Implementing the IBM Storwize V7000 V7.4

Discover the exciting new version of the IBM virtualization family

Become familiar with the leading-edge intuitive GUI

See how simple IBM virtualization really is

Jon Tate
Frank Enders
Torben Jensen
Hartmut Lonzan
Libor Miklas
Marcin Tabinowski
Note: Before using this information and the product it supports, read the information in “Notices” on page xi.

Fourth Edition (March 2015)

This edition applies to IBM Storwize software version 7.4 for the IBM Storwize V7000, both Gen1 and Gen2 versions.

© Copyright International Business Machines Corporation 2015. 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

<table>
<thead>
<tr>
<th>Notices</th>
<th>xi</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trademarks</td>
<td>xii</td>
</tr>
<tr>
<td>IBM Redbooks promotions</td>
<td>xiii</td>
</tr>
<tr>
<td>Preface</td>
<td>xv</td>
</tr>
<tr>
<td>Authors</td>
<td>xv</td>
</tr>
<tr>
<td>Now you can become a published author, too!</td>
<td>xvii</td>
</tr>
<tr>
<td>Comments welcome</td>
<td>xviii</td>
</tr>
<tr>
<td>Stay connected to IBM Redbooks</td>
<td>xviii</td>
</tr>
<tr>
<td>Summary of changes</td>
<td>xix</td>
</tr>
<tr>
<td>March 2015, Fourth Edition</td>
<td>xix</td>
</tr>
</tbody>
</table>

## Chapter 1. IBM Storwize V7000 system overview

1.1 Storage virtualization ........................................... 2
1.2 IBM Storwize V7000 overview ....................................... 3
1.3 IBM Storwize V7000 terminology ................................... 4
   1.3.1 IBM Storwize V7000 models .................................... 5
   1.3.2 IBM Storwize V7000 functions .................................. 7
   1.3.3 IBM Storwize V7000 licensing .................................. 9
1.4 IBM Storwize V7000 hardware ..................................... 10
1.5 IBM Storwize V7000 components ................................... 11
   1.5.1 Hosts ....................................................... 11
   1.5.2 Nodes ....................................................... 11
   1.5.3 I/O groups .................................................. 11
   1.5.4 Clustered system ............................................. 12
   1.5.5 RAID ......................................................... 13
   1.5.6 Managed disks ................................................ 14
   1.5.7 Quorum disks ................................................ 14
   1.5.8 Storage pools ................................................. 14
   1.5.9 Volumes ...................................................... 17
   1.5.10 Thin-provisioned volumes .................................... 19
   1.5.11 Mirrored volumes ............................................ 20
   1.5.12 Easy Tier .................................................. 21
   1.5.13 iSCSI ....................................................... 21
   1.5.14 Real-time Compression ....................................... 23
   1.5.15 IP replication .............................................. 24
1.6 Advanced copy services .......................................... 24
   1.6.1 Synchronous or asynchronous remote copy ...................... 25
   1.6.2 FlashCopy ................................................... 26
   1.6.3 Copy Services configuration limits ........................... 27
1.7 Management and support tools .................................... 27
   1.7.1 IBM Assist On-site and remote service ....................... 27
   1.7.2 Event notifications .......................................... 28
   1.7.3 SNMP traps .................................................. 28
   1.7.4 Syslog messages .............................................. 28
   1.7.5 Call Home email ............................................. 29
1.8 Useful Storwize V7000 websites .................................. 29
Chapter 2. IBM Storwize V7000 Initial configuration
2.1 Network requirements for the IBM Storwize V7000 Gen1 and Gen2
2.2 Initial configuration of IBM Storwize V7000 Gen2
2.3 Initial configuration of IBM Storwize V7000 Gen1
2.4 Secure Shell overview
2.4.1 Generating public and private SSH key pairs by using PuTTY
2.4.2 Uploading the SSH public key to the SAN Volume Controller cluster
2.4.3 Configuring the PuTTY session for the CLI
2.4.4 Starting the PuTTY CLI session
2.4.5 Configuring SSH for AIX clients
2.5 Using IPv6
2.5.1 Migrating a cluster from IPv4 to IPv6
2.5.2 Migrating a cluster from IPv6 to IPv4

Chapter 3. Graphical user interface
3.1 Overview of IBM Storwize V7000 management software
3.1.1 Access to the storage management software
3.1.2 System panel layout
3.1.3 Navigation
3.1.4 Multiple selection
3.1.5 Status indicators area
3.2 Overview panel
3.3 Monitoring menu
3.3.1 System overview
3.3.2 System Details option
3.3.3 Events
3.3.4 Performance
3.4 Pools menu
3.4.1 Volumes by pool
3.4.2 Internal storage
3.4.3 External storage
3.4.4 MDisks by pools
3.4.5 System migration
3.5 Volumes menu
3.5.1 All volumes
3.5.2 Volumes by pool
3.5.3 Volumes by host
3.6 Hosts menu
3.6.1 Hosts
3.6.2 Ports by host
3.6.3 Host mappings
3.6.4 Volumes by host
3.7 Copy Services
3.7.1 FlashCopy
3.7.2 Consistency Group
3.7.3 FlashCopy mappings
3.7.4 Remote Copy
3.7.5 Partnerships
3.8 Access menu
3.8.1 Users

1.8.1 IBM Storwize V7000 learning videos on YouTube
3.8.2 Audit Log menu ......................................................... 103
3.9 Settings menu ........................................................... 104
  3.9.1 Notifications ......................................................... 104
  3.9.2 Network ............................................................... 105
  3.9.3 Security ............................................................... 105
  3.9.4 System ............................................................... 107
  3.9.5 Support .............................................................. 108
  3.9.6 GUI preferences .................................................... 108

Chapter 4. Host configuration ........................................... 111
  4.1 Host attachment overview .......................................... 112
  4.2 Preparing the host operating system ............................ 113
    4.2.1 Windows 2008 R2 and 2012 R2: Preparing for Fibre Channel attachment .... 113
    4.2.2 Windows 2008 R2 and Windows 2012 R2: Preparing for iSCSI attachment .... 119
    4.2.3 VMware ESXi: Preparing for Fibre Channel attachment ....................... 121
    4.2.4 VMware ESXi: Preparing for iSCSI attachment .................................. 125
  4.3 Creating hosts using the GUI ....................................... 132
    4.3.1 Creating Fibre Channel hosts .................................. 133
    4.3.2 Creating iSCSI hosts ............................................ 137

Chapter 5. Basic volume configuration ............................... 145
  5.1 Provisioning storage from IBM Storwize V7000 to hosts .... 146
    5.1.1 Creating generic volumes ..................................... 147
    5.1.2 Creating thin-provisioned volumes ........................... 149
    5.1.3 Creating mirrored volumes .................................... 152
    5.1.4 Creating a thin-mirrored volume ............................. 155
    5.1.5 Creating compressed volumes ................................ 157
  5.2 Mapping a volume to the host ..................................... 160
    5.2.1 Mapping newly created volumes to the host using the wizard .............. 160
  5.3 Discovering volumes on hosts and multipathing .............. 161
    5.3.1 Windows 2008 Fibre Channel volume attachment ....................... 163
    5.3.2 Windows 2008 iSCSI volume attachment ........................ 167
    5.3.3 VMware ESX Fibre Channel attachment .......................... 176
    5.3.4 VMware ESX iSCSI attachment ................................ 180

Chapter 6. Migration wizard ............................................ 189
  6.1 Preparing for data migration ....................................... 190
  6.2 Data migration using the migration wizard .................... 190

Chapter 7. Storage pools ............................................... 203
  7.1 Working with internal drives ..................................... 204
    7.1.1 Actions on internal drives .................................... 206
    7.1.2 Configuring internal storage .................................. 210
  7.2 Working with MDisks ............................................... 220
    7.2.1 Adding MDisks to storage pools ................................ 222
    7.2.2 Importing MDisks ............................................... 226
    7.2.3 RAID action for MDisks ......................................... 230
    7.2.4 Selecting the tier for MDisks .................................. 232
    7.2.5 Additional actions on MDisks .................................. 234
  7.3 Working with storage pools ......................................... 237
    7.3.1 Child storage pools ............................................. 241

Chapter 8. Advanced host and volume administration ............ 247
  8.1 Advanced host administration ..................................... 248
11.4.6 Interaction and dependency between Multiple Target FlashCopy mappings . . . . . . . . . . . . . . . . 394
11.4.7 Summary of the FlashCopy indirection layer algorithm. .................................................. 395
11.4.8 Interaction with the cache ........................................................................................................ 396
11.4.9 FlashCopy and image mode volumes ....................................................................................... 397
11.4.10 FlashCopy mapping events ....................................................................................................... 399
11.4.11 FlashCopy mapping states ....................................................................................................... 400
11.4.12 Thin-provisioned FlashCopy ..................................................................................................... 402
11.4.13 Background copy .................................................................................................................... 403
11.4.14 Synthesis .................................................................................................................................. 404
11.4.15 Serialization of I/O by FlashCopy .......................................................................................... 404
11.4.16 Event handling ........................................................................................................................ 405
11.4.17 Asynchronous notifications ..................................................................................................... 406
11.4.18 Interoperation with Metro Mirror and Global Mirror ............................................................... 406
11.4.19 FlashCopy presets ................................................................................................................... 407
11.5 Volume mirroring and migration options ...................................................................................... 408
11.6 Native IP replication ...................................................................................................................... 410
11.6.1 Native IP replication technology ............................................................................................... 410
11.6.2 IP partnership limitations ......................................................................................................... 412
11.6.3 VLAN support ........................................................................................................................... 413
11.6.4 IP partnership and Storwize V7000 terminology ......................................................................... 413
11.6.5 States of IP partnership ............................................................................................................ 414
11.6.6 Remote copy groups .................................................................................................................. 415
11.6.7 Supported configurations .......................................................................................................... 416
11.7 Remote Copy .................................................................................................................................. 428
11.7.1 Multiple Storwize V7000 System Mirroring .............................................................................. 429
11.7.2 Importance of write ordering ..................................................................................................... 432
11.7.3 Remote copy intercluster communication .................................................................................. 434
11.7.4 Metro Mirror overview ............................................................................................................. 435
11.7.5 Synchronous remote copy ......................................................................................................... 436
11.7.6 Metro Mirror features ................................................................................................................ 437
11.7.7 Metro Mirror attributes ............................................................................................................. 437
11.7.8 Global Mirror ............................................................................................................................. 438
11.7.9 Asynchronous remote copy ....................................................................................................... 438
11.7.10 Storwize V7000 Global Mirror features .................................................................................. 440
11.7.11 Using Change Volumes with Global Mirror ............................................................................. 442
11.7.12 Distribution of work among nodes ........................................................................................... 445
11.7.13 Background copy performance ............................................................................................... 445
11.7.14 Thin-provisioned background copy .......................................................................................... 446
11.7.15 Methods of synchronization .................................................................................................... 446
11.7.16 Practical use of Metro Mirror .................................................................................................. 447
11.7.17 Practical use of Global Mirror .................................................................................................. 448
11.7.18 Valid combinations of FlashCopy, Metro Mirror, and Global Mirror ........................................ 448
11.7.19 Remote Copy configuration limits ........................................................................................... 448
11.7.20 Remote Copy states and events ............................................................................................... 449
11.8 Remote Copy commands ............................................................................................................... 456
11.8.1 Remote Copy process ................................................................................................................ 456
11.8.2 Listing available Storwize V7000 system partners .................................................................... 457
11.8.3 Changing the system parameters ............................................................................................... 457
11.8.4 Storwize V7000 system partnership ............................................................................................ 459
11.8.5 Creating an MM/GM Consistency Group .................................................................................... 460
11.8.6 Creating an MM/GM relationship ............................................................................................... 460
11.8.7 Changing an MM/GM relationship ............................................................................................. 460
11.8.8 Changing an MM/GM Consistency Group ................................................................................ 461
Chapter 12. Encryption

12.1 Encryption using the Storwize V7000 Gen2 .................................................. 468
12.2 Encryption activation .......................................................... 472
12.3 Automatic activation of the Encryption feature ......................... 474
12.4 Manual activation of the Encryption feature .............................. 478
12.5 Enabling encryption .................................................... 482
12.5.1 CLI commands to enable encryption ........................................ 491
12.6 Rekeying an encryption-enabled system .................................... 492
12.7 Disabling encryption ..................................................... 494

Chapter 13. RAS, monitoring, and troubleshooting

13.1 Reliability, availability, and serviceability ................................. 496
13.1.1 Node canisters ................................................... 496
13.1.2 Expansion canisters ............................................ 498
13.1.3 Enclosure SAS cabling ........................................ 499
13.1.4 Power ......................................................... 500
13.2 Configuration backup .................................................. 501
13.2.1 Backup using CLI ............................................. 502
13.2.2 Backup using GUI ............................................ 503
13.3 Software upgrade ...................................................... 504
13.3.1 Precautions before the upgrade .................................. 504
13.3.2 Storwize V7000 upgrade test utility .................................. 505
13.3.3 Upgrade procedure Version 7.3.x.x to 7.4.x.x .......................... 505
13.3.4 Upgrade procedure Version 7.4.x.x to 7.4.y.y ......................... 512
13.4 Critical Fix Notification feature ........................................... 519
13.5 Monitoring host paths to IBM Storwize V7000 ......................... 520
13.6 Troubleshooting and fix procedures ........................................ 521
13.6.1 Managing event log ............................................ 523
13.6.2 Running a fix procedure ....................................... 524
13.6.3 Event log details ................................................ 529
13.7 Monitoring .............................................................. 530
13.7.1 Email notifications and Call Home .................................. 531
13.8 Audit log .............................................................. 536
13.9 Collecting support information ........................................... 538
13.10 Shutting down IBM Storwize V7000 .................................... 541
13.10.1 Shutting down a node canister .................................. 541
13.10.2 Shutting down a system ......................................... 543
13.10.3 Shutting down and powering on an IBM Storwize V7000 infrastructure . . 544

Appendix A. CLI setup and SAN Boot

Command-line interface .......................................................... 546
Basic setup ................................................................. 546
Notices

This information was developed for products and services offered in the U.S.A.

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, Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: 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 states 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 supply in any way it believes appropriate without incurring any obligation to you.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

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.

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 the names and addresses used by an actual business enterprise 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.
Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at http://www.ibm.com/legal/copytrade.shtml.

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

- AIX®
- DB2®
- DS6000™
- DS8000®
- Easy Tier®
- FlashCopy®
- Global Technology Services®
- IBM®
- POWER®
- Real-time Compression™
- Redbooks®
- Redbooks (logo) ®
- Storwize®
- System Storage®
- Tivoli®
- XIV®

The following terms are trademarks of other companies:

Intel, 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.

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.

Other company, product, or service names may be trademarks or service marks of others.
Find and read thousands of IBM Redbooks publications

- Search, bookmark, save and organize favorites
- Get up-to-the-minute Redbooks news and announcements
- Link to the latest Redbooks blogs and videos

Get the latest version of the Redbooks Mobile App

Promote your business in an IBM Redbooks publication

Place a Sponsorship Promotion in an IBM Redbooks publication, featuring your business or solution with a link to your web site.

Qualified IBM Business Partners may place a full page promotion in the most popular Redbooks publications. Imagine the power of being seen by users who download millions of Redbooks publications each year!
Preface

Continuing its commitment to developing and delivering industry-leading storage technologies, IBM® introduces the IBM Storwize® V7000 solution, an innovative storage offering that delivers essential storage efficiency technologies and exceptional ease of use and performance, all integrated into a compact, modular design that is offered at a competitive, midrange price.

The IBM Storwize V7000 solution incorporates some of the top IBM technologies typically found only in enterprise-class storage systems, raising the standard for storage efficiency in midrange disk systems. This cutting-edge storage system extends the comprehensive storage portfolio from IBM and can help change the way organizations address the ongoing information explosion.

This IBM Redbooks® publication introduces the features and functions of the IBM Storwize V7000 system through several examples. This book is aimed at pre-sales and post-sales technical support and marketing, storage administrators, and will help you understand the architecture of the Storwize V7000, how to implement it, and take advantage of the industry-leading functions and features.

Authors

This book was produced by a team of specialists from around the world working at the IBM 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 joining the ITSO in 1999, he worked in the IBM Technical Support Center, providing Level 2 support for IBM storage products. Jon has 28 years of experience in storage software and management, services, and support, and 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.
Frank Enders has worked for the last seven years for EMEA IBM System Storage SAN Volume Controller and V7k Level 2 support in Mainz, Germany, and his duties include pre-sales and post-sales support. He has worked for IBM Germany for 20 years and started as a technician in disk production for IBM Mainz, and changed to magnetic head production four years later. When IBM closed disk production in Mainz in 2001, he also changed his role and continued working for IBM within ESCC Mainz as a team member of the Installation Readiness team for products such as the DS8000®, DS6000™, and the IBM System Storage SAN Volume Controller. During that time, he studied for four years to gain a diploma in Electrical Engineering.

Torben Jensen is an IT Specialist at IBM Global Technology Services®, Copenhagen, Denmark. He joined IBM in 1999 for an apprenticeship as an IT-System Supporter. From 2001 to 2005, he was the client representative for IBM Internal Client platforms in Denmark. Torben started work with the SAN/DISK for open systems department in March 2005. He provides daily and ongoing support to customers as well as colleagues, and works with SAN designs and solutions for customers. Torben has written several Redbooks publications within the Storage virtualization area in the past.

Hartmut Lonzer is a Client Technical Specialist for IBM Germany. He works in the IBM Germany headquarters in Ehningen. Hartmut is a member of the Technical Sales Support for Storage in Germany, Austria, and Switzerland. His main focus is on IBM SAN Volume Controller and the IBM Storwize Family. 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 37 years.

Libor Miklas is an IT Architect working at the IBM Integrated Delivery Center in the Czech Republic. He demonstrates ten years of extensive experience within the IT industry. During the last eight years, his main focus has been on the data protection solutions and on storage management and provisioning. Libor and his team design, implement, and support midrange and enterprise storage environments for various global and local clients, worldwide. He is an IBM Certified Deployment Professional for the IBM Tivoli® Storage Manager family of products and holds a Master's degree in Electrical Engineering and Telecommunications.
Thanks to the following people for their contributions to this project:

Trevor Boardman
Paul Cashman
John Fairhurst
Katja Gebuhr
Paul Merrison
Nicholas Sunderland
Barry Whyte
Stephen Wright

**IBM Hursley, UK**

Nick Clayton

**IBM Systems & Technology Group, UK**

Helen Burton

**IBM Systems & Technology Group, Boulder, CO, US**

Special thanks to the Brocade Communications Systems staff in San Jose, California for their unparalleled support of this residency in terms of equipment and support in many areas:

Jim Baldyga
Silviano Gaona
Sangam Racherla
Brian Steffler
Marcus Thordal

**Brocade Communications Systems**

---

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 found at:
  ibm.com/redbooks
- Send your comments in an email to:
  redbooks@us.ibm.com
- Mail your comments to:
  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:
  https://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:
- 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 SG24-7938-03
for Implementing the IBM Storwize V7000 V7.4
as created or updated on July 16, 2015.

March 2015, Fourth Edition

This revision includes the following new and changed information.

**New information**
- Encryption
- New software and hardware description

**Changed information**
- GUI and CLI
- Compression
IBM Storwize V7000 system overview

This chapter provides an overview of the IBM Storwize V7000 architecture with benefits of IBM Storwize Software Family V7.4. It helps you understand the differences between the first and second generations of IBM Storwize V7000.

At first, it briefly explains the concept of storage virtualization and how IBM implements it in the Storwize family of products.
### 1.1 Storage virtualization

Storage virtualization, like server virtualization, is one of the foundations of building a flexible and reliable infrastructure solution that allows companies to better align their business and IT needs.

Storage virtualization allows an organization to implement pools of storage across physically separate disk systems (which might be from different vendors). Storage can then be deployed from these pools and can be migrated between pools without any outage of the attached host systems. Storage virtualization provides a single set of tools for advanced functions, such as instant copy and remote mirroring solutions, which means that deploying storage can be performed by using a single tool regardless of the underlying storage hardware.

Figure 1-1 shows a storage virtualization scenario.

![Figure 1-1 Storage virtualization](image)

For a more detailed explanation of storage virtualization, see *Implementing the IBM System Storage SAN Volume Controller V7.4*, SG24-7933.
1.2 IBM Storwize V7000 overview

The IBM Storwize V7000 solution provides a modular storage system that includes the capability to virtualize external SAN-attached storage and its own internal storage. The IBM Storwize V7000 solution is built upon the IBM SAN Volume Controller technology and uses components from the IBM System Storage DS8000 family.

An IBM Storwize V7000 system provides several configuration options that are aimed at simplifying the implementation process. It also provides automated wizards, called Directed Maintenance Procedures (DMP) to assist in resolving any events that might occur. An IBM Storwize V7000 system is a clustered, scalable, and midrange storage system, and an external virtualization device.

Figure 1-2 shows the IBM Storwize V7000 high-level overview.

![IBM Storwize V7000 overview](image)

Included with an IBM Storwize V7000 system is a graphical user interface (GUI) that allows storage to be deployed quickly and efficiently. The GUI runs on the IBM Storwize V7000 system, so there is no need for a separate console. The management GUI contains a series of preestablished configuration options that are called *presets* and that use common settings to quickly configure objects on the system. Presets are available for creating volumes and IBM FlashCopy® mappings and for setting up a Redundant Array of Independent Disks (RAID) configuration.

The IBM Storwize V7000 solution provides a choice of up to 480 x 3.5 inch or 1056 x 2.5 inch serial-attached SCSI (SAS) drives for the internal storage in a clustered system and uses SAS cables and connectors to attach to the optional expansion enclosures. In a clustered system, the V7000 can provide up to 6 PiB raw capacity (with 6 TB nearline SAS disks).

When virtualizing external storage arrays, IBM Storwize V7000 system can provide up to 32 PiB of usable capacity. An IBM Storwize V7000 system supports a range of external disk systems similar to what the SAN Volume Controller supports today. See Figure 1-3 on page 4 for a view of the Storwise V7000 control enclosure.
IBM Storwize V7000 solution consists of one to four control enclosures and optionally, up to 36 expansion enclosures (and supports the intermixing of the different expansion enclosures). Within each enclosure are two canisters. Control enclosures contain two node canisters, and expansion enclosures contain two expansion canisters.

### 1.3 IBM Storwize V7000 terminology

IBM Storwize V7000 system uses the terminology defined in Table 1-1.

<table>
<thead>
<tr>
<th>IBM Storwize V7000 term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chain</td>
<td>A set of enclosures that provide redundant access to the drives that are inside the enclosures. Each control enclosure can have one or more chains.</td>
</tr>
<tr>
<td>Clone</td>
<td>A copy of a volume on a server at a particular point in time. The contents of the copy can be customized while the contents of the original volume are preserved.</td>
</tr>
<tr>
<td>Control enclosure</td>
<td>A hardware unit that includes the chassis, node canisters, drives, and power sources that include batteries.</td>
</tr>
<tr>
<td>Event</td>
<td>An occurrence that is significant to a task or system. Events can include completion or failure of an operation, a user action, or the change in the state of a process.</td>
</tr>
<tr>
<td>Expansion canister</td>
<td>A hardware unit that includes the serial-attached SCSI (SAS) interface hardware that enables the node hardware to use the drives of the expansion enclosure.</td>
</tr>
<tr>
<td>Expansion enclosure</td>
<td>A hardware unit that includes expansion canisters, drives, and power sources that do not include batteries.</td>
</tr>
<tr>
<td>External storage</td>
<td>Managed disks (MDisks) that are Small Computer Systems Interface (SCSI) logical units presented by storage systems that are attached to and managed by the clustered system.</td>
</tr>
<tr>
<td>Host mapping</td>
<td>The process of controlling which hosts have access to specific volumes within a clustered system.</td>
</tr>
</tbody>
</table>
### IBM Storwize V7000 term | Definition
--- | ---
Internal storage | Array managed disks (MDisks) and drives that are held in enclosures and nodes that are part of the clustered system.
Lane | A single SAS channel.
Managed disk (MDisk) | A component of a storage pool that is managed by a clustered system. An MDisk is either part of a RAID array of internal storage or a Small Computer System Interface (SCSI) logical unit (LU) for external storage. An MDisk is not visible to a host system on the storage area network.
Node canister | A hardware unit that includes the node hardware, fabric and service interfaces, and serial-attached SCSI (SAS) expansion ports.
PHY | A single SAS lane. There are four PHYs in each SAS cable.
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. Quorum disks can either be MDisks or internal drives.
Snapshot | An image backup type that consists of a point-in-time view of a volume.
Storage pool | A collection of storage capacity that provides the capacity requirements for a volume.
Strand | The serial-attached SCSI (SAS) connectivity of a set of drives within multiple enclosures. The enclosures can be either 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 assigned to that storage unit.
Volume | A discrete unit of storage on disk, tape, or other data recording medium that supports some form of identifier and parameter list, such as a volume label or input/output control.
Discovery | A process by which two Storwize clusters exchange information about their IP address configuration.
Remote Copy Port Group | The set of local and remote Ethernet ports that can access each other via a long-distance IP link.
Remote Copy Port Group ID | A numeric value that indicates to which group the port belongs.
RC login | A bidirectional full-duplex data path between two Storwize clusters that are remote copy (RC) partners.
Path configuration | The act of setting up RC logins for two partnered Storwize systems.

### 1.3.1 IBM Storwize V7000 models

The IBM Storwize V7000 platform consists of enclosures and drives. Each enclosure contains two canisters that are seen as part of the enclosure, although they can be replaced independently.
**Additional IBM Storwize V7000 information:** For the most up-to-date information about the features, benefits, and specifications of IBM Storwize V7000 models, see the following address:


The information in this book is valid at the time of writing, but as the IBM Storwize V7000 matures, expect to see new features and enhanced specifications.

The IBM Storwize V7000 models are described in Table 1-2.

**Table 1-2 IBM Storwize V7000 models**

<table>
<thead>
<tr>
<th>Model</th>
<th>Cache</th>
<th>FC / iSCSI / SAS ports</th>
<th>Drive slots</th>
<th>Power supply</th>
</tr>
</thead>
<tbody>
<tr>
<td>2076-112 (with two node canisters)</td>
<td>16 GB</td>
<td>8x8 Gb / 4x1 Gb / 4x6 Gb</td>
<td>12 x 3.5-inch</td>
<td>Integrated dual power supplies with battery backup</td>
</tr>
<tr>
<td>2076-124 (with two node canisters)</td>
<td>16 GB</td>
<td>8x8 Gb / 4x1 Gb / 4x6 Gb</td>
<td>24 x 2.5-inch</td>
<td>Integrated dual power supplies with battery backup</td>
</tr>
<tr>
<td>2076-312 (with two node canisters)</td>
<td>16 GB</td>
<td>8x8 Gb / 4x1 Gb+4x10Gb / 4x6 Gb</td>
<td>12 x 3.5-inch</td>
<td>Integrated dual power supplies with battery backup</td>
</tr>
<tr>
<td>2076-324 (with two node canisters)</td>
<td>16 GB</td>
<td>8x8 Gb / 4x1 Gb+4x10Gb / 4x6 Gb</td>
<td>24 x 2.5-inch</td>
<td>Integrated dual power supplies with battery backup</td>
</tr>
<tr>
<td>2076-524 (with two node canisters Gen2)</td>
<td>32/64 GB</td>
<td>4x16Gb / 4x1 Gb+4x10Gb / 4x6 Gb</td>
<td>24 x 2.5-inch</td>
<td>Integrated dual power supplies with battery backup</td>
</tr>
<tr>
<td>2076-212 (with two expansion canisters)</td>
<td>Not applicable (N/A)</td>
<td>-- / -- / 4x6 Gb</td>
<td>12 x 3.5-inch</td>
<td>Integrated dual power supplies</td>
</tr>
<tr>
<td>2076-224 (with two expansion canisters)</td>
<td>N/A</td>
<td>-- / -- / 4x6 Gb</td>
<td>24 x 2.5-inch</td>
<td>Integrated dual power supplies</td>
</tr>
<tr>
<td>2076-12F (with two expansion canisters Gen2)</td>
<td>N/A</td>
<td>-- / -- / 4x12 Gb</td>
<td>12 x 3.5-inch</td>
<td>Integrated dual power supplies. Attaches to 2076-524 only</td>
</tr>
<tr>
<td>2076-24F (with two expansion canisters Gen2)</td>
<td>N/A</td>
<td>-- / -- / 4x12 Gb</td>
<td>24 x 2.5-inch</td>
<td>Integrated dual power supplies. Attaches to 2076-524 only</td>
</tr>
</tbody>
</table>
1.3.2 IBM Storwize V7000 functions

The following functions are available with IBM Storwize V7000:

- **Thin provisioning (no license required)**
  
  Traditional fully allocated volumes allocate real physical disk capacity for an entire volume even if that capacity is never used. Thin-provisioned volumes allocate real physical disk capacity only when data is written to the logical volume.

- **Volume mirroring (no license required)**
  
  Provides a single volume image to the attached host systems while maintaining pointers to two copies of data in separate storage pools. Copies can be on separate disk storage systems that are being virtualized. If one copy is failing, IBM Storwize V7000 provides continuous data access by redirecting I/O to the remaining copy. When the copy becomes available, automatic resynchronization occurs.

- **FlashCopy (licensed per enclosure)**
  
  Provides a volume level point-in-time copy function for any storage being virtualized by IBM Storwize V7000. This function creates copies for backup, parallel processing, testing, and development, and have the copies available almost immediately.

  IBM Storwize V7000 includes the following FlashCopy functions:
  - Full or incremental copy
    
    This function copies only the changes from either the source or target data since the last FlashCopy operation and enables completion of point-in-time online backups much more quickly than using traditional FlashCopy.
  - Multitarget FlashCopy
    
    IBM Storwize V7000 supports copying of up to 256 target volumes from a single source volume. Each copy is managed by a unique mapping and in general, each mapping acts independently and is not affected by other mappings sharing the source volume.
  - Cascaded FlashCopy
    
    This function is used to create copies of copies and supports full, incremental, or nocopy operations.
  - Reverse FlashCopy
    
    This function allows data from an earlier point-in-time copy to be restored with minimal disruption to the host.
  - FlashCopy nocopy with thin provisioning
    
    This function provides a combination of using thin-provisioned volumes and FlashCopy together to help reduce disk space requirements when making copies. This option has two variations:
    - Space-efficient source and target with background copy
      
      Copies only the allocated space.
    - Space-efficient target with no background copy
      
      Copies only the space used for changes between the source and target and is generally referred to as snapshots.

    This function may be used with multitarget, cascaded, and incremental FlashCopy.
  - Consistency groups
    
    Consistency groups address the issue where application data is on multiple volumes. By placing the FlashCopy relationships into a consistency group, commands can be
issued against all of the volumes in the group. This action enables a consistent point-in-time copy of all of the data, even though it might be on a physically separate volume.

FlashCopy mappings can be members of a consistency group, or they can be operated in a stand-alone manner, that is, not as part of a consistency group. FlashCopy commands can be issued to a FlashCopy consistency group, which affects all FlashCopy mappings in the consistency group, or to a single FlashCopy mapping if it is not part of a defined FlashCopy consistency group.

- Metro Mirror (licensed based on the number of enclosures)
  Provides a synchronous remote mirroring function up to approximately 300 km between sites. As the host I/O only completes after the data is cached at both locations, performance requirements might limit the practical distance. Metro Mirror provides fully synchronized copies at both sites with zero data loss after the initial copy is completed.
  Metro Mirror can operate between multiple IBM Storwize systems.

- Global Mirror (licensed based on the number of enclosures)
  Provides long-distance asynchronous remote mirroring function up to approximately 8,000 km between sites. With Global Mirror, the host I/O completes locally and the changed data is sent to the remote site later. This maintains a consistent recoverable copy of data at the remote site, which lags behind the local site.
  Global Mirror can operate between multiple IBM Storwize systems.

- External Virtualization (license per external enclosure, no charge for data migration)
  IBM Storwize V7000 provides a data migration function that can be used to import external storage systems into the IBM Storwize V7000 system. You can do these tasks:
  - Move volumes nondisruptively onto a newly installed storage system
  - Move volumes to rebalance a changed workload
  - Migrate data from other back-end storage to IBM Storwize V7000 managed storage

- Encryption
  IBM Storwize V7000 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 only be enabled and configured on enclosures that support encryption.

- IBM Easy Tier® (licensed per enclosure)
  Provides a mechanism to seamlessly migrate hot spots to the most appropriate tier within the IBM Storwize V7000 solution. This migration could be to internal drives within IBM Storwize V7000 or to external storage systems that are virtualized by IBM Storwize V7000. Independently on Easy Tier, IBM Storwize V7000 provides an automatic storage pool balancing that is enabled by default and requires no license. This function is shown in Figure 1-4 on page 9.
IBM Real-time Compression™ (licensed per enclosure, specific hardware (HW) required)

IBM Real-time Compression technology is based on the Random Access Compression Engine (RACE). RACE is an integral part of the software stack of IBM Storwize V7000 version 6.4.0 and later. This integration does not alter the behavior of the system so that previously existing features are supported for compressed volumes. IBM Storwize V7000 V7.4 uses two software-enabled RACE engines sharing HW resources for compression.

1.3.3 IBM Storwize V7000 licensing

With the broad range of technical features and capabilities of the IBM Storwize V7000 Gen2, including Copy Services, External Virtualization, Easy Tier, and Real-time Compression, IBM simplified the licensing model to include these new features. IBM Storwize V7000 offers two ways of license procurement:

- Fully flexible
- Bundled (license packages)

The license model is based on license-per-enclosure concept known from the first generation of IBM Storwize V7000, however the second generation offers more flexibility that exactly matches your needs.

Upgrade: Installation/upgrade of the code on first generation of IBM Storwize V7000 to the version V7.4 does not change your existing license model or license needs.
The conceptual model of the licensing in the second generation of IBM Storwize V7000 is depicted in Figure 1-5.

IBM Storwize V7000 V7.4 introduces new software licenses for Storwize V7000 Gen2. This new license and pricing structure provides intuitive procurement based on the functions clients want to enable and use the most.

The base module is represented by IBM Storwize Family Software V7000 and is mandatory for every controller, enclosure, or externally managed controller unit. Additional licensed features can be purchased on-demand either as a full software bundle or each feature separately. Again, these additional licenses need to be procured per every existing enclosure where they are planned to be used.

1.4 IBM Storwize V7000 hardware

In conjunction with the previous release of the 7.3 code, IBM introduced a hardware refresh for the IBM Storwize V7000 platform. These improvements are further multiplied with the latest code release V7.4. In this section, we introduce these hardware changes and software improvements associated with the 2076-524 model. These include:

- New internal component layout, such as canister, power supplies, ports
- Integrated battery pack within node canisters
- Enhanced scalability and flexibility with 16 Gbps I/O adapters
- Improved Real-time Compression engine with hardware assistance
- Extended disk drive support

To meet these objectives, the base hardware configuration of the Storwize V7000 Gen2 was improved substantially to support more advanced processors, more memory, and faster interconnects.

The detailed hardware description is not primarily intended for this publication. All technical details are available in Implementing the IBM Storwize V7000 Gen2, SG24-8244.
1.5 IBM Storwize V7000 components

The IBM Storwize V7000 is a midrange virtualization RAID storage subsystem. It has the following benefits:

- Brings enterprise technology to midrange storage.
- Specialty administrators are not required.
- Client setup and service can be easy.
- The system can grow incrementally as storage capacity and performance needs change.
- Multiple storage tiers are in a single system with nondisruptive migration between them.
- Simple integration can be done into the server environment.

The IBM Storwize V7000 subsystem consists of a set of drive enclosures. Control enclosures contain disk drives and two nodes (an I/O group), which are attached to the SAN fabric. Expansion enclosures contain drives and are attached to control enclosures.

The simplest use of the IBM Storwize V7000 is as a traditional RAID subsystem. The internal drives are configured into RAID arrays and virtual disks created from those arrays.

The IBM Storwize V7000 can also be used to virtualize other storage controllers, as described in Chapter 9, “External storage virtualization” on page 323.

The IBM Storwize V7000 supports spinning disks and flash drives and uses IBM System Storage Easy Tier to automatically place volume hot spots on better performing storage. Even without Easy Tier license, the Automatic Storage Pool balancing is available and enabled by default to distribute workloads across all drives in the array.

1.5.1 Hosts

A host system is a server that is connected to IBM Storwize V7000 through a Fibre Channel connection, Fibre Channel over Ethernet (FCoE), or through an iSCSI connection.

Hosts are defined to IBM Storwize V7000 by identifying their worldwide port names (WWPNs) for Fibre Channel hosts. For iSCSI hosts, they are identified by using their iSCSI names. The iSCSI names can either be iSCSI qualified names (IQNs) or extended unique identifiers (EUIs).

1.5.2 Nodes

IBM Storwize V7000 can have two or up to eight hardware components, called nodes or node canisters, which provide the virtualization of internal and external volumes, and cache and copy services (remote copy) functions. A clustered system consists of one or four node pairs.

One of the nodes within the system is known as the configuration node and it is the node that manages configuration activity for the clustered system. If this node fails, the system nominates another node to become the configuration node.

1.5.3 I/O groups

Within IBM Storwize V7000, there are one to four pairs of node canisters known as I/O groups. The IBM Storwize V7000 supports eight node canisters in the clustered system, which provides four I/O groups.
When a host server performs I/O to one of its volumes, all 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.

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 2048 volumes per I/O group). However, both nodes also act as a failover node for its partner node within the I/O group, so a node takes over the I/O workload from its partner node, if required, with no impact to the server's application.

In a Storwize V7000 environment, using active/active architecture, the I/O handling for a volume can be managed by both nodes of the I/O group. Therefore, it is mandatory for servers that are connected through Fibre Channel connectors to use multipath device drivers to be able to handle this capability.

The Storwize V7000 I/O groups are connected to the SAN so that all application servers accessing volumes from the I/O group have access to them. Up to 2048 host server objects can be defined in four I/O groups.

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

### 1.5.4 Clustered system

A clustered system consists of one to four pairs of nodes. All configuration, monitoring, and service tasks are performed at the system level and 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 system.

A process is provided to back up the system configuration data to disk so that the clustered system can be restored in the event of a disaster. This method does not back up application data, only Storwize V7000 system configuration information.

**System configuration backup:** After backing up the system configuration, save the backup data on your hard disk (or at the least outside of the SAN). If you are unable to access the Storwize V7000, you do not have access to the backup data if it is on the SAN.

For the purposes of remote data mirroring, two or more clustered systems (IBM Storwize systems or SAN Volume Controller systems starting from Version 6.3) must form a partnership before creating relationships between mirrored volumes.

**Important:** IBM Storwize V7000 6.3 introduced the layer parameter. It can be changed by running chsystem using only the command-line interface (CLI). The default is the storage layer, and you must change it to replication if you need to set up a copy services relationship between the IBM Storwize Family and the SAN Volume Controller.

One node is designated as the configuration node canister and it is the only node that activates the system IP address. If the configuration node canister fails, the system chooses a new configuration node and the new configuration node takes over the system IP addresses.
The system can be configured using either the IBM Storwize V7000 management software, the CLI, or through an application that uses the IBM Storwize V7000 CIMOM (IBM Tivoli Storage Productivity Center). IBM Systems Director also provides flexible server and storage management capability.

1.5.5 RAID

The Storwize V7000 setup contains several internal drive objects, but these drives cannot be directly added to storage pools. The drives need to 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, and have various restrictions regarding the number of members in the array.

IBM Storwize V7000 supports 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.

Each array has a set of goals that describe the location and performance of each array. A sequence of drive failures and hot spare takeovers can leave an array unbalanced, that is, with members that do not match these goals. The system automatically rebalances such arrays when the appropriate drives are available.

The following RAID levels are available:

- **RAID 0** (striping, no redundancy)
  RAID 0 arrays stripe data across the drives. The system supports RAID 0 arrays with just one member, which is similar to traditional just a bunch of disks (JBOD) attach. RAID 0 arrays have no redundancy, so they do not support hot spare takeover or immediate exchange. A RAID 0 array can be formed by one to eight drives.

- **RAID 1** (mirroring between two drives)
  RAID 1 arrays stripe data over mirrored pairs of drives. A RAID 1 array mirrored pair is rebuilt independently. A RAID 1 array can be formed by two drives only.

- **RAID 5** (striping, can survive one drive fault)
  RAID 5 arrays stripe data over the member drives with one parity strip on every stripe. RAID 5 arrays have single redundancy. The parity algorithm means that an array can tolerate no more than one member drive failure. A RAID 5 array can be formed by 3 to 16 drives.

- **RAID 6** (striping, can survive two drive faults)
  RAID 6 arrays stripe data over the member drives with two parity stripes (known as the P-parity and the Q-parity) on every stripe. The two parity strips are calculated using different algorithms, which give the array double redundancy. A RAID 6 array can be formed by 5 to 16 drives.

- **RAID 10** (RAID 0 on top of RAID 1)
  RAID 10 arrays have single redundancy. Although they can tolerate one failure from every mirrored pair, they cannot tolerate two-disk failures. One member out of every pair can be rebuilding or missing at the same time. A RAID 10 array can be formed by 2 to 16 drives.
1.5.6 Managed disks

A managed disk (MDisk) is the unit of storage that IBM Storwize V7000 virtualizes. This unit could be a logical volume on an external storage array presented to the IBM Storwize V7000 or a RAID array consisting of internal drives. The IBM Storwize V7000 can then allocate these MDisks into various storage pools. An MDisk is not visible to a host system on the storage area network because it is internal or zoned only to the IBM Storwize V7000 system.

An MDisk has four modes:

- **Array**
  Array mode MDisks are constructed from drives using the RAID function. Array MDisks are always associated with storage pools.

- **Unmanaged**
  Unmanaged MDisks are not being used by the system. This situation might occur when an MDisk is first imported into the system, for example.

- **Managed**
  Managed MDisks are assigned to a storage pool and provide extents so that volumes can use it.

- **Image**
  Image MDisks are assigned directly to a volume with a one-to-one mapping of extents between the MDisk and the volume. This situation is normally used when importing logical volumes into the clustered system that already have data on them, which ensures that the data is preserved as it is imported into the clustered system.

1.5.7 Quorum disks

A quorum disk is an MDisk that contains a reserved area for use exclusively by the system. In the Storwize V7000, 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.

If the environment has multiple storage systems, to avoid the possibility of losing all of the quorum disks because of a failure of a single storage system, you should allocate the quorum disk on different storage systems. It is possible to manage the quorum disks by using the CLI.

1.5.8 Storage pools

A storage pool is a collection of MDisks (up to 4096) that are grouped to provide capacity for volumes. All MDisks in the pool are split into extents with the same size. Volumes are then allocated out of the storage pool and are mapped to a host system.

**Storwize V7000 object names:** The names must begin with a letter, which cannot be numeric. The name can be a maximum of 63 characters. Valid characters are uppercase (A-Z), lowercase letters (a-z), digits (0 - 9), underscore (_), period (.), hyphen (-), and space. The names must not begin or end with a space.
MDisks can be added to a storage pool at any time to increase the capacity of the storage pool. MDisks can belong in only one storage pool and only MDisks in unmanaged mode can be added to the storage pool. When an MDisk is added to the storage pool, the mode changes from unmanaged to managed and vice versa when you remove it.

Each MDisk in the storage pool is divided into several extents. The size of the extent is selected by the administrator at creation time of the storage pool and cannot be changed later. The size of the extent ranges from 16 MB up to 8 GB. The default extent size in v7.4 is 1024 MB.

The extent size has a direct impact on the maximum volume size and storage capacity of the clustered system. A system can manage 4 million (4 x 1024 x 1024) extents. For example, a system with a 16 MB extent size can manage up to 16 MB x 4 MB = 64 TB of storage.

The effect of extent size on the maximum volume size is shown in Table 1-3, which lists the extent size and the corresponding maximum clustered system size.

### Table 1-3 Maximum volume capacity by extent size

<table>
<thead>
<tr>
<th>Extent size (MB)</th>
<th>Maximum generic volume size (GiB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>2048 (2 TiB)</td>
</tr>
<tr>
<td>32</td>
<td>4096 (4 TiB)</td>
</tr>
<tr>
<td>64</td>
<td>8192 (8 TiB)</td>
</tr>
<tr>
<td>128</td>
<td>16,384 (16 TiB)</td>
</tr>
<tr>
<td>256</td>
<td>32,768 (32 TiB)</td>
</tr>
<tr>
<td>512</td>
<td>65,536 (64 TiB)</td>
</tr>
<tr>
<td>1024</td>
<td>131,072 (128 TiB)</td>
</tr>
<tr>
<td>2048</td>
<td>262,144 (256 TiB)</td>
</tr>
<tr>
<td>4096</td>
<td>528,288 (512 TiB)</td>
</tr>
<tr>
<td>8192</td>
<td>1,056,576 (1 PiB)</td>
</tr>
</tbody>
</table>

Table 1-4 compares the maximum volume, MDisk, and system capacity for each extent size.

### Table 1-4 Maximum volume, MDisk, and system capacity for each extent size

<table>
<thead>
<tr>
<th>Extent size (MB)</th>
<th>Maximum non-thin-provisioned volume capacity in GiB (and TiB)</th>
<th>Maximum thin-provisioned volume capacity in GiB (and TiB)</th>
<th>Maximum MDisk capacity in GiB (and TiB)</th>
<th>Total storage capacity manageable per systema</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>2048 (2 TiB)</td>
<td>2000</td>
<td>2048 (2 TiB)</td>
<td>64 TiB</td>
</tr>
<tr>
<td>32</td>
<td>4096 (4 TiB)</td>
<td>4000</td>
<td>4096 (4 TiB)</td>
<td>128 TiB</td>
</tr>
<tr>
<td>64</td>
<td>8192 (8 TiB)</td>
<td>8000</td>
<td>8192 (8 TiB)</td>
<td>256 TiB</td>
</tr>
<tr>
<td>128</td>
<td>16,384 (16 TiB)</td>
<td>16,000</td>
<td>16,384 (16 TiB)</td>
<td>512 TiB</td>
</tr>
<tr>
<td>256</td>
<td>32,768 (32 TiB)</td>
<td>32,000</td>
<td>32,768 (32 TiB)</td>
<td>1 PiB</td>
</tr>
<tr>
<td>512</td>
<td>65,536 (64 TiB)</td>
<td>65,000</td>
<td>65,536 (64 TiB)</td>
<td>2 PiB</td>
</tr>
<tr>
<td>1024</td>
<td>131,072 (128 TiB)</td>
<td>130,000</td>
<td>131,072 (128 TiB)</td>
<td>4 PiB</td>
</tr>
<tr>
<td>2048</td>
<td>262,144 (256 TiB)</td>
<td>260,000</td>
<td>262,144 (256 TiB)</td>
<td>8 PiB</td>
</tr>
</tbody>
</table>
Implementing the IBM Storwize V7000 V7.4

Use the same extent size for all storage pools in a clustered system, which is a prerequisite for supporting volume migration 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 7, “Storage pools” on page 203.

For most clustered systems, a capacity of 1 PiB is sufficient. Use a value of 256 MB.

Default extent size: The GUI of IBM Storwize V7000 has a default extent size value of 256 MB when you define a new storage pool.

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

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

- They should 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 providing the MDisks must 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 need to check the distribution of the volumes’ extents in that storage pool.

Multitiered storage pool
A multitiered storage pool has a mix of MDisks with more than one type of disk tier attribute, for example, a storage pool containing a mix of generic_hdd AND generic_ssd MDisks.

A multitiered storage pool contains MDisks with different characteristics as opposed to the single-tiered storage pool. However, each tier should have MDisks of the same size and MDisks that provide the same number of extents.

A multitiered storage pool is used to enable automatic migration of extents between disk tiers using the IBM Storwize V7000 Easy Tier function, as described in Chapter 10, “Advanced features for storage efficiency” on page 335.

Figure 1-6 on page 17 shows these components.
1.5.9 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 mapped to it by IBM Storwize V7000. IBM Storwize V7000 translates this volume into several 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 the extents without impacting 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.

There are three types of volumes:

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

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

  Figure 1-7 on page 18 shows how a striped volume is allocated, assuming 10 extents are required.
Sequential
A sequential volume is where the extents are allocated one after the other, from one MDisk to the next MDisk (Figure 1-8).

Image mode
Image mode volumes are special volumes that have a direct relationship with one MDisk. They are used to migrate existing data into and out of the clustered system.
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 (Figure 1-9).

Some 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 add an MDisk containing data to a storage pool, any data on the MDisk is lost. Ensure that you create image mode volumes from MDisks that contain data before adding MDisks to the storage pools.

### 1.5.10 Thin-provisioned volumes

Volumes can be configured to either be thin provisioned or fully allocated. A thin-provisioned volume behaves regarding application reads and writes as though they were fully allocated. 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 reported to IBM Storwize V7000 and to the host servers.

The real capacity is used to store both 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 overallocated volumes, or it can be used with FlashCopy. Thin-provisioned volumes can be used with the mirrored volume feature also.
A thin-provisioned volume can be configured to autoexpand, which causes IBM Storwize V7000 to automatically expand the real capacity of a thin-provisioned volume as its real capacity is used. Autoexpand attempts to maintain a fixed amount of unused real capacity on the volume. This amount 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 with a zero contingency capacity goes offline as soon as it needs to expand. A volume with a nonzero contingency capacity stays online until it is used up.

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 warning capacity. When the used free capacity of the group exceeds the warning capacity, a warning is logged. For example, if a warning of 80% has been specified, the warning is logged when 20% of the free capacity remains.

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

### 1.5.11 Mirrored volumes

IBM Storwize V7000 provides a function that is called volume mirroring, which enables a volume to have two physical copies. Each volume copy can belong to a different storage pool and can be on different physical storage systems, and provides a high-availability solution.

When a host system issues a write to a mirrored volume, IBM Storwize V7000 writes the data to both copies. When a host system issues a read to a mirrored volume, IBM Storwize V7000 places it into the primary copy. If one of the mirrored volume copies is temporarily unavailable, the IBM Storwize V7000 automatically uses the alternative copy without any outage for the host system. When the mirrored volume copy is repaired, IBM Storwize V7000 resynchronizes the data.

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

The mirrored volume copy can be any type: image, striped, sequential, and thin provisioned or not. The two copies can be different volume types.

Using mirrored volumes can also assist with migrating volumes between storage pools that have different extent sizes and can provide a mechanism to migrate fully allocated volumes to thin-provisioned volumes without any host outages.

If needed, you can change the timeout value, latency, which prioritizes low host latency (default) or redundancy, which prioritizes redundancy (longer timeout).

**Unmirrored volumes:** An unmirrored volume can be migrated from one location to another by adding a second copy to the destination that you want, waiting for the two copies to synchronize, and then removing the original copy. This operation can be stopped at any time.
1.5.12 Easy Tier

IBM Easy Tier is a performance function that automatically migrates or moves extents off a volume to, or from, flash storage to HDD storage. Easy Tier monitors the host I/O activity and latency on the extent of all volumes with the Easy Tier function turned on in a multitiered storage pool over a 24-hour period. It then creates an extent migration plan based on this activity and then dynamically moves high activity or hot extents to a higher disk tier within the storage pool. It also moves extent activity that has dropped off or cooled from the high tiered MDisk back to lower tiered MDisk.

The Easy Tier function may be turned on or off at the storage pool and volume level. The Automatic Storage Pool balancing is an integrated part of Easy Tier engine and enabled by default on all pools. It does not require an Easy Tier license.

It is possible to demonstrate the potential benefit of Easy Tier in your environment before installing flash drives. By turning on the Easy Tier function for a single level storage pool and the Easy Tier function for the volumes within that pool, Easy Tier creates a migration report every 24 hours on the number of extents it would move if the pool was a multitiered pool. Easy Tier statistics measurement is enabled.

Using Easy Tier might make it more appropriate to use smaller storage pool extent sizes.

The usage statistics file can be offloaded from IBM Storwize V7000 nodes and then an IBM Storage Tier Advisor Tool (STAT) can be used to create a summary report. You can find the STAT tool at the following website:

http://www.ibm.com/support/docview.wss?uid=ssg1S4000935

Contact your IBM representative or IBM Business Partner for more information about the Storage Tier Advisor Tool. Easy Tier is described in more detail in Chapter 10, “Advanced features for storage efficiency” on page 335.

1.5.13 iSCSI

iSCSI is an alternative means of attaching hosts to the IBM Storwize V7000. All communications with back-end storage subsystems, and with other IBM Storwize V7000s, only occur through a Fibre Channel connection.

The iSCSI function is a software function that is provided by the IBM Storwize V7000 code, not hardware.

In the simplest terms, iSCSI allows the transport of SCSI commands and data over an Internet Protocol network, based on IP routers and Ethernet switches. iSCSI is a block-level protocol that encapsulates SCSI commands into TCP/IP packets and uses an existing IP network, instead of requiring expensive FC HBAs and a SAN fabric infrastructure.

A pure SCSI architecture is based on the client/server model. A client (for example, server or workstation) initiates read or write requests for data from a target server (for example, a data storage system).

Commands, which are sent by the client and processed by the server, are put into the Command Descriptor Block (CDB). The server runs a command, and completion is indicated by a special signal alert.

The major functions of iSCSI include encapsulation and the reliable delivery of CDB transactions between initiators and targets through the Internet Protocol network, especially over a potentially unreliable IP network.
The concepts of names and addresses have been 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 not only the iSCSI name of an iSCSI node, but also 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 V7000 node represents an iSCSI node and provides statically allocated IP addresses.

Each iSCSI node, that is, an initiator or target, has a unique iSCSI qualified name (IQN), which can have a size of up to 255 bytes. The IQN is formed according to the rules adopted for Internet nodes.

The iSCSI qualified name format is defined in RFC3720 and contains (in order) the following elements:

- The “iqn” string.
- A date code specifying the year and month in which the organization registered the domain or subdomain name used as the naming authority string.
- The organizational naming authority string, which consists of a valid, reversed domain or a subdomain name.
- Optionally, a colon (:), followed by a string of the assigning organization’s choosing, which must make each assigned iSCSI name unique.

For IBM Storwize V7000, the IQN for its iSCSI target is specified as follows:

```
 iqn.1986-03.com.ibm:2145.<clusternname>.<nodename>
```

On a Windows server, the IQN, that is, the name for the iSCSI initiator, can be defined as follows:

```
 iqn.1991-05.com.microsoft:<computer name>
```

The IQNs can be abbreviated by using a descriptive name, known as an alias. An alias can be assigned to an initiator or a target. The alias is independent of the name and does not need to be unique. Because it is not unique, the alias must be used in a purely informational way. It cannot be used to specify a target at login or used during authentication. Both targets and initiators can have aliases.

An iSCSI name provides the correct identification of an iSCSI device irrespective of its physical location. Remember, the IQN is an identifier, not an address.

Changing names: Before changing system or node names for an IBM Storwize V7000 clustered system that has servers connected to it using SCSI, be aware that because the system and node name are part of the IQN for the IBM Storwize V7000, you can lose access to your data by changing these names. The IBM Storwize V7000 GUI shows a specific warning, but the CLI does not.

The iSCSI session, which consists of a login phase and a full feature phase, is completed with a special command.
The login phase of the iSCSI is identical to the FC port login process (PLOGI). It is used to adjust various parameters between two network entities and to confirm the access rights of an initiator.

If the iSCSI login phase is completed successfully, the target confirms the login for the initiator; otherwise, the login is not confirmed and the TCP connection breaks.

As soon as the login is confirmed, the iSCSI session enters the full feature phase. If more than one TCP connection was established, iSCSI requires that each command/response pair goes through one TCP connection. Thus, each separate read or write command is carried out without the necessity to trace each request for passing separate flows. However, separate transactions can be delivered through separate TCP connections within one session.

For further details about configuring iSCSI, see Chapter 4, “Host configuration” on page 111.

### 1.5.14 Real-time Compression

The IBM Real-time Compression solution addresses the challenges listed in the previous section because it was designed for primary storage. Implementing Real-time Compression in Storwize V7000 or SAN Volume Controller offers the following benefits:

- **Compression for active primary data:** IBM Real-time Compression can be used with active primary data. Therefore, it supports workloads that are not candidates for compression in other solutions. The solution supports online compression of existing data. It allows storage administrators to regain free disk space in an existing storage system without requiring administrators and users to clean up or archive data. This configuration significantly enhances the value of existing storage assets, and the benefits to the business are immediate. The capital expense of upgrading or expanding the storage system is delayed.

- **Compression for replicated/mirrored data:** Remote volume copies can be compressed in addition to the volumes at the primary storage tier. This process reduces storage requirements in Metro Mirror and Global Mirror destination volumes also.

- **No changes to the existing environment are required:** IBM Real-time Compression is part of the storage system. It was designed for transparency so that it can be implemented without changes to applications, hosts, networks, fabrics, or external storage systems. The solution is not apparent to hosts, so users and applications continue to work as-is. Compression occurs within the Storwize V7000 or SAN Volume Controller system itself.

- **Overall savings in operational expenses:** More data is stored in a rack space, so fewer storage expansion enclosures are required to store a data set. This reduced rack space has the following benefits:
  - Reduced power and cooling requirements: More data is stored in a system, therefore requiring less power and cooling per gigabyte or used capacity.
  - Reduced software licensing for additional functions in the system: More data stored per enclosure reduces the overall spending on licensing.

- **Disk space savings are immediate:** The space reduction occurs when the host writes the data. This process is unlike other compression solutions in which some or all of the reduction is realized only after a post-process compression batch job is run.
1.5.15 IP replication

Starting with Version 7.2, one of the most important functions in the Storwize family is 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 FC-based mirroring is. All remote mirroring modes (Metro Mirror, Global Mirror, and Global Mirror with changed volumes) are supported.

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

IP replication includes Bridgeworks SANSlide network optimization technology and 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 utilization as low as 20% (based on IBM measurements). And this scenario gets worse the longer the latency. Bridgeworks SANSlide technology that is integrated with IBM Storwize family requires no separate appliances and so 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 utilization up to 3x so customers can deploy a less costly network infrastructure or take advantage of faster data transfer to speed replication cycles, improve remote data currency, and enjoy faster recovery.

1.6 Advanced copy services

IBM Storwize V7000 supports the following copy services:

- Synchronous remote copy
- Asynchronous remote copy
- FlashCopy

Starting with SAN Volume Controller V6.3, copy services functions are implemented within a single IBM Storwize V7000 or between multiple members of the IBM Storwize family. The Copy Services layer sits above and operates independently of the function or characteristics of the underlying disk subsystems used to provide storage resources to an IBM Storwize V7000. Figure 1-10 on page 25 shows an example of copy services with SAN Volume Controllers and IBM Storwize V7000 (they must be at Version 6.3 or later). The layer parameter of the clustered system properties is introduced in v6.3 and can be changed from storage (default) to replication if you need to make a new relationship between the IBM Storwize V7000 and SAN Volume Controller systems. You can change it by using only the chsystem command. In SAN Volume Controller systems, this parameter is fixed to replication and cannot be changed.
1.6.1 Synchronous or asynchronous remote copy

The general application of remote copy seeks to maintain two copies of data. Often, the two copies are separated by distance, but not always.

The remote copy can be maintained in one of two modes: synchronous or asynchronous.

With the IBM Storwize V7000, Metro Mirror and Global Mirror are the IBM branded terms for the functions that are synchronous remote copy and asynchronous remote copy.

Synchronous remote copy ensures that updates are committed at both the primary and the secondary before the application considers the updates complete; therefore, the secondary is fully up to date if it is needed in a failover. However, the application is fully exposed to the latency and bandwidth limitations of the communication link to the secondary. In a truly remote situation, this extra latency can have a significant adverse effect on application performance.

Special configuration guidelines exist for SAN fabrics that are used for data replication. You must consider the distance and available bandwidth of the intersite links (see Chapter 11, “Advanced Copy Services” on page 379).

