IBM PowerHA SystemMirror for i: Using DS8000
(Volume 2 of 4)

David Granum

Edward Grygierczyk

Sabine Jordan

Peter Mayhew

David Painter

Power Systems

IBM

International Technical Support Organization

IBM PowerHA SystemMirror for i: Using DS8000

June 2016

> **Note:** Before using this information and the product it supports, read the information in "Notices" on page vii.

First Edition (June 2016)

This edition applies to Version 7, Release 2 of PowerHA SystemMirror for IBM 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.

AIX®	POWER6®	System z®
DS8000®	PowerHA®	SystemMirror®
FlashCopy®	PowerVM®	Tivoli®
HyperSwap®	Redbooks®	XIV®
IBM®	Redbooks (logo) ®	z/OS®
POWER®	Storwize®	
Power Systems™	System Storage®	

The following terms are trademarks of other companies:

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

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

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

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 PowerHA for i is combined with IBM i clustering technology, PowerHA for i delivers a complete HA and disaster-recovery (DR) solution for business applications that run in an IBM i environment. With PowerHA for i, you can support HA capabilities with either native disk storage, IBM DS8000® storage servers, or IBM Storwize® storage servers.

This IBM Redbooks® publication helps you to install, tailor, and configure IBM PowerHA SystemMirror for i with the IBM System Storage® DS8000 series. This publication provides you with planning information to prepare to use the various PowerHA offerings for the IBM DS8000 family. It also provides implementation and management information. It provides guidance about troubleshooting these solutions and identifies the documentation and data that you need to capture before you call IBM Support.

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

► *IBM PowerHA SystemMirror for i: Preparation*, SG24-8400
► *IBM PowerHA SystemMirror for i: Using IBM Storwize*, SG24-8402
► *IBM PowerHA SystemMirror for i: Using Geographic Mirroring*, 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, 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 part of the IBM Lab Services team and based in Poland. He joined IBM in 2008, working for IBM STG Lab Services as a Consultant in IBM Power Systems™. Edward has worked in the IBM i area since 1991. He built a highly successful IBM Business Partner company in Poland before he joined IBM. Edward has a background in systems, availability, integration, and implementation. Edward was also an author of IBM i courses. Edward can be reached at edward.grygierczyk@pl.ibm.com.

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 that use geographic mirroring, IBM DS8000, and IBM SAN Volume Controller (SVC) 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. In addition, 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 worked 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, Peter has maintained his skills on the IBM i, specializing in hardware and operating system installations, upgrades and migrations, and HA planning and implementation, supporting clients throughout the United States. Peter can be reached at Pete@itsolutions-group.com.

David Painter is part of the IBM Rochester Lab Services team and 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 David joined 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 in clients that range from small and medium business (SMB) accounts to large enterprise accounts. David regularly teaches courses in IBM i HA and external storage for IBM i administrators. He also conducts many availability workshops and holds numerous certifications. David can be reached at davpaint@uk.ibm.com.

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
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

Introduction to IBM PowerHA SystemMirror on i with the DS8000

This chapter describes the various IBM PowerHA SystemMirror for i solutions that are available when you use the IBM System Storage DS8000 series servers.

The following topics are described in this chapter:

- ► 1.1, "Value proposition" on page 2
- ► 1.2, "DS8000 terminology" on page 3
- ► 1.3, "Metro Mirror" on page 11
- ► 1.4, "Global Mirror" on page 15
- ► 1.5, "LUN-level switching" on page 21
- ► 1.6, "FlashCopy" on page 22

> **Note:** It is not the intention of this chapter to explain how to configure the DS8000 storage server. The goal of this chapter is to explain how to configure and manage PowerHA in a DS8000 environment.

For more information about the IBM DS8000 series, see this website:

https://ibm.biz/Bd4ukG

1.1 Value proposition

The 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 are exposed to negative business effects if they did not plan effectively for unplanned outage, never established a high-availability (HA) solution, or never tested a recovery plan.

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.

The DS8000 series can bring simplicity to your systems of insight with self-optimizing performance for the Power Systems server platforms. With the agility and automation to manage all workloads simultaneously, DS8000 can greatly reduce the challenge of managing your complex IT environment as your business requirements change.

To support your most demanding, business-critical applications, DS8000 combines rock-solid resiliency and intelligent flash performance to deliver consistent microsecond application response times. The ability to select from all-flash, hybrid-flash, or traditional disk configuration options gives you the necessary confidence to support dynamic applications.

The DS8000 delivers a range of self-tuning features, such as intelligent caching algorithms, automated quality-of-service (QoS) management, and advanced storage tiering that can optimize data placement between the DS8000 and the enterprise servers that are attached to it. Additionally, with a new generation of all-flash and hybrid-flash systems, the DS8000 deliver another level of extraordinary performance for your most time-sensitive analytics. For organizations that seek an ideal combination of performance, HA, resilience, and agility, DS8000 is the logical choice.

For more information about the IBM DS8000 series, see the following website:

https://ibm.biz/Bd4ukG

1.2 DS8000 terminology

Before you plan and configure a PowerHA solution by using the DS8000 storage server series, it is important to learn about the terminology and the functionality that is available with this family of storage. The following topics are covered in this section:

- ► 1.2.1, "Volumes" on page 3
- ► 1.2.2, "Volume groups" on page 4
- ► 1.2.3, "Resource group" on page 4
- ► 1.2.4, "Host connections" on page 4
- ► 1.2.5, "Logical subsystems" on page 8
- ► 1.2.6, "IBM PowerVM" on page 9
- ► 1.2.7, "Virtual I/O Server considerations" on page 9

1.2.1 Volumes

IBM i volumes are fixed-block (FB) volumes that consist of many 1 GB extents. However, for the DS8000 native or N_Port ID Virtualization (NPIV) attachment, the IBM i operating system supports only certain fixed volume sizes that are the same as the supported physical disks drives. Table 1-1 shows the available volume sizes for the IBM i system.

Table 1-1   Capacity and models of disk volumes for IBM i hosts that are running IBM i operating system

Size	Type	Protected model	Unprotected model
8.5 GB	242X	A01	A81
17.5 GB	242X	A02	A82
35.1 GB	242X	A05	A85
70.5 GB	242X	A04	A84
141.1 GB	242X	A06	A86
282.2 GB	242X	A07	A87
1 GB - 2000 GB	242X	099	050

With the IBM i, the minimum load source requirements changed for external and virtual load source logical unit numbers (LUNs). The minimum load source sizes are listed:

- ► Minimum of 35 GB for 520-byte sectors (DS8000 or virtual disks with Virtual I/O Server (VIOS) as the server)

- ► Minimum of 41 GB (to get 35 GB of usable space) for 512-byte sectors (IBM SAN Volume Controller (SVC), IBM Storwize storage servers, or virtual disks with VIOS virtual Small Computer System Interface (VSCSI) or IBM i as the server)

- ► Minimum of 70 GB for native-attached serial-attached SCSI (SAS) (both 520-byte or 4 K sectors)

The DS8000 provides the IBM i operating system with a specific serial number for every volume (same as physical disk drives). The disks are presented as 50*XXXXYYY,* where *XXXX* is the volume ID and *YYY* is the IBM i suffix number.

1.2.2 Volume groups

A *volume group* is a set of volumes. The volumes are presented together by the DS8000 to the host operating system through the volume group. The volume group type relates to the operating system that uses it. The types of volumes that are included into a volume group must match the volume group type. For example, you cannot include volumes that are not IBM i type models in an IBM i volume group.

> **Note:** The IBM i operating system does not support sharing a volume with another operating system.

1.2.3 Resource group

Resource groups in DS8000 code level 6.1 or higher protect LUNs in the DS8000 from accidental Copy Services actions. You can assign all of the LUNs for one environment to a resource group, which prevents IBM FlashCopy® or Peer-to-Peer Remote Copy (PPRC) actions from outside of the resource group from accidentally overwriting data on these LUNS.

Also, the independent auxiliary storage pool (IASP) LUNS for any partition in the environment need to be in a resource group to enable Copy Services between each other, while they are protected against accidental copying from another environment.

1.2.4 Host connections

A *host connection* is an object that represents the Fibre Channel (FC) adapter port of a host on DS8000.

The main attributes of a host connection are its type, which relates to the volume group type, and its worldwide port name (WWPN), which is a unique 16-hexadecimal-character address.

Figure 1-1 shows an example of host connection relationships to volume groups. IBM i Partition 1 has two Fibre Channel adapters, WWPN 1 and WWPN 2. IBM i Partition 2 also has two Fibre Channel adapters, WWPN 3 and WWPN 4. Inside the DS8000, WWPN 1 and WWPN 2 are assigned to the same volume group, Volume Group 1. WWPN 3 and WWPN 4 are assigned to the same volume group, Volume Group 2. With this configuration, IBM i Partition 1 uses volumes that are included in volume group 1. IBM i Partition 2 uses volumes that are included in Volume Group 2. Both partitions have a dual path to the volumes due to dual Fibre Channel attachments and assignment to the same volume group.
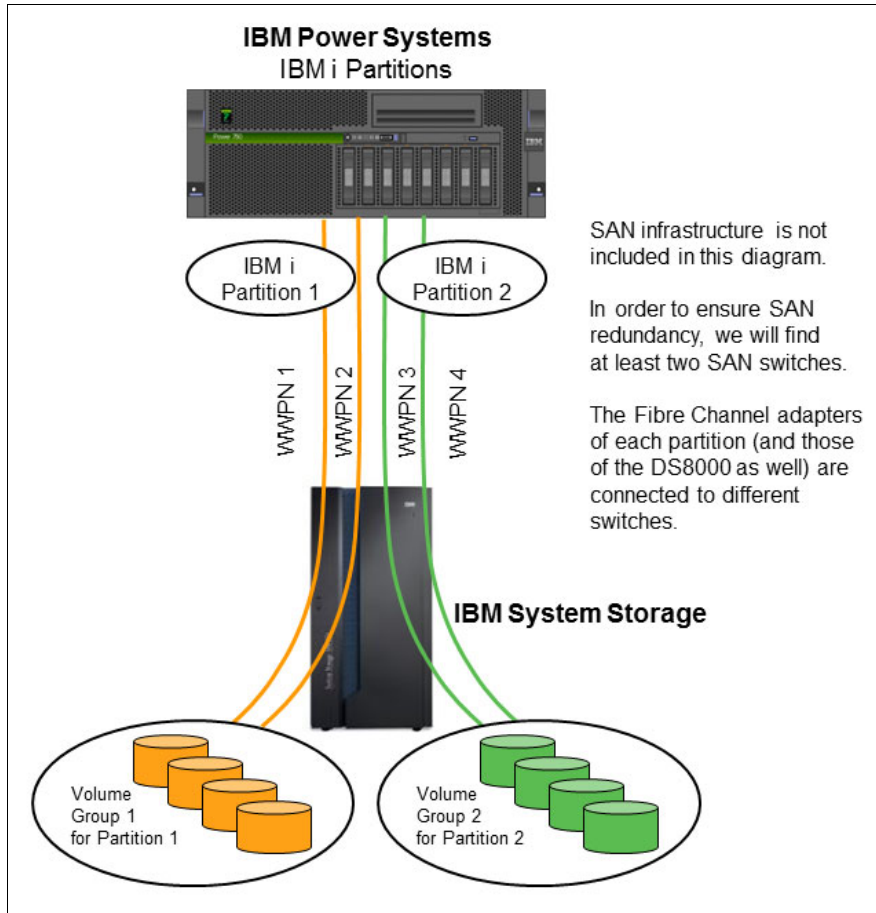
Figure 1-1   Example of host connection relationships to volume groups

A 1-to-*n* relationship exists between a volume group and host connections.

The host can connect by direct attachment, switch attachment, or with a storage area network (SAN) switch attachment. See Figure 1-2.
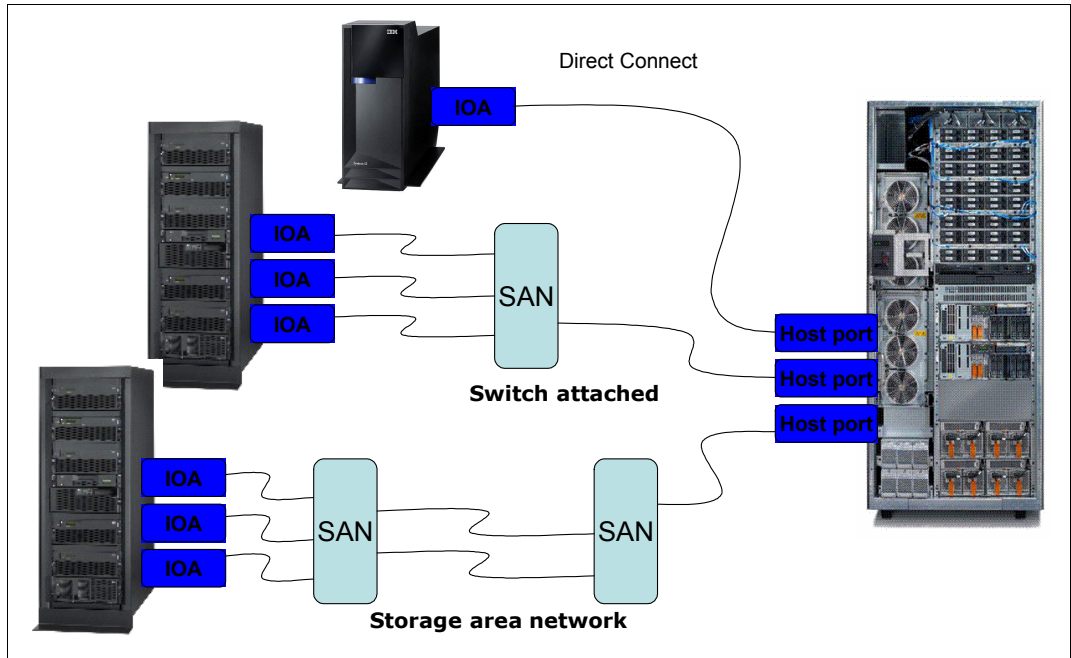
Figure 1-2   Host connection topologies

It is important for redundancy and performance to ensure that each host connection connects to only a single host port in the storage unit. Use multiple zones to separate the ports as shown in Figure 1-3.
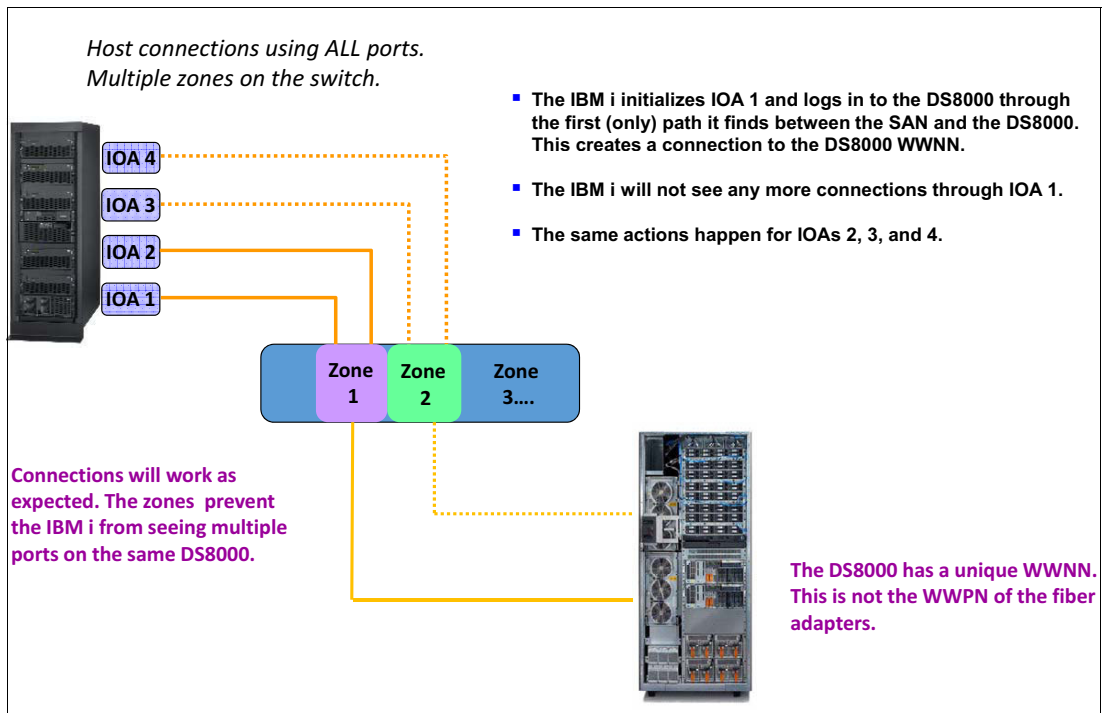
Figure 1-3   Host connection that uses all ports

The use of the multiple zones configuration prevents the IBM i from seeing any port except the ports to which it is authorized and isolates SAN traffic within the switch. This configuration is the most efficient configuration.

Additionally, by connecting IBM i to a storage system disk, you can connect each host connection to separate a volume group as shown in Figure 1-4.
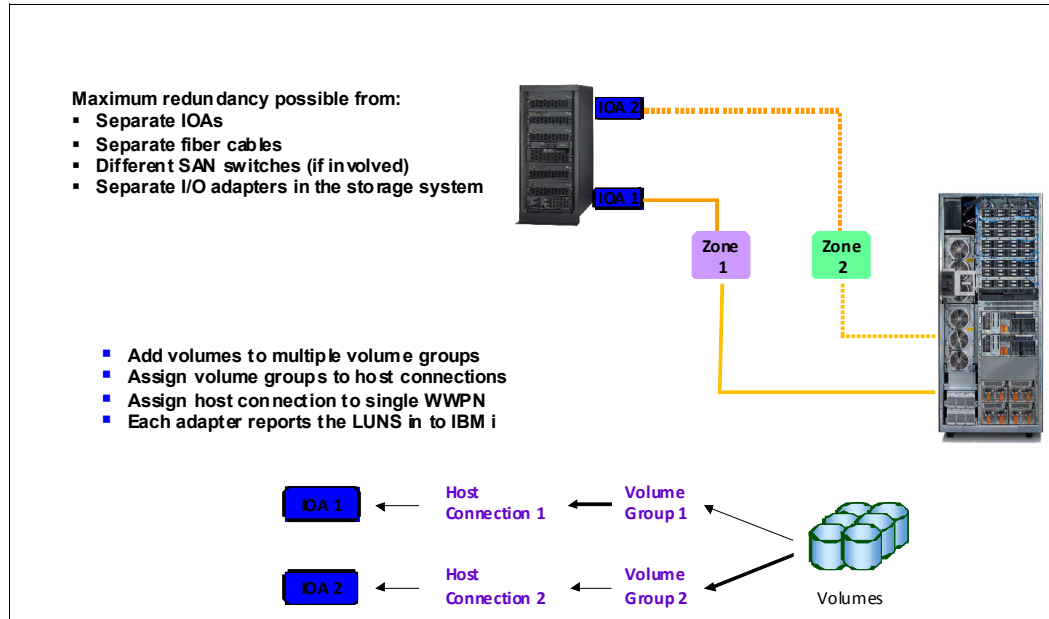
Figure 1-4   Multipath on IBM i host connections to different volume groups

Another approach is to connect two or more host connections to one volume group as shown in Figure 1-5.
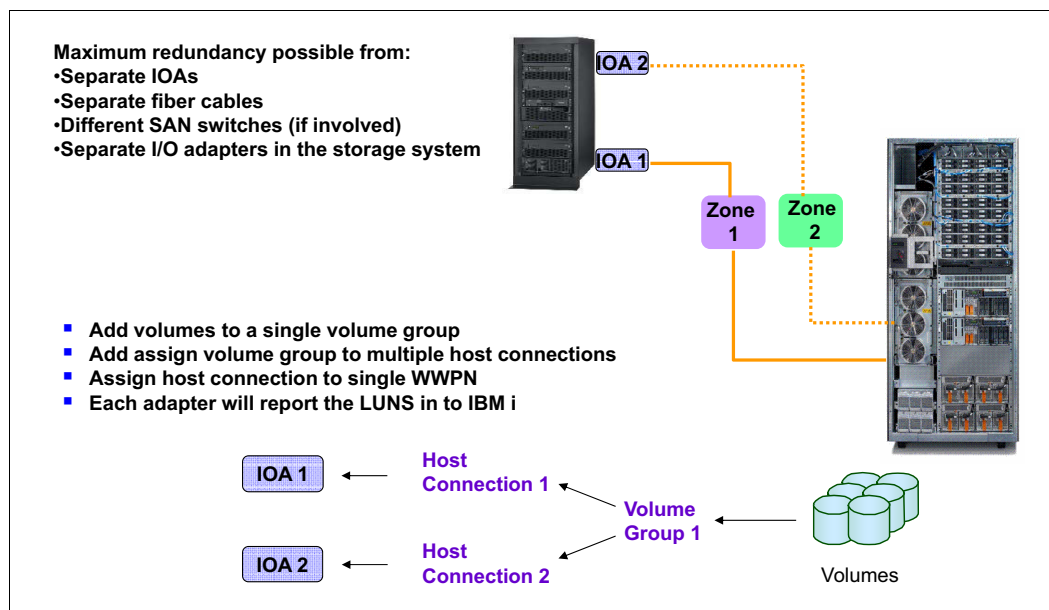
Figure 1-5   Multipath on IBM i host connections to one volume group

Multipath considerations

Consider these important points before you plan your configuration:

► A single volume (LUN) can have up to a maximum of eight paths.

► A single port can have a maximum of 64 LUNs.

► Two ports are available for each I/O card adapter.

► A simple round-robin algorithm is used for load-balancing.

► Multiple concurrent I/O requests are performed to a single volume.

► You can view the configuration through either IBM Navigator for i or dedicated service tools (DST) or system service tools (SST).

► IBM i changes the volume resource name to reflect multipath (DMPxxx) as soon as a second path is seen.

> **Note:** This name does not change back to a DDxxx device if mulitpath is removed.

► Use DST to reset paths if an input/output adapter (IOA) is moved, a bus number is changed, and so on.

Multipath for an external load source is handled in the same way as any other LUN as shown in Figure 1-6.
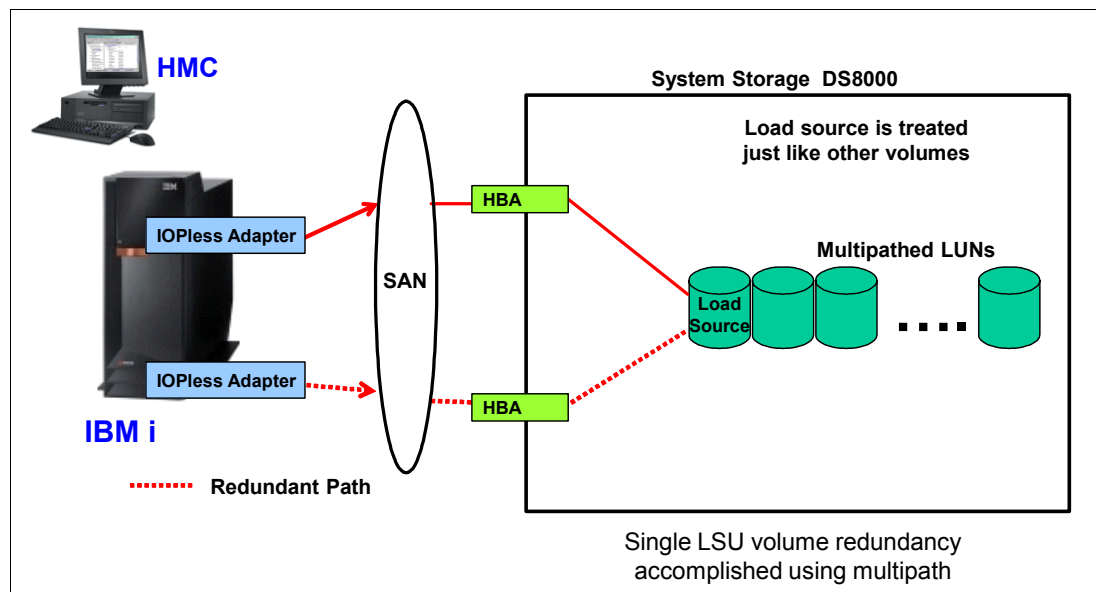
Figure 1-6 Multipath and external load source

1.2.5 Logical subsystems

Another logical group of volumes is the *logical subsystem* (LSS). The LSS is automatically created the first time that a volume with a new logical subsystem on DS8000 is created. An LSS has an affinity with one of the internal servers on DS8000.

The LSS is important for part of the Copy Services usage.

> **Important:** You need to use a unique LSS pair for each workload. For example, on a system with an IASP, the SYSBAS is one LSS pair and the IASP is another LSS pair.

For more information about LSS, see *IBM System Storage DS8000: Architecture and Implementation*, SG24-8886.

1.2.6 IBM PowerVM

IBM PowerVM® is a virtualization technology for IBM AIX®, IBM i, and Linux environments on IBM POWER® processor-based systems. The PowerVM VIOS that is included in the PowerVM provides a virtualization environment of storage and network I/O resources. With PowerHA, you can share hardware resources.

Virtualization technology is offered in three editions on Power Systems:

► PowerVM Express Edition
► PowerVM Standard Edition
► PowerVM Enterprise Edition

These editions provide logical partitioning (LPAR) technology by using either the Hardware Management Console (HMC) or the integrated Virtualization Manager (IVM) and Dynamic LPAR.

1.2.7 Virtual I/O Server considerations

Virtual I/O Server (VIOS) is virtualization software that runs in a separate partition on your Power system. Its purpose is to provide virtual storage and networking resources to one or more client partitions. An IBM i partition can be hosted by VIOS.

VIOS owns the physical I/O resources, such as Ethernet and SAS or Fibre Channel adapters. It virtualizes those resources for its client LPARs to share them remotely by using the built-in hypervisor services. These client LPARs can be created quickly, typically owning only physical memory and shares of processors without any physical disks or physical Ethernet adapters.

Traditionally, IBM i supported 520 bytes per sector storage. Therefore, restrictions exist to attaching a common storage system to IBM i directly for this reason. However, VIOS offers the IBM i client partition the ability to use 512 bytes per storage sector by using the sector conversion of IBM PowerVM Hypervisor, which converts IBM i traditional 8 x 520-byte sectors into 9 x 512-byte sectors of a 4 KB memory page. Therefore, open storage volumes (or logical units (LUNs)) are physically attached to VIOS through a Fibre Channel or an SAS connection and then available to IBM i.

> **Note:** When you configure the IBM i client partition with VIOS, consider a dependency on VIOS. If the VIOS partition fails, IBM i on the client loses contact with the virtualized open storage LUNs. The LUNs also are unavailable if VIOS is brought down for scheduled maintenance or a release upgrade. To remove this dependency, you can configure *Redundant VIOS* so that two or more VIOS partitions can be used to simultaneously provide virtual storage to one or more IBM i client partitions by using IBM i multipathing.

See Figure 1-7.
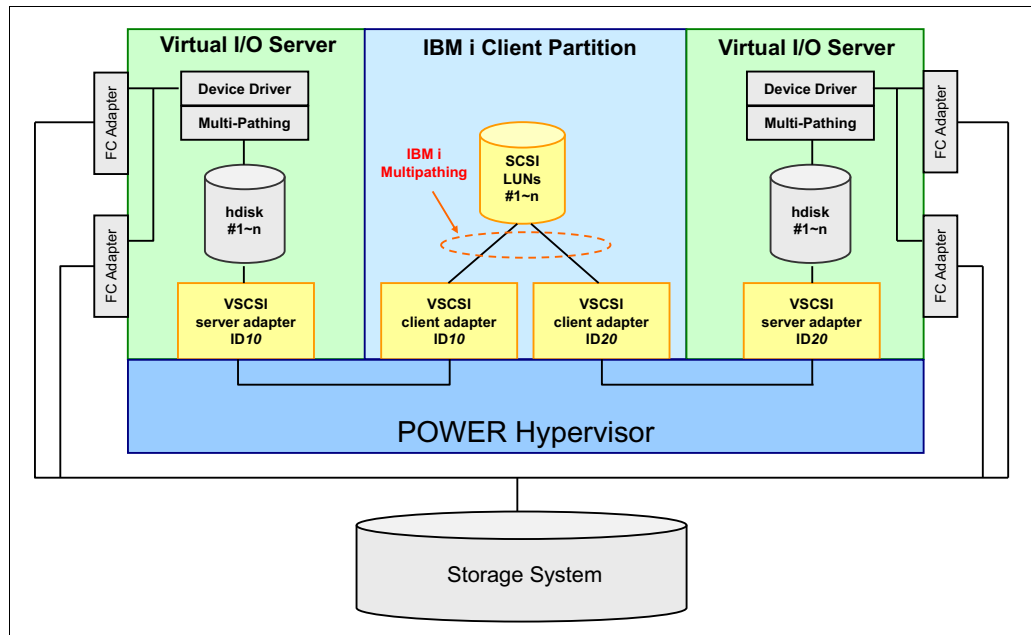
Figure 1-7   VSCSI connection by using VIOS with multipathing

Virtual SCSI

The IBM i Virtual SCSI (VSCSI) client driver supports Multi-Path I/O (MPIO) through two or more VIOS partitions to a single set of LUNs. This multipath configuration allows a VIOS partition to fail or be brought down for service without IBM i losing access to the disk volumes because the other VIOS partitions remain active.

If you do not want to use native IBM MPIO, you can install the Subsystem Device Driver Path Control Module (SDDPCM) on VIOS. However, IBM MPIO with SDDPCM does not support DS8000.

For more information about the Multipath Subsystem Device Driver, see the *Multipath Subsystem Device Driver User's Guide*:

https://ibm.biz/Bd4uNN

N_Port ID Virtualization (NPIV) with IBM System Storage DS8000

N_Port ID Virtualization (NPIV) is an industry-standard Fibre Channel protocol that allows VIOS to share a single Fibre Channel adapter among multiple client LPARs directly, acting as a Fibre Channel pass-through.

Unlike VSCSI, NPIV does not map a LUN to a virtual target device in VIOS, which the client LPAR can then access as a generic SCSI disk. Instead, a port on the physical Fibre Channel adapter is mapped to a virtual Fibre Channel server adapter in VIOS, which in turn is connected to a virtual Fibre Channel client adapter in IBM i. See Figure 1-8 on page 11.

When the virtual Fibre Channel client adapter is created, two unique worldwide port names (WWPNs) are generated for it. Through the link to the server virtual Fibre Channel adapter and then the physical adapter in VIOS, those WWPNs become available on the SAN, and storage can be mapped to them as with any other Fibre Channel host ports.

Note: WWPNs are unique, not only within the Power server, but globally on the SAN. By default, the PowerVM Hypervisor can create 32,000 WWPNs for virtual Fibre Channel client adapters.

The minimum requirements for an IBM i DS8000 NPIV attachment are listed:

► Eight Gb input/output processor (IOP)-less Fibre Channel IOA (custom card identification number (CCIN) 577D) or 10 Gb Fibre Channel over Ethernet (FCoE) Peripheral Component Interconnect Express (PCIe) Dual Port Adapter (CCIN 2B3B)

► NPIV-capable SAN switch

► IBM i

► Hardware Management Console (HMC) V7R3.5.0 or later

► IBM POWER6® Firmware 350_038 or later

► VIOS 2.1 Fix Pack (FP) 22.1 or later

Figure 1-8 shows NPIV for IBM i.
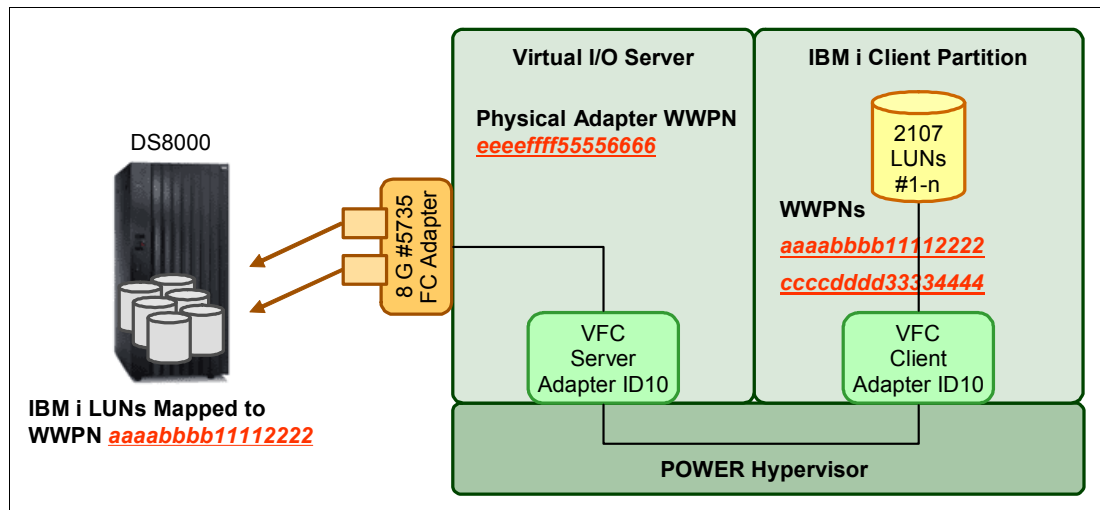
Figure 1-8   NPIV for IBM i

For more information about connecting IBM i through VIOS to DS8000 by using VSCSI and NPIV, see the IBM Redbooks publication *DS8000 Copy Services for IBM i with VIOS*, REDP-4584.

