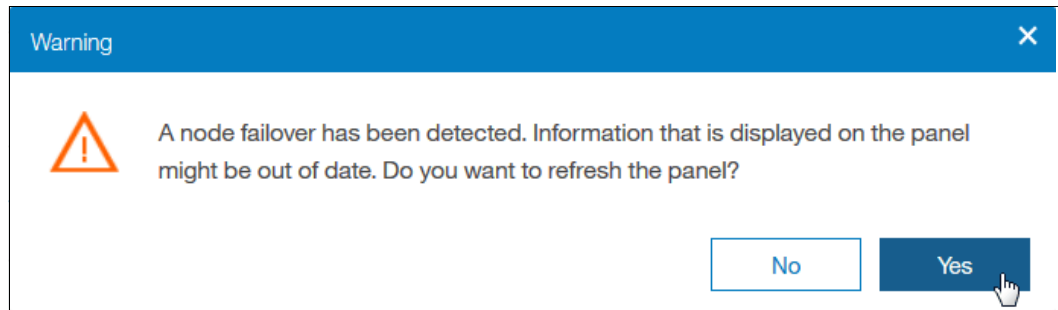


Figure 12-58 Configuration node failover warning

Important: The configuration node remains in the Service state when it is added to the cluster again. Therefore, you must exit the Service state manually.

You can see the progress of the update at the top of the window, as shown in Figure 12-59.

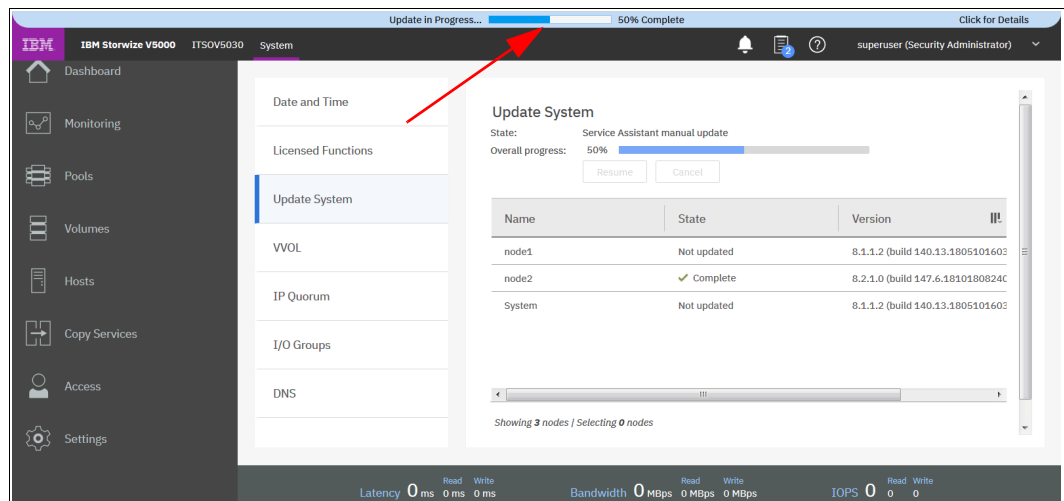


Figure 12-59 Progress of update

15. To exit Service state, browse to the Home window of the Service Assistant Tool and open the Actions menu. Select **Exit Service State** and click **Go** (see Figure 12-60).

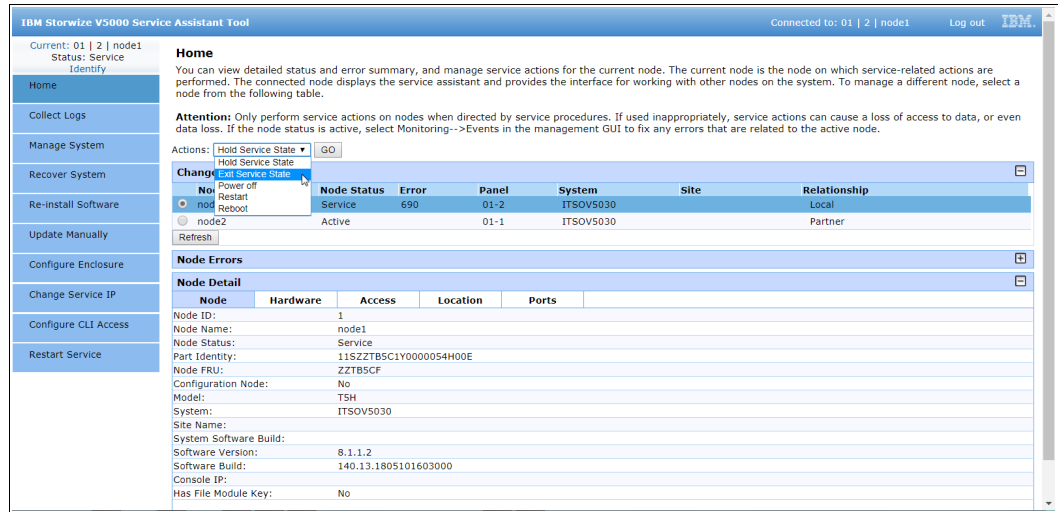


Figure 12-60 Exiting the Service state in the Service Assistant Tool

12.4.2 Updating the drive firmware

Drive firmware can be updated for all drives at the same time or individually.

For more information about the latest drive update package, see the [Supported Drive Types and Firmware Levels for the IBM Storwize V5000 web page](#).

Note: Find the download link for the drive firmware at the top of the Web page under **Downloads** → **Fix Central**.

Updating the firmware on individual drives

To update an individual drive, complete the following steps:

1. Navigate to **Pools** → **Internal Storage**, right-click the drive to update, and select **Upgrade** from the Actions menu, as shown in Figure 12-61.

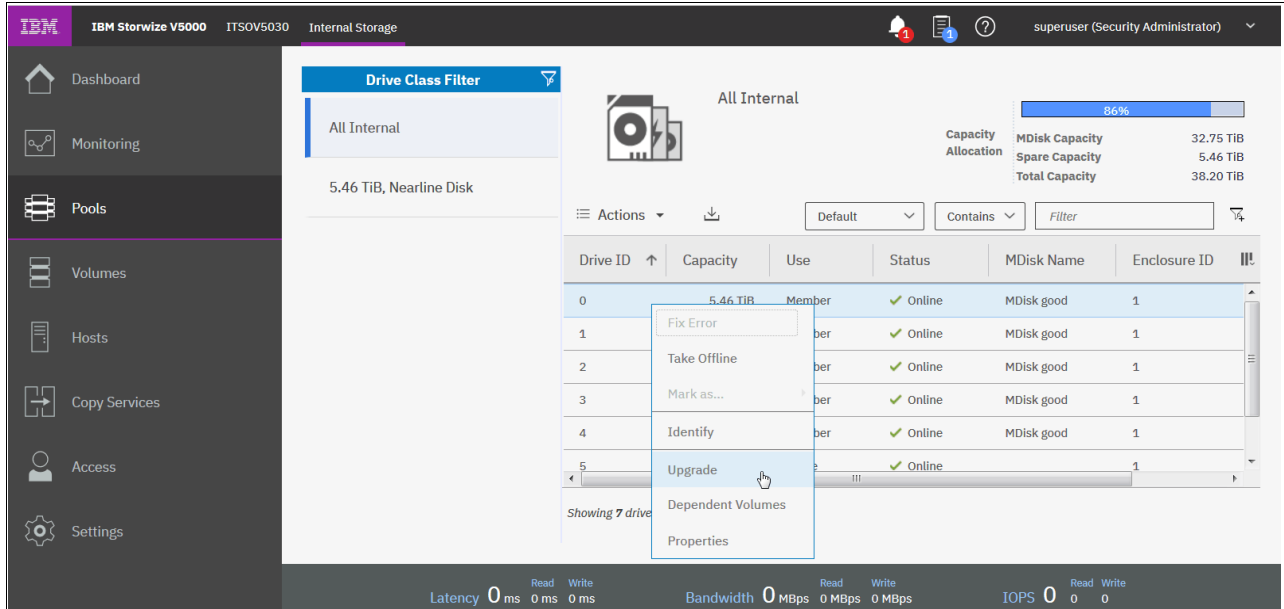


Figure 12-61 Individual drive update

2. Select the upgrade package, which was downloaded from the IBM Support site, by clicking the folder icon, and click **Upgrade**, as shown in Figure 12-62.

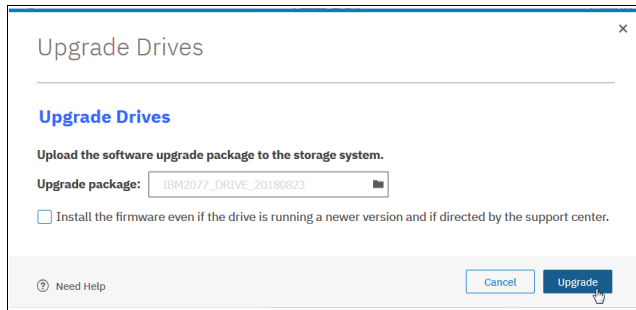


Figure 12-62 Individual drive update file selection

The drive firmware update takes about 2 - 3 minutes for each drive.

3. To verify the new firmware level, right-click the drive and select **Properties**, as shown in Figure 12-63.

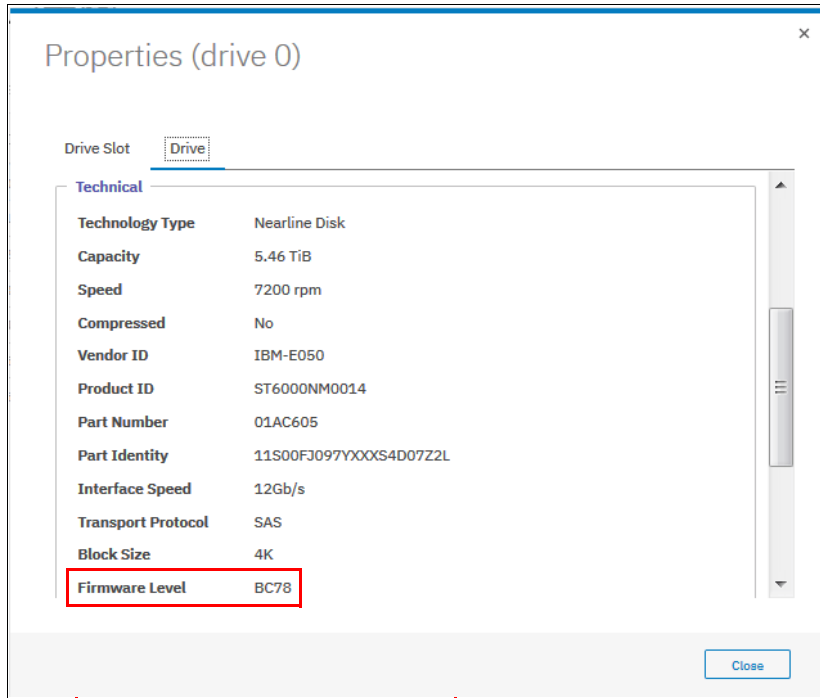


Figure 12-63 Individual drive update result

Updating the firmware on all drives

Complete the following steps to update all of the drives in an IBM Storwize V5000 Gen2 by using the management GUI:

1. Go to **Pools** → **Internal Storage**.

Figure 12-64 on page 668 shows how to update all drives through the Actions menu in the Internal Storage window.

2. Under Drive Class Filter, click **All Internal**.
3. In the Actions menu, click **Upgrade All**.

Note: If any drives are selected, the Actions menu displays actions for the selected drives and the Upgrade All option does not appear. If a drive is selected, clear it by pressing the Ctrl key and clicking the drive.

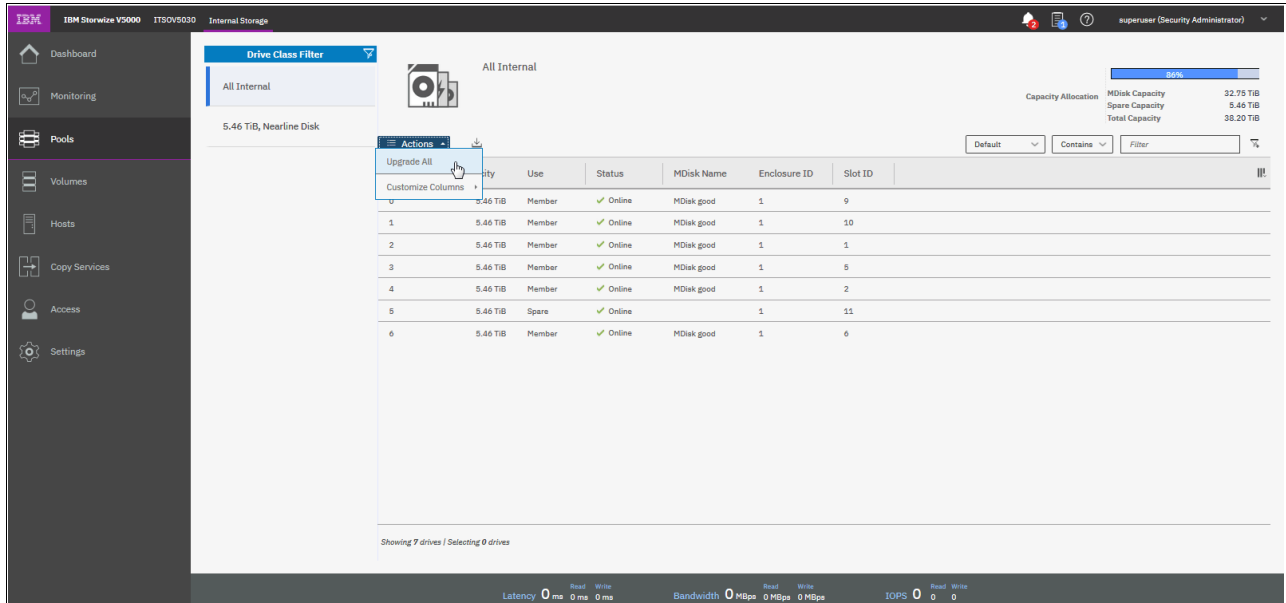


Figure 12-64 Update of multiple drives in the Internal Storage window

- After you start the drive upgrade process, the window that is shown in Figure 12-65 is displayed. Select the drive upgrade package, which was downloaded from the IBM Support site, by clicking the folder icon, and click **Upgrade**. You can also override newer versions of firmware when you select the option **Install the firmware even if the drive is running a newer version and if directed by the support center**.

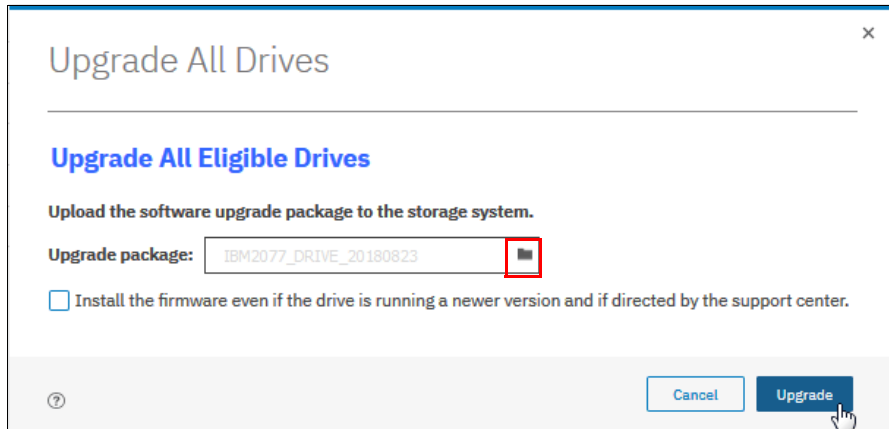


Figure 12-65 Upload the software upgrade package for multiple drives

All drives that require an update can now be updated.

12.5 Monitoring

Any issue that is reported by your IBM Storwize V5000 Gen2 system must be fixed as soon as possible. Therefore, it is important to configure the system to send automatic notifications when a new event is reported. You can select the type of event for which you want to be notified. For example, you can restrict notifications to only events that require immediate action.

Several event notification mechanisms are available:

- Email** Email notifications can be configured to send emails to one or more email addresses. With this mechanism, individuals can receive notifications wherever they have email access, including mobile devices.
- SNMP** SNMP notifications can be configured to send a Simple Network Management Protocol (SNMP) traps report to a data center management system that consolidates SNMP reports from multiple systems. With this mechanism, you can monitor your data center from a single workstation.
- Syslog** Syslog notifications can be configured to send a syslog report to a data center management system that consolidates syslog reports from multiple systems. With this mechanism, you can monitor your data center from a single location.

If your system is within warranty, or you have a hardware maintenance agreement, configure your IBM Storwize V5000 Gen2 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 an event is received, IBM automatically opens a problem report and, if appropriate, contacts you to verify whether replacement parts are required.

Important: If you set up Call Home to the IBM Support Center, ensure that the contact details that you configured are correct and kept up-to-date when personnel changes.

12.5.1 Email notifications and Call Home

The Call Home function of the IBM Storwize V5000 Gen2 uses the email notification mechanism to send emails to the specific IBM Support Center. To configure Call Home and other optional email addresses, complete the following steps:

1. Browse to **Settings** → **Support** → **Call Home** and select **Enable Notifications**, as shown in Figure 12-66.

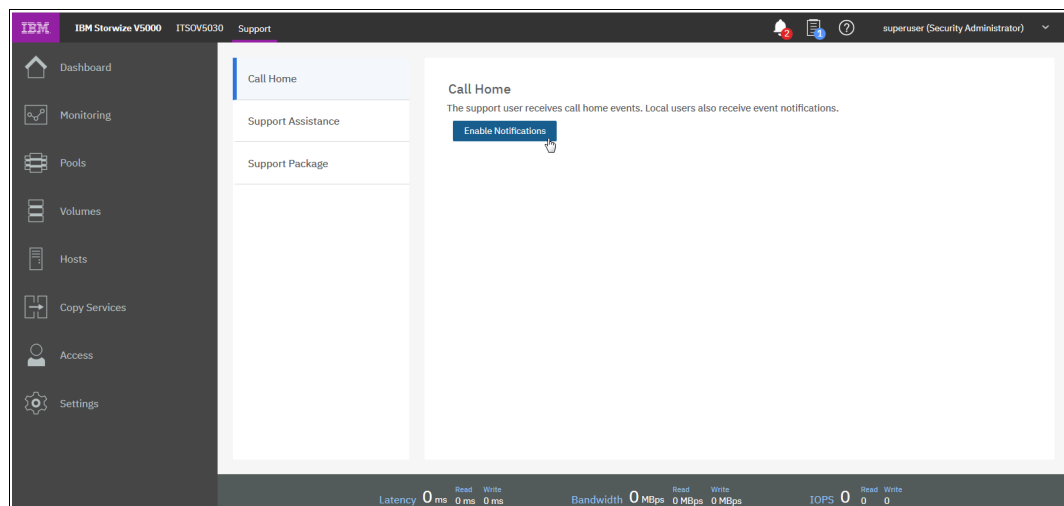


Figure 12-66 Enabling email notifications

- For the correct functionality of email notifications, ensure that Simple Mail Transfer Protocol (SMTP) is enabled on the management network and not, for example, blocked by firewalls.

If Email Notification is not enabled, you see a periodic warning that is similar to the warning that is shown in Figure 12-67.

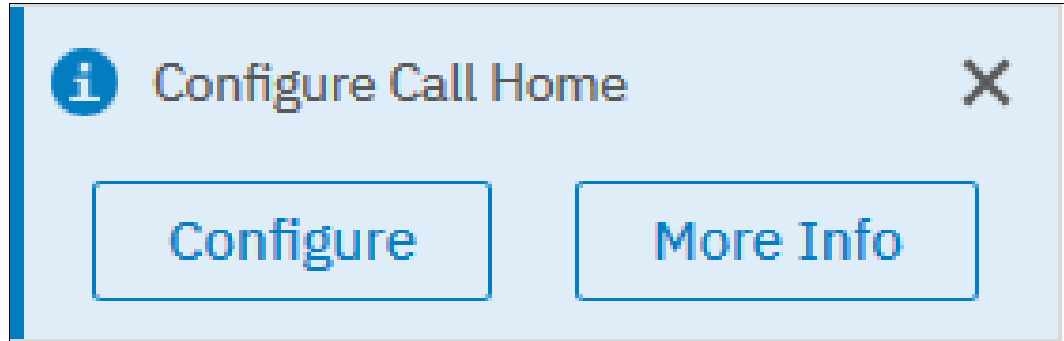


Figure 12-67 Configure Call Home information

If you start configuring the Call Home feature, you see the Welcome window that is shown in Figure 12-68.

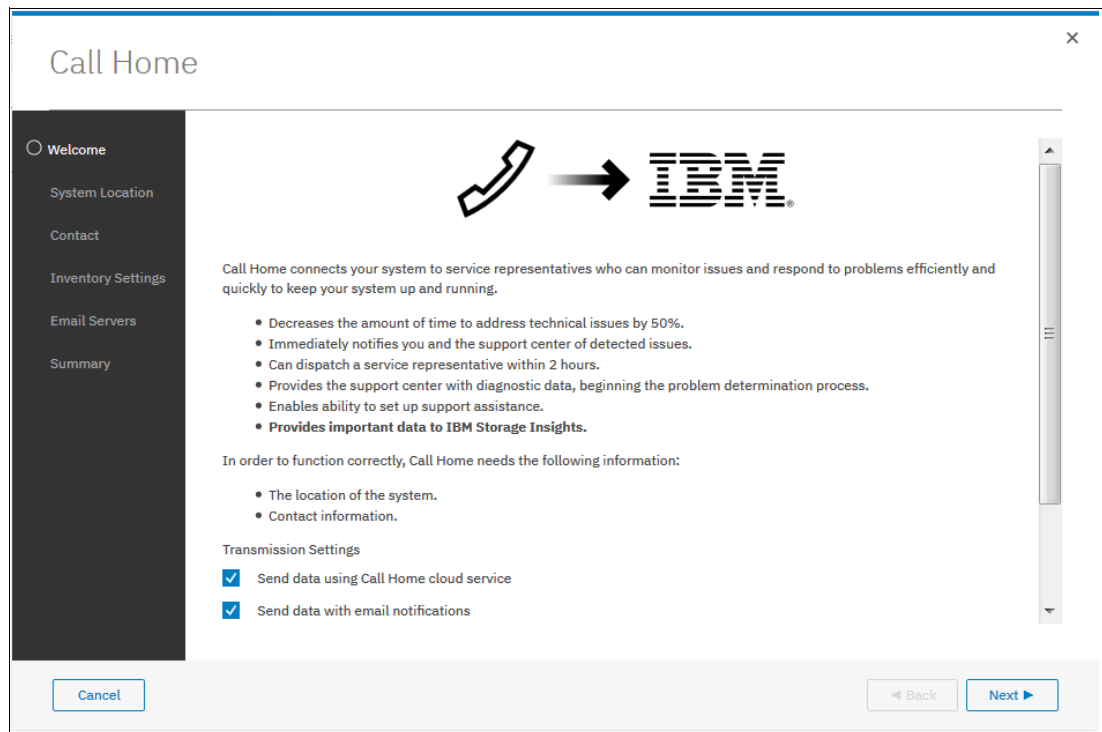


Figure 12-68 Call Home Welcome window

The following Transmission Settings are available:

- Send data using Call Home cloud service

Call Home with cloud services sends notifications directly to a centralized file repository that contains troubleshooting information that is gathered from customers. Support personnel can access this repository and be assigned issues automatically as problem reports. This method of transmitting notifications from the system to support removes the need for customers to create problem reports manually.

Call Home with cloud services also eliminates email filters dropping notifications to and from support which can delay resolution of problems on the system. Call Home with cloud services uses Representational State Transfer (RESTful) APIs, which are a standard for transmitting data through web services.

For new system installations, Call Home with cloud services is configured as the default method to transmit notifications to support. When you update the system software, Call Home with cloud services is also set up automatically. You need to ensure that network settings are configured to allow connections to IBM Support Center. This method sends notifications to the predefined support center only.

- Send data with email notifications

Call Home with email notification sends notifications through a local email server to support and local users or services that monitor activity on the system. With email notifications, you can send notifications to support and designate internal distribution of notifications as well, which alerts internal personnel of potential problems. Call Home with email notifications require configuration of at least one email server and local users.

Note: For new system installations, Call Home with cloud services is configured as the default method to transmit notifications to support.

Select the service that you prefer. Clicking **Next** checks the connection to IBM Support Center, as shown in Figure 12-69.

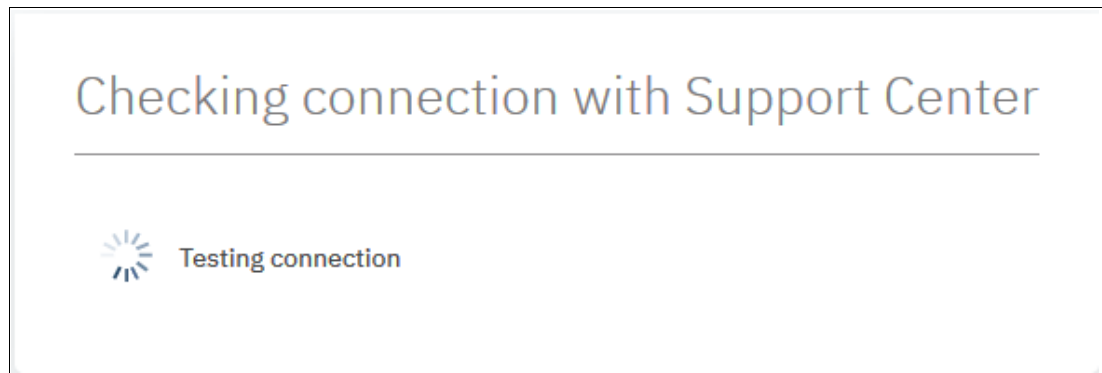


Figure 12-69 Checking connection with Support Center

3. You are taken to the next window (see Figure 12-70 on page 672) where you enter information regarding the System Location. This information is used by the support personnel to send the Support Representative to the failing system.

If only a minor problem exists, this information is used to send the CRU parts to the specific address. Ensure that you always keep this information updated. Figure 12-70 on page 672 shows how to enter the information.

Figure 12-70 Setup System Location

4. Clicking **Next** opens the Contact section. It is important to add an email contact who is responsible for this system. It is important to add an email contact who is responsible for this system. Provide the contact information of the system owner who can be contacted by the IBM Support Center when necessary. Ensure that you always keep this information updated. Figure 12-71 shows such an entry.

Figure 12-71 Contact information

In Figure 12-72, you can see a successful setup of the Call Home setup.

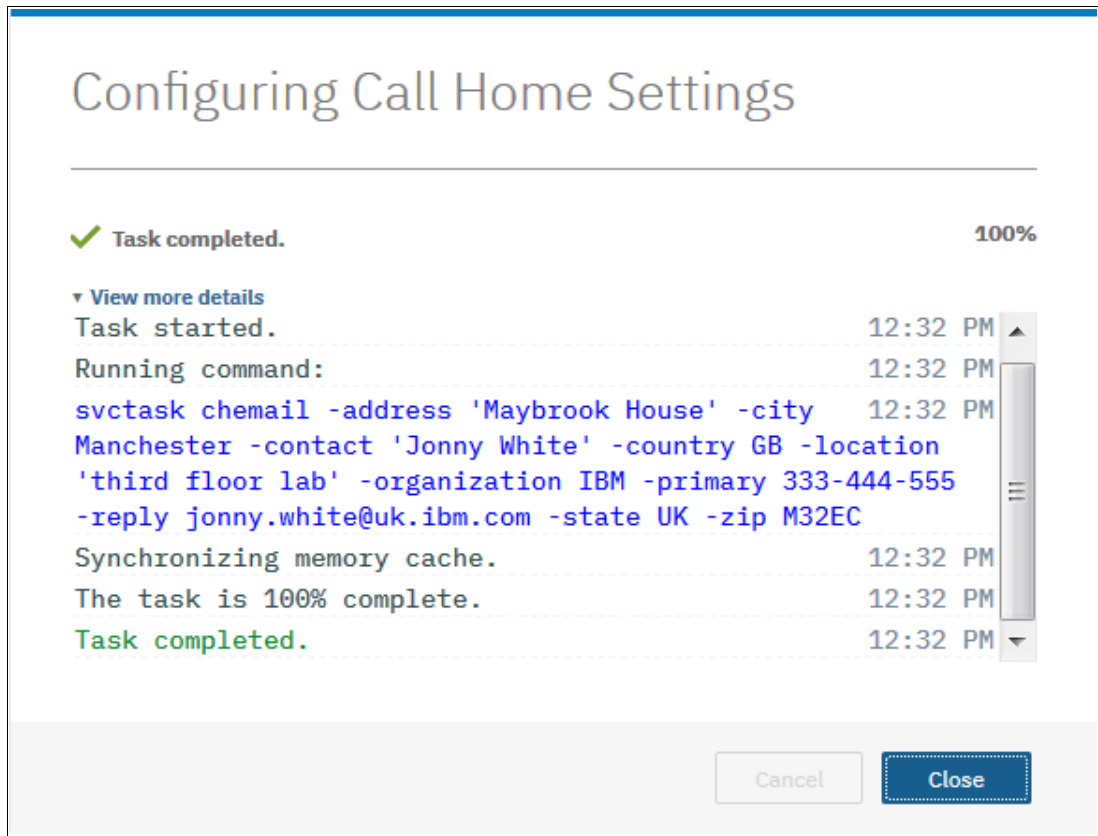


Figure 12-72 CLI of Call Home setup

The following optional settings are available, as shown in Figure 12-73:

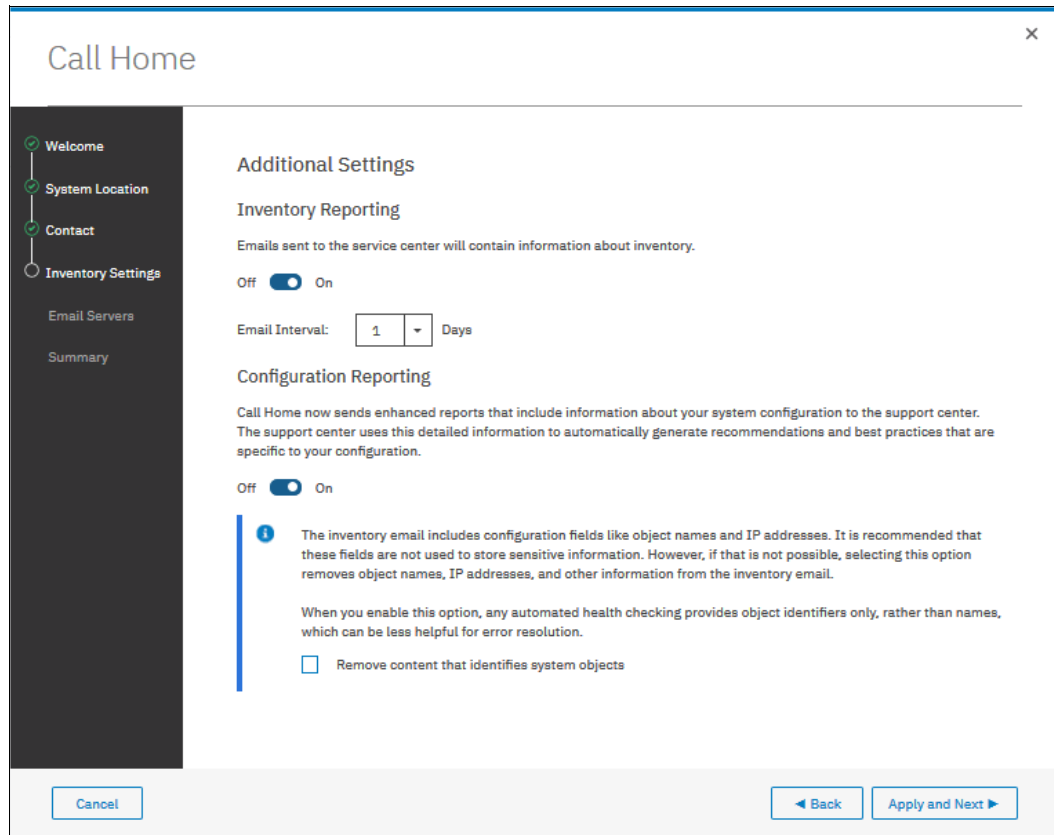


Figure 12-73 Additional Settings

- The first optional setting is to choose if you want to include an inventory file into your email to check the actual inventory of your system. Figure 12-73 shows the location where you can set the slider to indicate that you want to receive inventory details.

The emails include an inventory report that describes the system hardware and critical configuration information. Object names and other information, such as IP addresses, are not sent. Based on the information that is received, IBM can inform you whether the hardware or software that you are using requires an upgrade because of a known issue. Choose how often you want to get these emails.

- The second optional setting is to include inventory information. This email summarizes the hardware components and configuration of a system. Service personnel can use this information to contact you when relevant software updates are available or when an issue that can affect your configuration is discovered.

It is a good practice to enable inventory reporting. Because inventory information is sent by using the Call Home email function, you must meet the Call Home function requirements and enable the Call Home email function before you can attempt to send inventory information email.

You can adjust the contact information, adjust the frequency of inventory email, or manually send an inventory email using the management GUI or the command-line interface.

The Call Home function sends enhanced reports that include specific configuration information to IBM Support Center. IBM Support Center can use this information to automatically generate recommendations that are based on your configuration.

The inventory email includes the following information about the clustered system on which the Call Home function is enabled. You can select if sensitive information, such as IP addresses, is included or not:

- Licensing information
- Drives
- External storage systems
- Hosts
- MDisks
- Volumes
- Array types and levels
- Easy Tier
- FlashCopy
- Metro Mirror and Global Mirror
- HyperSwap

5. Configure the email servers. You can add several servers by clicking the plus (+) sign, as shown in Figure 12-74. Test the connection to the email server by clicking **Ping**.

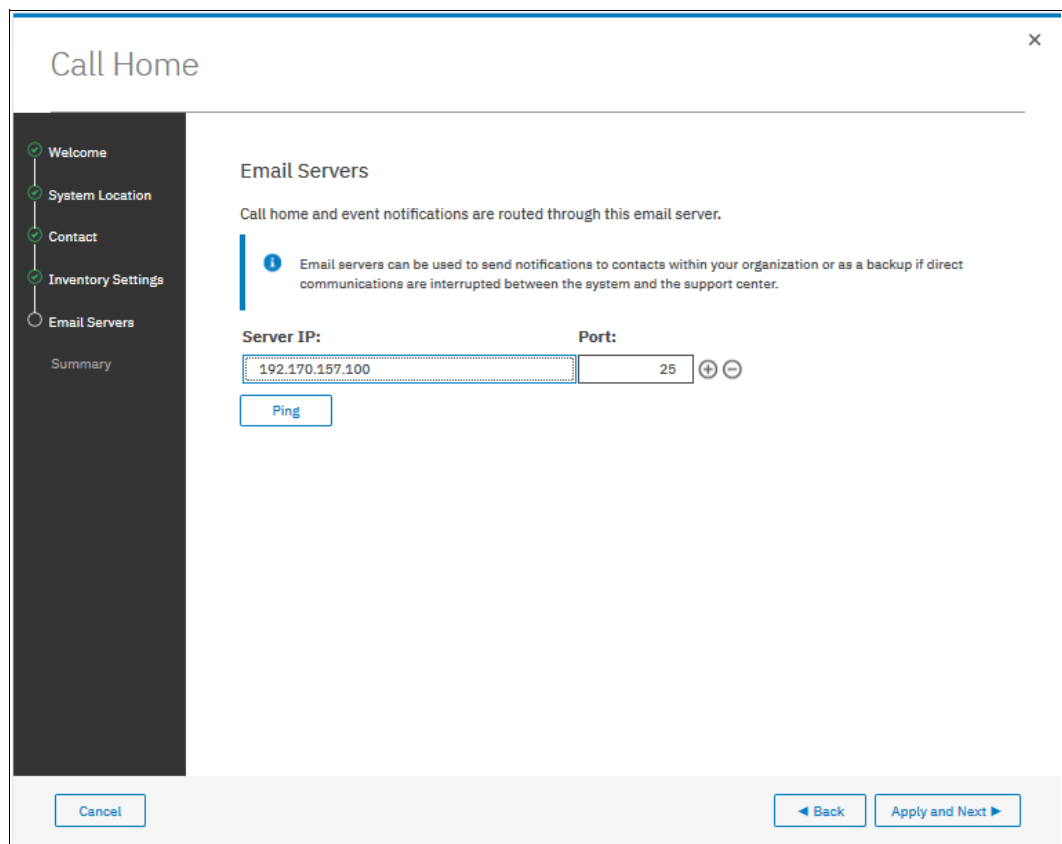


Figure 12-74 Email Servers

6. A summary overview is displayed, as shown in Figure 12-75. If all of the information is correct, click **Finish**. If changes must be made, correct your entries.

Call Home

Summary

Contact

Contact name: Jonny White
Email address: jonny.white@uk.ibm.com
Telephone (primary): 333-444-555
Telephone (alternate):

System Location

Company name: IBM
Street address: Maybrook House
City: Manchester
State or province: UK
Postal code: M32EC
Country or region: United Kingdom
Comment: third floor lab

Email Servers

Server IP: 192.170.157.100
Port: 25

Call Home

Transmission setting: Cloud, Email
Support center: callhome0@de.ibm.com
Alerts: Errors, Inventory
Inventory Reporting: On
Email Interval: Every day
Configuration Reporting: On
Remove Sensitive: Off
Information:

Cancel Back Finish

Figure 12-75 Summary of your settings for Call Home

- Click **Finish** to enable the Call Home function (see Figure 12-76).

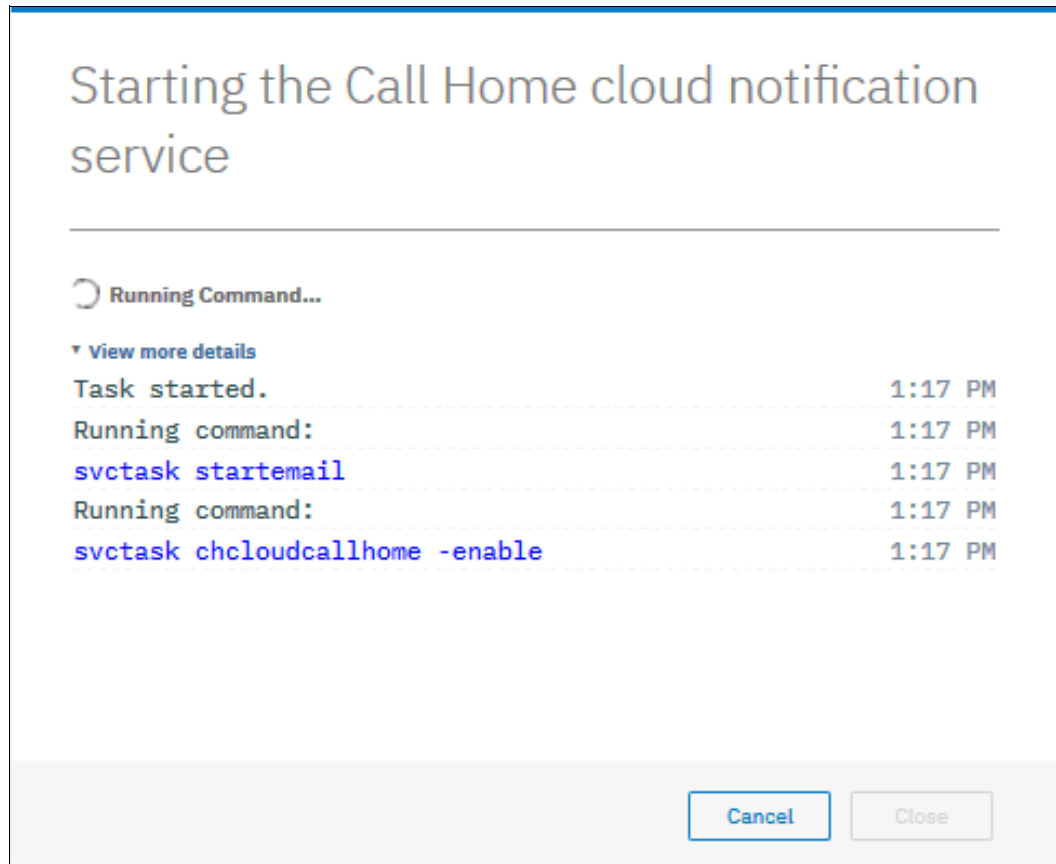


Figure 12-76 CLI to enable Call Home

Call Home is not set up. Figure 12-77 shows an overview of the Call Home settings.

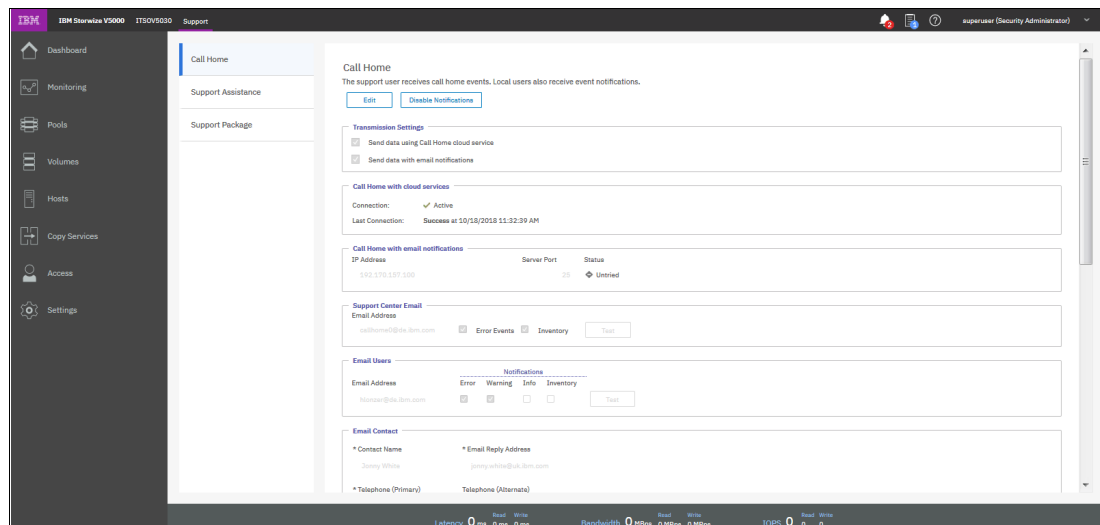


Figure 12-77 Overview of the Call Home settings

- After finalizing the setup, you can now add an email receiver, if needed. Click **Edit**, as shown in Figure 12-78.

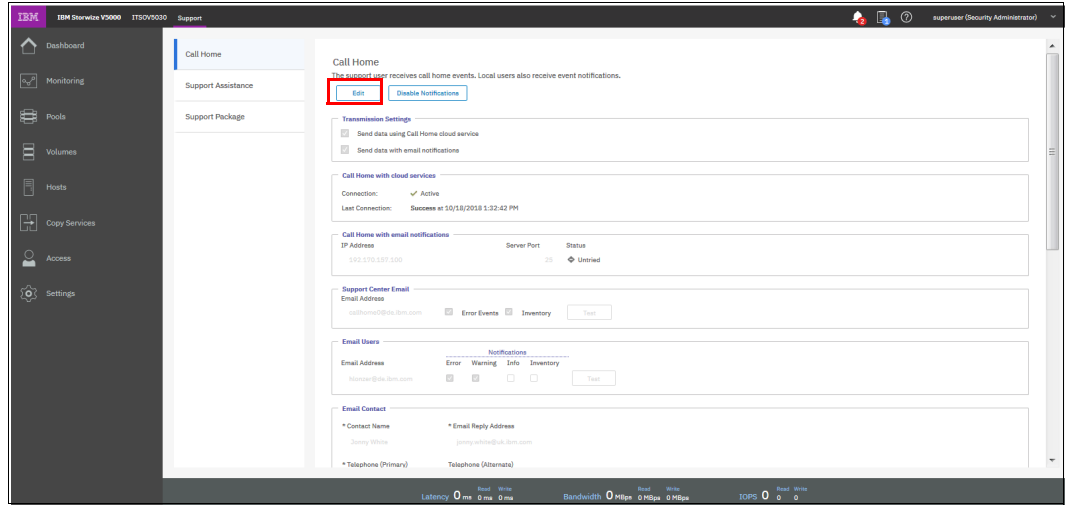


Figure 12-78 Edit Call Home

- In the Email Users section, click **+**, as shown in Figure 12-79.

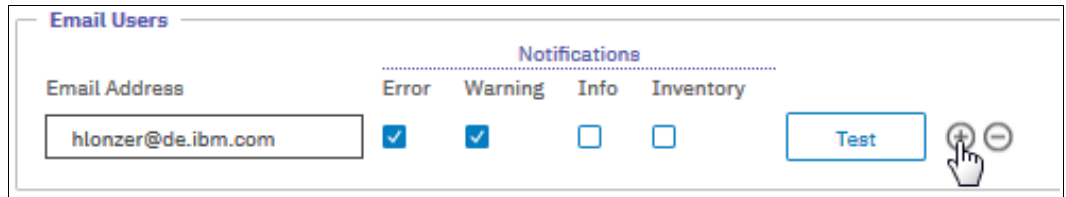


Figure 12-79 Add Email User

- Enter the email address of the new user (see Figure 12-80).

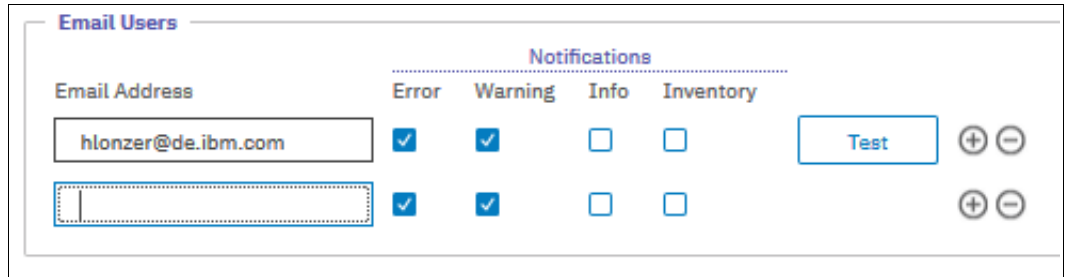


Figure 12-80 Add Email address

11. Save your changes (see Figure 12-81).

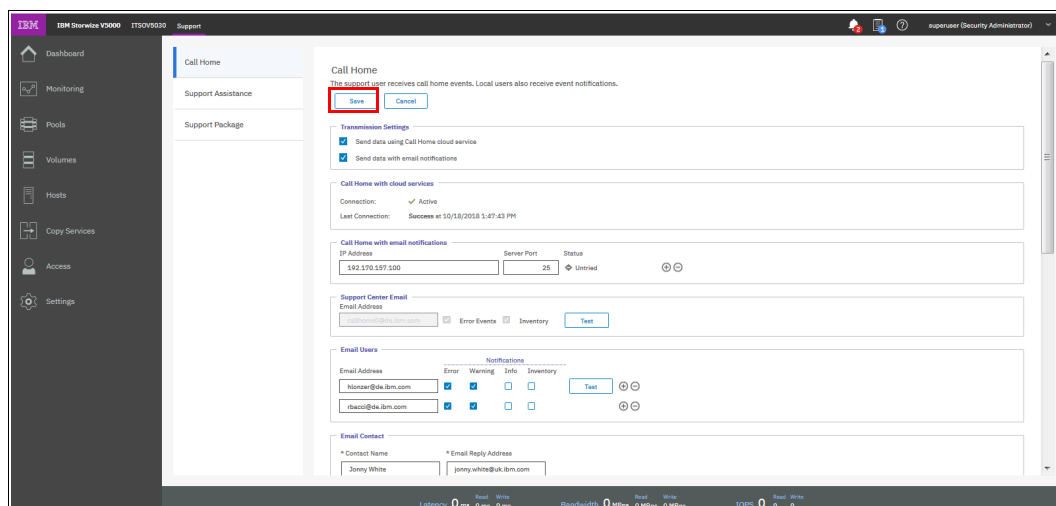


Figure 12-81 Save changes

Disabling and enabling notifications

Email notifications can be temporarily or permanently disabled at any time, as shown in Figure 12-82. Disabling email notifications is a preferred practice when you run maintenance tasks, such as upgrading code or replacing parts. After the maintenance operation, remember to re-enable the email notification function.

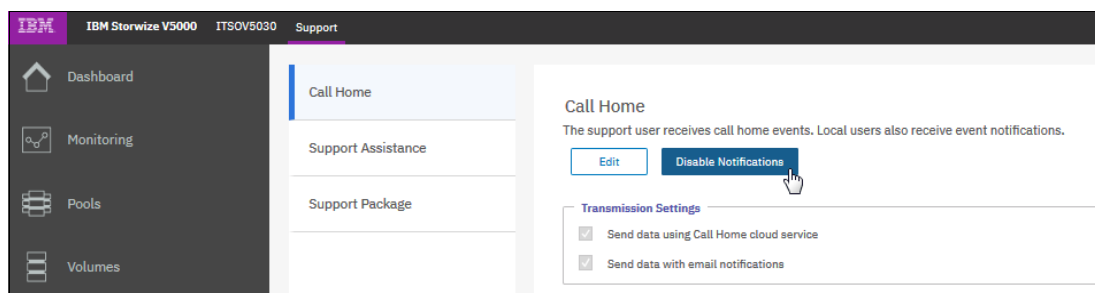


Figure 12-82 Disabling email notifications

The same results can be achieved by using the CLI and entering the `svctask stopemail` and `svctask startemail` commands.

12.6 Audit log

The *audit log* is useful when you analyze past configuration events, especially when you try to determine, for example, how a volume ended up being shared by two hosts, or why the volume was overwritten. The audit log is included in the support package to aid in problem determination.

The audit log tracks action commands that are issued through the CLI or the management GUI. It provides the following entries:

- ▶ Name 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

Failed commands and view commands are not logged in the audit log. Certain service commands also are not logged. The `svcconfig backup`, `cpdumps`, and `ping` service commands are not logged.

To access the audit log by using the GUI, browse to **Access** → **Audit Log**, as shown in Figure 12-83.

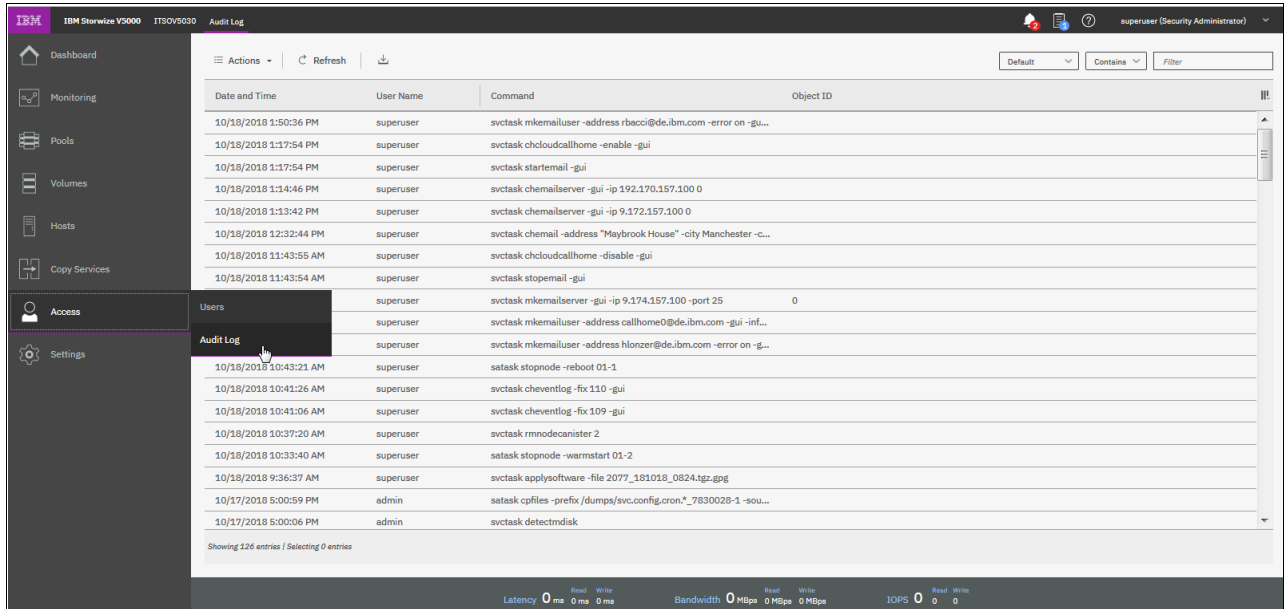


Figure 12-83 Audit log window

Right-clicking any column header opens the option menu in which you can select columns that are shown or hidden. It is also possible to click the Column icon on the far right of the column headers to open the option menu.

Figure 12-84 shows all of the possible columns that can be displayed in the audit log view.

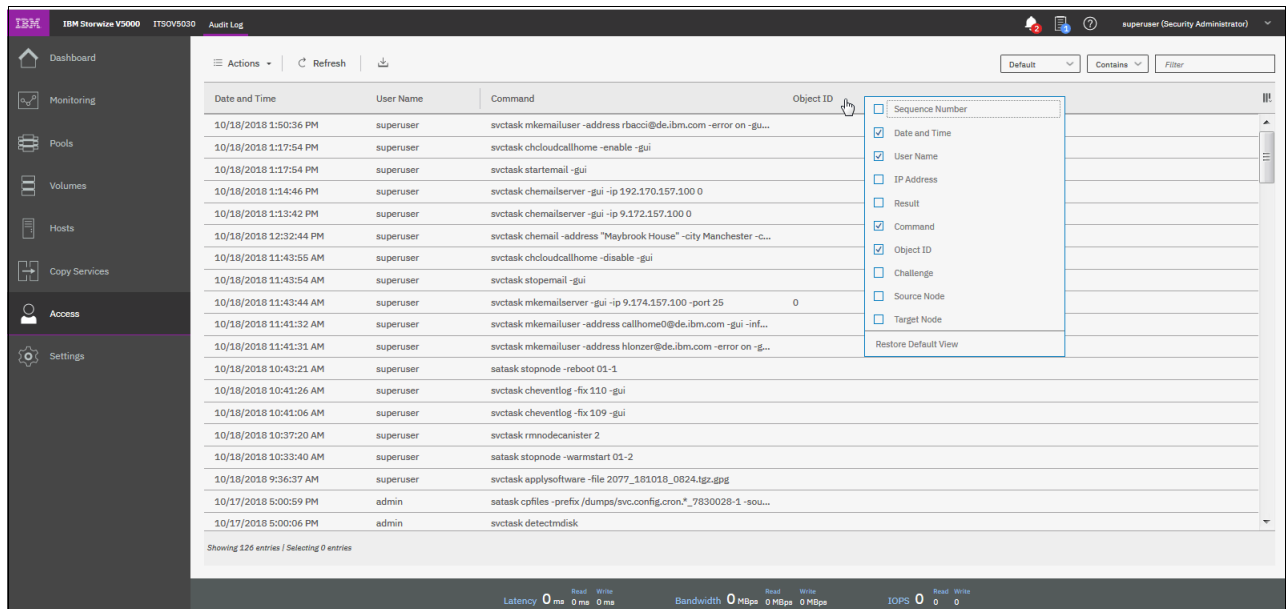


Figure 12-84 Possible audit log columns

12.7 Event log

Whenever a significant change in the status of the IBM Storwize V5000 Gen2 is detected, an event is submitted to the *event log*. All events are classified as *alerts* or *messages*.

An alert is logged when the event requires action. Certain alerts have an associated error code that defines the service action that is required. The service actions are automated through the fix procedures. If the alert does not have an error code, the alert represents an unexpected change in the state. This situation must be investigated to see whether it is expected or represents a failure. Investigate an alert and resolve it when it is reported.

A message is logged when a change that is expected is reported; for example, an IBM FlashCopy operation completes.

To check the event log, browse to **Monitoring** → **Events**, as shown in Figure 12-85.

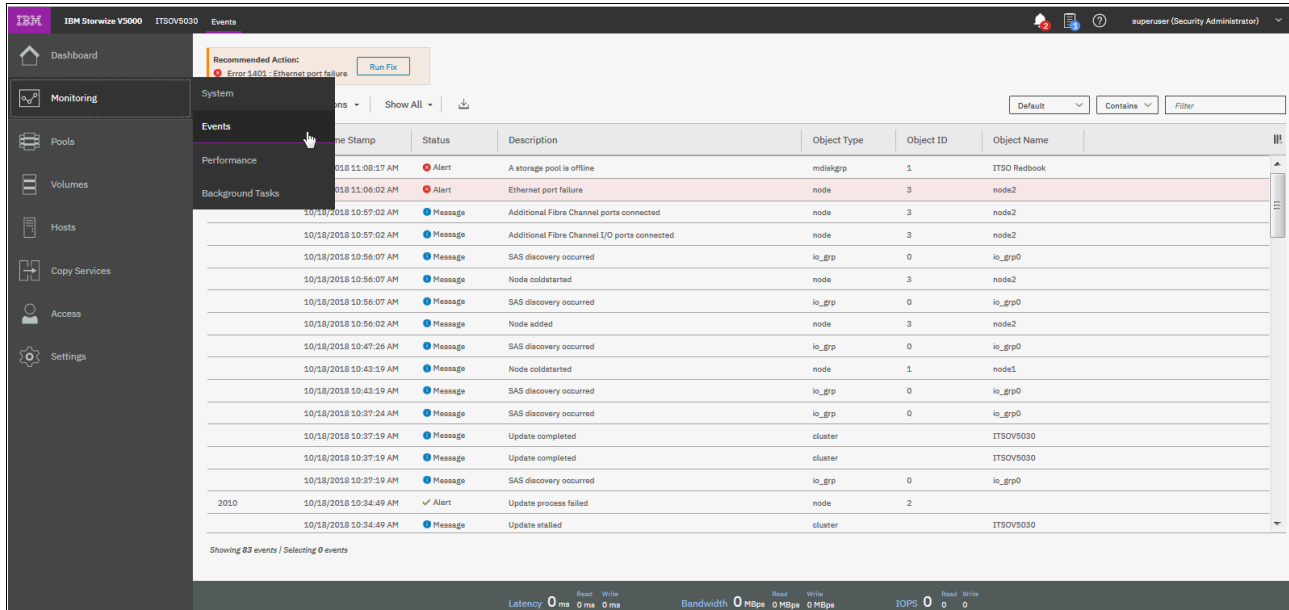


Figure 12-85 Event log

12.7.1 Managing the event log

The event log features a size limit. After the event log is full, newer entries replace the older entries, which are not required. To avoid a repeated event that fills the event log, certain records in the event log refer to multiple occurrences of the same event. When event log entries are coalesced in this way, the time stamp of the first occurrence of the problem and the time stamp the last occurrence of the problem are saved in the log entry. A count of the number of times that the error condition occurred is also saved in the log entry. Other data refers to the last occurrence of the event.

Event log window columns

Right-clicking any column header opens the option menu in which you can select columns that are shown or hidden. It is also possible to click the Column icon on the far right of the column headers to open the option menu.

Figure 12-86 shows all of the possible columns that can be displayed in the error log view.

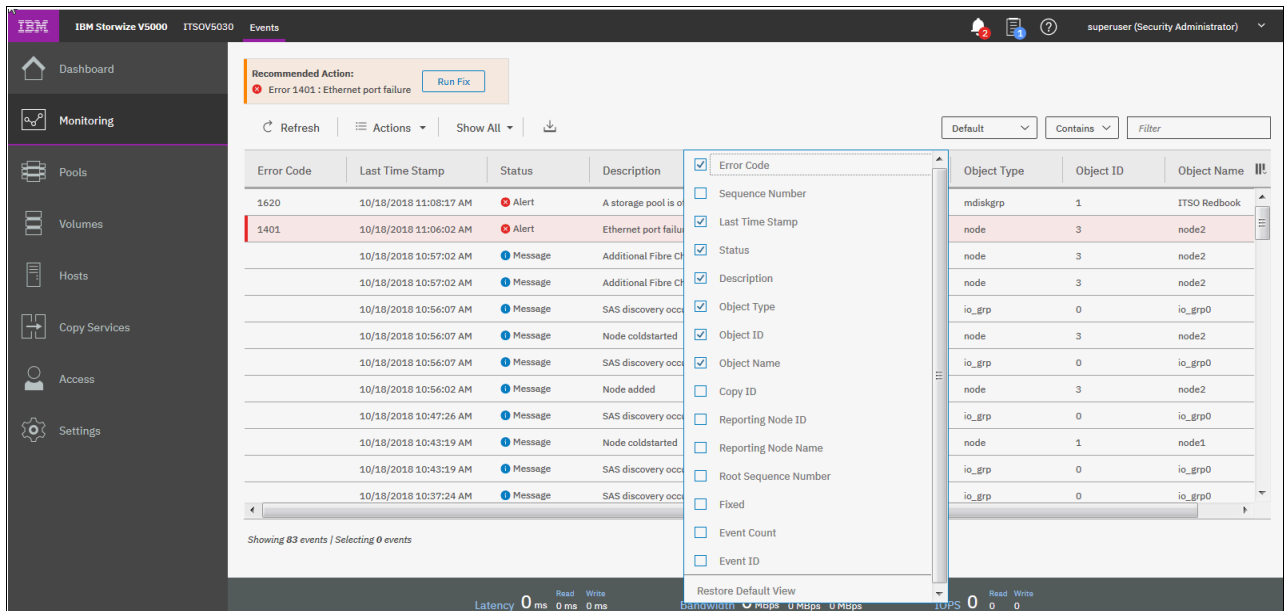


Figure 12-86 Possible event log columns

Event log filter options

The event log can be filtered by using the options that are shown in Figure 12-87.