A function of Global Mirror designed for low bandwidth has been introduced in IBM Storwize V7000. It uses change volumes that are associated with the primary and secondary volumes. These volumes are used to record changes to the remote copy volume, the FlashCopy relationship that exists between the secondary volume and the change volume, and between the primary volume and the change volume. This function is called Global Mirror cycling mode. Figure 1-11 on page 26 shows an example of this function where you can see the relationship between volumes and change volumes.
In asynchronous remote copy, the application acknowledges that the write is complete before
the write is committed at the secondary. Hence, on a failover, certain updates (data) might be
missing at the secondary. The application must have an external mechanism for recovering
the missing updates, if possible. This mechanism can involve user intervention. Recovery on
the secondary site involves starting the application on this recent “backup” and then rolling
forward or backward to the most recent commit point.

1.6.2 FlashCopy

FlashCopy makes a copy of a source volume on a target volume. The original content of
the target volume is lost. After the copy operation has started, the target volume has the
contents of the source volume as they existed at a single point in time. Although the copy
operation takes time, 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 a 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 V7000 also permits multiple target volumes to be FlashCopied 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 thin-provisioned volumes.

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 V7000 supports multiple targets and thus multiple
rollback points.

Most clients aim to integrate the FlashCopy feature for point in time copies and quick recovery
of their applications and databases. IBM Support is provided by Tivoli Storage FlashCopy
Manager, which is described at the following website:


You can read a detailed description about the FlashCopy copy services in Chapter 11,
“Advanced Copy Services” on page 379.
1.6.3 Copy Services configuration limits

Table 1-5 lists the Copy Services configuration limits. For the most up-to-date list of these limits, see the following website:

http://www.ibm.com/support/docview.wss?uid=ssg1S1004511

<table>
<thead>
<tr>
<th>Properties</th>
<th>Maximum number</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>Remote copy (Metro Mirror and Global Mirror)</td>
<td>4096</td>
<td>This can be any mix of Metro Mirror and Global Mirror</td>
</tr>
<tr>
<td>relationships per system</td>
<td></td>
<td>relationships.</td>
</tr>
<tr>
<td>Remote copy relationships per consistency group</td>
<td>-</td>
<td>No limit is imposed beyond the remote copy relationships per system limit</td>
</tr>
<tr>
<td>Remote copy consistency groups per system</td>
<td>256</td>
<td>-</td>
</tr>
<tr>
<td>Total Metro Mirror and Global Mirror</td>
<td>1024 TB</td>
<td>This limit is the total capacity for all master and auxiliary volumes in the I/O group.</td>
</tr>
<tr>
<td>relationships per I/O group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total number of Global Mirror with Change</td>
<td>256</td>
<td>-</td>
</tr>
<tr>
<td>Volumes relationships per system</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FlashCopy mappings per system</td>
<td>4096</td>
<td>-</td>
</tr>
<tr>
<td>FlashCopy targets per source</td>
<td>256</td>
<td>-</td>
</tr>
<tr>
<td>FlashCopy mappings per consistency group</td>
<td>512</td>
<td>-</td>
</tr>
<tr>
<td>FlashCopy consistency groups per system</td>
<td>127</td>
<td>-</td>
</tr>
<tr>
<td>Total FlashCopy volume capacity per I/O group</td>
<td>1024 TB</td>
<td>-</td>
</tr>
<tr>
<td>Cascaded FlashCopy maps</td>
<td>256</td>
<td>A volume can be the source of up to 256 FlashCopy maps.</td>
</tr>
</tbody>
</table>

1.7 Management and support tools

The IBM Storwize V7000 system can be managed through the native management software that runs in the hardware.

1.7.1 IBM Assist On-site and remote service

The IBM Assist On-site tool is a remote desktop-sharing solution that is offered through the IBM website. With it, the IBM service representative can remotely view your system to troubleshoot a problem.
You can maintain a chat session with the IBM service representative so that you can monitor this activity and either understand how to fix the problem yourself or allow the representative to fix it for you.

To use the IBM Assist On-site tool, the master console must be able to access the Internet. The following website provides further information about this tool:

http://www.ibm.com/support/assistonsite

When you access the website, you sign in and enter a code that the IBM service representative provides to you. This code is unique to each IBM Assist On-site session. A plug-in is downloaded on to your master console to connect you and your IBM service representative to the remote service session. The IBM Assist On-site tool contains several layers of security to protect your applications and your computers. The plug-in is removed after the next reboot.

You can also use security features to restrict access by the IBM service representative. Your IBM service representative can provide you with more detailed instructions for using the tool.

1.7.2 Event notifications

IBM Storwize V7000 can use Simple Network Management Protocol (SNMP) traps, syslog messages, and a 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.

Each event that IBM Storwize V7000 detects is assigned a notification type of Error, Warning, or Information. You can configure IBM Storwize V7000 to send each type of notification to specific recipients.

1.7.3 SNMP traps

Simple Network Management Protocol (SNMP) is a standard protocol for managing networks and exchanging messages. IBM Storwize V7000 can send SNMP messages that notify personnel about an event. You can use an SNMP manager to view the SNMP messages that IBM Storwize V7000 sends. You can use the management GUI or the IBM Storwize V7000 command-line interface 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 V7000. This file can be used with SNMP messages from all versions of the IBM Storwize V7000 Software.

1.7.4 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 either IPv4 or IPv6. IBM Storwize V7000 can send syslog messages that notify personnel about an event. IBM Storwize V7000 can transmit syslog messages in either expanded or concise format. You can use a syslog manager to view the syslog messages that IBM Storwize V7000 sends. IBM Storwize V7000 uses the User Datagram Protocol (UDP) to transmit the syslog message. You can use the management GUI or the IBM Storwize V7000 command-line interface to configure and modify your syslog settings.
1.7.5 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 Storwize V7000 is within the warranty period.

To send email, you must configure at least one SMTP server. You can specify as many as five additional SMTP servers for backup purposes. The SMTP server must accept the relaying of email from the IBM Storwize V7000 clustered system IP address. You can then use the management GUI or the IBM Storwize V7000 command-line interface 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 IBM Storwize V7000 command-line interface.

1.8 Useful Storwize V7000 websites

See the following IBM Storwize V7000 web pages for more information:

- Support page:
  http://ibm.co/1nyBrTn
- IBM Storwize V7000 Unified and Storwize V7000 Disk Systems:
- List of supported hardware:
  http://www.ibm.com/support/docview.wss?uid=ssg1S1004450
- Configuration Limits and Restrictions:
  http://www.ibm.com/support/docview.wss?uid=ssg1S1004511
- Documentation:
  http://ibm.co/1nxgxFF
- IBM Knowledge Center:

IBM Redbooks publications about IBM Storwize V7000 are available at the following website:
1.8.1 IBM Storwize V7000 learning videos on YouTube

Several helpful videos that describe the IBM Storwize V7000 system are available on YouTube at the links that are listed in Table 1-6.

<table>
<thead>
<tr>
<th>Description of the video</th>
<th>Web address</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBM Storwize V7000 Storage Virtualization Terminology overview</td>
<td><a href="http://www.youtube.com/watch?v=I2rzt3m2gP0">http://www.youtube.com/watch?v=I2rzt3m2gP0</a></td>
</tr>
<tr>
<td>IBM Storwize V7000 Interface Tour</td>
<td><a href="http://www.youtube.com/watch?v=FPbNRs9HacQ">http://www.youtube.com/watch?v=FPbNRs9HacQ</a></td>
</tr>
<tr>
<td>IBM Storwize V7000 Volume Management</td>
<td><a href="http://www.youtube.com/watch?v=YXeKqH8I9D9o">http://www.youtube.com/watch?v=YXeKqH8I9D9o</a></td>
</tr>
<tr>
<td>IBM Storwize V7000 Migration</td>
<td><a href="http://www.youtube.com/watch?v=dXxnUN6d974">http://www.youtube.com/watch?v=dXxnUN6d974</a></td>
</tr>
<tr>
<td>IBM Storwize V7000 Introduction to FlashCopy</td>
<td><a href="http://www.youtube.com/watch?v=MXWgGWjBzG4">http://www.youtube.com/watch?v=MXWgGWjBzG4</a></td>
</tr>
<tr>
<td>IBM Storwize V7000 Introduction to Remote Copy Part I: Creating Relationships and Consistency Groups</td>
<td><a href="http://www.youtube.com/watch?v=koFYm--gnEc">http://www.youtube.com/watch?v=koFYm--gnEc</a></td>
</tr>
<tr>
<td>IBM Storwize V7000 Introduction to Remote Copy Part II: Recovery and Return</td>
<td><a href="http://www.youtube.com/watch?v=kMOFFS3NYas">http://www.youtube.com/watch?v=kMOFFS3NYas</a></td>
</tr>
<tr>
<td>IBM Storwize V7000 Rapid Application Storage Solution: Tivoli Storage FlashCopy Manager</td>
<td><a href="http://www.youtube.com/watch?v=1zgEljRVDsc">http://www.youtube.com/watch?v=1zgEljRVDsc</a></td>
</tr>
<tr>
<td>IBM SAN Volume Controller and Storwize V7000 Performance Panel Sped-up! (HD)</td>
<td><a href="http://www.youtube.com/watch?v=7noC71tLkWs">http://www.youtube.com/watch?v=7noC71tLkWs</a></td>
</tr>
<tr>
<td>IBM Real-time Compression and Storwize V7000 and SAN Volume Controller demonstration</td>
<td><a href="http://www.youtube.com/watch?v=rgKj75kn2J0">http://www.youtube.com/watch?v=rgKj75kn2J0</a></td>
</tr>
<tr>
<td>Storage efficiency with IBM Easy Tier</td>
<td><a href="https://www.youtube.com/watch?v=-UEvEV61K90">https://www.youtube.com/watch?v=-UEvEV61K90</a></td>
</tr>
</tbody>
</table>
IBM Storwize V7000 Initial configuration

In this chapter, we discuss the following topics:

- Managing the IBM Storwize V7000 Gen2 and Gen1
  - Network Requirements
  - Prerequisites
- Storwize V7000 Gen2 and Gen1 initial configuration
  - How-to make the first connection to the IBM Storwize V7000 Gen2
  - System Setup Wizard
  - Secure Shell overview
  - Using IPv6

2.1 Network requirements for the IBM Storwize V7000 Gen1 and Gen2

To plan your installation, you need to consider the TCP/IP address requirements of the Storwize V7000 Gen2, and the requirements for the Storwize v7000 Gen2 to access other services. You must also plan the address allocation and the Ethernet router, gateway, and firewall configuration to provide the required access and network security.

Figure 2-1 on page 32 shows the TCP/IP ports and services that are used by the Storwize V7000 Gen2.
### 2.1.1 Prerequisites

Ensure that the Storwize V7000 Gen2 or Gen1 has been physically installed and that Ethernet and Fibre Channel (FC) connectivity has been correctly configured.

Before configuring the Storwize V7000 Gen2 and Gen1, ensure that the following information is available:

- **Licenses**
  
  The licenses indicate whether the client is permitted to use Easy tier, FlashCopy, External Virtualization, Remote Copy and Real-time Compression. For IPv4 addressing:
  
  - Cluster IPv4 address: This address includes one IP address for management.
  - Service IPv4 addresses: These addresses include two IPv4 addresses for the service interfaces.
  - IPv4 subnet mask.
  - Gateway IPv4 address.

- For IPv6 addressing:
  
  - Cluster IPv6 address: This address includes one address for management.
  - Service IPv6 addresses: These addresses include two IPv6 addresses for the service interface (one for each node).
  - IPv6 prefix.
  - Gateway IPv6 address.
2.2 Initial configuration of IBM Storwize V7000 Gen2

For our initial configuration, we are using the following hardware:

- 1 x IBM Storwize V7000 Gen2 model 2076-524.
- 2 x 4-port 8 Gbps Fibre Channel host interface card (one per node).
- 1 x IBM Storwize V7000 Expansion Enclosure.
- 2 x SAN Switches (for a redundant SAN fabric).

The first step is to connect a PC or notebook to the technician port on the rear of the V7000 Gen2 node. See Figure 2-2 for the technician port. The technician port provides a DHCP IP address V4, so you must ensure that your PC or notebook is configured for DHCP. The “default” IP address for a new node is 192.168.0.1. You can, however, also use a static IP, which should be set to 192.168.0.2 on your PC or notebook.

The Storwize V7000 Gen2 does not provide IPv6 IP addresses for the technician port.

When your PC or notebook is connected to the technician port, and you have validated that you have an IP v4 DCHP address, for example, 192.168.0.12 (first IP address the Storwize V7000 Gen2 node will assign), open a supported browser, which should automatically redirect you to 192.168.0.1 and the initial configuration of the cluster can start.

Figure 2-3 on page 34 shows the Log In and Legal Notice and will start the wizard that allows you to configure a new system.
Read the License Agreement and then click **Accept** which takes you to the login window, as shown in Figure 2-4.

You must type the default password for the account superuser, which is passw0rd (zero instead of o). When you click Log in, it takes you to the change password window. See Figure 2-5 on page 35.
You must type in a new password and confirm it. The password length is 6 - 63 characters, which cannot begin or end with a space. When you have typed in the password twice, click **Log in** again, and you will see the Welcome to System Setup window, as shown in Figure 2-6.

Click **Next**, and you can choose to give the cluster a new name, which we have given the system the name ITSO_V7000Gen2_2. See Figure 2-7 on page 36.
Click **Apply and Next** when you have changed the name to the wanted name of the system. The next steps are to set the time and date. See Figure 2-8.

---

**Figure 2-8  Setting date and time**
In this case, we set the date and time manually. Currently, you cannot choose to use the 24-hour clock, but that can be changed when the initial configuration is completed. However, it is highly recommended to use an NTP Server so you have a common time stamp for troubleshooting for all of your SAN/Storage devices.

Click **Apply and Next** and you will go to the Licensed Functions window, as shown in Figure 2-9.

![Licensed Functions window](image)

**Figure 2-9** Licensed Functions window

Enter the purchased amount of licenses (above are examples). Click **Apply and Next**, which takes you to the Encryption page, as shown in Figure 2-10 on page 38.
We selected “no” in this scenario. For details about encryption, refer to Chapter 12, “Encryption” on page 467.

Press **Next** and you are taken to the Detected Enclosures window, as shown in Figure 2-11 on page 39.
Simply click **Apply and Next**. This takes you to the configuration of the Email Event Notifications window, as shown in Figure 2-12.
You can choose not to set up Email Event Notifications, however this is not recommended. The next windows show how to set up Email Event Notifications.

**Note:** You need to have a valid SMTP server IP address available to complete this step.

On the first window that is shown in Figure 2-13, you have to set the System Location information.

Click **Next** and enter the contact details, as shown in Figure 2-14 on page 41.
Click **Apply and Next**, and you will now have to type the IP address of the email server. See Figure 2-15.
You can click the **Ping** button, which then verifies if there is network access to the email server (SMTP server).

Next, you choose whom the Email Notifications will be sent to. See Figure 2-16.

callhome@de.ibm.com is a default user, and cannot be deleted.

You can add more users to receive email notifications by clicking the **+** button, and you can select the wanted notifications that the defined users should receive, where the options are as follows:

- Errors
- Events
- Notifications
- Inventory

Click the **Apply and Next** button and you will be taken to the Configure Storage window, as shown in Figure 2-17 on page 43.
We have chosen No and will configure the storage when the initial setup is completed, which you can see how to beginning with Figure 2-21 on page 45.

Press Next and you are directed to the Summary window, as shown in Figure 2-18 on page 44.

Figure 2-17 Configure Storage window
Implementing the IBM Storwize V7000 V7.4

Click Finish, which completes the initial configuration and takes you to the System Overview, as shown in Figure 2-19.

We show how you can configure the internal storage on this system. This can also be performed by using the command-line interface.

Go to the Pools menu, as shown in Figure 2-20 on page 45.
Chapter 2. IBM Storwize V7000 Initial configuration

Click **Internal Storage** and then click **Configure Storage**, as shown in Figure 2-21.

Figure 2-21  Configure storage

Figure 2-22 on page 46 gives you the option to either choose:

- Use the recommended configuration
- Select a different configuration
Figure 2-22 Configure Internal Storage window

Figure 2-23 shows that we selected **Use the recommended configuration**.

Figure 2-23 Use the recommended configuration

When you press **Finish**, you return to the Internal Storage overview menu, as shown in Figure 2-24 on page 47.
As you can see, all drives are now configured according to the selected configuration.

This concludes the initial setup of the Storwize V7000 system. Now you can create volumes, hosts, and so on.

### 2.3 Initial configuration of IBM Storwize V7000 Gen1

For our initial configuration, we are using the following hardware:

- 1 x IBM Storwize V7000 Gen1 model 2076-124.
- 2 x 4-port 8 Gbps Fibre Channel host interface card (one per node).
- 2 x SAN Switches (for a redundant SAN fabric).

IBM Storwize V7000 Gen1 (2076-124) uses an initial setup that is contained within a USB key. The USB key is delivered with each storage system and has the initialization application, which is the `InitTool.exe` file. A system management IP address, the subnet mask, and the network gateway address are required. The initialization application creates a configuration file on the USB key.

The IBM Storwize V7000 starts the initial setup as soon as you plug in the USB key with the newly created file to the storage system.

**USB key:** If you are unable to find the official USB key that is supplied with IBM Storwize V7000, you can use any USB key that you have. Then, download and copy the package with the `InitTool.exe` application, from the IBM Storwize V7000 Support website at the following address, onto your USB stick:

```
http://www.ibm.com/storage/support/Storwize/V7000
```

The USB key contains the `InitTool.exe` file (Figure 2-25 on page 48).
To complete the initial setup using the USB key, follow these steps:

1. Plug the USB key into a Windows system and start the initialization tool. If the system is configured to autorun USB keys, the initialization tool starts automatically; otherwise, open My Computer and double-click the `InitTool.exe` file. The opening window of the tool is shown in Figure 2-26. After the tool is started, select **Initialize a new system using the USB Key** and click **Next**.
2. Enter the IPv4 or IPv6 address, subnet mask, and network gateway address, and then click **Next** (Figure 2-27).

![Figure 2-27 Initialization application: Network configuration](image)

Click **Finish** (Figure 2-28).

![Figure 2-28 Initialization application: Finalizing the initial setup](image)

The application creates a **satask.txt** file on the USB key (Figure 2-29).

![Figure 2-29 The satask.txt file created during the InitTool procedure](image)
The contents of this file is similar to the following command:

```bash
satask mkcluster -clusterip 192.168.70.123 -gw 192.168.70.1 -mask 255.255.255.0
```

3. Unplug the USB key from your Windows system and plug it into one of the IBM Storwize V7000 USB connector slots. The storage system automatically does these steps:

   a. Detect the key.
   b. Read the `satask.txt` file and run its contents.
   c. Delete the `satask.txt` file.
   d. Write the response to a new file called `satask_result.html` (Figure 2-30).

![Figure 2-30 The satask_result.html file created during the InitTool procedure](image)

Clustered system creation: While the clustered system is being created, the amber fault LED on the node canister flashes. When the amber fault LED stops flashing, remove the USB key from IBM Storwize V7000 and insert it in your system to check the results.

After this task completes successfully, the initial setup is done. The IBM Storwize V7000 is available for further configuration changes using the newly defined configuration address.

To complete the initial configuration see Figure 2-6 on page 35.

This concludes the initial setup of the IBM Storwize Gen2 and Gen1.

### 2.4 Secure Shell overview

Secure Shell (SSH) can be used to access the SAN Volume Controller with a command-line interface (CLI) like PuTTY. You can choose between password or SSH key authentication, or you can choose both password and SSH key authentication for the SAN Volume Controller CLI. We describe SSH in the following sections.

**Tip:** If you choose not to create an SSH key pair, you can still access the SAN Volume Controller cluster by using the SAN Volume Controller CLI, if you have a user password. You are authenticated through the user name and password.

The connection is secured by using a private key and a public key pair. Securing the connection includes the following steps:

1. A public key and a private key are generated together as a pair.
2. A public key is uploaded to the SSH server (SAN Volume Controller cluster).
3. A private key identifies the client. The private key is checked against the public key during the connection. The private key must be protected.
4. Also, the SSH server must identify itself with a specific host key.
5. If the client does not have that host key yet, it is added to a list of known hosts.
SSH is the communication vehicle between the management system (the System Storage Productivity Center or any workstation) and the SAN Volume Controller cluster.

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.

SSH keys are generated by the SSH client software. The SSH keys include a public key, which is uploaded and maintained by the cluster, and a private key that is kept private to the workstation that is running the SSH client. These keys authorize specific users to access the administrative and service functions on the cluster. Each key pair is associated with a user-defined ID string that can consist of up to 40 characters. Up to 100 keys can be stored on the cluster. 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 SAN Volume Controller clusters.

You must preinstall the freeware implementation of SSH-2 for Microsoft Windows (which is called PuTTY) on the System Storage Productivity Center or any other workstation. This software provides the SSH client function for users who are logged in to the SAN Volume Controller Console and who want to start the CLI to manage the SAN Volume Controller cluster.

### 2.4.1 Generating public and private SSH key pairs by using PuTTY

Complete the following steps to generate SSH keys on the SSH client system:

1. Start the PuTTY Key Generator to generate public and private SSH keys. From the client desktop, select **Start → Programs → PuTTY PuTTYgen**.

2. In the PuTTY Key Generator GUI window (see Figure 2-31 on page 52), complete the following steps to generate the keys:
   a. Select **SSH-2 RSA**.
   b. Leave the number of bits in a generated key value at 1024.
   c. Click **Generate**.
3. Move the cursor onto the blank area to generate the keys.

**To generate keys:** The blank area is the large blank rectangle on the GUI inside the section of the GUI labeled Key (see Figure 2-31 on page 52). Continue to move the mouse pointer over the blank area until the progress bar reaches the far right. This action generates random characters to create a unique key pair.

4. After the keys are generated, save them for later use by completing the following steps:
   a. Click **Save public key**, as shown in Figure 2-32 on page 53.
b. You are prompted for a name, for example, `pubkey`, and a location for the public key, for example, `C:\Support\Utils\PuTTY`. Click **Save**.

If another name and location are chosen, ensure that you maintain a record of the name and location. You must specify the name and location of this SSH public key in the steps that are described in 2.4.2, “Uploading the SSH public key to the SAN Volume Controller cluster” on page 54.

**Tip:** The PuTTY Key Generator saves the public key with no extension, by default. Use the string `pub` in naming the public key; for example, `pubkey`, to differentiate the SSH public key from the SSH private key easily.

c. In the PuTTY Key Generator window, click **Save private key**.
d. You are prompted with a warning message, as shown in Figure 2-33. Click Yes to save the private key without a passphrase.

![Figure 2-33 Saving the private key without a passphrase](image)

Figure 2-33   Saving the private key without a passphrase

e. When prompted, enter a name, for example, icat, and a location for the private key, for example, C:\Support.Utils\PuTTY. Click Save.

We suggest that you use the default name icat.ppk because this key was used for icat application authentication and must have this default name in SAN Volume Controller clusters that are running on versions before SAN Volume Controller 5.1.

| Private key extension: The PuTTY Key Generator saves the private key with the PPK extension. |

5. Close the PuTTY Key Generator GUI.

6. Browse to the directory, for example, C:\Support.Utils\PuTTY, where the private key was saved.

### 2.4.2 Uploading the SSH public key to the SAN Volume Controller cluster

After you create your SSH key pair, you must upload your SSH private key onto the SAN Volume Controller cluster. Complete the following steps:

1. From your browser, enter https://svcclusteripaddress/.

   Alternatively, from the GUI interface, you can go to the Access Management interface and select Users.

2. In the next window, as shown in Figure 2-34 on page 55, select Create User to create a user.
3. From the window to create a user, as shown in Figure 2-34 on page 56, enter the name (user ID) that you want to create and then enter the password twice. Select the access level that you want to assign to your user. The Security Administrator (SecurityAdmin) is the maximum access level. Select the location from which you want to upload the SSH Public Key file that you created for this user. Click Create.
You completed the user creation process and uploaded the users’ SSH public key that is paired later with the users’ private.ppk key, as described in 2.4.3, “Configuring the PuTTY session for the CLI” on page 56. Figure 2-38 on page 58 shows the successful upload of the SSH admin key.

You now completed the setup requirements for the SAN Volume Controller cluster by using the SAN Volume Controller cluster web interface.

### 2.4.3 Configuring the PuTTY session for the CLI

Before the CLI can be used, you must configure the PuTTY session by using the SSH keys that were generated in 2.4.1, “Generating public and private SSH key pairs by using PuTTY” on page 51, or by user name if you configured the user without an SSH key.

Complete the following steps to configure the PuTTY session on the SSH client system:

1. From the System Storage Productivity Center on a Microsoft Windows desktop, select Start → Programs → PuTTY PuTTY to open the PuTTY Configuration GUI window.

2. From the Category pane on the left in the PuTTY Configuration window (see Figure 2-36 on page 57), click Session if it is not selected.

**Tip:** The items that you select in the Category pane affect the content that appears in the right pane.
3. Under the “Specify the destination you want to connect to” section in the right pane, select **SSH**. Under the “Close window on exit” section, select **Only on clean exit**, which ensures that if any connection errors occur, they are displayed in the user’s window.

4. From the Category pane on the left, select **Connection → SSH** to display the PuTTY SSH connection configuration window, as shown in Figure 2-37.
5. In the right pane, for the Preferred SSH protocol version, select 2.

6. From the Category pane on the left side of the PuTTY Configuration window, select Connection → SSH → Auth.

7. As shown in Figure 2-38, in the “Private key file for authentication:” field under the Authentication parameters section in the right pane, browse to or enter the fully qualified directory path and file name of the SSH client private key file (for example, C:\Support\Utils\PuTTY\icat.PPK) that was created earlier.

You can skip the Connection → SSH → Auth part of the process if you created the user only with password authentication and no SSH key.

![PuTTY Configuration: Private key file location for authentication](image)

8. From the Category pane on the left side of the PuTTY Configuration window, click Session.

9. In the right pane, complete the following steps, as shown in Figure 2-39 on page 59:
   a. Under the “Load, save, or delete a stored session” section, select Default Settings, and then click Save.
   b. For the Host name (or IP address) field, enter the IP address of the SAN Volume Controller cluster.
   c. In the Saved Sessions field, enter a name (for example, SVC) to associate with this session.
   d. Click Save.
2.4.4 Starting the PuTTY CLI session

The PuTTY application is required for all CLI tasks. If it was closed for any reason, restart the session by completing the following steps:

1. From the SAN Volume Controller Console desktop, open the PuTTY application by selecting Start → Programs → PuTTY.

2. In the PuTTY Configuration window (see Figure 2-40 on page 60), select the session that was saved earlier (in our example, ITSO-SVC1) and click Load.

3. Click Open.
4. If this is the first time that you use the PuTTY application since you generated and uploaded the SSH key pair, a PuTTY Security Alert window with a prompt opens that warns that a mismatch exists between the private and public keys, as shown in Figure 2-41. Click Yes. The CLI starts.

5. As shown in Example 2-1, the private key that is used in this PuTTY session is now authenticated against the public key that was uploaded to the SAN Volume Controller cluster.

Example 2-1  Authenticating

Using username "admin".
Authenticating with public key "rsa-key-20100909"
IBM_2145:ITSO_SVC1:admin>
You completed the required tasks to configure the CLI for SAN Volume Controller administration from the SAN Volume Controller Console. You can close the PuTTY session.

2.4.5 Configuring SSH for AIX clients

To configure SSH for AIX® clients, complete the following steps:

**Note:** You must reach the SAN Volume Controller cluster IP address successfully by using the `ping` command from the AIX workstation from which cluster access is wanted.

1. OpenSSL must be installed for OpenSSH to work. Complete the following steps to install OpenSSH on the AIX client:
   a. You can obtain the installation images from the following websites:
      - [http://sourceforge.net/projects/openssh-aix](http://sourceforge.net/projects/openssh-aix)
   b. Follow the instructions carefully because OpenSSL must be installed before SSH is used.

2. Complete the following steps to generate an SSH key pair:
   a. Run the `cd` command to browse to the `/ssh` directory.
   b. Run the `ssh-keygen -t rsa` command. The following message is displayed:
      
      Generating public/private rsa key pair. Enter file in which to save the key (/ssh/id_rsa)
      
   c. Pressing Enter uses the default file that is shown in parentheses. Otherwise, enter a file name (for example, `aixkey`), and then press Enter. The following prompt is displayed:
      
      ![](Enter a passphrase (empty for no passphrase)
      
      d. When you use the CLI interactively, enter a passphrase because no other authentication exists when you are connecting through the CLI. After you enter the passphrase, press Enter. The following prompt is displayed:
      
      Enter same passphrase again:
      
      Enter the passphrase again. Press Enter.
      
   e. A message is displayed indicating that the key pair was created. The private key file has the name that was entered previously; for example, `aixkey`. The public key file has the name that was entered previously with an extension of `.pub`; for example, `aixkey.pub`.

**The use of a passphrase:** If you are generating an SSH keypair so that you can use the CLI interactively, use a passphrase so that you must authenticate whenever you connect to the cluster. It is possible to have a passphrase-protected key for scripted usage, but you must use the `expect` command or a similar command to have the passphrase parsed into the `ssh` command.
2.5 Using IPv6

You can use IPv4 or IPv6 in a dual-stack configuration. Migrating to (or from) IPv6 can be done remotely and is nondisruptive.

**Using IPv6:** To remotely access the SAN Volume Controller clusters that are running IPv6, you are required to run a supported web browser and have IPv6 configured on your local workstation.

2.5.1 Migrating a cluster from IPv4 to IPv6

As a prerequisite, enable and configure IPv6 on your local workstation. In our case, we configured an interface with IPv4 and IPv6 addresses on the System Storage Productivity Center, as shown in Example 2-2.

**Example 2-2  Output of ipconfig on the System Storage Productivity Center**

```plaintext
C:\Documents and Settings\Administrator>ipconfig
Windows IP Configuration
Ethernet adapter IPv6:
Connection-specific DNS Suffix . :
   IP Address. . . . . . . . . . . . : 10.0.1.115
   Subnet Mask . . . . . . . . . . . : 255.255.255.0
   IP Address. . . . . . . . . . . . : 2001:610::115
   IP Address. . . . . . . . . . . . : fe80::214:5eff:fecd:9352%5
   Default Gateway . . . . . . . . . :
```

To update a cluster, complete the following steps:

1. Select **Configuration** → **Network**, as shown in Figure 2-42.
2. Select **Management IP Addresses**. Click port 1 of one of the nodes, as shown in Figure 2-43.

![Management IP Addresses window](image)

**Figure 2-43 Management IP Addresses window**

3. In the window that is shown in Figure 2-44, complete the following steps:
   a. Select **Show IPv6**.
   b. Enter an IPv6 address in the IP Address field.
   c. Enter an IPv6 gateway in the Gateway field.
   d. Enter an IPv6 prefix in the Subnet Mask/Prefix field. The Prefix field can have a value of 0 - 127.
   e. Click **OK**.

![Ethernet Port 1 (Primary)](image)

**Figure 2-44 Modifying the IP addresses: Adding IPv6 addresses**

4. A confirmation window opens, as shown in Figure 2-45 on page 64. Click **Apply Changes**.
5. The Change Management task is started on the server, as shown in Figure 2-46. Click **Close** when the task completes.

6. Test the IPv6 connectivity by using the **ping** command from a cmd.exe session on your local workstation, as shown in Example 2-3 on page 65.
Example 2-3 Testing IPv6 connectivity to the SAN Volume Controller cluster

C:\Documents and Settings\Administrator>ping

Pinging 2001:610::119 from 2001:610::115 with 32 bytes of data:

Reply from 2001:610::119: time=3ms
Reply from 2001:610::119: time<1ms
Reply from 2001:610::119: time<1ms
Reply from 2001:610::119: time<1ms

Ping statistics for 2001:610::119:
   Packets: Sent = 4, Received = 4, Lost = 0 (0% loss),
Approximate round-trip times in milliseconds:
   Minimum = 0ms, Maximum = 3ms, Average = 0ms

7. Test the IPv6 connectivity to the cluster by using a compatible IPv6 and SAN Volume Controller web browser on your local workstation.

8. Remove the IPv4 address in the SAN Volume Controller GUI that is accessing the same windows, as shown in Figure 2-44 on page 63. Validate this change by clicking OK.

2.5.2 Migrating a cluster from IPv6 to IPv4

The process of migrating a cluster from IPv6 to IPv4 is identical to the process that we described in 2.5.1, “Migrating a cluster from IPv4 to IPv6” on page 62, except that you add IPv4 addresses and remove the IPv6 addresses.
Graphical user interface

This chapter provides an overview of the graphical user interface (GUI) of the IBM Storwize V7000 and shows how to use the navigation tools.
3.1 Overview of IBM Storwize V7000 management software

A GUI can simplify storage management and provide a fast and more efficient management tool. It is loosely based on the IBM System Storage XIV® software and has a similar look and feel.

Starting with version 7.4, the GUI layout has been enhanced and rebuilt in terms of a unified and common look as with other IBM products such as IBM Flash Systems, IBM SAN Volume Controller, and IBM System Storage XIV. It offers similar ways of maintenance and operational procedures using GUI.

3.1.1 Access to the storage management software

To log on to the management software, type 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 (Figure 3-1).

We suggest that each user who operates the IBM Storwize V7000 has an account that is not shared with someone else. The default user accounts should be disabled for remote access or the passwords changed from the default and known only to the system owner, or kept secured for emergency purposes only. Doing this helps to identify the personnel working on the device and to track all the important changes in the systems. The Superuser account should be used only for initial configuration.
After successful login, the IBM Storwize V7000 System panel opens (Figure 3-2).

![Figure 3-2 Getting started with the IBM Storwize V7000 Overview panel](image)

The IBM Storwize V7000 System panel is an important user interface. Throughout this chapter, we refer to it as the *V7000 System panel* or just the *System panel*. In the remaining chapters, we do not explain how to access it each time.

### 3.1.2 System panel layout

This System panel has three main sections for navigating through the management tool. Figure 3-3 shows those three areas.

![Figure 3-3 Main areas](image)
The left side of the window shows seven function icons. We refer to them collectively as a dynamic menu. The dynamic menu includes these function icons:

- Monitoring menu
- Pools menu
- Volumes menu
- Hosts menu
- Copy Services menu
- Access menu
- Settings menu

The middle of the window shows a component model of the existing configuration. Hovering the mouse cursor over each component and its part highlights that part and provide a pop-up menu with a description and identify important parameters and functions of this element. In order to see the rear of the component, you can dynamically rotate them as in a typical 360° view. A right-click on a component or its part opens its context menu with actions, normally available from the Dynamic menu on the left.

The bottom of the window shows three status indicators. Clicking any of them provides more detailed information about the existing configuration of the IBM Storwize V7000 solution. Click any of these function icons to expand them and minimize them as required or to switch between different information, for example virtual or allocated storage. In case of any error or warning, those three indicators are extended by the Status alerts icon in the lower right corner (Figure 3-3 on page 69).

3.1.3 Navigation

Navigating in the management tool is simple. You can hover the cursor over one of the seven function icons to emphasize that icon and display a menu of options. You can move the cursor to an option and click it. Figure 3-4 shows how to access, for example, the Volumes option.
Figure 3-5 shows a list of the IBM Storwize V7000 Software function icons and the associated menu options.

### 3.1.4 Multiple selection

With the improved management tool, you can select multiple items by using a combination of Shift or Ctrl keys. To select multiple items in a display, click the first item, press and hold the Shift key, and click the last item in the list you require. All rows between those two items are selected and highlighted in light blue (Figure 3-6).

![IBM Storwize V7000 System panel and its menu options](image)

**Figure 3-6** Multiple selections 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 the Ctrl key, and click the other items that you need.
3.1.5 Status indicators area

The status indicators area at the bottom of the System panel (Figure 3-7) show a brief status about the IBM Storwize V7000 storage system. The four indicators are as follows (from left to right):

- Capacity overview
- Running tasks
- Health status
- Status alerts

Click each icon or hover the mouse cursor over the icon to see more details about status of the system or task indicated by the icon.

Capacity overview

The leftmost indicator shows information about the overall physical capacity (the initial amount of storage that was allocated) and virtual capacity (thin provisioned storage; volume size is dynamically changed as data grows or shrinks, but you still see a fixed capacity). Click the indicator to switch between physical and virtual capacity (Figure 3-8).

![Figure 3-8 Storage capacity overview](image)

**Important:** Since version 7.4, the capacity units use the binary prefixes defined by International Electrotechnical Commission (IEC) representing a multiplication by 1024 with symbols GiB (gibibyte), TiB (tibibyte), PiB (pebibyte).

Running tasks

The Running tasks indicator provides the information about jobs initiated by the user or system and that typically run for a longer period, such array initialization, synchronization, or rebuilt. Hover the mouse cursor over the circular arrows icon on the indicator to open the context menu, which lists all running tasks (Figure 3-9).

![Figure 3-9 Running tasks status](image)
Health status
The Health status indicator warns administrators about critical errors or alerts that need greater attention. In a critical alert, the green color of the indicator becomes red. You can then click the X icon (Figure 3-10) to open the context menu to see the status alerts. If the status is healthy (green, not red), no context menu is available. Yellow color indicates warning only.

![Figure 3-10 Health status of the system](image)

Status alerts
The indicator for status alerts is at the lower right corner only in case of an important error or warning, which is not yet acknowledged or fixed in the Event log. Hover the mouse cursor over the icon to see the list of important system messages (Figure 3-11). After all alerts are cleared from the system, the Status alerts icon automatically disappears from the view.

![Figure 3-11 IBM Storwize V7000 status alerts](image)

Help
Another useful interface feature is integrated help. You can access help for some fields and objects by hovering the mouse cursor over the question mark icon, next to the field or object (Figure 3-12). Panel-specific help is available by clicking Need Help or by using the Help link in the upper-right corner of the GUI.

![Figure 3-12 Access to panel-specific help](image)
3.2 Overview panel

In version 7.4, the welcome panel of the GUI has changed from the well known former Overview panel to the new System panel as shown in Figure 3-2 on page 69. Clicking the Overview button in the upper-right corner of the System panel opens a modified Overview panel with similar meaning as in previous versions of the software (Figure 3-13).

![Figure 3-13 Opening Overview panel](image)

See 3.1.2, “System panel layout” on page 69 to understand the structure of the panel and how to navigate to various system components and manage it more efficient and quickly.

3.3 Monitoring menu

Hover the cursor over the Monitoring function icon to open the Monitoring menu (Figure 3-14 on page 75). The Monitoring menu offers these navigation directions:

- System: Opens the general overview of your IBM Storwize V7000 system including the depiction of all devices in a rack and the storage capacity. See 3.3.1, “System overview” on page 75.
- Events: Tracks all informational, warning, and error messages that occurred in the system. You can apply various filters to sort them according to your needs or export them to the external comma-separated values (CSV) file. See 3.3.3, “Events” on page 79.
- Performance: Reports the general system statistics with regards to the processor (CPU) utilization, host and internal interfaces, volumes, and managed disks (MDisks). You can switch between MBps or input/output operations per second (IOPS). See 3.3.4, “Performance” on page 80.

The formerly known option System Details has been in V7.4 integrated into the device overview on the general System panel, available after login or when clicking the option System from the Monitoring menu. Details are shown in 3.3.2, “System Details option” on page 77.
In the following sections, we describe each option on the Monitoring menu. Detailed procedures of how to best use some of these items are described in other chapters (for example, Events in Chapter 13, “RAS, monitoring, and troubleshooting” on page 495).

### 3.3.1 System overview

The System option on the Monitoring menu provides the general overview about your IBM Storwize V7000 system including the depiction of all devices in a rack and the allocated or physical storage capacity. When thin-provisioned volumes are enabled, the virtual capacity is also shown by hovering your mouse over the capacity indicator. See Figure 3-15 on page 76 for details.
When you click a specific component in an enclosure, a pop-up window indicates the details of disk drives in the unit. By right-clicking and selecting Properties, you see detailed technical parameters such as capacity, interface, rotation speed, and the drive status (online or offline).

See Figure 3-16 for details of Drive 4 in an enclosure.

In an environment with multiple IBM Storwize V7000 systems, you can easily navigate the onsite personnel or technician to the correct device by enabling the identification LED on the front panel. Click Identify in the pop-up window shown in Figure 3-16 and wait for the confirmation of the technician that the device in the data center was properly identified.
After confirmation, click **Turn LED Off** (Figure 3-17).

Alternatively, you can use the IBM Storwize V7000 command-line interface (CLI) to get the same results. Type the following sequence of commands:

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

Each system shown in the Dynamic system view in the middle of a System panel, can be rotated by 180° to see its rear side. Click the rotation arrow in the right-bottom corner of the device as illustrated in Figure 3-18.

### 3.3.2 System Details option

The System Details option has been removed from the Monitoring menu, however its modified information is still available directly from the System panel. It provides the extended level of parameters and technical details related to the system, including the integration of
each element with an overall system configuration. Right-click the wanted enclosure and navigate to the Properties option to obtain detailed information.

Figure 3-19   System details

The output is shown in Figure 3-20. Using this context menu, you can also power off the machine (without the option for remote start), remove the node or enclosure from the system, or for example, list all volumes associated with the system.

Figure 3-20   Enclosure technical details
In addition, from the System panel, you can get the overview of three important statuses and parameters of FC and SAS ports and system fans (Figure 3-21).

By choosing for example **Fibre Channel Ports**, you can see the list and status of available FC ports with their worldwide port name (WWPN) (Figure 3-22).

### 3.3.3 Events

The Events option, selected from the Monitoring menu, 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 comma-separated values (CSV) file. A CSV file can be created from the information that is here. Figure 3-23 on page 80 shows the display after the Event log is selected from the menu.
The exact procedures for how to work with the Event log and how to run various fix procedures using the Events option are described in Chapter 13, “RAS, monitoring, and troubleshooting” on page 495.

3.3.4 Performance

The Performance panel reports the general system statistics related to processor (CPU) utilization, host and internal interfaces, volumes, and MDisks. You can switch between MBps or IOPS or even drill down the statistics to the node level. This might be useful when comparing the performance of each node in the system if problems exist after a node failover occurs. See Figure 3-24.
The performance statistics in the GUI show, by default, the latest five minutes of data. To see details of each sample, click the graph and select the time stamp, as shown in Figure 3-25.

![Sample details](image)

Charts above represent five minutes of the data stream. For in-depth storage monitoring and performance statistics of your IBM Storwize V7000 system with historical data, use the IBM Tivoli Storage Productivity Center for Disk or IBM Virtual Storage Center.
3.4 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. Volumes are created from the extents that are available in the pool. You can add MDisks to a storage pool at any time, either to increase the number of extents that are available for new volume copies or to expand existing volume copies.

Hover the cursor over the Pools function icon to display the Pools menu options (Figure 3-26).

The Pools menu has the following options:

- Volumes by Pool: Applies the high-level filter, listing all defined volumes per pool. It also provides a capacity overview, valid only for a specific, selected pool. This view is excellent when planning a migration of a volume to another pool, so you have a common overview of all pools and their associated volumes. Unused volumes are not listed.

- Internal Storage: Provides the overview about all disk drives that are installed in the IBM Storwize V7000 system including its enclosures. You can filter based on disk type and capacity and see also unused volumes not assigned to any pool.

- External Storage: Shows all pools and their volumes created from the systems connected to the IBM Storwize V7000 externally and integrated into the system repository. It does not show any internal pools or volumes. This is also called external virtualization.

- MDisks by Pools: Provides the list of all managed disks either 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 migrating data from an old external storage to the IBM Storwize V7000 non-disruptively to the hosts.

Child pools

Before version 7.4, the disk space of a storage pool is provided from MDisks, so the capacity of a storage pool depends on the MDisks’ capacity. Creating or splitting a storage pool is not
possible and a user cannot freely make a storage pool with a particular capacity they want. A child pool is a new logical object that is created from a physical storage pool and provides most of the functions that mdiskgrps have (for example, volume creation), but the user can specify the capacity of the child pool at creation. A child pool is an object similar to a storage pool, and can be used interchangeably with storage pool. It supports volume copy and migration. However, there are some limitations and restrictions:

- Maximum capacity cannot exceed parent pool's size
- Capacity can be allocated at creation (thick) or flexible (thin)
- Parent storage pool must be always specified; child pool does not own any MDisks
- Child pools can be created only from CLI, but are visible in GUI
- Maximum number of child pools in one parent pool is 127
- Restricted to migrate image-mode volume to child pool
- Volume extents cannot be migrated out of the child pool
- Forbidden to shrink capacity below its real capacity

Child pools are visible in the GUI, however any manipulation of their parameters can be done only by using the CLI. Volumes defined in a child pool can be managed from the GUI in the standard way. The list of child pools is reachable from menu option MDisks by Pools by a right-click on the parent pool. From the context menu, choose option Child Pools as shown in Figure 3-27.

![Access to child pools](image)

Figure 3-27 Access to child pools

The opened dialog box provides the list of child pools that are associated with selected parent pool (Figure 3-28). Child pool parameters are locked from editing in a GUI.

![List of child pools](image)

Figure 3-28 List of child pools
3.4.1 Volumes by pool

This menu option lists all defined volumes, sorted by their pool assignment (Figure 3-29). Unassigned volumes are not visible in this menu. By using this menu you can, for example, create volumes, map or unmap volumes to and from hosts, migrate volumes to another pool, and rename, shrink, or expand volumes.

![Figure 3-29 Listing volumes by pool](image)

In addition, you can rename the pool or choose a different icon that represents this pool (Figure 3-30). Choose the appropriate icon.

![Figure 3-30 Changing the pool icon](image)

When the pools are defined and the volumes are assigned, the pool can have one of the following four operational status values:

**Online**

The storage pool is online and available. All the MDisks in the storage pool are available.

**Degraded path**

One or more nodes in the clustered system cannot access all the MDisks in the group. A degraded path state is most likely the result of incorrect configuration of either the storage system or the Fibre Channel fabric. However, hardware failures in the storage system, Fibre Channel fabric, or node can also be a contributing factor to this state.

**Degraded ports**

One or more 1220 errors have been logged against the MDisks in the storage pool. The 1220 error indicates that the remote Fibre Channel
port has been 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 **Troubleshooting Recommended Actions** in the management GUI. This action displays a list of unfixed errors that are currently in the event log. For these unfixed errors, select the error name to begin a guided maintenance procedure to resolve them. Errors are listed in descending order with the highest priority error listed first. Resolve 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:** Using this view, volumes from child pools are shown the same way as volumes from standard pools. Relationships between the child and parent pools are not visible.

### 3.4.2 Internal storage

Click the **Internal Storage** option in the Pools menu to open a window similar to Figure 3-31. 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 3-31: Internal storage window](image)

You can configure a storage array and assign it to the pool by first clicking **Configure Storage** (circled in Figure 3-31). A wizard opens where you choose the recommended array level or select a different one. The IBM Storwize V7000 suggests configuring all available drives, based on recommended values for the RAID level and drive class. However, you can select a different configuration from presets by using the wizard.
Figure 3-32 illustrates how to define the array. To understand the complete procedure for how to configure internal storage, see Chapter 7, “Storage pools” on page 203.

For the best performance of your system, always configure and use all disk drives in the IBM Storwize V7000, but with respect to a recommended number of spare drives for the array rebuild process, in case of drive failure.

3.4.3 External storage

Clicking the **External Storage** option opens the window as shown in Figure 3-33. It provides the list of all externally connected (storage area network (SAN)-attached) disk systems to the IBM Storwize V7000.

When the new external storage system is properly zoned to the IBM Storwize V7000, run the Detect MDisks procedure either from the Actions menu in the table header or by right-clicking any of the existing MDisks in the list (Figure 3-33).
A new storage controller (external storage system) is listed automatically when the SAN zoning configuration is done, but typically without detected disk drives (Figure 3-34).

![Figure 3-34  Automatically detected new external storage system](image)

By right-clicking a newly detected storage system, you can rename the controller’s default name to reflect the real type of the storage device, in our case the IBM System Storage DS3400. We suggest using a simple naming convention, in our case DS3400 (Figure 3-35).

![Figure 3-35  Renaming detected system](image)

When the new external storage system has the appropriate name, detect all disks configured on that external storage, in our case IBM System Storage DS3400. Because this IBM entry-level disk storage subsystem is withdrawn from marketing, it shows a good example of a migration scenario from an old storage system to your new IBM Storwize V7000 system. You can also detect new MDisks from the CLI by using the `svctask detectmdisk` command or just `detectmdisk`. Figure 3-36 shows details about detected managed disks.

![Figure 3-36  Newly discovered managed disks](image)

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.

**Attention:** The MDisks are not physical disk drives, but storage arrays configured on external systems.

If you add a managed disk that contains existing data to a managed disk group, you will lose the data that it contains. The `image` mode is the only mode that preserves its data.
3.4.4 MDisks by pools

This option on the Pools menu provides the list of all managed disks and arrays of disks either internally or externally connected and associated with one of the defined pools. It also lists all unassigned MDisks separately. An example of the panel is shown in Figure 3-37.

![Figure 3-37 List of managed disks sorted within pools](image)

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 (clicking the Name header of the table). Right-click a specific disk to open a context menu, where you can assign selected unmanaged disks to the pool.

From the same panel, you can define a new storage pool by using the Create Pool button in the upper left corner of the table (highlighted in Figure 3-37). The wizard window opens and you need to specify 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 just skip this task and add MDisks later.

For detailed instructions about how to define pools and assign MDisks, follow the guidance in Chapter 7, “Storage pools” on page 203.

**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.

3.4.5 System migration

Migrating data from older storage systems to the IBM Storwize V7000 storage system allows applications to benefit from the new features, such as Easy Tier, an intuitive management GUI, and advanced storage replication functions that better support applications.

To migrate existing data, use the IBM Storwize V7000 storage migration wizard to guide you through the procedure. This wizard is available by selecting Pools → System Migration as shown in Figure 3-38 on page 89.

The migration of external volumes into the IBM Storwize V7000 system is one of the key benefits and features of external storage virtualization provided by this product. Therefore, we dedicate the whole chapter to this topic. See Chapter 6, “Migration wizard” on page 189 for detailed steps of the migration process.
3.5 Volumes menu

A volume is a logical disk that the system presents to the attached host. Application servers access volumes, not MDisks or drives. To keep a volume accessible even when an MDisk on which it depends has become unavailable, a mirrored copy can be added to a selected volume. Each volume can have a maximum of two copies. Each volume copy is created from a set of extents in a storage pool.

The three types of volumes are striped, sequential, and image. Apart of standard Generic volumes you can also configure volumes with specific advanced features:

- **Compressed**: This is a special type of volume where data is compressed as it is written to disk, saving additional space. To use the compression function, you must obtain the IBM Real-time Compression license.

- **Mirrored**: 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.

- **Thin-provisioned**: When you create a volume, you can designate it as thin-provisioned. A thin-provisioned volume has a virtual capacity and a real capacity. Virtual capacity is the volume storage capacity that is available to a host. Real capacity is the storage capacity that is allocated to a volume copy from a storage pool. In a fully allocated volume, the virtual capacity and real capacity are the same. In a thin-provisioned volume, however, the virtual capacity can be much larger than the real capacity. Finally, **Thin-Mirrored** volumes combine the characteristics of mirrored and thin-provisioned volumes.

Hover the cursor over the Volumes function icon to display the Volumes menu options (Figure 3-39 on page 90).
3.5.1 All volumes

Select **Volumes** as shown in Figure 3-39. A list of all defined volumes, alphabetically sorted by the volume name (by default), is displayed. 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** (highlighted in Figure 3-39). The wizard opens and the list of volume types is displayed (Figure 3-40).

The description of each type of volume and procedures for how to effectively create these volumes are described in Chapter 5, “Basic volume configuration” on page 145.

In addition to the new volume creation, other direct volume functions are available, such as mapping and unmapping volumes to hosts; renaming, shrinking, or expanding existing volumes; migration to a different pool; or defining a volume copy. All of these tasks are available when you select a specific volume and click **Actions** (Figure 3-41 on page 91).
When moving a volume to another I/O group (a different IBM Storwize V7000 system or pair of IBM SAN Volume Controller nodes), be sure that the appropriate host zoning is in place. The target host must have access to both systems: source and target.

### 3.5.2 Volumes by pool

This menu is identical to the one described in 3.4.1, “Volumes by pool” on page 84 in a section Pools. See that section for details.

### 3.5.3 Volumes by host

Click **Volumes by Host** to open the window shown in Figure 3-42. This window shows the volumes that have been mapped to a given host. You can perform the same actions with volumes as in all previous views, either by clicking **Actions** or by using the context menu that opens after you right-click a specific volume. See Figure 3-42 for details.

In addition, you can rename the host by clicking its name and directly typing the new name into the context line (Figure 3-43 on page 92).
In a SAN environment, host systems are application servers that access data from the storage controllers that are connected to the SAN. Hosts that are running in several operating environments can connect to the storage using the IBM Storwize V7000. Host connections to the IBM Storwize V7000 are either Small Computer System Interface (SCSI) over the Fibre Channel SAN, Fibre Channel over Ethernet SAN, or Internet Small Computer System Interface (iSCSI) over an Ethernet network.

You can use several tools to manage hosts, including the management GUI, the CLI, and specialized utilities for working with host bus adapters. To work with hosts in the management GUI, select **Hosts**. When you hover the cursor over the Host function icon, the Hosts menu opens, which has the following options:

- Hosts
- Ports by Host
- Host Mappings
- Volumes by Host

### 3.6.1 Hosts

This option provides an overview about all hosts that are connected (zoned) to the system, detected, and configured to be ready for storage allocation. This overview shows the name of the host as defined in the IBM Storwize V7000, the type of the host, its status of accessibility, the number of ports used for host mapping, and whether host mapping is active. From the same panel, you can create a new host, rename a host, delete a host, or modify a host mapping. The output of the menu selection is shown in Figure 3-44.

For example, when you click **Add Host** in a panel header, a wizard opens where you define either an iSCSI or Fibre Channel host (Figure 3-45 on page 93).
To rename multiple hosts in a single step, mark all hosts that you want by using the Control or Shift key, right-click, and then from the opened context menu select **Rename**. The window shown in Figure 3-46 opens.

Some of the actions described here are available from different menus. For example, you can select **Volumes** and its option **Volumes by Hosts**, where you can also rename hosts. This is one of the advantages of the enhanced, redesigned IBM Storwize V7000 management GUI.

### 3.6.2 Ports by host

Click **Ports by Hosts** to open the panel shown in Figure 3-47 on page 94. This panel lists the Fibre Channel and iSCSI ports that are assigned to a particular host.
This overview shows hosts with active, inactive, or degraded ports. You can delete or add a port, or modify its characteristics. Also in this panel, you can create a new host or rename the existing one.

To do any of the tasks shown in Figure 3-48, click **Actions** and select a menu item.

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

Click Host Mappings to open the window shown in Figure 3-49. It lists the host name, SCSI identifier, volume name, and volume identifier for all mapped volumes.

![Figure 3-49 Host mappings](image)

From this window you can view the host properties, obtain the list of mapped volumes, or work with port definitions. Right-click the specific host and select Properties (Host) from the opened context menu. A window similar to the one in Figure 3-50 opens.

![Figure 3-50 Host properties](image)

With enabled details, you can modify host name, host type, I/O group assignment, or iSCSI CHAP Secret by clicking Edit and then the Save button as highlighted in Figure 3-50.

3.6.4 Volumes by host

This option is identical to the option that is available in the dynamic menu Volumes. For a description, see 3.5.3, “Volumes by host” on page 91.
3.7 Copy Services

The IBM Storwize V7000 Copy Services are part of the IBM Replication Family Services available on IBM SAN Volume Controller, IBM Storwize V7000, and IBM Storwize V5000 products. It consists of the following functions:

- IBM FlashCopy
- Metro Mirror and Global Mirror
- Global Mirror with Changed Volumes
- Volume Mirroring function (Volume Copy)
- Stretched Cluster (formerly Split Cluster) volume mirroring

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

3.7.1 FlashCopy

IBM FlashCopy is a function that you use to create a point-in-time copy of one of your IBM Storwize V7000 disks. This might be helpful when doing backups or application testing. These copies can be cascaded one on another, read from, written to, and even reversed.

FlashCopy snapshots are able to conserve storage, if needed, by being space-efficient copies (instead of full copies) that record only items that have changed from the originals. Select FlashCopy from the dynamic menu to open a panel similar to what is shown in Figure 3-51.

![FlashCopy operations](image)

Figure 3-51 FlashCopy operations

If you need to create a FlashCopy of an additional volume, right-click the volume and the list of available functions is listed. From here, you can do tasks such as initiate a new snapshot, and clone or back up a volume.
Clicking the volume name opens the window shown in Figure 3-52. From here, you can click the tabs at the top of the window to display additional information, such as the hosts that the volume or FlashCopy volume is mapped to and its dependant MDisks.

![Figure 3-52 FlashCopy volume details](image)

### 3.7.2 Consistency Group

FlashCopy Consistency Groups can be used to create a consistent point-in-time copy across multiple volumes, and even across multiple managed storage systems, thus managing the consistency of dependent writes.

Click Consistency Group to open the window shown in Figure 3-53. Here, 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 3-53 FlashCopy Consistency Group window](image)

When any FlashCopy Consistency Group is available, either empty or with existing relations, you can move an existing relation to that group. Right-click a relation and select **Move to Consistency Group** as shown in Figure 3-54 on page 98.

Other actions on the same context menu include remove from consistency group, start (resume) or stop that FlashCopy operation, rename a target volume or FlashCopy mapping, and delete a mapping.
Figure 3-54  Moving relation to the Consistency Group

From the drop-down menu, select the appropriate group (in our case the only one available) and confirm the selection (Figure 3-55).

Figure 3-55  Assigning the Consistency Group

The result of the operation is similar to the result shown in Figure 3-56.

Figure 3-56  Consistency groups

3.7.3 FlashCopy mappings

To create a new FlashCopy mapping, click Create FlashCopy Mapping (shown in Figure 3-57 on page 99) 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 will fail.

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, thus enabling multiple volumes to be copied at the same time. This 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 reside on separate disks.

If any 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.

The example of the wizard for FlashCopy mapping creation is shown in Figure 3-58. Select source and target volumes from the drop-down menu of the wizard.

You can select the Snapshot (copy-on-write), Clone (replica of the volume without impact 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.
3.7.4 Remote Copy

Click **Remote Copy** to open the window shown in Figure 3-59. 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.

![Remote Copy window](Image)

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

Metro Mirror makes *synchronous* copies, which means that the original write operations are not considered complete until the write operation to the destination disk has been confirmed. The distance between your two sites is usually determined by how much latency your applications can handle.

Global Mirror makes *asynchronous* copies of your disk. This means that 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 greatly reduces the latency 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 have the most recent changes committed to the local disk.

Global Mirror with Changed Volumes can be best described as “Continuous Remote FlashCopy.” If you use this feature, the IBM Storwize V7000 essentially takes periodic FlashCopies of a disk and writes them 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.
3.7.5 Partnerships

Click **Partnerships** to open the window shown in Figure 3-60. You can use this window to set up a new partnership or delete an existing partnership with another IBM Storwize V7000 system for the purposes of remote mirroring.

![Figure 3-60 Creating the partnership](image)

From this window, you can also set the background copy rate. This rate specifies the bandwidth, in megabytes per second (MBps), that is used by the background copy process between the clusters (Figure 3-61).

![Figure 3-61 Partnership properties](image)
3.8 Access menu

The Access menu has two options: Users (for user management) and Audit Log.

3.8.1 Users

Figure 3-62 shows the Users panel. From here, you can create and delete new users, change and remove passwords, and add and remove Secure Shell (SSH) keys.

Click **Create User** to open the panel shown in Figure 3-63. Use this panel to specify the name and password of the user, and load the SSH key (if the SSH key has been generated). Starting with SAN Volume Controller V6.3, the SSH key is not required for CLI access, and you can choose to use either SSH or a password for CLI authentication.
3.8.2 Audit Log menu

Click **Audit Log** to open the window shown in Figure 3-64. The cluster maintains an audit log of successfully executed commands, indicating which users performed particular actions at certain times.

![Audit Log entries](image)

**Figure 3-64 Audit Log entries**

You can filter audit log records by date or within a specific time frame, defined either by minutes, hours, or days (Figure 3-65).

![Filtering the records](image)

**Figure 3-65 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
### 3.9 Settings menu

The settings menu provides various options to adjust your system parameters according to your needs. You can configure these options:

- Notifications
- Network
- Security (Remote authentication)
- System
- Support
- GUI preferences

These options are described in the following sections. Details about how to use each of them are presented in Chapter 13, “RAS, monitoring, and troubleshooting” on page 495.