1.3 Metro Mirror

Metro Mirror provides the real-time copy of volumes between two DS8000 storage servers. It is a synchronous copy where write operations are completed on both DS8000 storage servers before write acknowledgements are sent back to the source operating system. This design ensures that no data is lost in a failure.

However, due to the synchronous copy, distance considerations between the local and the remote DS8000 storage servers apply. No more than 300 km (186 miles) are allowed.

> **Note:** The performance of the LUNs (as seen by the operating system) decreases as the distance between the storage units increases.

The following steps outline this write sequence when a host requests a write operation:

1. Data is stored in the source DS8000 cache and non-volatile storage (NVS) to be destaged to the current production volume later.

2. The same data is sent to the target DS8000 where it is stored in the target DS8000 cache and NVS to be destaged to the current backup volume later.

3. A write acknowledgment is sent by the target DS8000 to the source DS8000.

4. A write acknowledgment is sent by source DS8000 to the host.

1.3.1 Metro Mirror operations

The basic operations that can be performed with Metro Mirror are described:

► Establishing a Metro Mirror pair

Establishing a Metro Mirror pair operation establishes the copy relationship between the source (or local) volume and the target (or remote) volume. Normally, those two volumes reside on two DS8000s, but it is possible to use the same DS8000 for testing.

During the initial copy, the target volume is in a *target copy pending* state until all tracks are copied. At the end of the initial copy, the target volume is in a *target full duplex* state.

The following options are possible at establishment time:

– No copy. No data is copied from the source volume to the target volume. You must ensure that data is synchronized already.

– Target read. The target volumes can be read by a host while replication is active.

> **Note:** This option is not supported by IBM i, which must be able to write to any disk at any time.

– Suspend after synchronization. As soon as the target volume goes into the *target full duplex* state, the replication is suspended.

– Reset reserve on target. The Metro Mirror replication starts even if the target volume is reserved by a host.

> **Note:** A PPRC path must exist from the source to the target site before you create the Metro Mirror pair. For more information, see 1.3.2, "PPRC paths and links" on page 13.

► Suspending a Metro Mirror pair

Suspending a Metro Mirror pair operation stops sending data from the source volume to the target volume. However, the source DS8000 keeps a record of updated tracks on the source volume.

► Resuming a Metro Mirror pair

Resuming a Metro Mirror pair operation releases sending data from the source volume to the target volume. It uses updated track records during the suspended time to send only updated data.

► Terminating a Metro Mirror pair

Terminating a Metro Mirror pair operation ends the replication between the source volume and the target volume and removes this relationship. If you need the replication again, you must restart it from the beginning.

► Failover

Failover relates to the actions that you take to activate target volumes in a switch from a production site to a backup site. From the DS8000 standpoint, actions are the same for a planned switch and for an unplanned switch. Failover operations start after a failure or maintenance on the current production site. After a failover, current production can run on preferred backup volumes.

► Failback

Failback relates to the actions that you take to return to the nominal situation with production that is running on preferred source volumes. Failback operations can start after the preferred source site failure is fixed or maintenance is finished.

1.3.2 PPRC paths and links

The peer-to-peer remote copy (PPRC) path is a logical construct on the DS8000 that defines the physical links that are used by a Metro Mirror or Global Mirror relationship from one DS8000 to another DS8000. This path is built at the logical subsystem (LSS) level. Therefore, Metro Mirror relationships for all volumes in the same LSS use the same physical links. From the DS8000 standpoint, the physical link is from one of its Fibre Channel adapters to one of the target DS8000 Fibre Channel adapters, and the physical link can include supported channel extenders or routers.

The path creates a *unidirectional* association on the source DS8000 between the following components:

► Target DS8000 worldwide node name (WWNN)
► Source LSS
► Target LSS
► Source I/O port
► Target I/O port

For bandwidth and redundancy, a path can have up to eight links. The DS8000 handles failures and balances the workload across the available paths. As shown in Figure 1-9, a link can handle several paths.
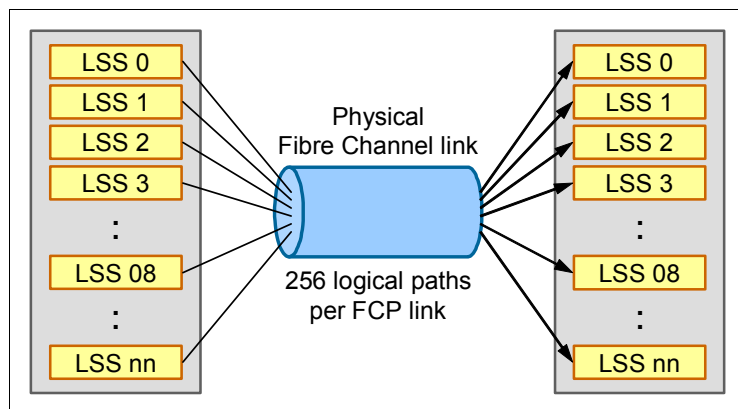
Figure 1-9   PPRC paths and links

Tip: The paths must be in both directions, from production to backup DS8000 and from backup to production DS8000. Otherwise, a failback operation is not possible. The paths from backup to production can use any of the available links. No constraint exists to use the same paths from production to backup.

Note: Although a DS8000 I/O port can handle replication traffic and host connections at the same time, for performance reasons, we do not consider it a preferred practice to mix them. Instead, dedicate I/O ports to either replication traffic or host connections.

1.3.3 Metro Mirror and IBM PowerHA SystemMirror for i

Within the IBM PowerHA SystemMirror for i product, Metro Mirror is used to replicate data that is contained in an IASP from the current production system to the current backup system. Both systems are in a cluster, and each partition uses a DS8000 external storage device for the IASP disks. No rule exists about SYSBAS disks. They can be internal drives or external storage.

Figure 1-10 shows an overview of a Metro Mirror installation setup for an IASP. Both IBM i partitions are members of a cluster and Metro Mirror replication is in place for the IASP volumes.
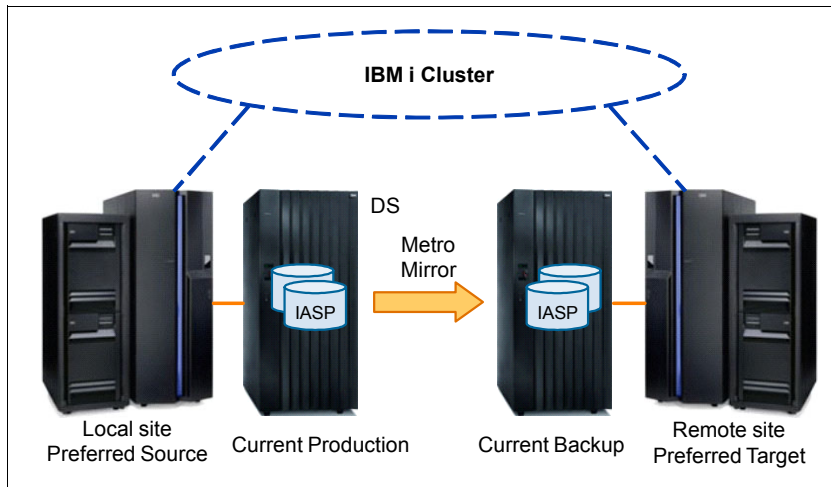
Figure 1-10   Metro Mirror implementation

When a switch occurs, either for a planned outage or an unplanned one, the current backup on the preferred target system starts to be current production. IASP volumes become available due to Metro Mirror failover operations and the partition can use them for production activity. The replication direction is reversed and can be resumed.

When it is possible to return to a normal situation, failback operations can start and the partition in the preferred source becomes the current production and the disks can start production activity. The replication direction changes to the normal direction and replication can be resumed to the current backup in the preferred target system.

1.4 Global Mirror

Global Mirror occurs when the distance from the source site to the target site is longer than the distance that is supported for Metro Mirror replication or when synchronous replication is not an option due to its impact on application performance. However, simple asynchronous replication (global copy) is unable to handle data consistency on the target site. To ensure consistency, Global Mirror also uses FlashCopy.

Global Mirror combines two DS8000 techniques: Global Copy and FlashCopy. Global Copy itself is not covered in this book because it is not an HA/DR solution. However, it can be used for storage migration. Global Mirror uses almost the same technique as Metro Mirror, with two main differences:

► The replication is *asynchronous*. A delay occurs before writes are effective on the target site. The recovery point objective (RPO) is not zero in this case. It mainly depends on network bandwidth and latency between the source and target sites. Figure 1-11 illustrates how Global Copy operates when a host requires a write operation:

   a. Data is stored in the source DS8000 cache and NVS to be destaged to the current production volumes later.

   b. A write acknowledgment is sent by the source DS8000 to the host.

   c. Later (that is, in an asynchronous manner), the source DS8000 sends the necessary data so that the updates are reflected on the current backup volumes. The updates are grouped in batches for efficient transmission.

   d. The target DS8000 returns write complete to the source DS8000 when the updates are written to the target DS8000 cache and NVS.
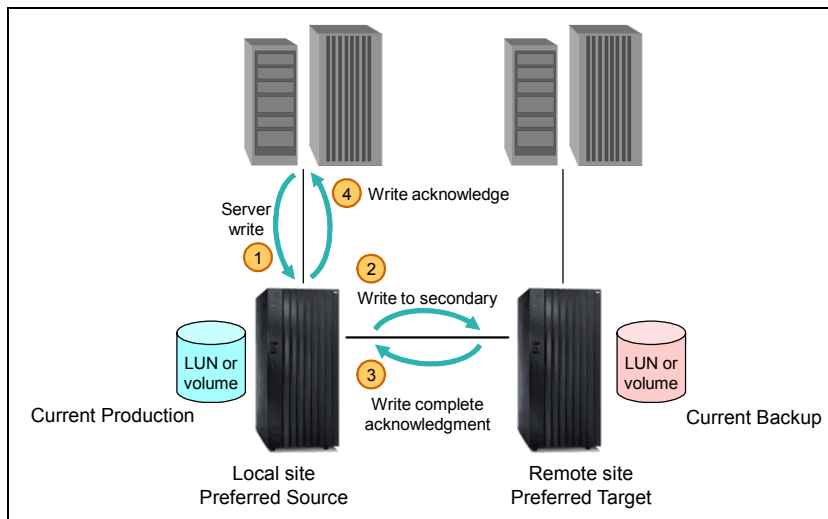
Figure 1-11   Global Copy write sequence

► No guarantee exists that the write sequence on the target site is the same as it is on the source site. Therefore, *data is not consistent* on the current backup. It becomes consistent when the application or host that uses the current production volumes is quiesced. So data migration is the main usage of Global Copy.

FlashCopy occurs to help maintain consistency volumes, as described in Figure 1-12. The following specific FlashCopy attributes are required for Global Mirror:

► Inhibit target write: Protect FlashCopy target volume C from an update by any user or product other than Global Mirror.

► Enable change recording: Apply changes only from the source volume to the target volume that occurred to the source volume in between FlashCopy establish operations, except for the first time that FlashCopy is established.

► Make relationship persistent: Keep the FlashCopy relationship until explicitly or implicitly terminated.

► Nocopy: Do not initiate a background copy from the source to the target, but keep the set of FlashCopy bitmaps that are required for tracking the source and target volumes. These bitmaps are established the first time that a FlashCopy relationship is created with the nocopy attribute. Before a track in source volume B is modified, between consistency group creations, the track is copied to target volume C to preserve the previous point-in-time copy. This copy includes updates to the corresponding bitmaps to reflect the new location of the track that belongs to the point-in-time copy.

Note: Each Global Copy write to its target volume within the window of two adjacent consistency groups can cause FlashCopy I/O operations.
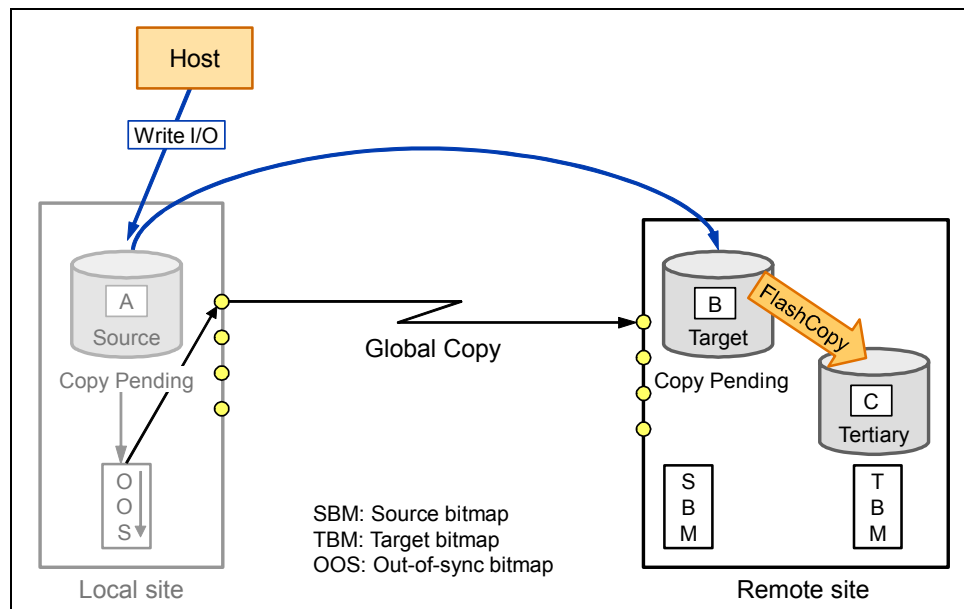
Figure 1-12 Global Mirror implementation

Consistency group creation requires three steps that are automatically processed by the DS8000. See Figure 1-13. After step 3 is complete, the C volumes represent the consistency group and the data is consistent on these volumes.

When consistency group creation is triggered, always by the source site, the following three steps occur:

1. Serialize all Global Copy source volumes. This action imposes a brief hold on all incoming write I/Os to all involved Global Copy source volumes. After all source volumes are serialized, the pause on the incoming write I/O is released and all further write I/Os are noted in the change recording bitmap. They are not replicated until step 3 is finished, but application write I/Os can continue immediately.

2. The drain process includes the process to replicate all remaining data that is indicated in the out-of-sync bitmap and still not replicated. After all out-of-sync bitmaps are empty, the third step is triggered by the microcode from the local site.

3. Now, the B volumes contain all data as a point-in-time copy and are consistent due to the serialization process in step 1 and the completed replication or drain process in step 2. Step 3 is now a FlashCopy that is triggered by the local system microcode as an inband FlashCopy command to volume B as the FlashCopy source, and volume C as the FlashCopy target volume. This FlashCopy is a two-phase process. First, the FlashCopy command runs to all involved FlashCopy pairs. Then, the master collects the feedback and all incoming FlashCopy completion messages. When all FlashCopy operations are successfully completed, the master concludes that a new consistency group was successfully created.

FlashCopy applies here only to data that changed since the last FlashCopy operation because the "enable change recording" property was set at the time that the FlashCopy relationship was established. The FlashCopy relationship does not end due to the nocopy property, which is also assigned at FlashCopy establish time. The nocopy attribute results in B volumes that are not fully replicated to the C volumes by a background process. Bitmaps are maintained and updated instead.
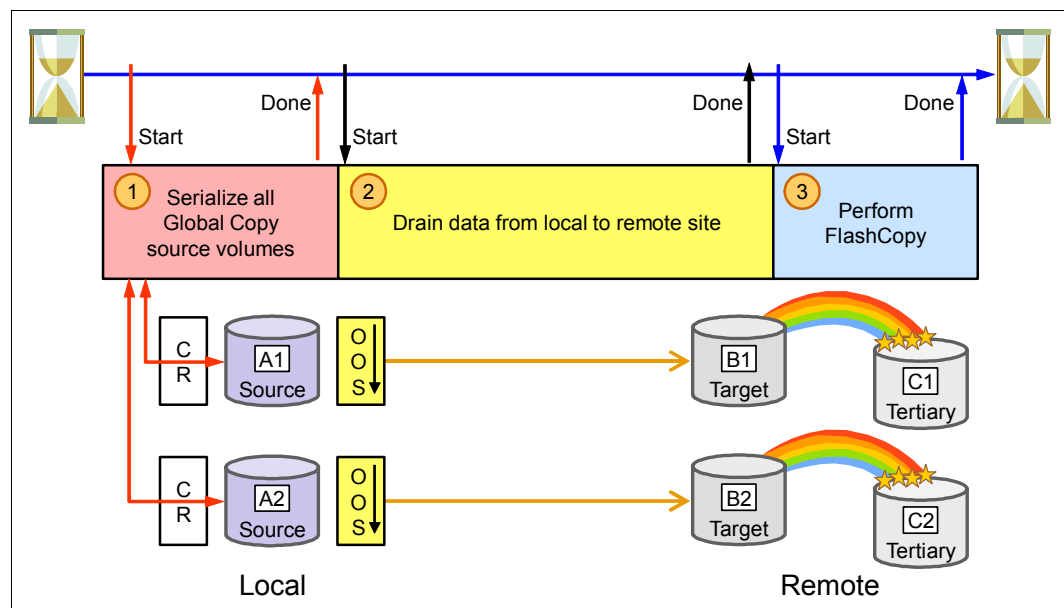
Figure 1-13   Global Mirror consistency group formation

Several parameters are used for Global Mirror tuning. These parameters determine how long to wait before it starts the formation of a new consistency group (for example, performing the three steps described previously). The default is 0 seconds. Therefore, consistency group creation happens constantly. As soon as a consistency group is created, the creation of another consistency group starts immediately.

1.4.1 Global Mirror operations

This section briefly reviews the following basic operations that can be performed with Global Mirror:

► Establishing Global Copy and FlashCopy relationships

   The following operations are needed before Global Mirror activation:

   – Global Copy

      This operation establishes the copy relationship between the source (or local) volume and the target (or remote) volume. During the copy, the target volume gets a *target copy pending* state. For Global Mirror purposes, the Global Copy command with default parameters values can be used.

      Note: A PPRC path must exist from the source to the target site before the Global Copy pair is created.

   – FlashCopy

      Before you create the FlashCopy relationship on the target site, we suggest that you wait until the end of the Global Copy initial copy (as soon as the initial copy step is finished or the *out of sync tracks* value is close to zero).

      Specific FlashCopy settings (Target inhibit, Record, No copy, and Persistent (which is automatically set with Record)) are to be specified.

► Creating a Global Mirror session

   The *session* is an object that contains a reference to LSS, which is part of a Global Mirror relationship through its volumes. The session creation operation must be performed for each LSS by using a unique session ID for the entire storage image.

► Adding Global Copy volumes to the Global Mirror session

   Global Copy source volumes can be added at any time to the Global Mirror session. The added volumes must match LSS.

► Starting a Global Mirror for a session

   Starting a Global Mirror requests that the source DS8000 start to create the consistency groups.

► Ending a Global Mirror for a session

   Ending a Global Mirror requests that the source DS8000 stop the creation of the consistency groups. This operation tries to keep the last consistency group available.

► Pausing and resuming a Global Mirror for a session

   Pausing and resuming a Global Mirror can be used in complex situations, for example, to run tests on target sites and keep normal business on source sites. This kind of scenario involves the use of a fourth set of volumes (which are called *D volumes*) on the target site.

► Monitoring a Global Mirror for a session

Commands and graphical user interface (GUI) windows exist to help monitor the Global Mirror status. For example, the `showgmiroos` command shows how many tracks are not synchronized from the source to the target.

► Terminating a Global Mirror environment

Terminating a Global Mirror environment requires that you request several steps in a specific order:

a. End Global Mirror.

b. Remove Global Copy volumes from the session.

c. Remove the Global Mirror session.

d. Terminate the FlashCopy pairs.

e. Terminate the Global Copy pairs.

f. Remove the paths.

► Failover

Failover operations for Global Mirror involve more steps than for Metro Mirror. The first part of the operation is to unlock the target LUNs. They will then have a status of Suspended. But, those volumes are not consistent. It is necessary to use FlashCopy target volumes (C volumes), which are consistent.

A failover uses the following steps:

a. Global Copy failover.

The situation is shown in Figure 1-14.
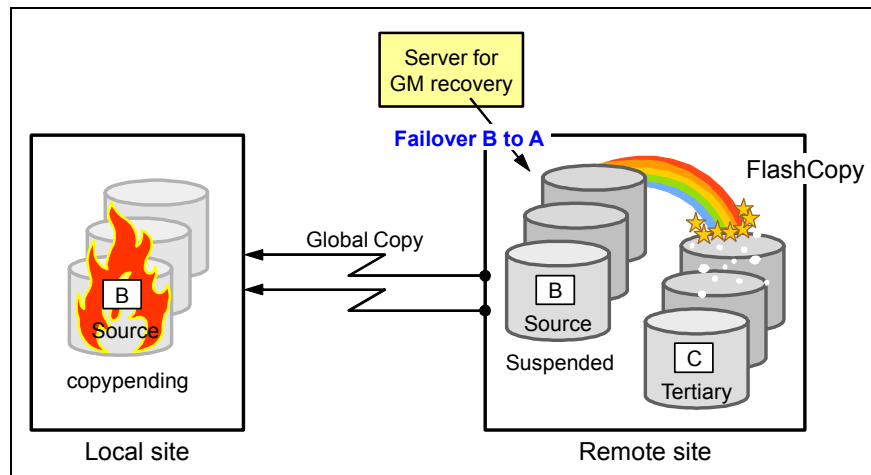
Figure 1-14   Global Mirror Global Copy failover

b. Set consistent data on the B volumes.

To make the B volumes consistent, you must use FlashCopy reverse, which copies content back to the B volumes from the C volumes. This background copy updates all changes on the B volumes since the last consistency group formation with the previous data that was stored on the C volumes. This operation terminates the FlashCopy relationship between the B volumes and the C volumes.

The situation is shown in Figure 1-15, but at the end, no FlashCopy relationships exist.
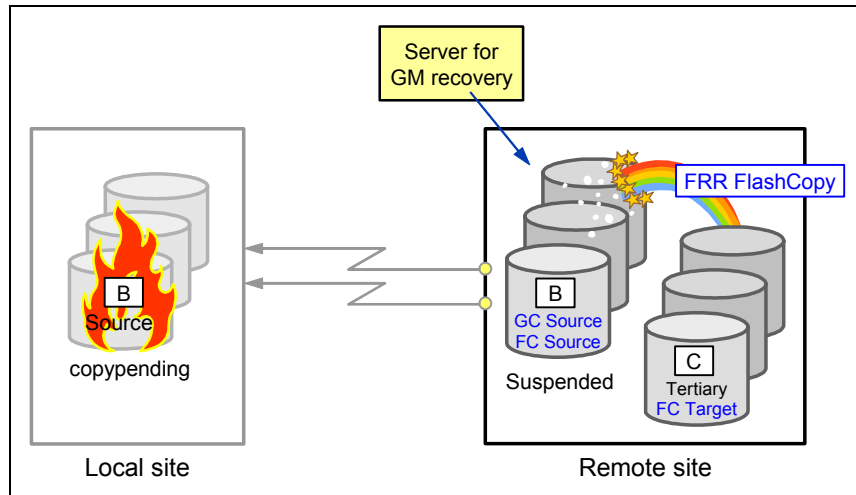
Figure 1-15   Global Mirror reverse FlashCopy

    c. Reestablish a FlashCopy relationship between volumes B and C.

       Because of the termination of the FlashCopy relationship, you must re-create it with the same settings that were used at the initial creation. At the end of this step, the situation is in the same state as it was after the first step, but with consistent data on the B volumes, and you can start host connections to these volumes.

► Failback

    When the source site becomes available in a Global Mirror environment, the first step is to run a failback operation from the target B volumes to synchronize the content of volume A with the content of volume B.

    The second step is to make volume A the source again and to start the synchronization of the B volumes with the A volumes.

    After Global Copy and FlashCopy are working again correctly, the last operation is to restart the Global Mirror session.

> **Note:** Each time that a failover operation is performed, applications can be quiesced before the failover operation for consistent data with no difference between the source and target data.

1.4.2 Global Mirror and IBM PowerHA SystemMirror for i

Within the IBM PowerHA SystemMirror for i product, Global Mirror is used to replicate data from an IASP to another IASP in a remote site when the distance does not allow synchronous replication with Metro Mirror.

As with Metro Mirror replication, the installation consists of a preferred source with an IBM i current production partition and a preferred target system with an IBM i current backup partition. Both partitions are in a cluster, and each partition uses a DS8000 external storage device for IASP disks. No rule exists about SYSBAS disks. They can be internal drives or use external storage.

Figure 1-16 shows an overview of a Global Mirror installation setup for an IASP. Both IBM i partitions are members of a cluster, and the Global Mirror replication is in place for the IASP volumes.
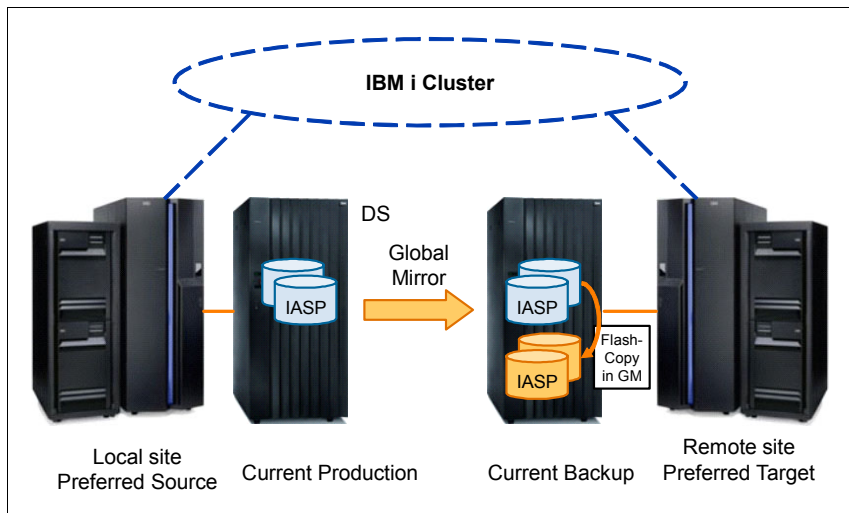
Figure 1-16   Global Mirror implementation

When a switch occurs, either in a planned outage or an unplanned outage, the current backup IASP volumes are available due to the Global Mirror switchover/failover. The partition on the preferred target can become the current production partition, and it can use them for production activity. The direction of replication is reversed and ready to resume.

When it is possible to return to a normal situation, failback operations can start, and the partition on preferred source becomes the current production partition and can restart production activity. The replication is reversed, and it can resume to the current backup on the preferred target location.

IBM PowerHA SystemMirror for i handles all setup, switch, and switchback tasks but one. It is not able to establish the Global Mirror configuration between the source and target volumes. The DS8000 setup is performed either through the data storage command-line interface (DSCLI) or Storage Manager GUI.

1.5 LUN-level switching

LUN-level switching that is provided by IBM PowerHA SystemMirror for i offers a local HA solution with the IBM System Storage DS8000 series.

With LUN-level switching, a single-copy of an IASP can be switched between IBM i systems in a cluster (Figure 1-17).
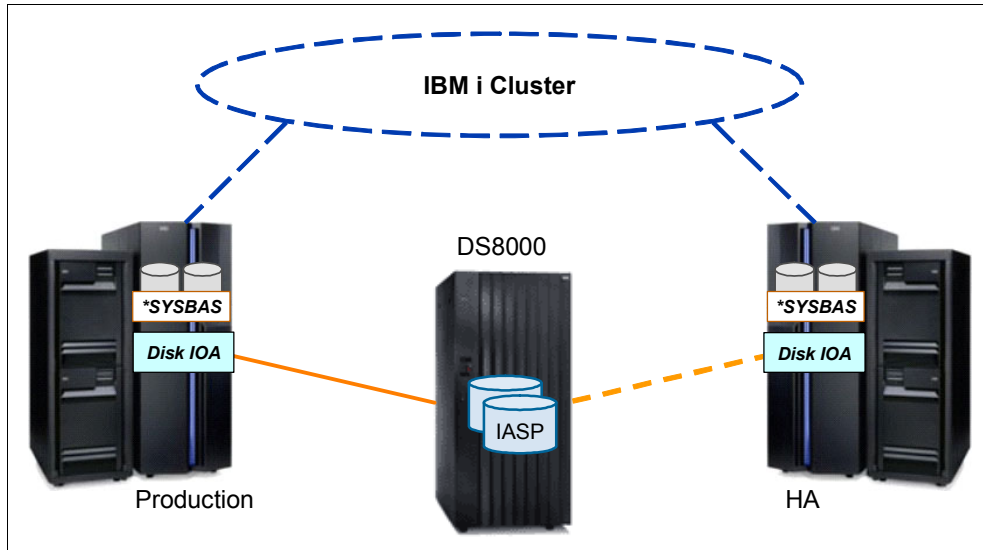
Figure 1-17   LUN-level switching between servers

A typical implementation scenario for LUN-level switching is at the production site for use for local HA protection, eliminating the requirement for a site-switch in potential IBM i server outages.

1.6 FlashCopy

With FlashCopy, you can create a point-in-time copy of the logical volume. By performing a FlashCopy, a relationship is established between a source volume and a target volume. Both volumes are considered to form a *FlashCopy pair*. As a result of the FlashCopy, either all physical blocks from the source volume are copied or, when you use the *nocopy* option, only those parts that changed are copied.

The target volume needs to be the same size as the source volume whenever FlashCopy is used to flash an entire volume.

While the IBM i spreads its data across several volumes, all of the volumes in an IASP must be FlashCopied at the same point-in-time. FlashCopy offers consistency groups so that multiple volumes can be FlashCopied at the same time.

1.6.1 Establishing a FlashCopy relationship

When the FlashCopy starts, the relationship between the source and the target is established within seconds by creating a pointer table, which is called a *bitmap* for the target.

If all bits for the bitmap of the target are set to their initial values, it means that no data block was copied so far. The data in the target is not modified during the setup of the bitmaps. During this first step, the bitmap and the data look as illustrated in Figure 1-18.

Note: The target volume in the following figures can be a normal volume or a virtual volume (space-efficient (SE) volume). In both cases, the logic is the same.
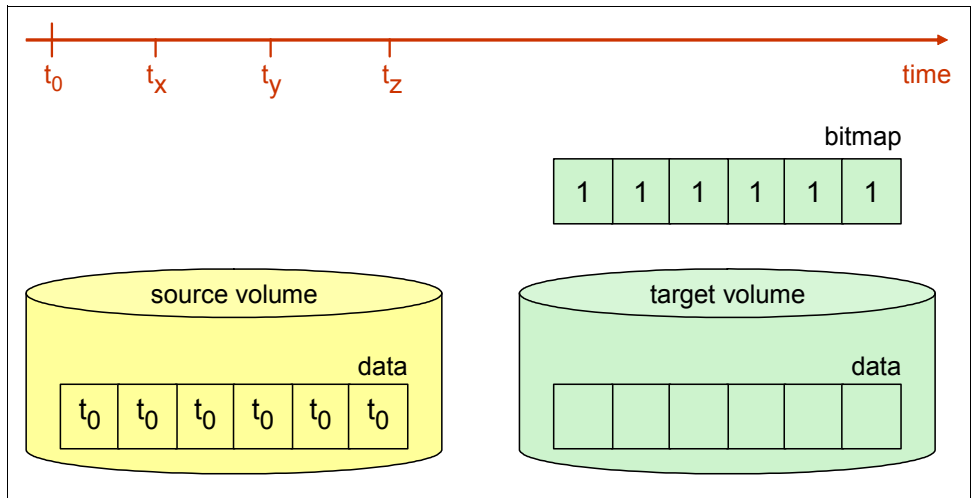
Figure 1-18   FlashCopy at time t_0

After the relationship is established, it is possible to perform read and write operations on both the source and the target. Assuming that the target is used for reads only while production is ongoing, the process works as illustrated in Figure 1-19.
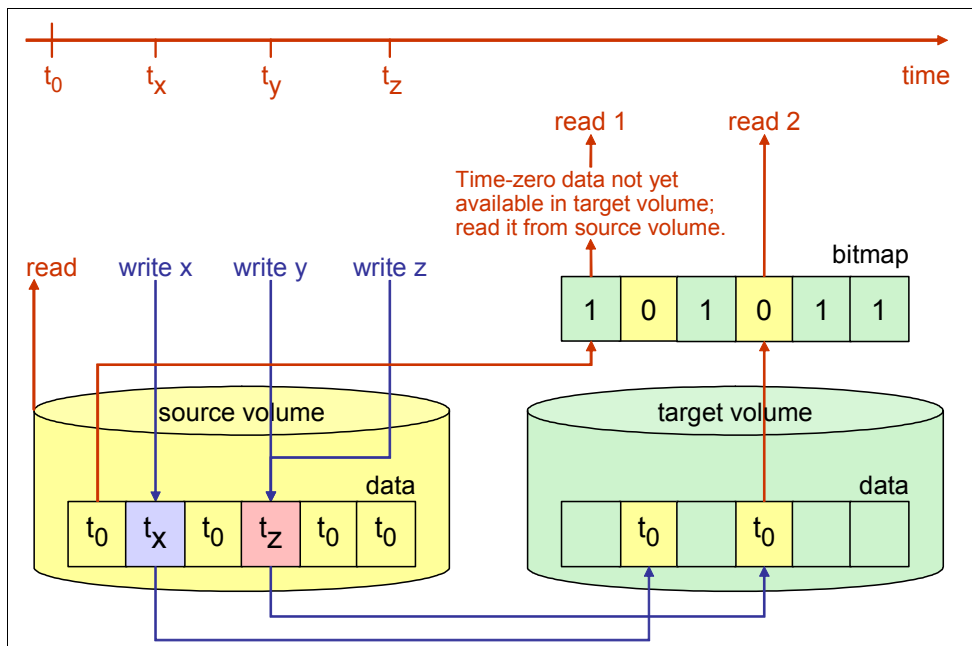
Figure 1-19   Reads from source and target volumes and writes to source volume

Reading from the source

The data is read immediately.