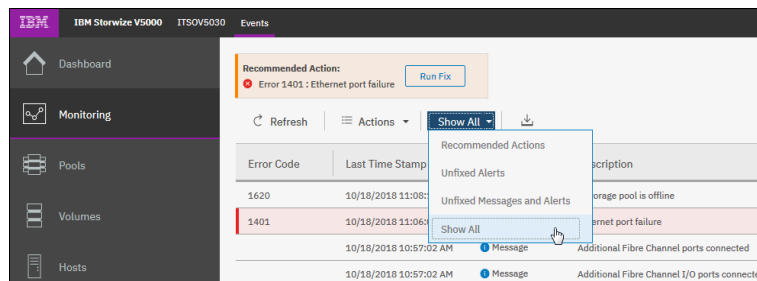


Figure 12-87 Event log filter options

The following options are available:

- ▶ **Recommended Actions (default)**

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

- Run a fix procedure
- View the properties

- ▶ **Unfixed Alerts**

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

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

► Unfixed Messages and Alerts

This option lists unfixed events. This option is useful to find events that must be handled, but no actions are required or recommended. For each entry that is selected, you can perform the following tasks:

- Run a fix procedure on any alert with an error code
- 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

This option lists all available events. For each entry that is selected, you can perform the following tasks:

- Run a fix procedure on any alert with an error code
- 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 specific 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.

You can also sort events by time or error code. When you sort by error code, the most serious events (those with the lowest numbers) are displayed first. You can select any event that is listed and select **Actions** → **Properties** to view details about the event.

Important: Check for this filter option if no event is listed. Events might exist that are not associated with recommended actions.

Figure 12-88 shows an event log with no items when the Recommended Actions filter was selected, which does not necessarily mean that the event log is clear. To check whether the log is clear, click **Show All**.

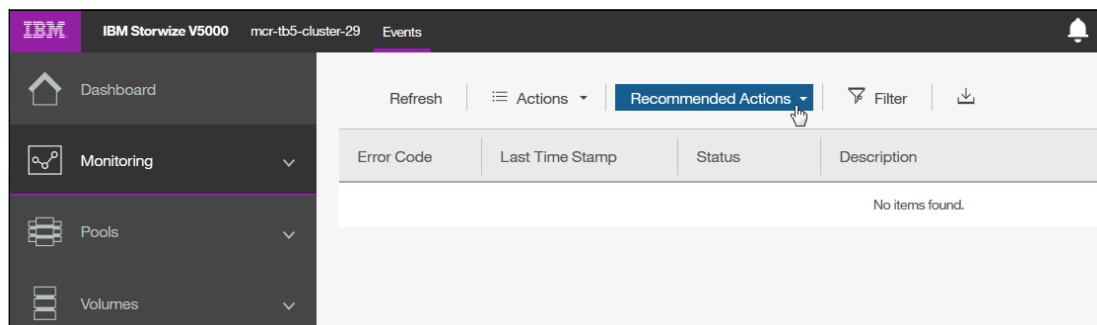


Figure 12-88 Event log with no recommended actions

Actions on a single event

Right-clicking a single event gives options that might be used for that specific event, as shown in Figure 12-89.

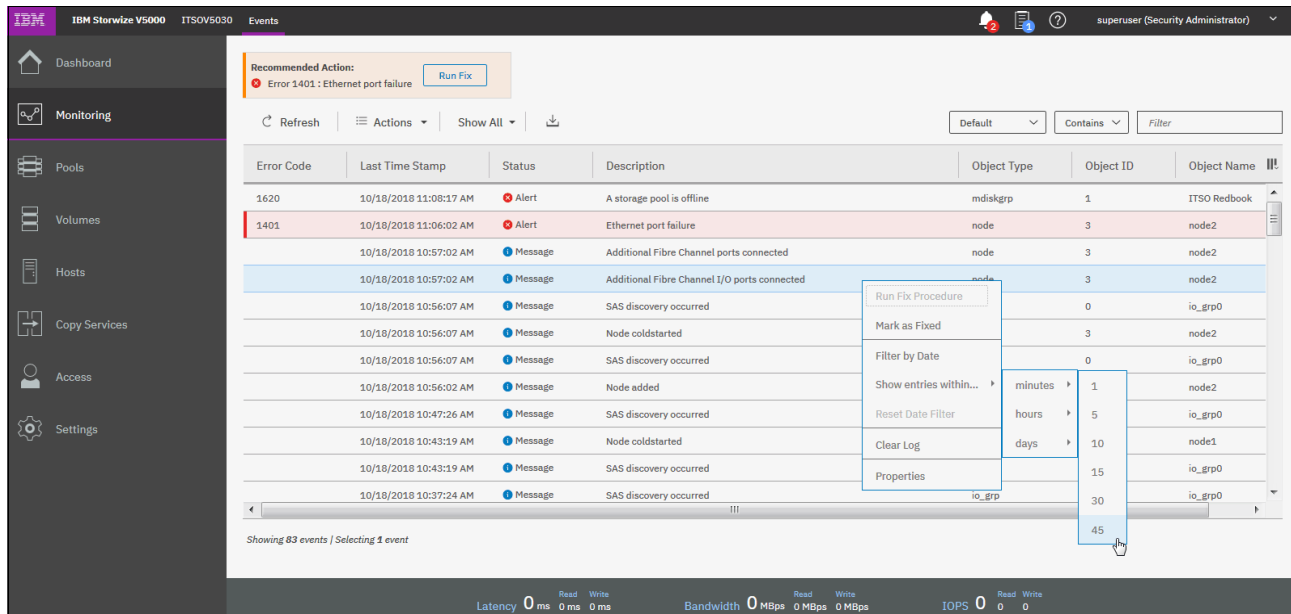


Figure 12-89 Possible actions on a single event

The following options are available:

- ▶ Run Fix Procedure

This option starts the fix procedure for this specific event. You can start a fix procedure, even if the procedure is not the recommended next action. However, we advise that you fix the error with the highest priority first.

- ▶ Mark as Fixed

This option marks this specific event as fixed. Message events must be marked as fixed to stop them from showing in the event log.

- ▶ Filter by Date

This option limits the event log entries to the events that occurred between an interval that is defined by the user.

- ▶ Show entries within (minutes/hours/days)

This option limits the event log entries to the events that occurred within the last period:

- 1, 5, 10, 15, 30, or 45 minutes
- 1, 2, 5, or 12 hours
- 1, 4, 7, 15, or 30 days

- ▶ Reset Date Filter

This option clears the Filter by Date.

- ▶ Clear Log

This option clears the complete event log, even if only one event was selected.

Important: These actions cannot be undone and might prevent the system from being analyzed when severe problems occur.

► Properties

This option provides more information for the selected event that is shown in the list.

Recommended actions

A fix procedure starts a wizard that is known as a Directed Maintenance Procedure (DMP) that helps to troubleshoot and correct the cause of an error. Certain DMPs reconfigure the system based on your responses, ensure that actions are carried out in the correct sequence, and prevent or mitigate the loss of data. For this reason, you must always run the fix procedure to fix an error, even if the fix might seem obvious.

To run the fix procedure for the error with the highest priority, go to the Recommended Action pane at the top of the Events page and click **Run Fix**, as shown in Figure 12-90. When you fix higher-priority events first, the system often can automatically mark lower-priority events as fixed.

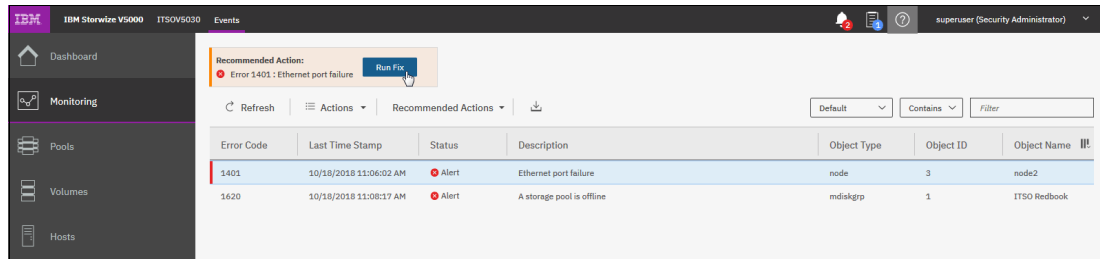


Figure 12-90 Next recommended action

12.7.2 Alert handling and recommended actions

All events that are in Alert status require attention. Alerts are listed in priority order. Alerts must be fixed sequentially by using the available fix procedures.

Example: Array mdisk not protected by sufficient spares

For example, look at an error that was raised by taking a drive offline in an array with redundancy of one.

This example can show how faults are represented in the error log, how information about the fault can be gathered, and how the Recommended Action (DMP) can be used to fix the error:

► Detecting the alert

The Health Status indicator shows a red alert. The Status Alerts indicator (on top of the GUI) shows two alerts. Click one alert's **Details** to retrieve the specific information, as shown in Figure 12-91.

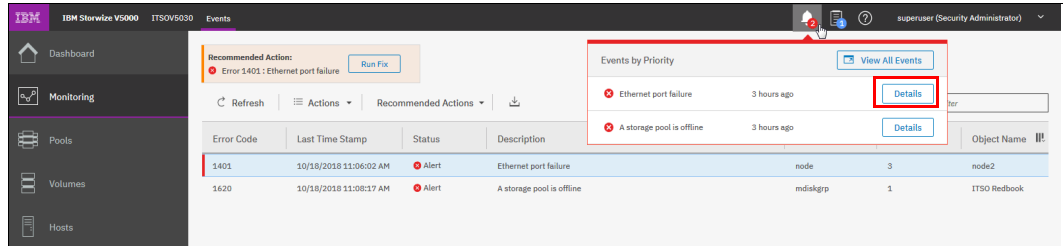


Figure 12-91 Status alert for an individual entry

Review the event log for more information.

► Gathering additional information

More information about the event is available by clicking the event and selecting **Details**. This information might help you fix a problem or analyze a root cause. Figure 12-92 shows the properties for the previous event.

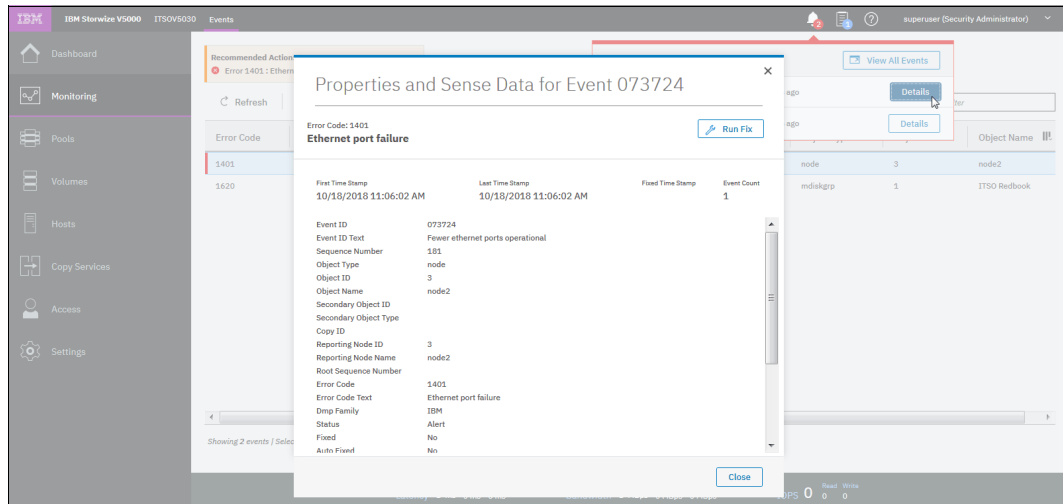


Figure 12-92 Alert properties

► Run the Recommended Action (DMP)

We recommend that you use the DMP to fix any alerts. You can miss tasks that are running in the background when you bypass the DMP. Not all alerts have available DMPs.

Figure 12-93 shows how to start the DMP by selecting **Run Fix** at the top of the window. This option always runs the recommended action.

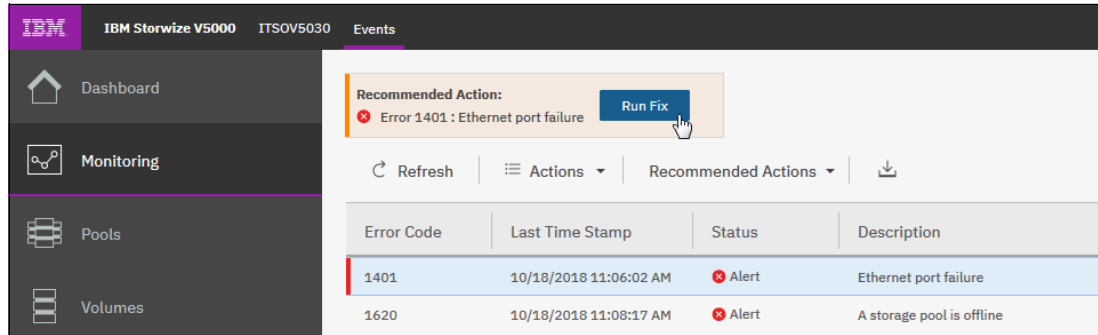


Figure 12-93 Starting the DMP (first option)

Figure 12-94 shows how to start the DMP by right-clicking the alert record and selecting **Run Fix Procedure**. You can use this option to run a fix procedure that might not be the recommended action.

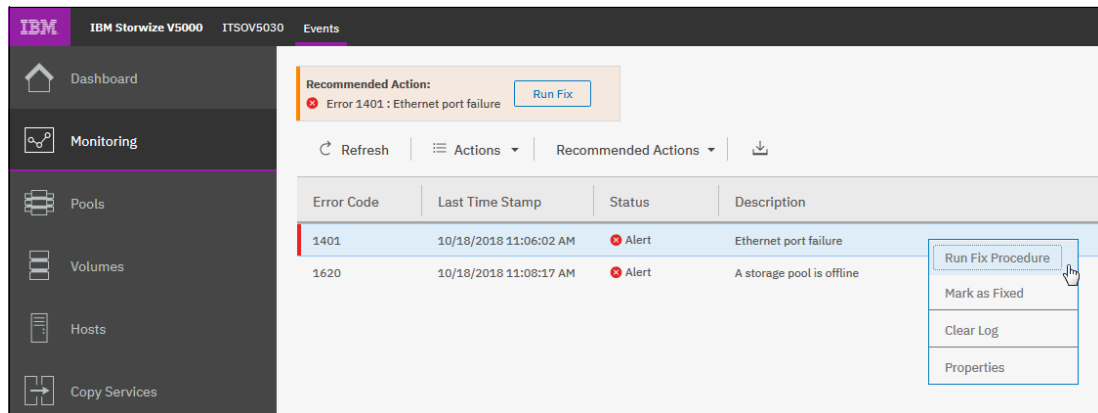


Figure 12-94 Starting the DMP (second option)

The third option that is shown in Figure 12-95 enables you to select the Run Fix Procedure from the Details window.

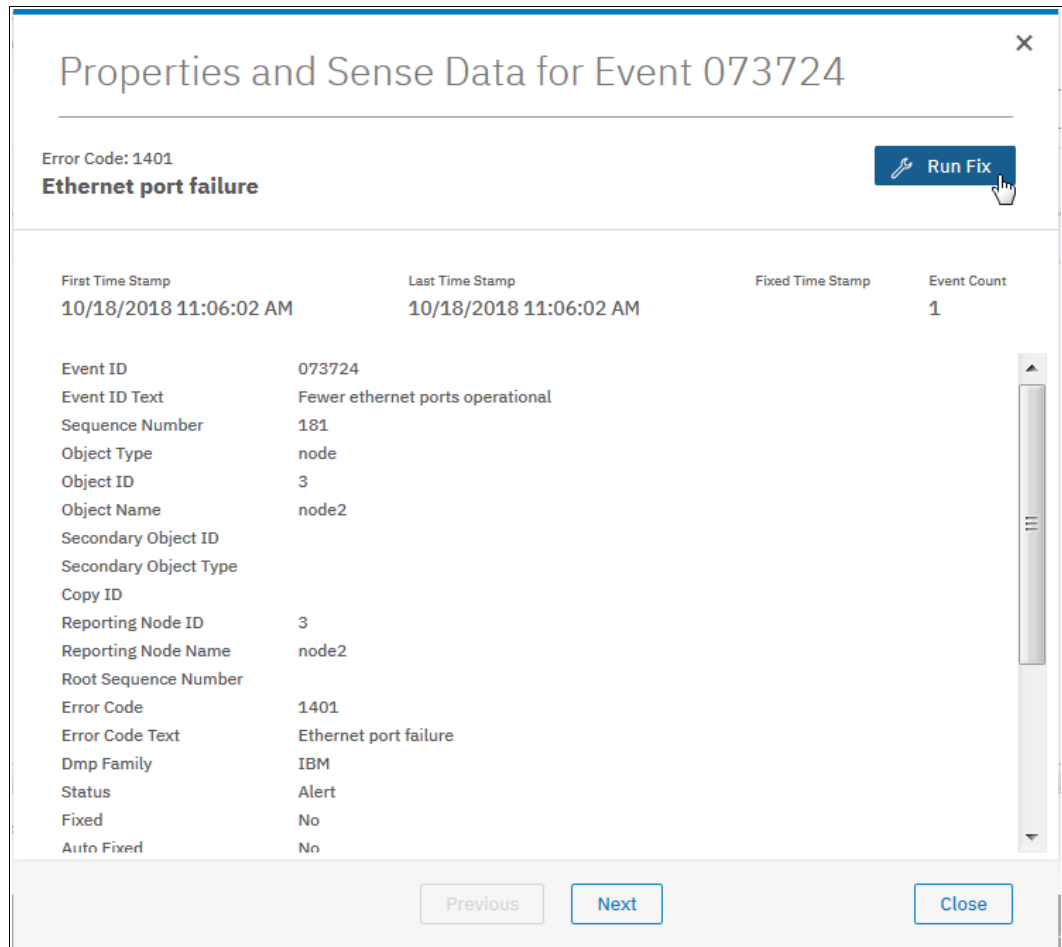


Figure 12-95 Starting the DMP (third option)

The steps and windows of a DMP are specific to the error. When all of the steps of the DMP are processed successfully, the recommended action is complete and the problem is usually fixed.

Figure 12-96 shows that the Health Status changed to green and both the Status Alerts indicator and the Recommended Action box disappeared, which implies that no other actions need to be taken.

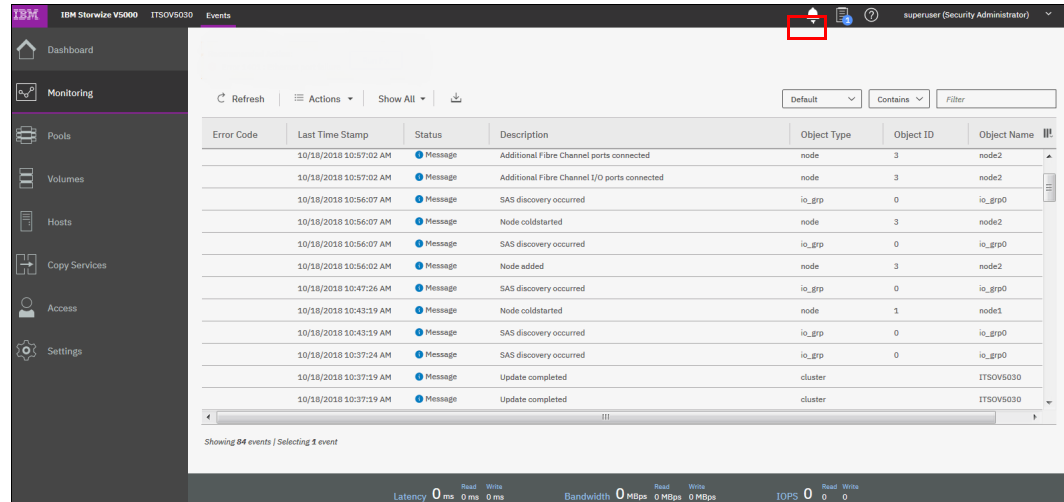


Figure 12-96 Event log with no outstanding recommended action

Handling multiple alerts

Figure 12-97 shows the event log with multiple alerts.

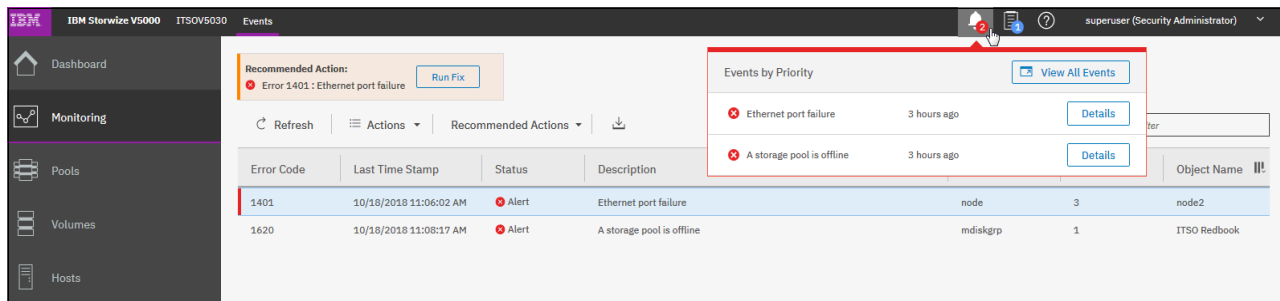


Figure 12-97 Multiple alert events that are displayed in the event log

The Recommended Action function orders the alerts by severity and displays the events with the highest severity first. If multiple events have the same severity, they are ordered by date and the oldest event is displayed first.

Events are ordered by severity. The first event is the most severe. Events are presented in the following severity order:

- ▶ Unfixed alerts (sorted by error code). The lowest error code has the highest severity.
- ▶ Unfixed messages.
- ▶ Monitoring events (sorted by error code). The lowest error code has the highest severity.
- ▶ Expired events.
- ▶ Fixed alerts and messages.

The less severe events are often fixed with the resolution of the most severe events.

12.8 Support assistance

Support assistance enables support personnel to access the system to complete troubleshooting and maintenance tasks. You can configure local support assistance, where support personnel visit your site to fix problems with the system, or remote support assistance.

Both local and remote support assistance use secure connections to protect data exchange between IBM Support Center and system. Access controls can be added by the system administrator. Assistance can be provided at your location or through a remote connection to your system.

Local support assistance

Use local support assistance if you have restrictions that require onsite support only. Unlike other authentication methods, you can audit all actions that support personnel conduct on the system when local support assistance is configured. Support personnel can log on to your system by using a console or over your intranet. These users can be authenticated only by a challenge-response mechanism.

Support personnel obtain the challenge-response access either through virtual private network (VPN) or over a telephone call with another support person or the administrator at IBM Support Center.

Note: If you want to enable remote support assistance or use the Assist On-Site tool, you must configure local support assistance.

Remote support assistance

With remote support assistance, support personnel can visit on site and they can also access the system remotely through a secure connection from IBM Support Center. However, before you enable remote support assistance between the system and support, you first need to configure local support assistance. You also must ensure that Call Home is configured and a valid email server is specified.

Call Home automatically contacts support when critical errors occur on the system. Call Home sends a return email that communicates information back to the system, such as a Problem Management Report (PMR) number that tracks the problem until it is resolved.

Note: You cannot enable remote support assistance and use the Assist On-Site tool at the same time.

In addition, a service IP address must be configured before you set up remote support assistance. During system initialization, you can optionally set up a service IP address and remote support assistance. If you did not configure a service IP address, go to **Settings** → **Network** → **Service IPs** to configure a service IP for each node on the system. Optionally, you must configure a proxy server if you use a firewall to protect your internal network.

When you enable remote support assistance, a shared-token is also generated by the system and sent to IBM Support Center. If the system needs support services, support personnel can be authenticated onto the system with a challenge-response mechanism. Use the **chsra** CLI command to enable remote support assistance on the system.

After support personnel obtain the response code, it is entered to gain access to the system. Service personnel have three attempts to enter the correct response code. After three failed attempts, the system generates a new random challenge and support personnel must obtain a new response code.

Support roles

When you enable local support assistance, support personnel are assigned either the Monitor role or the Restricted Administrator role. The Monitor role can view, collect, and monitor logs and errors to determine the solution to problems on the system. The Restricted Administrator role gives support personnel access to administrator tasks to help solve problems on the system. However, this role restricts these users from deleting volumes or pools, unmapping hosts, or creating, deleting, or changing users.

Roles limit access of the assigned user to specific tasks on the system. Users with the service role can set the time and date on the system, delete dump files, add and delete nodes, apply service, and shut down the system. They can also view objects and system configuration but cannot configure, modify, or manage the system or its resources. They also cannot read user data.

12.8.1 Configuring support assistance

You find the Support Assistance window under **Settings** → **Support** → **Support Assistance**, as shown in Figure 12-98.

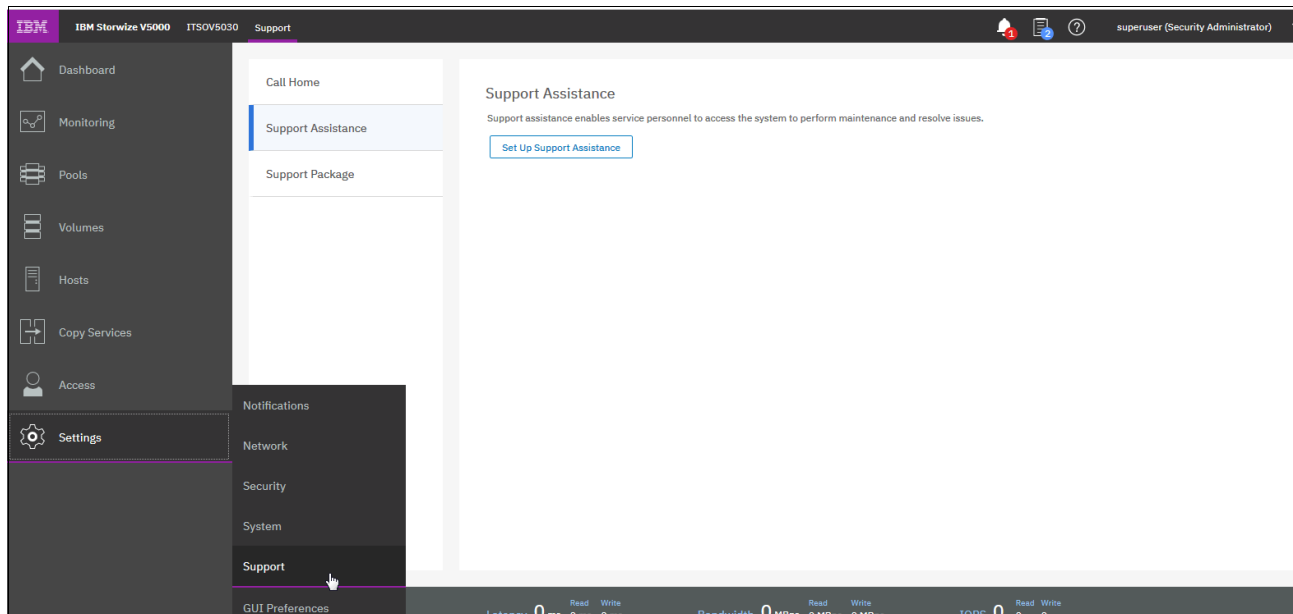


Figure 12-98 Support Assistance window

12.8.2 Setting up support assistance

You can configure either local or remote support assistance with the management GUI.

Use local support assistance if you have restrictions that require onsite support only. Unlike other authentication methods, you can audit all actions that support personnel conduct on the system when local support assistance is configured.

Support personnel can log on to your system by using a console or over your intranet. These users can be authenticated only by a challenge-response mechanism. Support personnel obtain the challenge-response access through a virtual private network (VPN) or over a telephone call with another support person or the administrator at IBM Support Center.

With remote support assistance, support personnel can visit on site and they can also access the system remotely through a secure connection from IBM Support Center. However, before you enable remote support assistance between the system and support, you first need to configure local support assistance.

With remote support assistance, support personnel can visit on site and they can also access the system remotely through a secure connection from IBM Support Center. However, before you enable remote support assistance between the system and support, you first need to configure local support assistance.

You must ensure that Call Home is configured and a valid email server is specified. Call Home automatically contacts support when critical errors occur on the system. Call Home sends a return email that communicates information back to the system such as a Problem Management Report (PMR) number that tracks the problem until it is resolved.

In addition, a service IP address must be configured before you set up remote support assistance. During system initialization, you can optionally set up a service IP address and remote support assistance. If you did not configure a service IP address, go to **Settings** → **Network** → **Service IPs** to configure a service IP for each node on the system.

Optionally, you need to configure a proxy server if you use a firewall to protect your internal network.

Prerequisites

If you are configuring remote support assistance, the following prerequisites must be met for all configurations:

- ▶ If your system can be configured for direct access to the internet, request that your network administrator allow these connections:
 - IP addresses 129.33.206.139 and 204.146.30.139 on port 22. These IP addresses and port are used to connect to support servers for remote support assistance.
 - IP addresses 129.42.56.189, 129.42.54.189, and 129.42.60.189 on port 443. These IP addresses and port are used to upload support packages to support from the system.
 - IP addresses 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 on port 22. These IP addresses and port are used to download software to the system from support.
- ▶ A DNS server is defined on your system. A DNS is required to upload support packages and to download software.

If your system is not configured to directly access the internet, or if you want to route traffic from multiple storage systems to the same place, you must configure a Remote Support Proxy server. You cannot upload support packages or download software with a Remote Support Proxy server.

To configure support assistance, complete the following steps:

1. In the management GUI, select **Settings** → **Support** → **Support Assistance** → **Set Up Support Assistance** (see Figure 12-99).

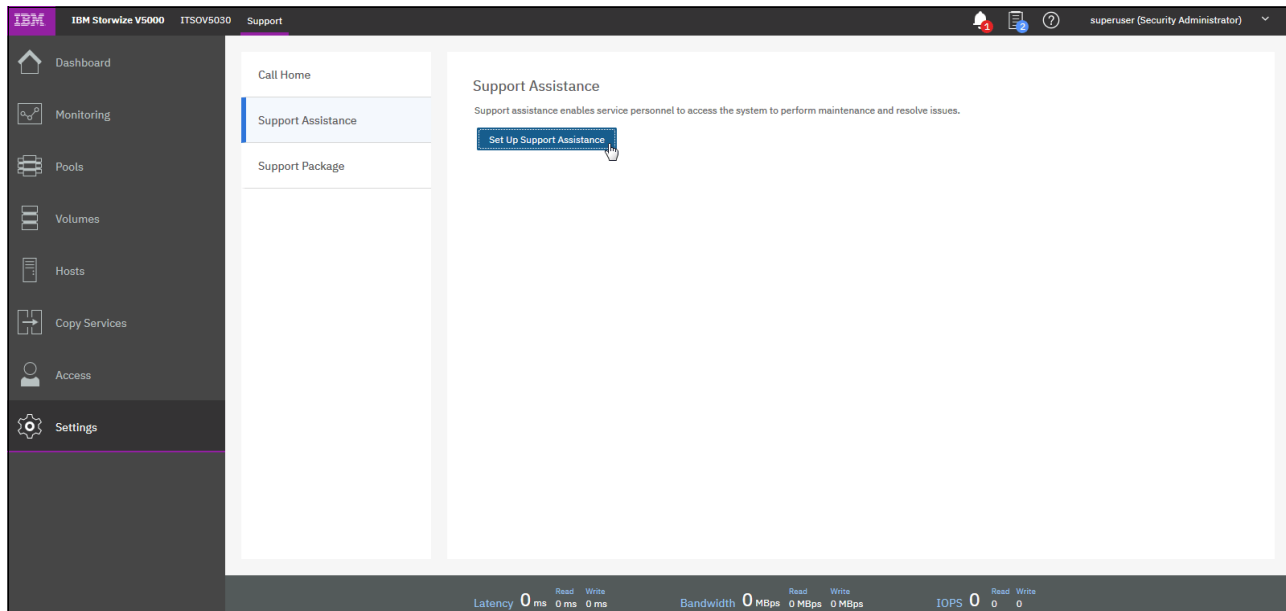


Figure 12-99 Support Assistance

2. If you selected to configure both local and remote support assistance, verify the pre-configured support centers.

Optionally, enter the name, IP address, and port for the proxy server on the Remote Support Centers page. A proxy server is used in systems where a firewall is used to protect your internal network or if you want to route traffic from multiple storage systems to the same place

Enable local support

To enable local support, complete the following steps:

1. Select the **I want support personnel to work on-site only** option. Figure 12-100 shows how to enable local support.

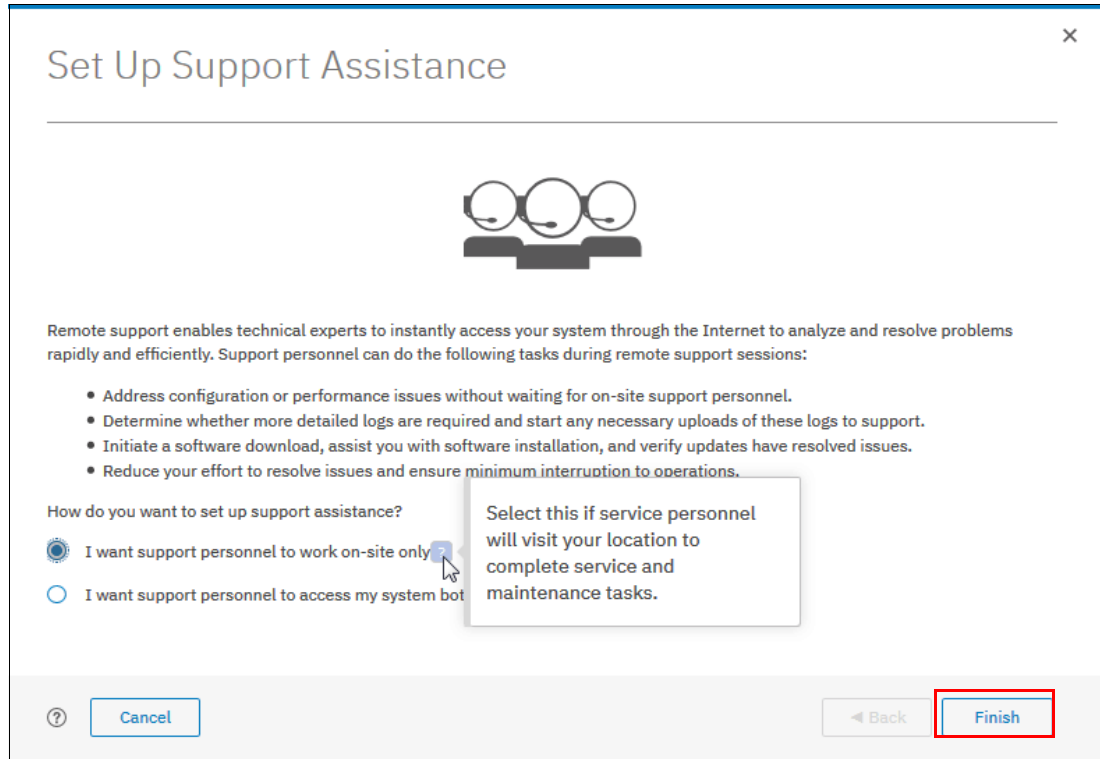


Figure 12-100 Enable local support

2. Select this option to configure local support assistance. Use this option if your system includes certain restrictions that require onsite maintenance. If you select this option, click **Finish** to finalize setting up local support assistance.

Enable remote and local support

To enable remote and local support select the **I want support personnel to access my system both on-site and remotely** option (see Figure 12-101).

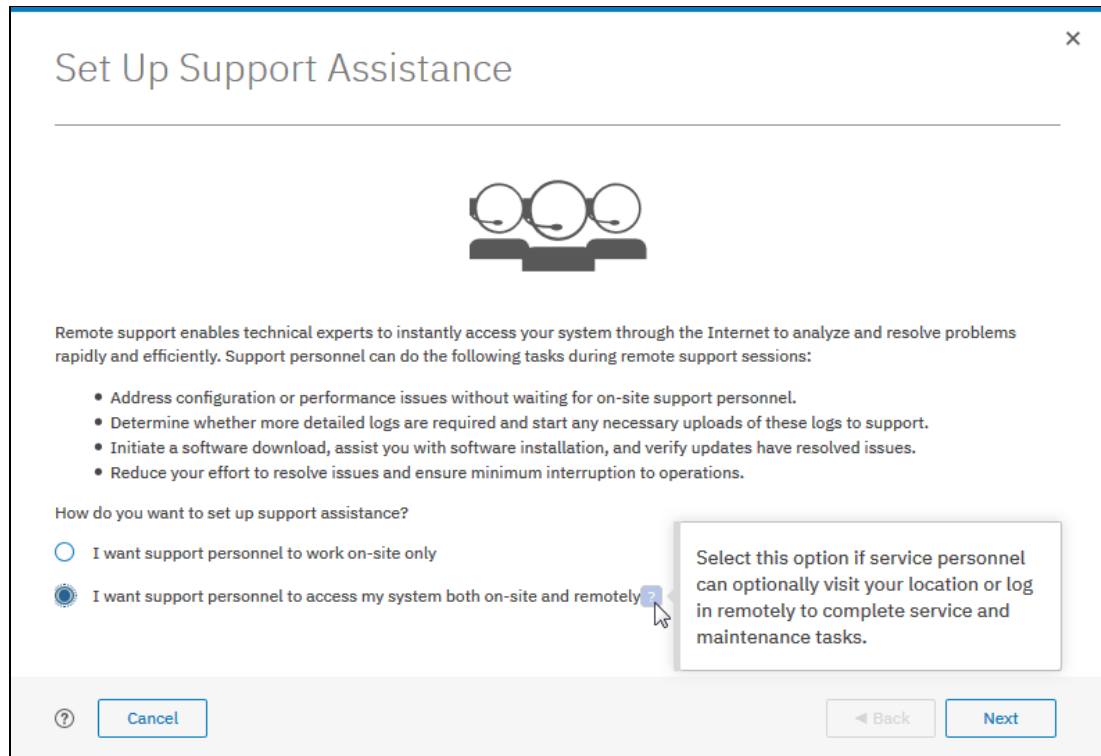


Figure 12-101 Remote and local support

Click **Next**. The window that is shown in Figure 12-102 on page 697 opens. Consider the following points about this window:

- ▶ Ensure that Call Home is configured with a valid email server.
- ▶ Ensure that a valid service IP address is configured on each node on the system.
- ▶ If your 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 set up process for support assistance, specify the IP address and the port number for the proxy server on the Remote Support Centers page.
- ▶ If you do not have firewall restrictions and the storage nodes 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.
- ▶ Both uploading support packages and downloading software require direct connections to the internet. A DNS server must be defined on your system for both of these functions to work.
- ▶ 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.
- ▶ 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.

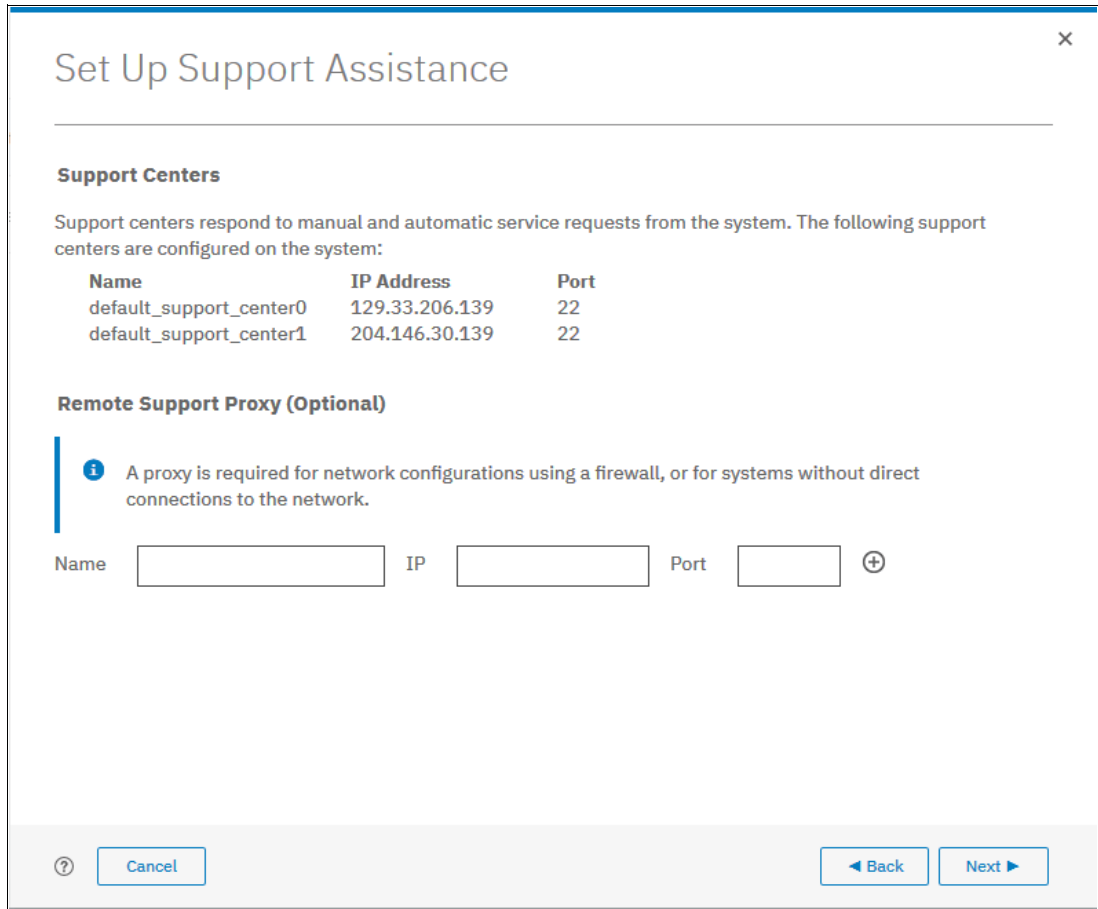


Figure 12-102 Support Centers

The IPs for the default Support Centers are given. Check your network to determine whether you can receive the IP numbers. Optional, you can add a Remote Support Proxy.

Remote Support Assistance for the system requires TCP/IP communication between the system and IBM Support Center. When a system does not have direct access to the internet (for example, because of a firewall), you can use the Remote Support Proxy to facilitate that connection.

The Remote Support Proxy utility creates a network proxy that connects one or more systems to remote support servers that are at IBM Support Center. It establishes a service on a Linux system that has internet connectivity to IBM Support Center and local network connectivity to the system.

When remote support assistance is configured on the system, the connection to IBM Support Center is started by the system through the management GUI. If a proxy server is necessary, it must be installed and configured, and defined in the management GUI. For the complete and up-to-date information about the compatibility and requirements of the Remote Support Proxy utility, refer to its latest release notes.

The communication between the system and the Remote Support Proxy uses Secure Shell (SSH). The communication between the Remote Support Proxy and IBM Support Center is encrypted with an extra layer of Secure Sockets Layer (SSL).

Set Up Support Assistance

Remote Support Access Settings

When do you want service personnel to complete maintenance and service tasks remotely? You can change these settings at any time.

At Any Time
The support center can start remote support sessions any time

On Permission Only
The support center can start a remote support session only if permitted by an admin. A time limit can be configured for the session.

? Cancel Back Finish

Figure 12-103 Allowed Service time

The following Remote Support Access settings are available (see Figure 12-103):

- ▶ **At Any Time**
Support personnel can access the system at any time. For this option, remote support session does not need to be started manually and sessions remain open continuously.
- ▶ **On Permission Only**
The system administrator must grant permission to support personnel before they can access the system (see Figure 12-104 on page 699).

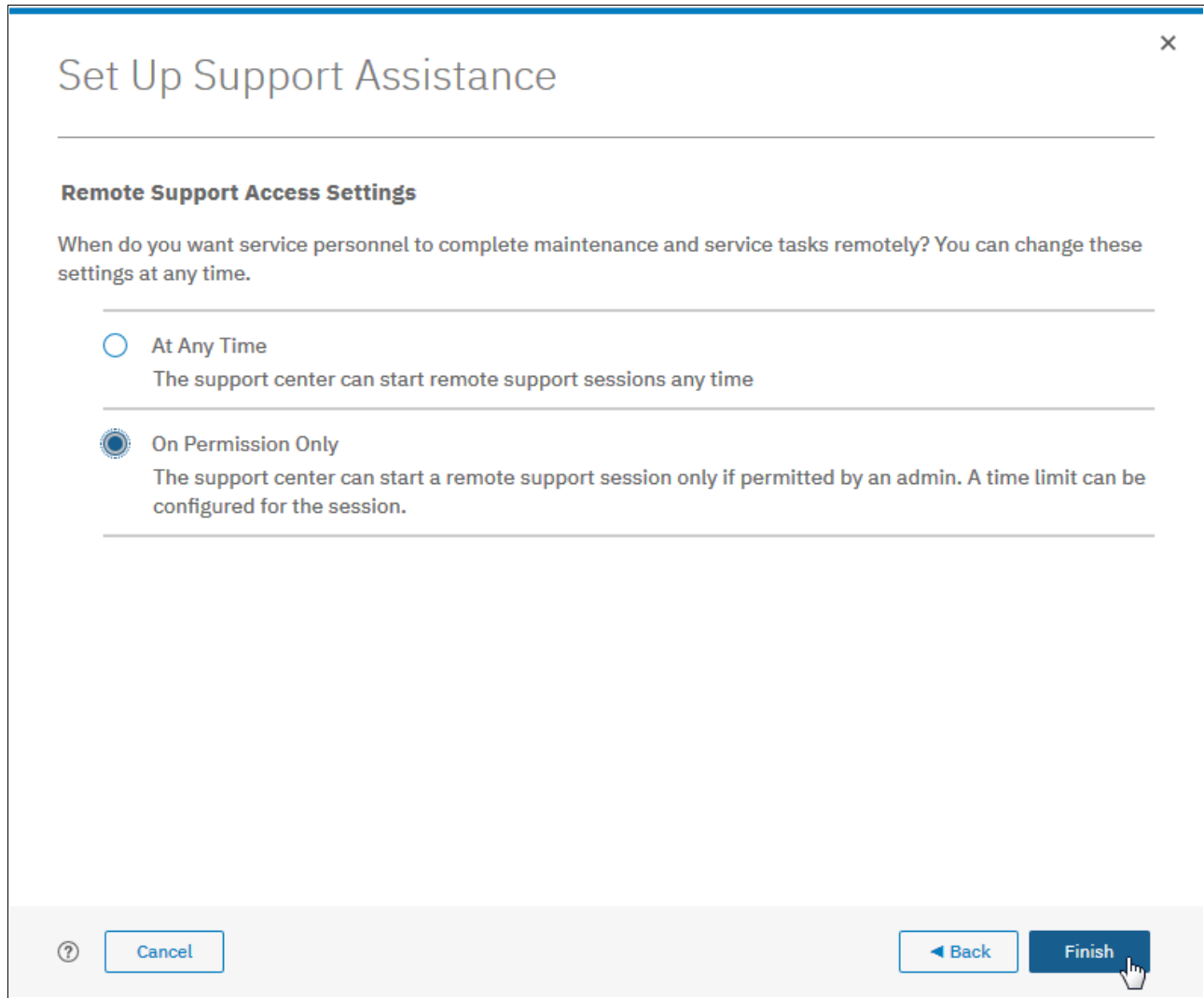


Figure 12-104 Remote Support Access Settings

3. Click **Finish**. After you configure remote support assistance with permission only, you can start sessions between IBM Support Center and the system.

As shown in Figure 12-105 the system is ready to start a new remote session.

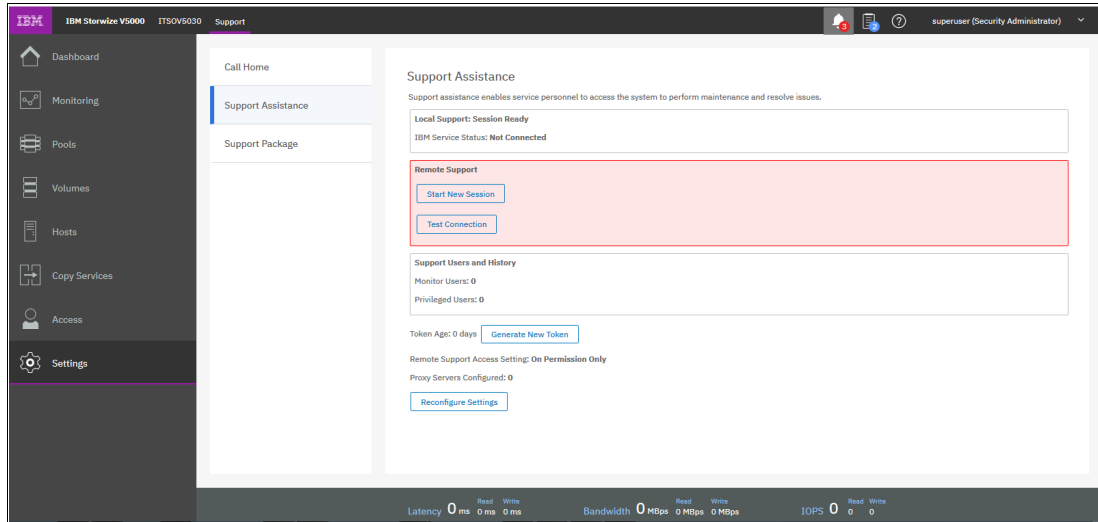


Figure 12-105 Ready to start a new session

4. Under **Support assistance** → **Start new Session** (marked) you can select the time that the Remote support Session can be idle before the system disconnects the line (see Figure 12-106).

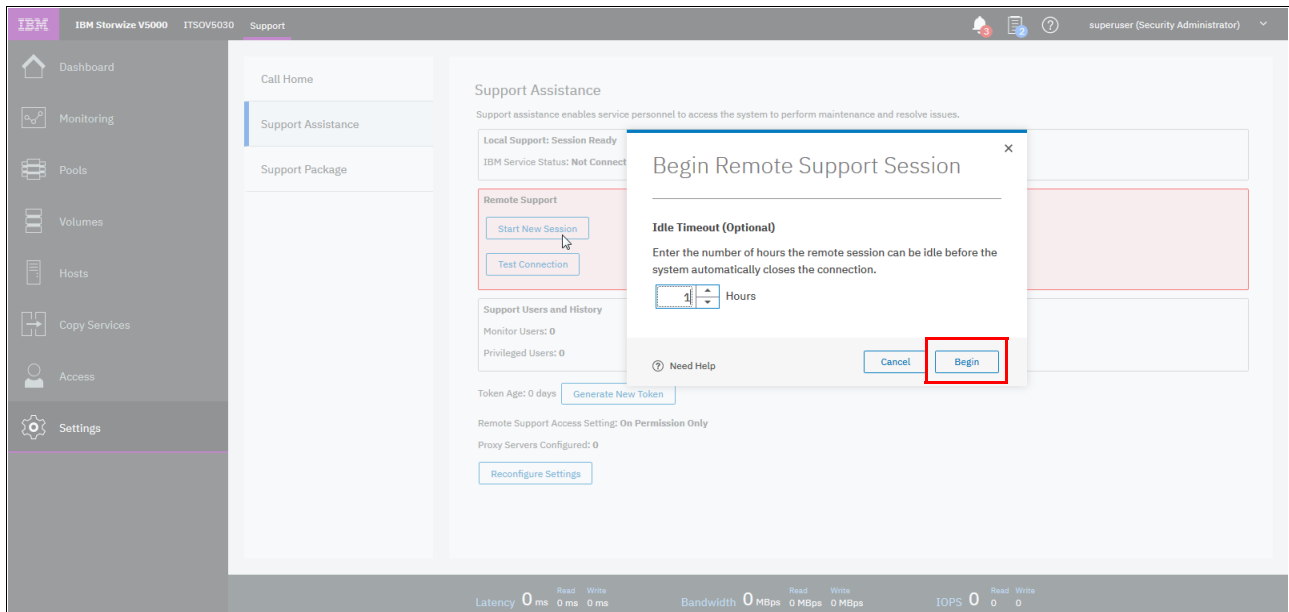


Figure 12-106 Set Idle time before the line will be disconnected

5. Click **Begin** to begin a new session.

6. Select **Test Connection** to test the connectivity, as shown in Figure 12-107.

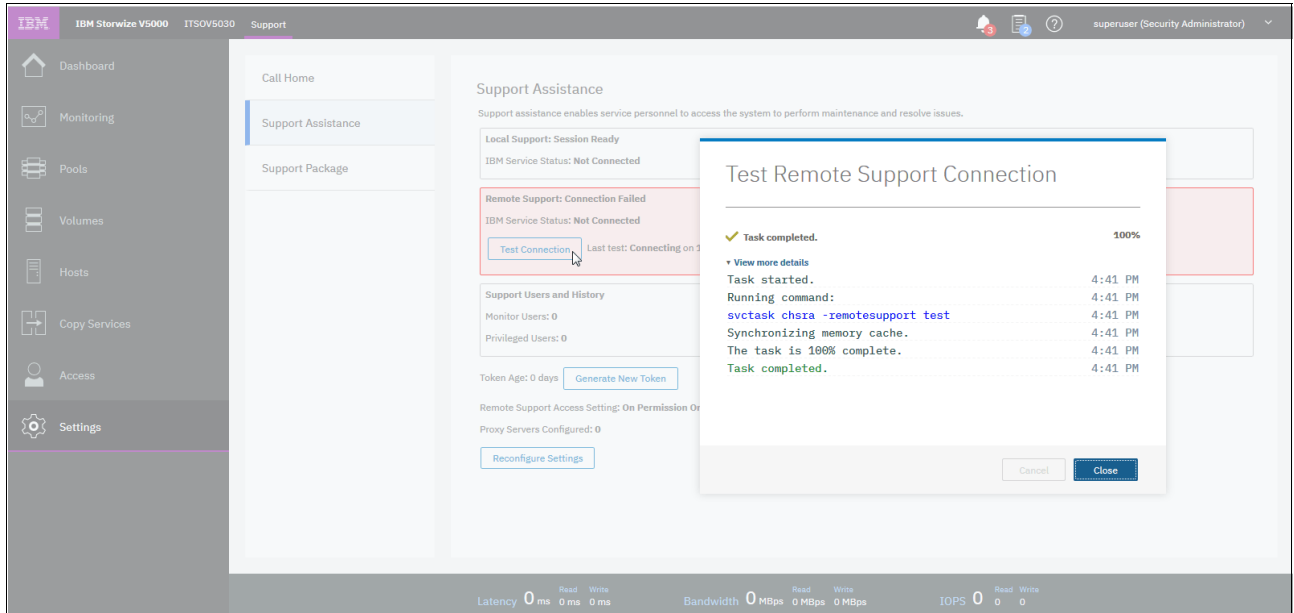


Figure 12-107 Test connection

Figure 12-108 shows the testing status message.

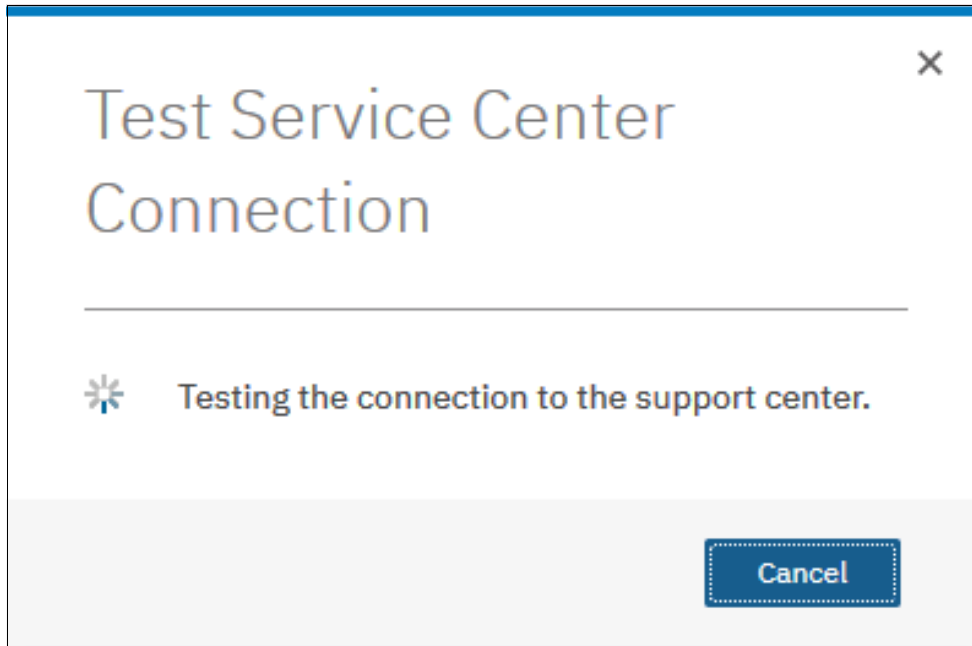


Figure 12-108 Test Service Center Connection

If a failure occurs, you see the status, as shown in Figure 12-109.

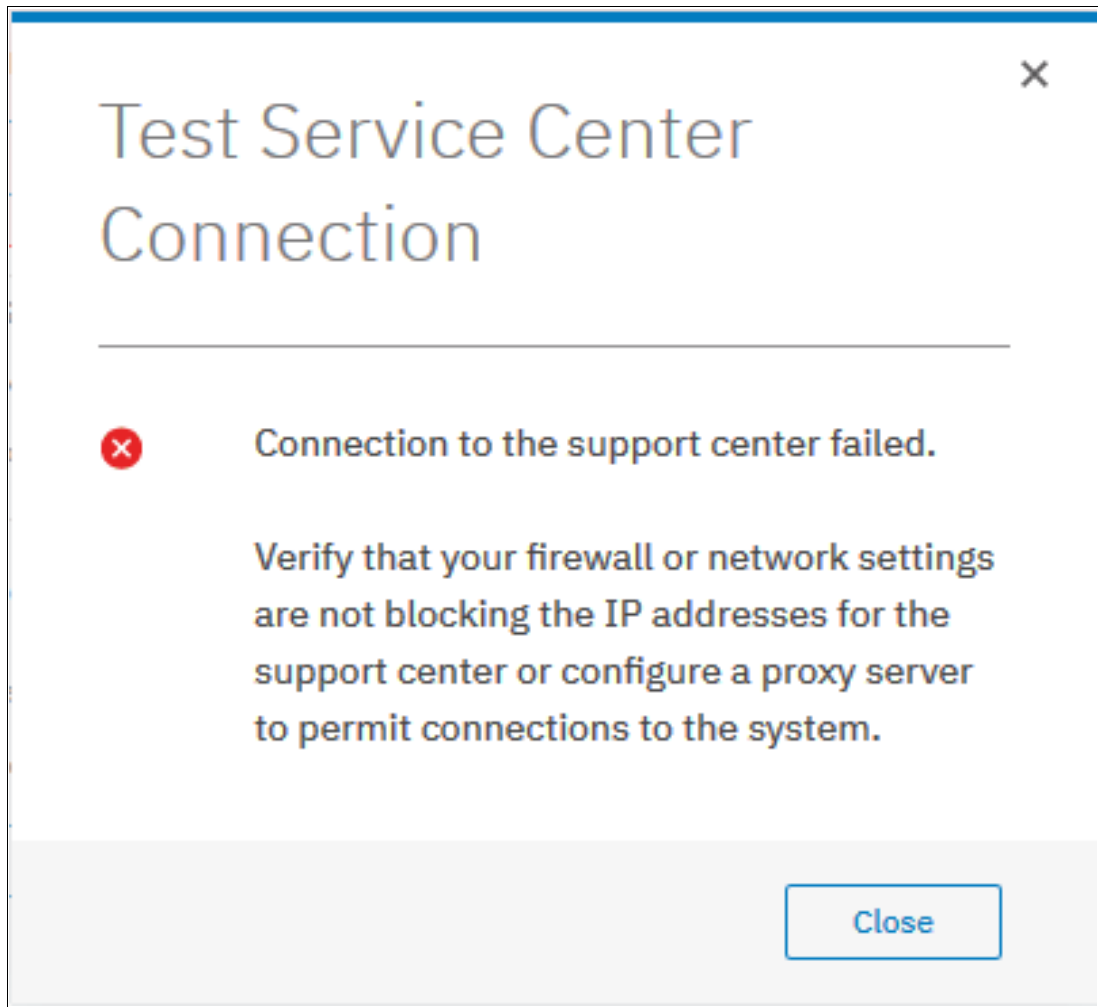


Figure 12-109 Connection error status

If no errors occur, you see an overview of the remote users, as shown in Figure 12-110.

