Redbooks
ibm.com/redbooks

IBM PowerHA SystemMirror for i: Using IBM Storwize
(Volume 3 of 4)

David Granum
Edward Grygierczyk
Sabine Jordan
Peter Mayhew
David Painter

Power Systems

IBM

Redbooks

International Technical Support Organization

IBM PowerHA SystemMirror for i: Using IBM Storwize (Volume 3 of 4)

June 2016

SG24-8402-00

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

First Edition (June 2016)

This edition applies to Version 7, Release 2 of IBM PowerHA SystemMirror for i (product number 5770-HAS).

Contents

Notices

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

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

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

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

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

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

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

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

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

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

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

COPYRIGHT LICENSE:

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

Trademarks

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

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

DS8000®
Easy Tier®
FlashCopy®
HyperSwap®
IBM®
IBM Spectrum™
Power Systems™
PowerHA®
PowerVM®
Real-time Compression™
Redbooks®
Redbooks (logo)  ®
Storwize®
System i®
System Storage®
SystemMirror®
Tivoli®

The following terms are trademarks of other companies:

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 personalized notifications of new content
- Link to the latest Redbooks blogs and videos

Get the latest version of the Redbooks Mobile App

iOS **Download Now** Android

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!

ibm.com/Redbooks
About Redbooks → Business Partner Programs

THIS PAGE INTENTIONALLY LEFT BLANK

Preface

IBM® PowerHA® SystemMirror® for i is the IBM high-availability (HA), disk-based clustering solution for the IBM i operating system. When combined with IBM i clustering technology, PowerHA for i delivers a complete HA and disaster recovery (DR) solution for business applications running in an IBM i environment. You can use PowerHA for i to support HA capabilities with either native disk storage, IBM DS8000® storage servers, or IBM Storwize® storage servers.

Use this IBM Redbooks® publication to help you install, tailor, and configure IBM PowerHA SystemMirror for i with the IBM Storwize storage servers. This publication provides you with planning information to prepare for using the various PowerHA offerings for the IBM Storwize storage family. It also provides implementation and managing information. Finally, it provides guidance on troubleshooting these solutions and identifies the documentation that you must capture before calling support.

This book is part of a four-book volume set that gives you a complete understanding of PowerHA for i by using native disk storage, IBM DS8000 storage servers, or IBM Storwize storage servers. The following publications are part of this PowerHA for i volume set:

- *IBM PowerHA SystemMirror for i: Preparation (Volume 1 of 4)*, SG24-8400
- *IBM PowerHA SystemMirror for i: Using DS8000 (Volume 2 of 4)*, SG24-8403
- *IBM PowerHA SystemMirror for i: Using Geographic Mirroring (Volume 4 of 4)*, SG24-8401

Important: The information that is presented in this volume set is for technical consultants, technical support staff, IT architects, and IT specialists who are responsible for providing HA and support for IBM i solutions. If you are new to HA, you should first review the information that is presented in the first book of this volume set, *IBM PowerHA SystemMirror for i: Preparation (Volume 1 of 4)*, SG24-8400, to get a general understanding of clustering technology, independent auxiliary storage pools (IASPs), and the PowerHA architecture.

Authors

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

David Granum is an Advisory Software Engineer at IBM Rochester. He has been with the IBM i Global Support Center (iGSC) throughout his 18 years at IBM. His areas of expertise include PowerHA, performance, communications, and TCP/IP applications on the IBM i. He holds a Computer Science degree from South Dakota State University and a masters degree in Telecommunications from Saint Mary's University of Minnesota.

Edward Grygierczyk is a part of the IBM Lab Services team and is based in Poland. He joined IBM in 2008, working for IBM STG Lab Services as a Consultant in IBM Power Systems™. Before joining IBM (he has worked for IBM i since 1991), he built a highly successful IBM Business Partner company in Poland. Edward has a background in systems, availability, integration, and implementation. Edward is an author of IBM courses in IBM i area.

Sabine Jordan is a Consulting IT Specialist working in IBM Germany. She has worked as a Technical Specialist in the IBM i area for more than 20 years, specializing in HA. She has worked on IBM PowerHA SystemMirror for i implementations for both SAP and non-SAP environments by using geographic mirroring, IBM DS8000, and SAN Volume Controller remote copy services. Among these implementations, she created concepts for the design and implemented the entire project (cluster setup and application changes), in addition to performing customer education and testing. Sabine presents and delivers workshops (internally and externally) on IBM PowerHA SystemMirror for i and HA and DR.

Peter Mayhew is an IBM i Technical Services Consultant working for IT Solutions Group, an IBM Business Partner. He has been working with midrange systems for almost 40 years, beginning as an IBM hardware service representative on the 3x systems, AS/400 CISC, and early RISC iSeries models. Leaving IBM after 23 years, he has maintained his skills on the IBM i, specializing in hardware and operating system installations, upgrades, and migrations, and HA planning and implementation supporting customers throughout the United States.

David Painter is a part of the IBM Rochester Lab Services team and is based in the United Kingdom. He joined IBM in 2006, working for IBM STG Lab Services as a Senior IT Specialist in IBM Power Systems, HA, and external storage. Before joining IBM, he set up and built a highly successful IBM Business Partner company in the UK. David has a background in systems, availability, integration, and implementation. Today, David implements HA solutions across the world for clients that range from SMB accounts to large Enterprise accounts. David also regularly teaches courses in IBM i HA and external storage for IBM i administrators. He also conducts many availability workshops and holds numerous certifications.

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

Philip Albu, Jenny Dervin, Steven Finnes, Scott Helt, Ben Rabe, Curt Schemmel, James Tilbury
IBM Rochester, IBM i Development and Support teams

Christian Aasland, Laural Bauer, Selwyn Dickey
IBM Systems Lab Services

Tracy Eidenschink, Cindy Mestad, Michael K Meyers
IBM i Performance and Scalability Services (PaSS) Center

Debbie Landon and Ann Lund
International Technical Support Organization

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:

  http://twitter.com/ibmredbooks

- Look for us on LinkedIn:

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

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

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

- Stay current on recent Redbooks publications with RSS Feeds:

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

1

Introduction to using IBM Storwize with PowerHA solutions on IBM i

This chapter discusses the various IBM PowerHA SystemMirror for i solution options that are available by using the IBM Storwize family of products.

The following topics are described in this chapter:

For more information about Storwize replication options, see *IBM System Storage SAN Volume Controller and Storwize V7000 Replication Family Services*, SG24-7574.

1.1 Value proposition

Increasing business demands for application availability require more clients of any size to look for a solution that can help eliminate planned and unplanned downtimes for their IT services.

An unplanned outage can have severe implications if the duration of the outage or recovery time exceeds business expectations. These implications include unexpected loss of reputation, client loyalty, and revenue. Companies who did not effectively plan for the risk of an unplanned outage, never fully completed their installation of a high-availability (HA) solution, or did not have a tested recovery plan in place, are especially exposed to negative business effects.

The IBM PowerHA SystemMirror for i solution offers a complete end-to-end integrated clustering solution for HA and disaster recovery (DR). PowerHA provides a data and application resiliency solution that is an integrated extension of the IBM i operating system and storage management architecture. It also features the design objective of providing application HA through planned and unplanned outages.

Built with IBM Spectrum™ Virtualize software, the IBM Storwize family provides hybrid solutions with common functions, management, and flexibility. It includes built-in functions such as IBM Real-time Compression™ and IBM Easy Tier® technology optimizing flash and hard disk drives (HDDs), and Remote Copy and IBM FlashCopy® functions to deliver extraordinary levels of efficiency and high performance. Available in a wide range of storage systems, the Storwize family delivers sophisticated capabilities that are easy to deploy, and help control costs for growing businesses.

Storwize family systems come in a range of offerings to meet the differing needs of organizations, but they are all built on a common platform. Shared technologies and common management features mean that the Storwize family is the correct system to choose today for use tomorrow, as your storage requirements grow.

Designed to deliver the benefits of enterprise-class storage virtualization to large and small organizations alike, IBM System Storage® SAN Volume Controller provides a single point of control to support improved application availability and greater resource utilization.

Beginning with entry-level storage systems and extending through midrange block and unified storage systems, virtualization and enterprise-level systems, the Storwize family delivers innovative built-in capabilities that are ready to use from day one:

- Entry storage system

  IBM Storwize V3700 is an easy-to-use, efficient, and affordable entry storage system that is designed to address the growing data requirements and infrastructure consolidation needs of small and midsize businesses with sophisticated capabilities that are unusual for a system of this class.

- Midrange, highly flexible storage system

  IBM Storwize V5000 is a highly flexible, easy to use, and virtualized storage system that enables midsize organizations to overcome their storage challenges with advanced functions.

- Midrange, block and unified storage systems

  IBM Storwize V7000 and IBM Storwize V7000 Unified are highly scalable midrange, virtualized storage systems that are designed to consolidate workloads into a single system for simplicity of management, reduced cost, superb performance, and HA.

- Storage virtualization system

  SAN Volume Controller is a leading-edge storage virtualization system that enhances existing storage to help improve productivity and availability while it helps reduce cost.

For more information about the IBM Storwize family, see the IBM Storwize family website:

http://www.ibm.com/systems/storage/storwize/

1.2 Prerequisites

This section describes the prerequisites that are needed to implement an IBM PowerHA SystemMirror for i solution together with IBM Storwize. Before installing the IBM PowerHA SystemMirror for i licensed product (5770-HAS), check whether the following requirements are in place:

- IBM i 7.2 is installed on all system nodes (servers or LPARs) that are part of your HA or disaster-recovery solution.
- HA Switchable Resources (option 41 of 5770-SS1) is installed on all system nodes (server or LPARs) that are part of your HA or disaster-recovery solution.

  Note: Licensing for HA Switchable Resources is included with 5770-HAS PowerHA SystemMirror for i 7.2. You do not have to order it separately, but must be installed separately.

- Portable Application Solution Environment (5770-SS1 option 33) is installed on all system nodes.
- IBM Portable Utilities for i and OpenSSH, OpenSSL, zlib (5733-SC1 base and option 1) is installed on all system nodes.

There are three editions of PowerHA SystemMirror for i:

- Express Edition. The Express Edition is intended to be the foundation for a class of High Availability / Disaster Recovery (HA/DR) offerings that are based on restarting (IPL) the logical partition (LPAR) onto a backup system for HA operations.
- Standard Edition. The Standard Edition, which is 5770-HAS option 2, is targeted at a data center HA solution.
- Enterprise Edition. The Enterprise Edition, which is 5770-HAS option 1, adds support for a multi-site HA and DR solution. Standard Edition must also be installed along with this option.

Table 1-1 provides an overview of the functions that are available with the different editions of IBM PowerHA SystemMirror for i.

Table 1-1 IBM PowerHA SystemMirror for i editions

IBM i HA/DR clustering	Express Edition	Standard Edition	Enterprise Edition
Centralized cluster management		✓	✓
Cluster resource management		✓	✓
Centralized cluster configuration		✓	✓
Automated cluster validation		✓	✓

IBM i HA/DR clustering	Express Edition	Standard Edition	Enterprise Edition
Cluster admin domain		✓	✓
Cluster device domain		✓	✓
Integrated heartbeat		✓	✓
Application monitoring		✓	✓
IBM i event/error management		✓	✓
Automated planned failover		✓	✓
Managed unplanned failover		✓	✓
Centralized FlashCopy		✓	✓
LUN-level switching		✓	✓
Geomirror Sync mode		✓	✓
Geomirror Async mode			✓
Multi-Site HA/DR management			✓
Metro Mirror			✓
Global Mirror			✓
IBM HyperSwap®	✓	✓	✓

1.2.1 Preparing for an SSH connection between IBM i and Storwize system

Communication between PowerHA and the Storwize system is done by using SSH. You must create SSH key pairs and attach the SSH public key to a user on the Storwize system. The corresponding private key file is specified in the creation of the auxiliary storage pool (ASP) copy descriptions. The private key must be distributed to all nodes in the cluster and stored in the same directory on all nodes. Generation of the SSH key pair is done on IBM i from QSHELL.

Example 1-1 on page 5 shows the QSHELL commands that you use to generate the SSH keys and shows a list of the files that are created. The directory that is used in this example to store the SSH key does not exist on the system but was created before.

Example 1-1 Generate SSH keys on IBM i

```
> cd /QIBM/UserData/HASM/hads/.ssh/
   $
> ssh-keygen -t rsa -f id_rsa -N ''
  Generating public/private rsa key pair.
  Your identification has been saved in id_rsa.
  Your public key has been saved in id_rsa.pub.
  The key fingerprint is:
  76:49:e1:9e:04:0a:c5:e2:68:3a:d0:6b:0b:4b:e2:2e powerha@CTCIHA9V.rchland.ibm.com
  $
> ls -la
  total 64
  drwxrwsrwx    2 powerha  0              8192 Sep 21 17:02 .
  drwx--Sr-x    3 qsys     0              8192 Sep 21 15:19 ..
  -rw-------    1 powerha  0              1679 Sep 21 17:02 id_rsa
  -rw-r--r--    1 powerha  0               414 Sep 21 17:02 id_rsa.pub
  $
```

Note: The SSH key pair that is generated here and used by PowerHA is in OpenSSH key format. It cannot be used by PuTTY, as PuTTY expects SSH2 keys.

After generating the SSH key pair, you must import the `id_rsa.pub` file as a key into a user on the Storwize system. This user must have a minimum role of Copy Operator to perform the functions that are used by PowerHA. If you want to use LUN-level switching, make sure that the user has a minimum role of Administrator because PowerHA must change host attachments when switching the independent auxiliary storage pools (IASP) to the secondary system.

Figure 1-1 shows transferring the `id_rsa.pub` file to the PC and importing it to the Storwize user. Make sure to distribute the `id_rsa` file to all nodes in your cluster to the same directory. The user profile QHAUSRPRF must have at least *R data authority to the key file on each system.

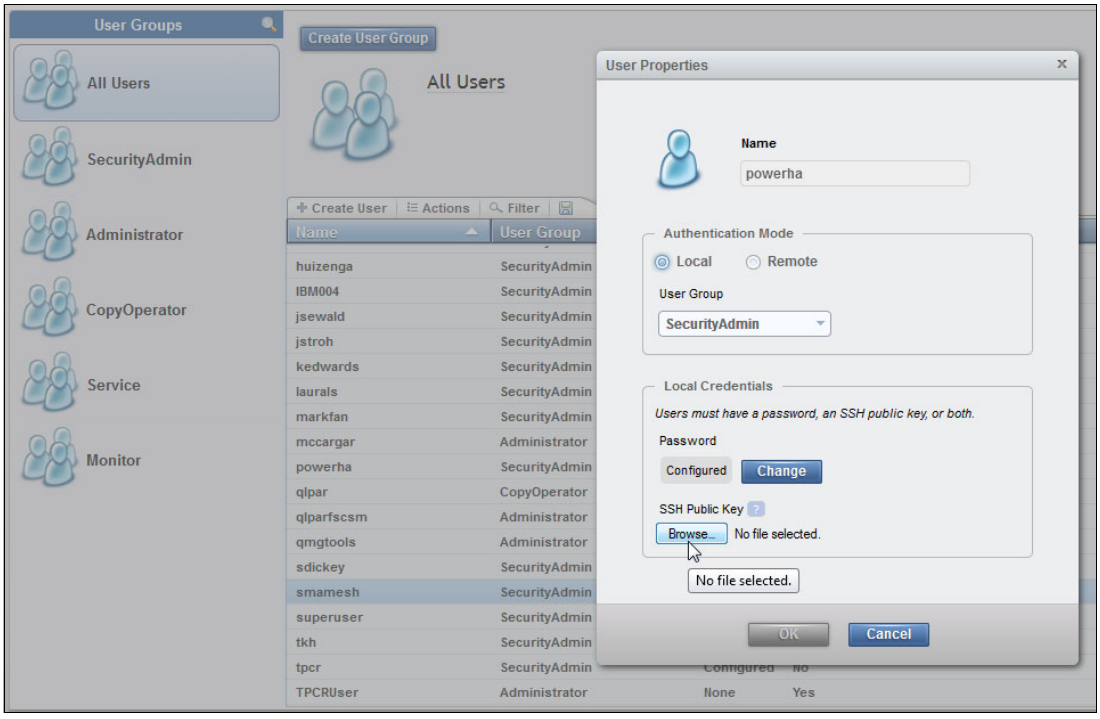

Figure 1-1 Import the SSH public key to a Storwize user

1.2.2 Initializing IBM i disk units on the backup nodes

Before setting up Copy Services on Storwize, you must initialize and format the read-protected DPHxxx disk units to become usable for the IASP on the IBM i *backup* nodes.

This task can be done in System Service Tools (SST) by choosing option 3 (Working with disk units), then selecting option 3 (Work with disk unit recovery), then selecting option 2 (Disk unit problem recovery procedure), and then selecting option 1 (Initialize and format disk unit). Failing to do so can result in IASP disk units not showing up after a switchover or failover to the secondary system.

1.3 LUN-level switching

LUN-level switching, as shown in Figure 1-2 on page 7, supports an automated IBM i HA solution by using a single copy of an IASP group that can be switched between two IBM i cluster nodes for a local HA solution.

LUN-level switching is supported for NPIV attachment and for native attachment of the IBM Storwize family. It is also supported for a SAN Volume Controller split cluster environment, which can make this an attractive solution, especially in heterogeneous environments where the SAN Volume Controller split cluster is used as the basis for a cross-platform, two-site, HA solution.

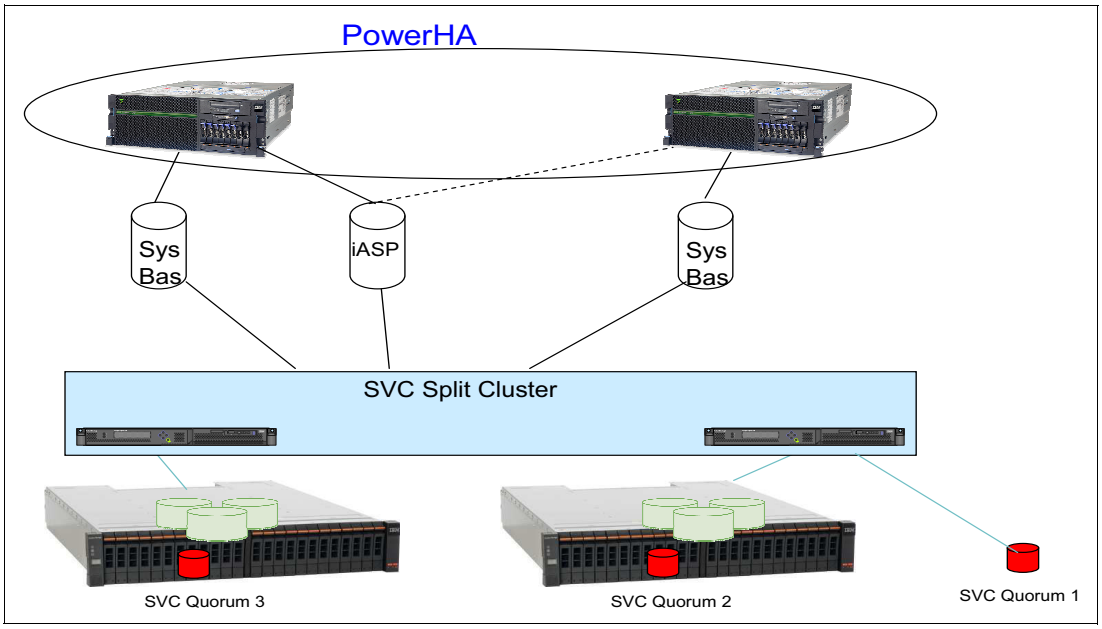

Figure 1-2 LUN-level switching

1.3.1 SAN Volume Controller split cluster

In a standard SAN Volume Controller configuration, all nodes physically are within the same rack. Beginning with Version 5.1, support was provided for split-cluster configurations where nodes within an I/O group can be physically separated from one another by up to 10 km. This capability allows nodes to be placed in separate failure domains, which provide protection against failures that affect a single failure domain.

The initial support for split cluster that was delivered in Version 5.1 contained the restriction that all communication between the SAN Volume Controller node ports cannot traverse Inter-Switch Links (ISLs). This limited the maximum supported distance between failure domains. Starting with SAN Volume Controller Version 6.3, the ISL restriction was removed, which allowed the distance between failure domains to be extended to 300 km. Additionally, in SAN Volume Controller Version 6.3, the maximum supported distance for non-ISL configurations was extended to 40 km.

Important: Make sure that you understand the influence of latency on your applications when using larger distances between the cluster nodes.

The SAN Volume Controller split-cluster configuration provides a continuous availability platform where host access is maintained if any single failure domain is lost. This availability is accomplished through the inherent active architecture of the SAN Volume Controller with the use of volume mirroring. During a failure, the SAN Volume Controller nodes and associated mirror copy of the data remain online and available to service all host I/O.

The split-cluster configuration uses the SAN Volume Controller volume mirroring function. Volume mirroring allows the creation of one volume with two copies of MDisk extents. The two data copies, if placed in different MDisk groups, allow volume mirroring to eliminate the effect to volume availability if one or more MDisks fails. The resynchronization between both copies is incremental and is started by the SAN Volume Controller automatically. A mirrored volume has the same functions and behavior as a standard volume. In the SAN Volume Controller software stack, volume mirroring is below the cache and copy services. Therefore, FlashCopy, Metro Mirror, and Global Mirror have no awareness that a volume is mirrored. All operations that can be run on non-mirrored volumes can also be run on mirrored volumes.

LUN-level switching is commonly used with a SAN Volume Controller split cluster setup. This setup provides storage HA through the SAN Volume Controller setup and server HA through PowerHA. Another common setup is a three-node cluster where local HA is achieved by LUN-level switching and DR is possible on the third node in the cluster by using Metro or Global Mirror. In this three-node setup, you often use a SAN Volume Controller split cluster on the primary site and a Storwize V7000 or V3700 at the secondary site. Within the Storwize family, remote copy operations, such as Global Mirror and Metro Mirror, are possible between the different members of the family.

Important: Although SAN Volume Controller split cluster is a supported environment with PowerHA, there is no automatic change of the SAN Volume Controller preferred node when a failover or switchover is done. Because of the latency that is involved when you are reading from the remote SAN Volume Controller node, make sure that the distance between the SAN Volume Controller nodes is close enough to meet your disk response time expectations. In addition, you must have separate Fibre Channel adapters for the system ASP and the IASP. Also, ensure that you set the preferred node correctly when you are creating the vDisks on the SAN Volume Controller because changing this setting later requires the attached IBM i LPAR to be powered down.

1.4 Storwize family Copy Services

The Storwize family offers a common platform and single point of control for regular provisioning and management of heterogeneous storage and for advanced functions, such as Copy Services, that are enabled by the Storwize family virtualization-layer between storage systems of different architectures or from different vendors.

The following Copy Services functions are available for the Storwize family:

- Metro Mirror for synchronous remote replication (see 1.4.1, “Metro Mirror” on page 9)
- Global Mirror for asynchronous remote replication (see 1.4.2, “Global Mirror” on page 13)
- FlashCopy for point-in-time volume copies (see 1.4.3, “FlashCopy” on page 15)

For more information about the Storwize family, see these publications:

- *Implementing the IBM System Storage SAN Volume Controller V7.4*, SG24-7933
- *IBM SAN Volume Controller 2145-DH8 Introduction and Implementation*, SG24-8229
- *Implementing the IBM Storwize V7000 and IBM Spectrum Virtualize V7.6*, SG24-7938
- *Implementing the IBM Storwize V5000*, SG24-8162
- *Implementing the IBM Storwize V3700*, SG24-8107

1.4.1 Metro Mirror

Metro Mirror is a synchronous remote copy relationship between two Storwize volumes (VDisks) of equal (virtual) size. When a remote copy relationship (either Metro Mirror or Global Mirror) is established, the preferred primary volume is designated as the *master* volume and the preferred secondary volume as the *auxiliary* volume. While the secondary volume is available, every host write I/O sent to the Metro Mirror primary volume is acknowledged back to the host only after it is committed to the write cache of the primary Storwize and the secondary Storwize system (Figure 1-3).

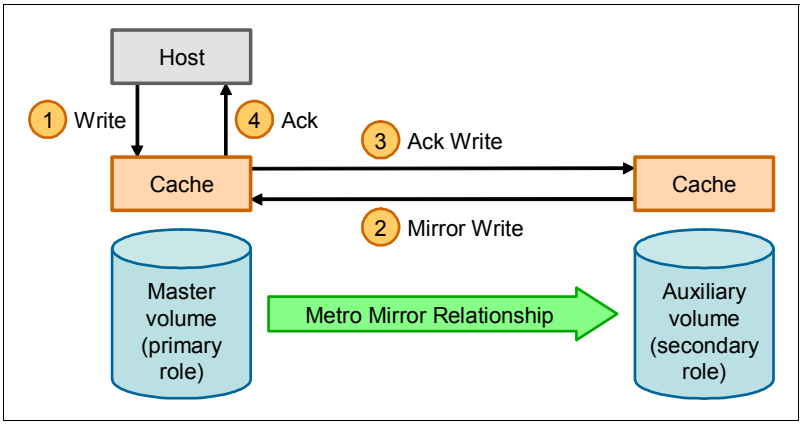

Figure 1-3 Metro Mirror Storwize family write I/O processing

The *role* of a master or auxiliary volume is either primary or secondary, depending on the direction or failover state of the current remote copy relationship. Up to 2048 remote copy relationships are supported in a two-node Storwize cluster.

With the Storwize family, establishing a Copy Services relationship is done in two phases of creating the relationship first before starting it in a second step. This is different from, for example, IBM System Storage DS8000 Copy Services, for which establishing a Copy Services relationship is done in a single step by creating the out-of-sync bitmaps and starting the relationship automatically at its creation.

When creating the Metro Mirror relationship, the user can specify whether the auxiliary volume is already in sync with the master volume, and the background copy process is then skipped. The in-sync option (path 1a in Figure 1-4 on page 10) is intended to be used when the volumes that were created with the format option should not be used for IBM i volumes because IBM i specially formats the volumes itself when they are configured (that is, added to an IBM i ASP). Hence, using the Storwize format option at volume creation and the in-sync option when creating a Metro Mirror relationship does not make sense.

Bandwidth thresholds

At initial synchronization, data is copied in data chunks, called *grains*, of 256 KB by a background copy process from the primary to secondary remote copy volume with a default bandwidth limit of 50 MBps between both Storwize clusters. This *partnership bandwidth limit* is evenly divided by the nodes in the cluster and is an attribute of the remote copy partnership between both Storwize systems. The bandwidth should be less than (when still accounting for host write I/O updates during synchronization) or equal to the available replication link bandwidth between both systems, when no relevant host write update are expected during synchronization.

Also, this overall cluster partnership bandwidth should be chosen deliberately to not exceed the capabilities of the primary and auxiliary storage systems to prevent performance impacts for the foreground host I/O.

Additionally, there is also a *relationship bandwidth limit* for the maximum background copy rate for each remote copy relationship hat defaults to 25 MBps and is an attribute of the Storwize cluster configuration.

Remote copy relationship states

A Metro Mirror or Global Mirror remote copy volume relationship can be in one of these states:

- Consistent stopped
- Inconsistent stopped
- Consistent synchronized
- Inconsistent copying
- Idling

Figure 1-4 shows an overview of these states as they apply to a connected remote copy relationship and the conditions that cause a state transition.

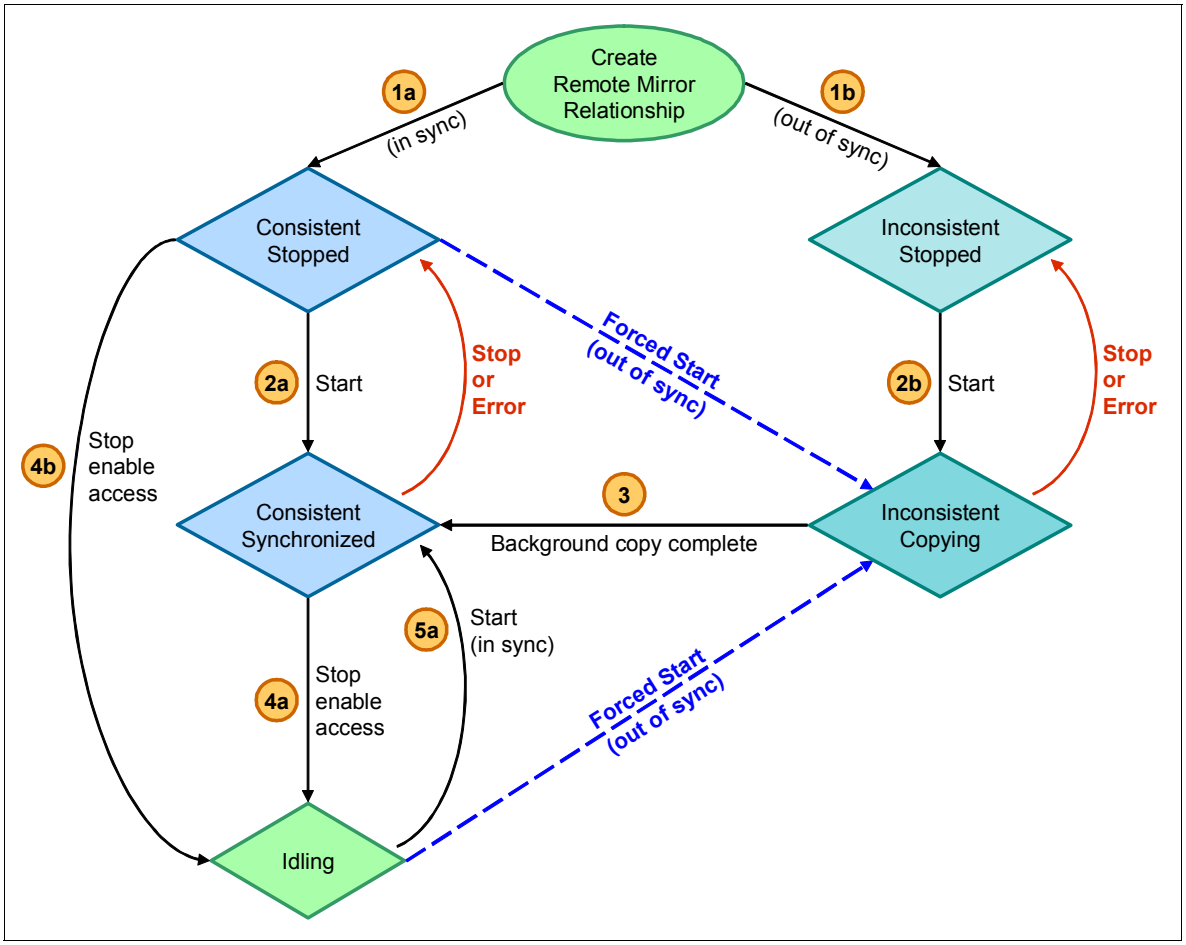

Figure 1-4 SAN Volume Controller/V7000 remote copy volume states and transitions

The remote copy states can be described as follows:

Inconsistent stopped State after creating a remote copy relationship (without using the in sync option) or after a failure condition that occurred while the relationship was in inconsistent copying state. Secondary volume data is not consistent with primary volume data and due to the risk of (undetected) data inconsistency should not be accessed by an application server.

Inconsistent copying State after starting an inconsistent stopped or idling relationship with changes to be synchronized with a background copy process running to copy data from the primary to the secondary volume. The primary is accessible for read and write I/O, but the secondary is offline (that is, not accessible for either read or write I/O) while the background copy is running and the relationship is not consistent.

Consistent synchronized State of an inconsistent copying relationship after completion of the background copy process or of a restarted consistent stopped relationship. The primary volume is accessible for read and write I/O, but the secondary volume is accessible only for read I/O. A switch of the remote copy direction does not change this state.

Stopping the relationship takes it to the consistent stopped state.

Stopping the relationship with the **-access** parameter takes it to the idling state.

Switching the relationship leaves it in the consistent synchronized state but reverses the primary and secondary roles.

Consistent stopped State after stopping a consistent synchronized relationship or after it encountered an error that forced a consistency freeze. The secondary contains a consistent image but might be out-of-date regarding the primary, which might have received write updates from the host after the relationship entered this state. Restarting on a non-synchronized relationship that has had changes requires the **-force** CLI command parameter.

Idling State after stopping a consistent synchronized relationship with enabling write access to the secondary volume. Both master and auxiliary volumes operate in the primary role so that both master and auxiliary volumes are accessible for read and write I/O.

Changes are tracked for both the master and auxiliary volumes so that when starting the remote copy relationship again in the wanted direction, which is specified by the required **-primary** CLI command parameter, only a partial synchronization for the changed grains is needed. Restarting on a non-synchronized relationship that has had changes requires the **-force** CLI command parameter.

In addition to these states that are valid for a connected remote copy relationship (that is, one where the primary system can communicate with the secondary system), there is also a disconnected state of remote copy relationships where the primary system can no longer communicate with the secondary system. When the clusters can communicate again, the relationships automatically become connected again.

If the relationship or consistency group becomes disconnected, the primary volumes transition to *inconsistent disconnected*. The master side moves to *idling disconnected*.

The Storwize system logs informational events like remote copy relationship changes, loss of synchronization, or remote cluster communication errors in an error log for which SNMP traps, email notification, or syslog server messages can be configured to trigger either automation or alert the user for manual intervention.

Consistency groups

Consistency is a concept of the storage system ensuring write I/O processing in the same order in which write updates are received by the host and maintaining this order even with Copy Services relationships. It applies to a single relationship, but can also be applied to a set of relationships spanning multiple volumes by using consistency groups.

For Metro Mirror and Global Mirror remote copy relationships, maintaining the correct write order processing requires that in case of an error event causing loss of replication for only a subset of remote copy relationships of application servers, remote write update processing for the non-affected remote copy relationships is automatically stopped to ensure application server data consistency at the secondary site.

For FlashCopy point-in-time volume copies, maintaining consistency and correct write order processing requires that write I/O to all volumes of an application server in a consistency group is temporarily put on hold until all FlashCopy volume relationships for the consistency group are started. The storage system depends on the concept of dependent writes being implemented in the application logic to ensure consistency across multiple volumes in a consistency group (for example, that a journal is updated with the intended database update before the database itself is updated). This application logic write dependency ensures that when a SCSI queue full status is set as part of the consistency group formation for a volume, further dependent application writes are put on hold by the application so that the storage system can proceed setting SCSI queue full status for all remaining volumes and ensure dependent write data consistency for all volumes in the consistency group. This write dependency concept still applies for IBM i with its single-level storage architecture, as IBM i SLIC storage management holds off all I/O to a disk unit in a SCSI queue full condition, but does not stop the I/O to other disk units that are still available for I/O operations.

Up to 127 FlashCopy consistency groups with up to 512 FlashCopy volume relationships in a consistency group are supported in a Storwize system. For Metro Mirror and Global Mirror, up to 256 remote mirror consistency groups are supported by no limit imposed for the number of either Metro Mirror or Global Mirror remote copy relationships other than the limit of 2048 volumes supported per I/O node pair.

PowerHA SystemMirror for i inherently uses consistency groups for Storwize FlashCopy relationships and requires them to be configured for Metro Mirror or Global Mirror relationships. Due to the IBM i single-level storage architecture, which stripes the data across all disk units of an ASP, consistency groups should be defined on an IASP group level.

Note: Stand-alone volume copy relationships and consistency groups share a common configuration and state model. All volume copy relationships in a consistency group that is not empty have the same state as the consistency group.

1.4.2 Global Mirror

Global Mirror is an asynchronous remote copy relationship between two Storwize volumes (VDisks) of equal (virtual) size. When a remote copy relationship (either Metro Mirror or Global Mirror) is established, the preferred primary volume is designated as the *master* volume and the preferred secondary volume as the *auxiliary* volume. Every host write I/O sent to the Global Mirror primary volume is acknowledged back to the host after it is committed to the write cache of both nodes for the corresponding I/O group of the primary Storwize system. Later (that is, asynchronously), this write update is sent by the primary Storwize to the secondary Storwize system (Figure 1-5). Global Mirror provides the capability to perform remote copy over long distances, up to the maximum supported round-trip latency of 80 ms, exceeding the performance-related limitations of synchronous remote copy without host write I/O performance impacts caused by remote replication delays.

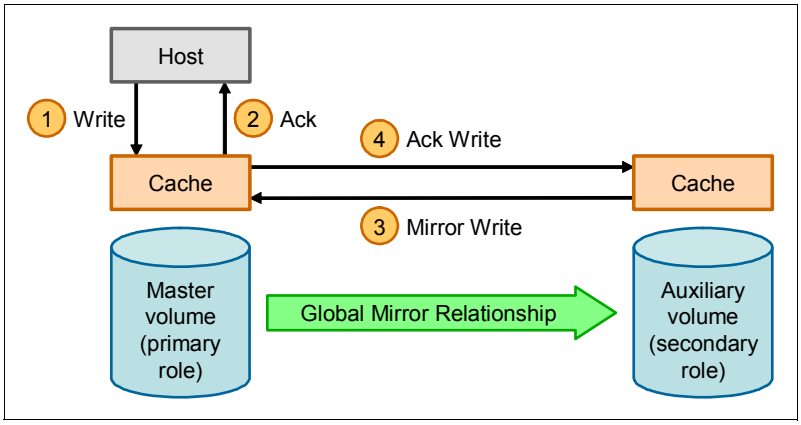

Figure 1-5 Global Mirror Storwize family write I/O processing

Though the data is sent asynchronously from the primary to the secondary Storwize system, the write ordering is maintained by sequence numbers that are assigned to acknowledged host write I/Os and with the secondary applying writes in order by their sequence number. Consistency of the data at the remote site is maintained always. However, during a failure condition, the data at the remote site might be missing recent updates that have not been sent or that were in-flight when a replication failure occurred, so using journaling to allow for proper crash consistent data recovery is of key importance.

Global Mirror volume relationship states and transitions are identical to those for Metro Mirror, as described in “Remote copy relationship states” on page 10.

A log file is used by the Storwize system for Global Mirror to maintain write ordering and help prevent host write I/O performance impacts when the host writes to a disk sector that is either in the process of being transmitted or due to bandwidth limits, is still waiting to be transmitted to the remote site. The Storwize system also uses shared sequence numbers to aggregate multiple concurrent (and dependent) write I/Os to minimize its Global Mirror processing impact.

Global Mirror link tolerance

Global Mirror uses a special *link tolerance* parameter that is defined at the cluster level that specifies the duration with a default of 300 seconds, for which inadequate intercluster link performance with write response times above 5 ms is tolerated. If this tolerated duration of degraded performance where the Storwize system must hold off writes to the primary volumes with the effect of synchronous replication-like degraded performance is exceeded, it stops the most busy active Global Mirror relationship consistency group to help protect the application server's write I/O performance and logs an event with error code 1920. The link tolerance can be disabled by the user setting its value to 0. However, this function provides no protection for the application server's write I/O performance in cases where there is congestion on either the replication link or the auxiliary storage system. Although you can use the link tolerance setting to define a period of accepted performance degradation, it is important to size correctly the remote copy replication bandwidth for the peak write I/O throughput and possible resync workload, and to help prevent longer production workload performance impacts.

The concept of consistency groups to ensure write-dependent data consistency applies to Global Mirror the same as previously described for Metro Mirror. Consistency groups are required to be configured for PowerHA SystemMirror for i.