#### 3.9.1 Notifications

It is important to correct any issues reported by your IBM Storwize V7000 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 just events that require immediate action.

You can use email, SNMP, or Syslog types of notifications. If your system is within warranty, or you have a hardware maintenance agreement, configure your IBM Storwize V7000 system to send email events to IBM directly, if an issue that requires hardware replacement is detected. This mechanism is called *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. The configuration window for email notifications is shown in Figure 3-66.

![Figure 3-66   Email event notifications](image)

The procedure for how to enable email notifications is described in Chapter 13, “RAS, monitoring, and troubleshooting” on page 495.
3.9.2 Network

Click Network to open the window shown in Figure 3-67. From here, you can update the network configuration, set up iSCSI definitions, and view information about the Fibre Channel connections.

![Network window](image)

Figure 3-67   Network window

When you click Fibre Channel Connectivity (Figure 3-68), useful information is displayed. In this example, we click Hosts from the menu and then select to display the details for one specific host, Hurricane, from the list of host systems. Other options that are available from the menu include displaying Fibre Channel details for all devices, for clusters, for nodes, for storage systems, or for hosts.

![Fibre Channel connectivity](image)

Figure 3-68   Fibre Channel connectivity

3.9.3 Security

Security and its directory services allow the user to remotely authenticate to the IBM Storwize V7000 without a need for a local account. This means that when you log on, you authenticate with your domain user ID and password instead of a locally created user ID and password.
The benefits of remote authentication are as follows:

- You do not have to configure every user on every SAN Volume Controller or Storwize V7000. If you have multiple machines, you can more efficiently set up authentication.
- When commands are executed on the SAN Volume Controller or Storwize V7000, the audit log will show the domain user name that issued that command, instead of a local user name, or worse just superuser (that is, 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 just remove access at the domain controller, meaning that orphan accounts will not remain on your storage equipment.

The access panel to configure remote authentication is shown in Figure 3-69.

![Figure 3-69 Configuring remote authentication](image)

The detailed steps to configure remote logon are described at the following addresses:

3.9.4 System

The System window provides three key options:

- Set system date and time
- Manage licenses
- Upgrade software

The System window opens (Figure 3-70) when you select System from the Settings menu.

![General menu](image)

Figure 3-70  General menu

You can also update the license information for specific features as shown in Figure 3-71.

![Licensing options](image)

Figure 3-71  Licensing options

To upgrade your IBM Storwize V7000 software, use the procedure that is described in Chapter 13, “RAS, monitoring, and troubleshooting” on page 495.
3.9.5 Support

The support menu provides access to features that help IBM to analyze the critical and uncertain issues with the system, based on downloaded and provided support packages. Figure 3-72 shows how to download support logs.

![Figure 3-72 Downloading support packages](image)

More details are available in Chapter 13, “RAS, monitoring, and troubleshooting” on page 495.

3.9.6 GUI preferences

Using this menu, you can configure the appearance and behavior of the GUI. Click GUI Preferences in the Settings option of the Dynamic menu. The layout of the dialog window and available options are shown in Figure 3-73 on page 109.
GUI Preferences

You can configure preferences for the management GUI.

- Refresh GUI Objects
- Restore Default Browser Preferences

Information Center

Web Address:

- Save
- Cancel
- Restore System Default

Extent Size:

- Allow extent size selection during pool creation.

Accessibility:

- Enable low graphics mode.
Host configuration

This chapter describes how to use the IBM Storwize V7000 GUI to create hosts, and how to prepare a host to access the volumes that are created. (Volume creation is described in Chapter 5, “Basic volume configuration” on page 145).
4.1 Host attachment overview

A host system is an open-systems computer that is connected to the switch through a Fibre Channel (FC) or an Internet Small Computer System Interface (iSCSI) interface.

This chapter describes the following topics:

► Preparing the host operating system:
  – Windows
    • FC
    • iSCSI
  – VMware
    • FC
    • iSCSI

► Creating hosts using the Storwize V7000 graphical user interface (GUI)
  – Creating FC hosts
  – Creating iSCSI hosts

In this chapter, we assume that your hosts are connected to your FC or IP network and you have completed the steps described in Chapter 2, “IBM Storwize V7000 Initial configuration” on page 31. 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 is connected to both canisters. This setup assures four paths for failover and failback purposes.

Before mapping the newly created volumes on the host of your choice, a little preparation goes a long way towards ease of use and reliability. There are several steps required on a host in preparation for mapping new IBM Storwize V7000 volumes to the host. Use the IBM System Storage Interoperation Center (SSIC) to check which code levels are supported to attach your host to your storage. SSIC is a web tool that checks the interoperation of host, storage, switches, and multipathing drivers. It can be found at the following address:

http://ibm.com/systems/support/storage/ssic/interoperability.wss

The complete support matrix is listed in the IBM Storwize V7000 Supported Hardware List, Device Driver, Firmware, and Recommended Software Levels V7.4 document, which is available at the following address:

http://www.ibm.com/support/docview.wss?uid=ssg1S1004941

This chapter focuses on Windows and VMware. If you must attach any other hosts, for example, IBM AIX, Linux, or an Apple system, you can find the required information at the following address:

4.2 Preparing the host operating system

In this section, we describe how to prepare Windows and VMware hosts for attachment to an IBM Storwize V7000, using either Fibre Channel or iSCSI to connect.

4.2.1 Windows 2008 R2 and 2012 R2: Preparing for Fibre Channel attachment

Complete the following steps to prepare a Windows 2008 R2 or Windows 2012 R2 host to connect to an IBM Storwize V7000 using Fibre Channel:

- Make sure that the latest OS service pack and test fixes are applied to your server.
- Use the latest firmware and driver levels on your host system.
- Install host bus adapter (HBA) or HBAs on the Windows server using the latest basic input/output system (BIOS) and drivers.
- Connect the FC Host Adapter ports to the switches.
- Configure the switches (zoning).
- Configure the HBA for hosts running Windows.
- Set the Windows timeout value.
- 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 will now be maintained from Qlogic. So check the Qlogic Web pages to find appropriate support for your Brocade Adapters. Also check and update the BIOS (firmware) level of the HBA using the manufacturer's provided tools. Check the readme file to see if there are Windows registry parameters that should be set for the HBA driver.

The latest supported levels are at the following addresses:

- For Windows 2008 R2:
- For Windows 2012 R2:
  http://www.ibm.com/support/docview.wss?uid=ssg1S1004941#_Win2012FC

Configuring Brocade HBAs for Windows

This section applies to Windows hosts that have Brocade HBAs installed. After installing the device driver and firmware, you must configure the HBAs. To perform this task, either use the Brocade HCM software or reboot 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)
- Queue depth: 4

Configuring QLogic HBAs for Windows

This section applies to Windows hosts that have QLogic HBAs installed.
After installing the device driver and firmware, you must configure the HBAs. To perform this task, either use the QCC software or reboot 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

**Configuring Emulex HBAs for Windows**

This section applies to Windows hosts that have Emulex HBAs installed.

After installing the device driver and firmware, you must configure the HBAs. To perform this task, either use the Emulex HBAnyware software or reboot 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 should be set to 60 seconds. To verify this setting, complete the following steps:

1. Click **Start** → **Run**. Or open a Power Shell window.
2. In the dialog box, type `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 necessary, change the value to 60 (Figure 4-1).

![Figure 4-1 Windows timeout value](image)

**Installing the multipathing software**

Microsoft Multipath Input/Output (MPIO) solutions work with device-specific modules (DSMs) written by vendors, but the MPIO driver package does not, by itself, form a complete solution. This joint solution allows the storage vendors to design device-specific solutions that are tightly integrated with the Windows operating system. MPIO is not shipped 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, and it allows the use of multipath in the SAN infrastructure during the boot process for SAN Boot hosts.
To ensure correct multipathing with IBM Storwize V7000, 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:
   
   http://www.ibm.com/support/docview.wss?uid=ssg1S7001350#WindowsSDDDSM

2. Extract the package to your hard disk drive and run `setup.exe` to install SDDDSM. A command prompt window opens (Figure 4-2). Confirm the installation by entering `Yes`.

   ![Figure 4-2 SDDDSM setup](image)

3. During the setup, SDDDSM will also check if an older version is installed and if you want to upgrade to the actual version.

4. After the setup completes, you are prompted to restart the system. Confirm this restart by typing `yes` and pressing Enter (Figure 4-3).

   ![Figure 4-3 Answer yes to restart the host](image)

You have now successfully installed IBM SDDDSM. To check if IBM SDDDSM is installed correctly, see the following chapter for Windows 2008 R2 and Windows 2012 R2.

**Windows 2008 R2**

To check the installed driver version, select `Start → All Programs → Subsystem Device Driver DSM → Subsystem Device Driver DSM`. A command prompt opens; run `datapath query version` to determine the version currently installed (Example 4-1 on page 116) for this Windows 2008 R2 host.
Example 4-1  datapath query version

C:\Program Files\IBM\SDDDSMD>datapath.exe query version
IBM SDDDSM Version 2.4.3.4-4
Microsoft MPIO Version 6.1.7601.17514

This command can also be used to determine the worldwide port names (WWPNs) of the host. Run datapath query wwpn (Example 4-2) and note the WWPNs of your host, as you need them later.

Example 4-2  datapath query wwpn

C:\Program Files\IBM\SDDDSMD>datapath.exe query wwpn
<table>
<thead>
<tr>
<th>Adapter Name</th>
<th>PortWWN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scsi Port 7</td>
<td>100000051EC76B89</td>
</tr>
<tr>
<td>Scsi Port 7</td>
<td>100000051EC76B8A</td>
</tr>
</tbody>
</table>

If you need more detailed information about SDDDSM, see *Multipath Subsystem Device Driver User’s Guide*, GC52-1309.

Windows 2012 R2

To check the installed driver version, first select the Windows start button in the lower left corner. See Figure 4-4.

![Figure 4-4  Windows 2012 R2 start button](image)

After this, expand the view, pressing the downward arrow shown in Figure 4-5 on page 117.
Search for the section *Subsystem Device Driver DSM*. See Figure 4-6.

Press the Subsystem Device Driver DSM command-line button. See Figure 4-7.

A command prompt opens; run `datapath query version` to determine the version currently installed. See Figure 4-8 on page 118.
Implementing the IBM Storwize V7000 V7.4

Figure 4-8  Datapath query version

See Example 4-3 for Windows 2012 R2 host.

Example 4-3  datapath query version

Microsoft Windows [Version 6.3.9600]
(c) 2013 Microsoft Corporation. All rights reserved.

C:\Program Files\IBM\SDDDSM>datapath query version
IBM SDDDSM Version 2.4.3.5-3
Microsoft MPIO Version 6.3.9600.16384
C:\ProgramFiles\IBM\SDDDSM>

This command can also be used to determine the WWPNs of the host. Run datapath query wwpn (Example 4-4) and note the WWPNs of your host, as you need them later.

Example 4-4  datapath query wwpn

C:\Program Files\IBM\SDDDSM>datapath query wwpn
Adapter Name    PortWWN
Scsi Port  7    100000051EC76B89
Scsi Port  7    100000051EC76B8A

If you need more detailed information about SDDDSM, see Multipath Subsystem Device Driver User’s Guide, GC52-1309.

A PDF Version of the Multipath Subsystem Device Driver User’s Guide can be found here:

https://ibm.biz/BdEqeZ

The Windows host has been prepared to connect to the IBM Storwize V7000 and you know the WWPNs of the host. The next step is to configure a host object for the WWPNs by using the IBM Storwize V7000 GUI. This task is explained in 4.3.1, “Creating Fibre Channel hosts” on page 133.

SAN Boot hosts are beyond the intended scope of this book. For more information about that topic, search for SAN Boot at the following address:

4.2.2 Windows 2008 R2 and Windows 2012 R2: Preparing for iSCSI attachment

In Windows 2008 R2 and 2012, the Microsoft iSCSI software initiator is preinstalled. Enter iscsi in the search field of the Windows 2008 R2 start menu (Figure 4-9) and click iSCSI Initiator.

![Figure 4-9 Windows 2008 R2 iSCSI Initiator](image)

For Windows 2012 R2, go to the all programs menu and enter iSCSI into the search field at the top of the window. See Figure 4-10.

![Figure 4-10 iSCSI Initiator Windows 2012 R2](image)

Confirm the automatic start of the iSCSI Service (Figure 4-11 on page 120).
Figure 4-11  Automatic start of the iSCSI service

The iSCSI Configuration window opens. Select the **Configuration** tab (Figure 4-12). Write down the initiator name of your Windows host, as you need it later.

![iSCSI Initiator Properties window](image)

Figure 4-12  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. More details are available at the IBM Storwize V7000 IBM Knowledge Center at the following address:

Setting the Windows registry keys
Modify the system registry as follows so that your iSCSI operations are more reliable:

1. In the search field of the Windows Start menu, type regedit and click regedit.exe.

2. 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 necessary, change the value to 120.

3. 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 necessary, 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 necessary, 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 necessary, change the value to 60.

6. Reboot your host for these changes to take effect.

These basic steps are to prepare a Windows 2008 R2 or Windows 2012 R2 host for iSCSI attachment. To configure the IBM Storwize V7000 for iSCSI connections, see 4.3.2, “Creating iSCSI hosts” on page 137.

4.2.3 VMware ESXi: Preparing for Fibre Channel attachment

Complete the following steps to prepare a VMware ESXi host to connect to an IBM Storwize V7000 using Fibre Channel:

1. Install HBA or HBAs on the ESXi server.
2. Make sure that the latest firmware levels are applied on your host system.
3. Update and configure the HBA for hosts 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 additional drivers if required.

Downloading and installing the supported firmware
Install the latest firmware levels to your host server. For the HBAs, check the IBM Storwize V7000 Supported Hardware List, Device Driver, Firmware, and Recommended Software Levels V7.4 document for VMware at the following address:
http://www.ibm.com/support/docview.wss?uid=ssg1S1004941#_VMwareFC

Download the latest supported HBA firmware for your configuration and apply it to your system. Some HBAs and especially the new CNA Adapters require an additional driver to be
loaded into ESXi. Check the VMware Compatibility Guide to see if there are any requirements for your configuration by going to the following address:

http://www.vmware.com/resources/compatibility/search.php

Configuring Brocade HBAs for VMware ESXi
This section applies to ESXi hosts that have Brocade HBAs installed. After installing the firmware, load the default settings of all your adapters installed on the host system and make sure that the Adapter BIOS is disabled, unless you are using SAN Boot.

Configuring QLogic HBAs for VMware ESXi
This section applies to ESXi hosts that have QLogic HBAs installed. After installing the firmware, you must configure the HBAs. To perform this task, either use the QLogic Converged Console, QCC software or the 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 (only enable it for debugging)
  – RIO Operation Mode: 0
  – Interrupt Delay Timer: 0

The QLogic Converged Console (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) are available. A VMware vCenter plug-in is also available. You can find the QCC for Windows here:

https://ibm.biz/BdEfsj

Configuring Emulex HBAs for VMware ESXi
This section applies to ESXi hosts that have Emulex HBAs installed. After installing the firmware, load the default settings of all your adapters installed on the host system and make sure that the Adapter BIOS is disabled, unless you are using SAN Boot.
VMware ESXi installation
Install your VMware ESXi server and load any additional drivers and patches if required. If you are not familiar with the procedure, see the installation guide at this address:


After you complete your ESXi installation, connect to your ESXi Server using the vSphere client and go to the Configuration tab, click Storage Adapters, and scroll down to your FC HBAs (Figure 4-13). Note the WWPNs of the installed adapters for later use.

![Figure 4-13  Show WWPNs in VMware ESXi](image)

VMware ESXi multipathing
The ESXi server has its own multipathing software. You do not need to install a multipathing driver, either on the ESXi server or on the guest operating systems. The ESXi multipathing policy supports three operating modes:

- Round Robin
- Fixed
- Most Recently Used (MRU)

The IBM Storwize V7000 is an active/active storage device. Since VMware ESX 4.0 and later, the suggested multipathing policy is Round Robin. Round Robin performs static load balancing for I/O. If you do not want to have 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 attaching IBM Storwize V7000 LUNs to the ESXi host (see Chapter 5, “Basic volume configuration” on page 145 for information). 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, discovered at system boot 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 is the default policy for logical unit numbers (LUNs) presented from an Active/Passive array. ESXi/ESX does not return to the previous path if, or when, it returns; it remains on the working path until it, for any reason, fails.
After all these steps are completed, the ESXi host is prepared to connect to the IBM Storwize V7000. Go to 4.3.1, “Creating Fibre Channel hosts” on page 133 to create the ESXi FC host in the IBM Storwize V7000 GUI.
4.2.4 VMware ESXi: Preparing for iSCSI attachment

This section describes how to enable iSCSI on VMware ESXi hosts. We focus on vSphere (ESX 4.0 and later) because the complete iSCSI stack has been rewritten in this level, offers improved performance, and supports useful features, such as jumbo frames and TCP Segmentation Offload. We focus on the basic ESXi iSCSI setup; more detailed information is provided in the VMware vSphere Documentation Center, which is available at the following address:

http://pubs.vmware.com/vsphere-55/index.jsp

Complete the following steps to prepare a VMware ESXi host to connect to an IBM Storwize V7000 using iSCSI:

1. Ensure that the latest firmware levels are applied on your host system.
2. Install VMware ESXi and load additional drivers if required.
3. Connect the ESXi server to your network. You should 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, and you only have to enable it. To enable it, complete the following steps:

1. Connect to your ESXi server using the vSphere Client, go to Configuration, and select Networking (Figure 4-14).

![Figure 4-14 Select VMware networking](image)

2. Click Add Networking to start the Add Network Wizard (Figure 4-15 on page 126). Select VMkernel and click Next.
3. Select one or more network interfaces that you want to use for iSCSI traffic and click Next (Figure 4-16).
4. Enter a meaningful Network Label and click **Next** (Figure 4-17).

![Figure 4-17 VMware: Enter a Network Label](image1.png)

5. Enter an IP address for your iSCSI network. You should use a dedicated network for iSCSI traffic (Figure 4-18).

![Figure 4-18 VMware: Enter iSCSI Network IP](image2.png)
6. Click **Finish** to complete the setup. See Figure 4-19.

![Image](image1.png)

**Figure 4-19** VMware: finish iSCSI setup

7. Check whether an iSCSI Software Adapter is available. Select for this **Storage Adapters** in the Hardware section. See Figure 4-20.

![Image](image2.png)

**Figure 4-20** VMware select new iSCSI Software Adapter

8. Press **Add** to add a new iSCSI Software Adapter. See Figure 4-21.

![Image](image3.png)

**Figure 4-21** VMware: Press Add button
9. The following window opens. See Figure 4-22.

![VMware: Add Software Adapter](image)

Figure 4-22 VMware: Add Software Adapter

10. After pressing **OK**, a message occurs to configure the adapter after it has been added. See Figure 4-23.

![VMware: Configure iSCSI Software Adapter after completion](image)

Figure 4-23 VMware: Configure iSCSI Software Adapter after completion

11. A new iSCSI Adapter is added to the Storage Adapter menu. See Figure 4-24.

![VMware: New iSCSI Software Adapter](image)

Figure 4-24 VMware: New iSCSI Software Adapter

12. Select **Storage Adapters** and scroll to iSCSI Software Adapter (Figure 4-25). Highlight it and click **Properties**.

![VMware: iSCSI Software Adapter](image)

Figure 4-25 VMware: iSCSI Software Adapter
13. The iSCSI Software Adapter Properties window opens. Figure 4-26 shows that the initiator is disabled by default. To change this setting, click **Configure**.

![Figure 4-26   VMware: iSCSI Software Adapter properties](image)

14. Select the **Enabled** check box and click **OK** (Figure 4-27).

![Figure 4-27   VMware: Select the Enabled check box](image)
15. The VMware ESX iSCSI initiator is now successfully enabled (Figure 4-28). Note the initiator name for later use.

![Figure 4-28 VMware iSCSI Initiator enabled](image)

Your VMware ESXi host is now prepared to connect to the IBM Storwize V7000. Go to 4.3.2, “Creating iSCSI hosts” on page 137 to create the ESXi iSCSI host using the IBM Storwize V7000 GUI.
4.3 Creating hosts using the GUI

This section describes how to create Fibre Channel and iSCSI hosts using the IBM Storwize V7000 GUI. We assume that the hosts are prepared for attachment, as described in 4.2, “Preparing the host operating system” on page 113, and that you know the host WWPNs and their iSCSI initiator names.

To create a host, complete the following steps:

1. Open the host configuration window by clicking **Hosts** (Figure 4-29).

2. To create a host, click **Add Host** to start the wizard (Figure 4-30).

3. After pressing **Add Host**, the host selection menu opens as shown in Figure 4-31 on page 133.

If you want to create a Fibre Channel host, continue with 4.3.1, “Creating Fibre Channel hosts” on page 133. To create iSCSI hosts, go to 4.3.2, “Creating iSCSI hosts” on page 137.

3. After pressing **Add Host**, the host selection menu opens as shown in Figure 4-31 on page 133.
4.3.1 Creating Fibre Channel hosts

To create Fibre Channel hosts, complete the following steps:

1. Click **Fibre Channel Host** (Figure 4-32). The Fibre Channel configuration wizard opens (Figure 4-32).

2. Enter a host name and click the **Fibre Channel Ports** drop-down menu to get a list of all known WWPNs (Figure 4-33 on page 134).
The IBM Storwize V7000 has the host port WWPNs available if you prepared the hosts, as described in 4.2, “Preparing the host operating system” on page 113. 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 and repeat the scanning.

3. Select the WWPN for your host and click **Add Port to List** (Figure 4-34).
4. Add all ports that belong to the host (Figure 4-35).

Creating offline hosts: If you want to create hosts that are offline, or not connected at the moment, it is also possible to enter the WWPNs manually. Type them into the Fibre Channel Ports Box and add them to the list also.
5. If you are creating an HP/UX or TPGS host, select the **Advanced** check box, and more options appear (Figure 4-36). Select your host type.

![Add Host Window](image)

*Figure 4-36  Create Host: Advanced Settings*

6. Click **Add Host** and the wizard creates the host (Figure 4-37 on page 137).
7. Click **Close** to return to the host window. Repeat these steps for all of your Fibre Channel hosts. Figure 4-38 shows the **All Hosts** window after creating a second host.

After you complete the adding of Fibre Channel hosts, go to Chapter 5, “Basic volume configuration” on page 145 to create volumes and map them to the created hosts.

### 4.3.2 Creating iSCSI hosts

To create iSCSI hosts, complete the following steps:

1. Click **iSCSI Host** (Figure 4-39 on page 138) and the iSCSI configuration wizard opens (Figure 4-39 on page 138).
2. Enter a host name, type the iSCSI initiator name into the iSCSI Ports box, and click **Add Port to List**. If you want to add several initiator names to one host, repeat this step.
3. If you are connecting an HP/UX or TPGS host, select the **Advanced** check box (Figure 4-40) and select the correct host type.

*Figure 4-40  Create an iSCSI host: Advanced Settings*
4. Click **Add Host** and the wizard completes (Figure 4-41). Click **Close**.

![Figure 4-41   Add an iSCSI host: Complete](image)

Figure 4-41  Add an iSCSI host: Complete
5. Repeat these steps for every iSCSI host that you want to create. Figure 4-42 shows the All Hosts window after creating two Fibre Channel and two iSCSI hosts.

The iSCSI hosts are now configured on the IBM Storwize V7000. To provide connectivity, the iSCSI Ethernet ports also must be configured. Complete the following steps to enable iSCSI connectivity:

1. Switch to the Configuration window and select **Network** (Figure 4-43).
2. Select **iSCSI** and the iSCSI Configuration window opens (Figure 4-44).

![iSCSI Configuration window](image)

**Figure 4-44**  iSCSI Configuration window

3. The system waits until you apply the changes that you made. Press **Apply Changes**. All changes are applied, as shown in Figure 4-45.

![Applying all iSCSI changes](image)

**Figure 4-45**  Applying all iSCSI changes

In the configuration, you have an overview of all the iSCSI settings for the IBM Storwize V7000. You can configure iSCSI Alias, iSNS Addresses, and Chap Authentication Configuration on this window, and the iSCSI IP address, which we also edit in the basic setup.
4. Click **Ethernet Ports** to enter the iSCSI IP address (Figure 4-46). Repeat this step for each port that you want to use for iSCSI traffic.

![Figure 4-46 Enter an iSCSI IP address](image)

5. After you enter the IP address for each port, click **Modify** to enable the configuration.

6. After the changes are successfully applied, click **Close** (Figure 4-47).

![Figure 4-47 iSCSI IP successfully modified](image)
7. Under **Actions**, you can check if all Hosts you want are iSCSI enabled. See Figure 4-48.

![Figure 4-48  Action menu to modify iSCSI hosts](image1)

8. By default all iSCSI hosts are enabled. (Figure 4-49).

![Figure 4-49  iSCSI host enabled](image2)

The IBM Storwize V7000 is now configured and ready for iSCSI use. Note the initiator names of your storage canisters (Figure 4-44 on page 142) because you need them later. To create volumes and map them to a host, go to Chapter 5, “Basic volume configuration” on page 145.
Basic volume configuration

This chapter describes how to use the IBM Storwize V7000 to create a volume and map a volume to a host. A volume is a logical disk on the IBM Storwize V7000 that is provisioned out of a storage pool and is recognized by a host with an identifier UID field and a parameter list.

The first part of the chapter describes how to create volumes and map them to defined hosts.

The second part of this chapter covers how to discover those volumes (5.1, “Provisioning storage from IBM Storwize V7000 to hosts” on page 146). After you finish this chapter, your basic configuration is done and you are able to store data on the IBM Storwize V7000.

Advanced host and volume administration, such as volume migration, creating volume copies, and so on, is described in Chapter 8, “Advanced host and volume administration” on page 247.
5.1 Provisioning storage from IBM Storwize V7000 to hosts

This section describes the setup process and shows how to create volumes and make them accessible from the host. The described steps for completing the basic setup in your environment are as follows:

1. Define new volumes in IBM Storwize V7000.
2. Map volumes to the host.
3. Discover mapped volumes from the host and configure multipathing.

From the dynamic menu (function icons), open the Volumes menu and click the Volumes option of the IBM Storwize V7000 graphical user interface (GUI) (Figure 5-1) to start the process of creating new volumes.

![Volumes menu](image)

A list of existing volumes that are defined in specific storage pools is displayed. To define a new volume, click Create Volumes on the tab header (Figure 5-2).

![New Volume window](image)
The wizard opens (Figure 5-3).

By default, all volumes that you create are striped across all available managed disks (MDisks) in one storage pool. The GUI for the IBM Storwize V7000 provides the following preset selections:

- **Generic**: A striped volume that is fully provisioned, as described in 5.1.1, “Creating generic volumes” on page 147.
- **Thin-provisioned**: A striped volume that is space-efficient. Click Advanced to see more to help you determine how much space is allocated initially and how large the volume is able to grow, as described in 5.1.2, “Creating thin-provisioned volumes” on page 149.
- **Mirror**: The striped volume consists of two striped copies and is synchronized to protect against loss of data if the underlying storage pool of one copy is lost, as described in 5.1.3, “Creating mirrored volumes” on page 152.
- **Thin-mirror**: Two synchronized copies. Both are thin-provisioned, as described in 5.1.4, “Creating a thin-mirrored volume” on page 155.
- **Compressed**: The IBM Real-time Compression engine, available in IBM Storwize V7000 v6.4 and later, enables data compression for storage efficiency. More details about compression can be found in *Real-time Compression in SAN Volume Controller and Storwize V7000*, REDP-4859.

### 5.1.1 Creating generic volumes

The most commonly used type of volume is the **generic** volume. This type of volume is fully provisioned, with the entire size dedicated to the defined volume. The host and the IBM Storwize V7000 see the fully allocated space.

Select a generic volume by clicking the leftmost icon (Figure 5-3) and choose the pool in which the volume should be created. In our example, we selected DS3400_pool.

Enter the size and name of the volume.
The result is shown in Figure 5-4. We suggest using an appropriate naming convention of volumes to help you easily identify the associated host or group of hosts. At a minimum, it should contain the name of the pool or at least some tag that identifies the underlying storage subsystem (in our case, the IBM System Storage DS3400 subsystem). It can also contain the host name that the volume is mapped to or content of this volume, for example, name of applications to be installed.

You can preformat the volume before its use by clicking Advanced and selecting Format Before Use (Figure 5-5) on the Characteristics tab.

After you enter a volume name and size, click Create and Map to Host.
The new generic volume is created (Figure 5-6). Click Continue. A new wizard opens so you can assign a volume to the host as described in 5.2.1, “Mapping newly created volumes to the host using the wizard” on page 160.

![Create Volumes]

If you do not want to map the volumes now, click Create to complete the task. Volumes can also be mapped later, as described in 5.2, “Mapping a volume to the host” on page 160.

### 5.1.2 Creating thin-provisioned volumes

Volumes can be configured to be thin-provisioned. With respect to application reads and writes, thin-provisioned volumes behave as though they were fully allocated. When creating a thin-provisioned volume, you may specify two capacities:

- The real physical capacity 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 Volumes selection in the dynamic menu, choose **Create Volumes** and select **Thin-Provision** (Figure 5-7). Enter the required capacity and volume name.

2. Click the **Advanced** button to manage the real and virtual capacity of the volume (Figure 5-8 on page 151).
For the Thin-Provisioned selection on the Capacity Management tab, several advanced options are available:

- **Real**: Specify the size of the real capacity space used during creation. The default is 2% of virtual capacity.
- **Automatically Expand**: This option enables the automatic expansion of real capacity, if additional capacity is to be allocated.
- **Warning Threshold**: Enter a threshold for receiving capacity alerts.
- **Thin-Provisioned Grain Size**: Specify the grain size for real capacity. This describes the size of chunk of storage to be added to used capacity. For example, when the host writes 1 MB of new data the capacity will be increased by adding four chunks of 256 KB each.

**Important**: If you do not use the autoexpand feature, the volume will go offline after reaching real capacity.
3. Apply all required changes and click **Create** to define the volume (Figure 5-9).

![Create Volumes](image)

**Figure 5-9** Create a thin-provisioned volume

4. Again, you can directly start the wizard for mapping this volume to the host by clicking **Create and Map to Host**.

### 5.1.3 Creating mirrored volumes

IBM Storwize V7000 offers the capability to mirror volumes, which means a single volume is presented to the host, but two copies exist in the storage back end, usually in different storage pools (all reads are handled by the primary copy). You can choose which copy is primary and switch between copies online, without interruption to the host I/O. This feature is similar to host-based software mirroring such as Logical Volume Manager (LVM), but it provides a single point of management for all operating systems, and provides storage high availability to operating systems that do not support software mirroring. With this setup, you can protect against array failures (for example, multiple disk failures), and it offers you more advanced features.

It also helps you to greatly improve availability, but it is not a disaster recovery solution, because both copies are accessed by the same node pair and addressable by only a single cluster.

To create a mirrored volume, complete the following steps:

1. Click **Mirror** (Figure 5-10 on page 153) and choose the primary and secondary pool by clicking their names. Although the mirrored volume can be created in the same pool, this is not typical. We suggest that you keep mirrored volumes on a separate set of physical disks (separate managed disks (MDisks)).
2. Select primary and secondary pool by clicking it, and then enter a volume name and the required size (Figure 5-10).

![Create Volumes](image1)

**Figure 5-10** Select primary and secondary pool, volume name, and size

The summary shows capacity information about the pool.

3. If you want to select advanced settings, click **Advanced** and click the **Mirroring** tab (Figure 5-11).

![Advanced mirroring features](image2)

**Figure 5-11** Advanced mirroring features

In the advanced mirroring settings, you can specify a synchronization rate. Enter a Mirror Sync Rate in the range of 1 - 100%. This option sets the priority of copy synchronization progress, allowing you to prefer more important volumes for synchronization that is faster than other mirrored volumes. By default, the rate is set to 50% for all volumes, which equals to 2 MBps.
Table 5-1 shows the synchronization rate value and corresponding data copy rate. Click **OK** to return to the previous panel.

**Table 5-1  Volume copy synchronization rate**

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

4. Click **Create and Map to Host**; the mirrored volume is created (Figure 5-12).

5. Click **Continue**. At this point, you can see the progress of a volume synchronization as shown in Figure 5-13 on page 155. Go to 5.2.1, “Mapping newly created volumes to the host using the wizard” on page 160.
5.1.4 Creating a thin-mirrored volume

With a thin-mirror volume, you can allocate the required space on demand (described in 5.1.2, “Creating thin-provisioned volumes” on page 149) and have several copies of a volume (described in 5.1.3, “Creating mirrored volumes” on page 152) available.

To create a thin-mirrored volume, select primary and secondary pool and make any changes that you want to the volume attributes (Figure 5-14 on page 156).
The summary shows you the capacity information and the allocated space. You can click **Advanced** and customize the thin-provision settings or the mirror synchronization rate. After you create the volume, the confirmation window opens (Figure 5-15 on page 157).
The initial synchronization of thin-mirrored volumes is fast when a small amount of real and virtual capacity is used.

5.1.5 Creating compressed volumes

IBM Real-time Compression is an advanced software function for storage efficiency, enabled by an IBM SAN Volume Controller family of products, including the IBM Storwize V7000 storage subsystem.

The configuration is similar to thin-provisioned volumes and is transparent to users. From the Volumes selection in the dynamic menu, click **Create Volume** and click **Compressed** (Figure 5-16 on page 158). Choose a storage pool and enter the required capacity and volume name.
Figure 5-16  Creating a compressed volume

The summary at the bottom of the wizard shows information about allocated (virtual) capacity and the real capacity that data consumes on this volume. In our example, we defined a 5 GB volume, but the real capacity is only 102.4 MB (because there is no data from the host).

After clicking Advanced, you can see options similar to those in case of thin-provisioned volumes, as shown in Figure 5-17 on page 159. This is because a compressed volume is a type of thin-provisioned volume and is at the same software stack level as thin-provisioned volumes.
As with all other volume types, a confirmation window opens when you click Create.
5.2 Mapping a volume to the host

You can map the newly created volume to the host at creation time or map it later. If you did not click Create and Map to Host when you created the volume, follow the steps in 5.2.1, “Mapping newly created volumes to the host using the wizard” on page 160.

5.2.1 Mapping newly created volumes to the host using the wizard

We continue to map the volume that was created in 5.1, “Provisioning storage from IBM Storwize V7000 to hosts” on page 146. We assume that you followed that procedure and clicked Continue as, for example, shown in Figure 5-6 on page 149.

To map the volumes, complete the following steps:

1. Select a host to which the new volume should be attached (Figure 5-19).

   ![Figure 5-19 Choose a host](image)

2. The Modify Host Mappings window opens, and your host and the newly created volume are already selected. Click Map Volumes to map the volume to the host (Figure 5-20).

   ![Figure 5-20 Modify mappings](image)
3. The confirmation window shows the result of mapping volume task (Figure 5-21).

![Figure 5-21 Confirmation of volume to host mapping](image)

4. After the task completes, the wizard returns to the Volumes window. By double-clicking the volume, you can see the host maps (Figure 5-22).

![Figure 5-22 Host maps](image)

The host is now able to access the volumes and store data on them. See 5.3, “Discovering volumes on hosts and multipathing” on page 161 for information about discovering the volumes on the host and making additional host settings, if required.

You can also create multiple volumes in preparation for discovering them later, and customize mappings. Advanced host configuration is described in Chapter 8, “Advanced host and volume administration” on page 247.

### 5.3 Discovering volumes on hosts and multipathing

This section explains how to discover the volumes that were created and mapped in 5.1, “ Provisioning storage from IBM Storwize V7000 to hosts” on page 146 and 5.2, “ Mapping a volume to the host” on page 160, and how to configure additional multipath settings, if required.
We assume that you have completed all previous steps in this book so that the hosts and the IBM Storwize V7000 are prepared:

- Prepare your operating systems for attachment (Chapter 4, “Host configuration” on page 111).
- Create hosts using the GUI (Chapter 4, “Host configuration” on page 111).
- Perform basic volume configuration and host mapping.

Our examples illustrate how to discover Fibre Channel and Internet Small Computer System Interface (iSCSI) volumes on Microsoft Windows 2008 and VMware ESX 4.x hosts.

From the dynamic menu of the IBM Storwize V7000 GUI, click the **Hosts** icon to open the Hosts menu, and click the **Hosts** option (Figure 5-23).

![Figure 5-23 Navigate to hosts menu](image)

An overview of the configured and mapped hosts is displayed (Figure 5-24).

![Figure 5-24 The existing hosts](image)
5.3.1 Windows 2008 Fibre Channel volume attachment

To complete Fibre Channel volume attachment in Windows 2008, use the following steps:

1. Right-click your Windows 2008 Fibre Channel host in the Hosts view (Figure 5-25) and select Properties.

![Host properties](image)

Navigate to the Mapped Volumes tab (Figure 5-26).

![Mapped volumes to a host](image)

The host details show you which volumes are currently mapped to the host, and you also see the volume UID and the SCSI ID. In our example, four volumes with SCSI ID 0-3 are mapped to the host.

2. Log on to your Microsoft host and click Start → All Programs → Subsystem Device Driver DSM → Subsystem Device Driver DSM. A command-line interface opens. Type the datapath query device command and press Enter to display IBM Storwize V7000 disks that are connected to this host (Example 5-1).

Example 5-1  datapath query device

```
C:\Program Files\IBM\SDDDSM>datapath query device

Total Devices : 4
```

Table of displayed devices:

```
DEV#: 0  DEVICE NAME: Disk5 Part0  TYPE: 2145  POLICY: OPTIMIZED
SERIAL: 600507640082000838000000000000019    LUN SIZE: 50.0GB
============================================================================
Path#  Adapter/Hard Disk  State  Mode  Select  Errors
```
The output provides information about mapped volumes. In our example, four disks are connected, Disk5, Disk6, Disk7, Disk8, and eight paths to the disks are available (State indicates OPEN).

3. Open the Windows Disk Management window (Figure 5-27 on page 165) by clicking Start → Run, and then type diskmgmt.msc, and click OK.
Chapter 5. Basic volume configuration

In our example, three of four disks are already initialized. We will use the fourth, unknown, 1 GB disk as an example for the next initialization steps.

4. Right-click the disk in the left pane and select **Online** (Figure 5-28).

---

**Figure 5-27**  Windows Disk Management

**Windows device discovery:** Usually, Windows discovers new devices, such as disks, by itself (Plug&Play function). If you completed all the steps but do not see any disks, click **Actions → Rescan Disk** in Disk Management to discover potential volumes.

---

**Figure 5-28**  Place a disk online
5. Right-click the disk again, select **Initialize Disk** (Figure 5-29), and click **OK**.

![Figure 5-29 Initialize Disk menu](image1)

6. Right-click in the right pane and select **New Simple Volume** (Figure 5-30).

![Figure 5-30 New Simple Volume](image2)

7. Follow the wizard and the volume is ready to use from your Windows host (Figure 5-31 on page 167).
Figure 5-31   Volume is ready to use

The basic setup is now complete. The IBM Storwize V7000 is configured. And the host is prepared to access the volumes over several paths and is able to store data on the storage subsystem.

5.3.2 Windows 2008 iSCSI volume attachment

To perform iSCSI volume attachment in Windows 2008, complete the following steps:

1. Right-click your Windows 2008 iSCSI host in the **Hosts** view, click **Properties**, and click the **Port Definitions** tab to see the defined host iSCSI address (Figure 5-32).
Clicking the **Mapped Volumes** tab shows you which volumes are currently mapped to the host, and you also see the volume UID and the SCSI ID. In our example, there are no mapped volumes so far (Figure 5-33).

![Figure 5-33 Volumes mapped to iSCSI host](image)

2. Log on to your Windows 2008 host and click **Start → Administrative Tools → iSCSI Initiator** to open the iSCSI Configuration tab (Figure 5-34).

![Figure 5-34 Windows iSCSI Configuration tab](image)
3. Enter the IP address of one of the IBM Storwize V7000 iSCSI ports and click **Quick Connect** (Figure 5-35).

**iSCSI IP addresses:** The iSCSI IP addresses are different from the cluster and canister IP addresses, and they are configured in Chapter 4, “Host configuration” on page 111.

![Figure 5-35  iSCSI Quick Connect](image)

The IBM Storwize V7000 initiator is discovered and connected (Figure 5-36).

![Figure 5-36  iSCSI Initiator target is connected](image)
Now you have completed the steps to connect the storage disk to your iSCSI host, but you are using only a single path at the moment. To enable multipathing for iSCSI targets, more actions are required. Complete the following steps:

1. Click **Start → Run** and type `cmd` to open a command prompt. Run the following command and press Enter (Example 5-2):
   
   ```
   ServerManagerCMD.exe -install Multipath-I0
   ```
   
   **Example 5-2   Installing MPIO**
   
   ```
   C:\Users\Administrator>ServerManagerCmd.exe -Install Multipath-I0
   ```
   
   Start Installation...
   <100/100>
   
   Success: Installation succeeded.

2. Click **Start → Administrative Tools → MPIO**, click the **Discover Multi-Paths** tab, and select the **Add support for iSCSI devices** check box (Figure 5-37).

   ![Figure 5-37   Enable iSCSI MPIO](image)

3. Click **Add** and at the prompt, confirm to reboot your host.
4. After reboot, select **Start → Administrative Tools → iSCSI Initiator** to open the iSCSI Initiator Properties window (Configuration tab). Click the **Discovery** tab (Figure 5-38).

![Figure 5-38 iSCSI Properties Discovery tab](image)

5. Click **Discover Portal**, enter the IP address of another IBM Storwize V7000 iSCSI port (Figure 5-39), and click **OK**.

![Figure 5-39 Discover Target Portal window](image)
6. Return to the Targets tab (Figure 5-40); the new connection is listed there as Inactive.

![Inactive target ports](image)

7. Highlight the inactive port and click **Connect**. The Connect to Target window opens (Figure 5-41).

![Connect to a target](image)
8. Select the **Enable Multipath** check box and click **OK**. The status of the second port now indicates Connected (Figure 5-42).

![Figure 5-42  Second target port connected](image)

Repeat this step for each IBM Storwize V7000 port you want to use for iSCSI traffic. You may have up to four port paths to the system.
9. Click **Devices → MPIO** to ensure that the multipath policy for Windows 2008 is set to the default, which is Round Robin with Subset (Figure 5-43), and click **OK** to close this view.

![Figure 5-43  Round Robin with Subset](image1)

10. Map volume to the iSCSI host if you have not done it already. In our example, we use 2 GB disk.

11. Open the Windows Disk Management window (Figure 5-44 on page 175) by clicking **Start → Run**, and then type `diskmgmt.msc`, and click **OK**.
12. Set the disk online, initialize it, create a file system on it, and then it is ready to use. The detailed steps of this process are the same as described in 5.3.1, “Windows 2008 Fibre Channel volume attachment” on page 163.

Now the storage disk is ready for use (Figure 5-45). In our example, we mapped a 2 GB disk, from IBM Storwize V7000 Generation 2, to a Windows 2008 host using iSCSI protocol.

![Figure 5-44  Windows Disk Management window](image)

![Figure 5-45  Windows Disk Management: Disk is ready to use](image)
5.3.3 VMware ESX Fibre Channel attachment

To do the VMware ESX Fibre Channel attachment, complete the following steps:

1. Right-click your VMware ESX Fibre Channel host in the Hosts view, select Properties, and then click the Mapped Volumes tab (Figure 5-46).

![Host Details: esx01](image)

**Figure 5-46** Mapped volume to ESX Fibre Channel (FC) host

The Host Details window shows that one volume is connected to the ESX FC host using SCSI ID 0. The UID of the volume is also displayed.

2. Connect to your VMware ESX Server using the vSphere client, navigate to the Configuration tab, and select Storage Adapters or Storage view (Figure 5-47).

![vSphere Client: Storage adapters](image)

**Figure 5-47** vSphere Client: Storage adapters

3. Select Rescan All and click OK (Figure 5-48 on page 177) to scan for new storage devices.
4. Select **Storage** and click **Add Storage** (Figure 5-49).

5. The Add Storage wizard opens. Click **Select Disk/LUN** and click **Next**. The IBM Storwize V7000 disk is displayed (Figure 5-50 on page 178). Select it and click **Next**.
6. Follow the wizard to complete the attachment of the disk. After you click **Finish**, the wizard closes and you return to the storage view.

   Figure 5-51 shows that the new volume is added to the configuration.

7. Highlight the new data store and click **Properties** to see the details of it (Figure 5-52 on page 179).
8. Click **Manage Paths** to customize the multipath settings. Select **Round Robin** (Figure 5-53) and click **Change**.
The storage disk is available and ready to use for your VMware ESX server using Fibre Channel attachment.

5.3.4 VMware ESX iSCSI attachment

To do a VMware ESX iSCSI attachment, complete the following steps:

1. Right-click your VMware ESX Fibre iSCSI host in the Hosts view and select Properties. Click the Mapped Volumes tab (Figure 5-54).

![Image](image1.png)

Figure 5-54 iSCSI ESX host properties

The Host Details window shows that one volume is connected to the ESX iSCSI host using SCSI ID 0. The UID of the volume is also displayed.

2. Connect to your VMware ESX Server using the vSphere Client, click the Configuration tab (Figure 5-55), and select Storage Adapters.

![Image](image2.png)

Figure 5-55 vSphere Client: Storage

3. Select iSCSI Software Adapter and click Properties. The iSCSI initiator properties window opens. Select the Dynamic Discovery tab (Figure 5-56 on page 181) and click Add.
4. To add a target, enter the target IP address (Figure 5-57). The target IP address is the IP address of a node canister in the I/O group from which you are mapping the iSCSI volume. Keep the IP port number at the default value of 3260, and click OK. The connection between the initiator and target is established.
5. After you have added all the ports required, close the iSCSI Initiator properties by clicking Close (Figure 5-56 on page 181).

You are prompted to rescan for new storage devices. Confirm the scan by clicking Yes (Figure 5-58).

![Confirm the rescan](image)

6. Go to the storage view shown in Figure 5-59 and click Add Storage.

![Click the Add Storage menu](image)
7. The Add Storage wizard opens (Figure 5-60). Select **Disk/LUN** and click **Next**.

8. The new iSCSI logical unit number (LUN) displays. Select it and click **Next** (Figure 5-61 on page 184).
184  Implementing the IBM Storwize V7000 V7.4

Figure 5-61  Select iSCSI LUN menu

9. Select the file system version (Figure 5-62).

Figure 5-62  Select VMFS version
10. Review the disk layout and click **Next** (Figure 5-63).

![Figure 5-63  Current Disk Layout](image)

11. Enter a name for the data store and click **Next** (Figure 5-64).

![Figure 5-64  Enter a data store name](image)

12. Select the maximum file system size and click **Next** (Figure 5-65).

![Figure 5-65  Select maximum file system size](image)
13. Review your selections and click **Finish** (Figure 5-66).

![Finish the wizard](image)

The new iSCSI LUN is now in the process of being added. After the task completes, the new data store is listed in the storage view (Figure 5-67).

![New data store available](image)
14. Highlight the new data store and click **Properties** to open and review the data store settings (Figure 5-68).

![Figure 5-68  iSCSI data store properties](image)
15. Click **Manage Paths**, select **Round Robin** as the multipath policy (Figure 5-69), and then click **Change**.

![IBM iSCSI Disk (nea.6005076-i000289800036000000000027) Manage Paths](image)

16. Click **Close** twice to return to the storage view, and now the storage disk is available and ready to use for your VMware ESX server using an iSCSI attachment.
Migration wizard

In this chapter, we describe how to migrate existing data from an older storage system to the IBM Storwize V7000 storage system. Because of this migration, applications can benefit from the new features, such as Easy Tier, Thin Provisioning, Real-time Compression, an intuitive management graphical user interface (GUI), and advanced storage replication functions that better support applications.

IBM Storwize V7000 provides a storage migration wizard to guide you through the entire procedure.
6.1 Preparing for data migration

In this section, we show you how to prepare external storage systems and IBM Storwize V7000 for data migration. Data migration is always an important step when setting up new storage systems in an existing storage infrastructure.

Before attaching any external storage systems to the IBM Storwize V7000, see the IBM Storwize V7000 support matrix at the following address:

http://www.ibm.com/support/docview.wss?uid=ssg1S1004941

When migrating data from an external storage system to IBM Storwize V7000 (external storage system is removed when completed), you can temporarily configure the external virtualization license setting. Configuring the external license setting prevents messages from being sent that indicate that you are in violation of the license agreement. When the migration is complete, the external virtualization license must be reset to its original limit.

External storage: When external storage is virtualized by IBM Storwize V7000, a per-enclosure external virtualization license is required. Contact your IBM account team or IBM Business Partner for further assistance, if required.

You can temporarily set the license without any charge only during the migration process.

To prepare the data migration, external storage systems must be configured so they are under IBM Storwize V7000 control. Complete the following steps:

1. Stop host I/O to the external storage logical unit numbers (LUNs) that need to be migrated.
2. Remove zones between the hosts and the storage system from which you are migrating.
3. Update your host device drivers, including your multipath driver, and configure them for attachment to the IBM Storwize V7000 system.
4. Create a storage system zone between the storage system being migrated and IBM Storwize V7000 system, and the host zones for the host attachment.
5. Unmap the LUNs in the external storage system to the host and map them to the IBM Storwize V7000 system.
6. Verify that the IBM Storwize V7000 has discovered the LUNs as unmanaged managed disks (MDisks).

6.2 Data migration using the migration wizard

In this section, we describe how to migrate existing data on external storage systems using the IBM Storwize V7000 storage migration wizard.

We use the wizard that was designed specifically for this scenario to guide you through the process:

- Click the Pools icon and click System Migration (Figure 6-1 on page 191).
Whichever method you use, the storage migration window opens (Figure 6-2). Click **Start New Migration** to start the storage migration wizard.

With the IBM Storwize V7000 storage migration wizard, you can more easily migrate your existing data. Complete the following steps:

1. When the wizard starts, confirm the restrictions and prerequisites by selecting the check boxes, and click **Next** (Figure 6-3 on page 192).
Avoiding data loss: To avoid any potential data loss, back up all the data stored on your external storage before using the wizard.

The restrictions and prerequisites for using the storage migration wizard are as follows:

– Restrictions:
  
  • Do not use the storage migration wizard to migrate cluster hosts, including clusters of VMware hosts and Virtual I/O Server (VIOS).
  
  • Do not use the storage migration wizard to migrate SAN Boot images.

If you have either of these two environments, you must migrate them outside of this wizard. You can find more information about this topic at the following address:


The VMware ESX Storage vMotion feature might be an alternative to migrating VMware clusters. For more information about this topic, see the following address:

– Prerequisites:

- Ensure that the IBM Storwize V7000 Fibre Channel ports have been connected to the storage area network (SAN) fabrics to which the external disk controller and hosts you want to migrate from are connected.

- If you have VMware ESX hosts in the migration, ensure that the VMware ESX hosts are set to allow volume copies to be recognized.

Click **Next** to go to the next step.

2. In the second step of the storage migration wizard, prepare the environment for migration, and click **Next** (Figure 6-4).

![Storage Migration Wizard](Image)

**Figure 6-4  Prepare the environment for migration**

3. Figure 6-5 on page 194 shows the third step (panel) of the storage migration wizard. Do everything that is listed as part of step 3 and complete. Be sure to record the information mentioned in this step because you need it for later steps. Click **Next**.

**SCSI ID:** You might need to record the Small Computer System Interface (SCSI) ID of the volume to be mapped to the host. Some operating systems do not support changing the SCSI ID during the migration.
After you click **Next**, the IBM Storwize V7000 starts to discover external devices (if you correctly zoned the external storage systems with the IBM Storwize V7000 and mapped the logical unit numbers (LUNs); see Figure 6-6 on page 195).
4. When discovery completes, the IBM Storwize V7000 shows the MDisks found (Figure 6-7). Select the MDisks you want to migrate and click **Next**.

If the MDisks that are the subject of the migration are present in the list, select them and click **Next**. The IBM Storwize V7000 starts to import the MDisks that you selected. If the MDisks that need migration are not in the list, check your zone configuration and LUN.
mapping, and reissue the **Detect MDisks** to trigger the IBM Storwize V7000 discovery procedure again.

You can select one or more MDisks as required, and detailed information about the MDisk is visible by double-clicking it.

Figure 6-7 on page 195 shows that two LUNs are discovered as MDisks that are candidates for migration. In your particular situation, you might need to reference the information you recorded earlier to identify these MDisks. In our example, MDisk0 and Mdisk1 have been planned for migration to the new Storwize V7000.

When you click **Next** in this step, the IBM Storwize V7000 completes the import of the MDisks with the host's data and a storage pool is created (Figure 6-8). The MDisks are added to the pool and image mode of the volumes (with the same size as the MDisks) is enabled. These MDisks are ready for mapping back to the original hosts.

![Figure 6-8 Importing discovered MDisks](image_url)
5. Configure the host that needs to access the data after the migration, or create new hosts as needed, and click **Next** (Figure 6-9). In our case, we plan to assign the migrated volumes to the host **Win_2008**.

![Configure Hosts window](image)

**Figure 6-9  Configure Hosts window**

Before you configure any hosts, ensure that the appropriate drivers have been installed on the host and that the host zones have been zoned correctly.

If the host that needs to access the data on the volume after the migration is complete is in the list, click **Next**.

If the host has not been created on the IBM Storwize V7000 storage system, click **Add Host** to create it as required.

6. Map the newly migrated volumes to the host, and when the mapping is complete, click **Next** (Figure 6-10 on page 198).
In step 6 of the migration wizard, the volumes from the imported MDisks that need to be migrated have been listed. The names of the volumes are assigned automatically by the IBM Storwize V7000 storage system. You can change the names to any words that would be meaningful to you by selecting the volume and clicking **Rename** in the **Actions** drop-down menu.

**Names:** Each name must begin with a letter. The name can be a maximum of 63 characters. Valid characters are uppercase letters (A - Z), lowercase letters (a - z), digits (0 - 9), underscore (_), period (.), and space. The names must not begin or end with a space.

To map the volumes to the hosts, select the volumes and click **Map to Host**. A window opens with a drop-down list of the hosts. Now, the IBM Storwize V7000 lets you choose which host to which you need to map the volumes. Choose the correct host and click **Next**.

**Volume mapping:** As a preferred practice, map the volume to the host with the same SCSI ID before the migration, which you should have recorded in step 3.

Figure 6-11 on page 199 shows the menu for choosing the host to which you map the volumes.
Figure 6-11  Choose the host to which you map the volumes

After choosing the host, the Modify Host Mappings window opens. On the right, you can find your newly mapped volumes highlighted in yellow. You can see the SCSI ID of the new mappings. Click Apply to complete the mapping.

Figure 6-12 shows the Modify Host Mappings window in the migration wizard.

When the mapping completes, you see that the Host Mappings column of the volumes changed from No to Yes. A scan can be performed to discover the new devices on the host for verification. Click Next to go to the next step of the storage migration wizard.
7. Select the destination storage pool for data migration and click **Next** (Figure 6-13).

![Select Pool (optional) screen](image)

Figure 6-13  Select a storage pool

The destination storage pool of the data migration might be an external storage pool or an internal storage pool. Ensure that there is enough space in the storage pools.

When you click **Next**, the migration begins (Figure 6-14). The migration runs in the background and results in a copy of the data being placed on the MDisks in the storage pool selected. The process uses the volume mirroring function included with the IBM Storwize V7000, and when complete, the volumes have pointers to both the new copy on the storage pool selected and on the original external storage system.

![Start Migration screen](image)

Figure 6-14  Starting migration

8. Click **Finish** to end the storage migration wizard (Figure 6-15 on page 201).
Figure 6-15  End of the storage migration wizard

The end of the storage migration wizard is not the end of the data migration. The data migration has begun, and after clicking **Finish** in step 8, you can find the migration progress in the migration window. You can also find the target storage pool to which your volumes are being migrated, along with the status of the volumes.

Figure 6-16 shows the data migration progress in the migration window.

![Migration progress in the migration window](image)

**Figure 6-16  Migration progress in the migration window**

When the migration progress reaches 100%, select the volumes and click **Finalize** in the **Actions** drop-down menu in the migration window (Figure 6-17). The image mode copy of the volumes on the original external storage system is deleted and the associated MDisks from the storage pool are removed; the status of those MDisks is unmanaged.

![Finalize the migration](image)

**Figure 6-17  Finalize the migration**
The IBM Storwize V7000 requests that you confirm the migration of the volumes. Verify the volume name and the number of migrations, and click **OK**. Figure 6-18 shows the Finalize Volume Migrations window.

![Finalize Volume Migrations](image)

**Figure 6-18 Verify the migration finalization**

When the finalization completes, the data migration to the IBM Storwize V7000 is done. You can unzone and remove the older storage system from the IBM Storwize V7000.

For more information about the advanced migration function, see Chapter 7, “Storage pools” on page 203 and Chapter 9, “External storage virtualization” on page 323.
Storage pools

This chapter explains how IBM Storwize V7000 manages physical storage resources. All storage resources under IBM Storwize V7000 control are managed using storage pools. Storage pools aggregate disk capacity on a storage area network (SAN) and provide the containers in which volumes can be created. Storage pools make it easy to dynamically allocate resources, maximize productivity, and reduce costs. Advanced internal storage, managed disks (MDisks), and storage pool management are covered in this chapter. External storage is covered in Chapter 9, “External storage virtualization” on page 323.
7.1 Working with internal drives

This section describes how to configure the internal storage disk drives using different Redundant Array of Independent Disks (RAID) levels and different optimization strategies.

In this section, we start with the environment shown in Figure 7-1. Most of the internal drives are unused. Currently, the existing MDisks come from internal storage, and sample arrays, storage pools, and volumes were created for use.

You can learn how to manage MDisks in 7.2, “Working with MDisks” on page 220, how to manage storage pools in 7.3, “Working with storage pools” on page 237, how to work with external storage in Chapter 9, “External storage virtualization” on page 323, how to create volumes in Chapter 5, “Basic volume configuration” on page 145, and how to create hosts in Chapter 4, “Host configuration” on page 111.
The IBM Storwize V7000 storage system provides an individual Internal Storage panel for managing all internal drives. To access the Internal Storage panel, hover the mouse cursor over the **Pools** selection in the dynamic menu and select **Internal Storage**. (Figure 7-2).

The Internal Storage panel (Figure 7-3) gives an overview of the internal drives installed in the IBM Storwize V7000 storage system. Selecting **All Internal** in the Drive Class Filter shows all drives installed in the managed system, including attached expansion enclosures. Alternatively, you can filter the drives by their size, RPM speed or class, for example, you can choose to show only SAS, SATA, or SSD drives. In our example, we have only one type of drives, which is 600 GB SAS 10 K RPM, so the **All Internal** menu will show the same output as filtered by drive.

The right side of the Internal Storage panel lists the internal disk drives of the selected type. By default, the following information is listed with it:

- Logical drive ID
- Drive's capacity
- Current type of use (unused, candidate, member, spare, or failed)
- Status (online, offline, and degraded)
- MDisk's name that the drive is a member of
- Enclosure ID that it is installed in
- Physical Drive Slot ID of the enclosure that it is installed in
More details, for example, the drive’s RPM speed or its MDisk member ID, are available by right-clicking the blue header bar of the table (Figure 7-4). Also, if you click the table headers, the table content can be ascending or descending, sorted by the value represented in that column.

![Figure 7-4 Internal Storage panel details selection](image)

**Note:** This view can also be customized by right-clicking any drive and selecting **Customize Columns**.

In addition, you can find the current internal storage capacity allocation indicator at the upper right. The Total Capacity shows the overall capacity of the internal storage that is installed in this IBM Storwize V7000 storage system. The MDisk Capacity shows the internal storage capacity that is assigned to the MDisks, which means capacity already configured in RAID arrays. The Spare Capacity shows total capacity of all drives that were designated as spare drives.

The percentage bar indicates the percentage of the total capacity that is allocated to the MDisk capacity, with the grayed area showing the capacity of the spare disks (Figure 7-5).