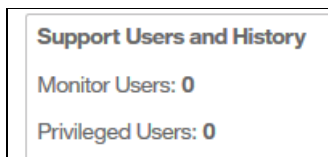


Figure 12-110 Support Users

7. A new Token can be generated by clicking **Generate New Token**, as shown in Figure 12-111.

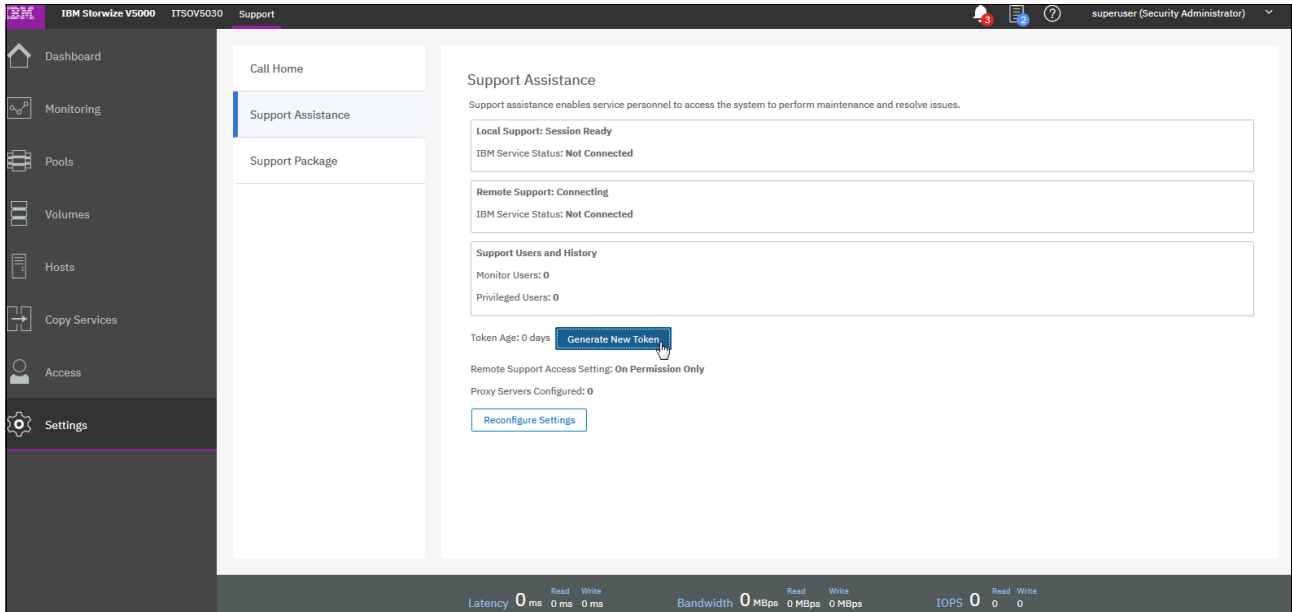


Figure 12-111 Generate New Token

When you enable remote support assistance, the system generates a support assistance token. This shared security token is sent to IBM Support Center and is used for authentication during support assistance sessions.

Updating a token is essentially overwriting the existing token, then sending it securely to the support assistance administration server in an email message. You specify the email addresses of the support assistance administration servers when you configure support assistance.

If the email is not received in time for a support incident or cannot be sent for some reason, a service engineer can manually add the token to the administration server. Before you can update a token, you must enable the support assistance feature. You can update the token periodically as a security practice, similar to how you update passwords.

8. To update a shared support assistance token, enter the following command:
`svctask chsra -updatetoken`

9. If settings change over time, you can reconfigure your settings by clicking **Reconfigure Settings**, as shown in Figure 12-112.

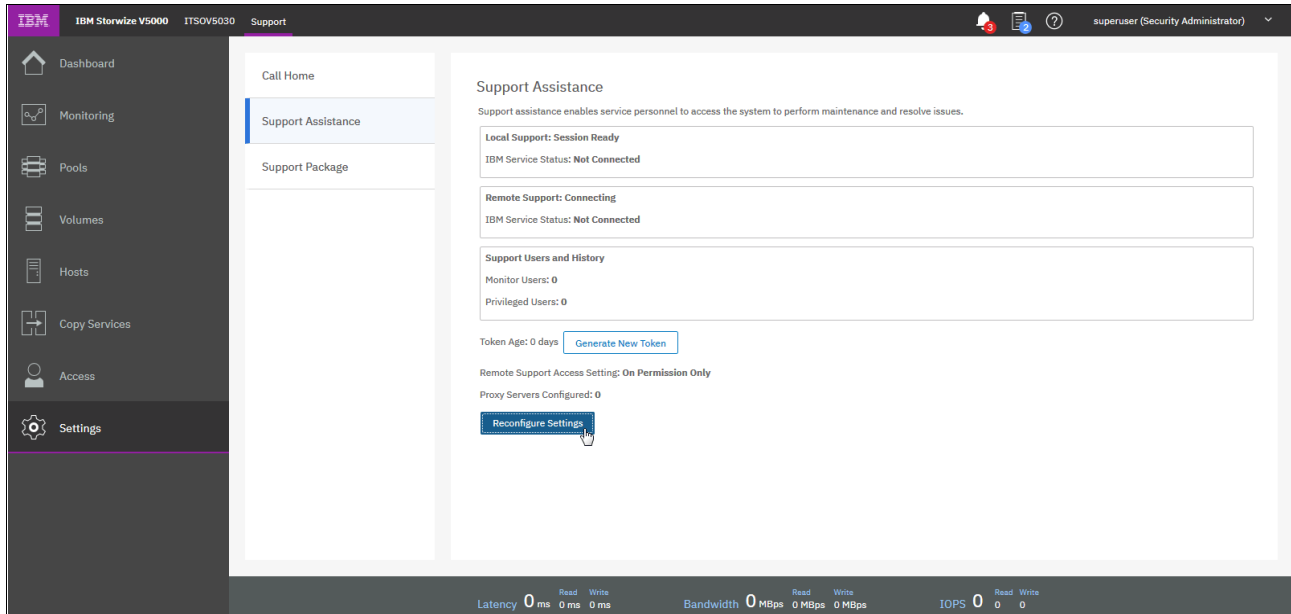


Figure 12-112 Reconfigure Settings

10. If you plan to use the command-line interface to configure assistance, use the following commands:
- For local support assistance, enter the following command:
`chsra -enable`
 - To configure remote support assistance, enter the following command:
`chsra -remotesupport enable`

12.8.3 Disabling support assistance

You can disable support assistance by using the command-line interface (CLI). When you disable support assistance, the support assistance token is deleted. All active secure remote access user sessions are closed immediately and a secure email message is sent to the administration server to indicate that secure remote access is disabled on the system.

To disable support assistance completely, enter the following command:

```
svctask chsra -disable
```

To disable remote support assistance only, enter the following command:

```
svctask chsra -remotesupport disable
```

12.9 Collecting support information

If you have an issue with an IBM Storwize V5000 Gen2 and contact the IBM Support Center, you might be asked to provide support data as described in the next section.

12.9.1 Collecting support information by using the GUI

To reach the Support Package window, click **Settings** → **Support** → **Support Package** and the window that is shown in Figure 12-113 opens.

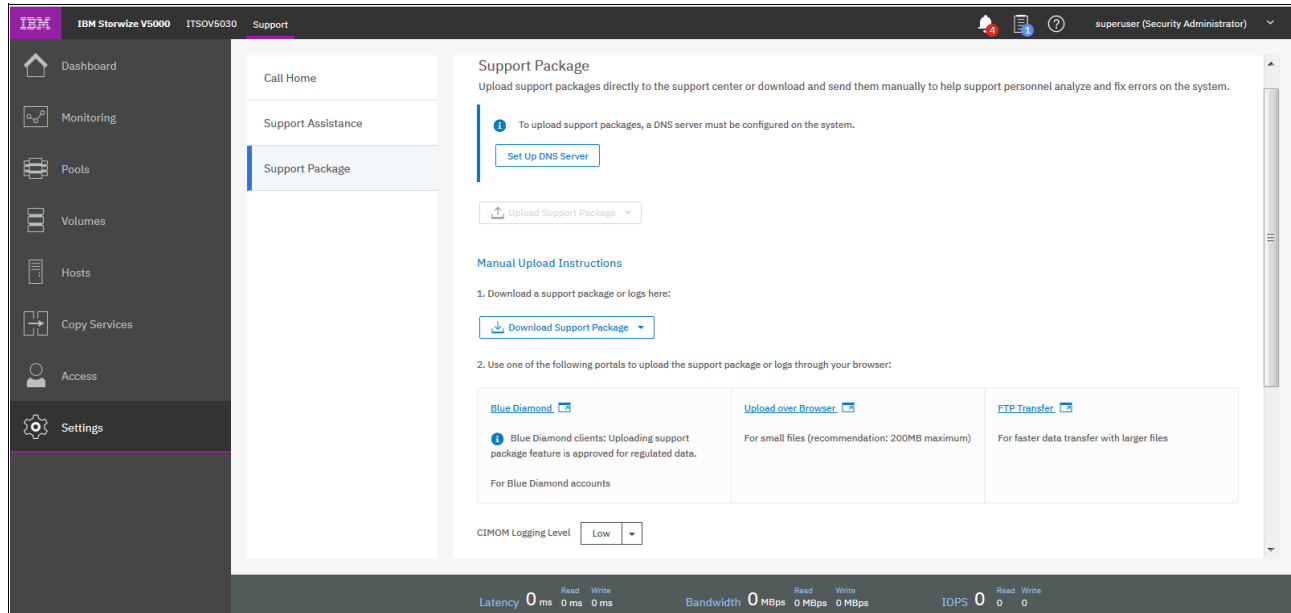


Figure 12-113 Support Package

12.9.2 Automatic upload of support packages

You can use the management GUI or the command-line interface to upload support packages to the IBM Support Center. If support assistance is configured on your systems, you can automatically or manually upload new support packages to the IBM Support Center to help analyze and resolve errors on the system. You can select individual logs to download to review or send directly to the IBM Support Center for analysis.

Before automatically uploading a support package, ensure that the following prerequisites are configured on the system:

- ▶ All the nodes on the system have internet access.
- ▶ A valid service IP address is configured on each node on the system.

- Configure at least one valid DNS server for domain name resolution. To configure a DNS server on the system, click **Set Up DNS Server** (see Figure 12-114). You can also use the `mkdnserver` command to configure DNS servers.

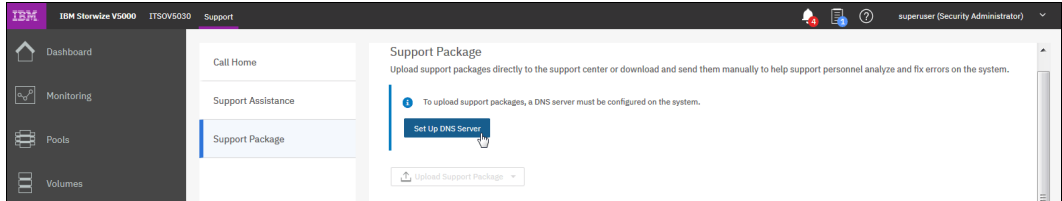


Figure 12-114 Set up DNS Server

The Settings Page for DNS then opens (see Figure 12-115).

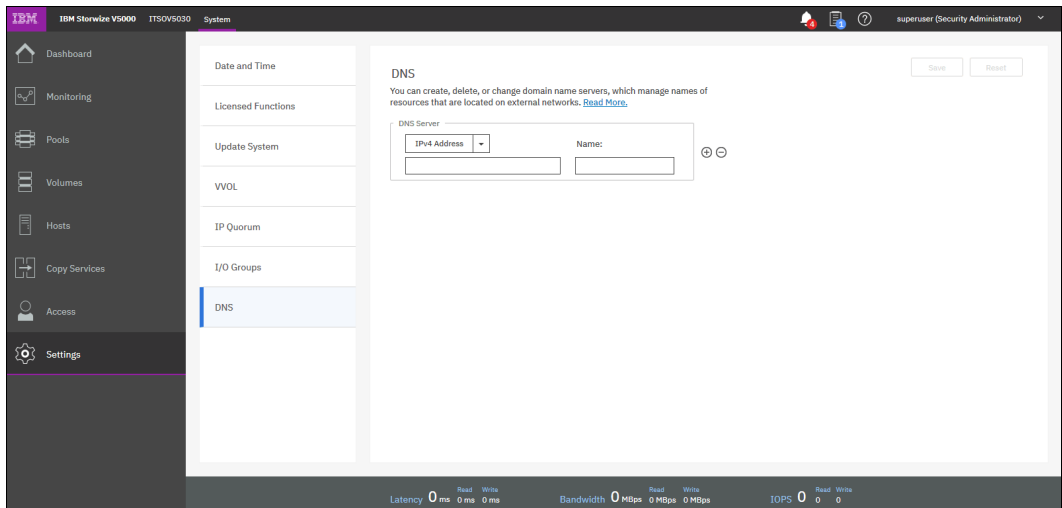


Figure 12-115 Set Up DNS

11. Enter the DNS IP address and click **Save** (see Figure 12-116).

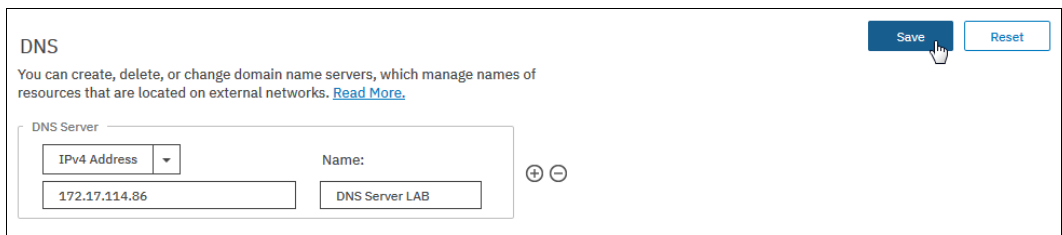


Figure 12-116 DNS IP Address

12. 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.

To test connections to IBM Support Center, select **Settings** → **Support** → **Support Assistance**. On the Support Assistance page, select **Test Connection** to verify connectivity between the system and IBM Support Center.

The management GUI supports uploading new or existing support packages to support automatically (see Figure 12-117).

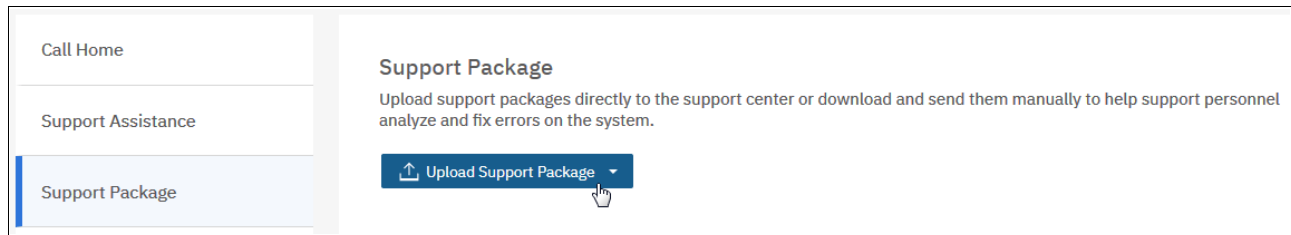


Figure 12-117 Upload Support Package

13. When you click **Upload Support Package**, the selection window opens, as shown in Figure 12-118. Here, you can choose to create a package or upload a package.

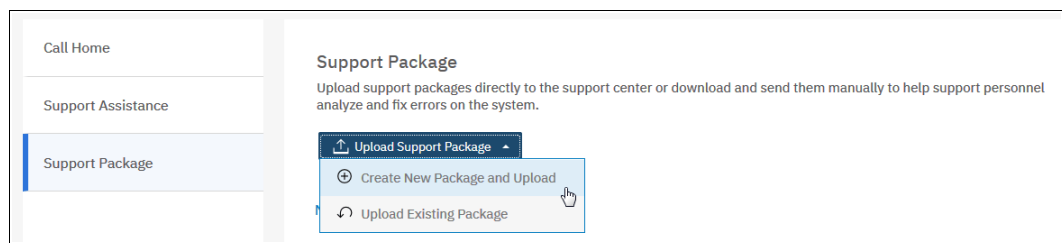


Figure 12-118 Create and Upload Support package

14. For our example, we create a package (see Figure 12-119).

Upload Support Package

Your system will generate and upload a new package to the IBM support center.

PMR Number: [Don't have PMR?](#)

ppppp,bbb,ccc

Select the type of new support package to generate and upload to the IBM support center:

- Snap Type 1: Standard logs
Contains the most recent logs for the system, including the event and audit logs.
- Snap Type 2: Standard logs plus one existing statesave
Contains all the standard logs plus one existing statesave from any of the nodes in the system.
- Snap Type 3: Standard logs plus most recent statesave from each node
Contains all the standard logs plus each node's most recent statesave.
- Snap Type 4: Standard logs plus new statesaves
Contains all the standard logs and generate a new statesave on each node in the system.

Need Help Cancel Upload

Figure 12-119 Upload Support package

15. On the Upload Support Package page, enter the Problem Management Report (PMR) number that is associated with the support package that you are uploading. If you do not have a PMR number, click **Don't have a PMR?** to open the Service Request (SR) tool to generate a PMR. You need an IBM Business Partner ID to register.

Note: If you are not sure if a PMR exists or do not want to create a PMR, the package can still be sent to IBM Support Center. The machine serial number and type are used to route the package to IBM Support Center. However, specifying a PMR number can decrease response time for Support personnel. You can call the IBM Support Line or use the IBM Support Portal to open a call.

For more information, see this [IBM Support web page](#).

16. Specify the type of package that you want to generate and upload to IBM Support Center by selecting one of the following options:

– Type 1: Standard logs

This option contains different log files for the underlying operating system and the software that includes the following critical log files:

- Event log
- Audit logs
- Linux based logs that include `/var/log/messages`

These logs are sufficient to diagnose many problems, especially simple hardware replacement procedures.

– Type 2: Standard logs plus one existing statesave

This option contains the standard logs for the system and an existing statesave from any of the nodes in the system. Statesaves can be a dump or a livedump. A dump file is collected when the software restarts for any reason. This file can be used to understand why the software restarted. A livedump collects the current state of the software with minimal impact to I/O operations. The contents of a livedump are similar to the contents of a dump.

– Type 3: Standard logs plus the most recent statesave from each node

Contains the standard logs for the system and the most recent statesave from each of the nodes on the system.

– Type 4: Standard logs plus new statesave

This option contains the standard logs and a new statesave (livedump) for all the nodes in the system and packages them with the most recent logs.

17. When you are deciding what type of package to send, consider the speed of the upload and if more information is necessary later to resolve the issue. The standard logs are uploaded more rapidly because the support package is smaller than the other options.

However, the standard logs might not include all the information that is needed to resolve the problem on the first attempt. Support personnel can require more data to resolve the problem.

If a new statesave (livedump) is not generated shortly after the standard log, the livedump might not contain the data necessary to resolve the issue. When you generate the standard logs with the new statesave, you have the greatest chance of resolving the issue without more information. However, the support package is larger and can take longer to upload to IBM Support Center.

You can also set CIMOM logging levels. CIMOM logging provides more information for advanced error resolution; however, increasing the level can cause performance issues on the system.

Figure 12-120 shows the four types of support packages with the approximate size for a support package for basic configuration. It also describes common troubleshooting scenarios that require that type of support package. Actual sizes of support packages vary based on the number of volumes and MDisks, software level, and compressibility of the logs.

Description	Approximate Support Package Size: System with 1 I/O group and 30 volumes)	Approximate Support Package Size: System with 4 I/O groups and 250 volumes	Common Scenarios
Standard logs	10 MB	340 MB	<ul style="list-style-type: none"> • Simple hardware replacement • Critical performance issues¹ • All other issues
Standard logs plus one existing statesave	50 MB	520 MB	
Standard logs plus most recent statesave from each node	90 MB	790 MB	2030, 1196 or 1195 errors
Standard logs plus new statesaves	90 MB	790 MB	<ul style="list-style-type: none"> • Compressed volumes issues • Host issues • External storage issues • Metro Mirror or Global Mirror issues² • General performance issues • Critical performance issues¹

1. Critical performance issues require that both standard logs and a new statesave are generated and uploaded to the support center. Generate and upload the standard logs first and then create the new statesaves.

2. Generate a support package on both the source and target systems in the remote-copy relationship.

Figure 12-120 Support Package Size

IBM Support Center notifies you know which package is needed.

18. Click **Upload**. After the new support package is generated, a summary window displays the progress of the upload. If the upload is unsuccessful or encounters errors, verify the connection between the system and IBM Support Center and retry the upload.

19. If you decide that you want to upload the support package later, you can use the function that is shown in Figure 12-121.

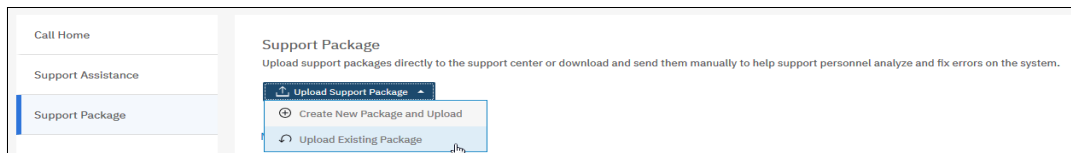


Figure 12-121 Upload Existing Package

20. Click **Upload Existing Package** and a window opens, as shown in Figure 12-122.

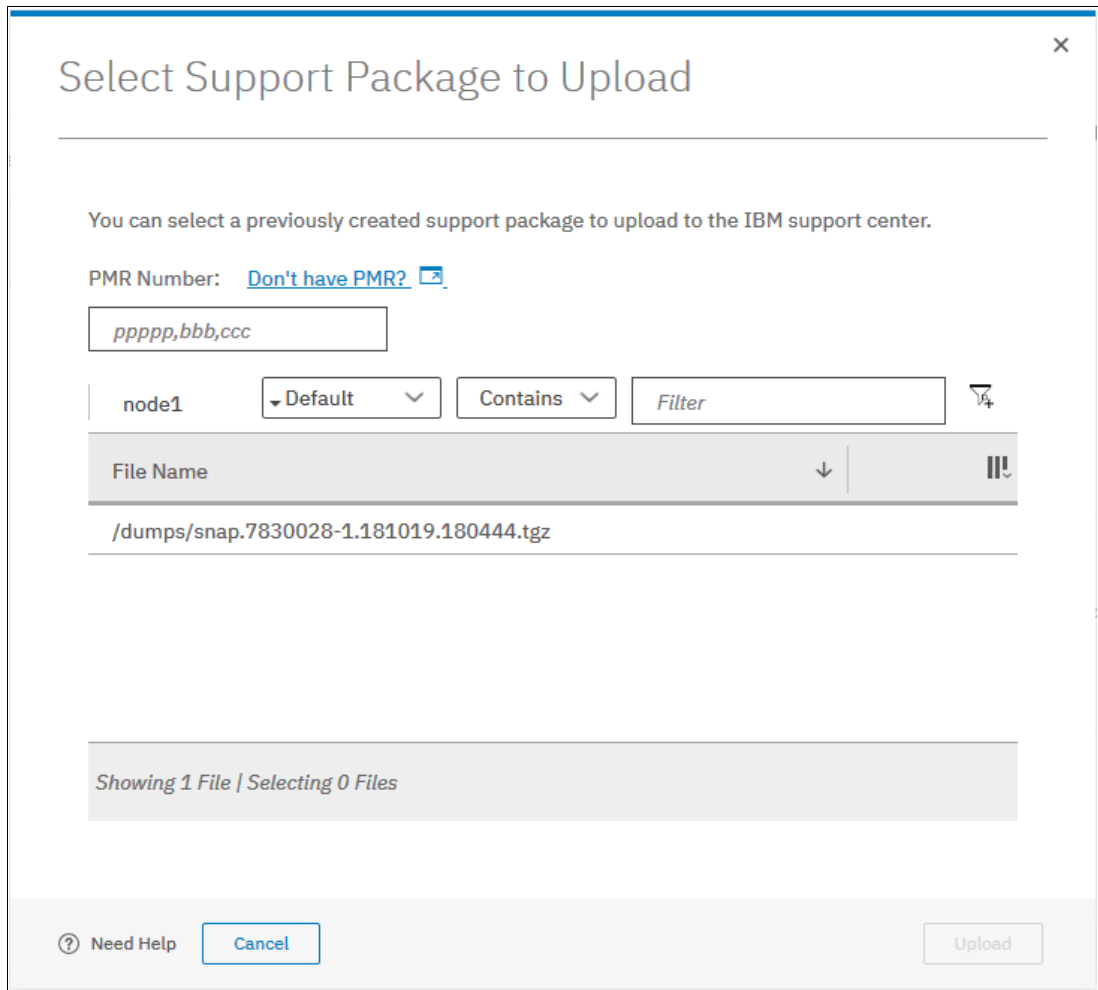


Figure 12-122 Select Support Package to Upload

Using the command-line interface

To upload a support package or other file with the CLI, complete the following steps:

1. Enter the following command:

```
satask supportupload -pmr pmr_number -filename fullpath/filename
```

Where the `pmr_number` is the number of an existing PMR and `fullpath/filename` is the full path and the name of the file that you are uploading (see Example 12-3). The `-pmr` and `-filename` parameters are not required. If you do not specify a PMR number, the file is uploaded by using the machine serial and type to route the file to IBM Support Center. If you do not specify a file name, the latest support package is uploaded (see Example 12-3).

Example 12-3 Using the CLI

```
satask supportupload -pmr 79556,019,866 -filename /dumps/snap.single.7830619-1.161219.161359.tgz
```

2. To verify the progress of the upload to IBM Support Center, enter the following command:

```
lscmdstatus
```

In the results of this command, verify that the **supportupload_status** is Complete, which indicates that the upload is successfully completed. Other possible values for this parameter include Active, Wait, Abort, and Failed. If the upload is Active, you can use the **supportupload_progress_percent** parameter to view the progress for the upload.

If you want to generate a new support package, complete the following steps:

1. Enter the following command in the command-line interface:

```
satask snap -upload -pmr pmr_number
```

where the `pmr_number` is the number of an existing PMR. The command generates a support package and uploads it to IBM Support Center with the identifying PMR number. If you do not have a PMR number that corresponds with support package, use the following command:

```
satask snap -upload
```

The command generates a support package and uploads it to IBM Support Center by using the machine type and serial to route the package.

2. To verify the progress of the upload to IBM Support Center, enter the following command:

```
lscmdstatus
```

In the results of this command, verify that the **supportupload_status** is Complete, which indicates that the upload is successfully completed. Other possible values for this parameter include Active, Wait, Abort, and Failed. If the upload is Active, you can use the **supportupload_progress_percent** parameter to view the progress for the upload.

12.9.3 Manual upload of Support Packages

You can use the management GUI or CLI to manually upload support packages to IBM Support Center. If support assistance is configured on your systems, you can manually upload new support packages to IBM Support Center to help analyze and resolve errors on the system. You can select individual logs to download to review or send directly to IBM Support Center for analysis.

Using the management GUI

The management GUI supports manually uploading support packages. Manually uploading support packages require that you download a new support package or an existing support package to your system and then upload the file to support directly.

To manually upload a new support package to IBM Support Center, complete the following steps:

1. In the management GUI, select **Settings** → **Support** → **Support Package**.

2. On the Support Package page, see the section **Manual Upload Instructions**, as shown in Figure 12-123.

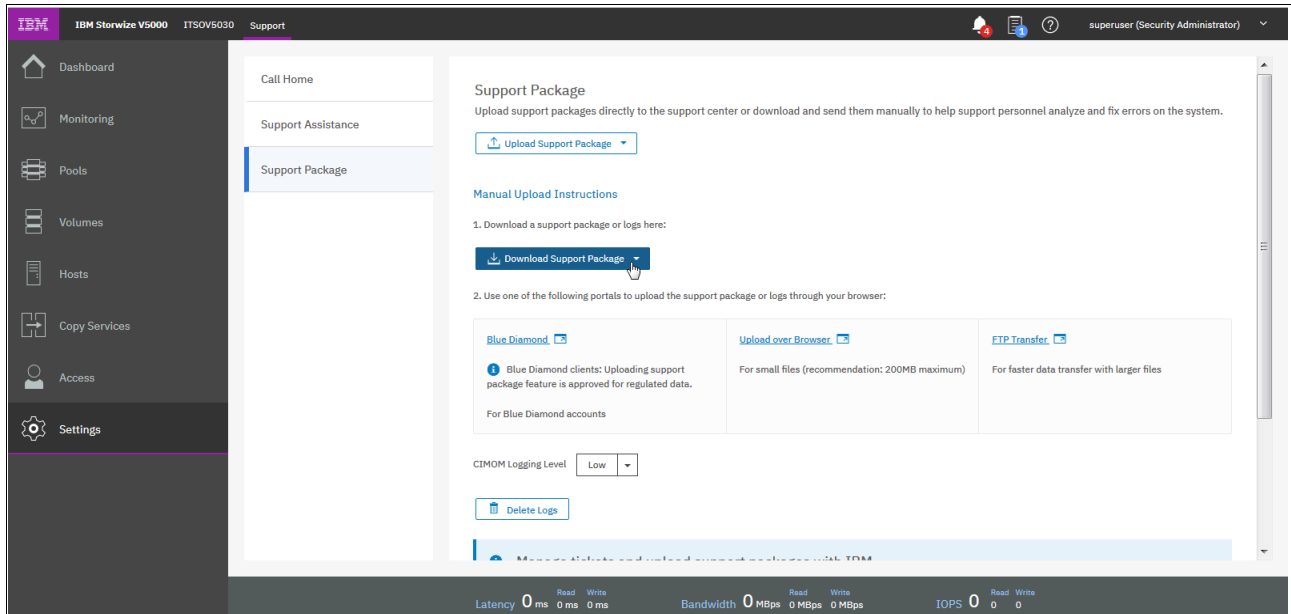


Figure 12-123 Manual Upload Instructions

3. In the Manual Upload Instructions section, click **Download Support Package** and select **Create New Package and Download** (see Figure 12-124).

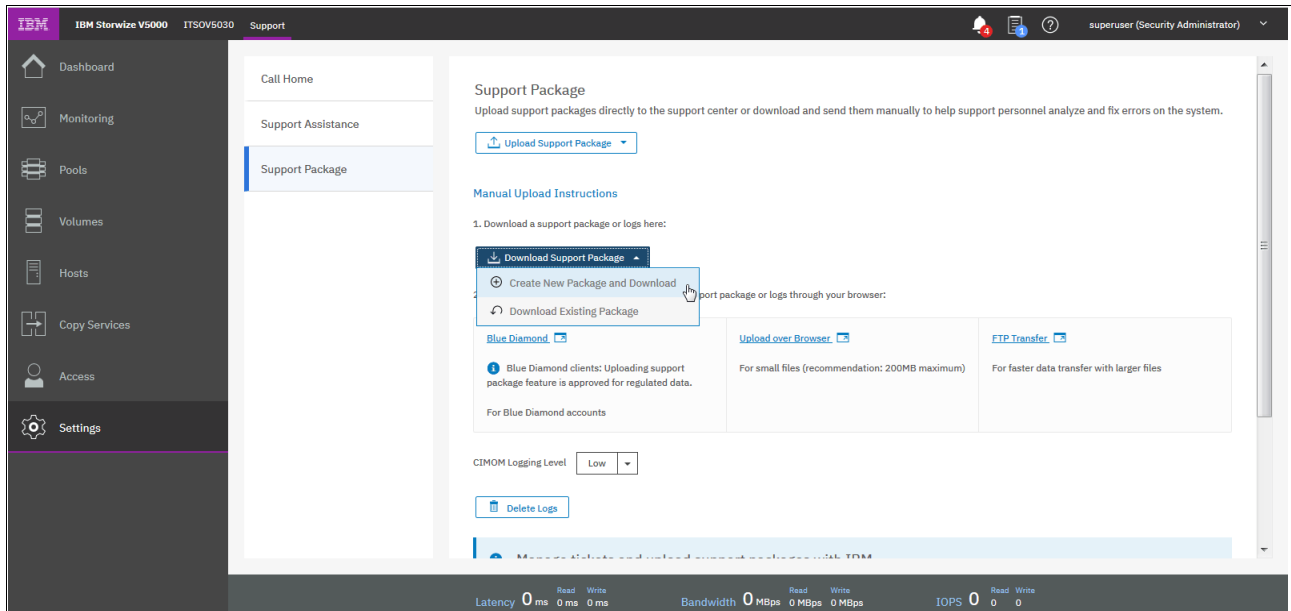


Figure 12-124 Download Support Package

4. In the Download New Support Package or Log File window, select one a type of support package to download, as shown in Figure 12-125.

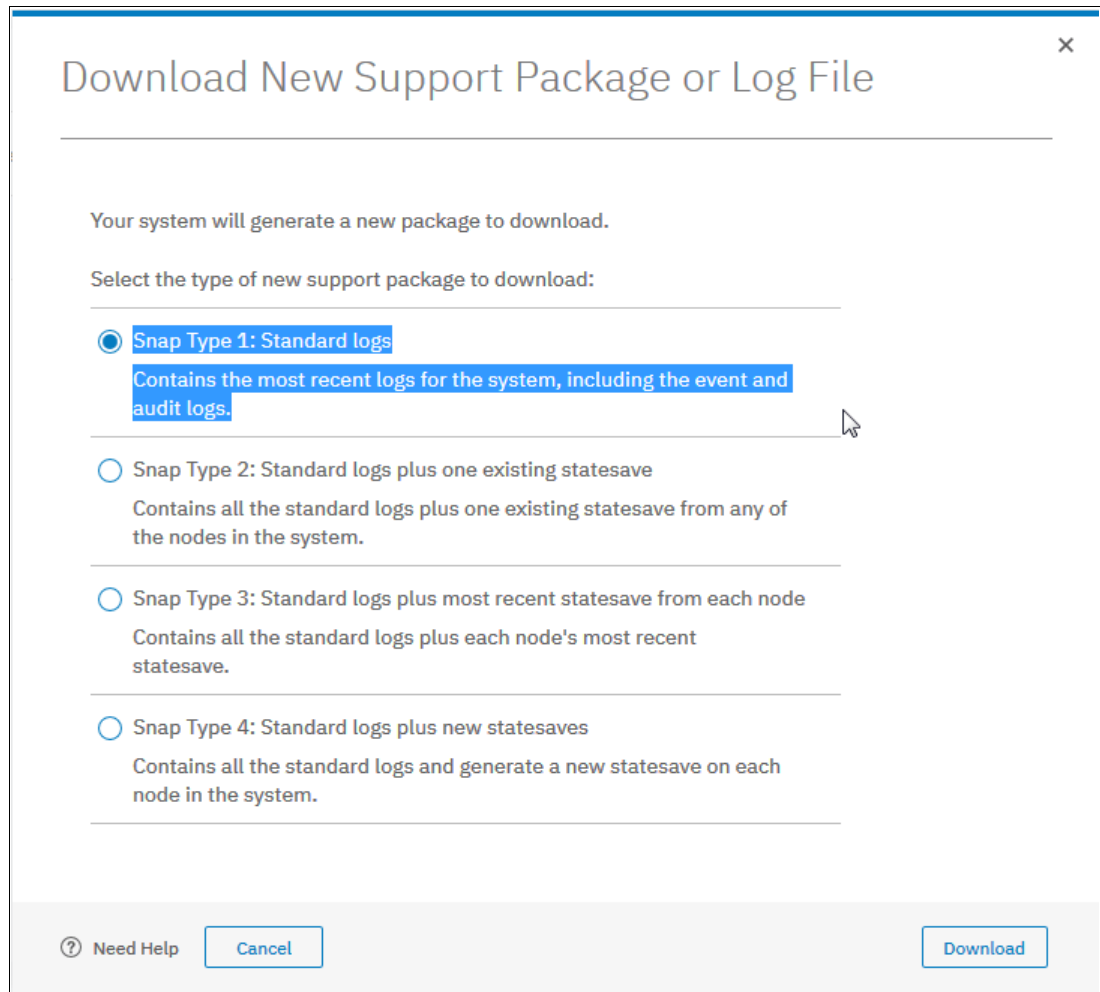


Figure 12-125 Download support package or log file

The type to select depends on the event that is being investigated. For example, if you notice that a node is restarted, capture the snap file with the latest existing statesave. If needed, the IBM Support Center can notify you of the package that is required.

The following components are included in each type of support package:

– Type 1: Standard logs

This option contains many different log files for both the underlying operating system and the software that includes these critical log files:

- Event log
- Audit logs
- Linux based logs that include `/var/log/messages`

These logs are sufficient to diagnose many problems, especially simple hardware replacement procedures.

– Type 2: Standard logs plus one existing statesave

This option contains the standard logs for the system and an existing statesave from any of the nodes in the system. Statesaves can either be a dump or a livedump. A dump file is collected when the software restarts for any reason.

This file can be used to understand why the software restarted. A livedump collects the current state of the software with minimal impact to I/O operations. The contents of a livedump are similar to the contents of a dump.

- Type 3: Standard logs plus the most recent statesave from each node

Contains the standard logs for the system and the most recent statesave from each of the nodes on the system.

- Type 4: Standard logs plus new statesave

This option contains the standard logs and a new statesave (livedump) for all the nodes in the system and packages them with the most recent logs.

For more information about the time it takes to upload your package, see step 17 on page 709.

5. Click **Download** to download the support package to your local computer.
6. After the download completes to your local computer, you can upload the package to IBM Support Center by using one of the methods that are shown in Figure 12-126.

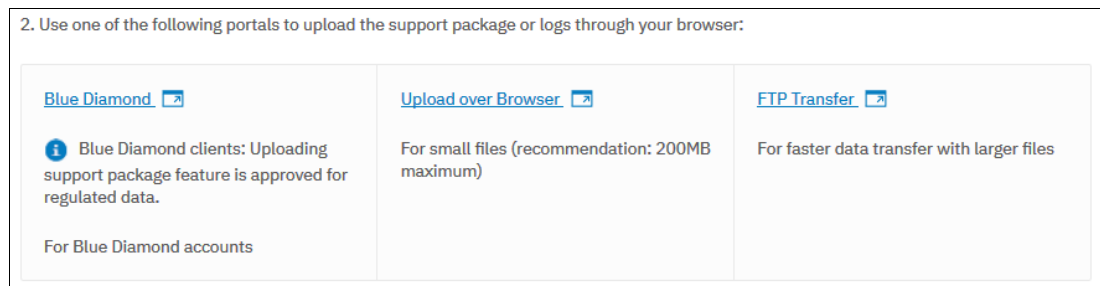


Figure 12-126 Download portals

The following options are available:

- Blue Diamond

Select the link to log in to the BlueDiamond portal. BlueDiamond provides enhanced security and support for healthcare clients. You must be a registered BlueDiamond client to use this option. After you accept the terms of service for the upload, log in to the BlueDiamond portal with your user name and password.

- Upload over Browser

Use this option for small files under 200 MB. Select the link to upload the support package to the support website through the web browser. On the support website, complete the following steps:

- Enter a valid PMR number that is associated with this support package.
- In the Upload is for field, select **Other**.
- Enter a valid email address for the contact for this package.

- FTP Transfer

Use this option for larger files. Select the link to send the package to support with file transfer protocol (FTP). You can send packages to support with standard FTP (non-secure), secure FTP, or with SFTP, which is FTP over Secure Shell protocol (SSH). On the support port for FTP transfers, select the type of FTP you want to use and follow the instructions for that method.

Choose the CIMOM Logging Level you prefer (see Figure 12-127).

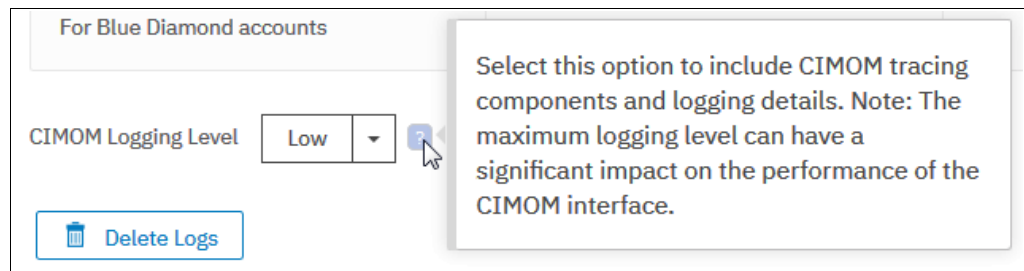


Figure 12-127 CIMOM tracing

A new support offering is available with V8.2.1 called Storage Insights (see Figure 12-128).

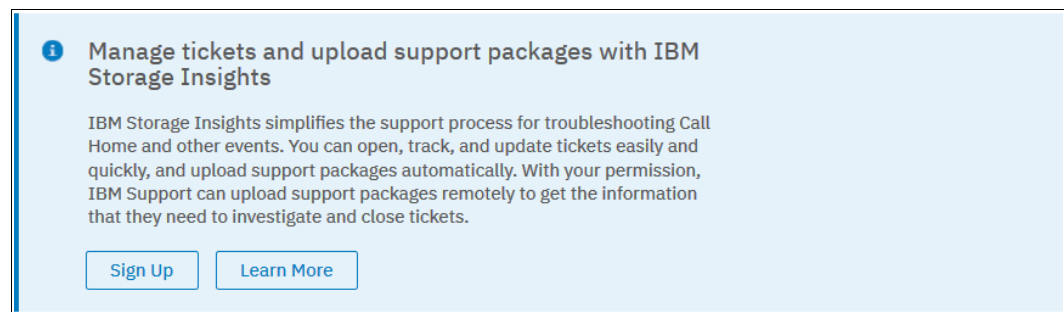


Figure 12-128 Storage Insights

For more information about the offering, see this IBM Storage Insights [Fact Sheet](#).

Using the command-line interface

To upload a support package or other file by using the CLI, complete the following steps:

1. Enter the following command:

```
satask supportupload -pmr pmr_number -filename fullpath/filename
```

Where the **pmr_number** is the number of an existing PMR and **fullpath/filename** is the full path and the name of the file that you are uploading (see Example 12-4).

Example 12-4 Upload support package

```
satask supportupload -pmr 79556,019,866 -filename  
/dumps/snap.single.7830619-1.161219.161359.tgz
```

The **-pmr** and **-filename** parameters are not required. If you do not specify a PMR number, the file is uploaded by using the machine serial and type to route the file to IBM Support Center. If you do not specify a file name, the latest support package is uploaded.

2. To verify the progress of the upload to IBM Support Center, enter the following command:

```
lscmdstatus
```

In the results of this command, verify that the **supportupload_status** is Complete, which indicates that the upload is successfully completed. Other possible values for this parameter include Active, Wait, Abort, and Failed. If the upload is Active, you can use the **supportupload_progress_percent** parameter to view the progress for the upload.

If you want to generate a support package, complete the following steps:

1. Enter the following command in the command-line interface:

```
satask snap -upload -pmr pmr_number
```

Where the `pmr_number` is the number of an existing PMR. The command generates a support package and uploads it to IBM Support Center with the identifying PMR number. If you do not have a PMR number that corresponds with support package, you can use the following command:

```
satask snap -upload
```

The command generates a new support package and uploads it to IBM Support Center by using the machine type and serial to route the package.

2. To verify the progress of the upload to IBM Support Center, enter the following command:

```
lscmdstatus
```

In the results of this command, verify that the `supportupload_status` is `Complete`, which indicates that the upload is successfully completed. Other possible values for this parameter include `Active`, `Wait`, `Abort`, and `Failed`. If the upload is `Active`, you can use the `supportupload_progress_percent` parameter to view the progress for the upload.

12.9.4 Collecting support information by using the Service Assistant Tool

IBM Storwize V5000 Gen2 management GUI collects information from all of the components in the system. The Service Assistant Tool (SAT) collects information from all node canisters. The `snap file` is the information that is collected and packaged in a single file.

If the package is collected by using the SAT, ensure that the node from which the logs are collected is the current node, as shown in Figure 12-129.

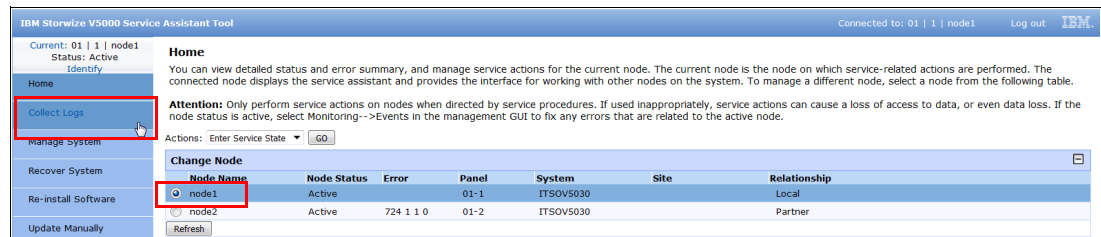


Figure 12-129 Accessing the Collect Logs window in the SAT

Support information can be downloaded with or without the latest statesave, as shown in Figure 12-130.

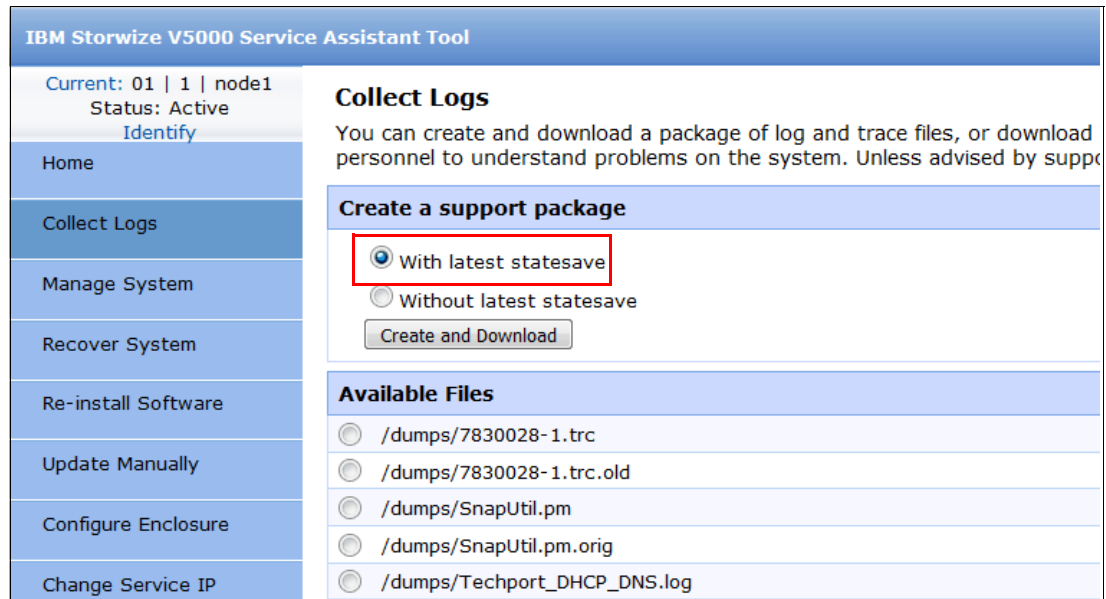


Figure 12-130 Collect Logs window in the Service Assistance Tool

Accessing the SAT by using the technician port

If your system or one of your node canisters is inaccessible through the administrative network, you can connect a personal computer directly to the technician port on the node canister to access the SAT.

Note: This procedure starts the initialization tool if the node canister that is being serviced is in the candidate state, if no system details are configured, and if the partner node is not in the active state.

Complete the following steps:

1. Configure Dynamic Host Configuration Protocol (DHCP) on the Ethernet port of the personal computer to connect to the node canister.
Alternatively, if the personal computer does not support DHCP, configure the static IPv4 address 192.168.0.2 on the port.
2. On the Storwize V5010 system or Storwize V5020 system, re-enable the technician port by completing the following steps:
 - a. Create a text file with the `satask chserviceip -techport enable -force` command.
 - b. Save the file as `satask.txt` in the root directory of the Universal Serial Bus (USB) stick.
 - c. Insert the USB stick in the USB port of the node that you want to service.
 - d. Wait until no write activity is recognized and remove the USB stick.

Note: The Storwize V5030 systems have a dedicated technician port that is always enabled, so this step is unnecessary.

3. Connect an Ethernet cable between the port on the personal computer and the technician port. The technician port is labeled with a T on the rear of the node canister.

4. Open a supported web browser on the personal computer and browse to the `http://192.168.0.1` URL.

Note: If the cluster is active and you connect to the configuration node, this URL opens the management GUI. If you want to access the SAT in this case, browse to `http://192.168.0.1/service`.

5. Complete the correct procedure to service the canister.
6. Log out of the Service Assistant Tool and disconnect the Ethernet cable from the technician port.
7. On the Storwize V5010 system or Storwize V5020 system, disable the technician port by running the command that is shown in Example 12-5.

Example 12-5 Disabling the technician port

```
>satask chserviceip -techport disable
```

SAS port 2 can then be used again to provide extra Ethernet connectivity for system management, iSCSI, and IP replication.

12.10 Powering off the system and shutting down the infrastructure

The following sections describe the process to power off the system and to shut down and start an entire infrastructure that contains an IBM Storwize V5000 Gen2.

12.10.1 Powering off

Important: Never power off your IBM Storwize V5000 Gen2 system by powering off the power supply units (PSUs), removing both PSUs, or removing both power cables from a running system. It can lead to inconsistency or loss of the data that is staged in the cache.

You can power off a node canister or the entire system. When you power off only one node canister for each I/O group, all of the running tasks remain active while the remaining node takes over.

Powering off the system is typically planned in site maintenance (power outage, building construction, and so on) because all components of the IBM Storwize V5000 Gen2 are redundant and replaceable while the system is running.

Important: If you are powering off the entire system, you lose access to all volumes that are provided by this system. Powering off the system also powers off all IBM Storwize V5000 Gen2 nodes. All data is flushed to disk before the power is removed.

Before you power off the system, stop all hosts with volumes that are allocated to this system. This step can be skipped for hosts with volumes that are provisioned with mirroring (host-based mirror) from different storage systems. However, skipping this step means that errors that relate to lost storage paths and disks can be logged on the host error log.

Note: If a canister or the system is powered off, a local visit can be required to either reseal the canister or power cycle the enclosures.

Powering off a node canister

To power off a canister by using the GUI, complete the following steps:

1. Browse to **Monitoring** → **System** as shown in Figure 12-131.

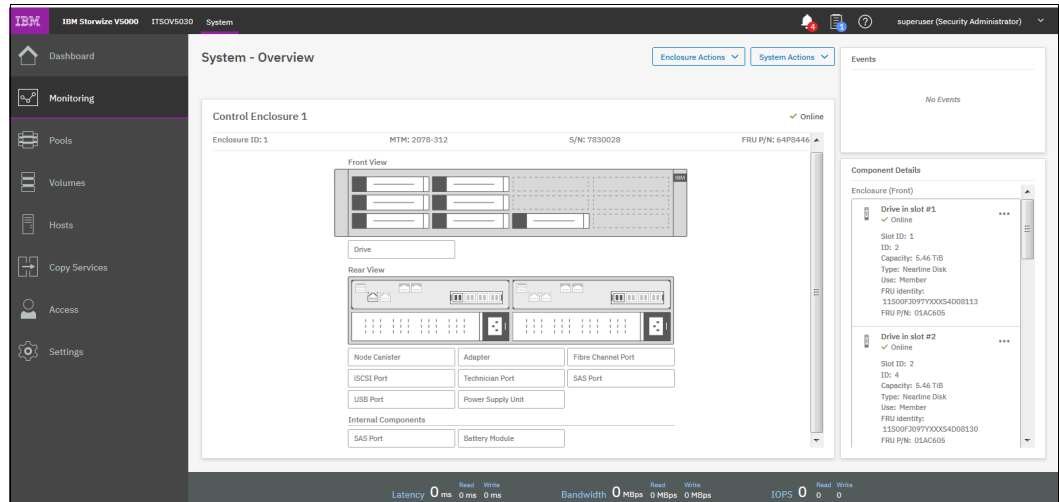


Figure 12-131 System Overview

2. Right-click the required canister and select **Power Off Canister**, as shown in Figure 12-132.

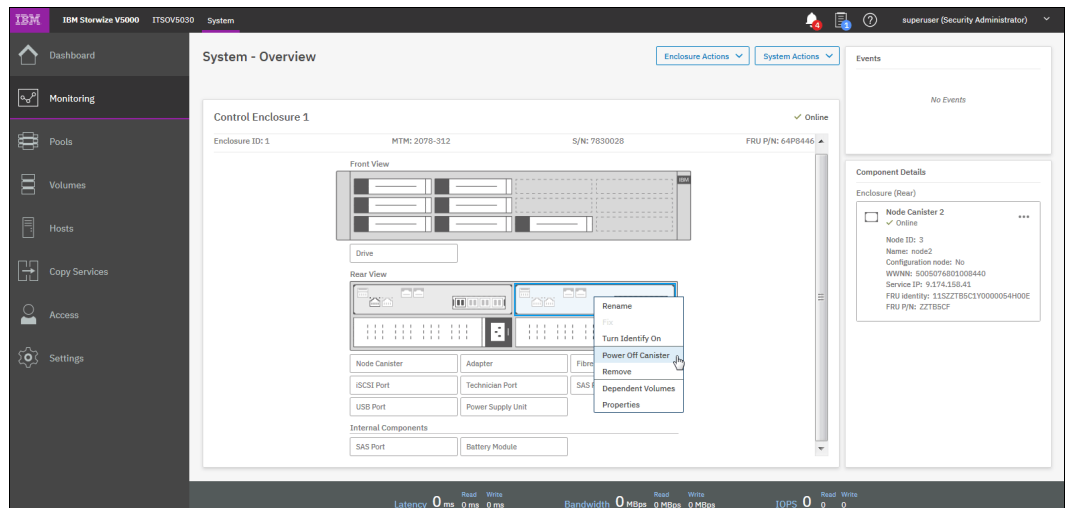


Figure 12-132 Powering off the canister

3. Confirm that you want to power off the canister by entering the confirmation code and clicking **OK**, as shown in Figure 12-133.

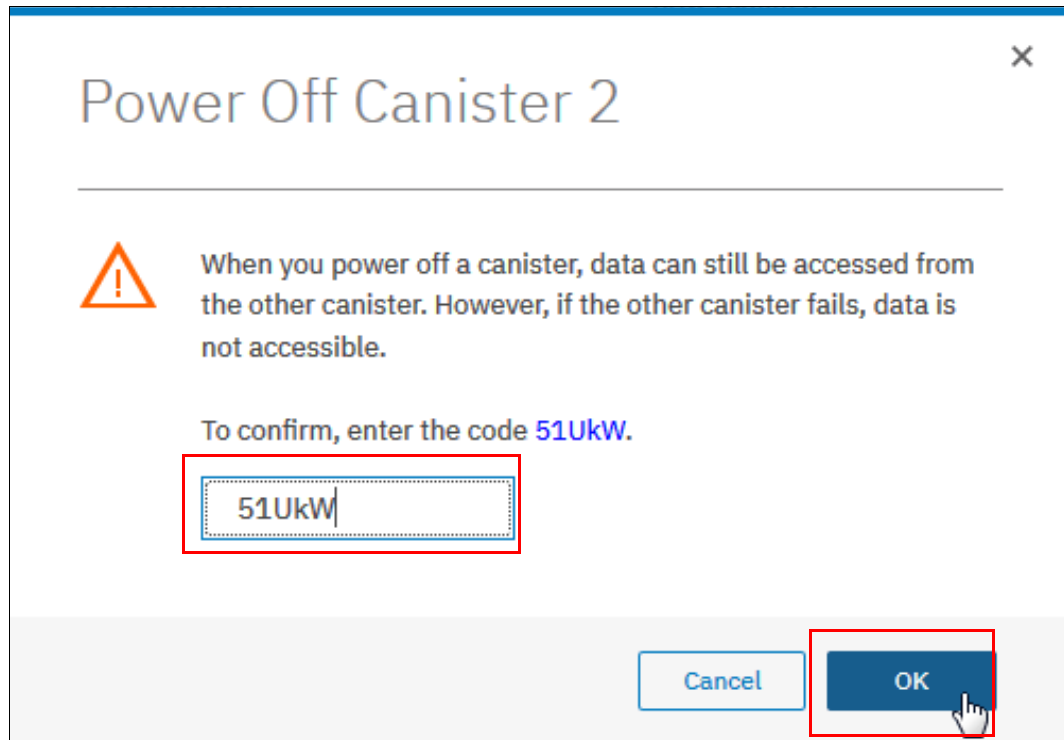


Figure 12-133 Canister power off confirmation window

4. After the node canister is powered off, you can confirm that it is offline in the System window, as shown in Figure 12-134.

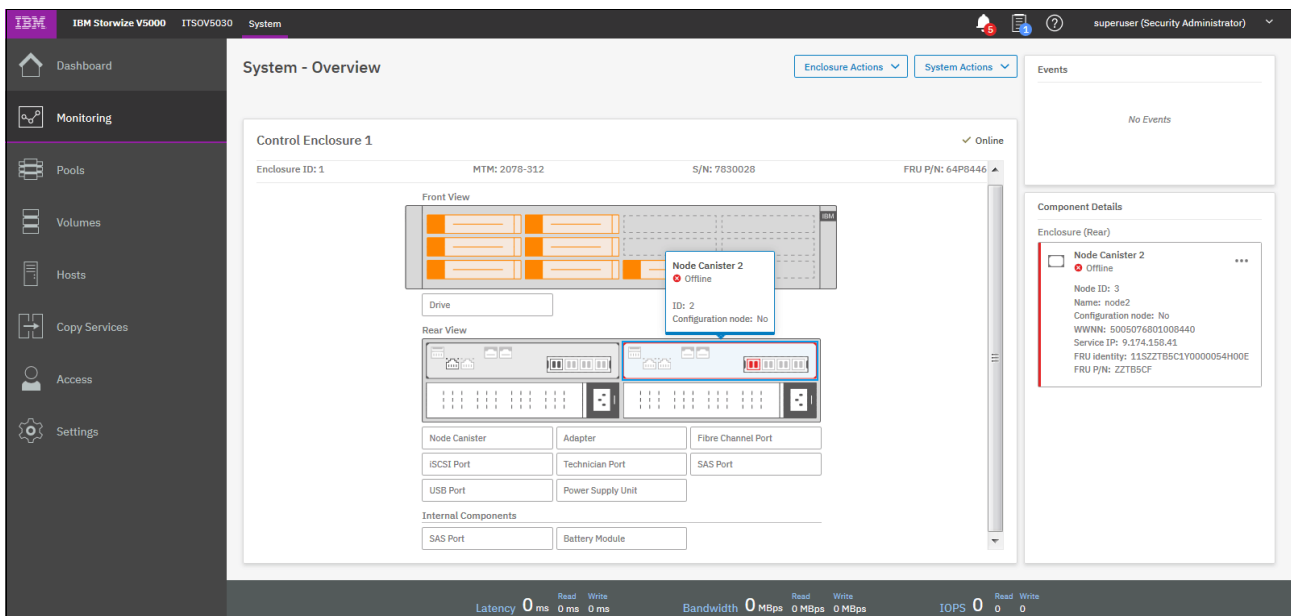


Figure 12-134 Checking the canister state

To power off a node canister by using the CLI, use the command that is shown in Example 12-6.

Example 12-6 Powering off a canister by using the CLI

```
>svctask stopsystem -node 2
```

```
Are you sure that you want to continue with the shut down? (y/yes to confirm)
```

If you power off the canister from the GUI, Figure 12-135 shows the system progress to power off this canister.

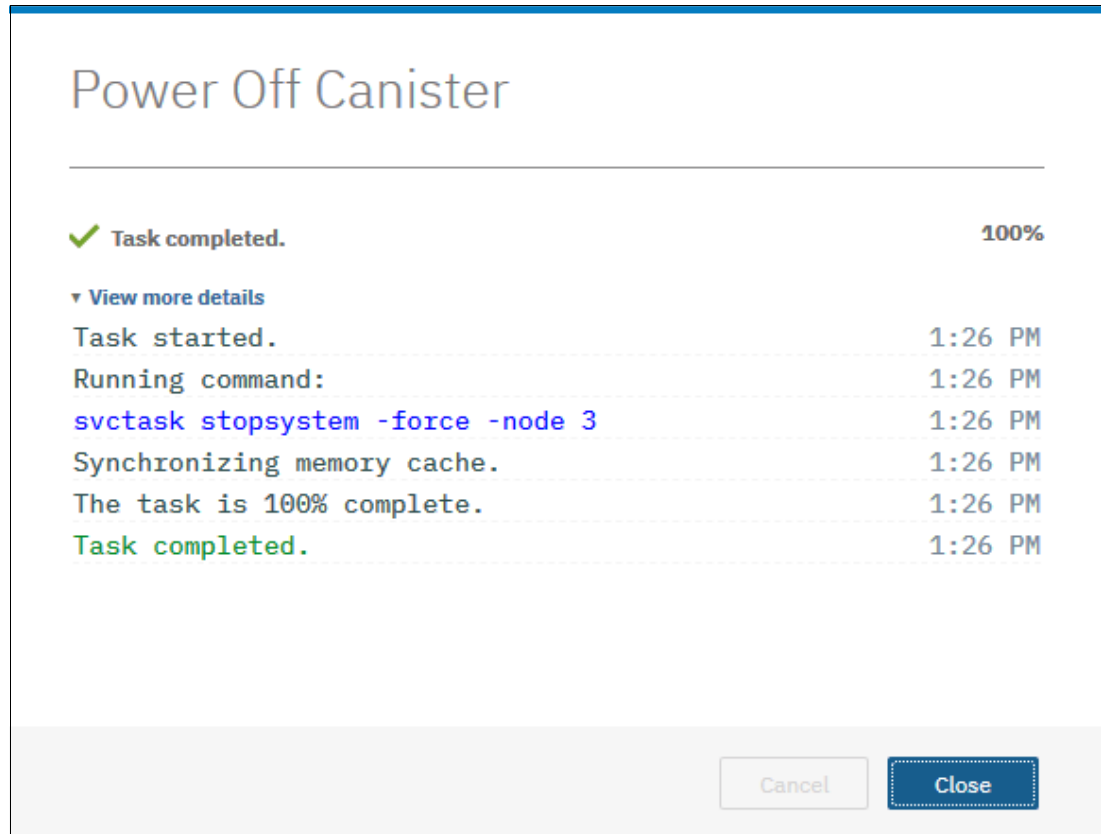


Figure 12-135 Power Off Canister

Powering off the system

To power off the entire system by using the GUI, complete the following steps:

1. Browse to **Monitoring** → **System**, click **Actions** → **Power Off System**, as shown in Figure 12-136 on page 723.