Global Mirror with Change Volumes

Global Mirror with Change Volumes (GMCV) provides asynchronous replication based on point-in-time copies of data. It allows for effective replication over lower bandwidth networks and reduces any impact on production hosts. PowerHA supports GMCV.

Before the release of GMCV, Global Mirror ensured a consistent copy on the target side, but the recovery point objective (RPO) was not tunable, and usually within seconds. This solution required enough bandwidth to support the peak workload.

GMCV uses FlashCopy internally to ensure the consistent copy, but offers a tunable RPO, called a *cycling period*. GMCV might be appropriate when bandwidth is an issue, although if bandwidth cannot support the replication, the cycling period might need to be adjusted from seconds up to 24 hours.

Figure 1-6 shows a high-level conceptual view of GMCV. GMCV uses FlashCopy to maintain image consistency and to isolate host volumes from the replication process.

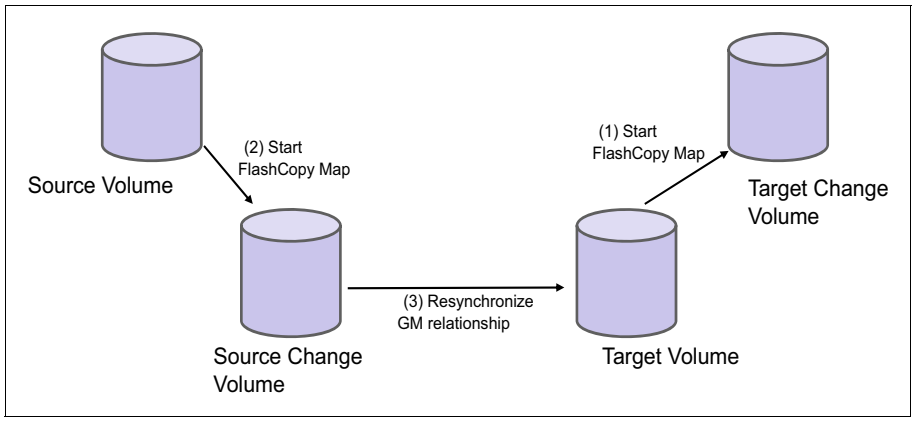

Figure 1-6 Global Mirror with Change Volumes

A FlashCopy mapping, called a *change volume*, exists on both the source and the target side. When replication begins, all data is sent from the source to the target, and then changes are tracked on the source change volume. At the end of each cycle period, the changes that are accumulated in the source change volume are sent to the target volume, which then stores that set of data as a consistent copy. If it takes longer than the cycling period to send the changes, then the next cycle period starts when the previous one finishes. The cycling period default is 300 seconds (5 minutes), but it can be adjusted by the user up to a maximum of 24 hours.

Change volumes hold point-in-time copies of 256 KB grains. If there is a change to any of the disk blocks in a given grain, that grain is copied to the change volume to preserve its contents. Change volumes are also maintained at the secondary site so that a consistent copy of the volume is always available even when the secondary volume is being updated.

GMCV also sends only one copy of a changed grain, which might be rewritten many times within the given cycle period. If the primary volume fails for any reason, GMCV ensures that the secondary volume holds the same data that the primary did at a given point. That period of data loss might be 5 minutes - 24 hours, but varies according to the design choices that you make.

Primary and change volumes are always in the same I/O group and the Change Volumes are always thin-provisioned. Change Volumes cannot be mapped to hosts and used for host I/O, and they cannot be used as a source for any other FlashCopy or Global Mirror operation.

With GMCV, a FlashCopy mapping, called a change volume, exists on both the source and target. When replication begins, all data is sent from the source to the target, and then changes are tracked on the source change volume. At the end of each cycle period, the changes that are accumulated in the source change volume are sent to the target change volume, which then stores that set of data as a consistent copy. If it takes longer than the cycling period to send the changes, then the next cycle period starts when the previous one finishes. The cycling period default is 300 seconds (5 minutes), but this can be adjusted by the user up to a maximum of 24 hours.

The RPO is determined by how long it takes for the cycle to complete. If the cycle completes within the configured cycle time, then the maximum RPO is 2 * cycle time. If all the changes cannot be written within the cycle period, then the maximum RPO is the sum of the previous and current cycle times. The cycle time should be configured in a way that it matches your RPO expectations (cycle time should be not more than half of your RPO) but that also matches the available bandwidth so that you do not regularly run into a situation where the source change volumes cannot be transferred to the target change volumes during the cycle time.

1.4.3 FlashCopy

The Storwize family FlashCopy function provides the capability to perform a point-in-time copy of one or more volumes (VDisks). In contrast to the remote copy functions of Metro Mirror and Global Mirror, which are intended primarily for DR and HA purposes, FlashCopy is typically used for online backup or creating a clone of a system or IASP for development, testing, reporting, or data mining purposes.

FlashCopy is supported only within the same Storwize system (though that Storwize system may consist of internal disks and other external storage systems). Up to 4096 FlashCopy relationships are supported per Storwize system, and up to 256 copies are supported per FlashCopy source volume. With Storwize V6.2 and later, a FlashCopy target volume can also be a non-active remote copy primary volume, which eases restores from a previous FlashCopy in a remote copy environment by using the FlashCopy reverse function.

PowerHA SystemMirror i supports these Storwize family FlashCopy functions, which are described in more detail below:

- FlashCopy no-copy and background copy
- Thin-provisioned (space-efficient) FlashCopy targets
- Incremental FlashCopy
- Reverse FlashCopy
- Multi-target FlashCopy (by using separate ASP copy descriptions for each target)

I/O indirection

The FlashCopy *indirection layer*, which is logically below the Storwize cache, acts as an I/O traffic director for active FlashCopy relationships. To preserve the point-in-time copy nature of a FlashCopy relationship, the host 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 whether the corresponding grain already is copied. Figure 1-7 and Figure 1-8 illustrate the different processing of read and write I/O for active FlashCopy relationships by the indirection layer.

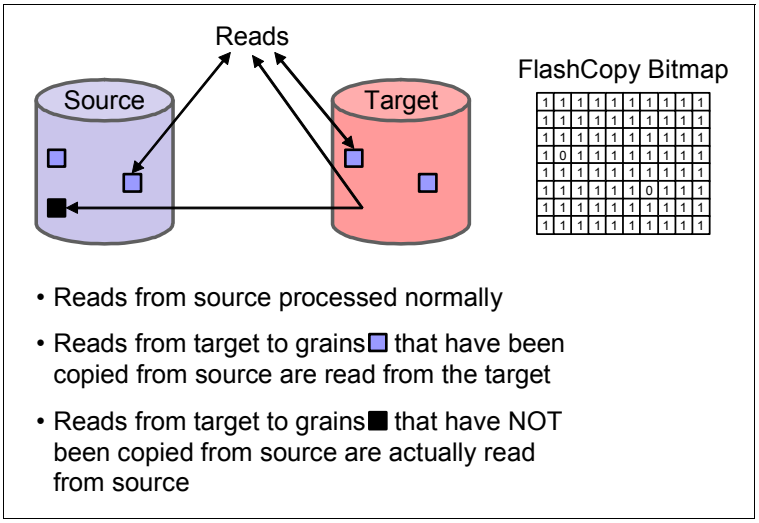

Figure 1-7 Storwize family FlashCopy read processing

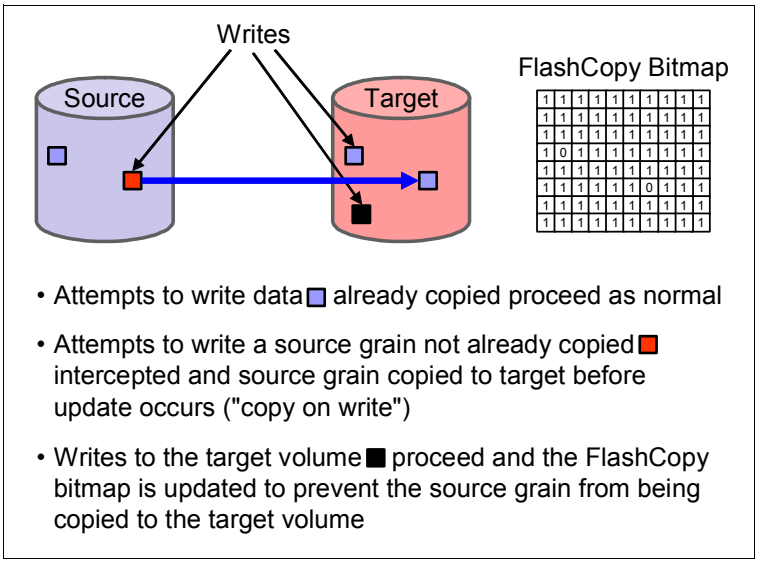

Figure 1-8 Storwize family FlashCopy write processing

Although a fixed grain size of 256 KB is used for remote mirror volume relationships for FlashCopy, you can choose from the default grain size of 256 KB or alternatively from the smaller grain size of 64 KB as the granularity for tracking and managing out-of-sync data of a FlashCopy relationship.The concept of consistency groups to ensure that dependent write data consistency across multiple volume copy relationships applies to FlashCopy (see “Consistency groups” on page 12).

Background copy

A FlashCopy relationship can either be a no-copy or a background copy relationship. With a *background copy* relationship, any grain from the source volume is copied to the target volume. By default (that is, if not specifying the autodelete option), the relationship is retained even after all grains are copied. For a *no-copy* relationship, only grains that are modified on the source after starting the FlashCopy relationship are copied from the source volume to the target volume before the source grain is allowed to be updated (*copy on write* processing), if the corresponding grain on the target volume is been updated already by the host accessing the target volumes.

An option for FlashCopy is creating an *incremental FlashCopy* relationship, which uses background copy to copy all of the data from the source to the target for the first FlashCopy and then only the changes that occurred since the previous FlashCopy for all subsequent FlashCopy copies being started for the relationship. When creating a FlashCopy relationship, you can specify a *copy rate* for the background copy process, which can either be 0 (meaning that a FlashCopy no-copy relationship without a background copy is established) or any value 1 - 100, which converts to the background copy throughputs (Table 1-2).

Table 1-2 FlashCopy background copy rates

Copy rate value	Data copied	256 KB grains/s	64 KB grains/s
1 - 10	128 KBps	0.5	2
11 - 20	256 KBps	1	4
21 - 30	512 KBps	2	8
31 - 40	1 MBps	4	16
41 - 50[a]	2 MBps	8	32
51 - 60	4 MBps	16	64
61 - 70	8 MBps	32	128
71 - 80	16 MBps	64	256
81 - 90	32 MBps	128	512
91 - 100	64 MBps	256	1024

a. Default value

A FlashCopy relationship is established on a Storwize system in three steps:

1. Creating a FlashCopy relationship

   This action triggers the internal creation of a FlashCopy out-of-sync bitmap that is used by the Storwize system for tracking the grains needing to be copied.

2. Preparing a FlashCopy relationship or consistency group

   This action achieves consistency for the volumes by destaging the source volume's modified data to disk, putting it in write-through mode, discarding the target volume's cache data, and rejecting any I/O to the target volume.

3. Starting a FlashCopy relationship or consistency group

   This action briefly pauses the I/O to the source volumes until all reads and writes below the Storwize cache layer complete and starts the actual FlashCopy relationship. The logical dependency between the source and target volume is established in the Storwize indirection layer.

FlashCopy relationship states

A FlashCopy volume relationship can be in any of the following states:

- Idle
- Copied
- Copying
- Stopped
- Stopping
- Suspended
- Preparing
- Prepared

The FlashCopy states are described here:

Idle or copied Mapping between source and target volume exists, but the source and the target behave as independent volumes.

Copying Background copy process is copying grains from the source to the target. Both the source and the target are available for read and write I/O, but the target depends on the source for grains that are not copied yet.

Stopped FlashCopy relationship was stopped either by the user or by an I/O error. The source volume is still assessable for read and write I/O, but the target volume is taken offline because data integrity is not provided. From a stopped state, the relationship can either be started again with the previous point-in-time image or lost or deleted if it is not needed anymore.

Stopping The relationship is in the process of transferring data to a dependent relationship. The source volume remains accessible for I/O, but the target volume remains online if the background copy process completed or is put offline if the background copy process is not completed while the relationship was in the copying state. Depending on whether the background copy completed, the relationship moves either to the idle/copied state or the stopped state.

Suspended This is the state of a relationship when access to metadata that is used by the copy process is lost. Both the source and the target are taken offline and the background copy is put on hold. When metadata becomes available again, the relationship returns to the copying state or the stopping state.

Preparing The source volume is placed in write-through mode and modified data of the source volume is destaged from the SAN Volume Controller/V7000 cache to create a consistent state of the source volume on disk in preparation for starting the relationship. Any read or write data that is associated with the target volume is discarded from the cache.

Prepared The relationship is ready to be started with the target volume in an offline state. Write performance for the source volume can be degraded, as it is in write-through mode.

Thin-provisioned FlashCopy

In addition to regular, full-provisioned volumes (VDisks), which at their creation have the full physical storage capacity allocated corresponding to their volume capacity, the Storwize family also supports *thin-provisioned* or *space-efficient* volumes, which are created with a virtual capacity that is reported to the host higher than the physical capacity pre-allocated to the volume. If the thin-provisioned volume is created with the **`autoexpand`** option, additional extents up to the virtual capacity of the volume are automatically allocated from the storage pool (MDisks group) as needed when the currently allocated physical capacity is exhausted.

Note: Using thin-provisioned Storwize volumes for IBM i is supported only for FlashCopy *target* volumes, not for volumes that are directly assigned to an IBM i host.

Using thin-provisioned volumes also does not make sense for remote copy secondary volumes, which become fully allocated at initial synchronization. Similarly, using thin-provisioned volumes for FlashCopy targets makes sense only for FlashCopy no-copy relationships that are used for a limited duration and have limited changes.

For optimal performance of thin-provisioned FlashCopy, the grain size for the thin-provisioned volume that is used as the FlashCopy target should match the grain size of the FlashCopy relationship. Additionally, for space-efficiency reasons, to help minimize physical storage allocations on the thin-provisioned target, consider also using the small grain size of 64 KB.

Multi-target and reverse FlashCopy

As previously mentioned, a single FlashCopy source volume supports up to 256 target volumes. Creating and maintaining multiple targets from a FlashCopy source volume from different times might be useful, for example, to have multiple points to restore from by using the reverse FlashCopy function (Figure 1-9).

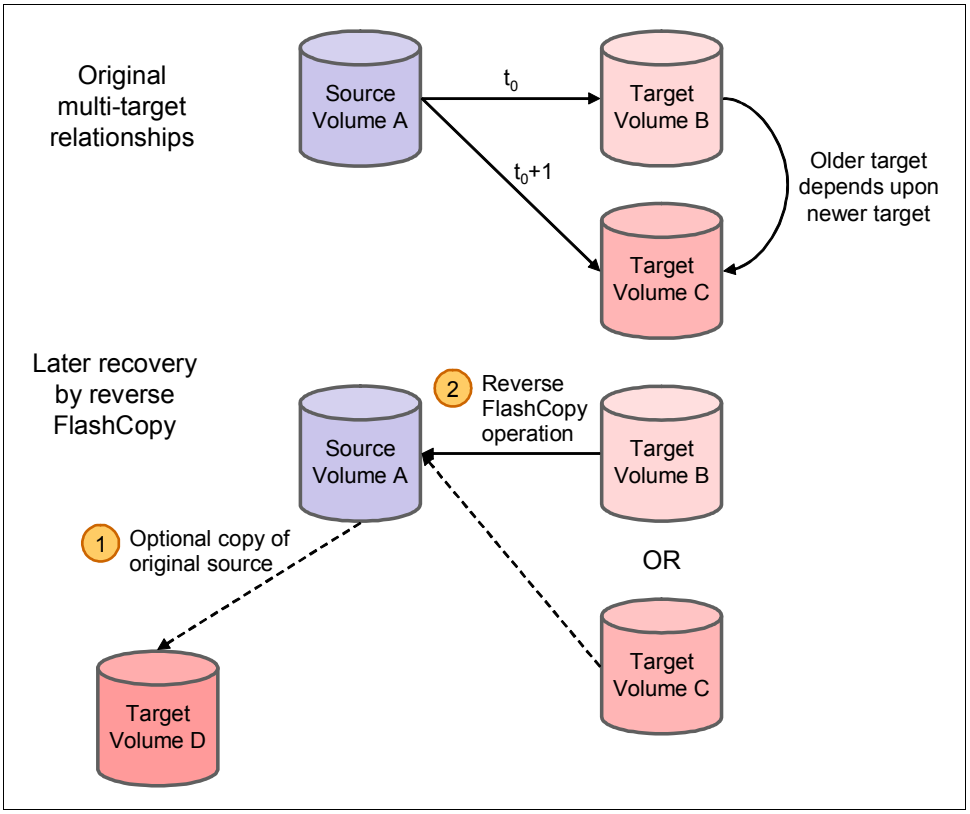

Figure 1-9 Storwize family multi-target and reverse FlashCopy

A key advantage of the Storwize reverse FlashCopy function is that it does not destroy the original source and target relationship, so any processes that use the target (such as tape backup jobs) can continue to run uninterrupted. It does not require a possible background copy process to have completed, and regardless of whether the initial FlashCopy relationship is incremental, the reverse FlashCopy function copies data only from the original target to the source for grains that were modified on the source or target.

Consistency groups cannot contain more than one FlashCopy relationship with the same target volume, so they must be reversed by creating a set of *reverse* FlashCopy relationships and adding them to the new reverse consistency group.

The Storwize family performs the *copy on write* processing for multi-target relationships in a way that data is not copied to all targets, but only once from the source volume to the newest target so that older targets refer to newer targets first before referring to the source. Newer targets together with the source can be regarded as composite source for older targets.

2

Planning

This chapter describes critical planning steps that are required to implement an IBM PowerHA SystemMirror for i solution in an IBM Storwize family environment. Before implementing such a solution, theses items should be considered:

- Sizing the backup system
- Synchronous and asynchronous mirroring considerations
- Thin provisioning
- Compression

The following topics are described in this chapter:

2.1 Sizing the backup system

Make sure that your backup system is sized so that it can run the production workload if you switch or fail over to the system. You should regularly test your HA solution, which is best done by making the current backup system the production system and then running your production workload on that system. There should not be a reason to do only a switchover test, start the applications, and then immediately switch back to the old production system.

Companies who regularly switch between the systems and consider them to be equal instead of having one designated production and one designated backup system are most likely the ones who know the following:

- Know exactly how to switch or fail over
- Have deep trust in their HA solution and do not waste time with lengthy decision processes when a system outage happens

If your backup system is sized smaller than your production system (whether it is on the processor or the storage configuration), then make sure that you understand what that means to your production workload. Giving everybody access to an undersized backup system in case of a production failure might lead to a situation where you easily overload the backup system, which makes it more or less unusable. Instead, you should have a plan in place that clearly establishes which workloads can be run on the smaller system and which workloads and users do not get access to that system.

The same situation is true if you are running development and test environments on your backup system. Make sure that you do not need HA for these environments. Points to consider include:

- Is it feasible to send home many external developers or consultants because your development system is not available?
- What is the business impact of not meeting project milestones because your development or test systems experience a longer outage?
- If your production system fails, can you run your production and development / test workload concurrently on the other system or do you have a plan in place on what workloads you do not run in this case?

Answering these questions can lead to the realization that your test and development systems require high availability (HA), whether permanently or during periods of heavy development activities.

2.2 Synchronous or asynchronous mirror

The Storwize family offers synchronous (Metro Mirror) and asynchronous (Global Mirror with or without Change Volumes) mirroring. With PowerHA supporting all three options, you can decide which setup best fits your requirements and available infrastructure.

Synchronous mirroring ensures that every write operation that reaches the source storage system is also present on the target storage system. Therefore, you can achieve an RPO of zero (no data loss because all transactions that completed on the source system also are available on the target system). Network latency and available bandwidth have a direct impact on I/O and application response times.

With asynchronous mirroring, the source storage system sends the I/O confirmation to the server before sending data to the target storage system. Therefore, network latency and available bandwidth do not impact I/O or application response times. There is no guarantee that the data is identical between source and target. However, the data on LUNs in one consistency group always has a consistent state on the target volumes. Depending on the available bandwidth in comparison to the write workload, be prepared to lose some data when the source system fails. Therefore, an RPO of zero cannot be met. Make sure that you understand what that means to your business:

- Can the business discover which data did not make it to the backup system?
- Does the business have a way of recovering this data automatically or manually?
- How much time and effort does it take to recover this data?
- Can you give access to the system to users while you are still recovering this data?

In addition, when using Global Mirror with Change Volumes (GMCV), make sure that you understand the consequences regarding failover times (recovery time objective (RTO)), as described in 5.2.1, “Special considerations when failing over in a Global Mirror setup” on page 109.

When considering Global Mirror, make sure that spikes in your write characteristics do not lead to a situation where the connection cannot catch up on data that is still waiting to be sent to the target site. Here is a simple approach to calculate this situation:

- Collect write data rates over a representative amount of time and use small collection intervals (for example, 1 minute) to make sure that you do not even out your results.
- Assume 1 byte is equal to 10 bits (to accommodate for protocol impact).
- For each period:
  - If data from the previous period was sent over (the write data rate in bits is smaller or equal to available bandwidth in bits), then consider only the current period write data rate.
  - If data from the previous period was not sent over (write data rate in bits is larger than available bandwidth in bits), then add the difference between the bandwidth and the write data rate from the previous period to the write data rate of the current period.

Using a spreadsheet tool, you can quickly calculate whether your RPO keeps growing infinitely or whether you can catch up after spikes in a reasonable amount of time. Make sure that you also include the specific amount of bandwidth that is required for intersystem heartbeat traffic into your calculation.

2.2.1 Determining write data rates

To determine the required bandwidth for synchronous or asynchronous replication, you must know the write data rate to the independent auxiliary storage pool (IASP). This rate can be calculated by sampling the collection services data on the production copy system.

Using Performance Data Investigator

This section assumes that Collection Services data has been running for at least the previous 5 days. Also, the RETPERIOD value in the Configure Performance Collection (**CFGPFRCOL**) command is set to a minimum of 120 hours. To view historical performance data, Performance Data Investigator (PDI) requires that all of the collections must be in a single library, and there can be no more than five collections graphed at the same time.

To set up your environment for the data to be viewed in the PDI tool, complete the following steps:

1. On the IBM i command line, enter **CFGPFRCOL** and press F4. Note the Collection Library value. This is typically QPFRDATA or QMPGDATA. This set of instructions assumes that the collection services library is QPFRDATA for the remainder of these steps.
2. Create a library to copy the management collection objects for this exercise. This library can be named anything. For this example, the library is called BWDATA:

   ```
   CRTLIB BWDATA
   ```

3. Get a list of all of the management collection objects in the performance data library by running the following command:

   ```
   WRKOBJPDM LIB(QPFRDATA) OBJ(*ALL) OBJTYPE(*MGTCOL)
   ```

   Note: The management collection objects naming convention begins with the letter “Q” followed by the Julian date and then the time stamp. For example, Q247000032 is the management collection object from the 247th day of the year, and the collection started at 32 seconds after midnight on that day.

4. From the Work with Objects using PDM window, specify option 3 (Copy) for up to five collections, which in most collection services configurations represent 5 days of system performance data. Press Enter.
5. From the Copy Objects window, specify a To library of BWDATA and press Enter or F19 to submit the copying of the objects to batch. After the process completes, the collections are placed in the new library.
6. In order for the PDI tool to graph the collections in a single graph, the performance data files must be created. For *each* management collection object, run the following command, where Qxxxxxxxxx is the name of the management collection object:

   ```
   CRTPFRDTA FROMMGTCOL(BWDATA/Qxxxxxxxxx)
   ```

7. Open a web browser on the workstation and specify the address or host name of the system for port 2001:

   ```
   http://<host_name_or_ip_address>:2001
   ```

8. When prompted, log in with a user that has *ALLOBJ special authority.
9. Once signed on, the IBM Navigator for i Welcome window opens (Figure 2-1).

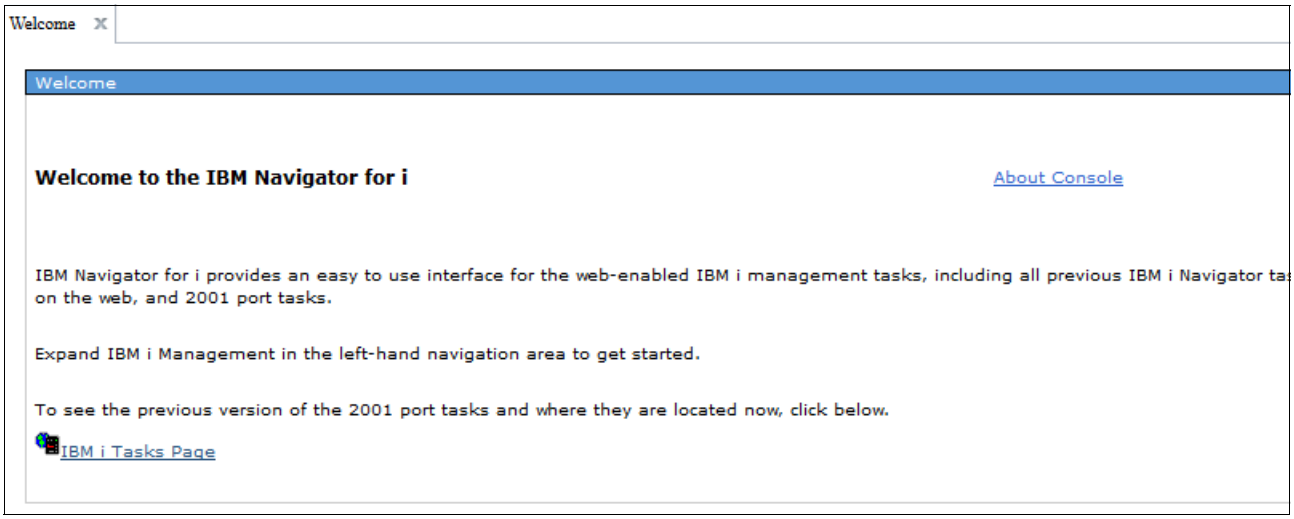

Figure 2-1 IBM Navigator for i Welcome window

10. To view the data in the new library in PDI, the collection table must be rebuilt. In the left pane of the Welcome window, expand **Performance** and click **Manage Collections**, as shown in Figure 2-2.

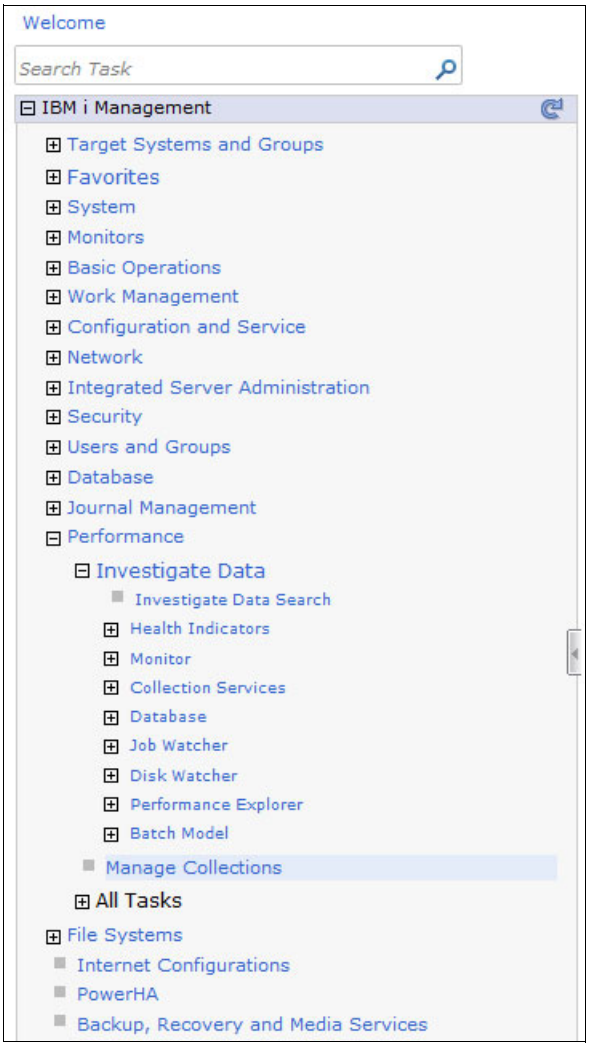

Figure 2-2 Select the PDI Manage Collections option

11. In the right pane, the list of collections is shown. Click **Actions** → **Maintain Collections** → **Rebuild Collection Table**, as shown in Figure 2-3.

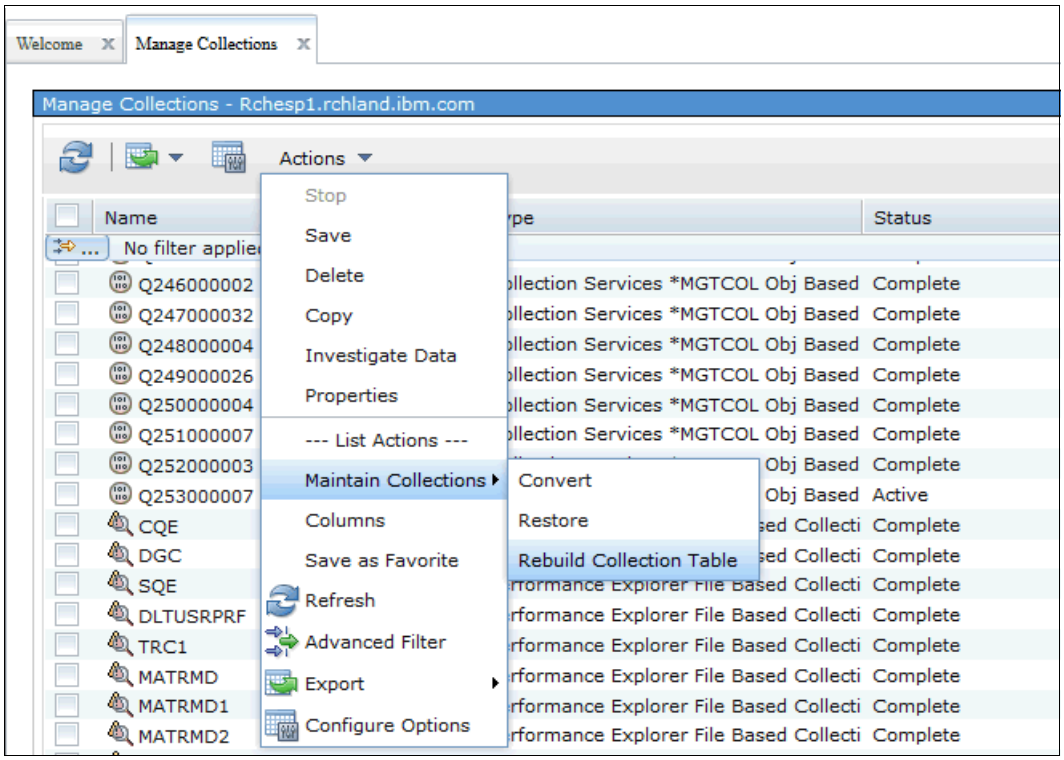

Figure 2-3 Select the PDI option to rebuild the collection table

12. The rebuild of the collection table takes a few moments to complete. After it is complete, the Manage Collections list shows the collections in the BWDATA library. In the left pane, expand **Performance** → **Investigate Data**, as shown in Figure 2-4.

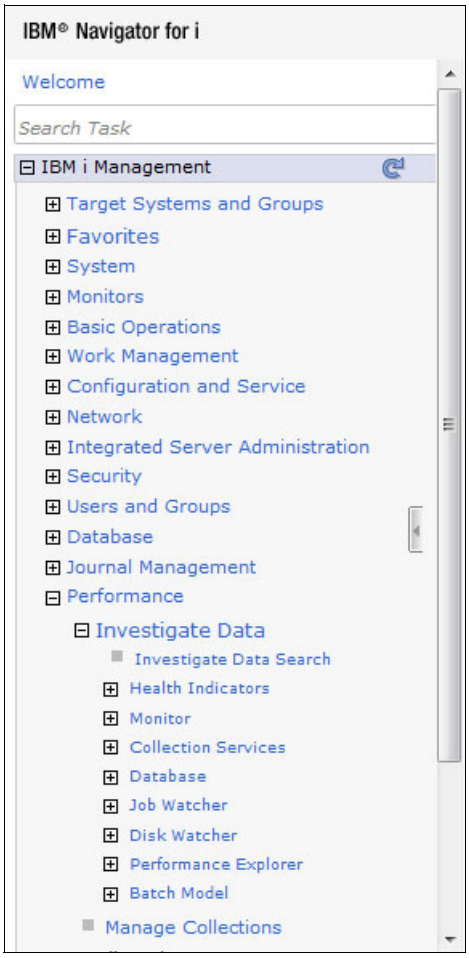

Figure 2-4 Select the PDI Investigate Data option

13. In the left pane, expand **Collection Services** → **Disk** and then click **Disk Throughput Overview for Disk Pools**, as shown in Figure 2-5.

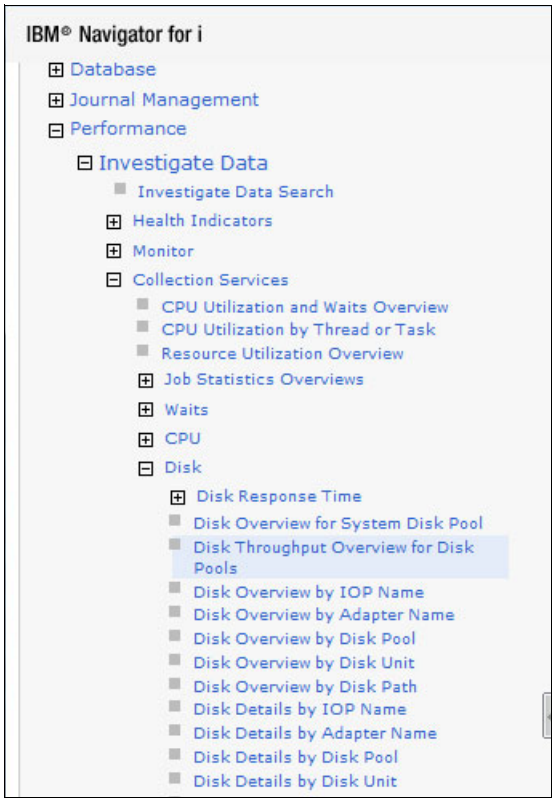

Figure 2-5 Select a PDI Collection Services Disk option

14. In the right pane, change the Collection Library to **BWDATA** and the Collection Name to **All**, then click **Display**, as shown in Figure 2-6.

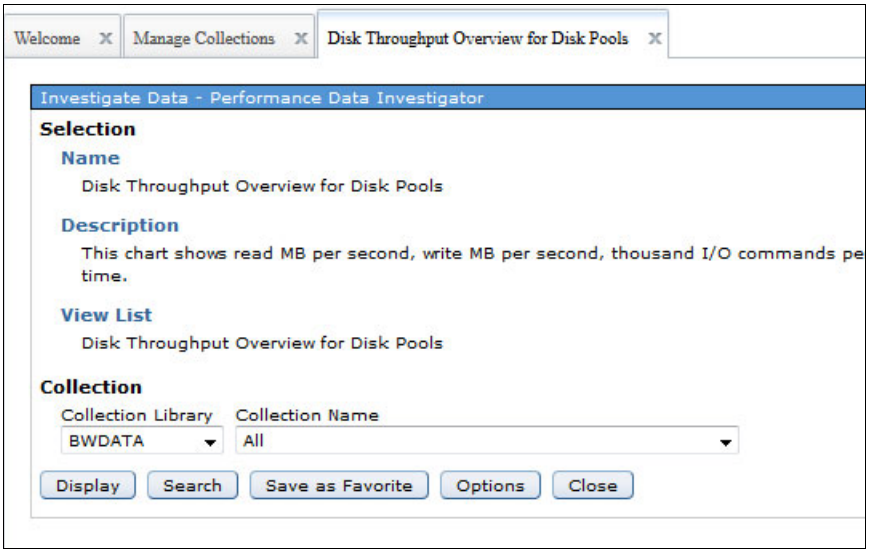

Figure 2-6 PDI Disk Throughput Overview for Disk Pools collection criteria

15. The resulting graph is displayed. The Full Zoom Out tool, as shown in Figure 2-7, can be used to view the graph during the number of collections.

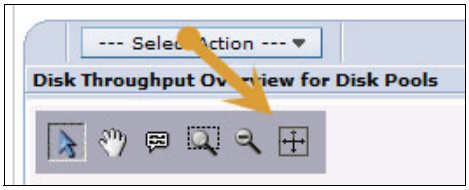

Figure 2-7 Select the PDI full zoom out tool

16. The resulting graph shows the IASP disk activity throughout the duration of the specified collections. The example that is shown in Figure 2-8 shows data for only a couple of days. However, it is considered a preferred practice that this entire process is performed several times to view several weeks of data, which helps ensure the amount of the system's IASP write activity.

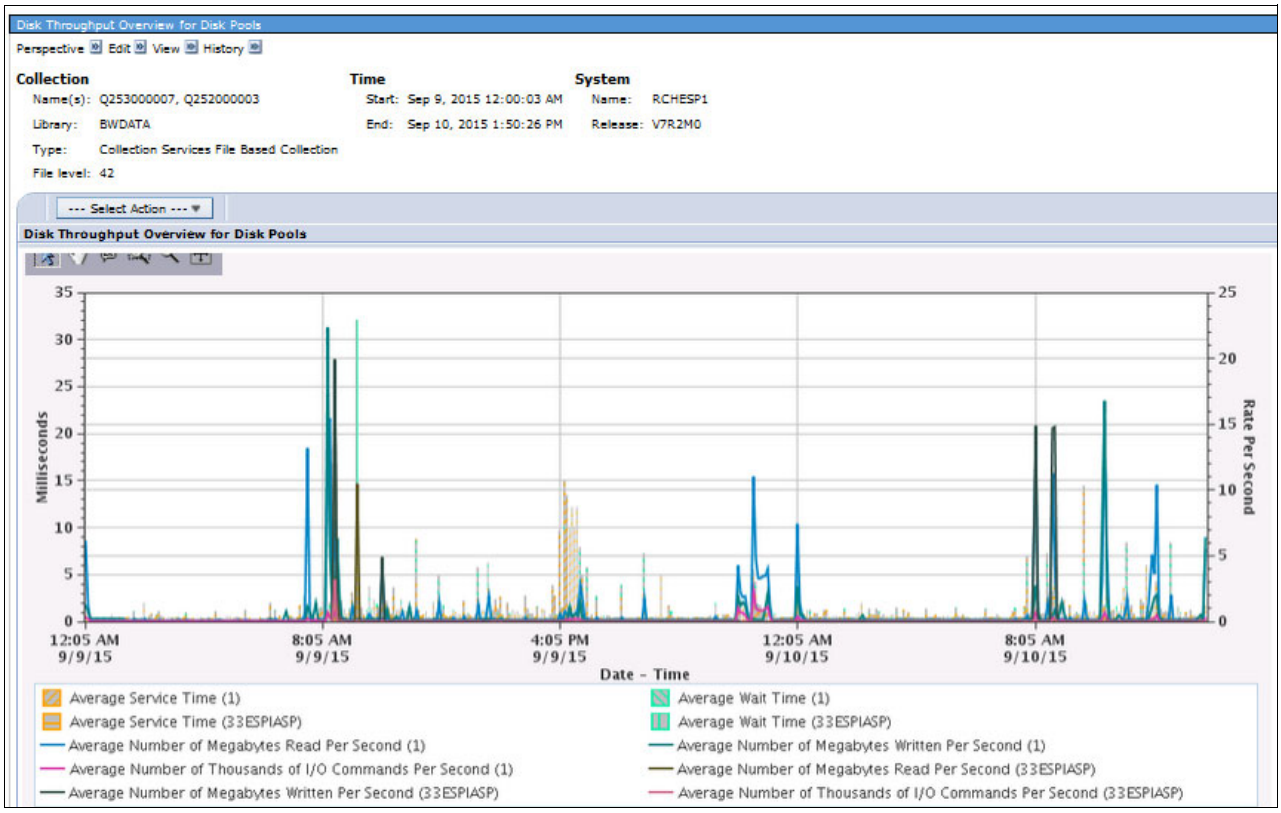

Figure 2-8 PDI Disk Throughput Overview graph

The number to focus on in the graph that is shown in Figure 2-8 is the Average number of Megabytes Written Per Second (for the IASP to be used in the PowerHA environment).