Writing to the source

Whenever data is written to the source volume while the FlashCopy relationship exists, the storage system ensures that the time-zero (t_0) data is copied to the target volume before the storage system overwrites it in the source volume. When the target volume is a space-efficient volume, the data is written to a repository.

To identify whether the data of the physical track on the source volume needs to be copied to the target volume, the bitmap is analyzed. If it identifies that the t_0 data is not available on the target volume, the data is copied from source to target. If it states that the t_0 data was copied to the target volume, no further action is taken.

Reading from the target

Whenever a read request goes to the target while the FlashCopy relationship exists, the bitmap is used to identify whether the data must be retrieved from the source or from the target. If the bitmap states that the t_0 data was not copied to the target, the physical read is directed to the source. If the t_0 data was copied to the target, the read is performed immediately against the target (Figure 1-19 on page 23).

Writing to the target

Whenever data is written to the target volume while the FlashCopy relationship exists, the storage system ensures that the bitmap is updated. This way, the t_0 data from the source volume never overwrites updates that were made directly to the target volume. Figure 1-20 illustrates the concept of writes to the target volume.
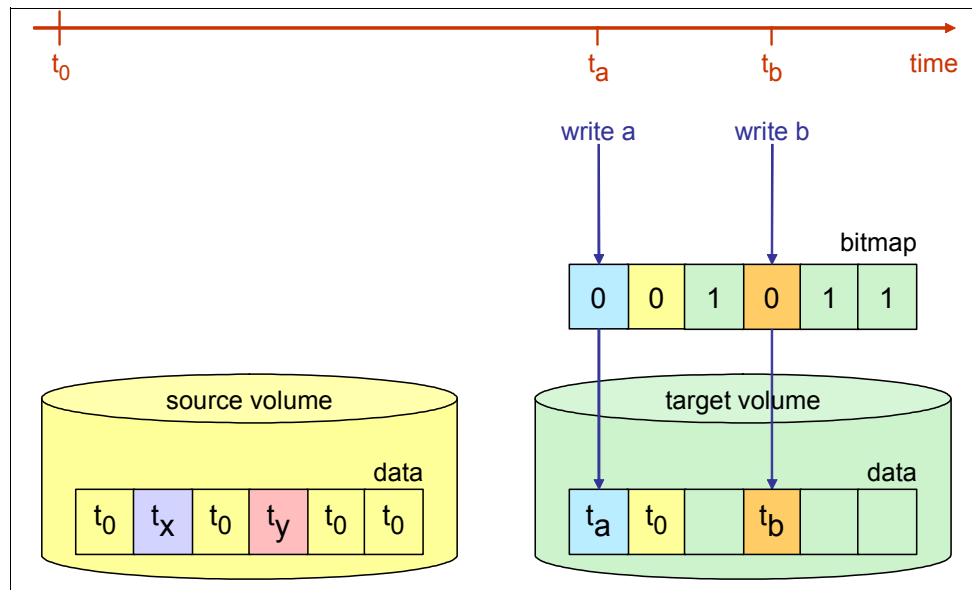
Figure 1-20   Writes to the target volume

1.6.2 Terminating the FlashCopy relationship

The FlashCopy relationship is ended automatically when all tracks were copied from the source volume to the target volume if the FlashCopy option was to copy the data.

The relationship can also be withdrawn explicitly by issuing the corresponding commands.

A FlashCopy space-efficient relationship ends when it is withdrawn. When the relationship is withdrawn, an option is available to release the allocated space of the space-efficient volume.

Full volume copy

When the copy option is invoked and the establish process completes, a background process starts that copies all data from the source to the target. If the FlashCopy relationship is not defined explicitly as *persistent*, the FlashCopy relationship ends as soon as all data is copied.

Nocopy option

If FlashCopy is established by using the nocopy option, the relationship lasts until it is withdrawn explicitly or until all data in the source volume is modified. Blocks for which no write occurred on the source or on the target stay as they were at the time that the FlashCopy was established.

If the persistent FlashCopy option was specified, the FlashCopy relationship must be withdrawn explicitly.

1.6.3 FlashCopy SE

IBM PowerHA SystemMirror for i supports space-efficient FlashCopy of the IBM System Storage DS8000 series.

The IBM System Storage DS8000 series FlashCopy SE licensed feature allows the creation of space-efficient FlashCopy target volumes that can help to reduce the required physical storage space for the FlashCopy target volumes. A space-efficient FlashCopy target volume has a virtual storage capacity that is reported to the host that matches the physical capacity of the fully provisioned FlashCopy source volume, but no physical storage space is ever allocated.

Physical storage space for space-efficient FlashCopy target volumes is allocated in 64 KB track granularity. This allocation is performed on demand for host write operations from a configured repository volume that is shared by all space-efficient FlashCopy target volumes within the same DS8000 extent pool.

See Figure 1-21.
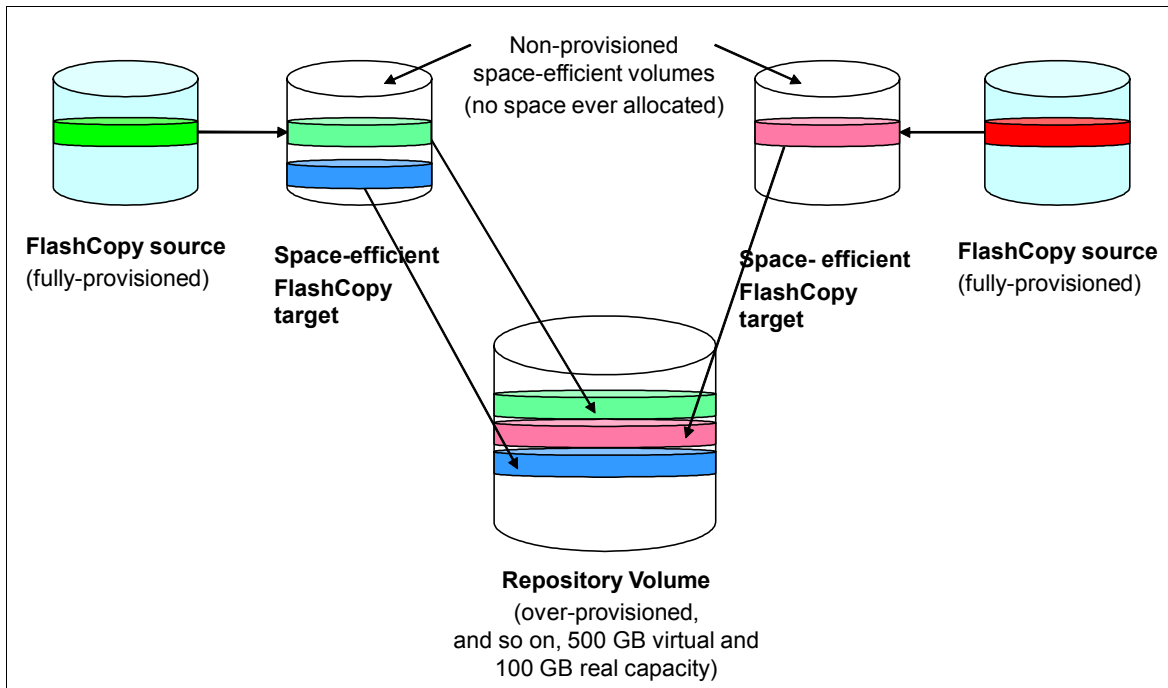
Figure 1-21   DS8000 space-efficient FlashCopy

Planning

This chapter describes the required planning steps to implement IBM PowerHA SystemMirror for i with the DS8000 storage server series.

The following topics are described in this chapter:

27

2.1 IBM PowerHA SystemMirror for i preparation

To set up a PowerHA environment on the DS8000 series, various hardware and software components need to be prepared. This section provides information about these requirements.

2.1.1 Hardware

The following hardware requirements must be met:

► Minimum of one storage Hardware Management Console (HMC)

> **Note:** We consider it a preferred practice to use two storage HMCs for each storage system for redundancy.

► Two Power Systems with latest available firmware
► Two DS8000 storage devices

Use two DS8000 family storage devices with adequate resources, depending on the required storage and I/O requirements. Each storage device needs an HMC. For FlashCopy with DS8000 system storage devices, you need one set of disks for the source, an equal set of disk units for the target, and another set for each consistency group. In certain cases, space-efficient FlashCopy (FlashCopy SE) can be used to reduce this disk requirement.

► Heartbeat interfaces

Heartbeat interfaces are used to ensure efficient communication between all cluster nodes. They do not need high bandwidth because they exchange only a small amount of information.

To avoid situations in which the users can no longer connect to the node and at the same time the nodes can no longer exchange heartbeat information, consider the use of separate network devices for both user access and heartbeat access. Conversely, heartbeat exchanges can activate redundancy by using virtual IP addressing, Etherchannel, or two IP interfaces at the cluster definition level. If you use multiple physical links, the links need to use separate adapters and ideally connect to a redundant network infrastructure.

► Communication with an HMC or Virtual I/O Server (VIOS)

When you use advanced node failure detection, the HMC or VIOS must be able to send information to the cluster nodes in a partition failure. This communication uses the Internet Protocol (IP) routes that are available on the system. Ensure that no single point of failure (SPOF) can affect this communication.

If a firewall exists between the HMC or VIOS and the cluster nodes, ensure that the ports are open. Advanced node failure detection uses the Common Information Model (CIM) object manager (CIMOM) services, which listen on Transmission Control Protocol (TCP) port 5989 (unencrypted communication) and port 5990 (Secure Sockets Layer (SSL)-encrypted communication). These ports must be open in both directions.

- ► Communication with the DS8000

  When you use the DS8000, the cluster nodes need to establish IP communication with the DS8000 to check the replication status and start the correct actions, such as failover or failback. This communication uses the data storage command-line interface (DSCLI) that comes with the DS8000 storage unit, and it must be maintained at the same level as the microcode of the storage units that you are communicating with.

  If a firewall exists between the cluster nodes and the DS8000, ensure that the flow is authorized. Many ports are required for DSCLI to communicate and these ports are documented in the IBM DS8000 Series V7.2 Knowledge Center:

  `http://www.ibm.com/support/knowledgecenter/HW213_7.2.0/com.ibm.storage.ssic.help.doc/f2c_cliesscli_1kx2so.html?lang=en`

  Each system/partition has its own Fibre Channel attachment cards, or access to N_Port ID Virtualization (NPIV) adapters. The independent auxiliary storage pool (IASP) and SYSBAS logical unit numbers (LUNs) must connect to different physical or virtual adapters. The logical subsystem (LSS) must be unique for any specific workload because many commands in the storage system operate on the LSS level. Failing to separate the workloads can lead to unpredictable results.

- ► Storage inter-site link

  The DS8000 communicates over a fiber connection to its replication partner. It is also possible to use IP-based traffic between storage area network (SAN) switches.

  > **Important:** Ensure that the bandwidth and latency of the link are suitable for the planned usage. Based this choice on many factors, including the following areas:
  >
  > - ► Volume and size of transactions to be transmitted
  > - ► Response time requirements
  > - ► Distance between the nodes
  > - ► Line availability
  > - ► Cost considerations

2.1.2 Software

The following software is required:

- ► IBM i with the latest program temporary fixes (PTFs)
- ► IBM PowerHA SystemMirror for i 5770-HAS (The edition depends on your requirements.)
- ► 5770-SS1 option 30 Qshell
- ► 5770-SS1 option 33 Portable Application Solution Environment
- ► 5770-SS1 option 34 Digital Certificate Manager (if required)
- ► 5770-SS1 option 35 Common Cryptographic Architecture (CCA) Cryptographic Service Provider (if required)
- ► 5770-SS1 option 41 HA Switchable Resources
- ► 5770-DG1 Base IBM HTTP Server for i
- ► 5773-SC1 Base IBM Portable Utilities for i
- ► 5733-SC1 option 1 OpenSSH, OpenSSL, zlib functions

- ► 5770-UME Base IBM Universal Manageability Enablement for i
- ► DS command-line interface (DSCLI), of the same level as the microcode that is installed on your DS8000 storage. For more information, see 3.1, "Setting up DS8000 Copy Services" on page 44.

> **Important:** If you upgrade your DS8000 microcode, you also must upgrade the DSCLI that is installed on your IBM i partitions to the new version that came with the new microcode level.

2.2 Sizing the backup system

Ensure that your backup system is sized so that it can run the production workload if you switch or fail over to the system. You need to test your high availability (HA) solution regularly by making the current backup system the production system and then running your production workload on that system. Do not wait for a reason to test the switchover. Start the applications and immediately switch back to the old production system.

Companies that regularly switch between the systems and consider them to be equal instead of having one that is designated the production system and one that is designated the backup system are most likely the companies that know how to switch or fail over. Also, they have a deep trust in their HA solution and do not waste time with lengthy decision processes when a system outage happens.

However, if your backup system is sized smaller than your production system (either on the processor configuration or on the storage configuration), ensure that you understand what the size difference that means to your production workload. Giving everybody access to an undersized backup system in a production failure might lead to a situation where you overload the backup system, making it more or less unusable. Instead, you need a plan in place that clearly establishes the workloads that can be run on the smaller system and the workloads and users that do not get access to that system.

The same is true if you are running development and test environments on your backup system. Ensure that you do not need HA for these environments. Consider the following points:

- ► 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 either run your production, development, and test workloads concurrently on the other system, or do you have a plan in place about the workloads that you do not run in this case?

Answering these questions can lead to a decision that your test and development systems require HA too, either permanently or during periods of heavy development activities.

2.3 Synchronous versus asynchronous replication

The DS8000 family offers synchronous (Metro Mirror) and asynchronous (Global Mirror with consistency groups) mirroring. With PowerHA that supports both options, you need to decide the setup that best fits your requirement and the available infrastructure.

Synchronous mirroring guarantees 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 and all transactions that were completed on the source system are also available on the target system). Also, the network latency and available bandwidth directly affect I/O response times and application response times.

With asynchronous mirroring, the source storage system sends back the I/O confirmation to the server before it sends data over to the target storage system. Therefore, network latency and available bandwidth do not affect I/O or application response times. However, no guarantee exists that the data is identical between the source and the target. A guarantee exists though that 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, you must be prepared to lose part of the data when the source system fails. Therefore, an RPO of zero cannot be met. Ensure that you understand what this loss means to your business:

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

When you consider Global Mirror, ensure that spikes in your write characteristics do not lead to a situation where the connection is never able to catch up on data that is still waiting to be sent to the target site. A simple approach to calculate this size difference is shown:

► Collect write data rates over a representative amount of time and use small collection intervals (for example, 1 minute) to ensure that you do not even out your results.

► Assume 1 byte is 10 bits (to accommodate for protocol overhead).

► For each period:

  – If the complete data from the previous period was sent over (write data rate in bits is smaller or equal to the available bandwidth in bits), consider the current period write data rate only.

  – If incomplete data from the previous period was sent over (write data rate in bits is larger than the available bandwidth in bits), add the difference between the bandwidth and write data rate from the previous period to the write data rate of the current period.

► By 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. Ensure that you also include the specific amount of required bandwidth for intersystem heartbeat traffic in your calculation.

2.4 Space-efficient LUNs

With a FlashCopy SE space-efficient relationship, disk space is only allocated for the target when a write changes a sector on the source, or when a write is directed to the target. For this reason, the use of FlashCopy SE requires less disk capacity than the use of standard FlashCopy, which can help lower the amount of physical storage that is needed.

FlashCopy SE is designed for temporary copies, so FlashCopy SE is optimized for use cases where a small percentage of the source volume is updated during the life of the relationship. If more than 20% of the source is expected to change, a trade-off might exist in terms of performance versus space efficiency. Also, the copy duration must generally not last longer than 24 hours unless the source and target volumes have little write activity.

The DS8000 FlashCopy SE repository is an object within an extent pool, and it provides the physical disk capacity that is reserved for space-efficient target volumes. When you provision a repository, storage pool striping is automatically used with a multirank extent pool to balance the load across the available disks. FlashCopy SE is optimized to work with repository extent pools that consist of Redundant Array of Independent Disks (RAID) arrays. After the repository is defined in the extent pool, it cannot be expanded, so planning is important to ensure that it is configured to be large enough. If the repository is full, the FlashCopy SE relationships fail. After the relationship fails, the target becomes unavailable for reads or writes, but the source volumes are not affected.

You can estimate the physical space that is needed for a repository by using historical performance data for the source volumes with knowledge of the duration of the FlashCopy SE relationship. Each write to a source volume consumes one track of space on the repository (57 KB for count-key data (CKD) and 64 KB for fixed-block (FB)).

For more information about FlashCopy SE on a DS8000 configuration, see the following resources:

► IBM Redbooks website:

http://www.redbooks.ibm.com/redbooks.nsf/searchsite?SearchView&query=DS8000

► IBM Knowledge Center:

http://www.ibm.com/support/knowledgecenter/HW213_7.2.0/com.ibm.storage.ssic.help.doc/f2c_ichomepage.htm?lang=en

> **Tip:** You can monitor whether the repository reached a specified threshold by using Simple Network Management Protocol (SNMP) traps. By using the Lab Services IASP Copy Services Manager, you can convert and send these messages to the QSYSOPR message queue.

2.5 I/O adapter virtualization

Another consideration with this kind of provisioning is I/O adapter virtualization. By virtualizing the adapters, you can present disks to more partitions than you have physical adapters for.

In a VIOS environment, the preferred method is to use NPIV adapters because NPIV adapters offer benefits if you need to debug a connection. But, rules must be followed:

► Each NPIV adapter within the partition must be mapped to a separate physical adapter port.

► SYSBAS LUNs and IASP LUNs must not share adapters.

2.5.1 N_Port ID Virtualization

N_Port ID Virtualization (NPIV) is a standardized method for virtualizing a physical Fibre Channel (FC) port. An NPIV-capable FC host bus adapter (HBA) can have multiple N_Ports, each with a unique identity. NPIV with VIOS adapter-sharing capabilities allows a physical FC HBA to be shared across multiple guest operating systems.

The PowerVM implementation of NPIV enables POWER logical partitions (LPARs) to have virtual FC HBAs, each with a dedicated worldwide port name. Each virtual FC HBA has a unique SAN identity that is similar to the SAN identity of a dedicated physical HBA.

See Figure 2-1.

DS8000 Access through NPIV for IBM i

IBM DS8000

VIOS Host

IBM i Client

Physical adapter WWPN:
eeeeffff55556666

8Gb

VFC Server Adapter

Virtual FC connection

WWPNs:
aaaabbbb11112222
ccccddddd33334444

VFC Client Adapter

IBM i 520B LUNs mapped to
WWPN aaaabbbb11112222

DD001-DD004

Power Server

- Step 1: Configure physical and virtual FC adapters
  - Gives you virtual WWPNs
- Step 2: Configure SAN fabric and storage/tape
  - Using the virtual WWPNs

Figure 2-1   NPIV

You can provide resources to multiple IBM i client partitions as shown in Figure 2-2.

IBM i + NPIV (Virtual Fiber Channel)

System 1

System 2

System 3

VIOS

8Gbs HBA

IBM i Client (System 1)

IBM i Client (System 1)

IBM i Client (System 1)

•Hypervisor assigns 2 unique WWPNs to each virtual fiber.

•Hostconnect is created as an IBM i hosttype.

•Requires 520 byte per sector LUNs to be assigned to the IBM i hostconnect on DS8000.

Virtual address example C001234567890001

Hypervisor

POWER Systems

Note: An NPIV (N_port) capable switch is required to connect the VIOS to the DS8000 to use virtual fiber.

Figure 2-2   NPIV for multiple IBM i clients

> **Note:** NPIV disks are not seen by VIOS. The LUNs are presented directly to the client partition.

You can also use NPIV for tape media libraries, but you must use separate NPIV adapters for tape and disk.

You can see the physical mapping of NPIV adapters to the physical adapters in VIOS by selecting the VIOS partition on the HMC and then taking the Configuration and Virtual Storage Management option as shown in Figure 2-3.
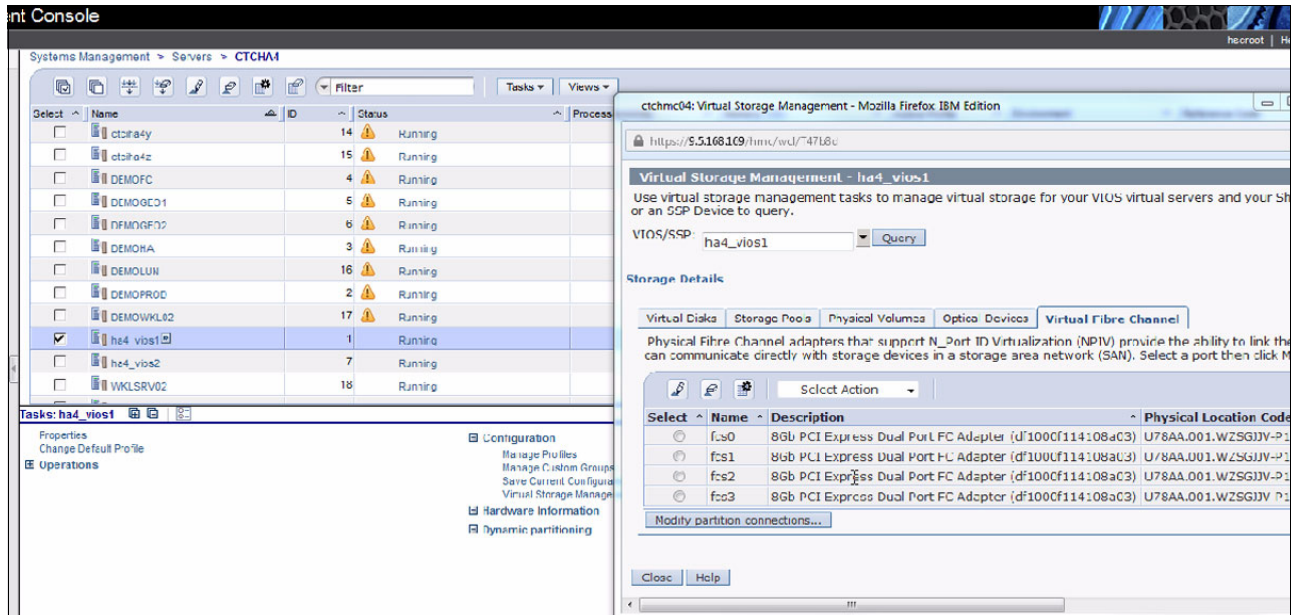
Figure 2-3   NPIV mapping to physical adapters

2.5.2 Virtual SCSI

With a virtual SCSI (VSCSI) connection, the storage is configured to VIOS and then presented to the IBM i client partition, as shown in Figure 2-4.
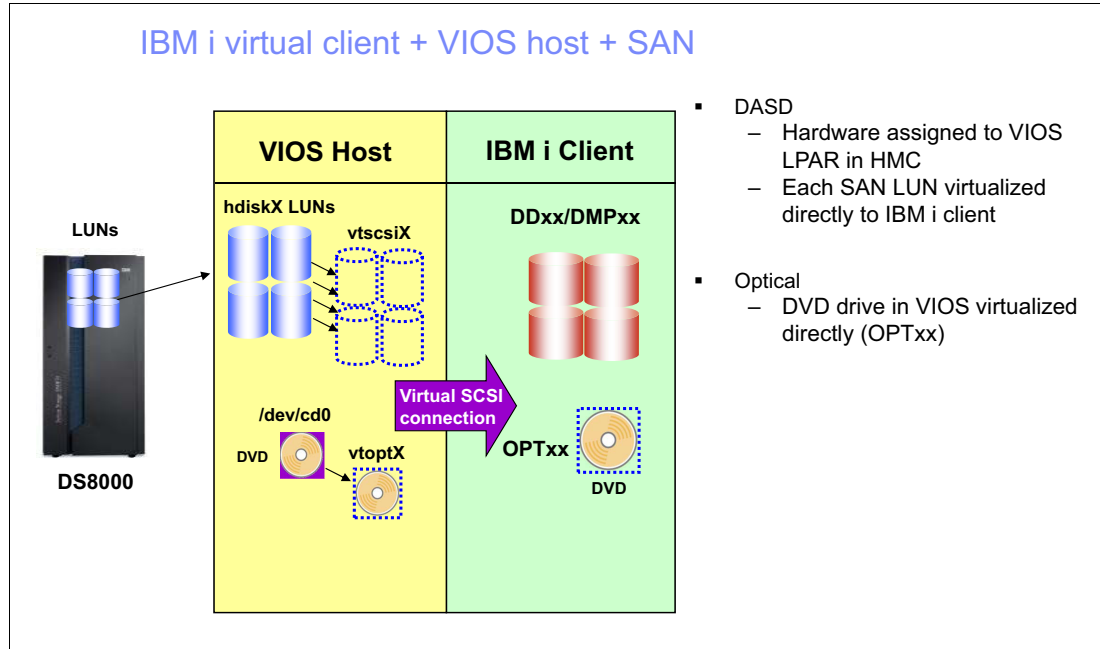
Figure 2-4   Disks and I/O adapter (IOA) virtualization by using the VIOS partition

The LUNs are connected to VIOS and then presented to the IBM i client partition. By using multiple SCSI connection IDs, you can present disks to multiple IBM i client partitions. See Figure 2-5.
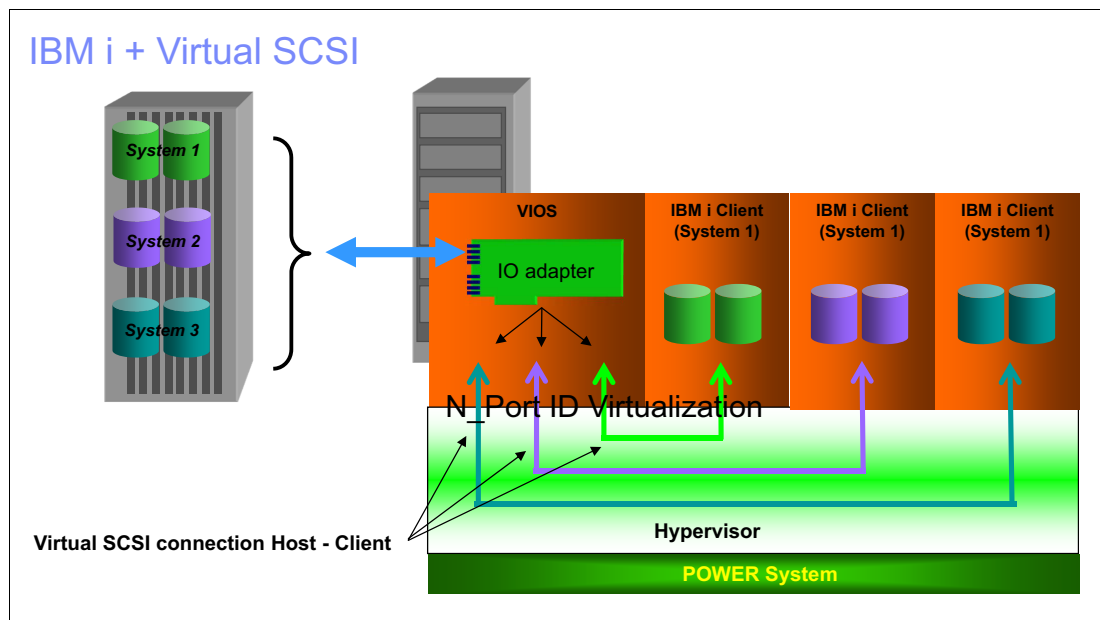
Figure 2-5   Virtual SCSI for multiple IBM i client partitions

2.6 Calculating full synchronization time bandwidth

Writes to the IASP that are geographically mirrored to the mirror copy system and that occur during normal runtime operations must also be considered when you determine bandwidth requirements. The bandwidth requirements can be calculated by sampling the collection services data on the production copy system.

Using Performance Data Investigator (PDI)

This section assumes that Collection Services data ran 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, PDI requires that all of the collections are in a single library, and no more than five collections can be graphed at one time.

The following steps set up the environment for the data to be viewed in the PDI tool:

1. On the IBM i command line, enter the `CFGPFRCOL` command and prompt it with by pressing F4. Make a note of the Collection Library value, which 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 to for this exercise. This library can be named anything. For the example in this example, it is named `BWDAT`:

   `CRTLIB BWDATA`

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

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

   > **Note:** The management collection object 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 (to copy) for up to five collections, which in most collection services configurations represent 5 days of system performance data, and 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 now placed in the new library.