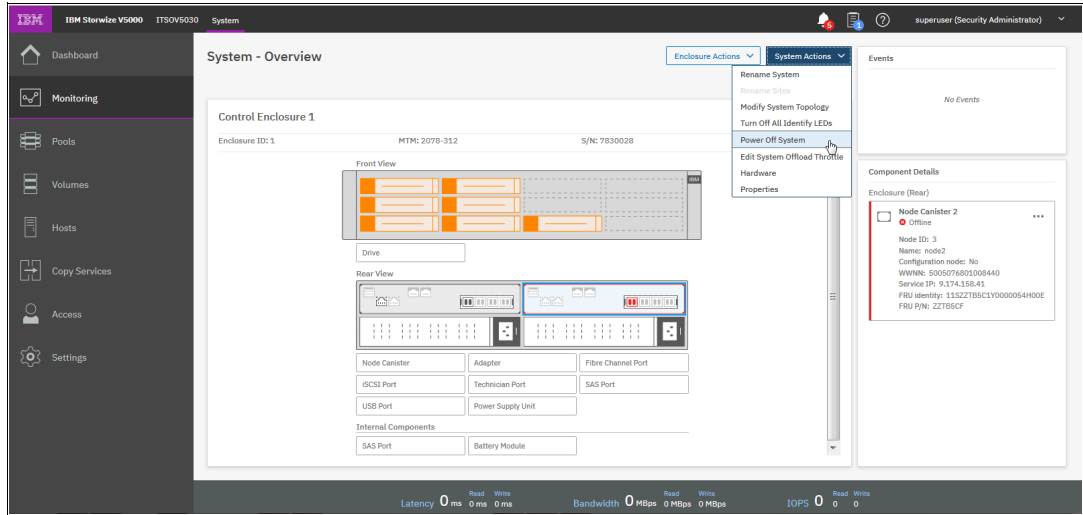


Figure 12-136 Powering off the system

2. Confirm that you want to power off the system by entering the confirmation code and clicking **OK**, as shown in Figure 12-137. Ensure that all FlashCopy, Metro Mirror, Global Mirror, data migration operations, and forced deletions are stopped or allowed to complete before you continue.

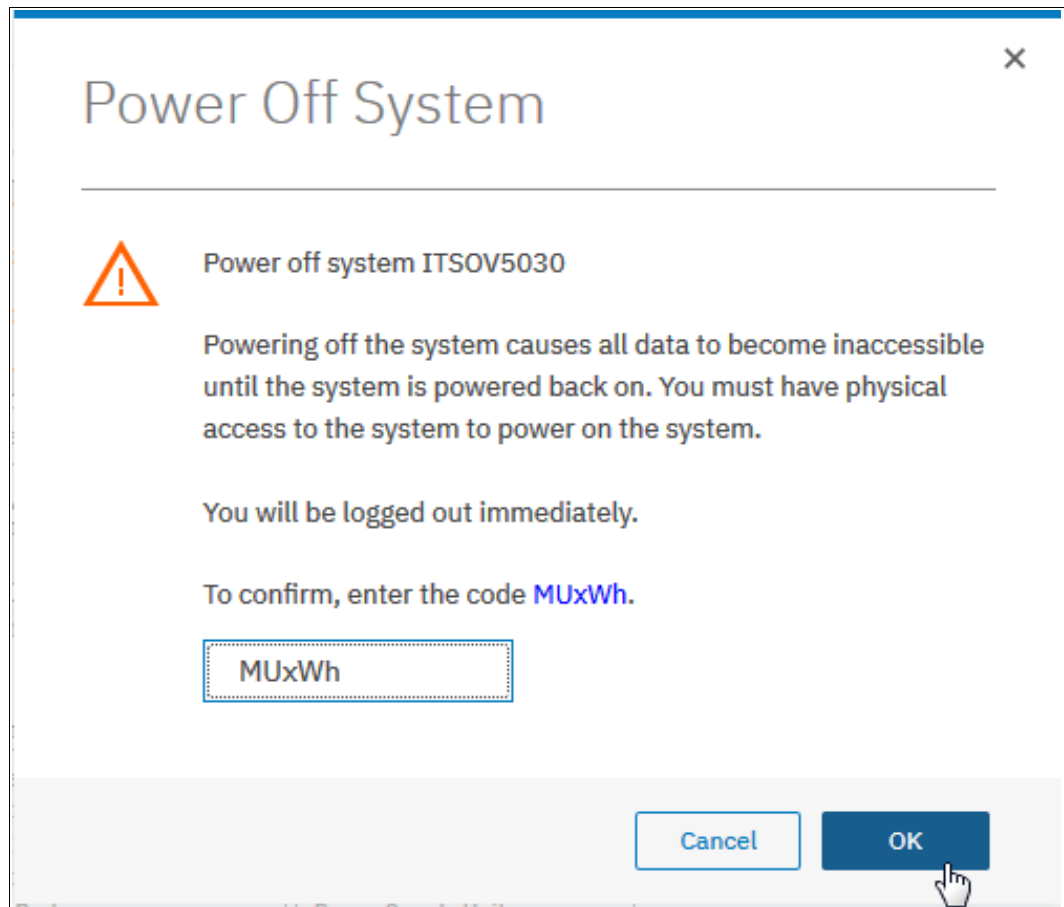


Figure 12-137 Power Off System confirmation window

3. To power off the system by using the CLI, use the command that is shown in Example 12-7. Ensure that all FlashCopy, Metro Mirror, Global Mirror, data migration operations, and forced deletions are stopped or allowed to complete before you continue.

Example 12-7 Powering off the system by using the CLI

```
>svctask stopsystem
```

```
Are you sure that you want to continue with the shut down? (y/yes to confirm)
```

4. Wait for the power LED on the node canisters to flash slowly, which indicates that the power off operation completed.

Note: When you power off an IBM Storwize V5000 Gen2, it does not automatically restart. You must manually restart the system by removing and reapplying the power or power cords.

12.10.2 Shutting down and starting up the infrastructure

To shut down an entire infrastructure (storage, servers, and applications), complete the following steps:

1. Power off your servers and all applications.
2. Power off your IBM Storwize V5000 Gen2 system by using the GUI or the CLI.
3. Remove the power cords that are connected to both power supplies in the rear of the enclosure on every control and expansion enclosure.
4. Power off your storage area network (SAN) switches.

To start an entire infrastructure, complete the following steps:

1. Power on your SAN switches and wait until the start completes.
2. Power on any expansion enclosures by connecting the power cord to both power supplies in the rear of the enclosure or by turning on the power circuit.
3. Power on the control enclosures by connecting the power cords to both power supplies in the rear of the enclosure and by turning on the power circuits.

The system starts. The system starts successfully when the status LEDs of all node canisters in the control enclosure are permanently on, which takes no longer than 10 minutes.

Power on your servers and start all applications.



Encryption

Encryption protects against the potential exposure of sensitive user data that is stored on discarded, lost, or stolen storage devices. IBM Storwize V5020, IBM Storwize V5030F, and IBM Storwize V5030 support optional encryption of data-at-rest.

IBM Storwize V5010 does not support encryption.

The chapter includes the following topics:

- ▶ 13.1, “Planning for encryption” on page 726
- ▶ 13.2, “Defining encryption of data-at-rest” on page 726
- ▶ 13.3, “Activating encryption” on page 731
- ▶ 13.4, “Enabling encryption” on page 742
- ▶ 13.5, “Configuring more providers” on page 770
- ▶ 13.6, “Migrating between providers” on page 776
- ▶ 13.7, “Recovering from a provider loss” on page 780
- ▶ 13.8, “Using encryption” on page 780
- ▶ 13.9, “Rekeying an encryption-enabled system” on page 789
- ▶ 13.10, “Disabling encryption” on page 795

13.1 Planning for encryption

Data-at-rest encryption is a powerful tool that can help organizations protect confidentiality of sensitive information. However, encryption, like any other tool, needs to be used correctly to fulfill its purpose.

Multiple drivers exist for an organization to implement data-at-rest encryption. These can be internal, such as protection of confidential company data, and ease of storage sanitization, or external, such as compliance with legal requirements or contractual obligations.

Therefore, before configuring encryption on the storage, the organization should define its needs and, if it is decided that data-at-rest encryption is required, include it in the security policy. Without defining the purpose of the particular implementation of data-at-rest encryption, it is difficult or impossible to choose the best approach to implement encryption and verify whether the implementation meets the set of goals.

The following items are worth considering during the design of a solution that includes data-at-rest encryption:

- ▶ Legal requirements
- ▶ Contractual obligations
- ▶ Organization's security policy
- ▶ Attack vectors
- ▶ Expected resources of an attacker
- ▶ Encryption key management
- ▶ Physical security

Multiple regulations mandate data-at-rest encryption, from processing of Sensitive Personal Information to the guidelines of the Payment Card Industry. If any regulatory or contractual obligations govern the data that is held on the storage system, they often provide a wide and detailed range of requirements and characteristics that need to be realized by that system. Apart from mandating data-at-rest encryption, these documents might contain requirements concerning encryption key management.

Another document that should be consulted when planning data-at-rest encryption is the organization's security policy.

The outcome of a data-at-rest encryption planning session should reply to the following questions:

- ▶ What are the goals that the organization wants to realize by using data-at-rest encryption?
- ▶ How will data-at-rest encryption be implemented?
- ▶ How can it be demonstrated that the proposed solution realizes the set of goals?

13.2 Defining encryption of data-at-rest

Encryption is the process of encoding data so that only authorized parties can read it. Secret keys are used to encode the data according to well-known algorithms.

Encryption of data-at-rest as implemented in IBM Spectrum Virtualize is defined by the following characteristics:

- ▶ *Data-at-rest* means that the data is encrypted on the end device (drives).
- ▶ The algorithm that is used is the Advanced Encryption Standard (AES) US government standard from 2001.

- ▶ Encryption of data at-rest complies with the Federal Information Processing Standard 140 (FIPS-140) standard, but is not certified.
- ▶ Ciphertext stealing XTS-AES-256 is used for data encryption.
- ▶ AES 256 is used for master access keys.
- ▶ The algorithm is public. The only secrets are the keys.
- ▶ A symmetric key algorithm is used. The same key is used to encrypt and decrypt data.

The encryption of system data and metadata is not required, so they are not encrypted.

13.2.1 Encryption methods

There are two types of encryption on devices running IBM Spectrum Virtualize: hardware encryption and software encryption. Both methods of encryption protect against the potential exposure of sensitive user data that is stored on discarded, lost, or stolen media. Both can also facilitate the warranty return or disposal of hardware.

Which method is used for encryption is chosen automatically by the system based on the placement of the data:

- ▶ Hardware encryption: Data is encrypted by using serial-attached SCSI (SAS) hardware. It is used only for internal storage (drives).
- ▶ Software encryption: Data is encrypted by using nodes' CPU (encryption code uses AES-NI CPU instruction set). It is used only for external storage.

Note: Software encryption is available in IBM Spectrum Virtualize code V7.6 and later.

Both methods of encryption use the same encryption algorithm, key management infrastructure, and license.

Note: The design for encryption is based on the concept that a system should be encrypted or not encrypted. Encryption implementation is intended to encourage solutions that contain only encrypted volumes or only unencrypted volumes. For example, after encryption is enabled on the system, all new objects (for example, pools) are by default created as encrypted.

13.2.2 Encrypted data

IBM Spectrum Virtualize performs data-at-rest encryption, which is the process of encrypting data that is stored on the end devices, such as physical drives.

Data is encrypted or unencrypted when it is written to or read from internal drives (hardware encryption) or external storage systems (software encryption).

So, data is encrypted when transferred across the SAN only between IBM Spectrum Virtualize systems and external storage. Data is *not* encrypted when transferred on SAN interfaces under the following circumstances:

- ▶ Server to storage data transfer
- ▶ Remote copy (for example, Global or Metro Mirror)
- ▶ Intracluster communication

Note: Only data-at-rest is encrypted. Host to storage communication and data sent over links used for Remote Mirroring are not encrypted.

Figure 13-1 shows an encryption example. Encrypted disks and encrypted data paths are marked in blue. Unencrypted disks and data paths are marked in red. In this example, the server sends unencrypted data to a SAN Volume Controller 2145-DH8 system, which stores hardware-encrypted data on internal disks. The data is mirrored to a remote Storwize V5000 Gen 1 system by using Remote Copy. The data flowing through the Remote Copy link is not encrypted. Because the Storwize V5000 Gen1 is unable to perform any encryption activities, data on the Storwize V5000 Gen1 is not encrypted.

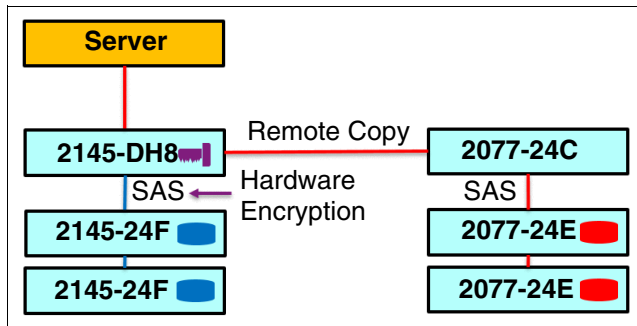


Figure 13-1 Encryption on single site

The data is mirrored to a remote Storwize V5000 Gen 1 system by using Remote Copy. The data flowing through the Remote Copy link is not encrypted. Because the Storwize V5000 Gen1 is unable to perform any encryption activities, data on the Storwize V5000 Gen1 is not encrypted.

To enable encryption of both data copies, the Storwize V5000 Gen1 must be replaced by an encryption capable IBM Spectrum Virtualize system, as shown in Figure 13-2. After such replacement, both copies of data are encrypted, but the Remote Copy communication between both sites remains unencrypted.

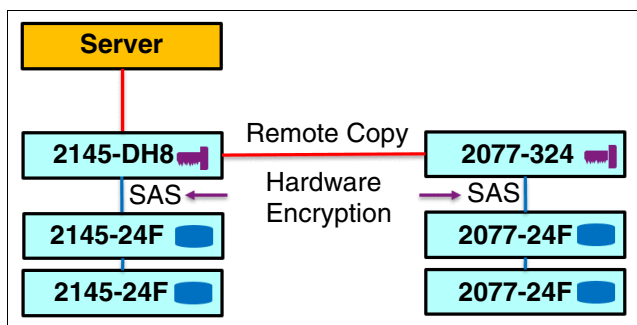


Figure 13-2 Encryption on both sites

Figure 13-3 shows an example configuration that uses both software and hardware encryption. Software encryption is used to encrypt an external virtualized storage system (20777-24C in Figure 13-3). Hardware encryption is used for internal, SAS-attached disk drives.

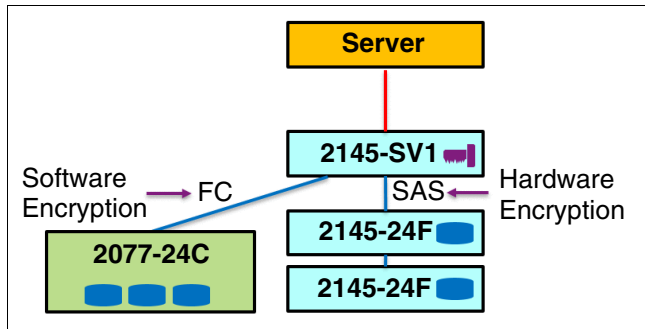


Figure 13-3 Example of software encryption and hardware encryption

Placement of hardware encryption and software encryption in the Storwize code stack are shown in Figure 13-4. The functions that are implemented in software are shown in blue. The external storage system is shown in yellow. The hardware encryption on the SAS chip is marked in pink. Compression is performed before encryption. Therefore, it is possible to realize benefits of compression for the encrypted data.

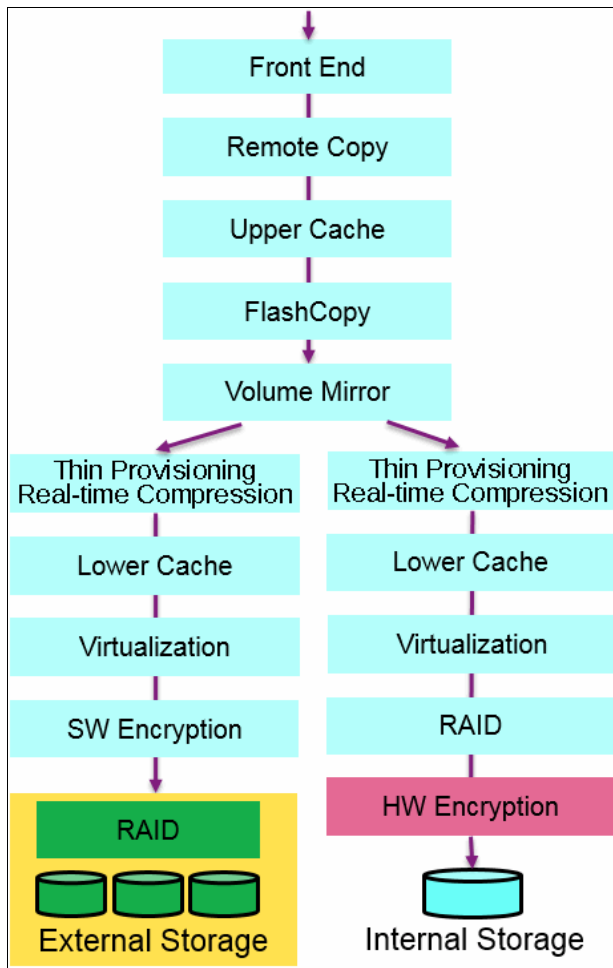


Figure 13-4 Encryption placement in the IBM Spectrum Virtualize software stack

Each volume copy can use different encryption methods (hardware, software). It is also allowed to have volume copies with different encryption status (encrypted versus unencrypted). The encryption method depends only on the pool that is used for the specific copy. You can migrate data between different encryption methods by using volume migration or volume mirroring.

13.2.3 Encryption keys

Hardware and software encryption use the same encryption key infrastructure. The only difference is the object that is encrypted by using the keys. The following objects can be encrypted:

- ▶ Pools (software encryption)
- ▶ Child pools (software encryption)
- ▶ Arrays (hardware encryption)

Consider the following points regarding encryption keys:

- ▶ Keys are unique for each object, and they are created when the object is created.
- ▶ Two types of keys are defined in the system:
 - Master access key:
 - The master access key is created when encryption is enabled.
 - The master access key can be stored on USB flash drives or key servers. One master access key is created for each enabled encryption key provider.
 - It can be copied or backed up as necessary.
 - It is *not* permanently stored anywhere in the system.
 - It is required at boot time to unlock access to encrypted data.
 - Data encryption keys (one for each encrypted object):
 - Data encryption keys are used to encrypt data. When an encrypted object (such as an array, a pool, or a child pool) is created, a new data encryption key is generated for this object.
 - Managed disks (MDisk) that are not self-encrypting are automatically encrypted by using the data encryption key of the pool or child pool to which they belong.
 - MDisk that are self-encrypting are not reencrypted by using the data encryption key of the pool or child pool they belong to by default. You can override this default by manually configuring the MDisk as not self-encrypting.
 - Data encryption keys are stored in secure memory.
 - During cluster internal communication data encryption keys are encrypted with the master access key.
 - Data encryption keys cannot be viewed.
 - Data encryption keys cannot be changed.
 - When an encrypted object is deleted, its data encryption key is discarded (*secure erase*).

Important: If all master access key copies are lost and the system must cold restart, all encrypted data is gone. No method exists, even for IBM, to decrypt the data without the keys. If encryption is enabled and the system cannot access the master access key, all SAS hardware is offline, including unencrypted arrays.

Note: A self-encrypting MDisk is an MDisk from an encrypted volume in an external storage system.

13.2.4 Encryption licenses

Encryption is a licensed feature that uses key-based licensing.

No trial licenses for encryption exist on the basis that when the trial ends, the access to the data is lost. Therefore, you must purchase an encryption license before you activate encryption. Licenses are generated by IBM Data storage feature activation (DSFA) based on the serial number (S/N) and the machine type and model number (MTM) of the nodes.

You can activate an encryption license during the initial system setup (in the Encryption window of the initial setup wizard) or later on, in the running environment.

Contact your IBM marketing representative or IBM Business Partner to purchase an encryption license.

13.3 Activating encryption

Encryption is enabled at a system level and all of the following prerequisites must be met *before* you can use encryption:

- ▶ You must purchase an encryption license before you activate the function.
If you did not purchase a license, contact an IBM marketing representative or IBM Business Partner to purchase an encryption license.
- ▶ At least three USB flash drives are required if you plan not to use a key management server. They are available as a feature code from IBM (see the note on 745).
- ▶ You must activate the license that you purchased.
- ▶ Encryption must be enabled.

After purchase an encryption license, the first step to use encryption is to retrieve it and activate the license.

Activation of the license can be performed in one of two ways:

- ▶ Automatically activation
This method is used when you have the authorization code and the workstation that is being used to activate the license can access the external network. In this case, you must enter only the authorization code, and the license key is automatically obtained from the internet and activated in the IBM Spectrum Virtualize system.
- ▶ Manually activation
If you cannot activate the license automatically because any of the requirements are not met, you can follow the instructions that are provided in the GUI to obtain the license key from the web and activate in the IBM Spectrum Virtualize system.

Both methods are available during the initial system setup and when the system is in use.

13.3.1 Obtaining an encryption license

You must purchase an encryption license before you activate encryption. If you did not purchase a license, contact an IBM marketing representative or IBM Business Partner to purchase an encryption license.

When you purchase a license, you receive a function authorization document that includes an authorization code printed. This code allows you to proceed using the automatic activation process.

If the automatic activation process fails or if you prefer using the manual activation process, see [this web page](#) to retrieve your license keys.

Ensure that you have the following information:

- ▶ Machine type (MT)
- ▶ Serial number (S/N)
- ▶ Machine signature
- ▶ Authorization code

For more information about how to retrieve the machine signature of a node, see 13.3.5, “Activating the license manually” on page 740.

13.3.2 Start activation process during initial system setup

One of the steps in the initial setup enables encryption license activation. The system asks “Was the encryption feature purchased for this system?”.

To activate encryption at this stage, complete the following steps:

1. Select **Yes**, as shown in Figure 13-5.

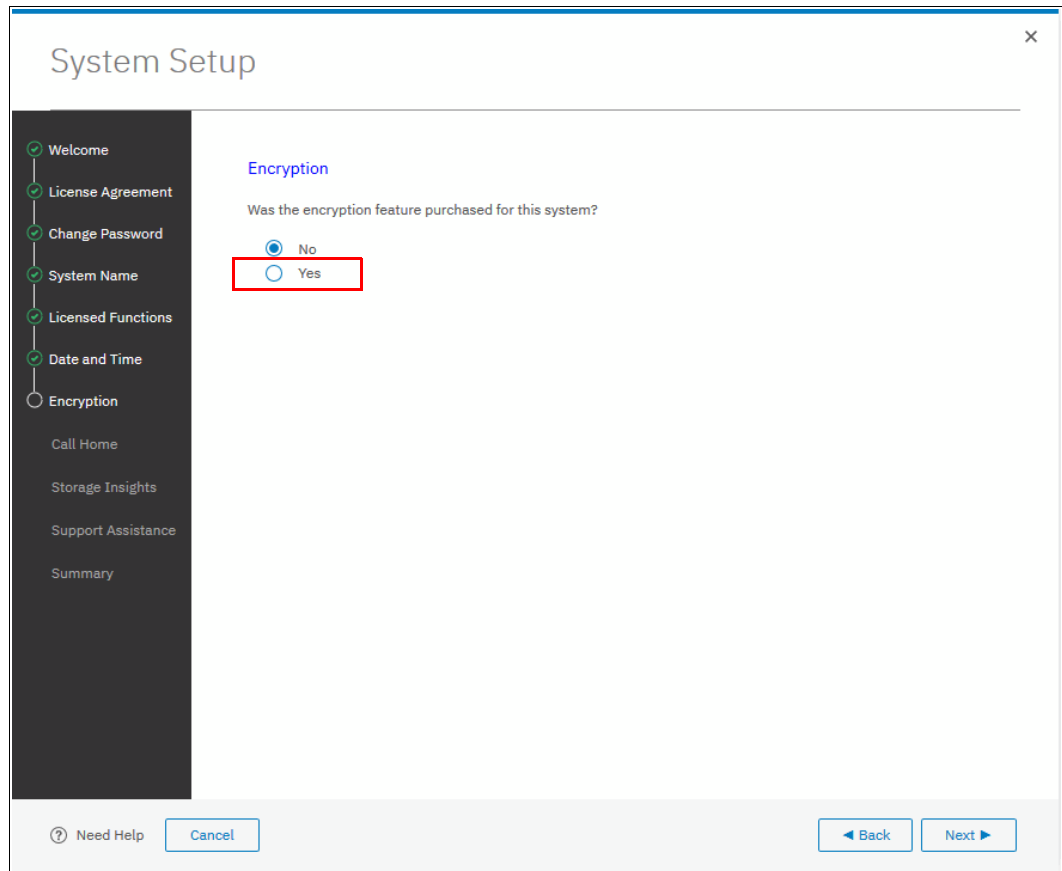


Figure 13-5 Encryption activation during initial system setup

- The Encryption window displays information about your storage system, as shown in Figure 13-6.

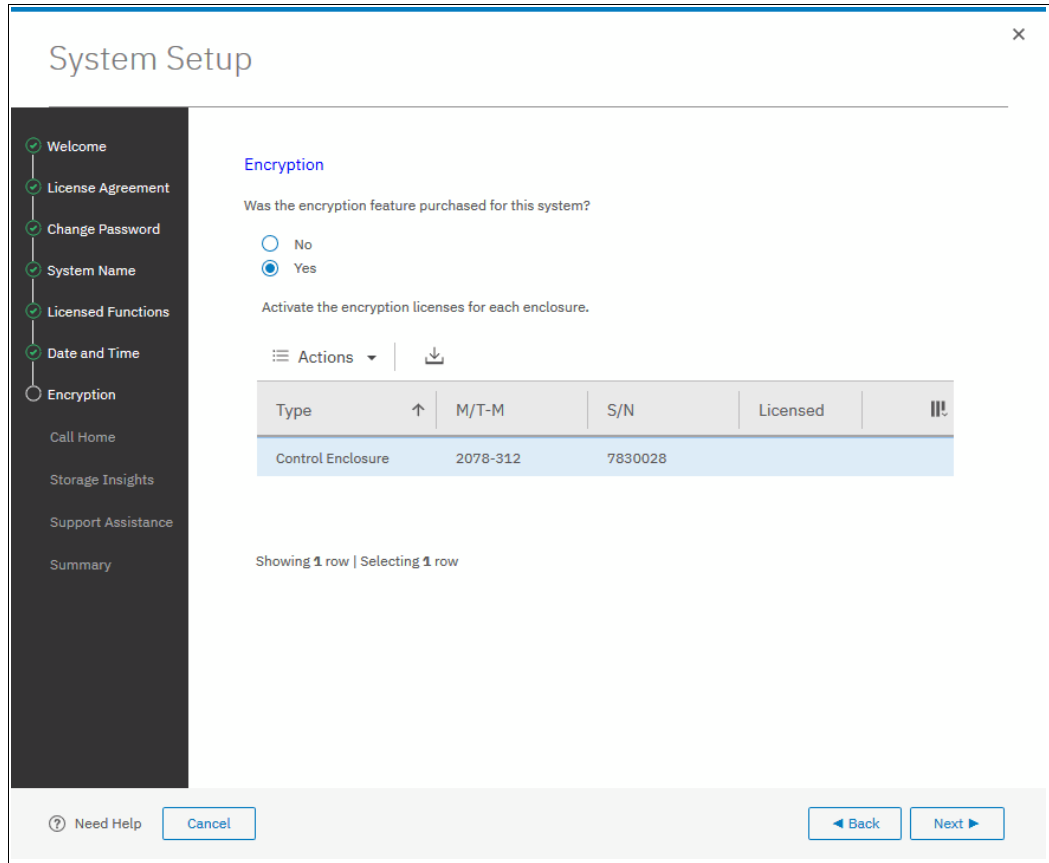


Figure 13-6 Information storage system during initial system setup

- Right-clicking the node opens a menu with two license activation options (Activate License Automatically and Activate License Manually), as shown in Figure 13-7. Use either option to activate encryption.

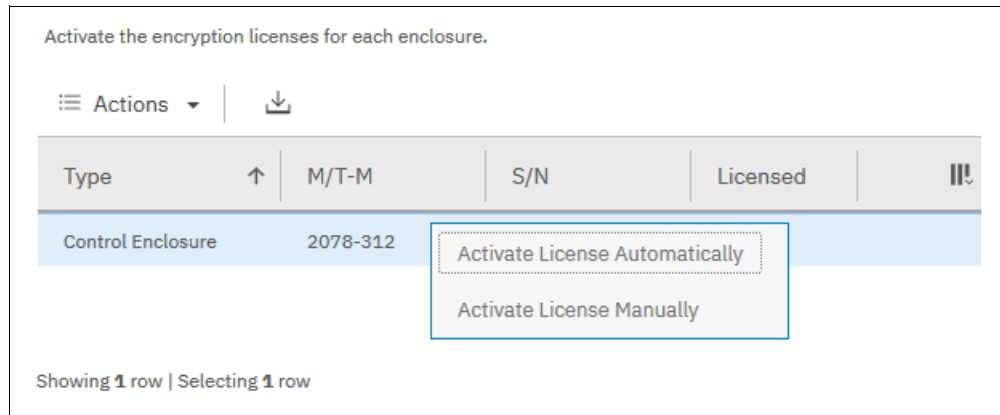


Figure 13-7 Selecting license activation method

For more information about how to complete an automatic activation process, see 13.3.4, “Activate the license automatically” on page 737.

For more information about how to complete a manual activation process, see “Activating the license manually” on page 740.

After the activation process is complete, you see a green check mark in the column labeled Licensed next to a node for which the license was enabled. You can proceed with the initial system setup by clicking **Next**, as shown in Figure 13-8.

Note: Every enclosure needs an active encryption license before you can enable encryption on the system. Attempting to add a non-licensed enclosure to an encryption-enabled system fails.

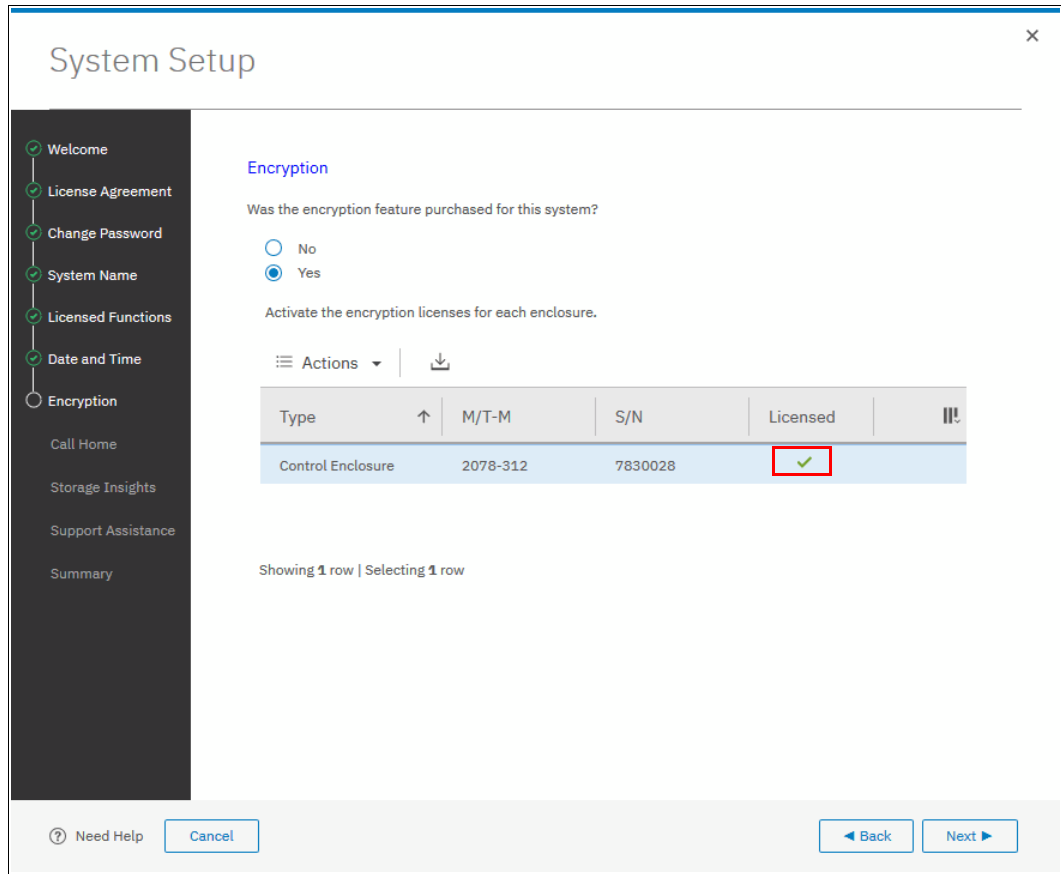


Figure 13-8 Successful encryption license activation during initial system setup

13.3.3 Start activation process on a running system

To activate encryption on a running system, complete the following steps:

1. Click **Settings** → **System** → **Licensed Functions**.
2. Click **Encryption Licenses**, as shown in Figure 13-9.

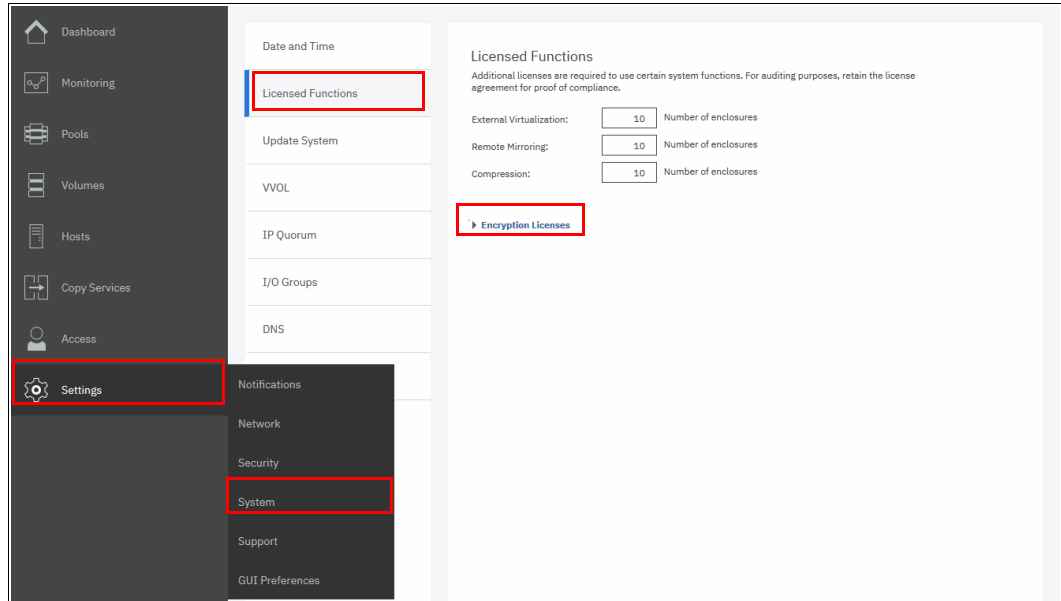


Figure 13-9 Expanding Encryption Licenses section on the Licensed Functions window

3. The Encryption Licenses window displays information about your control enclosures. Right-click the enclosure on which you want to install an encryption license. This action opens a menu with two license activation options (Activate License Automatically and Activate License Manually), as shown in Figure 13-10. Use either option to activate encryption.

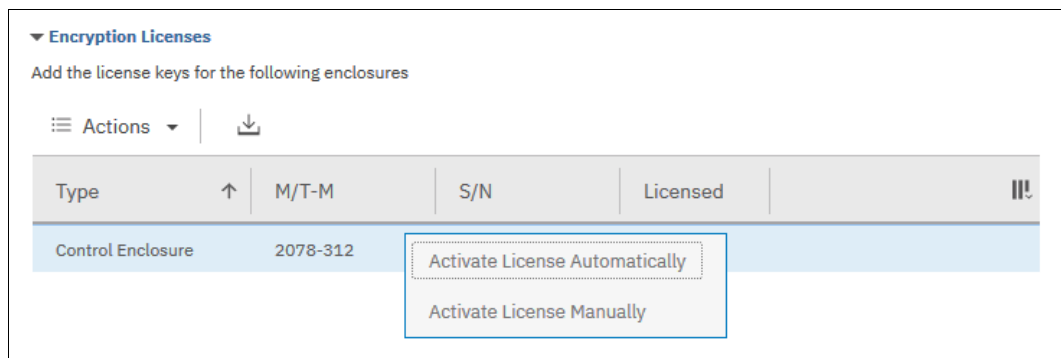


Figure 13-10 Select the Control Enclosure on which you want to enable the encryption

For more information about how to complete an automatic activation process, see 13.3.4, “Activate the license automatically” on page 737.

For more information about how to complete a manual activation process, see “Activating the license manually” on page 740.

4. After the activation process is complete, you see a green check mark in the column labeled Licensed for the control enclosure, as shown in Figure 13-11.

Type	M/T-M	S/N	Licensed
Control Enclosure	2078-312	7830028	✓

Figure 13-11 Successful encryption license activation on a running system

13.3.4 Activate the license automatically

Automatic license activation is the faster method to activate the encryption license for IBM Spectrum Virtualize. You need the authorization code and the workstation that is used to access the GUI to access to the external network.

Important: To perform this operation, the personal computer that is used to connect to the GUI and activate the license must connect to the internet.

To activate the encryption license for a control enclosure automatically, complete the following steps:

1. Select **Activate License Automatically** to open the Activate License Automatically window, as shown in Figure 13-12.

Activate License Automatically

Enter the authorization code for node:

Type or paste the authorization code here

Cancel Activate

Figure 13-12 Encryption license Activate License Automatically window

2. Enter the authorization code that is specific to the control enclosure that you selected, as shown in Figure 13-13. Click **Activate**.

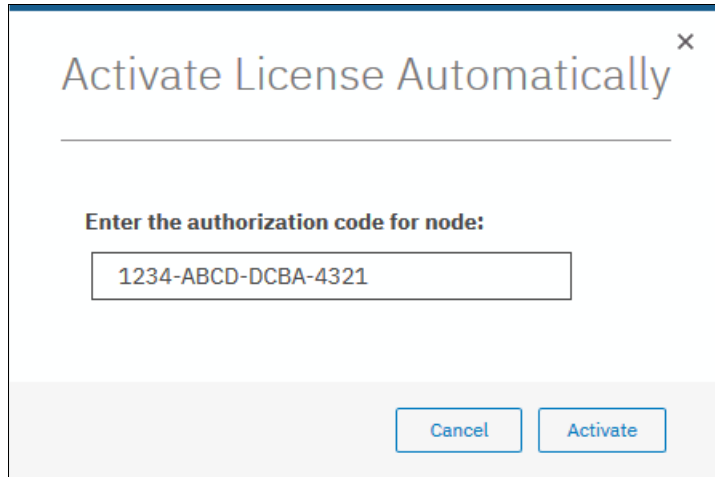


Figure 13-13 Entering an authorization code

The system connects to IBM to verify the authorization code and retrieve the license key. Figure 13-14 shows a window that is displayed during this connection. If the process is successful, the procedure takes less than a minute.

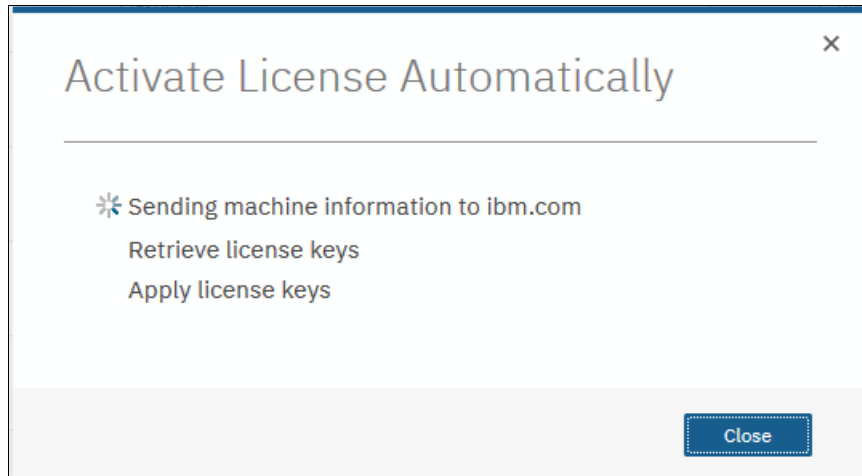


Figure 13-14 Activating encryption

After the license key is retrieved, it is automatically applied, as shown in Figure 13-15.



Figure 13-15 Successful encryption license activation

Problems with automatic license activation

If connections problems occur with the automatic license activation procedure, the system times out after 3 minutes with an error.

Check whether the personal computer that is used to connect to the Storwize V5000 GUI and activate the license can access the internet. If you are unable to complete the automatic activation procedure, use the manual activation procedure that is described in 13.3.5, “Activating the license manually” on page 740.

Although authorization codes and encryption license keys use the same format (four groups of four hexadecimal digits), you can only use each of them in the appropriate activation process. If you use a license key when the system expects an authorization code, the system displays an error message, as shown in Figure 13-16.

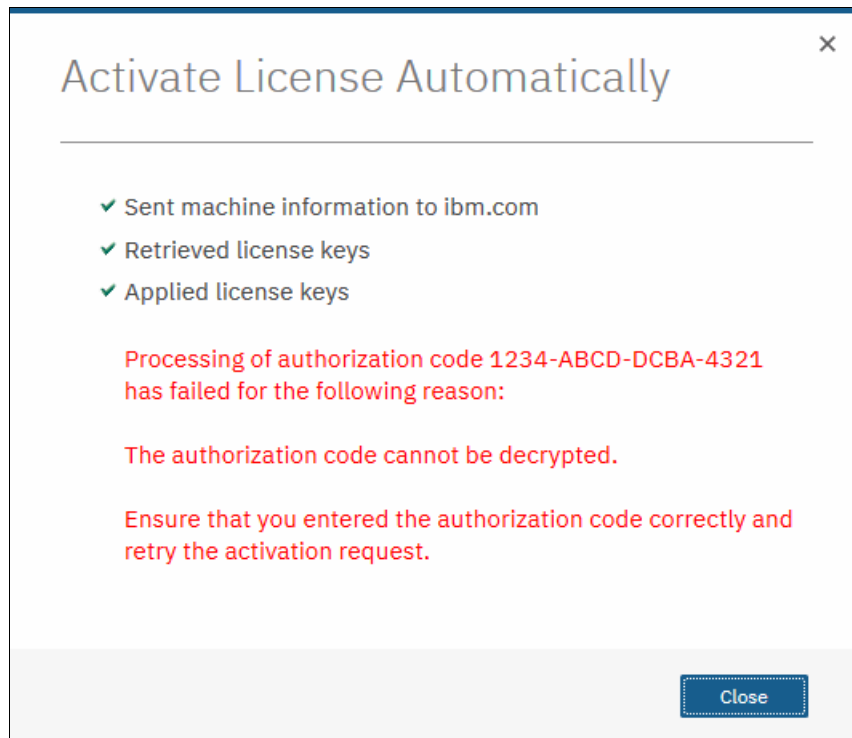


Figure 13-16 Authorization code failure

13.3.5 Activating the license manually

To manually activate the encryption license for a control enclosure, complete the following steps:

1. Select **Activate License Manually** to open the Manual Activation window, as shown in Figure 13-17.

Manual Activation

Enter the license key:

Type or paste a license key of any format here

1. Go to <https://www.ibm.com/storage/dsfa>
2. Select Storwize
3. Enter the following information:
 - o Machine type: 2078
 - o Serial number: 7830028
 - o Machine signature: 2710-9AE8-A266-D386
4. Enter the authorization codes that were sent with your purchase agreement for the licensed function.
5. Copy or download the keys.

Cancel Activate

Figure 13-17 Manual encryption license activation window

2. If you have not done so already, obtain the encryption license for the control enclosure. The information that is required to obtain the encryption license is displayed in the Manual Activation window. Use this data to follow the instructions in 13.3.1, “Obtaining an encryption license” on page 732.

3. You can enter or paste the license key or click the folder icon and upload it to the storage system that the license key file downloaded from DSFA. In Figure 13-18, the sample key is already entered. Click **Activate**.

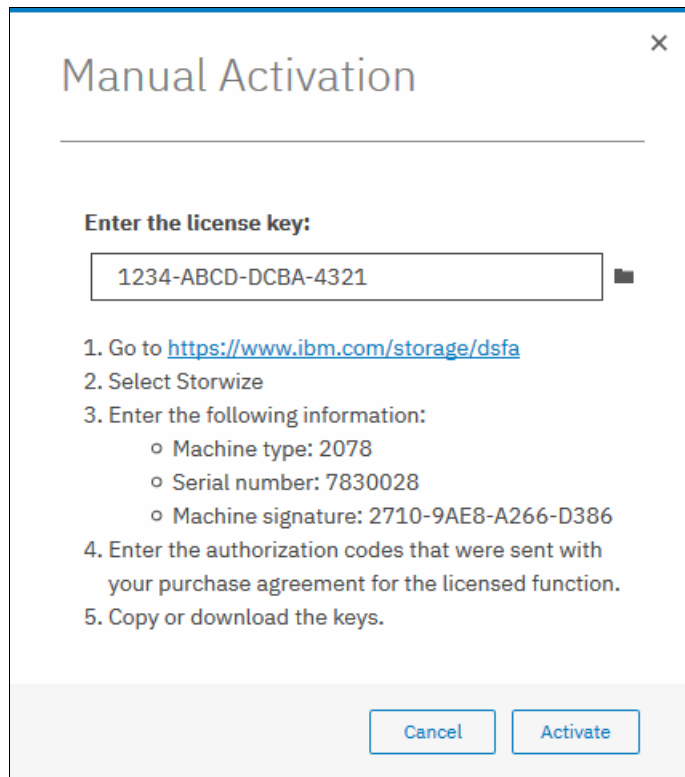


Figure 13-18 Entering an encryption license key

After the task completes successfully, the GUI shows that encryption is licensed for the specified control enclosure, as shown in Figure 13-19.

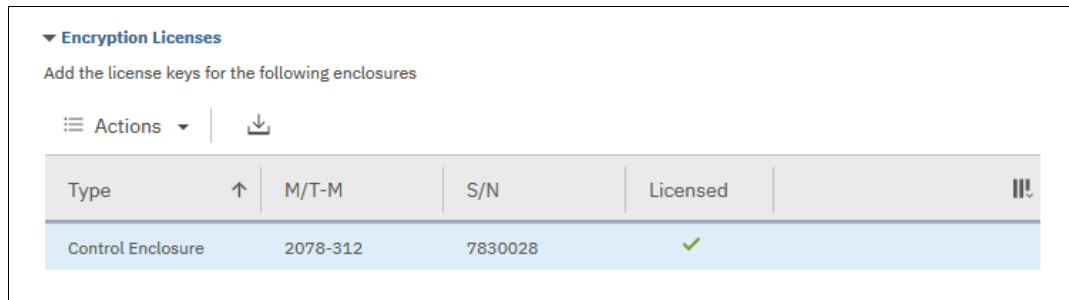


Figure 13-19 Successful encryption license activation

Problems with manual license activation

Although authorization codes and encryption license keys use the same format (four groups of four hexadecimal digits), you can only use each of them in the appropriate activation process. If you use an authorization code when the system expects a license key, the system displays an error message, as shown in Figure 13-20.

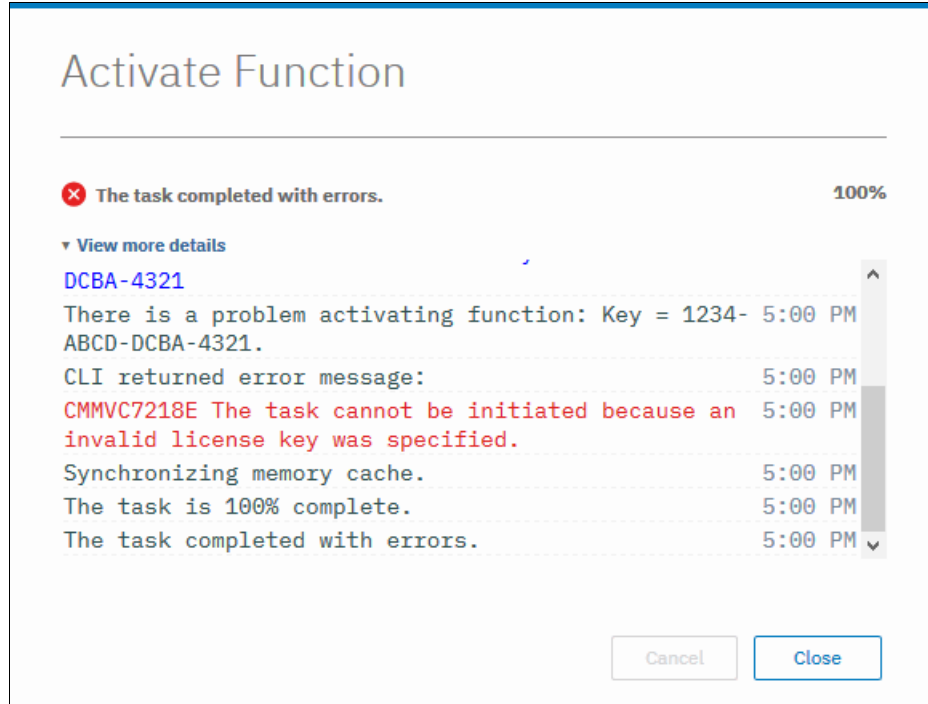


Figure 13-20 License key failure

13.4 Enabling encryption

This section describes the process to create and store system master access key copies, also referred to as *encryption keys*. These keys can be stored on any or both of two key providers: USB flash drives or a key server.

The following types of key servers are supported by IBM Spectrum Virtualize:

- ▶ IBM Security Key Lifecycle Manager (SKLM), introduced in IBM Spectrum Virtualize V7.8.
- ▶ Gemalto SafeNet KeySecure, introduced in IBM Spectrum Virtualize V8.2.

IBM Spectrum Virtualize code V8.1 introduced the ability to define up to four encryption key servers, which is a preferred configuration because it increases key provider availability. In this version, support for simultaneous use of both USB flash drives and key server was added.

Organizations that use encryption key management servers might consider parallel use of USB flash drives as a backup solution. During normal operation, such drives can be disconnected and stored in a secure location. However, during a catastrophic loss of encryption servers, the USB drives can still be used to unlock the encrypted storage.

The characteristics of key servers and USB flash drive might help you to choose the type of encryption key provider that you want to use.

Key servers can have the following characteristics:

- ▶ Physical access to the system is not required to perform a rekey operation.
- ▶ Support for businesses that have security requirements that preclude use of USB ports.
- ▶ Possibility to use hardware security modules (HSMs) for encryption key generation.
- ▶ Ability to replicate keys between servers and perform automatic backups.
- ▶ Implementations follow an open standard (Key Management Interoperability Protocol [KMIP]) that aids in interoperability.
- ▶ Ability to audit operations related to key management.
- ▶ Ability to separately manage encryption keys and physical access to storage systems.

USB flash drives have the following characteristics:

- ▶ Physical access to the system might be required to process a rekey operation.
- ▶ No moving parts with almost no read or write operations to the USB flash drive.
- ▶ Inexpensive to maintain and use.
- ▶ Convenient and easy to have multiple identical USB flash drives available as backups.

Important: Maintaining confidentiality of the encrypted data depends on security of the encryption keys. Pay special attention to ensure secure creation, management, and storage of the encryption keys.

13.4.1 Starting the Enable Encryption wizard

After the license activation step is successfully completed on IBM Storwize V5000 control enclosures, you can now enable encryption. You can enable encryption after completing the initial system setup by using the GUI or command-line interface (CLI).

The Enable Encryption wizard can be started in the GUI by clicking **Run Task** next to Enable Encryption on the Suggested Tasks window, as shown in Figure 13-21.

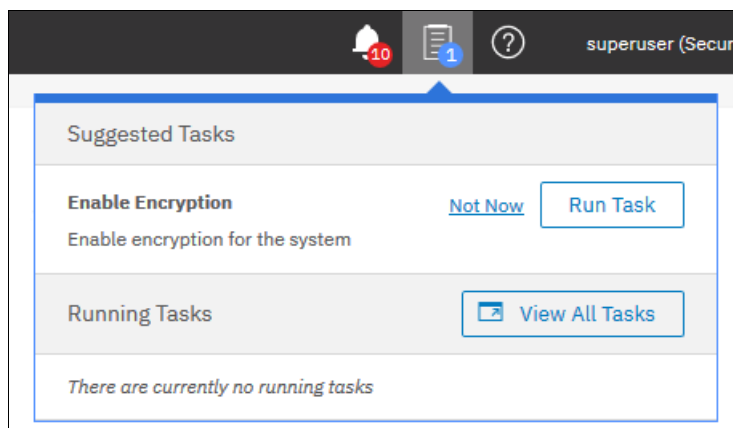


Figure 13-21 Enable Encryption from the Suggested Tasks window

The wizard also can be started by clicking **Settings** → **Security** → **Encryption** and then, clicking **Enable Encryption**, as shown in Figure 13-22.

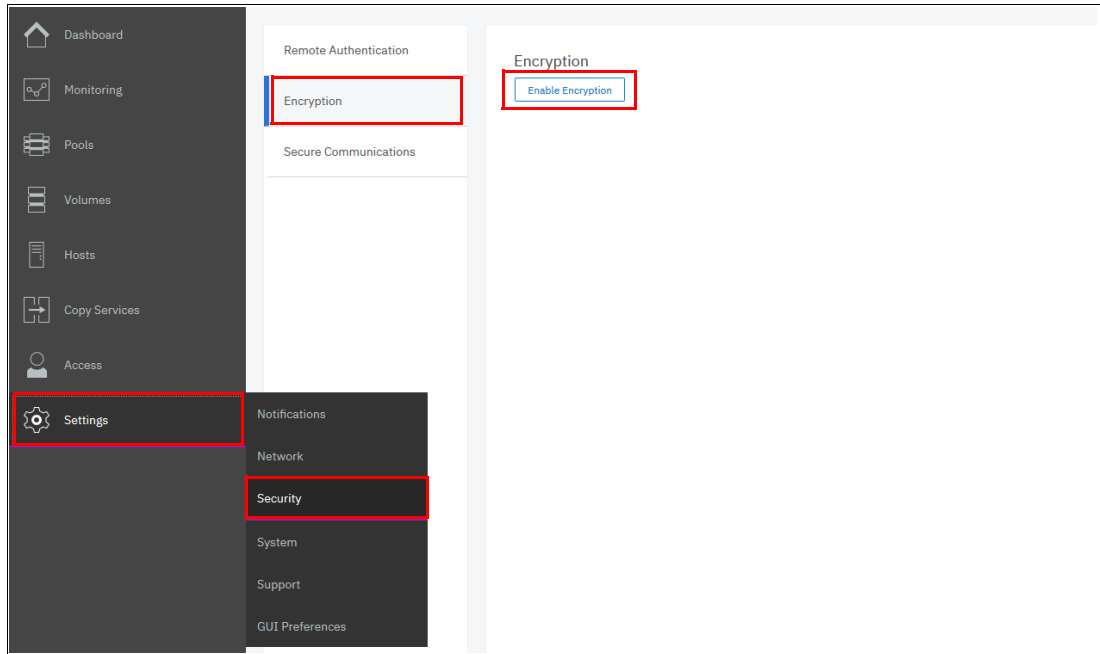


Figure 13-22 Enable Encryption from the Security pane

The Enable Encryption wizard starts by asking which encryption key provider to use for storing the encryption keys, as shown in Figure 13-23. You can enable either or both providers.

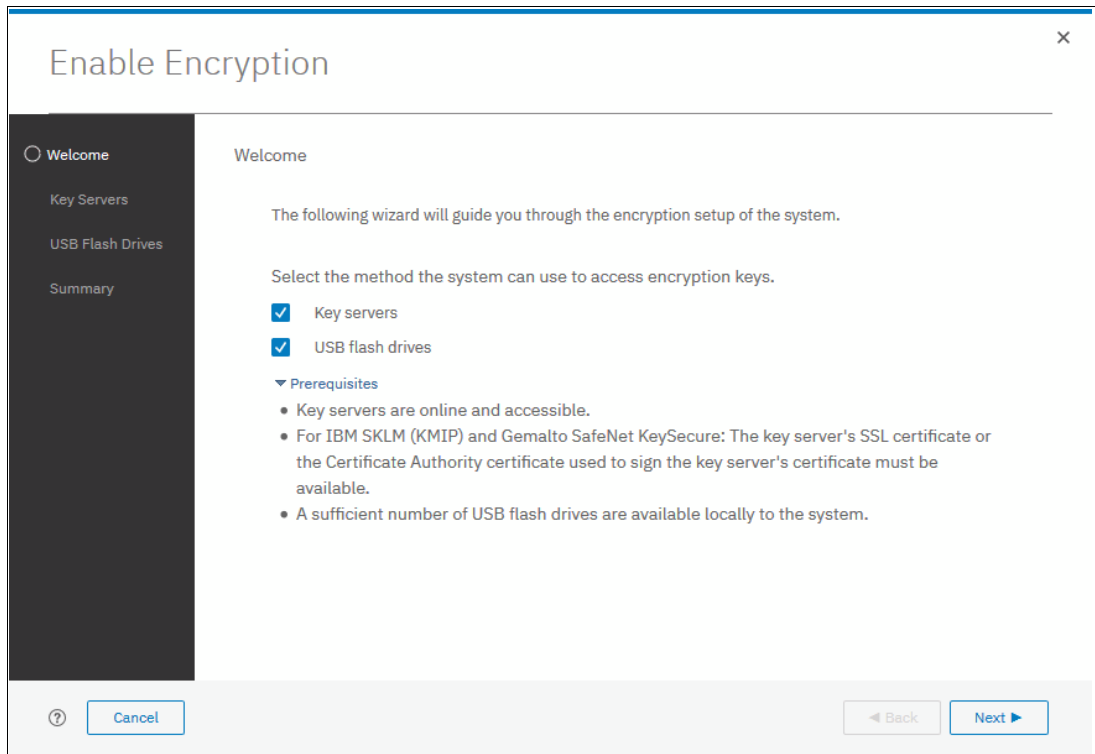


Figure 13-23 Enable Encryption wizard Welcome window

The next section presents a scenario in which both encryption key providers are enabled at the same time.

For more information about for instructions on how to enable encryption by using only USB flash drives, see 13.4.2, “Enabling encryption using USB flash drives”.

For more information about how to enable encryption by using key servers as the sole encryption key provider, see 13.4.3, “Enabling encryption using key servers” on page 750.

13.4.2 Enabling encryption using USB flash drives

Note: The system needs at least three USB flash drives to be present before you can enable encryption by using this encryption key provider. IBM USB flash drives are preferred and can be obtained from IBM with the feature name Encryption USB Flash Drives (Four Pack). But, other flash drives also might work. You can use the USB port in any node of the cluster.

Using USB flash drives as the encryption key provider requires a minimum of three USB flash drives to store the generated encryption keys. Because the system attempts to write the encryption keys to any USB key inserted into a node port, it is critical to maintain physical security of the system during this procedure.

While the system enables encryption, you are prompted to insert two USB flash drives into the system, each drive in the USB port of each node. After the encryption key is copied to the first two USB flash drives, the system prompts you to remove the two flash drives and to insert the third USB flash drive. When the final copy completes, you can create any other backup copies by repeating the process. When the system detects the USB flash drives, the encryption key is automatically copied to the USB flash drives.

Ensure that each copy of the encryption key is valid before you write any user data to the system. The system validates any key material on a USB flash drive when it is inserted into the canister. If the key material is not valid, the system logs an error. If the USB flash drive is unusable or fails, the system does not display it as output. Figure 13-26 on page 748 shows an example where the system detected and validated three USB flash drives.

If your system is in a secure location with controlled access, one USB flash drive for each canister can remain inserted in the system. If there is a risk of unauthorized access, all USB flash drives with the master access keys must be removed from the system and stored in a secure place.

Securely store all copies of the encryption key. For example, any USB flash drives that hold an encryption key copy that are not left plugged into the system can be locked in a safe. Similar precautions must be taken to protect any other copies of the encryption key that are stored on other media.

Notes: Generally, create at least one other copy on another USB flash drive for storage in a secure location. You can also copy the encryption key from the USB drive and store the data on other media, which can provide more resilience and mitigate risk that the USB drives used to store the encryption key come from a faulty batch.

Every encryption key copy must be stored securely to maintain confidentiality of the encrypted data.

A minimum of one USB flash drive with the correct master access key is required to unlock access to encrypted data after a system restart such as a system-wide restart or power loss. No USB flash drive is required during a warm restart, such as a node exiting service mode or a single node restart. The data center power-on procedure needs to ensure that USB flash drives containing encryption keys are plugged into the storage system before it is powered on.

During power-on, insert USB flash drives into the USB ports on two supported canisters to safeguard against failure of a node, node's USB port, or USB flash drive during the power-on procedure.

To enable encryption by using USB flash drives as the only encryption key provider, complete the following steps:

1. In the Enable Encryption wizard Welcome tab, select **USB flash drives** and click **Next**, as shown in Figure 13-24.