![Figure 7-5 The percentage bar for internal storage capacity](image)

### 7.1.1 Actions on internal drives

There are a few actions that can be taken on internal drives when you select the drive and right-click it, or click the **Actions** drop-down menu (Figure 7-6 on page 207).
Chapter 7. Storage pools

Options on the Actions menu are as follows:

- **Fix Error**: If the physical drive is in an error condition, the Fix Error action starts the Directed Maintenance Procedure (DMP) for a defective drive. For more information about this topic, see Chapter 13, “RAS, monitoring, and troubleshooting” on page 495.

- **Take Offline**: When problems occur on the drives, the internal drives can be taken offline by selecting this option in the Actions drop-down menu. A confirmation window opens (Figure 7-7). The IBM Storwize V7000 storage system prevents the drive from being taken offline if the result might produce data loss. A drive should be taken offline only if a spare drive is available.

- **Mark as**: The internal drives in the IBM Storwize V7000 storage system can be assigned to several usage roles. These roles are as follows:
  - Unused: The drive is not a member of an MDisk (array).
  - Candidate: The drive is available for use in an array.
  - Spare: The drive can be used as a hot spare if required.

Select **Mark as** from the Actions drop-down menu, and select the role you want to assign to the drive (Figure 7-8 on page 208).
Identify: Use the Identify action to turn on the LED light so you can easily identify a drive that needs to be replaced, or that you want to troubleshoot.

Figure 7-9 shows the information that is displayed when you click the Identify action.

Click Turn LED Off when you are done.

Show Dependent Volumes: Shows you the volumes that are dependent on the drives (Figure 7-10). This view will show volumes only if removal or failure of the drive they are dependent on would cause volumes to become unavailable. For example, this view does not show any drive if you chose Show Dependent Volumes on one of RAID5 or RAID10 drives, because removal of one drive from RAID5 or RAID10 will not cause volumes to go offline. By holding the ctrl key, you can select more than one drive and show volumes dependent on those drives. This will show volumes that would go offline in case selected drives failed or were removed. In our example on Figure 7-10, we selected two drives and showed volume dependent on those two drives.

Upgrade: This lets you upgrade the drive firmware from the GUI. The drive firmware package is separate from Storwize V7000 firmware package but you can download it the same way from the IBM Support and Downloads web page. After clicking the Upgrade
option, the following window appears where you can choose and upload to Storwize the drive firmware package (Figure 7-11). By holding the Ctrl key, you can select many drives and upgrade them all at the same time.

![Upgrade Drive Firmware Window](image)

**Figure 7-11  Upgrade drive firmware window**

- Properties: Clicking **Properties** in the Actions drop-down menu, or double-clicking the drive, provides more information about the drives (Figure 7-12).

![Properties Tab with Default Format](image)

**Figure 7-12  Properties tab with default format**
If you select the **Show Details** check box, you can discover more detailed information, including vendor ID, FRU Part Number, and ID (Figure 7-13).

![Properties tab with details](image)

**Figure 7-13  Properties tab with details**

### 7.1.2 Configuring internal storage

To configure internal storage for use with hosts, move the mouse cursor over the **Pools** selection and click **Internal Storage**, and then click **Configure Storage**.

A configuration wizard opens and guides you through the process of configuring internal storage (Figure 7-14 on page 211).
Chapter 7. Storage pools

211

Figure 7-14 Configure Internal Storage wizard

The wizard shows all internal drives with a status of candidate available for configuration. If there are internal drives with a status of unused, a window opens, offering to include them in the RAID configuration.

The storage configuration wizard simplifies the initial disk drive setup and offers two options:

- Use the recommended configuration
  
  This option guides you through the wizard described in “Using the recommended configuration” on page 213.

- Select a different configuration
  
  This option uses the wizard described in “Selecting a different configuration” on page 215.

Before going through the storage configuration wizard, we first describe IBM Storwize V7000 RAID configuration presets.

RAID configuration presets

RAID configuration presets are used to configure internal drives based on recommended values for the RAID level and drive class. Each preset has a specific goal for the number of drives per array, the number of spare drives to maintain redundancy, and whether the drives in the array are balanced across enclosure chains, thus protecting the array from enclosure failures.

Table 7-1 describes the presets that are used for solid-state drives (SSDs) for the IBM Storwize V7000 storage system.

Table 7-1 SSD RAID presets

<table>
<thead>
<tr>
<th>Preset</th>
<th>Purpose</th>
<th>RAID level</th>
<th>Drives per array goal</th>
<th>Spare drive goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSD RAID 5</td>
<td>Protects against a single drive failure. Data and one stripe of parity are striped across all array members.</td>
<td>5</td>
<td>8</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 7-2 describes the RAID presets that are used for hard disk drives (HDDs) for the IBM Storwize V7000 storage system.

<table>
<thead>
<tr>
<th>Preset</th>
<th>Purpose</th>
<th>RAID level</th>
<th>Drives per array goal</th>
<th>Spare drive goal</th>
<th>Chain balance</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSD RAID 6</td>
<td>Protects against two drive failures. Data and two stripes of parity are striped across all array members.</td>
<td>6</td>
<td>12</td>
<td>1</td>
<td>All drives in the array are from the same chain wherever possible.</td>
</tr>
<tr>
<td>SSD RAID 10</td>
<td>Protects against at least one drive failure. All data is mirrored on two array members.</td>
<td>10</td>
<td>8</td>
<td>1</td>
<td>All drives in the array are from the same chain wherever possible.</td>
</tr>
<tr>
<td>SSD RAID 0</td>
<td>Provides no protection against drive failures.</td>
<td>0</td>
<td>8</td>
<td>0</td>
<td>All drives in the array are from the same chain wherever possible.</td>
</tr>
<tr>
<td>SSD Easy Tier</td>
<td>Mirrors data to protect against drive failure. The mirrored pairs are spread between storage pools to be used for the Easy Tier function.</td>
<td>10</td>
<td>2</td>
<td>1</td>
<td>All drives in the array are from the same chain wherever possible.</td>
</tr>
</tbody>
</table>

SSD RAID instances: In all SSD RAID instances, drives in the array are balanced across enclosure chains if possible.
Using the recommended configuration

As shown in Figure 7-15, when you click Use the recommended configuration, the wizard offers a recommended storage configuration at the bottom of the window.

The recommended RAID presets for different drive classes are as follows:

- SSD Easy Tier preset for solid-state drives
- Basic RAID 5 for SAS drives
- Basic RAID 6 for Nearline SAS drives

Using the recommended configuration, spare drives are also automatically created to meet the spare goals according to the preset chosen. One spare drive is created out of every 24 disk drives of the same drive class on a single chain.

For example, if you have 20 x 600 GB 10 K SAS drives on one chain, one drive in these 20 drives is dedicated as a spare drive. If you have 20 x 600 GB 10 K SAS drives on both chains, which means that there are 10 drives in each chain, then one spare drive on each chain is created. So, if you have 40 x 600 GB 10k SAS drives on both chains, then two spare drives on each chain are created and you have a total of 36 drives that can be array members for the RAID setup.

Spare drives in the IBM Storwize V7000 are global spares, which means that any spare drive having at least the same capacity as the drive that needs to be replaced can be used in any array. Thus, an SSD array with no SSD spare available will use an HDD spare instead.

In our example of using the recommended configuration, one array that uses the Basic RAID 5 preset is proposed with no hot spare, because the hot spare is already assigned (see Figure 7-3 on page 205). If the proposed configuration meets your requirements, click Finish, and the system automatically creates the array MDisk with a size according to the chosen RAID level.
Storage pools are also automatically created to contain the MDisks with similar performance characteristics, including the consideration of RAID level, number of member drives, drive class, and so on (Figure 7-16).

![Array and storage pool creation by wizard](image1)

Figure 7-16   Array and storage pool creation by wizard

After an array is created, the Array MDisk members are synchronized with each other through a background initialization process. To monitor the progress of the initialization process, click the icon at the left of the **Running Tasks** status bar and then click the displayed task (Figure 7-17).

![Running Tasks status bar](image2)

Figure 7-17   Running Tasks status bar

The progress window opens (Figure 7-18). The array is available for I/O during this process. The initialization does not affect the availability because of possible member drive failures.

![Array initialization](image3)

Figure 7-18   Array initialization
The capacity allocation indicator shows that allocation capacity has reached 92% after the configuration is set (Figure 7-19).

If the proposed configuration does not meet your requirements, click Select a different configuration on the initial window of the Configure Storage wizard and continue with the more flexible setup shown in “Selecting a different configuration” on page 215.

Selecting a different configuration

The option Select a different configuration offers a more flexible way for the configuration of the internal storage as compared to the Use the recommended configuration preset in terms of drive selection, RAID level, and storage pool to be used.

Complete the following steps to select a different configuration:

1. Click Select a different configuration and select the Drive Class you want to configure.
   Drive class selection is disabled if you have only one disk drive class, as is in our case.
   Then select the appropriate RAID preset (Figure 7-20 on page 216).
2. Select the **Number of drives to provision** in the configuration. By default, all drives with the Candidate status are included, and drives with the Unused status are made available for use in the array also. The wizard offers to automatically configure spares to match the spare goal, as explained in “RAID configuration presets” on page 211. Disabling this option enables you to define spare drives manually by marking drives with the status Candidate as Spare. Then, decide which of the following two optimization options to use:

- **Optimize for Performance** (see “Optimize for Performance” on page 216)
- **Optimize for Capacity** (see “Optimize for Capacity” on page 219)

**Optimize for Performance**

The goal of this algorithm is to create arrays with identical capacity and performance characteristics to achieve the best possible performance for the volumes to be created. In a performance optimized setup, the IBM Storwize V7000 provisions eight physical disk drives in a single array MDisk, except for the following situations:

- RAID 6 uses 12 disk drives.
- SSD Easy Tier uses two disk drives.

Therefore, creating an Optimized for Performance configuration is possible only if enough drives are available to match your needs. As a consequence, all arrays with similar physical disks feature the same performance characteristics. Because of the defined presets, this setup might leave drives unused. The remaining unconfigured drives can be used then in another array.

Figure 7-21 on page 217 shows an example of the performance optimized setup.
3. Choose the storage pool that you want to assign the capacity to (Figure 7-22 on page 218). Select either an existing storage pool that does not contain MDisks, or a pool that contains MDisks with similar performance characteristics, which are listed automatically.
4. Alternatively, create a new storage pool (Figure 7-23).

Figure 7-22  Assign capacity to an existing storage pool

Figure 7-23  Create new storage pool
5. Click **Finish** to finalize the wizard. After the wizard completes, the configuration changes are reflected at the upper right of the Internal Storage panel (Figure 7-24).

With the performance optimized setup, three array MDisks are created, each containing two drives and one drive remains unconfigured. The progress of the array initialization process is displayed in the Running Tasks status indicator, and the capacity allocation indicator has been updated to 83% also.

**Optimize for Capacity**

The goal of this algorithm is to create a setup with the maximum usable capacity, depending on the selected RAID level.

While creating arrays, the IBM Storwize V7000 system attempts to fulfill the *width goal* for each array before creating another one. This setup is an example with 7 *unused* drives. **Select a different configuration** was chosen with the following options:

- Preset: Basic RAID-5.
- Automatically configure spares was checked.
- Optimize for capacity.

The results are as follows:

- No spare drive is defined because it is already assigned on the chain.
- One basic RAID-5 array is created. This array contains all seven drives that are not fully width-goal matched.
- No unused drives remain.

The *width goals* for the array levels are shown in Table 7-3.

**Table 7-3  Array width goals**

<table>
<thead>
<tr>
<th>RAID level</th>
<th>Array width goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAID 5, 10, or 0</td>
<td>eight disks</td>
</tr>
<tr>
<td>RAID 6</td>
<td>12 disks</td>
</tr>
<tr>
<td>SSD Easy Tier</td>
<td>two disks</td>
</tr>
</tbody>
</table>

![Configuration complete with performance optimized setup](image)
If configuring one fully populated IBM Storwize V7000 SFF disk enclosure with 24 disk drives in the Capacity Optimized setup, all available Candidate disk drives are used, no “unconfigured drives” remain, unlike in the Performance Optimized setup. Figure 7-25 shows an example of the capacity optimized setup.

![Capacity optimized setup](image)

**Provisioned drives:** You can choose the number of drives to be provisioned. The wizard proposes that you create arrays from all available candidate drives. If you want to configure only a subset of drives now, you can change the number in the **Drives to provision** field to the value that you want. If you want to manually select the drives to provision, assign Candidate status to only the drives that you want. Alternatively, you can use the CLI for full control over the drive configuration, which is beyond the intended scope of this book.

### 7.2 Working with MDisks

After the configuration is completed for the internal storage, you can find the MDisks that are created on the internal arrays in the MDisks by Pools panel.
To access the MDisks by Pools panel, hover over the Pools selection in the dynamic menu on the left of the IBM Storwize V7000 GUI and select MDisks by Pools (Figure 7-26).

The MDisks by Pools panel (Figure 7-27) lets you manage all MDisks made of both internal and external storage.

The panel provides the following information:

- MDisk name
- Numerical ID
- Status
- Capacity
- Mode
- Name of the storage pool it belongs to
- Name of the back-end storage system for MDisk on external storage
- MDisk’s LUN ID from external storage systems
- Assigned storage tier

**Note:** As with *Internal Storage*, this panel view can be customized to show more or less than default information, as shown on Figure 7-4 on page 206.

IBM Storwize V7000 has four modes for MDisks:

- Array: Array mode MDisks are constructed from internal drives using the RAID functionality. Array MDisks are always associated with storage pools (also known as *MDisk Groups*). This means you cannot have an array MDisk that is not in a storage pool.
- Unmanaged: The MDisk is not a member of any storage pools, which means it is not being used by the IBM Storwize V7000 storage system. LUNs presented by external storage systems to IBM Storwize V7000 are discovered as unmanaged MDisks.
- Managed: The MDisk is assigned to a storage pool and provides extents to be used by volumes.
- Image: The MDisk is assigned directly to a volume with a one-to-one mapping of extents between the MDisk and the volume.

### 7.2.1 Adding MDisks to storage pools

By adding *unmanaged MDisks* to a pool, their status changes to *managed MDisks*. Managed MDisks can belong to only one pool. Unmanaged MDisks can either be added to a newly created pool or to an existing pool to expand its capacity. Pools are commonly used to group MDisks from the same storage subsystem.

A new pool can be created in the MDisks by Pools panel by clicking the **Create Pool** icon. Optionally, assign a name to the pool and choose an icon (Figure 7-28).

![Create Pool](image)

*Figure 7-28  Create Pool: part 1 of 2*
In the Create Pool window (Figure 7-29), you can include unmanaged MDisks in the new pool. Several filter options are available so you can limit the selection, for example, by storage subsystem, capacity, and so on. To select multiple MDisks, press the Ctrl key while using the mouse to click the MDisks that you want from the list.

Alternatively, if you did not assign any MDisks to a storage pool at pool creation time, you can assign an unmanaged MDisk later, to an existing pool. To do so, select and right-click an unmanaged MDisk, and click Actions → Assign to Pool (Figure 7-30).

Choose the storage pool to which you want to add the MDisk and click Add to Pool (Figure 7-31 on page 224).

**Existing data:** If you need to preserve existing data on the unmanaged MDisks, do not select Assign to Pool on this LUN because this action deletes the data. Use Import instead, which is described in 7.2.2, “Importing MDisks” on page 226.
Implementing the IBM Storwize V7000 V7.4

Figure 7-31 Add MDisk to Pool

After the IBM Storwize V7000 system completes this action, the MDisk shows up in the pool that it was added to (Figure 7-32).

Figure 7-32 MDisk added to pool
In some cases, you might want to remove MDisks from storage pools to reorganize your storage allocation. You can remove MDisks from storage pools by selecting the MDisks and clicking **Unassign from Pool** from the **Actions** drop-down menu (Figure 7-33).

![Figure 7-33 Unassign an MDisk from the storage pool](image)

Confirm the number of MDisks that you want to remove (Figure 7-34). If you have data on the MDisks, and you still need to remove the MDisks from the pool, select the following check box: **Remove the MDisk from the pool even if it has data on it.** The system migrates the data to other MDisks in the pool.

![Figure 7-34 Confirm the removal of MDisk from the pool](image)

**Available capacity:** The data on the MDisk that is being unassigned from the storage pool will be migrated to available space on the remaining MDisks. Ensure that you have enough available capacity remaining in the storage pool for the data on the removed MDisk to be migrated or else this command will fail.
After you click **Delete**, the **Warning** window appears as shown in Figure 7-35. You have to leave the MDisk that is to be removed accessible to the system while all data is copied to other MDisks in the same storage pool. If the MDisk is unmapped from the Storwize V7000 before all data is copied, all volumes created in this storage pool go offline and will remain in this state until the removed MDisk is connected again.

![Figure 7-35 Delete MDisk warning window](image)

If you are sure you want to remove MDisks from the pool, click **Yes** and then data migration from the MDisks to be removed starts. You can find the migration progress in the Running Tasks status indicator (Figure 7-36).

![Figure 7-36 Data migration progress when removing MDisks from the pool](image)

### 7.2.2 Importing MDisks

LUNs that are hosted on external storage systems can be migrated into IBM Storwize V7000 internal storage. Also, hosts that were previously directly attached to those external storage systems can continue to use their storage that is now presented by the IBM Storwize V7000 instead.

To achieve this configuration, the existing external LUNs must be imported as an *image-mode volume* using the Import option. This action is possible for unmanaged MDisks only. Those disks must not have been added to a pool, as described in 7.2.1, “Adding MDisks to storage pools” on page 222.

If the Import option is used and no existing storage pool is chosen, a temporary *migration pool* is created to hold the new image-mode volume. This image-mode volume has a direct block-for-block translation from the imported MDisk to the volume presented from now on by the IBM Storwize V7000 and existing data is being preserved.

Figure 7-37 on page 227 shows an example of how to import an unmanaged MDisk. Select the unmanaged MDisk and click **Import** from the **Actions** drop-down menu.
Figure 7-37  Import MDisk as image-mode volume

Figure 7-38 shows that the import wizard is activated; the wizard guides you through the import process:

1. Select whether to enable or disable caching on the image-mode volume, and then click Next. Clear the Enable Caching check box if you use copy services on the external storage system hosting the LUN. A preferred practice is to use the copy services of IBM Storwize V7000 for volumes that are virtualized.

Figure 7-38  Import Wizard, enable caching on image-mode volume

2. Choose to import the volume into the existing storage pool or use a temporary pool for migration (Figure 7-39). If you choose to migrate to a temporary pool, select the extent size characteristics of the storage pool. Select the extent size and click Finish.

Figure 7-39  Import Wizard, select target storage pool
3. You can check the Task completion window on the action performed. Note that an image-mode volume from exactly one MDisk was created and placed into a new storage pool. See Example 7-1.

Example 7-1  Task window showing CLI command on Migrate MDisk action

```
svctask mkmdiskgrp -ext 1024 -name MigrationPool_1024
svctask mkvdisk -iogrp io_grp0 -mdisk mdisk0 -mdiskgrp MigrationPool_1024 -name controller0_0000000000000000 -syncrate 80 -vtype image
```

4. This volume is now presented through IBM Storwize V7000 pool and can be mapped to the original host in the pass through mode. The data is still physically present on the physical disk of the original storage controller system and no automatic migration process is running at this time. If needed, the image-mode volume can be migrated manually into an internal pool (array) by selecting either Migration to Another Pool or Volume Copy Actions on the Volumes panel of IBM Storwize V7000 GUI. More information about these volume actions can be found in Chapter 8, “Advanced host and volume administration” on page 247.

Figure 7-40 shows the Volume panel with the imported image-mode MDisk.

```
Figure 7-40 Imported image-mode MDisk presented as an image-mode volume
```

5. If you want to migrate the data from the external volume (that is being imported) to existing V7000 internal MDisks (arrays), choose an existing destination storage pool in step 2 on page 227. Only pools with sufficient free extent capacity are listed (Figure 7-41). The data migration begins automatically after the MDisk is imported successfully.

```
Figure 7-41 Import Wizard: Select existing storage pool
```

You can check the migration progress in the Running Tasks status indicator (Figure 7-42 on page 229) or by selecting Pools → System Migration.
6. After the migration completes, you can find the volume in the chosen destination pool (Figure 7-43).

![Figure 7-42 Migration progress in the status indicator of Running Tasks](image)

**Note:** Notice the difference between importing an MDisk into a new temporary storage pool and importing an MDisk into an existing storage pool (array). In the first case, we only hide the original storage controller LUN behind IBM Storwize V7000 system without moving any of its data blocks. In the latter case, we directly migrate the data from the original storage controller LUN to IBM Storwize V7000 internal storage pool (array).

![Figure 7-43 Imported Image MDisk migrated to destination pool as a striped volume](image)

All data has been migrated off the source MDisk and striped to MDisks in the destination storage pool. The source MDisk status is changed to managed and is associated with an automatically created migration pool. It can be further used as a regular MDisk to host volumes (Figure 7-44).

![Figure 7-44 MDisk mode changed to managed](image)

Alternatively, the migration into another pool can be done by selecting **Pools → System Migration**. Migration is described in more detail in Chapter 6, “Migration wizard” on page 189.

This feature is normally used as a vehicle to migrate data from existing external LUNs into storage pools, either located internally or externally, on the IBM Storwize V7000. You should not use image-mode volumes as a **long-term solution** for reasons of performance and reliability.
7.2.3 RAID action for MDisks

Internal drives in the IBM Storwize V7000 are managed as Array mode MDisks, on which several RAID actions can be performed. Choose the appropriate Array MDisk by selecting **Pools → MDisks by Pools** from the dynamic menu, and then click **Actions → RAID Actions** (Figure 7-45).

![Figure 7-45  RAID actions on Array MDisks](image)

You can choose the following RAID actions:

- **Set Spare Goal**: You can set the number of spare drives that is required to protect the array from drive failures. If the number of drives assigned as Spare does not meet the configured spare goal, an error is logged in the event log that reads “Array MDisk is not protected by sufficient spares.” This error can be fixed by either replacing defunct drives, which occupy a spare drive, or by adding additional drives as spare. Remember, during the internal drive configuration, spare drives are automatically assigned according to the chosen RAID preset’s spare goals, as described in 7.1.2, "Configuring internal storage" on page 210.

- **Swap Drive**: This action can be used to replace a drive in the array with another drive with the status of Candidate or Spare. This action is used to replace a drive that has failed, or that is expected to fail soon, for example, as indicated by an error message in the event log. Select an MDisk that contains the drive to be replaced and click **RAID Actions → Swap Drive**. In the Swap Drive window (Figure 7-46 on page 231), select the member drive to be replaced and click **Next**.
As shown in Figure 7-47, select an available Candidate or Spare drive to swap into the Array MDisk and click **Finish**.

The exchange of the drives starts and runs in the background, and the volumes on the affected MDisk remain accessible.

- **Delete**: An Array MDisk can be deleted from the MDisks by Pools panel by clicking **RAID Actions → Delete**. You must confirm the deletion by entering the correct number of array MDisks to be deleted (Figure 7-48 on page 232). If data is on the MDisk, force the deletion
by selecting the check box (Delete the RAID array MDisk even if it has data on it. The system migrates the data to other MDisks in the pool).

![Delete Array MDisk](image)

Figure 7-48 Delete Array MDisks window

**Available capacity:** Ensure that you have enough available capacity left in the storage pool for the data on the MDisks to be removed.

Data on MDisks is migrated to other MDisks in the pool, assuming enough space is available on the remaining MDisks in the pool. After deleting an MDisk from a pool, RAID is deleted and all its member drives return to candidate mode.

### 7.2.4 Selecting the tier for MDisks

The IBM Storwize V7000 Easy Tier feature is described in Chapter 10, “Advanced features for storage efficiency” on page 335. In this section, we show how to adjust the tier settings.

Three tiers exist:

- Flash tier is for storage made of solid-state drives or flash arrays, which is the faster-performing storage.
- Enterprise tier is for SAS storage. This is the default for external MDisks.
- Nearline is for NL-SAS storage.

Internal drives have their proper tier assigned automatically by the IBM Storwize V7000 because Storwize knows the type of its internal drives. MDisks on the external storage systems are assigned the Enterprise tier by default. This setting can be changed manually by the user. To assign a specific tier to an MDisk, select **Pools → MDisks by Pool** from the dynamic menu and select **Actions → Select Tier** (Figure 7-49 on page 233).
We assign the **Flash** tier to the external mdisk1 (Figure 7-50). We can change the MDisk tier even if it is not correct, like in this example, we change the MDisk type to `flash` even if drives that this MDisk is made of are SAS drives.

**Note:** You cannot change the Storwize V7000 internal MDisk type in the GUI. This can be done only in CLI using the `chmdisk` command.

After the action completes successfully, the mdisk1 can be found in the flash tier (Figure 7-51 on page 234).
Several additional actions can be performed on MDisks:

- **Detect MDisks**: Selecting **Actions** → **Detect MDisks** is useful if you have external storage controllers in your environment (for more information, see Chapter 9, “External storage virtualization” on page 323). The Detect MDisks action initiates a rescan of the Fibre Channel network. It discovers any new MDisks that are mapped to the IBM Storwize V7000 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 V7000 configuration to reflect any changes.

When external storage controllers are added to the IBM Storwize V7000 environment, the IBM Storwize V7000 automatically discovers the controllers and the LUNs that are presented by those controllers and lists them as unmanaged MDisks. However, if you attached new storage and the IBM Storwize V7000 does not detect it, you might need to use the Detect MDisks action before the system automatically detects the new LUNs. If the configuration of the external controllers is modified after, the IBM Storwize V7000 might be unaware of these configuration changes. Use the Detect MDisks action to rescan the Fibre Channel network and update the list of unmanaged MDisks.

Although the name of this task is Detect MDisks, it can also detect Fibre Channel host ports, additional Storwize V7000 controller enclosures, or SVC/Storwize systems to create partnership with. Basically, Storwize detects changes in the fabric automatically but you can force it to scan the Fibre Channel network by issuing Detect MDisk action.

**MDisks detection**: The Detect MDisks action is asynchronous. Although the task appears to be finished, it still might be running in the background.

- Include Excluded MDisk: An MDisk might be excluded from the IBM Storwize V7000 because of multiple I/O failures. These failures might be caused, for example, by link errors. After a fabric-related problem is corrected, the excluded disk can be added back into the IBM Storwize V7000 by selecting the MDisk and clicking **Actions** → **Include Excluded MDisk**.

- Rename: MDisks can be renamed by selecting the MDisk and clicking **Actions** → **Rename**. Input the new name of your MDisk and click **Rename** (Figure 7-52 on page 235).
Figure 7-52  Rename MDisk

- Show Dependent Volumes: The volumes that depend on a particular MDisk can be displayed by selecting the MDisk and clicking \textbf{Actions} \rightarrow \textbf{Show Dependent Volumes}. This information shows volumes that will go offline in case of MDisk failure, goes offline, or is excluded. The volumes are listed with general information (Figure 7-53).

Figure 7-53  Show dependent volumes

From the Volumes panel of the IBM Storwize V7000 GUI, several actions can be done to volumes by selecting the volume and choosing an action from the \textbf{Actions} menu (Figure 7-54). For more information about volume actions, see Chapter 5, “Basic volume configuration” on page 145.

Figure 7-54  Volume actions

To view information about MDisks that are related to the specific volume, select volume \textbf{Properties}, and on the \textbf{Member MDisks} tab, right-click an MDisk and select \textbf{Properties}. A window opens.
This information window has three tabs:

- **Overview**: Lists information about the MDisk. To show more details, click **Show Details** (Figure 7-55).

![Figure 7-55 MDisk overview with detailed information](image)

- **Member Drives**: Is used only with array MDisks and lists all the member drives of this MDisk (Figure 7-56). Also, all actions described in 7.1.1, “Actions on internal drives” on page 206 can be done on the drives listed here.

![Figure 7-56 Member Drives tab](image)
Dependent Volumes: Lists all volumes that use extents on this MDisk (Figure 7-57).

7.3 Working with storage pools

Storage pools aggregate a disk capacity into a single container for MDisks and provision the storage space for volumes that are allocated to storage hosts. They are the same as storage pools in an IBM SAN Volume Controller when they are using external storage for their capacity, or similar to RAID arrays on a traditional storage subsystem when they are built by using internal drives. IBM Storwize V7000 organizes storage in storage pools to ease storage management, making it more efficient. Storage pools and MDisks are managed in the MDisks by Pools panel. You can access the MDisks by Pools panel from the IBM Storwize V7000 Overview panel by clicking the Pools icon.

Figure 7-58 shows how to access the Pools panel from the Storwize V7000 Overview panel.

The other way to access the Pools panel is to move the mouse cursor over the Pools selection in the dynamic menu and click MDisks by Pools (Figure 7-59 on page 238).
The MDisk by Pools panel (Figure 7-60) allows you to manage storage pools, both internal and external. All existing storage pools are displayed row-by-row. The first row is the Unassigned MDisks item, containing all unmanaged MDisks, if any exist.
To make the view clearer and structured, filters can be used to show only those storage pools that match conditions that are defined in the filter field in the upper part of the panel (Figure 7-61).

![Image of storage pool view]

Figure 7-61   Filter the pools view

A few actions can be performed on storage pools using the **Actions** menu. A pool can be **renamed**, its **icon** can be changed, and it can be deleted from here also.

When you expand a pool's entry by clicking the plus (+) sign to the left of the pool's icon, you get access to the MDisks that are associated with this pool. You can perform all actions on them, as described in 7.2, “Working with MDisks” on page 220.

In its default view, each defined storage pool is displayed along with its assigned icon and name, status, graphical indicator showing the pool's capacity allocated to volumes, mode of the containing MDisks, storage system, and MDisk LUN ID.

To create a storage pool, click the **+Create Pool** button in the upper right corner of the MDisks by Pools view and follow the steps in 7.2.1, “Adding MDisks to storage pools” on page 222.

**Note:** If you want to have more control over storage pool creation and set different than default parameters, such as extent size or Easy Tier setting, you have to use CLI command `mkmdiskgrp`. 
To remove a storage pool that is not needed, click **Actions → Delete Pool** (Figure 7-62).

The confirmation window opens (Figure 7-63).

**Note:** If there are volumes with their only copy on MDisk in the pool to be deleted, you will not be able to remove the pool using the GUI and you will need to use the CLI command `rmmdiskgrp` with the `-force` parameter instead.

If the pool contains MDisk made of internal drives, Storwize informs that MDisk will be destroyed and all physical disk drives will become candidates.

**Attention:** After you delete the pool, all the data in the pool is lost except for the image-mode MDisk. Their volume definition is deleted, yet the data on the imported MDisk remains untouched.

After you delete the pool, all the associated volumes and their host mappings are removed. All the managed or image-mode MDisk in the pool return to a status of unmanaged after the pool is deleted. If the pool is deleted, all the array mode MDisk in the pool are removed and all the member drives return to candidate status.
7.3.1 Child storage pools

In version 7.4 of Storwize V7000, a new storage pool option was introduced called child storage pool. A child storage pool is a smaller storage pool within a bigger storage pool, called the parent storage pool. The parent storage pool is a normal storage pool and the child storage pool is created within the parent storage pool and uses the storage of the parent storage pool.

Unlike the parent (or normal) storage pool, child storage pool does not contain MDisks but uses only extents. The size of the parent storage pool is defined by the sum of MDisks included in the pool and can only be changed by adding or removing an MDisk to or from the pool. On other hand, the size of the child storage pool is set at the time of child pool creation and can be modified later. Because the child storage pool does not contain any MDisks, it does not fall into MDisk boundaries. The size of the child storage pool is a multiple of parent pool extents and can be set to a value between zero and the parent pool size.

The child storage pool is built out of extents, which means it uses some extents of all MDisks included in the parent storage pool. This means that the child storage pool might be much smaller than the parent pool, and it can use extents from all MDisks contained in the parent pool. This, of course, means the child pool can benefit from using all physical drives forming MDisks in the parent pool.

The number of storage pools is constant and equals 128 in v7.4 of Storwize V7000 code. This number can be divided by parent and child pools or parent pools only. You can have only one parent pool and 127 child pools, or you can have 128 parent pools, or 16 parent pools each with 8 child pools if there is a need for such configuration.

Child storage pools do not cause cache partitioning. This may be desirable in environments where users want to have many pools, but at the same time they want to avoid cache partitioning, which is a case in configuration of many normal storage pools. To avoid cache partitioning and still benefit from using all possible MDisks, simply create one big parent pool and many child pools inside the parent pool. In such a configuration, cache partitioning will not activate because there is only one parent pool and it can use all the cache available in Storwize V7000. At the same time, all child pools will use all parent pools MDisks so performance will not be compromised.

Child storage pools can be created only by using the CLI. Although you can view child pools from the GUI, creation and modification using the GUI are not supported. After creation of a child storage pool inside a parent pool, you can create volumes inside the child pool as though it was a normal pool. In case you have volumes already created in other storage pools, you can move them to the storage pool by using the volume mirroring feature, in the same way you mirror normal volumes between normal pools. Migration of volumes to or from storage pools is not supported. Use volume mirroring instead.

In the following sections, we show how to manage child storage pools.

Creating child storage pool

As stated before, child storage pools can be created only by using the CLI. We start by creating a normal storage pool that will be used as the parent storage pool. If you have a normal storage pool already, you can use it as a parent storage pool as well:

1. Log in to Storwize V7000 via Secure Shell (SSH) and list unmanaged MDisks (Example 7-2 on page 242).
Example 7-2 Login

```
IBM_Storwize:V7000 Gen2_2:superuser>lsmdisk
id  name   status mode     mdisk_grp_id mdisk_grp_name capacity ctrl_LUN_#
controller_name tier       encrypt
0  mdisk0 online unmanaged                             64.0GB   0000000000000000
controller0     enterprise no
1  mdisk1 online unmanaged                             64.0GB   0000000000000001
controller0     enterprise no
2  mdisk2 online unmanaged                             64.0GB   0000000000000002
controller0     enterprise no
3  mdisk3 online unmanaged                             64.0GB   0000000000000003
controller0     enterprise no
4  mdisk5 online array     0            Encryption     558.4GB
enterprise yes
5  mdisk4 online array     0            Encryption     558.4GB
enterprise yes
```

We have four unmanaged MDisks mapped from the external storage controller to Storwize V7000. We use them to create a parent storage pool.

2. Create a normal storage pool out of unmanaged MDisks. Use this pool as the parent pool (Example 7-3).

Example 7-3 Create normal storage pool

```
IBM_Storwize:V7000 Gen2_2:superuser>mkmdiskgrp -name test_parent_pool -mdisk 0:1:2:3 -ext 1024
MDisk Group, id [1], successfully created
```

You can use the `lsmdiskgrp` command or the GUI to see the properties of a created storage pool as shown in Figure 7-64.

![Figure 7-64 Created storage pool to be used as parent storage pool](image)

As you can see, this new pool has four MDisks, which equates to 256 GB. Zero bytes are allocated so far.

3. Create a child storage pool within a previously created parent storage pool (Example 7-4).

Example 7-4 Create child pool

```
IBM_Storwize:V7000 Gen2_2:superuser>mkmdiskgrp -name test_child_pool -size 100 -unit gb -parentmdiskgrp test_parent_pool
MDisk Group, id [2], successfully created
```
As before, you can use the `lsmdiskgrp` command to list the new pool via the CLI. But you will not see the child pool using the GUI by default. 100 GB is now allocated from the parent pool for use by the child pool, but there is no child pool in default view.

4. To see the child pool, right-click the parent pool and select **Child Pools** as shown in Figure 7-65.

![Figure 7-65 Listing child pools menu](image)

After selecting this option, you can see all created child pools as shown on Figure 7-66.

![Figure 7-66 Listing available child pools](image)

The GUI provides basic information about child storage pools. For more details, use the `lsmdiskgrp name_or_ID_of_the_child_storage_pool` command in the CLI.

**Modifying child storage pool**

With child pools you can change all the settings like name, Easy Tier, or warning threshold the same way you change it with normal storage pools. The one difference is you can change the size of the child pool, which is not possible in normal or parent pool. To increase or decrease the size of the storage pool use the `chmdiskgrp -size` command (Example 7-5 on page 244).
Example 7-5  Change size

IBM_Storwize:V7000 Gen2_2:superuser>chmdiskgrp -size 150 -unit gb test_child_pool

We increased the size of the child pool so that now 150 GB is allocated from the parent pool and is available for use by volumes created in the child storage pool (Figure 7-67).

Figure 7-67  Increased child pool size

In the same way, you can decrease the child pool size. In the following example, we shrunk our child pool by 100 GB to 50 GB (Example 7-6).

Example 7-6  Change size

IBM_Storwize:V7000 Gen2_2:superuser>chmdiskgrp -size 50 -unit gb test_child_pool

You can see the changes in the CLI or the GUI as shown in Figure 7-68.

Figure 7-68  Decreased child pool size

You cannot increase the child pool size beyond the parent pool size. Also, if you create some volumes inside the child pool you cannot decrease the child pool size below the used capacity. To show this, we increased the child pool size again to 150 GB and created a volume of 100 GB in size, so the used capacity is 100 GB and free capacity is 50 GB as seen on Figure 7-69 on page 245.
Figure 7-69  100 GB Volume created inside child pool

Then, we tried to shrink the child pool by 100 GB as shown in Example 7-7.

Example 7-7  Shrink child pool size

IBM_Storwize:V7000 Gen2_2:superuser>chmdiskgrp -size 50 -unit gb test_child_pool
  CMMVC8524E The command failed because the child mdiskgrp can not be shrunk below its used capacity.

Deleting a child pool
Deleting a child pool is a task very similar to deleting a normal pool. You have to use the rmmdiskgrp command as shown in Example 7-8.

Example 7-8  Delete child pool

IBM_Storwize:V7000 Gen2_2:superuser>rmmdiskgrp test_child_pool
IBM_Storwize:V7000 Gen2_2:superuser>

You cannot delete a child pool when there are still some volumes present. You have to remove, or move, by using volume mirroring, the volumes before the child pool can be deleted. Failing to do so results in the following error (Example 7-9).

Example 7-9  Example error message

IBM_Storwize:V7000 Gen2_2:superuser>rmmdiskgrp test_child_pool
  CMMVC5818E The managed disk group was not deleted because there is at least one MDisk in the group.
Chapter 8. Advanced host and volume administration

The IBM Storwize V7000 offers many functions for volume and hosts creation and configuration. In Chapter 4, “Host configuration” on page 111 and in Chapter 5, “Basic volume configuration” on page 145, the basic host and volumes features of IBM Storwize V7000 are covered. Those chapters show how to create hosts and volumes, and how to map volumes to a host. This chapter describes other advanced host and volume administration tasks apart from Easy Tier, Compression, or Encryption, which are described in Chapter 10, “Advanced features for storage efficiency” on page 335 and Chapter 12, “Encryption” on page 467.

This chapter contains the following topics:

- Advanced host administration
- Advanced volume administration
8.1 Advanced host administration

This section covers host administration, including topics such as host modification, host mappings, and deleting hosts. Basic host creation using Fibre Channel (FC) and Internet Small Computer System Interface (iSCSI) connectivity is described in Chapter 4, “Host configuration” on page 111.

We assume that you created several hosts in your IBM Storwize V7000 GUI and that some volumes are already mapped to them. We describe the three functions that are covered in the Hosts section of the IBM Storwize V7000 GUI (Figure 8-1):

- Hosts (8.1.1, “Modifying Mappings menu” on page 250).
- Ports by Host (8.2, “Adding and deleting host ports” on page 267).
- Host Mappings (8.3, “Host mappings overview” on page 275).

Figure 8-1   IBM Storwize V7000 Hosts menu

If you hover the mouse cursor over the Hosts selection in the IBM Storwize V7000 dynamic menu, the Hosts panel opens (Figure 8-2 on page 249).
As our example shows, six hosts are created and volumes are already mapped to some of them. We use these hosts to show the modification possibilities.

If you highlight a host, you can either click **Actions** (Figure 8-3) or right-click the host (Figure 8-4) to see all the available tasks.
8.1.1 Modifying Mappings menu

On the Hosts panel, select a host and click **Actions → Modify Mappings** (Figure 8-3 on page 249). The window shown in Figure 8-5 opens. At the upper left, you see that the highlighted host is selected. The two list boxes show all available unmapped and mapped volumes. The left list includes the volumes that are ready for mapping to this host; the right list includes the volumes already mapped. In our example, one volume with SCSI ID 0 is mapped to the host win_iscsi, and 21 more volumes are available (unmapped).

![Modify Host Mappings window](image)

**Figure 8-5  Modify Host Mappings window**

You can easily identify if the volume you want to map is already mapped to another host. The already mapped volumes on the left pane have a green arrow and host icon by them. When you hover a mouse cursor over this icon, the little box with the **Mapped** sign shows as seen in Figure 8-6 on page 251.
Figure 8-6  Volumes already mapped to other hosts

To map a volume, select it in the left pane, and then click the right arrow (>>) to move the volume to the list on the right, where mapped volumes are placed (Figure 8-7).

Figure 8-7  Modify Host Mappings: map volume
The changes are marked in yellow and now the Map Volumes and Apply buttons are enabled. If you click **Map Volumes**, the changes are applied (Figure 8-8) and the window closes. If you click **Apply**, the changes are submitted to the system (Figure 8-8), but the window remains open for further changes (Figure 8-9).

![Modify Mappings task completion window](image1)

**Figure 8-8** Modify Mappings task completion window

![Modify Mappings: applied changes](image2)

**Figure 8-9** Modify Mappings: applied changes
You can now select another host in the Host drop-down menu (Figure 8-10) to modify the host settings for it or continue working with the one that is selected (Figure 8-9 on page 252).

![Figure 8-10 Modify another hosts mappings](image)

Select the volume to be modified, again, and click the right arrow to move the volume to the list in the right pane. The changes are shown in yellow there. If you right-click the highlighted volume, you are able to change the SCSI ID, which is used for the host mapping (Figure 8-11). Select **Edit SCSI ID**.

![Figure 8-11 Edit the SCSI ID](image)

Enter a SCSI ID and click **OK** to change the ID (Figure 8-12 on page 254).
Implementing the IBM Storwize V7000 V7.4

Figure 8-12  Enter a new SCSI ID

The changes are shown in the Modify Host Mappings window (Figure 8-13).

Figure 8-13  Modify Host Mappings

Click **Apply** to submit the changes. The resulting output is shown in Figure 8-14.

Figure 8-14  Modify Mappings task completion window

**Note:** You can change the volume SCSI ID only before it is mapped to a host. Changing it after is not possible unless you unmap it again.
If you want to remove a host mapping, the steps are similar, except that you select a volume in the right pane and click the left arrow (<<) button to remove the mapping (Figure 8-15).

![Figure 8-15  Modify Host Mappings: remove mapping](image)

Figure 8-16 shows that the selected volume has been moved to the left pane for unmapping.

![Figure 8-16  Modify Host Mappings: mapping removed](image)

Click **Apply** to submit the changes to the system (Figure 8-17 on page 256).
Figure 8-17  Modify Host Mappings: removal complete

After you are done with all host mapping modifications, click **Close** to return to the Modify Mappings window (Figure 8-5 on page 250).

### 8.1.2 Unmapping all volumes from a host

A host is able to access only those volumes on your IBM Storwize V7000 system that are mapped to it. If you want to remove access to all volumes for one host, regardless of how many volumes are mapped to it, you can do this task in one step. From the Hosts panel, select the host, click **Actions → Unmap All Volumes** to remove all access that this host has to its volumes (Figure 8-18).

Figure 8-18  Unmap All Volumes action

You are prompted about the number of mappings you want to remove. Enter the number and click **Unmap** (Figure 8-19 on page 257). In our example, we remove two mappings.
**Unmapping:** If you click **Unmap**, all access for this host to volumes that are controlled by IBM Storwize V7000 system is removed. Ensure that you run the required procedures on your host operating system, such as unmounting file system, taking disk offline, or disabling volume group, before removing the volume mappings from your host object on IBM Storwize V7000.

Figure 8-19 Enter the number of mappings to be removed

The changes are applied to the system (Figure 8-20). Click **Close** after you review the output.

Figure 8-20 Unmap all Volumes from Host task completion window

Figure 8-21 on page 258 shows that the selected host does not have host mappings anymore.
8.1.3 Duplicating and importing mappings

Volumes that are assigned to a host can be mapped to another host object. You can do this for example when you add a new node to the host cluster and want to ensure that the new host node has access to the same set of volumes as the source host.

Verify the mappings on the existing source host object. From the Hosts Actions menu (Figure 8-3 on page 249), right-click the host, select Properties, and click the Mapped Volumes tab (Figure 8-22).

Select a host whose mappings you want to duplicate, and then click Actions → Duplicate Mappings (Figure 8-23).

The Duplicate Mappings window opens. Select a target host object to which you want to add all the existing source host mappings and click Duplicate (Figure 8-24 on page 259).
After the task completion is displayed (Figure 8-25), verify the new mappings on the target host object. From the Hosts Actions menu (Figure 8-3 on page 249), right-click the target host, select **Properties**, and click the **Mapped Volumes** tab (Figure 8-26 on page 260).
You can perform the same action from the Actions menu of the target host object. You can also import existing source host mappings provided the target host has no existing mappings defined.

Verify that no mappings are on the target host object. From the Hosts Actions menu (Figure 8-3 on page 249), right-click the host, select Properties, and then click the Mapped Volumes tab (Figure 8-27).

You can duplicate mappings only to a host that has no volumes mapped.

From the Hosts Actions menu, select a host to which you want to import existing mappings, and then click Actions → Import Mappings (Figure 8-28 on page 261).
Chapter 8. Advanced host and volume administration

The Import Mappings window opens. Select the source host object from the drop-down list and click Import (Figure 8-29).

After the task completion window is displayed (Figure 8-30), verify the new mappings on the target host object. From the Hosts Actions menu (Figure 8-3 on page 249), right-click the target host, select Properties, and click the Mapped Volumes tab (Figure 8-31 on page 262).
8.1.4 Renaming

To rename a host, select it, and then right-click and select Rename (Figure 8-32).

Enter a new name and click Rename (Figure 8-33). If you click Reset, the changes are reset to the original host name.

After the changes are applied to the system, click Close (Figure 8-34 on page 263).
8.1.5 Removing a host

To remove a host, from the IBM Storwize V7000 Hosts panel, select the host and right-click it or click **Actions → Remove** (Figure 8-35).

Confirm the number of hosts that you want to remove and click **Delete** (Figure 8-36 on page 264).
If you want to remove a host that has volumes mapped, you must force the removal by selecting the check box in the lower part of the window. If you select this check box, the host is removed and it no longer has access to this system.

After the task is completed, click **Close** (Figure 8-37) to return to the mappings window.

8.1.6 Host properties

From the IBM Storwize V7000 Hosts panel, select the host and right-click it or click **Actions → Properties** (Figure 8-38 on page 265).
Chapter 8. Advanced host and volume administration

Figure 8-38 Host properties

The Host Details window opens (Figure 8-39).

Figure 8-39 Host properties overview

The Host Details window shows an overview of your host properties. It has three tabs: Overview, Mapped Volumes, and Port Definitions.

The Overview tab is shown in Figure 8-39. Select the **Show Details** check box to see more information about the host (Figure 8-40 on page 266).
Click **Edit** if you want to change some host properties (Figure 8-41).
In this window, you can modify the following items:

- **Host Name**: Change the host name.
- **Host Type**: Change this setting if you are going to attach HP/UX, OpenVMS, or TPGS hosts.
- **I/O Group**: Host will have access to volumes mapped from selected I/O Groups.
- **iSCSI Challenge Handshake Authentication Protocol (CHAP) Secret**: Enter or change the iSCSI CHAP secret for this host.

When you finish making changes (if required), click **Save** to apply them (Figure 8-41 on page 266). The editing window closes.

The Mapped Volumes tab shows an overview of which volumes are currently mapped with which SCSI ID and UID to this host (Figure 8-42). The Show Details check box does not show any additional information.

![Figure 8-42 Mapped volumes tab](image)

The Port Definitions tab shows the configured host ports of a host and gives you status information about them (Figure 8-43).

![Figure 8-43 Port definitions](image)

This window offers you the option to start Add and Delete Port actions, as described in 8.2, “Adding and deleting host ports” on page 267.

Click **Close** to close the Host Details window.

### 8.2 Adding and deleting host ports

To configure host ports, from the IBM Storwize V7000 dynamic menu hover over **Hosts** and select **Ports by Host** to open the associated panel (Figure 8-44 on page 268).
The left pane lists all the hosts; the function icons indicate whether the host is Fibre Channel (orange cable) or iSCSI (blue cable). The properties of the highlighted host are shown in the right pane. If you click **New Host**, the wizard starts; it is described in Chapter 4, “Host configuration” on page 111. If you click **Actions** (Figure 8-45), the tasks described in 8.1.1, “Modifying Mappings menu” on page 250 can be started from this location.

### 8.2.1 Adding a Fibre Channel or iSCSI host port

To add a host port, highlight the host, click **Add** (Figure 8-46), and select Fibre Channel Port (see “Adding a Fibre Channel port”) or an iSCSI Port (see “Adding an iSCSI host port”).

**Adding a Fibre Channel port**

Click **Fibre Channel Port** (Figure 8-46). The Add Fibre Channel Ports window opens (Figure 8-47 on page 269).
If you click the drop-down menu, a list of all known Fibre Channel host ports (Figure 8-48) is displayed. If the worldwide port name (WWPN) of your host is not available in the menu, check your SAN zoning and rescan the SAN from the host. Then, click Rescan; the new port is now available in the menu.

Select the WWPN you want to add and click Add Port to List (Figure 8-49 on page 270).
You can repeat this step to add more ports to a host.

If you want to add an offline port (if the WWPN of your host is not available in the drop-down menu), manually enter the WWPN of the port into the Fibre Channel Ports field and click Add Port to List. The port is unverified (Figure 8-50) because it is not logged on to the IBM Storwize V7000. The first time that it logs on, its state is automatically changed to online and the mapping is applied to this port.
To remove a port from the list, click the red X next to the port (Figure 8-51). In this example, delete the manually added FC port so only the detected port remains.

![Figure 8-51  Remove a port from a list](image)

Click **Add Ports to Host** and the changes are applied (Figure 8-52).

![Figure 8-52  Add Ports to Host task completion window](image)

Click **Close** to return to the Ports to Host window.
Adding an iSCSI host port
To add an iSCSI host port, click **iSCSI Port** (Figure 8-46 on page 268). The Add iSCSI Ports window opens (Figure 8-53).

Enter the initiator name of your host (Figure 8-54) and click **Add Port to List**.

---

**Figure 8-53**  Add iSCSI host ports

**Figure 8-54**  Enter the initiator name
Click **Add Ports to Host** (Figure 8-55).

![Add iSCSI Ports](image)

**Figure 8-55**  Add ports to the ports definitions

The tasks are completed and changes to the system are applied (Figure 8-56). Click **Close** to return to the Ports by Host window.

![Add Ports to Host](image)

**Figure 8-56**  Add Ports to Host task completion window
8.2.2 Deleting a host port

To delete a host port, highlight it and right-click it or click **Delete Port** (Figure 8-57).

![Figure 8-57 Delete host port](image)

You can also press the Ctrl key while you select several host ports to delete (Figure 8-58).

![Figure 8-58 Delete several host ports](image)

Click **Delete** and enter the number of host ports you want to remove (Figure 8-59).

![Figure 8-59 Enter the number of host ports to delete](image)
Click **Delete** to apply the changes to the system (Figure 8-60).

![Image of Delete Host Ports task completion window]

**Figure 8-60  Delete Host Ports task completion window**

Click **Close** to return to the Host by Ports window.

**Note:** Deleting FC and iSCSI ports is done the same way.

### 8.3 Host mappings overview

Hover mouse cursor over the host menu and select **Host Mappings** (Figure 8-1 on page 248) to open the host mappings panel (Figure 8-61 on page 276).
This panel lists all hosts and volumes. Our example shows that the host ITSO_SVC_CF8 has three mapped volumes, and their associated SCSI ID, Volume Name, and Volume Unique Identifier (UID). If you have more than one caching I/O group, you also see which volume is handled by which I/O group.

If you select one line and click **Actions** (Figure 8-62), the following tasks are available:

- Unmap Volumes
- Properties (Host)
- Properties (Volume)

### 8.3.1 Unmapping a volume

Select one or more lines with the Ctrl key, click **Unmap Volumes**, enter the number of entries to remove (Figure 8-63 on page 277), and then click **Unmap**.
This action removes the mappings for all selected entries (Figure 8-64).

8.3.2 Properties (Host)

Selecting an entry and clicking **Properties (Host)**, as shown in Figure 8-62 on page 276, opens the host properties window. The contents of this window are described in 8.1.6, “Host properties” on page 264.

8.3.3 Properties (Volume)

Selecting an entry and clicking **Properties (Volume)**, as shown in Figure 8-62 on page 276, opens the volume properties view. The contents of this window are described in Chapter 5, “Basic volume configuration” on page 145.
8.4 Advanced volume administration

This section covers volume administration, such as volume modification and the migration or creation of new volume copies. Basic volume creation is covered in Chapter 5, “Basic volume configuration” on page 145. In this section, we assume that you already created some volumes in your IBM Storwize V7000 and that you are familiar with generic, thin-provisioned, and mirrored volumes.

Figure 8-65 shows that three volumes options are available to administer advanced features:
- Volumes (8.4.1, “Advanced volume functions” on page 278 and 8.6, “Advanced volume copy functions” on page 305)
- Volumes by Pool (8.7, “Volumes by storage pool” on page 315)
- Volumes by Host (8.8, “Volumes by Host” on page 320)

8.4.1 Advanced volume functions

Click Volumes (Figure 8-65); the Volumes panel opens (Figure 8-66).
This window lists all configured volumes on the system and provides the following information:

- **Name**: Shows the name of the volume. A plus sign (⁺) next to the name means that several copies of this volume exist. Click it to expand the view and list the copies. Primary copy is indicated by a little star sign (Figure 8-67).

- **State**: Gives you information about the volume state, which can be online, offline, or degraded.

- **Capacity**: The capacity that is presented to the host is listed here. A blue volume icon listed next to the capacity means that this volume is thin-provisioned, and that the listed capacity is the virtual capacity, which might be more than the real capacity on the system. Similarly, the compressed volume icon indicates that the data on the volume is compressed by the Real-time Compression feature.

- **Storage Pool**: Shows in which storage pool the volume is stored. If you have several volume copies, it shows you the pool of the primary copy.

- **Host Mappings**: Shows in which storage pool the volume is stored. If you have several volume copies, it shows you the pool of the primary copy.

- **UID**: This is the volume unique identifier.

To create a volume, click **New Volume** and complete the steps described in Chapter 5, “Basic volume configuration” on page 145.

Select a volume and right-click it or select **Actions** to see the available actions for a volume (Figure 8-68).

The following volume options are available:

- **Create Volumes** (Chapter 5, “Basic volume configuration” on page 145)
- **Map to Host** (“Mapping volume to a host” on page 280)
8.4.2 Creating a volume

Volume creation is described in detail in Chapter 5, “Basic volume configuration” on page 145.

8.4.3 Mapping volume to a host

Mapping volumes to hosts and modifying those mappings is described in Chapter 5, “Basic volume configuration” on page 145 and 8.1.1, “Modifying Mappings menu” on page 250.

8.4.4 Unmapping volumes from all hosts

Unmapping volumes from all hosts is similar to unmapping all volumes from a host described in 8.1.2, “Unmapping all volumes from a host” on page 256. But instead of unmapping many volumes from one host, you can unmap one volume from many hosts.

8.4.5 Viewing a host mapped to a volume

If you want to know which host mappings are currently configured, highlight a volume and click View Mapped Host (Figure 8-68 on page 279). This action opens the Host Maps tab of the Volume Details window (Figure 8-69 on page 281). This example shows one existing host mapping for the test_vol1 volume.
To remove a mapping, select the host and click Unmap from Host, which removes the access for the selected host (after you confirm it). If several hosts are mapped to this volume (for example, in a cluster), only the highlighted host is removed.

8.4.6 Duplicating a volume

You can create a new volume using the same presets and parameters as an existing volume. These parameters are as follows:

- Volume preset (generic, thin-provision, and compressed)
- Volume size
- Storage pool
- Access and Caching I/O group
- Caching mode
- Easy Tier status
- Virtualization type

Note: Duplicating a volume creates a new volume with the same preset and volume parameters as the source volume. Duplicating a volume does not duplicate volume data. Duplicating mirrored and image-mode volumes is not supported.

To duplicate a volume, select Duplicate Volume from the Actions menu (Figure 8-68 on page 279). The Duplicate Volume window opens (Figure 8-70). By default, a sequence integer is appended to the name of the volume you are duplicating. You can change this name as shown in this figure. Click Duplicate.

Figure 8-70 Duplicate Volume window

Close the task window after the volume duplicate task is finished (Figure 8-71 on page 282).
8.4.7 Renaming a volume

To rename a volume, select Rename (Figure 8-68 on page 279). A window opens.

Enter the new name (Figure 8-72). If you click Reset, the name field is always reset to the currently active name of the volume. Click Rename to apply the changes.

Click Close when you are done (Figure 8-73 on page 283).
8.4.8 Shrinking a volume

IBM Storwize V7000 offers you the ability to shrink volumes. However, you should only use this feature if your host operating system supports this feature. Before shrinking a volume, complete the preparation required in your host operating system to shrink a volume on the storage system. After you have prepared your OS, click Shrink (Figure 8-68 on page 279). You can either enter the new size, or enter how much the volume should shrink. If you enter a value, the other line updates itself (Figure 8-74). Click Shrink to start the process.

Click Close (Figure 8-75 on page 284) to return to the Volumes window.
Implementing the IBM Storwize V7000 V7.4

8.4.9 Expanding a volume

To expand a volume, click **Expand** (Figure 8-68 on page 279); the Expand Volume window opens. Before you continue, determine whether your operating system supports online volume expansion. Enter the new volume size and click **Expand** (Figure 8-76).

After the tasks are completed, click **Close** (Figure 8-77 on page 285) to return to the Volumes panel.
8.4.10 Migrating a volume to another storage pool

IBM Storwize V7000 provides online volume migration while applications are running. Storage pools are managed disk groups, as described in Chapter 7, “Storage pools” on page 203. With volume migration, you can move the data between these 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 itself is a low priority process that does not affect the performance of the IBM Storwize V7000. 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, click Migrate to Another Pool (Figure 8-68 on page 279). The Migrate Volume Copy window opens. If your volume consists of more than one copy, select the copy (from the menu shown in Figure 8-78 on page 286) that you want to migrate to another storage pool. If the selected volume consists of one copy, this selection menu is not available.
Implementing the IBM Storwize V7000 V7.4

Select the new target storage pool and click **Migrate** (Figure 8-79).

The volume copy migration starts (Figure 8-80 on page 287). Click **Close** to return to the Volumes panel.
Depending on the size of the volume, the migration process takes some time, but you can monitor the status of the migration in the running tasks bar at the bottom of the GUI (Figure 8-81).

After the migration is completed, the volume is shown in the new storage pool. Figure 8-82 shows that it was moved from the DS3400_pool to the test_pool.

The volume copy has now been 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 using the volume copies, as described in 8.6.6, “Migrating volumes using the volume copy features” on page 314.

**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.

### 8.4.11 Exporting to an image mode volume

Exporting to an image mode volume provides a direct block-for-block translation from the MDisk to the volume with no virtualization. An image mode MDisk is associated with exactly one volume. This feature can be used to export a volume to a non-virtualized disk and to remove the volume from storage virtualization, for example, to map it directly from external storage system to host. If you have two copies of a volume, you can choose one to export to image mode. To export a volume to an image mode volume, click **Export to Image Mode** (Figure 8-68 on page 279), and the window opens (Figure 8-83). Select the MDisk to which you want to export the volume and click **Next**.