6. 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 Q*xxxxxxxxx* 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 you are prompted, log in with a user that has *ALLOBJ special authority. The IBM Navigator for i Welcome window opens as shown in Figure 2-6.
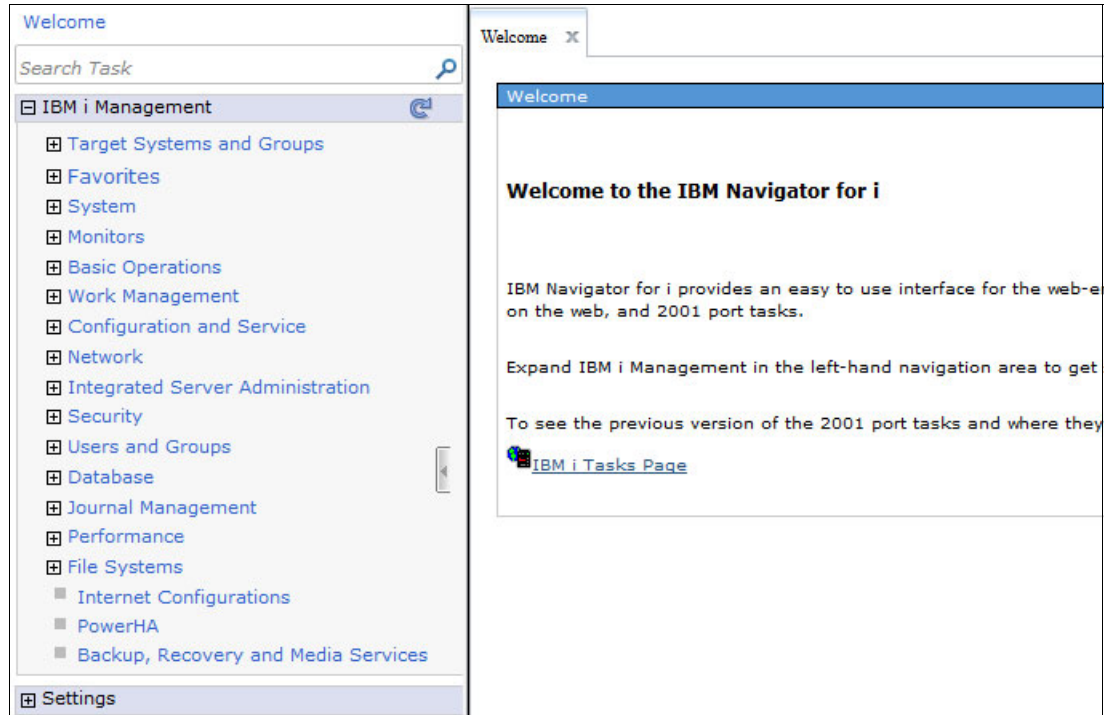
Figure 2-6   IBM Navigator for i Welcome window

9. To view the data in the new library in PDI, the collection table must be rebuilt. In the left pane, expand **Performance** → **Manage Collections** as shown in Figure 2-7.
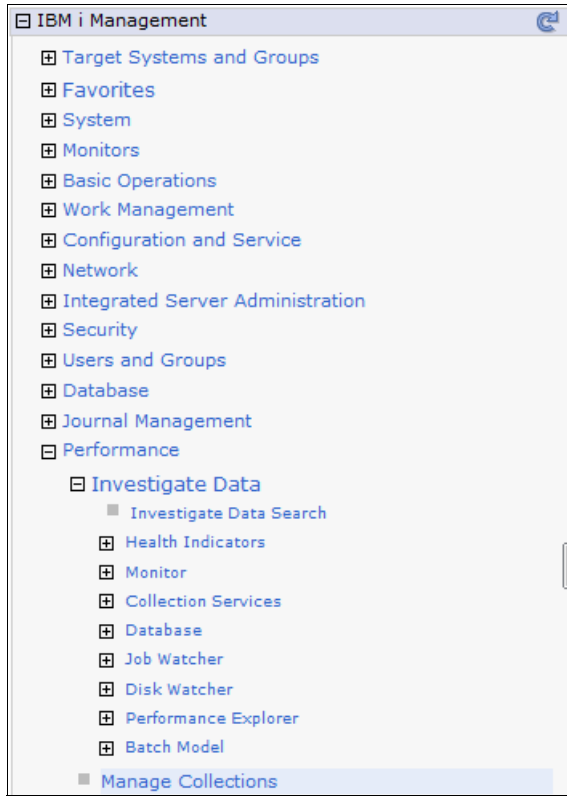
Figure 2-7   PDI Performance, Manage Collections

10. In the right pane of the Manage Collections window, the list of collections is shown. Click **Actions** → **Maintain Collections** → **Rebuild Collection Table** as shown in Figure 2-8.
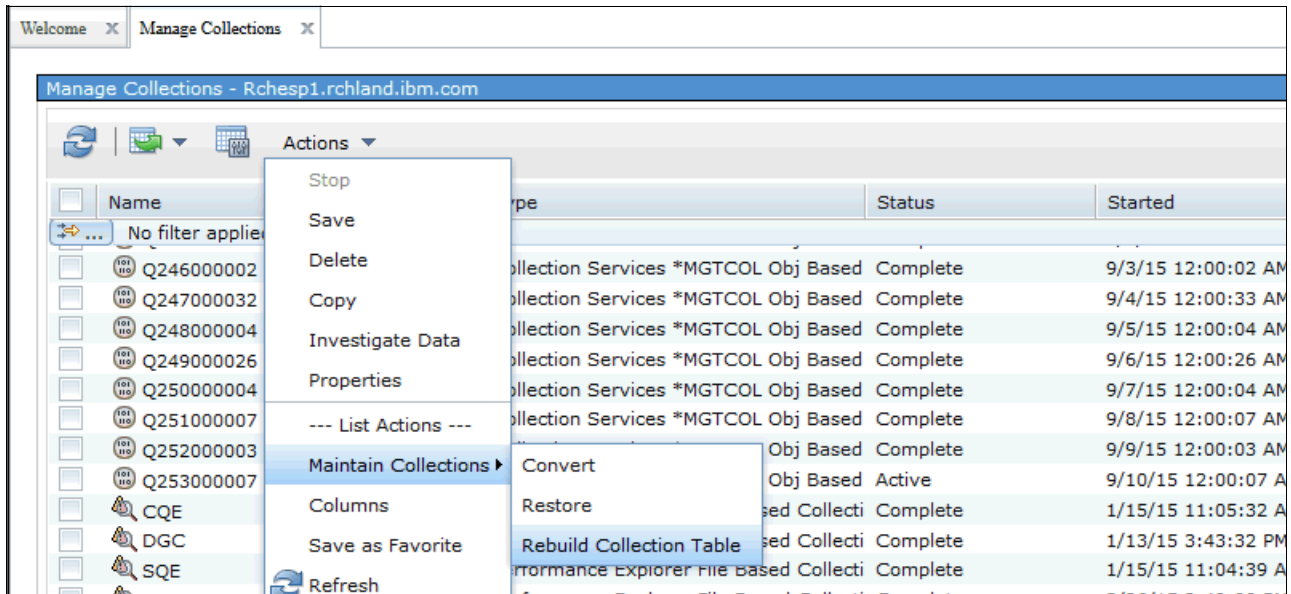
Figure 2-8   PDI, Rebuild Collection Table

11. The rebuild of the collection table takes a few moments to complete. After the rebuild is complete, the Manage Collections list shows the collections in the `BWDATA` library.

In the left pane, expand **Performance** → **Investigate Data** → **Collection Services** → **Disk** → **Disk Throughput Overview for Disk Pools** as shown in Figure 2-9.
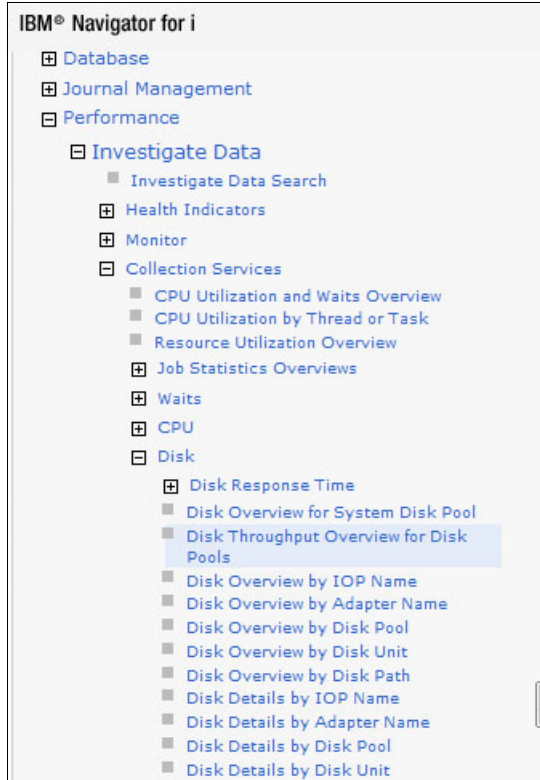
Figure 2-9   PDI Collection Services, Disk

12. In the right pane, change the Collection Library to **BWDATA** and the Collection Name to **All**, then click **Display**. See Figure 2-10.
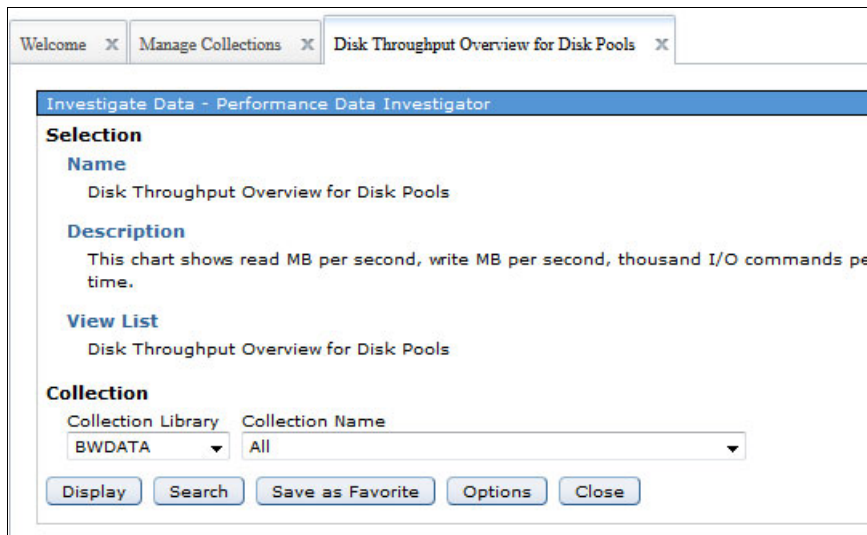
Figure 2-10   Disk Throughput Overview for Disk Pools collection criteria

13. The resulting graph is displayed. You can use the Full Zoom Out tool (see Figure 2-11) to view the graph during the collections.
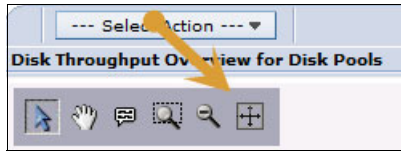
Figure 2-11   PDI Full Zoom Out tool

14. The resulting graph (Figure 2-12) shows the IASP disk activity throughout the duration of the specified collections. In this example, only data for a couple of days exists, but we suggest that you perform this entire process several times to view several weeks of data. This analysis helps you to understand the amount of system IASP write activity.
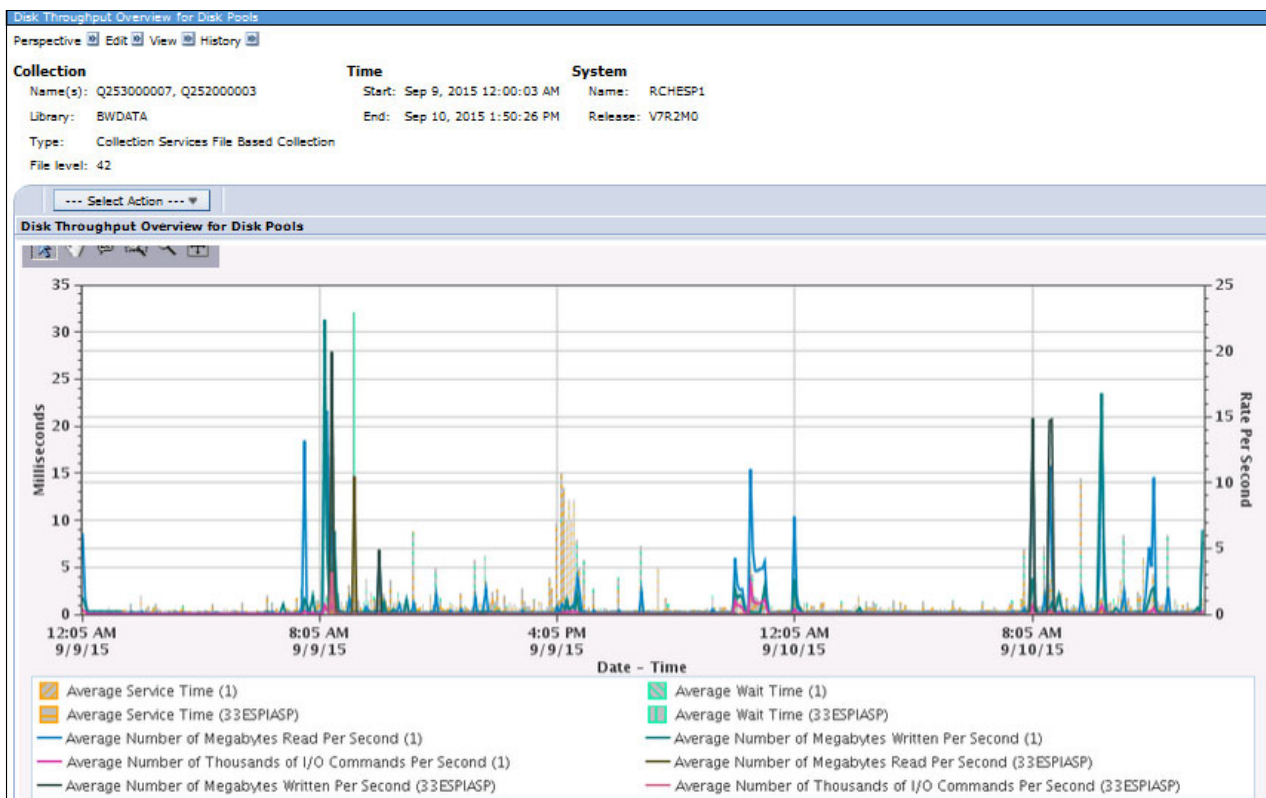
Figure 2-12   PDI-Disk Throughput Overview for Disk Pools graph

The number to focus on in the graph that is shown in Figure 2-12 is the Average number of Megabytes Written Per Second (for the IASP to be used in the PowerHA environment).

Note: If you did not migrate to an IASP yet, you can still use this tool to get a general idea of the number of megabytes that were written by looking at the Average Number of Megabytes Written Per Second for ASP 1.

Review this graph for the peaks in the write rates to the IASP. The Zoom Region tool in the toolbar can be used to view a specific range of time, which helps to make the lines on the graph more readable.

The write rates to the IASP during a certain time frame correspond to the amount of bandwidth that will be needed for geographic mirroring traffic to sufficiently mirror the writes to the preferred target system during that same time frame.

Throughout the days of data that needs to be analyzed, peaks and valleys occur in the write rates to the IASP. If sufficient bandwidth is not available to geographic mirroring during any of these times, performance issues can result on the preferred source system, an automatic suspension of geographic mirroring can occur, or both. The importance of the bandwidth analysis planning step cannot be stressed enough.

> **Note:** In addition to the normal runtime disk writes, the full or partial synchronization process consumes additional network bandwidth and system resources.

3

Implementation

This chapter describes the implementation of DS8000 Copy Services with IBM PowerHA SystemMirror for i.

The following scenarios are described in this chapter:

- ► 3.1, "Setting up DS8000 Copy Services" on page 44
- ► 3.2, "Configuring Metro Mirror" on page 45
- ► 3.3, "Configuring FlashCopy" on page 54
- ► 3.4, "Configuring Global Mirror" on page 57
- ► 3.5, "Configuring IBM i LUN-level switching" on page 67
- ► 3.6, "Full System scenarios" on page 72
- ► 3.7, "HyperSwap" on page 78
- ► 3.8, "IASP Copy Services Manager" on page 84
- ► 3.9, "Multiple target replication" on page 87

3.1 Setting up DS8000 Copy Services

IBM PowerHA SystemMirror for i controls IBM i clustering functions, such as switchover and failover. When IBM PowerHA SystemMirror for i is used with DS8000 Copy Services, it also needs to be able to control the Copy Services functions, such as Peer-to-Peer Remote Copy (PPRC) failover/failback or FlashCopy. This control requires the installation of the data storage command-line interface (DSCLI) on the IBM i partitions.

You can download the DSCLI installation package from the following URL. Use the version that matches your installed DS8000 microcode level:

`ftp://ftp.software.ibm.com/storage/ds8000/updates/DS8K_Customer_Download_Files/CLI`

The installation packages are provided as International Organization for Standardization (ISO) image files, which can be used by any virtual CD drive tool on a personal computer, or burned to a physical CD media.

For DSCLI installation procedures, refer to the following documentation:

► *IBM System Storage DS Command-Line Interface User's Guide for the DS6000 series and DS8000 series*, GC53-1127, which is included in the installation package.

► Section 8.3 in the IBM Redbooks publication *IBM i and IBM System Storage: A Guide to Implementing External Disks on IBM i*, SG24-7120.

You need to provide a DS storage user profile and its password for the IBM i partitions to manage DS8000 Copy Services. Whenever you change the password on the DS8000, you must change the configuration on the IBM i partition at the same time. For more information, see 9.5, "User management," in *IBM System Storage DS8000: Architecture and Implementation*, SG24-8886.

> **Note:** User profiles and passwords are case-sensitive on the DS8000.

The communication through DSCLI requires an Internet Protocol (IP) connection between the IBM i partition and the DS8000. This connection is initiated from the IBM i to the DS8000, which listens on Transmission Control Protocol (TCP) port 1750.

> **Important:** If a firewall exists between the IBM i partition and the DS8000, TCP port 1750 must be open.

3.2 Configuring Metro Mirror

This section describes the configuration of IBM PowerHA SystemMirror for i with DS8000 Metro Mirror. Figure 3-1 shows an example of this kind of environment.

The configuration is divided into two parts. The first part is performed on the DS8000, and the second part is performed on the IBM i.
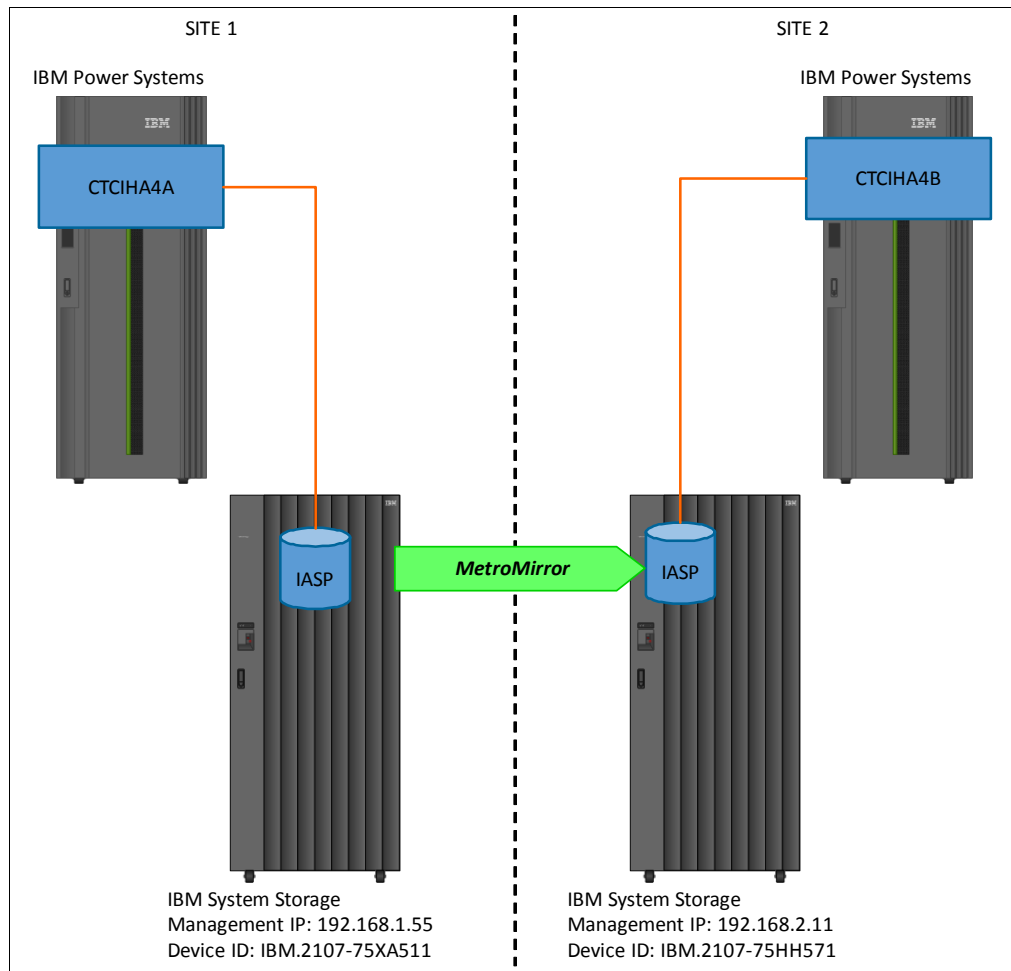
Figure 3-1   Metro Mirror environment

3.2.1 Configuring Metro Mirror: DS8000 steps

The following steps might be performed already by a storage specialist. These steps assume that you created the Metro Mirror/Global Mirror target logical unit numbers (LUNs) on the secondary DS8000 system and that storage area network (SAN) zoning is set up so that the two DS8000 systems can see each other.

When you configure the DS8000, you can choose between a graphical user interface (GUI) or the data storage command-line interface (DSCLI). If you are not familiar with the DSCLI commands, the GUI might be easier to use.

Therefore, the GUI path is documented in the example that is shown in the following steps:

1. Log in to the DS8000 GUI of your primary DS system.

2. If you did not set up any remote copy connections between the two DS8000 systems, you need to create a partnership between these two systems first. Select **Copy Services** and then select **Mirroring Connectivity** as shown in Figure 3-2.
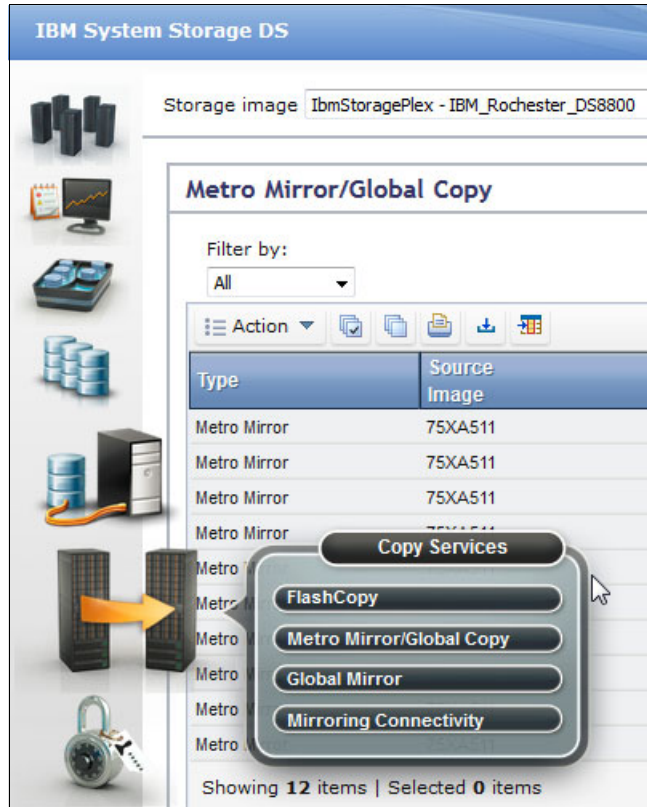
Figure 3-2   Starting Mirroring Connectivity

3. On the Create Mirroring Connectivity window, you must choose both the storage units and the logical subsystem (LSS) and I/O port. See Figure 3-3. After you select the correct options, click **Add** and then click **Create**.

> **Important:** You must also configure the reverse path or you will not be able to replicate back after a role swap.

Figure 3-3   Create Mirroring Connectivity window

4. Select **Copy Services** (Figure 3-2 on page 46) and then select **Metro Mirror/Global Copy**. From the Action menu, select **Create**.

5. On the Create Metro Mirror window, select your LSS and volumes from each DS storage that you want to be in a replication relationship. Click **Create**. See Figure 3-4.
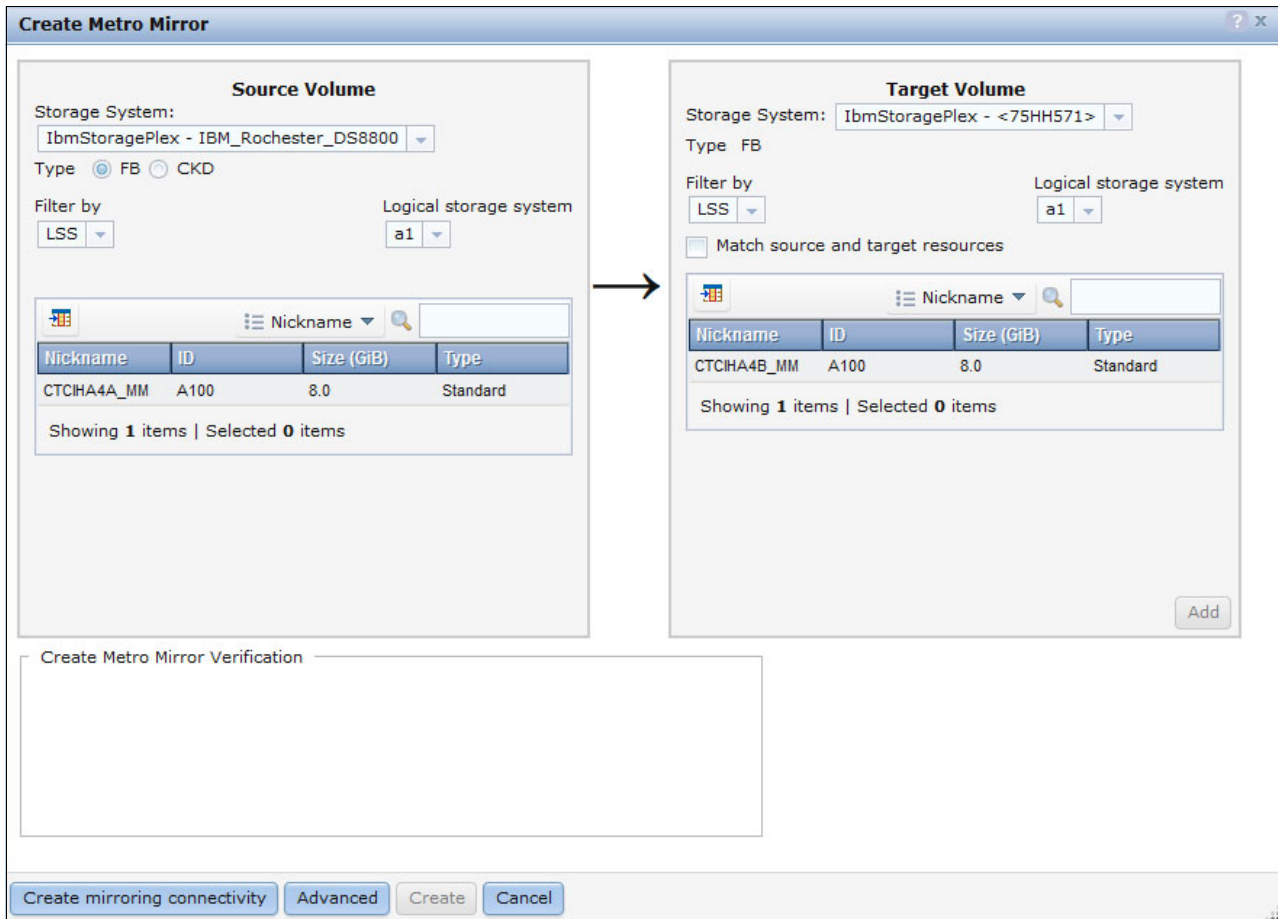
Figure 3-4   Create the Metro Mirror relationship

6. You now see your LUNs in a Metro Mirror relationship.

3.2.2 Configuring Metro Mirror: IBM i steps

The following steps for setting up the IBM PowerHA SystemMirror for i configuration for DS8000 Metro Mirror assume that you created in advance a two-node cluster environment, an administrative domain (if needed), and an independent auxiliary storage pool (IASP) on your production site. The cluster was started. If this situation is not the case, see *IBM PowerHA SystemMirror for i: Preparation (Volume 1 of 4)*, SG24-8400.

Follow these steps:

1. On the backup system, you need to create the device description for your IASP with the same IASP name that you use on your production site:

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

   If you chose a different relational database (RDB) name when you created the IASP on the production site, ensure that you specify the same value on the backup site. Otherwise, you can use the default value of *GEN. Alternatively, you can add the IASP to the cluster administrative domain.

2. Create the cluster resource group (CRG) by using the Create Cluster Resource Group (**CRTCRG**) command that is shown in Figure 3-5.

```
                    Create Cluster Resource Group (CRTCRG)

 Type choices, press Enter.

 Cluster  . . . . . . . . . . . . > CLUSTER        Name
 Cluster resource group . . . . . > PWRHAMM        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  . . . . . . . > CTCIHA4A      Name
    Node role  . . . . . . . . . . > *PRIMARY      *CRGTYPE, *PRIMARY...
    Backup sequence number . . . . > *LAST         1-127, *LAST
    Site name  . . . . . . . . . . > CTCIHA4A      Name, *NONE
    Data port IP address . . . . . > *NONE

                + for more values
    Node identifier  . . . . . . . > CTCIHA4B      Name
    Node role  . . . . . . . . . . > *BACKUP       *CRGTYPE, *PRIMARY...
    Backup sequence number . . . . > *LAST         1-127, *LAST
    Site name  . . . . . . . . . . > CTCIHA4B      Name, *NONE
    Data port IP address . . . . . > *NONE

                + for more values
 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 . . . . . > *NONE Name, *NONE
    Configuration object type  . .   *DEVD         *DEVD, *CTLD, *LIND, *NWSD
    Configuration object online  . > *OFFLINE      *OFFLINE, *ONLINE, *PRIMARY
    Server takeover IP address . .   *NONE

                + for more values
 Text description . . . . . . . .  *BLANK
 Failover message queue . . . . .   *NONE          Name, *NONE
   Library  . . . . . . . . . . .                  Name
```

Figure 3-5   Creating the CRG

The following descriptions relate to the `CRTCRG` command that is shown in Figure 3-5 on page 49:

– The cluster resource group has a type of *DEV because this CRG controls a switchable device (the IASP).

– You can specify a CRG exit program that is called whenever an event occurs in the cluster or the CRG status changes.

– The recovery domain node list defines the primary node in the cluster nodes and the order of the backup nodes in the cluster.

– Ensure that you specify different site names for the different nodes to represent the locations of the two replicating systems.

– The configuration object list defines the objects that are controlled by the CRG. In this example, this object is PWRHAMM, which has an object type of *DEVD. You can also specify whether the IASP must be varied on after a failover or a switchover and the server takeover IP address to activate.

– The failover message queue defines a message queue on all nodes for this specific CRG that receives a message when a failover is about to happen and that can be configured to wait for an answer before it proceeds with the failover. In this example, the failover message queue was already configured on the cluster level. Therefore, it is set to *NONE in the CRG.

3. Start the CRG by using the following command:

```
STRCRG CLUSTER(CLUSTER) CRG(PWRHAMM)
```

4. Add the ASP copy description by using the Add ASP Copy Description (**ADDASPCPYD**) command that is shown in Figure 3-6 and Figure 3-7 on page 52. You must perform this step for each copy of the IASP on both sites, so two ASP copy descriptions are required.

```
                  Add ASP Copy Description (ADDASPCPYD)

 Type choices, press Enter.

 ASP copy . . . . . . . . . . . . > METRO4A        Name
  ASP device . . . . . . . . . . . > PWRHAMM        Name
  Device domain  . . . . . . . . .   *              Name, *
  Cluster resource group . . . . . > PWRHAMM        Name, *NONE
  Cluster resource group site  . . > CTCIHA4A       Name, *NONE
  Storage host: Storage host:
     User name  . . . . . . . . . . > QLPAR
     Password . . . . . . . . . . .   xxxx
     Internet address . . . . . . . > '192.168.1.58'

 Location . . . . . . . . . . . .   *NONE          Name, *DEFAULT, *NONE
 Logical unit name:
    IBM System Storage device  . . > 'IBM.2107-75XA511'
    Logical unit range . . . . . . > 'a100-a106'   Character value
              + for more values
    Consistency group range  . . .                 Character value
              + for more values
  Recovery domain:
    Cluster node . . . . . . . . .   *NONE          Character value, *NONE
    Host identifier  . . . . . . .                  Character value
              + for more values
    Volume group . . . . . . . . .                  Character value
              + for more values
              + for more values

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

Figure 3-6   Adding the ASP copy description (part 1 of 2)

Figure 3-7 shows the second part of adding the ASP copy description.

```
                    Add ASP Copy Description (ADDASPCPYD)

 Type choices, press Enter.

 ASP copy . . . . . . . . . . . . > METRO4B        Name
 ASP device . . . . . . . . . . . > PWRHAMM        Name
 Device domain  . . . . . . . . .   *              Name, *
 Cluster resource group . . . . . > PWRHAMM        Name, *NONE
 Cluster resource group site  . . > CTCIHA4B       Name, *NONE
 Storage host:
   User name  . . . . . . . . . . > QLPAR
   Password . . . . . . . . . . .
   Internet address . . . . . . . >'192.168.2.11'

 Location . . . . . . . . . . . .   *NONE          Name, *DEFAULT, *NONE

 Logical unit name:
   IBM System Storage device  . . > 'IBM.2107-75HH571'
   Logical unit range . . . . . . > 'a100-a106'    Character value
             + for more values
   Consistency group range  . . .                  Character value
             + for more values
 Recovery domain:
   Cluster node . . . . . . . . .   *NONE          Character value, *NONE
   Host identifier  . . . . . . .                  Character value
             + for more values
   Volume group . . . . . . . . .                  Character value
             + for more values
             + for more values
```

Figure 3-7   Adding the ASP copy description (part 2 of 2)

The following descriptions relate to the **ADDASPCPYD** command that is shown in Figure 3-6 on page 51 and Figure 3-7:

– In this example, the unit range is only six LUNs. However, you can specify multiple LUN ranges, depending on your configuration requirements.

– The user profile (in this example QLPAR) that is used to communicate with the DS8000 depends on your configuration.

> **Important:** Enter the DS storage name correctly, for example, `'IBM.2107-75XA511'`. Errors are common when people enter the DS storage name.

5. Add the IASP device entry to the cluster resource group by using the Add CRG Device Entry (**ADDCRGDEVE**) command as shown in Figure 3-8.