Note: If the migration to an IASP has not occurred, you still can use this tool to get a general idea of the number of megabytes that is written by looking at the Average Number of Megabytes Written Per Second for auxiliary storage pool (ASP) 1.

Review this graph for the peaks in the write rates to the IASP. Use the Zoom Region tool in the toolbar to view a specific range of time, which helps make the lines on the graph more readable. The write rates are needed to mirror sufficiently the writes to the preferred target system during that same time frame.

Throughout the days of data that should be analyzed, there are peaks and valleys in the write rates to the IASP. If sufficient bandwidth is not available during any of these times, this situation can result in performance issues on the preferred source system when using synchronous replication or extended RPOs when using asynchronous replication.

Note: The importance of the bandwidth analysis planning step cannot be stressed enough. Also, in addition to the normal runtime disk writes, a full or partial synchronization process consumes additional network bandwidth and system resources.

2.3 Thin provisioning

Volumes can be configured to be thin-provisioned or fully allocated. A thin-provisioned volume behaves as though application reads and writes were fully allocated. When a thin-provisioned volume is created, the user specifies two capacities: the real physical capacity that is allocated to the volume from the storage pool, and its virtual capacity that is available to the host. In a fully allocated volume, these two values are the same.

Therefore, the real capacity determines the quantity of MDisk extents that is initially allocated to the volume. The virtual capacity is the capacity of the volume that is reported to all other SAN Volume Controller components (for example, FlashCopy, cache, and remote copy) and to the host servers.

The real capacity is used to store the user data and the metadata for the thin-provisioned volume. The real capacity can be specified as an absolute value or a percentage of the virtual capacity. Thin-provisioned volumes can be used as volumes that are assigned to the host, by FlashCopy to implement thin-provisioned FlashCopy targets, and with the mirrored volumes feature.

When a thin-provisioned volume is initially created, a small amount of the real capacity is used for initial metadata. I/Os are written to grains of the thin volume that were not previously written, which causes grains of the real capacity to be used to store metadata and the user data. I/Os are written to grains that were previously written, which updates the grain where data was previously written.

The grain size is defined when the volume is created. The grain size can be 32 KB, 64 KB, 128 KB, or 256 KB. The default grain size is 256 KB, which is the recommended option. If you select 32 KB for the grain size, the volume size cannot exceed 260,000 GB. The grain size cannot be changed after the thin-provisioned volume is created. Generally, smaller grain sizes save space, but they require more metadata access, which can adversely affect performance. If you do not use the thin-provisioned volume as a FlashCopy source or target volume, use 256 KB to maximize performance. If you use the thin-provisioned volume as a FlashCopy source or target volume, specify the same grain size for the volume and for the FlashCopy function.

Thin-provisioned volumes store user data and metadata. Each grain of data requires metadata to be stored. Therefore, the I/O rates that are obtained from thin-provisioned volumes are less than the I/O rates that are obtained from fully allocated volumes.

The metadata storage impact is never greater than 0.1% of the user data. The impact is independent of the virtual capacity of the volume. If you are using thin-provisioned volumes in a FlashCopy map, use the same grain size as the map grain size for the best performance. If you are using the thin-provisioned volume directly with a host system, use a small grain size.

The real capacity of a thin volume can be changed if the volume is not in image mode. Increasing the real capacity allows a larger amount of data and metadata to be stored on the volume. Thin-provisioned volumes use the real capacity that is provided in ascending order as new data is written to the volume. If the user initially assigns too much real capacity to the volume, the real capacity can be reduced to free storage for other uses.

A thin-provisioned volume can be configured to autoexpand. This feature causes the SAN Volume Controller to automatically add a fixed amount of more real capacity to the thin volume as required. Therefore, autoexpand attempts to maintain a fixed amount of unused real capacity for the volume, which is known as the *contingency capacity*.

The contingency capacity is initially set to the real capacity that is assigned when the volume is created. If the user modifies the real capacity, the contingency capacity is reset to be the difference between the used capacity and real capacity.

A volume that is created without the autoexpand feature and has a zero contingency capacity, goes offline when the real capacity is used and it must expand.

Autoexpand does not cause the real capacity to grow much beyond the virtual capacity. The real capacity can be manually expanded to more than the maximum that is required by the current virtual capacity, and the contingency capacity is recalculated.

To support the auto expansion of thin-provisioned volumes, the storage pools from which they are allocated have a configurable capacity warning. When the used capacity of the pool exceeds the warning capacity, a warning event is logged. For example, if a warning of 80% is specified, the event is logged when 20% of the free capacity remains.

A thin-provisioned volume can be converted nondisruptively to a fully allocated volume (or vice versa) by using the volume mirroring function. For example, you can add a thin-provisioned copy to a fully allocated primary volume and then remove the fully allocated copy from the volume after it is 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.

Important: For IBM i, thin-provisioned volumes are recommended only for FlashCopy targets. The reason for that is that if a thin-provisioned volume reaches its capacity, an LPAR that uses this thin-provisioned volume fails and all data of that LPAR becomes unusable.

2.4 Compression

IBM Real-time Compression is available for the Storwize V7000 and the SAN Volume Controller. It improves storage efficiency by compressing data through supported real-time compression for block storage. With real-time compression, you save a considerable amount of space. Real-time Compression is used with primary data and in production environments with database and email systems. Real-time Compression operates as data is written to disk, which avoids the need to store uncompressed data when you are waiting for compression.

On the SAN Volume Controller, Real-time Compression is licensed by storage space (in terabytes); on the other Storwize V7000, it is licensed by enclosures.

Important: A system with compressed volumes dedicates processor and memory resources for running the Real-time Compression function. Clients must evaluate the impact to existing workloads on systems that are heavily used before the Real-time Compression function is enabled.

Storwize V7000 and SAN Volume Controller customers can evaluate the Real-time Compression capability at no cost. Clients have 45 days to evaluate the Real-time Compression function. At the end of this trial period, you must purchase the required licenses for Real-time Compression or disable the function. For more information about this program, see *IBM V7.2.0 Real-time Compression Evaluation User Guide for IBM System Storage SAN Volume Controller and Storwize V7000*, which is available at the following website:

http://www.ibm.com/support/docview.wss?uid=ssg1S7003988&myns=s028&mynp=OCST3FR7&mync=E

2.4.1 Using the IBM Comprestimator Utility

The IBM Comprestimator Utility is a host-based utility that can be used to provide an estimate for the achievable capacity savings by using Storwize compression and thin-provisioning. This utility is delivered as a binary executable file and is run from a host with access to the Storwize volumes to be analyzed. It performs read operations only.

The Comprestimator tool can analyze and provide estimated compression rate for data. By using this tool, clients can determine the data that can be a good candidate for Real-time Compression. You can download this tool here:

http://www.ibm.com/support/customercare/sas/f/comprestimator/home.html

Important: The IBM Comprestimator Utility does not run on IBM i. Therefore, it can be used only with Virtual I/O Server (VIOS) attached storage that uses vSCSI. It cannot be used for direct-attached storage or NPIV-attached storage. If you are using NPIV- or direct-attached storage, creating another copy of your data by using FlashCopy (provided there is enough space available) and attaching it to VIOS by using vSCSI is an option for using the Comprestimator Utility. If you already own Real-time Compression licenses, you can use FlashCopy to create a compressed copy and get information about achievable compression rates from that copy.

Example 2-1 shows the use of this utility on an IBM PowerVM® Virtual I/O Server to provide an estimate for the compression savings for an IBM i load source volume that is mapped to the VIOS.

Example 2-1 IBM Comprestimator Utility output

```
# ./comprestimator -d /dev/hdisk1
Version: 1.2.01 (Build u0004)
Start time: 30/07/2012 03:50:58
Device name: /dev/hdisk1
Device size: 100.0 GB
Number of processes: 10
Exhaustive: no
Sample# | Device      | Size(GB) | Compressed | Total       | Total       | Thin Provisioning | Compression | Compression
        | Name        |          | Size(GB)   | Savings(GB) | Savings(%)  | Savings(%)        | Savings(%)  | Accuracy Range(%)
--------+-------------+----------+------------+-------------+-------------+-------------------+-------------+------------------
3526    |/dev/hdisk1  | 100.0    | 19.5       | 80..5       | 80.6%       | 22.9%             | 74.9%       | 4.7%
```

In this example of an IBM i load source of a 100 GB volume on a Storwize system, the utility estimates thin-provisioning savings of 22.9% and a compression savings of 74.9%. This results in a compressed size of 19.5 GB, which is a total capacity savings of 80.6%. This result is the required physical storage capacity that is predicted to be reduced by 80.6% as a result of the combined savings from thin provisioning and compression.

Note: A Storwize compressed volume is always a thin-provisioned volume from an architectural perspective when it is created, and has a real size of 100%. Therefore, the reported compression savings are in addition to the savings that are provided by thin provisioning, which stores areas of zeros with minimal capacity.

Important: Make sure that you understand the impact that using compression on Storwize has on your applications.

3

Implementation

This chapter provides detailed step-by-step instructions about how to configure different scenarios by using IBM PowerHA SystemMirror for i in a Storwize environment.

The following topics are described in this chapter:

- 3.1, “Configuring PowerHA for i by using Storwize remote Copy Services” on page 36
- 3.2, “Configuring PowerHA for i by using Storwize FlashCopy” on page 46
- 3.3, “Configuring PowerHA for i by using LUN-level switching” on page 51
- 3.4, “Configuring PowerHA for i by using LUN-level switching with remote Copy Services” on page 55
- 3.5, “Full-system scenarios” on page 59, including full-system FlashCopy and full-system replication

3.1 Configuring PowerHA for i by using Storwize remote Copy Services

This section describes the configuration of IBM PowerHA SystemMirror for i independent auxiliary storage pool (IASP) replication by using Storwize Metro Mirror or Global Mirror.

Important: After you create the LUNs that make up the Metro Mirror or Global Mirror targets and create a host connection on the Storwize system for them, initialize these LUNs on the target LPAR by using SST before setting up Metro Mirror / Global Mirror on the Storwize system so that the LUNs correctly enlist on the target LPAR.

The following topics are described in this section:

3.1.1 Setting up an IBM i Storwize remote Copy Services environment

To set up Storwize remote Copy Services, complete the following steps. These steps assume that you have already created the Metro Mirror or Global Mirror target LUNs on the secondary Storwize system and that SAN zoning is set up in a way that the two Storwize systems can see each other.

1. Log in to the Storwize GUI of your primary Storwize system.
2. If you have not set up any remote copy connections between the two Storwize systems, you must create a partnership between these two systems. Click **Copy Services** → **Partnerships**.

Figure 3-1 Storwize - Partnerships

3. Click **Create Partnership**, choose whether the partnership uses a type Fibre Channel or IP, and enter the required parameters, as shown in Figure 3-2.

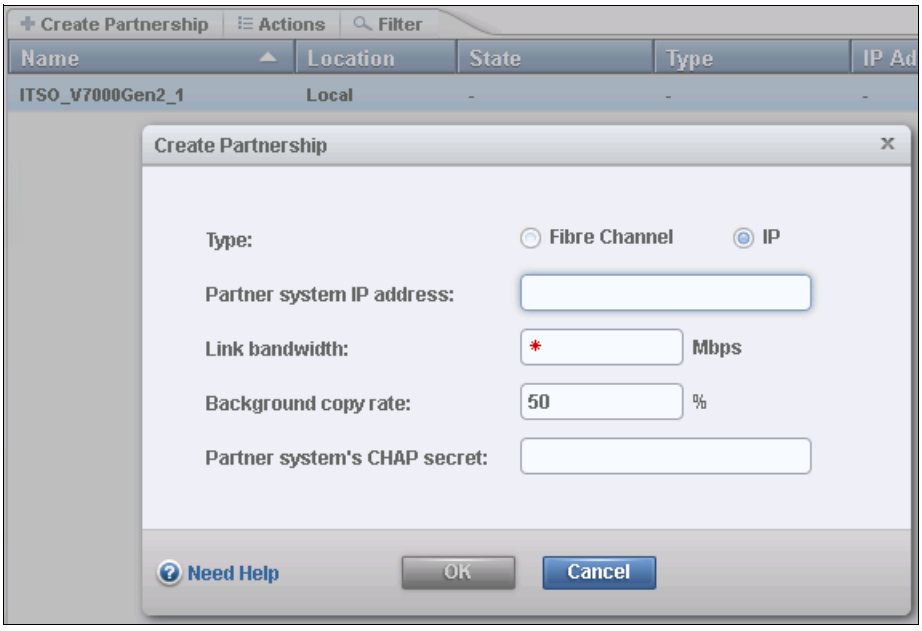

Figure 3-2 Storwize - Create Partnership

4. Create a remote copy consistency group on your source Storwize system by clicking **Copy Services** → **Remote Copy**, as shown in Figure 3-3.

Figure 3-3 Storwize - Remote Copy Services

5. Create a consistency group, as shown in Figure 3-4.

Figure 3-4 Remote Copy Services - Create Consistency Group

6. In the Create Consistency Group window, enter the name of the new Consistency Group, as shown in Figure 3-5. Click **Next**.

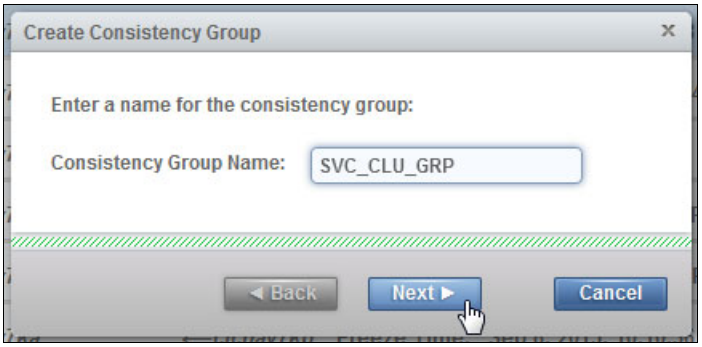

Figure 3-5 Remote Copy Services - name of consistency group

7. Specify where the target LUNs are. Select **On another system** and then select the correct system from the drop-down menu, as shown in Figure 3-6. Click **Next**.

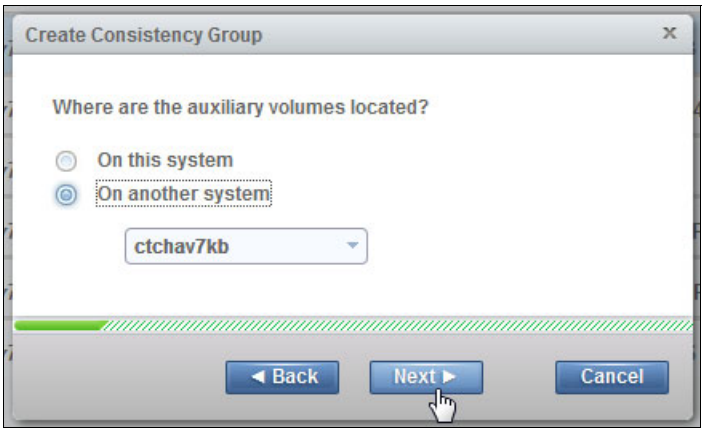

Figure 3-6 Remote Copy Services - select a remote system

8. Specify that you want to add relationships to the new consistency group, as shown in Figure 3-7. Click **Next**.

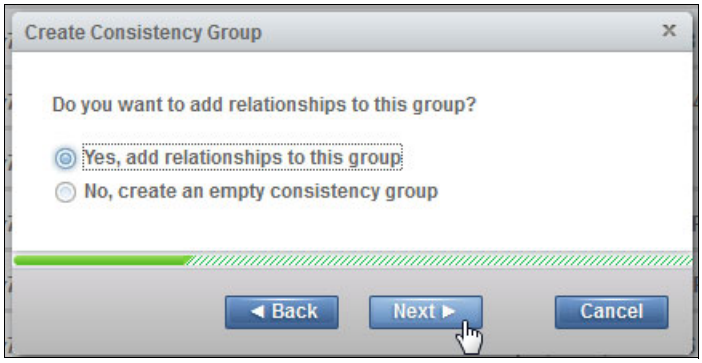

Figure 3-7 Remote Copy Services - add relationships

9. Select the type of relationship that you want to use. In this example, it is a Metro Mirror relationship, as shown in Figure 3-8. Click **Next**.

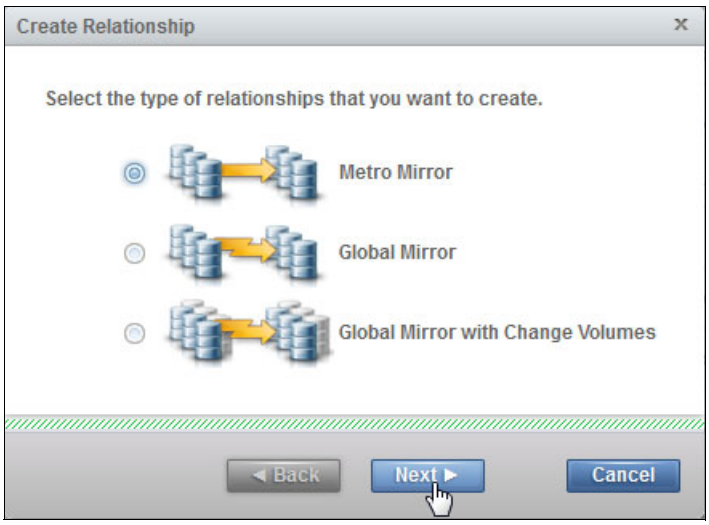

Figure 3-8 Remote Copy Services - Metro Mirror

10. If you already defined Metro Mirror relationships that are not yet part of a Consistency Group, you can add them to the Consistency Group in the next step. Otherwise, click **Next**, as shown in Figure 3-9.

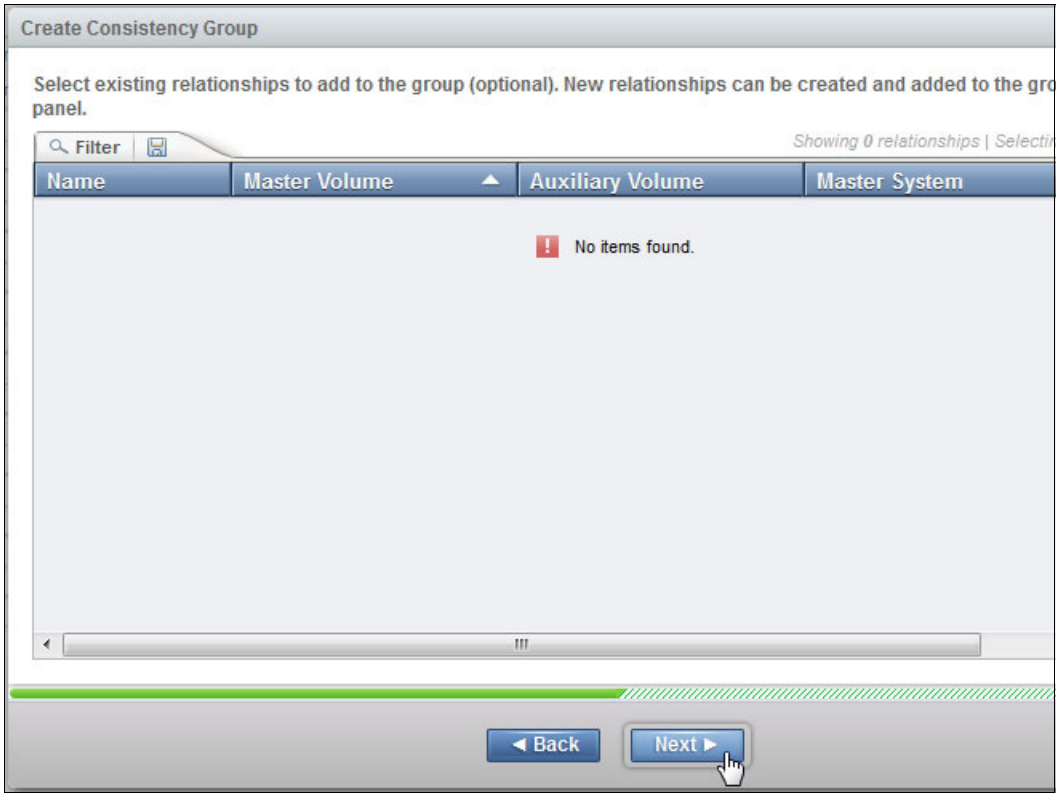

Figure 3-9 Remote Copy Services - add existing relationships

11. As there are no existing Metro Mirror relationships in this example, set them up as shown in Figure 3-10. Select the pairs of Master and Auxiliary LUNs from the drop-down menu and click **Add** for each individual pair. When finished, click **Next**.

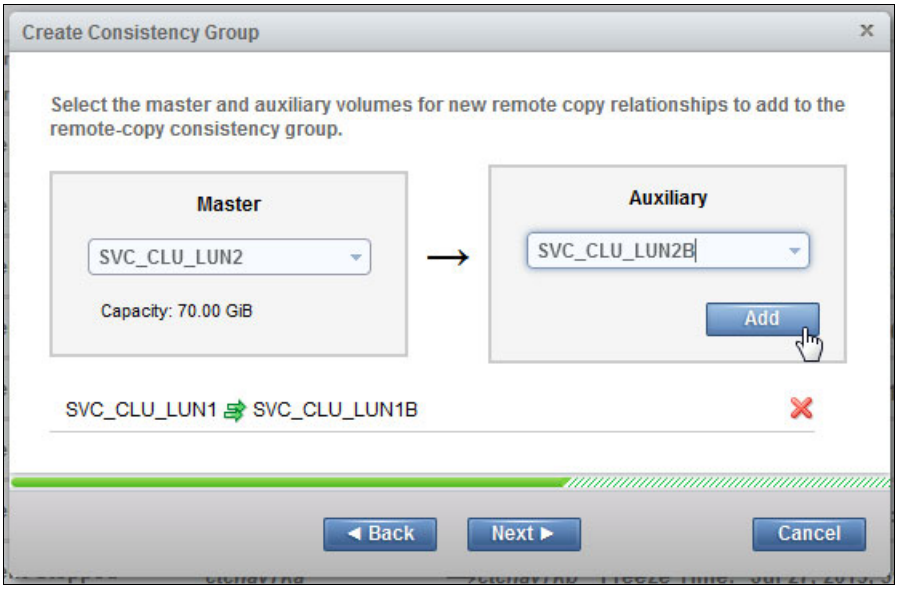

Figure 3-10 Remote Copy Services - add new relationships

12. Specify that the volumes in the new relationships are not synchronized, as shown in Figure 3-11. Click **Next**.

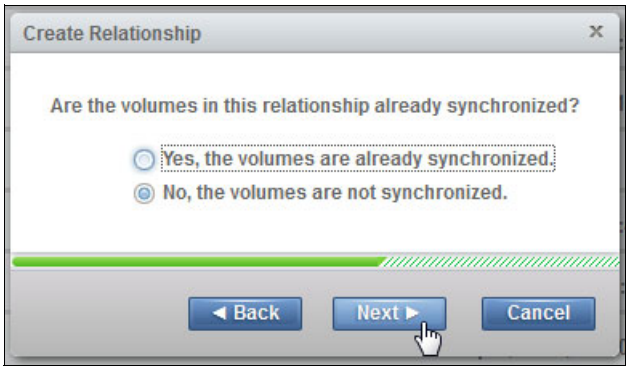

Figure 3-11 Remote Copy Services - volumes not synchronized

13. Specify that you want to start copying immediately, as shown in Figure 3-12.

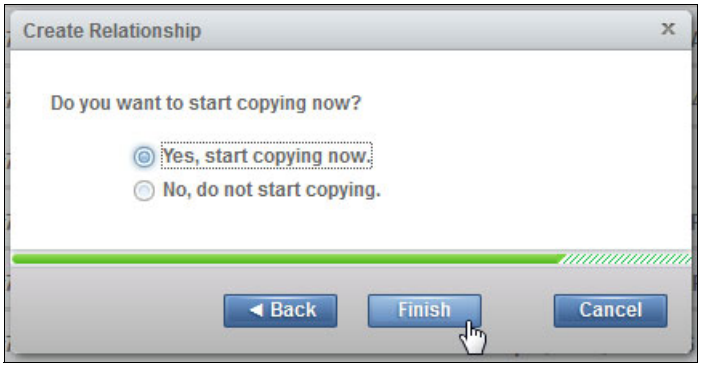

Figure 3-12 Remote Copy Services - start copying

14. After the initial synchronization is finished, the consistency group and the individual remote copy relationships are shown as **Consistent synchronized** (Figure 3-13).

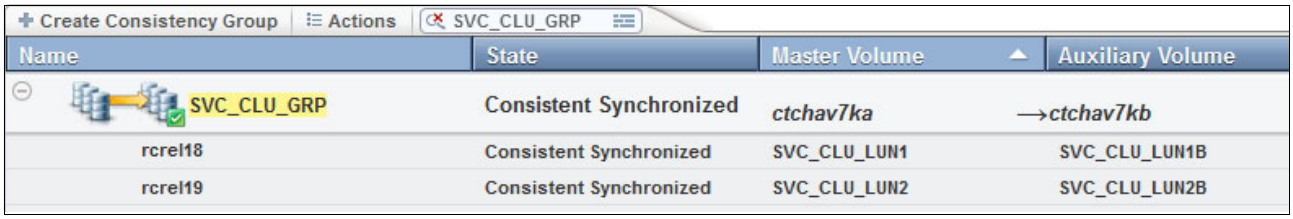

Figure 3-13 Remote Copy Services - Synchronized

For more information about how to configure remote copy services, see *IBM System Storage SAN Volume Controller and Storwize V7000 Replication Family Services*, SG24-7574.

3.1.2 Configuring PowerHA for i with Storwize remote Copy Services

This section provides the steps for setting up the PowerHA configuration for Storwize Metro Mirror.

These steps assume that you already created a two node cluster environment, an administrative domain (if needed), and an IASP on your production site. These steps also assume that the cluster is already started. If not, see *IBM PowerHA SystemMirror for i: Preparation (Volume 1 of 4)*, SG24-8400.

Table 3-1 provides an overview of the naming conventions and parameter values that are used for this specific setup.

Table 3-1 Settings for the Storwize remote Copy Services scenario

Parameter	Preferred primary	Preferred backup
System name	CTCIHA9V	CTCIHA9W
Cluster name	PWRHA_CLU	
Cluster resource group (CRG)	SVC_MM_CRG	
Device domain	PWRHA_DMN	
Administrative domain	PWRHA_CAD	
IASP name	IASP1	IASP1
CRG site name	SITE1	SITE2
Takeover IP	10.0.0.1	
Heartbeat cluster IP	10.10.10.1	10.10.10.2
Management access IP	9.5.167.53	9.5.167.54
ASP Copy description	SVC_MM_S	SVC_MM_T
SAN Volume Controller Session	SVV_MM	
SAN Volume Controller IP address	9.5.168.218	9.5.168.220
SSH key file location	/QIBM/UserData/HASM/hads/.ssh/id_rsa	
Volumes IDs[a]	15-18	0-3

a. Volumes IDs do not need to be the same on the source and target of remote copy relationships.

To configure a PowerHA configuration for Storwize Metro Mirror, complete the following steps:

1. You must have a device description for the IASP on the preferred target system. To do so, put the device description into the Cluster Administrative Domain as a Monitored Resource Entry.

   Alternatively, on the backup system, create a device description for the IASP with the same IASP name that you used on your preferred source system. If you choose a different RDB name than the IASP name on the preferred source system, make sure to specify the same value on the preferred target system. Otherwise, you can use the default value of `*GEN`.

   ```
   CRTDEVASP DEVD(IASP1) RSRCNAME(IASP1) RDB(*GEN)
   ```

2. Create the CRG by running the Create Cluster Resource Group (**CRTCRG**) command that is shown in Figure 3-14.

```
                    Create Cluster Resource Group (CRTCRG)

Type choices, press Enter.

Cluster  . . . . . . . . . . . . > PWRHA_CLU     Name
Cluster resource group . . . . . > SVC_MM_CRG    Name
Cluster resource group type  . . > *DEV          *DATA, *APP, *DEV, *PEER
CRG exit program . . . . . . . . > *NONE         Name, *NONE
  Library  . . . . . . . . . . .                 Name
User profile . . . . . . . . . . > *NONE          Name, *NONE

Recovery domain node list:
  Node identifier  . . . . . . . > CTCIHA9V      Name
  Node role  . . . . . . . . . . > *PRIMARY      *CRGTYPE, *PRIMARY...
  Backup sequence number . . . .   *LAST         1-127, *LAST
  Site name  . . . . . . . . . . > SITE1         Name, *NONE
  Data port IP address . . . . .   *NONE

  Node identifier  . . . . . . . > CTCIHA9W      Name
  Node role  . . . . . . . . . . > *BACKUP       *CRGTYPE, *PRIMARY...
  Backup sequence number . . . .   *LAST         1-127, *LAST
  Site name  . . . . . . . . . . > SITE2         Name, *NONE
  Data port IP address . . . . .   *NONE

Exit program format name . . . .   EXTP0100      EXTP0100, EXTP0101...
Exit program data  . . . . . . .   *NONE
Distribute info user queue . . .   *NONE         Name, *NONE
  Library  . . . . . . . . . . .                 Name
Job  . . . . . . . . . . . . . .   *JOBD          Name, *JOBD, *CRG

Configuration object list:
  Configuration object . . . . . > IASP1         Name, *NONE
  Configuration object type  . .   *DEVD        *DEVD, *CTLD, *LIND, *NWSD
  Configuration object online  . > *ONLINE       *OFFLINE, *ONLINE,
  Server takeover IP address . . > '10.0.0.10'

Text description . . . . . . . .   *BLANK
Failover message queue . . . . .   *NONE         Name, *NONE
  Library  . . . . . . . . . . .                 Name


F3=Exit   F4=Prompt   F5=Refresh   F12=Cancel   F13=How to use this display
F24=More keys
```

Figure 3-14 Create a cluster resource group for Storwize Metro Mirror

Note the following items for the **CRTCRG** command:

- The CRG type is *DEV because this CRG controls an IASP.
- The CRG exit program can specify an exit program that is activated whenever there is a change in the CRG status.

- The recovery domain node list defines which of the cluster nodes is the primary node and the order of the backup nodes in the cluster. Make sure to specify different site names for the different nodes.
- The configuration object list defines which objects are controlled by the CRG. In this example, this is IASP1, which has an object type of *DEVD.
- Configuration object online defines whether the IASP and the server takeover IP address should be varied on after a failover or a switchover.
- The failover message queue defines a message queue on the backup node for this specific CRG that receives a message when a failover is about to happen and can be configured to wait for an answer before proceeding with the failover. In this example, the failover message queue was configured on the cluster level and is set to *NONE in the CRG.

3. Start the CRG by running the following command:

```
STRCRG CLUSTER(PWRHA_CLU) CRG(SVC_MM_CRG)
```

4. Add the auxiliary storage pool (ASP) copy descriptions by running the Add SVC ASP Copy Description (**ADDSVCCPYD**) command that is shown in Figure 3-15. This task must be done for the IASP configuration on both the production site and on the backup site. Therefore, you must create two ASP copy descriptions.

```
                Add SVC ASP Copy Description (ADDSVCCPYD)

Type choices, press Enter.

ASP copy . . . . . . . . . . . . > SVC_MM_S      Name
ASP device . . . . . . . . . . . > IASP1         Name
Cluster resource group . . . . . > SVC_MM_CRG    Name, *NONE
Cluster resource group site  . . > SITE1         Name, *NONE
Node identifier  . . . . . . . . > *CRG          Name, *CRG, *NONE
Storage host:
  User name  . . . . . . . . . . > admin
  Secure Shell key file  . . . . > '/QIBM/UserData/HASM/hads/.ssh/id_rsa'

  Internet address . . . . . . . > '9.5.168.218'

Virtual disk range:
  Range start  . . . . . . . . . > 15            0-8191
  Range end  . . . . . . . . . . > 18            0-8191
             + for more values
  Host identifier  . . . . . . .  *ALL          0-512, *ALL
              + for more values
              + for more values
Device domain  . . . . . . . . .  *             Name, *
Recovery domain:
  Cluster node . . . . . . . . .  *NONE         Name, *NONE
  Host identifier  . . . . . . .                0-512
              + for more values
              + for more values
                                                                      Bottom
F3=Exit   F4=Prompt   F5=Refresh   F12=Cancel   F13=How to use this display
 F24=More keys
```

Figure 3-15 Add the SAN Volume Controller auxiliary storage pool copy description

Note the following items on the **ADDSVCCPYD** command:

- Unless you are on Storwize V6.2 or later, the user name that you specify here must be `admin`. This profile has no relationship to the user that is used on the Storwize system. The user is chosen based on the SSH key file pair only. Starting with Storwize V6.2, you can either specify `admin` as the user or the name of the Storwize user to which the SSH keyfile belongs.

  Important: The user name field is case-sensitive and must match what is configured on the SAN Volume Controller.

- The virtual disk range is the Storwize volume IDs of the disks in your IASP.
- Device domain specifies the name of the device domain to which you want to issue the command. Either enter a specific name or enter * to use the device domain that the node you are issuing the command on is in.
- Recovery domain specifies the connection information for each node in the CRG site recovery domain. This field is required only for switchable logical units. Therefore, the default value of `*NONE` is used here.

5. Start the ASP session by running the Start SVC Session (**STRSVCSSN**) command, as shown in Figure 3-16.

```
                      Start SVC Session (STRSVCSSN)

Type choices, press Enter.

Session  . . . . . . . . . . . . > SVC_MM        Name
Session type . . . . . . . . . . > *METROMIR     *METROMIR, *GLOBALMIR...
Device domain  . . . . . . . . .   *             Name, *
Cluster resource group . . . . .  > SVC_MM_CRG    Name
Switchover reverse replication     *YES          *YES, *NO
Failover reverse replication . .   *YES          *YES, *NO

                                                                      Bottom
F3=Exit   F4=Prompt   F5=Refresh   F12=Cancel   F13=How to use this display
F24=More keys
```

Figure 3-16 Start the SAN Volume Controller session

Note the following items on the **STRSVCSSN** command:

- For the session type, you must specify whether the underlying remote copy function that is used in the Storwize setup is Metro Mirror (specify session type as `*METROMIR`) or Global Mirror (specify session type as `*GLOBALMIR`).
- Device domain specifies the name of the device domain to which you want to issue the command. Either enter a specific name or enter * to use the device domain that the node you are issuing the command on is in.

– The two parameters, Switchover reverse replication and Failover reverse replication, determine whether reverse replication in the Storwize environment should be started automatically after a switchover or failover occurs or whether the Storwize session should stay in suspended mode.
– You do not have to specify explicitly the ASP copy descriptions in the **STRSVCSSN** command. They are chosen by PowerHA from the CRG name that is provided in the session description and the information about IASP name and site name in the CRG.

Your IBM PowerHA SystemMirror i Storwize remote copy configuration is now complete and your IASP is highly available for planned and unplanned site switches, as described in Chapter 5, “Switching and IBM FlashCopy operations” on page 93.

3.2 Configuring PowerHA for i by using Storwize FlashCopy

You can use PowerHA SystemMirror for i to create a FlashCopy of an IASP in a non-replicated IASP environment and from a replicated IASP with Metro Mirror or Global Mirror. You can take point-in-time copy of either the primary or the secondary site of a remote replication environment. The steps that are described in this section assume that you configured the FlashCopy target LUNs and have specified a host connection for them.

Important: After you create the FlashCopy target LUNs and create a host connection on the Storwize system for them, make sure to initialize these LUNs on the target LPAR by using SST before starting FlashCopy to these LUNs so that the LUNs can correctly enlist on the target LPAR.

The following steps describe the setup for an example scenario of configuring FlashCopy on a Metro Mirror secondary site based on the Storwize environment that is created in 3.1.2, “Configuring PowerHA for i with Storwize remote Copy Services” on page 42, taking the FlashCopy from the Metro Mirror secondary volumes. Table 3-2 provides an overview of the naming conventions and parameter values that are used for this specific setup.

Table 3-2 Settings for the Storwize FlashCopy from Metro Mirror target scenario

Parameter	Preferred primary	Preferred backup	FlashCopy
System name	CTCIHA9V	CTCIHA9W	CTCIHA9Y
Cluster name	PWRHA_CLU		
CRG	SVC_MM_CRG / SVC_GM_CRG		
Device domain	PWRHA_DMN		
Administrative domain	PWRHA_CAD		
IASP name and number	IASP1,	IASP1	IASP1
CRG site name	SITE1	SITE2	
Takeover IP	10.0.0.10		
Heartbeat cluster IP	10.10.10.1	10.10.10.2	10.10.10.4
Management access IP	9.5.167.53	9.5.167.54	9.5.167.60
ASP Copy Description	SVC_MM_S	SVC_MM_T	SVC_FLC_T

Parameter	Preferred primary	Preferred backup	FlashCopy
SAN Volume Controller Session for Metro Mirror	SVC_MM		
SAN Volume Controller Session for FlashCopy		SVC_FLC	
SAN Volume Controller IP address	9.5.168.218	9.5.168.220	9.5.168.220
SSH key file location	/QIBM/UserData/HASM/hads/.ssh/id_rsa		
Volumes IDs[a]	15-18	0-3	23-26

a. Volumes IDs do not need to be the same on the source and target of remote copy relationships.

To configure a Storwize FlashCopy from Metro Mirror target scenario, complete the following steps:

1. The IBM i partition CTCIHA9Y is not yet part of the cluster and device domain. Add it to the existing cluster environment by using the commands that are shown in Example 3-1.

   Example 3-1 Add the FlashCopy node into the cluster and device domain

   ```
   ADDCLUNODE CLUSTER(PWRHA_CLU) NODE(CTCIHA9Y ('10.10.10.4'))
   ADDDEVDMNE CLUSTER(PWRHA_CLU) DEVDMN(PWRHA_DMN) NODE(CTCIHA9Y)
   ```

2. If you have set up an Administrative Domain, add the additional cluster node to it by running the Add CAD Node Entry (**ADDCADNODE**) command.

   ```
   ADDCADNODE CLUSTER(PWRHA_CLU) ADMDMN(PWRHA_DMN) NODE(CTCIHA9Y)
   ```

3. You must have a device description for the IASP on the preferred target system, which can be achieved by putting the device description into the Cluster Administrative Domain as a Monitored Resource Entry.