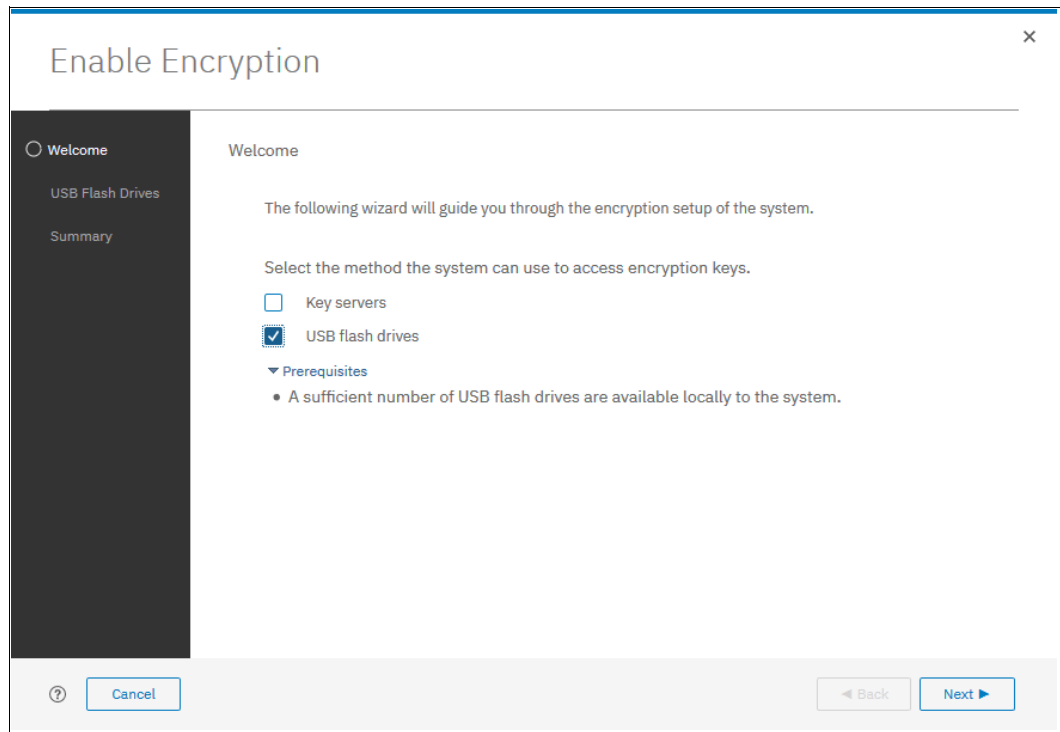


Figure 13-24 Selecting USB flash drives in the Enable Encryption wizard

2. If there are fewer than two USB flash drives inserted into the system, you are prompted to insert more drives, as shown in Figure 13-25. The system reports how many more drives need to be inserted.

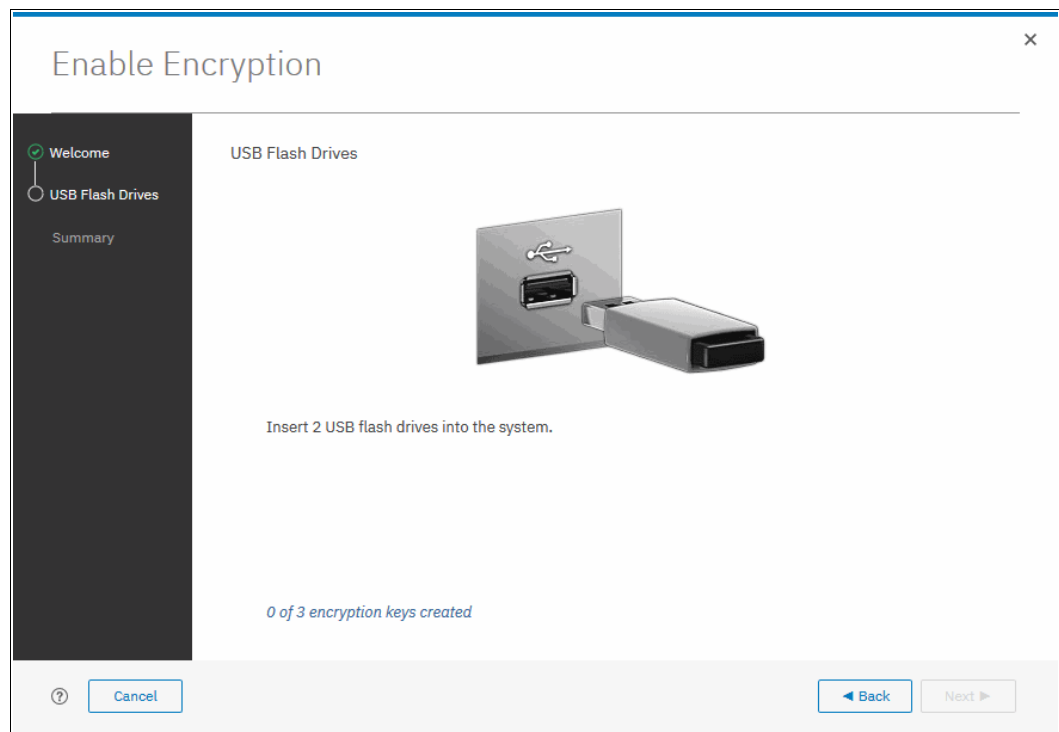


Figure 13-25 Waiting for USB flash drives to be inserted

Note: The **Next** option remains disabled until at least two USB flash drives are detected.

3. Insert the USB flash drives into the USB ports as requested.

4. After the encryption key is copied to the first two USB flash drives, the management GUI prompts you to remove the two flash drives. After you remove the flash drives, insert the last required flash drive into the system. Figure 13-26 shows the wizard after three encryption keys were created and copied to USB flash drives.

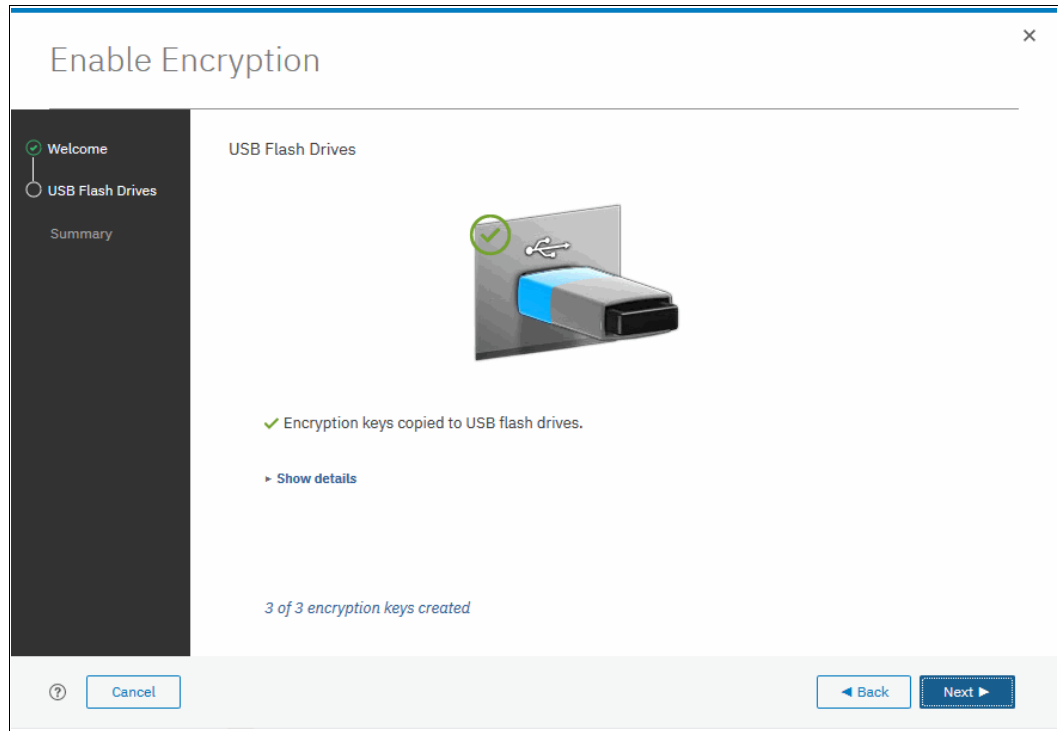


Figure 13-26 Writing the master access key to USB flash drives

You can keep adding USB flash drives or replacing the ones that are plugged in to create copies. When done, click **Next**.

- The number of keys that were created is shown in the Summary tab, as shown in Figure 13-27. Click **Finish** to finalize the encryption enablement.

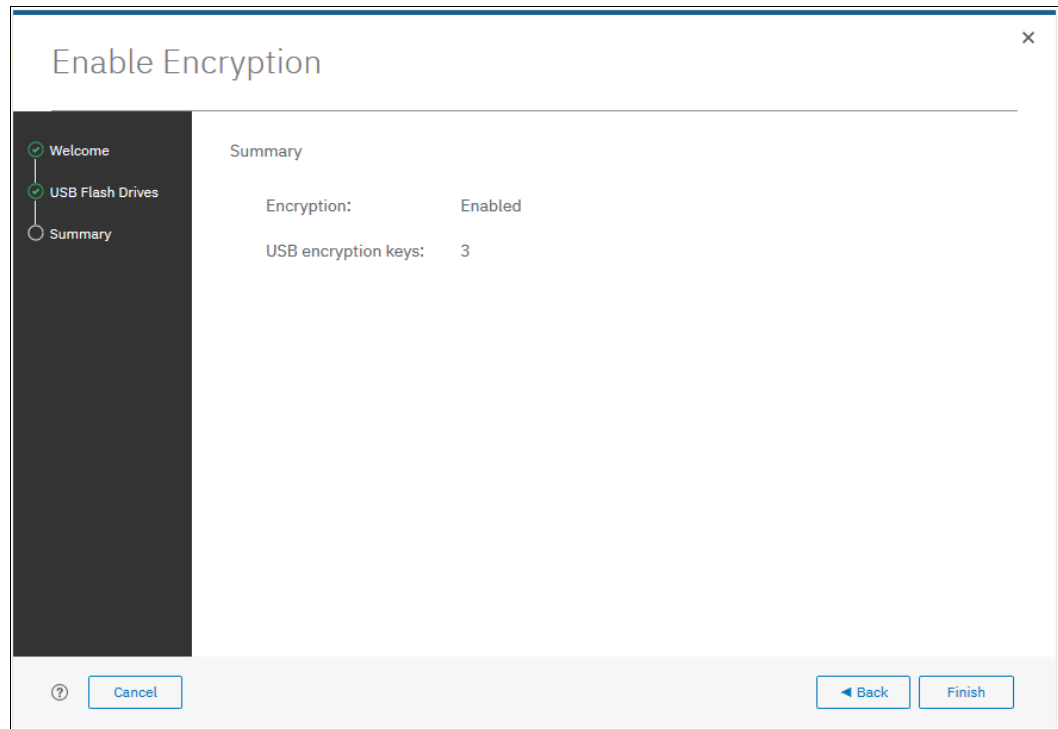


Figure 13-27 Commit the encryption enablement

- You receive a message confirming that the encryption is now enabled on the system, as shown in Figure 13-28.



Figure 13-28 Encryption enabled message using USB flash drives

7. You can confirm that encryption is enabled and verify which key providers are in use by going to **Settings** → **Security** → **Encryption**, as shown in Figure 13-29.

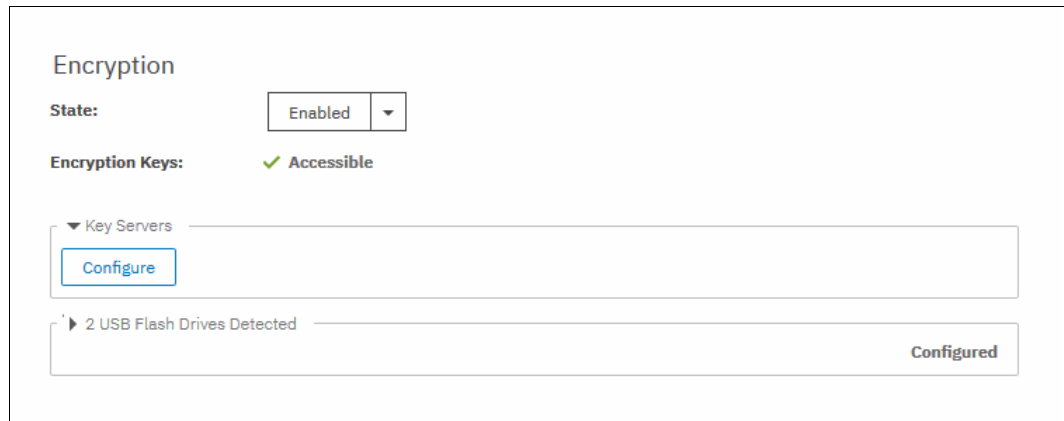


Figure 13-29 Encryption view showing using USB flash drives as the enabled provider

13.4.3 Enabling encryption using key servers

A key server is a centralized system that receives and then distributes encryption keys to its clients, including IBM Spectrum Virtualize systems.

IBM Spectrum Virtualize supports use of the following key servers as encryption key providers:

- ▶ IBM Security Key Lifecycle Manager (SKLM)
- ▶ Gemalto SafeNet KeySecure

Note: Support for IBM Security Key Lifecycle Manager was introduced in IBM Spectrum Virtualize V7.8. Support for Gemalto SafeNet KeySecure was introduced in IBM Spectrum Virtualize V8.2.1.

SKLM and KeySecure SafeNet support Key Management Interoperability Protocol (KMIP), which is a standard for management of cryptographic keys.

Note: Make sure that the key management server functionality is fully independent from encrypted storage, which has encryption managed by this key server environment. Failure to observe this requirement might create an encryption deadlock. An encryption deadlock is a situation in which none of key servers in the environment can become operational because some critical part of the data in each server is stored on a storage system that depends on one of the key servers to unlock access to the data.

IBM Spectrum Virtualize code V8.1 and later supports up to four key server objects defined in parallel. However, only one key server type (SKLM or KeySecure) can be enabled at one time.

Another characteristic when working with key servers is that it is not possible to migrate from one key server type directly to another. If you want to migrate from one type to another, you first need to migrate from your current key server to USB encryption, and then, migrate from USB to the other type of key server.

Enabling encryption using SKLM

Before you create a key server object in the storage system, the key server must be configured. Ensure that you complete the following tasks on the SKLM server before you enable encryption on the storage system:

- ▶ Configure the SKLM server to use Transport Layer Security version 1.2. The default setting is TLSv1, but IBM Spectrum Virtualize supports only version 1.2. Therefore, set the value of security protocols to SSL_TLSv2 (which is a set of protocols that includes TLSv1.2) in the SKLM server configuration properties.
- ▶ Ensure that the database service is started automatically on startup.
- ▶ Ensure that there is at least one Secure Sockets Layer (SSL) certificate for browser access.
- ▶ Create a SPECTRUM_VIRT device group for IBM Spectrum Virtualize systems.

For more information about completing these tasks, see SKLM documentation at [IBM Knowledge Center](#).

Access to the key server storing the correct master access key is required to enable encryption for the cluster after a system restart, such as a system-wide restart or power loss. Access to the key server is not required during a warm restart, such as a node exiting service mode or a single node restart. The data center power-on procedure must ensure key server availability before the storage system using encryption is booted.

To enable encryption by using an SKLM key server, complete the following steps:

1. Ensure that you have service IPs configured on all your nodes.
2. In the Enable Encryption wizard Welcome tab, select **Key servers** and click **Next**, as shown in Figure 13-30.

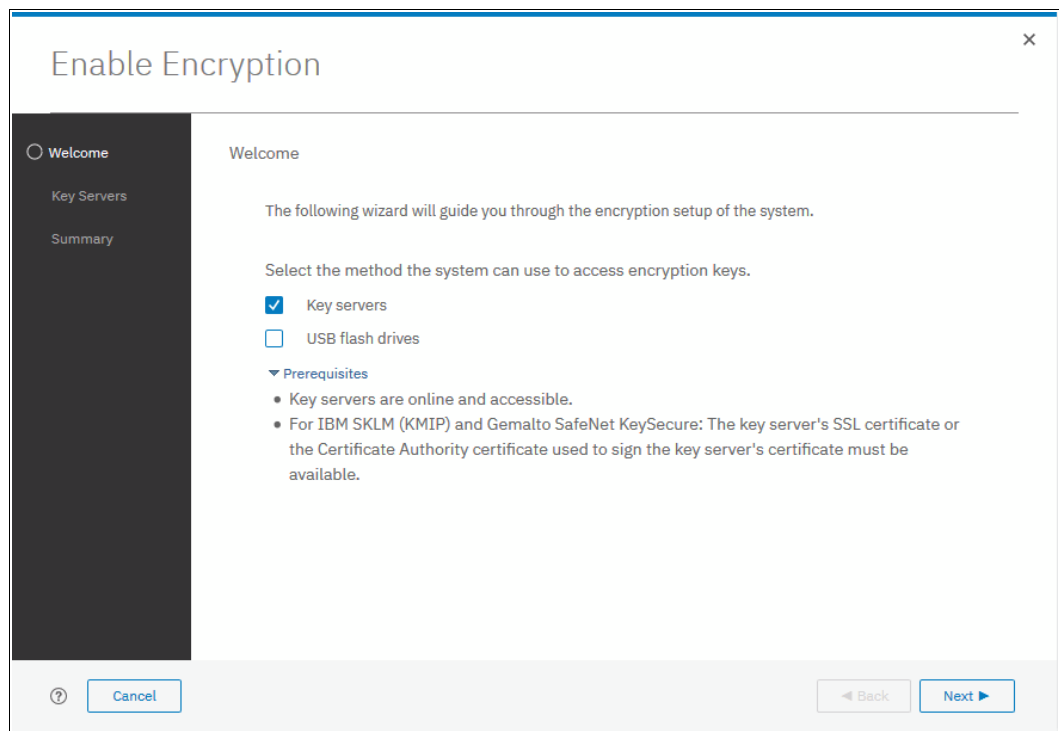


Figure 13-30 Selecting Key server as the only provider in the Enable Encryption wizard

3. Select IBM SKLM (with KMIP) as the key server type, as shown in Figure 13-31.

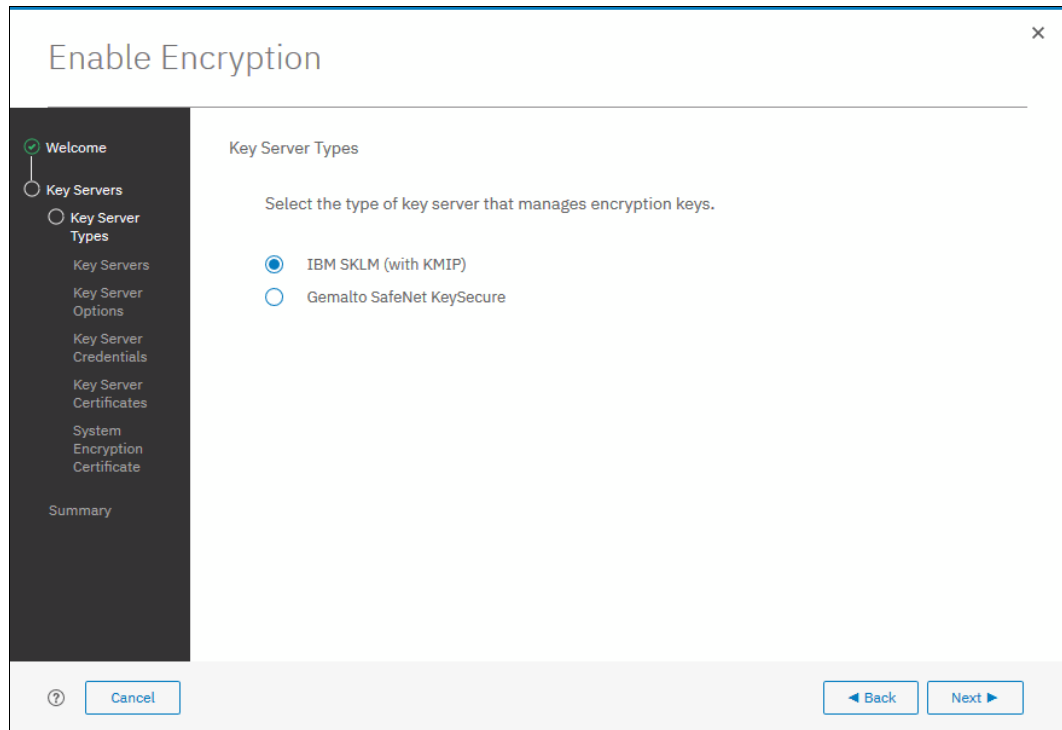


Figure 13-31 Selecting SKLM as key server type

4. The wizard moves to the Key Servers tab, as shown in Figure 13-32 on page 753. Enter the name and IP address of the key servers. The first key server that is specified must be the primary SKLM key server.

Note: The supported versions of IBM Security Key Lifecycle Manager (up to V3.0, which was the latest code version available at the time of this writing) differentiate between the primary and secondary key server role. The Primary SKLM server as defined on the Key Servers window of the Enable Encryption wizard must be the server defined as the primary by SKLM administrators.

The key server name serves only as a label. Only the provided IP address is used to contact the server. If the key server's TCP port number differs from the default value for the KMIP protocol (that is, 5696), enter the port number. An example of a complete primary SKLM configuration is shown in Figure 13-32 on page 753.

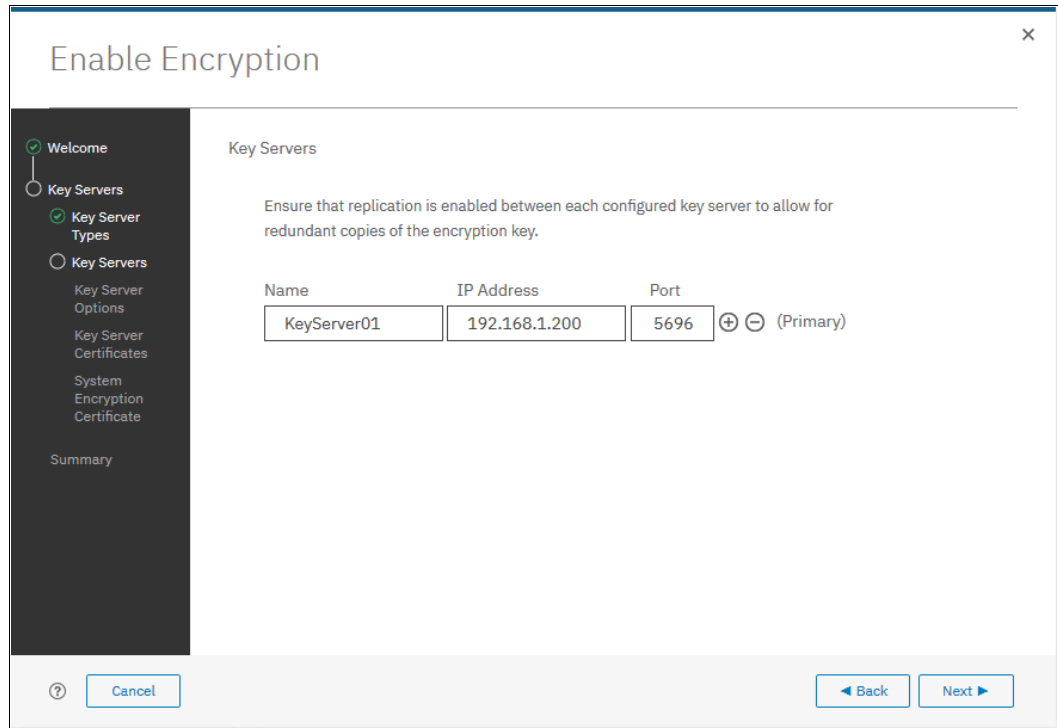


Figure 13-32 Configuration of the primary SKLM server

5. If you want to add more secondary SKLM servers, click + and enter the data for the secondary SKLM servers, as shown on Figure 13-33. You can define up to four SKLM servers. Click **Next** when you are done.

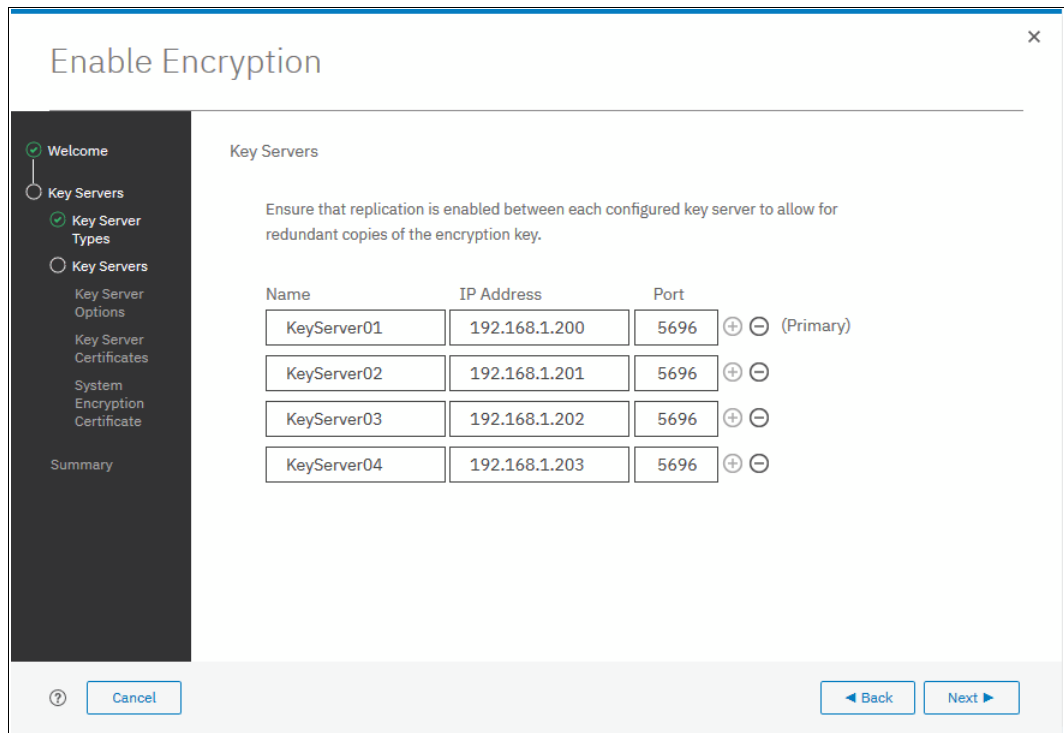


Figure 13-33 Configuring multiple SKLM servers

- The next window in the wizard is a reminder that SPECTRUM_VIRT device group that is dedicated for IBM Spectrum Virtualize systems must exist on the SKLM key servers. Make sure that this device group exists and click **Next** to continue (see Figure 13-34).

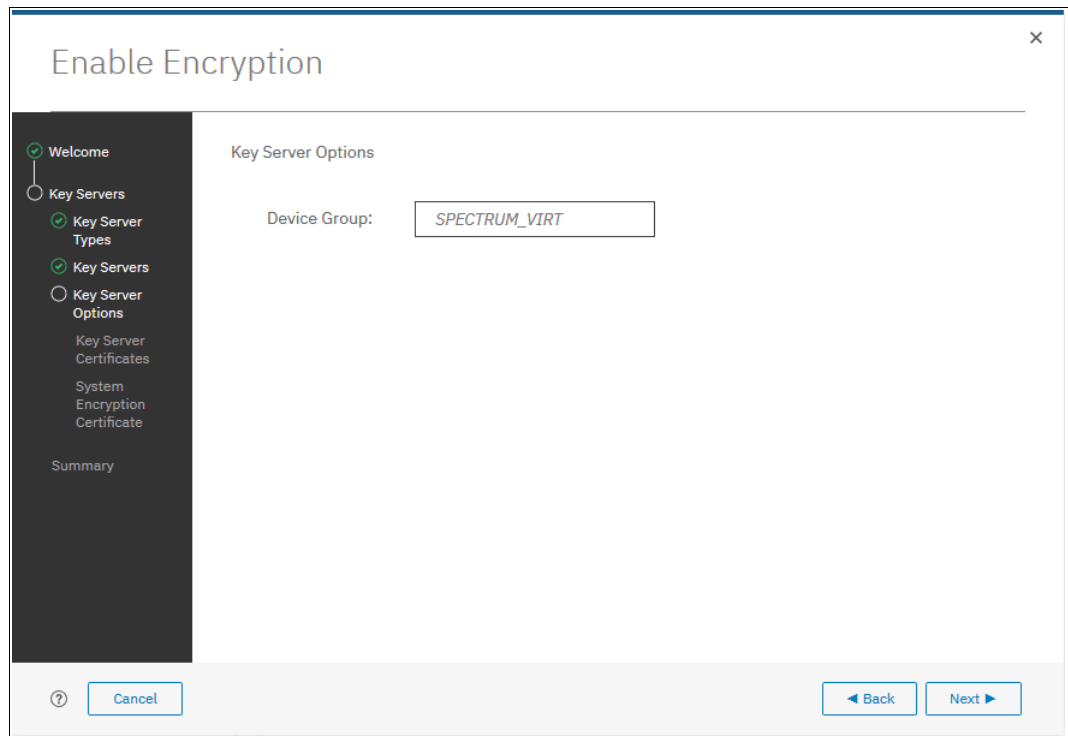


Figure 13-34 Checking key server device group

- Enable secure communication between the IBM Spectrum Virtualize system and the SKLM key servers by uploading the key server certificate (from a trusted third party or a self-signed certificate), or by uploading the public SSL certificate of each key server directly.

After uploading any of the certificates in the window that is shown in Figure 13-35 on page 755, click **Next** to proceed to the next step.

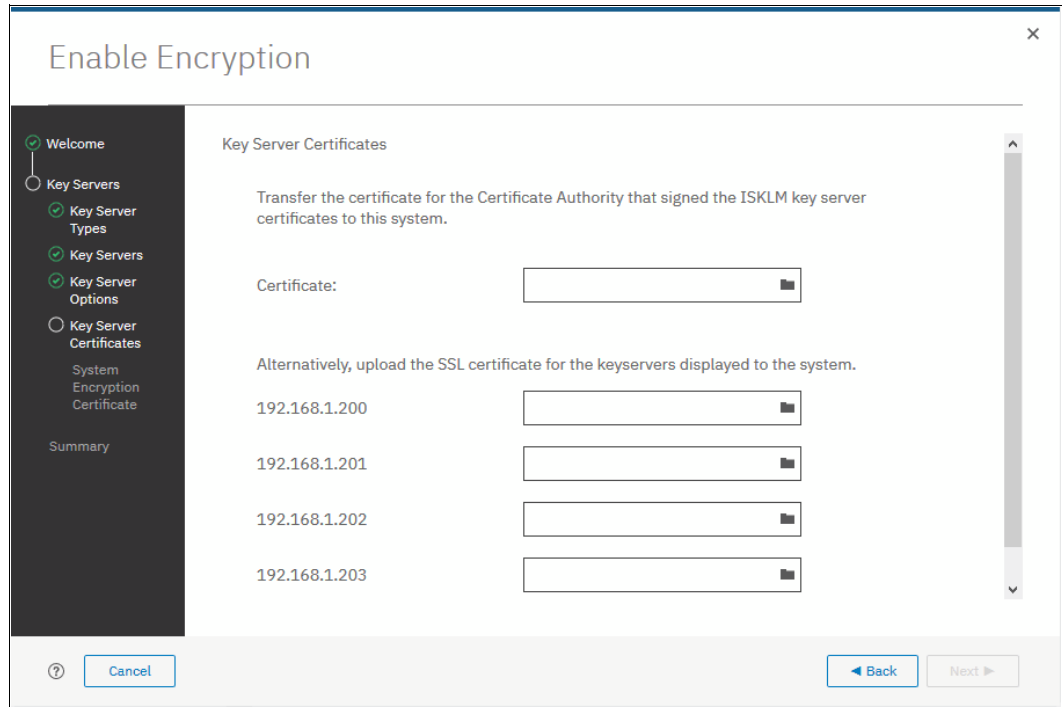


Figure 13-35 Uploading key servers or Certificate Authority SSL certificate

8. Configure the SKLM key server to trust the public key certificate of the IBM Spectrum Virtualize system. You can download the IBM Spectrum Virtualize system public SSL certificate by clicking **Export Public Key**, as shown in Figure 13-36. Install this certificate in the SKLM key server in the SPECTRUM_VIRT device group.

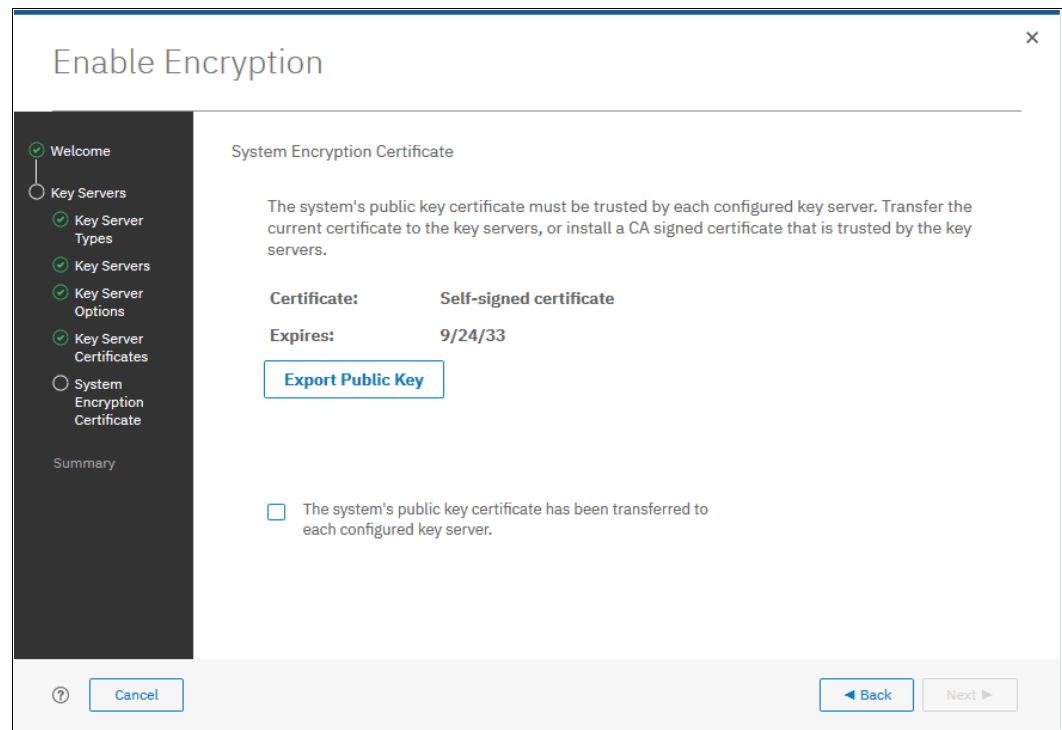


Figure 13-36 Downloading the IBM Spectrum Virtualize SSL certificate

9. When the IBM Spectrum Virtualize system public key certificate is installed on the SKLM key servers, acknowledge this installation by selecting the box that is indicated in blue Figure 13-37 and click **Next**.

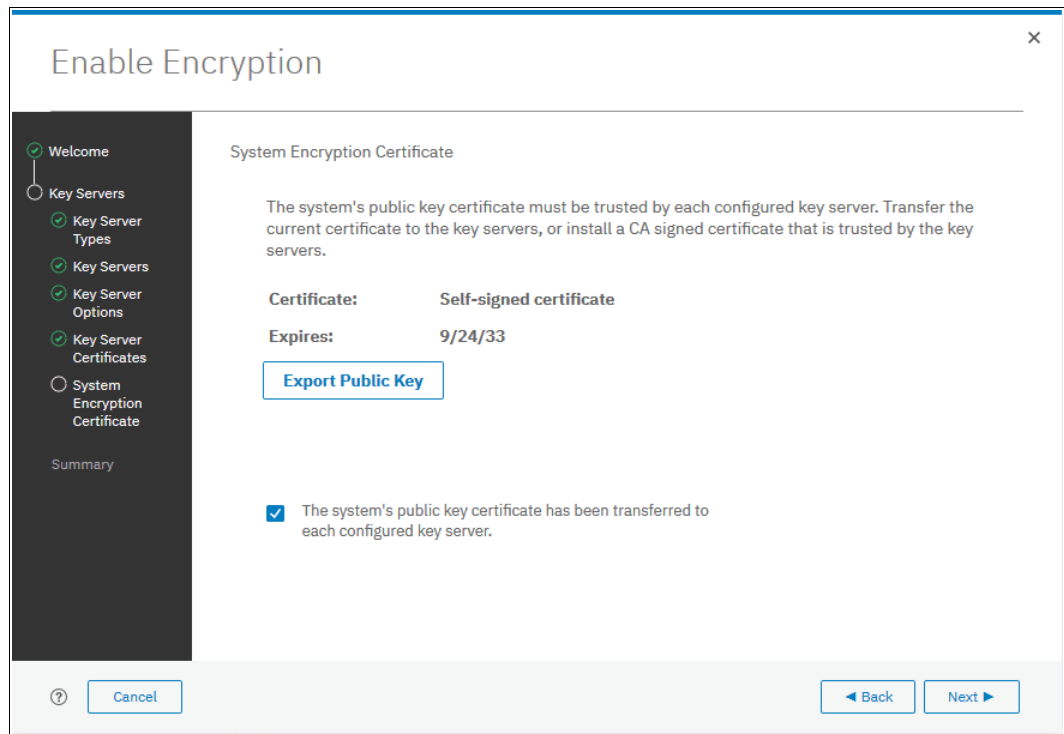


Figure 13-37 Acknowledge IBM Spectrum Virtualize public key certificate transfer

10. The key server configuration is shown in the Summary tab, as shown in Figure 13-38. Click **Finish** to create the key server object and finalize the encryption enablement.

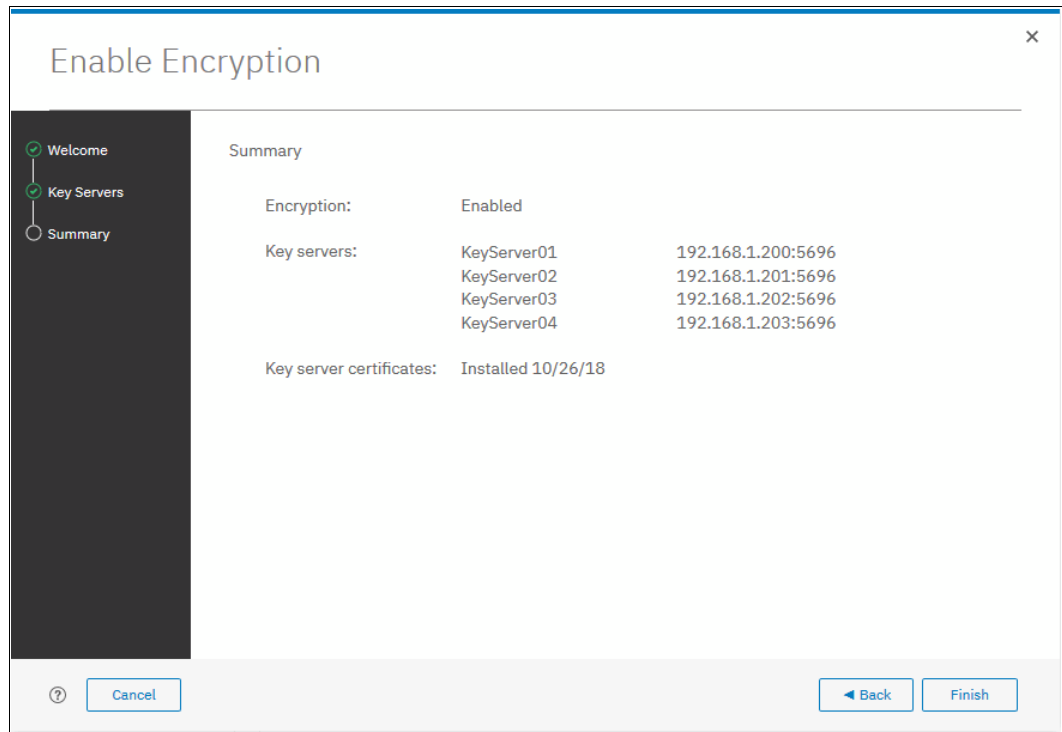


Figure 13-38 Finish enabling encryption using SKLM key servers

11. If no errors occurred while the key server object was created, you receive a message that confirms that the encryption is now enabled on the system, as shown in Figure 13-39. Click **Close**.

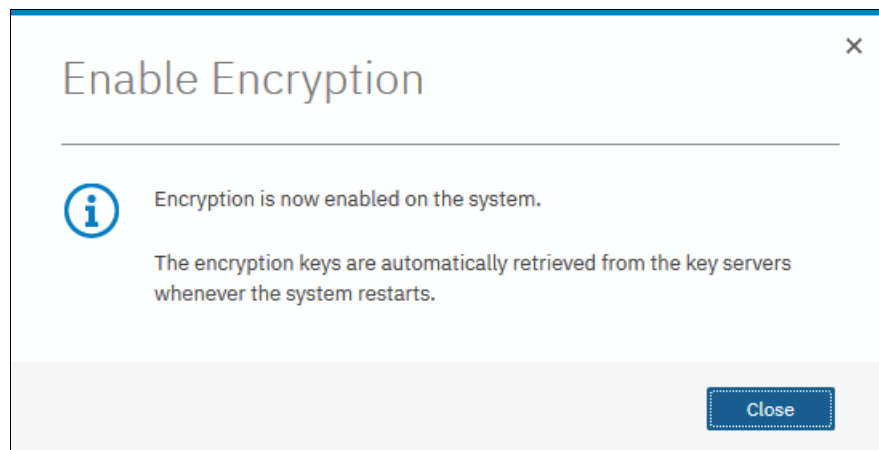


Figure 13-39 Encryption enabled message using an SKLM key server

12. Confirm that encryption is enabled by clicking **Settings** → **Security** → **Encryption**, as shown in Figure 13-40. The “Online” state indicates which SKLM servers are detected as available by the system.

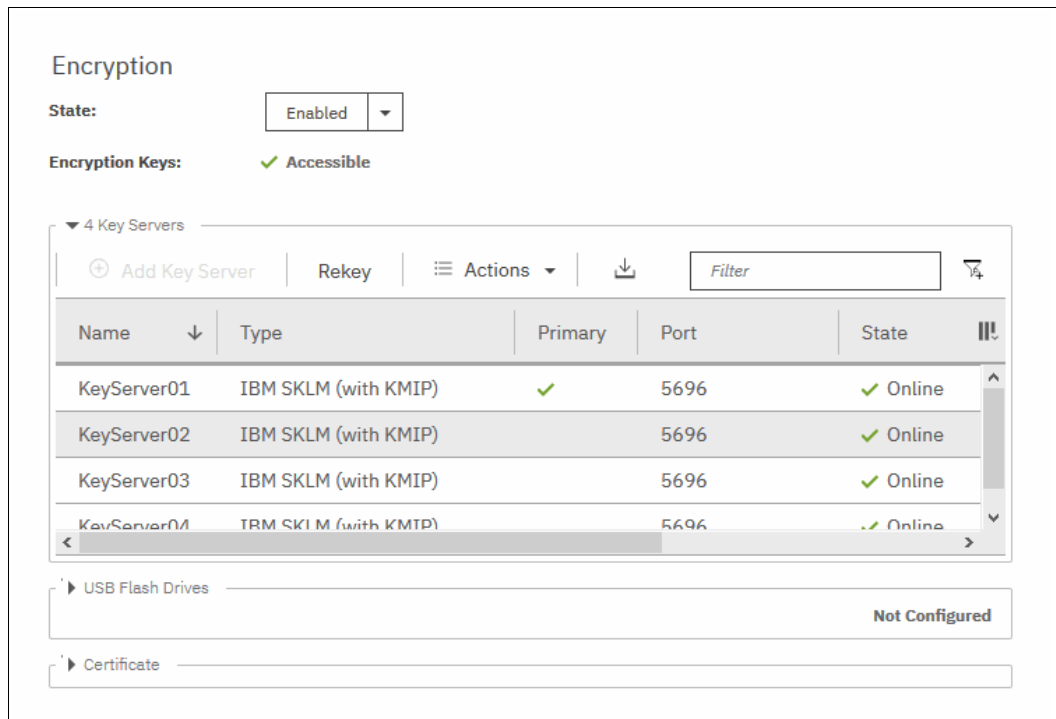


Figure 13-40 Encryption enabled with only SKLM servers as encryption key providers

Enabling encryption using SafeNet KeySecure

IBM Spectrum Virtualize V8.2.1 introduces support for Gemalto SafeNet KeySecure, which is a third-party key management server. It can be used as an alternative to IBM Security Key Lifecycle Manager (SKLM).

IBM Spectrum Virtualize supports Gemalto SafeNet KeySecure version 8.3.0 and later, and only using KMIP protocol. It is possible to configure up to four SafeNet KeySecure servers in IBM Spectrum Virtualize for redundancy, and they can coexist with USB flash drive encryption.

It is not possible to have both SafeNet KeySecure and SKLM key servers configured at the same time in IBM Spectrum Virtualize. It is also not possible to migrate directly from one type of key server to another (from SKLM to SafeNet KeySecure or vice versa). If you want to migrate from one type to another, first migrate to USB flash drives encryption, and then, migrate to the other type of key servers.

KeySecure uses an active-active clustered model. All changes to one key server are instantly propagated to all other servers in the cluster.

Although KeySecure uses KMIP protocol as IBM SKLM does, an option is available to configure user name and password for IBM Spectrum Virtualize and KeySecure server authentication, which is not possible when performing the configuration with SKLM.

The certificate for client authentication in SafeNet KeySecure can be self-signed or signed by a Certificate Authority.

To enable encryption in IBM Spectrum Virtualize by using a Gemalto SafeNet KeySecure key server, complete the following steps:

1. Ensure that you have service IPs configured on all your nodes.
2. In the Enable Encryption wizard Welcome tab, select **Key servers** and click **Next**, as shown in Figure 13-41.

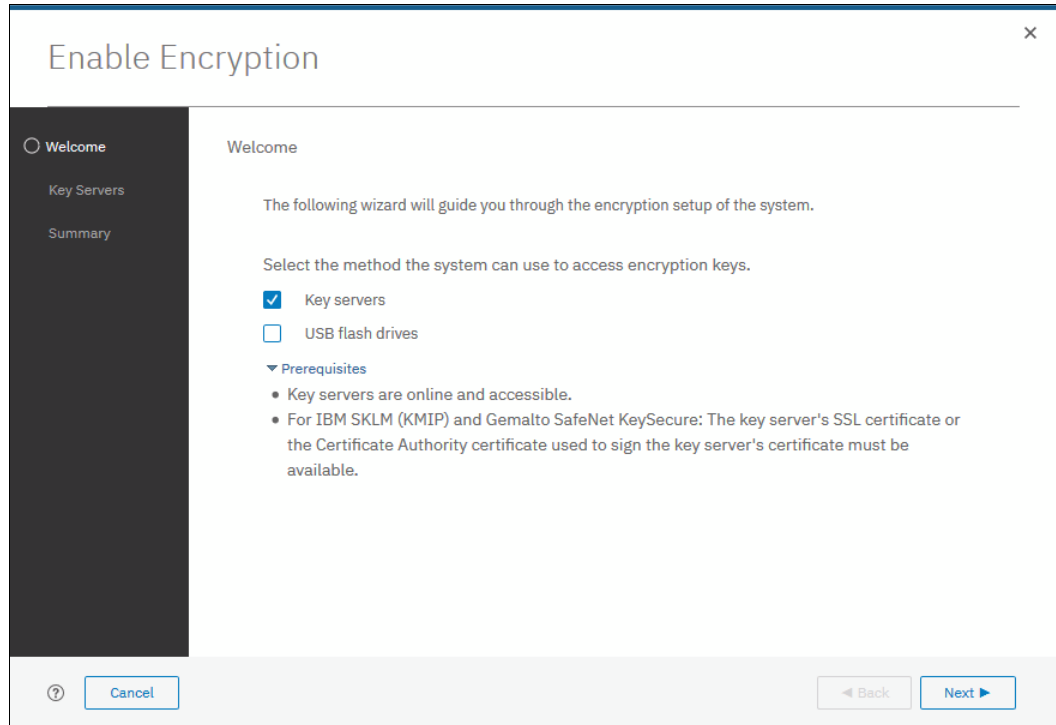


Figure 13-41 Selecting Key servers as the only provider in the Enable Encryption wizard

3. In the next window, you can choose between IBM SKLM or Gemalto SafeNet KeySecure server types, as shown in Figure 13-42. Select **Gemalto SafeNet KeySecure** and click **Next**.

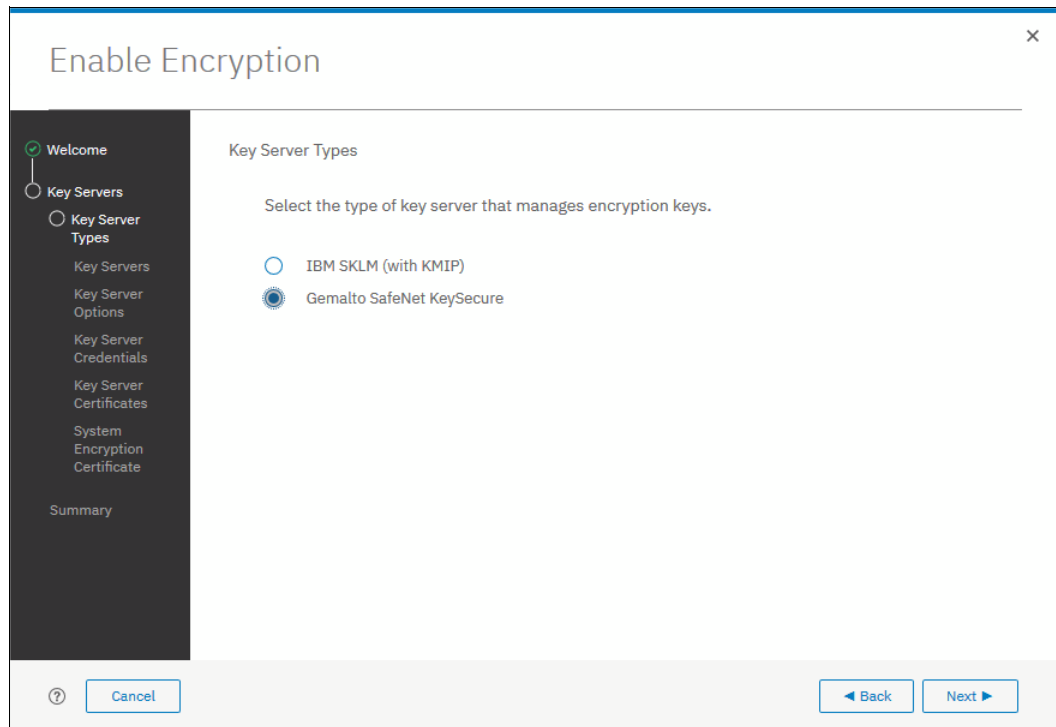


Figure 13-42 Selecting Gemalto SafeNet KeySecure as key server type

4. Add up to four SafeNet KeySecure servers in the next wizard window, as shown in Figure 13-43 on page 761. For each key server, enter the name, IP address, and TCP port for KMIP protocol (default value is 5696). Because the server name is only a label, it does not need to be the real host name of the server.

Although Gemalto SafeNet KeySecure uses an active-active clustered model, IBM Spectrum Virtualize asks for a primary key server. The primary key server represents only the KeySecure server, which is used for key create and rekey operations. Therefore, any of the clustered key servers can be selected as the primary.

Selecting a primary key server is beneficial for load balancing, and any four key servers can be used to retrieve the master key.

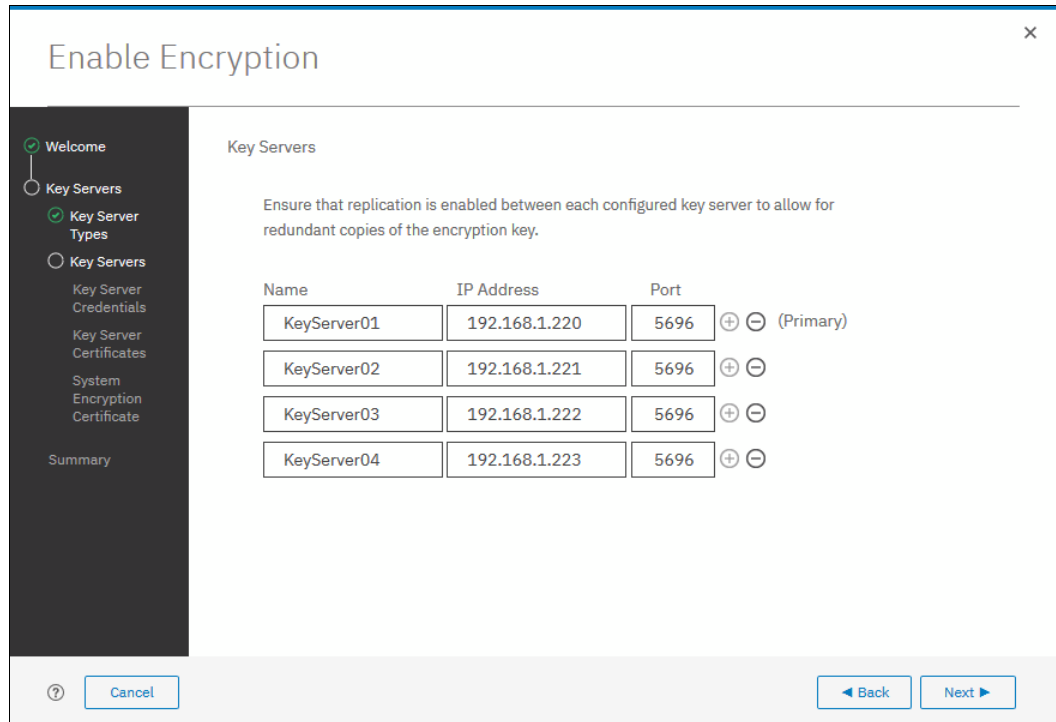


Figure 13-43 Configuring multiple SafeNet KeySecure servers

- The next window in the wizard prompts for key servers credentials (user name and password), as shown in Figure 13-44. This setting is optional because it depends on how SafeNet KeySecure servers are configured.

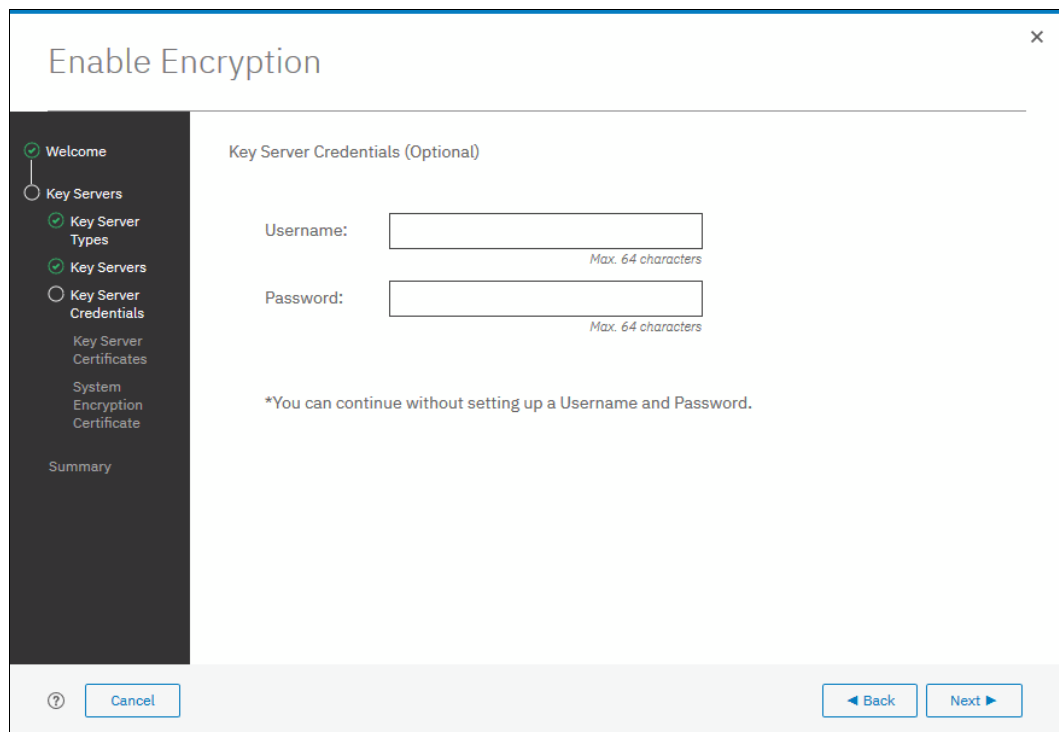


Figure 13-44 Key server credentials input (optional)

6. Enable secure communication between the IBM Spectrum Virtualize system and the SafeNet KeySecure key servers by uploading the key server certificate (from a trusted third party or a self-signed certificate), or by uploading the SSL certificate of each key server directly. After uploading any of the certificates in the window that is shown in Figure 13-45, click **Next** to proceed to the next step.

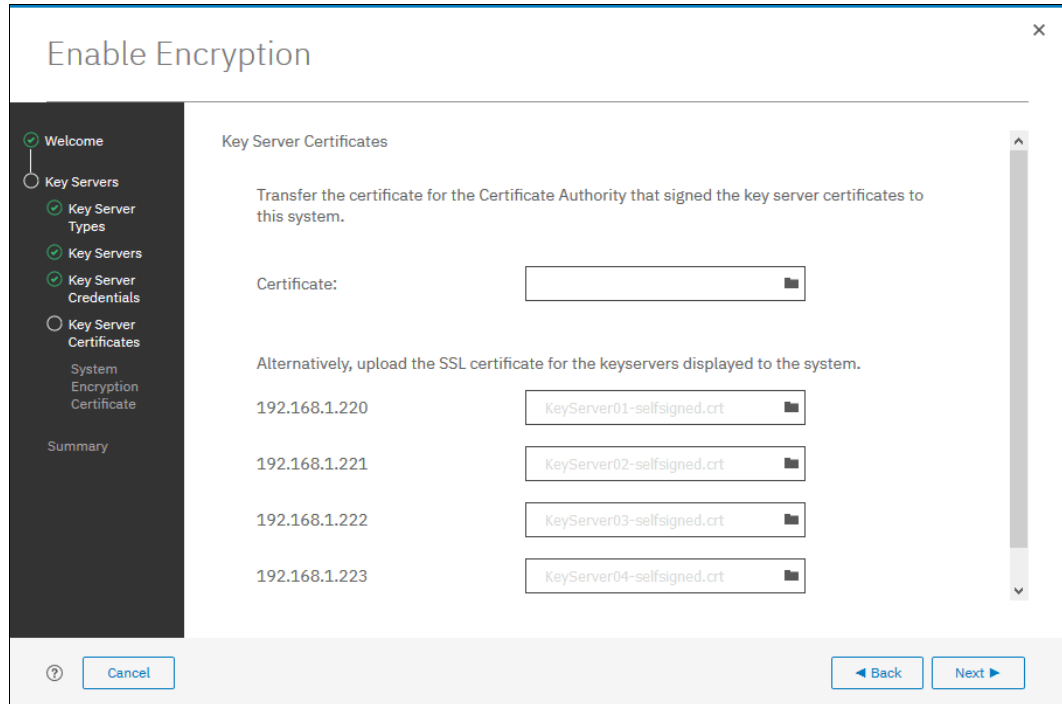


Figure 13-45 Uploading SafeNet KeySecure key servers certificate

7. Configure the SafeNet KeySecure key servers to trust the public key certificate of the IBM Spectrum Virtualize system. You can download the IBM Spectrum Virtualize system public SSL certificate by clicking **Export Public Key**, as shown in Figure 13-46 on page 763. After adding the public key certificate to the key servers, select the box and then, click **Next**.

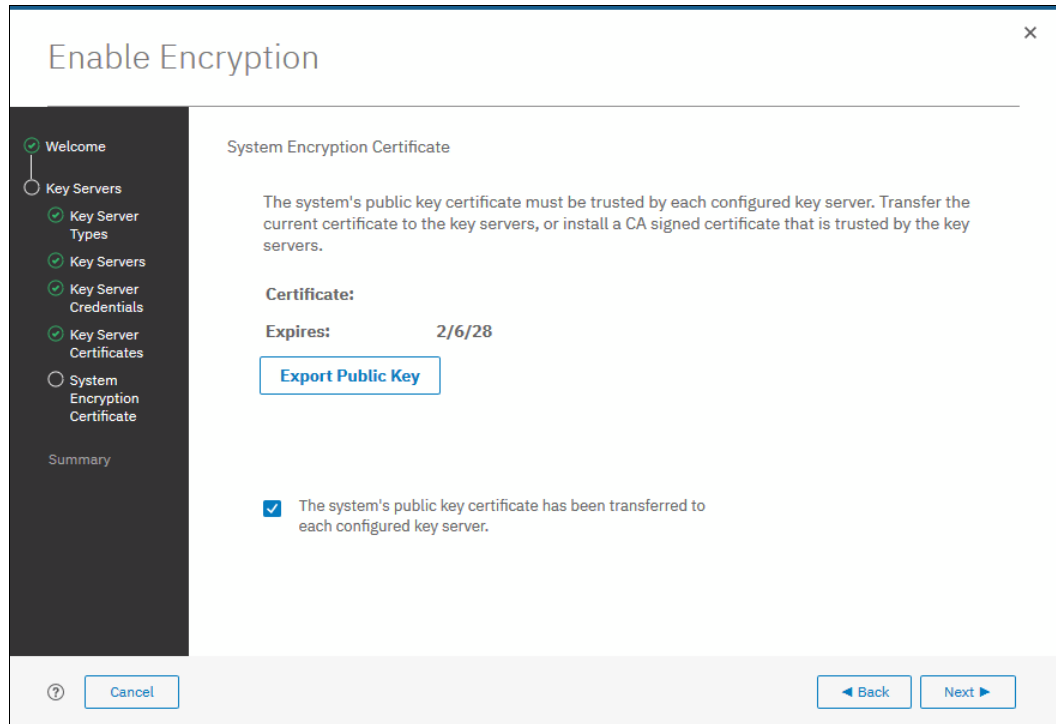


Figure 13-46 Downloading the IBM Spectrum Virtualize SSL certificate

- The key server configuration is shown in the Summary tab, as shown in Figure 13-47. Click **Finish** to create the key server object and finalize the encryption enablement.