```
                  Add CRG Device Entry (ADDCRGDEVE)

 Type choices, press Enter.

 Cluster  . . . . . . . . . . . . > CLUSTER        Name
 Cluster resource group . . . . . > PWRHAMM        Name
 Configuration object list:
   Configuration object . . . . .   pwrhamm        Name
   Configuration object type  . .   *DEVD          *DEVD, *CTLD, *LIND, *NWSD
   Configuration object online  .   *online        *OFFLINE, *ONLINE, *PRIMARY
   Server takeover 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 3-8   Adding the CRG device entry

6. Start the ASP session by using the Start ASP Session (**STRASPSNN**) command as shown in Figure 3-9.

```
                   Start ASP Session (STRASPSSN)

 Type choices, press Enter.

 Session  . . . . . . . . . . . . > METRO          Name
 Session type . . . . . . . . . . > *METROMIR      *GEOMIR, *METROMIR...
 ASP copy:
   Preferred source . . . . . . . > METRO4A        Name
   Preferred target . . . . . . . > METRO4B        Name
                 + for more values
 Device domain  . . . . . . . . .   *              Name, *
```

Figure 3-9   Starting the ASP session

Your Metro Mirror configuration is now complete, and your IASP is available for planned and unplanned site switches.

3.3 Configuring FlashCopy

This section describes the configuration of FlashCopy with the DS8000 series. Figure 3-10 shows the environment's layout.
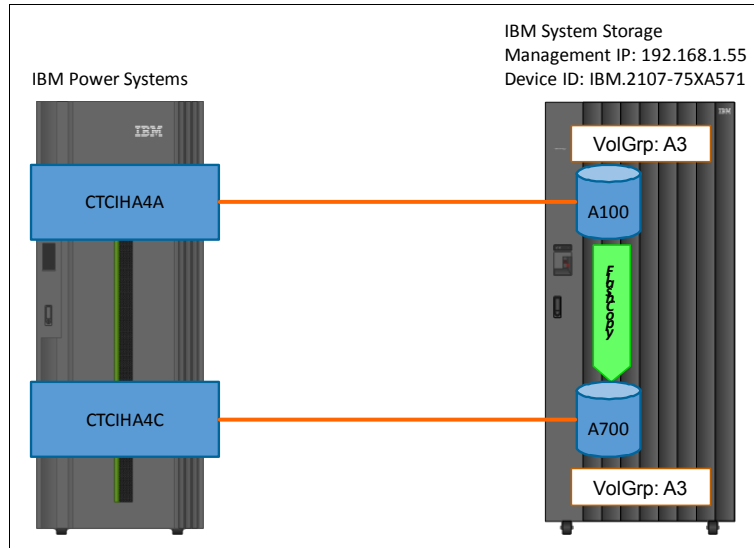
Figure 3-10   FlashCopy environment

With IBM PowerHA SystemMirror for i, you can create a FlashCopy of an IASP in a non-replicated IASP environment but also from a replicated IASP environment with Metro Mirror or Global Mirror. This capability allows a point-in-time copy to be taken of either the primary or the secondary site of a remote replication environment.

The steps that are described in the following setup procedure assume that you configured the FlashCopy target LUNs and specified a host connection for them in advance. The node that the FlashCopy of the IASP connects to must be added to the cluster and the cluster device domain. The IASP device must exist on the source system in advance.

The following steps describe the setup for an example scenario of configuring FlashCopy. The FlashCopy is performed from the primary site volumes on the preferred source system:

1. On the Flash partition, create a device description for the IASP with the same IASP name that you use on your production site. If you choose a different RDB name, ensure that you specify the same value on the backup site. Otherwise, you can use the default value of *GEN. Alternatively, you can use the cluster administrative domain to ensure that the ASP device description is on all of the nodes in the cluster. Use this command:

   ```
   CRTDEVASP DEVD(PWRHAMM) RSRCNAME(PWRHAMM)
   ```

2. An ASP copy description for the source volumes on the preferred source site might exist (if created for replication purposes) already. If this description exists, you do not need to create a second one for the source LUNs. You can use the existing ASP copy description for source volumes unless the existing replication is for geographic mirroring. If the existing replication is for geographic mirroring, you need to unconfigure the geographic mirroring and reconfigure with a DS8000 copy description for the source LUNs.

3. Create an ASP copy description for the FlashCopy target volumes as shown in Figure 3-11.

For the target copy description to use for FlashCopy, the cluster resource group and cluster resource group site must be *NONE. The node identifier must be set to the cluster node name that will own the target copy of the independent ASP.

```
                    Add ASP Copy Description (ADDASPCPYD)

 Type choices, press Enter.

 ASP copy . . . . . . . . . . . > FLASH          Name
 ASP device . . . . . . . . . . > PWRHAMM        Name
 Device domain  . . . . . . . . .   *            Name, *
 Cluster resource group . . . . .   *NONE        Name, *NONE
 Cluster resource group site  . .   *NONE        Name, *NONE
 Storage host:
   User name  . . . . . . . . . > admin
   Password . . . . . . . . . . .
   Internet address . . . . . . .>'192.168.1.58'

 Location . . . . . . . . . . . > CTCIHA4C       Name, *DEFAULT, *NONE

 Logical unit name:
   IBM System Storage device  . . > 'IBM.2107-75XA511'
   Logical unit range . . . . . . > 'a700-a706'  Character value
               + for more values
   Consistency group range  . . .              Character value
               + for more values
 Recovery domain:
   Cluster node . . . . . . . . .   *NONE        Character value, *NONE
   Host identifier  . . . . . . .              Character value
               + for more values
   Volume group . . . . . . . . .              Character value
               + for more values
               + for more values

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

Figure 3-11   Creating the ASP copy description for FlashCopy

4. Now, you can start a session for FlashCopy with the Start ASP Session (**STRASPSSN**) command as shown in Figure 3-12.

```
                        Start ASP Session (STRASPSSN)

 Type choices, press Enter.

 Session  . . . . . . . . . . . . > FLASH          Name
 Session type . . . . . . . . . . > *FLASHCOPY     *GEOMIR, *METROMIR...
 ASP copy:
   Preferred source . . . . . . . > METRO4A        Name
   Preferred target . . . . . . . > FLASH          Name
                  + for more values
 Device domain  . . . . . . . . .   *              Name, *
 FlashCopy type . . . . . . . . .   *NOCOPY        *COPY, *NOCOPY
 Persistent relationship  . . . .   *NO            *YES, *NO


                         Additional Parameters

 Switchover reverse replication     *YES           *YES, *NO
 Failover reverse replication . .   *NO            *YES, *NO



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

Figure 3-12   Starting an ASP FlashCopy session

Example 3-1 shows a CL script to run from the production node to automate a FlashCopy backup, including quiescing the IASP on the source node before it starts the FlashCopy session and varying on the IASP on the FlashCopy target node.

Also, you need to register a "Post Vary on exit" program for the IASP that performs the backup. When it is complete, it varies off and then ends the session.

Example 3-1   Running CHGASPACT from the FlashCopy target node to quiesce an IASP

```
PGM

CHGASPACT  ASPDEV(PWRHAMM) OPTION(*SUSPEND) SSPTIMO(30)

STRASPSSN  SSN(FLASH) TYPE(*FLASHCOPY) ASPCPY((METRO4A FLASH))

CHGASPACT  ASPDEV(PWRHAMM) OPTION(*RESUME)

RUNRMTCMD  CMD('VRYCFG CFGOBJ(PWRHAMM) CFGTYPE(*DEV) +
           STATUS(*ON)') RMTLOCNAME(CTCIHA4C *IP) +
           RMTUSER(POWERHA) RMTPWD(REDBOOK)

ENDPGM
```

Displaying an ASP session for FlashCopy

By using the Display ASP Session (`DSPASPSSN`) command, you can display the new ASP session for FlashCopy (Figure 3-13). For the FlashCopy target node, the ASP status shows AVAILABLE because the IASP is varied on on the target node.

```
                          Display ASP Session                        CTCIHA4C
                                                            09/17/15  17:15:20
 Session . . . . . . . . . . . . . . . . . :    FLASH
 Type  . . . . . . . . . . . . . . . . . . :    *FLASHCOPY

 Source node . . . . . . . . . . . . . . . :    CTCIHA4A
 Target node . . . . . . . . . . . . . . . :    CTCIHA4C
 Persistent  . . . . . . . . . . . . . . . :    *NO
                                                                      More...
                            Copy Descriptions

 ASP
 device          Name                  Role           State
 PWRHAMM         METRO4A               SOURCE         VARYOFF
                 FLASH                 TARGET         AVAILABLE


                                                                      Bottom
 Press Enter to continue

 F3=Exit    F5=Refresh    F11=View 2    F12=Cancel    F19=Automatic refresh
```

Figure 3-13   Displaying an ASP session for FlashCopy by using DSPASPSSN

3.4 Configuring Global Mirror

The configuration of Global Mirror is divided into two parts. First, you perform steps on the DS8000. Then, you perform the second PowerHA part on the IBM i.

This section describes the configuration of replication by using DS8000 series Global Mirror. Figure 3-14 shows the environment.
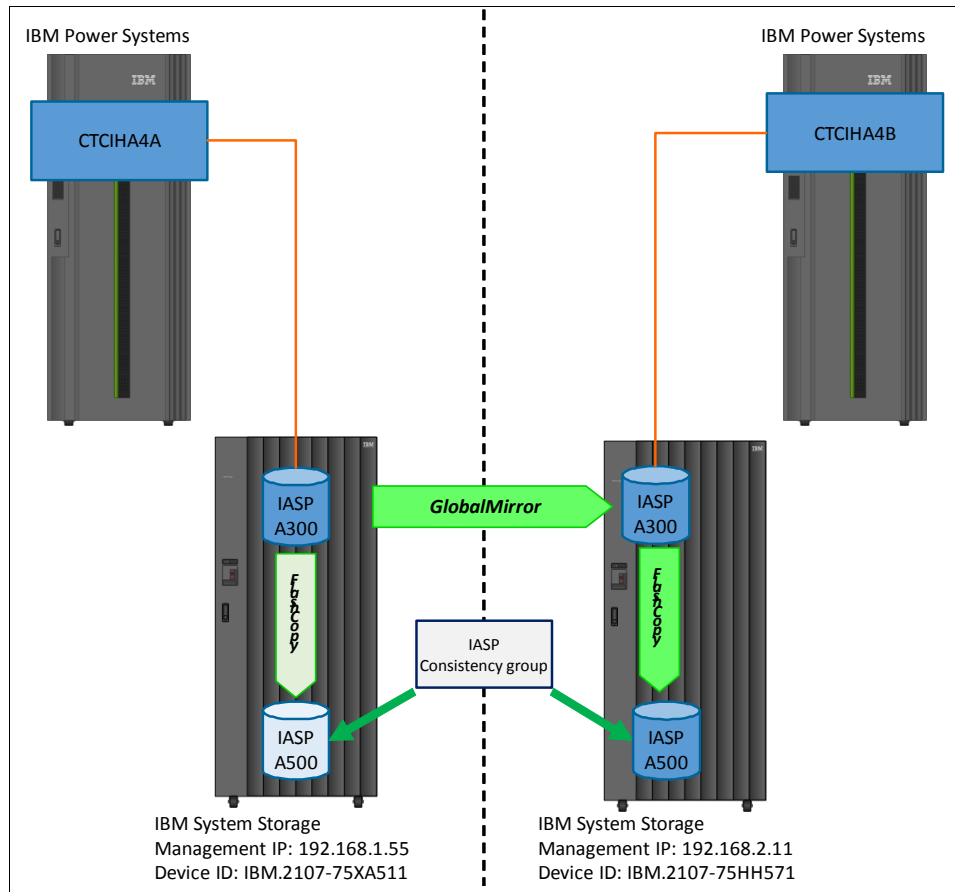
Figure 3-14   Global Mirror environment

3.4.1 Configuring Global Mirror: DS8000 steps

Perform the following steps to set up DS8000 Copy Services if DS8000 Copy Services was not set up by a storage specialist earlier. The steps assume that you created the Global Mirror target LUNs on the secondary DS8000 system and that SAN zoning is set up so that the two DS8000 systems can see each other. You can perform all of these steps by using DSCLI. However, if you are not familiar with DSCLI, you can use GUI.

Follow these steps:

1. Log in to the DS8000 GUI of your primary DS system.

2. If remote copy connections between the two DS8000 systems were not set up earlier, you need to create a partnership between these two systems. Select **Copy Services** and then select **Mirroring Connectivity**. See Figure 3-15.
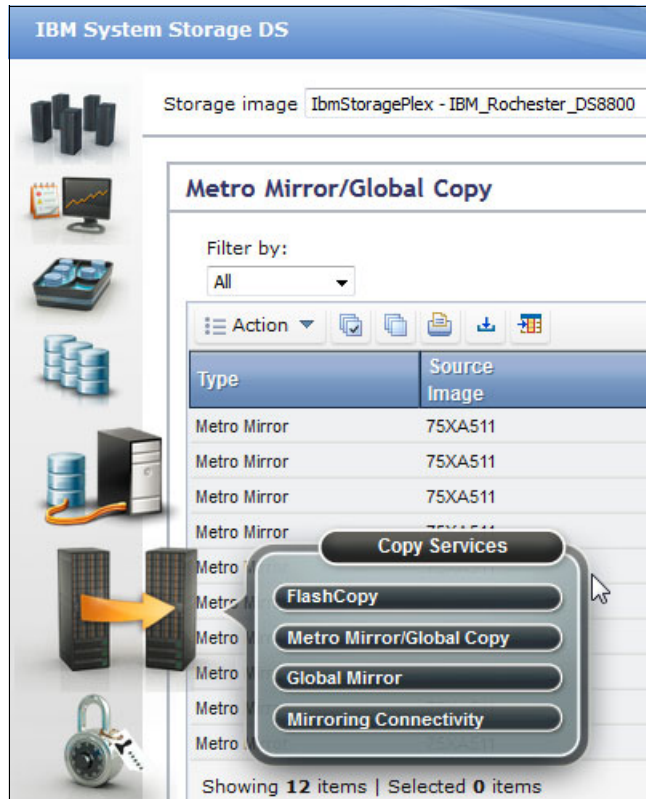
Figure 3-15   Starting Mirroring Connectivity

3. On the Create Mirroring Connectivity window, select an available LSS and I/O port. See Figure 3-16. Click **Add** and then click **Create**.
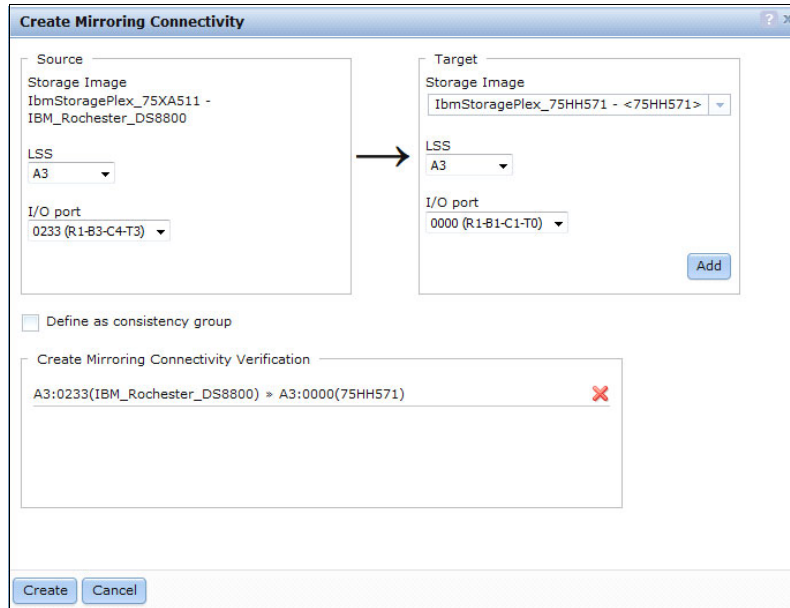
Figure 3-16   Creating connectivity for Global Mirror

4. Create the Global Copy relationship. See Figure 3-17. Select your LSS that was prepared for Global Copy, click **Add**, and click **Create**.
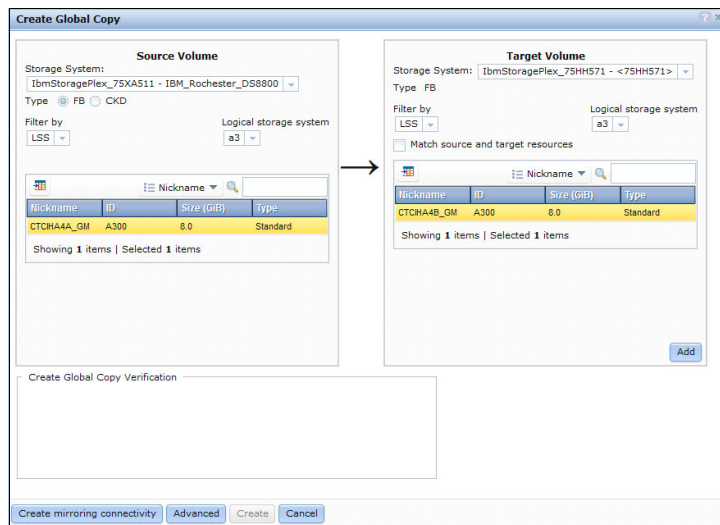
Figure 3-17   Creating the Global Copy relationship

5. After the Global Copy is established between the production DS8000 and the backup DS8000 and the copy catches up with production, you can establish the FlashCopy relationships (between B backup and C consistency group volumes) on the backup DS8000. Specify the following parameters for a FlashCopy relationship that is used in a Global Mirror session:

– Persistent (persist)
– Target write inhibit (tgtinhibit)
– Record changes (record)
– No full copy (nocp)

To specify these parameters, you must click **Advanced**. See Figure 3-18.
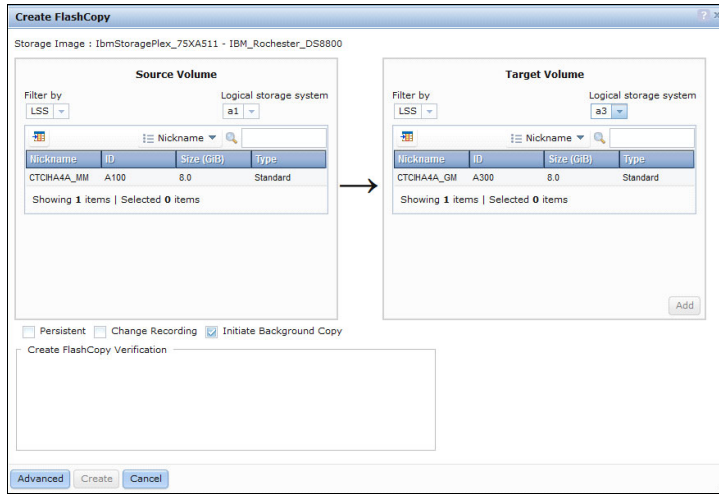
Figure 3-18   Creating FlashCopy for the Global Mirror consistency group

6. Create the Global Mirror session. From the **Copy Services** menu (Figure 3-15 on page 59), click **Global Mirror** and enter the LSS and session ID parameters. See Figure 3-19.
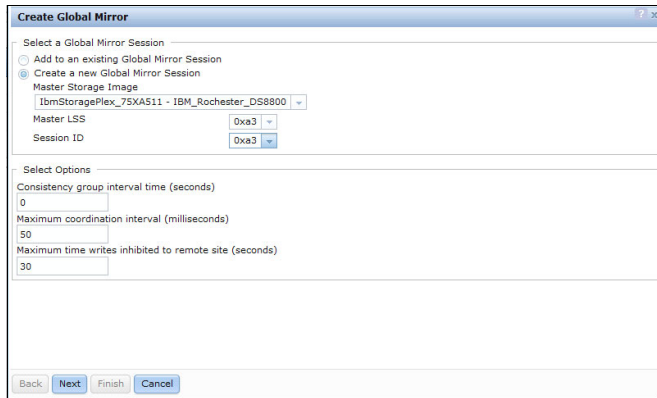
Figure 3-19   Creating a Global Mirror session

7. Add an existing Global Copy as shown in Figure 3-20.
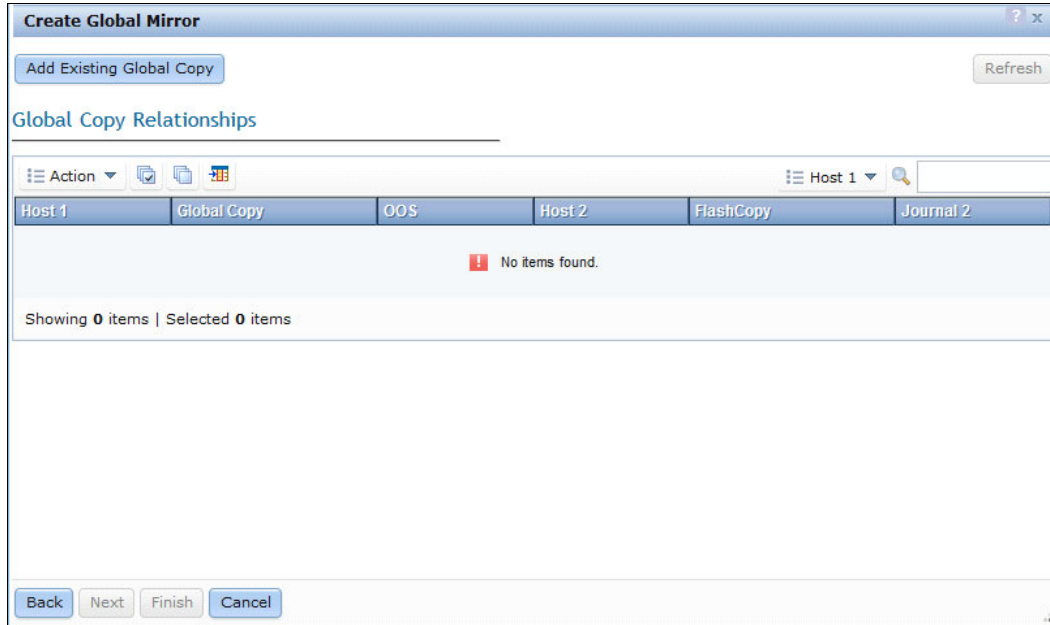
Figure 3-20   Adding an existing Global Copy

8. Your Global Mirror session is running as shown in Figure 3-21.
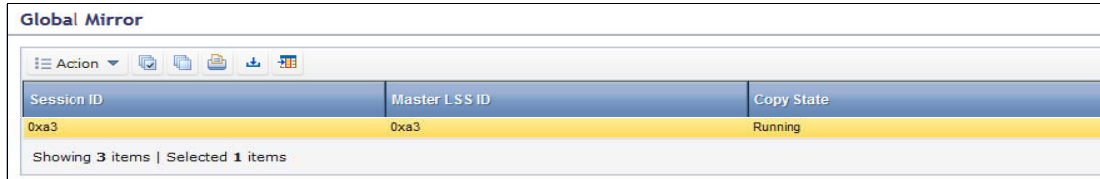
Figure 3-21   Global Mirror session that is running

3.4.2 Configuring Global Mirror: IBM i steps

The following steps for setting up the PowerHA configuration for DS8000 Global Mirror assume that you created a two-node cluster environment, an administrative domain (if needed), and an IASP on your production site in advance. The cluster is started. If this situation is not the case, see Chapter 4, "Creating the cluster framework," in the IBM Redbooks publication *IBM PowerHA SystemMirror for i: Preparation (Volume 1 of 4)*, SG24-8400.

Run the following steps:

1. On the backup system, create a device description for the IASP with the same IASP name that you use on your production site. If you choose a different RDB name than the IASP name on the production site, ensure that you specify the same value on the backup site. Otherwise, you can use the default value of *GEN:

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

2. Create the CRG by using the Create Cluster Resource Group (`CRTCRG`) command that is shown in Figure 3-22.

```
                      Create Cluster Resource Group (CRTCRG)

 Type choices, press Enter.

 Cluster  . . . . . . . . . . . . > CLUSTER        Name
 Cluster resource group . . . . . > PWRHAGM        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  . . . . . . . > CTCIHA4A       Name
    Node role  . . . . . . . . . . > *PRIMARY       *CRGTYPE, *PRIMARY...
    Backup sequence number . . . . > *LAST          1-127, *LAST
    Site name  . . . . . . . . . . > SITE1          Name, *NONE
    Data port IP address . . . . . > *NONE

                + for more values

    Node identifier  . . . . . . . > CTCIHA4B       Name
    Node role  . . . . . . . . . . > *BACKUP        *CRGTYPE, *PRIMARY...
    Backup sequence number . . . . > *LAST          1-127, *LAST
    Site name  . . . . . . . . . . > SITE2          Name, *NONE
    Data port IP address . . . . . > *NONE

                + for more values

                + for more values
  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 . . . . .   *NONE          Name, *NONE
    Configuration object type  . .   *DEVD          *DEVD, *CTLD, *LIND, *NWSD
    Configuration object online  .   *OFFLINE       *OFFLINE, *ONLINE, *PRIMARY
    Server takeover IP address . .   *NONE

                + for more values
 Text description . . . . . . . .   *BLANK

 Failover message queue . . . . .   *NONE          Name, *NONE
    Library  . . . . . . . . . . .                  Name

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

Figure 3-22   Creating the cluster resource group

The following descriptions describe the creation of the CRG by using the **CRTCRG** command that is shown in Figure 3-22 on page 63:

– The cluster resource group has a type of *DEV because this CRG controls a switchable device (the IASP).

– You can specify a CRG exit program that is called whenever an event occurs in the cluster or the CRG status changes.

– The recovery domain node list defines the primary node of the cluster nodes and the order of the backup nodes in the cluster.

– Ensure that you specify different site names for the different nodes to represent the locations of the two replicating systems.

– The configuration object list defines the objects that are controlled by the CRG. In this example, this name is PWRHAGM, which has an object type of *DEVD. Also, you can specify whether the IASP will be varied on after a failover or a switchover and whether the server takeover IP address will be activated.

– The failover message queue defines a message queue on all nodes for this specific CRG that receives a message when a failover is about to happen and that can be configured to wait for an answer before it proceeds with the failover. In this example, the failover message queue was already configured on the cluster level. Therefore, it is set to *NONE in the CRG, now.

3. Add the resource group device IASP (in this case, PWRHAGM) by using the Add CRG Device Entry (**ADDCRGDEVE**) command that is shown in Figure 3-23.

```
                   Add CRG Device Entry (ADDCRGDEVE)

 Type choices, press Enter.

 Cluster  . . . . . . . . . . .   CLUSTER        Name
 Cluster resource group . . . . .   PWRHAGM        Name
 Configuration object list:
   Configuration object . . . . .   PWRHAGM        Name
   Configuration object type  . .   *DEVD          *DEVD, *CTLD, *LIND, *NWSD
   Configuration object online  .   *ONLINE        *OFFLINE, *ONLINE,
 *PRIMARY
   Server takeover IP address . .   *SAME

               + for more values

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

Figure 3-23   Adding the CRG device entry

4. Start the CRG by using the following command:

STRCRG CLUSTER(CLUSTER) CRG(PWRHAGM)

5. Add the ASP copy descriptions by using the **ADDASPCPYD** command that is shown in Figure 3-24 and Figure 3-25 on page 66.

> **Important:** You must perform this task for the IASP configuration on the production site and also for the IASP configuration on the backup site, so you need to create two ASP copy descriptions.

The copy descriptions for Global Mirror include two LUN ranges, the actual LUNs that are used by the associated partitions, and also the LUNs that are used for the consistency group.

```
                    Add ASP Copy Description (ADDASPCPYD)

 Type choices, press Enter.

 ASP copy . . . . . . . . . . . . > GLOBAL4A      Name
 ASP device . . . . . . . . . . . > PWRHAGM       Name
 Device domain  . . . . . . . . .   *             Name, *
 Cluster resource group . . . . . > PWRHAGM       Name, *NONE
 Cluster resource group site  . .   *NONE         Name, *NONE
 Storage host:
   User name  . . . . . . . . . . > QLPAR
   Password . . . . . . . . . . . > XXX
   Internet address . . . . . . . > '192.168.1.58'

 Location . . . . . . . . . . . . > *DEFAULT      Name, *DEFAULT, *NONE
 Logical unit name:
   IBM System Storage device  . . > 'IBM.2107-75XA511'
   Logical unit range . . . . . . > 'A300-A306'   Character value
               + for more values
   Consistency group range  . . . > 'A500-A506'   Character value
               + for more values
 Recovery domain:
   Cluster node . . . . . . . . . > *NONE         Character value, *NONE
   Host identifier  . . . . . . .                 Character value
               + for more values
   Volume group . . . . . . . . .                 Character value
               + 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 3-24   Adding the ASP copy description (part 1 of 2)

Figure 3-25 shows the second part of adding the ASP copy description.

```
                      Add ASP Copy Description (ADDASPCPYD)

 Type choices, press Enter.

 ASP copy . . . . . . . . . . . . > GLOBAL4B      Name
 ASP device . . . . . . . . . . . > PWRHAGM       Name
 Device domain  . . . . . . . . .   *             Name, *
 Cluster resource group . . . . . > PWRHAGM       Name, *NONE
 Cluster resource group site  . .   *NONE         Name, *NONE
 Storage host:
   User name  . . . . . . . . . . > QLPAR
   Password . . . . . . . . . . . > XXX
   Internet address . . . . . . . > '192.168.2.11'

 Location . . . . . . . . . . . . > *DEFAULT      Name, *DEFAULT, *NONE
 Logical unit name:
   IBM System Storage device  . . > 'IBM.2107-75HH571'
   Logical unit range . . . . . . > 'A300-A306'   Character value
             + for more values
   Consistency group range  . . . > 'A500-A506'   Character value
             + for more values
 Recovery domain:
   Cluster node . . . . . . . . . > *NONE         Character value, *NONE
   Host identifier  . . . . . . .                 Character value
             + for more values
   Volume group . . . . . . . . .                 Character value
             + 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 3-25   Adding the ASP copy description (part 2 of 2)

The following descriptions refer to the **ADDASPCPYD** command that is shown in Figure 3-24 on page 65 and Figure 3-25:

– In this example, the unit range consists of only six LUNs. However, you can specify multiple LUN ranges, depending on your configuration requirements.

– The user profile (in this example, QLPAR) that is used to communicate with the DS8000 depends on your configuration.

– It is important to enter the correct DS storage name, for example, `'IBM.2107-75XA511'`. Errors are common when people enter this field.

6. Start the ASP session by using the Start ASP Session (**STRASPSNN**) command as shown in Figure 3-26.

```
                     Start ASP Session (STRASPSSN)

 Type choices, press Enter.

 Session  . . . . . . . . . . . . > GLOBAL        Name
 Session type . . . . . . . . . . > *GLOBALMIR    *GEOMIR, *METROMIR...
 ASP copy:
   Preferred source . . . . . . . > CTCIHA4A      Name
   Preferred target . . . . . . . > CTCIHA4B      Name
                 + for more values
 Device domain  . . . . . . . .    *             Name, *



 F3=Exit   F4=Prompt   F5=Refresh   F10=Additional parameters   F12=Cancel
  F13=How to use this display       F24=More keys
```

Figure 3-26   Starting the ASP session

Now, your IBM PowerHA SystemMirror for i DS8000 remote copy configuration is complete and your IASP is available for planned and unplanned site switches.

3.5 Configuring IBM i LUN-level switching

This section describes the configuration of a LUN-level switching environment. Figure 3-27 shows the environment.
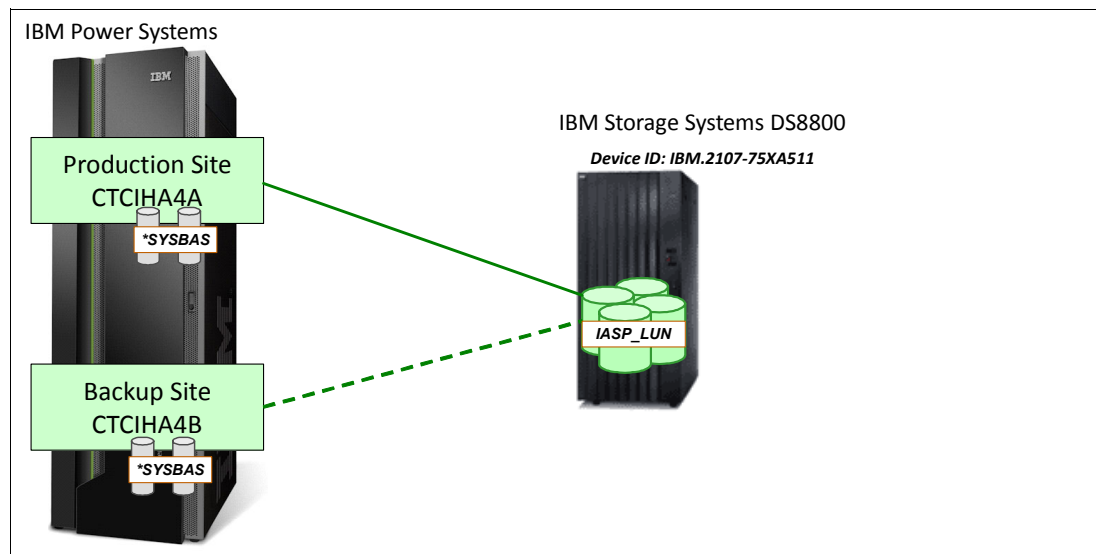
Figure 3-27   LUN-level switching environment

Important: At no time can the LUNs connect to both partitions. If the LUNs connect to both partitions by mistake, data loss is likely.