   Alternatively, on the backup system, create a device description for the IASP with the same IASP name that you use on your preferred source system. If you choose a different RDB name than the IASP name on the preferred source system, make sure to specify the same value on the preferred target system; otherwise, you can use the default value of *GEN.

   ```
   CRTDEVASP DEVD(IASP1) RSRCNAME(IASP1) RDB(*GEN)
   ```

4. An ASP copy description for the Metro Mirror volumes on the secondary site exists, so there is no need to create a one for FlashCopy. You can use the existing ASP copy description from the Metro Mirror volumes to describe the FlashCopy source volumes. If you do not have an existing ASP copy description already from a remote copy relationship, run the **ADDSVCCPYD** command to add a corresponding Storwize copy description for the FlashCopy source volumes.

5. Create an ASP copy description for the FlashCopy target volumes by running the Add SVC ASP Copy Description (**ADDSVCCPYD**) command, as shown in Figure 3-17 on page 48, by specifying information about the IASP being replicated and the details of the virtual disks (VDisks) within the Storwize system that are used as the FlashCopy target volumes for the IASP.

   Note the following items on the **ADDSVCCPYD** command:

   – For the target copy description that is to be used for FlashCopy, the CRG and CRG site must be *NONE.

   – The node identifier must be set to the cluster node name that owns the target copy of the IASP.

- The parameters for storage host provide the information (user name, SSH key file, and IP address of the Storwize system) that is needed for PowerHA to talk to the Storwize system.
- Device domain specifies the name of the device domain to which you want to issue the command. Either enter a specific name or enter * to use the device domain that the node you are issuing the command on is in.
- Recovery domain specifies the connection information for each node in the CRG site recovery domain. This field is required only for switchable logical units. Therefore, the default value of *NONE is used here.

```
                  Add SVC ASP Copy Description (ADDSVCCPYD)

Type choices, press Enter.

ASP copy . . . . . .   SVC_FLC_T     Name
ASP device . . . . . .   IASP1         Name
Cluster resource group . . . . .   *NONE         Name, *NONE
Cluster resource group site  . .   *NONE         Name, *NONE
Node identifier  . . . . . .    CTCIHA9Y     Name, *CRG, *NONE
Storage host:
  User name  . . . . . .   admin
  Secure Shell key file  . . . .  '/QIBM/UserData/HASM/hads/.ssh/id_rsa'

  Internet address . . . . . .   '9.5.168.220'

Virtual disk range:
  Range start  . . . . . .   8            0-8191
  Range end  . . . . . .   11           0-8191
             + for more values
  Host identifier  . . . . . .  *ALL         0-512, *ALL
              + for more values
              + for more values
Device domain  . . . . . .  *            Name, *
Recovery domain:
  Cluster node . . . . . .  *NONE        Name, *NONE
  Host identifier  . . . . . .             0-512
              + for more values
              + for more values
                                                                  Bottom

F3=Exit   F4=Prompt   F5=Refresh   F12=Cancel   F13=How to use this display
 F24=More keys
```

Figure 3-17 Add Storwize ASP Copy Description to the FlashCopy target volumes

6. After you create the two ASP copy descriptions for the FlashCopy source and target volumes, create a FlashCopy for the IASP by running the Start SVC Session (**STRSVCSSN**) command, as shown in Figure 3-18 on page 49.

   PowerHA requires the FlashCopy relationships to be included in a consistency group. Therefore, by using the default value of *GEN, PowerHA creates the consistency group and the reverse consistency group on the Storwize system when a FlashCopy session is started. Alternatively, the user can specify the name for an existing FlashCopy consistency and reverse consistency group.

You can use run the **STRSVCSSN** command to create an incremental FlashCopy relationship and specify a FlashCopy background copy rate and grain size.

For more information about these parameters for the **STRSVCSSN** command, see 1.4.3, “FlashCopy” on page 15.

For information about managing a FlashCopy environment with Storwize, see 5.4, “Working with PowerHA FlashCopy” on page 111.

```
                      Start SVC Session (STRSVCSSN)

Type choices, press Enter.

Session  . . . . . . . . . . . . > SVC_FLC       Name
Session type . . . . . . . . . .> *FLASHCOPY    *METROMIR, *GLOBALMIR...
Device domain  . . . . . . . . .  *             Name, *
ASP copy:
  Preferred source . . . . . . .   SVC_MM_T     Name
  Preferred target . . . . . . .   SVC_FLC_T     Name
             + for more values
Incremental flash  . . . . . . .   *NO          *NO, *YES
Copy rate  . . . . . . . . . . .   0            0-100
Cleaning rate  . . . . . . . . .   0            0-100
Grain size . . . . . . . . . . .   256          256, 64
Consistency group  . . . . . . .  *GEN
Reverse consistency group  . . .  *GEN
                                                                      Bottom
F3=Exit   F4=Prompt   F5=Refresh   F12=Cancel   F13=How to use this display
F24=More keys
```

Figure 3-18 Start a Storwize FlashCopy session

Example 3-2 shows a CL script to be run from the FlashCopy target node for automating a FlashCopy backup. This automation includes quiescing the IASP on the production node CTCIHA9V before starting the FlashCopy session, varying on the IASP on the FlashCopy target node CTCIHA9X for doing the backup to tape before varying off the IASP on the FlashCopy node, and removing the FlashCopy session again.

Example 3-2 CHGASPACT running from the FlashCopy target node for quiescing an IASP

```
PGM
RUNRMTCMD  CMD('CHGASPACT ASPDEV(IASP1) +
             OPTION(*SUSPEND) SSPTIMO(30)') +
             RMTLOCNAME(CTCIHA9V *IP) RMTUSER(POWERHA) +
             RMTPWD(REDBOOK)
STRSVCSSN  SSN(SVC_FLC) TYPE(*FLASHCOPY) +
             ASPCPY((SVC_MM_T SVC_FLC_T))
RUNRMTCMD  CMD('CHGASPACT ASPDEV(IASP1) +
             OPTION(*RESUME)') RMTLOCNAME(CTCIHA9V +
             *IP) RMTUSER(POWERHA) RMTPWD(REDBOOK)
VRYCFG     CFGOBJ(IASP1) CFGTYPE(*DEV) STATUS(*ON)
/* INSERT CALL OF YOUR BACKUP PROGRAMS HERE */
VRYCFG     CFGOBJ(IASP1) CFGTYPE(*DEV) STATUS(*OFF)
ENDSVCSSN  SSN(SVC_FLC)
ENDPGM
```

Displaying an auxiliary storage pool session for FlashCopy

By running the **DSPSVCSSN SSN(SVC_FLC)** command, you can display the newly created ASP session for FlashCopy, as shown in Figure 3-19.

Note: The ASP status for the IASP FlashCopy source node is always UNKNOWN, as the FlashCopy target node cannot determine the ASP state for the FlashCopy source node. For the FlashCopy target node, the ASP status shows AVAILABLE because the IASP is varied on at the target node CTCIHA9Y.

```
                          Display SVC Session                    CTCIHA9Y
                                                              09/27/11  17:40:26
Session . . . . . :   SVC_FLC
Type  . . . . . :   *FLASHCOPY

Incremental flash . . . . . :   *NO
Copy rate . . . . . :   0
Cleaning rate . . . . . :   0
Grain size (KB) . . . . . :   256
Consistency group . . . . . :   fccstgrp0
Reverse consistency group . . . . . :   fccstgrp1
Storage cluster name  . . . . . :   ctchav7kb
                                                                       Bottom
                             Copy Descriptions

ASP         ASP copy              ASP         Replication
device      name        Role      Status         state     Node
IASP1       SVC_MM_T    SOURCE    AVAILABLE   ACTIVE       CTCIHA9W
            SVC_FLC_T   TARGET    ACTIVE                   CTCIHA9Y

                                                                       Bottom
                             Copy Descriptions

ASP                               Copy
device      Role      Node      progress  Storage state
IASP1       SOURCE    CTCIHA9W 1          Copying
            TARGET    CTCIHA9Y

Press Enter to continue

F3=Exit   F5=Refresh   F11=View 2   F12=Cancel
```

Figure 3-19 Display an auxiliary storage pool session for FlashCopy by running DSPSVCSSN

The PowerHA SystemMirror for i log file `/QIBM/UserData/HASM/hads/xsm.log` shows the actions that are performed by PowerHA for starting the ASP session, which includes the Storwize CL commands that were run for creating the FlashCopy consistency group and mappings. See Example 3-3 on page 51.

Example 3-3 /QIBM/UserData/HASM/hads/xsm.log

```
09/27/2011 17:31:34 950F29B47F165001 <INFO> : YaspPluginSession : start : Start of plugin session start for ASP session: SVC_FLC
09/27/2011 17:31:34 950F29B47F165001 <INFO> : YaspSession : deleteIAspObjects : Start of delete IASP Objects (Perm: 0) : 144
09/27/2011 17:31:34 950F29B47F165001 <INFO> : YaspSession : deleteIAspObjects : delete IASP objects succeeded
09/27/2011 17:31:34 950F29B47F165001 <INFO> : YaspSession : iapEnlistReject : Start of IASP enlist reject: 5
144
09/27/2011 17:31:34 950F29B47F165001 <INFO> : YaspSession : iapEnlistReject : iaspEnlistReject succeeded
09/27/2011 17:31:34 950F29B4CA3B3001 <INFO> : YaspSVCSession : startFlashCopy : Start of FlashCopy start for session SVC_FLC
09/27/2011 17:31:35 950F29B4CA3B3001 /QOpenSys/usr/bin/ssh -i "/QIBM/UserData/HASM/hads/.ssh/id_rsa" -o UserKnownHostsFile=/dev/null -o
StrictHostKeyChecking=no admin@9.5.168.220 "svctask mkfcconsistgrp"

FlashCopy Consistency Group, id [2], successfully created

09/27/2011 17:31:35 950F29B4CA3B3001 /QOpenSys/usr/bin/ssh -i "/QIBM/UserData/HASM/hads/.ssh/id_rsa" -o UserKnownHostsFile=/dev/null -o
StrictHostKeyChecking=no admin@9.5.168.220 "svcinfo lsfcconsistgrp -delim \"&\" -nohdr -filtervalue \"id=2\""

2&fccstgrp2&empty

09/27/2011 17:31:36 950F29B4CA3B3001 /QOpenSys/usr/bin/ssh -i "/QIBM/UserData/HASM/hads/.ssh/id_rsa" -o UserKnownHostsFile=/dev/null -o
StrictHostKeyChecking=no admin@9.5.168.220 "svctask mkfcmap -source 0 -target 8 -consistgrp fccstgrp2 -copyrate 0 -grainsize 256
-cleanrate 0"

FlashCopy Mapping, id [0], successfully created

09/27/2011 17:31:36 950F29B4CA3B3001 /QOpenSys/usr/bin/ssh -i "/QIBM/UserData/HASM/hads/.ssh/id_rsa" -o UserKnownHostsFile=/dev/null -o
StrictHostKeyChecking=no admin@9.5.168.220 "svctask mkfcmap -source 1 -target 9 -consistgrp fccstgrp2 -copyrate 0 -grainsize 256
-cleanrate 0"

FlashCopy Mapping, id [1], successfully created

09/27/2011 17:31:37 950F29B4CA3B3001 /QOpenSys/usr/bin/ssh -i "/QIBM/UserData/HASM/hads/.ssh/id_rsa" -o UserKnownHostsFile=/dev/null -o
StrictHostKeyChecking=no admin@9.5.168.220 "svctask mkfcmap -source 2 -target 10 -consistgrp fccstgrp2 -copyrate 0 -grainsize 256
-cleanrate 0"

FlashCopy Mapping, id [2], successfully created

09/27/2011 17:31:38 950F29B4CA3B3001 /QOpenSys/usr/bin/ssh -i "/QIBM/UserData/HASM/hads/.ssh/id_rsa" -o UserKnownHostsFile=/dev/null -o
StrictHostKeyChecking=no admin@9.5.168.220 "svctask mkfcmap -source 3 -target 11 -consistgrp fccstgrp2 -copyrate 0 -grainsize 256
-cleanrate 0"

FlashCopy Mapping, id [3], successfully created

09/27/2011 17:31:38 950F29B4CA3B3001 /QOpenSys/usr/bin/ssh -i "/QIBM/UserData/HASM/hads/.ssh/id_rsa" -o UserKnownHostsFile=/dev/null -o
StrictHostKeyChecking=no admin@9.5.168.220 "svctask startfcconsistgrp -prep fccstgrp2"

09/27/2011 17:31:39 950F29B4CA3B3001 <INFO> : YaspSVCSession : startFlashCopy : FlashCopy session SVC_FLC    started successfully
09/27/2011 17:31:39 950F29B4CA3B3001 <INFO> : yaspSVCActionPgm : doSessionAction : Session action completed with return code: 1
09/27/2011 17:31:39 950F29B47F165001 <INFO> : YaspSession : iapEnlistReject : Start of IASP enlist reject: 6
144
09/27/2011 17:31:39 950F29B47F165001 <INFO> : YaspSession : iapEnlistReject : iaspEnlistReject succeeded
09/27/2011 17:31:39 950F29B47F165001 <INFO> : YaspSession : waitForUnitsToEnlist : Start of wait for units to enlist for session: SVC_FLC
09/27/2011 17:31:47 950F29B47F165001 <INFO> : YaspSession : waitForUnitsToEnlist : Disk units for all ASPs in session: SVC_FLC    have
enlisted
09/27/2011 17:31:47 950F29B47F165001 <INFO> : YaspSession : resetMultipath : Start of reset multipath: 144
09/27/2011 17:31:47 950F29B47F165001 <INFO> : YaspSession : resetMultipath : Reset multipath succeeded
09/27/2011 17:31:47 950F29B47F165001 <INFO> : YaspPluginSession : start : Plugin session start for ASP Session: SVC_FLC    completed
successfully.
```

3.3 Configuring PowerHA for i by using LUN-level switching

This section provides an example of configuring SAN Volume Controller split-cluster and LUN-level switching.

In the example that is shown in this section, it is assumed that SAN Volume Controller split cluster was set up correctly. Also, ensure that SAN zoning is set up so that the IASP can be reached from the primary cluster node and from the secondary cluster node.

The steps in this section assume that you created a two-node cluster environment, an administrative domain (if needed), and an IASP on your production site. These steps also assume that the cluster is started. If not, see *IBM PowerHA SystemMirror for i: Preparation (Volume 1 of 4)*, SG24-8400.

Table 3-3 provides an overview of the naming conventions and parameter values that are used for this specific setup.

Table 3-3 Settings for the Storwize LUN-level switching scenario

Parameter	Preferred primary	Preferred backup
System name	CTCIHA9V	CTCIHA9W
Cluster name	PWRHA_CLU	
CRG	SVC_LL_CRG	
Device domain	PWRHA_DMN	
Administrative domain	PWRHA_CAD	
IASP name	IASP1	IASP1
CRG site name	SITE1	SITE1
Takeover IP	10.0.0.1	
Heartbeat cluster IP	10.10.10.1	10.10.10.2
Management access IP	9.5.167.53	9.5.167.54
ASP Copy description	SVC_LL	
SAN Volume Controller IP address	9.5.168.218	
SSH key file location	/QIBM/UserData/HASM/hads/.ssh/id_rsa	
Volumes IDs	15-18	

To configure PowerHA for i to use LUN-level switching, complete the following steps:

1. You must have a device description for the IASP on the preferred target system. Put the device description into the Cluster Administrative Domain as a Monitored Resource Entry.

   Alternatively, on the backup system, create a device description for the IASP with the same IASP name that you use on your preferred source system. If you choose a different RDB name than the IASP name on the preferred source system, make sure to specify the same value on the preferred target system; otherwise, you can use the default value of `*GEN`.

   ```
   CRTDEVASP DEVD(IASP1) RSRCNAME(IASP1) RDB(*GEN)
   ```

2. Create the CRG by running the Create Cluster Resource Group (**CRTCRG**) command, as shown in Figure 3-20.

```
                    Create Cluster Resource Group (CRTCRG)

Type choices, press Enter.

Cluster  . . . . . . . . . . . . > PWRHA_CLU     Name
Cluster resource group . . . . . > SVC_LL_CRG    Name
Cluster resource group type  . . > *DEV          *DATA, *APP, *DEV, *PEER
CRG exit program . . . . . . . . > *NONE         Name, *NONE
  Library  . . . . . . . . . . .                 Name
User profile . . . . . . . . . . > *NONE          Name, *NONE

Recovery domain node list:
  Node identifier  . . . . . . . > CTCIHA9V      Name
  Node role  . . . . . . . . . . > *PRIMARY      *CRGTYPE, *PRIMARY...
  Backup sequence number . . . .   *LAST         1-127, *LAST
  Site name  . . . . . . . . . . > SITE1         Name, *NONE
  Data port IP address . . . . .   *NONE

  Node identifier  . . . . . . . > CTCIHA9W      Name
  Node role  . . . . . . . . . . > *BACKUP       *CRGTYPE, *PRIMARY...
  Backup sequence number . . . .   *LAST         1-127, *LAST
  Site name  . . . . . . . . . . > SITE1         Name, *NONE
  Data port IP address . . . . .   *NONE

Exit program format name . . . .   EXTP0100      EXTP0100, EXTP0101...
Exit program data  . . . . . . .   *NONE
Distribute info user queue . . .   *NONE         Name, *NONE
  Library  . . . . . . . . . . .                 Name
Job  . . . . . . . . . . . . . .   *JOBD          Name, *JOBD, *CRG

Configuration object list:
  Configuration object . . . . . > IASP1         Name, *NONE
  Configuration object type  . .   *DEVD         *DEVD, *CTLD, *LIND, *NWSD
  Configuration object online  . > *ONLINE       *OFFLINE, *ONLINE,
  Server takeover IP address . . > '10.0.0.1'

Text description . . . . . . . .   *BLANK
Failover message queue . . . . .   *NONE         Name, *NONE
  Library  . . . . . . . . . . .                 Name


F3=Exit   F4=Prompt   F5=Refresh   F12=Cancel   F13=How to use this display
F24=More keys
```

Figure 3-20 Create a cluster resource group for Storwize Metro Mirror

Note the following items on the **CRTCRG** command:

– The CRG type is *DEV because this CRG controls an IASP.

– The CRG exit program can specify an exit program that is activated whenever there is a change in the CRG status.

- The recovery domain node list defines which of the cluster nodes is the primary node and the order of the backup nodes in the cluster. Make sure to specify the same site names for both cluster nodes because in this setup you work with only one copy of the IASP that is switched between two systems.
- The configuration object list defines which objects are controlled by the CRG. In this example, this is IASP1, which has an object type of *DEVD.
- Configuration object online defines whether the IASP and the server takeover IP address should be varied on after a failover or a switchover.
- The failover message queue defines a message queue on the backup node for this specific CRG that receives a message when a failover is about to happen and can be configured to wait for an answer before proceeding with the failover. In this example, the failover message queue was already configured on the cluster level and is set to *NONE in the CRG.

3. Create an ASP copy description by running the Add SVC ASP Copy Description (**ADDSVCCPYD**) command, as shown in Figure 3-21. In this example, there is only one ASP copy description because there is only one set of LUNs from an IBM i point of view.

```
                Add SVC ASP Copy Description (ADDSVCCPYD)

Type choices, press Enter.

ASP copy . . . . . . . . . . . . > SVC_LL        Name
ASP device . . . . . . . . . . . > IASP1         Name
Cluster resource group . . . . . > SVC_LL_CRG    Name, *NONE
Cluster resource group site  . . > SITE1         Name, *NONE
Node identifier  . . . . . . . . > *CRG          Name, *CRG, *NONE
Storage host:
  User name  . . . . . . . . . . > admin
  Secure Shell key file  . . . . > '/QIBM/UserData/HASM/hads/.ssh/id_rsa'

  Internet address . . . . . . . > '9.5.168.218'

Virtual disk range:
  Range start  . . . . . . . . . > 15            0-8191
  Range end  . . . . . . . . . . > 18            0-8191
             + for more values
  Host identifier  . . . . . . .  *ALL          0-512, *ALL
               + for more values
               + for more values
Device domain  . . . . . . . . .  *             Name, *
Recovery domain:
  Cluster node . . . . . . . . . > CTC9IHA9V     Name, *NONE
  Host identifier  . . . . . . . > 4              0-512
  Cluster node . . . . . . . . . > CTC9IHA9W     Name
  Host identifier  . . . . . . . > 5              0-512
               + for more values
               + for more values
                                                                      Bottom
F3=Exit   F4=Prompt   F5=Refresh   F12=Cancel   F13=How to use this display
 F24=More keys
```

Figure 3-21 Add a SAN Volume Controller auxiliary storage pool Copy Description

Note the following items on the **ADDSVCCPYD** command:

- Unless you are on Storwize V6.2 or later, the user name that you specify here must be `admin`. This profile has no relationship to the user that is used on the Storwize system. The user is chosen based on the SSH key file pair only. Starting with Storwize V6.2, you can either specify `admin` as the user or the name of the Storwize user to which the SSH keyfile belongs.
- The virtual disk range is the Storwize volume IDs of the disks in your IASP. Device domain specifies the name of the device domain to which you want to issue the command. Either enter a specific name or enter * to use the device domain that the node you are issuing the command on is in.
- Recovery domain specifies the connection information for each node in the CRG site recovery domain. Enter the cluster node names here. The host identifier for each cluster node can be found in the Storwize GUI.

4. Start the CRG by running the following command:

```
STRCRG CLUSTER(PWRHA_CLU) CRG(SVC_LL_CRG)
```

3.4 Configuring PowerHA for i by using LUN-level switching with remote Copy Services

LUN-level switching can also be used together with Metro Mirror or Global Mirror in a three-node cluster. In this setup, LUN-level switching provides local high availability (HA). In addition, if the whole local site goes down, Metro Mirror or Global Mirror provide a disaster recovery (DR) option to a remote site.

For this scenario, it is assumed that the following steps are complete:

- The example setup is done for Global Mirror. If you want to use Metro Mirror, make sure to change the **STRSVCSSN** command.
- The PowerHA setup for the local LUN-level switching is complete, as described in 3.3, “Configuring PowerHA for i by using LUN-level switching” on page 51.
- Global Mirror for all volumes of the IASP is configured.
- The private SSH key that is used for communication between PowerHA and the SAN Volume Controller is distributed to the third system and the public SSH key is assigned to a user on the secondary SAN Volume Controller. For more information, see 1.2.1, “Preparing for an SSH connection between IBM i and Storwize system” on page 4.

Table 3-4 provides an overview of the naming conventions and parameter values that are used for this specific setup.

Table 3-4 Settings for the LUN-level switching and Global Mirror combined scenario

Parameter	Preferred primary	Preferred backup 1	Preferred Backup2
System name	CTCIHA9V	CTCIHA9W	CTCIHA9X
Cluster name	PWRHA_CLU		
CRG	SVC_LL_CRG		
Device domain	PWRHA_DMN		
Administrative domain	PWRHA_CAD		

Parameter	Preferred primary	Preferred backup 1	Preferred Backup2
IASP name and number	IASP1	IASP1	IASP1
CRG site name	SITE1		SITE3
Takeover IP	10.0.0.1		
Heartbeat cluster IP	10.10.10.1	10.10.10.2	10.10.10.3
Management access IP	9.5.167.53	9.5.167.54	9.5.167.60
ASP Copy Description	SVC_LL		SVC_GM_T
SAN Volume Controller Session for Global Mirror		SVC_FLC	
SAN Volume Controller IP address	9.5.168.218		9.5.168.220
SSH key file location	/QIBM/UserData/HASM/hads/.ssh/id_rsa		
Volumes IDs[a]	15-18	15-18	23-26

a. Volumes IDs do not need to be the same on the source and target of remote copy relationships.

For this the LUN-level switching and Global Mirror combined scenario, complete the following steps:

1. The IBM i partition CTCIHA9X is not yet part of the cluster and device domain. Add it to the existing cluster environment by using the commands that are shown in Example 3-4.

   Example 3-4 Add the FlashCopy node into the cluster and device domain

   ```
   ADDCLUNODE CLUSTER(PWRHA_CLU) NODE(CTCIHA9X ('10.10.10.3'))
   ADDDEVDMNE CLUSTER(PWRHA_CLU) DEVDMN(PWRHA_DMN) NODE(CTCIHA9X)
   ```

2. If you set up an Administrative Domain, add the additional cluster node to it by running the following Add CAD node Entry (**ADDCADNODE**) command:

   ```
   ADDCADNODE CLUSTER(PWRHA_CLU) ADMDMN(PWRHA_DMN) NODE(CTCIHA9X)
   ```

3. You must have a device description for the IASP on the preferred target system. To do so, put the device description into the Cluster Administrative Domain as a Monitored Resource Entry.

   Alternatively, on the backup system, create a device description for the IASP with the same IASP name that you use on your preferred source system. If you choose a different RDB name than the IASP name on the preferred source system, make sure to specify the same value on the preferred target system; otherwise, you can use the default value of *GEN.

   ```
   CRTDEVASP DEVD(IASP1) RSRCNAME(IASP1) RDB(*GEN)
   ```

4. Add the third node into the recovery domain of the existing CRG for LUN-level switching as a second backup node by running the Add CRG Node Entry (**ADDCRGNODE**) command, as shown in Figure 3-22.

Note: Make sure that you specify a different site name from the one that you used for the pure LUN-level switching.

```
                    Add CRG Node Entry (ADDCRGNODE)

Type choices, press Enter.

Cluster  . . . . . . . . . . . . CLUSTER        PWRHA_CLU
Cluster resource group . . . . . CRG            SVC_LL_CRG
Recovery domain node list:       RCYDMN
  Node identifier  . . . . . . .                CTCIHA9X
  Node role  . . . . . . . . . .                *BACKUP
  Backup sequence number . . . .                *LAST
  Site name  . . . . . . . . . .                SITE3
  Data port IP address . . . . .                *NONE

                         + for more values

                                                                      Bottom
F3=Exit   F4=Prompt   F5=Refresh   F12=Cancel   F13=How to use this display
F24=More keys
```

Figure 3-22 Use the ADDCRGNODE command to add a third node into the recovery domain

5. Create an additional ASP copy description for the Global Mirror target LUNs by running the Add SVC ASP Copy Description (**ADDSVCCPYD**) command, as shown in Figure 3-23.

```
                Add SVC ASP Copy Description (ADDSVCCPYD)

Type choices, press Enter.

ASP copy . . . . . . . . . . . . > SVC_GM_T     Name
ASP device . . . . . . . . . . . > IASP1        Name
Cluster resource group . . . . . > SVC_LL_CRG   Name, *NONE
Cluster resource group site  . . > SITE3         Name, *NONE
Node identifier  . . . . . . . .  > *CRG        Name, *CRG, *NONE
Storage host:
  User name  . . . . . . . . . . > admin
  Secure Shell key file  . . . . > '/QIBM/UserData/HASM/hads/.ssh/id_rsa'
  Internet address . . . . . . . > '9.5.168.220'

Virtual disk range:
  Range start  . . . . . . . . .> 23 0-8191
  Range end  . . . . . . . . . .> 26 0-8191
              + for more values
  Host identifier  . . . . . . .  *ALL         0-512, *ALL
               + for more values
               + for more values
Device domain  . . . . . . . . .  *            Name, *
Recovery domain:
  Cluster node . . . . . . . . . > *NONE       Name, *NONE
  Host identifier  . . . . . . . >             0-512
               + for more values
                                                                      Bottom
F3=Exit   F4=Prompt   F5=Refresh   F12=Cancel   F13=How to use this display
F24=More keys
```

Figure 3-23 Add SAN Volume Controller independent auxiliary storage pool Copy Description

Note the following items on the **ADDSVCCPYD** command:

- For the target copy description that is used for Global Mirror, specify the CRG you used for LUN-level switching. Make sure that the CRG site matches the value from the CRG (third node site name in the recovery domain).
- The node identifier must be set to `*CRG`. Unless you are on Storwize V6.2 or later, the user name that you specify here must be `admin`. This profile has no relationship to the user that is used on the Storwize system. The user is chosen based on the SSH key file pair only. Starting with Storwize V6.2, you can either specify `admin` as the user or the name of the Storwize user to which the SSH keyfile belongs.
- The virtual disk range is the Storwize volume IDs of the disks in your IASP.

6. Start the SAN Volume Controller session by running the following command:

```
STRSVCSSN SSN(GLOBALMIR) TYPE(*GLOBALMIR) CRG(SVC_LL_CRG)
```

3.5 Full-system scenarios

This section describes the following full-system scenarios that are an extension to the options that are offered by PowerHA by using IASP technologies:

- 3.5.1, “Full-system FlashCopy by using the Full System Copy Services Manager” on page 59
- 3.5.2, “PowerHA Full-system Replication for Storwize” on page 64

3.5.1 Full-system FlashCopy by using the Full System Copy Services Manager

STG Lab Services offer a Full System Copy Services Manager (FSCSM). FSCSM is installed on a production partition and is cloned by using FlashCopy. A managing partition controls the entire FlashCopy process by using IP communication to the production and FlashCopy backup partition and includes SSH communication through the HMC.

Figure 3-24 shows an overview of FSCSM.

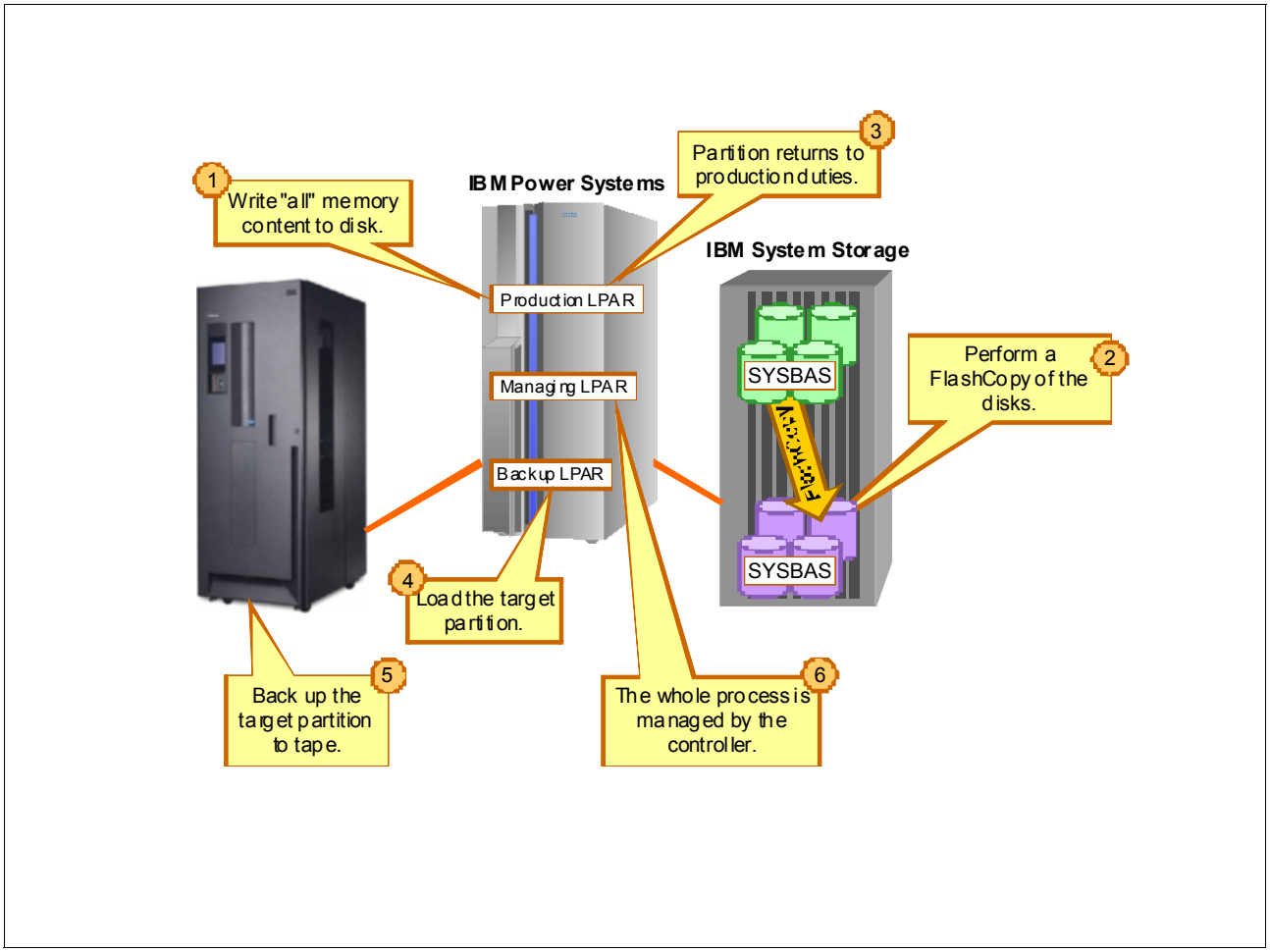

Figure 3-24 Full System Copy Services Manager from STG Lab Services

FSCSM is licensed separately from Advanced Copy Services for PowerHA, but uses the same Copy Services Environment menus for setting up the profiles and scripts to communicate with the Storwize storage system. FSCSM must be installed on the managing and production partition in the QZRDIASH5 library.

FSCSM communicates with the HMC to perform IPLs for and monitor partitions. It also can change host connections on the Storwize system so that one backup LPAR can be used for several production LPARs if the saves of all the LPARs finish in time before a new cycle of FlashCopy operations must be started.

FSCSM supports space-efficient, thin-provisioned, and incremental FlashCopy. The FlashCopy can be taken at the source side of a Metro Mirror, Global Mirror, or Global Mirror with Change Volumes (GMCV) setup. In addition, the FlashCopy can be taken at the target side of a Metro Mirror or GMCV setup.

For ease of use, a single command is used to initiate and monitor the process, from quiesce to backups to BRMS information transfer. Administrative and operator tasks are separated for ease of use and security. In addition, it prevents the production IP addresses from coming online at the backup LPAR.

Important: While taking a warm FlashCopy with the full system online, even when quiesced, any object that is resident in memory and not journaled, such as files, data queues, and data areas, is subject to potential damage due to object inconsistencies from parts of the object that are not on disk. Journaling allows the target partition to recover from possible damaged journaled objects by applying journal entries to journaled files during the IPL. It is only with a full-system FlashCopy that there is a chance that on taking a warm FlashCopy that some IBM i internal objects are corrupted at IPL time. This might lead to a situation where the backup LPAR cannot successfully finish the IPL. You might need a SLIP installation, or another FlashCopy might be needed for a recovery.

In addition to controlling the IBM i full-system FlashCopy process itself, FSCSM also provides a high level of integration with IBM i Backup Recovery and Media Services (BRMS) to make sure that the backup by BRMS that is run from the FlashCopy backup partition appears in the BRMS history containing all the backup and media information, as was done from the production partition with the BRMS history (QUSRBRM) automatically transferred to the production partition after the backup finished. See Figure 3-25.

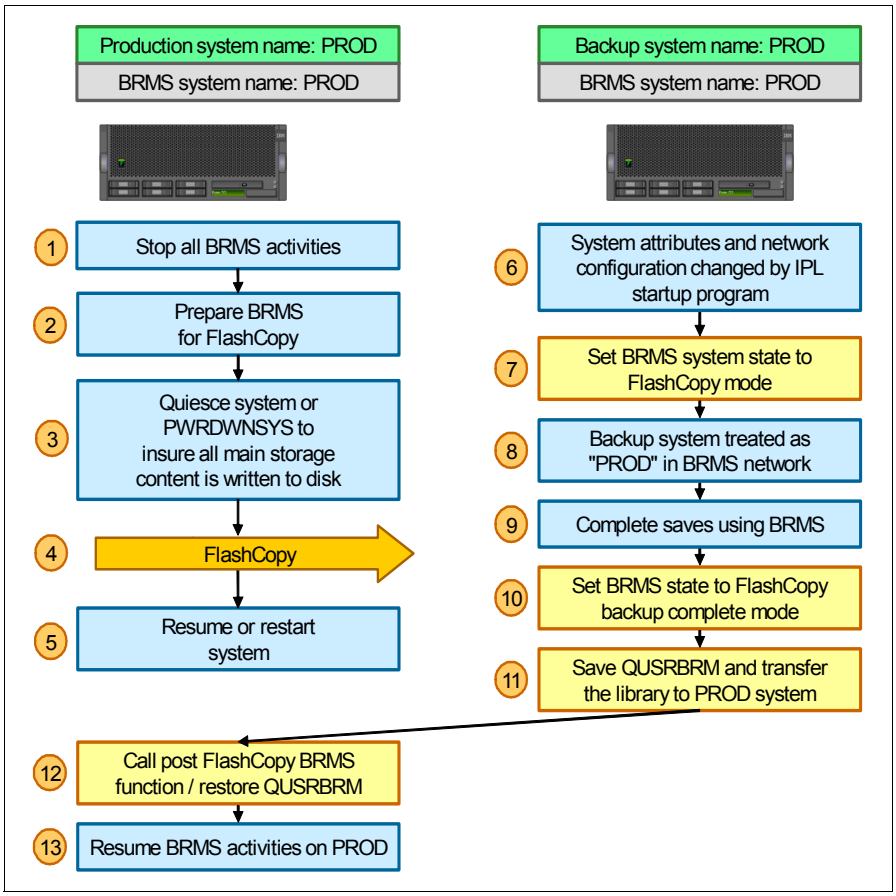

Figure 3-25 Full-system FlashCopy toolkit integration with BRMS

The FSCSM full-system FlashCopy functions are configured on the managing partition by running the Create Full System FlashCopy (**CRTSYSCPY**) command, as shown in Figure 3-26.

```
                    Create Full Sys FlashCopy (CRTSYSCPY)

 Type choices, press Enter.

 Configuration Name . . . . . .   FLC_HA7A      Character value
 Environment name . . . . . .     HA7A            F4 to prompt
 Source partition host name . . .  CTCIHA7A.RCHLAND.IBM.COM
 Source HMC partition name  . . .  CTCIHA7A
 Source partition profile . . . .  DEFAULT
 Source managing system . . . . .  CTCIHA7
 Source HMC1 address  . . . . . .  9.5.168.169
 Source HMC2 address  . . . . . .  *NONE
 Shutdown target  . . . . . .      *YES          *YES, *NO
 Restart target partition . . . .  *YES          *YES, *NO, *INQ, *OFF
 Target LPAR IPL source . . . . .  *PANEL        *PANEL, A, B, D
 Target LPAR keylock position . .  *PANEL        *PANEL, *AUTO, *MANUAL
 Target LPAR pre-activation pgm    *NONE         *NONE, name
   Library . . . . . .  *LIBL  *LIBL, library name
                                                                    More...
 Target partition host name ... CTCIHA7C.RCHLAND.IBM.COM
 Target HMC partition name  . . .  CTCIHA7C
 Target partition profile . . . .  DEFAULT
 Target managing system . . . . .  *SOURCE
 Target HMC1 address  . . . . . .  *SOURCE
 Target HMC2 address  . . . . . .  *NONE
 Backup Application . . . . . .    *BRMS         *NATIVE, *BRMS
 Lock BRMS  . . . . . .            *BOTH         *BOTH, *NO, *SRCONLY...
 Restricted BRMS media classes  .  *NONE           F4 to prompt
               + for more values
                                                                    More...
Target LPAR Device Config:
   Backup device name . . . . . .  TS3400        *NONE, device name
   Backup device serial number  .  78-1011003    *NONE, serial number
   Robot host . . . . . .          *NOCHANGE
   Local internet address . . . .  *NOCHANGE
               + for more values
 Program to move QUSRBRM  . . . .  *SAVF_SOCK    *SAVF_SOCK,
*VRTTAP,*NONE,name
   Library  . . . . . .              *LIBL       *LIBL, library name
 Save compression for QUSRBRM . .  *DEV          *DEV, *YES, *NO, *LOW...
 Operation type . . . . . .        *QUIESCE      *QUIESCE, *IPL, *NOIPL
 Storage Type . . . . . .          *DS8K         *DS5K, *DS8K, *SVC
 Issue IPL confirmation message    *NO           *YES, *NO
 Source LPAR shutdown command . .
                                                                    More...
 F3=Exit   F4=Prompt   F5=Refresh   F12=Cancel   F13=How to use this display
 F24=More keys
```

Figure 3-26 Full-system FlashCopy toolkit CRTSYSCPY command panel

This figure shows an example of a FlashCopy configuration where CTCIHA7A is the production partition to be flashed by using the managing partition to the target partition CTCIHA7C with quiescing the production partition (OPTYPE *QUIESCE) before taking the FlashCopy, resuming the production partition, and activating the backup partition by starting its specified IP interface and BRMS backup job.