![Figure 8-83 Export to Image Mode: select copy](image)

**Note:** The target MDisk must be the same size or larger than source volume.

Select a storage pool for the new image mode volume and click **Finish** (Figure 8-84 on page 289). As stated before in 8.4.10, “Migrating a volume to another storage pool” on page 285, target storage pool has to have the same extent size as the source storage pool. Also, the target storage pool may be an empty pool, so the selected MDisks will be the target pool's only member at the end of migration procedure. But the target storage pool does not have to be empty. It can have other image mode or striped MDisks. In case you have image and striped MDisks in the same pool, volumes created in this pool will only use striped
MDisks because MDisks that are in image mode already have image mode volume created on top of them and cannot be used as an extent source for other volumes.

**Note:** In case of migration, create an empty migration pool just for migration purposes and delete this pool after all migration tasks complete. Using one pool for striped, sequential and image mode volumes is possible but not recommended.

The migration starts (Figure 8-85). Click **Close** to return to the Volumes panel.
8.4.12 Deleting a volume

To delete a volume, select **Delete** (Figure 8-68 on page 279). Enter the number of volumes you want to delete and select the check box if you want to force the deletion (Figure 8-86). You must force the deletion if the volume has host mappings or is used in FlashCopy mappings or Remote Copy relationships.

![Delete Volume](image)

**Figure 8-86  Delete Volume**

Click **Delete** and the volume is removed from the system (Figure 8-87 on page 291).

**Attention:** This action removes all copies from your storage system and the data on the volume is lost. Before you perform this step, ensure that you do not need the volume anymore.
8.4.13 Protecting volume on deletion

Storwize v7.4 introduces a new function when it comes to volume deletion called *Volume protection*. This new function prevents volume deletion if the volume in question has received any I/O during a specified period. This period can be tuned and set between 15 and 1440 minutes (24 hours).

This function not only protects volumes from deletion but also from unmapping from the host. For example, if your threshold is set to 30 minutes and the volume you want to delete has received any I/O within the preceding 30 minutes, you will not be able to delete the volume or unmap the volume from the host. This function was introduced to prevent any accidental removing or unmapping of volumes. While this is not so much the case when using the GUI, as the GUI always asks the administrator to confirm deletion, it could potentially be a problem in the CLI where you can remove any volume just by adding the `-force` parameter as the CLI does not prompt for confirmation.

Volume protection is a system-wide parameter, which means when turned on it will affect all volumes created in the Storwize V7000. It can be enabled or disabled online at any time but only by using the CLI; there is no GUI support to enable this feature. Additionally, there is a new field in the detailed volume view called `last_access_time`, which shows when the volume was last accessed.

To change volume protection settings, use the `chsystem` command (Example 8-1). The output is truncated for readability.

Example 8-1  Change volume protection

```
IBM_Storwize:ITSO_V7000Gen2_1:superuser>lssystem
id 000001002080020E
name ITSO_V7000Gen2_1
location Local
```
partnership
.
.
.
topology standard
topology_status
rc_auth_method none
vdisk_protection_time 15
vdisk_protection_enabled no
product_name IBM Storwize V7000

IBM_Storwize:ITSO_V7000Gen2_1:superuser>chsystem -vdiskprotectionenabled yes
-vdiskprotectiontime 20(Example )

IBM_Storwize:ITSO_V7000Gen2_1:superuser>lssystem
id 000001002080020E
name ITSO_V7000Gen2_1
location Local
partnership
.
.
.
topology standard
topology_status
rc_auth_method none
vdisk_protection_time 20
vdisk_protection_enabled yes
product_name IBM Storwize V7000

Now volume protection is enabled on the entire system. To see when the volume was last accessed, type the lsvdisk command (Example 8-2).

Example 8-2 Checking when volume was last accessed

IBM_Storwize:ITSO_V7000Gen2_1:superuser>lsvdisk test_vol1
id 14
name test_vol1
I0_group_id 0
I0_group_name io_grp0
status online
mdisk_grp_id 5
mdisk_grp_name DS3400_pool
capacity 1.00GB
type striped
formatted no
mdisk_id
mdisk_name
FC_id
FC_name
RC_id
RC_name
vdisk_UID 60050764008200083800000000000024
throttling 0
preferred_node_id 5
In our example, the volume `test_vol1` was last accessed 7th of December 2014 at 10:42:28, so if the system parameter `vdisk_protection_time` is set to 20, we will be able to delete or unmap this volume at the earliest at 11:02:28 assuming it will not receive any I/O during this 20-minute period.

If you try to delete or unmap a protected volume that is still receiving an I/O, the following error appears (Figure 8-88).

![Figure 8-88 Trying to unmap a protected volume](image_url)
8.5 Volume properties

To open the advanced view of a volume, select **Properties** (Figure 8-68 on page 279); the Volume Details window opens (Figure 8-89). In this window, three tabs are available:

- Overview tab
- Host Maps tab
- Member MDisks tab

![Figure 8-89 Volume properties](image)

8.5.1 Overview tab

The Overview tab (Figure 8-89) shows a basic overview of the volume properties. In the left part of the window, you find common volume properties, and in the right part, you see information about the volume copies (Figure 8-90 on page 295).
The following details are available:

- **Volume Properties:**
  - Volume Name: Shows the name of the volume.
  - Volume ID: Shows the ID of the volume. Every volume has a system-wide unique ID.
  - Status: Gives status information about the volume, which can be online, offline, or degraded.
  - Capacity: Shows the capacity of the volume. If the volume is thin-provisioned, this number is the virtual capacity; the real capacity is displayed for each copy separately.
  - # of FlashCopy Mappings: Number of existing FlashCopy relationships. A detailed description about FlashCopy is covered in Chapter 11, “Advanced Copy Services” on page 379.
  - Volume UID: The volume unique identifier.
  - Caching I/O Group: Specifies the I/O group to which the volume belongs.
  - Accessible I/O Groups: Specifies the I/O groups the volume can be moved to.
  - Preferred Node: Specifies the ID of the preferred node for the volume.
  - I/O Throttling: You can set a maximum rate at which the volume processes I/O requests. The limit can be set in IOPS or MBps. This is an advanced feature and you can enable it only through the CLI, as described in Appendix A, “CLI setup and SAN Boot” on page 545.
  - Mirror Sync Rate: After creation, or if a volume copy is offline, the mirror sync rate weights the synchronization process. Volumes with a high sync rate (100%) complete the synchronization faster than volumes with a lower priority. By default, the rate is set...
to 50% for all volumes. For percentage sync rate and corresponding MBps copying speed, see Table 5-1 on page 154.

- Cache Mode: Shows whether the cache is enabled or disabled for this volume.
- Cache State: Gives you feedback if open I/O requests are inside the cache that has not been destaged to the disks.
- UDID (OpenVMS): The unit device identifiers are used by OpenVMS host to access the volume.

- **Copy Properties:**
  - Pool: Gives you information about in which storage pool the copy rests, what copy type it is (generic, thin-provisioned, or compressed), and gives you volume and EasyTier status information.
  - Capacity: Shows the allocated (used) and the virtual (real) capacity, the warning threshold, and the grain size.

To edit any of these settings, click **Edit**; the window changes to the modify mode (Figure 8-91).

![Figure 8-91   Edit volume details](image)

In the Volume Details window, you can change the following properties:

- **Volume Name**
- **Accessible I/O Group**
- **Mirror Sync Rate**
- **Cache Mode**
- **UDID**
Make any required changes and click **Save** (Figure 8-92). For this example, we changed the Mirror Sync Rate.

![Volume overview after change](image)

**Figure 8-92** Volume overview after change

**Mirror Sync Rate:** Setting the Mirror Sync Rate to 0% disables synchronization.

### 8.5.2 Host Maps tab

The second tab of the volume properties is Host Maps (Figure 8-93). All hosts that are currently mapped to this volume are listed in this view.

![Volume Details: Host Maps tab](image)

**Figure 8-93** Volume Details: Host Maps tab

If you want to unmap a host from this volume, select it and click **Unmap from Host**. Confirm the number of mappings to remove and click **Unmap** (Figure 8-94 on page 298).
The changes are applied to the system (Figure 8-95). The selected host no longer has access to this volume. Click **Close** to return to the Volumes panel.

**Note:** If you have problems unmapping the volume, check your volume protection settings and volume last access time, as explained in 8.4.13, “Protecting volume on deletion” on page 291.
8.5.3 Member MDisks tab

The third tab is Member MDisks, which lists all MDisks on which the volume is located. If a volume has two copies, you can select a copy and see which MDisks each copy is based on. Select a copy and the associated MDisks are then listed in the window (Figure 8-96).

![Volume Details: Member MDisks](image)

Select an MDisk and click **Actions** to view the available tasks (Figure 8-97 on page 300). The tasks are described in Chapter 7, “Storage pools” on page 203.
Click Close to return to the Volumes panel.

### 8.5.4 Adding a mirrored volume copy

If you have an existing volume that consists of only one copy, you can add a mirrored volume copy to it (Figure 8-98 on page 301). This action creates an extra copy of your volume online. The second copy can be generic, thin-provisioned, or compressed, and it can be created in any storage pool. The process of adding a copy of each type is described in Chapter 5, “Basic volume configuration” on page 145.
8.5.5 Editing thin-provisioned or compressed volume properties

From a host’s perspective, the virtual capacity expansion and shrinkage of a volume affects the host access. The real capacity expansion or shrinkage of a volume, which is described in this section, is transparent to the hosts.

**Note:** In the following sections, we demonstrate real capacity operations; we use a thin-provisioned volume as an example. However, the same actions apply to the compressed volume preset.

Modifying the volume size that is presented to a host is covered in 8.4.8, “Shrinking a volume” on page 283 and 8.4.9, “Expanding a volume” on page 284. However, if you have a thin-provisioned or compressed volume, you can also edit the allocated (real) size and the warning thresholds. To accomplish this task, select the thin-provisioned or compressed volume, click **Actions → Volume Copy Actions → Thin Provisioned (or Compressed)** (Figure 8-99 on page 302). Three more options are available, as shown in Figure 8-99 on page 302:

- **Shrink:** See “Shrinking thin-provisioned or compressed volume space” on page 302.
- **Expand:** See “Expanding thin-provisioned or compressed volume space” on page 303.
- **Edit Properties:** See “Editing thin-provisioned or compressed volume properties” on page 304.

These changes are only to the internal storage, so you do not have to make any changes on your host.
Figure 8-99  Thin-provisioned actions

Shrinking thin-provisioned or compressed volume space
Select **Shrink** (Figure 8-99) to reduce the allocated space of a thin-provisioned or compressed volume. Enter either how much the volume should shrink or the new final size (Figure 8-100), and click **Shrink**.

Figure 8-100  Shrink Volume real capacity window

After the task completes, click **Close** (Figure 8-101 on page 303).
The allocated space of the thin-provisioned volume is now reduced.

**Deallocation extents**: You can only deallocate extents that do not have stored data on them. If the space is allocated because there is data on them, you are not able to shrink the allocated space.

**Expanding thin-provisioned or compressed volume space**

To expand the allocated space of a thin-provisioned or compressed volume, select **Expand** (Figure 8-99 on page 302). Enter either how much space should be allocated or enter the new final size and click **Expand** (Figure 8-102).

The new space is now allocated. Click **Close** (Figure 8-103 on page 304).
Editing thin-provisioned or compressed volume properties

To edit thin-provisioned or compressed volume properties, select Edit Properties (Figure 8-99 on page 302). Edit the settings if required (Figure 8-104) and click OK to apply the changes.

After the task completes, click Close (Figure 8-105 on page 305) to return to the Volumes panel.
8.6 Advanced volume copy functions

In 8.4.1, “Advanced volume functions” on page 278, we described all the available actions at a volume level and how to create a second volume copy. In this section, we focus on volumes consisting of two volume copies. If you expand the volume, highlight a copy and right-click it. More volume copy actions become available (as shown in Figure 8-106 on page 306):

- Create volumes
- Thin-provisioned or compressed (when selecting thin-provisioned or compressed volume)
- Make Primary (when selecting secondary copy)
- Split into New Volume
- Validate Volume Copies
- Duplicate Volume Copy
- Delete this Copy
If you look at the volume copies shown in Figure 8-106, you notice that one of the copies has an asterisk (*) next to its name (Figure 8-107).

Each volume has a primary and a secondary copy, and the star indicates the primary copy. After creation and initial synchronization, the two copies are always synchronized, which means that all writes are destaged to both copies, but all reads are always done from the primary copy. Copies are always called Copy0 and Copy1. You can easily switch between both copies online, without the host noticing that you have changed the physical storage that it reads from.

To accomplish this task, right-click the secondary copy and select Make Primary (Figure 8-108 on page 307).
Click **Make Primary** and the role of the copy is changed to online. Click **Close** when the task completes (Figure 8-109).

Usually, a preferred practice is to place the volume copies on storage pools with similar performance, although this is not necessary and depends mostly on configuration and wanted setup. Each volume has a parameter called *mirrorwritepriority*, which can have two settings: *latency* or *redundancy*.

Latency means a copy that is slow to respond to a write I/O that becomes unsynchronized, and write I/O competes if the other copy successfully writes the data. In other words, when one of the volume copies is slower than the other copy, the faster copy will process I/Os and acknowledge host completion without waiting for the slower copy.

Redundancy means that copies will always be synchronized and completes I/O when both copies finish writing the data. Latency provides a faster response to the host, and redundancy
makes sure that copies are always synchronized. Latency is the system default setting and can be changed only by using the CLI (Example 8-3).

Example 8-3  Changing volume parameter

```
IBM_Storwize:ITSO_V7000Gen2_1:superuser>lsvdisk test_voll
id 14
name test_voll
IO_group_id 0
IO_group_name io_grp0
status online
mdisk_grp_id many
mdisk_grp_name many
capacity 1.00GB
type many
formatted no
mdisk_id many
mdisk_name many
FC_id
FC_name
RC_id
RC_name
vdisk_UID 6005076400820083800000000000002A
throttling 0
preferred_node_id 5
fast_write_state empty
cache readwrite
uid
fc_map_count 0
sync_rate 90
copy_count 2
se_copy_count 1
filesystem
mirror_write_priority latency
RC_change no
compressed_copy_count 0
access_IO_group_count 1
last_access_time 141207115228
parent_mdisk_grp_id many
parent_mdisk_grp_name many

IBM_Storwize:ITSO_V7000Gen2_1:superuser>chvdisk -mirrorwritepriority redundancy
test_voll

IBM_Storwize:ITSO_V7000Gen2_1:superuser>lsvdisk test_voll
id 14
name test_voll
IO_group_id 0
IO_group_name io_grp0
status online
mdisk_grp_id many
mdisk_grp_name many
capacity 1.00GB
type many
formatted no
mdisk_id many
```
If you have both copies on disks with similar performance, which is the recommended setup, the mirrorwritepolicy setting is not so critical.

If you need only high-read performance, you might want to place the primary copy in an SSD pool, and the secondary copy in a normal disk pool and change the mirrorwritepriority to the redundancy setting. This action maximizes the read performance of the volume and makes sure that you have a synchronized second copy in your slower performing disk pool.

But if you need both high read and write performance and one copy is on SSD but the other is on slower disks, you might want to leave mirrorwritepriority at the default latency setting. This provides maximum read and write performance and at the same time makes sure that all data is consistent. It is also possible to migrate the copies online between storage pools; in 8.4.10, “Migrating a volume to another storage pool” on page 285, you can select which copy you want to migrate.

The volume copy feature is also a powerful option for migrating volumes, as described in 8.6.6, “Migrating volumes using the volume copy features” on page 314.
8.6.1 Thin-provisioned or compressed

These functions (Figure 8-110) are the same as described in the following sections. You can specify the same settings for each volume copy.

► “Shrinking thin-provisioned or compressed volume space” on page 302
► “Expanding thin-provisioned or compressed volume space” on page 303
► “Editing thin-provisioned or compressed volume properties” on page 304

Figure 8-110   Thin-provisioned volume copy tasks

8.6.2 Splitting into a new volume

If your two volume copies are synchronized, you can split one of the copies to a new volume and map this new volume to another host. From a storage point of view, this procedure can be performed online, which means you could split one copy from the volume, and create a copy from the remaining one without any host impact. However, if you want to use the split copy for testing or backup purposes, you must make sure that the data inside the volume is consistent. Therefore, you must flush the data to storage to make the copies consistent. See your operating system documentation for details about flushing the data. The easiest way to flush the data is to shut down the hosts before splitting a copy or close the application and unmount file systems.

To split a copy, click Split into New Volume (Figure 8-106 on page 306). If you perform this action on the primary copy, the remaining secondary copy automatically becomes the primary for the source volume. This happens online without interruption to host I/Os. Enter a name for the new volume and click Split Volume Copy (Figure 8-111).
After the task completes, click Close to return to the Volumes panel, where the copy appears as a new volume that can be mapped to a host (Figure 8-112).

Figure 8-112  All Volumes: New volume from split copy

Note: After splitting the volume copies, the new volume presents itself with a new UID. Compare the UIDs in Figure 8-110 on page 310 and Figure 8-112.

8.6.3 Validate Volume Copies

You must validate the volume copies in certain cases. An example is when one of the storage back-end arrays that is hosting one of the copies becomes unavailable or lost because of the planned or unplanned outage, storage array maintenance, or any other situation that causes the two copies to become out of sync.

To validate the copies of a mirrored volume, complete the following steps:

1. Select Validate Volume Copies (Figure 8-106 on page 306). The Validate Volume Copies window opens (Figure 8-113).

The following options are available:

- Generate Event of Differences: Use this option only if you want to verify that the mirrored volume copies are identical. If any difference is found, 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 being 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 either 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 volumes copies that 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. The volume is now being checked. Click **Close** (Figure 8-114).

![Figure 8-114  Repair Volume Copy task completion window](image)

The validation process takes some time depending on the volume size. You can check the status in the Running Tasks window (Figure 8-115 on page 313).
8.6.4 Duplicate Volume copy

You can create a new volume by using the same presets and parameters as an existing volume copy. The parameters are as follows:

- Volume preset (generic, thin-provision, and compressed)
- Volume size
- Storage pool
- Access and Caching I/O group
- Caching mode
- Easy Tier status
- Virtualization type

See 8.4.6, “Duplicating a volume” on page 281 for more information.

**Note:** Duplicating a volume creates a new volume with the same preset and volume parameters as the source volume. Duplicating a volume does not duplicate volume data. Duplicating mirrored and image-mode volumes is not supported. Duplicating a mirrored copy of a volume is supported.

8.6.5 Delete this Copy

To delete a volume copy, select **Delete this Copy** (Figure 8-106 on page 306). The copy is deleted, but the volume remains online by using the remaining copy. Confirm the deletion process by clicking **Yes** (Figure 8-116 on page 314).
Implementing the IBM Storwize V7000 V7.4

8.6.6 Migrating volumes using the volume copy features

The previous sections show that creating, synchronizing, splitting, and deleting volume copies are possible. 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 described in 8.4.10, “Migrating a volume to another storage pool” on page 285. If you use this feature, one extent after another is migrated to the new storage pool. However, using volume copies provides another possibility to migrate volumes. To migrate a volume, complete the following steps:

1. Create a second copy of your volume in the target storage pool (see 8.5.4, “Adding a mirrored volume copy” on page 300).
2. Wait until the copies are synchronized.
3. Change the role of the copies and make the new copy of the primary copy (see 8.6, “Advanced volume copy functions” on page 305).
4. Split or delete the old copy from the volume (see 8.6.2, “Splitting into a new volume” on page 310 or 8.6.5, “Delete this Copy” on page 313).
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.

8.7 Volumes by storage pool

To see an overview of which volumes are in which storage pool, hover the mouse cursor over the Volumes selection in the IBM Storwize V7000 dynamic menu and click Volumes by Pool (Figure 8-65 on page 278). The Volumes by Pool panel opens (Figure 8-118).

![Figure 8-118  Volumes by Pool panel](image)

The left pane lists all of your existing storage pools. You can filter them by using the Pool Filter.

In the upper right, you see information about the pool that you selected in the pool filter, and the following information is also there:

- Pool Icon: Storage Pools have different characteristics. It is possible to change the pool function icon to identify the pool type (see 8.7.1, “Changing the Storage Pool function icon” on page 317).
- Pool Name: This name is the name of the storage pool given during creation. It is possible to change this name from this window (see 8.7.2, “Changing the storage pool name” on page 319).
- Pool Details: Gives you status information about the pool, such as the number of MDisks and volume copies, and the Easy Tier status.
- Volume allocation: Provides details about the available, allocated, compressed, and virtual space in this pool.

The lower right lists all volumes that have at least one copy in this storage pool and provides the following information about them:

- Name: Shows the name of the volume.
实施 IBM Storwize V7000 V7.4

- **State**: 提供有关卷的状态信息。
- **容量**: 显示分配给卷的容量。一个**绿色**卷标志表示该卷是Thin-provisioned的，并且显示的容量是虚拟大小。同样，一个**压缩**图标表示卷使用了Real-time Compression特性。
- **主机映射**: 显示至少存在一个主机映射。
- **UID**: 显示卷的唯一标识符。

**注意**: 其他列可以右击蓝色标题行添加到此视图。

你也可以在选定的池中创建一个卷。点击**创建卷**。向导在第5章，“基本卷配置”第145页上描述。

如果你选择一个卷并点击**操作**，或者右击该卷（图8-119），从卷面板（第8.4.1节，“高级卷功能”第278页）中列出的相同选项也会显示。每个任务的详细说明都在该节中。

![Figure 8-119 Volumes by Pool actions](image)

如果你选择一个卷副本并点击**操作**，或者右击此卷（图8-120第317页），相同的选项也会显示。每个任务的详细说明也都在第8.6节，“高级卷副本功能”第305页中。

316  实施 IBM Storwize V7000 V7.4
8.7.1 Changing the Storage Pool function icon

Usually, storage pools have different characteristics; you can change the Storage Pool function icon to reflect these differences. Click the pool function icon; the Change Icon window opens (Figure 8-121).

Use the left and right arrows to select a new function icon (Figure 8-122 on page 318). Several options are available. Click OK.
The changes are applied to the system (Figure 8-123). Click Close.

The function icon has been changed so that identifying the pool is easier (Figure 8-124).
8.7.2 Changing the storage pool name

To change the given name of a storage pool, click the name. You can then edit it (Figure 8-125).

![Figure 8-125 Change pool name](image)

Enter a new name (Figure 8-126) and press Enter.

![Figure 8-126 Type new pool name](image)

The changes are applied to the system (Figure 8-127).

![Figure 8-127 Rename pool task completion window](image)

The name for the storage pool is now changed (Figure 8-128).

![Figure 8-128 Change pool name: new name](image)
8.8 Volumes by Host

To see an overview of which volumes a host is able to access, hover the mouse cursor over Volumes in the IBM Storwize V7000 dynamic menu and click Volumes by Host (Figure 8-65 on page 278). The Volumes by Host panel opens (Figure 8-129).

![Figure 8-129 Volumes by Host panel](image)

The left pane lists all host objects that are defined on the IBM Storwize V7000 system. You can use Host Filter to search for a specific host. When you select a host, its properties are listed in the right pane of the view. The hosts with the orange cable represent Fibre Channel hosts; the blue cable represents iSCSI hosts. The upper right side shows the host function icon, the host name, the number of host ports, and the host type. The volumes that are mapped to this host are listed in the pane below it.

If you want to create a volume for this host, click Create Volume. The wizard starts (this is the same wizard described in Chapter 5, “Basic volume configuration” on page 145).

Select a volume and click Actions or right-click the volume (Figure 8-130); the same options from the Volumes panel appear. Details about each task are covered in 8.4.1, “Advanced volume functions” on page 278.

![Figure 8-130 Volume actions](image)
If the volume owns more than one copy, you can select a volume copy and click Actions or right-click the copy (Figure 8-131), and the same options from the Volumes panel appear. Details about each task are covered in 8.6, “Advanced volume copy functions” on page 305.

8.8.1 Renaming a host

To rename a host in the Volumes by Host panel, click it and then edit the name (Figure 8-132).

Enter a new name (Figure 8-133) and press Enter.

The changes to the system are shown in Figure 8-134 on page 322.
Figure 8-134 Rename Host task completion window

Click Close to return to the Volumes by Host panel (Figure 8-135).

Figure 8-135 Volumes by Host panel
External storage virtualization

In this chapter, we describe how to incorporate external storage systems into the virtualized world of the IBM Storwize V7000. A key feature of IBM Storwize V7000 is its ability to consolidate disk controllers from various vendors into pools of storage. In this way, the storage administrator can, from a single user interface, manage and provision storage to applications, and use a common set of advanced functions across all the storage systems under the control of the IBM Storwize V7000.
9.1 Planning for external storage virtualization

In this section, we describe how to plan for virtualizing external storage with IBM Storwize V7000. Virtualizing the storage infrastructure with IBM Storwize V7000 makes your storage environment more flexible, cost-effective, and easy to manage. The combination of IBM Storwize V7000 and an external storage system allows more storage capacity benefits from the powerful software function within the IBM Storwize V7000.

The external storage systems that are incorporated into the IBM Storwize V7000 environment can be new or existing systems. The data on existing storage systems can be easily migrated to the IBM Storwize V7000 managed environment, as described in Chapter 6, “Migration wizard” on page 189 and Chapter 7, “Storage pools” on page 203.

The following web page has the latest supported hardware list, device drivers, firmware, and recommended software levels for IBM Storwize V7000:
http://www.ibm.com/support/docview.wss?uid=ssg1S1004450#_Controllers

9.1.1 License for external storage virtualization

From a licensing standpoint, when external storage systems are virtualized by IBM Storwize V7000, a per-enclosure External Virtualization license is required. You can find more licensing information in Chapter 1, “IBM Storwize V7000 system overview” on page 1, and contact your IBM account team or IBM Business Partner for further assistance, if required.

Migration: If the IBM Storwize V7000 is used as a general migration tool, the appropriate External Virtualization licenses must be ordered. The only exception is if you want to migrate existing data from external storage systems to IBM Storwize V7000 internal storage, as you can temporarily configure your External Storage license within 45 days. For a more than 45-day migration requirement from external storage to IBM Storwize V7000 internal storage, an appropriate External Virtualization license must be ordered.
You can configure the IBM Storwize V7000 licenses by going to the Settings section in the IBM Storwize V7000 dynamic menu and then clicking **Settings** (Figure 9-1).

![Figure 9-1   Settings, General panel](image)

In the Settings on the left, click System and then **Licensing**. The Update License view opens in the right pane (Figure 9-2). This Update License view shows three license options that you can set. Set these options to the limit you obtain from IBM:

- External Virtualization
- Remote Copy
- Real-time Compression

![Figure 9-2   Update License window](image)

For assistance with licensing questions or to purchase an External Virtualization, Remote Copy, or Real-time Compression license, contact your IBM account team or IBM Business Partner. Encryption Licenses are not discussed in this chapter because it is not applicable for the external storage. For more about Encryption, refer to Chapter 12, “Encryption” on page 467.
9.1.2 SAN configuration planning

External storage controllers virtualized by IBM Storwize V7000 must be connected through SAN switches. A direct connection between the IBM Storwize V7000 and storage controllers or hosts ports is not supported.

Ensure that the switches or directors are at the firmware levels supported by the IBM Storwize V7000 and that the IBM Storwize V7000 port login maximums listed in the restriction document will not be exceeded. The configuration restrictions can be found by navigating to the Support home page at the following address:


The recommended SAN configuration is composed of a minimum of two fabrics. The ports on external storage systems that will be virtualized by the IBM Storwize V7000 and the IBM Storwize V7000 ports are evenly split between the two fabrics to provide redundancy if one of the fabrics goes offline.

After the IBM Storwize V7000 and external storage systems are connected to the SAN fabrics, SAN zoning must be implemented. In each fabric, create a zone with the four IBM Storwize V7000 port worldwide port names (WWPNs), two from each node canister, along with up to a maximum of eight port WWPNs from each external storage system.

**Ports:** IBM Storwize V7000 supports a maximum of 16 ports or WWPNs from a given external storage system that will be virtualized.

Figure 9-3 is an example of how to cable the devices to the SAN. Refer to this example as we describe the zoning.

![SAN cabling and zoning example diagram](image)

Create an IBM Storwize V7000 and external storage zone for each storage system to be virtualized as in the following example:

- Zone DS5100 controller ports A1 and B1 with all node ports 1 and 3 in the RED fabric
- Zone DS5100 controller ports A2 and B2 with all node ports 2 and 4 in the BLUE fabric
9.1.3 External storage configuration planning

External storage systems provide redundancy through various Redundant Array of Independent Disks (RAID) levels, which prevents a single physical disk failure from causing a managed disk (MDisk), storage pool, or associated host volume, from going offline. To minimize the risk of data loss, only virtualize storage systems where logical unit numbers (LUNs) are configured using a RAID level other than RAID 0 (for example RAID 1, RAID 10, RAID 0+1, RAID 5, or RAID 6).

Verify that the storage controllers to be virtualized by IBM Storwize V7000 meet the requirements, and that you found the configuration restrictions by navigating at the IBM Support Portal website:

http://www.ibm.com/support

Ensure that the firmware or microcode levels of the storage controllers to be virtualized are supported by IBM Storwize V7000.

IBM Storwize V7000 must have exclusive access to the LUNs from the external storage system mapped to it. LUN access cannot be shared between IBM Storwize V7000s or between an IBM Storwize V7000 and other storage virtualization platforms or between an IBM Storwize V7000 and hosts. However, different LUNs could be mapped from one external storage system to an IBM Storwize V7000 and other hosts in the SAN through different storage ports.

Make sure to configure the storage subsystem LUN masking settings to map all LUNs to all the WWPNs in the IBM Storwize V7000 storage system.

Be sure to go to the IBM Storwize V7000 page and review the Configuring and servicing external storage system topic before you prepare the external storage systems for discovery by the IBM Storwize V7000 system. This website is at the following location:


You can also go to this link:


9.1.4 Guidelines for virtualizing external storage

When virtualizing external storage with the IBM Storwize V7000, follow these guidelines:

- 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 V7000.
- Except for Easy Tier, do not mix MDisks that vary in performance or reliability in the same storage pool. Always put similarly sized MDisks into one storage pool. For more information about Easy Tier, see Chapter 10, “Advanced features for storage efficiency” on page 335.
- Do not leave volumes in image mode. Only use image mode to import or export existing data into or out of the IBM Storwize V7000. Migrate such data from image mode MDisks to other storage pools to benefit from storage virtualization.
Using the copy services in Storwize V7000 gives you a unified method to manage data integrity across heterogeneous storage systems.

The Easy Tier function is included with the IBM Storwize V7000 system, and the external storage system could benefit from this powerful storage tiering function to remove hot spots and improve overall performance.

9.2 Working with external storage

In this section, we describe how to manage external storage using an IBM Storwize V7000.

The basic concepts of managing an external storage system are the same as internal storage. IBM Storwize V7000 discovers LUNs from the external storage system as one or more MDisks. These MDisks are ultimately added to a storage pool in which volumes are created and mapped to hosts as needed.

9.2.1 Adding external storage

To add new external storage systems to the IBM Storwize V7000 virtualized environment, complete the following steps:

1. Zone a minimum of two and a maximum of 16 Fibre Channel ports from the external storage system with all eight Fibre Channel ports on the IBM Storwize V7000 system. As a preferred practice, have two fabrics for redundancy in the SAN. Then, in each fabric zone two ports from each node canister in the IBM Storwize V7000 system with half the ports from the external system. As the IBM Storwize V7000 is virtualizing your storage, hosts should be zoned with the V7000 controller port WWPNs.

2. Using the storage partitioning or LUN masking feature of the external storage system, create a group that includes all eight IBM Storwize V7000 WWPNs.

3. Create equal size arrays on the external system using any RAID level but zero.

4. Create a single LUN per RAID array.

5. Map the LUNs to all eight Fibre Channel ports on the IBM Storwize V7000 system by assigning them to the group created in step 2.

6. Verify that IBM Storwize V7000 discovered the LUNs as unmanaged MDisks. If they do not show up automatically, click the Detect MDisk option from the MDisk window of the GUI, as described in Chapter 7, “Storage pools” on page 203. You should see the MDisks mapped to the IBM Storwize V7000 under the respective Storage system.

7. Select the storage tier for the MDisks, as described in Chapter 7, “Storage pools” on page 203.

8. Create a storage pool, as described in Chapter 7, “Storage pools” on page 203.

9. Add the MDisks to the pool, as described in Chapter 7, “Storage pools” on page 203.

10. Create volumes and map them to hosts as needed, as described in Chapter 4, “Host configuration” on page 111 and Chapter 5, “Basic volume configuration” on page 145.

If the external storage systems are not the new systems, that is, existing data on the LUNs must be kept after virtualization, complete the steps in Chapter 6, “Migration wizard” on page 189 to prepare the environment. Then, you can migrate the existing data (with or without using wizard) to IBM Storwize V7000 internal storage or another external storage system.
To manually import MDisks and migrate the data to other storage pools, see Chapter 7, “Storage pools” on page 203. Whether you migrate the data with the wizard or not, you can select your destination storage pools to be internal storage pools or external storage pools.

9.2.2 Managing external storage

The IBM Storwize V7000 provides an individual External Storage panel for managing external storage systems.

You can access the External Storage pane by opening the IBM Storwize V7000 Home pane and clicking the External Storage System function icon. Extended help information for external storage is displayed. Click **Pools** and the External Storage pane opens.

Figure 9-4 shows how to access the External Storage pane from the IBM Storwize V7000 Home pane.

![Figure 9-4 Access the External Storage pane from the IBM Storwize V7000 panel](image-url)
The External Storage window (Figure 9-5) shows an overview of all your external storage systems. To see a list of managed disks (MDisks) presented by each storage system to your IBM Storwize V7000, expand each external storage system by clicking the plus sign (+).

You can change the name of an external storage system by selecting the storage system and clicking **Actions** or right-clicking the storage system. The status of the storage system, its type, and WWPN are also listed in the table.

If you select the storage system and click **Actions**, you can do these tasks (Figure 9-6):

- **Show Dependent Volumes**
- **Rename the storage system**
Click **Show Dependent Volumes** to view all the volumes in this external storage system (Figure 9-7).

![Volumes Dependent on Storage System DS3400](image)

**Figure 9-7**  Volumes dependent on external storage

One of the features of the IBM Storwize V7000 storage system is that it can be used as a data migration tool. In the IBM Storwize V7000 virtualization environment, you can migrate your application data nondisruptively from one internal or external storage system to another storage system, which makes storage management much simpler and reduces risk.

Volume copy is another key feature that you can benefit from by using IBM Storwize V7000 virtualization. Two copies could be applied to your data to enhance the availability for a critical application. A volume copy could be also used to generate test data or data migration.

Returning to the External Storage window, you can perform various MDisk tasks. In the list, you can find the name of an MDisk, its capacity and mode, the storage pool, and the storage system it belongs to in the list. The actions on MDisks can also be made through the Actions menu (Figure 9-8 on page 332), including Detect MDisks, Assign to Pool, Import, and so on. This menu is the same as the one in the MDisks window, which is described in Chapter 7, “Storage pools” on page 203.
9.2.3 Removing external storage

If you want to remove the external storage systems from the IBM Storwize V7000 virtualized environment, you have several options:

- If you want to remove the external storage systems and discard the data on it, complete the following steps:
  a. Stop any host I/O on the volumes.
  b. Remove the volumes from the host file system, logical volume, or volume group, and remove the volumes from the host device inventory.
  c. Remove the host mappings of volumes and the volumes themselves on IBM Storwize V7000, as described in Chapter 8, “Advanced host and volume administration” on page 247.
  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, as described Chapter 7, “Storage pools” on page 203.
  e. Unzone and disconnect the external storage systems from the IBM Storwize V7000.
  f. Click Detect MDisks to make IBM Storwize V7000 discover the removal of the external storage systems, as described in Chapter 7, “Storage pools” on page 203.

- If you want to remove the external storage systems and keep the volumes and their data on the IBM Storwize V7000, complete the following steps:
  a. Migrate volumes and their data to the other storage pools that are on IBM Storwize V7000 internal storage or other external storage systems, as described in Chapter 5, “Basic volume configuration” on page 145.
  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, as described in Chapter 7, “Storage pools” on page 203.
  c. Unzone and disconnect the external storage systems from the IBM Storwize V7000.
  d. Click Detect MDisks to make IBM Storwize V7000 discover the removal of the external storage systems, as described in Chapter 7, “Storage pools” on page 203.
If you want to remove the external storage systems from IBM Storwize V7000 control and keep the volumes and their data on external storage systems, complete the following steps:

a. Migrate volumes and their data to the other storage pools that are on IBM Storwize V7000 internal storage or other external storage systems, as described in Chapter 5, “Basic volume configuration” on page 145.

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, as described in Chapter 7, “Storage pools” on page 203.

c. Export volumes to image mode with the MDisks on the external storage systems. The restrictions and prerequisites for migration can be found in Chapter 6, “Migration wizard” on page 189. You also need to record pre-migration information, for example, the original Small Computer System Interface (SCSI) IDs the volumes used to be mapped to hosts. Some operating systems do not support changing the SCSI ID during the migration. More information about migration is at the IBM Storwize V7000 welcome page:
   

d. Unzone and disconnect the external storage systems from the IBM Storwize V7000.

e. Click **Detect MDisks** to make IBM Storwize V7000 discover the removal of the external storage systems, as described Chapter 7, “Storage pools” on page 203.
Advanced features for storage efficiency

In this chapter, we introduce the basic concepts of dynamic data relocation and storage optimization features. The IBM Storwize V7000 offers the following software functions for storage efficiency:

- Easy Tier
- Thin provisioning
- Real-time compression

We provide only a basic technical overview and benefits of each feature. For more information about planning and configuration, see the following IBM Redbooks publications:

- Easy Tier:
  - Implementing IBM Easy Tier with IBM Real-time Compression, TIPS1072
  - IBM System Storage SAN Volume Controller and Storwize V7000 Best Practices and Performance Guidelines, SG24-7521-03
  - IBM DS8000 Easy Tier, REDP-4667 (this concept is similar to SAN Volume Controller Easy Tier)

- Thin Provisioning:
  - Thin Provisioning in an IBM SAN or IP SAN Enterprise Environment, REDP-4265
  - DS8000 Thin Provisioning, REDP-4554 (similar concept to SAN Volume Controller Thin provisioning)

- Real-Time Compression:
  - Real-time Compression in SAN Volume Controller and 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:

- Introduction
- Easy Tier
- Thin provisioning
- Real-time Compression Software

10.1 Introduction

In modern and complex application environments, the increasing and often unpredictable demands for storage capacity and performance lead to issues of planning and optimization of storage resources.

Consider the following typical storage management issues:

- Usually when a storage system is implemented, only a portion of the configurable physical capacity is deployed. When the storage system runs out of the installed capacity and more capacity is requested, a hardware upgrade is implemented to add physical resources to the storage system. This new physical capacity can hardly be configured to keep an even spread of the overall storage resources. Typically, the new capacity is allocated to fulfill only new storage requests. The existing storage allocations do not benefit from the new physical resources. Similarly, the new storage requests do not benefit from the existing resources; only new resources are used.

- In a complex production environment, it is not always possible to optimize storage allocation for performance. The unpredictable rate of storage growth and the fluctuations in throughput requirements, which are I/O per second (IOPS), often lead to inadequate performance. Furthermore, the tendency to use even larger volumes to simplify storage management works against the granularity of storage allocation, and a cost-efficient storage tiering solution becomes difficult to achieve. With the introduction of high performing technologies, such as solid-state drives (SSD) or all flash arrays, this challenge becomes even more important.

- The move to larger and larger physical disk drive capacities means that previous access densities that were achieved with low-capacity drives can no longer be sustained.

- Any business has applications that are more critical than others, and there is a need for specific application optimization. Therefore, there is a need for the ability to relocate specific application data to faster storage media.

- Although more servers are purchased with local SSDs attached for better application response time, the data distribution across these direct-attached SSDs and external storage arrays must be carefully addressed. An integrated and automated approach is crucial to achieve performance improvement without compromise to data consistency, especially in a disaster recovery situation.

All of these issues deal with data placement and relocation capabilities or data volume reduction. Most of these challenges can be managed by having spare resources available and by moving data, and by the use of data mobility tools or operating systems features (such as host level mirroring) to optimize storage configurations. However, all of these corrective actions are expensive in terms of hardware resources, labor, and service availability. Relocating data among the physical storage resources that dynamically or effectively reduces the amount of data, that is, transparently to the attached host systems, is becoming increasingly important.
10.2 Easy Tier

In today's storage market, SSDs and flash arrays are emerging as an attractive alternative to hard disk drives (HDDs). Because of their low response times, high throughput, and IOPS-energy-efficient characteristics, SSDs and flash arrays have the potential to allow your storage infrastructure to achieve significant savings in operational costs. However, the current acquisition cost per GB for SSDs or flash array is higher than for HDDs. SSD and flash arrays performance depends greatly on workload characteristics; therefore, they should be used with HDDs for optimal performance.

It is critical to choose the right mix of drives and the right data placement to achieve optimal performance at low cost. Maximum value can be derived by placing “hot” data with high I/O density and low response time requirements on SSDs or flash arrays, while targeting HDDs for “cooler” data that is accessed more sequentially and at lower rates.

Easy Tier automates the placement of data among different storage tiers and it can be enabled for internal and external storage. This SAN Volume Controller feature boosts your storage infrastructure performance to achieve optimal performance through a software, server, and storage solution. Additionally, the new, no charge feature called Storage Pool Balancing, introduced in 7.3 SAN Volume Controller firmware version, automatically moves extents within the same storage tier, from overloaded to less loaded managed disks (MDisks). This assures that your data is optimally placed among all disks within storage pools.

10.2.1 Easy Tier concepts

Storwize V7000 implements Easy Tier enterprise storage functions, originally available on IBM DS8000 and IBM XIV enterprise class storage systems. It enables automated subvolume data placement throughout different or within the same storage tiers to intelligently align the system with current workload requirements and to optimize the usage of SSDs or flash arrays. This functionality includes the ability to automatically and non-disruptively relocate data (at the extent level) from one tier to another tier or even within the same tier, in either direction to achieve the best available storage performance for your workload in your environment. Easy Tier reduces the I/O latency for hot spots, but it does not replace storage cache. Both Easy Tier and storage cache solve a similar access latency workload problem, but these two methods weigh differently in the algorithmic construction that is based on “locality of reference”, recency, and frequency. Because Easy Tier monitors I/O performance from the device end (after cache), it can pick up the performance issues that cache cannot solve and complement the overall storage system performance. Figure 10-1 on page 338 shows placement of Easy Tier engine within Storwize V7000 software stack.
In general, the storage environment's I/O is monitored on volume level and the entire volume is always placed inside one appropriate storage tier. Determining the amount of I/O, moving part of the underlying volume to an appropriate storage tier and reacting to workload changes is too complex for manual operation. This is where the Easy Tier feature can be used.

Easy Tier is a performance optimization function, as it automatically migrates (or moves) extents that belong to a volume between different storage tiers (see Figure 10-2 on page 339) or the same storage tier (see Figure 10-4 on page 341). As this migration works at the extent level, it is often referred to as sub-LUN migration. Movement of the extents is done online and is unnoticed from host point of view. As a result of extent movement, the volume no longer has all its data in one tier but rather in two or three tiers. Figure 10-2 on page 339 shows the basic Easy Tier principle of operation.

![Figure 10-1 Easy Tier in Storwize V7000 software stack](image)
You can enable Easy Tier on a volume basis. It monitors the I/O activity and latency of the extents on all Easy Tier enabled volumes over a 24-hour period. Based on the performance log, it creates an extent migration plan and dynamically moves high activity or hot extents to a higher disk tier within the same storage pool. It also moves extents whose activity dropped off, or cooled, from higher disk tier MDisks back to a lower tier MDisk. When running in storage pool rebalance mode, it moves extents from busy MDisks to less busy MDisks of the same type.

10.2.2 SSD arrays and flash MDisks

The SSDs or flash arrays are treated no differently by the Storwize V7000 than HDDs regarding RAID arrays or MDisks.

The individual SSDs in the storage enclosures are combined into an array, usually in RAID 10 or RAID 5 format. It is unlikely that RAID6 SSD arrays are used because of the double parity overhead, with two logical SSDs used for parity only. As with usual HDDs, RAID is an MDisk of an array type and after creation is then managed the same way the HDD MDisks are.
As is the case for HDDs, the SSD RAID array format helps to protect against individual SSD failures. Depending on your requirements, you can achieve more high availability protection above the RAID level by using volume mirroring.

The internal storage configuration of flash arrays can differ depending on an array vendor. But regardless of methods used to configure flash-based storage, the flash system maps a volume to host, in this case, Storwize V7000. From Storwize perspective, the volume presented from flash memory is seen also as normal managed disk.

As mentioned before, after creation of an SSD RAID array it appears as a usual MDisk but with a type of ssd. This differs from MDisks presented from external storage systems or RAID arrays made of HDDs. Because Storwize does not know from what kind of physical disks external MDisks are formed, the default MDisk type that Storwize adds to each external MDisk is enterprise. It is up to the user or administrator to change the type of MDisks to ssd, enterprise, or nearline.

**Note:** It is possible to change the type of MDisks made of internal Storwize drives even if the type of MDisk will not fit the type of physical drives MDisk is made of. Storwize knows the type of drives it has in disk enclosures and selects the MDisk type accordingly to the drive type but this selection can be overridden by a user or administrator.

To change a type of external or internal MDisk in CLI, use the `chmdisk` command as in Example 10-1.

```
Example 10-1 changing MDisk tier

IBM_Storwize:ITSO_V7000Gen2_1:superuser>lsmdisk -delim " "
 id name status mode mdisk_grp_id mdisk_grp_name capacity ctrl_LUN_# controller_name UID tier encrypt
0 testmdisk1 online array 0 INFRA 2.2TB enterprise no

IBM_Storwize:ITSO_V7000Gen2_1:superuser>chmdisk -tier ssd testmdisk1

IBM_Storwize:ITSO_V7000Gen2_1:superuser>lsmdisk -delim " 
 id name status mode mdisk_grp_id mdisk_grp_name capacity ctrl_LUN_# controller_name UID tier encrypt
0 testmdisk1 online array 0 INFRA 2.2TB ssd no
```

**Note:** The type of the MDisk can also be changed in the GUI. To do that, from the animated menu on the left side of the window, hover over Pools and select External Storage or MDisks by Pools. Click the little + sign next to the storage controller name or storage pool name, depending if you chose external storage or MDisks by pools to expand MDisks. Next, right-click MDisk and choose Select Tier. Then, you have three options to choose from to select the proper tier for your MDisks.

If you do not see the Tier column in the External Storage or MDisks by Pools view, right-click the blue title row and select the Tier check box as presented on Figure 10-3 on page 341.
10.2.3 Disk tiers

The internal or external MDisks (LUNs) are likely to have different performance attributes because of the type of disk or RAID array on which they reside. The MDisks can be on 15 K RPM Fibre Channel or SAS disk, Near-line SAS or SATA, or even SSDs or flash storage systems.

As mentioned in 10.2.2, “SSD arrays and flash MDisks” on page 339, Storwize V7000 does not automatically detect the type of external MDisks. Instead, all external MDisks initially are put into the enterprise tier by default. Then, the administrator must manually change the tier of MDisks and add them to storage pools. Depending on what type of disks are gathered to form a storage pool, we distinguish two types of storage pools: single-tier and multi-tier.

Single-tier storage pools

Figure 10-4 shows a scenario in which a single storage pool is populated with MDisks that are presented by an external storage controller. In this solution, the striped or volumes can be measured by Easy Tier and can benefit from Storage Pool Balancing mode, which will move extents between MDisks of the same type.
MDisks that are used in a single-tier storage pool should have the same hardware characteristics; for example, the same RAID type, RAID array size, disk type, disk revolutions per minute (RPM), and controller performance characteristics.

**Multitier storage pools**

A multitier storage pool has a mix of MDisks with more than one type of disk tier attribute; for example, a storage pool that contains a mix of enterprise and ssd MDisks, or enterprise and nl-sas MDisk. Figure 10-5 shows a scenario in which a storage pool is populated with three different MDisk types: one belonging to an SSD array, one belonging to SAS HDD array, and one belonging to an NL-SAS HDD array. Although this example shows RAID 5 arrays, other RAID types can be used as well.

Adding SSDs to the pool means that more space also is now available for new volumes or volume expansion.

**Attention:** Image mode and sequential volumes are not candidates for Easy Tier automatic data placement because for those types of volumes all extents must reside on one, specific MDisk and cannot be moved.

The Easy Tier setting can be changed on a storage pool and volume basis. Depending on the Easy Tier setting and number of tiers in storage pool Easy Tier services may function in a different way. Table 10-1 shows possible combinations of Easy Tier setting.

**Table 10-1 EasyTier settings**

<table>
<thead>
<tr>
<th>Storage pool Easy Tier setting</th>
<th>Number of tiers in the storage pool</th>
<th>Volume copy Easy Tier setting</th>
<th>Volume copy Easy Tier status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Off</td>
<td>One</td>
<td>off</td>
<td>inactive (see note 2)</td>
</tr>
<tr>
<td>Off</td>
<td>One</td>
<td>on</td>
<td>inactive (see note 2)</td>
</tr>
<tr>
<td>Off</td>
<td>Two or three</td>
<td>off</td>
<td>inactive (see note 2)</td>
</tr>
</tbody>
</table>
Figure 10-6 on page 344 shows the naming convention and all supported combinations of storage tiering used by Easy Tier.

<table>
<thead>
<tr>
<th>Storage pool Easy Tier setting</th>
<th>Number of tiers in the storage pool</th>
<th>Volume copy Easy Tier setting</th>
<th>Volume copy Easy Tier status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Off</td>
<td>Two or three</td>
<td>on</td>
<td>inactive (see note 2)</td>
</tr>
<tr>
<td>Measure</td>
<td>One</td>
<td>off</td>
<td>measured (see note 3)</td>
</tr>
<tr>
<td>Measure</td>
<td>One</td>
<td>on</td>
<td>measured (see note 3)</td>
</tr>
<tr>
<td>Measure</td>
<td>Two or three</td>
<td>off</td>
<td>measured (see note 3)</td>
</tr>
<tr>
<td>Measure</td>
<td>Two or three</td>
<td>on</td>
<td>measured (see note 3)</td>
</tr>
<tr>
<td>Auto</td>
<td>One</td>
<td>off</td>
<td>measured (see note 3)</td>
</tr>
<tr>
<td>Auto</td>
<td>One</td>
<td>on</td>
<td>balanced (see note 4)</td>
</tr>
<tr>
<td>Auto</td>
<td>Two or three</td>
<td>off</td>
<td>measured (see note 3)</td>
</tr>
<tr>
<td>Auto</td>
<td>Two or three</td>
<td>on</td>
<td>active (see note 5)</td>
</tr>
<tr>
<td>On</td>
<td>One</td>
<td>off</td>
<td>measured (see note 3)</td>
</tr>
<tr>
<td>On</td>
<td>One</td>
<td>on</td>
<td>balanced (see note 4)</td>
</tr>
<tr>
<td>On</td>
<td>Two or three</td>
<td>off</td>
<td>measured (see note 3)</td>
</tr>
<tr>
<td>On</td>
<td>Two or three</td>
<td>on</td>
<td>active (see note 5)</td>
</tr>
</tbody>
</table>

**Note:**

1. If the volume copy is in image or sequential mode or is being migrated, the volume copy Easy Tier status is measured instead of active.

2. When the volume copy status is inactive, no Easy Tier functions are enabled for that volume copy.

3. When the volume copy status is measured, the Easy Tier function collects usage statistics for the volume but automatic data placement is not active.

4. When the volume copy status is balanced, the Easy Tier function enables performance-based pool balancing for that volume copy.

5. When the volume copy status is active, the Easy Tier function operates in automatic data placement mode for that volume.

6. The default Easy Tier setting for a storage pool is auto, and the default Easy Tier setting for a volume copy is on. This means that Easy Tier functions except pool performance balancing are disabled for storage pools with a single tier, and that automatic data placement mode is enabled for all striped volume copies in a storage pool with two or more tiers.
10.2.4 Easy Tier process

The Easy Tier function includes the following four main processes:

- **I/O Monitoring**
  - This process operates continuously and 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.
  - Easy Tier makes allowances for large block I/Os; therefore, it considers only I/Os of up to 64 KB as migration candidates.
  - This process is efficient and adds negligible processing overhead to the SAN Volume Controller nodes.

- **Data Placement Advisor**
  - The Data Placement Advisor uses workload statistics to make a cost benefit decision as to which extents are to be candidates for migration to a higher performance tier.
  - This process also identifies extents that must be migrated back to a lower tier.

- **Data Migration Planner**
  - By using the extents that were previously identified, the Data Migration Planner builds the extent migration plans for the storage pool. DMP builds two plans:
    - ADR plan to migrate extents across adjacent tiers
    - RB plan to migrate extents within the same tier

- **Data Migrator**
  - This process involves the actual movement or migration of the volume’s extents up to, or down from, the higher disk tier. The extent migration rate is capped so that a maximum of up to 30 MBps is migrated, which equates to approximately 3 TB per day that is migrated between disk tiers.

When enabled, Easy Tier performs the following actions between three tiers presented in Figure 10-6:

- **Promote**
  - Moves the relevant hot extents to higher performing tier

- **Swap**
  - Exchange cold extent in upper tier with hot extent in lower tier
- **Warm Demote**
  - Prevents performance overload of a tier by demoting a warm extent to the lower tier
  - Triggered when bandwidth or IOPS exceeds predefined threshold

- **Demote or Cold Demote**
  - Coldest data is moved to lower HDD tier. Only supported between HDD tiers

- **Expanded Cold Demote**
  - Demotes appropriate sequential workloads to the lowest tier to better utilize Nearline disk bandwidth

- **Storage Pool Balancing**
  - Redistribute extents within a tier to balance utilization across MDisks for maximum performance
  - Moves hot extents from high utilized MDisks to low utilized MDisks
  - Exchanges extents between high utilized MDisks and low utilized MDisks

- **It attempts to migrate the most active volume extents up to SSD first**

- **A previous migration plan and any queued extents that are not yet relocated are abandoned**

**Note:** Extent migration occurs only between adjacent tiers. In a three-tiered storage pool, Easy Tier will not move extents from SSD directly to NL-SAS and vice versa without moving them first to SAS drives.

Easy Tier extent migration types are presented on Figure 10-7 on page 346.
10.2.5 Easy Tier operating modes

Easy Tier includes the following main operating modes:

- Off
- Evaluation or measurement only
- Automatic Data Placement or extent migration
- Storage Pool Balancing

Easy Tier off mode

With Easy Tier turned off, no statistics are recorded and no cross tier extent migration occurs. In this mode, only storage pool balancing is active, which means extents are migrated within the same storage pool.

Evaluation or measurement only mode

Easy Tier Evaluation or measurement-only mode collects usage statistics for each extent in a single tier storage pool where the Easy Tier value is set to On for both the volume and the pool. This collection is typically done for a single-tier pool that contains only HDDs so that the benefits of adding SSDs to the pool can be evaluated before any major hardware acquisition.

A dpa_heat.nodeid.yymmd.hhmss.data statistics summary file is created in the /dumps directory of the Storwize V7000 node canisters. This file can be offloaded from the Storwize...
node canisters with PSCP -load or by using the GUI, as described in IBM System Storage SAN Volume Controller and Storwize V7000 Best Practices and Performance Guidelines, SG24-7521-03. A web browser is used to view the report that is created by the tool.

**Automatic Data Placement or extent migration mode**

In Automatic Data Placement or extent migration operating mode, the storage pool parameter -easytier on or auto must be set, and the volumes in the pool must have -easytier on. The storage pool must also contain MDisks with different disk tiers, which makes it a multitiered storage pool.

Dynamic data movement is not apparent to the host server and application users of the data, other than providing improved performance. Extents are automatically migrated as explained in “Implementation rules” on page 348.

The statistic summary file is also created in this mode. This file can be offloaded for input to the advisor tool. The tool produces a report on the extents that are moved to higher tier and a prediction of performance improvement that can be gained if more higher tier disks are available.

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

**Storage Pool Balancing**

Storage Pool Balancing is a new feature within the 7.3 code, which while associated with Easy Tier, operates independently of Easy Tier and does not require an Easy Tier license. This feature assesses the extents that are written in a pool and balances them automatically across all MDisks within the pool. This process works with Easy Tier when multiple classes of disks exist in a single pool.

The process will automatically balance existing data when new MDisks are added into an existing pool even if the pool only contains a single type of drive. This does not mean it will migrate extents from existing MDisks to achieve even extent distribution among all, old and new, MDisks in the storage pool. Easy Tier RB within tier migration plan is based on performance and not capacity of underlying MDisks.

**Note:** Storage Pool Balancing can be used to balance extents when mixing different size disks of the same performance tier. For example, when adding larger capacity drives to pool with smaller capacity drives of the same class, Storage Pool Balancing redistributes the extents to take advantage of the additional performance of the new MDisks.

**10.2.6 Implementation considerations**

Easy Tier is a licensed feature for storage pool balancing, which is a no-charge feature and is enabled by default. Easy Tier comes as part of the Storwize V7000 code. For Easy Tier to migrate extents between different tier disks, you must have disk storage available that has different tiers; for example, a mix of SSD and HDD. Easy Tier will use Storage Pool Balancing if you have only single tier pool.
**Implementation rules**
Keep in mind the following implementation and operation rules when you use the IBM System Storage Easy Tier function on the Storwize V7000:

- Easy Tier automatic data placement is not supported on image mode or sequential volumes. I/O monitoring for such volumes is supported, but you cannot migrate extents on such volumes unless you convert image or sequential volume copies to striped volumes.
- Automatic data placement and extent I/O activity monitors are supported on each copy of a mirrored volume. Easy Tier works with each copy independently of the other copy.

**Volume mirroring consideration:** 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. Thus, the number of extents that Easy Tier migrates to the SSD tier might be different for each copy.

- If possible, the Storwize V7000 creates volumes or volume expansions by using extents from MDisks from the HDD tier. However, it uses extents from MDisks from the SSD tier, if necessary.

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 if it is between pools that both have Easy Tier automatic data placement enabled. Automatic data placement for the volume is re-enabled when the migration is complete.

**Limitations**
When you use Easy Tier on the Storwize V7000, keep in mind the following limitations:

- Removing an MDisk by using the `-force` parameter
  When an MDisk is deleted from a storage pool with the `-force` parameter, 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.

- Migrating extents
  When Easy Tier automatic data placement is enabled for a volume; you cannot use the `svctask migrateexts` CLI command on that volume.

- Migrating a volume to another storage pool
  When the Storwize V7000 migrates a volume to a new storage pool, Easy Tier automatic data placement between the two tiers is temporarily suspended. After the volume is migrated to its new storage pool, Easy Tier automatic data placement between the generic SSD tier and the generic HDD tier resumes for the moved volume, if appropriate.

  When the Storwize V7000 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. In several cases, such as where a target tier is unavailable, the other tier is used. For example, the generic SSD tier might be unavailable in the new storage pool.

- Migrating a volume to image mode.
  Easy Tier automatic data placement does not support image mode. When a volume with Easy Tier automatic data placement mode active is migrated to image mode, Easy Tier automatic data placement mode is no longer active on that volume.

- Image mode and sequential volumes cannot be candidates for automatic data placement; however, Easy Tier supports evaluation mode for image mode volumes.
10.2.7 Modifying the Easy Tier setting

The Easy Tier setting for storage pools and volumes can be changed only via command line. Use the `chvdisk` command to turn off or turn on Easy Tier on selected volumes and `chmdiskgrp` to change status of Easy Tier on selected storage pools as shown in Example 10-2.