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

If you want to add LUN-level switching to an existing Metro or Global Mirror, you do not need to create the cluster, IASP, or change the cluster administrative domain. You must create the IASP device description on the backup system from the LUN switching perspective. Before you start this process, you need to know the volume groups that are involved in the LUN switch and the host connect IDs for each of the partitions in the LUN switching arrangement that they relate to.

> **Important:** Configure the volume groups to the source system host connections only.

The following steps show you how to implement a LUN switching environment:

1. On the new target partition, create a device description for the IASP with the same IASP name that you use on your production site. If you choose a different RDB name, ensure that you specify the same value on the backup site. Otherwise, you can use the default value of *GEN. Alternatively, you can use the cluster administrative domain to ensure that the ASP device description is on all of the nodes in the cluster. Use this command:

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

2. If you are adding to an existing replication, use the Work with Auxiliary Storage Pool Copy Descriptions (**WRKASPCPYD**) command, option 24 to end any sessions that exist that use the copy descriptions that are involved.

3. If you are adding to an existing replication, use the Work with Cluster (**WRKCLU**) command and specify option 9 to work with cluster resource groups (CRGs), and option 9 to end your CRG.

4. If you are adding to an existing replication, delete the existing CRG.

5. Create a CRG by using the Create Cluster Resource Group (**CRTCRG**) command as shown in Figure 3-28 on page 69.

```
                    Create Cluster Resource Group (CRTCRG)

 Type choices, press Enter.

 Cluster  . . . . . . . . . . . . > CLUSTER        Name
 Cluster resource group . . . . . > PWRHALUN       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  . . . . . . . > CTCIHA4A       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  . . . . . . . > CTCIHA4B       Name
   Node role  . . . . . . . . . . > *BACKUP        *CRGTYPE, *PRIMARY...
   Backup sequence number . . . . > *LAST          1-127, *LAST
   Site name  . . . . . . . . . . > SITE2          Name, *NONE
   Data port IP address . . . . . > *NONE

   Node identifier  . . . . . . . > CTCIHA4C        Name
   Node role  . . . . . . . . . . > *BACKUP        *CRGTYPE, *PRIMARY...
   Backup sequence number . . . . > *LAST          1-127, *LAST
   Site name  . . . . . . . . . . > SITE1          Name, *NONE
   Data port IP address . . . . . > *NONE

                + for more values
 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 . . . . .   *NONE          Name, *NONE
   Configuration object type  . .   *DEVD          *DEVD, *CTLD, *LIND, *NWSD
   Configuration object online  .   *OFFLINE       *OFFLINE, *ONLINE, *PRIMARY
   Server takeover IP address . .   *NONE

                + for more values
 Text description . . . . . . . .   *BLANK

 Failover message queue . . . . .   *NONE          Name, *NONE
   Library  . . . . . . . . . . .                  Name

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

Figure 3-28   Creating an ASP copy description for a LUN-level switch

6. Either change the existing ASP copy description or create a new one. Add the primary and backup host connection IDs and volume group IDs for your additional recovery domain node. (Obtain this information from DS or your storage specialist.) See Figure 3-29.

```
                     Change ASP Copy Description (CHGASPCPYD)

 Type choices, press Enter.

 ASP copy . . . . . . . . . . . . > METRO4A       Name
 Device domain  . . . . . . . . .   *             Name, *
 ASP device . . . . . . . . . . .   PWRHAMM       Name, *SAME
 Cluster resource group . . . . .   PWRHAMM       Name, *SAME, *NONE
 Cluster resource group site  . .   CTCIHA4A      Name, *SAME, *NONE
 Storage host:
   User name  . . . . . . . . . .   QLPAR
   Password . . . . . . . . . . .   *SAME
   Internet address . . . . . . .   '192.168.1.58'



 Location . . . . . . . . . . . .   *DEFAULT      Name, *SAME, *DEFAULT, *NONE
 Logical unit name:
   IBM System Storage device  . .   'IBM.2107-75XA511'
   Logical unit range . . . . . .   'A106-A106'   Character value, *SAME
             + for more values
   Consistency group range  . . .                 Character value, *SAME
             + for more values

 Recovery domain:
   Cluster node . . . . . . . . . > CTCIHA4A      Character value, *SAME, *NONE
   Host identifier  . . . . . . . > 000C          Character value
             + for more values > 000F
   Volume group . . . . . . . . . > V12           Character value
             + for more values

   Cluster node . . . . . . . . .   CTCIHA4C      Character value
   Host identifier  . . . . . . .   0015          Character value
             + for more values   0016
   Volume group . . . . . . . . .   V12           Character value
             + for more values
```

Figure 3-29   Changing the ASP copy description

7. Add the device IASP entry into the cluster resource group by using the Add CRG Device Entry (**ADDCRGDEVE**) command. See Figure 3-30.

```
                    Add CRG Device Entry (ADDCRGDEVE)

 Type choices, press Enter.

 Cluster  . . . . . . . . . . . . > CLUSTER      Name
 Cluster resource group . . . . . > PWRHAMM      Name
 Configuration object list:
   Configuration object . . . . .   pwrhamm      Name
   Configuration object type  . .   *DEVD        *DEVD, *CTLD, *LIND, *NWSD
   Configuration object online  .   *online      *OFFLINE, *ONLINE, *PRIMARY
   Server takeover 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 3-30   Adding the CRG device entry

8. Start the ASP session by using the Start ASP Session (**STRASPSSN**) command (Figure 3-31).

```
                      Start ASP Session (STRASPSSN)

 Type choices, press Enter.

 Session  . . . . . . . . . . . . > METRO      Name
 Session type . . . . . . . . . . > *METROMIR    *GEOMIR, *METROMIR...
 ASP copy:
   Preferred source . . . . . . . > CTCIHA4A      Name
   Preferred target . . . . . . . > CTCIHA4B      Name
              + for more values
 Device domain  . . . . . . . . .   *            Name, *




                                                                    Bottom
 F3=Exit    F4=Prompt    F5=Refresh    F10=Additional parameters    F12=Cancel
 F13=How to use this display        F24=More keys
```

Figure 3-31   Starting the ASP session

Your LUN-level switching environment is complete, and you can switch your LUNs between nodes.

Note: If you use a three-site solution, you need to manage the recovery domain to control where PowerHA switches to or fails over to.

3.6 Full System scenarios

This section describes Full System scenarios that are an extension of the PowerHA with IASP technology options. The following topics are covered in this section:

- ► 3.6.1, "Full System FlashCopy by using Full System Copy Services Manager" on page 72
- ► 3.6.2, "PowerHA Full System replication for DS8000" on page 76

3.6.1 Full System FlashCopy by using Full System Copy Services Manager

STG Lab Services offer a Full System Copy Services Manager (FSCSM) that is installed on a production partition that is cloned through FlashCopy, a managing partition that controls the entire FlashCopy process through IP communication to the production partition, and a FlashCopy backup partition that includes Secure Shell (SSH) communication through the Hardware Management Console (HMC). See Figure 3-32.
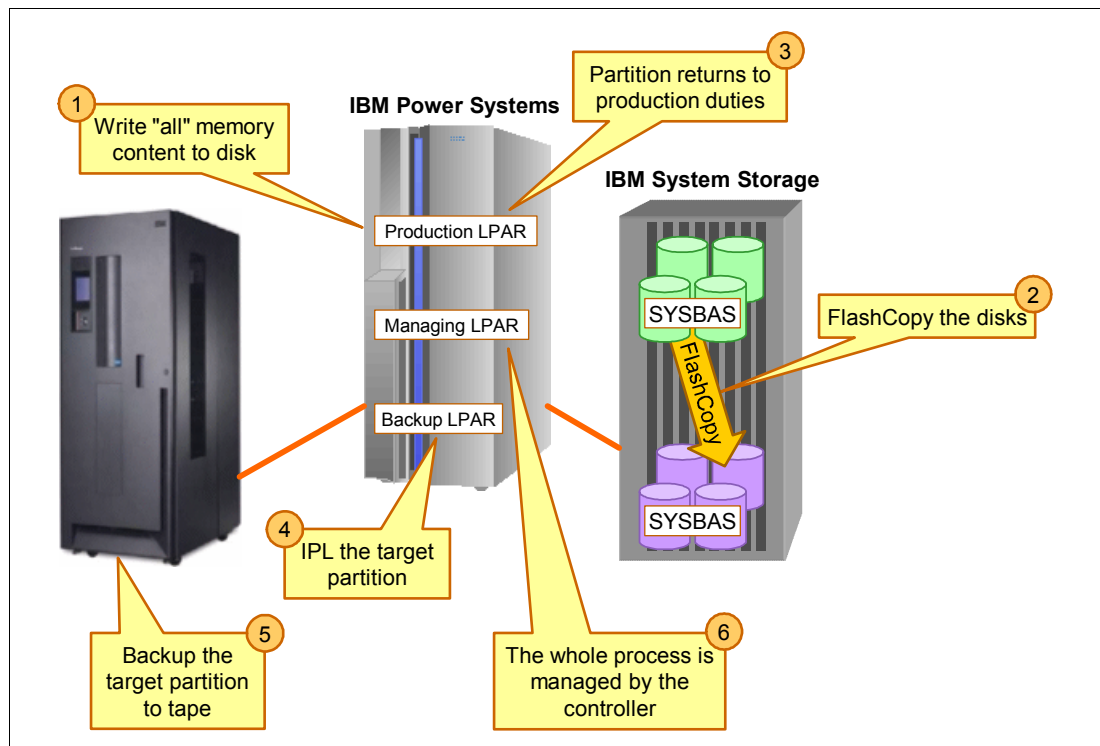
Figure 3-32   Full System Copy Services Manager

FSCSM is licensed separately from Advanced Copy Services for PowerHA, but it uses the same copy services environment menus to set up the profiles and scripts to communicate with the DS8000 storage system. FSCSM requires DSCLI. You need to install FSCSM on the managing and production partitions in the QZRDIASH5 library. FSCSM communicates with the HMC to IPL and monitor partitions. FSCSM can change host connections on the DS8000 system so that one backup LPAR is used for several production LPARs, if the saves of all LPARs finish before a new cycle of FlashCopy operations is started.

FSCSM supports space efficient FlashCopy (FlashCopy SE) and incremental FlashCopy. A FlashCopy can be taken at the source side of a Metro Mirror, Global Mirror, or Global Mirror setup. Also, a FlashCopy can be taken at the target site of a Metro Mirror or Global Mirror setup.

With FSCSM, you control the replication by using the suspend replication, resume replication, and change replication direction functions. FSCSM does not provide automation to check the status of the production partition, power off the production partition, or IPL the replicated partition in the target site. FSCSM is not an HA solution. It is a DR solution only.

For ease of use, you use one command to initiate and monitor the process, from quiesce to backups to IBM i Backup, Recovery, and Media Services (BRMS) information transfer. Administrative and operator tasks are separate for ease of use and security.

FSCSM prevents the production IP addresses from coming online at the backup LPAR. With FSCSM, you can create a full-system backup (opt 21 (Entire System) on the Backup menu).

In addition to controlling the IBM i Full System FlashCopy process, FSCSM provides high-level integration with BRMS to ensure that the BRMS backup that runs from the FlashCopy backup partition appears in the BRMS history and contains all backup and media information. This function is similar to the way that the backup from the production partition with the BRMS history (QUSRBRM) is transferred automatically to the production partition after the backup completes. See Figure 3-33.

Production System Name: PROD

1. Stop all BRMS activities
2. Prepare BRMS for FlashCopy
3. Quiesce system or PWRDWNSYS to ensure all main storage content written to disk
4. FlashCopy
5. Resume or Restart System

Backup System Name: PROD

6. Change system attributes and Network configuration during IPL startup program
7. Set BRMS system state to FlashCopy mode
8. Backup System treated as "PROD" in BRMS Network
9. Complete saves using BRMS
10. Set BRMS state to FlashCopy backup complete mode
11. Save QUSRBRM and transfer the library to PROD system
12. Call post FlashCopy BRMS function /restore QUSRBRM
13. Resume BRMS activities on PROD

Figure 3-33   Full System FlashCopy toolkit integration with BRMS

The FSCSM Full System FlashCopy functions are configured on the managing partition by using the Create Full System Flash Copy (`CRTSYSCPY`) command. See Figure 3-34 on page 74. This example shows a FlashCopy configuration where CTCIHA7A is the production partition to be flashed through the managing partition to the target partition CTCIHA7C. It quiesces the production partition (OPTYPE *QUIESCE) before it takes the FlashCopy, resumes the production partition, and activates the backup partition by starting its specified IP interface and BRMS backup job.

```
                   Create Full Sys Flash Copy (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 3-34   Full System FlashCopy toolkit CRTSYSCPY command panel (page 1 of 2)

Figure 3-35 shows the remaining pages from the **CRTSYSCPY** command panel.

```
Minutes to wait for power down    60            minutes
Force *SYSBAS resume . . . . . .  60            seconds
Force flash copy . . . . . . . .  *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-EC7560A    *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 3-35   Full System FlashCopy toolkit CRTSYSCPY command panel (page 2 of 2)

After the copy services environment is created by using **WRKCSE** and the configuration by using **CRTSYSCPY**, a Full System FlashCopy can be started by using **MKSYSCPY** (Figure 3-36).

```
                  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 3-36   Full System FlashCopy toolkit MKSYSCPY command panel

For more information about the Full System FlashCopy FSCSM service offering, contact STG Lab Services on the following website:

https://ibm.biz/Bd4Lbs

3.6.2 PowerHA Full System replication for DS8000

PowerHA Full System Replication for DS8000 (PowerHA FSR) is the latest toolkit offering that was developed by IBM Lab Services Power Systems Delivery Practice. PowerHA FSR is a solution to provide a full current copy of a production system's disks at another "site".

This solution is not considered an HA solution. This solution is a DR solution only. PowerHA FSR has a few limitations when you compare it to the use of PowerHA with an IASP implementation:

► The DR system must be powered down during normal operation.

► Only one copy exists of the operating system, so any changes to the operating system that require an IPL require a production outage versus PowerHA where you can switch to the other system.

► In a system failure that requires a main storage dump, you cannot switch to the DR site before you finish the dump (unless you interrupt the dump and then you cannot analyze why the problem happened). You can switch to the DR site before you finish the dump with PowerHA because you can fail over to the backup system and perform analysis on the failed system because that system has its own copy of the operating system.

PowerHA FSR automates the correction of resources when you attach a full-system copy to a different hardware system. Likely, the resource names for communication lines, tapes, and so on differ on the new system when they are compared to the old system. Therefore, the corresponding line descriptions, device descriptions, and so on need to be changed before those resources can be used. PowerHA FSR performs the following tasks:

► 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 the Transmission Control Protocol (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 Global Mirror with Change Volumes.

As shown in Figure 3-37, a preferred setup for the use of PowerHA FSR is an environment with two controlling LPARs: one LPAR on the production site and one LPAR on the DR site.

If only one controlling LPAR is used, the controlling LPAR must be located in the DR site. The controlling LPARs do not need to be dedicated to controlling the environment. They can be used for other work, such as development or testing, in parallel. They need to be able to talk to both HMCs, both storage systems, the production node, and the DR node. All communications between the LPARs, the HMCs, and the storage systems are encrypted.
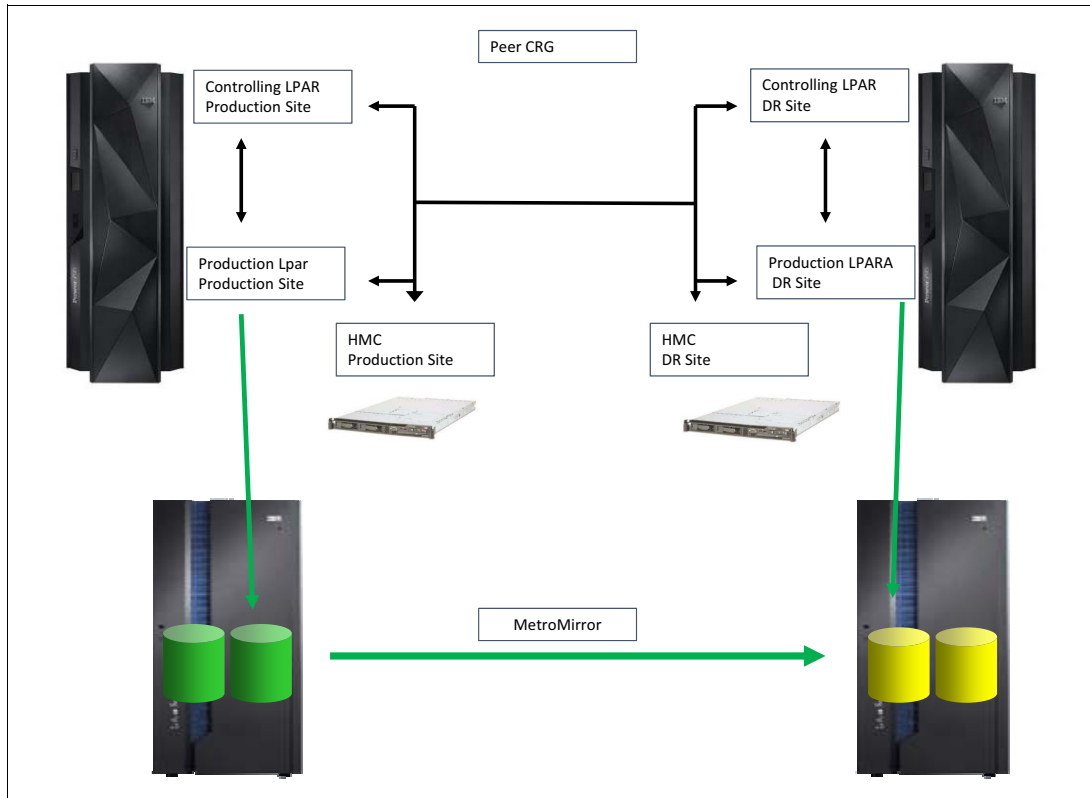
Figure 3-37   PowerHA FSR for DS8000

> **Important:** IBM PowerHA SystemMirror for i Enterprise Edition needs to be installed and licensed on the controlling LPARs.

PowerHA FSR requires that the DS8000 storage system is either natively attached (without the use of Virtual I/O Server (VIOS)) or through VIOS with N_Port ID Virtualization (NPIV). Virtual Small Computer System Interface (VSCSI) attachment through VIOS is not supported.

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. The switch must be triggered manually by using the `SWCSE` command.

For more information about the PowerHA Full System Copy Services Manager Replication offering, contact STG Lab Services on the following website:

https://ibm.biz/Bd4Lbs

3.7 HyperSwap

You can migrate from one DS8000 storage system to another DS8000 storage system with IBM HyperSwap®.

HyperSwap is an HA solution for DS8000 storage. Metro Mirror copies of the same disk from multiple DS8000s report into IBM i simultaneously. IBM i recognizes the disk units as HyperSwap devices by querying the Peer-to-Peer Remote Copy (PPRC) relationships in the DS.

HyperSwap support is available for both full system environments as well as IASP environments.

The following topics are covered in this section:

► 3.7.1, "Full System HyperSwap" on page 78
► 3.7.2, "HyperSwap and Live Partition Mobility" on page 81
► 3.7.3, "HyperSwap and IASP" on page 82

3.7.1 Full System HyperSwap

You use the IBM i commands to display and manage Full System HyperSwap. In the case of a fault on the source DS, an automatic failover to the target DS is triggered. Replication is not restarted automatically on a failover, but replication can be restarted automatically on a planned swap. See Figure 3-38.
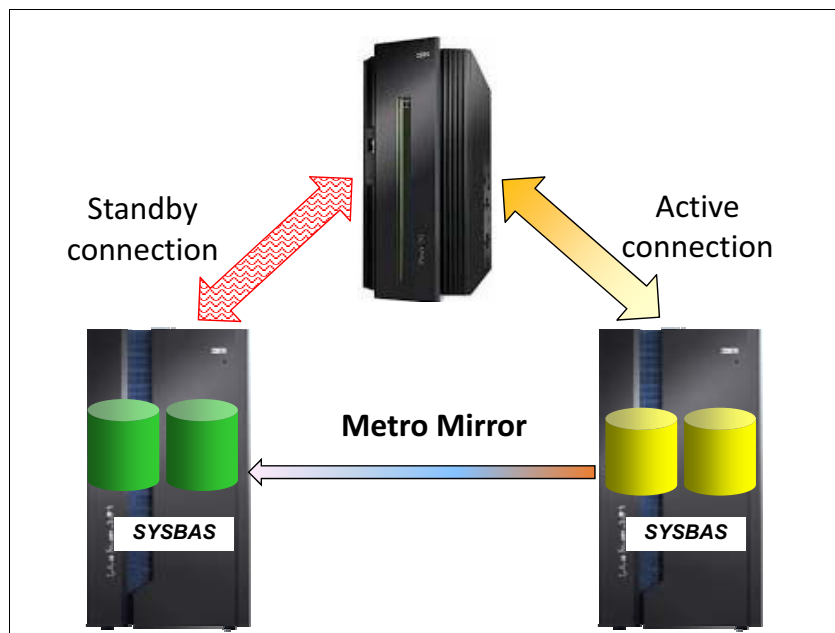
Figure 3-38   Full System HyperSwap

The following steps are the configuration steps for Full System HyperSwap. We assume that the system is already running on current disk.

1. Zone the second DS8000 to partition (or it can be direct connect if enough partitions have enough adapters).

   Important: You must not present the secondary LUNs to the system at this time.

2. Logically configure the secondary disk in the second DS8000. Do not report the volumes in to the system (that is, do not attach volgrps to hostconnects).

3. Configure the PPRC paths between the DSs.

4. Make the PPRC relationships on all of the LUNs.

5. Present the secondary volumes into the partition. PPRC does not need to have completed the synchronization.

6. PowerHA automatically recognizes the secondary disk and detects the PPRC relationship.

7. As soon as the relationship is established, you can use the Display HyperSwap Status (`DSPHYSSTS`) command to view the status. See Figure 3-39.

```
                         Display HyperSwap Status                    CTCHYPER
                                                           08/27/14  12:36:38
                Volume                   Copy              Volume
   Resource     Identifier               Status  Resource  Identifier
   DMP045       IBM.2107-75ZF751/8800    >>>>    DMP007    IBM.2107-75XA511/8800
   DMP050       IBM.2107-75ZF751/8801    >>>>    DMP010    IBM.2107-75XA511/8801
   DMP051       IBM.2107-75ZF751/8802    >>>>    DMP011    IBM.2107-75XA511/8802
   DMP052       IBM.2107-75ZF751/8900    >>>>    DMP013    IBM.2107-75XA511/8900
   DMP054       IBM.2107-75ZF751/8901    >>>>    DMP015    IBM.2107-75XA511/8901
   DMP055       IBM.2107-75ZF751/8902    >>>>    DMP017    IBM.2107-75XA511/8902
```

Figure 3-39   Displaying the HyperSwap status

Note: Ensure that you count the number of relationships on the panel that is shown in Figure 3-39 and ensure that the number on the panel matches the expected number. HyperSwap ignores volumes where PPRC is not turned on.

8. You can use press F11 for more information. See Figure 3-40. The synchronization progress goes blank when it is complete.

```
                            Display HyperSwap Status
CTCHYPER

08/27/14   12:36:38
             Copy                   Synchronization
  Resource   Status   Resource         Progress
  DMP087     >>>>     DMP007            91%
  DMP088     >>>>     DMP010            67%
  DMP089     >>>>     DMP011            82%
  DMP090     >>>>     DMP013            98%
  DMP091     >>>>     DMP077            86%
  DMP092     >>>>     DMP079            80%
```

Figure 3-40   Synchronization progress

9. The Copy Status column shows a different status as shown in Figure 3-41.

```
>>>>     Volume in currently syncing from left DS to right DS
<<<<     Volume in currently syncing from right DS to left DS
--->     Volume is in sync from left DS to right DS
<---     Volume is in sync from right DS to left DS
-\->     Volume is suspended from left DS to right DS
<-\-     Volume is suspended from left DS to right DS
XXXX     Volume is suspended.  Primary undetermined by
            DSPHYSSTS.. ....   but direction may be known in
SLIC                    if no IPL has been performed.
```

Figure 3-41   HyperSwap status value

10. The `DSPHYSSTS` command shows you the status after all of the volumes are synchronized. See Figure 3-42.

```
                  Display HyperSwap Status                    CTCHYPER
                                                   09/03/14
08:12:08
          Volume              Copy              Volume
Resource  Identifier          Status  Resource  Identifier
DMP087    IBM.2107-75ZF751/8800  --->   DMP007    IBM.2107-75XA511/8800
DMP088    IBM.2107-75ZF751/8801  --->   DMP010    IBM.2107-75XA511/8801
DMP089    IBM.2107-75ZF751/8802  --->   DMP011    IBM.2107-75XA511/8802
DMP090    IBM.2107-75ZF751/8900  --->   DMP013    IBM.2107-75XA511/8900
DMP091    IBM.2107-75ZF751/8901  --->   DMP015    IBM.2107-75XA511/8901
DMP092    IBM.2107-75ZF751/8902  --->   DMP017    IBM.2107-75XA511/8902
```

Figure 3-42   Ready to swap status

11. The QSYSOPR message queue also reports the status change to fully synchronized, as shown in Figure 3-43.

```
                     Additional Message Information
Message ID . . . . . . . :   CPI09B2                               Date sent
. . . . . . :   09/02/14       Time sent  . . . . . . :    15:02:04
Message . . . . :   HyperSwap replication is fully synchronized.
Cause . . . . . :   All of the data on the disk units in the configuration has   been
fully synchronized. No data has been lost. The partition continues        normal
processing.
```

Figure 3-43   QSYSOPR message

Now, you are ready for HyperSwap.

3.7.2 HyperSwap and Live Partition Mobility

When you want to use HyperSwap in a Live Partition Mobility (LPM) environment as shown in Figure 3-44, you must perform a DS assignment for HyperSwap.
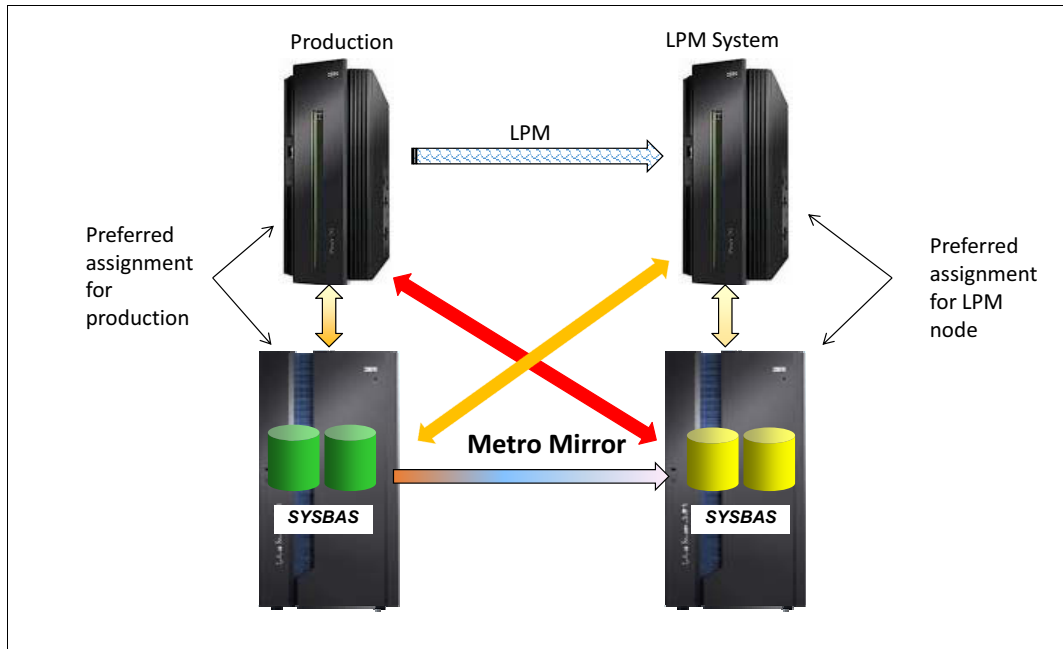
Figure 3-44   HyperSwap in an LPM environment

To perform a DS assignment for HyperSwap, you use the Add HyperSwap Storage Description (**ADDHYSSTGD**) command to describe a preferred system to DS8000 assignment for HyperSwap. See Figure 3-45. By performing an LPM migration, you automatically switch the HyperSwap to the preferred DS8000.

```
             Add HyperSwap Storage Desc (ADDHYSSTGD)

Type choices, press Enter.

System Serial Number . . . . . .     1016B3P
Character value      IBM System Storage device  . . .    IBM.2107-75ZF751
```

Figure 3-45   ADDHYSSTGD command

Add a storage description for both the DS8000 and systems in the LPM to ensure the correct assignment on the switch. If the preferred storage is not available, the switch still completes with the current assignment. LPM can complete if one DS is offline or disconnected.

> **Note:** LPM occurs first. Then, the HyperSwap occurs. If any configuration issue exists from the LPM target disk and the preferred DS8000, HyperSwap detects this issue after the LPM switch and logs a HyperSwap error in the QSYSOPR message queue.

> **Important:** Full System HyperSwap requires PowerHA Express Edition (5770-HAS option 3).

3.7.3 HyperSwap and IASP

HyperSwap support is also available for an IASP environment as shown in Figure 3-46, which was made available in the latest technology refresh for Version 7 Release 2 of PowerHA. As a part of this update, modifications were made to several commands that are used for HyperSwap so these commands will look different than the commands that are shown above.
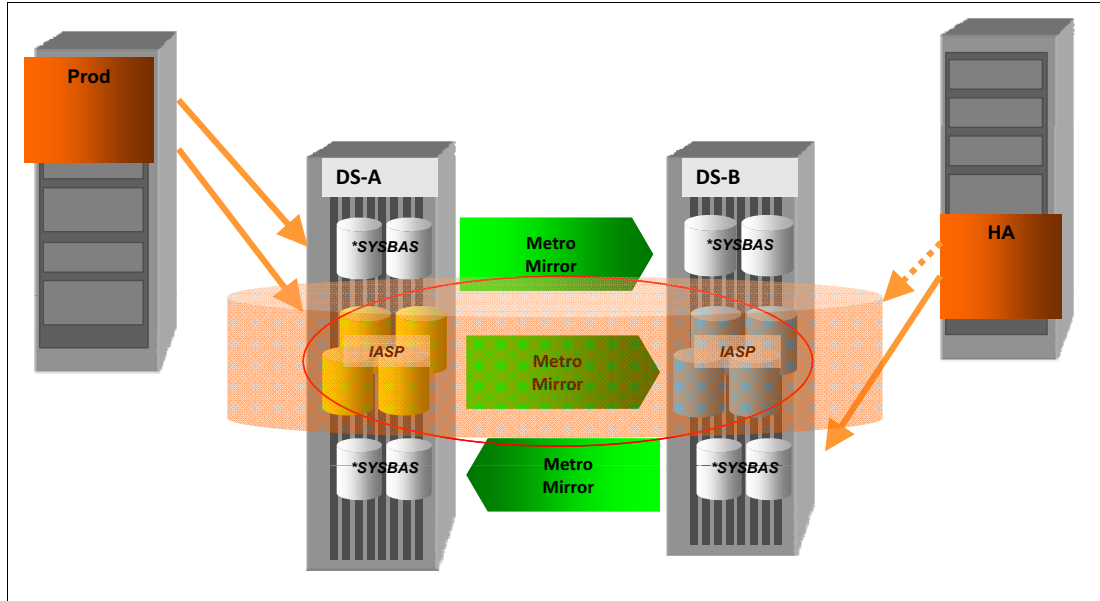
Figure 3-46   IASP HyperSwap

You use the HyperSwap functionality for the SYSBAS LUNs of both the source and target systems. This method protects you from a storage failure if the storage fails. If SYSBAS is not protected this way, it defeats the purpose of using IASP HyperSwap.

Additionally, you also configure HyperSwap for the IASP LUNs. But, you also configure a PowerHA environment by using a variation of the LUN-level switching support.

PowerHA has the capability to declare multiple storage units for a single copy description as shown in Example 3-2.

After you set up the copy description, switching the IASP works the same way as it is described in Chapter 5, "Switching and FlashCopy operations" on page 121 for other PowerHA scenarios.

Example 3-2   ADDASPCPYD command for IASP HyperSwap

```
                    Add ASP Copy Description (ADDASPCPYD)

 Type choices, press Enter.