Figure 3-27 shows the remaining pages from the CRTSYSCPY command panel.

```
Minutes to wait for power down     60            minutes
Force *SYSBAS resume . . . . . .   60            seconds
Force FlashCopy . . . . . .       *NO           *YES, *NO
FlashCopy exit program . . . . .   *NONE         *NONE, name
  Library  . . . . . .  . . . .     *LIBL        *LIBL, library name
Minutes to wait for power up . .   600           minutes
Target LPAR startup program  . .   *NONE         *ORIG, *NONE, name
  Library  . . . . . .  . . . .     *LIBL        *LIBL, library name
Startup program job queue  . . .   QCTL          *NONE, name
  Library  . . . . . .  . . . .     QSYS         *LIBL, library name
Hold scheduled jobs on target  .   *YES          *YES, *NO
Target LPAR default route:
  Binding interface  . . . . . .   *NOCHANGE
  Next hop . . . . . .  . . . .
                                                                    More...
Target Comm Interfaces:
  IO card serial number  . . . .   00-EC756OA    *NONE, Serial number
  Line Description . . . . . .     FSFCL         *NEW, line name
  IO card port number  . . . . .   0             0-32
  IO card IP address . . . . . .   9.5.168.117
  Network Mask . . . . . .  . .    '255.255.255.0'
              + for more values
Target partition backup cmd  . .   STRBKUBRM USERDATA SBMJOB(*NO)
              + for more values
Wait for final notification  . .   *NO           *YES, *NO
                                                                   More...
Stop target after process  . . .   *RMV          *YES, *NO, *RMV
```

Figure 3-27 Full-system FlashCopy toolkit CRTSYSCPY command panel (continued)

After the Copy Services environment is created by running **WRKCSE** and the configuration by running **CRTSYSCPY**, a full-system FlashCopy can be started by running the Make Full System Copy (**MKSYSCPY**) command, as shown in Figure 3-28.

```
                      Make Full System Copy (MKSYSCPY)

Type choices, press Enter.

Configuration Name . . . . . .   FLC_HA7A        F4 to prompt
User Profile . . . . . .  . . .   POWERHA       *DDM, User Profile
Password . . . . . .  . . . . .                 Character value


                                                                  Bottom
F3=Exit   F4=Prompt   F5=Refresh   F12=Cancel   F13=How to use this display
F24=More keys
```

Figure 3-28 Full-system FlashCopy toolkit STRFLASH command panel

For more information about the full-system FlashCopy toolkit service offering, contact STG Lab Services at the following website:

http://www.ibm.com/systems/services/labservices/contact.html

3.5.2 PowerHA Full-system Replication for Storwize

PowerHA Full System Replication for Storwize (PowerHA FSR) is the current toolkit offering that is developed by IBM Lab Services Power Systems Delivery Practice. PowerHA FSR is a solution to provide a full, current copy of a production system disks at another “site”.

This solution should not be considered as HA but as DR. It does have some limitations when compared to using PowerHA together with an IASP implementation:

- The DR system must be powered down during normal operation.
- There is only one copy of the operating system. Therefore, any changes to the operating system that require an IPL require a production outage, but with PowerHA you can switch to the other system.
- If there is a system failure that requires a main storage memory dump, you cannot switch to the DR site before finishing the dump (unless you interrupt the memory dump and then cannot analyze why the problem happened). With PowerHA, you can fail over to the backup system and still do analysis on the failed system because that system does have its own copy of the operating system.

PowerHA FSR automates the correction of resources. When attaching a full-system copy to a different hardware system, chances are that resource names for communication lines, tapes, and so on are different on the new system compared to the old system. Therefore, corresponding line descriptions, device descriptions, and so on must be changed before those resources can be used.

PowerHA FSR does the following:

- Associates specific storage resources based on the system serial number and LPAR number.
- Associates specific IP addresses based on the system serial number and LPAR number.
- Associates communication resources based on the system serial number and LPAR number.
- Fixes TCP routes based on the system serial number and LPAR number.
- Adjusts the BRMS location for devices.

In addition, PowerHA FSR can perform a switch or a detach for Metro Mirror, Global Mirror, and GMCV.

As shown in Figure 3-29 on page 65, a preferred setup for using PowerHA FSR is to have two controlling LPARs, one on the production site and one on the DR site. If only one controlling LPAR is used, then that controlling LPAR must be in the DR site. The controlling LPARs do not need to be dedicated to controlling the environment; they can be used for other work (for example, testing or development) in parallel. They must be able to talk to both HMCs, both Storwize storage systems, the production node, and the DR node. All communication between LPARs, HMSC, and Storwize systems is encrypted.

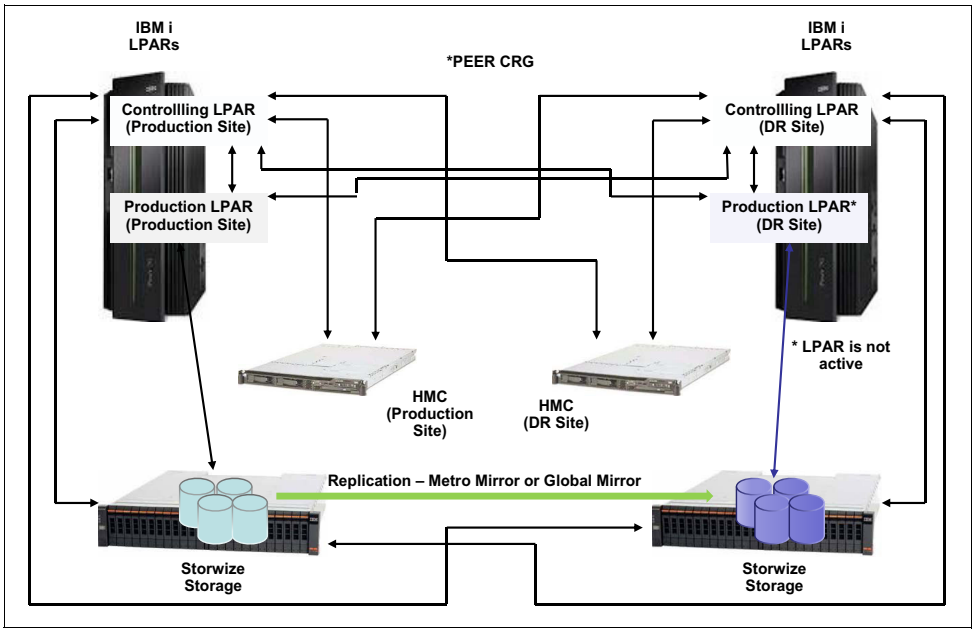

Figure 3-29 PowerHA FSR preferred setup

Important: PowerHA SystemMirror for IBM i Enterprise Edition must be installed and licensed on the controlling LPARs.

PowerHA FSR requires that the Storwize storage system is either natively attached (without using VIOS) or through VIOS by using NPIV. vSCSI attachment through VIOS is not supported.

Figure 3-30 shows an example of the commands that are available with PowerHA FSR. The Switch Copy Services Environment (**SWCSE**) command is used to switch from the production to the DR site. A switch is not performed automatically if the current production system fails, but must be triggered manually by running the **SWCSE** command.

```
                    Switch Copy Services Environ. (SWCSE)

Type choices, press Enter.

Environment name . . . . . .                 Name
Switch type  . . . . . .  .  .   *SCHEDULED    *SCHEDULED, *UNSCHEDULED
Type . . . . . .  .  .  .  .  .  .  .  .  .  .  *            *, *GMIR, *MMIR
Auto replicate . . . . . .  .  .  .  .  .  .  .  *DFT         *DFT, *YES, *NO
```

Figure 3-30 Switch Copy Services Environment command

The **SWCSE** command performs the following functions:

- If the Production node is running, sends an *INQ message to QSYSOPR requesting permission to proceed.
- Calls a customer-specified power-down command.
- Waits until the production node LPAR is deactivated.
- Switches replication direction between the SAN Volume Controllers.

- Performs an IPL on the DR node.
- Calls the customer startup program:
  - FIXSTRPRSC processes the information for *CMN, STG, *BRM, and *RTE differences that are entered with the **WRKSTRPRSC** command.
  - Changes license keys for other programs that depend upon the system serial number.
  - Performs other customer-specified processing.
- Starts clustering if clustering was set up and IP addresses for both the production and DR nodes are specified.

For more information about the PowerHA FSCSM Replication for Storwize offering, contact STG Lab Services at the following website:

http://www.ibm.com/systems/services/labservices/contact.html

4

Monitoring and managing your environment

This chapter describes methods that help you monitor and manage your PowerHA for i by using Storwize environment.

The following topics are described in this chapter:

Note: All of the sample CL programs that are shown in this chapter are available as a save file. For more information about how to access and download these save files, see Appendix B, “Additional material” on page 127.

4.1 Monitoring your environment

Monitoring of environment is the most important task that you must undertake in a high-availability (HA) environment.

Historically, it was deemed sufficient to check your environment once a day. However, the problem with checking only once a day is that all you discover is that your availability solution is unusable at the time you run the check, and that your environment might have been unusable for anything from a must to assume that the solution was not available for use for an average of 12 hours. Putting that in perspective, your availability solution is usable only approximately 50% of the time.

This situation is easily prevented by implementing an automated, proactive monitoring regime. By using tools and automation to do the monitoring on a regular and automatic basis, you are alerted when there are situations that need a resolution. You have a better understanding of the current ability to switch if there is a problem with your production partition.

Important: It is imperative that you implement effective and complete monitoring processes for your specific environment. These monitoring tools should be as resilient as the systems they are monitoring. Failing to monitor correctly might lead to problems later.

There are many areas that you should consider when you are monitoring your environment, which include monitoring the areas that are described in these sections:

- 4.1.1, “QSYSOPR message queue” on page 69
- 4.1.2, “Monitoring the status of your replication” on page 69
- 4.1.3, “Monitoring the status of cluster nodes” on page 73
- 4.1.4, “Monitoring the Cluster Administrative Domain status” on page 76
- 4.1.5, “Monitoring the Cluster Resource Group status” on page 78
- 4.1.6, “Monitoring the IASP status” on page 81
- 4.1.7, “Other monitoring considerations” on page 85

Different environments have different characteristics and different potential points of failure. Therefore, you should make sure that any monitoring processes that you implement can identify any of these potential points of failure to allow you time to make informed decisions about the resolution of any issues that might arise. What you do not want to do is discover that you have a problem when you come to the point of needing to switch in a hurry.

Note: The example CL programs that are provided in this chapter are designed for use with an IBM PowerHA SystemMirror for i (PowerHA) environment. If you have an IBM Systems Lab Services Toolkit, they have other monitoring methods that complement or replace the basic PowerHA methods. These methods are identified as part of the installation service of the toolkit.

4.1.1 QSYSOPR message queue

The first place that you should always look for information about the health of your replication solution is the QSYSOPR message queue. PowerHA periodically checks the health of the environment and sends messages to indicate any problems that it finds to the QSYSOPR message queue.

4.1.2 Monitoring the status of your replication

The most important area to monitor is the actual replication. It is imperative that any issues here are identified as early as possible. If the replicated copy of data is unusable, the whole solution is put at risk.

Traditionally, on a logical replication solution, this task is done once a day. With the hardware solutions, it is a much simpler process, so it can be done more frequently, but not *too* frequently.

Note: A preferred practice might be to monitor 2 - 3 times a day. Possible times might be first thing in the morning, the middle of the day, and just before major batch processes are scheduled to run. However, your environment has its own specific requirements. Do not be tempted to run every half hour because it is excessive.

To monitor for any replication problems, you can write a simple CL program that checks the current state of the replication by running the Retrieve Auxiliary Storage Pool Session (**RTVASPSSN**) command and alert your system operators by sending a message to the QSYSOPR message queue or by running the Send SMTP Email (**SNDSMTPEMM**) command to send an email message to an email recipient.

Example 4-1 shows a sample CL program that you can use to build your own monitoring tool for replication sessions. This CL program is available as a save file. For more information about how to access and download this save file, see Appendix B, “Additional material” on page 127.

Tip: If you want to use the sample code that is shown in Example 4-1, specify the name of an appropriate email address if you want to use that method of notification.

Example 4-1 Sample CL program to monitor replication status

```
/*****************************************************************************/
/* MONASPSSN - Monitor the replication status                                */
/*                                                                           */
/*             This program is intended to provide a simple way of           */
/*             monitoring the replication status.                            */
/*                                                                           */
/*             This is achieved by using the RTVASPSSN command to obtain     */
/*             information about the replication environment and based       */
/*             upon the type of environment it will identify the status of   */
/*             the environment.                                              */
/*                                                                           */
/*             If the replication is not in a fully consistent status, it    */
/*             will send a message to the system operator and also send an    */
/*             email to the named user to alert the administrators to the    */
/*             fact that they should investigate the problems.               */
/*                                                                           */
```

```
/*            ============================================================  */
/*           I NOTE:  This sample code assumes that there are only 2 (TWO)  I */
/*           I copy descriptions involved in the session. It is possible    I */
/*           I that more could exist in which case this sample code would   I */
/*           I  need to be reworked to allow for this possibility. I */
/*            ============================================================  */
/*                                                                            */
/******************************************************************************/
/* Written by  : David Painter                                                */
/* Date Written: Sep 3, 2015                                                  */
/* Date Changed: None                                                         */
/******************************************************************************/
pgm
             DCL        VAR(&SSN) TYPE(*CHAR) LEN(10) VALUE('GEOMIR    ')
             DCL        VAR(&TYPE) TYPE(*CHAR) LEN(10)
             DCL        VAR(&ASPCPYLST) TYPE(*CHAR) LEN(4472)
             DCL        VAR(&DELIVERY) TYPE(*CHAR) LEN(8)
             DCL        VAR(&MODE) TYPE(*CHAR) LEN(8)
             DCL        VAR(&INTRANSIT) TYPE(*DEC) LEN(10 0)
             DCL        VAR(&TIMEOUT) TYPE(*DEC) LEN(4 0)
             DCL        VAR(&PRIORITY) TYPE(*CHAR) LEN(8)
             DCL        VAR(&TRACKSPACE) TYPE(*DEC) LEN(3 0)
             DCL        VAR(&PERSISTENT) TYPE(*CHAR) LEN(8)
             DCL        VAR(&FLASHTYPE) TYPE(*CHAR) LEN(8)
             DCL        VAR(&ASPCPY) TYPE(*CHAR) LEN(31232)

/* Variables that are used to extract the copy information */
             DCL        VAR(&OFFIASP1) TYPE(*INT) STG(*DEFINED) +
                          LEN(4) DEFVAR(&ASPCPY 1)
             DCL        VAR(&LENIASP) TYPE(*INT) STG(*DEFINED) +
                          LEN(4) DEFVAR(&ASPCPY 5)
             DCL        VAR(&NUMIASP) TYPE(*INT) STG(*DEFINED) +
                          LEN(4) DEFVAR(&ASPCPY 9)

/* Variables for Copy information 1 */
             DCL        VAR(&ASPCPY1) TYPE(*CHAR) LEN(70)
             DCL        VAR(&ASPCPYD1) TYPE(*CHAR) STG(*DEFINED) +
                          LEN(10) DEFVAR(&ASPCPY1 1)
             DCL        VAR(&ASPDEV1) TYPE(*CHAR) STG(*DEFINED) +
                          LEN(10) DEFVAR(&ASPCPY1 11)
             DCL        VAR(&ASPROLE1) TYPE(*CHAR) STG(*DEFINED) +
                          LEN(10) DEFVAR(&ASPCPY1 21)
             DCL        VAR(&ASPSTATE1) TYPE(*CHAR) STG(*DEFINED) +
                          LEN(10) DEFVAR(&ASPCPY1 31)
             DCL        VAR(&ASPDSTATE1) TYPE(*CHAR) STG(*DEFINED) +
                          LEN(10) DEFVAR(&ASPCPY1 41)
             DCL        VAR(&ASPNODE1) TYPE(*CHAR) STG(*DEFINED) +
                          LEN(8) DEFVAR(&ASPCPY1 51)
             DCL        VAR(&ASPPCTTRK1) TYPE(*int) STG(*DEFINED) +
                          LEN(4) DEFVAR(&ASPCPY1 59)
             DCL        VAR(&ASPOOS1) TYPE(*int) STG(*DEFINED) +
                          LEN(4) DEFVAR(&ASPCPY1 63)
             DCL        VAR(&ASPSYNCP1) TYPE(*INT) STG(*DEFINED) +
                          LEN(4) DEFVAR(&ASPCPY1 67)
```

```
/* Variables for Copy information 2 */
             DCL        VAR(&ASPCPY2) TYPE(*CHAR) LEN(70)
             DCL        VAR(&ASPCPYD2) TYPE(*CHAR) STG(*DEFINED) +
                          LEN(10) DEFVAR(&ASPCPY2 1)
             DCL        VAR(&ASPDEV2) TYPE(*CHAR) STG(*DEFINED) +
                          LEN(10) DEFVAR(&ASPCPY2 11)
             DCL        VAR(&ASPROLE2) TYPE(*CHAR) STG(*DEFINED) +
                          LEN(10) DEFVAR(&ASPCPY2 21)
             DCL        VAR(&ASPSTATE2) TYPE(*CHAR) STG(*DEFINED) +
                          LEN(10) DEFVAR(&ASPCPY2 31)
             DCL        VAR(&ASPDSTATE2) TYPE(*CHAR) STG(*DEFINED) +
                          LEN(10) DEFVAR(&ASPCPY2 41)
             DCL        VAR(&ASPNODE2) TYPE(*CHAR) STG(*DEFINED) +
                          LEN(8) DEFVAR(&ASPCPY2 51)
             DCL        VAR(&ASPPCTTRK2) TYPE(*int) STG(*DEFINED) +
                          LEN(4) DEFVAR(&ASPCPY2 59)
             DCL        VAR(&ASPOOS2) TYPE(*INT) STG(*DEFINED) +
                          LEN(4) DEFVAR(&ASPCPY2 63)
             DCL        VAR(&ASPSYNCP2) TYPE(*INT) STG(*DEFINED) +
                          LEN(4) DEFVAR(&ASPCPY2 67)
/* Work variables */
             DCL        VAR(&X) TYPE(*DEC) LEN(5 0)

/* Step 1 - Use RTVASPSSN to obtain the session data */
             RTVASPSSN  SSN(&SSN) TYPE(&TYPE) ASPCPYLST(&ASPCPYLST) +
                          DELIVERY(&DELIVERY) MODE(&MODE) +
                          INTRANSIT(&INTRANSIT) TIMEOUT(&TIMEOUT) +
                          PRIORITY(&PRIORITY) +
                          TRACKSPACE(&TRACKSPACE) +
                          PERSISTENT(&PERSISTENT) +
                          FLASHTYPE(&FLASHTYPE) ASPCPY(&ASPCPY)
             MONMSG     MSGID(HAE0000 CPF0000) EXEC(DO)
             SNDPGMMSG  MSGID(CPF9898) MSGF(QCPFMSG) MSGDTA('An +
                          error occurred with the MONASPSSN command, +
                          check job log for more information') +
                          MSGTYPE(*ESCAPE)
             RETURN
             ENDDO

/* Step 2 - If there are more than 2 copy descriptions involved issue a warning */
             IF         COND(&NUMIASP *NE 2) THEN(DO)
             SNDPGMMSG  MSGID(CPF9898) MSGF(QCPFMSG) MSGDTA('There +
                          are more than 2 copy descriptions +
                          involved, this sample is invalid') +
                          MSGTYPE(*ESCAPE)
             RETURN
             ENDDO

/* Step 3 - Extract the two data about the source and target ASP copies */
             CHGVAR     VAR(&X) VALUE(&OFFIASP1 + 1)
             CHGVAR     VAR(&ASPCPY1) VALUE(%SST(&ASPCPY &X &LENIASP))
             CHGVAR     VAR(&X) VALUE(&X + &LENIASP)
             CHGVAR     VAR(&ASPCPY2) VALUE(%SST(&ASPCPY &X &LENIASP))
```

```
/* Step 4 - Based on the type of session, calculate what is a success state */
/* FlashCopy */
             IF         COND(&TYPE *EQ '*FLASHCOPY') THEN(DO)
             SNDPGMMSG  MSGID(CPF9898) MSGF(QCPFMSG) MSGDTA('A +
                          FlashCopy session is not supported for +
                          this function') MSGTYPE(*ESCAPE)
             RETURN
             ENDDO

/* Geographic Mirroring */
             IF         COND(&TYPE *EQ '*GEOMIR') THEN(DO)
             IF         COND((&ASPROLE1 *EQ 'MIRROR') *AND +
                          ((&ASPDSTATE1 *EQ 'USABLE') *OR +
                          (&ASPDSTATE1 *EQ 'INSYNC'))) THEN(DO)
             SNDPGMMSG  MSGID(CPF9898) MSGF(QCPFMSG) MSGDTA('The +
                          IASP replication is in the correct status +
                          for this setup') MSGTYPE(*COMP)
             RETURN
             ENDDO
             IF         COND((&ASPROLE2 *EQ 'MIRROR') *AND +
                          ((&ASPDSTATE2 *EQ 'USABLE') *OR +
                          (&ASPDSTATE2 *EQ 'INSYNC'))) THEN(DO)
             SNDPGMMSG  MSGID(CPF9898) MSGF(QCPFMSG) MSGDTA('The +
                          IASP replication is in the correct status +
                          for this setup') MSGTYPE(*COMP)
             RETURN
             ENDDO
             ENDDO

/* Metro Mirror */
             IF         COND(&TYPE *EQ '*METROMIR') THEN(DO)
             IF         COND((&ASPROLE1 *EQ 'TARGET') *AND +
                          (&ASPSTATE1 *NE 'ACTIVE') *AND +
                          (&ASPSTATE1 *NE 'AVAILABLE')) THEN(DO)
             SNDPGMMSG  MSGID(CPF9898) MSGF(QCPFMSG) MSGDTA('The +
                          IASP replication is in the correct status +
                          for this setup') MSGTYPE(*COMP)
             RETURN
             ENDDO
             IF         COND((&ASPROLE2 *EQ 'TARGET') *AND +
                          (&ASPSTATE2 *NE 'ACTIVE') *AND +
                          (&ASPSTATE2 *NE 'AVAILABLE')) THEN(DO)
             SNDPGMMSG  MSGID(CPF9898) MSGF(QCPFMSG) MSGDTA('The +
                          IASP replication is in the correct status +
                          for this setup') MSGTYPE(*COMP)
             RETURN
             ENDDO
             ENDDO

/* Global Mirror */
             IF         COND(&TYPE *EQ '*GLOBALMIR') THEN(DO)
             IF         COND((&ASPROLE1 *EQ 'TARGET') *AND +
                          (&ASPSTATE1 *NE 'ACTIVE') *AND +
                          (&ASPSTATE1 *NE 'AVAILABLE')) THEN(DO)
             SNDPGMMSG  MSGID(CPF9898) MSGF(QCPFMSG) MSGDTA('The +
```

```
                    IASP replication is in the correct status +
                    for this setup') MSGTYPE(*COMP)
          RETURN
          ENDDO
          IF         COND((&ASPROLE2 *EQ 'TARGET') *AND +
                    (&ASPSTATE2 *NE 'ACTIVE') *AND +
                    (&ASPSTATE2 *NE 'AVAILABLE')) THEN(DO)
          SNDPGMMSG  MSGID(CPF9898) MSGF(QCPFMSG) MSGDTA('The +
                    IASP replication is in the correct status +
                    for this setup') MSGTYPE(*COMP)
          RETURN
          ENDDO
          ENDDO

/* Step 5 - Send an email to an admin user if there is an error */
          SNDSMTPEMM RCP((SYSTEMADMINISTRATOR@YOURCOMPANY)) +
                    SUBJECT('IASP replication problems have +
                    been found.') NOTE('Some problems were +
                    found when checking the IASP +
                    replication.  You should use the +
                    DSPASPSSN command to review the situation +
                    and take any necessary actions to correct +
                    the errors.')
          MONMSG     MSGID(TCP5090 TCP5092)

/* Step 6 - Send a message to QSYSOPR if there is an error */
          SNDPGMMSG  MSGID(CPF9898) MSGF(QCPFMSG) MSGDTA('Errors +
                    were found with the IASP replication, +
                    review') TOMSGQ(*SYSOPR)

/* Step 7 - Send an escape message to indicate that there is an error */
          SNDPGMMSG  MSGID(CPF9898) MSGF(QCPFMSG) MSGDTA('Errors +
                    were found with the IASP replication, +
                     review') MSGTYPE(*ESCAPE)
          RETURN
endpgm
```

4.1.3 Monitoring the status of cluster nodes

For PowerHA to be functional, you must ensure that the cluster nodes are ACTIVE under normal conditions. Therefore, you must put in place monitoring to verify this situation.

When verifying the status of the cluster nodes, you must do this from the perspective of a single node. You retrieve the local node status and then, from the local node, you also obtain the remote node status. This is important because failing to do this correctly can result in you believing that the cluster nodes are both active when in reality they are in a PARTITION status.

Figure 4-1 shows an example of a partitioned cluster. If you check the partition LONDON, it states that it is ACTIVE, and if you check the partition FRANKFURT, it also states that it is ACTIVE. Only by performing the check on a single partition are you aware of the fact that the cluster is in a PARTITION state.

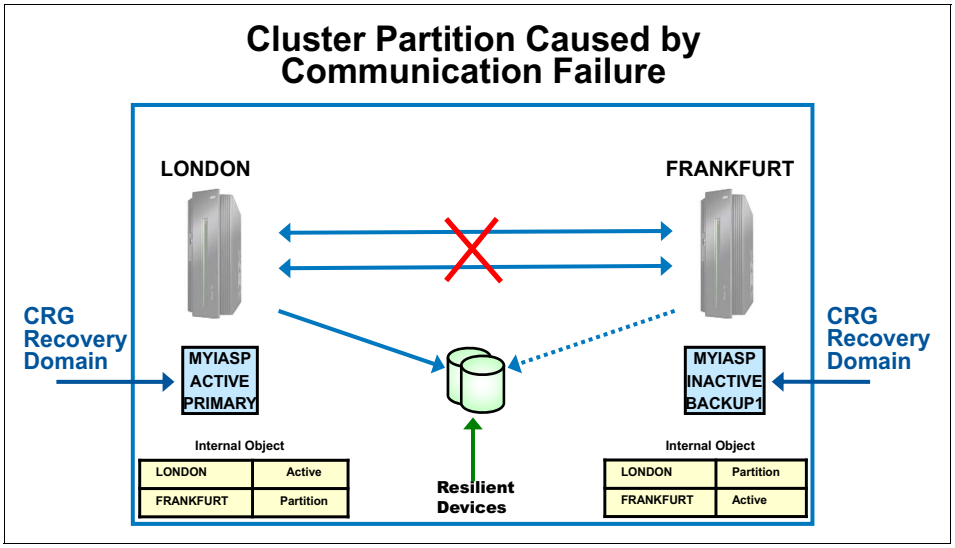

Figure 4-1 A partitioned cluster

Important: Do not use remote commands to query other nodes in the cluster because you do not get the correct information.

Example 4-2 shows a sample CL program that can you can use to build your own monitoring tool for the cluster node status. This CL program is available as a save file. For more information about how to access and download this save file, see Appendix B, “Additional material” on page 127.

Example 4-2 Sample CL program to monitor the cluster node status

```
/*****************************************************************************/
/* MONNODSTS - Monitor the cluster node status                               */
/*                                                                           */
/*             This program is intended to provide a simple way of           */
/*             monitoring the status of the cluster nodes                    */
/*                                                                           */
/*             This is achieved by using system APIs to obtain the current   */
/*             status of the nodes, it should be run on a single node only   */
/*             to avoid missing a partition condition.                       */
/*****************************************************************************/
/* Written by  : David Painter                                               */
/* Date Written: Sep 3, 2015                                                 */
/* Date Changed: None                                                        */
/*                                                                           */
/* Required service programs: QHASM/QHAAPI, QCSTCTL1                         */
/*****************************************************************************/
pgm
DCL        VAR(&CLUNAME) TYPE(*CHAR) LEN(10)

/* API Variables */
```

```
             DCL        VAR(&APIRCVR) TYPE(*CHAR) LEN(18)
             DCL        VAR(&APIRCVRLEN) TYPE(*INT) LEN(4) VALUE(18)
             DCL        VAR(&APIUSRSPC) TYPE(*CHAR) LEN(20)

/* User space header fields and pointers */
             DCL        VAR(&PTR_USRSPC) TYPE(*PTR)
             DCL        VAR(&HEADER) TYPE(*CHAR) STG(*BASED) +
                          LEN(192) BASPTR(&PTR_USRSPC)
             DCL        VAR(&OFFLIST) TYPE(*UINT) STG(*DEFINED) +
                          LEN(4) DEFVAR(&HEADER 125)
             DCL        VAR(&NUMLIST) TYPE(*UINT) STG(*DEFINED) +
                          LEN(4) DEFVAR(&HEADER 133)
             DCL        VAR(&LENENTRY) TYPE(*UINT) STG(*DEFINED) +
                          LEN(4) DEFVAR(&HEADER 137)

/* User space list data fields and pointers */
             DCL        VAR(&PTR_LIST) TYPE(*PTR)
             DCL        VAR(&LISTDATA) TYPE(*CHAR) STG(*BASED) +
                          LEN(512) BASPTR(&PTR_LIST)
             DCL        VAR(&LENNODE) TYPE(*UINT) STG(*DEFINED) +
                          LEN(4) DEFVAR(&LISTDATA 1)
             DCL        VAR(&NODE) TYPE(*CHAR) STG(*DEFINED) LEN(8) +
                          DEFVAR(&LISTDATA 5)
             DCL        VAR(&NODESTS) TYPE(*UINT) STG(*DEFINED) LEN(4) +
                          DEFVAR(&LISTDATA 13)

/* Work variables */
             DCL        VAR(&ERRFLG) TYPE(*CHAR) LEN(1) VALUE('N')
             DCL        VAR(&X) TYPE(*UINT) LEN(4)

/* Step 1 - Use the QhaRetrieveHAInfo API to obtain the cluster name */
             CALLPRC    PRC('QhaRetrieveHAInfo') PARM((&APIRCVR) +
                          (&APIRCVRLEN) ('RHAI0100') (X'00000000'))
             CHGVAR     VAR(&CLUNAME) VALUE(%SST(&APIRCVR 9 10))

/* Step 2 - Use the QcstListClusterInfo API to obtain the cluster node status */
             CHGVAR     VAR(&APIUSRSPC) VALUE('CLUAPI    QGPL      ')
             CALL       PGM(QUSCRTUS) PARM(&APIUSRSPC 'SAMPLE    ' +
                          X'00000001' X'00' '*ALL      ' 'Temporary +
                          user space                              ')
             MONMSG     MSGID(CPF9870)
             CALLPRC    PRC('QcstListClusterInfo') PARM((&APIUSRSPC) +
                          (&CLUNAME) ('LCTI0100') ('*ALL    ') +
                          (X'00000000'))

/* Step 3 - Get a pointer to the USRSPC, and then get pointer to the first list
data */
             CALL       PGM(QUSPTRUS) PARM(&APIUSRSPC &PTR_USRSPC)
             CHGVAR     VAR(&PTR_LIST) VALUE(&PTR_USRSPC)
             CHGVAR     VAR(%OFFSET(&PTR_LIST)) VALUE(&OFFLIST)
             CHGVAR     VAR(&X) VALUE(1)

/* Step 4 - If this node is not active, then there is an error */
 TESTSTS:    IF         COND(&NODESTS *NE 2) THEN(DO)
             SNDPGMMSG  MSGID(CPF9898) MSGF(QCPFMSG) MSGDTA('Node: ' +
```

```
                          *CAT &NODE *TCAT ' is not active') +
                          MSGTYPE(*DIAG)
             CHGVAR     VAR(&ERRFLG) VALUE('Y')
             ENDDO

/* Step 5 - Have we processed all the entries,  if not increment and test again */
             IF         COND(&X *LT &NUMLIST) THEN(DO)
             CHGVAR     VAR(&X) VALUE(&X + 1)
             CHGVAR     VAR(%OFFSET(&PTR_LIST)) +
                          VALUE(%OFFSET(&PTR_LIST) + &LENNODE)
             GOTO       CMDLBL(TESTSTS)
             ENDDO

/* Step 6 - Send an email to an admin user if there is an error */
             IF         COND(&ERRFLG *NE 'Y') THEN(DO)
             SNDPGMMSG  MSGID(CPF9898) MSGF(QCPFMSG) MSGDTA('All +
                          cluster nodes are active') MSGTYPE(*COMP)
             RETURN
             ENDDO

/* Step 7 - Send an email to an admin user if there is an error */
             SNDSMTPEMM RCP((SYSTEMADMINISTRATOR@YOURCOMPANY)) +
                          SUBJECT('Not all cluster nodes are +
                          active, investigate.') NOTE('Not +
                          all of the cluster nodes are active, +
                          while this could be deliberate it should +
                          be investigated to ensure that there is +
                          no exposure to the availability or +
                          replication solution.)')
             MONMSG     MSGID(TCP5090 TCP5092)

/* Step 6 - Send a message to QSYSOPR if there is an error */
             SNDPGMMSG  MSGID(CPF9898) MSGF(QCPFMSG) MSGDTA('Not all +
                          cluster nodes are active, +
                          investigate') TOMSGQ(*SYSOPR)

/* Step 7 - Send an escape message to indicate that there is an error */
             SNDPGMMSG  MSGID(CPF9898) MSGF(QCPFMSG) MSGDTA('Not all +
                          cluster nodes are active, +
                          investigate') MSGTYPE(*ESCAPE)
             RETURN
endpgm
```

4.1.4 Monitoring the Cluster Administrative Domain status

The cluster administrative domain is used to monitor multiple types of resources, and although it is reliable, there might be occasions when something gets out of step between the nodes involved in it. For example, you might have deleted a user profile and forgotten to remove the monitored resource entry from the cluster administrative domain, which results in an inconsistency. Another example is adding a user that has a specific job description that does not exist on one of the other nodes.

To monitor for these problems, you can write a simple CL program to monitor for any entries that are not CONSISTENT and alert your system operators or monitoring tools as your business needs dictate.

You can monitor the cluster administrative domain by running the Print Cluster Administrative Domain Monitored Resource Entry (**PRTCADMRE**) command to produce an output file of monitored resource entries (MREs) that are not in a global consistent state. You can then read how many entries are in the file and if it is not zero, you can then alert the system operators by sending a message to the QSYSOPR message queue or by running the Send SMTP Email (**SNDSMTPEMM**) command to send an email message to an email recipient.

Note: If you use the sample code that is shown in Example 4-3, you must remember to specify the name of the cluster administration domain and an appropriate email address if you plan to use that method of notification.

Example 4-3 shows a sample CL program that can assist you in building your own monitoring tool for the cluster administrative domain. This CL program is available as a save file. For more information about how to access and download this save file, see Appendix B, “Additional material” on page 127.

Example 4-3 Sample CL program to monitor the cluster administrative domain

```
/*****************************************************************************/
/* MONADMDMN - Monitor the cluster administrative domain                     */
/*                                                                           */
/*             This program is intended to provide a simple way of           */
/*             monitoring the Cluster Administrative Domain                  */
/*                                                                           */
/*             This is achieved by simply generating a file of Monitored     */
/*             Resource Entries that are not in a consistent state           */
/*                                                                           */
/*             If this file has any records in it then there are issues so a */
/*             message is sent to the QSYSOPR message queue and an email     */
/*             is sent to a named user as well.  This alerts the admin team  */
/*             that they should investigate the problems.                    */
/*                                                                           */
/*****************************************************************************/
/* Written by  : David Painter                                               */
/* Date Written: Sep 2, 2015                                                 */
/* Date Changed: None                                                        */
/*****************************************************************************/
pgm
             DCL        VAR(&NBRCURRCD) TYPE(*DEC) LEN(10 0)

/* Step 1 - Use the PRTCADMRE to generate a file of errors */
             PRTCADMRE  ADMDMN(ADMDMN) RSCGLBSTS(*ADDED *ENDED +
                          *FAILED *INCONSISTENT *PENDING) +
                          DETAIL(*RSC) OUTPUT(*OUTFILE) +
                          OUTFILE(QTEMP/ADMDMNERR)
             MONMSG     MSGID(HAE0000 CPF0000) EXEC(DO)
             SNDPGMMSG  MSGID(CPF9898) MSGF(QCPFMSG) MSGDTA('An +
                          error occurred with the MONADMDMN command, +
                          check job log for more information') +
                          MSGTYPE(*ESCAPE)
             RETURN
```

```
             ENDDO

/* Step 2 - Retrieve the number of records in the error file */
             RTVMBRD    FILE(QTEMP/ADMDMNERR) NBRCURRCD(&NBRCURRCD)

/* Step 3 - If the number of records is 0 then terminate normally */
             IF         COND(&NBRCURRCD *EQ 0) THEN(DO)
             SNDPGMMSG  MSGID(CPF9898) MSGF(QCPFMSG) MSGDTA('The +
                          Cluster Administrative Domain is all +
                          synchronized') MSGTYPE(*COMP)
             RETURN
             ENDDO

/* Step 4 - Send an email to an admin user if there is an error */
             SNDSMTPEMM RCP((SYSTEMADMINISTRATOR@YOURCOMPANY)) +
                          SUBJECT('Cluster Administrative Domain +
                          errors have been found.') NOTE('Some +
                          errors were found when checking the +
                          Cluster Administrative Domain.  You +
                          should use the WRKCADMRE command to +
                          review the situation and take any +
                          necessary actions to correct the errors.')
             MONMSG     MSGID(TCP5090 TCP5092)

/* Step 5 - Send a message to QSYSOPR if there is an error */
             SNDPGMMSG  MSGID(CPF9898) MSGF(QCPFMSG) MSGDTA('Errors +
                          were found in the Cluster Administrative +
                          Domain, review') TOMSGQ(*SYSOPR)

/* Step 6 - Send an escape message to indicate that there is an error */
             SNDPGMMSG  MSGID(CPF9898) MSGF(QCPFMSG) MSGDTA('Errors +
                          were found in the Cluster Administrative +
                          Domain, review') MSGTYPE(*ESCAPE)
             RETURN
endpgm
```

4.1.5 Monitoring the Cluster Resource Group status

Cluster resource groups (CRGs) are used to define replication characteristics for the PowerHA configuration. They are also effectively an “ON/OFF” switch to tell PowerHA if you are expecting it to take charge in a failover situation. It is important that your CRGs are ACTIVE under normal operations.