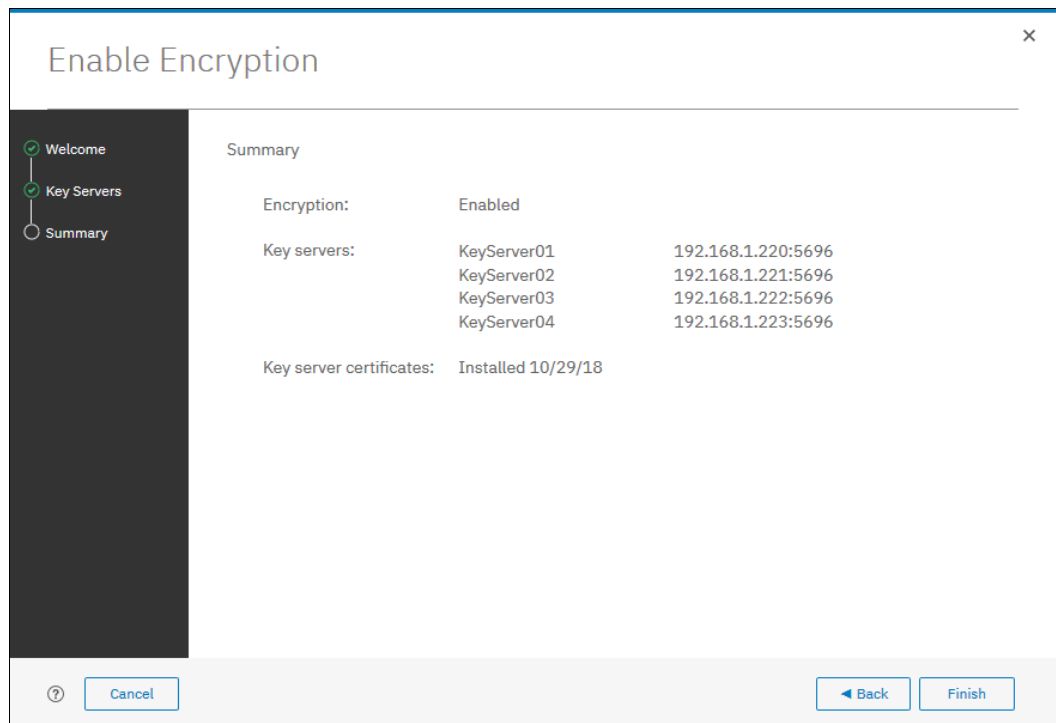


Figure 13-47 Finish enabling encryption using SafeNet KeySecure key servers

- If no errors occur while creating the key server object, you receive a message that confirms that the encryption is now enabled on the system, as shown in Figure 13-48. Click **Close**.

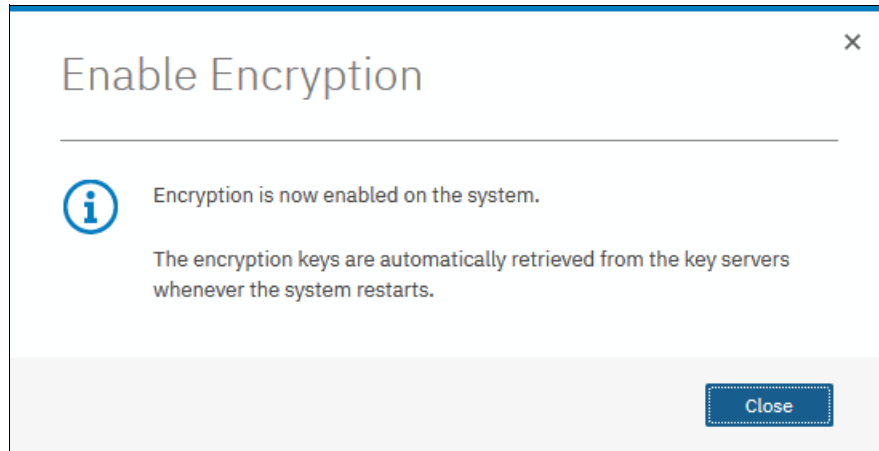


Figure 13-48 Encryption enabled using SafeNet KeySecure key servers

- Confirm that encryption is enabled by clicking **Settings** → **Security** → **Encryption**, as shown in Figure 13-49. Check whether the four servers are shown as online, which indicate that all four SafeNet KeySecure servers are detected as available by the system.

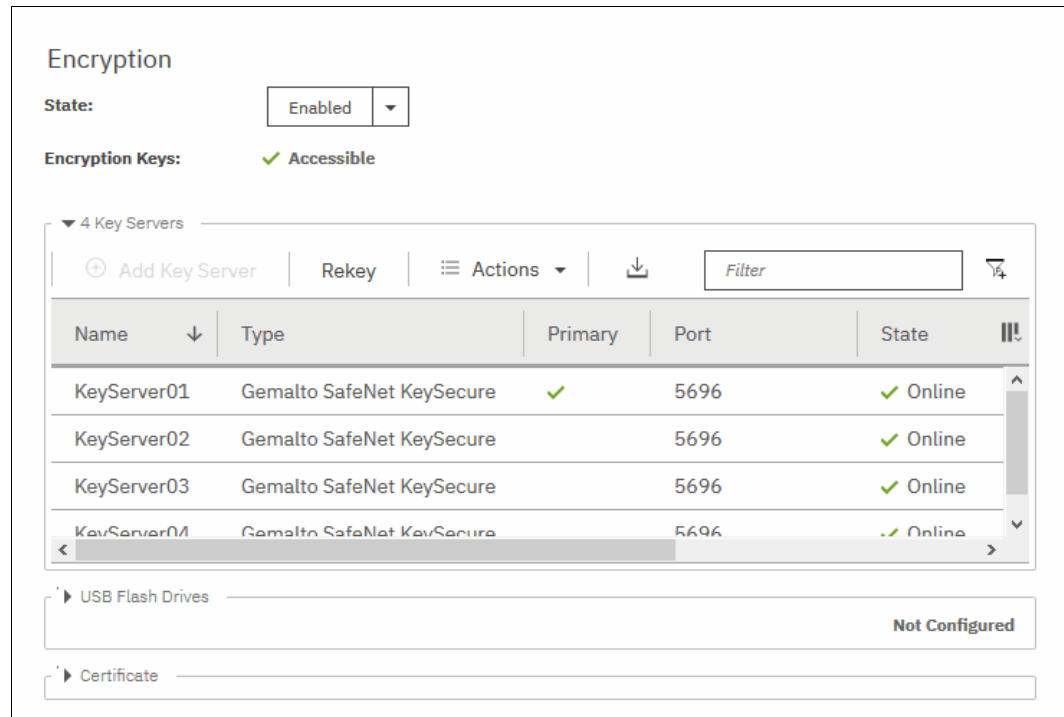


Figure 13-49 Encryption enabled with 4 SafeNet KeySecure key servers

13.4.4 Enabling encryption by using both providers

IBM Spectrum Virtualize allows parallel use of both USB flash drive and one type of key server (SKLM or SafeNet KeySecure) as encryption key providers. It is possible to configure both providers in a single run of encryption enable wizard. To perform this configuration, the system must meet requirements of both key server (SKLM or SafeNet KeySecure) and USB flash drive encryption key providers.

Note: Make sure that the key management server functionality is fully independent from an encrypted storage, which has encryption managed by this key server environment. Failure to observe this requirement might create an encryption deadlock. An encryption deadlock is a situation in which none of key servers in the environment can become operational because some critical part of the data in each server is stored on an encrypted storage system that depends on one of the key servers to unlock access to the data.

IBM Spectrum Virtualize code V8.1 and later supports up to four key server objects that are defined in parallel.

Before you start to enable encryption by using both USB flash drives and a key servers, confirm that the requirements that are described in 13.4.2, “Enabling encryption using USB flash drives” on page 745, and section 13.4.3, “Enabling encryption using key servers” on page 750 are met.

To enable encryption by using a key server and USB flash drive, complete the following steps:

1. Ensure that you have service IPs configured on all your nodes.
2. In the Enable Encryption wizard Welcome tab, select **Key servers** and **USB flash drives** and click **Next**, as shown in Figure 13-50.

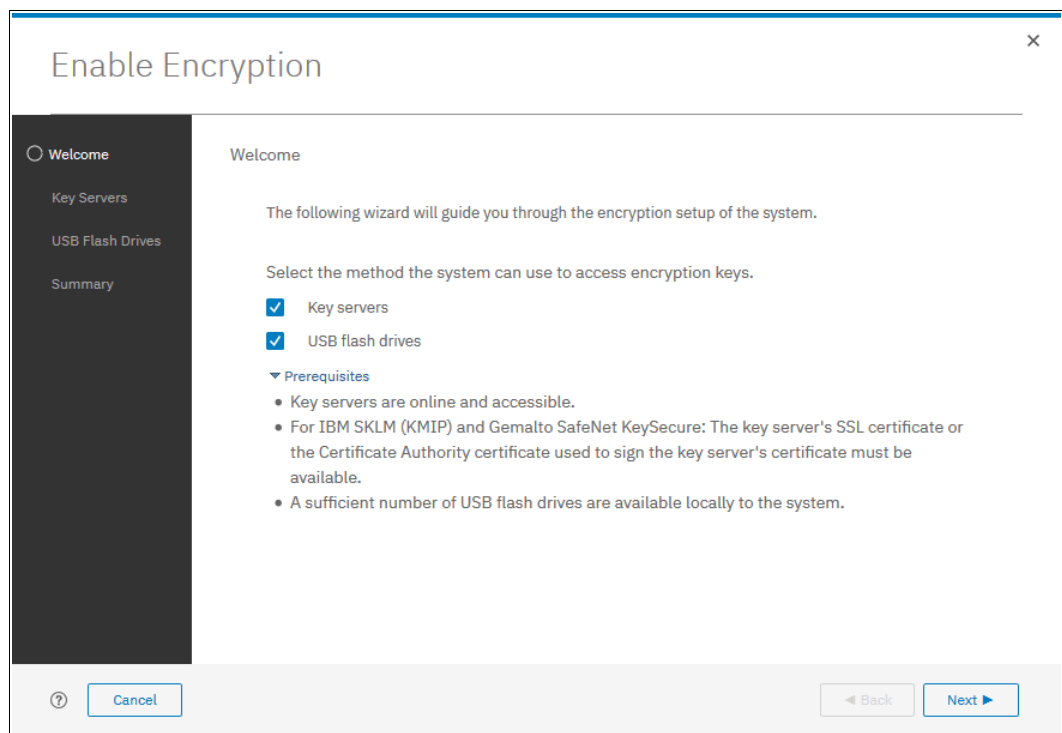


Figure 13-50 Selecting Key servers and USB flash drives in the Enable Encryption wizard

3. The wizard moves to the Key Server Types window, as shown in Figure 13-51. Then, select the key server type that will manage the encryption keys.

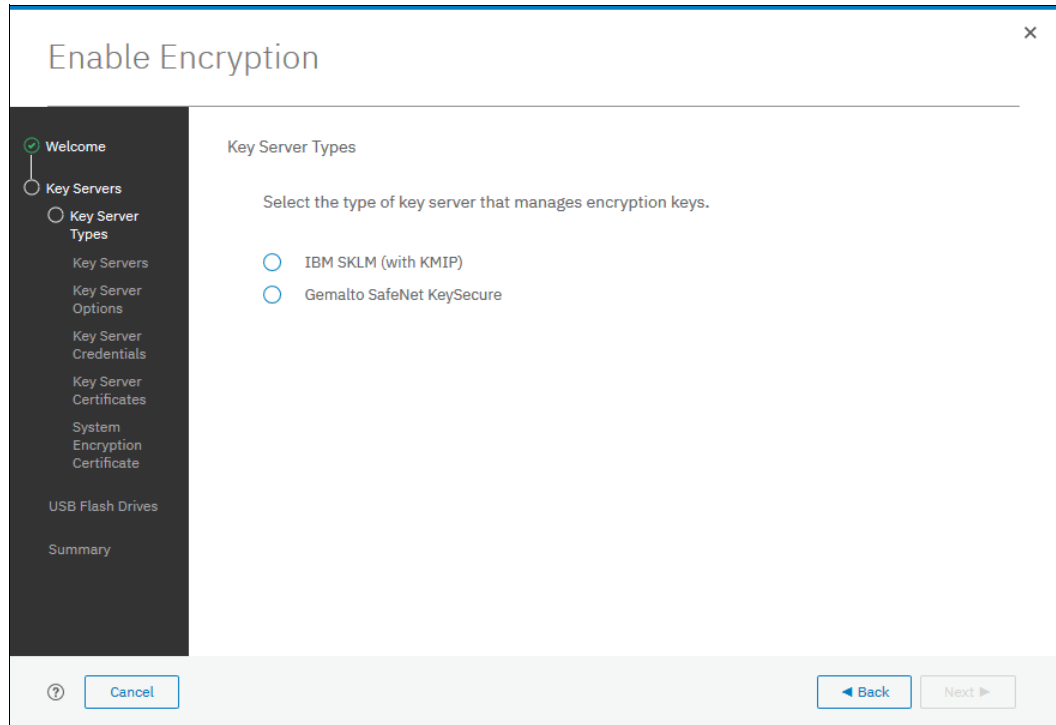


Figure 13-51 Selecting the key server type

4. The next windows that are displayed are the same windows that are shown in section 13.4.3, “Enabling encryption using key servers” on page 750 (depending on the type of key server selected).

When all of the key servers details are entered, the USB flash drive encryption configuration is displayed. In this step, master encryption key copies are stored in the USB flash drives.

If there are fewer than two drives detected, the system requests plugging more USB flash drives, as shown on Figure 13-52. You cannot proceed until the required minimum number of USB flash drives is detected by the system.

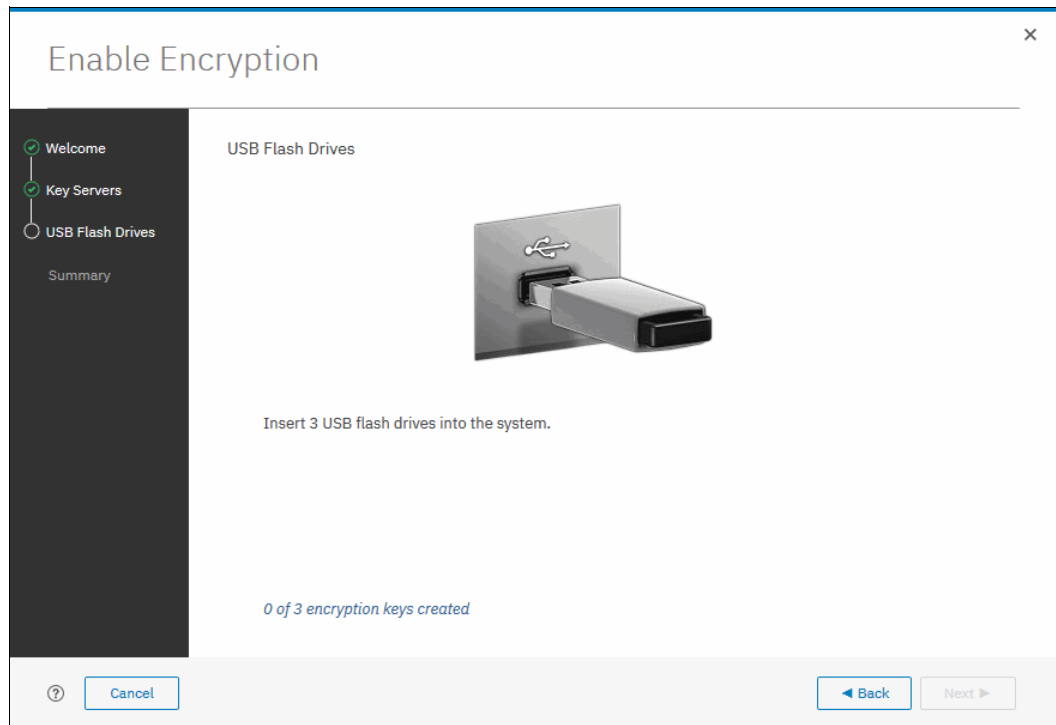


Figure 13-52 Prompt to insert USB flash drives

After the encryption key is copied to the first two USB flash drives, the management GUI prompts you to remove the two flash drives. After you remove the flash drives, insert the last required flash drive into the system. The system attempts to write the encryption key to any flash drive it detects. Therefore, it is crucial to maintain physical security of the system during this procedure. After the keys are successfully copied to at least three USB flash drives, the system displays a window as shown on Figure 13-53.



Figure 13-53 Master Access Key successfully copied to USB flash drives

5. After copying encryption keys to USB flash drives, the next window is shown with the summary of the configuration that is implemented on the system (see Figure 13-54). Click **Finish** to create the key server object and finalize the encryption enablement.

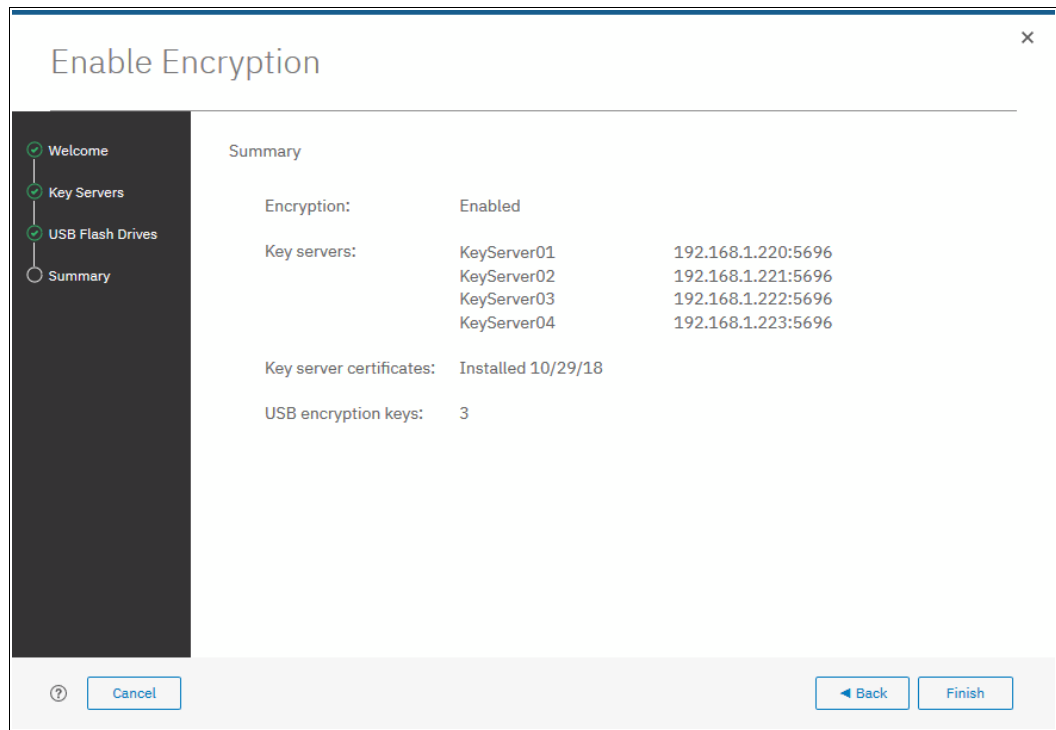


Figure 13-54 Encryption configuration summary in two providers scenario

6. If no errors occurred while creating the key server object, the system displays a window that confirms that the encryption is now enabled on the system, and that both encryption key providers are enabled (see Figure 13-55).

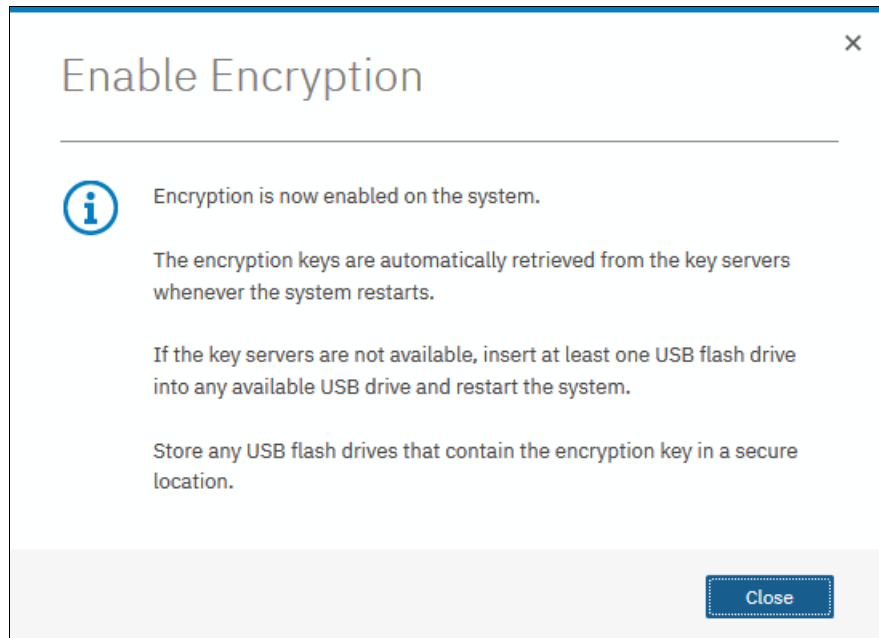


Figure 13-55 Encryption enabled message using both encryption key providers

- You can confirm that encryption is enabled and verify which key providers are in use by clicking **Settings** → **Security** → **Encryption**, as shown in Figure 13-56. Note the *Online* state of key servers and *Validated* state of USB ports where USB flash drives are inserted to make sure they are properly configured

The screenshot shows the 'Encryption' configuration page. At the top, the 'State' is set to 'Enabled'. Below it, 'Encryption Keys' are marked as 'Accessible'. There are two sections: '4 Key Servers' and '2 USB Flash Drives Detected'.

4 Key Servers

Name	Type	Primary	Port	State
KeyServer01	Gemalto SafeNet KeySecure	✓	5696	✓ Online
KeyServer02	Gemalto SafeNet KeySecure		5696	✓ Online
KeyServer03	Gemalto SafeNet KeySecure		5696	✓ Online
KeyServer04	Gemalto SafeNet KeySecure		5696	✓ Online

2 USB Flash Drives Detected

ID	USB Port	State
0	1	✓ Validated
9	1	✓ Validated

Figure 13-56 Encryption enabled with both USB flash drives and key servers

13.5 Configuring more providers

When the system is configured with a single encryption key provider, it is possible to add a provider.

Note: If you set up encryption of your storage system when it was running IBM Spectrum Virtualize code version earlier than V7.8.0, when you upgrade to code version V8.1 or later, you must rekey the master encryption key before you can enable second encryption provider.

13.5.1 Adding key servers as a second provider

If the storage system is configured with the USB flash drive provider, it is possible to configure SKLM or SafeNet KeySecure servers as a second provider. To enable key servers as a second provider, complete the following steps:

1. Click **Settings** → **Security** → **Encryption**, expand the Key Servers section and then, click **Configure**, as shown in Figure 13-57. To enable key server as a second provider, the system must detect at least one USB flash drive with a current copy of the master access key.

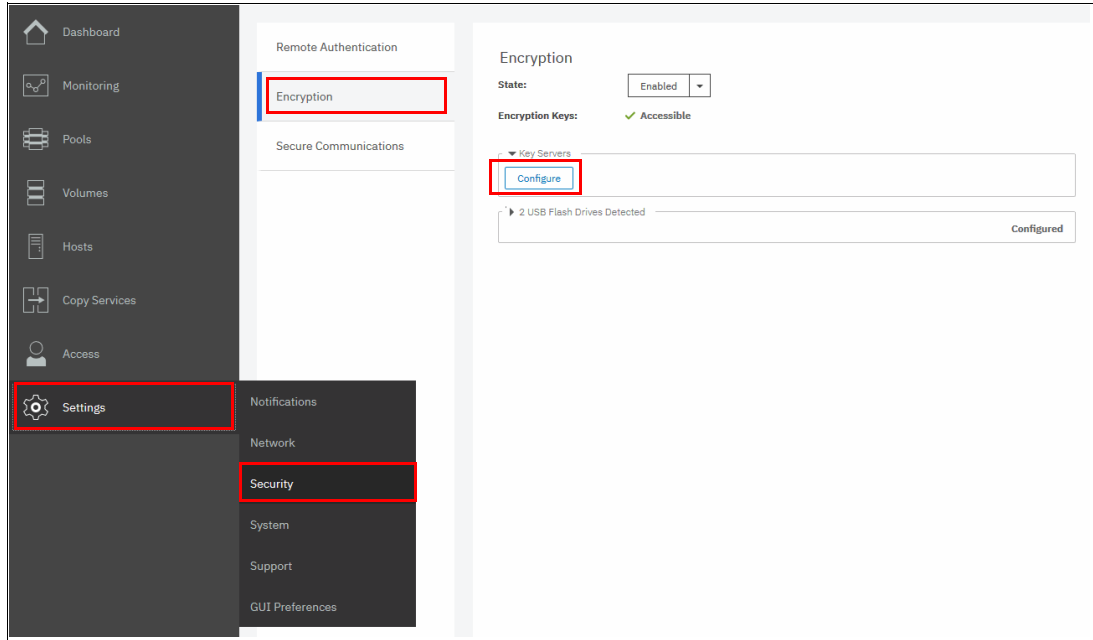


Figure 13-57 Enable key servers as a second provider

2. Complete the required steps to configure the key server provider, as described in 13.4.3, “Enabling encryption using key servers” on page 750.

The difference in the process described in that section is that the wizard gives you an option to disable USB flash drive encryption, which aims to migrate from the USB flash drive to key server provider.

Select **No** to enable both encryption key providers, as shown in Figure 13-58 on page 772.

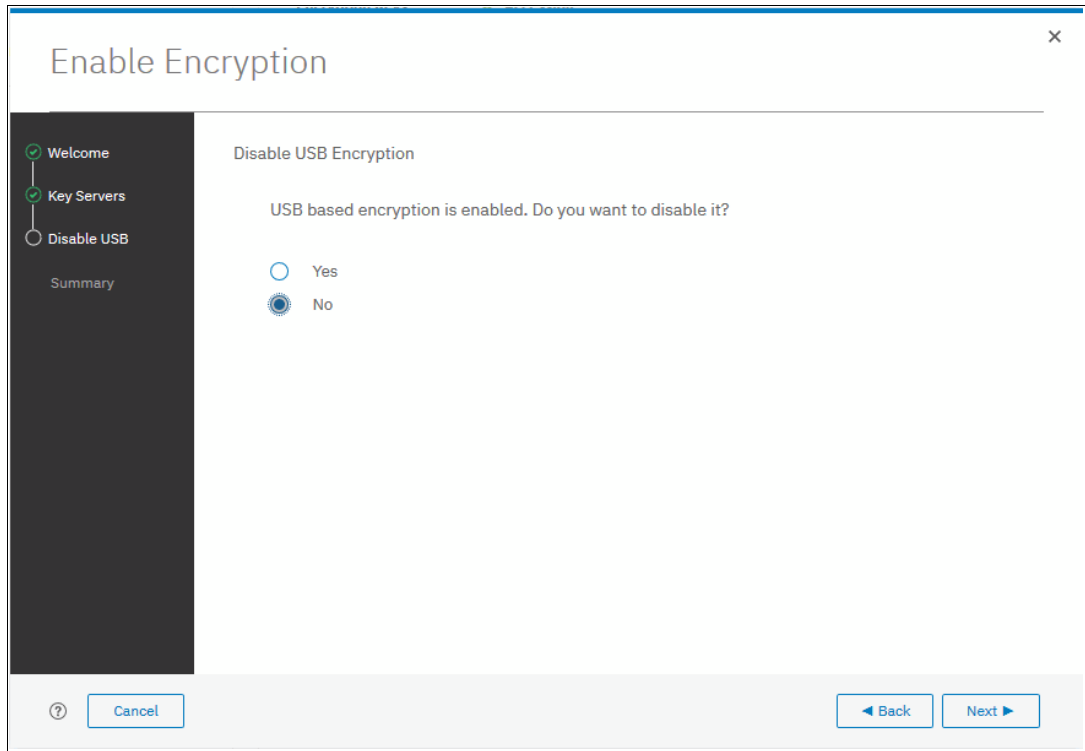


Figure 13-58 Do not disable USB flash drive encryption key provider

This choice is confirmed on the summary window before the configuration is committed, as shown in Figure 13-59.

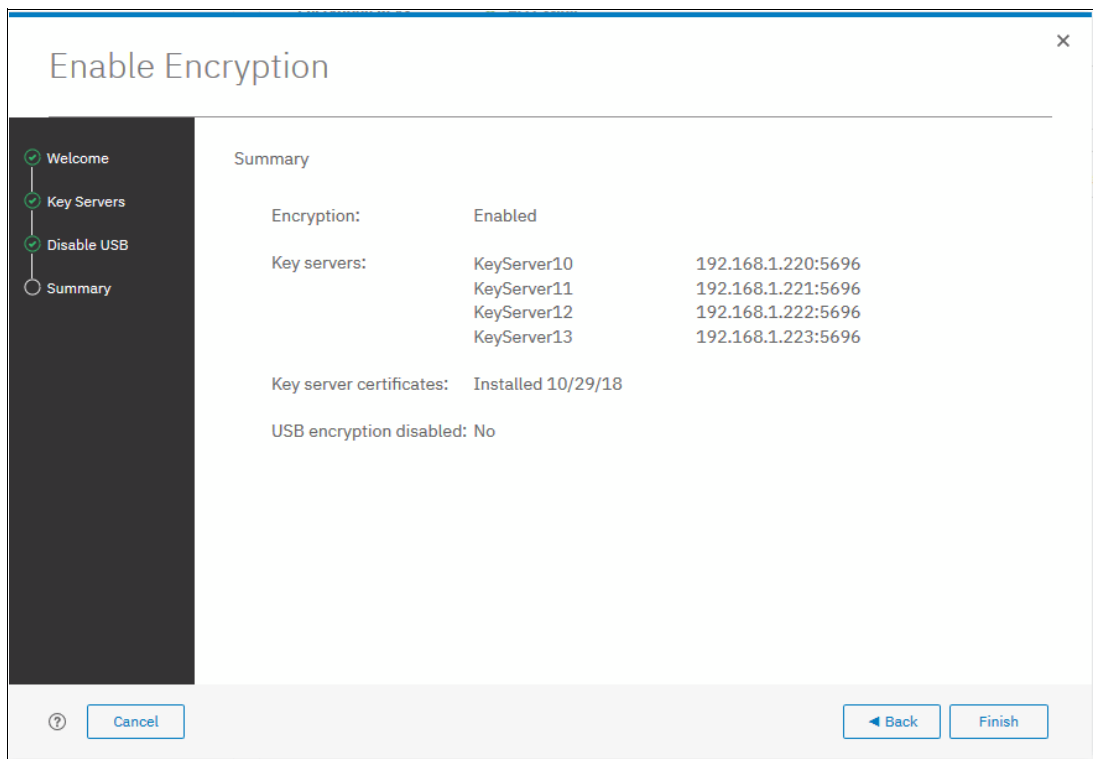


Figure 13-59 Configuration summary before committing

3. Click **Finish**. The system configures keys servers as a second encryption key provider. Successful completion of the task is confirmed by a message, as shown in Figure 13-60. Click **Close**.



Figure 13-60 Confirmation of successful configuration of two encryption key providers

- You can confirm that encryption is enabled and verify which key providers are in use by clicking **Settings** → **Security** → **Encryption**, as shown in Figure 13-61. Note the *Online* state of key servers and *Validated* state of USB ports where USB flash drives are inserted to make sure they are properly configured.

The screenshot shows the 'Encryption' configuration page. At the top, the 'State' is set to 'Enabled'. Below it, 'Encryption Keys' are marked as 'Accessible'. There are two main sections: '4 Key Servers' and '2 USB Flash Drives Detected'.

4 Key Servers

Name	Type	Primary	Port	State
KeyServer01	Gemalto SafeNet KeySecure	✓	5696	✓ Online
KeyServer02	Gemalto SafeNet KeySecure		5696	✓ Online
KeyServer03	Gemalto SafeNet KeySecure		5696	✓ Online
KeyServer04	Gemalto SafeNet KeySecure		5696	✓ Online

2 USB Flash Drives Detected

ID	USB Port	State
0	1	✓ Validated
9	1	✓ Validated

Figure 13-61 Encryption enabled with two key providers available

13.5.2 Adding USB flash drives as a second provider

If the storage system is configured with an SKLM or SafeNet KeySecure encryption key provider, it is possible to configure USB flash drives as a second provider. To enable USB flash drives as a second provider, complete the following steps:

- Click **Settings** → **Security** → **Encryption**, expand the USB Flash Drives section and then, click **Configure**, as shown in Figure 13-62 on page 775. To enable USB flash drives as a second provider, the system must access key servers with the current master access key.

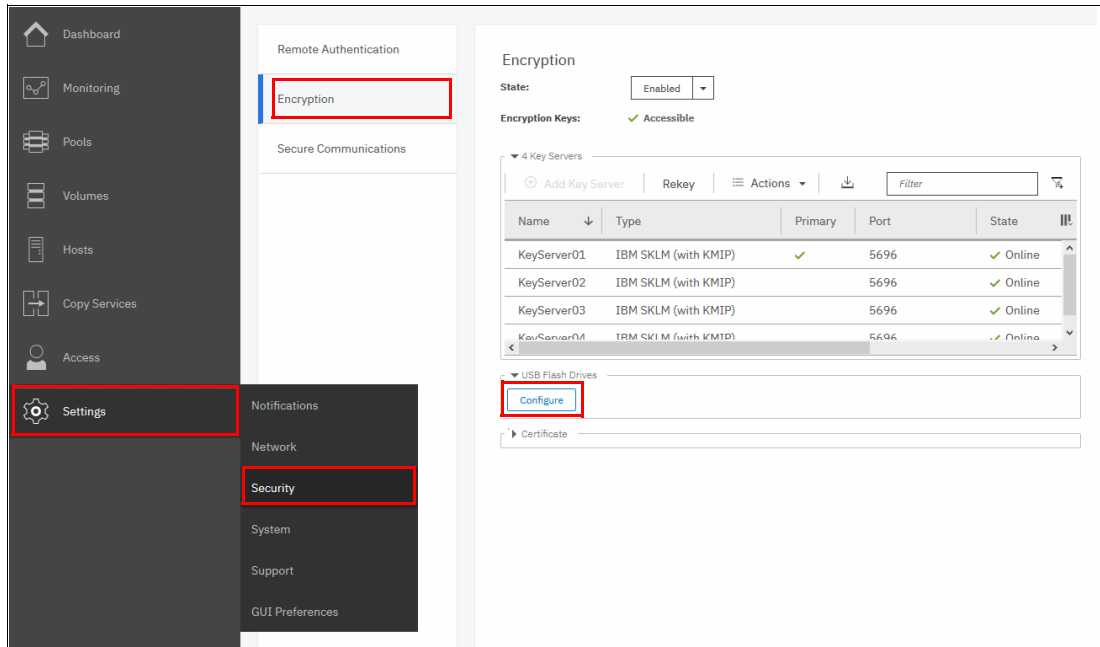


Figure 13-62 Enable USB flash drives as a second encryption key provider

2. Click **Configure**. You are presented with a wizard that is similar to the one that is described in 13.4.2, “Enabling encryption using USB flash drives” on page 745. You *cannot* disable key server provider during this process. After successful completion of the process, you are presented with a message confirming that both encryption key providers are enabled, as shown in Figure 13-63.

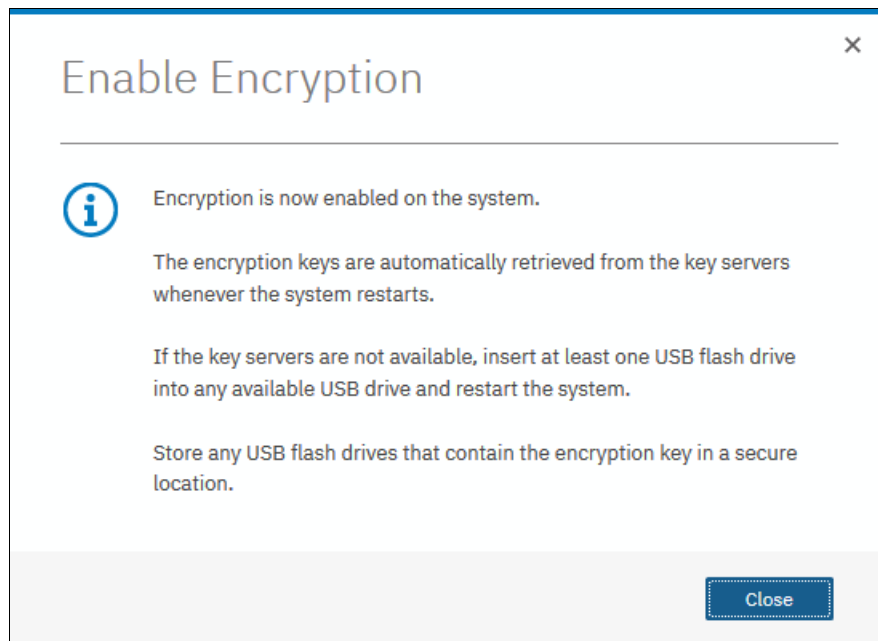


Figure 13-63 Confirmation of successful configuration of two encryption key providers

- You can confirm that encryption is enabled and verify which key providers are in use by clicking **Settings** → **Security** → **Encryption**, as shown in Figure 13-64. Note the state *Online* state of key servers and *Validated* state of USB ports where USB flash drives are inserted to make sure they are properly configured.

Encryption

State: Enabled

Encryption Keys: ✓ Accessible

▼ 4 Key Servers

Name	Type	Primary	Port	State
KeyServer01	Gemalto SafeNet KeySecure	✓	5696	✓ Online
KeyServer02	Gemalto SafeNet KeySecure		5696	✓ Online
KeyServer03	Gemalto SafeNet KeySecure		5696	✓ Online
KeyServer04	Gemalto SafeNet KeySecure		5696	✓ Online

▼ 2 USB Flash Drives Detected

ID	USB Port	State
0	1	✓ Validated
9	1	✓ Validated

Figure 13-64 Encryption enabled with two key providers available

13.6 Migrating between providers

IBM Spectrum Virtualize V8.1 introduced support for simultaneous use of USB flash drives and a key server as encryption key providers. The system also allows migration from configuration using only USB flash drive provider to key servers provider, and vice versa.

If you want to migrate from one key server type to another (for example, migrating from SKLM to SafeNet KeySecure or vice versa), direct migration is not possible. In this case, it is required first to migrate from the current key server type to a USB flash drive, and then migrate to the other type of key server.

13.6.1 Migrating from USB flash drive provider to encryption key server

The system is designed to facilitate changing from USB flash drives encryption key provider to encryption key server provider. Follow the steps that are described in 13.5.1, “Adding key servers as a second provider” on page 771, but when completing step 2 on page 771, select **Yes** instead of **No** (see Figure 13-65). This action causes de-activation of the USB flash drives provider. The procedure completes with only key servers configured as key provider.

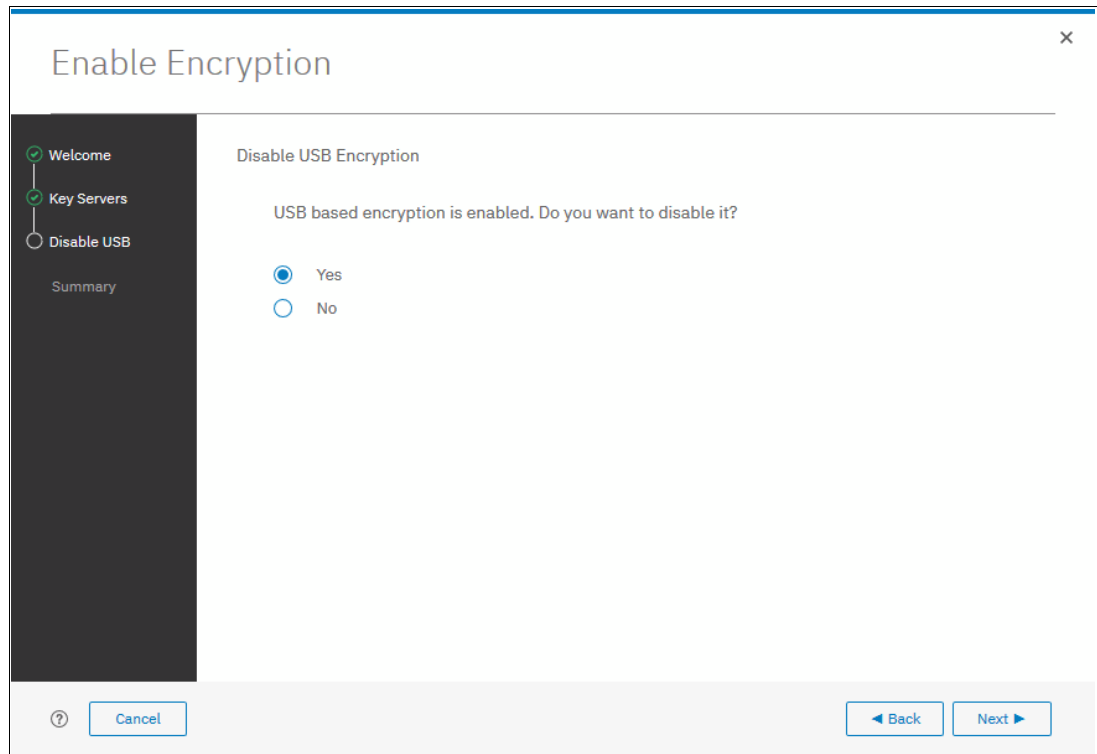


Figure 13-65 Disable USB flash drive provider while changing to SKLM provider

13.6.2 Migrating from encryption key server to USB flash drive provider

Migration in the other direction (that is, from using encryption key servers provider to USB flash drives provider) is not possible using only the GUI.

To perform the migration, add USB flash drives as a second provider. You can do this by completing the steps that are described in 13.5.2, “Adding USB flash drives as a second provider” on page 774. Then, issue the following command in the CLI:

```
chencryption -usb validate
```

To make sure that USB drives contain the correct master access key, disable the encryption key server provider by issuing the following command:

```
chencryption -keyserver disable
```

This command disables the encryption key server provider, which effectively migrates your system from the encryption key server to USB flash drive provider.

13.6.3 Migrating between different key server types

The migration between different key server types cannot be performed directly, from one type of key server to another. USB flash drives encryption needs to be used to facilitate this.

So, if you want to migrate from one type of key server to another, you first need to migrate from your current key servers to USB encryption, and then migrate from USB to the other type of key servers.

The procedure to migrate from one key server type to another is shown here. In this example, we migrate an IBM Spectrum Virtualize system that is configured with IBM SKLM key servers (see Figure 13-66) to SafeNet KeySecure servers.

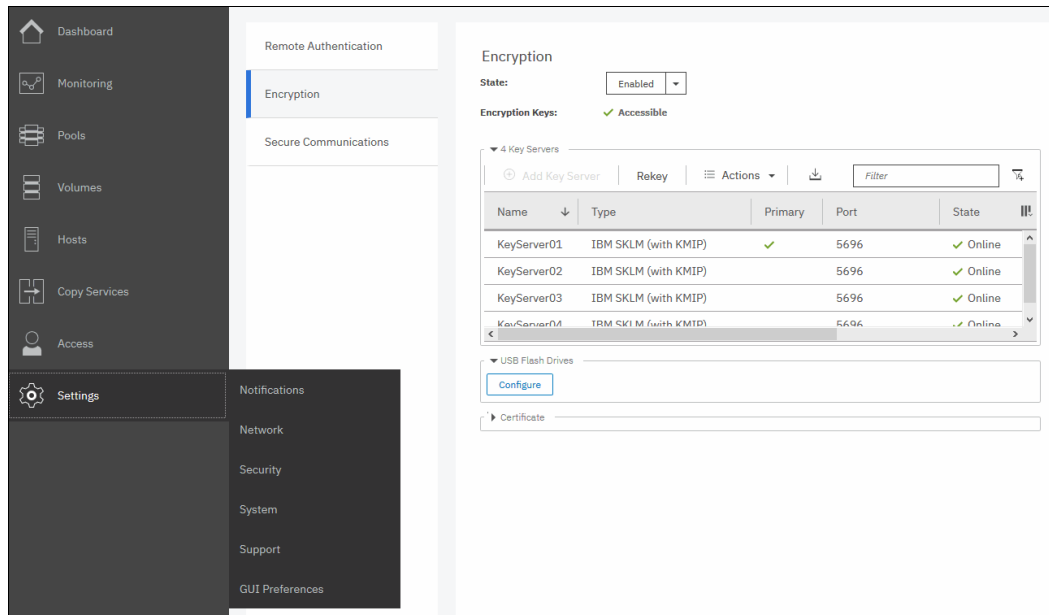


Figure 13-66 IBM Storwize V5000 encryption configured with IBM SKLM servers

To migrate to Gemalto SafeNet KeySecure, complete the following steps:

1. Migrate from key server encryption to USB flash drives encryption, as described in 13.6.2, “Migrating from encryption key server to USB flash drive provider” on page 777.

After this step, only USB flash drives encryption is configured, as shown in Figure 13-67 on page 779.

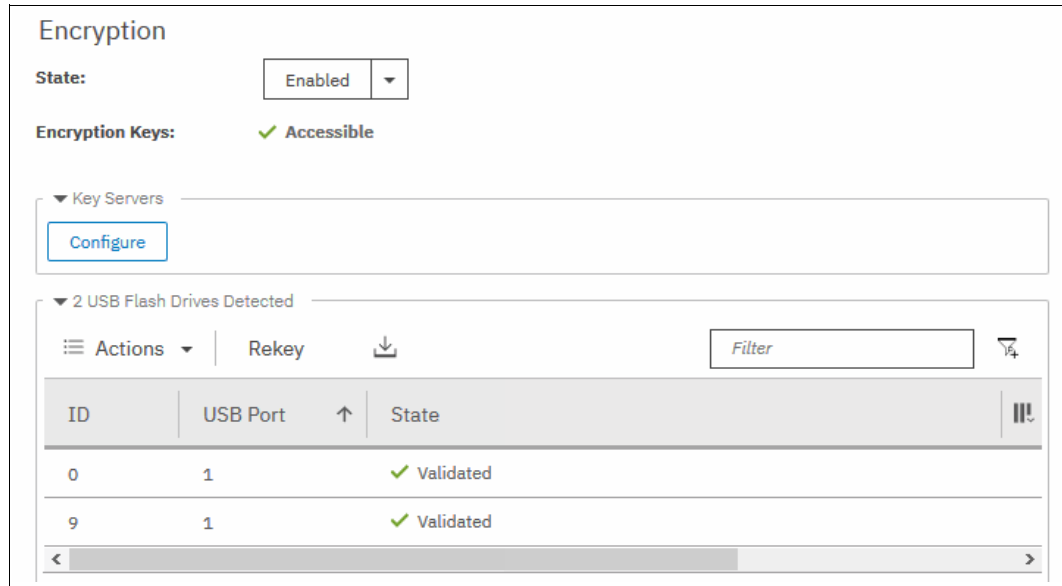


Figure 13-67 IBM Storwize V5000 encryption configured with USB Flash Drives

2. Migrate from USB flash drives encryption to the other key server type encryption (in this example, Gemalto SafeNet KeySecure), following the steps that are described in 13.6.1, “Migrating from USB flash drive provider to encryption key server” on page 777.

After completing this step, the other key server type is configured as encryption provider in IBM Spectrum Virtualize, as shown in Figure 13-68.

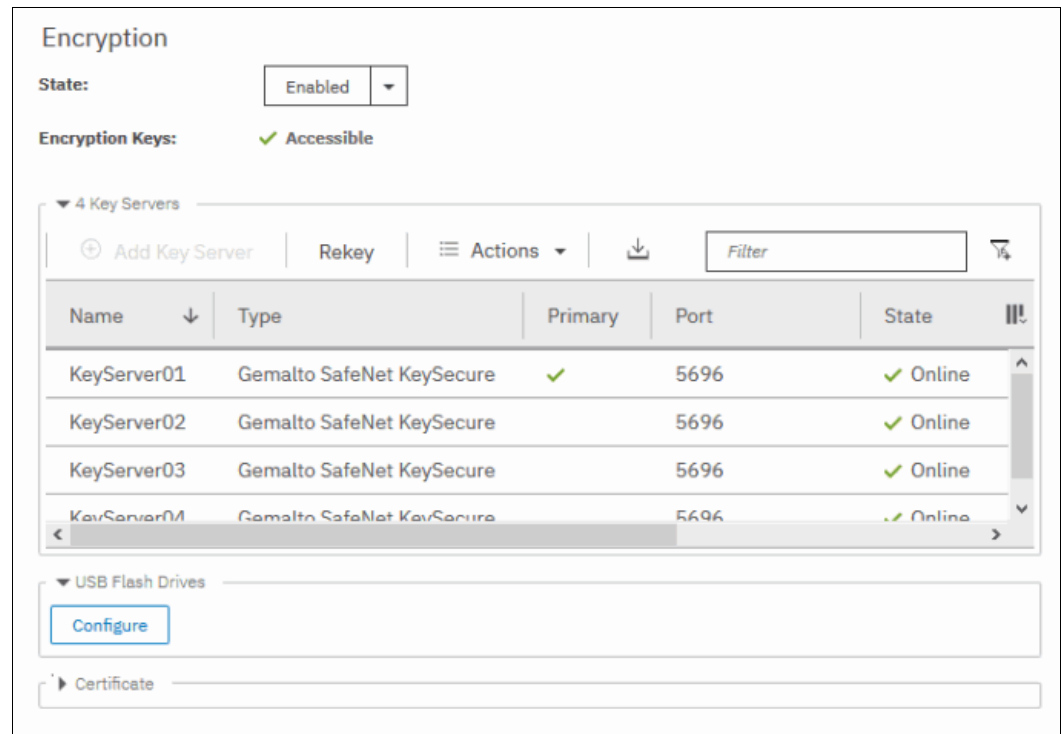


Figure 13-68 IBM Storwize V5000 encryption configured with SafeNet KeySecure

13.7 Recovering from a provider loss

If both encryption key providers are enabled and you lose one of them (by losing all copies of the encryption key kept on the USB flash drives or by losing all SKLM servers), you can recover from this situation by disabling the provider to which you lost the access. To disable the unavailable provider, you must have access to a valid master access key on the remaining provider.

If you have lost access to the encryption key server provider, then issue this command:

```
chencryption -keyserver disable
```

If you have lost access to the USB flash drives provider, then issue this command:

```
chencryption -usb disable
```

If you want to restore the configuration with both encryption key providers, follow the instructions in 13.5, “Configuring more providers” on page 770.

Note: If you lose access to all encryption key providers that are defined in the system, no method is available to recover access to the data that is protected by the master access key.

13.8 Using encryption

The design for encryption is based on the concept that a system is fully encrypted or not encrypted. Encryption implementation is intended to encourage solutions that contain only encrypted volumes or only unencrypted volumes.

For example, after encryption is enabled on the system, all new objects (for example, pools) are by default created as encrypted. Some unsupported configurations are actively policed in code. For example, no support exists for creating unencrypted child pools from encrypted parent pools. However, the following exceptions exist:

- ▶ During the migration of volumes from unencrypted to encrypted volumes, a system might report encrypted and unencrypted volumes.
- ▶ It is possible to create unencrypted arrays from CLI by manually overriding the default encryption setting.

Notes: Encryption support for Distributed RAID is available in IBM Spectrum Virtualize code V7.7 and later.

You must decide whether to encrypt or not encrypt an object when it is created. You cannot change this setting later. To change the encryption state of stored data, you must migrate from an encrypted object (for example, pool) to an unencrypted one, or vice versa. Volume migration is the only way to encrypt any volumes that were created before enabling encryption on the system.

13.8.1 Encrypted pools

For more information about how to open the Create Pool window, see Chapter 4, “Storage pools” on page 159. After encryption is enabled, any new pool is by default created as encrypted, as shown in Figure 13-69.

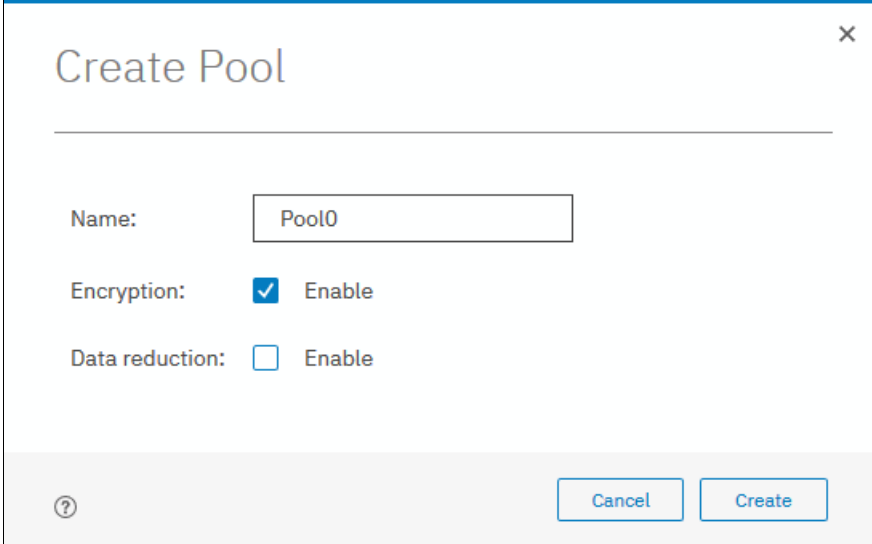
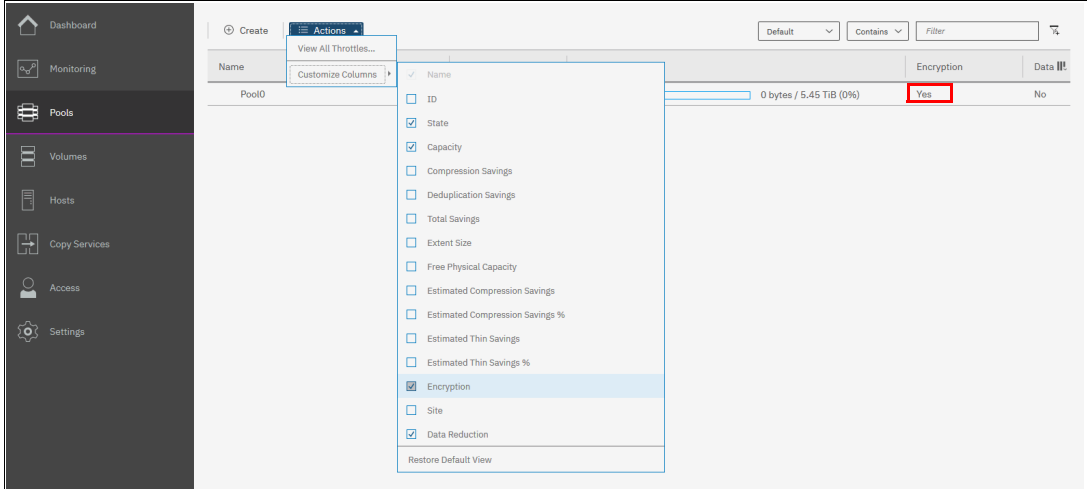


Figure 13-69 Create Pool window basic

You can click **Create** to create an encrypted pool. All storage that is added to this pool is encrypted.

You can customize the Pools view in the management GUI to show pool encryption status. Click **Pools** → **Pools**, and then, click **Actions** → **Customize Columns** → **Encryption**, as shown in Figure 13-70.



Name	Encryption	Data
Pool0	Yes	No

Figure 13-70 Pool encryption state

If you create an unencrypted pool but you add only encrypted arrays or self-encrypting MDisks to the pool, the pool is reported as encrypted because all extents in the pool are encrypted. The pool reverts to the unencrypted state if you add an unencrypted array or MDisk.

More information about how to add encrypted storage to encrypted pools is presented in the following sections. You can mix and match storage encryption types in a pool. Figure 13-71 shows an example of an encrypted pool that contains storage using different encryption methods.

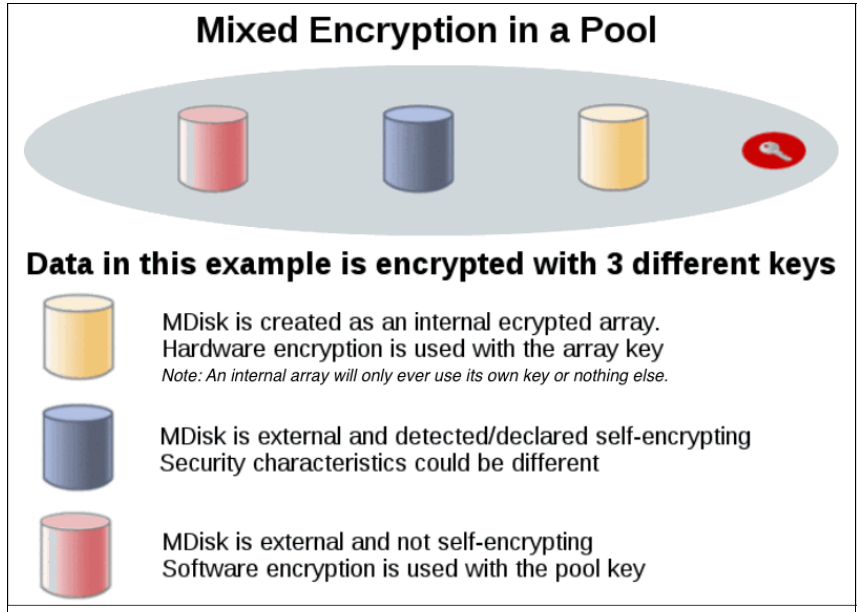


Figure 13-71 Mix and match encryption in a pool

13.8.2 Encrypted child pools

For more information about how to open the Create Child Pool window, see Chapter 4, “Storage pools” on page 159. If the parent pool is encrypted, every child pool also must be encrypted. The GUI enforces this requirement by automatically selecting **Encryption Enabled** in the Create Child Pool window and preventing changes to this setting, as shown in Figure 13-72.

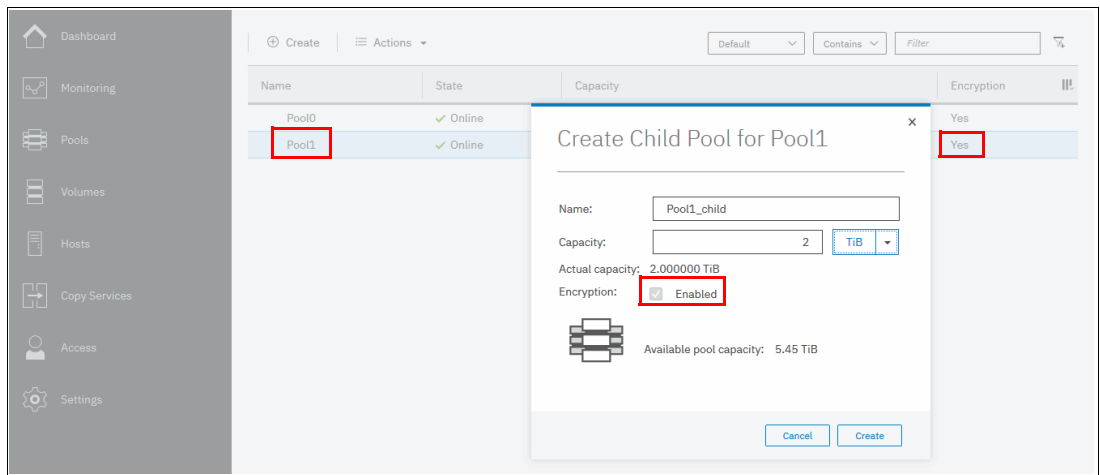


Figure 13-72 Create a child pool of an encrypted parent pool

However, if you want to create encrypted child pools from an unencrypted storage pool containing a mix of internal arrays and external MDisks, the following restrictions apply:

- ▶ The parent pool must not contain any unencrypted internal arrays. If there is any unencrypted internal array in the unencrypted pool, when you try to create a child pool and select the option to set as encrypted, it is created as unencrypted.
- ▶ All V5000 Storwize control enclosures in the system must support software encryption and have the encryption license activated.

Note: An encrypted child pool created from an unencrypted parent storage pool reports as unencrypted if the parent pool contains any unencrypted internal arrays. Remove these arrays to ensure that the child pool is fully encrypted.

If you modify the Pools view, you see the encryption status of child pools, as shown in Figure 13-73. The example shows an encrypted child pool with non-encrypted parent pool.

Name	State	Capacity	Encryption
Pool0	Online	0 bytes / 5.45 TiB (0%)	Yes
Pool1	Online	0 bytes / 5.45 TiB (0%)	Yes
Pool1_child	Online	0 bytes / 2.00 TiB (0%)	Yes
Pool2	Online	0 bytes / 5.45 TiB (0%)	No
Pool2_child	Online	0 bytes / 2.00 TiB (0%)	Yes

Figure 13-73 Child pool encryption state

13.8.3 Encrypted arrays

For more information about how to add internal storage to a pool, see Chapter 4, “Storage pools” on page 159. After encryption is enabled, all newly built arrays are hardware encrypted by default. In this case, the GUI does not allow you to create an unencrypted array. To create an unencrypted array, the CLI needs to be used. Example 13-1 shows how to create an unencrypted array by using the CLI.

Example 13-1 Creating an unencrypted array by using CLI

```
IBM_Storwize:ITS0-V5k:superuser>svctask mkarray -drive 6:4 -level raid1 -sparegoal
0 -strip 256 -encrypt no Pool2
MDisk, id [2], successfully created
IBM_Storwize:ITS0-V5k:superuser>
```

Note: It is not possible to add unencrypted arrays to an encrypted pool.