Example 10-2 Changing Easy Tier setting

```
IBM_Storwize:ITSO_V7000Gen2_2:superuser>lsvdisk Generic_0
id 0
name Generic_0
IO_group_id 0
IO_group_name io_grp0
status online
mdisk_grp_id 0
mdisk_grp_name V7000_Pool1
capacity 10.00GB
type striped
formatted no
mdisk_id
mdisk_name
FC_id
FC_name
RC_id
RC_name
vdisk_UID 60050764008200078000000000000000
throttling 0
preferred_node_id 2
fast_write_state empty
cache readwrite
uid
fc_map_count 0
sync_rate 99
copy_count 1
se_copy_count 0
filesystem
mirror_write_priority latency
RC_change no
compressed_copy_count 0
access_IO_group_count 1
last_access_time
parent_mdisk_grp_id 0
parent_mdisk_grp_name V7000_Pool1

copy_id 0
status online
sync yes
primary yes
mdisk_grp_id 0
mdisk_grp_name V7000_Pool1
type striped
mdisk_id
mdisk_name
fast_write_state empty
used_capacity 10.00GB
```
real_capacity 10.00GB
free_capacity 0.00MB
overallocation 100
autoexpand
warning
grainsize
se_copy no
easy_tier off
easy_tier_status measured
tier ssd
tier_capacity 0.00MB
tier enterprise
tier_capacity 10.00GB
tier nearline
tier_capacity 0.00MB
compressed_copy no
uncompressed_used_capacity 10.00GB
parent_mdisk_grp_id 0
parent_mdisk_grp_name V7000_Pool1
IBM_Storwize:ITSO_V7000Gen2_2:superuser>

IBM_Storwize:ITSO_V7000Gen2_2:superuser>chvdisk -easytier on Generic_0

IBM_Storwize:ITSO_V7000Gen2_2:superuser>lsvdisk Generic_0
id 0
name Generic_0
IO_group_id 0
IO_group_name io_grp0
status online
mdisk_grp_id 0
mdisk_grp_name V7000_Pool1
capacity 10.00GB
type striped
formatted no
mdisk_id
disk_name
FC_id
FC_name
RC_id
RC_name
vdisk_UID 60050764008200078000000000000000
throttling 0
preferred_node_id 2
fast_write_state empty
cache readwrite
udid
fc_map_count 0
sync_rate 99
copy_count 1
se_copy_count 0
filesystem
mirror_write_priority latency
RC_change no
compressed_copy_count 0
access_IO_group_count 1
last_access_time
parent_mdisk_grp_id 0
parent_mdisk_grp_name V7000_Pool1

copy_id 0
status online
sync yes
primary yes
mdisk_grp_id 0
mdisk_grp_name V7000_Pool1
type striped
mdisk_id
mdisk_name
fast_write_state empty
used_capacity 10.00GB
real_capacity 10.00GB
free_capacity 0.00MB
overallocation 100
autoexpand
warning
grainsize
se_copy no
easy_tier on
easy_tier_status balanced
tier ssd
tier_capacity 0.00MB
tier enterprise
tier_capacity 10.00GB
tier nearline
tier_capacity 0.00MB
compressed_copy no
uncompressed_used_capacity 10.00GB
parent_mdisk_grp_id 0
parent_mdisk_grp_name V7000_Pool1

IBM_Storwize:ITSO_V7000Gen2_2:superuser>lsmdiskgrp V7000_Pool1
id 0
name V7000_Pool1
status online
mdisk_count 2
vdisk_count 14
capacity 4.90TB
extent_size 1024
free_capacity 4.55TB
virtual_capacity 356.00GB
used_capacity 356.00GB
real_capacity 356.00GB
overallocation 7
warning 80
easy_tier auto
easy_tier_status balanced
tier ssd
tier_mdisk_count 0
IBM Storwize:ITSO_V7000Gen2_2:superuser>chmdiskgrp -easytier off V7000_Pool1

IBM Storwize:ITSO_V7000Gen2_2:superuser>lsmdiskgrp V7000_Pool1
id 0
name V7000_Pool1
status online
mdisk_count 2
vdisk_count 14
capacity 4.90TB
extent_size 1024
free_capacity 4.55TB
virtual_capacity 356.00GB
used_capacity 356.00GB
real_capacity 356.00GB
overallocation 7
warning 80
easy_tier off
easy_tier_status inactive
tier ssd
tier_mdisk_count 0
tier_capacity 0.00MB
tier_free_capacity 0.00MB
tier enterprise
tier_mdisk_count 2
tier_capacity 4.90TB
tier_free_capacity 4.55TB
tier nearline
tier_mdisk_count 0
tier_capacity 0.00MB
tier_free_capacity 0.00MB
compression_active no

IBM Storwize:ITSO_V7000Gen2_2:superuser>chmdiskgrp -easytier off V7000_Pool1

IBM Storwize:ITSO_V7000Gen2_2:superuser>lsmdiskgrp V7000_Pool1
id 0
name V7000_Pool1
status online
mdisk_count 2
vdisk_count 14
capacity 4.90TB
extent_size 1024
free_capacity 4.55TB
virtual_capacity 356.00GB
used_capacity 356.00GB
real_capacity 356.00GB
overallocation 7
warning 80
easy_tier off
easy_tier_status inactive
tier ssd
tier_mdisk_count 0
tier_capacity 0.00MB
tier_free_capacity 0.00MB
tier enterprise
tier_mdisk_count 2
tier_capacity 4.90TB
tier_free_capacity 4.55TB
tier nearline
tier_mdisk_count 0
tier_capacity 0.00MB
tier_free_capacity 0.00MB
compression_active no
10.2.8 Monitoring tools

The IBM Storage Tier Advisor Tool (STAT) is a Windows console application that analyzes heat data files produced by Easy Tier and produces a graphical display of the amount of “hot” data per volume and predictions of how additional flash drive (SSD) capacity, Enterprise Drive, and Nearline Drive could benefit performance for the system and by storage pool.

Heat data files are produced approximately once a day (that is, every 24 hours) when Easy Tier is active on one or more storage pools and summarizes the activity per volume since the prior heat data file was produced. On SAN Volume Controller and Storwize serial products, the heat data file is in the /dumps directory on the configuration node and is named “dpa_heat.node_name.time_stamp.data”.

Any existing heat data file is erased when it has existed for longer than seven days. The file must be offloaded by the user and Storage Tier Advisor Tool invoked from a Windows command prompt console with the file specified as a parameter. The user can also specify the output directory. The Storage Tier Advisor Tool creates a set of HTML files and the user can then open the resulting “index.html” in a browser to view the results.

Updates to the STAT tool for SAN Volume Controller and Storwize V7000 7.3 have added additional capability for reporting. As a result, when the STAT tool is run on a heat map file an additional three CSV files are created and placed in the Data_files directory.

The IBM STAT tool can be downloaded from the IBM Support website:

http://www.ibm.com/support/docview.wss?uid=ssg1S4000935

Figure 10-8 on page 354 shows the CSV files highlighted in the Data_files directory after running the stat tool over an SVC/Storwize V7000 heatmap.
In addition to the STAT tool, SAN Volume Controller 7.3 code now has an additional utility, which is a Microsoft SQL file for creating additional graphical reports of the workload that Easy Tier is performing. The IBM STAT Charting Utility takes the output of the three CSV files and turns them into graphs for simple reporting.

The three new graphs display:

- **Workload Categorization**

  New workload visuals help you compare activity across tiers within and across pools to help determine optimal drive mix for current workloads. The output is illustrated in Figure 10-9.
Daily Movement Reporting

A new Easy Tier summary report every 24-hours illustrating data migration activity (5-min. intervals) can help visualize migration types and patterns for current workloads. The output is illustrated in Figure 10-10.

![Figure 10-10  STAT Charting Utility Daily Summary report](image)

Workload Skew

This shows skew of all workloads across the system in a graph to help clients visualize and accurately tier configurations when adding capacity or a new system. The output is illustrated in Figure 10-11.

![Figure 10-11  STAT Charting Utility Workload Skew report](image)
The STAT Charting Utility can be downloaded from the IBM support website:

### 10.2.9 More information

For more information about planning and configuration considerations, best practices, and monitoring and measurement tools, see *IBM System Storage SAN Volume Controller and Storwize V7000 Best Practices and Performance Guidelines, SG24-7521-03*, and *Implementing IBM Easy Tier with IBM Real-time Compression, TIPS1072*.

### 10.3 Thin provisioning

In a shared storage environment, thin provisioning is a method for optimizing the usage of available storage. It relies on allocation of blocks of data on demand versus the traditional method of allocating all of the blocks up front. This methodology 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. The IBM Storwize V7000 has this capability for Fibre Channel and iSCSI provisioned volumes.

An example of thin provisioning is when a storage system contains 5000 GB of usable storage capacity, but the storage administrator mapped volumes of 500 GB each to 15 hosts. In this example, the storage administrator makes 7500 GB of storage space visible to the hosts, even though the storage system has only 5000 GB of usable space, as shown in Figure 10-12. In this case, all 15 hosts cannot use immediately all 500 GB that is provisioned to them. The storage administrator must monitor the system and add storage as needed.

![Figure 10-12 Concept of thin provisioning](image-url)
You can imagine thin provisioning as the same process as when airlines sell more tickets on a flight than there are available physical seats, assuming that some passengers do not appear at check-in. They do not assign actual seats at the time of sale, which avoids each client having a claim on a specific seat number. The same concept applies to thin provisioning (airline) Storwize V7000 (plane) and its volumes (seats). The storage administrator (airline ticketing system) must closely monitor the allocation process and set proper thresholds.

### 10.3.1 Configuring a thin-provisioned volume

Volumes can be configured as **thin-provisioned** or **fully allocated**. Thin-provisioned volumes are created with real and virtual capacities. You can still create volumes by using a striped, sequential, or image mode virtualization policy, as you can with any other volume.

**Real capacity** defines how much disk space is allocated to a volume. **Virtual capacity** is the capacity of the volume that is reported to other Storwize V7000 components (such as FlashCopy or remote copy) and to the hosts. For example, you can create a volume with real capacity of only 100 GB but virtual capacity of 1 TB. The actual space used by the volume on SAN Volume Controller will be 100 GB, but hosts will see 1 TB volume.

A directory maps the virtual address space to the real address space. The directory and the user data share the real capacity.

Thin-provisioned volumes are available in two operating modes: autoexpand and non-autoexpand. You can switch the mode at any time. If you select the autoexpand feature, the Storwize V7000 automatically adds a fixed amount of more real capacity to the thin volume as required. Therefore, the autoexpand feature attempts to maintain a fixed amount of unused real capacity for the volume. This amount 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 thus has a zero contingency capacity, goes offline when the real capacity is used and must expand.

**Warning threshold**: Enable the warning threshold (by using email or an SNMP trap) when you are working with thin-provisioned volumes. You can enable it on the volume, and on the storage pool side, especially when you do not use the autoexpand mode. Otherwise, the thin volume goes offline if it runs out of space.

Autoexpand mode does not cause 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.

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 Storwize is to detect zeros on the volume you have to use some kind of software on the host side to write zeros to all unused space on the disk or file system.
Space allocation

When a thin-provisioned volume is created, a small amount of the real capacity is used for initial metadata. Write I/Os to the grains of the thin volume (that were not previously written to) cause 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) update the grain where data was previously written.

**Grain definition:** The grain is defined when the volume is created and can be 32 KB, 64 KB, 128 KB, or 256 KB.

Smaller granularities can save more space, but they have larger directories. When you use thin-provisioning with FlashCopy, specify the same grain size for the thin-provisioned volume and FlashCopy.

To create a thin-provisioned volume, choose Create Volume from the Volumes in a dynamic menu and select Thin Provision, as shown in Figure 10-13. Enter the required capacity and volume name.

![Create Volumes](image)

**Figure 10-13** Thin-provisioned volume creation

In the Advanced menu of this wizard, you can set virtual and real capacity, warning thresholds, and grain size, as shown in Figure 10-14 on page 359.
10.3.2 Performance considerations

Thin-provisioned volumes save capacity only if the host server does not write to whole volumes. Whether the thin-provisioned volume works well partly depends on how the file system allocated the space.

Some file systems (for example, New Technology File System [NTFS]) write to the whole volume before overwriting deleted files. Other file systems reuse space in preference to allocating new space.

File system problems can be moderated by tools, such as “defrag,” or by managing storage by using host Logical Volume Managers (LVMs).

The thin-provisioned volume also depends on how applications use the file system. For example, some applications delete log files only when the file system is nearly full.

**Important:** Do not use defrag on thin-provisioned volumes. Defragmentation process can write data to different areas of a volume, which can cause thin-provisioned volume to grow up to its virtual size.

There is no recommendation for thin-provisioned volumes. As explained previously, the performance of thin-provisioned volumes depends on what is used in the particular
environment. For the best performance, use fully allocated volumes instead of thin-provisioned volumes.

**Note:** Starting with Storwize firmware v7.3, all cache subsystems were rewritten and now thin-provisioned volumes can benefit from lower cache functions like coalescing writes or prefetching, which greatly improves performance.

### 10.3.3 Limitations of virtual capacity

A few factors (extent and grain size) limit the virtual capacity of thin-provisioned volumes beyond the factors that limit the capacity of regular volumes. Table 10-2 shows the maximum thin-provisioned volume virtual capacities for an extent size.

**Table 10-2  Maximum thin volume virtual capacities for an extent size**

<table>
<thead>
<tr>
<th>Extent size in MB</th>
<th>Maximum volume real capacity in GB</th>
<th>Maximum thin virtual capacity in GB</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>2048</td>
<td>2000</td>
</tr>
<tr>
<td>32</td>
<td>4096</td>
<td>4000</td>
</tr>
<tr>
<td>64</td>
<td>8192</td>
<td>8000</td>
</tr>
<tr>
<td>128</td>
<td>16,384</td>
<td>16,000</td>
</tr>
<tr>
<td>256</td>
<td>32,768</td>
<td>32,000</td>
</tr>
<tr>
<td>512</td>
<td>65,536</td>
<td>65,000</td>
</tr>
<tr>
<td>1024</td>
<td>131,072</td>
<td>130,000</td>
</tr>
<tr>
<td>2048</td>
<td>262,144</td>
<td>260,000</td>
</tr>
<tr>
<td>4096</td>
<td>262,144</td>
<td>262,144</td>
</tr>
<tr>
<td>8192</td>
<td>262,144</td>
<td>262,144</td>
</tr>
</tbody>
</table>

Table 10-3 on page 360 shows the maximum thin-provisioned volume virtual capacities for a grain size.

**Table 10-3  Maximum thin volume virtual capacities for a grain size**

<table>
<thead>
<tr>
<th>Grain size in KB</th>
<th>Maximum thin virtual capacity in GB</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>260,000</td>
</tr>
<tr>
<td>64</td>
<td>520,000</td>
</tr>
<tr>
<td>128</td>
<td>1,040,000</td>
</tr>
<tr>
<td>256</td>
<td>2,080,000</td>
</tr>
</tbody>
</table>

For more information and detailed performance considerations for configuring thin provisioning, see *IBM System Storage SAN Volume Controller and Storwize V7000 Best Practices and Performance Guidelines*, SG24-7521-03. You can also visit SAN Volume Controller 7.4 Infocenter available here:

10.4 Real-time Compression Software

The IBM Real-time Compression Software that is embedded in the IBM Storwize V7000 solution addresses the requirements of primary storage data reduction, including performance. It does so by using a purpose-built technology that is called Real-time Compression that uses the Random Access Compression Engine (RACE) engine. It offers the following benefits:

- Compression for active primary data
  IBM Real-time Compression can be used with active primary data. Therefore, it supports workloads that are not candidates for compression in other solutions. The solution supports online compression of existing data. Storage administrators can regain free disk space in an existing storage system without requiring administrators and users to clean up or archive data. This configuration significantly enhances the value of existing storage assets, and the benefits to the business are immediate. The capital expense of upgrading or expanding the storage system is delayed.

- Compression for replicated or mirrored data
  Remote volume copies can be compressed in addition to the volumes at the primary storage tier. This process reduces storage requirements in Metro Mirror and Global Mirror destination volumes as well.

- No changes to the existing environment are required
  IBM Real-time Compression is part of the storage system. It was designed with transparency in mind so that it can be implemented without changes to applications, hosts, networks, fabrics, or external storage systems. The solution is not apparent to hosts, so users and applications continue to work as-is. Compression occurs within the Storwize V7000 system.

- Overall savings in operational expenses
  More data is stored in a rack space, so fewer storage expansion enclosures are required to store a data set. This reduced rack space has the following benefits:
  - Reduced power and cooling requirements. More data is stored in a system, which requires less power and cooling per gigabyte or used capacity.
  - Reduced software licensing for more functions in the system. More data that is stored per enclosure reduces the overall spending on licensing.

Tip: Implementing compression in Storwize V7000 provides the same benefits to internal SSDs and externally virtualized storage systems.

- Disk space savings are immediate
  The space reduction occurs when the host writes the data. This process is unlike other compression solutions in which some or all of the reduction is realized only after a post-process compression batch job is run.

10.4.1 Common use cases

This section addresses the most common use cases for implementing compression:

- General-purpose volumes
- Databases
- Virtualized infrastructures
**General-purpose volumes**

Most general-purpose volumes are used for highly compressible data types, such as home directories, CAD/CAM, oil and gas geo seismic data, and log data. Storing such types of data in compressed volumes provides immediate capacity reduction to the overall consumed space. More space can be provided to users without any change to the environment.

There can be many file types stored in general-purpose servers. However, for practical information, the estimated compression ratios are based on actual field experience. Expected compression ratios are 50% to 60%.

File systems that contain audio, video files, and compressed files are not good candidates for compression. The overall capacity savings on these file types are minimal.

**Databases**

Database information is stored in table space files. It is common to observe high compression ratios in database volumes. Examples of databases that can greatly benefit from Real-Time Compression are IBM DB2®, Oracle, and Microsoft SQL Server. Expected compression ratios are 50% to 80%.


**Virtualized infrastructures**

The proliferation of open systems virtualization in the market has increased the use of storage space, with more virtual server images and backups kept online. The use of compression reduces the storage requirements at the source.

Examples of virtualization solutions that can greatly benefit from Real-time Compression are VMware, Microsoft Hyper-V, and KVM. Expected compression ratios are 45% to 75%.

| Tip: Virtual machines with file systems that contain compressed files are not good candidates for compression as described in “General-purpose volumes”. |

**10.4.2 Real-time Compression concepts**

RACE technology is based on over 50 patents that are not primarily about compression. Instead, they define how to make industry standard Lempel-Ziv (LZ) compression of primary storage operate in real-time and allow random access. The primary intellectual property behind this is the RACE engine.

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 in-line compression technology, meaning that each host write is compressed as it passes through the Storwize V7000 software to the disks. This 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 in a real-time method. When a host sends a write request, it is acknowledged by the write cache of the system, and then staged to the storage pool. As part of its staging, it 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 Real-time Compression is a self-tuning solution. It is adapting to the workload that runs on the system at any particular moment.

10.4.3 Random Access Compression Engine

To understand why RACE is unique, you need to review the traditional compression techniques. This description is not about the compression algorithm itself, that is, how the data structure is reduced in size mathematically. Rather, the description is about how the data is laid out within the resulting compressed output.

Compression utilities

Compression is probably most known to users because of the widespread use of compression utilities such as the zip and gzip utilities. At a high level, these utilities take a file as their input, and parse the data by using a sliding window technique. Repetitions of data are detected within the sliding window history, most often 32 KB. Repetitions outside of the window cannot be referenced. Therefore, the file cannot be reduced in size unless data is repeated when the window “slides” to the next 32 KB slot.

Figure 10-15 on page 364 shows compression that uses a sliding window, where the first two repetitions of the string “ABCDEF” fall within the same compression window, and can therefore be compressed using the same dictionary. Note, the third repetition of the string falls outside of this window, and therefore cannot be compressed using the same compression dictionary as the first two repetitions, reducing the overall achieved compression ratio.
Implementing the IBM Storwize V7000 V7.4

Traditional data compression in storage systems

The traditional approach taken to implement data compression in storage systems is an extension of how compression works in the compression utilities previously mentioned. Similar to compression utilities, the incoming data is broken into fixed chunks, and then each chunk is compressed and extracted independently.

However, there are drawbacks to this approach. An update to a chunk requires a read of the chunk followed by a recompression of the chunk to include the update. The larger the chunk size chosen, the heavier the I/O penalty to recompress the chunk. If a small chunk size is chosen, the compression ratio is reduced because the repetition detection potential is reduced.

Figure 10-16 on page 365 shows an example of how the data is broken into fixed size chunks (in the upper-left side of the figure). It also shows how each chunk gets compressed independently into variable length compressed chunks (in the upper-right side of the figure). The resulting compressed chunks are stored sequentially in the compressed output.

Although this approach is an evolution from compression utilities, it is limited to low performance use cases. This limitation is mainly because it does not provide real random access to the data.
Random Access Compression Engine

The IBM patented Random Access Compression Engine implements an inverted approach when compared to traditional approaches to compression. RACE uses variable-size chunks for the input, and produces fixed-size chunks for the output.

This method enables an efficient and consistent method to index the compressed data because it is stored in fixed-size containers.

Figure 10-17 on page 366 shows Random Access Compression.
Both compression utilities and traditional storage systems compression compress data by finding repetitions of bytes within the chunk that is being compressed. The compression ratio of this chunk depends on how many repetitions can be detected within the chunk. The number of repetitions is affected by how much the bytes stored in the chunk are related to each other. The relation between bytes is driven by the format of the object. For example, an office document might contain textual information, and an embedded drawing (like this page). Because the chunking of the file is arbitrary, it has no notion of how the data is laid out within the document. Therefore, a compressed chunk can be a mixture of the textual information and part of the drawing. This process yields a lower compression ratio because the different data types mixed together cause a suboptimal dictionary of repetitions. That is, fewer repetitions can be detected because a repetition of bytes in a text object is unlikely to be found in a drawing.

This traditional approach to data compression is also called location-based compression. The data repetition detection is based on the location of data within the same chunk.

This challenge was also addressed with the predecide mechanism that was introduced in version 7.1.
Predecide mechanism
Some data chunks have a higher compression ratio than others. Compressing some of the chunks will save very little space, but still will require resources such as CPU and memory. To avoid spending resources on uncompressible data, and to provide the ability to use a different, more effective (in this particular case) compression algorithm, IBM has invented a predecide mechanism that was first introduced in version 7.1.

The chunks that are below a given compression ratio are skipped by the compression engine thus saving CPU time and memory processing. Chunks that are decided not to be compressed with the main compression algorithm, but that still can be compressed well with the other, will be marked and processed accordingly. The result may vary since predecide does not check the entire block, only a sample of it.

Figure 10-18 shows how the detection mechanism works.

Temporal compression
RACE offers a technology leap beyond location-based compression, temporal compression.

When host writes arrive to RACE, they are compressed and fill up fixed size chunks also called compressed blocks. Multiple compressed writes can be aggregated into a single compressed block. A dictionary of the detected repetitions is stored within the compressed block. When applications write new data or update existing data, it is typically sent from the host to the storage system as a series of writes. Because these writes are likely to originate from the same application and be from the same data type, more repetitions are usually detected by the compression algorithm.

This type of data compression is called temporal compression because the data repetition detection is based on the time the data was written into the same compressed block. Temporal compression adds the time dimension that is not available to other compression
Implementing the IBM Storwize V7000 V7.4

algorithms. It offers a higher compression ratio because the compressed data in a block represents a more homogeneous set of input data.

Figure 10-19 shows (in the upper part) how three writes sent one after the other by a host end up in different chunks. They get compressed in different chunks because their location in the volume is not adjacent. This yields a lower compression ratio because the same data must be compressed non-natively by using three separate dictionaries. When the same three writes are sent through RACE (in the lower part of the figure), the writes are compressed together by using a single dictionary. This yields a higher compression ratio than location-based compression.

10.4.4 Random Access Compression Engine in Storwize V7000 software stack

It is important to understand where the RACE technology is implemented in the Storwize V7000 software stack. This location determines how it applies to other Storwize components.

RACE technology is implemented into the Storwize thin provisioning layer, and is an organic part of the stack. The Storwize V7000 software stack is shown in Figure 10-20 on page 369. Compression is transparently integrated with existing system management design. All of the Storwize advanced features are supported on compressed volumes. You can create, delete,
migrate, map (assign), and unmap (unassign) a compressed volume as though it were a fully allocated volume. In addition, you can use Real-time Compression along with Easy Tier on the same volumes. This compression method provides nondisruptive conversion between compressed and decompressed volumes. This conversion provides a uniform user-experience and eliminates the need for special procedures when dealing with compressed volumes.

![Diagram of I/O flow](image)

**Figure 10-20  RACE integration within Storwize V7000 7.4 software stack**

### 10.4.5 Data write flow

When a host sends a write request to Storwize V7000, it reaches the upper cache layer. The host is immediately sent an acknowledgment of its I/Os.

When the upper cache layer destages to the RACE, the I/Os are sent to the thin-provisioning layer. They are then sent to RACE, and if necessary, to the original host write or writes. The metadata that holds the index of the compressed volume is updated if needed, and is compressed as well.

### 10.4.6 Data read flow

When a host sends a read request to the Storwize V7000 for compressed data, it is forwarded directly to the Real-time Compression (RtC) component:

- If the RtC component contains the requested data, the Storwize V7000 cache replies to the host with the requested data without having to read the data from the lower-level cache or disk.
If the RtC component does not contain the requested data, the request is forwarded to the Storwize V7000 lower-level cache.

If the lower-level cache contains the requested data, it is sent up the stack and returned to the host without accessing the storage.

If the lower-level cache does not contain the requested data, it sends a read request to the storage for the requested data.

10.4.7 Compression of existing data

In addition to compressing data in real time, it is also possible to compress existing data sets. To do that, you need to add a compressed mirrored copy of an existing volume by right-clicking a particular volume and selecting Add Mirrored Copy from Volume Copy Actions as shown on Figure 10-21.

Then, you select a compressed type of volume and storage pool when you want to place the new copy. If you do not want to move volume to different storage, select the same storage pool as one for existing, original volume copy.
After the copies are fully synchronized, you can delete original, non-compressed copy as shown on Figure 10-23 on page 372.
As a result of this, you end up having a compressed data on existing volume as shown on Figure 10-24. This process is nondisruptive, so the data remains online and accessible by applications and users.

This capability enables customers to regain space from the storage pool, which can then be reused for other applications.
With virtualization of external storage systems, the ability to compress already stored data significantly enhances and accelerates the benefit to users. It allows them to see a tremendous return on their Storwize V7000 investment. On initial purchase of a Storwize V7000 with Real-time Compression, customers can defer their purchase of new storage. As new storage needs to be acquired, IT purchases a lower amount of the required storage before compression.

**Important:** Remember that Storwize will reserve some of its resources like CPU cores and RAM memory after you create just one compressed volume or volume copy. This may impact your system performance if not planned accordingly in advance.

### 10.4.8 Configuring compressed volumes

To use compression on the Storwize V7000, licensing is required. With the Storwize V7000, Real-time Compression is licensed by capacity, per terabyte of virtual data.

The configuration is similar to generic volumes and not apparent to users. From Volumes in the dynamic menu, choose **Create Volume** and select **Compressed**, as shown in Figure 10-25. Choose the wanted storage pool and enter the required capacity and volume name.

![Create Volume Interface](image)

**Figure 10-25** Configuring compressed volume

The summary at the bottom of the wizard provides the information that is about allocated (virtual) capacity and the real capacity that data uses on this volume. In our example, we defined a 15 GB volume, but the real capacity is only 307 MB because there is no data from the host.
When the compressed volume is configured, you can directly map it to the host or do so later on request. For more information about compressed volumes, see 10.4.8, “Configuring compressed volumes” on page 373.

10.4.9 Storwize V7000 Generation 2 software and hardware updates related to Real-time Compression

The Storwize V7000 Generation 2 and its software version 7.3 introduce significant software and hardware improvements that enhance and extend the applicability of the Real-time Compression feature. In this section, we provide an overview of these enhancements:

- Software enhancements
  - Storwize V7000 cache rewrite
- Hardware enhancements
  - Enhanced CPU options
  - Increased memory options
  - Optional Intel Assist Acceleration Technology (Coleto Creek) compression acceleration cards

10.4.10 Software enhancements

Storwize V7000 software version 7.3 introduces an enhanced, dual-level caching model. This model differs from the single-level cache model of previous software versions.

In the previous model, the Real-time Compression software component sits below the single-level read/write cache. The benefit of this model is that the upper-level read/write cache masks from the host any latency introduced by the Real-time Compression software component. However, in this single-level caching model, the destaging of writes for compressed I/Os to disk might not be optimal for certain workloads because the RtC component is interacting direct with uncached storage. Figure 10-26 on page 375 depicts compression code in the current Storwize software stack.
In the new, dual-level caching model, the Real-time Compression software component sits below the upper-level fast write cache and above the lower-level advanced read/write cache. There are several advantages to this dual-level model regarding Real-time Compression:

- Host writes, whether to compressed or decompressed volumes, are still serviced directly via the upper-level write cache, preserving low host write I/O latency. Response time can improve with this model as the upper cache flushes less data to RACE more frequently.
- The performance of the destaging of compressed write I/Os to storage is improved because these I/Os are now destaged via the advanced lower-level cached, as opposed to directly to storage.
- The existence of a lower-level write cache below the Real-time Compression component in the software stack allows for the coalescing of compressed writes, and as a result, a reduction in back-end I/Os due to the ability to perform full-stride writes for compressed data.
- The existence of a lower-level read cache below the Real-time Compression component in the software stack allows the temporal locality nature of RtC to benefit from pre-fetching from the backend storage.
- The main (lower level) cache now stores compressed data for compressed volumes, increasing the effective size of the lower-level cache.
- Support for larger numbers of compressed volumes.

*Figure 10-26  Real Time Compression code in Storwize V7000 software stack*
10.4.11 Hardware updates

Storwize V7000 Gen2 introduces numerous hardware enhancements. Several of these enhancements relate directly to the Real-time Compression feature and offer significant performance and scalability improvements over previous hardware versions.

Enhanced CPU options
The Storwize V7000 Generation 2 node canisters offer an updated primary CPU that contains eight cores per node canister as compared to the four core CPUs available in previous hardware versions. When using compression, four of those nodes are dedicated to processing compression traffic and the rest four are dedicated to process normal, decompressed traffic. This is a major enhancement comparing to previous generation where three of four CPU cores were dedicated for compression.

Increased memory options
The second generation of Storwize V7000 offers the option to increase the node canister memory from the base 32 GB to 64 GB, for use with Real-time Compression. This additional, compression dedicated memory allows for improved overall system performance when utilizing compression over previous hardware models.

On-board Intel Quick Assist Acceleration Technology (Coleto Creek) compression acceleration card
The second generation of Storwize V7000 introduces an on-board compression acceleration card. This card is standard equipment for second-generation hardware and comes installed with every Storwize V7000 Gen2. Its purpose is to support central CPU cores when using compressed volumes. To use this on-board compression acceleration card, a special compression pass-through adapter is installed in the first PCI slot of each node canister.

Optional Intel Quick Assist Acceleration Technology (Coleto Creek) compression acceleration card
The Storwize V7000 Generation 2 models offers the option to include one additional Intel Quick Assist compression acceleration card based on the Coleto Creek chipset. The introduction of this Intel based compression acceleration card in the Storwize V7000 Generation 2 models is an industry first, providing dedicated processing power and greater throughout over previous models. An additional compression acceleration card is installed in first PCI slot of each node canister and replaces default pass-through adapter.

Note: With a single on-board card, the maximum number of compressed volumes per I/O group is 200. With the addition of a second Quick Assist card, the maximum number of compressed volumes per I/O group is 512.

10.4.12 Dual RACE engine
In the 7.4 Storwize firmware version compression code was enhanced by addition of second RACE engine per Storwize node canister. This feature takes advantage of multi-core controller architecture and makes better use of compression accelerator cards. Basically, what the dual RACE engine does is it adds the second RACE instance, which works in parallel with the first instance, as shown on Figure 10-27 on page 377.
Thanks to dual RACE enhancement, the compression performance can be boosted up to two times for compressed workloads, comparing to earlier Storwize code.

To take advantage of dual RACEs, several software and hardware requirements must be met:

- Storwize V7000 software must be at level 7.4
- Only Storwize V7000 Generation 2 systems are supported
- Additional 32 GB must be installed per Storwize node canister
- Second Coleto Creek acceleration card must be installed per Storwize node canister

**Note:** We recommend using compression with dual RACEs and two acceleration cards for best performance.

When using the dual RACE feature, the acceleration cards are shared between RACE engines, which means acceleration cards are used simultaneously by both RACE engines. The rest of resources, like CPU cores and RAM memory are evenly divided between the RACE engines. There is no need to manually enable dual RACEs; it triggers automatically when all minimal software and hardware requirements are met. In case Storwize is compression-capable but the minimal requirements for dual RACE are not met, only one RACE instance is used as it was in the previous versions of Storwize code.

Figure 10-28 on page 378 shows how Storwize resources are split when using compression.
Figure 10-28  Storwize CPU cores and RAM memory allocation

For more information about Real-time Compression and its deployment in IBM Storwize V7000, see Real-time Compression in SAN Volume Controller and Storwize V7000, REDP-4859.
Advanced Copy Services

Before you proceed with this chapter, review the content of the Advanced Copy Services Overview in Chapter 1, “IBM Storwize V7000 system overview” on page 1, where we first describe these functions at a high-level view.

In this chapter, we describe the Advanced Copy Services functions that are also available for the SAN Volume Controller and most of these functions are also available for the IBM Storwize product family.

In this chapter, we also describe the native IP Replication.

In Chapter 9, “External storage virtualization” on page 323, we describe how to use the command-line interface (CLI) and Advanced Copy Services.

This chapter includes the following topics:

- FlashCopy
- Reverse FlashCopy
- FlashCopy functional overview
- Implementing Storwize V7000 FlashCopy
- Volume mirroring and migration options
- Native IP replication
- Remote Copy
- Remote Copy commands
- Metro Mirror process
- Metro Mirror commands
- Global Mirror process
- Global Mirror commands
- Troubleshooting remote copy
11.1 FlashCopy

By using the FlashCopy function of the Storwize V7000, you can perform a *point-in-time copy* of one or more volumes. In this section, we describe the inner workings of FlashCopy and provide details of 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 cache and therefore, 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 frozen briefly to initialize the FlashCopy bitmap and then 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 a while to complete, the resulting data on the target volume is presented so that the copy appears to complete immediately. This process is done by using a bitmap (or bit array), which tracks changes to the data after the FlashCopy is started and an indirection layer, which allows data to be read from the source volume transparently.

### 11.1.1 Business requirements for FlashCopy

When you are deciding whether FlashCopy addresses your needs, you must adopt a combined business and technical view of the problems that you want to solve. First, determine the needs from a business perspective. Then, determine whether FlashCopy can address the technical needs of those business requirements.

The business applications for FlashCopy are wide-ranging. 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 Storwize V7000 is flexible and has a broad feature set, which makes it applicable to many scenarios.

### 11.1.2 Back up 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 or 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 complete, 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.

11.1.3 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. (It might be easier to think of this method 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.
- It 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 target volume to modified blocks on the source volume, the target can be used to perform a restore of individual files; you make it available on a host. We suggest that you do not make it available to a source host because seeing doubles of disks causes problems for most host operating systems. Copy the files to the source via normal host data copy methods for your environment.

11.1.4 Moving and migrating data with FlashCopy

FlashCopy can be used to facilitate the movement or migration of data between hosts while minimizing downtime for applications. By using FlashCopy, application data can be copied from source volumes to new 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.
11.1.5 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.

11.1.6 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 Storwize V7000 and therefore, 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.

Note: Although the best way to perform FlashCopy is to flush host cache first, some companies, like Oracle, support using snapshots without it, as stated in Metalink note 604683.1.

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 opposed to flushing data from both the application and the backing database, which is always the recommended method because it is safer. However, this method can be used when facilities do not exist or your environment includes time sensitivity.

11.1.7 FlashCopy attributes

The FlashCopy function in Storwize V7000 features the following attributes:

- The target is the time-zero copy of the source, which is known as FlashCopy mapping targets.
FlashCopy produces an exact 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 Storwize V7000 clustered 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 in 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 nodes of the Storwize V7000 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 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 Storwize V7000 system.

The size of the source and target volumes cannot be altered (increased or decreased) while a FlashCopy mapping is defined.

11.2 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 thus multiple rollback points.

A key advantage of the Storwize V7000 Multiple Target Reverse FlashCopy function is that the reverse FlashCopy does not destroy the original target, which allows processes by using the target, such as a tape backup, to continue uninterrupted.
Storwize V7000 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 11-1 shows an example of 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.
11.2.1 FlashCopy and Tivoli Storage FlashCopy Manager

The management of many large FlashCopy relationships and Consistency Groups is a complex task without a form of automation for assistance.

IBM Tivoli Storage FlashCopy Manager provides fast application-aware backups and restores using advanced point-in-time image technologies in the Storwize V7000. In addition, it provides an optional integration with IBM Tivoli Storage Manager for the long-term storage of snapshots. Figure 11-2 shows the integration of Tivoli Storage Manager and FlashCopy Manager from a conceptual level.

![Figure 11-2 Tivoli Storage Manager for Advanced Copy Services features](image)

Tivoli FlashCopy Manager provides many of the features of Tivoli Storage Manager for Advanced Copy Services without the requirement to use Tivoli Storage Manager. With Tivoli FlashCopy Manager, you can coordinate and automate host preparation steps before you issue FlashCopy start commands to ensure that a consistent backup of the application is made. You can put databases into hot backup mode and flush file system cache before starting the FlashCopy.

FlashCopy Manager also allows for easier management of on-disk backups that use FlashCopy, and provides a simple interface to perform the “reverse” operation.

Figure 11-3 on page 386 shows the FlashCopy Manager feature.
Released December 2013, IBM Tivoli FlashCopy Manager V4.1 adds support for VMware 5.5 and vSphere environments with Site Recovery Manager (SRM), along with instant restore for virtual machine file system (VMFS) data stores. This release also integrates with IBM Tivoli Storage Manager for Virtual Environments, and it allows backup of point-in-time images into the Tivoli Storage Manager infrastructure for long-term storage.

The addition of VMware vSphere brings support and application awareness for FlashCopy Manager up to the following applications:

- Microsoft Exchange and Microsoft SQL Server, including SQL Server 2012 Availability Groups
- IBM DB2® and Oracle databases, for use either with or without SAP environments
- IBM GPFS software snapshots for DB2 IBM pureScale®
- Other applications can be supported via script customizing

For more information about IBM Tivoli FlashCopy Manager, see this website:

FlashCopy Manager integration with Remote Copy Services
One of the interesting features of FlashCopy Manager is the capability of creating FlashCopy backups from remote copy target volumes. This means that backup does not have to be copied from the primary site to secondary site because it is already copied via MM/GMMM/GM or Global Mirror. Applications running in the primary site can have its backup taken in the secondary site, where the source of this backup is target remote copy volumes. This concept is presented on Figure 11-4 on page 387.
## 11.3 FlashCopy functional overview

FlashCopy works by defining a FlashCopy mapping that consists of one source volume with one target volume. Multiple FlashCopy mappings (source-to-target relationships) 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 391.

Before you start a FlashCopy (regardless of the type and options specified), you must issue a `prestartfcmap` or `prestartfconsistgrp`, which puts the Storwize V7000 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).

**Note:** Instead of using `prestartfcmap` or `prestartfconsistgrp`, you can also use the `-prep` parameter in the `startfcmap` or `startfconsistgrp` command to prepare and start FlashCopy in one step.

Immediately following the FlashCopy operation, the source and target volumes are available for use. 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 target and updates are made to the source.

For more information about background copy, see 11.4.5, “Grains and the FlashCopy bitmap” on page 393. Figure 11-5 shows the redirection of the host I/O toward the source volume and the target volume.

Figure 11-5  Redirection of host I/O

11.4 Implementing Storwize V7000 FlashCopy

In this section, we describe how FlashCopy is implemented in the Storwize V7000.

11.4.1 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 Storwize V7000 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 Storwize V7000 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 either a stand-alone mapping or a Consistency Group.

Figure 11-6 shows the concept of FlashCopy mapping.

![FlashCopy mapping](image)

### 11.4.2 Multiple Target FlashCopy

The Storwize V7000 supports up to 256 target volumes from a single source volume. Each copy is managed by a unique mapping. Figure 11-7 shows the Multiple Target FlashCopy implementation.

**Important:** Older copies must be complete for independent FlashCopy mappings.

![Multiple Target FlashCopy implementation](image)

Figure 11-7 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 11-7). The older targets refer to new targets first before referring to the source.
From the point of view of an intermediate target disk (neither 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 11.4.6, “Interaction and dependency between Multiple Target FlashCopy mappings” on page 394.

### 11.4.3 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 11-8 shows a Consistency Group that includes two FlashCopy mappings.

![FlashCopy Consistency Group](image)

*Figure 11-8 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 Storwize V7000 supports the concept of Consistency Groups.

A FlashCopy Consistency Group can contain up to 512 FlashCopy mappings. The maximum number of FlashCopy mappings that is supported by the Storwize V7000 system v7.4 is 4096. 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 11-1 on page 392 lists the FlashCopy properties and maximum configurations.
11.4.4 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 that is 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

<table>
<thead>
<tr>
<th>Table 11-1 FlashCopy properties and maximum configurations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FlashCopy property</strong></td>
</tr>
<tr>
<td>------------------------</td>
</tr>
<tr>
<td>FlashCopy targets per source</td>
</tr>
<tr>
<td>FlashCopy mappings per system</td>
</tr>
<tr>
<td>FlashCopy Consistency Groups per system</td>
</tr>
<tr>
<td>FlashCopy volume capacity per I/O Group</td>
</tr>
<tr>
<td>FlashCopy mappings per Consistency Group</td>
</tr>
</tbody>
</table>
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.

### 11.4.5 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 KB or 256 KB grain sizes (256 KB 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 KB grain size uses bitmap space at a rate of four times the default 256 KB 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:
- If the grain was not copied to the target yet, the grain is copied before actual write is performed to source. Bitmap is updated indicating this grain is already copied to the target.
- If the grain was already 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:
- If the grain was not copied from source to target, the grain is copied from the source to the target before the actual write is performed to source. Bitmap is updated indicating 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, 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 11-9 on page 394 shows how the background copy runs while I/Os are handled according to the indirection layer algorithm.
11.4.6 Interaction and dependency between Multiple Target FlashCopy mappings

Figure 11-10 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.
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 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 the source if none are 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 are 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.

For example, if the mapping that is associated with Target 0 was issued a stopfcmap or stopfconsistgrp 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.

### 11.4.7 Summary of the FlashCopy indirection layer algorithm

Table 11-2 on page 396 summarizes the indirection layer algorithm.
11.4.8 Interaction with the cache

Starting with Storwize firmware v7.3 the entire cache subsystem was redesigned and changed accordingly. Cache has been 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 11-11 shows the new Storwize cache architecture.

![New Cache Architecture Diagram]

**Figure 11-11  New cache architecture**

This copy-on-write process introduces significant latency into write operations. To isolate the active application from this additional latency, the FlashCopy indirection layer is placed logically between upper and lower cache. Therefore, the additional latency that is introduced

<table>
<thead>
<tr>
<th>Accessed volume</th>
<th>Was the grain copied?</th>
<th>Host I/O operation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source</td>
<td>No</td>
<td>Read from the source volume.</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>Read from the source volume.</td>
</tr>
<tr>
<td>Target</td>
<td>No</td>
<td>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.</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>Read from the target volume.</td>
</tr>
</tbody>
</table>

### Table 11-2  Summary table of the FlashCopy indirection layer algorithm

<table>
<thead>
<tr>
<th>Accessed volume</th>
<th>Was the grain copied?</th>
<th>Host I/O operation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source</td>
<td>No</td>
<td>Read from the source volume.</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>Read from the source volume.</td>
</tr>
<tr>
<td>Target</td>
<td>No</td>
<td>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.</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>Read from the target volume.</td>
</tr>
</tbody>
</table>
by the copy-on-write process is encountered only by the internal cache operations and not by the application.

The logical placement of the FlashCopy indirection layer is shown in Figure 11-12.

![Figure 11-12 Logical placement of the FlashCopy indirection layer](image)

Introduction of the two level cache provides additional performance improvements to FlashCopy mechanism. Because now the FlashCopy layer is above lower cache in Storwize 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. Additionally, in the multitarget FlashCopy the target volumes of the same image share cache data, which is opposite to previous Storwize code versions where each volume had its own copy of cached data.

### 11.4.9 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 `Storwizeinfo lsvdisk -bytes volumeName` 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. In Example 11-1 on page 398, we list the size of the `ds_3400_img_vol` volume. The `ds_3400_img_vol_copy` volume is then created, which specifies the same size.
Example 11-1  Listing the size of a volume in bytes and creating a volume of equal size

IBM_2145:ITSO_Storwize2:superuser>lsvdisk -bytes ds_3400_img_vol
id 22
name ds_3400_img_vol
IO_group_id 0
IO_group_name io_grp0
status online
mdisk_grp_id 5
mdisk_grp_name MigrationPool_1024
capacity 7516192768
type image
formatted no
mdisk_id 15
mdisk_name mdisk15
FC_id
FC_name
RC_id
RC_name
vdisk_UID 6005076801FD802840000000000000038
throttling 0
preferred_node_id 2
fast_write_state not_empty
cache readwrite
uuid
fc_map_count 0
sync_rate 80
copy_count 1
se_copy_count 0
filesystem
mirror_write_priority latency
RC_change no
compressed_copy_count 0
access_IO_group_count 1
last_access_time
parent_mdisk_grp_id 5
parent_mdisk_grp_name MigrationPool_1024

IBM_2145:ITSO_Storwize2:superuser>mkvdisk -iogrp 0 -mdiskgrp test_pool_01 -size 7516192768 -unit b -name ds_3400_img_vol_copy
Virtual Disk, id [23], successfully created

IBM_2145:ITSO_Storwize2:superuser>lsvdisk -delim " 
22 ds_3400_img_vol 0 io_grp0 online 5 MigrationPool_1024 7.00GB image 6005076801FD802840000000000000038 0 1 not_empty 0 no 0 5 MigrationPool_1024 23 ds_3400_img_vol_copy 0 io_grp0 online 1 test_pool_01 7.00GB striped 6005076801FD802840000000000000039 0 1 not_empty 0 no 0 1 test_pool_01

Tip: Alternatively, you can use the expandvolumesize and shrinkvolumesize volume commands to modify the size of the volume.

You can use an image mode volume as a FlashCopy source volume or target volume.
11.4.10 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 11-3.

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>
<thead>
<tr>
<th>Mapping event</th>
<th>Description</th>
</tr>
</thead>
</table>
| Create        | 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:  
   - 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       | 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.  
   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. |
| Flush done    | 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. |


<table>
<thead>
<tr>
<th>Mapping event</th>
<th>Description</th>
</tr>
</thead>
</table>
| Start                  | 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. The following actions occur during the running of the `startfcmap` command or the `startfcconsistgrp` command:  
  - 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 allow 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.  
  As part of the `startfcmap` or `startfcconsistgrp` command, read and write caching is enabled for the source and target volumes.                                                                 |
| Modify                 | The following FlashCopy mapping properties can be modified:  
  - 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                   | The following separate mechanisms can be used to stop a FlashCopy mapping:  
  - Issue a command  
  - An I/O error occurred |
| Delete                 | 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. |
| Flush failed           | If the flush of data from the cache cannot be completed, the FlashCopy mapping enters the Stopped state. |
| Copy complete          | 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. |
| Bitmap online/offline  | The node failed. |

### 11.4.11 FlashCopy mapping states

In this section, we describe 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 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, Storwize V7000 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 11-4 lists the various FlashCopy mapping states and the corresponding states of the source and target volumes.

<table>
<thead>
<tr>
<th>State</th>
<th>Source</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Online/Offline</td>
<td>Cache state</td>
</tr>
<tr>
<td>Idling/Copied</td>
<td>Online</td>
<td>Write-back</td>
</tr>
<tr>
<td>Copying</td>
<td>Online</td>
<td>Write-back</td>
</tr>
<tr>
<td>Stopped</td>
<td>Online</td>
<td>Write-back</td>
</tr>
<tr>
<td>Stopping</td>
<td>Online</td>
<td>Write-back</td>
</tr>
<tr>
<td>Suspended</td>
<td>Offline</td>
<td>Write-back</td>
</tr>
<tr>
<td>Preparing</td>
<td>Online</td>
<td>Write-through</td>
</tr>
<tr>
<td>Prepared</td>
<td>Online</td>
<td>Write-through</td>
</tr>
</tbody>
</table>

**11.4.12 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.

11.4.13 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 the use of 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 11-5 shows the relationship of the background copy rate value to the attempted number of grains to be copied per second.

<table>
<thead>
<tr>
<th>Value</th>
<th>Data copied per second</th>
<th>Grains per second (256 KB grain)</th>
<th>Grains per second (64 KB grain)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - 10</td>
<td>128 KB</td>
<td>0.5</td>
<td>2</td>
</tr>
<tr>
<td>11 - 20</td>
<td>256 KB</td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 11-5  Background copy rate
The grains per second numbers represent the maximum number of grains that the Storwize V7000 copies per second, assuming that the bandwidth to the managed disks (MDisks) can accommodate this rate.

If the Storwize V7000 cannot achieve these copy rates because of insufficient bandwidth from the Storwize V7000 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.

### 11.4.14 Synthesis

The FlashCopy functionality in Storwize V7000 creates copies of the volumes. All of the data in the source volume is copied to the destination volume, including operating system, logical volume manager, and application metadata.

**Synthesis:** Certain operating systems cannot use FlashCopy without another step, which is called *synthesis*. Synthesis performs a type of transformation on the operating system metadata that is on the target volume so that the operating system can use the disk.

### 11.4.15 Serialization of I/O by FlashCopy

In general, the FlashCopy function in the Storwize V7000 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.

11.4.16 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 Storwize V7000 nodes to each other and to the MDisk is made up of many independent links, it is possible for a subset of the nodes to be temporarily isolated from several of the MDisk. 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.

When an MDisk enters the Path Offline state on a Storwize V7000 node, all of the volumes that have extents on the MDisk also become Path Offline. Again, this situation happens only on the affected nodes. When a volume is Path Offline on a particular Storwize V7000 node, the host access to that volume through the node fails with the SCSI check condition indicating Offline.

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.

11.4.17 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.

11.4.18 Interoperation with Metro Mirror and Global Mirror

FlashCopy can work with Metro Mirror and Global Mirror to provide better protection of the data. For example, we can perform a Metro Mirror copy to duplicate data from Site_A to Site_B and then perform a daily FlashCopy to back up the data to another location.

Table 11-6 lists the supported combinations of FlashCopy and remote copy. In the table, remote copy refers to Metro Mirror and Global Mirror.

Table 11-6  FlashCopy and remote copy interaction

<table>
<thead>
<tr>
<th>Component</th>
<th>Remote copy primary site</th>
<th>Remote copy secondary site</th>
</tr>
</thead>
<tbody>
<tr>
<td>FlashCopy Source</td>
<td>Supported</td>
<td>Supported</td>
</tr>
<tr>
<td></td>
<td></td>
<td>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.</td>
</tr>
</tbody>
</table>
11.4.19 FlashCopy presets

The Storwize V7000 GUI interface provides three FlashCopy presets (Snapshot, Clone, and Backup) to simplify the more common FlashCopy operations.

Although these presets meet most FlashCopy requirements, they do not provide support for 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.

In this section, we describe the three 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, 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: 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.

<table>
<thead>
<tr>
<th>Component</th>
<th>Remote copy primary site</th>
<th>Remote copy secondary site</th>
</tr>
</thead>
<tbody>
<tr>
<td>FlashCopy Target</td>
<td>This is a supported combination and has the following restrictions:</td>
<td>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.</td>
</tr>
<tr>
<td></td>
<td>- Issuing a <code>stop -force</code> might cause the remote copy relationship to be fully resynchronized.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Code level must be 6.2.x or higher.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- I/O Group must be the same.</td>
<td></td>
</tr>
</tbody>
</table>

FlashCopy Target

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.
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 the loss of the underlying physical controller. The user plans to periodically update the secondary copy and does not want to suffer the overhead of creating a new 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 thin-provisioned on 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.

11.5 Volume mirroring and migration options

Volume mirroring is a simple RAID 1-type function that allows a volume to remain online even when the storage pool 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
With volume mirroring, host I/O completes when both copies are written. Before version 6.3.0, this feature took a copy offline when it had an I/O timeout, and then resynchronized with the online copy after it recovered. With 6.3.0, 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 \textit{Latency} or \textit{Redundancy}.

The Latency tuning option, which is set with \texttt{Storwizetask chvdisk -mirrowritepriority latency}, is the default. This behavior was available in releases before 6.3.0. It prioritizes host I/O latency, which yields a preference to host I/O over availability.

However, you might have a need in your environment to give preference to Redundancy when availability is more important than I/O response time. Use \texttt{Storwizetask chvdisk -mirrowritepriority redundancy}.

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 Storwize V7000 as a migration device. For example, vendor A's product cannot communicate with vendor B's product, but you must migrate existing 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 an existing storage MDisk or logical unit number (LUN) with its existing data from an external storage system without putting metadata on it so that the existing 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 to migrate data between storage pools and devices. You might use this option if you want to move volumes to multiple storage pools. Volume can have two copies at a time, which means you can add only one copy to the original volume and then you have to 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.

\textbf{Migration}: While these migration methods do not disrupt access, you must take a brief outage to install the host drivers for your Storwize V7000 if you do not already have them installed.

With volume mirroring, you can move data to different MDisks within the same storage pool or move data between different storage pools. There is a benefit when using volume mirroring...
over volume migration because with volume mirroring storage pools do not need to have the same extent size as is a 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 firmware 7.3 and the introduction of the new cache architecture, mirrored volume performance has been significantly improved. Now, lower cache is beneath the volume mirroring layer, which means both copies have its own cache. This helps in cases of having copies of different types, for example generic and compressed, because now both copies use its independent cache and each copy can do its own read prefetch now. 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 utilization and this can be done for each copy independently.

**Managing Volume Mirror and migration with the GUI**

For more information about volume mirroring and migration operations by using the Storwize V7000 GUI, see Chapter 8, “Advanced host and volume administration” on page 247.

## 11.6 Native IP replication

Before we describe Remote Copy features that benefit from the use of multiple Storwize V7000 systems, it is important to describe the new partnership option that is introduced with the 7.2 version of Storwize V7000 code: native IP replication.

### 11.6.1 Native IP replication technology

Remote Mirroring over IP communication is now supported on IBM SAN Volume Controller and Storwize Family systems by using Ethernet communication links. Storwize V7000 IP replication uses innovative Bridgeworks SANSlide technology to optimize network bandwidth and utilization. This new function enables the use of lower speed and lower-cost networking infrastructure for data replication. Bridgeworks’ SANSlide technology that is integrated into IBM SAN Volume Controller and Storwize Family Software uses artificial intelligence to help optimize network bandwidth utilization and adapt to changing workload and network conditions. This technology can improve remote mirroring network bandwidth utilization 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.
In a typical Ethernet network data flow, the data transfer slows down over time. This condition occurs because of the latency that is caused by waiting for 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 11-13.

By using the Bridgeworks SANSlide technology, this typical behavior can be eliminated with enhanced parallelism of the data flow by using multiple virtual connections (VCs) 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 it 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 11-27 on page 430.

For more information about Bridgeworks SANSlide technology, see IBM Storwize V7000 and SANSlide Implementation, REDP-5023.
With native IP partnership, the following Copy Services features are supported:

- **Metro Mirror**
  Referred to as *synchronous replication*, Metro Mirror 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 and Global Mirror Change Volumes**
  Referred to as *asynchronous replication*, Global Mirror 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 operation is performed. An added extension to Global Mirror is Global Mirror with Change Volumes. Global Mirror with Change Volumes is the preferred method for use with native IP replication.

### 11.6.2 IP partnership limitations

The following prerequisites and assumptions must be considered before IP partnership between two Storwize V7000 systems can be established:

- Storwize V7000 systems are successfully installed with the Storwize V7000 7.2.0 or later code levels.
- Storwize V7000 systems have the necessary licenses that allow remote copy partnerships to be configured between two systems. No separate license is required to enable IP partnership.
- The storage SAN configurations are properly done and the infrastructure to support Storwize V7000 systems in remote copy partnerships over IP links is properly in place.
- The two systems should be able to ping each other and perform the discovery.
- The maximum partnerships between the local and remote systems including both IP and FC partnerships is limited to the current maximum supported; that 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. FC zones between two systems must be removed before an IP partnership is configured.
- IP partnerships are supported on both 10 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 Fibre Channel-based partnerships to IP partnerships are not supported.
- IP partnerships between the two systems can be over IPv4 or IPv6 only and not both.
- VLAN tagging of the IP addresses configured for remote copy is supported starting with Storwize code version 7.4.
- Management IP and iSCSI IP on the same port can be in a different network starting with Storwize code 7.4.
An added layer of security is provided by using Challenge Handshake Authentication Protocol (CHAP) authentication.

TCP ports 3260 and 3265 are used for IP partnership communications; therefore, these ports must be open in firewalls between the systems.

The following maximum throughput is restricted based on use of 1 Gbps or 10 Gbps ports:

- One 1 Gbps port might transfer up to 110 Mbps
- Two 1 Gbps ports might transfer up to 220 Mbps
- One 10 Gbps port might transfer up to 190 Mbps
- One 10 Gbps ports might transfer up to 280 Mbps

Note: The Bandwidth setting definition when the IP partnerships are created changed. Previously, the bandwidth setting defaults to 50 MBs and it was the maximum transfer rate from primary to secondary site for initial sync/resyncs of volumes.

The Link Bandwidth setting is now configured by using Mbits not MBs and you set this to a value that the communication link can sustain or what is allocated for replication. The Background Copy Rate setting is now a percentage of the Link Bandwidth and it determines the bandwidth available for initial sync and resyncs or for Global Mirror with Change Volumes.

11.6.3 VLAN support

Starting with version 7.4 of Storwize code VLAN tagging is supported for both Internet Small Computer System Interface (iSCSI) host attachment and IP replication. Hosts and remote-copy operations can connect to the system through Ethernet ports. Each of these traffic types has different bandwidth requirements, which can interfere with each other if they share 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 VLAN ID is configured for the IP addresses used for either iSCSI host attach or IP replication on Storwize, appropriate VLAN settings on the Ethernet network and servers must also be properly configured in order not to experience connectivity issues. Once VLANs are configured, changes to VLAN settings disrupt iSCSI and IP replication traffic to and from Storwize.

During VLAN configuration for each IP address individually, the user must be aware that if VLAN settings for the local and failover ports on two nodes of an iogroup are different, then switches must be configured such that failover VLANs are configured on the local switch ports as well such that failover of IP addresses from failing node to surviving node succeeds. In cases where this is not done, the user experiences loss of paths to Storwize/Storwize storage during a node failure.

11.6.4 IP partnership and Storwize V7000 terminology

The IP partnership terminology and abbreviations that are used are listed in Table 11-7 on page 414.
**Table 11-7 Terminology**

<table>
<thead>
<tr>
<th>IP partnership and Storwize V7000 terminology</th>
<th>Description</th>
</tr>
</thead>
</table>
| Remote copy group or Remote copy port group | 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:  
  - 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  
  Each IP address can be shared for iSCSI host attach and remote copy functionality. Therefore, appropriate settings must be applied to each IP address. |
| IP partnership | Two Storwize V7000 systems that are partnered to perform remote copy over native IP links. |
| FC partnership | Two Storwize V7000 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. |
| linkbandwidthm| 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. |

**11.6.5 States of IP partnership**

The different partnership states in IP partnership are listed in Table 11-8.

**Table 11-8 States of IP partnership**

<table>
<thead>
<tr>
<th>State</th>
<th>Systems connected</th>
<th>Support for active remote copy I/O</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partially_Configured_Local</td>
<td>No</td>
<td>No</td>
<td>This state indicates that the initial discovery is complete.</td>
</tr>
<tr>
<td>Fully_Configured</td>
<td>Yes</td>
<td>Yes</td>
<td>Discovery successfully completed between two systems and the two systems can establish remote copy relationships.</td>
</tr>
<tr>
<td>Fully_ConfiguredStopped</td>
<td>Yes</td>
<td>Yes</td>
<td>The partnership is stopped on the system.</td>
</tr>
<tr>
<td>Fully_Configured_Remote_Stopped</td>
<td>Yes</td>
<td>No</td>
<td>The partnership is stopped on the remote system.</td>
</tr>
<tr>
<td>Not_Present</td>
<td>Yes</td>
<td>No</td>
<td>The two systems cannot communicate with each other. This state is also seen when data paths between the two systems are not established.</td>
</tr>
</tbody>
</table>
The following steps must be completed to establish two systems in 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. If required, the administrator configures the system IP addresses on both local and remote systems so that they can discover each other over the network.