To monitor for these problems, you can write a simple CL program to alert your system operators or monitoring tools when they are not in an ACTIVE status, as your business needs dictate.

You can monitor the status of CRGs by using system application programming interface (API) programs to provide you with the status.

Note: If you use the sample code in Example 4-4, you must remember to specify the name of an appropriate email address if you plan to use that method of notification.

Example 4-4 shows a sample CL program that can assist you in building your own monitoring tool for the cluster resource group status. This CL program is available as a save file. For more information about how to access and download this save file, see Appendix B, “Additional material” on page 127.

Example 4-4 Sample CL program to monitor the Cluster Resource Group status

```
/*****************************************************************************/
/* MONCRGSTS - Monitor the cluster resource group status                     */
/*                                                                           */
/*             This program is intended to provide a simple way of           */
/*             monitoring the status of the cluster resource groups          */
/*                                                                           */
/*             This is achieved by using system APIs to obtain the current   */
/*             status of the CRGs.                                           */
/*                                                                           */
/*****************************************************************************/
/* Written by  : David Painter                                               */
/* Date Written: Sep 9, 2015                                                 */
/* Date Changed: None                                                        */
/*                                                                           */
/* Required service programs: QCSTCRG3                                       */
/*****************************************************************************/
pgm
DCL        VAR(&CLUNAME) TYPE(*CHAR) LEN(10)

/* API Variables */
             DCL        VAR(&APIRCVR) TYPE(*CHAR) LEN(18)
             DCL        VAR(&APIRCVRLEN) TYPE(*INT) LEN(4) VALUE(18)
             DCL        VAR(&APIUSRSPC) TYPE(*CHAR) LEN(20)

/* User space header fields and pointers */
             DCL        VAR(&PTR_USRSPC) TYPE(*PTR)
             DCL        VAR(&HEADER) TYPE(*CHAR) STG(*BASED) +
                          LEN(192) BASPTR(&PTR_USRSPC)
             DCL        VAR(&OFFLIST) TYPE(*UINT) STG(*DEFINED) +
                          LEN(4) DEFVAR(&HEADER 125)
             DCL        VAR(&NUMLIST) TYPE(*UINT) STG(*DEFINED) +
                          LEN(4) DEFVAR(&HEADER 133)
             DCL        VAR(&LENENTRY) TYPE(*UINT) STG(*DEFINED) +
                          LEN(4) DEFVAR(&HEADER 137)

/* User space list data fields and pointers */
             DCL        VAR(&PTR_LIST) TYPE(*PTR)
             DCL        VAR(&LISTDATA) TYPE(*CHAR) STG(*BASED) +
                          LEN(512) BASPTR(&PTR_LIST)
             DCL        VAR(&CRGNAME) TYPE(*CHAR) STG(*DEFINED) +
                          LEN(10) DEFVAR(&LISTDATA 1)
             DCL        VAR(&CRGTYPE) TYPE(*UINT) STG(*DEFINED) +
                          LEN(2) DEFVAR(&LISTDATA 11)
             DCL        VAR(&CRGSTS) TYPE(*UINT) STG(*DEFINED) LEN(4) +
                          DEFVAR(&LISTDATA 13)
             DCL        VAR(&CRGPRI) TYPE(*CHAR) STG(*DEFINED) LEN(8) +
                          DEFVAR(&LISTDATA 17)
             DCL        VAR(&CRGAPPID) TYPE(*CHAR) STG(*DEFINED) +
                          LEN(2) DEFVAR(&LISTDATA 25)
```

```
/* Work variables */
             DCL        VAR(&ERRFLG) TYPE(*CHAR) LEN(1) VALUE('N')
             DCL        VAR(&X) TYPE(*UINT) LEN(4)

/* Step 1 - Use the QhaRetrieveHAInfo API to obtain the cluster name */
             CALLPRC    PRC('QhaRetrieveHAInfo') PARM((&APIRCVR) +
                          (&APIRCVRLEN) ('RHAI0100') (X'00000000'))
             CHGVAR     VAR(&CLUNAME) VALUE(%SST(&APIRCVR 9 10))

/* Step 2 - Use the QcstListClusterResourceGroups API to obtain the cluster node
status */
             CHGVAR     VAR(&APIUSRSPC) VALUE('CLUAPI    QGPL      ')
             CALL       PGM(QUSCRTUS) PARM(&APIUSRSPC 'SAMPLE    ' +
                          X'00000001' X'00' '*ALL      ' 'Temporary +
                          user space                          ')
             MONMSG     MSGID(CPF9870)
             CALLPRC    PRC('QcstListClusterResourceGroups') +
                          PARM((&APIUSRSPC) ('CRGL0100') (&CLUNAME) +
                          (X'00000000'))

/* Step 3 - Get a pointer to the USRSPC, and then get pointer to the first list
data */
             CALL       PGM(QUSPTRUS) PARM(&APIUSRSPC &PTR_USRSPC)
             CHGVAR     VAR(&PTR_LIST) VALUE(&PTR_USRSPC)
             CHGVAR     VAR(%OFFSET(&PTR_LIST)) VALUE(&OFFLIST)
             CHGVAR     VAR(&X) VALUE(1)

/* Step 4 - If this CRG is not active, then there is an error */
 TESTSTS:    IF         COND(&CRGSTS *NE 10) THEN(DO)
             SNDPGMMSG  MSGID(CPF9898) MSGF(QCPFMSG) MSGDTA('CRG: ' +
                          *CAT &CRGNAME *TCAT ' is not active') +
                          MSGTYPE(*DIAG)
             CHGVAR     VAR(&ERRFLG) VALUE('Y')
             ENDDO

/* Step 5 - Have we processed all the entries,  if not increment and test again */
             IF         COND(&X *LT &NUMLIST) THEN(DO)
             CHGVAR     VAR(&X) VALUE(&X + 1)
             CHGVAR     VAR(%OFFSET(&PTR_LIST)) +
                          VALUE(%OFFSET(&PTR_LIST) + &LENENTRY)
             GOTO       CMDLBL(TESTSTS)
             ENDDO

/* Step 6 - Send an email to an admin user if there is an error */
             IF         COND(&ERRFLG *NE 'Y') THEN(DO)
             SNDPGMMSG  MSGID(CPF9898) MSGF(QCPFMSG) MSGDTA('All +
                          cluster resource groups are active') +
                          MSGTYPE(*COMP)
             RETURN
             ENDDO

/* Step 7 - Send an email to an admin user if there is an error */
             SNDSMTPEMM RCP((SYSTEMADMINISTRATOR@YOURCOMPANY)) +
                          SUBJECT('Not all cluster resource groups +
```

```
                          are active, review.') NOTE('Not +
                          all of the cluster resource groups are +
                          active, while this could be deliberate it +
                          should be investigated to ensure that +
                          there is no exposure to the availability +
                          or replication solution.)')
             MONMSG     MSGID(TCP5090 TCP5092)

/* Step 6 - Send a message to QSYSOPR if there is an error */
             SNDPGMMSG  MSGID(CPF9898) MSGF(QCPFMSG) MSGDTA('Not all +
                          cluster resource groups are active, +
                           reviews') TOMSGQ(*SYSOPR)

/* Step 7 - Send an escape message to indicate that there is an error */
             SNDPGMMSG  MSGID(CPF9898) MSGF(QCPFMSG) MSGDTA('Not all +
                          cluster resource groups are active, +
                           review') MSGTYPE(*ESCAPE)
             RETURN
endpgm
```

4.1.6 Monitoring the IASP status

One of the important considerations for the IASP is that it cannot overflow into SYSBAS, so the current utilization is something that you must keep under review.

Although you can set a threshold within System Service Tools (SST), this lets you know only when you exceed the threshold. You can also write a program to use the system API **QYASPOL**, which can provide you with the total and current capacity. It is also possible to obtain the number of LUNs that are involved in the IASP configuration.

By using the information from the system API, it is possible to raise an alert when the utilization grows by more than a certain percentage since the last time the check was run, or if the number of disks in the IASP changed, which might require a corresponding change in the PowerHA configuration.

Note: If you use the sample code in Example 4-5, you must remember to specify the name of an appropriate email address if you plan to use that method of notification.

Example 4-5 shows a sample CL program that can assist you in building your own monitoring tool to verify the IASP utilization status. This CL program is available as a save file. For more information about how to access and download this save file, see Appendix B, “Additional material” on page 127.

Example 4-5 Sample CL program to verify the IASP utilization status

```
/*****************************************************************************/
/* MONIASPSTS- Monitor the IASP status                                       */
/*                                                                           */
/*             This program is intended to provide a simple way of           */
/*             monitoring the status of the IASP, specifically the level     */
/*             of utilisation, and number of disks.                          */
/*                                                                           */
/*             This is achieved by using system APIs to obtain information   */
/*             about the IASP and to alert the operations should any         */
```

```
/*             potential problems exist.                                    */
/*                                                                          */
/*             If problems are identified, then the program                 */
/*             will send a message to the system operator and also send an   */
/*             email to the named user to alert the administrators to the    */
/*             fact that they should investigate the problems.              */
/*                                                                          */
/*            ============================================================  */
/*           I NOTE:  This sample code assumes that there are only 2 (TWO)  I */
/*           I copy descriptions involved in the session. It also assumes   I */
/*           I that only a single IASP exists. It is possible that more     I */
/*           I could exist in which case this sample code would need to be  I */
/*           I be reworked to allow for this possibility. I */
/*            ============================================================  */
/****************************************************************************/
/* Written by  : David Painter                                              */
/* Date Written: Sep 10, 2015                                               */
/* Date Changed: None                                                       */
/****************************************************************************/
pgm

/* Variables for growth calculation, Note this is the percentage growth that
triggers an event */
             DCL        VAR(&GROWTH) TYPE(*DEC) LEN(5 1) VALUE(10.0)

/* QYASPOL variables */
             DCL        VAR(&YASPRCVR) TYPE(*CHAR) LEN(148)
             DCL        VAR(&YASPRCVLEN) TYPE(*UINT) LEN(4) VALUE(148)
             DCL        VAR(&YASPLINF) TYPE(*CHAR) LEN(80)
             DCL        VAR(&YASPREC) TYPE(*INT) LEN(4) VALUE(-1)
             DCL        VAR(&YASPFTRNUM) TYPE(*INT) LEN(4) VALUE(1)
             DCL        VAR(&YASPFTR) TYPE(*CHAR) LEN(16)
             DCL        VAR(&YASPFTRSIZ) TYPE(*UINT) STG(*DEFINED) +
                          LEN(4) DEFVAR(&YASPFTR 1)
             DCL        VAR(&YASPFTRKEY) TYPE(*UINT) STG(*DEFINED) +
                          LEN(4) DEFVAR(&YASPFTR 5)
             DCL        VAR(&YASPFTRDSZ) TYPE(*UINT) STG(*DEFINED) +
                          LEN(4) DEFVAR(&YASPFTR 9)
             DCL        VAR(&YASPFTRDTA) TYPE(*INT) STG(*DEFINED) +
                          LEN(4) DEFVAR(&YASPFTR 13)

/* Break out variables for &YASPRCVR */
             DCL        VAR(&ASPNUM) TYPE(*UINT) STG(*DEFINED) +
                          LEN(4) DEFVAR(&YASPRCVR 1)
             DCL        VAR(&DSKNUM) TYPE(*UINT) STG(*DEFINED) +
                          LEN(4) DEFVAR(&YASPRCVR 5)
             DCL        VAR(&ASPCAPT) TYPE(*UINT) STG(*DEFINED) +
                          LEN(4) DEFVAR(&YASPRCVR 9)
             DCL        VAR(&ASPCAPTA) TYPE(*UINT) STG(*DEFINED) +
                          LEN(4) DEFVAR(&YASPRCVR 13)
             DCL        VAR(&ASPCAPP) TYPE(*UINT) STG(*DEFINED) +
                          LEN(4) DEFVAR(&YASPRCVR 17)
             DCL        VAR(&ASPCAPPA) TYPE(*UINT) STG(*DEFINED) +
                          LEN(4) DEFVAR(&YASPRCVR 21)
             DCL        VAR(&ASPCAPU) TYPE(*UINT) STG(*DEFINED) +
```

```
                    LEN(4) DEFVAR(&YASPRCVR 25)
             DCL    VAR(&ASPCAPUA) TYPE(*UINT) STG(*DEFINED) +
                    LEN(4) DEFVAR(&YASPRCVR 29)
             DCL    VAR(&ASPSTGTHR) TYPE(*UINT) STG(*DEFINED) +
                    LEN(4) DEFVAR(&YASPRCVR 61)

/* Utilisation variables */
             DCL    VAR(&UTILPCT) TYPE(*DEC) LEN(5 1)  /* Percent Util      */
             DCL    VAR(&UTILPCTD) TYPE(*DEC) LEN(5 1) /* Difference in Util*/
             DCL    VAR(&UTILPCTL) TYPE(*DEC) LEN(5 1) /* Last Percent Util */
             DCL    VAR(&IASPUTIL) TYPE(*CHAR) LEN(32)
             DCL    VAR(&LSTUSED) TYPE(*UINT) STG(*DEFINED) +
                    LEN(4) DEFVAR(&IASPUTIL 1)
             DCL    VAR(&LSTCAP) TYPE(*UINT) STG(*DEFINED) +
                    LEN(4) DEFVAR(&IASPUTIL 5)
             DCL    VAR(&LSTDAT) TYPE(*DEC) STG(*DEFINED) LEN(6 +
                    0) DEFVAR(&IASPUTIL 9)
             DCL    VAR(&LSTTIM) TYPE(*DEC) STG(*DEFINED) +
                    LEN(6 0) DEFVAR(&IASPUTIL 15)
             DCL    VAR(&LSTDSK) TYPE(*UINT) STG(*DEFINED) +
                    LEN(4) DEFVAR(&IASPUTIL 21)
             DCL    VAR(&CURUSED) TYPE(*UINT) LEN(4)
             DCL    VAR(&QDATE) TYPE(*CHAR) LEN(6)
             DCL    VAR(&QTIME) TYPE(*CHAR) LEN(6)
             DCL    VAR(&ERRDSK) TYPE(*CHAR) LEN(1) VALUE('N')
             DCL    VAR(&ERRUTIL) TYPE(*CHAR) LEN(1) VALUE('N')

/* Step 1 - Use the QYASPOL YASP0200 to obtain the IASP data */
             CHGVAR     VAR(&YASPFTRSIZ) VALUE(16)
             CHGVAR     VAR(&YASPFTRKEY) VALUE(1)
             CHGVAR     VAR(&YASPFTRDSZ) VALUE(4)
             CHGVAR     VAR(&YASPFTRDTA) VALUE(-3)
             CALL       PGM(QYASPOL) PARM(&YASPRCVR &YASPRCVLEN +
                          &YASPLINF &YASPREC &YASPFTRNUM &YASPFTR +
                          'YASP0200' X'00000000')

/* Step 2 - Calculate the current percentage utilisation */
             CHGVAR     VAR(&CURUSED) VALUE(&ASPCAPT - &ASPCAPTA)
             CHGVAR     VAR(&UTILPCT) VALUE(100.0 * (&CURUSED / +
                          &ASPCAPT))

/* Step 3 - Retrieve the previous utilisation percentage */
             RTVSYSVAL  SYSVAL(QDATE) RTNVAR(&QDATE)
             RTVSYSVAL  SYSVAL(QTIME) RTNVAR(&QTIME)
             RTVDTAARA  DTAARA(MONIASPSTS *ALL) RTNVAR(&IASPUTIL)
             MONMSG     MSGID(CPF1015) EXEC(DO) /* *DTAARA does not +
                          exist, yet */
             CRTDTAARA  DTAARA(MONIASPSTS) TYPE(*CHAR) LEN(32)
             CHGVAR     VAR(&LSTUSED) VALUE(&ASPCAPT - &ASPCAPTA)
             CHGVAR     VAR(&LSTCAP) VALUE(&ASPCAPT)
             CHGVAR     VAR(&LSTDAT) VALUE(000101)
             CHGVAR     VAR(&LSTTIM) VALUE(000000)
             CHGVAR     VAR(&LSTDSK) VALUE(&DSKNUM)
             ENDDO
```

```
/* Step 4 - If the growth of the storage is an increase on the last utilisation,*/
/* then error. The logic here is that if the actual storage used has increased  */
/* by more than the wanted percentage since last run, then an alert is raised. */
/* If the actual capacity has changed. then the process is bypassed.            */
/* Check if installed capacity has changed */
             IF         COND(&ASPCAPT *NE &LSTCAP) THEN(DO)
             GOTO       CMDLBL(CHKDSK)
             ENDDO

             IF         COND((((&CURUSED - &LSTUSED) / &LSTUSED) * +
                          100) *LT &GROWTH) THEN(DO)
             GOTO       CMDLBL(CHKDSK)
             ENDDO

             CHGVAR     VAR(&ERRUTIL) VALUE('Y')

/* Step 5 - Has the number of disks changed */
 CHKDSK:
             IF         COND(&LSTDSK *NE &DSKNUM) THEN(CHGVAR +
                          VAR(&ERRDSK) VALUE('Y'))

/* Step 6 - Send an email to an admin user if there is an error */
             IF         COND(&ERRDSK *EQ 'Y') THEN(DO)
             SNDSMTPEMM RCP((SYSTEMADMINISTRATOR@YOURCOMPANY)) +
                          SUBJECT('The IASP disk config has +
                          changed, verify PowerHA config.') +
                          NOTE('The configuration of your IASP has +
                          changed, verify that your PowerHA +
                          configuration has been modified for these +
                          changes.')
             MONMSG     MSGID(TCP5090 TCP5092)
             ENDDO
             IF         COND(&ERRUTIL *EQ 'Y') THEN(DO)
             SNDSMTPEMM RCP((SYSTEMADMINISTRATOR@YOURCOMPANY)) +
                          SUBJECT('The IASP utilization has grown +
                          by more than expected.') NOTE('The IASP +
                          utilisation has grown by more than the +
                          growth percentage specified in the +
                          monitor program.  You should investigate +
                          the situation to ensure appropriate +
                          actions are taken.')
             MONMSG     MSGID(TCP5090 TCP5092)
             ENDDO

/* Step 7 - Send a message to QSYSOPR if there is an error */
             IF         COND(&ERRDSK *EQ 'Y') THEN(DO)
             SNDPGMMSG  MSGID(CPF9898) MSGF(QCPFMSG) MSGDTA('The +
                          disk configuration has changed for the +
                          IASP, review PowerHA settings') +
                          TOMSGQ(*SYSOPR)
             ENDDO
             IF         COND(&ERRUTIL *EQ 'Y') THEN(DO)
             SNDPGMMSG  MSGID(CPF9898) MSGF(QCPFMSG) MSGDTA('The +
                          IASP utilization has grown by more than +
                          the expected growth percentage,  +
```

```
                        review') TOMSGQ(*SYSOPR)
             ENDDO

/* Step 8 - Update data area with the current information */
             CHGVAR     VAR(&LSTUSED) VALUE(&ASPCAPT - &ASPCAPTA)
             CHGVAR     VAR(&LSTCAP) VALUE(&ASPCAPT)
             CHGVAR     VAR(&LSTDAT) VALUE(&QDATE)
             CHGVAR     VAR(&LSTTIM) VALUE(&QTIME)
             CHGVAR     VAR(&LSTDSK) VALUE(&DSKNUM)
             CHGDTAARA  DTAARA(MONIASPSTS *ALL) VALUE(&IASPUTIL)

/* Step 9 - Send an escape message to indicate that there is an error */
             IF         COND((&ERRDSK *EQ 'Y') *OR (&ERRUTIL *EQ +
                          'Y')) THEN(DO)
             SNDPGMMSG  MSGID(CPF9898) MSGF(QCPFMSG) MSGDTA('Errors +
                          were found with the IASP utilisation or +
                          configuration') MSGTYPE(*ESCAPE)
             ENDDO
             RETURN
endpgm
```

4.1.7 Other monitoring considerations

You should investigate other methods of monitoring. Here is a list of some typical areas to monitor:

- Network switches
- Power service
- SAN switches
- Storage subsystems

Any of these areas can cause system failures and impact your ability to use the systems.

Often, these areas (and others) can be monitored by the use of SNMP monitoring software. SNMP is an industry standard that allows a device to raise a trap, which is a notification that is broadcast on your network. An SNMP monitor can then receive the trap and decode it. This allows the monitor to raise alerts in real time.

For example, a storage subsystem typically raises a trap should it lose connectivity to a remote storage subsystem, or should a physical component fail internally.

4.2 Managing your environment

Although little management is required for a PowerHA environment, there are some areas that you must consider. The following areas are covered in this section:

4.2.1 Adding disks to an IASP in a PowerHA environment

You should be careful when adding logical unit numbers (LUNs) to a PowerHA environment because there is a particular sequence that must be followed.

Important: It is important that when you start this procedure that there is no replication in place for the new LUNs.

To add disks to an IASP in a PowerHA environment, complete the following steps:

1. Run the Work with Cluster (**WRKCLU**) command to ensure that all nodes are active in the cluster.
2. Configure the new volumes and attach them to the various nodes in the cluster by completing the following steps. The new volumes should appear as non-configured disk units in SST.
   a. Run **STRSST**.
   b. Sign on.
   c. Select option 3 - Work with disk units.
   d. Select option 1 - Display disk configuration.
   e. Select option 4 - Display non-configured units.
   f. Exit SST
3. To initialize and format the new volumes (DS8000 installations production node only; Storwize installations run on all nodes), complete the following steps:
   a. Run **STRSST**.
   b. Sign on.
   c. Select option 3 - Work with disk units.
   d. Select option 3 - Work with disk unit recovery procedures.
   e. Select option 2 - Disk unit problem recovery procedures.
   f. Select option 1 - Initialize and format disk unit.
   g. Select option 1 on all the new volumes, press Enter, and then press F10 to confirm. Wait for the `Initialize and format completed successfully` message.
   h. Exit SST.
4. Start replication on the LUNs as required (Metro Mirror or Global Mirror).
5. Wait for the replication to complete and then finish any Global Mirror configuration if needed.
6. Ensure that any FlashCopy is ended.
7. To add the new volumes to the IASP on the production node only, complete the following steps:
   a. Run **STRSST**.
   b. Sign on.
   c. Select option 3 - Work with disk units.
   d. Select option 2 - Work with disk configuration.
   e. Select option 2 - Add units to ASP.
   f. Select option 3 - Add units to existing ASP.

g. Specify the IASP number for each new volume and press Enter.

h. Press Enter to confirm your choice for Add units.

i. Wait for the `Selected units have been added successfully` message.

j. Exit SST.

8. To end any existing auxiliary storage pool (ASP) sessions, complete the following steps:

a. Run **WRKCLU**.

b. Select option 10 - Work with ASP Copy Descriptions.

c. Identify any Copy Descriptions that are being updated. If a session is present, select option 25 - Display Session. Make a note of the settings and then select option 24 - End Session.

d. Repeat until your Copy Descriptions no longer have any sessions active against them

9. To add the new LUN ranges to the production source and target copy descriptions, complete the following steps:

a. Run **WRKCLU**.

b. Select option 10 - Work with ASP Copy Descriptions.

c. Select option 2 - Change copy and press F4.

d. Add the new LUN ranges to the existing ranges.

e. Press Enter.

10. To add the new LUN ranges to any FlashCopy copy descriptions, complete the following steps:

a. Run **WRKCLU**.

b. Select option 10 - Work with ASP Copy Descriptions.

c. Select option 2 - Change copy and press F4.

d. Add the new LUN ranges to the existing ranges.

e. Press Enter.

11. To restart any ASP sessions that you ended, complete the following steps:

a. Run **WRKCLU**.

b. Select option 10 - Work with ASP Copy Descriptions.

c. Using the information that you gathered in step 8, select option 21 - Start Session.

d. Repeat these steps until your Copy Descriptions have the correct sessions active against them.

4.2.2 Removing disks from an IASP in a PowerHA environment

You should be careful when removing LUNs from a PowerHA environment because there is a particular sequence that must to be followed. Complete the following steps:

1. To ensure that all nodes are active in the cluster, run the **WRKCLU** command.
2. To remove certain volumes from the IASP on the production node only, complete the following steps:

a. Run **STRSST**.

b. Sign on.

c. Select option 3 - Work with disk units.

d. Select option 2 - Work with disk configuration.

e. Select option 12 - Work with removing units from configuration.

f. Select option 3 - Remove units from configuration.

g. Carefully select the disk units to remove with by selecting option 4 and pressing Enter.

h. Press Enter to confirm your choice for removing units.

i. Wait for the `Selected units have been removed successfully` message.

j. Exit SST.

3. Ensure that any FlashCopy operations are ended.

4. To end any existing ASP sessions, complete the following steps:

   a. Run **WRKCLU**.

   b. Select option 10 - Work with ASP Copy Descriptions.

   c. Identify any Copy Descriptions that are being updated. If a session is present, select option 25 - Display Session. Make a note of the settings and then select option 24 - End Session.

   d. Repeat until your Copy Descriptions no longer have any sessions active against them.

5. To remove the LUN ranges from the production source and target Copy descriptions, complete the following steps:

   a. Run **WRKCLU**.

   b. Select option 10 - Work with ASP Copy Descriptions.

   c. Select option 2 - Change copy and press F4.

   d. Remove the LUN ranges from the existing ranges.

   e. Press Enter.

6. To remove the LUN ranges from any FlashCopy copy descriptions, complete the following steps:

   a. Run **WRKCLU**.

   b. Select option 10 - Work with ASP Copy Descriptions.

   c. Select option 2 - Change copy and press F4.

   d. Remove the LUN ranges from the existing ranges.

   e. Press Enter.

7. To restart any ASP sessions that you ended, complete the following steps:

   a. Run **WRKCLU**.

   b. Select option 10 - Work with ASP Copy Descriptions.

   c. Using the information that you gathered in step 8 on page 87, select option 21 - Start Session.

   d. Repeat these steps until your Copy Descriptions have the correct sessions active against them.

8. End replication on the LUNs as required (Metro Mirror or Global Mirror).

9. To verify that the old volumes should appear as non-configured disk units in SST, complete the following steps:

   a. Run **STRSST**.

   b. Sign on.

c. Select option 3 - Work with disk units.

d. Select option 1 - Display disk configuration.

e. Select option 4 - Display non-configured units.

f. Exit SST.

10. Unconfigure the now removed LUNs as needed within the storage unit.

4.2.3 Adding users to the systems

When you must add new users to the systems, they must be also added to the Cluster Administrative Domain.

Depending on the frequency with which you add users, you can use either of the following options:

- To add a user to the Cluster Administrative Domain, run the Add Cluster Admin Domain MRE (**ADDCADMRE**) command.

  You must write a simple CL program that runs the **ADDCADMRE** command for you automatically.

- To automatically add the new user to the list of MREs, run the system-supplied exit point on the Create User Profile (**CRTUSRPRF**) command.

  The exit point is called QIBM_QSY_CRT_PROFILE. The documentation for this exit point can be found in the IBM Knowledge Center:

  http://www.ibm.com/support/knowledgecenter/ssw_ibm_i_72/apis/XCRTUP.htm?lang=en

  Important: When writing a program for the QIBM_QSY_CRT_PROFILE exit point, you should be aware that when the creation of the user profile occurs on other nodes in the partition, it also calls the exit program. Therefore, you should include a check to see whether the user that is running the program is QSYS. If so, you should terminate with no further action.

4.2.4 Adding network printers and other SYSBAS objects to the systems

There are many objects in SYSBAS that can be replicated within the Cluster Administrative Domain. How you manage these various objects varies depending on a number of factors.

For things such as network printers, these are not usually created frequently, so adding them manually after creation is often the sensible approach. For other objects, you might have a program that creates them in a specific way, in which case it would make sense to change that program to include the Add Cluster Admin Domain MRE (**ADDCADMRE**) command.

Whichever route you choose to go, keep in mind that some objects “happen to be the same” and should be managed under change control, rather than “must be the same” and should be replicated.

4.2.5 Upgrading your operating system

One of the benefits of PowerHA is the ability to upgrade your operating system, apply Program Temporary Fixes (PTFs) or Technology Refreshes with less production impact than without PowerHA. There are two different scenarios involved here.

- PTFs and Technology Refreshes
- Operating system upgrades

Although both scenarios benefit from the use of PowerHA, the way you manage the upgrades differs slightly.

PTFs and Technology Refreshes

With the PTFs and Technology Refreshes scenario, you can simply install the updates to the current non-production node. After you have correctly installed the updates, you can then schedule a switch at a convenient time so you can use the new updates. If you experience a problem, you can simply switch back again until you resolve the problem.

After you are comfortable that the updates are working well, you can perform the updates on the remaining node.

Important: If you are using geographic mirroring and the updates you want to apply require an IPL, you should suspend replication before starting the process.

Operating system upgrades

The operating system upgrades scenario is slightly more complex because once an IASP is updated to a new operating system level, it cannot be used on an older level partition. Therefore, it is necessary to use a different approach to the upgrade.

You can perform the operating system upgrade and associated PTFs on the non-production node without impacting production. Rather than using a switch to use the new code level, you should vary off the IASP and then issue a detach of the replication by running the following Change Auxiliary Storage Pool Session (**CHGASPSSN**) command, where *sessionname* is the name of the replication session.

```
CHGASPSSN SSN(sessionname) OPTION(*DETACH)
```

This makes the replication target usable, and you can vary on the IASP on the newly upgraded node and start testing.

The overall process consists of the following steps:

1. Vary off the production IASP copy.
2. Run the **CHGASPSSN** command with **OPTION(*DETACH)**.
3. Vary on the IASP copy on the upgraded system and test.

Restriction: If you are using geographic mirroring, you must complete the operating system upgrade on the original production system before attempting to restart the replication.

4. When ready to go into production, vary off the upgraded copy.
5. Run an End Cluster Resource Group (**ENDCRG**) command.
6. Run a Change Cluster Resource Group (**CHGCRG *CHGCUR**) command and change the primary and backup nodes.

7. Run a Start Cluster Resource Group (**STRCRG**) command.
8. Run a **CHGASPSSN** with **OPTION(*REATTACH)** command, which should be from the upgraded node to the old production node.
9. Vary on the IASP on the upgraded node and replication should restart. Under some conditions, you might need to resume the replication.

Important: After you upgrade the operating system, apply Technology Refreshes, or PowerHA group PTFs, you must check that the actual cluster version and PowerHA version match the new potential cluster version and PowerHA version. This is done by running the Change Cluster Version (**CHGCLUVER**) command, which can increment versions or modification levels if all of the nodes can operate at the new level. Failure to follow this process can ultimately result in being unable to start your cluster nodes and use PowerHA.

4.2.6 Powering down your system

When powering down a cluster node, follow the proper sequence to ensure that clustering is stopped correctly (in addition to any existing power-down procedures).

If a node is being powered down, the cluster recognizes the node as failed. If the node is the primary location for resources that are controlled by an active CRG, the point of access for the resources is moved to a backup node. Moving the point of access to a backup node might not be the wanted action. Therefore, it is preferable to establish a procedure for a PWRDWNSYS operation. Complete the following steps:

1. Notify users that a power down is imminent (existing process).
2. Take actions to prepare any IASPs for vary off / switch.
3. Vary off IASPs.
4. If the node being powered down is the primary node for an active CRG:
   a. Run a Change CRG Primary (**CHGCRGPRI**) command to move the active resources to a node that remains active in the cluster.
   b. Run an **ENDCRG** command to prevent the resources from moving to a node that remains active in the cluster.
5. Run an End Cluster Node (**ENDCLUNOD**) command for the node being powered down.
6. Power down the node.

4.2.7 Verifying the health of your replication solution

There are times when you want to look at the overall health of your replication solution. With a PowerHA environment, you can do so by using the GUI shown in Figure 4-2.

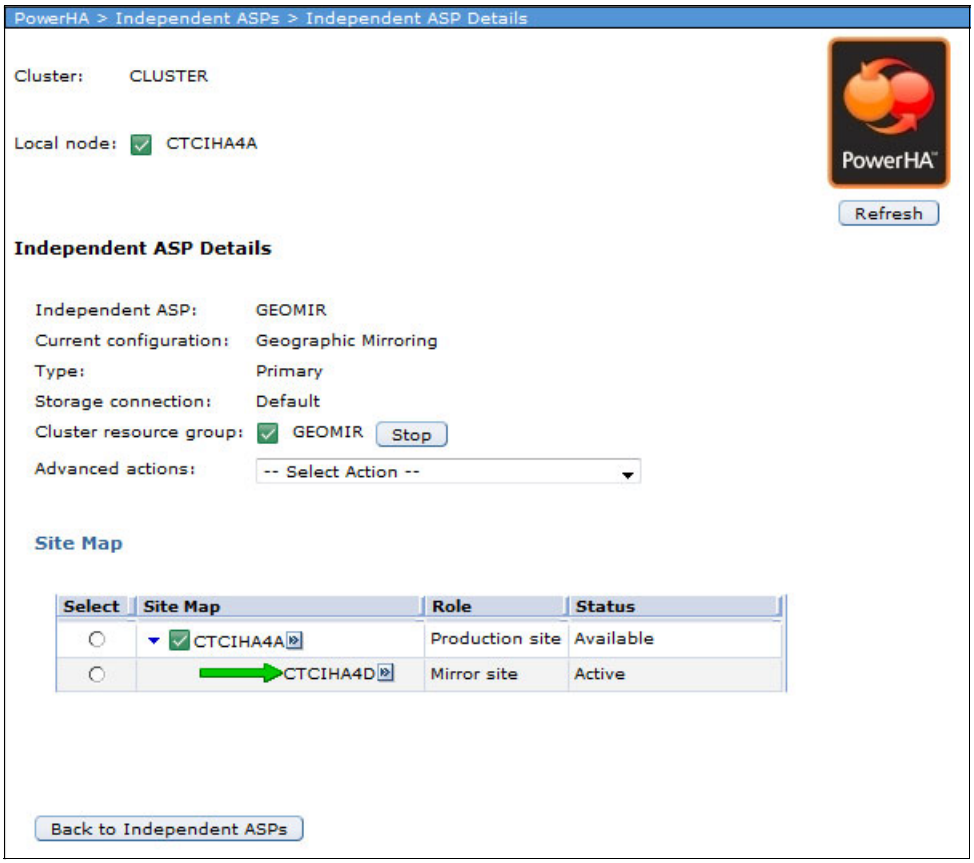

Figure 4-2 PowerHA GUI view of the health of the replication

As shown in Figure 4-2, there are multiple indicators of the health of your replication. At the top of the window, the local node shows a green check makr, which indicates that PowerHA is operational on this node. Further down, you can see that the CRG has a green check makr, which indicates that the CRG is active. Finally, you see the green arrow in the list of nodes showing that the replication is working and is going from CTCIHA4A to CTCIHA4D.

If anything is not green, this indicates that something is not correct and you should investigate.

Tip: You can hover your cursor over the color indicator and a message appears describing the condition.

5

Switching and IBM FlashCopy operations

This chapter describes the operation of a PowerHA for IBM i environment that uses the IBM Storwize family storage products and their remote copy and FlashCopy options, which includes switching to the current backup system, handling various failover scenarios, and using FlashCopy with Backup, Recovery, and Media Services for IBM i (BRMS).

The following topics are described in this chapter:

5.1 Planned switching

A planned switch is normally performed when you either want to do a maintenance action on your current primary system or as part of a regular role swap. The following steps should be taken to assure that the switch over runs as fast as possible:

1. In addition to your normal regular and automatic monitoring as described in 4.1, “Monitoring your environment” on page 68, you should also check that your cluster environment is running and that the data is in sync between the cluster nodes before you start the switchover. This task can be done by using the cluster GUI. Open a web browser session to the IBM Navigator for i and select **PowerHA**:

   `http://<yoursystemname>:2001`

   This action displays the PowerHA overview window that is shown in Figure 5-1.

   This view provides you with an overview of your entire cluster status. If any of the icons do not show a green check mark status, make sure to fix the potential problem before proceeding with the switch.

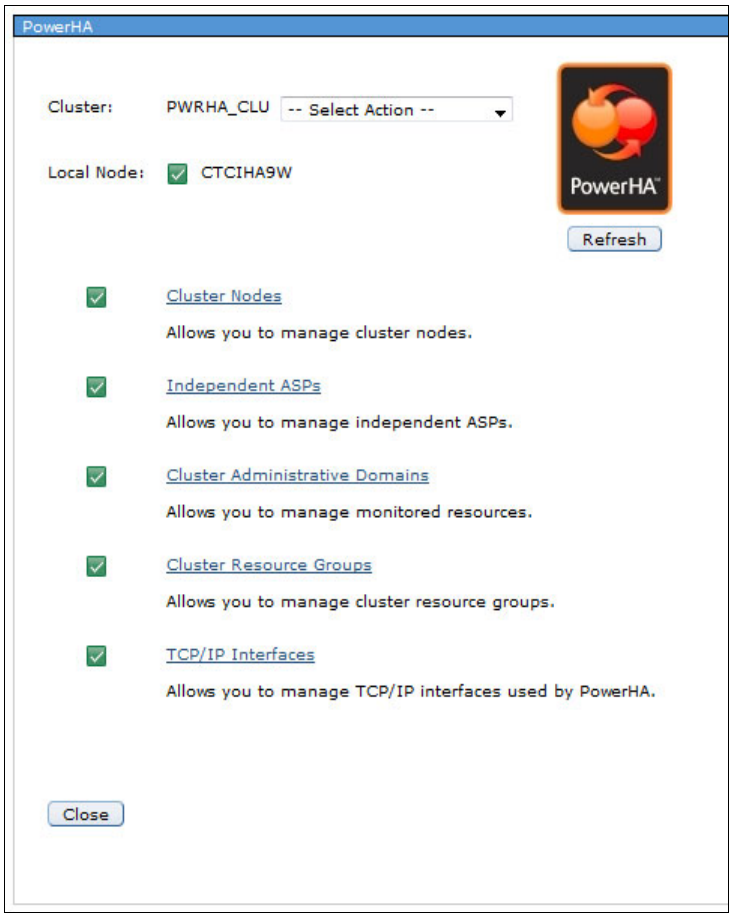

Figure 5-1 IBM Navigator for i - PowerHA overview window

Alternatively, you can use the following cluster CL commands:

- To check that clustering and your nodes are active, run the Display Cluster Information (**DSPCLUINF**) command.
- To check that the Administrative Domain is in a consistent status, run the Print Cluster Administrative Domain Monitored Resource Entry (**PRTCADMRE**) command.
- To check that the cluster resource group (CRG) is active and the nodes show up in the recovery domain as active, run the Display Cluster Resource Information (**DSPCRGINF**) command.
- To check that the SAN Volume Controller session shows a replication state of active, run the Display Storage Area Network (SAN) Volume Controller Auxiliary Storage Pool Session (**DSPSVCSSN**) command.