 ASP copy . . . . . . . . . . . . > HYPERSW       Name
 ASP device . . . . . . . . . . . > HYPERSW       Name
 Device domain  . . . . . . . . .   *             Name, *
 Cluster resource group . . . . . > HYPERSW       Name, *NONE
 Cluster resource group site  . . > SITE1         Name, *NONE
 Storage host:
   User name  . . . . . . . . . . > qlpar
   Password . . . . . . . . . . . > '********'
```

```
   Internet address . . . . . . . > '9.1.1.1'

 Location . . . . . . . . . . . .     *NONE         Name, *DEFAULT, *NONE

 Logical unit name:
   IBM System Storage device  . . > 'IBM.2107-7512341'
   Logical unit range . . . . . . > '8900-8902'   Character value
               + for more values > '8A00-8A02'
   Consistency group range  . . .               Character value
               + for more values
 Recovery domain:
   Cluster node . . . . . . . . . > HYPER1        Character value, *NONE
   Host identifier  . . . . . . . > 0004          Character value
               + for more values > 0005
   Volume group . . . . . . . . . > V39           Character value
               + for more values > V39

   Cluster node . . . . . . . . . > HYPER2        Character value, *NONE
   Host identifier  . . . . . . . > 0006          Character value
               + for more values > 0007
   Volume group . . . . . . . . . > V39           Character value
               + for more values > V39
               + for more values

 Second storage host:
   User name  . . . . . . . . . . > qlpar
   Password . . . . . . . . . . . > '********'
   Internet address . . . . . . . > '9.2.2.2'

 Second logical unit name:
   IBM System Storage device  . . > 'IBM.2107-7543211'
   Logical unit range . . . . . . > '8900-8902'   Character value
               + for more values > '8A00-8A02'
   Consistency group range  . . .               Character value
               + for more values

 Second recovery domain:
   Cluster node . . . . . . . . . > HYPER1        Character value, *NONE
   Host identifier  . . . . . . . > 0010          Character value
               + for more values > 0011
   Volume group . . . . . . . . . > V23           Character value
               + for more values > V23
   Cluster node . . . . . . . . . > HYPER2        Character value, *NONE
   Host identifier  . . . . . . . > 0012          Character value
               + for more values > 0013
   Volume group . . . . . . . . . > V23           Character value
               + for more values > V23

               + for more values

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

3.8 IASP Copy Services Manager

IASP Copy Services Manager (ICSM) is a set of tools that provide additional automation and extend the capabilities of IBM PowerHA SystemMirror for i. The use of ICSM is an alternative to the use of a stand-alone PowerHA implementation of Metro Mirror or Global Mirror.

ICSM consists of two separate components that work together:

► Clustering management

Clustering management uses a data type cluster resource group and device domain data (CRGDDD) to maintain the information about the cluster. It stores the information about the nodes, direction, state of PPRC, and state of FlashCopy. It triggers exit programs on all nodes in the cluster any time that a "cross-node" action is required. It also automatically starts cluster nodes if required. The associated commands are `CRTCSECRG` and `CHGCSEDTA`.

► Copy services environments

Copy services environments store information about particular copy environments in the DS8000, store most information in ASP copy descriptions, and store information in device domain data. Any inconsistency between the copy of data and the ASP copy descriptions can be detected as additional protection for the environment.

See Figure 3-47.
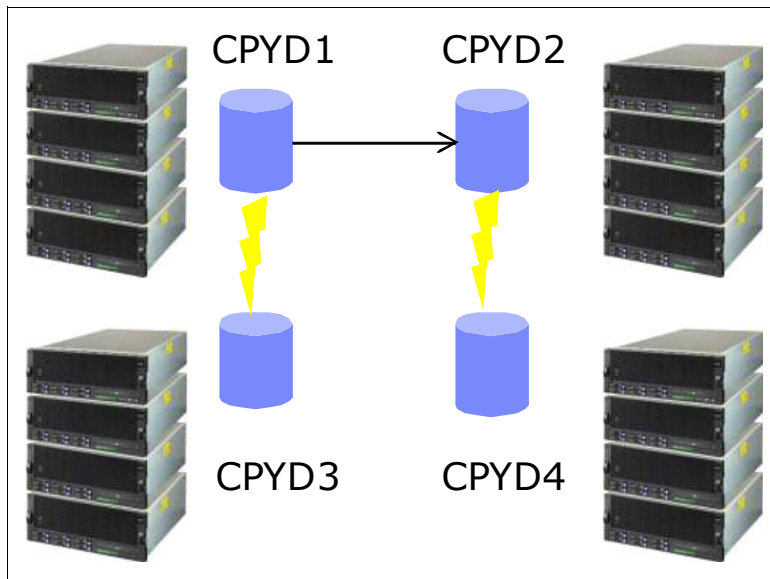
Figure 3-47   Copy services environments

ICSM setup

Run the following steps to set up ICSM:

1. Restore the `ICSM` library and licensing. The `QZRDHASM` library that is provided contains both ICSM and DS Management. Every version and fix is shipped as the entire library. Restore the library.

2. Use `RUNSETUP` to set up the environment. For new installations, this command creates the integrated file system (IFS) directories, restores the files, and so on.

3. Use the **ADDPRDACS** command to add the license keys. ICSM environments can have three keys:

   – A PPRC option:

     • Flash
     • PPRC
     • Flash and PPRC

   – DS management (for DS8000)

   – PowerHA tools

4. The QLPAR user must be created on all nodes.

5. Create a Copy Services Cluster Resource Group Cluster. The nodes must be in a device domain before you can set up the cluster resource group (CRG) to store data across all nodes. **CRTCSECRG** tries to initialize a CRG on all nodes, including running an exit program on every node to set up cluster USRQs (user queues).

6. Establish the connection to the Storage Management Console. Set up the DSCLI user and password files. Set up a `pwfile` (password file) on IBM i partitions.

7. Use the Change CSE CRG Data (**CHGCSEDTA**) command to set details and add additional FlashCopy environments. See Figure 3-48.

```
                        Change CSE CRG Data

 Supply all required values, press Enter.


 Independent ASP CRG Name . . . :    METROCRG
 Copy type  . . . . . . . . . . .    *PPRC          *FLASH, *PPRC, *BOTH
 Independent ASP Name . . . . . .    METROCRG       Name
 Preferred production node  . . .    SOURCE         Name


 MMIR PPRC information:
   Preferred backup node  . . . .    TARGET         Name
   PPRC status  . . . . . . . . .    *READY         *READY, *INCOMPLETE, number
   PPRC direction . . . . . . . .    *NORMAL        *NORMAL, *REVERSED
   Automatic PPRC Replicate . . .    *YES           *YES, *NO


 GMIR PPRC information:
   Preferred backup node  . . . .                   Name


 LUN PPRC information:
   Preferred backup node  . . . .                   Name



 Supply all required values, press Enter.


 Request type . . . . . . . . . .    0              Number
 Auto start cluster . . . . . . .    *YES           *YES, *NO
 Wait time  . . . . . . . . . . .    60             Number of seconds
 Suspend timeout  . . . . . . . .    2              Number of seconds
 Device CRG (optional)  . . . . .                   Name
 Message Queue  . . . . . . . . .    *SYSOPR        name, *SYSOPR
   Library  . . . . . . . . . . .                   library name
```

Figure 3-48   Changing CSE data

8. Each copy services environment gets a set of directories to store scripts and results. This set of directories includes one or more of the following directories:

   – /QIBM/QZRDHASM/SCRIPT/*<ENV>*_Flash/
   – /QIBM/QZRDHASM/SCRIPT/*<ENV>*_MMIR/
   – /QIBM/QZRDHASM/SCRIPT/*<ENV>*_GMIR/
   – /QIBM/QZRDHASM/SCRIPT/*<ENV>*_LUN/
   – /QIBM/QZRDHASM/profiles/……

See the example in Figure 3-49.

```
Work with Object Links


Directory . . . . :    /qibm/qzrdhasm/scripts/<IASPname>_<copytype>/


Type options, press Enter.
  2=Edit   3=Copy   4=Remove   5=Display   7=Rename   8=Display attributes
  11=Change current directory ...


Opt     Object link
__      failoverpprc_to_PS.result
__      failoverpprc_to_PS.script
__      failoverpprc_to_PT.script
__      lspprc_PS.result
__      lspprc_PS.script
__      lspprc_PT.result
__      lspprc_PT.script
__      lspprcpath_PS.script
__      lspprcpath_PT.result
                                                                 More...
Parameters or command
===> _____ .
F3=Exit   F4=Prompt   F5=Refresh   F9=Retrieve   F12=Cancel   F17=Position to
F22=Display entire field        F23=More options
```

Figure 3-49   Copy services environment scripts and results

9. The **WRKCSE** command, option 14 shows a customized subset of a scripts directory and a profile directory. See Figure 3-50.

```
                    Copy Services Environments

 Type options, press Enter.
   1=Add    2=Change       4=Delete   5=Display        12=Work with
   14=List Stream files   16=Define host connections   18=Make PPRC Paths


 Opt   Name          Type     Text


       IASP01        FLASH
       IASP01        GMIR
       IASP01        MMIR



                                                                 Bottom
 Command
 ===>
 F1=Help   F3=Exit   F4=Prompt   F9=Viewlog   F12=Cancel
```

Figure 3-50   Copy Services Environments

10.You can use the **WRKCSE** command to manage the copy services environment easily. See Figure 3-51.

```
 Add, Change or Delete Volumes

   Environment  . :   SEKIU              Source device  :   IBM.1750-ABC1234
   Type . . . . . :   FLASH              Target device  :   IBM.1750-ABC1234
   Volume sets  . :   3

 Type Volume options; 1=Add, 2=Change, 4=Delete, press Enter.
           <Source>        Flash
   Opt     Volumes         Volumes

    _     _____      _____
    _     1300-1302       1900-1902


                                                                 Bottom
 F1=Help   F3=Exit   F12=Cancel
```

Figure 3-51   Changing a copy services environment

11. Create the PPRC paths.

12. Start the replication with ICSM:
   – Metro Mirror
   – Global Mirror

For more information about the IASP Copy Services Manager (ICSM) service offering, contact STG Lab Services on the following website:

https://ibm.biz/Bd4Lbs

3.9 Multiple target replication

Metro Mirror is a synchronous transfer of data from the source DS8000 to the target DS8000, which means that the target data is always consistent while everything is operational. For many years, IBM i clients used Metro Mirror and switched sites successfully in an unplanned failure at the source site.

However, most clients do not plan for or anticipate a "dual failure" in their environments. The concept of a dual failure with Metro Mirror is when the first failure occurs with one or more LUNs in the Metro Mirror relationship being suspended. Other LUNs continue to copy between the source and target DS8000s, causing the data at the target to be inconsistent. Then, a second failure occurs where access to the source DS8000 is lost. When a dual failure occurs, it is not possible to switch to the target DS8000 because the data is inconsistent. Likely, the IBM i system will not be able to IPL or the IASP will not be able to vary on.

This scenario might occur for many reasons. A user error might cause LUNs to be suspended. Hardware issues can occur in the DS8000. A software error might occur on a specific LUN. If the source site cannot be recovered, the client must activate "Disaster mode" and recover the system from the latest backup. Likely, this backup is from the latest FlashCopy, but it might be from any backup and it might be up to 24 hours old.

The Metro Mirror consistency groups on the DS8000 are designed to provide a better solution in this type of a disaster instead of reverting to the last backup data. In a single LUN replication failure, the DS8000 halts I/O to the source LUN and issues a Simple Network Management Protocol (SNMP) trap to alert the IBM i that the replication is broken. For IBM i, this method works effectively because the loss of access to a single disk unit halts all system I/O until access to the LUN is recovered. Therefore, the target disk can be 100% consistent at that point in time.

However, IBM PowerHA SystemMirror for i does not take advantage of Metro Mirror consistency groups for a simple reason. The DS8000 can connect through SNMP to only two IP addresses, which limits its effectiveness to a single pair of systems in a PowerHA cluster. Most IBM i DS8000 users have tens of partitions that attach to their DS8000. Because each pair of nodes in Metro Mirror clusters does not know about the other pairs, no centralized management of multiple consistency groups for IBM i is possible.

Up to now, clients relied on their backup processes to recover from this kind of dual failure. One option was available to use Metro Mirror consistency groups. Clients bought IBM Tivoli® Productivity Center for the Replication Manager function only and used Lab Services integration with Tivoli Productivity Center for Replication.

However, a new cost-effective option is available that is called Cluster Systems Management (CSM). CSM provides a better recovery point in a dual failure. CSM offers the following benefits:

► The target partition might be an earlier level, but it is consistent and available to be brought online immediately.

► The recovery point of the data that was created by the freeze will be at the point where the first LUN stopped replicating, which normally is much later than the point of the last FlashCopy.

► The time of the freeze is available on the `lspprc` command so that clients know the recovery point.

CSM setup

IBM i required the use of Tivoli Productivity Center for Replication and now CSM for Metro Global Mirror since it was introduced. Lab Services provides the integration between PowerHA and CSM for this three-site solution, and Metro Mirror support is also available. In the future, IBM i Lab Services will require CSM for all future three-site solutions:

► Metro Global Mirror (MGM)
► Multiple target Metro/Metro Mirror
► Multiple target Metro Global Mirror

Metro Mirror and Global Mirror must have the same environment name in `WRKCSE`. Global Mirror target FlashCopy must have same name as Metro Mirror and Global Mirror.

Global Mirror target FlashCopy can be run on the FlashCopy partition only.

The number of volumes that are managed in the Tivoli Productivity Center for Replication session must match the number of volumes that are configured in `WRKCSE`. Otherwise, `CHKPPRC` fails.

If you use dual Tivoli Productivity Center for Replication servers, the user and password combination must be the same on each server.

Automatic class selection (ACS) controls the takeover and redefinition of the Tivoli Productivity Center for Replication HA server if or when takeover and redefinition are required.

If you look at the integration of CSM and ICSM, you see differences in the DS8000 connection method. See Figure 3-52.
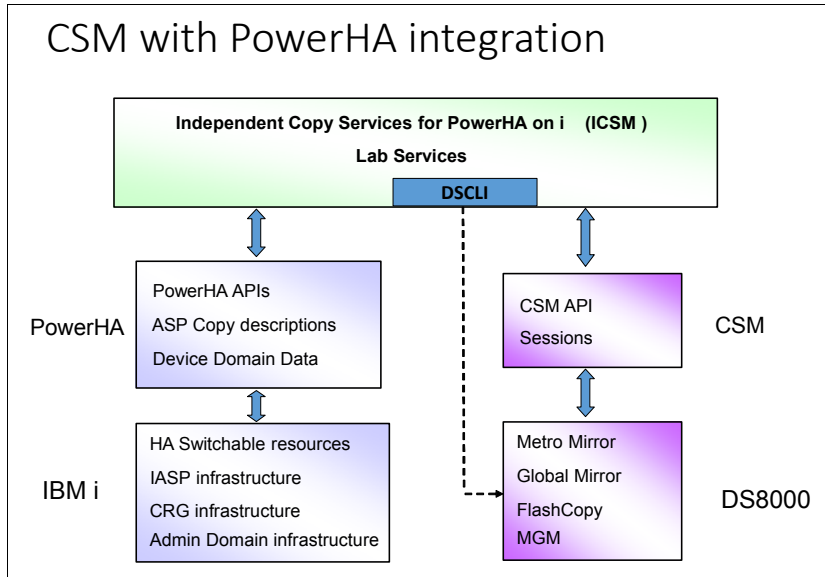
Figure 3-52   CSM integration with PowerHA

The CSM interface to manage the replication replaces the use of DSCLI to create and manage replication in most cases.

CSM provides different tiers of security: admin, ops, and monitor.

CSM is integrated with IASP Copy Services Manager (ICSM) to enable management from the IBM i cluster. Follow these steps:

1. Use `WRKCSE` on IBM i to configure the environments for PowerHA (Figure 3-53).

```
                    Copy Services Environments

 Type options, press Enter.
   1=Add    2=Change        4=Delete    5=Display         12=Work with
   14=List Stream files    16=Define host connections    18=Make PPRC Paths


 Opt    Name          Type     Text

        IASP01        FLASH
        IASP01        GMIR
        IASP01        MMIR




Bottom
 Command
 ===>
 F1=Help    F3=Exit    F4=Prompt    F9=Viewlog    F12=Cancel
```

Figure 3-53   Create the copy services environments

2. Add CSM information in the copy services environment. See Figure 3-54.

```
Change a MMIR Environment.
 Type choices, press Enter.


 Environment name . . . . :    IASP01

 Metro Mirroring Power HA, ASP information:
   Device name  . . . . . .   IASP01                   Name
   Source Copy Description    PROD                      Name
   Target Copy Description    HA                        Name

 TPC information:
   TPC Replication  . . . .   *YES                      *YES, *NO
   Metro-Global Mirroring     *YES                      *YES, *NO
   Primary server . . . . .   1.1.1.1                   IPv4
   Secondary server . . . .   2.2.2.2                   IPv4
   Session name . . . . . .   MGM-TEST                  Name
   User . . . . . . . . . .   Administrator             Profile name
   Password . . . . . . . .
```

Figure 3-54   Adding CSM information to the copy services environment

Custom actions from IBM i are built into the interface with CSM. A *CHK action performs the storage steps of a `CHKPPRC` command (as in ICSM) for the specified environment. It returns an XML document with the status information of the environment. Also, this *CHK action is integrated directly into the ICSM `CHKPPRC` and `SWPPRC` commands.

See Figure 3-55.

```
                      Run TPC Action (RUNTPCACT)


 Type choices, press Enter.

  Environment name . . . . . . . . . > IASP01        Name
  Type . . . . . . . . . . . . . . .   *MMIR         *GMIR, *MMIR
  Action . . . . . . . . . . . . . . > *CHK          *CHK, *RTV, *RUN, *ACS, *RCS




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

*Figure 3-55   *CHK action*

In addition to the *CHK action, CSM offers other actions:

► **`*RCS`**: Remove Copy Sets. This action performs a soft removal of all copy sets in an environment as part of the MGM DR test process.

► **`*ACS`**: Add Copy Sets. This action places all of the copy sets from a file into the CSM session.

► **`*RTV`**: Retrieve. Use this action to retrieve attributes from a CSM session, including the list of available commands.

IBM CSM was available as a request for price quotation (RPQ) in 2015. CSM was available generally at the beginning of 2016.

CSM provides the following benefits:

- ▶ Provides a central place for control of the replication environment.
- ▶ Simplifies and automates complex replication tasks.
- ▶ Removes the need for complicated error-prone scripting.
- ▶ Provides an intuitive web browser-based GUI for complex solutions.
- ▶ Provides a command-line interface (CLI) for easy automation.
- ▶ Notifies a contact when events and alerts occur.

CSM supports the following products and environments:

- ▶ All IBM Storage Systems and heterogeneous environments: DS8000, SVC, Storwize family, and IBM XIV®.

- ▶ All storage-based Copy Services features: FlashCopy, Metro Mirror, Global Copy, Global Mirror, Metro Global Mirror, and Multiple Target (Multi-target).

- ▶ Support for IBM z/OS® HyperSwap technology in two-site and three-site environments.

- ▶ CSM is data-agnostic and covers z/OS volumes (CKD) and open systems LUNs (FB).

- ▶ Installation on all major platforms: Microsoft Windows, AIX, Linux, Linux for System z®, and z/OS.

- ▶ Business continuity.

Multi-target is a significant improvement over traditional MGM because it eliminates all of the required full resynchronizations when the system switches to the third site. Multi-target allows any site of the environment to switch, whether the third DS8000 is active or not. Multi-target offers a "detach" capability, which is the ability to fail over and use either of the two target DS8000s without affecting the production system.

4

Monitoring and managing your environment

This chapter describes methods that help you monitor and manage your IBM PowerHA SystemMirror for i in an IBM System Storage DS8000 environment.

The following topics are described in this chapter:

► 4.1, "Monitoring your environment" on page 94
► 4.2, "Managing your environment" on page 114

Note: All of the example 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 163.

4.1 Monitoring your environment

Monitoring your environment is probably the most important task that you will undertake in a high availability (HA) environment.

Historically, it was sufficient to check your environment one time a day. However, the problem with checking only one time a day is that all you discover is that your availability solution is unusable at the time that you run the check, and that your environment might be unusable for anything from a minute to 24 hours ago. If you check only one time a day and discover a failure, on average you must assume that the solution was not available for you to use for an average of 12 hours. Putting that in perspective, your availability solution is only usable approximately 50% of the time.

This situation is prevented easily by implementing an automated, proactive, monitoring regime. By using tools and automation to monitor on a regular and automatic basis, you are alerted when situations need a resolution. Therefore, you understand the current ability to switch if a problem occurs with your production partition.

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

You need to consider many areas when you monitor your environment, including the following areas:

- ▶ 4.1.1, "QSYSOPR message queue" on page 94
- ▶ 4.1.2, "Monitoring the status of your replication" on page 95
- ▶ 4.1.3, "Monitoring the status of cluster nodes" on page 99
- ▶ 4.1.4, "Monitoring the cluster administrative domain status" on page 102
- ▶ 4.1.5, "Monitoring the Cluster Resource Group status" on page 104
- ▶ 4.1.6, "Monitoring the independent auxiliary storage pool status" on page 107
- ▶ 4.1.7, "Other monitoring considerations" on page 111
- ▶ 4.1.8, "IASP Copy Services Manager CHKPPRC command" on page 111

Different environments have different characteristics and different potential points of failure. Ensure that any monitoring processes that you implement can identify any of these potential points of failure so that you have enough time to make informed decisions about the resolution of any issues that arise. What you do not want to do is discover that you have a problem when you need to switch in a hurry.

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

4.1.1 QSYSOPR message queue

The first place to look for information about the health of your replication solution is always 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 check occurs once a day. With the hardware solutions, it is a simple process, so you can check more frequently. However, do not run a check *too* frequently.

> **Note:** A good policy might be to monitor two or three times a day. Possibly, your plan can include the 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 this frequency is excessive.

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

Example 4-1 shows a sample CL program to assist you in building 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 163.

> **Tip:** If you want to use the sample code that is shown in Example 4-1, you must remember to 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 that */
/*         I more could exist in which case this sample code would need to  */
/*         I be reworked to allow for this possibility.                   */
```

```
//*            ============================================================ */
/*                                                                          */
/****************************************************************************/
/* 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 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 the 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 joblog 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 upon 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, +
                          please review') TOMSGQ(*SYSOPR)

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

4.1.3 Monitoring the status of cluster nodes

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

When you verify the status of the cluster nodes, you must verify the status from the
perspective of a single node. You retrieve the local node status. Then, from the local node,
you also obtain the remote node status. This action is important because failing to verify the
status correctly can result in a belief 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 on the partition LONDON, it states that it is ACTIVE, and if you check on the partition FRANKFURT, it also states that it is ACTIVE. Only by performing the check on a single partition can you see that the cluster is in a PARTITION state.
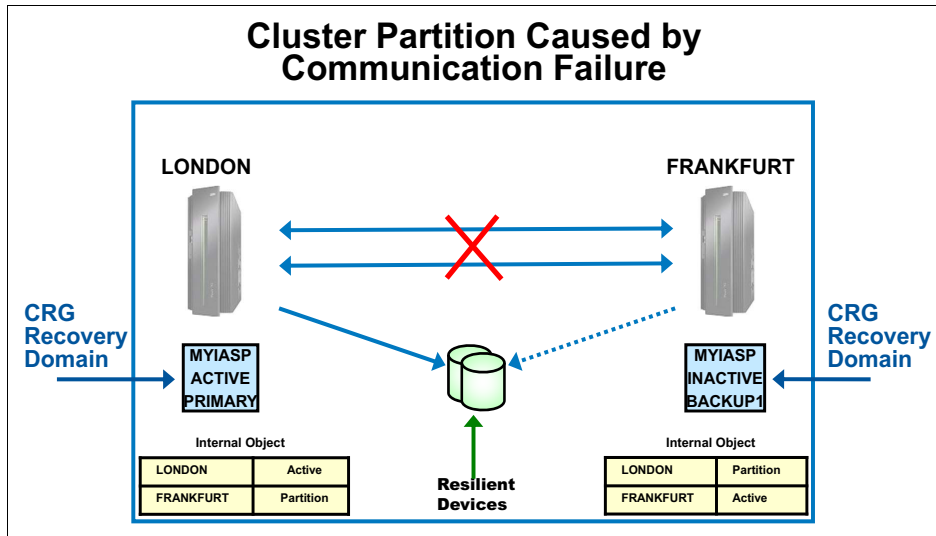
Figure 4-1   A partitioned cluster

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

Example 4-2 shows a sample CL program that can assist you in building 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 163.

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, please 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, please +
                       investigate') TOMSGQ(*SYSOPR)

/* Step 7 - Send an escape message to indicate there is an error */
           SNDPGMMSG  MSGID(CPF9898) MSGF(QCPFMSG) MSGDTA('Not all +
                       cluster nodes are active, please +
                       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. The cluster administrative domain is reliable, except when the nodes that are involved in the cluster administrative domain are inconsistent. For example, you deleted a user profile and forgot to remove the monitored resource entry (MRE) from the cluster administrative domain, resulting in an inconsistency. Another example is adding a user with 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 using the Print Cluster Administrative Domain Monitored Resource Entry (**PRTCADMRE**) command to produce an output file of monitored resource entries that are not in a global consistent state. You can then read how many entries are in the file and if the number is not zero, you can then alert the system operators by sending a message to the QSYSOPR message queue or by using 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 163.

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 joblog 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 we can 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, please review') TOMSGQ(*SYSOPR)

/* Step 6 - Send an escape message to indicate there is an error */
             SNDPGMMSG  MSGID(CPF9898) MSGF(QCPFMSG) MSGDTA('Errors +
                          were found in the Cluster Administrative +
                          Domain, please 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. Therefore, 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 the CRGs 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 that is shown 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 163.

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, please 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, +
                      please reviews') TOMSGQ(*SYSOPR)

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

4.1.6 Monitoring the independent auxiliary storage pool status

One of the important considerations for the independent auxiliary storage pool (IASP) is that it cannot overflow into SYSBAS. Therefore, you need to review the current utilization actively.

Although you can set a threshold within System Service Tools (SST), this threshold notifies you only when you exceed the threshold. You can also write a program to use the system API **QYASPOL**, which can provide the total and current capacity. It is also possible to obtain the number of logical unit numbers (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 changes, 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 163.

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 utilization, 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 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)
/* Utilization 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 utilization */
              CHGVAR      VAR(&CURUSED) VALUE(&ASPCAPT - &ASPCAPTA)
              CHGVAR      VAR(&UTILPCT) VALUE(100.0 * (&CURUSED / +
                            &ASPCAPT))

/* Step 3 - Retrieve the previous utilization 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 utilization,*/
/* then error. The logic here is that if the actual storage used has increased  */
/* by more than the desired 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, please 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 thanexpected.') NOTE('The IASP +
                             utilization 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, Please 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,  please +
                             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 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 utilization or +
                           configuration') MSGTYPE(*ESCAPE)
              ENDDO
              RETURN
endpgm
```

4.1.7 Other monitoring considerations

In addition to the monitoring that is described earlier in this chapter, you also need to
investigate other methods of monitoring. The following list shows other typical areas to
monitor:

► Network switches
► Power service
► Storage area network (SAN) switches
► Storage subsystems

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

Often, these areas (and other areas) can be monitored by the use of Simple Network
Management Protocol (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 capability allows the monitor to
raise alerts in real time.

For example, a storage subsystem typically raises a trap if it loses connectivity to a remote
storage subsystem or if a physical component fails internally.

4.1.8 IASP Copy Services Manager CHKPPRC command

If you are using a toolkit from IBM Systems Lab Services, different methods are available for
monitoring and managing your solution. These toolkits are developed to enhance and extend
the functionality that is available with PowerHA. An example is the Check PPRC (`CHKPPRC`)
command, which provides a simple way to verify multiple aspects of your environment in a
single command.

> **Note:** The `CHKPPRC` command is a monitoring command that is provided by IASP Copy
> Services Manager for use in an IBM Systems Lab Services Toolkit environment.

Part of enhancing the PowerHA experience is to make the solution even easier to use. To
meet this goal, the `CHKPPRC` command either completes normally if everything is OK, or it ends
with an escape message if any problem causes you to be unable to switch. This way, you can
use a simple CL program to run the `CHKPPRC` command and use MONMSG to identify an error
situation.