You can customize MDisks by Pools view to show array encryption status. Click **Pools** → **MDisk by Pools**, and then, click **Actions** → **Customize Columns** → **Encryption**. You also can right-click the table header to customize columns and select **Encryption**, as shown in Figure 13-74.

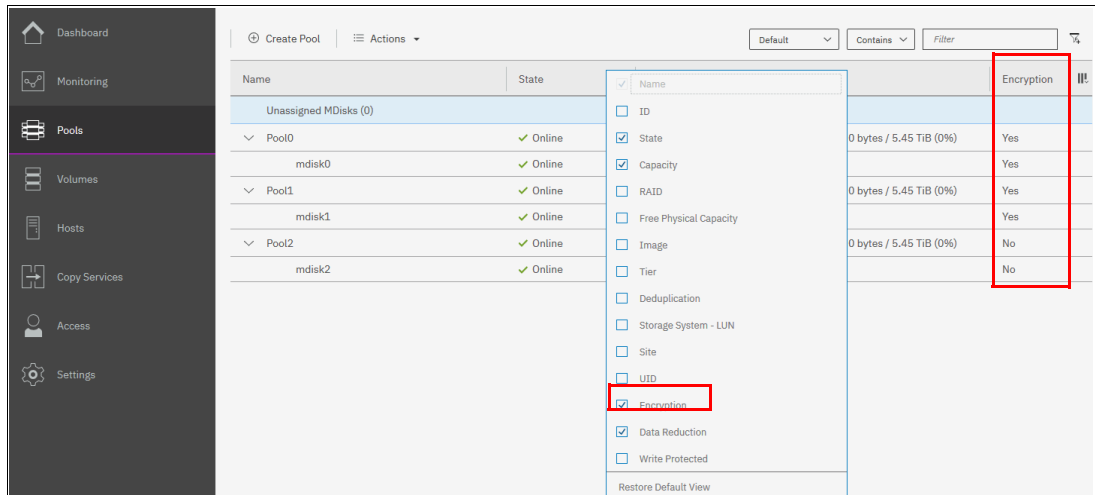


Figure 13-74 Array encryption state

You can also check the encryption state of an array by reviewing its drives in **Pools** → **Internal Storage** view. The internal drives that are associated with an encrypted array are assigned an encrypted property that can be seen as shown in Figure 13-75.

Drive ID	Capacity	Use	Status	MDisk Name	Slot ID	Encrypted
0	5.46 TiB	Member	Online	mdisk0	11	✓
1	5.46 TiB	Spare	Online		1	
2	5.46 TiB	Member	Online	mdisk0	9	✓
3	5.46 TiB	Member	Online	mdisk1	2	✓
4	5.46 TiB	Member	Online	mdisk2	6	
5	5.46 TiB	Member	Online	mdisk1	10	✓
6	5.46 TiB	Member	Online	mdisk2	5	

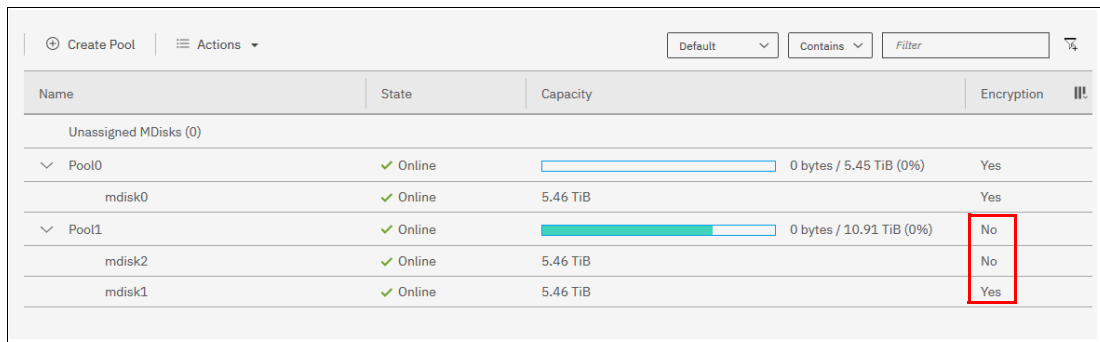
Figure 13-75 Drive encryption state

13.8.4 Encrypted MDisks

For more information about how to add external storage to a pool, see Chapter 4, “Storage pools” on page 159. Each MDisk that belongs to external storage that is added to an encrypted pool or child pool is automatically encrypted by using the pool or child pool key, unless the MDisk is detected or declared as self-encrypting.

The user interface gives no method to see which extents contain encrypted data and which do not. However, if a volume is created in a correctly configured encrypted pool, all data written to this volume is encrypted.

You can use the MDisk by Pools view to show the object encryption state by clicking **Pools** → **MDisk by Pools**. Figure 13-76 shows an example in which a self-encrypting MDisk is in an unencrypted pool.



The screenshot shows a table with columns: Name, State, Capacity, and Encryption. It lists two pools: Pool0 and Pool1. Pool0 contains mdisk0 (5.46 TiB, Yes encryption). Pool1 contains mdisk2 (5.46 TiB, No encryption) and mdisk1 (5.46 TiB, Yes encryption). The 'No' entry for mdisk2 is highlighted with a red box.

Name	State	Capacity	Encryption
Unassigned MDisks (0)			
Pool0	Online	0 bytes / 5.45 TiB (0%)	Yes
mdisk0	Online	5.46 TiB	Yes
Pool1	Online	0 bytes / 10.91 TiB (0%)	No
mdisk2	Online	5.46 TiB	No
mdisk1	Online	5.46 TiB	Yes

Figure 13-76 MDisk encryption state

When working with MDisks encryption, you *must* take extra care when configuring MDisks and pools.

If the MDisk was used earlier for storage of unencrypted data, the extents can contain stale unencrypted data. This is because file deletion only marks disk space as free. The data is not removed from the storage. Therefore, if the MDisk is not self-encrypting and was a part of an unencrypted pool and later was moved to an encrypted pool, it contains stale data from its previous life.

Another mistake that can happen is to misconfigure an external MDisk as self-encrypting, while in reality it is not self-encrypting. In that case, the data that is written to this MDisk is not encrypted by Storwize V5000 because Storwize V5000 is convinced that the MDisk encrypts the data. At the same time, the MDisk does not encrypt the data because it is not self-encrypting; therefore, unencrypted data is on an extent in an encrypted pool.

However, all data that is written to any MDisk that is a part of correctly configured encrypted pool is going to be encrypted.

Self-encrypting MDisks

When adding external storage to a pool, be exceptionally diligent when declaring the MDisk as self-encrypting. Correctly declaring an MDisk as self-encrypting avoids wasting resources, such as CPU time. However, when used improperly, it might lead to unencrypted data at-rest.

To declare an MDisk as self-encrypting, select **Externally encrypted** when adding external storage in the Assign Storage view, as shown in Figure 13-77.

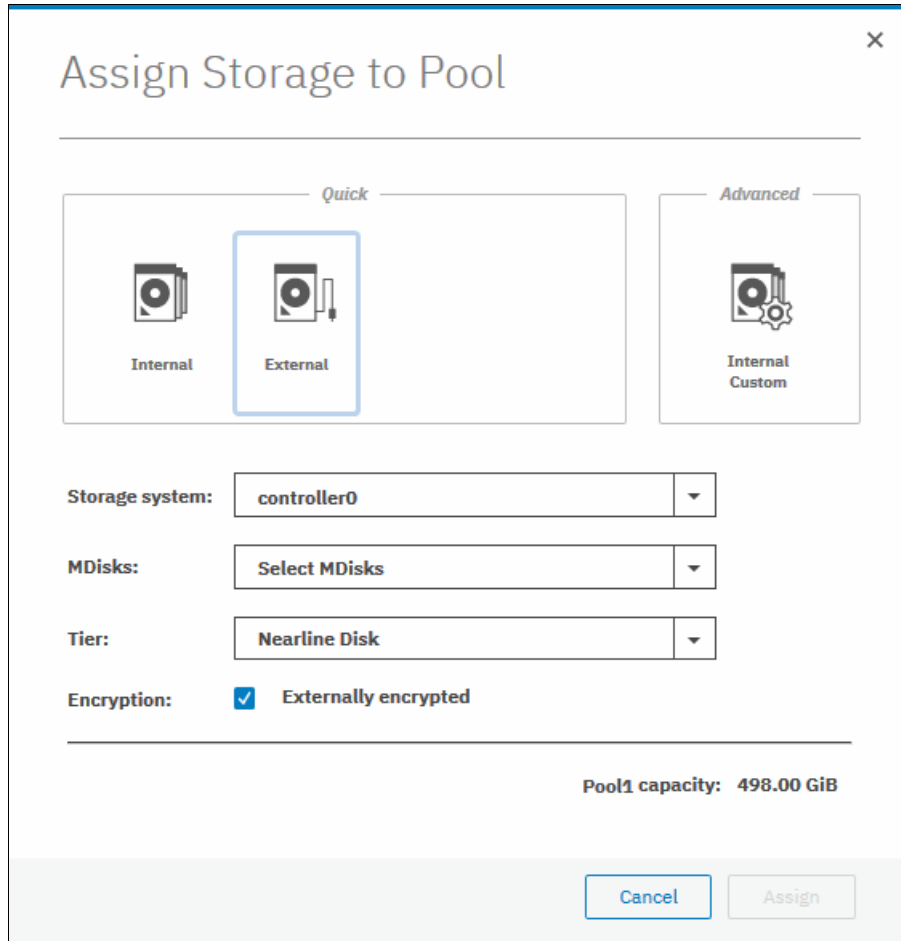


Figure 13-77 Declaring MDisk as externally encrypted

IBM Spectrum Virtualize products can detect that an MDisk is self-encrypting by using the SCSI Inquiry page C2. MDisks that are provided by other IBM Spectrum Virtualize products report this page correctly. For these MDisks, the Externally encrypted option that is shown in Figure 13-77 *not* selected. However, when added, they are still considered as self-encrypting.

Note: You can override external encryption setting of an MDisk detected as self-encrypting and configure it as unencrypted by using the CLI command `chmdisk -encrypt no`. However, only do so if you plan to decrypt the data on the backend or if the backend uses inadequate data encryption.

To check whether an MDisk was detected or declared as self-encrypting, click **Pools** → **MDisk by Pools** and verify in the Encryption column, as shown in Figure 13-78.

Name	State	Capacity	Encryption	!!!
Unassigned MDisks (1)				
Pool0	Online	24.00 GiB / 99.00 GiB (24%)	No	
mdisk0	Online	100.00 GiB	No	
Pool1	Online	0 bytes / 498.00 GiB (0%)	Yes	
mdisk1	Online	200.00 GiB	Yes	
mdisk2	Online	300.00 GiB	No	

Figure 13-78 MDisk self-encryption state

The value that is shown in the Encryption column shows the property of objects in respective rows. That means that in the configuration that is shown in Figure 13-78, Pool1 is encrypted, so every volume created from this pool is encrypted. However, that pool is formed by two MDisks, out of which one is self-encrypting and one is not. Therefore, a value of No next to mdisk2 does not imply that encryption of Pool1 is in any way compromised. It only indicates that encryption of the data placed on mdisk2 is done by using software encryption, while data placed on mdisk1 is encrypted by the back-end storage that is providing these MDisks.

Note: You can change the self-encrypting attribute of an MDisk that is unmanaged or member of an unencrypted pool. However, you cannot change the self-encrypting attribute of an MDisk after it is added to an encrypted pool.

13.8.5 Encrypted volumes

For more information about how to create and manage volumes, see Chapter 6, “Volume configuration” on page 309. The encryption status of a volume depends on the pool encryption status. Volumes that are created in an encrypted pool are automatically encrypted.

You can modify Volumes view to show if the volume is encrypted. Click **Volumes** → **Volumes**, then, click **Actions** → **Customize Columns** → **Encryption** to customize the view to show volumes encryption status, as shown in Figure 13-79.

Name	State	Synchronized	Pool	UID	Encryption	!!!
Volume000	Online (formatting)		Pool0	6005076801B807F934000000000000...	Yes	
Volume001	Online (formatting)		Pool0	6005076801B807F934000000000000...	Yes	
Volume002	Online (formatting)		Pool0	6005076801B807F934000000000000...	Yes	
Volume003	Online		Pool1	6005076801B807F934000000000000...	No	
Volume004	Online		Pool1	6005076801B807F934000000000000...	No	
Volume005	Online		Pool1	6005076801B807F934000000000000...	No	
Volume006	Online		Pool0	6005076801B807F934000000000000...	No	
Volume007	Online		Pool0	6005076801B807F934000000000000...	No	
Volume008	Online		Pool0	6005076801B807F934000000000000...	Yes	
Volume009	Online		Pool0	6005076801B807F934000000000000...	Yes	
Volume010	Online		Pool1	6005076801B807F934000000000000...	No	
Volume011	Online		Pool1	6005076801B807F934000000000000...	No	

Figure 13-79 Volume view customization

A volume is reported as encrypted only if all the volume copies are encrypted, as shown in Figure 13-80.

Name	State	Synchronized	Pool	Encryption
Volume003	Online		Pool0	Yes
Copy 0*	Online	Yes	Pool0	Yes
Copy 1	Online	Yes	Pool0	Yes
Volume004	Online		Pool1	No
Copy 0*	Online	Yes	Pool1	No
Copy 1	Online	Yes	Pool0	Yes

Figure 13-80 Volume encryption status depending on volume copies encryption

When creating volumes, make sure to select encrypted pools to create encrypted volumes, as shown in Figure 13-81.

Figure 13-81 Create an encrypted volume by selecting an encrypted pool

You cannot change an unencrypted volume to an encrypted version of itself dynamically. However, this conversion is possible by using one of the following migration options:

- ▶ Migrate a volume to an encrypted pool or child pool.
- ▶ Mirror a volume to an encrypted pool or child pool and delete the unencrypted copy.

For more information about these methods, see Chapter 6, “Volume configuration” on page 309.

13.8.6 Restrictions

The following restrictions apply to encryption:

- ▶ Image mode volumes cannot be in encrypted pools.
- ▶ You cannot add external non self-encrypting MDisks to encrypted pools unless all control enclosures in the system support encryption.

13.9 Rekeying an encryption-enabled system

Changing the master access key is a security requirement. *Rekeying* is the process of replacing current master access key with a newly generated one. The rekey operation works whether encrypted objects exist.

The rekeying operation requires access to a valid copy of the original master access key on an encryption key provider that you plan to rekey. Use the rekey operation according to the schedule that is defined in your organization’s security policy and whenever you suspect that the key might be compromised.

If USB and key server are both enabled, rekeying is done separately for each of the providers.

Important: Before you create a master access key, ensure that all nodes are online and that the current master access key is accessible.

There is no method to directly change data encryption keys. If you need to change the data encryption key that is used to encrypt data, the only available method is to migrate that data to a new encrypted object (for example, an encrypted child pool). Because the data encryption keys are defined per encrypted object, such migration forces a change of the key that is used to encrypt that data.

13.9.1 Rekeying using a key server

Ensure that all the configured key servers can be reached by the system and that service IPs are configured on all your nodes.

To rekey the master access key kept on the key server provider, complete the following steps:

1. Click **Settings** → **Security** → **Encryption**, ensure that Encryption Keys shows that all configured SKLM servers are reported as Accessible, as shown in Figure 13-82. Click **Key Servers** to expand the section.

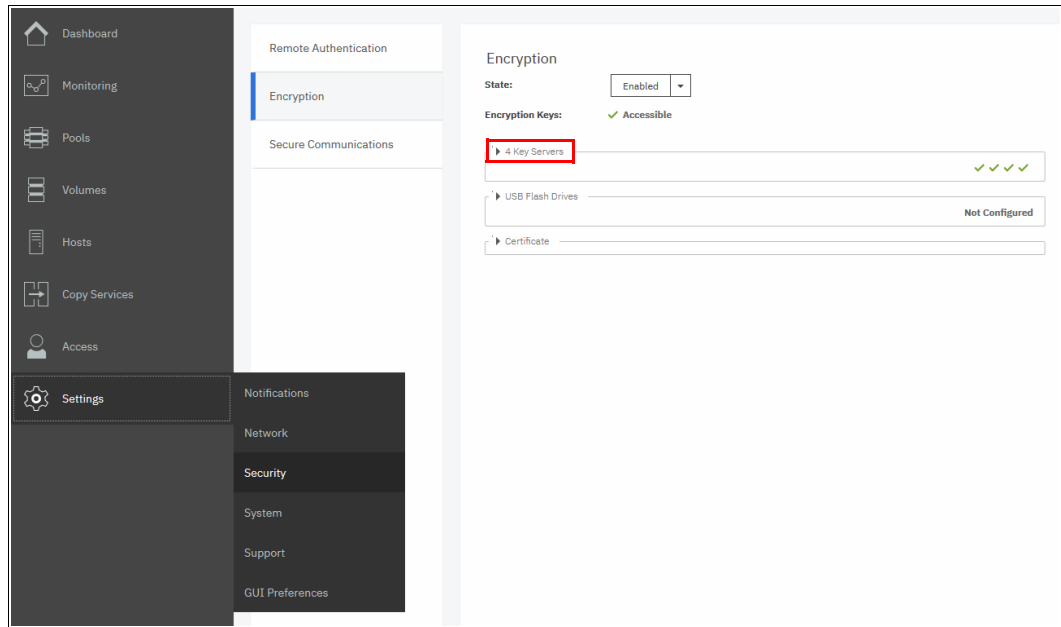


Figure 13-82 Locate Key Servers section on Encryption window

2. Click **Rekey**, as shown in Figure 13-83.

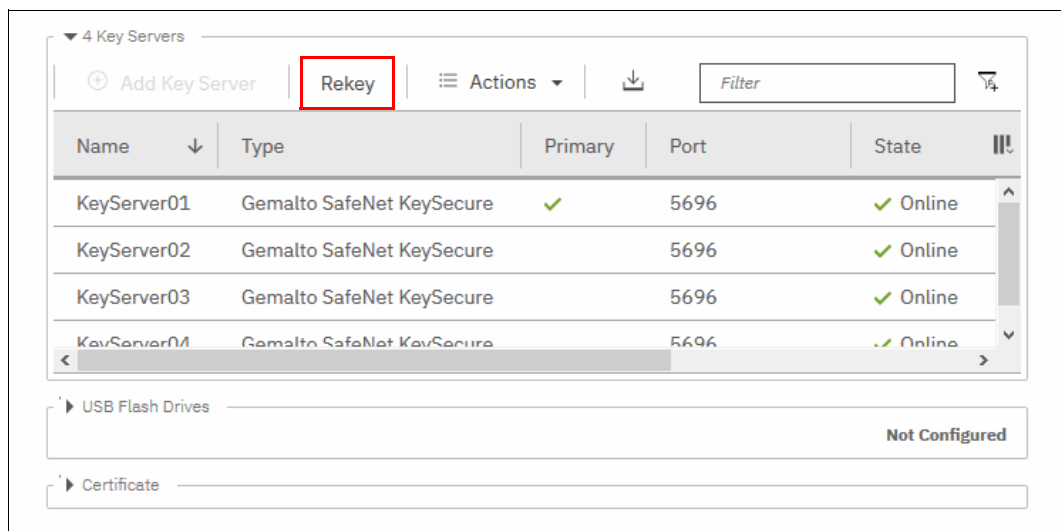


Figure 13-83 Start rekey on SKLM key server

3. Click **Yes** in the next window to confirm the rekey operation, as shown in Figure 13-84.

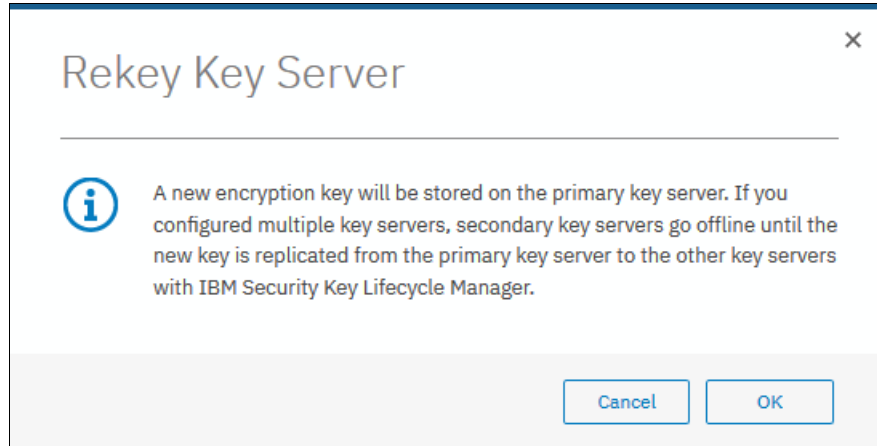


Figure 13-84 Confirm key server rekey operation

Note: The rekey operation is performed on only the primary key server that is configured in the system. If you have more key servers configured apart from the primary one, they do not hold the updated encryption key until they obtain it from the primary key server. To restore encryption key provider redundancy after a rekey operation, replicate the encryption key from the primary key server to the secondary key servers.

You receive a message confirming that the rekey operation was successful, as shown in Figure 13-85.

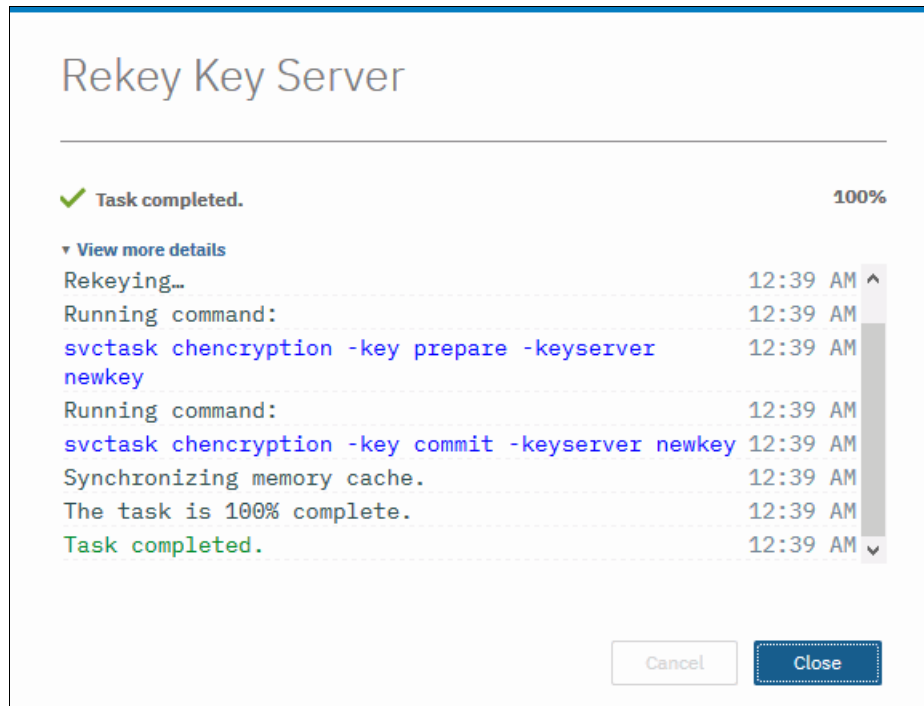


Figure 13-85 Successful key server rekey operation

13.9.2 Rekeying using USB flash drives

During the rekey process, new keys are generated and copied to the USB flash drives. These keys are then used instead of the current keys. The rekey operation fails if at least one of the USB flash drives does not contain the current key. To rekey the system, you need at least three USB flash drives to store the master access key copies.

After the rekey operation is complete, update all other copies of the encryption key, including copies stored on other media. Take the same precautions to securely store all copies of the new encryption key as when you were enabling encryption for the first time.

To rekey the master access key on USB flash drives, complete the following steps:

1. Click **Settings** → **Security** → **Encryption**. Click **USB Flash Drives** to expand the section, as shown in Figure 13-86.

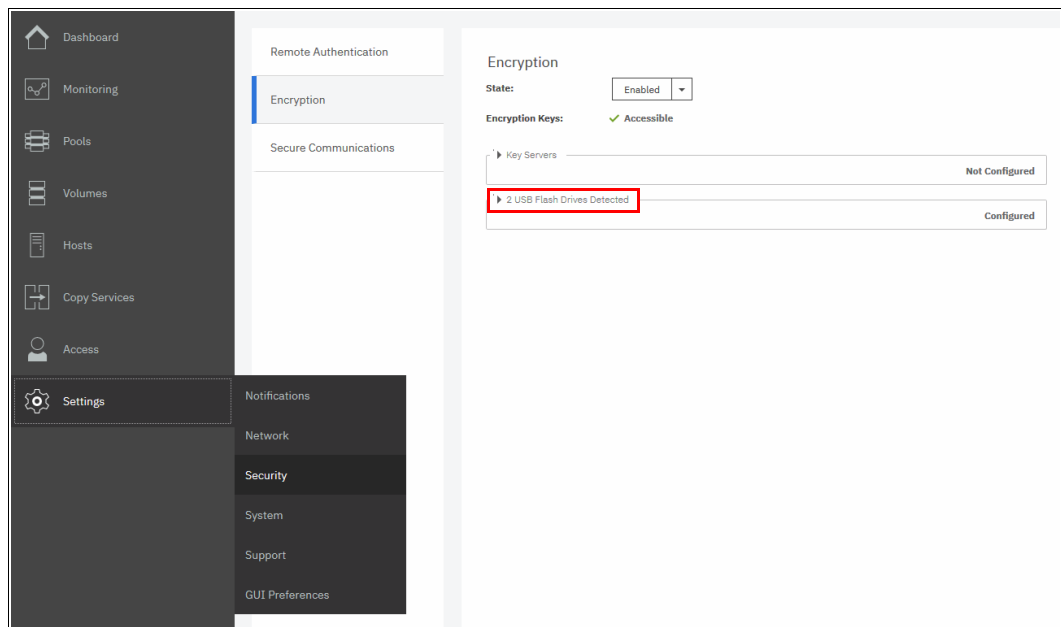


Figure 13-86 Locate USB Flash Drive section in the Encryption view

2. Verify that all USB drives plugged into the system are detected and show as Validated, as shown in Figure 13-87. Click **Rekey**. You need at least three USB flash drives, with at least one reported as Validated to process with rekey.

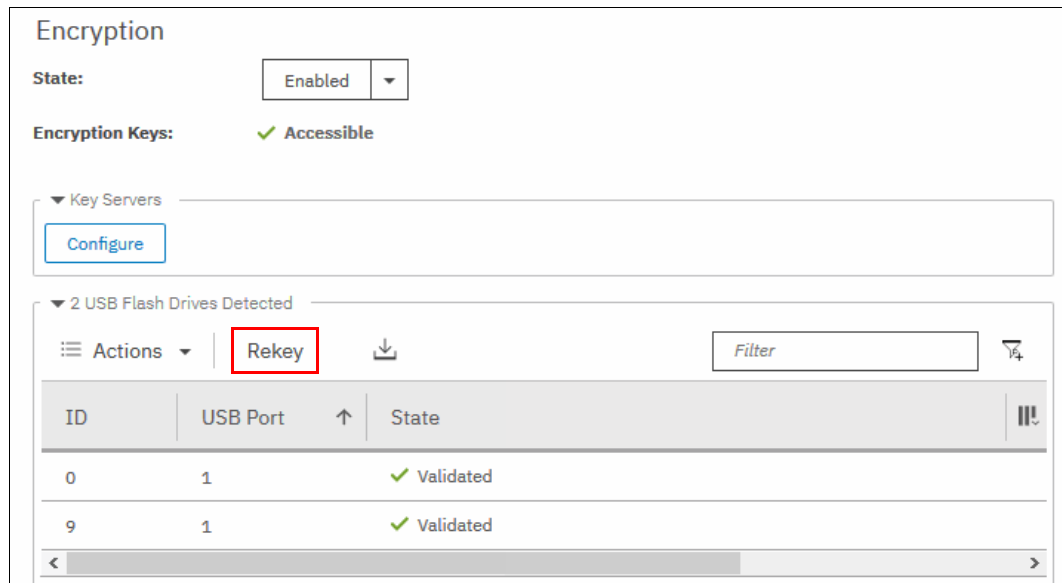


Figure 13-87 Start rekey on USB flash drives provider

3. If the system detects a validated USB flash drive and at least three available USB flash drives, new encryption keys are automatically copied on the USB flash drives, as shown in Figure 13-88 on page 794. Click **Commit** to finalize the rekey operation.

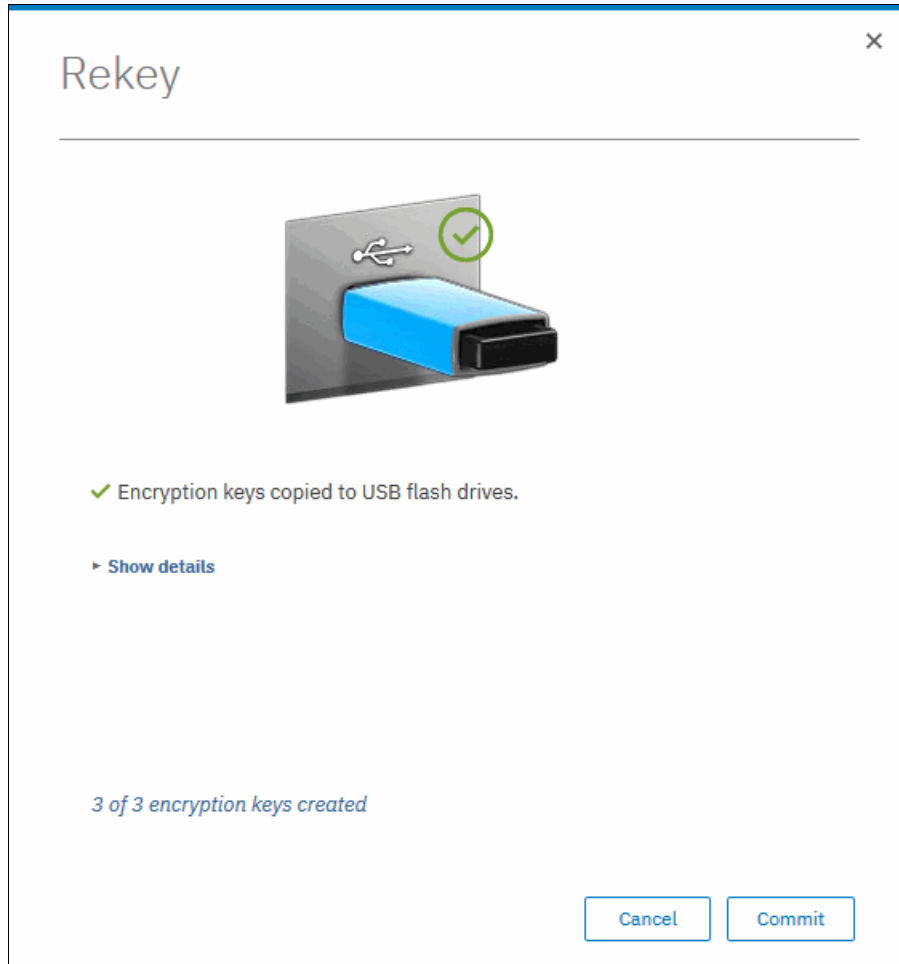


Figure 13-88 Writing new keys to USB flash drives

4. You should receive a message confirming the rekey operation was successful, as shown in Figure 13-89. Click **Close**.

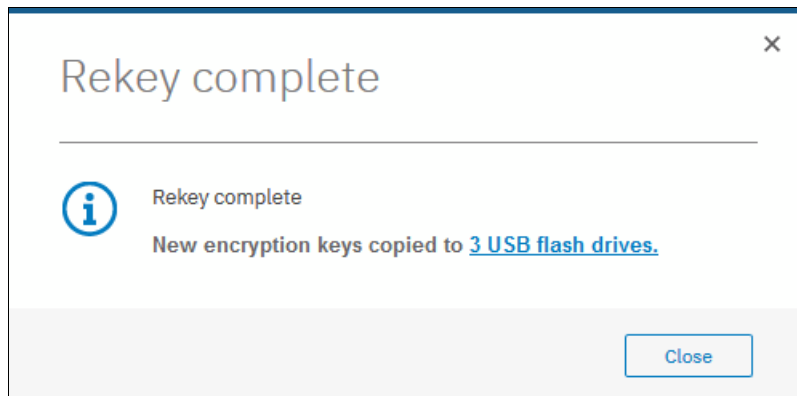


Figure 13-89 Successful rekey operation using USB flash drives

13.10 Disabling encryption

You are prevented from disabling encryption if any encrypted objects are defined apart from self-encrypting MDisks. You can disable encryption in the same way whether you use USB flash drives, key server, or both providers.

To disable encryption, complete the following steps:

1. Click **Settings** → **Security** → **Encryption** and click **Enabled**. A menu is displayed if no encrypted objects exist. Click **Disabled** to disable encryption on the system. Figure 13-90 shows an example for a system with both encryption key providers configured.

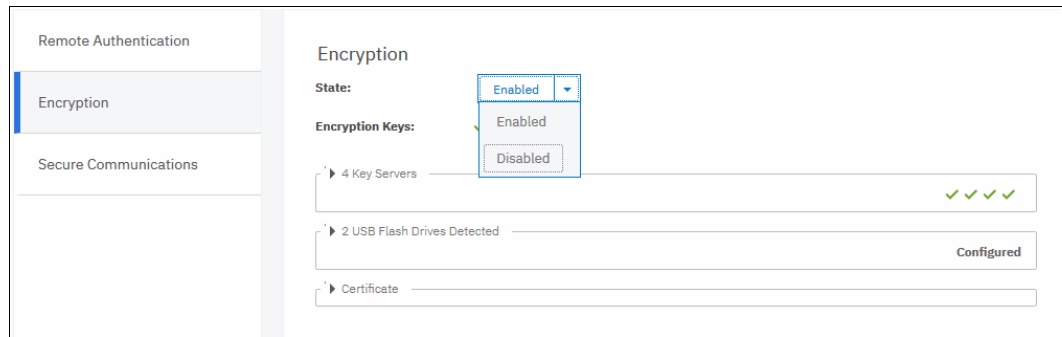


Figure 13-90 Disabling encryption on a system with both providers

2. You receive a message confirming that encryption was disabled. Figure 13-91 shows the message when a key server is used.

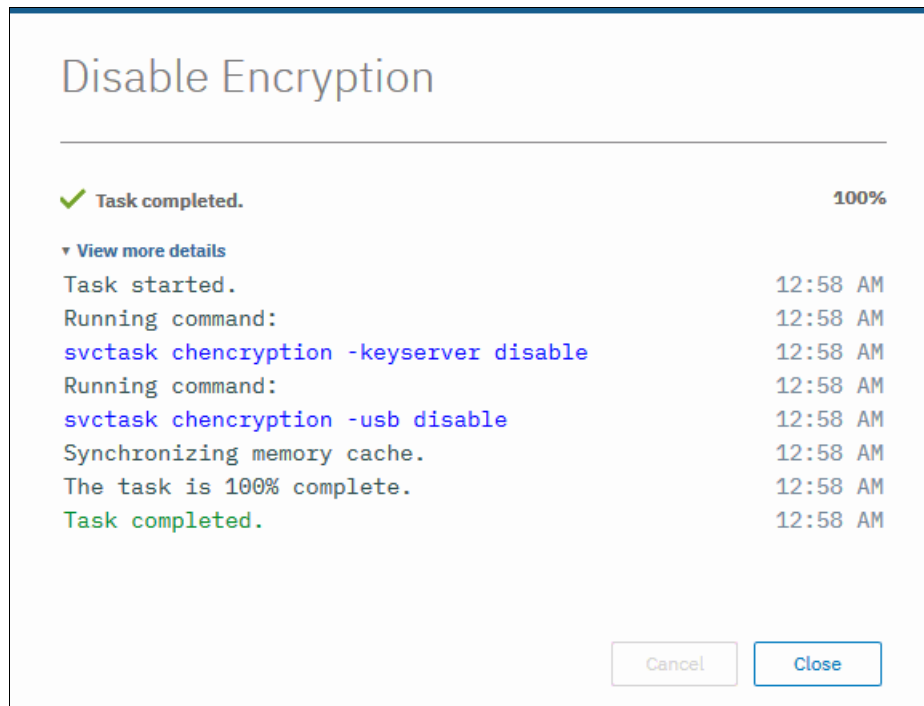


Figure 13-91 Encryption disabled



CLI setup and SAN Boot

This appendix describes the setup of the command-line interface (CLI) and provides information about the SAN Boot function.

The appendix includes the following topics:

- ▶ “Command-line interface” on page 798
- ▶ “SAN Boot” on page 812

Command-line interface

The IBM Storwize V5000 Gen2 system has a powerful CLI, which offers even more functions than the graphical user interface (GUI). This section is not intended to be a detailed guide to the CLI because that topic is beyond the scope of this book.

The basic configuration of the IBM Storwize V5000 Gen2 CLI is covered. Example commands are described. However, the CLI commands are the same in the IBM SAN Volume Controller and Storwize family, and more commands are available to manage internal storage. If a task completes in the GUI, the CLI command is always displayed in the details, as shown throughout this book.

For more information about CLI, see this IBM Storwize V5000 Gen2 [web page](#).

Basic setup

In the IBM Storwize V5000 Gen2 GUI, authentication is performed by using a user name and password. The CLI uses a Secure Shell (SSH) to connect from the host to the IBM Storwize V5000 Gen2 system. A private and a public key pair or user name and password combination is necessary.

The following steps are required to enable CLI access with SSH keys:

1. A public key and a private key are generated as a pair.
2. A public key is uploaded to the IBM Storwize V5000 Gen2 system through the GUI.
3. A client SSH tool must be configured to authenticate with the private key.
4. A secure connection can be established between the client and the IBM Storwize V5000 Gen2.

SSH is the communication vehicle between the management workstation and the IBM Storwize V5000 Gen2 system. The SSH client provides a secure environment from which to connect to a remote machine. It uses the principles of public and private keys for authentication.

The system supports up to 32 interactive SSH sessions on the management IP address simultaneously. After 1 hour, a fixed SSH interactive session times out, which means that the SSH session is automatically closed. This session timeout limit is not configurable.

SSH keys are generated by the SSH client software. The SSH keys include a public key, which is uploaded and maintained by the clustered system, and a private key, which is kept private on the workstation that is running the SSH client. These keys authorize specific users to access the administration and service functions on the system.

Each key pair is associated with a user-defined ID string that consists of up to 30 characters. Up to 100 keys can be stored on the system. New IDs and keys can be added, and unwanted IDs and keys can be deleted. To use the CLI, an SSH client must be installed on that system, the SSH key pair must be generated on the client system, and the client's SSH public key must be stored on the IBM Storwize V5000 Gen2.

The SSH client that is used in this book is PuTTY. Also, a PuTTY key generator can be used to generate the private and public key pair. The PuTTY client can be downloaded at no cost from [this website](#).

Download the following tools:

- ▶ PuTTY SSH client: `putty.exe`
- ▶ PuTTY key generator: `puttygen.exe`

Generating a public and private key pair

To generate a public and private key pair, complete the following steps:

1. Start the PuTTY key generator to generate the public and private key pair (see Figure A-1).

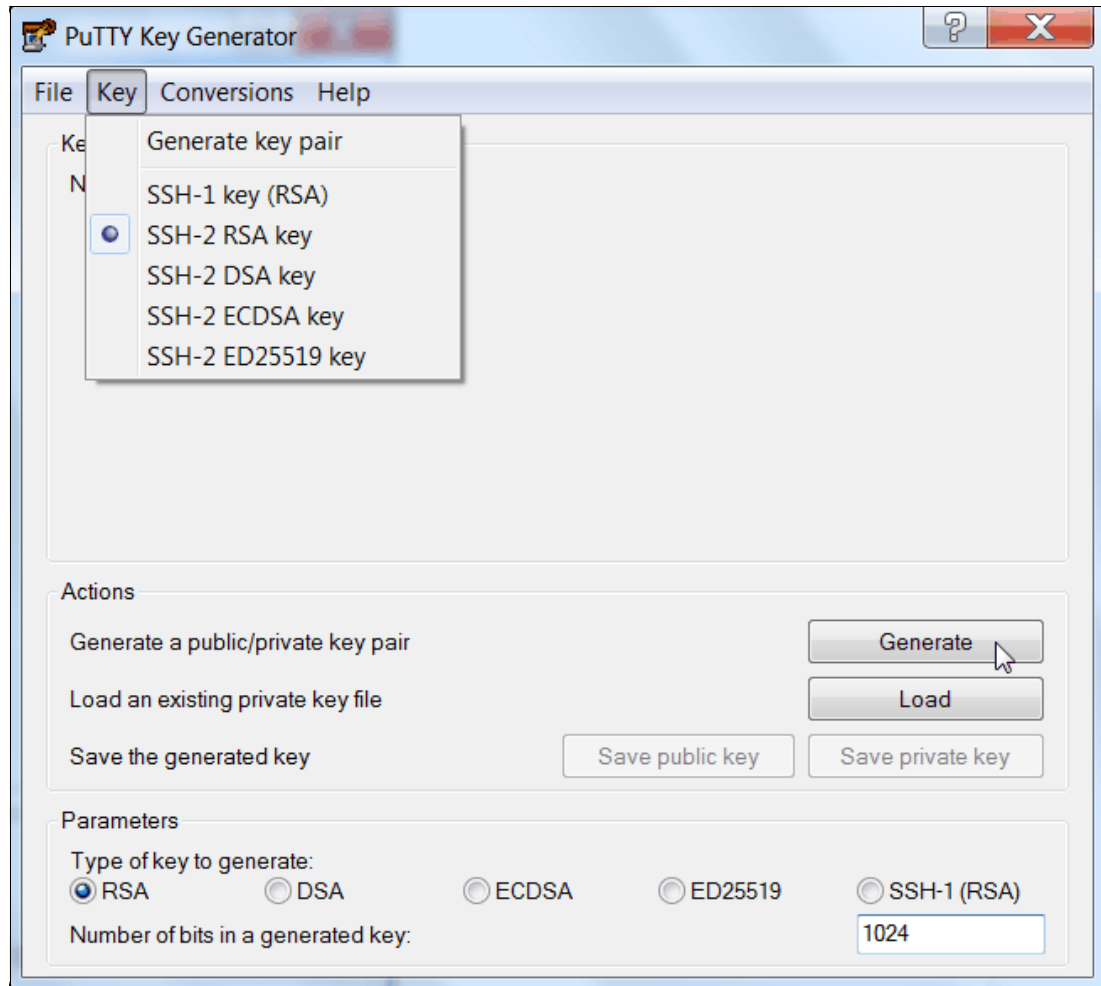


Figure A-1 PuTTY key generator

Ensure that the following options are used:

- SSH-2 RSA
- Number of bits in a generated key: 1024

Note: Ignore the warning not to create keys less than 2048 bits, as shown in Figure A-2 on page 800. Click **OK** to continue.

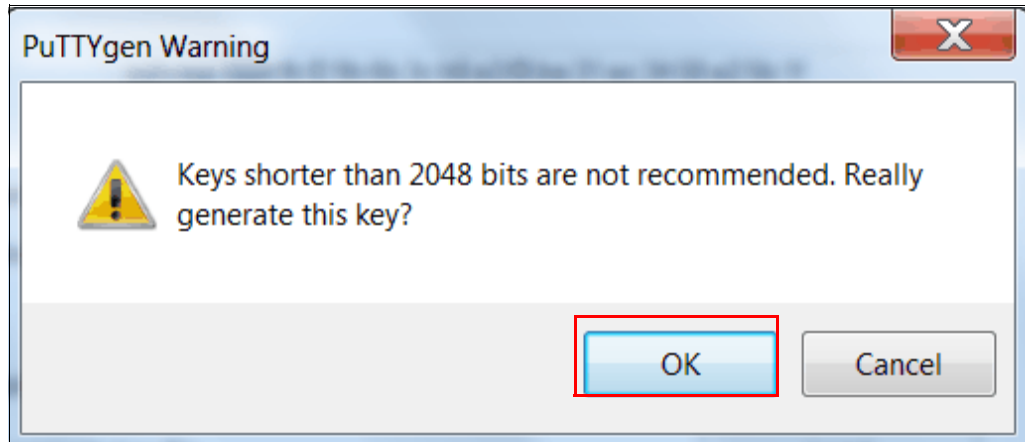


Figure A-2 Warning

2. Click **Generate** and move the cursor over the blank area to generate keys (see Figure A-3).

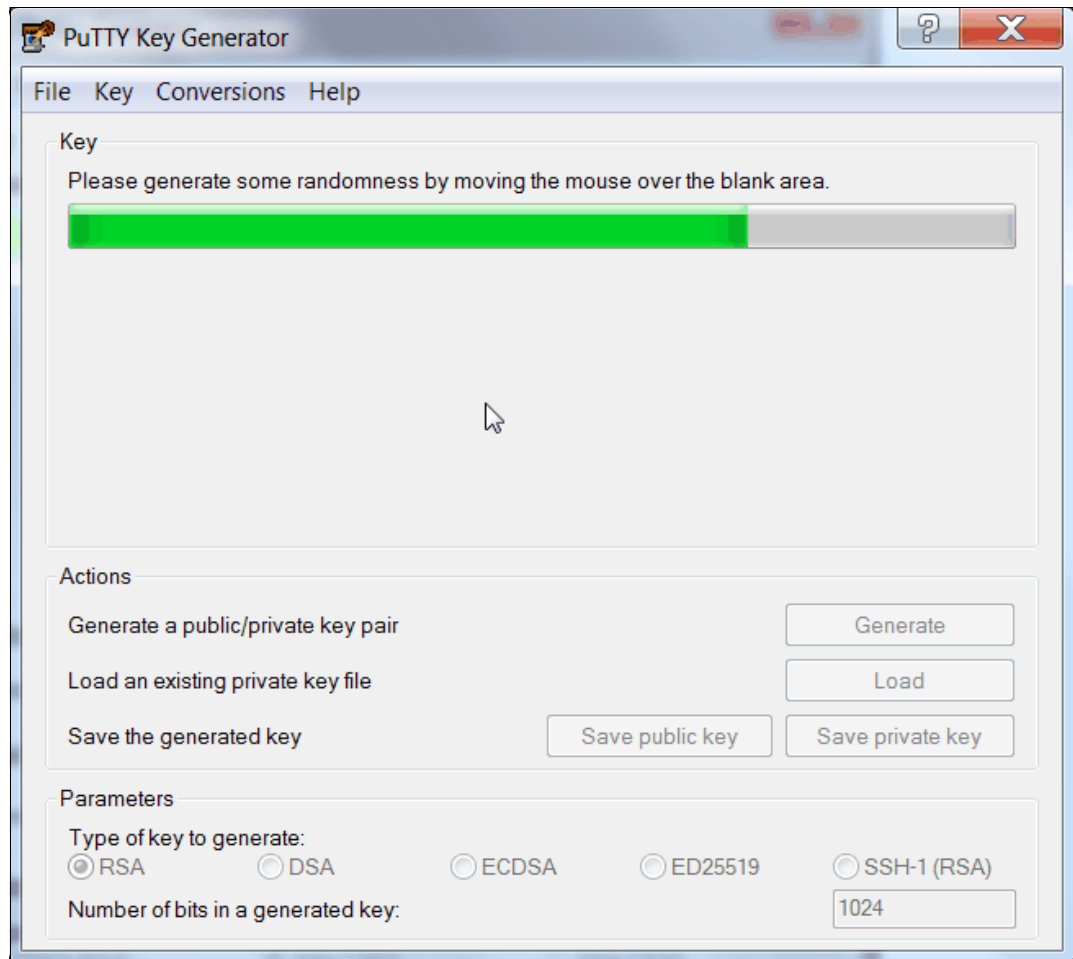


Figure A-3 Generate keys

Generating keys: The blank area that is indicated by the message is the large blank rectangle on the GUI inside the section of the GUI that is labeled Key. Continue to move the mouse pointer over the blank area until the progress bar reaches the far right side. This action generates random characters to create a unique key pair.

3. After the keys are generated, save them for later use. Click **Save public key** (see Figure A-4).

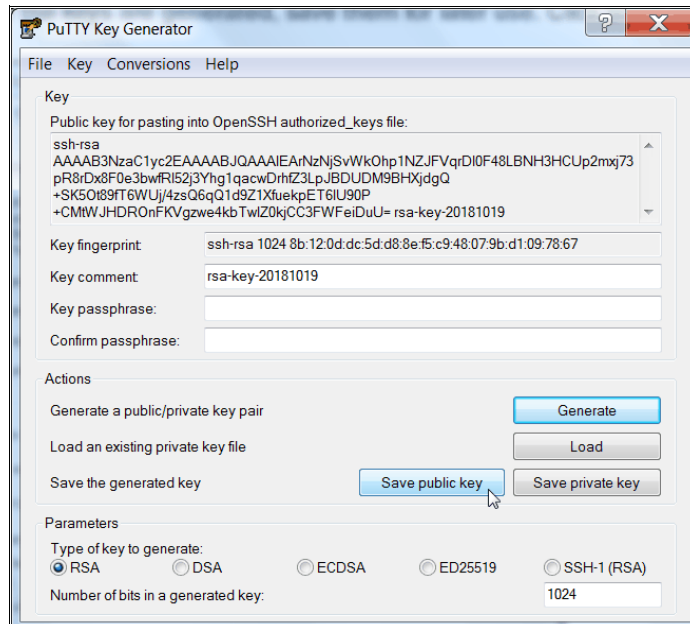


Figure A-4 Save public key

4. You are prompted for a name (for example, pubkey) and a location for the public key (for example, C:\Support Utils\PuTTY). Click **Save**.

Ensure that you record the name and location because the name and location of this SSH public key must be specified later.

Public key extension: By default, the PuTTY key generator saves the public key with no extension. Use the string “pub” for naming the public key (for example, superuser.pub) to easily differentiate the SSH public key from the SSH private key.

5. Click **Save private key** (see Figure A-5).

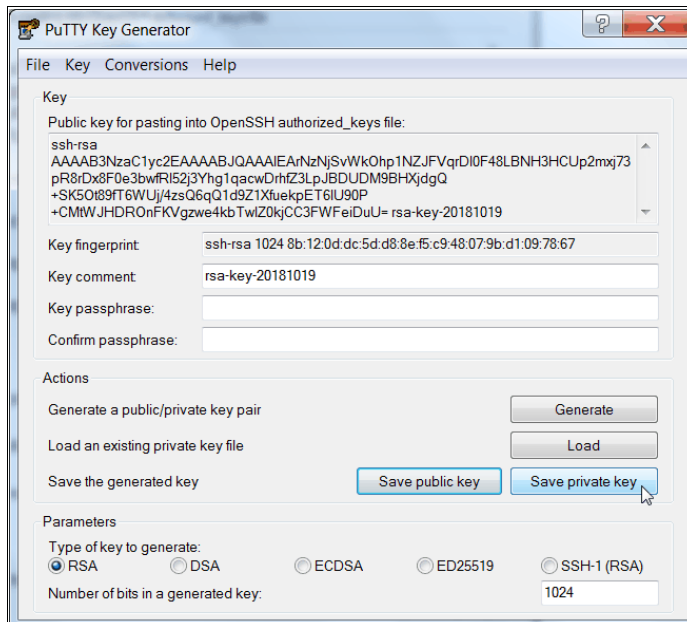


Figure A-5 Save private key

6. You are prompted with a warning message (see Figure A-6). Click **Yes** to save the private key without a passphrase.

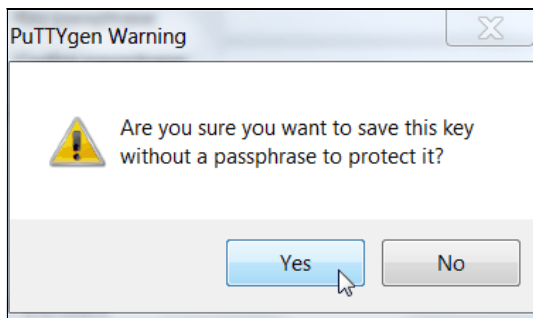


Figure A-6 Confirm the security warning

7. When you are prompted, enter a name (for example, i cat), select a secure place as location, and click **Save**.

Key generator: The PuTTY key generator saves the private key with the PPK extension.

8. Close the PuTTY key generator.

Uploading the SSH public key to the IBM Storwize V5000 Gen2

After you create your SSH key pair, upload your SSH public key onto the IBM V5000 Gen2 system. Complete the following steps:

1. On the System Overview, click the **Access** icon and select **Users** in the GUI menu (see Figure A-7).

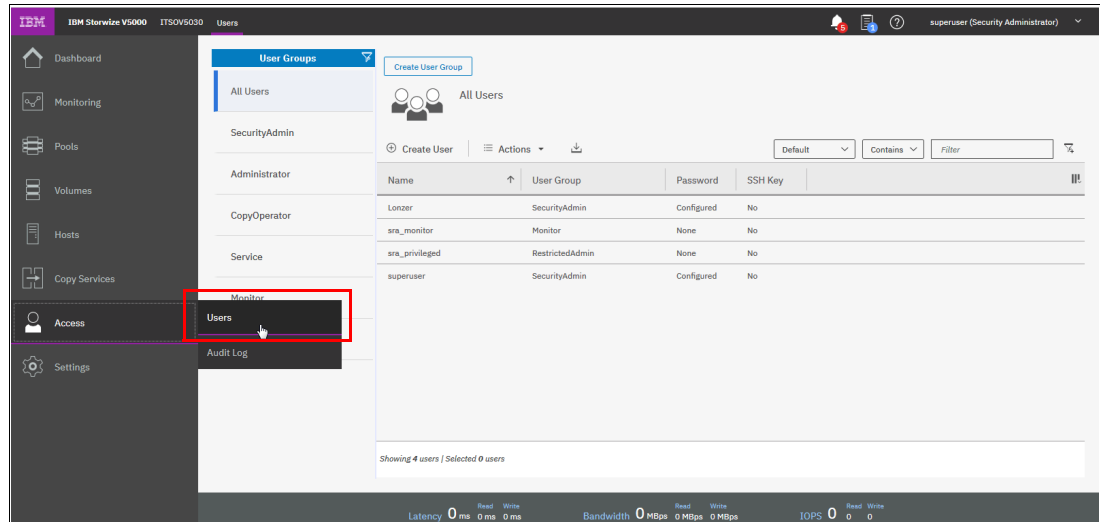


Figure A-7 Click Users on the Access menu

2. Under User Groups, select **All Users**. Right-click the user name for which you want to upload the key and click **Properties** (see Figure A-8).

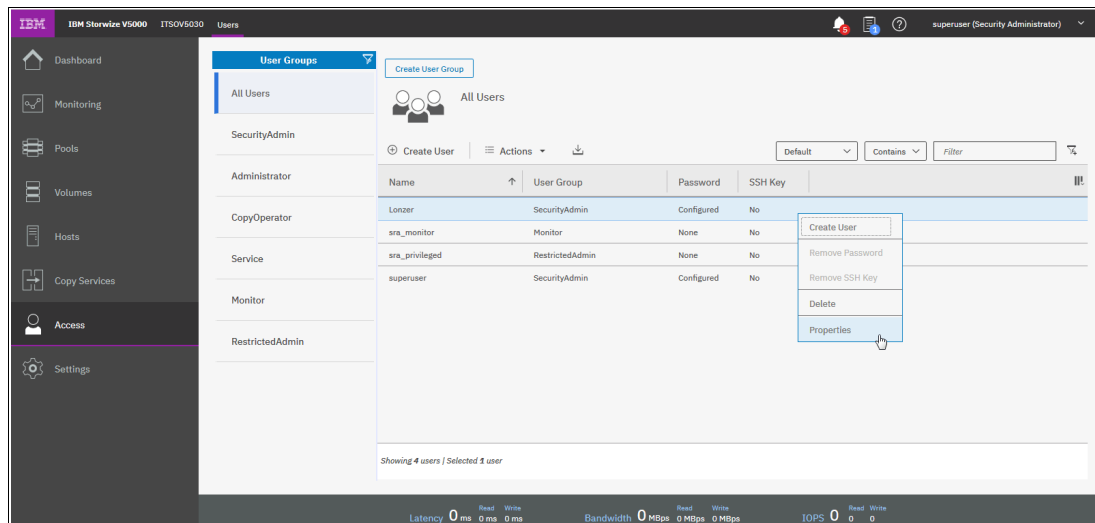


Figure A-8 User properties

3. To upload the public key, click **Browse**, and select the folder where you stored the public SSH key (see Figure A-9).