2. End your applications and get all the users off the system. You can check whether there are still jobs connected to the independent auxiliary storage pool (IASP) by running the Work with ASP Job (**WRKASPJOB**) command. If there are still jobs that are connected, the output looks like what is shown in Figure 5-2.

```
                          Work with ASP Jobs                      CTCIHA9V
                                                                09/10/15
11:46:35
 Type options, press Enter.
   4=End job   5=Work with   7=Send message

 Opt     ASP/Job          User           Number     Type        Status
         V7KNPIVGM
           QPADEV0001     TESTUSER       657148      INT        RUN



Bottom
 Parameters or command
 ===>
 F3=Exit   F4=Prompt   F5=Refresh   F9=Retrieve   F12=Cancel
```

Figure 5-2 WRKASPJOB output screen

3. Make sure that your job is not currently connected to the IASP by running the following command:

```
SETASPGRP ASPGRP(*NONE)
```

Note: If your job is connected to the IASP while it runs the switchover process, it gets disconnected during the switchover process, which might lead to unpredictable results for the switchover process. Specifically, Metro Mirror or Global Mirror tend to end up in a detached state as a result.

4. Vary off the IASP by running the following command:

```
VRYCFG CFGOBJ(IASP01) CFGTYPE(*DEV) STATUS(*OFF)
```

For special considerations when using a Global Mirror setup, see 5.1.1, “Special considerations when switching in a Global Mirror setup” on page 98.

5. Perform the switchover by using the IBM Navigator for i PowerHA GUI. From the main PowerHA overview window (Figure 5-1 on page 94), click **Cluster Resource Groups**.

6. In the **PowerHA** → **Cluster Resource Groups** window, select **Switch** for the correct CRG, as shown in Figure 5-3.

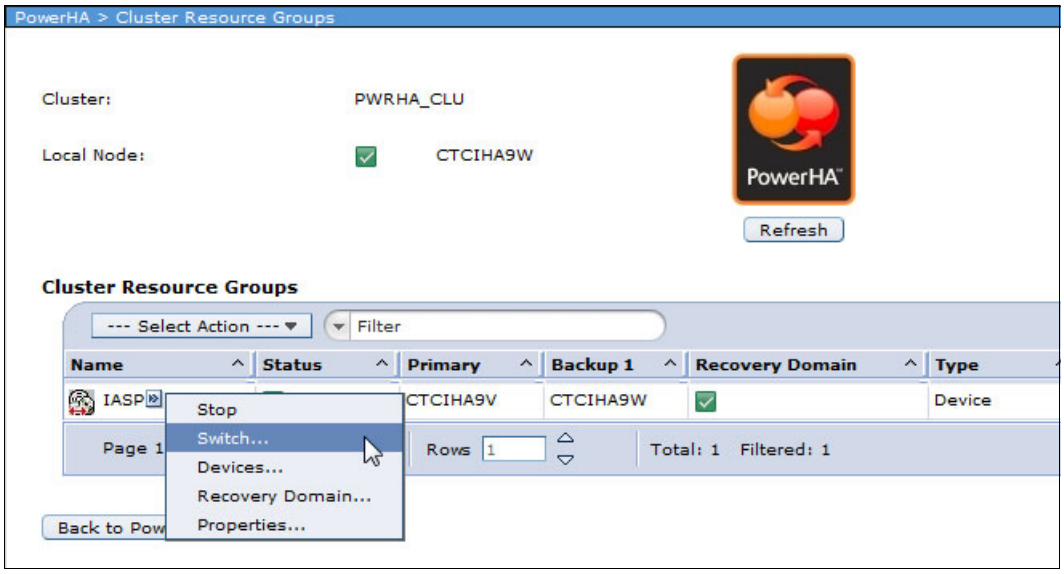

Figure 5-3 Performing the switchover from the IBM Navigator for i PowerHA GUI

7. A confirmation window opens. You can find information about the current recovery domain and the recovery domain after the switch and information about the IASP that is switched over, as shown in Figure 5-4. Click **OK** to initiate the switchover.

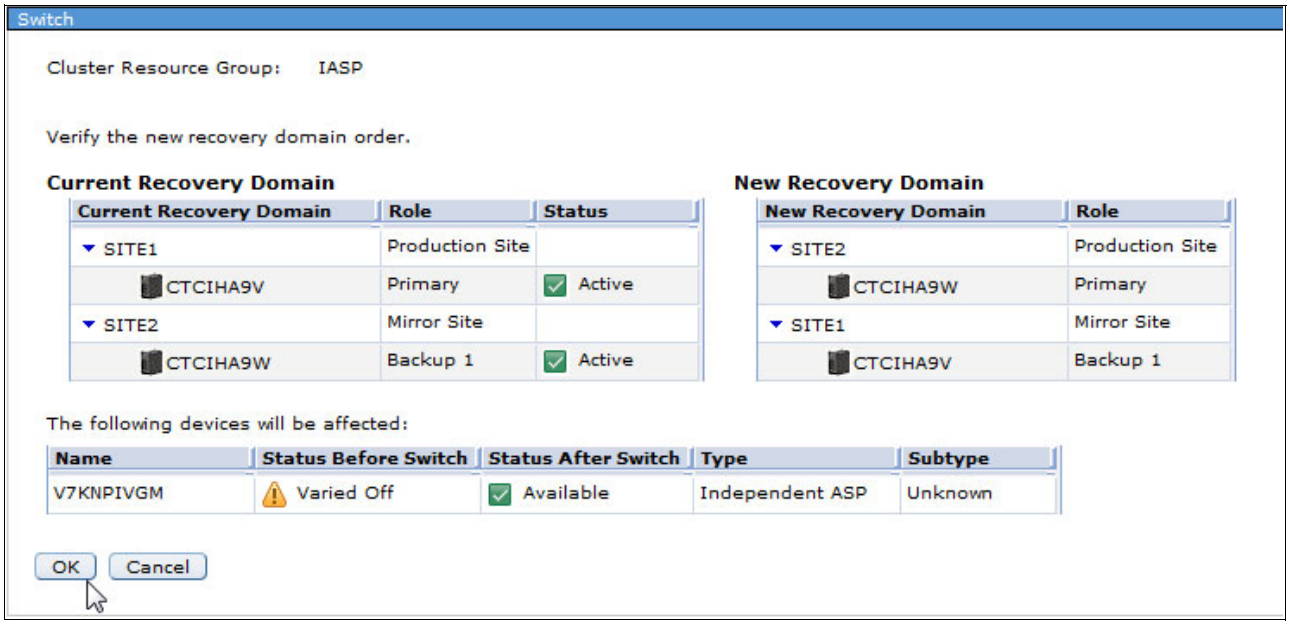

Figure 5-4 Confirmation of performing the switchover

Alternatively, you can use the following CL command to initiate the switchover:

```
CHGCRGPRI CLUSTER(PWRHA_CLU) CRG(IASP)
```

The switchover process performs the following functions:

a. Varies off the IASP on the primary cluster node, if it is not already varied off.

b. If a server takeover IP address is defined in the CRG, then this IP interface is ended on the current primary cluster node.

c. The recovery domain information in the CRG is updated to show the new current primary and backup node.

d. If switchover reverse replication is configured with the default value of *YES in the auxiliary storage pool (ASP) session, then the remote copy direction is turned on the Storwize system.

e. If switchover reverse replication is configured as *NO in the ASP session, then the remote copy is detached on the Storwize system. When you want to restart the replication, you must run the following command:

```
CHGSVCSSN <session name> OPTION(*RESUME)
```

f. On the new primary node, the IASP is varied on if this is configured in the CRG device entry.

g. On the new primary node, it starts the server takeover IP address if one is defined in the CRG.

8. You can check the status of the vary on process by running the Display ASP Vary Status (**DSPASPSTS**) command on the new primary system. During the vary on process, several steps are performed, as shown in Figure 5-5.

```
                          Display ASP Vary Status

ASP Device . . . . :   V7KNPIVGM         Step . . . . . . . . :   11 / 34
ASP Number . . . . :   183               Current time . . . :   00:00:04
ASP State  . . . . :   VARIED OFF        Previous time  . . :   00:02:06

      Step                                                    Elapsed time
      Scanning DASD pages
      Directory recovery - permanent directory
      Authority recovery                                        00:00:00
>     Context rebuild                                           00:00:02
      Journal recovery
      Database recovery
      Journal synchronization



Press Enter to continue

F3=Exit   F5=Refresh   F12=Cancel   F19=Automatic refresh
```

Figure 5-5 DSPASPSTS output

After the vary on process is finished, the ASP state changes to AVAILABLE.

The GUI interface also provides status updates, as shown in Figure 5-6.

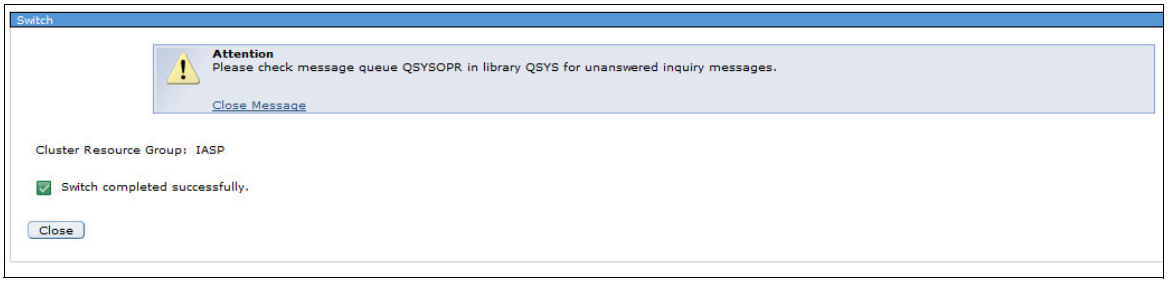

Figure 5-6 Switch completed successfully

9. If these steps are not done automatically, vary on the IASP and the IP addresses that are needed for user access manually, and start your applications.

5.1.1 Special considerations when switching in a Global Mirror setup

This section describes special considerations when switching in a Global Mirror setup. The following topics are described:

- "Global Mirror without change volumes" on page 98
- "Global Mirror with Change Volumes" on page 101

Global Mirror without change volumes

After you end your applications and vary off the IASP, make sure that all changes were synchronized between the source and the target volumes, which can be verified by running the following steps from the IBM Navigator for i PowerHA GUI:

1. From the main PowerHA GUI overview window (Figure 5-1 on page 94), click **Independent ASPs**.
2. In the Independent ASPs window (Figure 5-7), select **Details** for the IASP you want to switch to the current backup system.

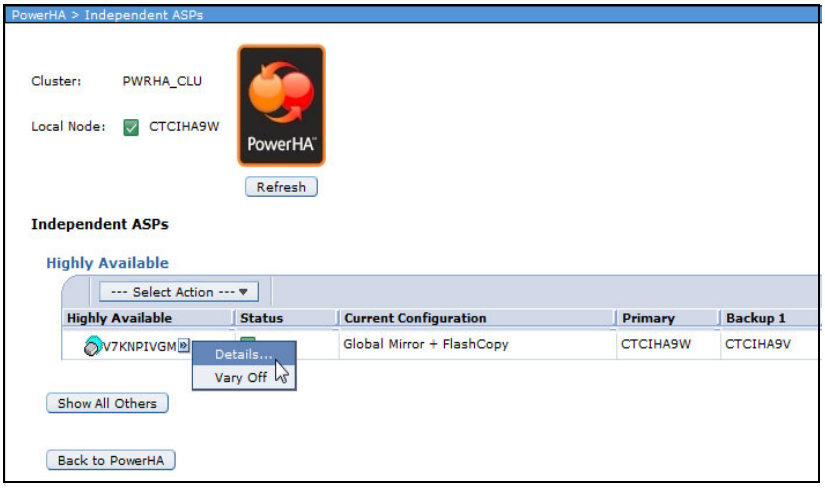

Figure 5-7 Select IASP details

3. In the Independent ASP Details window, from the Advanced Actions drop-down menu, select **ASP Sessions and Copy Descriptions**, as shown in Figure 5-9 on page 99.

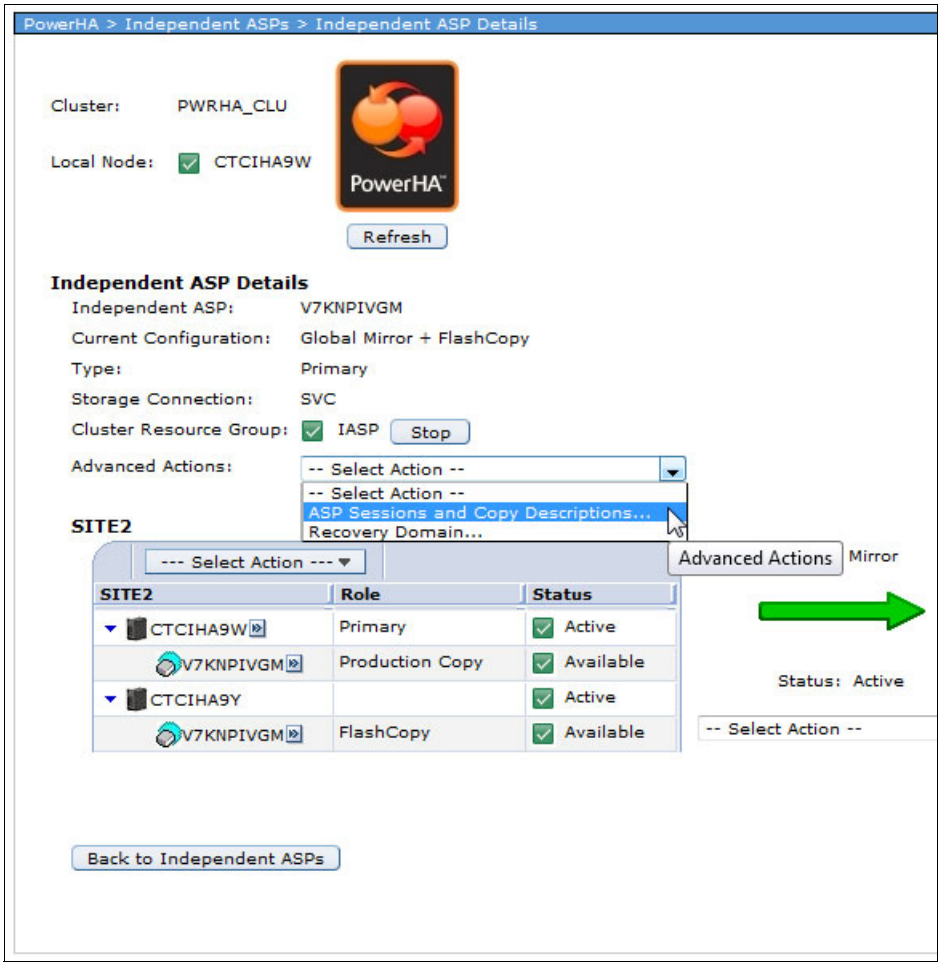

Figure 5-8 Independent ASP details - select ASP Session and Copy Descriptions

4. In the ASP Sessions and Copy Descriptions window, for the IASP session in question, select **Properties**, as shown in Figure 5-9.

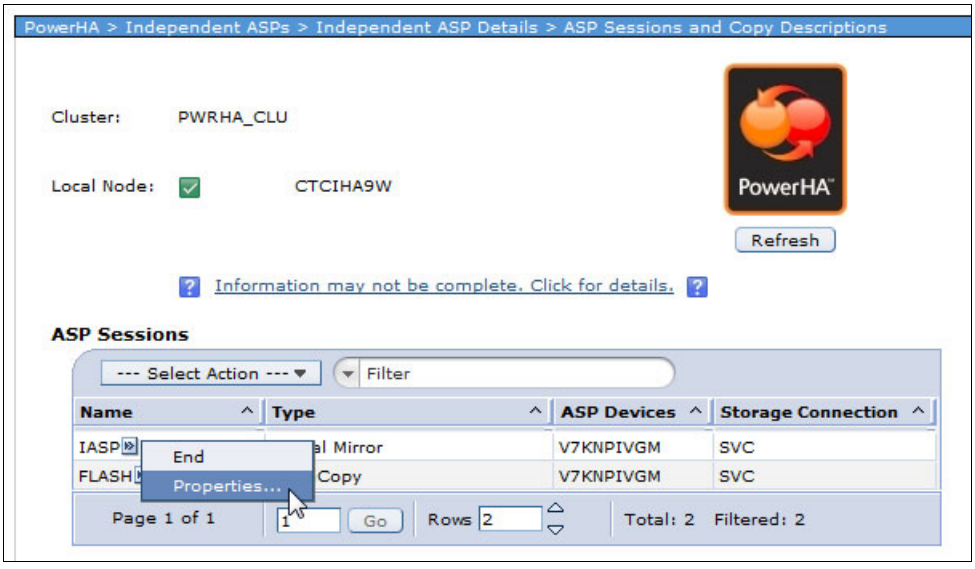

Figure 5-9 Select the IASP session properties

5. As shown in Figure 5-10, you can observe the copy progress. If the value that is shown here is 100, then all data was synchronized between the Global Mirror source and the Global Mirror target.

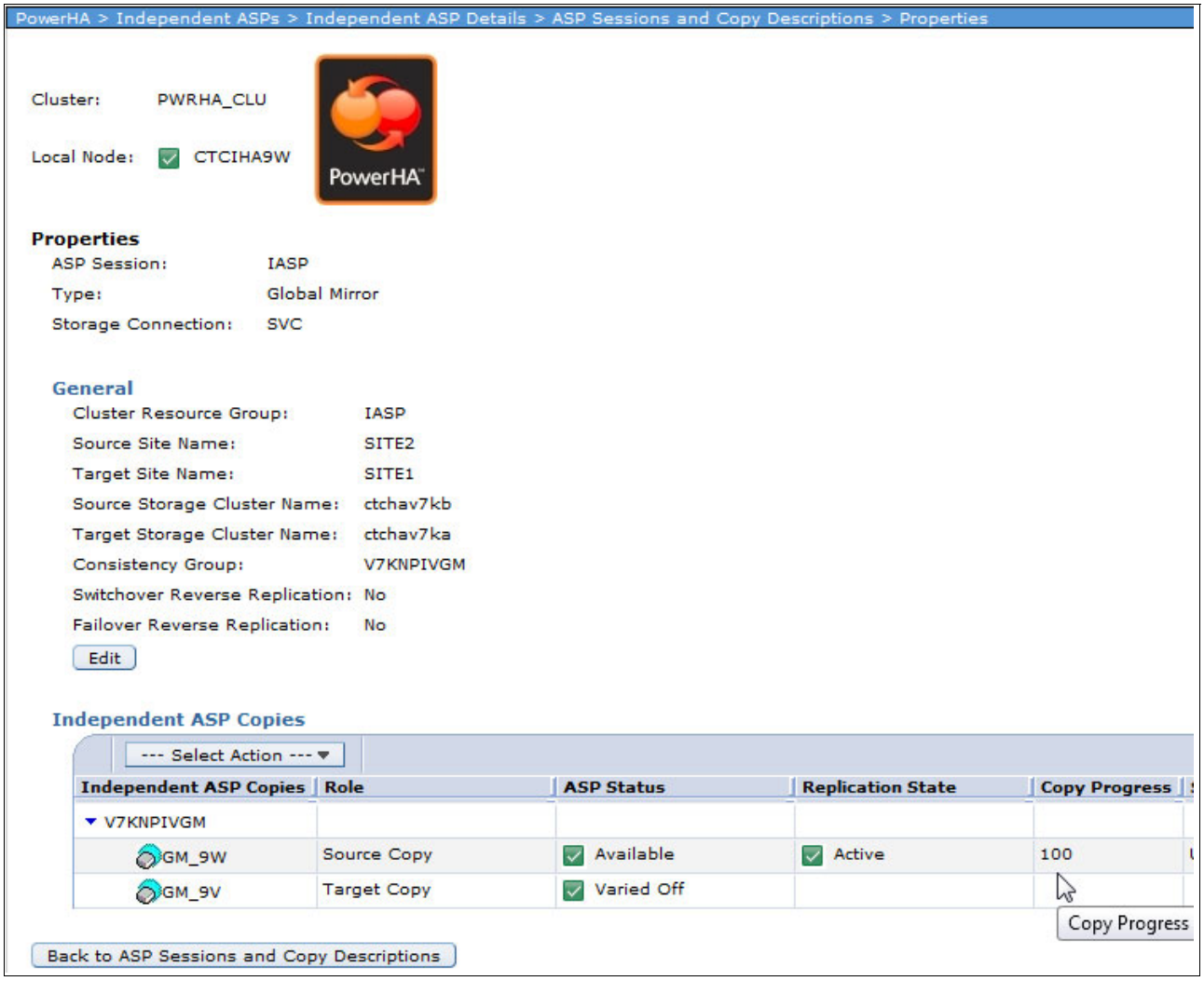

Figure 5-10 Independent auxiliary storage pool session properties showing the copy progress

Alternatively, you can run the Display SVC Session (**DSPSVCSSN**) command. Pressing F11 to open View 2 shows the copy progress. A value of 100, as shown in Figure 5-11 on page 101, means that the source volumes and the target volumes are identical.

```
                            Display SVC Session                  CTCIHA9W

Session . . . . . . . . . . . . . . . . . . . . :   IASP
Type  . . . . . . . . . . . . . . . . . . . . . :   *GLOBALMIR

Switchover reverse replication  . . . . . . . . :   *NO
Failover reverse replication  . . . . . . . . . :   *NO
Consistency group . . . . . . . . . . . . . . . :   V7KNPIVGM
Source storage cluster name . . . . . . . . . . :   ctchav7kb
Target storage cluster name . . . . . . . . . . :   ctchav7ka
                                                                    More...
                              Copy Descriptions

ASP                                 Copy
device      Role      Node      progress  Storage state
V7KNPIVGM   SOURCE    CTCIHA9W    100     UNKNOWN
            TARGET    CTCIHA9V



                                                                   Bottom
Press Enter to continue

F3=Exit   F5=Refresh   F11=View 1   F12=Cancel
```

Figure 5-11 Run the DSPSVCSSN command to check the copy progress

Global Mirror with Change Volumes

After the required PowerHA PTFs are applied, IBM PowerHA SystemMirror for i recognizes that Global Mirror with Change Volumes (GMCV) is configured and manages the replication environment.

Note: If GMCV is configured and the PTFs are not applied on both the source and target systems, the periodic health check fails, and any planned or unplanned switch also fails.

To ensure that no data is lost in a planned switch, the following restriction applies for a GMCV relationship:

If replication is active, PowerHA runs checks to ensure that the data on source and target is equivalent. There might be a delay in the switchover process as PowerHA waits for equivalence before completing the switch. If at some point in the switchover process the synchronization is suspended, the switchover fails.

You can run the **DSPSVCSSN** command to estimate the switchover/failover time. The **DSPSVCSSN** command prints a DIAG message to the job log, giving an approximation of the data out of sync and the time based on current estimated transfer rate. Because of the data that is returned from the storage server, these estimates tend to be more accurate for environments consisting of many small volumes, rather than a few large volumes.

5.1.2 Considerations for three node clusters

When working with a cluster setup with more than two nodes in the recovery domain, performing a switch might have unexpected results on the order of systems in the recovery domain.

Given the following setup:

- System A has a current role of primary.
- System B has a current role of backup 1.
- System C has a current role of backup 2.

The A switch in the recovery domain (**CHGCRGPRI**) promotes System B to become the new primary, but System A does not become the new backup 1. Instead, it is moved to the last position in the recovery domain. Therefore, after performing a role switch in this example, the new recovery domain is as follows:

- System B has a current role of primary.
- System C has a current role of backup 1.
- System A has a current role of backup 2.

This situation is important if you consider a setup where System A and System B are local systems (that use, for example, LUN-level switching) and System C is a remote system (for example, that uses Metro Mirror or Global Mirror from System B). After a failure of system A, you cannot simply switch back to system A by running another **CHGCRGPRI** command because this action promotes System C to the new primary.

Therefore, you might want to change the order of nodes in the recovery domain by completing the following steps:

1. From the main PowerHA overview window (Figure 5-1 on page 94), click **Cluster Resource Groups**.
2. In the Cluster Resource Groups window, select **Recovery Domain** for the CRG you are working with, as shown in Figure 5-12.

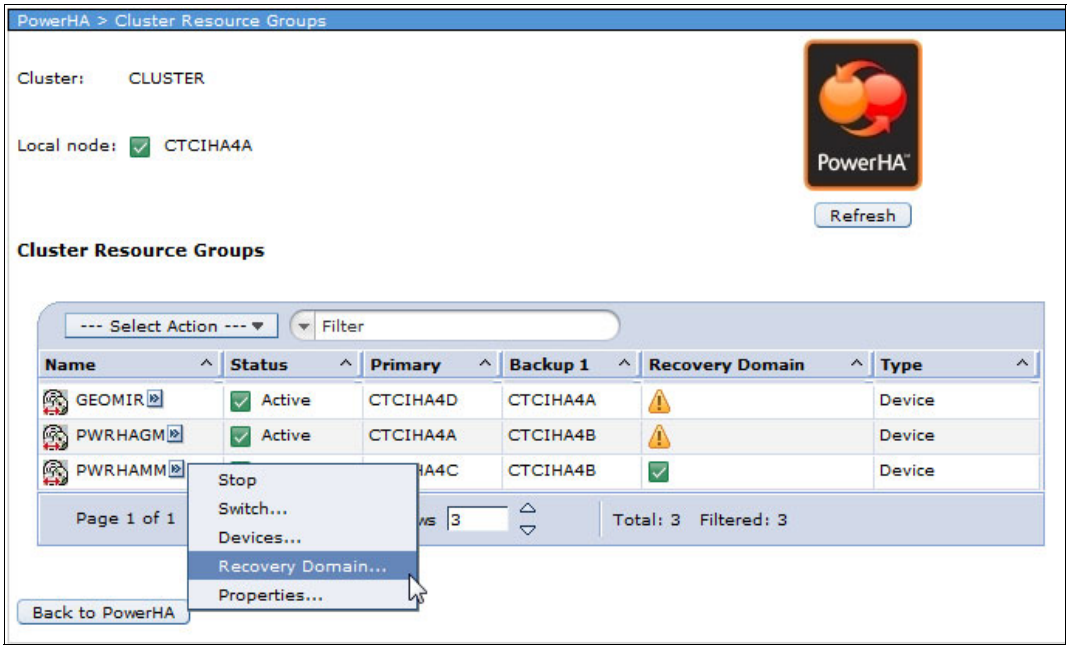

Figure 5-12 Select to change the recovery domain configuration

3. In the Recovery Domain window, select **Change Node Roles** from the Select Action drop-down menu, as shown in Figure 5-13.

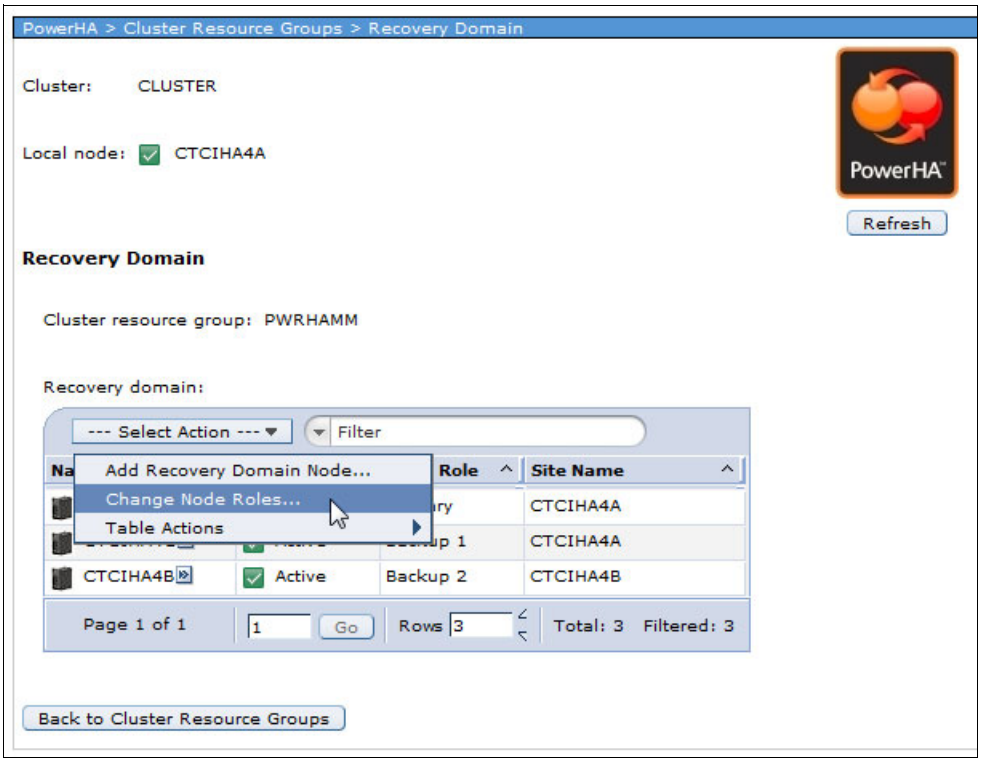

Figure 5-13 Select to change the recovery domain

4. Select the node that you want to change the node role for and click **Move Up** or **Move Down**, as shown in Figure 5-14. When the nodes are in the expected order in the recovery domain, click **Save** to save your changes.

Important: Make sure that you change only the order of the *backup* nodes by using this procedure. Changing the primary node role changes only the role of the system within the recovery domain and does not change the replication direction on the storage system and potentially rendering your solution unusable.

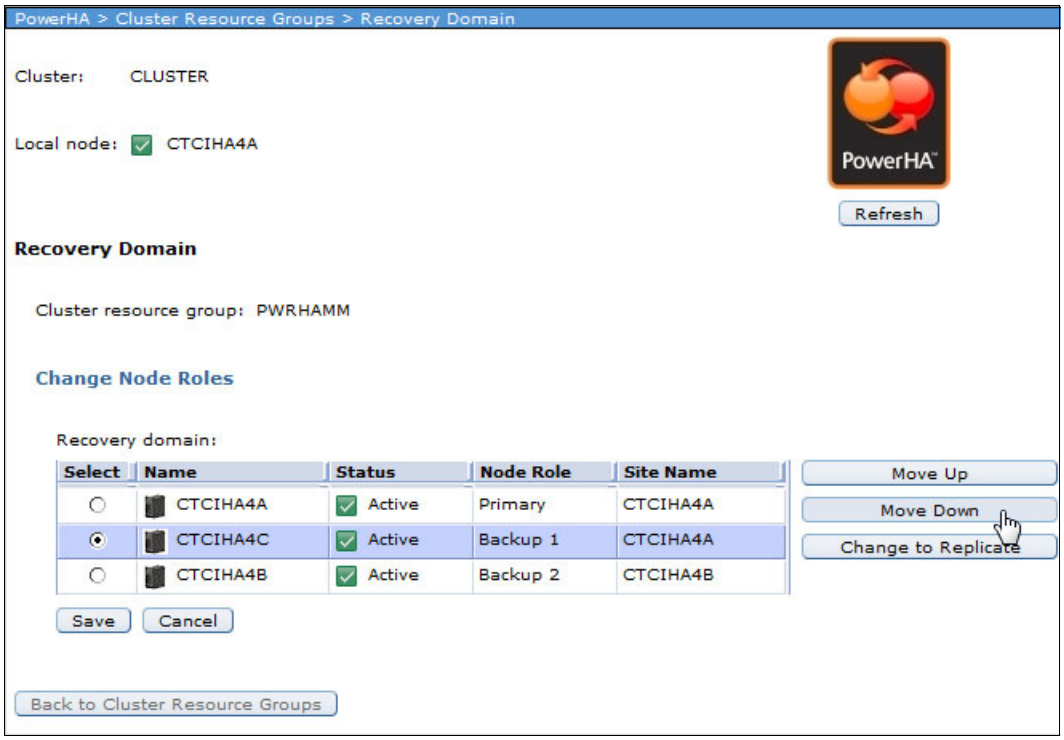

Figure 5-14 Change nodes roles

5.2 Failover after a production system outage

A failover occurs when the current production node unexpectedly fails in some manner and the backup node takes over. The default failover action or procedure depends on the cluster failover wait time and failover default action settings.

Important: As with a planned switch, the success of a failover operation also depends on prior testing and verification that business applications can run on the backup node and users can access those applications.

Consider the following failure scenarios:

- A primary node failure triggering an automatic failover, which can occur either by a panic message sent to the current backup node or cluster monitors in place.
- A primary node or cluster communications failure resulting in node partition status.

Each of these scenarios requires different failover and recovery actions, which are described in this section.

Primary node failure triggering an automatic failover event

An unplanned automatic failover event is triggered for a switchable IASP in a cluster resource group for a primary node failure event that is detected by cluster resource services either by a panic message that is sent by the failing primary node, as for an action of ending a cluster node or powering down the system, or by a power state change event that is sent by the HMC CIMOM server for a partition failure to the registered IBM i cluster nodes when using advanced node failure detection, as described in 4.3, “Configuring advanced node failure detection”, in *IBM PowerHA SystemMirror for i: Preparation (Volume 1 of 4)*, SG24-8400

For an unplanned automatic failover event, the following CPABB02 inquiry message is sent to the cluster or CRG message queue on the *backup* node if a failover message queue is defined for either the cluster or the CRG. If no failover message queue is defined, then the failover starts immediately without any message being posted.

```
Cluster resource groups are failing over to node <nodename> (C / G)
```

With the default cluster parameters of `FLVWAITTIM=*NOWAIT` and `FLVDFTACT=*PROCEED`, an automatic failover is triggered:

- Setting a *failover wait time* with the FLVWAITTIM cluster parameter, either specified with a duration in minutes or `*NOMAX`, allows the user to respond to the CPABB02 inquiry message to either proceed with or cancel the failover.
- The *failover default action* FLVDFTACT parameter setting determines IBM PowerHA SystemMirror for its behavior to either proceed automatically or cancel the failover processing after the specified failover wait time expired without getting a response for the inquiry message from the user.

In any case, the primary IASP is taken offline by IBM PowerHA SystemMirror for i for a failover event regardless of these cluster failover parameter settings.

When using the default ASP session setting parameter FLVRVSREPL=*NO for a Storwize remote copy ASP session, after a failover event, the session is detached, as shown in Figure 5-15, only for the ASP session for SAN Volume Controller Metro Mirror. Also, the remote copy relationships are not reversed, which allows the user to preserve the primary node IASP data for possible further failure analysis. The replication status is DETACHED and the storage state is IDLING.

```
                          Display SVC Session                    CTCIHA9W
                                                              09/28/11  18:23:47
 Session . . . . . . . . . . . . . . . . . . :   SVC_MM
 Type  . . . . . . . . . . . . . . . . . . . :   *METROMIR

 Switchover reverse replication  . . . . . . :   *YES
 Failover reverse replication  . . . . . . . :   *NO
 Consistency group . . . . . . . . . . . . . :   IASP1_MM
 Source storage cluster name . . . . . . . . :   ctcsvcclu2
 Target storage cluster name . . . . . . . . :   ctcsvcclu1
                                                                      More...
 CRG name  . . . . . . . . . . . . . . . . . :   SVC_MM_CRG
 Source site name  . . . . . . . . . . . . . :   SITE2
 Target site name  . . . . . . . . . . . . . :   SITE1

                                                                      Bottom
                            Copy Descriptions

 ASP         ASP copy              ASP         Replication
 device      name        Role      Status         state     Node
 IASP1       SVC_MM_T    SOURCE    AVAILABLE   DETACHED     CTCIHA9W
             SVC_MM_S    TARGET    UNKNOWN                  CTCIHA9V

ASP                               Copy
 device      Role      Node      progress  Storage state
 IASP1       SOURCE    CTCIHA9W    100     Idling
             TARGET    CTCIHA9V


                                                                      Bottom
 Press Enter to continue

 F3=Exit   F5=Refresh   F11=View 2   F12=Cancel
```

Figure 5-15 Display an auxiliary storage pool session for SAN Volume Controller Metro Mirror after an unplanned failover event

After the recovery of the preferred primary node occurs, including the restart of its cluster services, the detached ASP session should be reattached to resume remote copy replication and make the IASP highly available again. To do this task, run the following Change SVC Session (**CHGSVCSSN**) command:

```
CHGSVCSSN SSN(SVC_MM) OPTION(*REATTACH)
```

After resynchronization is finished, a planned switch to the preferred primary node can then be done, as described in 5.1, “Planned switching” on page 94, whenever a time window is available for doing so.

Primary node failure without an automatic failover event

Even with Advanced Node Failure Detection, there might be outages that do not cause an automatic failure. In this case, the current backup system puts the current production system into a partitioned status, which means that it cannot communicate with the current production system any more, and that it cannot determine whether that is because of a communication problem on the heartbeat connection or because of an actual failure of the current production system. In this case, an automatic failover does not take place to avoid situations where a failover occurs although the current production system is still running.

If a partition status occurs, the following CPFBB47 message is sent to QSYSOPR on the current backup node. If this happens, check whether the current production system experienced a failure.

```
Automatic fail over not started for cluster resource group <crgname> in cluster
<clustername>
```

If so, you can run the following steps to change the status of the current primary node from *partitioned* to *failed* by using the PowerHA GUI.

1. From the main PowerHA overview window (Figure 5-1 on page 94), click **Cluster Nodes**.
2. On the Cluster Nodes window, for the node in partitioned status, select **Change Status**, as shown in Figure 5-16.