3. In case you want to use VLANs, configure your LAN switches and Storwize Ethernet ports to use VLAN tagging.

4. The Administrator configures the Storwize V7000 ports on each node in both the systems by using the `Storwizetask cfgportip` 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 are allowed.

5. The administrator establishes partnership with the remote system from the local system where the partnership state then transitions to the `Partially_Configured_Local` state.

6. The administrator establishes partnership from remote system with local system and if successful, partnership state then transitions to the `Fully_Configured` state, which implies that the partnerships over IP network were successfully established. The partnership state momentarily remains in the `not_present` state before transitioning to `fully_configured`.

7. The administrator creates Metro Mirror, Global Mirror, and Global Mirror with Change Volume relationships.

**Partnership consideration**: When the partnership is created, there is no master or auxiliary status that is defined or implied. The partnership is equal and the concepts of master or auxiliary and primary or secondary apply to volume relationships only, not to system partnerships.

### 11.6.6 Remote copy groups

This section describes remote copy groups (or remote copy port groups) and the different ways in which the links between the two remote systems can be configured. The two Storwize V7000 systems can be connected to each other over one link or at most two links. To address the requirement to allow Storwize V7000 to be aware of the physical links between two sites, the concept of remote copy port groups was introduced.

### State | Systems connected | Support for active remote copy I/O | Comments
---|---|---|---
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.

<table>
<thead>
<tr>
<th>State</th>
<th>Systems connected</th>
<th>Support for active remote copy I/O</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fully_Configured_Exceeded</td>
<td>Yes</td>
<td>No</td>
<td>There are too many systems in the network, and the partnership from the local system to remote system is disabled.</td>
</tr>
<tr>
<td>Fully_Configured_Excluded</td>
<td>No</td>
<td>No</td>
<td>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.</td>
</tr>
</tbody>
</table>
Storwize V7000 IP addresses that are connected to the same physical link are designated with identical remote copy port groups. Storwize V7000 supports three remote copy groups: 0, 1, and 2. Storwize V7000 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 there is a single inter-site link or 1 and 2 if there are two inter-site links.

You can assign each one IPv4 address and one IPv6 address to each Ethernet port on Storwize V7000 platforms. Each of these IP addresses can be shared between iSCSI host attach and 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 IP partnership, both must have IPv6 addresses 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 Storwize V7000 node must have only a single remote copy port group configured; that is, 1 or 2. Remaining IP addresses must be in remote copy port group 0.

### 11.6.7 Supported configurations

The following supported configurations for IP partnership that were in the first release are described in this section:

- Two 2-node systems in IP partnership over a single inter-site link, as shown in Figure 11-15.

As shown in Figure 11-15, there are two systems: System A and System B. A single remote copy port group 1 is created on Node A1 on System A and on Node B2 on System
B (administrator might choose to configure the remote copy port group on Node B1 on System B instead of Node B2) as there is only a single inter-site link to facilitate the IP partnership traffic. At any time, only the IP addresses that are configured in remote copy port group 1 on the nodes in System A and System B participate in establishing data paths between the two systems after the IP partnerships are created. In this configuration, there are no failover ports that are configured on the partner node in the same I/O group.

This configuration has the following characteristics:

- Only one node in each system has remote copy port group that is configured and there are no failover ports configured.
- If the Node A1 in System A or the Node B2 in System B were to encounter some failure, the IP partnership stops and enters the not_present state until the failed nodes recover.
- After the nodes recover, the IP ports fail back, the IP partnership recovers, and the partnership state goes to the fully_configured state.
- If the inter-site system link fails, the IP partnerships transition to the not_present state.
- This configuration is not recommended because it is not resilient to node failures.

Two 2-node systems in IP partnership over a single inter-site link (with failover ports configured), as shown in Figure 11-16.

![Figure 11-16](image)

As shown in Figure 11-16, there are two systems: System A and System B. A single remote copy port group 1 is configured on two Ethernet ports, one each, on Node A1 and Node A2 on System A and similarly, on Node B1 and Node B2 on System B. Although there are two ports on each system that are configured for remote copy port group 1, only one Ethernet port in each system actively participates in the IP partnership process. This selection is determined by a path configuration algorithm that is designed to choose data paths between the two systems to optimize performance.

The other port on the partner node in the I/O Group behaves as a standby port that is used if there is a node failure. If Node A1 fails in System A, as failover port is configured on Node A2 on Ethernet Port 2, IP partnership continues servicing replication I/O from Ethernet Port 2. However, it might take some time for discovery and path configuration.
logic to re-establish paths post failover and this can cause partnerships to transition to `not_present` for that time. The details of the particular IP port that is actively participating in IP partnership is provided in the `Storwizeinfo lsportip` output (reported as `used`).

This configuration has the following characteristics:

- Each node in the I/O group has the same remote copy port group that is configured. However, only one port in that remote copy port group is active at any time at each system.

- If the Node A1 in System A or the Node B2 in System B fails in the respective systems, IP partnerships rediscovery is triggered and continues servicing the I/O from the failover port.

- The discovery mechanism that is triggered because of failover might introduce a delay where-in the partnerships momentarily transition to the `not_present` state and then recover.

- Two 4-node systems in IP partnership over a single inter-site link (with failover ports configured), as shown in Figure 11-17.

As shown in Figure 11-17, there are two 4-node systems: System A and System B. A single remote copy port group 1 is configured on nodes A1, A2, A3, and A4 on System A, Site A; and on nodes B1, B2, B3, and B4 on System B, Site B. Although there are four ports that are configured for remote copy group 1, only one Ethernet port in each remote copy port group on each system actively participates in the IP partnership process. Port selection is determined by a path configuration algorithm. The other ports play the role of standby ports.
If Node A1 fails in System A, the IP partnership selects one of the remaining ports that is configured with remote copy port group 1 from any of the nodes from either of the two I/O groups in System A. However, it might take some time (generally tens of seconds) for discovery and path configuration logic to re-establish paths post failover and this process can cause partnerships to transition to the `not_present` state. This result leads to remote copy relationships to stop and the administrator might need to manually verify the issues in the event log and start the relationships or remote copy consistency groups, if they do not autorecover. The details of the particular IP port actively participating in the IP partnership process is provided in the `Storwizeinfo lsportip` view (reported as `used`).

This configuration has the following characteristics:

- Each node has the remote copy port group that is configured in both I/O groups. However, only one port in that remote copy port group remains active and participates in IP partnership on each system.
- If the Node A1 in System A or the Node B2 in System B were to encounter some failure in the system, IP partnerships discovery is triggered and it continues servicing the I/O from the failover port.
- The discovery mechanism that is triggered because of failover might introduce a delay wherein the partnerships momentarily transition to the `not_present` state and then recover.
- The bandwidth of the single link is used completely.

- Eight-node system in IP partnership with four-node system over single inter-site link, as shown in Figure 11-18 on page 420.
As shown in Figure 11-18, there is an eight-node system (System A in Site A) and a four-node system (System B in Site B). A single remote copy port group 1 is configured on nodes A1, A2, A5, and A6 on System A at Site A and similarly, a single remote copy port group 1 is configured on nodes B1, B2, B3, and B4 on System B.

Although there are four I/O groups (eight nodes) in System A, any two I/O groups at maximum are supported to be configured for IP partnerships. If Node A1 fails in System A, IP partnership continues using one of the ports that is configured in remote copy port group from any of the nodes from either of the two I/O groups in System A. However, it might take some time for discovery and path configuration logic to re-establish paths post-failover and this delay might cause partnerships to transition to the not_present state. This process can lead to remote copy relationships to stop and the administrator must manually start them if the relationships do not auto-recover. The details of which particular IP port is actively participating in IP partnership process is provided in Storwizeinfo lsportip output (reported as used).
This configuration has the following characteristics:

- Each node has the remote copy port group that is configured in both the I/O groups that are identified for participating in IP Replication. However, only one port in that remote copy port group remains active on each system and participates in IP Replication.

- If the Node A1 in System A or the Node B2 in System B fails in the system, the IP partnerships trigger discovery and continue servicing the I/O from the failover ports.

- The discovery mechanism that is triggered because of failover might introduce a delay wherein the partnerships momentarily transition to the `not_present` state and then recover.

- The bandwidth of the single link is used completely.

Two 2-node systems with two inter-site links, as shown in Figure 11-19.

As shown in Figure 11-19, remote copy port groups 1 and 2 are configured on the nodes in System A and System B as there are two inter-site links available. In this configuration, the failover ports are not configured on partner nodes in the I/O group. Instead, the ports are maintained in different remote copy port groups on both of the nodes and they remain active and participate in IP partnership by using both of the links.

However, if either of the nodes in the I/O group fail (that is, if Node A1 on System A fails), the IP partnership continues only from the available IP port that is configured in remote copy port group 2. Therefore, the effective bandwidth of the two links is reduced to 50%; that is, only the bandwidth of a single link is available only until the failure is resolved.

This configuration has the following characteristics:

- There are two inter-site links and two remote copy port groups are configured.

- Each node has only one IP port in remote copy port group 1 or 2.

- Both the IP ports in the two remote copy port groups participate simultaneously in IP partnerships. Therefore, both of the links are used.

- During node failure or link failure, the IP partnership traffic continues from the other available link and the port group. Hence, if two links of 10 Mbps each are available and...
you have 20 Mbps of effective link bandwidth, bandwidth is reduced to 10 Mbps only during a failure.

- After the node failure or link failure is resolved and failback happens, the entire bandwidth of both of the links is available as before.

- Two 4-node systems in IP partnership with dual inter-site links, as shown in Figure 11-20.

As shown in Figure 11-20, there are two 4-node systems: System A and System B. This configuration is an extension of Configuration 5 to a multinode multi-I/O group environment. As seen in this configuration, there are two I/O groups and each node in the I/O group has a single port that is configured in remote copy port groups 1 or 2. Although there are two ports that are configured in remote copy port groups 1 and 2 on each system, only one IP port in each remote copy port group on each system actively participates in IP partnership. The other ports that are configured in the same remote copy port group act as standby ports in event of failure. Which port in a configured remote copy port group participates in IP partnership at any moment is determined by a path configuration algorithm.

In this configuration, if Node A1 fails in System A, IP partnership traffic continues from Node A2 (that is, remote copy port group 2) and at the same time the failover also causes discovery in remote copy port group 1. Therefore, the IP partnership traffic continues from...
Node A3 on which remote copy port group 1 is configured. The details of the particular IP port that is actively participating in IP partnership process is provided in the Storwizeinfo lsportip output (reported as used).

This configuration has the following characteristics:

– Each node has the remote copy port group that is configured in the I/O groups 1 or 2. However, only one port per system in both remote copy port groups remains active and participates in IP partnership.

– Only a single port per system from each configured remote copy port group participates simultaneously in IP partnership. Therefore, both of the links are used.

– During node failure or port failure of a node that is actively participating in IP partnership, IP partnership continues from the alternative port because another port is in the system in the same remote copy port group but in a different I/O Group.

– The pathing algorithm can start discovery of available port in the affected remote copy port group in the second I/O group and pathing is re-established, which restores the total bandwidth; that is, both of the links are available to support IP partnership.

▶ Eight-node system in IP partnership with a four-node system over dual inter-site links, as shown in Figure 11-21 on page 424.
As shown Figure 11-21, there is an eight-node System A in Site A and a four-node System B in Site B. Because a maximum of two I/O groups in IP partnership is supported in a system, although there are four I/O groups (eight nodes), nodes from only two I/O groups’ are configured with remote copy port groups in System A. The remaining or all of the I/O groups can be configured to be remote copy partnerships over Fibre Channel. In this configuration, there are two links and two I/O groups that are configured with remote copy port groups 1 and 2, but path selection logic is managed by an internal algorithm. Therefore, this configuration depends on the pathing algorithm to decide which of the nodes actively participates in IP partnership. Even if Node A5 and Node A6 are configured with remote copy port groups properly, active IP partnership traffic on both of the links might be driven from Node A1 and Node A2 only.
If Node A1 fails in System A, IP partnership traffic continues from Node A2 (that is, remote copy port group 2) and the failover also causes IP partnership traffic to continue from Node A5 on which remote copy port group 1 is configured. The details of the particular IP port actively participating in IP partnership process is provided in the Storwizeinfo lsportip output (reported as used).

This configuration has the following characteristics:

- There are two I/O Groups with nodes in those I/O groups that are configured in two remote copy port groups as there are two inter-site links for participating in IP partnership. However, only one port per system in a particular remote copy port group remains active and participates in IP partnership.
- One port per system from each remote copy port group participates in IP partnership simultaneously. Therefore, both of the links are used.
- If a node or port on the node that is actively participating in IP partnership fails, RC data path is established from that port because another port is available on an alternative node in the system with the same remote copy port group.
- The path selection algorithm starts discovery of available port in the affected remote copy port group in the alternative I/O groups and paths are re-established, which restores the total bandwidth across both links.
- The remaining or all of the I/O groups can be in remote copy partnerships with other systems.

▶ An example of unsupported configuration for single inter-site link, as shown in Figure 11-22.

As shown in Figure 11-22, this configuration is similar to Configuration 2, but differs because each node now has the same remote copy port group that is configured on more than one IP port.

On any node, only one port at any time can participate in IP partnership. Configuring multiple ports in the same remote copy group on the same node is not supported.
An example of an unsupported configuration for dual inter-site link, as shown in Figure 11-23.

As shown in Figure 11-23, this configuration is similar to Configuration 5, but differs because each node now also has two ports that are configured with remote copy port groups. In this configuration, the path selection algorithm can select a path in a manner such that at times this might cause partnerships to transition to the not_present state and then recover.

This result is a configuration restriction and the use of this configuration is not recommended until the configuration restriction is lifted in future releases.

Example deployment for the configuration presented on page 417 with dedicated inter-site link, as shown in Figure 11-24.

In this configuration, one port on each node in System A and System B is configured in remote copy group 1 to establish IP partnership and support remote copy relationships.
There is a dedicated inter-site link that is used for IP partnership traffic and iSCSI host attach is disabled on those ports.

The following configuration steps are used:

a. Configure system IP addresses properly; as such they can be reached over the inter-site link.

b. Qualify if the partnerships must be created over IPv4 or IPv6 and then assign IP addresses and open firewall ports 3260 and 3265.

c. Configure IP ports for remote copy on both the systems by using the following settings:
   - Remote copy group: 1
   - Host: No
   - Assign IP address

d. Check that the maximum transmission unit (MTU) levels across the network meet the requirements as set (default MTU is 1500 on Storwize V7000).

e. Establish IP partnerships from both of the systems.

f. After the partnerships are in the fully_configured state, you can create the remote copy relationships.

Example deployment for Configuration on page 421. Ports that are shared with host access are shown in Figure 11-25.

In this configuration, IP ports are to be shared by both iSCSI hosts and for IP partnership.

The following configuration steps are used:

a. Configure System IP addresses properly so that they can be reached over the inter-site link.

b. Qualify if the partnerships must be created over IPv4 or IPv6 and then assign IP addresses and open firewall ports 3260 and 3265.

c. Configure IP ports for remote copy on System A1 by using the following settings:
   - Node 1:
     - Port 1, remote copy port group 1
     - Host: Yes
- Assign IP address

- Node 2:
  - Port 4, Remote Copy Port Group 2
  - Host: Yes
  - Assign IP address

- Configure IP ports for remote copy on System B1 by using the following settings:

- Node 1:
  - Port 1, remote copy port group 1
  - Host: Yes
  - Assign IP address

- Node 2:
  - Port 4, remote copy port group 2
  - Host: Yes
  - Assign IP address

- Check the MTU levels across the network and meet the requirements as set (default MTU is 1500 on Storwize V7000).

- Establish IP partnerships from both systems.

- After the partnerships are in the fully_configured state, you can create the remote copy relationships.

### 11.7 Remote Copy

In this section, we describe the Remote Copy services, which are a synchronous remote copy called Metro Mirror (MM), asynchronous remote copy called Global Mirror (GM), and Global Copy with Changed Volumes. Remote Copy in the Storwize V7000 is similar to Remote Copy in the IBM System Storage DS8000® family at a functional level, but the implementation differs.

The Storwize V7000 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 Storwize V7000.

The general application of remote copy services is to maintain two real-time synchronized copies of a disk. Often, two copies are geographically dispersed between two Storwize V7000 systems, although it is possible to use Metro Mirror or Global Mirror 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. Licensing must also be doubled because source and target are within the same system.

Use intercluster Metro Mirror or Global Mirror when possible. For mirroring volumes in the same I/O group, it is better to use Volume Mirroring or FlashCopy feature.
A typical application of this function is to set up a dual-site solution that uses two Storwize V7000 systems. The first site is considered the primary 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.

### 11.7.1 Multiple Storwize V7000 System Mirroring

Each Storwize V7000 system can maintain up to three partner system relationships, which allows as many as four systems to be directly associated with each other. This Storwize V7000 partnership capability enables the implementation of disaster recovery (DR) solutions.

**Note:** For more information about restrictions and limitations of native IP replication, see 11.6.2, “IP partnership limitations” on page 412.

Figure 11-26 shows an example of a Multiple System Mirroring configuration.

![Multiple System Mirroring configuration example](image-url)
Multiple System Mirroring allows for various partnership topologies, as shown in the examples in Figure 11-27.

Example: A → B, A → C, and A → D

Figure 11-27 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 process that is described in the following example:

1. Suspend application at A.
2. Remove the A → B relationship.
3. Create the A → C relationship (or the B → C relationship).
4. Synchronize to system C, and ensure that $A \to C$ is established:
   - $A \to B, A \to C, A \to D, B \to C, B \to D, \text{ and } C \to D$
   - $A \to B, A \to C, \text{ and } B \to C$

Figure 11-28 shows an example of a triangle topology.

Example: $A \to B, A \to C, \text{ and } B \to C$

Figure 11-28    Storwize V7000 triangle topology

Figure 11-29 shows an example of a Storwize V7000 fully connected topology.

Example: $A \to B, A \to C, A \to D, B \to D, \text{ and } C \to D$

Figure 11-29    Storwize V7000 fully connected topology

Figure 11-29 is a fully connected mesh in which every system has a partnership to each of the three other systems. This topology allows volumes to be replicated between any pair of systems.

Example: $A \to B, A \to C, \text{ and } B \to C$

Figure 11-30 shows a daisy-chain topology.

Figure 11-30    Storwize V7000 daisy-chain topology
Although systems can have up to three partnerships, volumes can be part of only one remote copy relationship; for example, A → B.

**System partnership intermix:** All of the preceding topologies are valid for the intermix of the IBM SAN Volume Controller with the Storwize V7000 if the Storwize V7000 is set to the replication layer and running 6.3.0 code or later.

### 11.7.2 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 following 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 Storwize V7000 Metro Mirror and Global Mirror implementation operates in a manner that is designed to always keep a consistent image at the secondary site. The Storwize V7000 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, your Global Mirror target data is at least crash consistent and allow for quick recovery via your application crash recovery facilities.

For more information about dependent writes, see 11.4.3, “Consistency Groups” on page 390.

**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 Storwize V7000 system. MM/GM commands can be issued to a Remote Copy Consistency Group and therefore, 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 11-31 on page 433 shows the concept of Metro Mirror Consistency Groups. The same applies to Global Mirror Consistency Groups.
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 undesired 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 there is an error, there is a loss of synchronization and a background copy process is required to recover synchronization. While this process is in
progress, 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.

11.7.3 Remote copy intercluster communication

In the traditional Fibre Channel (FC), the intercluster communication between systems in a Metro Mirror and Global Mirror partnership is performed over the SAN. In the following section, we describe this communication path.

For more information about intercluster communication between systems in an IP partnership, see 11.6.5, “States of IP partnership” on page 414.

Zoning
The Storwize V7000 node ports on each Storwize V7000 system must communicate with each other for the partnership creation to be performed. Switch zoning is critical to facilitating intercluster communication.

Intercluster communication channels
When a Storwize V7000 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 Storwize V7000 systems is interrupted or lost, an event is logged (and the Metro Mirror and Global Mirror relationships stops).

Alerts: You can configure the Storwize V7000 to raise Simple Network Management Protocol (SNMP) traps to the enterprise monitoring system to alert on events that indicate that an interruption in internode communication occurred.

Intercluster links
All Storwize V7000 nodes 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 Storwize V7000 nodes are categorized according to the Storwize V7000 system to which they belong. Storwize V7000 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 Storwize V7000.

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: You can use chsystem with -partnerfcportmask to dedicate several Storwize FC ports only to system-to-system traffic to make sure that remote copy is not affected by other traffic such as host-to-node or node-to-node within the same system.

11.7.4 Metro Mirror overview

Metro Mirror establishes a synchronous relationship between two volumes of equal size. The volumes in a Metro Mirror relationship are referred to as the master (primary) volume and the auxiliary (secondary) volume. 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 the volumes have identical data when the copy completes. After the initial copy completes, the Metro Mirror function maintains a fully synchronized copy of the source data at the target site always.

Metro Mirror has the following characteristics:

- Zero recovery point objective (RPO)
- Synchronous
- Production application performance that is impacted 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 and Global Mirror Consistency Groups (FlashCopy Consistency Groups and Global Mirror Consistency are described in 11.4, “Implementing Storwize V7000 FlashCopy” on page 388).

The Storwize V7000 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 Storwize V7000 systems must be defined in a Storwize V7000 partnership, which must be performed on both Storwize V7000 systems to establish a fully functional Metro Mirror partnership.

By using standard single-mode connections, the supported distance between two Storwize V7000 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.

11.7.5 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 allows I/O to the master volume to continue to avoid affecting the operation of the master volumes.

Figure 11-32 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 11-32  Write on volume in Metro Mirror relationship](image)

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 Fibre Channel Metro Mirror has distance limitations that are based on your performance requirements. The Storwize V7000 does not support more than 300 km (186.4 miles).

### 11.7.6 Metro Mirror features

Storwize V7000 Metro Mirror supports the following features:

- Synchronous remote copy of volumes that are dispersed over metropolitan distances.
- The Storwize V7000 implements Metro Mirror relationships between volume pairs, with each volume in a pair that is managed by a Storwize V7000 system or IBM SAN Volume Controller system (requires code version or later 6.3.0).
- The Storwize V7000 supports intracluster Metro Mirror where both volumes belong to the same system (and I/O Group).
- The Storwize V7000 supports intercluster Metro Mirror where each volume belongs to a separate Storwize V7000 system. You can configure a specific Storwize V7000 system for partnership with another system. All intercluster Metro Mirror processing occurs between two Storwize V7000 systems that are configured in a partnership.
- Intercluster and intracluster Metro Mirror can be used concurrently.
- The Storwize V7000 does not require that a control network or fabric is installed to manage Metro Mirror. For intercluster Metro Mirror, the Storwize V7000 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 Storwize V7000 uses for Metro Mirror I/O.
- The Storwize V7000 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 Storwize V7000 allows 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.

### 11.7.7 Metro Mirror attributes

The Metro Mirror function in Storwize V7000 possesses the following attributes:

- A Storwize V7000 system partnership is created between two Storwize V7000 systems or a Storwize V7000 System and Storwize V7000 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.

11.7.8 Global Mirror

In the following topics, we describe the Global Mirror copy service, which is an asynchronous remote copy service. It 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 of equal size. 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 completing on the auxiliary volume.

Global Mirror has the following characteristics:

- Near-zero RPO
- Asynchronous
- Production application performance that is impacted 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 overhead. However, leaving this functionality in place simplifies testing and allows for 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 Storwize V7000 System.

**Intercluster Global Mirror**

Intercluster Global Mirror operations require a pair of Storwize V7000 systems that are connected by a number of intercluster links. The two Storwize V7000 systems must be defined in a Storwize V7000 system 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.

11.7.9 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 11-33 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.

![Global Mirror write sequence](image)

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, it 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.

In Storwize V7000 code 7.2, these algorithms were enhanced to optimize Global Mirror behavior and latency even further. Global Mirror write I/O from production to a secondary Storwize V7000 system requires serialization and sequence-tagging before being sent across network to remote site (to maintain a write-order consistent copy of data). Sequence-tagged Global Mirror writes on secondary system are processed without parallelism and management of write I/O sequencing imposes more latency on write I/Os in code versions before 7.2. As a result, high-bandwidth Global Mirror throughput environments can experience performance impacts on primary system during high I/O peak periods.

Starting with code V7.2, Storwize allows more parallelism in processing and managing Global Mirror writes on secondary system by using the following methods:

- Nodes on the secondary system 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

**Note:** The V7.2 enhancements of Global Mirror require no changes in administration and management. However, before you upgrade to Storwize V7000 V7.2, you must stop all Global Mirror relationships. The proper checks are provided in the latest Storwizeupgradetest utility.

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, FCIP, FCOE, and native IP connections. The maximum distance cannot exceed 80 ms round-trip, which is about 4000 km (2485 mi) between mirrored systems. But, starting with Storwize code V7.4, this distance was significantly increased for certain Storwize V7000 Gen2 and SAN Volume Controller configurations. Figure 11-34 shows the current supported distances for Global Mirror remote copy.

![Figure 11-34 Supported Global Mirror distances](image)

### 11.7.10 Storwize V7000 Global Mirror features

Storwize V7000 Global Mirror supports the following features:

- Asynchronous remote copy of volumes that are dispersed over metropolitan-scale distances.
- The Storwize V7000 implements the Global Mirror relationship between a volume pair, with each volume in the pair being managed by a SAN Volume Controller system or Storwize V7000 system. Storwize must be in replication layer running at least code v6.3.
- The Storwize V7000 supports intracluster Global Mirror where both volumes belong to the same system (and I/O Group).
- The Storwize V7000 intercluster Global Mirror in which each volume belongs to its separate Storwize V7000 system. A Storwize V7000 system can be configured for partnership with between one and three other systems. For more information about IP partnership restrictions, see 11.6.2, “IP partnership limitations” on page 412.
- Intercluster and intracluster Global Mirror can be used concurrently but not for the same volume.
- The Storwize V7000 does not require a control network or fabric to be installed to manage Global Mirror. For intercluster Global Mirror, the Storwize V7000 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 Storwize V7000 uses for Global Mirror I/O.

- The Storwize V7000 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 Storwize V7000 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 permits a delay simulation to be applied on writes that are sent to auxiliary volumes. It is useful in intracluster scenarios for testing purposes.
- As of Storwize V7000 6.3.0 and above, Global Mirror source and target volumes can be associated with Change Volumes.

**Colliding writes**

Before V4.3.1, the Global Mirror algorithm required that only a single write is active on any 512-byte LBA of a volume. If a further write is received from a host while the auxiliary write is still active (even though the master write might complete), the new host write is delayed until the auxiliary write is complete. This restriction is needed if a series of writes to the auxiliary must be tried again (which 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 serialized, 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 11-35 on page 442 shows a colliding write sequence example.
The following numbers correspond to the numbers that are shown in Figure 11-35:

- (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 permits a delay simulation to be applied on writes that are sent to auxiliary volumes. This feature allows 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.

### 11.7.11 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, or overloaded hosts, you might be affected by multiple congestion errors.

Congestion errors happen in the following primary situations:

- Congestion at the source site via the host or network
- Congestion in the network link or network path
- Congestion at the target site via the host or network

Global Mirror has functionality that is designed to address the following conditions, which could possibly negatively affect certain Global Mirror implementations:

- The estimation of the bandwidth requirements tends to be complex.
- It is often difficult to ensure that the latency and bandwidth requirements can be met.
- 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 special purpose only. Change Volumes replicates point-in-time images on a cycling period. The default 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, instead of 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 overhead because point-in-time copies are created locally

Figure 11-36 shows a simple Global Mirror relationship without Change Volumes.

Figure 11-36  Global Mirror without Change Volumes

With Change Volumes, this environment looks as it is shown in Figure 11-37 on page 444.
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.

How Change Volumes might save you replication traffic is shown in Figure 11-38.

In Figure 11-38, you can see a number of 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 11-39.
In Figure 11-39 on page 444, 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, your 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.

### 11.7.12 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.

In case this best practice is not maintained, for example, source volumes are assigned to only one node in I/O group, you can change the preferred node for each volume to distribute volumes evenly between nodes. Starting with firmware v7.3, the preferred node can be changed without changing the I/O group. This means it will not affect host I/O to particular volume. Additionally, now you can also change the preferred Storwize node for volumes that are in a remote copy relationship. The remote copy relationship type does not matter and can be any of Metro Mirror, Global Mirror, or Global Mirror with Change Volumes. Changing the preferred node can be done both to source and target volumes participating in remote copy relationship.

### 11.7.13 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 Storwize V7000 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 a while. 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 of the 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.

Default value of background copy is 25 MBps, per volume.

**Important:** Background copy value is a system-wide parameter and can be changed dynamically but only on system basis and not per relationship basis. This means all relationships will have their copy rate changed when this value is increased or decreased. In systems with many remote copy relationships, increasing this value may impact overall system or intercluster link performance. Background copy rate can be changed 1 - 1000 MBps.

### 11.7.14 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.

### 11.7.15 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 `startcrelationship` 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 `mkrelrelationship` with the `-sync` flag.
- Run `startrelrelationship` 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, therefore, creating a data loss or data integrity exposure for hosts accessing data on the auxiliary volume.

### 11.7.16 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 destroyed.

**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, much like FlashCopy's restore option. But while FlashCopy target volume can operate in read/write mode, the target volume of 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 Storwize V7000 allows read-only access to the auxiliary volume when it contains a consistent image. Storwize allows boot time operating system discovery to complete without 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 logical block address (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, there is no way in which coherency can 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 and that 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 Storwize V7000 provides SNMP traps and programming (or scripting) for the CLI to enable this automation.

11.7.17 Practical use of Global Mirror

Practical use of Global Mirror is similar to the Metro Mirror described in 11.7.16, “Practical use of Metro Mirror” on page 447. The main difference between those two remote copy modes is Global Mirror or Global Mirror with Change Volumes are mostly used on much larger distances than Metro Mirror. The weak link quality or not enough bandwidth between primary and secondary site can also be a reason to prefer asynchronous Global Mirror over synchronous Metro Mirror. Besides that, use cases for Metro Mirror and Global Mirror are the same.

11.7.18 Valid combinations of FlashCopy, Metro Mirror, and Global Mirror

Table 11-9 lists the combinations of FlashCopy and Metro Mirror or Global Mirror functions that are valid for a single volume.

<table>
<thead>
<tr>
<th>FlashCopy</th>
<th>Metro Mirror or Global Mirror source</th>
<th>Metro Mirror or Global Mirror target</th>
</tr>
</thead>
<tbody>
<tr>
<td>FlashCopy Source</td>
<td>Supported</td>
<td>Supported</td>
</tr>
<tr>
<td>FlashCopy Target</td>
<td>Supported</td>
<td>Not supported</td>
</tr>
</tbody>
</table>

11.7.19 Remote Copy configuration limits

Table 11-10 lists the Metro Mirror and Global Mirror configuration limits.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Metro Mirror or Global Mirror Consistency Groups per system</td>
<td>256</td>
</tr>
<tr>
<td>Number of Metro Mirror or Global Mirror relationships per system</td>
<td>8192</td>
</tr>
<tr>
<td>Number of Metro Mirror or Global Mirror relationships per Consistency Group</td>
<td>8192</td>
</tr>
<tr>
<td>Total volume size per I/O Group</td>
<td>There is a per I/O Group limit of 1024 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 MB of bitmap space for the I/O Group and allows 10 MB of space for all remaining copy services features.</td>
</tr>
</tbody>
</table>
11.7.20 Remote Copy states and events

In this section, we describe the various states of a MM/GM relationship and the conditions that cause them to change.

In Figure 11-40, the MM/GM relationship state diagram shows an overview of the states that can apply to a MM/GM relationship in a connected state.

![Figure 11-40 Metro Mirror or Global Mirror mapping state diagram](image)

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 especially useful when MM/GM relationships are created for volumes that were created with the format option.

The following step identifiers are shown in Figure 11-40:

- **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:
- When a MM/GM relationship is stopped in the ConsistentSynchronized state, the MM/GM relationship enters the Idling state when you specify the \texttt{-access} option, which enables write I/O on the auxiliary volume.
- When a MM/GM relationship is stopped in the ConsistentSynchronized state without \texttt{-access} parameter, the auxiliary volumes remain read-only and the state of relationship changes to ConsistentStopped.
- To enable write I/O on the auxiliary volume, when the MM/GM relationship is in the ConsistentStopped state, issue the command \texttt{Storwizetask stoprcrelationship}, which specifies the \texttt{-access} option, and the MM/GM relationship enters the Idling state.

Step 5:
- When a MM/GM relationship is started from the Idling state, you must specify the \texttt{-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.
- If write I/O was performed to the master or auxiliary volume, the \texttt{-force} option must be specified and the MM/GM relationship then enters the InconsistentCopying state while the background copy is started. The background copy copies only the data that changed on primary volume while the relationship was stopped.

Stop or Error
When a MM/GM relationship is stopped (intentionally or because an error), a state transition is applied.

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 Storwize V7000 systems that are in a partnership, all (intercluster) MM/GM relationships enter a Disconnected state. For more information, see “Connected versus disconnected” on page 450.

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 portrayed as having a changed state. The new states describe what is known about the relationship and what configuration commands are permitted.

When the systems can communicate again, the relationships are reconnected. MM/GM automatically reconciles the two state fragments, taking into account any configuration or other event that took place while the relationship was disconnected. As a result, the relationship can return to the state that it was in when it became disconnected or enter a new state.

Relationships that are configured between volumes in the same Storwize V7000 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).

For writes for which the host did not receive a successful completion (that is, it received bad completion or no completion at all), and 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 that 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 cope 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.

For more information about dependent writes, see 11.4.3, “Consistency Groups” on page 390.

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 application 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 a number of disks that are accessed through multiple systems; therefore, consistency must operate across all 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 and therefore, losing consistency.

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 freeze 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 regarding the master.

This state can arise when a relationship was in a ConsistentSynchronized state and suffers 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 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 it 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 already 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.

### InconsistentDisconnected

InconsistentDisconnected is a disconnected state. The target volumes in this half of the relationship or Consistency Group are all in the auxiliary role and do not accept read or write I/O.

Except for deletes, no configuration activity is permitted until the relationship becomes connected again.

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 become 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.

### 11.8 Remote Copy commands

In this section, we present commands that need to be issued to create and operate Remote Copy services.

#### 11.8.1 Remote Copy process

The MM/GM process includes the following steps:

1. A Storwize V7000 system partnership is created between two Storwize V7000 systems (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.

The remote host server is mapped to the auxiliary volume and the disk is available for I/O.


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

Where 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 that the command that sets systems into a MM/GM partnership. The mkfcpartnership and mkippartnership command 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

### 11.8.2 Listing available Storwize V7000 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.

### 11.8.3 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 Storwizetask chcluster command**

The 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 MBps - 1000 MBps. The partnership overall limit is controlled by the chpartnership -bandwidth 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.
-gminterdelaysimulation link_tolerance

This parameter specifies the number of milliseconds that I/O activity (intercluster copying to an auxiliary volume) is delayed. This parameter permits you to test performance implications before Global Mirror is deployed and a long-distance link is obtained. Specify a value of 0 - 100 milliseconds in 1-millisecond increments. The default value is 0. Use this argument to test each intercluster Global Mirror relationship separately.

-gmintradelaysimulation link_tolerance

This parameter specifies the number of milliseconds that I/O activity (intracluster copying to an auxiliary volume) is delayed. By using this parameter, you can test performance implications before Global Mirror is deployed and a long-distance link is obtained. Specify a value of 0 - 100 milliseconds in 1-millisecond increments. The default value is 0. Use this argument to test each intracluster Global Mirror relationship separately.

Use the Storwizetask chcluster command to adjust these values, as shown in the following example:

Storwizetask chcluster -gmlinktolerance 300

You can view all of these parameter values by using the Storwiseinfo lscluster <clustername> command.

gmlinktolerance

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 Global Mirror relationships are automatically stopped, which protects the application hosts at the primary site. During normal operations, application hosts experience a minimal effect from the response times because the Global Mirror feature uses asynchronous replication.

However, if Global Mirror 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 gmlinktolerance feature stops Global Mirror relationships and the application host's response time returns to normal. After a 1920 event occurs, the Global Mirror auxiliary volumes are no longer in the consistent_synchronized state until you fix the cause of the event and restart your Global Mirror relationships. For this reason, ensure that you monitor the system to track when these 1920 events occur.

You can disable the gmlinktolerance feature by setting the gmlinktolerance value to 0 (zero). However, the gmlinktolerance feature cannot protect applications from extended response times if it is disabled. It might be appropriate to disable the gmlinktolerance 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 Global Mirror volumes.

- During periods when application hosts can tolerate extended response times and it is expected that the gmlinktolerance feature might stop the Global Mirror relationships. For example, if you test by using an I/O generator that is configured to stress the back-end storage, the gmlinktolerance feature might detect the high latency and stop the Global Mirror relationships. Disabling the gmlinktolerance 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 Tivoli Storage Productivity Center, to help identify and resolve the problem.

### 11.8.4 Storwize V7000 system partnership

To create a Storwize V7000 system partnership, use the `Storwizetask mkfcpartnership` command for traditional Fibre Channel (FC or FCoE) connections, or, `Storwizetask mkippartnership` for IP-based connections.

**The Storwizetask mkfcpartnership command**

Use the `Storwizetask mkfcpartnership` command to establish a one-way MM/GM partnership between the local system and a remote system. Alternatively, use `Storwizetask 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 Storwize V7000 systems.

When the partnership is created, you can specify the bandwidth to be used by the background copy process between the local Storwize V7000 system and the remote Storwize V7000 system. If it is not specified, the bandwidth defaults to 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 overload the auxiliary storage and again delay the synchronous secondary writes of foreground I/Os.

To set the background copy bandwidth optimally, make sure 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 and then reducing the background copy to accommodate peaks in workload and another safety margin.
The Storwizetask chpartnership command
To change the bandwidth that is available for background copy in a Storwize V7000 system partnership, use the Storwizetask chpartnership command to specify the new bandwidth.

11.8.5 Creating an MM/GM Consistency Group

Use the mkrcconsistgrp command to create an empty MM/GM Consistency Group.

The MM/GM Consistency Group name must be unique across all Consistency Groups that are known to the systems owning this Consistency Group. 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 chrelationship command.

11.8.6 Creating an MM/GM relationship

Use the mkrcrelationship command to create a new 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 instead of 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 that exists or it can be a stand-alone MM/GM relationship if no Consistency Group is specified.

The Storwiseinfo 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.

11.8.7 Changing an MM/GM 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 using the -force flag.
11.8.8 Changing an MM/GM Consistency Group

Use the `chrcconsistgrp` command to change the name of an MM/GM Consistency Group.

11.8.9 Starting an MM/GM 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 already 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) takes place and therefore, is unusable for DR 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.

11.8.10 Stopping an MM/GM 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 (that is, in the ConsistentStopped, ConsistentSynchronized, or ConsistentDisconnected state), you can use the `-access` parameter with the `stoprcrelationship` command to enable write access to the auxiliary volume.

**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.
11.8.11 Starting an MM/GM Consistency Group

Use the `startrccconsistgrp` command to start an MM/GM Consistency Group. You can issue this command only to a Consistency Group that is connected.

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.

11.8.12 Stopping an MM/GM Consistency Group

Use the `Storwizetask startrccconsistgrp` 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.

If the Consistency Group is in an inconsistent state, any copy operation stops and does not resume until you issue the `Storwizetask startrccconsistgrp` 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.

When a Consistency Group is in a consistent state (for example, in the ConsistentStopped, ConsistentSynchronized, or ConsistentDisconnected state), you can use the `-access` parameter with the `Storwizetask stoprccconsistgrp` command to enable write access to the auxiliary volumes within that group.

11.8.13 Deleting an MM/GM relationship

Use the `Storwizetask rmrcrelationship` command to delete the relationship that is specified. Deleting a relationship deletes only the logical relationship between the two volumes. It does not affect the volumes themselves.

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 of the 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.

11.8.14 Deleting an MM/GM Consistency Group

Use the `Storwizetask rmrcc consistgrp` command to delete an MM/GM Consistency Group. This command deletes the specified Consistency Group. You can issue this command for any existing 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 `Storwizetask 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.

### 11.8.15 Reversing an MM/GM relationship

Use the `Storwizetask 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.

### 11.8.16 Reversing an MM/GM Consistency Group

Use the `Storwizetask 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, by reversing the roles your current source volumes become a target and target volumes become source, which means you lose write access to your current primary volumes.

### 11.9 Troubleshooting remote copy

Remote copy (Metro Mirror and Global Mirror) has two primary error codes that are displayed: 1920 or 1720. 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.

#### 11.9.1 1920 error

A 1920 error (event ID 050010) can have several triggers, including the following probable causes:

- Primary 2145 system or SAN fabric problem (10%)
- Primary 2145 system or SAN fabric configuration (10%)
- Secondary 2145 system or SAN fabric problem (15%)
- Secondary 2145 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 11-34 on page 440. A message that is sent from your source Storwize V7000 system to your target Storwize V7000 System and the accompanying acknowledgment must have a total time of 80 or 250-milliseconds round trip. In other words, it has to 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.

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 that 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. The FCIP routing layer can introduce latency if it is not properly configured. 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 allow you to check the RTT. When you are checking latency, remember that TCP/IP routing devices (including FCIP routers) report RTT or round-trip-time using standard 64-byte ping packets.

In Figure 11-41 on page 465, 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 overhead to be safe 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.

Before we proceed, we 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 11-41 on page 465 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 11-41, 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.

11.9.2 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 that is 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 HBA port for each node per I/O Group if your fabric has more than 64 host bus adapter (HBA) ports zoned. One port for each node per I/O Group per fabric that is associated with the host is the recommended zoning configuration for fabrics. For those fabrics with 64 or more host ports, this recommendation becomes a rule. This means you will see four paths to each volume discovered on the host. This is because host should have at least two FC ports from separate HBA cards, each in separate fabric. On each fabric, each host FC port is zoned to two of Storwize ports where each Storwize port comes from one Storwize node. 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 via IBM Tivoli Storage Productivity 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 can 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 could indicate a larger problem.

If you receive multiple 1720 errors, recheck your network connection and then check the Storwize 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.
Encryption

The Storwize V7000 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 only be enabled and configured on enclosures that support encryption.

Be aware that the IBM Storwize V7000 Gen1 does not support encryption.
12.1 Encryption using the Storwize V7000 Gen2

What does data at rest mean:
- Encryption is the process of encoding data so that only authorized parties can read it
- Uses secret keys to encode the data according to well known algorithms
- Data at rest means that the data is encrypted on the end device (drives)
- Algorithm being used is AES – US government standard from 2001
- Complies with FIPS-140 standard, but not certified
- XTS-AES 256 for data keys
- AES 256 for master keys
- Algorithm is public. The only secrets here are the keys
- Symmetric key algorithm (Same key used to encrypt and decrypt data)

Encryption of system data and system metadata is not required, so system data and metadata are not encrypted. You must have purchased an encryption license before you activate the function. If you have not purchased a license, contact a customer representative to purchase an encryption license.

The following is some important information about the implementation:
- Works on Storwize V7000 Gen2 only

**Note:** You can cluster a Storwize V7000 Gen1 and Storwize V7000 Gen2 and perform encryption.

- Compression is performed “before” encryption. Can use both compression and encryption for the same data.
- Encryption is performed by the SAS hardware. Therefore, encryption only works to internal drives.
- The object that the user can decide whether to be encrypted or not is the array. This can only be decided at create time, and it is not possible to change it once it has been created.

**Note:** Once encryption in the enclosure is enabled, all new build arrays are encrypted. If you want to disable encryption, you need to move the data on to a different system or you have to create new arrays in the same system. You can create a non-encrypted array via the CLI option `mkarray` with the `-encrypt no` parameter.

GUI default is: **encryption always on**

- Uses unique keys for each array, created at array create time. Creation of multiple keys is possible.
- Two types of keys
  - Master key (one per system)
    - Master key is created when encryption is enabled. It is stored on USB devices when encryption is enabled.
    - Can be copied or backed up as necessary
  - Data encryption key (one per encrypted array)
    - Data encryption key is used to encrypt data and is created automatically when an encrypted array/managed disk (MDisk) is created
    - Stored in secure memory in SAS controller hardware
• Stored encrypted with the master key
• No way to view data encryption key
• Cannot be changed
• Discarded when an array is deleted (secure erase)

**Note:** Have at least three USB flash drives available otherwise Encryption cannot be enabled.

We recommend IBM USB devices, although it is not a requirement, order them in eConfig as:

Feature code ACEA: Encryption USB Flash Drives (Four Pack).

► It is *not* stored in CSM memory, or any other non-volatile storage on the cluster.
► There are no “trial” licenses for encryption on the basis that when that trial runs out, access to the data would be lost.

**Note:** If all keys are lost and the system has to reboot, all data is gone. There is no way, even for IBM, to decrypt the data without the keys.

► Contact your IBM sales representative to obtain the authorization codes.

**Note:** There are two different keys used to activate and enable the Encryption feature. One is the **Activation key**, which you need to enable the Encryption feature automatically during initial System setup. The second one is the **License key**. If you want to enable the system later or automatic System setup fails, you can enable Encryption with the License key.

Ensure that you have the following information:
► Machine type and model
► Serial number
► Machine signature

The machine signature can be found in the help panel of the manual activation panel. The following steps show you how to obtain the machine signature. Therefore, you have to open **System → Licensing** as shown in Figure 12-1.

![Figure 12-1](image) **Encryption Licenses selection in the System/Licensing folder**
Press *Encryption Licenses* as marked in Figure 12-1 on page 469. The enclosure selection menu opens as shown in Figure 12-2.

![Figure 12-2](image)

**Figure 12-2** Select the enclosure where you want to enable the encryption

Under Actions, open the Activation menu. From here, select *Activate License Manually* (Figure 12-3).

![Figure 12-3](image)

**Figure 12-3** Activate license manually
The panel to enter the Encryption license opens. On Figure 12-4 at the lower left corner you find Need Help.

![Figure 12-4 Need Help button](image)

When you press the Need Help button, a panel opens where you can find the Machine signature (Figure 12-5).

![Figure 12-5 Machine signature of the machine](image)
12.2 Encryption activation

During system setup, you are asked to either manually or automatically activate the license on the system. Figure 12-6 shows the prompt during System Setup.

![Encryption activation during System setup](image)

When you press Yes in the panel shown in Figure 12-6, you get the Information panel (see Figure 12-7 on page 473) with your local machine. Under the Actions button, you find two options. You are asked either to activate the Encryption automatically or manually.
Figure 12-7  Information panel of the local machine
12.3 Automatic activation of the Encryption feature

In this section, we show how to activate the Encryption function automatically.

Figure 12-8 shows how to start with the Automatic activation. It requires that the workstation that is being used to activate the license is connected to an external network.

![Figure 12-8 Activate Encryption automatically](image)

Select your enclosure and press **Next**. Figure 12-9 on page 475 shows the pop-up menu to enter the purchased license key.
Figure 12-9  Enter authorization code

Figure 12-10 shows an example of the authorization code being entered.

![Activate License Automatically dialog box]

Figure 12-10  Sample of an authorization code

Enter the authorization code that is sent with the licensed function authorization documents that you receive after purchasing the license. These documents contain the authorization codes that are required to obtain keys for each licensed function that you purchased for your system. Click **Activate**. After this, the Storwize V7000 connects to IBM to verify the authorization code and downloads the license key.

In Figure 12-11 on page 476 you can see the process of the automatic encryption enablement.
Figure 12-11  Process of activating Encryption

Figure 12-12 shows the system connecting to IBM. If all works correctly the procedure takes less than a minute.

When Encryption is successfully enabled, you see the green check mark under the Licensed row as shown in Figure 12-13 on page 477.
Chapter 12. Encryption

Figure 12-13  Successfully activation of the Encryption feature

If there are problems with the activation procedure, the Storwize V7000 will timeout after a while (approximately 2:30 minutes) as shown in Figure 12-14.

In this case, check that you have a valid activation (not license) code, access to the Internet or other problems with the Storwize V7000 Gen2. If this still does not work, you can try to use the manual activation procedure, which is described in the next section. For this, you need a valid Encryption license key.

Figure 12-14  No connection possible
12.4 Manual activation of the Encryption feature

If you already set up your system, or automatic activation does not work, you have the ability to enable encryption manually. The following steps show you how to enable encryption manually during initial system setup. Figure 12-15 shows the panel when you open the drop-down menu on the Actions button.

![Figure 12-15  Selecting ‘Activating License Manually’ during System Setup](image)

You find the activation function on a running system under System → Licensing, as shown in Figure 12-16 on page 479.
Chapter 12. Encryption

Figure 12-16  Activate Encryption license panel on a running system

After pressing Activate License Manually, the panel to enter the license code opens. Figure 12-17 shows the Manual Activation panel.

Figure 12-17  Manually enter the license key

You can enter the license key either by typing it or by using cut/copy and paste. In Figure 12-18 on page 480 the sample key is already entered.
Press *Activate* after entering the license key. Figure 12-19 shows the completion of the activation process.

Now all is ready to use Encryption on the Storwize V7000 Gen2. Figure 12-20 on page 481 shows the check mark under the *Licensed* section.
Figure 12-20  Encryption activated
12.5 Enabling encryption

If the activation passed all the steps required, you need to enable encryption.

The process requires three USB flash drives to store the generated keys. If your system is in a secure location, two USB flash drives can remain inserted in two different node canisters. If the location is not secure, all USB flash drives with the keys can be removed from the system and be stored securely.

**Note:** Keep in mind that when you remove all USB keys, the system is not able to access the encrypted data after a system reboot. You have to insert at least one USB key with the actual encryption key, so bear this in mind if you want to avoid any delays however small.

Additional copies of the keys must be created and stored securely to ensure access to the system if the USB flash drives become damaged or stolen. You are required to insert USB flash drives into the canisters to enable encryption and to rekey the system. During these operations, you are responsible for ensuring the security of the system. During power-on, we recommend inserting USB flash drives into the USB ports on two supported canisters to guard against any unexpected failure of the node, node’s USB port, or the USB flash drive during power-on. Use these general guidelines when enabling encryption and managing USB flash drives that contain encryption keys. While the system is enabling encryption, you are prompted to insert the USB flash drives into the system. The system copies the encryption keys to these drives systematically. The system generates and copies the encryption keys to all available USB flash drives.

**Notes:**

- In addition to the copies that are generated on the USB flash drives when encryption is enabled on the system, make at least one additional copy on another USB flash drive and store in a secure location. Copy the encryption keys to other forms of storage to provide resiliency and to mitigate risk, if, for example, the USB flash drives are from a faulty batch of drives.

- Ensure that each copy of the encryption key is valid before writing 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 not usable or failed, the system does not display it as output.

- Securely store all copies of the encryption key. As an example, any USB flash drives that are not left inserted into the system could be locked in a safe. Comparable precautions should be taken to securely protect any other copies of the encryption key stored to other forms of storage.

We now show you how to proceed with the enabling. Figure 12-21 on page 483 shows the wizard, which appears after successful activation of encryption during system setup.
Pressing *Enable Encryption* makes the panel in Figure 12-22 appear.

![Enable Encryption wizard](image)
Insert the three USB flash drives into the USB Ports as requested in Figure 12-23, on the rear of the Storwize V7000 Gen2. You can insert up to four USB flash drives.

![Figure 12-23 Insert at least 3 USB flash drives](image)

Figure 12-23 Insert at least 3 USB flash drives

Figure 12-24 shows where on the Storwize V7000 node canister to put the USB flash drives.

![Figure 12-24 USB ports for the flash drives on one canister](image)

Figure 12-24 USB ports for the flash drives on one canister
The system recognizes the insertion of the three necessary keys and then you are able to press **Next**, as shown in Figure 12-25.

![Image: Two flash drives inserted](image)

*Figure 12-25  Two flash drives inserted*

The encryption keys are automatically copied on the USB flash drives as shown in Figure 12-26 on page 486. In our case, we use four flash drives.
Figure 12-26  Writing encryption key

When the process is complete, you see the panel shown in Figure 12-27 on page 487.
When you select the *Show details/Hide details* flip button, you can see, as shown in Figure 12-28, that all four keys are prepared and ready to use. If one USB flash drive fails, it is not shown.
The **Next** button takes you to the final Commit panel. Click **Commit** to finalize as shown in Figure 12-29.

![Figure 12-29 Commit the installation](image)

Encryption is now enabled as shown in Figure 12-30.

![Figure 12-30 Encryption is now enabled](image)

Figure 12-31 on page 489 shows the entry of enabling encryption in the Audit Log.
To verify the proper installation of Encryption, see the Security page of the Storwize V7000 Gen2 GUI shown in Figure 12-32. Here, press >4 USB Flash Drives Detected to get more details of the proper state of your USB flash drives.

Ensure that each copy of the encryption key is valid before writing 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 not usable or has failed, the system does not display it as output.

**Note:** In addition to the copies that are generated on the USB flash drives when encryption is enabled on the system, make at least one additional copy on another USB flash drive and store it in a secure location.

For each canister on the system, you can leave the USB flash drives inserted into the system. However, the area where the system is located must be secure to prevent keys being lost or stolen. If the area where the system is located is not secure, remove all the USB flash drives from the system and store them securely.

Figure 12-34 on page 490 shows what it looks like when there is no USB flash drive inserted.
After successfully enabling encryption, you can start to create an array.

When the array is created, you can see Figure 12-35 at the Filter bar that the newly created MDisks are encrypted.
Figure 12-36 shows an example of a newly created array with encryption enabled. On the right side, we opened the Filter bar and you can see that we activated the encryption view.

12.5.1 CLI commands to enable encryption

Run `chencryption -usb enable`

Insert three or more USB flash devices into the system.

Run `chencryption -usb newkey -key prepare`

This validates the USB flash devices. Then, it generates new master keys and writes them to a file on each USB device. It also writes keys to SAS hardware, but does not yet enable them. It stores the number of USB flash devices that the keys have been written to. Any failure here will fail this command.

Run `chencryption -usb newkey -key commit`

This will only work if the `-key prepare` step worked and enables the keys in SAS hardware.

If encryption is enabled, then arrays will automatically be created as encrypted if the control enclosure supports it.

To disable encryption, it requires that you must have no data on encrypted arrays.

Encryption can be disabled for a new array, running `mkarray` with the `-encrypt no` parameter.

There is no option in the GUI to create a non-encrypted array.
12.6 Rekeying an encryption-enabled system

*Rekeying* is the process of creating a new key for the system. To create new keys, encryption must be enabled on the system; however, the rekey operation works whether or not there are encrypted arrays.

**Note:** Before creating a new key, ensure that the canisters are online and at least one USB port contains a USB flash drive that contains the current key.

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 USB flash drive does not contain the current key. To rekey the system, you need at least three USB flash drives to store the copied key material.

To rekey the system in the management GUI, complete these steps:

In the management GUI, select **Settings Security Encryption** as shown in Figure 12-37.

![Figure 12-37](image-url)

*Figure 12-37  Overview of the actual inserted encryption keys*

Verify that the encryption key is accessible, which means at least one of the USB flash drives contains the current key. See Figure 12-37. Insert other USB flash drives into the remaining ports. Available ports are displayed to indicate which ports need USB flash drives.

After inserting the remaining USB flash drives into the system, select **Action → Rekey**, as shown in Figure 12-38 on page 493.
When the system detects the required number of the USB flash drives with at least one drive that contains an existing key, the new keys are generated and copied to the USB flash drives. Click **Commit** after the keys are created to complete the rekey operation. If errors occur during the rekey process, status messages display problems with the copy or creation of new keys. For example, if the minimum number of keys are inserted but none of them have an existing encryption key, the rekey operation fails. To determine and fix other possible errors, select **Monitoring → Events**.
12.7 Disabling encryption

The process of disabling encryption requires that you have no data on encrypted arrays as it will effectively be destroyed if there is.

Only then can you remove the encryption key from the SAS hardware, and go back to a known state. To disable the Encryption feature, go to the Security page and select the Enabled flip down arrow as shown in Figure 12-39.

![Encryption Disable Process](image1)

*Figure 12-39  To disable encryption, select the arrow at “Enabled” and then press “Disabled”*

The process starts to disable encryption and Figure 12-40 shows the output when the task is completed.

![Encryption Task Details](image2)

*Figure 12-40  Encryption is disabled*

Encryption is now disabled.
RAS, monitoring, and troubleshooting

There are many ways to manage, monitor, and troubleshoot the IBM Storwize V7000 system. In this chapter, we introduce useful, common procedures to do so according to the best practices. We present the following topics:

- Reliability, availability, and serviceability (RAS) options
- Hardware descriptions with status indications
- Monitoring from a host
- Monitoring from IBM Storwize V7000 system
- Backup procedure
- Software upgrade procedure
- Troubleshooting
- Recommended actions to solve events
- Event log navigation
- Audit log navigation
- Support options
- Shutting down IBM Storwize V7000 system
- Shutting down the infrastructure with installed IBM Storwize V7000 system
13.1 Reliability, availability, and serviceability

Reliability, availability, and serviceability (RAS) are important concepts in the design of IBM Storwize V7000 system. Hardware features, software features, design considerations, and operational guidelines all contribute to make the IBM Storwize V7000 system reliable.

Fault tolerance and high levels of availability are achieved by these methods:

- The Redundant Array of Independent Disks (RAID) capabilities of the underlying disks
- IBM Storwize V7000 nodes clustering using a Compass architecture
- Auto-restart of hung nodes
- Integrated UPS units to provide memory protection in the event of a site power failure
- Host system failover capabilities

High levels of serviceability are available through these methods:

- Cluster error logging
- Asynchronous error notification
- Dump capabilities to capture software detected failures
- Concurrent diagnostic procedures
- Directed maintenance procedures with simplified drive replacement process
- Concurrent log analysis and memory dump data recovery tools
- Concurrent maintenance of all IBM Storwize V7000 components
- Concurrent upgrade of IBM Storwize V7000 software and microcode
- Concurrent addition or deletion of node canisters in a clustered system
- Automatic software version correction when replacing a node
- Detailed status and error conditions displayed on the service panel
- Error and event notification through SNMP, syslog, and email

The heart of IBM Storwize V7000 system is a pair of node canisters. These two canisters share the data transmitting and receiving load between the attached hosts and the disk arrays. This section examines the RAS features of IBM Storwize V7000 system, monitoring, and troubleshooting.

13.1.1 Node canisters

The two node canisters are in the control enclosure and they work as a clustered system. Figure 13-1 shows the ports and indicator lights of a node canister. Rear canister LEDs are identical to front panel LEDs. The second canister is placed next to the first one in a side-by-side position.

![Figure 13-1 Ports and indicators of node canister (524 controller model)](image-url)
Host Interface Cards
Two Host Interface Cards (HIC) slots are vertically placed on the right side of the canister complemented by additional slot for Compression Accelerator card or compression pass-through module. HIC slots can accommodate up to two dual-port 16 Gbps Fibre Channel (FC) cards, or one four-port 10 GbE adapter stand-alone or with combination with FC card.

The 16 Gbps FC adapter is based on Emulex Lancer multiprotocol chip and allows FC or FCoE traffic using a single chip (Converged Network Adapter). It can be configured as dual-port 16 Gbps FC adapter or as a four-port 10 GbE with FCoE support. In case of FC configuration, the meaning of port LEDs is explained in Table 13-1.

<table>
<thead>
<tr>
<th>Port LED</th>
<th>Color</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Link status</td>
<td>Green</td>
<td>Link us up, connection established.</td>
</tr>
<tr>
<td>Speed</td>
<td>Amber</td>
<td>Link is not up or speed fault.</td>
</tr>
</tbody>
</table>

USB
Two active USB connectors are available horizontal position; they have no numbers and no indicators are associated with them.

Ethernet and LED status
Four 10/100/1000 Mbps Ethernet ports are side by side on the canister; they are marked as 1 and 3 on the left and 2 and T on the right. Each port has two LEDs and their status values are shown in Table 13-2. While ports one to three are available for management and Internet Small Computer System Interface (iSCSI) purposes, the T port is strictly dedicated to the technician actions (initial and emergency configuration by local support personnel).