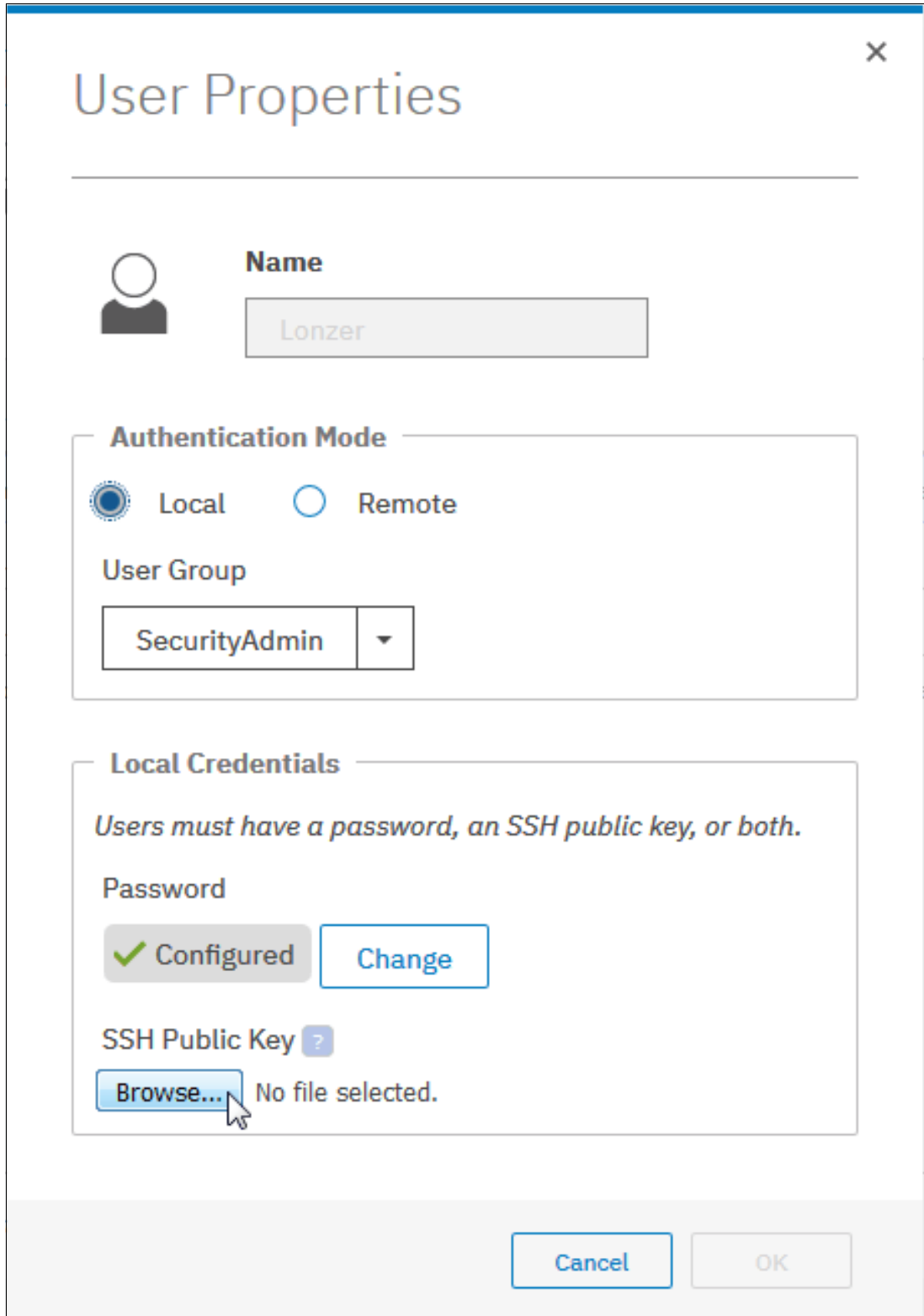


Figure A-9 Select public key

4. Select your public key, and click **Open** (see Figure A-10).

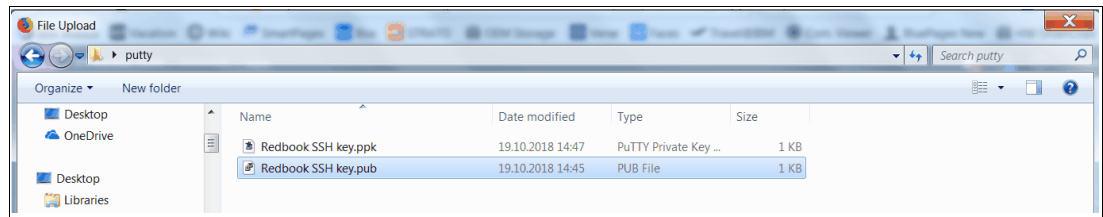


Figure A-10 Selection of the public SSH key

5. Click **OK**, as shown in Figure A-11. The key is uploaded.

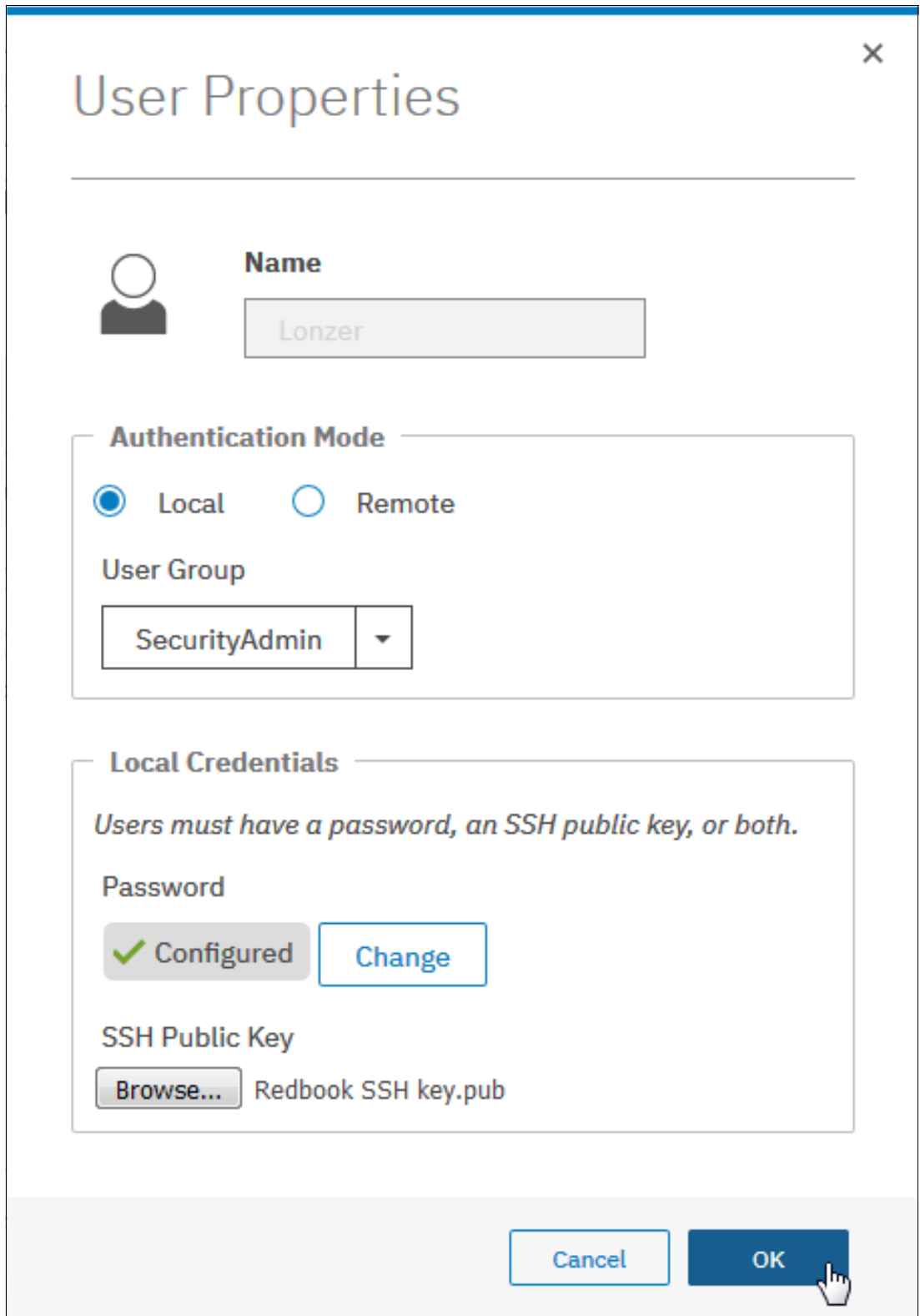


Figure A-11 Select the public key

6. Check the GUI to see whether the SSH key was successfully imported (see Figure A-12).

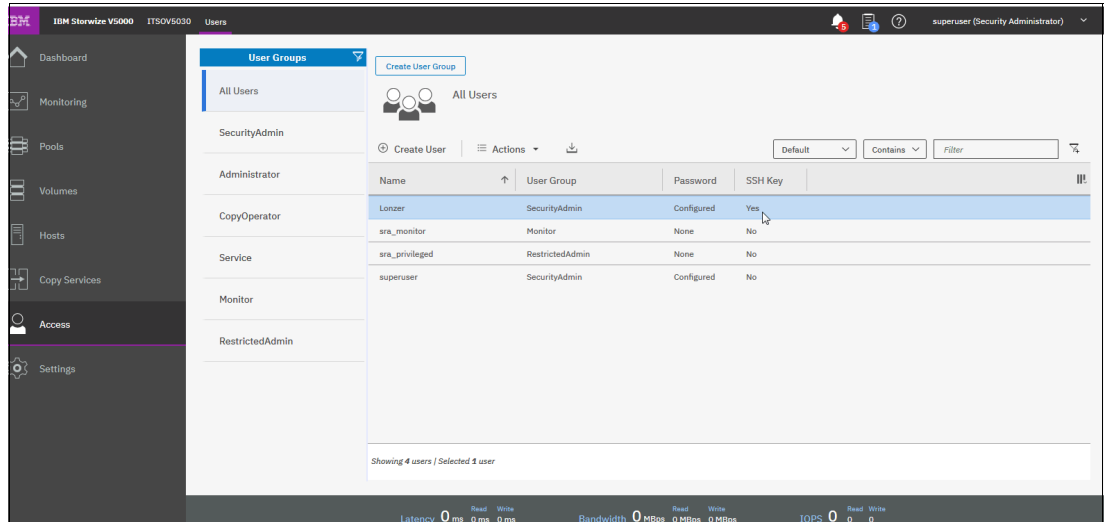


Figure A-12 SSH key was successfully imported

Configuring the SSH client

Before you can use the CLI, complete the following steps to configure the SSH client:

1. Start PuTTY. The PuTTY Configuration window opens (see Figure A-13).

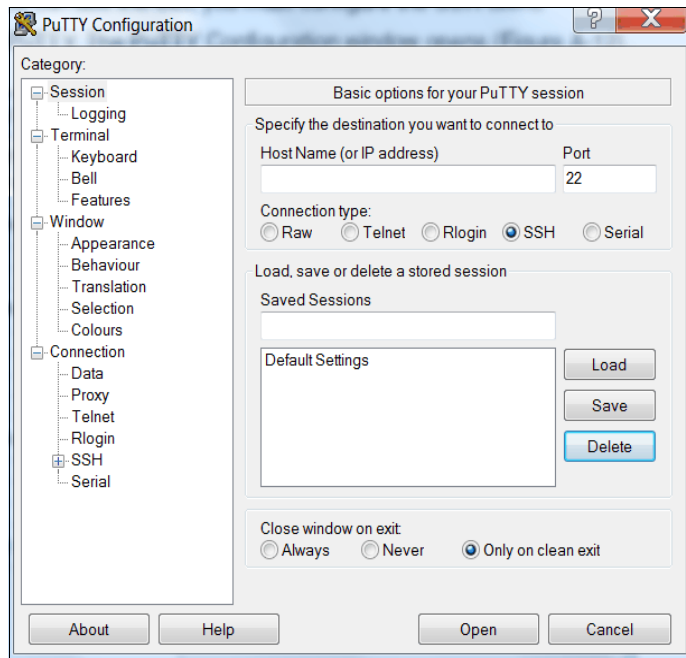


Figure A-13 PuTTY

- In the right side pane, select **SSH** as the connection type. Under the “Close window on exit” section, select **Only on clean exit**, which ensures that if any connection errors occur, they are displayed on the user’s window (see Figure A-14).

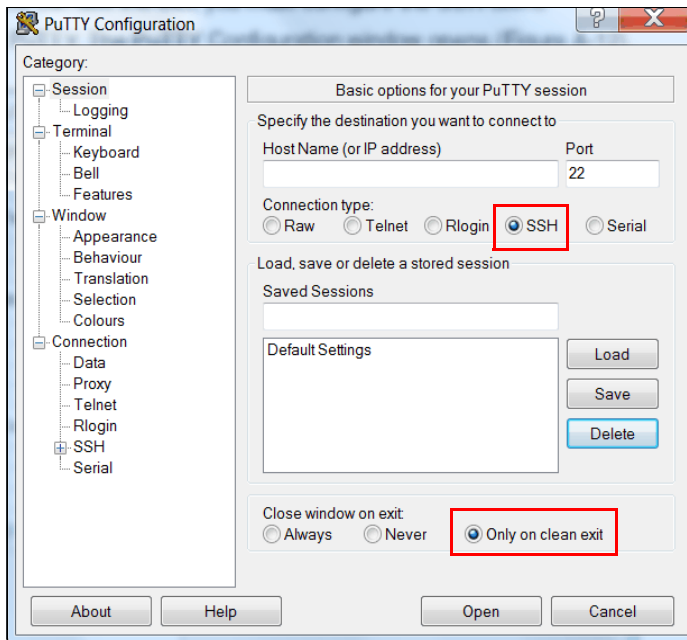


Figure A-14 Select SSH + Only on clean exit

- In the Category pane, on the left side of the PuTTY Configuration window (see Figure A-15), click **Connection** → **SSH** to open the PuTTY Configuration window Options controlling SSH connections view.

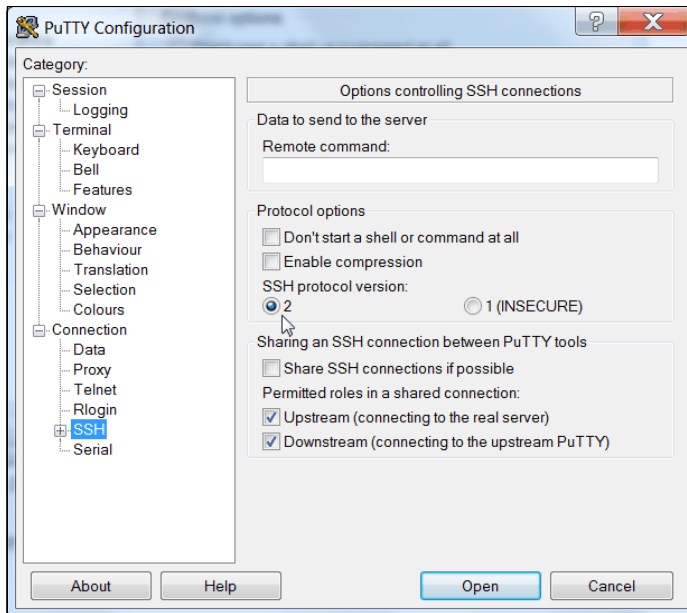


Figure A-15 SSH protocol version 2

Under Preferred SSH protocol version, select **2**.

4. In the Category pane on the left, click **Connection** → **SSH** → **Auth**, as shown in Figure A-16. More options are displayed for controlling SSH authentication.

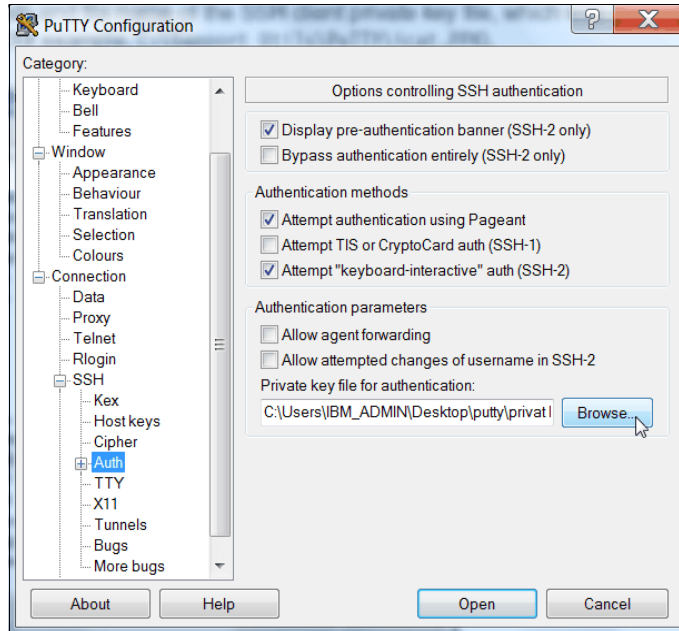


Figure A-16 SSH authentication

In the Private key file for authentication field, browse to or enter the fully qualified directory path and file name of the SSH client private key file, which was created previously (for example, C:\Support Utils\putty\privat.PPK).

5. In the Category pane, click **Session** to return to the PuTTY Configuration window Basic options for your PuTTY session view (see Figure A-17).

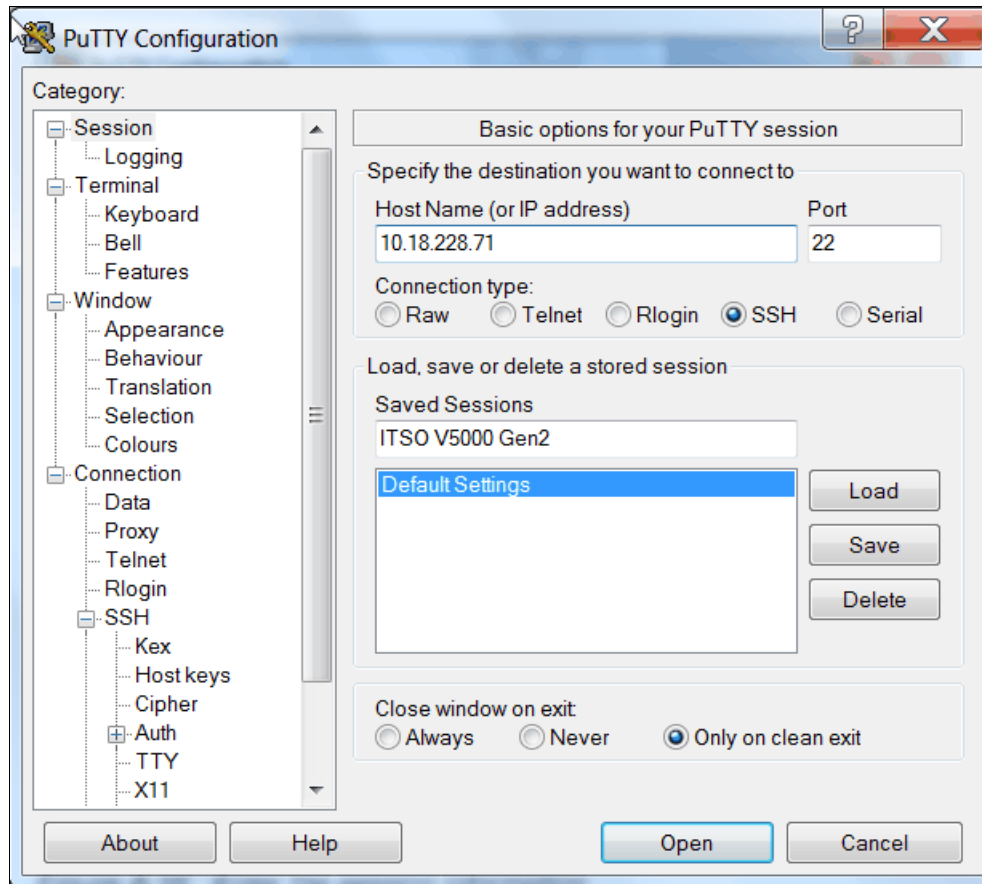


Figure A-17 Enter the session information

6. Complete the following fields in the right side pane:
- Host Name (or IP address): Specify the host name or system IP address of the IBM Storwize V5000 Gen2 clustered system.
 - Saved Sessions: Enter a session name.

7. Click **Save** to save the new session (see Figure A-18).

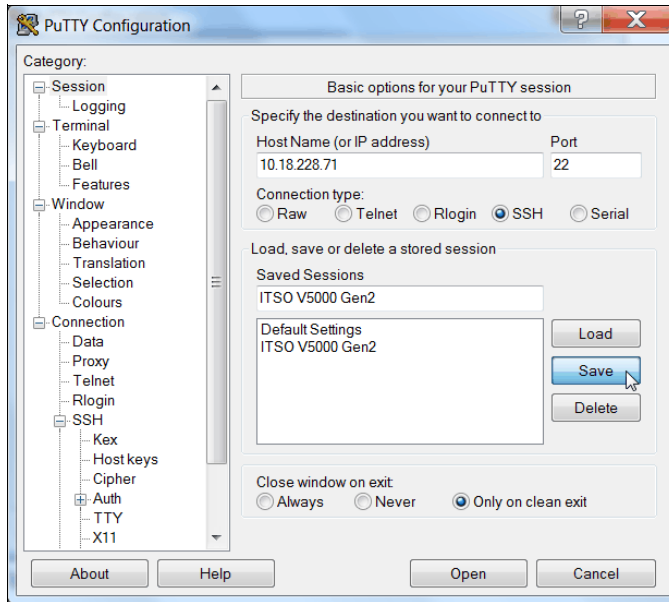


Figure A-18 Save the new session

8. Figure A-19 shows the saved PUTTY session. Select the new session and click **Open**.

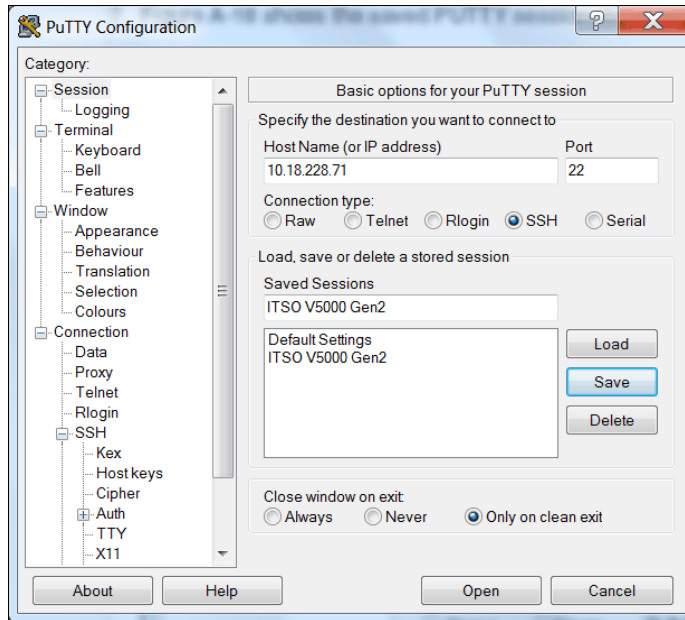


Figure A-19 Saved PUTTY session

9. If a PuTTY Security Alert window opens, confirm it by clicking **Yes** (see Figure A-20).

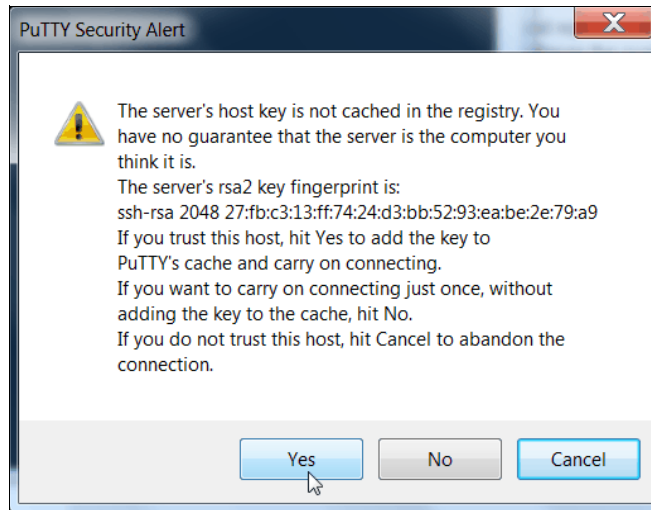


Figure A-20 Confirm the security alert

10. PuTTY now connects to the system and prompts you for a user name to log in as. Enter Superuser as the user name (see Example A-1) and click **Enter**.

Example A-1 Enter user name

```
login as: Lonzer
Authenticating with public key "rsa-key-20181019"
IBM_Storwize:ITS0V5030:Lonzer>
```

The CLI is now configured for the IBM Storwize V5000 Gen2 administration.

SAN Boot

The IBM Storwize V5000 Gen2 supports SAN Boot for Microsoft Windows, VMware, and many other operating systems. Because SAN Boot support can change, regularly check the [IBM Storwize V5000 Gen2 interoperability matrix](#).

The IBM Knowledge Center for Storwize V5000 Gen2 contains information about SAN Boot in combination with various operating systems. For more information, see [IBM Knowledge Center](#).

More information about SAN Boot is also available in [IBM Multipath Subsystem Device Driver User's Guide](#).

Enabling SAN Boot for Windows

Complete the following steps if you want to install a Windows host b using SAN Boot:

1. Configure the IBM Storwize V5000 Gen2 system so that only the boot volume is mapped to the host.

2. Configure the Fibre Channel storage area network (SAN) so that the host sees only one IBM Storwize V5000 Gen2 system node port. Multiple paths during installation are not supported.
3. Configure and enable the host bus adapter (HBA) BIOS.
4. Install the operating system by using the normal procedure, selecting the volume as the partition on which to install.

HBAs: You might need to load another HBA device driver during installation, depending on your Windows version and the HBA type.

5. Install Subsystem Device Driver Device Specific Module (SDDDSM) after the installation completes.
6. Modify your SAN zoning to allow multiple paths.
7. Check your host to see whether all paths are available.
8. Set redundant boot devices in the HBA BIOS to enable the host to boot when its original path fails.

Enabling SAN Boot for VMware

Complete the following steps if you want to install a VMware ESX host by using SAN Boot:

1. Configure the IBM Storwize V5000 Gen2 system so that only the boot volume is mapped to the host.
2. Configure the Fibre Channel SAN so that the host sees only one IBM Storwize V5000 Gen2 system node port. Multiple paths during installation are not supported.
3. Configure and enable the HBA BIOS.
4. Install the operating system by using the normal procedure, selecting the volume as the partition on which to install.

HBAs: You might need to load an extra HBA device driver during installation, depending on your ESX level and the HBA type.

5. Modify your SAN zoning to allow multiple paths.
6. Check your host to see whether all paths are available and modify the multipath policy, if required.

Windows SAN Boot migration

If your host runs the Windows Server 2008, Windows 2012, or Windows 2016 operating system and uses SAN Boot images that are controlled by storage controllers, you can migrate these images to image-mode volumes that are controlled by the IBM Storwize V5000 Gen2 system.

SAN Boot procedures: For more information about SAN Boot procedures for other operating systems, see [IBM Knowledge Center for the Storwize V5000](#).

Complete the following steps to migrate your SAN Boot images:

1. Shut down the host.

2. Complete the following configuration changes on the storage controller:
 - a. Remove all the image-to-host mappings from the storage controller.
 - b. Map the existing SAN boot image and any other disks to the system.
3. Zone one port of each host bus adapter (HBA) to one of the system ports that is associated with the I/O group for the target image-mode volume.
4. Complete the following configuration changes on the system:
 - a. Create an image-mode volume for the managed disk (MDisk) that contains the SAN boot image.
Use the MDisk unique identifier to specify the correct MDisk.
 - b. Create a host object and assign it to the HBA port that you zoned to the system port in step 3.
 - c. Map the image mode volume to the host.
For example, you might map the boot disk to the host with SCSI LUN ID 0.
 - d. Map the swap disk to the host, if required.
For example, you might map the swap disk to the host with SCSI LUN ID 1.
5. Change the boot address of the host by completing the following steps:
 - a. Restart the host and open the BIOS utility of the host during the booting process.
 - b. Set the BIOS settings on the host to find the boot image at the worldwide port name (WWPN) of the node that is zoned to the HBA port.
6. Boot the host in single-path mode.
7. Uninstall any multipathing driver that is not supported for system hosts that run the applicable Windows Server operating system.
8. Install a supported multipathing driver.
9. Restart the host in single-path mode to ensure that the supported multipath driver was properly installed.
10. Zone each HBA port to one port on each system node.
11. Add HBA ports to the host object that you created in step 4.b.
12. Configure the HBA settings on the host by using the following steps:
 - a. Restart the host and open the host's BIOS utility during the booting process.
 - b. Ensure that all HBA ports are boot-enabled and can see both nodes in the I/O group that contains the SAN boot image. Configure the HBA ports for redundant paths.
 - c. Exit the BIOS utility and finish booting the host.
13. Map any additional volume to the host as required.



B

Terminology

This appendix summarizes the IBM Spectrum Virtualize and IBM Storwize V5000 terms that are commonly used in this book.

To see the complete set of terms that relate to the IBM Storwize V5000, see [IBM Knowledge Center](#).

For general Terminology descriptions, see [this web page](#).

Commonly encountered terms

This book uses the common IBM Spectrum Virtualize and IBM Storwize V5000 terminology that is listed in this section.

Access mode

Access mode is one of the modes in which a logical unit (LU) in a disk controller system can operate. The three access modes are image mode, managed space mode, and unconfigured mode. See also “Image mode” on page 824, “Managed mode” on page 827, “Unique identifier” on page 834.

Activation key

See “License key” on page 826.

Array

Array is an ordered collection, or group, of physical devices (disk drive modules) that are used to define logical volumes or devices. An array is a group of drives designated to be managed with a Redundant Array of Independent Disks (RAID).

Asymmetric virtualization

Asymmetric virtualization is a virtualization technique in which the virtualization engine is outside the data path and performs a metadata-style service. The metadata server contains all the mapping and locking tables, and the storage devices contain only data. See also “Symmetric virtualization” on page 833.

Asynchronous replication

Asynchronous replication is a type of replication in which control is given back to the application as soon as the write operation is made to the source volume. Later, the write operation is made to the target volume. See also “Synchronous replication” on page 833.

Automatic data placement mode

Automatic data placement mode is an Easy Tier operating mode in which the host activity on all the volume extents in a pool are “measured,” a migration plan is created, and then automatic extent migration is performed.

Auxiliary volume

The auxiliary volume that contains a mirror of the data on the master volume. See also “Management node” on page 827, and “Relationship” on page 831.

Available (usable) capacity

See “Capacity” on page 817.

Back end

See “Front end and back end” on page 823.

Caching I/O Group

The caching I/O Group is the I/O Group in the system that performs the cache function for a volume.

Call home

Call home is a communication link that is established between a product and a service provider. The product can use this link to call IBM or another service provider when the product requires service. With access to the machine, service personnel can perform service tasks, such as viewing error and problem logs or initiating trace and dump retrievals.

Canister

A canister is a single processing unit within a storage system.

Capacity

The following definitions are applied to capacity by IBM:

- ▶ **Raw capacity**
The reported capacity of the drives in the system before formatting or RAID.
- ▶ **Usable capacity**
The amount of capacity after formatting and RAID available for storing data on a system, pool, array or MDisk. Usable capacity is the total of used and available capacity. For example, 50 TiB used, 50 TiB available is a usable capacity of 100 TiB.
- ▶ **Used capacity**
The amount of usable capacity taken up by data in a system, pool, array or MDisk after data reduction techniques have been applied.
- ▶ **Available capacity**
The amount of usable capacity that is not yet used in a system, pool, array or MDisk.
- ▶ **Effective capacity**
The amount of provisioned capacity that can be created in the system or pool without running out of usable capacity given the current data reduction savings being achieved. This capacity equals the physical capacity divided by the data reduction savings percentage.
- ▶ **Provisioned capacity**
Total capacity of all volumes in a pool or system.
- ▶ **Written capacity**
The amount of usable capacity that is used to store written data in a pool or system before data reduction is applied.
- ▶ **Overhead capacity**
The amount of usable capacity occupied by metadata in a pool or system and other data used for system operation.
- ▶ **Total capacity savings**
The total amount of usable capacity saved in a pool, system, or volume through thin-provisioning and data reduction techniques. This capacity saved is the difference between the used usable capacity and the provisioned capacity.
- ▶ **Data reduction**
The techniques used to reduce the size of data including deduplication and compression.
- ▶ **Data reduction savings**
The total amount of usable capacity saved in a pool, system or volume through the application of a compression or deduplication algorithm on the written data. This capacity saved is the difference between the written capacity and the used capacity.

- ▶ Thin provisioning savings

The total amount of usable capacity saved in a pool, system, or volume by using usable capacity when needed as a result of write operations. The capacity saved is the difference between the provisioned capacity minus the written capacity.
- ▶ Over provisioned

A storage system or pool where there is more provisioned capacity than there is usable capacity.
- ▶ Over provisioned ratio

The ratio of provisioned capacity to usable capacity in the pool or system.
- ▶ Provisioning limit, maximum provisioned capacity, over provisioning limit

In some storage systems, restrictions in the storage hardware or configured by the user that define a limit the maximum provisioned capacity allowed in a pool or system.

Capacity licensing

Capacity licensing is a licensing model that licenses features with a price-per-terabyte model. Licensed features are FlashCopy, Metro Mirror, Global Mirror, and virtualization. See also “FlashCopy” on page 823, “Metro Mirror” on page 827, and “Virtualization” on page 834.

Capacity recycling

Capacity recycling will take the amount of provisioned capacity that can be recovered without causing stress or performance degradation. This capacity identifies the amount of resources that can be reclaimed and provisioned to other objects in an environment.

Chain

A set of enclosures that are attached to provide redundant access to the drives inside the enclosures. Each control enclosure can have one or more chains.

Challenge Handshake Authentication Protocol

Challenge Handshake Authentication Protocol (CHAP) is an authentication protocol that protects against eavesdropping by encrypting the user name and password.

Change volume

A change volume is used in Global Mirror that holds earlier consistent revisions of data when changes are made.

Channel extender

A channel extender is a device that is used for long-distance communication that connects other storage area network (SAN) fabric components. Generally, channel extenders can involve protocol conversion to asynchronous transfer mode (ATM), Internet Protocol (IP), or another long-distance communication protocol.

Child pool

Administrators can use child pools to control capacity allocation for volumes that are used for specific purposes. Rather than being created directly from managed disks (MDisks), child pools are created from existing capacity that is allocated to a parent pool. As with parent pools, volumes can be created that specifically use the capacity that is allocated to the child pool. Child pools are similar to parent pools with similar properties. Child pools can be used for volume copy operation. Also, see “Parent pool” on page 828.

CLI

Command Line Interface is a computer interface in which the input and output are text based.

Cloud Container

Cloud Container is a virtual object that includes all of the elements, components or data that are common to a specific application or data.

Cloud Service Provider

Cloud Service Provider (CSP) is the company or organization that provides off- and on-premises cloud services such as storage, server, network, and so on. IBM Spectrum Virtualize has built in software capabilities to interact with Cloud Providers such as IBM SoftLayer®, Amazon S3 and deployments of OpenStack Swift.

Cloud Tenant

Cloud Tenant is a group or an instance that provides common access with the specific privileges to a object, software or data source.

Clustered system (Storwize V5000)

A clustered system, formerly known as a cluster, is a group of up to four IBM Storwize V5000 canisters (two in each system) that presents a single configuration, management, and service interface to the user.

Cold extent

A cold extent is an extent of a volume that does not get any performance benefit if it is moved from a hard disk drive (HDD) to a Flash disk. A cold extent also refers to an extent that needs to be migrated onto an HDD if it is on a Flash disk drive.

Compression

Compression is a function that removes repetitive characters, spaces, strings of characters, or binary data from the data that is being processed and replaces characters with control characters. Compression reduces the amount of storage space that is required for data. See also “RACE engine” on page 829.

Compression accelerator

A compression accelerator is hardware onto which the work of compression is off-loaded from the microprocessor.

Configuration node

While the cluster is operational, a single node in the cluster is appointed to provide configuration and service functions over the network interface. This node is termed the configuration node. This configuration node manages the data that describes the clustered-system configuration and provides a focal point for configuration commands. If the configuration node fails, another node in the cluster transparently assumes that role.

Consistency Group

A Consistency Group is a group of copy relationships between virtual volumes or data sets that are maintained with the same time reference so that all copies are consistent in time. A Consistency Group can be managed as a single entity.

Container

A container is a software object that holds or organizes other software objects or entities.

Contingency capacity

For thin-provisioned volumes that are configured to automatically expand, the unused real capacity that is maintained. For thin-provisioned volumes that are not configured to automatically expand, the difference between the used capacity and the new real capacity.

Copied state

Copied is a FlashCopy state that indicates that a copy was triggered after the copy relationship was created. The Copied state indicates that the copy process is complete and the target disk has no further dependency on the source disk. The time of the last trigger event is normally displayed with this status.

Counterpart SAN

A counterpart SAN is a non-redundant portion of a redundant SAN. A counterpart SAN provides all of the connectivity of the redundant SAN, but without 100% redundancy. IBM Storwize V5000 canisters are typically connected to a “redundant SAN” that is made up of two counterpart SANs. A counterpart SAN is often called a SAN fabric.

Cross-volume consistency

A consistency group property that ensures consistency between volumes when an application issues dependent write operations that span multiple volumes.

Data consistency

Data consistency is a characteristic of the data at the target site where the dependent write order is maintained to ensure the recoverability of applications.

Data deduplication

Data deduplication is a method of reducing storage needs by eliminating redundant data. Only one instance of the data is retained on storage media. Other instances of the same data are replaced with a pointer to the retained instance.

Data encryption key

The data encryption key is used to encrypt data and it is created automatically when an encrypted object, such as an array, a pool, or a child pool, is created. It is stored in secure memory and it cannot be viewed or changed. The data encryption key is encrypted using the master access key.

Data migration

Data migration is the movement of data from one physical location to another physical location without the disruption of application I/O operations.

Data reduction

Data reduction is a set of techniques that can be used to reduce the amount of physical storage that is required to store data. An example of data reduction includes data deduplication and compression. See also “Data reduction pool”. See also “Capacity” on page 817.

Data reduction pool

Data Reduction pools are specific types of pools where more control over volumes capacity is given to specific hosts (for example VMware VAAI/VASA/VVOL, Microsoft ODX). These hosts are able to return unused space for reuse. With standard pools, the system is not aware of any unused space on host-allocated volumes. See also “Data reduction”.

Data reduction savings

See “Capacity” on page 817.

Dependent write operation

A write operation that must be applied in the correct order to maintain cross-volume consistency.

Directed Maintenance Procedure

The fix procedures, which are also known as Directed Maintenance Procedures (DMPs), ensure that you fix any outstanding errors in the error log. To fix errors, from the Monitoring panel, click **Events**. The Next Recommended Action is displayed at the top of the Events window. Select **Run This Fix Procedure** and follow the instructions.

Discovery

The automatic detection of a network topology change, for example, new and deleted nodes or links.

Disk tier

MDisks (logical unit numbers (LUNs)) that are presented to the IBM Storwize V5000 likely have different performance attributes because of the type of disk or RAID array on which they are installed. MDisks can be on 15,000 revolutions per minute (RPM) Fibre Channel (FC) or serial-attached SCSI (SAS) disk, Nearline SAS, or Serial Advanced Technology Attachment (SATA), or even Flash Disks. Therefore, a storage tier attribute is assigned to each MDisk, and the default is *generic_hdd*.

Distributed RAID or DRAID

An alternative RAID scheme where the number of drives that are used to store the array can be greater than the equivalent, typical RAID scheme. The same data stripes are distributed across a greater number of drives, which increases the opportunity for parallel I/O and hence improves overall array performance. See also “Rebuild area” on page 831.

Easy Tier

Easy Tier is a volume performance function within the IBM Storwize family that provides automatic data placement of a volume’s extents in a multitiered storage pool. The pool normally contains a mix of Flash Disks and HDDs. Easy Tier measures host I/O activity on the volume’s extents and migrates hot extents onto the Flash Disks to ensure the maximum performance.

Effective capacity

See “Capacity” on page 817.

Encryption key

The encryption key, also known as master access key, is created and stored on USB flash drives or on a key server when encryption is enabled. The master access key is used to decrypt the data encryption key.

Encryption key server

An internal or external system that receives and then serves existing encryption keys or certificates to a storage system.

Encryption of data at rest

Encryption of data at rest is the inactive encryption data that is stored physically on the storage system.

Evaluation mode

The evaluation mode is an Easy Tier operating mode in which the host activity on all the volume extents in a pool are “measured” only. No automatic extent migration is performed.

Event (error)

An event is an occurrence of significance to a task or system. Events can include the completion or failure of an operation, user action, or a change in the state of a process.

Event code

An event code is a value that is used to identify an event condition to a user. This value might map to one or more event IDs or to values that are presented on the service panel. This value is used to report error conditions to IBM and to provide an entry point into the service guide.

Event ID

An event ID is a value that is used to identify a unique error condition that was detected by the Storwize V5000. An event ID is used internally in the cluster to identify the error.

Excluded condition

The excluded condition is a status condition. It describes an MDisk that the IBM Storwize V5000 has decided is no longer sufficiently reliable to be managed by the cluster. The user must issue a command to include the MDisk in the cluster-managed storage.

Extent

An extent is a fixed-size unit of data that is used to manage the mapping of data between MDisks and volumes. The size of the extent can range 16 MB - 8 GB in size.

External storage

External storage refers to managed disks (MDisks) that are SCSI logical units that are presented by storage systems that are attached to and managed by the clustered system.

Expansion Canister

A hardware unit that includes the serial-attached SCSI (SAS) hardware that enables the node hardware to use the drives of the expansion enclosure.

Expansion enclosure

A hardware unit that includes enclosure chassis, expansion canisters, drives, and function that allows extra drives to be connected.

Failback

Failback is the restoration of an appliance to its initial configuration after the detection and repair of a failed network or component.

Failover

Failover is an automatic operation that switches to a redundant or standby system or node in a software, hardware, or network interruption. See also Failback.

Feature activation code

An alphanumeric code that activates a licensed function on a product.

Fibre Channel port logins

Fibre Channel (FC) port logins refer to the number of hosts that can see any one Storwize V5000 port. The IBM Storwize V5000 has a maximum limit per node port of FC logins that are allowed.

Field-replaceable unit

Field-replaceable units (FRUs) are individual parts that are replaced entirely when any one of the unit's components fails. They are held as spares by the IBM service organization.

FlashCopy

FlashCopy refers to a point-in-time copy where a virtual copy of a volume is created. The target volume maintains the contents of the volume at the point in time when the copy was established. Any subsequent write operations to the source volume are not reflected on the target volume.

FlashCopy mapping

A FlashCopy mapping is a continuous space on a direct-access storage volume, which is occupied by or reserved for a particular data set, data space, or file.

FlashCopy relationship

See FlashCopy mapping.

FlashCopy service

FlashCopy service is a copy service that duplicates the contents of a source volume on a target volume. In the process, the original contents of the target volume are lost. See also "Point-in-time copy" on page 829.

Flash drive

A data storage device that uses solid-state memory to store persistent data.

Flash module

A modular hardware unit containing flash memory, one or more flash controllers, and associated electronics.

Front end and back end

The IBM Storwize V5000 takes MDisks to create pools of capacity from which volumes are created and presented to application servers (hosts). The MDisks are in the controllers at the back end of Storwize V5000 and in the Storwize V5000 to the back-end controller zones. The volumes that are presented to the hosts are in the front end of IBM Storwize V5000.

Global Mirror

Global Mirror (GM) is a method of asynchronous replication that maintains data consistency across multiple volumes within or across multiple systems. Global Mirror is generally used where distances between the source site and target site cause increased latency beyond what the application can accept.

Global Mirror with change volumes

Change volumes are used to record changes to the primary and secondary volumes of a remote copy relationship. A FlashCopy mapping exists between a primary and its change volume and a secondary and its change volume.

Grain

A grain is the unit of data that is represented by a single bit in a FlashCopy bitmap (64 KiB or 256 KiB) in the IBM Storwize V5000. A grain is also the unit to extend the real size of a thin-provisioned volume (32 KiB, 64 KiB, 128 KiB, or 256 KiB).

Hop

Hop is one segment of a transmission path between adjacent nodes in a routed network.

Host bus adapter

A host bus adapter (HBA) is an interface card that connects a server to the SAN environment through its internal bus system, for example, PCI Express. Typically it is referred to the Fibre Channel adapters.

Host ID

A host ID is a numeric identifier that is assigned to a group of host FC ports or Internet Small Computer System Interface (iSCSI) host names for LUN mapping. For each host ID, SCSI IDs are mapped to volumes separately. The intent is to have a one-to-one relationship between hosts and host IDs, although this relationship cannot be policed.

Host mapping

Host mapping refers to the process of controlling which hosts have access to specific volumes within a cluster (host mapping is equivalent to LUN masking).

Hot extent

A hot extent is a frequently accessed volume extent that gets a performance benefit if it is moved from an HDD onto a Flash Disk.

Hot Spare node (SAN Volume Controller only)

Hot Spare Node is an online SAN Volume Controller node defined in a cluster but not in any I/O group. In case of a failure of any of online nodes in any I/O group of cluster, it is automatically swapped by this Spare node. After the recovery of an original node finishes, the Spare node gets back to the standby spare status. This feature is not available for IBM Storwize V5000.

HyperSwap

Pertaining to a function that provides continuous, transparent availability against storage errors and site failures, and is based on synchronous replication.

Image mode

Image mode is an access mode that establishes a one-to-one mapping of extents in the storage pool (existing LUN or (image mode) MDisk) with the extents in the volume. See also "Access mode" on page 816, "Managed mode" on page 827, "Unique identifier" on page 834.

Image volume

An image volume is a volume in which a direct block-for-block translation exists from the managed disk (MDisk) to the volume.

I/O Group

Each pair of IBM Storwize V5000 canisters is known as an input/output (I/O) Group. An I/O Group has a set of volumes that are associated with it that are presented to host systems. Each Storwize V5000 canister is associated with exactly one I/O Group. The canister in an I/O Group provide a failover and failback function for each other. An IBM Storwize V5000 cluster consist of two I/O groups.

Internal storage

Internal storage refers to an array of managed disks (MDisks) and drives that are held in IBM Storwize V5000 enclosures.

Internet Small Computer System Interface qualified name

Internet Small Computer System Interface (iSCSI) qualified name (IQN) refers to special names that identify both iSCSI initiators and targets. IQN is one of the three name formats that is provided by iSCSI. The IQN format is `iqn.<yyyymm>.<reversed domain name>`. For example, the default for a Storwize V5000 canister can be in the following format:

```
iqn.1986-03.com.ibm:2076.<clustername>.<nodename>
```

Internet storage name service

The internet Storage Name Service (iSNS) protocol that is used by a host system to manage iSCSI targets and the automated iSCSI discovery, management, and configuration of iSCSI and FC devices. It was defined in Request for Comments (RFC) 4171.

Inter-switch link hop

An inter-switch link (ISL) is a connection between two switches and counted as one ISL hop. The number of hops is always counted on the shortest route between two N-ports (device connections). In an IBM Storwize V5000 environment, the number of ISL hops is counted on the shortest route between the pair of canister that are farthest apart. The Storwize V5000 supports a maximum of three ISL hops.

Input/output group

A collection of volumes and canister relationships that present a common interface to host systems. Each pair of canister is known as an input/output (I/O) group.

I/O throttling rate

The maximum rate at which an I/O transaction is accepted for a volume.

iSCSI initiator

An initiator functions as an iSCSI client. An initiator typically serves the same purpose to a computer as a SCSI bus adapter would, except that, instead of physically cabling SCSI devices (such as hard drives and tape changers), an iSCSI initiator sends SCSI commands over an IP network.

iSCSI session

An iSCSI Initiator and an iSCSI Target talk with each other. This conversation is called an iSCSI Session.

iSCSI target

An iSCSI target is a storage resource on an iSCSI server.

Just a bunch of disks (JBOD)

Hard disks that have not yet been configured according to the RAID system to increase fault tolerance and improve data access performance.

Latency

The time interval between the initiation of a send operation by a source task and the completion of the matching receive operation by the target task. More generally, latency is the time between a task initiating data transfer and the time that transfer is recognized as complete at the data destination.

Least recently used

Least recently used (LRU) pertains to an algorithm used to identify and make available the cache space that contains the data that was least recently used.

Licensed capacity

The amount of capacity on a storage system that a user is entitled to configure.

License key

An alphanumeric code that activates a licensed function on a product. Also known as activation key.

License key file

A file that contains one or more licensed keys.

Lightweight Directory Access Protocol

Lightweight Directory Access Protocol (LDAP) is an open protocol that uses TCP/IP to provide access to directories that support an X.500 model and that does not incur the resource requirements of the more complex X.500 Directory Access Protocol (DAP). For example, LDAP can be used to locate people, organizations, and other resources in an Internet or intranet directory.

Local and remote fabric interconnect

The local fabric interconnect and the remote fabric interconnect are the SAN components that are used to connect the local and remote fabrics. Depending on the distance between the two fabrics, they can be single-mode optical fibers that are driven by long wave (LW) gigabit interface converters (GBICs) or small form-factor pluggables (SFPs), or more sophisticated components, such as channel extenders or special SFP modules that are used to extend the distance between SAN components.

Local fabric

The local fabric is composed of SAN components (switches, cables, and so on) that connect the components (nodes, hosts, and switches) of the local cluster together.

Logical unit and logical unit number

The logical unit (LU) is defined by the SCSI standards as a logical unit number (LUN). LUN is an abbreviation for an entity that exhibits disk-like behavior, for example, a volume or an MDisk.

LUN masking

A process where a host object can detect more LUNs than it is intended to use, and the device-driver software masks the LUNs that are not to be used by this host.

Machine signature

A string of characters that identifies a system. A machine signature might be required to obtain a license key.

Managed disk

A managed disk (MDisk) is a SCSI disk that is presented by a RAID controller and managed by IBM Storwize V5000. The MDisk is not visible to host systems on the SAN.

Managed disk group (storage pool)

See “Storage pool (managed disk group)” on page 833.

Managed mode

An access mode that enables virtualization functions to be performed. See also “Access mode” on page 816, “Image mode” on page 824, “Unique identifier” on page 834.

Management node

Management node is a node that is used for configuring, administering, and monitoring a system.

Maximum replication delay

Maximum replication delay is the number of seconds that Metro Mirror or Global Mirror replication can delay a write operation to a volume.

Master volume

In most cases, the volume that contains a production copy of the data and that an application accesses. See also “Auxiliary volume” on page 816, and “Relationship” on page 831.

Metro Global Mirror

Metro Mirror Global is a cascaded solution where Metro Mirror synchronously copies data to the target site. This Metro Mirror target is the source volume for Global Mirror that asynchronously copies data to a third site. This solution has the potential to provide disaster recovery with no data loss at Global Mirror distances when the intermediate site does not participate in the disaster that occurs at the production site.

Metro Mirror

Metro Mirror (MM) is a method of synchronous replication that maintains data consistency across multiple volumes within the system. Metro Mirror is generally used when the write latency that is caused by the distance between the source site and target site is acceptable to application performance.

Mirrored volume

A mirrored volume is a single virtual volume that has two physical volume copies. The primary physical copy is known within the IBM Storwize V5000 as copy 0 and the secondary copy is known within the IBM Storwize V5000 as copy 1.

Nearline SAS drive

Nearline is a SAS drive that combines the high capacity data storage technology of a Serial Advanced Technology Attachment (SATA) drive with the benefits of a serial-attached SCSI (SAS) interface for improved connectivity.

Node canister

A node canister is a hardware unit that includes the node hardware, fabric and service interfaces, and serial-attached SCSI (SAS) expansion ports. Node canisters are specifically recognized on IBM Storwize products. In SVC all these components are spread within the whole system chassis, so we usually do not consider node canisters in SVC, but just the node as a whole.

Node rescue

The process by which a node that has no valid software installed on its hard disk drive can copy software from another node connected to the same Fibre Channel fabric.

NPIV

NPIV or N_Port ID Virtualization is a Fibre Channel feature whereby multiple Fibre Channel node port (N_Port) IDs can share a single physical N_Port.

Object Storage

Object storage is a general term that refers to the entity in which an Cloud Object Storage (COS) organize, manage and store with units of storage or just *objects*.

Overhead capacity

See “Capacity” on page 817.

Oversubscription

Oversubscription refers to the ratio of the sum of the traffic on the initiator N-port connections to the traffic on the most heavily loaded ISLs, where more than one connection is used between these switches. Oversubscription assumes a symmetrical network, and a specific workload that is applied equally from all initiators and sent equally to all targets. A symmetrical network means that all the initiators are connected at the same level, and all the controllers are connected at the same level.

Over provisioned

See “Capacity” on page 817.

Over provisioned ratio

See “Capacity” on page 817.

Parent pool

Parent pools receive their capacity from MDisks. All MDisks in a pool are split into extents of the same size. Volumes are created from the extents that are available in the pool. You can add MDisks to a pool at any time to increase the number of extents that are available for new volume copies or to expand volume copies. The system automatically balances volume extents between the MDisks to provide the best performance to the volumes. See also “Child pool” on page 818.

Partnership

In Metro Mirror or Global Mirror operations, the relationship between two clustered systems. In a clustered-system partnership, one system is defined as the local system and the other system as the remote system.

Point-in-time copy

A point-in-time copy is the instantaneous copy that the FlashCopy service makes of the source volume. See also “FlashCopy service” on page 823.

Preparing phase

Before you start the FlashCopy process, you must prepare a FlashCopy mapping. The preparing phase flushes a volume’s data from cache in preparation for the FlashCopy operation.

Primary volume

In a stand-alone Metro Mirror or Global Mirror relationship, the target of write operations issued by the host application.

Private fabric

Configure one SAN per fabric so that it is dedicated for node-to-node communication. This SAN is referred to as a private SAN.

Provisioned capacity

See “Capacity” on page 817.

Provisioning limit, maximum provisioned capacity, over provisioning limit

See “Capacity” on page 817.

Public fabric

Configure one SAN per fabric so that it is dedicated for host attachment, storage system attachment, and remote copy operations. This SAN is referred to as a public SAN. You can configure the public SAN to allow Storwize V5000 node-to-node communication also. You can optionally use the `-localportfcmask` parameter of the `chsystem` command to constrain the node-to-node communication to use only the private SAN.

Quorum disk

The Quorum disk contains a reserved area that is used exclusively for system management. The quorum disk is accessed when it is necessary to determine which half of the clustered system continues to read and write data. Quorum disks can be MDisks or drives.

Quorum index

The quorum index is the pointer that indicates the order that is used to resolve a tie. Nodes attempt to lock the first quorum disk (index 0), followed by the next disk (index 1), and finally the last disk (index 2). The tie is broken by the node that locks them first.

Quota

Quota is the amount of disk space and number of files and directories assigned as upper limits for a specified user, group of users, or file set.

RACE engine

The RACE engine compresses data on volumes in real time with minimal effect on performance. See “Compression” on page 819 or “Real-time Compression”.

Raw capacity

See “Capacity” on page 817.

Real capacity

Real capacity is the amount of storage that is allocated to a volume copy from a storage pool.

Real-time Compression

Real-time Compression is an IBM integrated software function for storage space efficiency. The RACE engine compresses data on volumes in real time with minimal effect on performance. See also “RACE engine”.

Redundant Array of Independent Disks

Redundant Array of Independent Disks (RAID) refers to two or more physical disk drives that are combined in an array in a certain way, which incorporates a RAID level for failure protection or better performance. The most common RAID levels are 0, 1, 5, 6, and 10. Some storage administrators refer to the RAID group as TRAIID - Traditional RAID.

RAID 0

RAID 0 is a data striping technique that is used across an array and no data protection is provided.

RAID 1

RAID 1 is a mirroring technique that is used on a storage array in which two or more identical copies of data are maintained on separate mirrored disks.

RAID 10

RAID 10 is a combination of a RAID 0 stripe that is mirrored (RAID 1). Therefore, two identical copies of striped data exist; no parity exists.

RAID 5

RAID 5 is an array that has a data stripe, which includes a single logical parity drive. The parity check data is distributed across all the disks of the array.

RAID 6

RAID 6 is a RAID level that has two logical parity drives per stripe, which are calculated with different algorithms. Therefore, this level can continue to process read and write requests to all of the array’s virtual disks in the presence of two concurrent disk failures.

Read intensive drives

The Read Intensive (RI) flash drives (SSD drives) that are available on IBM Storwize V5000 Gen2, IBM Storwize V7000 Gen2, and IBM SAN Volume Controller 2145-DH8/24F are one Drive Write Per Day (DWPD) Read Intensive drives.

Real capacity

Real capacity amount of storage that is allocated to a volume copy from a storage pool.

Rebuild area

Reserved capacity that is distributed across all drives in a redundant array of drives. If a drive in the array fails, the lost array data is systematically restored into the reserved capacity, returning redundancy to the array. The duration of the restoration process is minimized because all drive members simultaneously participate in restoring the data. See also “Distributed RAID or DRAID” on page 821.

Reclaimable (or reclaimed) capacity

Reclaimable Data is the capacity that is no longer needed. Reclaimable capacity is created when data is overwritten and the new data is stored in a new location, when data is marked as unneeded by a host using the SCSI unmap command, or when a volume is deleted.

Redundant storage area network

A redundant SAN is a SAN configuration in which there is no single point of failure (SPoF); therefore, data traffic continues no matter what component fails. Connectivity between the devices within the SAN is maintained (although possibly with degraded performance) when an error occurs. A redundant SAN design is normally achieved by splitting the SAN into two independent counterpart SANs (two SAN fabrics), so that if one path of the counterpart SAN is destroyed, the other counterpart SAN path keeps functioning.

Relationship

In Metro Mirror or Global Mirror, the association between a master volume and an auxiliary volume. These volumes also have the attributes of a primary or secondary volume. See also “Auxiliary volume” on page 816, “Management node” on page 827, primary volume, secondary volume.

Reliability, availability, and serviceability

Reliability, availability, and serviceability (RAS) are a combination of design methodologies, system policies, and intrinsic capabilities that, when taken together, balance improved hardware availability with the costs that are required to achieve it.

Reliability is the degree to which the hardware remains free of faults. Availability is the ability of the system to continue operating despite predicted or experienced faults. Serviceability is how efficiently and nondisruptively broken hardware can be fixed.

Remote fabric

The remote fabric is composed of SAN components (switches, cables, and so on) that connect the components (nodes, hosts, and switches) of the remote cluster together. Significant distances can exist between the components in the local cluster and those components in the remote cluster.

Remote Support Server and Client

Remote Support Client is a software toolkit that resides in IBM Storwize V5000 and opens a secured tunnel to the Remote Support Server. Remote Support Server resides in the IBM network and collects key health check and troubleshooting informations required by IBM support personnel.

SAN Volume Controller

The IBM SAN Volume Controller is an appliance that is designed for attachment to various host computer systems. The SAN Volume Controller performs block-level virtualization of disk storage. IBM Spectrum Virtualize is a software engine of SAN Volume Controller (and Storwize family) that performs block-level virtualization of disk storage.

Secondary volume

Pertinent to remote copy, the volume in a relationship that contains a copy of data written by the host application to the primary volume.

Secure Sockets Layer certificate

Secure Sockets Layer (SSL) is the standard security technology for establishing an encrypted link between a web server and a browser. This link ensures that all data passed between the web server and browsers remain private and to be able to create an SSL connection a web server requires an SSL Certificate.

Security Key Lifecycle Manager

Security Key Lifecycle Manager (SKLM) centralizes, simplifies, and automates the encryption key management process to help minimize risk and reduce operational costs of encryption key management.

Serial-attached SCSI

Serial-attached Small Computer System Interface (SAS) is a method that is used in accessing computer peripheral devices that employs a serial (one bit at a time) means of digital data transfer over thin cables. The method is specified in the American National Standard Institute standard called SAS. In the business enterprise, SAS is useful for access to mass storage devices, particularly external hard disk drives.

Service Location Protocol

The Service Location Protocol (SLP) is an Internet service discovery protocol that enables computers and other devices to find services in a local area network (LAN) without prior configuration. It was defined in the request for change (RFC) 2608.

Small Computer System Interface (SCSI)

Small Computer System Interface (SCSI) is an ANSI-standard electronic interface with which personal computers can communicate with peripheral hardware, such as disk drives, tape drives, CD-ROM drives, printers, and scanners, faster and more flexibly than with previous interfaces.

Snapshot

A snapshot is an image backup type that consists of a point-in-time view of a volume.

Solid-state disk

A solid-state disk (SSD) or Flash Disk is a disk that is made from solid-state memory and therefore has no moving parts. Most SSDs use NAND-based flash memory technology. It is defined to the Storwize V5000 as a disk tier generic_ssd.

Space efficient

See “Thin provisioning” on page 834.

Spare

An extra storage component, such as a drive or tape, that is predesignated for use as a replacement for a failed component.

Spare goal

The optimal number of spares that are needed to protect the drives in the array from failures. The system logs a warning event when the number of spares that protect the array drops below this number.

Space-efficient volume

For more information about a space-efficient volume, see “Thin-provisioned volume” on page 833.

Stand-alone relationship

In FlashCopy, Metro Mirror, and Global Mirror, relationships that do not belong to a consistency group and that have a null consistency-group attribute.

Statesave

Binary data collection that is used for a problem determination by IBM service support.

Storage area network or SAN

A storage area network (SAN) is a dedicated storage network that is tailored to a specific environment, which combines servers, systems, storage products, networking products, software, and services.

Storage pool (managed disk group)

A storage pool is a collection of storage capacity, which is made up of managed disks (MDisks), that provides the pool of storage capacity for a specific set of volumes. A storage pool can contain more than one tier of disk, which is known as a multitier storage pool and a prerequisite of Easy Tier automatic data placement.

Striped

Pertaining to a volume that is created from multiple managed disks (MDisks) that are in the storage pool. Extents are allocated on the MDisks in the order specified.

Support Assistant

A function that is used to provide support personnel access to the system to complete troubleshooting and maintenance tasks.

Symmetric virtualization

Symmetric virtualization is a virtualization technique in which the physical storage, in the form of a Redundant Array of Independent Disks (RAID), is split into smaller chunks of storage known as extents. These extents are then concatenated, by using various policies, to make volumes. See also “Asymmetric virtualization” on page 816.

Synchronous replication

Synchronous replication is a type of replication in which the application write operation is made to both the source volume and target volume before control is given back to the application. See also “Asynchronous replication” on page 816.

Thin-provisioned volume

A thin-provisioned volume is a volume that allocates storage when data is written to it.

Thin provisioning

Thin provisioning refers to the ability to define storage, usually a storage pool or volume, with a “logical” capacity size that is larger than the actual physical capacity that is assigned to that pool or volume. Therefore, a thin-provisioned volume is a volume with a virtual capacity that differs from its real capacity.

Thin provisioning savings

See “Capacity” on page 817.

Throttles

Throttling is a mechanism to control the amount of resources that are used when the system is processing I/Os on supported objects. The system supports throttles on hosts, host clusters, volumes, copy offload operations, and storage pools. If a throttle limit is defined, the system either processes the I/O for that object, or delays the processing of the I/O to free resources for more critical I/O operations.

Transparent Cloud Tiering

Transparent Cloud Tiering is a separately installable feature of IBM Spectrum Scale™ that provides a native cloud storage tier.

Total capacity savings

See “Capacity” on page 817.

T10 DIF

T10 DIF is a *Data Integrity Field* (DIF) extension to SCSI to enable end-to-end protection of data from host application to physical media.

Unique identifier

A unique identifier (UID) is an identifier that is assigned to storage-system logical units when they are created. It is used to identify the logical unit regardless of the logical unit number (LUN), the status of the logical unit, or whether alternate paths exist to the same device. Typically, a UID is used only once.

Unmanaged mode

A mode in which I/O operations cannot be performed. See also “Access mode” on page 816, “Image mode” on page 824, “Managed mode” on page 827.

Usable capacity or used (usable) capacity

See “Capacity” on page 817.

Virtualization

In the storage industry, virtualization is a concept in which a pool of storage is created that contains several storage systems. Storage systems from various vendors can be used. The pool can be split into volumes that are visible to the host systems that use them. See also “Capacity licensing” on page 818.

Virtualized storage

Virtualized storage is physical storage that has virtualization techniques applied to it by a virtualization engine.

Virtual local area network

Virtual local area network (VLAN) tagging separates network traffic at the layer 2 level for Ethernet transport. The system supports VLAN configuration on both IPv4 and IPv6 connections.

Virtual Storage Area Network

A virtual Storage Area Network (VSAN) is a logical fabric entity defined within the storage area network (SAN). It can be defined on a single physical SAN switch or across multiple physical switches or directors. In VMware terminology the vSAN is defined as a logical layer of storage capacity built from physical disk drives attached directly into the ESXi hosts. This solution is not considered for the scope of our publication.

Vital product data

Vital product data (VPD or VDP) is information that uniquely defines system, hardware, software, and microcode elements of a processing system.

Volume

A volume is an IBM Storwize V5000 logical device that appears to host systems that are attached to the SAN as a SCSI disk. Each volume is associated with exactly one I/O Group. A volume has a preferred node within the I/O Group.

Volume copy

A volume copy is a physical copy of the data that is stored on a volume. Mirrored volumes have two copies. Non-mirrored volumes have one copy.

Volume protection

To prevent active volumes or host mappings from inadvertent deletion, the system supports a global setting that prevents these objects from being deleted if the system detects that they have recent I/O activity. When you delete a volume, the system checks to verify whether it is part of a host mapping, FlashCopy mapping, or remote-copy relationship. In these cases, the system fails to delete the volume, unless the **-force** parameter is specified. Using the **-force** parameter can lead to unintentional deletions of volumes that are still active. Active means that the system detected recent I/O activity to the volume from any host.

Write-through mode

Write-through mode is a process in which data is written to a storage device at the same time that the data is cached.

Written capacity

See “Capacity” on page 817.

Related publications

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

IBM Redbooks

The following IBM Redbooks publications (or later versions as they become available) can provide more information about the topics in this book. Some publications that are referenced in the following list might be available in softcopy only:

- ▶ *IBM System Storage SAN Volume Controller and Storwize V7000 Best Practices and Performance Guidelines*, SG24-7521
- ▶ *IBM b-type Gen 5 16 Gbps Switches and Network Advisor*, SG24-8186
- ▶ *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

You can search for, view, download, or order these documents and other Redbooks publications, Redpaper publications, web docs, drafts, and other materials, at the following website:

<http://www.ibm.com/redbooks>

IBM Storwize V5000 publications and support

The IBM Storwize V5000 publications and support are available at IBM Knowledge Center:

<http://ibm.biz/BdsBsB>

Help from IBM

IBM Support and downloads:

ibm.com/support

IBM Global Services:

ibm.com/services



Implementing the IBM Storwize V5000 Gen2 (including the

SG24-8162-04
ISBN 0738457655



(1.5" spine)
1.5" <-> 1.998"
789 <-> 1051 pages



SG24-8162-04

ISBN 0738457655

Printed in U.S.A.

Get connected