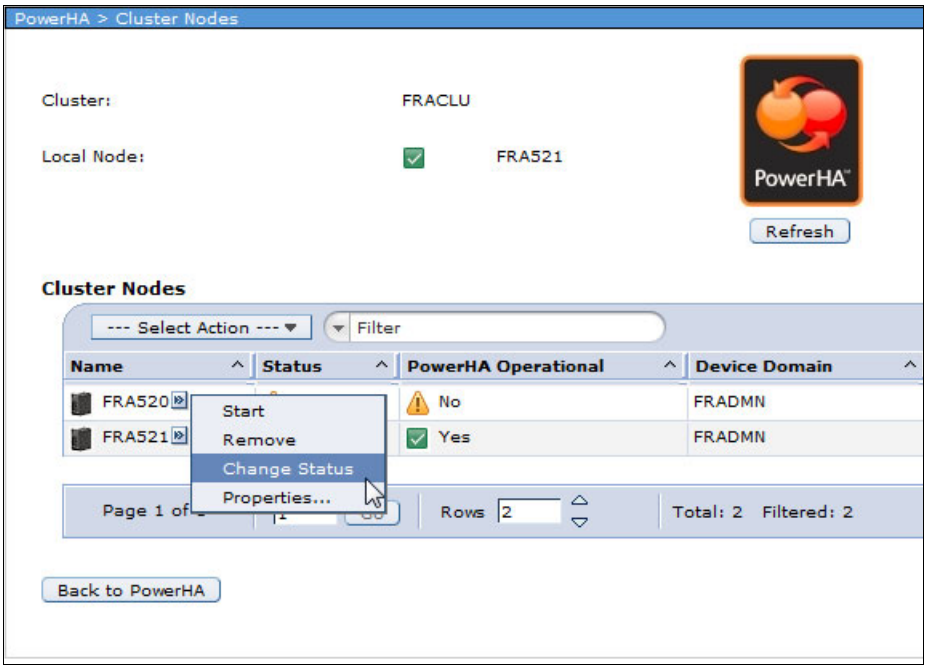

Figure 5-16 Change a cluster node status

3. You receive a confirmation message, as shown in Figure 5-17. Click **Yes** to continue.

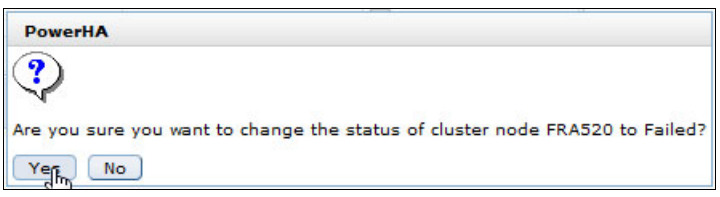

Figure 5-17 Change status from partitioned to failed - confirmation

4. The status is changed to failed, as shown in Figure 5-18.

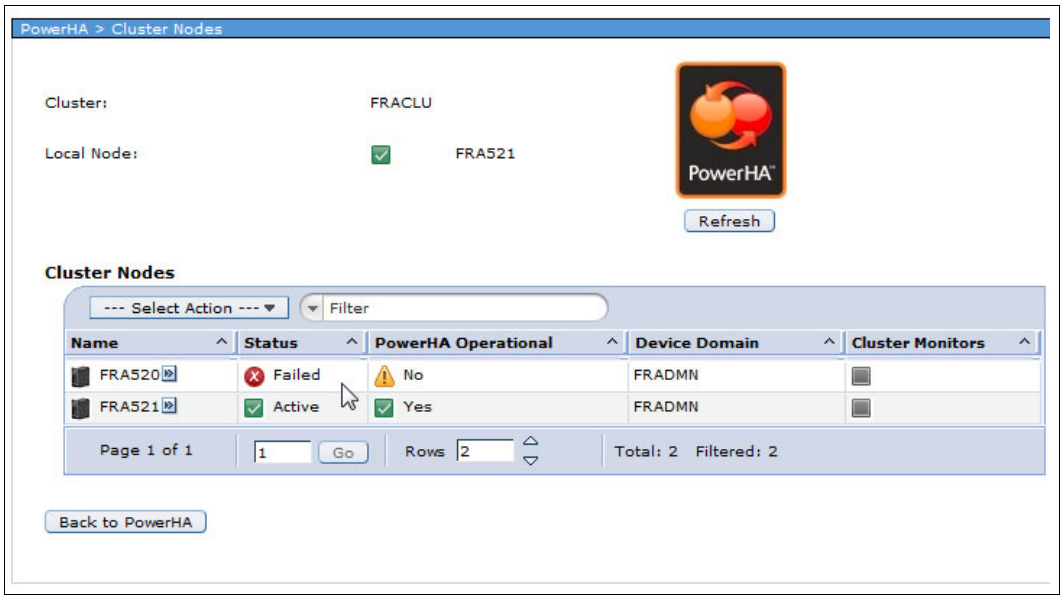

Figure 5-18 Cluster node status changed to failed

Alternatively, you can use the Change Cluster Node Entry (**CHGCLUNODE**) command:

```
CHGCLUNODE CLUSTER(<cluster_name>) NODE(<primary_node_name>) OPTION(*CHGSTS)
```

Make sure that the node name you specify is the node name of the current production system. The command can change only the status of a cluster node from partitioned to failed.

Note: The **CHGCLUNODE** command as shown triggers a cluster failover without a failover inquiry message and still requires the user to vary on manually the IASP on the new primary node and start the takeover IP interface.

After recovery of the preferred primary node, including a restart of its cluster services, the detached ASP session should be reattached to resume remote copy replication and make the IASP highly available again. To do this, run the following Change SVC Session (**CHGSVCSSN**) command:

```
CHGSVCSSN SSN(SVC_MM) OPTION(*REATTACH)
```

After resynchronization is finished, a planned switch to the preferred primary node can then be done, as described in 5.1, “Planned switching” on page 94, whenever a time window is available for doing so.

5.2.1 Special considerations when failing over in a Global Mirror setup

Because of the asynchronous nature of Global Mirror, there are some additional considerations to take into account when a failover occurs.

Global Mirror without change volumes

With Global Mirror being an asynchronous replication mechanism, there is no guarantee that the data on the current backup system is always the same as on the current production system. However, the data is consistent from a storage point of view (writes are done in the correct order on the target copy) and by using commitment control, the database can ensure that data is consistent from a database point of view. However, you might still be several transactions behind on your current backup system.

Make sure that you understand what must be done to discover which transactions have not made it to the new production system after a failover occurs. You should have a plan in place to recover those transactions.

Global Mirror with Change Volumes

The restrictions that are mentioned in “Global Mirror without change volumes” on page 109 also apply to GMCV. In addition, if replication is active, failover waits for the greater of 10 minutes or the configured cycle time plus 30 seconds for the data to become synchronized. If the current cycle is not complete in that period, then the failover fails and storage is not changed. If at some point during the failover process the synchronization is suspended, the failover also fails.

So, if you have situations where the cycle time is shorter than 9 minutes and 30 seconds and that cycle time tends to be too short to send over the changes that are accumulated in the source change volume to the target change volume, then you are at risk that in a failover situation that the failover cannot take place. The same is true if you configured a cycle time of more than 10 minutes and often need more than 10 minutes and 30 seconds to send over the changes that are accumulated in the source change volume to the target change volume.

You can run the Display Storage Area Network (SAN) Volume Controller Auxiliary Storage Pool Session (**DSPSVCSSN**) command to estimate switchover/failover time. The command sends a DIAG message to the job log giving an approximation of the data out of sync and the time based on the current estimated transfer rate. Because of the data that is returned from the storage server, these estimates tend to be more accurate for environments consisting of many small volumes, rather than a few large volumes.

Make sure that the settings for cycle time on the Storwize system are done in a way that does not prevent failovers from being successful. Your RTO must take the cycle time into consideration. Your expected RTO cannot be shorter than the cycle time that you set in the Storwize configuration because the failover process might need that amount of time to get the data synchronized between source and target copy of Global Mirror.

5.3 Working with SAN Volume Controller sessions

This section describes what can be done with the SAN Volume Controller session descriptions. The following topics are covered in this section:

- 5.3.1, “Suspending and resuming a SAN Volume Controller session” on page 110
- 5.3.2, “Detaching and reattaching a SAN Volume Controller session” on page 110
- 5.3.3, “Changing the attributes of a SAN Volume Controller session” on page 110

5.3.1 Suspending and resuming a SAN Volume Controller session

Suspending a SAN Volume Controller session stops Metro Mirror or Global Mirror without enabling access to the target data. On the Storwize system, the remote copy status is set to consistent stopped. To suspend the session, run the following command:

```
CHGSVCSSN SSN(<session-name>) OPTION(*SUSPEND)
```

After suspending the SAN Volume Controller session, remote copy is established again by running the following Change SVC Session (**CHGSVCSSN**) command with an option of RESUME. Using information about tracked changes, a partial resynchronization is started.

```
CHGSVCSSN SSN(<session-name>) OPTION(*RESUME)
```

5.3.2 Detaching and reattaching a SAN Volume Controller session

Detaching a SAN Volume Controller session stops Metro Mirror or Global Mirror from enabling access to the target data. On the Storwize system, the remote copy status is set to idling. To detach the session, run the following command:

```
CHGSVCSSN SSN(<session-name>) OPTION(*DETACH)
```

Note: For a detach, the same recommendations as for taking a FlashCopy exist. For an explanation of the quiesce function and the consequences of not doing a quiesce before doing a FlashCopy, see 5.4, “Working with PowerHA FlashCopy” on page 111.

After detaching a SAN Volume Controller session, remote copy is established again by running the following Change SVC Session (**CHGSVCSSN**) command with an option of REATTACH. Using information about tracked changes, a partial resynchronization from the current production copy to the current backup copy (as known by the recovery domain information in the CRG) is started. All changes that were made to the current backup copy are lost.

```
CHGSVCSSN SSN(<session-name>) OPTION(*REATTACH)
```

5.3.3 Changing the attributes of a SAN Volume Controller session

It is also possible to change some of the attributes of an existing SAN Volume Controller session by running the Change SVC Session (**CHGSVCSSN**) command with an option of CHGATTR:

```
CHGSVCSSN SSN(<session-name>) OPTION(*CHGATTR)
```

For a FlashCopy session, you can then change the copy rate and the cleaning rate. For a Metro Mirror or Global Mirror session, you can change whether replication should start in reverse direction after a switchover or after a failover.

5.4 Working with PowerHA FlashCopy

When working with FlashCopy in a PowerHA environment, make sure that you understand what is in the copy and what is not. FlashCopy creates a point-in-time copy of logical disk units. Therefore, all the data that still is in main memory on your source system is not copied to the FlashCopy target LUNs.

IBM i provides a quiesce function that suspends database transactions and database and Integrated File System (IFS) file change operations for the system and configured basic ASPs or IASPs. Already started transactions continue until they reach a commit boundary. Initiation of new transactions or new file (database or IFS) operations are temporarily halted.

In addition, the function forces to disk changed pages that are associated with the system and configured basic ASPs or IASPs. From a user perspective, the system seems to respond slower than expected during the suspend. Transactions should not break though unless the application uses timeouts.

Run the Change ASP Activity (**CHGASPACT**) command with the parameters that are shown in Figure 5-19:

- For the ASP device (ASPDEV) parameter, specify the name of the IASP.
- The following values are available for the OPTION parameter:
  - *SUSPEND starts the suspend.
  - *RESUME sets the ASP back to normal operation.
  - *FRCWRT forces main storage data for the ASP to disk without pausing transactions.
- The Suspend timeout (SSPTIMO) parameter specifies the amount of time in seconds that the suspend operation is allowed to take. If a complete suspend (all transactions reached a commit boundary) is achieved during this time frame, a message CPCB717 (Access to ASP suspended) is sent to message queue QSYSOPR. This message also contains data on how long it took to achieve the quiesce point.
- The Suspend timeout action (SSPTIMOACN) parameter defines what action should be taken if a full suspend of the ASP was not reached in the specified amount of time. The value of *CONT keeps the suspend active, and *END cancels the suspend and sets the ASP back to normal operation.

```
                      Change ASP Activity (CHGASPACT)

Type choices, press Enter.

ASP device . . . . . . . . . . . ASPDEV       > IASP1
Option . . . . . . . . . . . . . OPTION       > *SUSPEND
Suspend timeout  . . . . . . . . SSPTIMO      > 60
Suspend timeout action . . . . . SSPTIMOACN     *CONT



                                                                        Bottom
F3=Exit   F4=Prompt   F5=Refresh   F12=Cancel   F13=How to use this display
F24=More keys
```

Figure 5-19 CHGASPACT command

You also must understand the consequences of taking a FlashCopy from an IASP that is not quiesced. The FlashCopy operation itself finishes successfully (as it also would if you had not used the **CHGASPACT** command). The reason for this is that FlashCopy happens on the storage system, and the storage system is not aware of the interdependencies of activities happening on the server.

When you vary on the IASP that contains the FlashCopy to another cluster node, one of the steps during vary on is database recovery. This step ensures that the data in your IASP is in a consistent state. When open transactions that have not reached a commit boundary are found, these transactions are rolled back to the last commit point by using journal entries. You can then perform a save to tape from this consistent data.

Restoring this data and then running an Apply Journal Changes (**APYJRNCHG**) command by using journal entries from your production system to reach a more current data state does involve much manual work and might prove to be impossible within a reasonable amount of time. You must discover for which files a rollback took place during database recovery and find the individual starting point (last commit point) for each of these files manually from the journal entries. You then must run the **APYJRNCHG** command for each of these files with a different starting point. Chances are that you end up with inconsistent data when trying to do this task.

To make a FlashCopy of your IASP, complete the following steps:

1. On the production node, to quiesce the database, either vary off the IASP or run the **CHGASPACT** command.
2. To start the FlashCopy process, run the Start SVC Session (**STRSVCSSN)** command (Figure 5-20), making sure that you specify the correct source and target copy descriptions. You also can decide whether the IASP should be fully copied or only changes copied, and whether the relationship should persist after a copy is complete.

```
                    Start SVC Session (STRSVCSSN)

Type choices, press Enter.

Session  . . . . . . . . . . . . > SVC_FLC       Name
Session type . . . . . . . . . .> *FLASHCOPY    *METROMIR, *GLOBALMIR...
Device domain  . . . . . . . . .  *              Name, *
ASP copy:
  Preferred source . . . . . .   SVC_MM_T      Name
  Preferred target . . . . . .   SVC_FLC_T     Name
             + for more values
Incremental flash  . . . . . .   *NO           *NO, *YES
Copy rate  . . . . . . . . . .   0             0-100
Cleaning rate  . . . . . . . .   0             0-100
Grain size . . . . . . . . . .   256           256, 64
Consistency group  . . . . . .  *GEN
Reverse consistency group  . . .  *GEN

                                                                      Bottom
F3=Exit   F4=Prompt   F5=Refresh   F12=Cancel   F13=How to use this display
F24=More keys
```

Figure 5-20 STRSVCSSN command

3. When the **STRSVCSSN** command completes, run the **CHGASPACT *RESUME** command, if needed.
4. On the FlashCopy node, vary on the IASP and perform the backup or other processing that you want to perform.

5.4.1 Integration between PowerHA and Backup, Recovery, and Media Services

When doing a save from a FlashCopy, the original data and the original system do not know that a save took place, which means that, for example, the BRMS recovery reports do not show correct instructions for restoring data. BRMS has since been enhanced to support a FlashCopy environment.

Using the correct setup, the original system “thinks” that it did the save operations itself and the recovery reports on the original system reflect that a save was done from the FlashCopy. Using the advanced feature of BRMS, it is now possible to set the save time stamp to the time the FlashCopy was taken instead of the time that the actual save occurred.

Setting up the BRMS network group

To share the BRMS database information about the media classes, available storage locations, and media policies, all PowerHA nodes must be brought into a BRMS network configuration. Complete the following steps.

Important: These steps require that the BRMS network feature is installed.

1. Log on to the IBM i LPAR and open the BRMS policy menu by running the following Work with Policies using BRM (**WRKPCYBRM**) command:

   ```
   WRKPCYBRM *SYS
   ```

2. From the System Policy menu, select option 4 - Change network group.

3. In the Change Network Group screen (Figure 5-21), change the communication method to *IP and add your IBM i LPARs as nodes to the BRMS network. The host name must be reachable in the network. Add a host table entry, if necessary. Press Enter.

```
                          Change Network Group                CTCIHAV9 APPN

Network group . . . . :  *MEDINV          Position to  . . . . :
FlashCopy state . . . :  *ENDPRC
Communication method  :  *IP              *AVAIL, *IP, *SNA, *NONE
Notify period . . . . :    30             30-99999 seconds
Type options, press Enter.
  1=Add   4=Remove   8=Set time

     Remote Local    Remote     System            Network
Opt Location Name  Network ID   Status            Status
 1     CTCIHA9V     APPN
       CTCIHA9X     APPN        Online            Active

                                                                      Bottom

             F3=Exit    F5=Refresh    F11=BRMS Media Information   F12=Cancel
```

Figure 5-21 Add a node to the BRMS network

4. Import the BRMS database information from the leading BRMS system to the secondary partitions by using the following Initialize BRMS (**INZBRM**) command (ignore all messages):

   ```
   INZBRM OPTION(*NETSYS) FROMSYS(APPN.CTCIHA9V)
   ```

5. From the system that backs up the IASP, define which systems receive the save information and define that they also own the save history. If the option to change the save history owner is implemented, it appears as though the target system performed the backup. This option is useful for the Print recovery report parameter (PRTRCYRPT) and Start Recovery using BRM (**STRRCYBRM**) command. Complete the following steps:

   a. To add a specific system sync, run the following command to change the system name to make it look as though the backup was done by this system and synchronize the reference date and time. The command is run on the system that performs the actual save and the system name that is entered as the third parameter is the name of the production system.

      ```
      CALL QBRM/Q1AOLD PARM('HSTUPDSYNC' '*ADD' 'CTCIHA9V' 'APPN' 'IASP1'
      '*CHGSYSNAM')
      ```

   b. To add a specific system sync, run the following command to keep the name of who did the backup and synchronize the reference date and time. Again, the command is run on the system that performs the actual save and the system name that is entered as the third parameter is the name of the production system.

      ```
      CALL QBRM/Q1AOLD PARM('HSTUPDSYNC' '*ADD' 'CTCIHA9V' 'APPN' 'IASP1'
      '*NORMAL')
      ```

   c. To display what is set up, run the following command:

      ```
      CALL QBRM/Q1AOLD PARM('HSTUPDSYNC' '*DISPLAY')
      ```

   d. To remove a specific system, run the following command:

      ```
      CALL QBRM/Q1AOLD PARM('HSTUPDSYNC' '*REMOVE' 'CTCIHA9V' 'APPN' 'IASP1')
      ```

User-defined independent auxiliary storage pool time stamps

When using FlashCopy as the basis for your save operations, you must be aware that the save time differs from the time the FlashCopy was taken (and that represents the “real” save time from a data perspective). For example, if the FlashCopy occurred at 1 PM but the backups were not performed until 3 PM, BRMS, by default, uses 3 PM as the time of backup and reference point for future incremental backups. However, the FlashCopy time reflects when the IASP data was last changed.

The following commands allow users to define their own FlashCopy time stamps for specific IASPs. These commands are run on the system where the actual save operation takes place:

- To add or update a time stamp, run the following command, where IASPNAME is the name of the IASP, FILESYSTEM TYPE is either *QSYS or *IFS, and 1071213143302 is the time stamp in CYYMMDDHHMMSS format:

  ```
  CALL QBRM/Q1AOLD PARM('FLASHTIME' '*ENABLE' 'IASPNAME' 'FILESYSTEM TYPE'
  '1071213143302')
  ```

  If you want to have a time stamp for IASP ASP45 for both *IFS and *QSYS types, you must use *ENABLE twice: once for *QSYS and once for *IFS. These time stamps are independent of each other, and increment independently. *QSYS applies to libraries and objects in the IASP and *IFS refers to all link objects in the IASP.

- To view the current configurations, run the following command:

  ```
  CALL QBRM/Q1AOLD PARM('FLASHTIME' '*DISPLAY')
  ```

  When displayed, two time stamps are shown. The first is the user-specified time stamp, and the second time stamp is what BRMS used internally.

- To remove a configuration for a specific IASP, run the following command:

  ```
  CALL QBRM/Q1AOLD PARM('FLASHTIME' '*DISABLE' 'IASPNAME' 'FILESYSTEM TYPE')
  ```

- To remove all the user-defined time stamp configurations, run the following command:

  ```
  CALL QBRM/Q1AOLD PARM('FLASHTIME' '*DISABLE' '*ALL')
  ```

- Internally, there are places where BRMS requires a unique time stamp. In those cases, BRMS generates a new unique time stamp that is incremented from the user-defined time stamp. Save While Active sync points and incremental reference time stamps always use the value that is defined by the user. BRMS maintains a list of the time stamps that are used, and clears this list when the **STRMNTBRM** command with the RUNCLNUP(*YES) parameter is specified.

Important: The BRMS Advanced Feature (57xxBR1 option 1 (FC5102)) is required on the system where the FLASHTIME interface is used:

- User-defined time stamps are not used for archival backups.
- User-defined time stamps are not used for *LINK from IASP *ALLAVL or *CURASPGRP.
- Any SAVBRM or *LINK control group entry that saves from multiple IASPs does not use user-defined time stamps.
- Native backups (**SAVLIB**, **SAVOBJ,** and so on) are not affected.
- If a library with the same name exists in multiple IASPs and object detail is retained, use different user-defined time stamps for each IASP entry of the *QSYS type.
- Do not use time stamps later than the current system date and time.

5.4.2 Considerations when using FlashCopy on a Global Mirror target

A FlashCopy of a GMCV target can be up to two cycle periods behind the status of the source copy, depending on when the FlashCopy or detach is done in relationship to the cycle time. The calculation is the same as that to calculate the maximum RPO. Even if a vary off is done before the detach, to ensure an up to date detached copy, wait two cycle periods after the vary off before completing the detach. This situation assumes that your cycle time is set to a value that corresponds to your data change rate and to your available bandwidth.

5.4.3 Using FlashCopy when doing changes to your application

When doing larger changes to data in your IASP, using FlashCopy can provide an attractive way of creating a copy of the status before doing the changes, which provides a quick way of getting back to that status if the changes to the data in the IASP application fail.

The following steps can be taken:

1. If you want to use a FlashCopy as a “last save of valid data” before doing major changes, you want to be sure that this FlashCopy contains complete and valid data. Therefore, stop your applications and vary off the IASP on the current production system to finish all open transactions and write all IASP data from main memory to disk.
2. Establish a FlashCopy of the IASP, as shown in Figure 5-22. The Session type parameter must be *FLASHCOPY. Make sure that you specify the ASP copy description that describes the current production copy of your IASP as preferred source. The preferred target is the ASP copy description of the FlashCopy ASP.

```
                     Start SVC Session (STRSVCSSN)

Type choices, press Enter.

Session  . . . . . . . . . . . . SSN          > FLASH
Session type . . . . . . . . . . TYPE         > *FLASHCOPY
Device domain  . . . . . . . . . DEVDMN         *
ASP copy:                        ASPCPY
  Preferred source . . . . . . .              > SVC_MM_S
  Preferred target . . . . . . .              > SVC_FLC_T
                       + for more values
Incremental flash  . . . . . . . INCR           *NO
Copy rate  . . . . . . . . . . . CPYRATE        0
Cleaning rate  . . . . . . . . . CLNRATE        0
Grain size . . . . . . . . . . . GRAINSIZE      256
Consistency group  . . . . . . . CNSGRP         *GEN
Reverse consistency group  . . . RVSCNSGRP      *GEN




                                                                        Bottom
F3=Exit   F4=Prompt   F5=Refresh   F12=Cancel   F13=How to use this display
F24=More keys
```

Figure 5-22 Start FlashCopy

3. Vary on the IASP on your FlashCopy LPAR and make sure that the data is usable.

4. Vary on the production copy of your IASP again and perform the changes. If you discover that for some reason you must revert to the old status, complete the following steps.

   Important: These steps reverse *all* changes that you made to the production IASP data after you created the FlashCopy, which includes any normal production work that took place after you performed the FlashCopy.

   a. Vary off the production IASP.

   b. Suspend the ASP session between current production and current backup by running the following command:

      ```
      CHGSVCSSN SSN(SVC_MM) OPTION(*SUSPEND)
      ```

   c. Reverse the FlashCopy, which copies the data from the FlashCopy point in time over the current data of the IASP. This action eliminates *all* changes that were made to the data in the IASP after the FlashCopy was taken. To do this task, run the following command:

      ```
      CHGSVCSSN SSN(FLASH) OPTION(*REVERSE)
      ```

   d. Vary on the production IASP and resume the ASP session between current production and current backup by running the following command:

      ```
      CHGSVCSSN SSN(SVC_MM) OPTION(*RESUME)
      ```

5.4.4 Working with full-system FlashCopy

When using full-system FlashCopy, you must understand the restrictions in this setup. The only way to ensure a 100% save and consistent copy of a full system is to power down the source system. Doing a quiesce of the source system by running the Change ASP Copy Description command (**CHGASPACT**) only suspends database activities and file and IFS object change operations.

While taking a warm FlashCopy with the full system online, even when quiesced, any object that is resident in memory and not journaled, such as files, data queues, and data areas, is subject to potential damage due to object inconsistencies from parts of the object not being on disk. Journaling allows the target partition to recover from possible damaged journaled objects by applying journal entries to journaled files during the IPL.

It is only with a full system FlashCopy that there is a chance that on taking a warm FlashCopy that some IBM i internal objects can be found corrupted at IPL time, which is true even when doing a successful quiesce of the system ASP by running the Change ASP Activity command (**CHGASPACT**), which might lead to a situation where the backup LPAR cannot finish the IPL. You might need a SLIP installation, or another FlashCopy might be needed for a recovery.

6

Troubleshooting

This chapter covers some troubleshooting tips when using IBM PowerHA SystemMirror for i with the IBM Storwize family of products.

The following topics are described in this chapter:

For a complete description of IBM PowerHA for i troubleshooting procedures, see Chapter 6, “Troubleshooting and collecting data for problem determination”, of *IBM PowerHA SystemMirror for i: Preparation (Volume 1 of 4)*, SG24-8400.

6.1 Storwize Copy Services

When using Storwize for your independent auxiliary storage pool (IASP) storage and PowerHA for IBM System i®, you can log on to the Storwize system and check the status of the replication. Complete the following steps:

1. To check the Metro or Global Mirror status, select **Remote Copy** from the Storwize GUI, as shown in Figure 6-1.

Figure 6-1 Select to show the Remote Copy status

Figure 6-2 shows an example where you can see several remote copy consistency groups and their current state. You can also see in what direction mirroring is occurring.

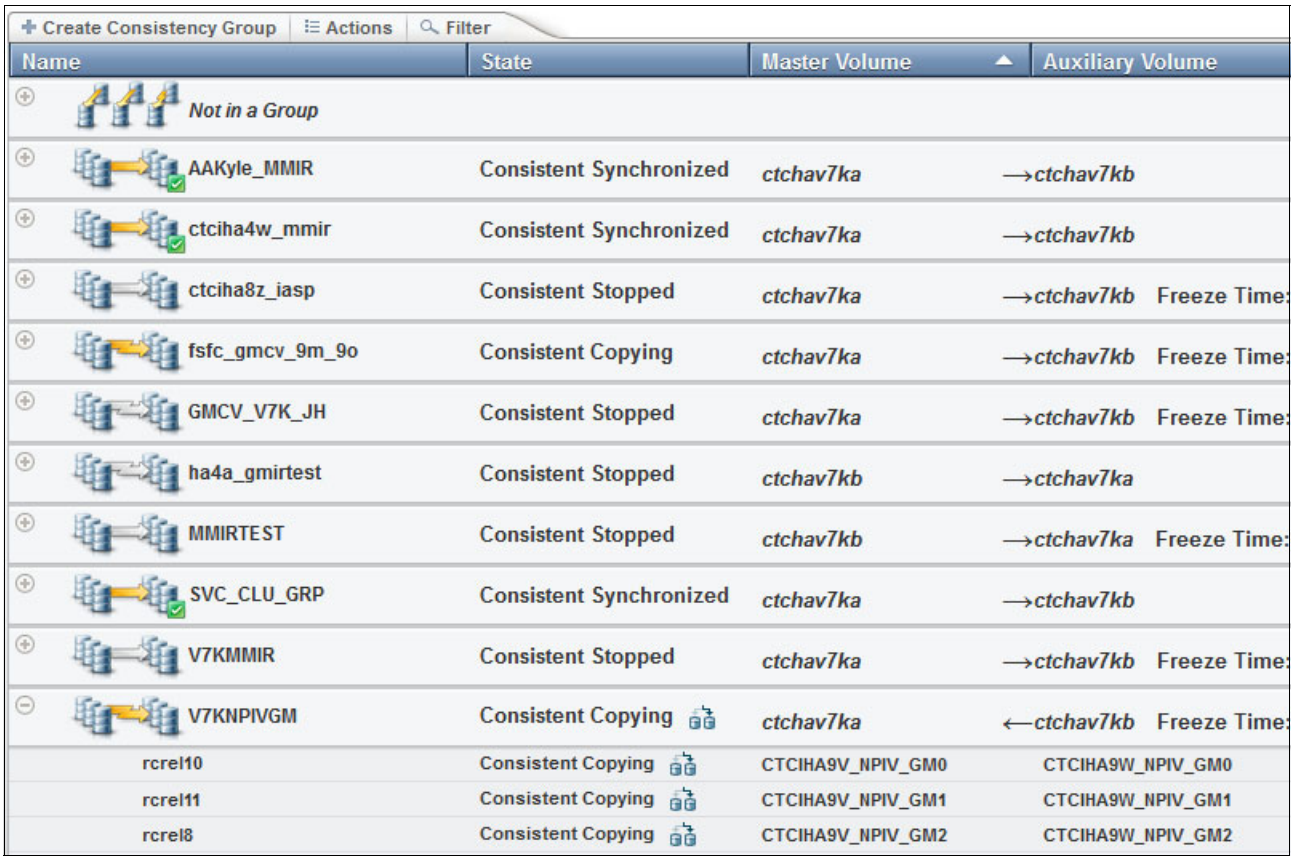

Figure 6-2 Remote Copy status - consistent

2. For FlashCopy relationships and their status, select **Consistency Groups** from the Storwize GUI, as shown in Figure 6-3.

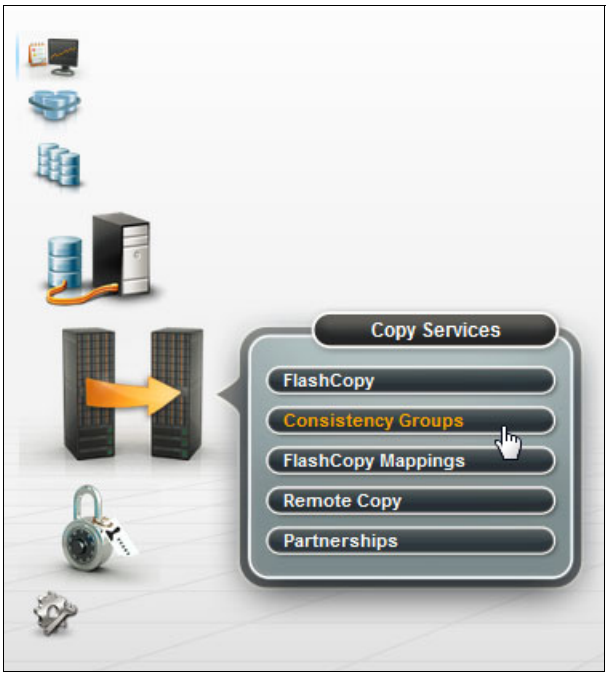

Figure 6-3 Select to show Consistency Groups status

Figure 6-4 shows an example of the FlashCopy Consistency Group status.

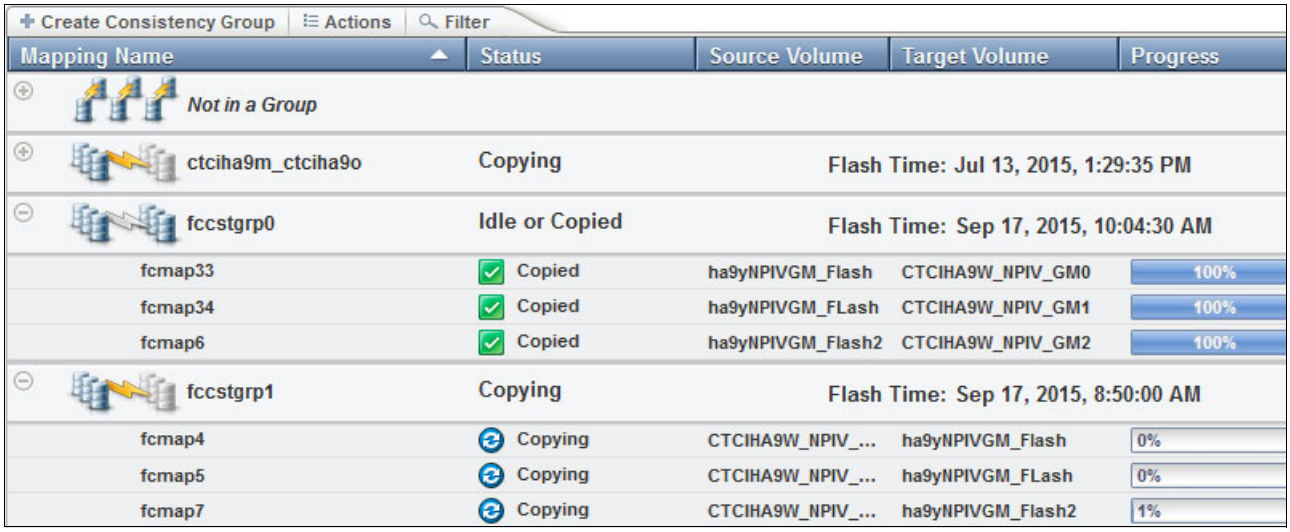

Figure 6-4 FlashCopy Consistency Group status

6.2 Support contact

If there is any problem that you cannot resolve, contact your country service support. For more information, see the IBM Service and Support Best Practices For Power Systems website:

http://www.software.ibm.com/webapp/set2/sas/f/best/home.html

A

PowerHA Tools for IBM i

This appendix describes the PowerHA Tools for IBM i offerings and services that are available from IBM Systems Lab Services.

The PowerHA Tools for IBM i complement and extend the PowerHA and IBM storage capabilities for high availability (HA) and disaster recover (DR).

The PowerHA Tools for IBM i provide the following benefits:

- Helps reduce business risk and improve resiliency for critical applications.
- Simplifies setup and automation of HA, DR, and backup solutions.
- Reduces the cost of maintaining and regular testing of an HA/DR environment.
- Facilitates flexible deployment options for single or multi-site protection.
- Assures consistent deployment by using preferred practices and experienced consultants.

For more information about PowerHA Tools for IBM i, see the following IBM Systems Lab Services and Training website:

http://www.ibm.com/systems/services/labservices

PowerHA Tools for IBM i

Table A-1 lists the PowerHA Tools for IBM i that are available from IBM Systems Lab Services.

Table A-1 PowerHA Tools for IBM i

PowerHA Tools for IBM i	Capability	Benefit	DS8000	Storwize	Internal
Smart Assist for PowerHA on IBM i	Provides operator commands and scripts to supplement PowerHA installation and ongoing operations for iASP enabled applications.	Simplifies deployment and ongoing management of HA for critical IBM i applications.	X	X	X
IASP Copy Services Manager (Automated recovery with faster independent auxiliary storage pool (IASP)-level vary on with no system IPL)					
FlashCopy	Automates FlashCopy of IASP for daily offline backup with seamless Backup, Recovery, and Media Services (BRMS) integration.	Increases application availability by reducing or eliminating the backup window for routine daily backups.	X	X	
LUN-level switching	Simplifies deployment and automates switching an IASP between IBM i cluster nodes in one data center.	Enables a business continuity manager to provide a simple, single-site HA solution.	X[1]	[2]	
Metro Mirror or Global Mirror	Simplifies initial deployment and automates ongoing server and storage management of two-site Metro Mirror or Global Mirror HA or DR solutions. Requires IASP-enabled applications.	Enables a business continuity manager to provide seamless operation of integrated server and storage operations for two-site HA and DR.	X		
Metro Global Mirror (MGM)	Extends PowerHA functions to provide a three-site server or storage replication solution containing Metro Mirror for HA with Global Mirror for DR. Requires IASP-enabled applications and IBM Tivoli® Productivity Center for Replication.	Enables a business continuity manager to further lower business risk and maximize business resilience for critical applications that require three-site HA and DR protection.	X		
Full System Copy Services Manager (FSCSM) (Automated recovery, and requires a full-system IPL)					
FlashCopy	Automates full-system FlashCopy for daily offline backup with integrated support for BRMS without IASP-enabled applications.	Increases application availability by reducing or eliminating backup window for online daily backups. Enables an entry solution while planning IASP enablement.	X	X	
Metro Mirror or Global Mirror	Simplifies initial deployment and automates ongoing server and storage management of two-site Metro Mirror or Global Mirror HA or DR solutions without IASP-enabled applications.	Enables a business continuity manager to provide seamless operation of integrated server and storage operations for HA and DR. Enables an entry solution while an IASP enablement is planned.	X	X[3]	

[1] DS8000 support is available with PowerHA Tools for IBM i V6.1 or earlier and included in PowerHA SystemMirror V7.1.

[2] V7000 support is included with PowerHA V7.1 TR6

[3] Storwize Full System Replication requires Full System Replication for PowerHA.

IBM Systems Lab Services offerings for PowerHA for IBM i

Table A-2 lists the PowerHA for IBM i service offerings that are available from IBM Systems Lab Services.

Table A-2 PowerHA for IBM i service offerings

PowerHA for IBM i service offering	Description
IBM i High Availability Architecture and Design Workshop	An experienced IBM i consultant conducts a planning and design workshop to review solutions and alternatives to meet HA and DR and backup and recovery requirements. The consultant provides an architecture and implementation plan to meet these requirements.
PowerHA for IBM i Bandwidth Analysis	An experienced IBM i consultant reviews network bandwidth requirements for implementing storage data replication. IBM reviews I/O data patterns and provides a bandwidth estimate to build into the business and project plan for clients who are deploying PowerHA for IBM i.
IBM i Independent Auxiliary Storage Pool (IASP) Workshop	An experienced IBM i consultant provides jumpstart services for migrating applications into an IASP. Training includes enabling applications for IASPs, clustering techniques, and managing PowerHA and HA and DR solution options with IASPs.
PowerHA for IBM i Implementation Services	An experienced IBM consultant provides services to implement an HA/DR solution for IBM Power Systems servers with IBM Storage. Depending on specific business requirements, the end-to-end solution implementation can include a combination of PowerHA for IBM i and PowerHA Tools for IBM i, and appropriate storage software, such as Metro Mirror, Global Mirror, or FlashCopy.

B

Additional material

This book refers to additional material that can be downloaded from the internet, as described in the following sections.

Locating the web material

The web material that is associated with this book is available in softcopy on the internet from the IBM Redbooks web server. Point your web browser at:

`ftp://www.redbooks.ibm.com/redbooks/SG248402`

Alternatively, you can go to the IBM Redbooks website at:

ibm.com/redbooks

Select the **Additional materials** link and open the directory that corresponds with the IBM Redbooks form number, SG248402.

Using the web material

The additional web material that accompanies this book includes the following IBM i save files:

File name	*Description*
MONASPSSN.mbr	Sample CL program to monitor replication status
MONNODSTS.mbr	Sample CL program to monitor the cluster node status
MONADMDMN.mbr	Sample CL program to monitor the cluster administrative domain
MONCRGSTS.mbr	Sample CL program to monitor the CRG status
MONIASPSTS.mbr	Sample CL program to verify the IASP utilization status

Related publications

The publications that are listed in this section are considered 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. Some publications that are referenced in this list might be available in softcopy only.

- *IBM PowerHA SystemMirror for i: Preparation (Volume 1 of 4)*, SG24-8400
- *IBM PowerHA SystemMirror for i: Using DS8000 (Volume 2 of 4)*, SG24-8403
- *IBM PowerHA SystemMirror for i: Using Geographic Mirroring (Volume 4 of 4)*, SG24-8401
- *IBM SAN Volume Controller 2145-DH8 Introduction and Implementation*, SG24-8229
- *IBM System Storage SAN Volume Controller and Storwize V7000 Replication Family Services*, SG24-7574
- *Implementing the IBM Storwize V5000*, SG24-8162
- *Implementing the IBM Storwize V3700*, SG24-8107
- *Implementing the IBM Storwize V7000 and IBM Spectrum Virtualize V7.6*, SG24-7938
- *Implementing the IBM System Storage SAN Volume Controller V7.4*, SG24-7933

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 publications

This publication is also relevant as a further information source:

- *IBM V7.2.0 Real-time Compression Evaluation User Guide for IBM System Storage SAN Volume Controller and Storwize V7000*, found at:

  http://www.ibm.com/support/docview.wss?uid=ssg1S7003988&myns=s028&mynp=OCST3FR7&mync=E

Online resources

These websites are also relevant as further information sources:

- Compresstimator Utility Version 1.5.2.2:
  http://www.ibm.com/support/customercare/sas/f/comprestimator/home.html
- Contact IBM Systems Lab Services:
  http://www.ibm.com/systems/services/labservices/contact.html
- IBM Service and Support Best Practices for Power Systems:
  http://www.software.ibm.com/webapp/set2/sas/f/best/home.html
- IBM Storwize family:
  http://www.ibm.com/systems/storage/storwize/Description2
- IBM Systems Lab Services and Training:
  http://www.ibm.com/systems/services/labservices

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

Redbooks

IBM PowerHA SystemMirror for i: Using IBM Storwize (Volume 3 of 4)

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

SG24-8402-00

ISBN 0738441724

Printed in U.S.A.