<table>
<thead>
<tr>
<th>LED</th>
<th>Color</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Link state</td>
<td>Green</td>
<td>It is on when there is an Ethernet link.</td>
</tr>
<tr>
<td>Activity</td>
<td>Amber</td>
<td>It is flashing when there is activity on the link.</td>
</tr>
</tbody>
</table>

Serial-attached SCSI ports
Two 12 Gbps serial-attached SCSI (SAS) ports are side by side on the canister with indicator LEDs below them. They are numbered 1 on the left and 2 on the right. Each port is associated with one green and one amber LED indicating its status of the operation as shown in Table 13-3.

<table>
<thead>
<tr>
<th>LED</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green</td>
<td>Link is connected and up.</td>
</tr>
<tr>
<td>Orange</td>
<td>Fault on the SAS link (disconnected, wrong speed, errors).</td>
</tr>
</tbody>
</table>

Node canister status LEDs
There are three LEDs in a row in the upper middle position of the canister that indicate the status and the functionality of the node (Table 13-4 on page 498).
### Table 13-4  Node canister LEDs

<table>
<thead>
<tr>
<th>Position</th>
<th>Color</th>
<th>Name</th>
<th>State</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left</td>
<td>Green</td>
<td>Power</td>
<td>On</td>
<td>The node is powered up and active. It might not be safe to remove the canister. If the fault LED is off, the node is an active member of a cluster or candidate. If the fault LED is also on, node is in service state or in error preventing the software to start.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Flashing (2 Hz)</td>
<td>Canister is powered up and in standby mode.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Blinking (4 Hz)</td>
<td>Node is running power-on self-test (POST).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Off</td>
<td>No power to the canister or it is running on battery.</td>
</tr>
<tr>
<td>Middle</td>
<td>Green</td>
<td>Status</td>
<td>On</td>
<td>The node is a cluster member.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Flashing (2 Hz)</td>
<td>The node is a candidate or in service state.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Blinking (4 Hz)</td>
<td>The node performs Firehose dump. Never unplug the canister at this time.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Off</td>
<td>No power to the canister or canister is in standby mode.</td>
</tr>
<tr>
<td>Right</td>
<td>Amber</td>
<td>Fault</td>
<td>On</td>
<td>The canister is in service state or in error preventing the software to start.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Blinking (2 Hz)</td>
<td>Canister is being identified.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Off</td>
<td>Node is either in candidate or active state.</td>
</tr>
</tbody>
</table>

### 13.1.2 Expansion canisters

As Figure 13-2 shows, two 12 Gbps SAS ports are side by side on the canister of every enclosure. They are numbered 1 on the left and 2 on the right. Similarly to the controller canisters, also expansion canisters are installed in the enclosure side-by-side in vertical position.

The interpretation of SAS status LED indicators has the same meaning as the LED indicators of SAS ports in control enclosure (Table 13-3 on page 497).
Table 13-5 shows the LED status values of the expansion canister.

<table>
<thead>
<tr>
<th>Position</th>
<th>Color</th>
<th>Name</th>
<th>State</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left</td>
<td>Green</td>
<td>Power</td>
<td>On</td>
<td>The canister is powered on.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Off</td>
<td>No power available to the canister.</td>
</tr>
<tr>
<td>Middle</td>
<td>Green</td>
<td>Status</td>
<td>On</td>
<td>The canister is operating normally.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Blinking</td>
<td>There is an error with the VPD.</td>
</tr>
<tr>
<td>Right</td>
<td>Amber</td>
<td>Fault</td>
<td>On</td>
<td>There is an error logged against the canister or the system is not running (OSES).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Blinking</td>
<td>Canister is being identified.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Off</td>
<td>No fault, canister is operating normally.</td>
</tr>
</tbody>
</table>

### 13.1.3 Enclosure SAS cabling

Expansion enclosures are attached to control enclosures using 12 Gbps SAS cables. IBM Storwize V7000 Gen2 Control Enclosure attaches up to 20 expansion enclosures; twice more than the previous generation.

A strand starts with an SAS initiator chip inside an IBM Storwize V7000 node canister and progresses through SAS expanders, which connect disk drives; each canister contains an expander. Figure 13-3 shows how the SAS connectivity works inside the node and expansion canisters. Each drive has two ports, each connected to a different expander and strand. This configuration assures that both nodes in the I/O group have direct access to each drive and there is no single point of failure.

![Figure 13-3 Concept of SAS chaining](image)
A chain consists of a set of enclosures, correctly interconnected (Figure 13-4). Chain 1 of an I/O group is connected to SAS port 1 of both node canisters; chain 2 is connected to SAS port 2. This configuration means that chain 2 includes the SAS expander and drives of the control enclosure.

![Figure 13-4 SAS cabling with numbering of enclosures](image)

At system initialization, when devices are added to or removed from strands, IBM Storwize V7000 performs discovery process to update the state of the drive and enclosure objects.

### 13.1.4 Power

All enclosures accommodate two power supply units (PSUs) for normal operation. A single PSU can supply the entire enclosure for redundancy. For this reason, it is highly recommended to supply AC power to each PSU from different power distribution units (PDUs).

There is a power switch on the power supply and indicator LEDs. The switch must be on for the PSU to be operational. If the power switch is turned off, the PSU stops providing power to the system. For control enclosure PSUs, the battery integrated in the node canister continues to supply power to the node. Fully charged battery is able to perform two firehose dumps. It supports the power outage 5 seconds before initiating safety procedures.

Figure 13-5 on page 501 shows two PSUs present in the control and expansion enclosure. The controller PSU has two green and one amber indication LEDs reporting the status of PSU.
Figure 13-5   Controller and expansion enclosure LED status indicator

Figure 13-6 presents the rear overview of the enclosure canister with power unit (PSU). In contrast to the control enclosure, these PSUs do not have a power switch. The enclosure is powered on by the direct attachment of power cable.

Figure 13-6   Expansion enclosure power supply unit

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 for less than 5 seconds, the canister will not perform firehose dump and continue operation from battery. This is useful for a case of, for example, maintenance of UPS systems in the datacenter or replugging the power to the different power source or PDU unit. A fully charged battery is able to perform two firehose dumps.

13.2 Configuration backup

You can download and save the configuration backup file using IBM Storwize V7000 GUI or command-line interface (CLI). On an ad hoc basis, we suggest manually doing this procedure because it is able to save the file directly to your workstation. The command-line option requires login to the system and downloading the dumped file using specific SCP protocol. The command-line option is a good practice for an automated backup of the configuration.

Important: Save configuration files of IBM Storwize V7000 regularly. The best approach is to do this daily and automate this task. Always perform the additional backup before any critical maintenance task such an upgrade of the microcode, software version, and so on.

The backup file is updated by the cluster every day. Saving it after any changes to your system configuration is also important. It contains configuration data of arrays, pools, volumes, and so on. The backup never contains any client data.
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.
- All objects should have non-default names.

Although objects should have non-default names at the time that the backup is taken, this prerequisite is not mandatory. The backup command reports an error when the default name is discovered, but the configuration is saved. However, the default object names for controllers, I/O groups, and managed disks (MDisks) do not restore correctly if the ID of the object differs from what is recorded in the current configuration data file. All other objects with default names are renamed during the restore process.

**Important:** Ad hoc backup of configuration can be done only from the CLI using the `svcconfig backup` command. Then, the output of the command can be downloaded from the graphical user interface (GUI).

### 13.2.1 Backup using CLI

You can use CLI to trigger configuration backup either manually on an ad hoc basis or by an automatic process regularly. The `svcconfig backup` command generates a new backup file. Triggering a backup using the GUI is not possible, but you can save the output from GUI.

Example 13-1 shows output of the `svcconfig backup` command.

```plaintext
Example 13-1   Saving configuration using CLI

IBM_Storwize:ITSO_V7000Gen2_2:ITSO_admin>svcconfig backup
..................................................................................
..................................................................................
...........................................................................
CMMVC6155I SVCCONFIG processing completed successfully
IBM_Storwize:ITSO_V7000Gen2_2:ITSO_admin>
```

The `svcconfig backup` command generates three files that provide information about the backup process and cluster configuration. These files are dumped into the `/tmp` directory on the configuration node. Use the `lsdumps` command to list them. They are typically at the bottom of the list (Example 13-2).

```plaintext
Example 13-2   Listing backup files in CLI

IBM_Storwize:ITSO_V7000Gen2_2:ITSO_admin>lsdumps
id  filename
0  snap.single.78N10WD-1.121221.164843.tgz
1  78N10WD-1.trc.old
2  dump.78N10WD-1.140826.170838
....
32 svc.config.backup.bak_78N10WD-1
33 svc.config.backup.xml_78N10WD-1
34 svc.config.backup.sh_78N10WD-1
35 svc.config.backup.log_78N10WD-1
36 dpa_heat.78N10WD-1.141107.130950.data
IBM_Storwize:ITSO_V7000Gen2_2:ITSO_admin>
```
Table 13-6 describes the three files that are created by the backup process.

<table>
<thead>
<tr>
<th>File name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>svc.config.backup.xml</td>
<td>This file contains your cluster configuration data.</td>
</tr>
<tr>
<td>svc.config.backup.sh</td>
<td>This file contains the names of the commands that were issued to create the backup of the cluster.</td>
</tr>
<tr>
<td>svc.config.backup.log</td>
<td>This file contains details about the backup, including any error information that might have been reported.</td>
</tr>
</tbody>
</table>

### 13.2.2 Backup using GUI

IBM Storwize V7000 does not offer an option to initiate a backup from the GUI, however you can download existing daily backups or those manual backups triggered from CLI. To download a backup of the configuration using GUI, complete the following steps:

1. Navigate to the **Settings** icon and click **Support** (Figure 13-7).

   ![Support option](image)

   **Figure 13-7 Support option**

2. The window shown in Figure 13-8 on page 504 opens. Click **Show full log listing** to show all log files.
3. Search for and right-click the following file name, and then select **Download** to transfer the file to your workstation:

/\dumps/svc.config.backup.xml_* 

### 13.3 Software upgrade

In this section, we describe the operations to upgrade your Storwize V7000 software from 7.3 to the 7.4 version.

**Note:** We also show how to upgrade from 7.4.0.x to 7.4.0.y later in this topic.

The format for the software upgrade package name ends in four positive integers that are separated by dots. For example, a software upgrade package might have the name that is shown in the following example:

IBM_2145_INSTALL_7.4.0.0

### 13.3.1 Precautions before the upgrade

In this section, we describe the precautions you should take before you attempt an upgrade.

**Important:** Before you attempt any IBM Storwize V7000 code update, read and understand the Storwize V7000 concurrent compatibility and code cross-reference matrix. For more information, see the following website and click **Latest Storwize V7000 code**:

http://www.ibm.com/support/docview.wss?uid=ssg1S1003705

During the upgrade, each node in your Storwize V7000 clustered system is automatically shut down and restarted by the upgrade process. Because each node in an I/O Group provides an
alternative path to volumes, use the Subsystem Device Driver (SDD) to make sure that all I/O paths between all hosts and SANs work.

If you do not perform this check, certain hosts might lose connectivity to their volumes and experience I/O errors when Storwize V7000 node that provides that access is shut down during the upgrade process. You can check the I/O paths by using SDD datapath query commands.

### 13.3.2 Storwize V7000 upgrade test utility

The Storwize V7000 software upgrade test utility is a Storwize V7000 software utility that checks for known issues that can cause problems during a Storwize V7000 software upgrade. More information about the utility is available at this website:

http://www.ibm.com/support/docview.wss?rs=591&uid=ssg1S4000585

Download the software upgrade utility from this page where you can also download the firmware. This ensures that you get the latest version of this utility. You can use the svcupgradetest utility to check for known issues that might cause problems during a Storwize V7000 software upgrade.

The software upgrade test utility can be downloaded in advance of the upgrade process, or it can be downloaded and run directly during the software upgrade, as guided by the upgrade wizard.

You can run the utility multiple times on the same Storwize V7000 system to perform a readiness check-in preparation for a software upgrade. We strongly advise running this utility for a final time immediately before you apply the Storwize V7000 upgrade to ensure that there were no new releases of the utility since it was originally downloaded.

The installation and use of this utility is nondisruptive and does not require restart of any Storwize V7000 node; therefore, there is no interruption to host I/O. The utility is only installed on the current configuration node.

System administrators must continue to check whether the version of code that they plan to install is the latest version. You can obtain the latest information at this website:

http://www.ibm.com/support/docview.wss?uid=ssg1S1003705

This utility is intended to supplement rather than duplicate the existing tests that are performed by the Storwize V7000 upgrade procedure (for example, checking for unfixed errors in the error log).

Concurrent software update of all components is supported through the standard Ethernet management interfaces. However, during the upgrade process, most of the configuration tasks are restricted.

### 13.3.3 Upgrade procedure Version 7.3.x.x to 7.4.x.x

To upgrade IBM Storwize V7000 software, complete the following steps:

1. Open a supported web browser and navigate to your cluster IP address. A login window opens (Figure 13-9 on page 506).
2. Log in with superuser rights. IBM Storwize V7000 management home window opens. Move the mouse cursor over **Settings → General** (Figure 13-10) and click **General**.

3. In the General menu, choose **Upgrade Software**. The Upgrade Software pane opens (Figure 13-11).
This window offers two options:

- **Check for updates**: This function checks with the IBM website about whether a version of IBM Storwize V7000 software is newer than the version you have installed. You must have an Internet connection to use this function.

  **My Notifications**: Use the My Notifications tool to receive notifications of new and updated support information to better maintain your system environment, especially in an environment where a direct Internet connection is not possible. Go to the following address (an IBM account is required) and add your IBM Storwize V7000 system to the notifications list to be advised of support information and to download the latest code to your workstation for later upload:


- **Launch Upgrade wizard**: This function starts the software upgrade process.

4. Click **Launch Upgrade Wizard** to start the upgrade process. You are redirected to the window shown in Figure 13-12.

![Figure 13-12: Upgrade Package window](image)

From this window, you can download the upgrade test utility. If you downloaded it previously, you can browse to the location where it is saved (Figure 13-13).

![Figure 13-13: Upload Test Utility](image)

**Naming**: The upgrade test utility is common for all products from the IBM Storwize family. Do not be confused by the IBM2145_ naming convention, which refers to SAN Volume Controller products (while IBM Storwize V7000 is marked as IBM2076_).
5. Click **Next**; the upgrade test utility is applied and the next window opens (Figure 13-14).

![Upgrade Software](image1)

*Figure 13-14  Upgrade Test Utility applied*

6. Click **Close**. The window shown in Figure 13-15 opens, where you can run the Upgrade Test Utility.

![Upgrade Package](image2)

*Figure 13-15  Run Upgrade Test Utility*

7. Click **Next**. You are redirected to the window shown in Figure 13-16. At this time, the Upgrade Test Utility runs and you are able to see the suggested actions, if any, in the window shown in Figure 13-16.

![Results of the Test Utility](image3)

*Figure 13-16  Results of the Test Utility*
8. Click **Next** to start the software upload procedure, and you are redirected to the window shown in Figure 13-17.

![Figure 13-17  Downloading software code](image)

9. Download IBM Storwize V7000 software upgrade package, or browse and upload the software upgrade package from the location where you saved it. Verify that the file is successfully loaded into the system (Figure 13-18).

![Figure 13-18  System ready to upgrade](image)

When prompted, choose **Automatic upgrade** rather than Service Assistant Manual upgrade. Manual upgrade is eligible for the cases when the action is suggested and monitored by IBM support personnel (Figure 13-19).

![Figure 13-19  The automatic upgrade selection](image)
10. Click **Finish**. The software upgrade starts. You are redirected to the window illustrated in Figure 13-20. Ignore the warning message that is related to the detection of cluster node, if it appears.

![Figure 13-20 Upgrade the software code](image)

11. Click **OK**. You have completed your task to upgrade the IBM Storwize V7000 software. You receive messages that inform you that first one node, then the other, has been upgraded. When both nodes have been rebooted, the system waits for your confirmation of the upgrade steps and commits the upgrade. The sequence of messages is illustrated Figure 13-21.

![Figure 13-21 Information messages during upgrade process](image)

During the upgrade process, a node failover occurs and you will temporarily lose connection to the GUI. The pop-up window appears, requesting a confirmation that you want to refresh the current session using a connection to the cluster node that failed over.
12. Once the code load and initial upgrade of both nodes and the system completes, the confirmation is required. Click **Confirm Upgrade** as shown in Figure 13-22.

![Update System](image1)

**Figure 13-22  Upgrade confirmation**

No other configuration activity is available until the upgrade process is confirmed and committed to the system. Before confirming upgrade, make sure that there is no unresolved critical HW error logged in the system. However, with SAN Volume Controller code version 7.4 the system allows the upgrade even if it has recognized a failed disk drive in any expansion unit. Figure 13-23 shows the CLI command.

![Update System](image2)

**Figure 13-23  Confirmation of software upgrade**
13. After successful confirmation, IBM Storwize V7000 will start committing upgrades to each node and system separately in a sequence as shown in Figure 13-24. You might temporarily lose connection to GUI due to failover.

![Figure 13-24 Upgrade commitment](image)

14. The upgrade process completes when both nodes and the system unit are confirmed. The final status indicates the level of the code installed in the system.

### 13.3.4 Upgrade procedure Version 7.4.x.x to 7.4.y.y

To upgrade IBM Storwize V7000 software, complete the following steps:

1. Open a supported web browser and navigate to your cluster IP address. A login window opens (Figure 13-25).

![Figure 13-25 IBM Storwize V7000 GUI login window](image)
2. Log in with superuser rights. IBM Storwize V7000 management home window opens. Move the mouse cursor over **Settings → System** (Figure 13-26).

![Figure 13-26 Settings menu](image)

3. From the window shown in Figure 13-26, you can now only select the following:
   - **Update**: This brings you to the firmware selection panel. Be prepared that you need the actual firmware already downloaded. You will find the firmware here:
     

4. In the System menu, choose **Update**. The Update Software pane opens (Figure 13-27).

![Figure 13-27 Update Software menu](image)

5. Select the Test utility and the Update package that you downloaded before to your local storage. If all selected firmware is valid, you see (as in Figure 13-28 on page 514) the test utility and package information, code level, and the update button changes color. Press **Update** to proceed.
**Naming:** The upgrade test utility is common for all products from the IBM Storwize family. Do not be confused by the IBM2145_ naming convention, which refers to SAN Volume Controller products (while IBM Storwize V7000 is marked as IBM2076_).

---

![Image of Update Package window](image)

**Figure 13-28  Update Package window**

6. The selection panel to choose either to update the system automatically or manually comes next. Below are the differences:

- **Updating the system automatically**

  During the automatic update process, each node in a system is updated one at a time, and the new code is staged on the nodes. While each node restarts, there might be some degradation in the maximum I/O rate that can be sustained by the system. After all the nodes in the system are successfully restarted with the new code level, the new level is automatically committed. During an automatic code update, each node of a working pair is updated sequentially. The node that is being updated is temporarily unavailable and all I/O operations to that node fail. As a result, the I/O error counts increase and the failed I/O operations are directed to the partner node of the working pair. Applications do not see any I/O failures. When new nodes are added to the system, the update package is automatically downloaded to the new nodes from the SAN Volume Controller system. The update can normally be done concurrently with normal user I/O operations. However, there is a possibility that performance could be impacted. If any restrictions apply to the operations that can be done during the update, these restrictions are documented on the product website that you use to download the update packages. During the update procedure, most of configuration commands are not available.

- **Updating the system manually**

  During an automatic update procedure, the system updates each of the nodes systematically. The automatic method is the preferred procedure for updating the code on nodes. However, to provide more flexibility in the update process, you can also update each node manually. During this manual procedure, you prepare the update, remove a node from the system, update the code on the node, and return the node to the system. You repeat this process for the remaining nodes until the last node is removed from the system. Every node must be updated to the same code level. You cannot interrupt the update and switch to installing a different level. After the all the nodes are updated, you must confirm the update to complete the process. The confirmation restarts each node in order and takes about 30 minutes to complete.
7. In Figure 13-29, we select **Automatic update**.

![Figure 13-29 The automatic upgrade selection](image)

8. When you click **Finish**, the Storwize V7000 software upgrade starts. The window that is shown in Figure 13-30 opens. The system starts with the upload of the test utility and the Storwize V7000 system firmware.

![Figure 13-30 Uploading of the test utility and the Storwize V7000 System firmware](image)

9. After a short while, the system starts automatically to run the update test utility (Figure 13-31 on page 516).
Implementing the IBM Storwize V7000 V7.4

Figure 13-31  Running update test utility

10. When the system detects an issue or an error, you are guided by the GUI. See Figure 13-32.

Figure 13-32  Issue detected

11. Close the window. You come back to the Update System panel. Here, press Read more (Figure 13-33).

Figure 13-33  Issues detected by the update test utility

12. The results panel opens and shows you what is wrong. Figure 13-34 on page 517.
13. In our case, the warning is only a warning that we did not enable email notification. Therefore, we can press Close and proceed with the update. For this, press Resume as shown in Figure 13-35.

14. Due to the error another warning comes up, as shown in Figure 13-36 on page 518. We proceed and press Yes.
15. Finally, now the Update process starts. Figure 13-37.

16. When the update for the first node is complete, the system paused for a while (approx. 30 minutes) to ensure that all paths are reestablished to the now updated node. See Figure 13-38.
17. After a while, a node failover happens and closes the web session. Press Yes to reestablish the web session. Figure 13-39.

![Node failover](image1.png)

**Figure 13-39** Node failover

18. After refresh, after a while you will see that the system is updated.

![System is updated](image2.png)

**Figure 13-40** System is updated

The update is complete.

### 13.4 Critical Fix Notification feature

Starting with IBM Storwize V7000 software V6.3.0.0, the new *Critical Fix Notification* function enables IBM to warn IBM Storwize V7000 and SAN Volume Controller users when a critical issue exists in the level of code that they are using. The system notifies users when they log on to the GUI using a web browser connected to the Internet.

Consider the following information about this function:

- It warns users only about critical fixes and does not warn them that they are running a previous version of the software.
- It works only if the browser also has access to the Internet (IBM Storwize V7000 and SAN Volume Controller systems themselves do not need to be connected to the Internet).
- The function cannot be disabled, and each time it displays a warning, it must be acknowledged (with the option to not warn the user again for that issue).
The decision about what is a “critical” fix is subjective and requires judgment, which is
exercised by the development team. As a result, clients might still encounter bugs in code that
were not deemed critical. They should continue to review information about new code levels
to determine if they are supposed to upgrade even without a critical fix notification.

13.5 Monitoring host paths to IBM Storwize V7000

In this section, we show how to monitor paths to IBM Storwize V7000 system. On host
systems, using the IBM multipath module, such as Windows and other operating systems, a
powerful CLI tool is available (see the basic Windows configuration in Chapter 4, “Host
configuration” on page 111 for information about monitoring IBM Storwize V7000 connectivity
from the host). Example 13-3 shows the output of the `datapath query device` command with
four paths that should be available, but two of the paths have been closed, indicating a
problem.

Example 13-3  `pcmpath query device` showing a problem with one canister

```
C:\Program Files\IBM\SDDDSM>datapath query device
Total Devices : 2
DEV#:   0  DEVICE NAME: Disk2 Part0  TYPE: 2145       POLICY: OPTIMIZED
SERIAL: 600507680280801AC8000000000001F
============================================================================
Path#            Adapter/Hard Disk        State  Mode       Select     Errors
0     Scsi Port5 Bus0/Disk2 Part0     OPEN   NORMAL         78          1
1     Scsi Port5 Bus0/Disk2 Part0     OPEN   NORMAL          0          0
2     Scsi Port6 Bus0/Disk2 Part0    CLOSE   NORMAL          0          0
3     Scsi Port6 Bus0/Disk2 Part0    CLOSE   NORMAL         25          1
DEV#:   1  DEVICE NAME: Disk3 Part0  TYPE: 2145       POLICY: OPTIMIZED
SERIAL: 600507680280801AC800000000000020
============================================================================
Path#            Adapter/Hard Disk        State  Mode       Select     Errors
0     Scsi Port5 Bus0/Disk3 Part0     OPEN   NORMAL          1          1
1     Scsi Port5 Bus0/Disk3 Part0     OPEN   NORMAL         78          1
2     Scsi Port6 Bus0/Disk3 Part0    CLOSE   NORMAL         28          1
3     Scsi Port6 Bus0/Disk3 Part0    CLOSE   NORMAL          1          1
```

The `datapath query adapter` command (Example 13-4) shows all IBM Storwize V7000 paths
that are available to the host. It shows that only Adapter 0 is available, and that the state of
Adapter 1 is FAILED.

Example 13-4  `datapath query adapter` command

```
C:\Program Files\IBM\SDDDSM>datapath query adapter
Active Adapters :2
Adpt#   Name       State Mode Select Errors Paths Active
0  Scsi Port5 Bus0 NORMAL ACTIVE 36    0  2   2
1  Scsi Port6 Bus0 FAILED ACTIVE 17    1  2   0
```
After the problem is fixed, scan for new disks on your host, and verify that all paths are available again (Example 13-5).

Example 13-5  datapath query device command

```
C:\Program Files\IBM\SDDDSM>datapath query device
Total Devices : 2

DEV#:   0  DEVICE NAME: Disk2 Part0  TYPE: 2145       POLICY: OPTIMIZED
SERIAL: 600507680280801AC8000000000001F
============================================================================
Path#            Adapter/Hard Disk        State  Mode       Select     Errors
0     Scsi Port5 Bus0/Disk2 Part0     OPEN   NORMAL         95          1
1     Scsi Port5 Bus0/Disk2 Part0     OPEN   NORMAL          0          0
2     Scsi Port6 Bus0/Disk2 Part0     OPEN   NORMAL          0          0
3     Scsi Port6 Bus0/Disk2 Part0     OPEN   NORMAL         28          1

DEV#:   1  DEVICE NAME: Disk3 Part0  TYPE: 2145       POLICY: OPTIMIZED
SERIAL: 600507680280801AC80000000000020
============================================================================
Path#            Adapter/Hard Disk        State  Mode       Select     Errors
0     Scsi Port5 Bus0/Disk3 Part0     OPEN   NORMAL          1          1
1     Scsi Port5 Bus0/Disk3 Part0     OPEN   NORMAL         95          1
2     Scsi Port6 Bus0/Disk3 Part0     OPEN   NORMAL         31          1
3     Scsi Port6 Bus0/Disk3 Part0     OPEN   NORMAL          1          1
```

You can also use the `datapath query adapter` command and check that the FAILED path is back online, where the State is indicated as NORMAL (Example 13-6).

Example 13-6  datapath query adapter to check if a failed path is online

```
C:\Program Files\IBM\SDDDSM>datapath query adapter
Active Adapters :2
Adpt#        Name        State    Mode       Select     Errors  Paths   Active
0  Scsi Port5 Bus0   NORMAL   ACTIVE         52          0      2        2
1  Scsi Port6 Bus0   NORMAL   ACTIVE         21          1      2        2
```

13.6 Troubleshooting and fix procedures

The management GUI of IBM Storwize V7000 is a browser-based GUI for configuring and managing all aspects of your system. It provides extensive facilities to help troubleshoot and correct problems. We explain how to effectively use its features to avoid service disruption of your IBM Storwize V7000.

Figure 13-41 on page 522 shows the menu to start the Monitoring menu for System information, viewing Events, or seeing real-time Performance statistics.
Use the management GUI to manage and service your system. Select **Monitoring → Events** to access problems that must be fixed and maintenance procedures that step you through the process of correcting problems. Information in the Events panel can be filtered in three ways:

- **Recommended Actions**
  
  Shows only the alerts that require attention. 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 do these tasks:
  
  - Run a fix procedure
  - View the properties

- **Unfixed Messages and Alerts**
  
  Displays only the alerts and messages that are not fixed. For each entry that is selected, you can do these tasks:
  
  - Run a fix procedure
  - Mark an event as fixed
  - Filter the entries to show them by specific minutes, hours, or dates
  - Reset the date filter
  - View the properties

- **Show All**
  
  Displays all event types whether they are fixed or unfixed. For each entry that is selected, you can do these tasks:
  
  - Run a fix procedure
  - Mark an event as fixed
  - Filter the entries to show them by specific minutes, hours, or dates
  - Reset the date filter
  - View the properties

Some events require a certain number of occurrences in 25 hours before they are displayed as unfixed. If they do not reach this threshold in 25 hours, they are flagged as **expired**. Monitoring events are below the coalesce threshold and are usually transient.
Important: The management GUI is the primary tool that is used to operate and service your system. The real-time monitoring should be established via SNMP traps, email notifications, or syslog messaging on an automatic manner.

13.6.1 Managing event log

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

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

1. Select Monitoring → Events to see all problems that exist on the system (Figure 13-42).

![Figure 13-42 Messages in the event log]

2. Select Show All → Recommended Actions to display the most important events to be resolved (Figure 13-43 on page 524). The Recommended Actions tab shows the highest priority maintenance procedure that must be run. Use the troubleshooting wizard so that IBM Storwize V7000 system can determine the proper order of maintenance procedures.
In this example, the number of device logins reduced is listed (service error code 1630). Review the physical FC cabling to determine the issue and then click Run Fix. At any time and from any GUI panel, you can directly navigate to this menu by using the Status Alerts icon at the right bottom corner of the GUI (Figure 13-44).

13.6.2 Running a fix procedure

If there is a service error code for the alert, you can run a fix procedure that assists you in resolving the problem. These fix procedures analyze the system and provide more information about the problem. They suggest actions to take and step you through the actions that automatically manage the system where necessary. Finally, they check that the problem is resolved.
If an error is reported, always use the fix procedures from the management GUI to resolve the problem. Always use the fix procedures for both software configuration problems and hardware failures. The fix procedures analyze the system to ensure that the required changes do not cause volumes to be inaccessible to the hosts. The fix procedures automatically perform configuration changes that are required to return the system to its optimum state.

**Good practice:** Before running a fix procedure for the most critical errors, take a backup of the system configuration, as suggested in 13.2, “Configuration backup” on page 501.

The fix procedure displays information that is relevant to the problem and provides various options to correct the problem. Where possible, the fix procedure runs the commands that are required to reconfigure the system. Always use the recommended action for an alert because these actions ensure that all required steps are performed. Use the recommended actions even in cases where the service action seems obvious, such as a drive showing a fault. In this case, the drive must be replaced and reconfiguration must be performed. The fix procedure performs the reconfiguration for you.

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

**Hint:** If possible, fix the alerts in the order shown to resolve the most serious issues first. Often, other alerts are corrected automatically because they were the result of a more serious issue.

The following example demonstrates how to clear the error that is related to the malfunctioning FC connectivity between control canisters of IBM Storwize V7000:

1. From the dynamic menu (the icons on the left) select **Monitoring** → **Events**, and then focus on the errors with the highest priority first. List only the recommended actions by selecting the filters in the **Actions** menu (Figure 13-45). Click **Run Fix**.

![Figure 13-45  Initiate Run Fix procedure from management GUI](image)
2. The pop-up window asks whether the issue was caused by a planned change or maintenance task or whether it appeared in an uncontrolled manner (Figure 13-46).

<table>
<thead>
<tr>
<th>Number of device logins reduced</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Device logins reduced</strong></td>
</tr>
<tr>
<td>The number of device logins is reduced.</td>
</tr>
</tbody>
</table>

Is the reduction of device logins an expected result caused by a change in the configuration?

If it is caused by configuration change, click Yes to manually mark the event as fixed, otherwise, click No to scan the SAN and try to recover the paths.

- **Yes**
- **No**

Click Next for more information.

*Figure 13-46  Determination of planned action*

If you answer Yes, the fix procedure finishes, assuming that all changes in the system are done on purpose and no other action is necessary. However, our example simulates a broken FC cable and follow the complete fix procedure. Select No and click Next.

3. In the next window (Figure 13-47), IBM Storwize V7000 lists suggested actions and which components must be checked to fix and close the error. When you are sure that all possible technical requirements are met (in our case we replaced a broken FC cable), click Next.

<table>
<thead>
<tr>
<th>Number of device logins reduced</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Check configuration</strong></td>
</tr>
</tbody>
</table>

To ensure that the configuration is correct, use the following configuration rules for connection redundancy:

- Minimum, non-redundant connectivity is achieved when an initiator port of each node in an I/O group can connect to a target port of an external storage system.
- Redundant connectivity is achieved when at least two initiator ports of each node in an I/O group can connect to at least two target ports of an external storage system. In this case, connection is provided by at least two separate networks. Each network provides connectivity between one of the target ports on the external storage system and one initiator port of each node.
- Connectivity can be by Fibre Channel or Fibre Channel over Ethernet, or a combination, depending on the hardware configuration of your system.
- Connectivity can be by directly-attached Serial Attached SCSI (SAS), in which case a cable attaches each initiator port on a node to a target port on the external storage system.

Does the current configuration have adequate connection redundancy?

Click Next to confirm that the current configuration has adequate redundancy. Click Cancel to re-verify the configuration.

*Figure 13-47  Verification steps to eliminate single point of failure*
The discovery of managed disks starts (Figure 13-48).

**Discovery start**

The managed disk discovery has started.
If discovery takes longer than 2 minutes, your system may have a problem.
The discovery progress checking will time-out after 15 minutes.
Click Next to check the discovery progress, or click Cancel to exit unfixed.

If no other important issue exists, it finishes maximally within 2 minutes, depending on the number of enclosures and installed disk drives (Figure 13-49).

**Discovery exit**

The managed disk discovery progress is now complete.
Click Next to continue.
4. An event has been marked as fixed and you can safely finish the fix procedure. Click Close and the event is removed from the list of events (Figure 13-50).

<table>
<thead>
<tr>
<th>Event has been marked as fixed</th>
</tr>
</thead>
<tbody>
<tr>
<td>This event relates to the storage system <strong>controller1</strong>.</td>
</tr>
<tr>
<td>The reduced device logins event is now fixed.</td>
</tr>
<tr>
<td>Click Close to exit.</td>
</tr>
</tbody>
</table>

*Figure 13-50  Correctly finished fix procedure*

**Resolve alerts in a timely manner**

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

Multiple views of the events and recommended actions are available. The GUI works like a typical Windows context menu, so the event log grid is manipulated through the row that contains the column headings (Figure 13-51). When you right-click a table heading, a menu for the column choices opens.

![Figure 13-51 Grid options of the event log]

Select or remove columns as needed. You can then also extend or shrink the width of the column to fit your screen resolution and size. This is the way to manipulate it for the majority of grids in the management GUI of IBM Storwize V7000, not only the events panel.

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

You might want to see more details about each critical event. Some details are not shown in the main grid. To access properties and sense data of a specific event, right-click the specific event (anywhere in its row) and choose Properties from the context menu.
The properties window opens (Figure 13-52) with all the relevant sense data such as first and last time of an event occurrence, worldwide port name (WWPN) and worldwide node name (WWNN), enabled or disabled automatic fix, and more.

![Properties and Sense Data for Event 010033]

For more details about troubleshooting options, see IBM Storwize V7000 Troubleshooting, Recovery, and Maintenance Guide, GC27-2291, which is available at the following location:


13.7 Monitoring

An important step is to correct any issues that are reported by your IBM Storwize V7000 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 for which you want to be notified. For example, restrict notifications to just events that require immediate action. Several event notification mechanisms exist:

**Email**

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

**SNMP**

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

**Syslog**

A syslog report can be sent to a data center management system that consolidates syslog reports from multiple systems. With this option, you can monitor your data center from a single location.
If your system is within warranty, or you have a hardware maintenance agreement, configure your IBM Storwize V7000 system to send email events directly to IBM, if an issue that requires hardware replacement is detected. This mechanism is known as *Call Home*. When this event is received, IBM automatically opens a problem report, and if appropriate, contacts you to verify whether replacement parts are required.

**Important:** If you set up Call Home to IBM, ensure that contact details that you configure are correct and kept up to date because personnel can change.

### 13.7.1 Email notifications and Call Home

The Call Home function of IBM Storwize V7000 utilizes the email notification being sent to the specific IBM support center, therefore the configuration is similar as in case of sending emails to the specific person or system owner. The following procedure summarizes how to configure email notifications and emphasizes what is specific to Call Home:

1. Prepare your contact information that you want to use for the email notification and verify the accuracy of the data. From the dynamic menu, select *Settings → Notifications* (Figure 13-53).

![Figure 13-53  Configuration of event notifications](image)

2. Select *Email* and then click *Enable Notifications* (Figure 13-54 on page 532). You can also access the IBM eLearning movie for more technical details:


For the correct functionality of email notifications, ask your network administrator if Simple Mail Transfer Protocol (SMTP) is enabled on the management network and is not, for example, blocked by firewalls. Be sure to test the accessibility to the SMTP server using the `telnet` command (port 25 for a non-secured connection, port 465 for SSL-encrypted communication) using any server in the same network segment.
3. Provide the information about the location of the system (Figure 13-55) and contact information of IBM Storwize V7000 owner (Figure 13-56 on page 533) in order to be reachable by IBM Support. *Always* keep this information current.
4. Configure the SMTP server according to the instruction in Figure 13-57. When the correct SMTP server is provided, you can test the connectivity using Ping to its IP address.

5. In the next step, verify email addresses of IBM Support (callhome0@de.ibm.com) and optionally local users who need to also receive notifications. See Figure 13-58 on page 534 for details. In this wizard you can set only one email address. Other recipients can be added afterward.

The default support email address callhome0@de.ibm.com is predefined by the system during initial configuration of email notifications and at this stage cannot be altered. You can modify it in Edit mode once the initial configuration is saved (see step 6 on page 534).

For additional users, you can also enable the Inventory Reporting function that is enabled by default for Home Call. Rather than reporting a problem, an email is sent to IBM that describes your system hardware and critical configuration information. Object names and other information, such as IP addresses, are not sent. The inventory email is sent regularly. Based on the information that is received, IBM can inform you if the hardware or software that you are using requires an upgrade because of a known issue.
6. Complete the configuration wizard and test the email function. To do so, you have to enter Edit mode as illustrated in Figure 13-59. In the same window, you can define the additional email recipient, either from IBM Support or local users.

We strongly suggest to keep sending inventory enabled to at least IBM support. However, it might be beneficial to do the same for local users. The email output can serve as a basis for the client’s inventory and asset management to keep track of all hardware devices installed in the environment.

7. In Edit mode, you are allowed to change any of the previously configured settings. When you are finished editing these parameters, have added more recipients, or just tested the connection, you can save the configuration to make the changes take effect (Figure 13-60 on page 535).
Disabling and enabling notifications

At any time, you can temporarily or permanently disable email notifications as shown in Figure 13-61. This is good practice when running maintenance tasks on your IBM Storwize V7000, such as code upgrade or replacement of malfunctioning parts. After the maintenance operation, remember to re-enable the email notification function. The same results can be achieved with the CLI `svctask stopmail` and `svctask startmail` commands.
13.8 Audit log

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

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

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

The following items are not documented in the audit log:

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

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

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

Figure 13-62 on page 537 shows the access to the audit log. Click Audit Log in the dynamic menu to see which configuration CLI commands have been run on IBM Storwize V7000 system.
Figure 13-62 Audit Log from Access Management window

Figure 13-63 shows an example of the audit log after creating FlashCopy volume and mapping it to hosts, with a command highlighted. The Running Tasks button is available at the bottom of the window in the status panel. If you click that button, the progress of the currently running tasks is displayed.

<table>
<thead>
<tr>
<th>Date and Time</th>
<th>User Name</th>
<th>Command</th>
</tr>
</thead>
<tbody>
<tr>
<td>11/7/14 8:31:15 AM</td>
<td>superuser</td>
<td>svctask startfmap -prep 0</td>
</tr>
<tr>
<td>11/7/14 8:32:14 AM</td>
<td>superuser</td>
<td>svctask mkudisk -name fctest01_01 loggrp io grp0 mdiskgrp INFRA -size 1073741824</td>
</tr>
<tr>
<td>11/7/14 8:33:14 AM</td>
<td>superuser</td>
<td>svctask mkfmap -source fctest01_target fctest01_01 copyrate 0 : deduprate 0</td>
</tr>
<tr>
<td>11/7/14 8:39:43 AM</td>
<td>superuser</td>
<td>svctask mkudisk -name fctest01_01 loggrp io grp0 mdiskgrp INFRA -size 1073741824</td>
</tr>
<tr>
<td>11/7/14 9:00:01 AM</td>
<td>superuser</td>
<td>svctask detectmdisk</td>
</tr>
<tr>
<td>11/6/14 12:03:41 PM</td>
<td>superuser</td>
<td>svctask mkudiskhostmap -host 0 : vscsi 7 : force 13</td>
</tr>
<tr>
<td>11/6/14 12:09:21 PM</td>
<td>superuser</td>
<td>svctask mkudisk -name image10GB loggrp io grp0 mdiskgrp INFRA -size 204703680</td>
</tr>
<tr>
<td>11/6/14 12:09:52 PM</td>
<td>superuser</td>
<td>svctask mkfmaphostmap -host 0 : vscsi 8 : force 12</td>
</tr>
<tr>
<td>11/6/14 12:48:10 PM</td>
<td>superuser</td>
<td>svctask mkudisk -name image10GB loggrp io grp0 mdiskgrp INFRA -size 1073741824</td>
</tr>
<tr>
<td>11/6/14 1:00:02 AM</td>
<td>superuser</td>
<td>svctask detectmdisk</td>
</tr>
<tr>
<td>11/5/14 10:00:01 AM</td>
<td>superuser</td>
<td>svctask mkudiskhostmap -host 0 : vscsi 7 : force 13</td>
</tr>
<tr>
<td>11/4/14 2:51:43 PM</td>
<td>superuser</td>
<td>svctask detectmdisk</td>
</tr>
</tbody>
</table>

Figure 13-63 Audit log

Changing the view of the Audit Log grid is also possible by right-clicking column headings (Figure 13-64 on page 538). The grid layout and sorting is completely under the user’s control, so you can view everything in the audit log, sort different columns, or reset the default grid preferences.
13.9 Collecting support information

Occasionally, if you have a problem and call the IBM Support Center, they might ask you to provide support data. You can find this data under the Support tab of the Troubleshooting navigation window.

Click Settings and then the Support tab to begin the procedure of collecting support data (Figure 13-65). Assuming that the node restarts, use the menu shown in Figure 13-66 on page 539 to collect the default logs plus all the existing statesaves to capture the maximum data for support.
Click **Download Support Package** (Figure 13-66). To list all individual log files, click the **Show All Individual Log Files** menu.

![Download Support Package window](image)

**Figure 13-66  Download Support Package window**

The window for collecting various versions of svc_snap opens (Figure 13-67); the version you download depends on the event that is investigated. For example, if you notice that a node was restarted in the event log, capture the snap with the latest existing statesave.

![Download support package choices](image)

**Figure 13-67  Download support package choices**

The procedure to create the snap on IBM Storwize V7000 system, including the latest statesave from each node canister, starts. This process might take a few minutes (Figure 13-68 on page 540).
A window opens that gives you the choice to save the file on your local Windows system (Figure 13-69).

Save the resulting snap file in a directory (Figure 13-70).

Before you open a call with IBM Support, be prepared to upload the resulting snap file to the IBM Support portal at the following address:

http://www.ecurep.ibm.com/app/upload

You are ready to call the IBM Support Line or use the IBM Support Portal to open a call. If you use the latter option, go to the following address:

http://www.ibm.com/support/entry/portal/Open_service_request?brandind=Hardware
13.10 Shutting down IBM Storwize V7000

You can safely shut down an IBM Storwize V7000 system using both the GUI or the CLI.

**Important:** Never shut down your IBM Storwize V7000 system by powering off the PSUs, removing both PSUs, or removing both power cables from a running system. It can lead to inconsistency or loss of the data staged in the cache.

Before shutting down IBM Storwize V7000, stop all hosts that have allocated volumes from the device. This step can be skipped for hosts that have volumes that are also provisioned with mirroring (host-based mirror) from different storage device. However, doing so incurs errors that are related to lost storage paths/disks on the host error log.

You can shut down only one node canister or the entire cluster. When you shut down only one node canister, all activities remain active. When you shut down the entire cluster, you need to power on locally to start the system.

13.10.1 Shutting down a node canister

To shut down a single node canister using the GUI, complete the following steps:

1. Hover the cursor over the **Monitoring** function icon and click **System** (Figure 13-71). Rotate the system to the rear side by the arrow in the right-bottom corner.

   ![Figure 13-71  System Device option of the Monitoring function icon](image)

2. Right-click the **Canisters** that you want to stop. Click **Power Off** in the opened context menu (Figure 13-72 on page 542).
3. The confirmation window opens (Figure 13-73). Confirm whether you want to shut down the node. Type the confirmation code and click **OK**. If the node is active as a Configuration node, the control is moved automatically to the second node canister. Your session to the GUI will probably interrupt. Re-establish it again from the browser after takeover happens.

![Figure 13-73 Confirm Shutdown window](image)

Shutdown is complete (Figure 13-74).

![Figure 13-74 Shutdown complete](image)
4. A look at the rear side of the enclosure after canister shutdown is indicated in Figure 13-75.

![Figure 13-75 Rear side of enclosure with powered off canister](image)

To shut down a node canister from the CLI, run the `svctask stopsystem -node 2` command.

### 13.10.2 Shutting down a system

The procedure to shut down a system is similar to shutting down a node canister. Instead of rotating of enclosure and selecting a specific canister in the menu, select the whole Storwize V7000 system from front side, and then from the right-click context menu, select **Power Off** option (Figure 13-76).

![Figure 13-76 Shutdown of IBM Storwize V7000](image)
Confirm the validity of your decision to shut down IBM Storwize V7000 clustered systems by typing the confirmation code in the pop-up window (Figure 13-77).

![Shutdown confirmation](image)

The whole system shutdown is typically planned in case of site maintenance (power outage, building construction, and so on), because all components of IBM Storwize V7000 are redundant and replaceable while the system is running. To start the device again after shutdown, you must have physical access to the system and then turn on the switches on the power supplies.

### 13.10.3 Shutting down and powering on an IBM Storwize V7000 infrastructure

When you shut down or power on the entire infrastructure (storage, servers, applications), follow a particular sequence for both the shutdown and the power-on actions. Here is an example sequence of a shutdown, and then a power-on, of an infrastructure that includes IBM Storwize V7000 system.

**Shutting down**
To shut down the infrastructure, complete the following steps:
1. Shut down your servers and all applications.
2. Shut down your IBM Storwize V7000 system:
   a. Shut down the cluster using either the GUI or CLI.
   b. Power off both switches of the controller enclosure.
   c. Power off both switches of all the expansion enclosures.
3. Shut down your SAN switches.

**Powering on**
To power on your infrastructure, complete the following steps:
1. Power on your SAN switches and wait until the boot completes.
2. Power on your storage systems and wait until the systems are up, and then:
   a. Power on both switches of all the expansion enclosures.
   b. Power on both switches of the controller enclosure.
3. Power on your servers and start your applications.
Appendix A. CLI setup and SAN Boot

This appendix describes the setup of the command-line interface (CLI) and provides extra information about the SAN Boot function.
Command-line interface

The IBM Storwize V7000 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, as that topic is beyond the scope of this book. The basic configuration of the IBM Storwize V7000 CLI and some example commands are covered. However, the CLI commands are the same as in the SAN Volume Controller, and in addition, there are more commands that 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.

Detailed CLI information is available at the IBM Storwize V7000 web page under the command-line section, which is at the following address:

http://www-01.ibm.com/support/knowledgecenter/ST3FR7_7.4.0/com.ibm.storwize.v7000.740.doc/svc_clicommandscontainer_229g0r.html

Implementing the IBM System Storage SAN Volume Controller V7.4, SG24-7933 provides more detailed information about using the CLI. The commands in that book also apply to the IBM Storwize V7000 system.

Basic setup

In the IBM Storwize V7000 GUI, authentication is done by using a user name and password. The CLI uses a Secure Shell (SSH) to connect from the host to the IBM Storwize V7000 system. As of IBM Storwize V7000 V6.3, either a private and a public key pair or user name and password is necessary. The following steps are required to enable CLI access with SSH keys:

- A public key and a private key are generated together as a pair.
- A public key is uploaded to the IBM Storwize V7000 system through the GUI.
- A client SSH tool must be configured to authenticate with the private key.
- A secure connection can be established between the client and IBM Storwize V7000.

Secure Shell is the communication vehicle between the management workstation and the IBM Storwize V7000 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.

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 can consist of up to 40 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 V7000 systems.

The SSH client 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 from the following address at no cost:

http://www.chiark.greenend.org.uk
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 (Figure A-1).

![PuTTY Key Generator](image)

*Figure A-1  PuTTY key generator*

   Ensure that the following options are selected:
   - SSH2 RSA
   - Number of bits in a generated key: 1024

2. Click **Generate** and move the cursor over the blank area to generate keys (Figure A-2 on page 548).
To generate keys: The blank area that is indicated by the message is the large blank rectangle on the GUI inside the section of the GUI labeled Key. Continue to move the mouse pointer over the blank area until the progress bar reaches the far right. 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** (Figure A-3 on page 549).
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, “pubkey”, to easily differentiate the SSH public key from the SSH private key.
5. Click **Save private key** (Figure A-4).

![PuTTY Key Generator](image)

**Figure A-4  Save private key**

6. You are prompted with a warning message (Figure A-5). Click **Yes** to save the private key without a passphrase.

![PuTTYgen Warning](image)

**Figure A-5  Confirm the security warning**

7. When prompted, enter a name (for example, “icat”), select a secure place as the 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 V7000

After you create your SSH key pair, upload your SSH public key onto the SAN Volume Controller system. Complete the following steps:

1. Open the user section in the GUI (Figure A-6).

2. Right-click the user name for which you want to upload the key and click Properties (Figure A-7).
c. To upload the public key, click **Browse**, and select the folder where you stored the public SSH key (Figure A-8).

![File Upload](image1)

*Figure A-8  Selection of the public SSH key*

3. Select your public key, and click **OK** (Figure A-9).

![User Properties](image2)

*Figure A-9  Select public key*

4. Click **OK** and the key is uploaded. Click **Close** to close the wizard and return to the GUI.
Configuring the SSH client

Before the CLI can be used, the SSH client must be configured as follows:

1. Start PuTTY. The PuTTY Configuration window opens (Figure A-10).

   ![PuTTY Configuration Window](image)

   

   Figure A-10   PuTTY

   In the right 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.

2. In the Category pane, on the left side of the PuTTY Configuration window (Figure A-11 on page 554), click **Connection → SSH** to open the PuTTY SSH Configuration window.
3. Under the “Preferred SSH protocol version” section, select 2.

4. In the Category pane on the left, click Connection → SSH → Auth. More options are displayed for controlling SSH authentication.
In the **Private key file for authentication** field (Figure A-12), either browse to or type 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\icat.PPK`).

![PuTTY Configuration](image)

*Figure A-12  SSH authentication*

5. In the Category pane, click **Session** to return to the Basic options for your PuTTY session view (Figure A-13 on page 556).
6. Enter the following information in these fields (Figure A-13) in the right pane:
   - Host Name: Specify the host name or system IP address of the IBM Storwize V7000 clustered system.
   - Saved Sessions, and enter a session name.

![PuTTY Configuration](image)

*Figure A-13  Enter session information*
7. Click **Save** to save the new session (Figure A-14).

![PuTTY Configuration](image)

**Figure A-14**   Save the new session

8. Select the new session and click **Open** to connect to the IBM Storwize V7000 system. A PuTTY Security Alert opens; confirm it by clicking **Yes** (Figure A-15).

![PuTTY Security Alert](image)

**Figure A-15**   Confirm Security Alert

9. PuTTY now connects to the system and prompts you for a user name to log in as. Enter **ITSO Admin** as the user name (Example A-1) and press Enter.

**Example: A-1**   Enter user name

```
login as: ITSO Admin
Authenticating with public key "putty public key"
IBM_2076:ITSO-Storwize-V7000Gen2-2:ITSO Admin>
```
You have now completed the tasks to configure the CLI for IBM Storwize V7000 administration.

**SAN Boot**

IBM Storwize V7000 supports SAN Boot for Windows, VMware, and many other operating systems. SAN Boot support can change, so regularly check the IBM Storwize V7000 interoperability matrix at this address:

http://www.ibm.com/systems/support/storage/ssic/interoperability.wss

The IBM Knowledge Center for Storwize V7000 has much information about SAN Boot in combination with various operating systems. For more information, go to this address:


More information about SAN Boot is also in the *IBM Multipath Subsystem Device Driver User’s Guide*, which is available at the following address:

http://www.ibm.com/support/docview.wss?rs=503&context=HW26L&uid=ssg1S7000303

**Enabling SAN Boot for Windows**

Complete the following procedure if you want to install Windows host using SAN Boot:

1. Configure the IBM Storwize V7000 system so that only the boot volume is mapped to the host.
2. Configure the Fibre Channel SAN so that the host only sees one IBM Storwize V7000 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 using the normal procedure, selecting the volume as the partition on which to install.

   **HBAs:** You might need to load an additional HBA device driver during installation, depending on your Windows version and the HBA type.

5. Install SDDDSM after the installation has completed.
6. Modify your SAN zoning to allow multiple paths.
7. Check your host to see if all paths are available.
8. Set redundant boot devices in the HBA BIOS to allow the host to boot when its original path has failed.

**Enabling SAN Boot for VMware**

Complete the following steps if you want to install a VMware ESX host using SAN Boot:

1. Configure the IBM Storwize V7000 system so that only the boot volume is mapped to the host.
2. Configure the Fibre Channel SAN so that the host only sees one IBM Storwize V7000 system node port. Multiple paths during installation are not supported.
3. Configure and enable the HBA BIOS.
4. Install the operating system using the normal procedure, selecting the volume as the partition on which to install.

**HBA:** You might need to load an additional 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 if all paths are available and modify the multipath policy if required.

**Windows SAN Boot migration**

If you have a host that runs a Windows 2000 Server, Windows Server 2003, Windows Server 2008, or Windows 2012 operating system, and have existing 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 V7000 system.

**SAN Boot procedures:** For SAN Boot procedures for other operating systems, check the IBM Knowledge Center for Storwize V7000:


Complete the following steps to migrate your existing SAN Boot images:

1. If the existing SAN Boot images are controlled by an IBM storage controller that uses the IBM Subsystem Device Driver (SDD) as the multipathing driver, you must use SDD V1.6 or later. Run the SDD `datapath set bootdiskmigrate 2076` command to prepare the host for image migration. See the Multipath Subsystem Device Driver documentation for more information.

2. Shut down the host.

3. Complete the following configuration changes on the storage controller:
   a. Write down the Small Computer System Interface (SCSI) logical unit number (LUN) ID that each volume is using (for example, boot LUN SCSI ID 0, Swap LUN SCSI ID 1, Database Lun SCSI ID 2, and so on).
   b. Remove all the image-to-host mappings from the storage controller.
   c. Map the existing SAN Boot image and any other disks to the IBM Storwize V7000 system.

4. Change the zoning so that the host is able to see the IBM Storwize V7000 I/O group for the target image mode volume.

5. Complete the following configuration changes on the IBM Storwize V7000 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 the host HBA ports.
   c. Map the image mode volume to the host using the same SCSI ID as before. 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.
6. Change the boot address of the host by completing the following steps:
   a. Restart the host and open the HBA 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.
7. If SDD V1.6 or later is installed and you ran `bootdiskmigrate` in step 1 on page 559, reboot your host, update SDDDSM to the latest level, and go to step 14. If SDD V1.6 is not installed, go to step 8.
8. Modify the SAN Zoning so that the host only sees one path to the IBM Storwize V7000.
9. Boot the host in single-path mode.
10. Uninstall any multipathing driver that is not supported for IBM Storwize V7000 system hosts that run the applicable Windows Server operating system.
11. Install SDDDSM.
12. Restart the host in single-path mode and ensure that SDDDSM was properly installed.
13. Modify the SAN Zoning to enable multipathing.
14. Rescan drives on your host and check that all paths are available.
15. Reboot your host and enter the HBA BIOS.
16. Configure the HBA settings on the host. 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.
17. Exit the BIOS utility and finish starting the host.
18. Map any additional volumes to the host as required.
Related publications

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

IBM Redbooks

The following IBM Redbooks publications provide additional information about the topic in this document. Note that some publications referenced in this list might be available in softcopy only.

- *Implementing the IBM System Storage SAN Volume Controller V7.2*, SG24-7933
- *Implementing the IBM Storwize V7000 V7.2*, SG24-7938
- *IBM b-type Gen 5 16 Gbps Switches and Network Advisor*, SG24-8186
- *Introduction to Storage Area Networks and System Networking*, SG24-5470
- *IBM SAN Volume Controller and IBM FlashSystem 820: Best Practices and Performance Capabilities*, REDP-5027
- *Implementing the IBM SAN Volume Controller and FlashSystem 820*, SG24-8172
- *Implementing IBM FlashSystem 840*, SG24-8189
- *IBM FlashSystem in IBM PureFlex System Environments*, TIPS1042
- *IBM FlashSystem 840 Product Guide*, TIPS1079
- *IBM FlashSystem 820 Running in an IBM Storwize V7000 Environment*, TIPS1101
- *Implementing FlashSystem 840 with SAN Volume Controller*, TIPS1137
- *IBM FlashSystem V840 Enterprise Performance Solution*, TIPS1158
- *IBM Midrange System Storage Implementation and Best Practices Guide*, SG24-6363
- *IBM System Storage b-type Multiprotocol Routing: An Introduction and Implementation*, SG24-7544
- *IBM Tivoli Storage Area Network Manager: A Practical Introduction*, SG24-6848
- *Tivoli Storage Productivity Center for Replication for Open Systems*, SG24-8149
- *Tivoli Storage Productivity Center V5.2 Release Guide*, SG24-8204
- *Implementing an IBM b-type SAN with 8 Gbps Directors and Switches*, SG24-6116

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

ibm.com/redbooks
Other resources

These publications are also relevant as further information sources:

- IBM System Storage Open Software Family SAN Volume Controller: CIM Agent Developers Reference, SC26-7545
- IBM System Storage SAN Volume Controller 6.2.0 Configuration Limits and Restrictions, S1003799
- IBM TotalStorage Multipath Subsystem Device Driver User’s Guide, SC30-4096
- IBM XIV and SVC Best Practices Implementation Guide
  http://ibm.co/lbk64gW
- Considerations and Comparisons between IBM SDD for Linux and DM-MPIO
  http://ibm.co/1CD1gxG

Referenced websites

These websites are also relevant as further information sources:

- IBM Storage home page
  http://www.storage.ibm.com
- SAN Volume Controller supported platform
  http://ibm.co/IFNjddm
- SAN Volume Controller IBM Knowledge Center
  http://www-01.ibm.com/support/knowledgecenter/STPVGU/welcome
- Cygwin Linux-like environment for Windows
  http://www.cygwin.com
Open source site for SSH for Windows and Mac
http://www.openssh.com/windows.html

Sysinternals home page
http://www.sysinternals.com

Download site for Windows SSH freeware
http://www.chiark.greenend.org.uk/~sgtatham/putty

Help from IBM

IBM Support and downloads
ibm.com/support

IBM Global Services
ibm.com/services
Implementing the IBM Storwize V7000 V7.4
Implementing the IBM Storwize V7000 V7.4
Implementing the IBM Storwize V7000 V7.4

Discover the exciting new version of the IBM virtualization family

Become familiar with the leading-edge intuitive GUI

See how simple IBM virtualization really is

Continuing its commitment to developing and delivering industry-leading storage technologies, IBM introduces the IBM Storwize V7000 solution, an innovative storage offering that delivers essential storage efficiency technologies and exceptional ease of use and performance, all integrated into a compact, modular design that is offered at a competitive, midrange price.

The IBM Storwize V7000 solution incorporates some of the top IBM technologies typically found only in enterprise-class storage systems, raising the standard for storage efficiency in midrange disk systems. This cutting-edge storage system extends the comprehensive storage portfolio from IBM and can help change the way organizations address the ongoing information explosion.

This IBM Redbooks publication introduces the features and functions of the IBM Storwize V7000 system through several examples. This book is aimed at pre-sales and post-sales technical support and marketing, storage administrators, and will help you understand the architecture of the Storwize V7000, how to implement it, and take advantage of the industry-leading functions and features.