An example of the **CHKPPRC** command is shown in Figure 4-2.

```
                          Check PPRC (CHKPPRC)

  Type choices, press Enter.


  Environment name . . . . . . . .   ICSMMM         Name
  Type . . . . . . . . . . . . . .   *MMIR          *, *GMIR, *LUN, *MMIR...




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

Figure 4-2   The CHKPPRC command

The **CHKPPRC** command is simple to use. You provide only the environment name and type of replication.

If you want to understand the activities that are performed by this command in more depth, look at the log of its activities as shown in Example 4-6.

Example 4-6   Log of CHKPPRC activities

```
2015-09-24 15:08:01  Start  CHKPPRC for ACSMMIR starting from job 372276/DPAINTER/QPADEV0010
                            IBM i Copy Services Manager version 4.1.0 built Thu Sep 24 09:28:38 2015.
2015-09-24 15:08:01          ACSMMIR is not PPRC Switch locked. (chkDDDlock)
2015-09-24 15:08:01  Error  More than 1 work record and type = *' (chkpprc)
2015-09-24 15:08:01          CHKPPRC/CHKCSE command was canceled. (ppckCancelHandler)

2015-09-24 15:08:12  Start  CHKPPRC for ACSMM starting from job 372276/DPAINTER/QPADEV0010
                            IBM i Copy Services Manager version 4.1.0 built Thu Sep 24 09:28:38 2015.
2015-09-24 15:08:12          Ignored CPF9801 message returned by the command:
                            CHKOBJ OBJ(QUSRSYS/ACSMM) OBJTYPE(*CRG) (doCmd)
2015-09-24 15:08:12  Error  CRG not found, ACSMM is a stand alone environment. (rtvCScrgData)
2015-09-24 15:08:12          CHKPPRC/CHKCSE command was canceled. (ppckCancelHandler)

2015-09-24 15:08:19  Start  CHKPPRC for ACSMMIR starting from job 372276/DPAINTER/QPADEV0010
                            IBM i Copy Services Manager version 4.1.0 built Thu Sep 24 09:28:38 2015.
2015-09-24 15:08:19          ACSMMIR is not PPRC Switch locked. (chkDDDlock)
2015-09-24 15:08:19          Some scripts not written, no ACS PPRC path data stored for ACSMMIR. (constructPathList)
2015-09-24 15:08:19         Some scripts not written, no ACS PPRC path data stored for ACSMMIR.
2015-09-24 15:08:19        Write scripts successful for environment ACSMMIR type MMIR. (writeAllScripts)
2015-09-24 15:08:19          Starting DS SMC checks. (chkpprc)
2015-09-24 15:08:19          Starting... (fillStatus)
2015-09-24 15:08:19          A MMIR  PPRC CRG is active, status = 0.
2015-09-24 15:08:19          Executing with Java Tunnel; /QIBM/Qzrdhasm/scripts/ACSMMIR_MMIR/lspprc_PS.script
(doPPRCScript)
2015-09-24 15:08:27          Tunnel to '9.5.168.55' opened, PID = 645, index = 0. (doJavaTunnelCmd)
2015-09-24 15:08:29      Processing file /QIBM/Qzrdhasm/scripts/ACSMMIR_MMIR/lspprc_PS.result. (checkThoseResults)
2015-09-24 15:08:29 Warning Strings Suspended | Duplex | Copy |  |  not found in any records.
2015-09-24 15:08:29 Warning Expected lspprc_PS.script results not found. (doPPRCScript)
2015-09-24 15:08:30          Executing with Java Tunnel; /QIBM/Qzrdhasm/scripts/ACSMMIR_MMIR/lspprc_PT.script
2015-09-24 15:08:31      Processing file /QIBM/Qzrdhasm/scripts/ACSMMIR_MMIR/lspprc_PT.result. (checkThoseResults)
2015-09-24 15:08:31 Warning Strings Suspended | Duplex | Copy |  |  not found in any records.
2015-09-24 15:08:31 Warning Expected lspprc_PT.script results not found. (doPPRCScript)
2015-09-24 15:08:31 Warning No SMCs are available. (fillStatus)
2015-09-24 15:08:31          Returning; direction = 0, statusGMIR = -1, statusMMIR = 0, statusPPRC = -1, cluActive =
0, smcsActive =
2015-09-24 15:08:31       listStorageResources: Name=DC05    , bus=255, slot=   . (listStorageResources)
2015-09-24 15:08:31          listStorageResources: Name=DC06    , bus=255, slot=C20 .
2015-09-24 15:08:31          listStorageResources: Name=DC06    , bus=255, slot=C20 .
2015-09-24 15:08:31          listStorageResources: Name=DC07    , bus=255, slot=C21 .
2015-09-24 15:08:31          listStorageResources: Name=DC01    , bus=255, slot=C25 .
2015-09-24 15:08:31          listStorageResources: Name=DC02    , bus=255, slot=C26 .
2015-09-24 15:08:31          listStorageResources: Name=DC03    , bus=255, slot=C23 .
2015-09-24 15:08:31          listStorageResources: Name=DC04    , bus=255, slot=C27 .
2015-09-24 15:08:31        PS Volumes: 6000 6001 6002 6100 6101 6102   (markStorageResources)
2015-09-24 15:08:31          Executing with Java Tunnel; /QIBM/Qzrdhasm/scripts/ACSMMIR_MMIR/lsfbvol_PS.script
(doAnyScr
```

```
2015-09-24 15:08:32          lsfbvol_PS.result size = 180 (markStorageResources)
2015-09-24 15:08:32          1 strlen r109 = 180
2015-09-24 15:08:32          Executing with Java Tunnel; /QIBM/Qzrdhasm/scripts/ACSMMIR_MMIR/lsallhostconn_PS.script
(do
2015-09-24 15:08:33          lsallhostconn_PS.result size = 180 (markStorageResources)
2015-09-24 15:08:33          2 strlen r109 = 180
2015-09-24 15:08:33 Error  Volume  6000  not found in the lsfbvol result.
2015-09-24 15:08:33 Error  Volume  6001  not found in the lsfbvol result.
2015-09-24 15:08:33 Error  Volume  6002  not found in the lsfbvol result.
2015-09-24 15:08:33 Error  Volume  6100  not found in the lsfbvol result.
2015-09-24 15:08:33 Error  Volume  6101  not found in the lsfbvol result.
2015-09-24 15:08:33 Error  Volume  6102  not found in the lsfbvol result.
2015-09-24 15:08:33 Error  Volume  6102  not found in the lsfbvol result.
2015-09-24 15:08:33          Volume groups:
2015-09-24 15:08:33          WWPNs:
2015-09-24 15:08:33          This is a MMIR solution. (chkpprc)
2015-09-24 15:08:33          MMIR direction is reversed.
2015-09-24 15:08:33          MMIR exit data status is 0
2015-09-24 15:08:33          The current MMIR node is DEMOHA  .
                             The current Production Node is DEMOPROD.
2015-09-24 15:08:33 Error  No DS SMC's are currently active.
2015-09-24 15:08:33          Starting checks of the cluster nodes.
2015-09-24 15:08:33          3 nodes are listed in the CRG ACSMMIR. (crgNodesAreActive)
2015-09-24 15:08:33          3 nodes are listed in the CRG ACSMMIR.
2015-09-24 15:08:33          Starting check of the Device Domains. (chkpprc)
2015-09-24 15:08:33          Executing with Java Tunnel; /QIBM/Qzrdhasm/scripts/ACSMMIR_MMIR/lspprcpath_PS.script (d
2015-09-24 15:08:34          Executing with Java Tunnel; /QIBM/Qzrdhasm/scripts/ACSMMIR_MMIR/lspprcpath_PT.script
2015-09-24 15:08:35          PPRC Path check for MMIR Environment ACSMMIR was successful. (chkpprc)
2015-09-24 15:08:35 Error  Check PPRC failed because one or more HMCs are not available.
2015-09-24 15:08:35          Ended tunnel to 9.5.168.55, PID = 645, index = 0. (closeJavaTunnel)
2015-09-24 15:08:35 Error  FAILED, refer to errors logged above. (chkpprc)
2015-09-24 15:08:35          CHKPPRC/CHKCSE command was canceled. (ppckCancelHandler)
```

The **CHKPPRC** command contains many tests to answer the question, "If I need to switch to this node, can I?"

The tests include the following checks:

► Does the CRG exist for the environment?

► Is the environment locked (in the middle of an operation)?

► Can a connection be made to the storage controllers?

► Can the storage controllers respond?

► What state does the source controller think the replication is in?

► What state does the target controller think the replication is in?

► Do both storage units agree on the replication state?

► Do all of the LUNs in the IASP exist?

► Are the LUNs in the configuration logically connected to this system?

► Can replication occur in both directions?

As you can see, the list is detailed. It is a good example of the benefits that the IBM Systems Lab Services toolkits can bring to your HA environments. For more information, see Appendix A, "PowerHA Tools for IBM i" on page 159.

4.2 Managing your environment

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

► 4.2.1, "Adding disks to an IASP in a PowerHA environment" on page 114
► 4.2.2, "Removing disks from an IASP in a PowerHA environment" on page 116
► 4.2.3, "Adding users to the systems" on page 117
► 4.2.4, "Adding network printers and other SYSBAS objects to the systems" on page 118
► 4.2.5, "Upgrading your operating system" on page 118
► 4.2.6, "Powering off your system" on page 119
► 4.2.7, "Verifying the health of your replication solution" on page 120

4.2.1 Adding disks to an IASP in a PowerHA environment

> **Important:** Be careful when you add logical unit numbers (LUNs) to a PowerHA environment because you must follow a particular sequence.
>
> Also, no replication can be in place for the new LUNs when you start this procedure.

Run the following steps to add disks to an IASP in a PowerHA environment:

1. Use the Work with Cluster (`WRKCLU`) command to ensure that all nodes are active in the cluster.

2. Run the following steps to configure the new volumes and attach them to the various nodes in the cluster. The new volumes must 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. Run the following steps to initialize and format the new volumes. (For DS8000 installations, run the steps on the production node only. For Storwize installations, run the steps on all nodes).

   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 of the new volumes, press Enter, and then press F10 to confirm. Wait for message that states "`Initialize and format completed successfully.`"

   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. Run the following steps to add the volumes to the IASP on the production node only:

   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 adding units.

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

   j. Exit SST.

8. Run the following steps to end any existing ASP sessions:

   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. Document the settings, and then, select option 24 - End Session.

   d. Repeat until your copy descriptions no longer show that any sessions are active against them.

9. Run the following steps to add the LUN ranges to the production source and target copy descriptions:

   a. Run `WRKCLU`.

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

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

   d. Add the LUN ranges to the existing ranges.

   e. Press Enter.

10. Run the following steps to add the LUN ranges to any FlashCopy copy descriptions:

    a. Run `WRKCLU`.

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

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

    d. Add the LUN ranges to the existing ranges.

    e. Press Enter.

11. Run the following steps to restart any ASP sessions that you ended:

    a. Run `WRKCLU`.

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

c. By using the information that you gathered in step 8 on page 115, select option 21 - Start Session.

d. Repeat these steps until your copy descriptions show that the correct sessions are active against them.

4.2.2 Removing disks from an IASP in a PowerHA environment

Be careful when you remove LUNs from a PowerHA environment because you must follow a particular sequence. Run the following steps:

1. Use the `WRKCLU` command to ensure that all nodes are active in the cluster.

2. Run the following steps to remove the volumes that you want to remove from the IASP on the production node only:

   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 by entering a 4 by each one that you want to remove, and press Enter.

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

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

   j. Exit SST.

3. Ensure that any FlashCopy is ended.

4. Run the following steps to end any existing ASP sessions:

   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. Document the settings, and then select option 24 - End Session.

   d. Repeat until your copy descriptions no longer show that any sessions are active against them.

5. Run the following steps to remove the LUN ranges from the production source and target copy descriptions:

   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. Run the following steps to remove the LUN ranges from any FlashCopy copy descriptions:

   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. Run the following steps to restart any ASP sessions that you ended:

      a. Run **WRKCLU**.

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

      c. By using the information that you gathered in step 4 on page 116, select option 21 - Start Session.

      d. Repeat these steps until your copy descriptions show that the correct sessions are active against them.

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

9. Run the following steps to verify that the old volumes 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.

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

4.2.3 Adding users to the systems

When you add users to the systems, you must also add them to the cluster administrative domain.

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

► Use the Add Cluster Admin Domain MRE (**ADDCADMRE**) command to add a user to the cluster administrative domain.

► Use the system-supplied exit point on the Create User Profile (**CRTUSRPRF**) command to add the user automatically to the list of monitored resource entries (MREs).

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

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

    You need to write a simple CL program that performs the **ADDCADMRE** command for you automatically.

> **Important:** When you write a program for the **QIBM_QSY_CRT_PROFILE** exit point, you need to know that when the creation of the user profile occurs on other nodes in the partition, it also calls the exit program. Therefore, you need to include a check to see whether the user who is running the program is QSYS. If the user who is running the program is QSYS, terminate with no further action.

4.2.4 Adding network printers and other SYSBAS objects to the systems

Many objects in SYSBAS can be replicated with the cluster administrative domain. How you manage these various objects varies, depending on several factors.

For devices, such as network printers, the network printers are not created frequently, so adding them manually after creation is often the sensible approach. For other objects, you might use a program that creates them in a specific way. Therefore, it makes sense to change that program to include the Add Cluster Admin Domain MRE (`ADDCADMRE`) command.

Whichever route you choose, certain objects "happen to be the same" and need to be managed under change control, rather than "must be the same" and need to 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 (TR) with less production impact than without PowerHA. Two scenarios are involved:

► PTFs and technology refreshes
► Operating system upgrades

Although both scenarios benefit from the use of PowerHA, the way that 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 correctly install the updates, you can then schedule a switch at a convenient time so that 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 that you want to apply require an IPL, you must suspend replication before you start the process.

Operating system upgrades

The operating system upgrades scenario is slightly more complex because after 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 affecting production. Rather than use a switch to use the new code level, you need to vary off the IASP. Then, issue a detach of the replication with the following Change Auxiliary Storage Pool Session (`CHGASPSSN`) command, where *sessionname* is the name of the replication session:

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

This command 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. Issue 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 you attempt to restart the replication.

4. When you are ready to go into production, vary off the upgraded copy.

5. Issue an End Cluster Resource Group (**ENDCRG**) command.

6. Issue a Change Cluster Resource Group (**CHGCRG *CHGCUR**) command, and change the primary and backup nodes.

7. Issue a Start Cluster Resource Group (**STRCRG**) command.

8. Issue a **CHGASPSSN** with OPTION(*REATTACH) command from the upgraded node to the old production node.

9. Vary on the IASP on the upgraded node, and replication restarts. Under certain 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. Use 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 result in being unable to start your cluster nodes and use PowerHA.

4.2.6 Powering off your system

When you power off a cluster node, follow the correct 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 action that you want. Therefore, it is best to establish a procedure for a PWRDWNSYS operation. Run the following steps:

1. Notify users that the power down is imminent (existing process).

2. Act to prepare any IASPs for the vary off and the switch.

3. Vary off the IASPs.

4. If the node that is powering down is the primary node for an active CRG, follow these steps:

   a. Issue a Change CRG Primary (**CHGCRGPRI**) command to move the active resources to a node that remains active in the cluster.

   b. Issue an **ENDCRG** command to prevent the resources from moving to a node that remains active in the cluster.

5. Issue an End Cluster Node (**ENDCLUNOD**) command for the node that is powering down.

6. Power off the node.

4.2.7 Verifying the health of your replication solution

At certain times, you will want to look at the overall health of your replication solution. With a PowerHA environment, you can use the graphical user interface (GUI) that is shown in Figure 4-3.

As shown in Figure 4-3, multiple indicators show the health of your replication. At the top of the window, the local node shows a green check mark in a box, which indicates that PowerHA is operational on this node. Further down the window, you see that the CRG shows a green check mark in a box, which indicates that the CRG is active. Also, you see the green arrow in the list of nodes that shows that the replication is working and is going from CTCIHA4A to CTCIHA4D.

If any box is not green, a problem exists and you need to investigate.

Tip: You can hover the mouse over the color indicator and a pop-up window opens that describes the condition that is indicated.
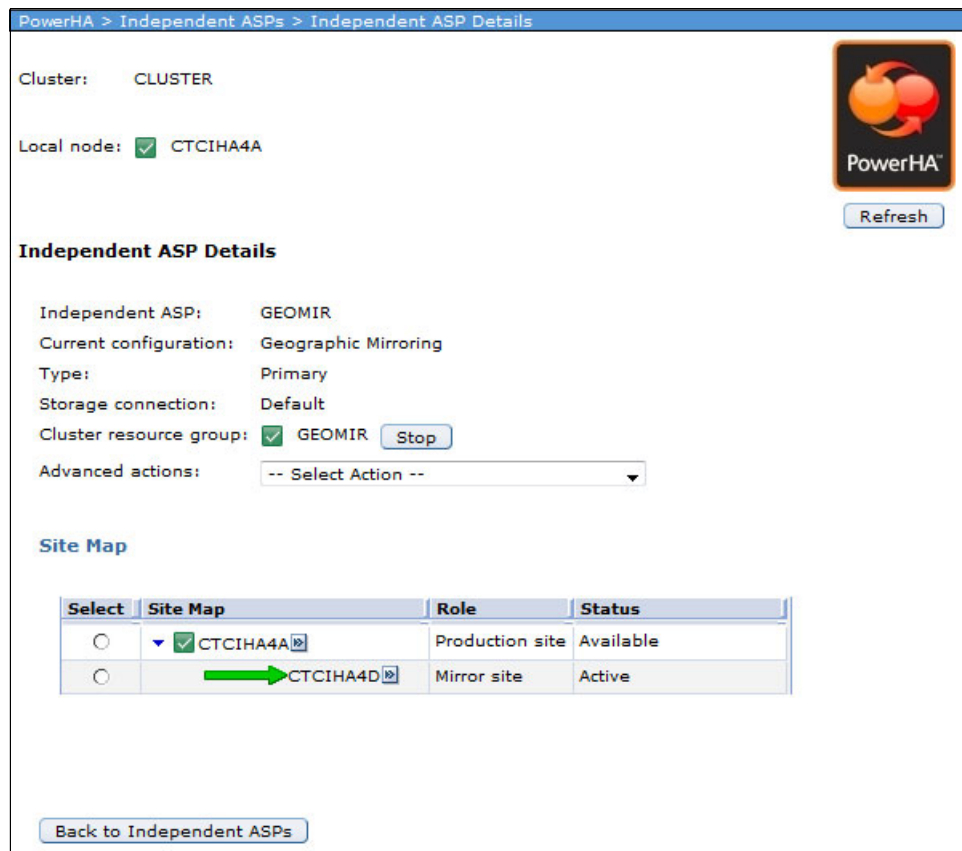
Figure 4-3   PowerHA GUI view of the health of the replication

5

Switching and FlashCopy operations

This chapter describes the operation of an IBM PowerHA SystemMirror for i environment that uses the IBM DS8000 series storage products and their remote copy and FlashCopy options. The operations include the switching between nodes in both a planned mode and an unplanned mode and also the use of the FlashCopy function.

The following topics are covered in this chapter:

- ► 5.1, "Planned switching" on page 122
- ► 5.2, "Failover after a production system outage" on page 133
- ► 5.3, "Scheduled or unscheduled switching with IASP Copy Services Manager" on page 139
- ► 5.4, "Working with PowerHA FlashCopy" on page 140
- ► 5.6, "Using FlashCopy when you change your application" on page 146
- ► 5.7, "Working with Full System FlashCopy" on page 147
- ► 5.8, "Full System Copy Services Manager (FSCSM)" on page 148
- ► 5.9, "HyperSwap" on page 148
- ► 5.10, "Full System replication switch" on page 155

5.1 Planned switching

A *planned switch* is normally performed when you either want to perform a maintenance action on your current primary system, any time you feel your production site might be at risk (for example, power or networking modifications), or as part of a routine planned role swap.

You need to follow these steps so that the switchover 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 94, you also might want to check that your cluster environment is up and running and that the data replicates correctly between the nodes before you start the switchover.

    You can use the cluster graphical user interface (GUI). Open a web browser session to the IBM Navigator for i and select **PowerHA**:

    `http://<yoursystemname>:2001`

    The PowerHA overview window opens as shown in Figure 5-1.

    This view provides you with an overview of your entire cluster's status. If any of the icons do not show a green check mark for the status, ensure that you fix the potential problem before you proceed with the switch.
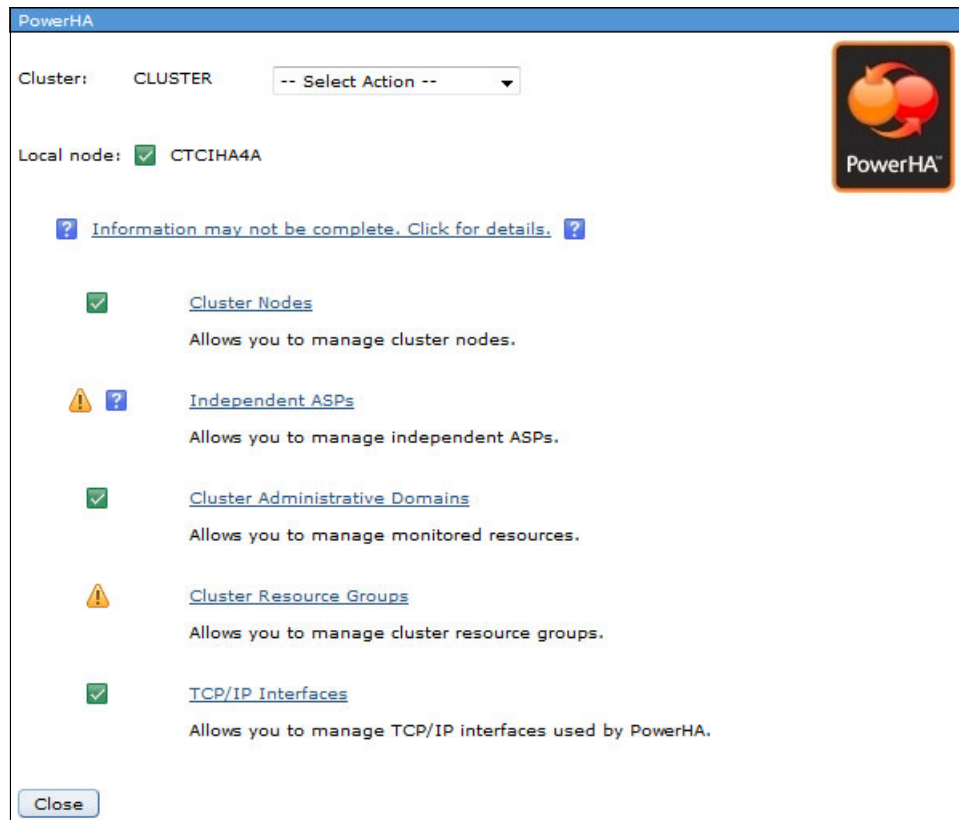
Figure 5-1   IBM Navigator for i, PowerHA overview window, checking the cluster's status

2. End your applications and get all of the users off the system. You can check whether jobs are connected to the independent auxiliary storage pool (IASP) with the Work with ASP Jobs (**WRKASPJOB**) command. If jobs are still connected, the output is similar to Figure 5-2.

```
                          Work with ASP Jobs                              CTCIHA4A
                                                                          09/22/15
 10:16:51
  Type options, press Enter.
    4=End job   5=Work with   7=Send message

  Opt     ASP/Job             User          Number      Type         Status
          PWRHAMM
            QPADEV000F       EGRYGIER        081481      INT            RUN




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

Figure 5-2   WRKASPJOB output window

3. Ensure that your job is not connected to the IASP by issuing the following command:

   SETASPGRP ASPGRP(*NONE)

   > **Important:** If your job is connected to the IASP while it runs the switchover process, your job is 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 of this issue.

4. Vary off the IASP:

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

   For special considerations when you use a Global Mirror setup, see 5.1.3, "Special considerations when you switch in a Global Mirror setup" on page 131.

5. Perform the actual switchover by using the IBM Navigator for i PowerHA GUI. From the main PowerHA overview window (Figure 5-1 on page 122), click **Cluster Resource Groups**.

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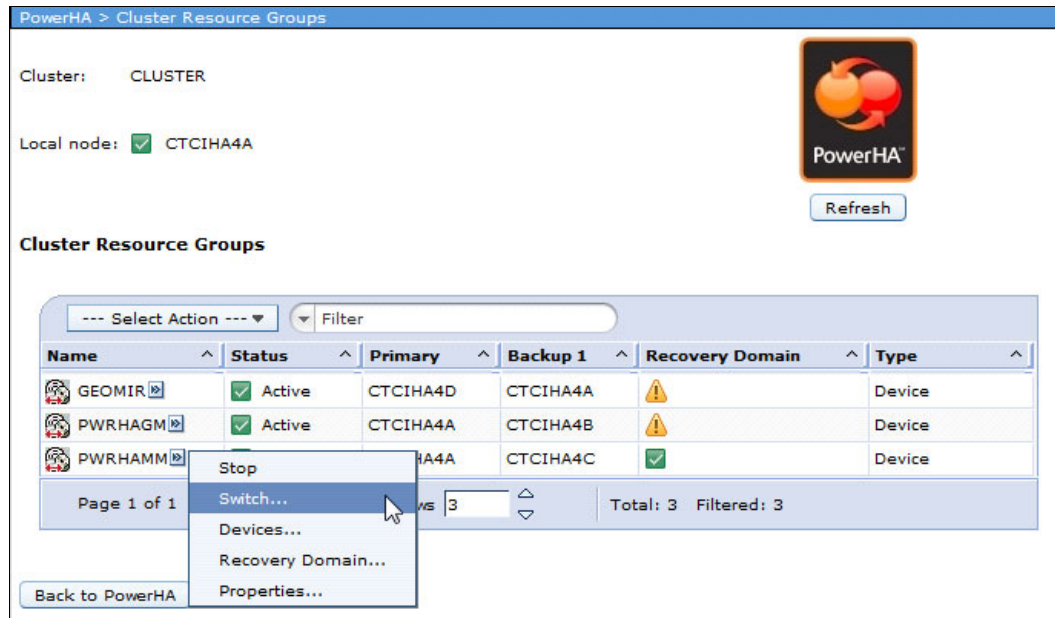

Figure 5-3   Performing the switchover from the IBM Navigator for i PowerHA GUI

7. A verification window opens. You can obtain information about the current recovery domain and the recovery domain after the switch. You can view information about the IASP that will switch over as shown in Figure 5-4. Click **OK** to initiate the switchover.
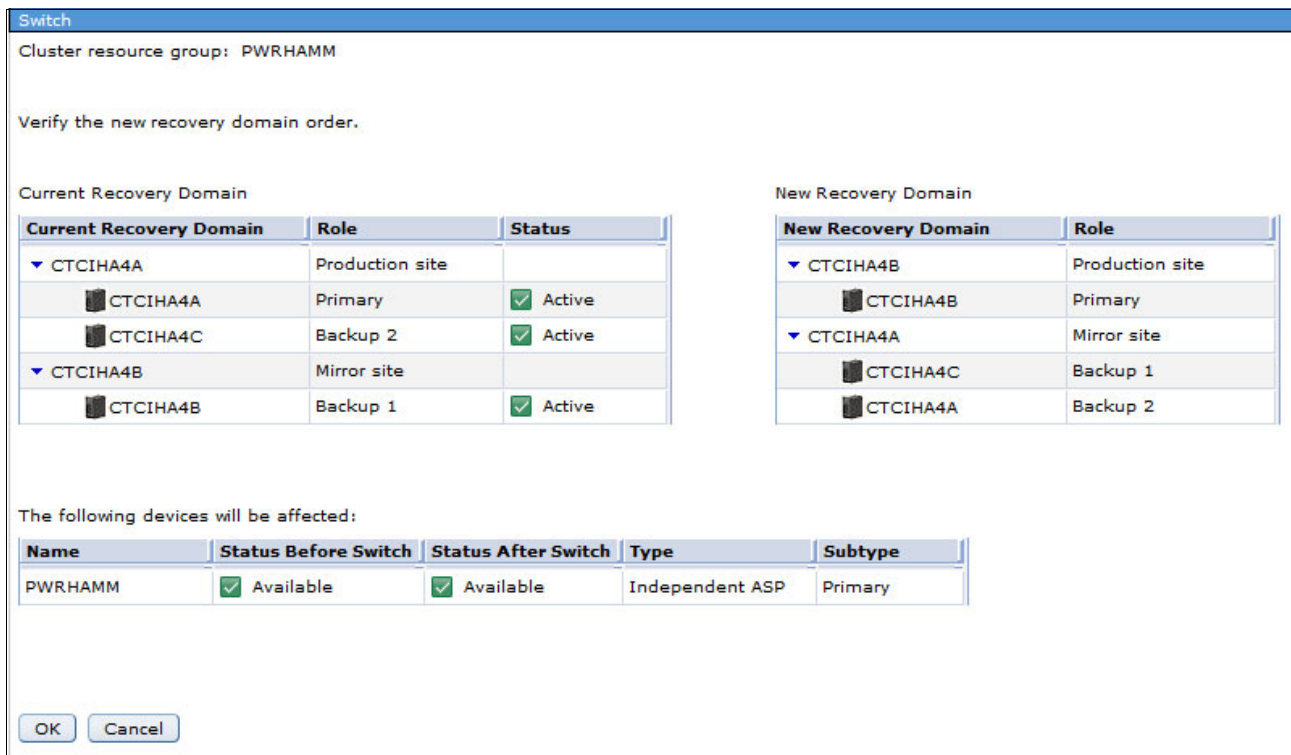

Figure 5-4   Verify the information about the switchover

8. Figure 5-5 shows the window that opens when you prepare to switch.
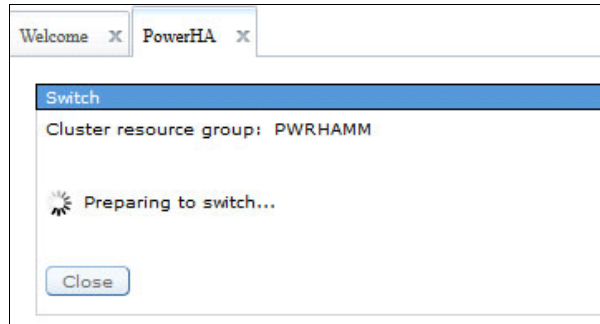

Figure 5-5   Preparing to switch

Figure 5-6 shows the window that opens when the switch completes. Click **Close**.
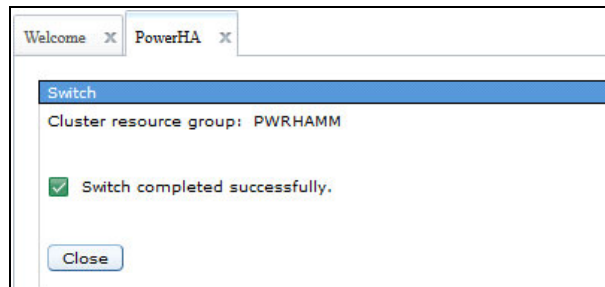

Figure 5-6   The switch completed successfully

9. You can see that the roles of the nodes changed as shown in Figure 5-7. The current primary (current production) node is now CTCIHA4B.
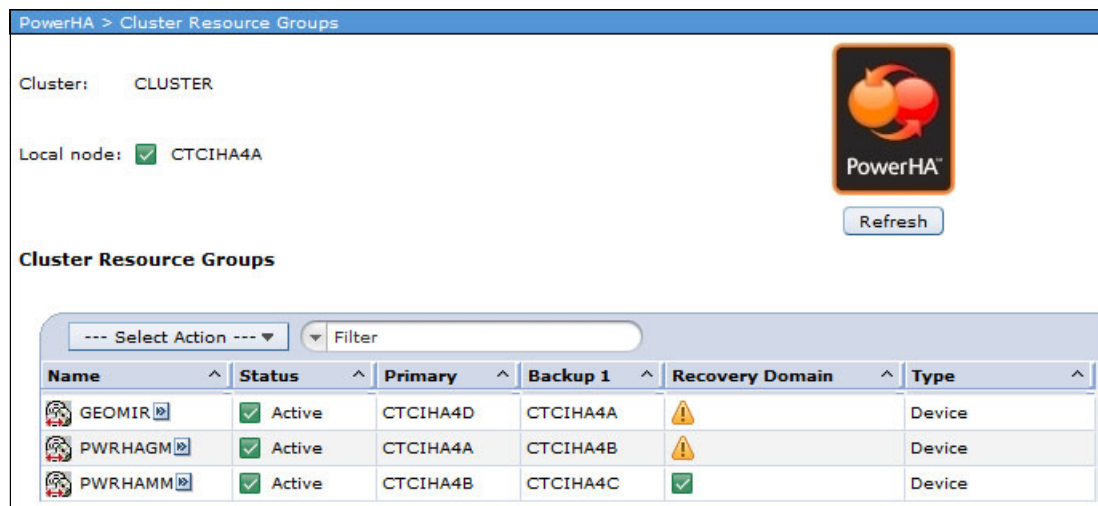

Figure 5-7   Cluster resource groups after the switch

5.1.1 Changing the node to switch to in a three-site configuration with the GUI

If you want to change the node that becomes the primary after a switch, run the following steps to change the cluster resource group (CRG) recovery domain:

1. On the PowerHA > Cluster Resource Groups window, click the arrow next to the CRG name and select **Recovery Domain** as shown in Figure 5-8.
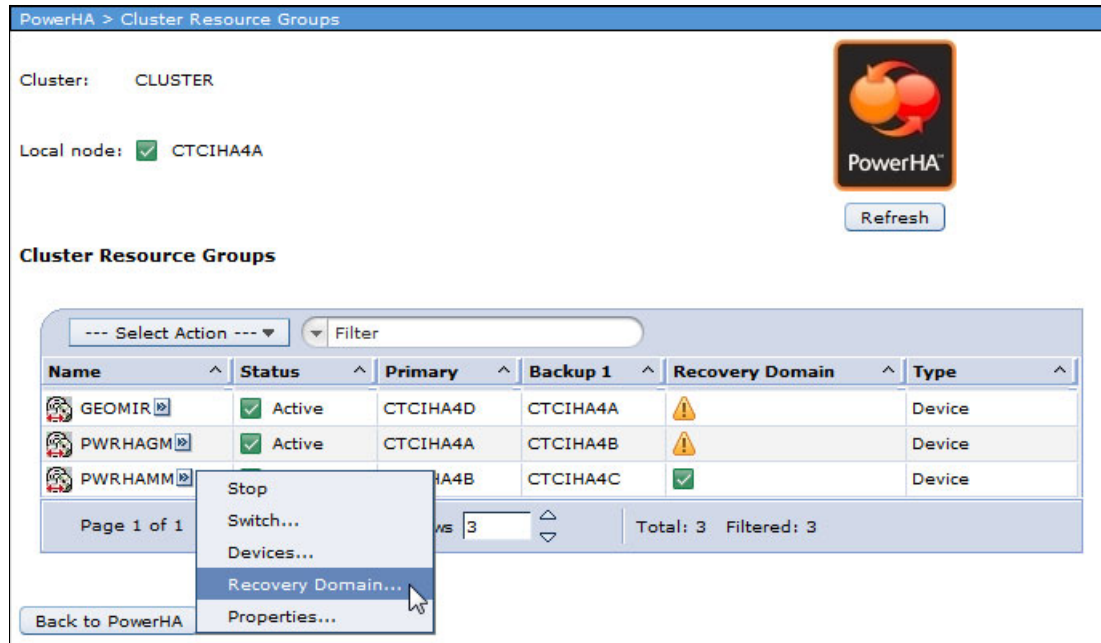

Figure 5-8   Changing the Recovery Domain configuration

2. In this example, the recovery domain has three nodes as shown in Figure 5-9. In this example, the current production (Primary) node is CTCIHA4B, and the first backup node is CTCIHA4C. If you want the next switch or failover to go to CTCIHA4A, you must change node roles. From the Select Action pull-down menu, select **Change Node Roles**.

See Figure 5-9.

Figure 5-9   Recovery Domain

3. Select **CTCIHA4C** and click **Move Up** so that CTCIHA4A becomes the first backup node as shown in Figure 5-10. Click **Save** to save your changes. Then, click **Back to Cluster Resource Groups**.
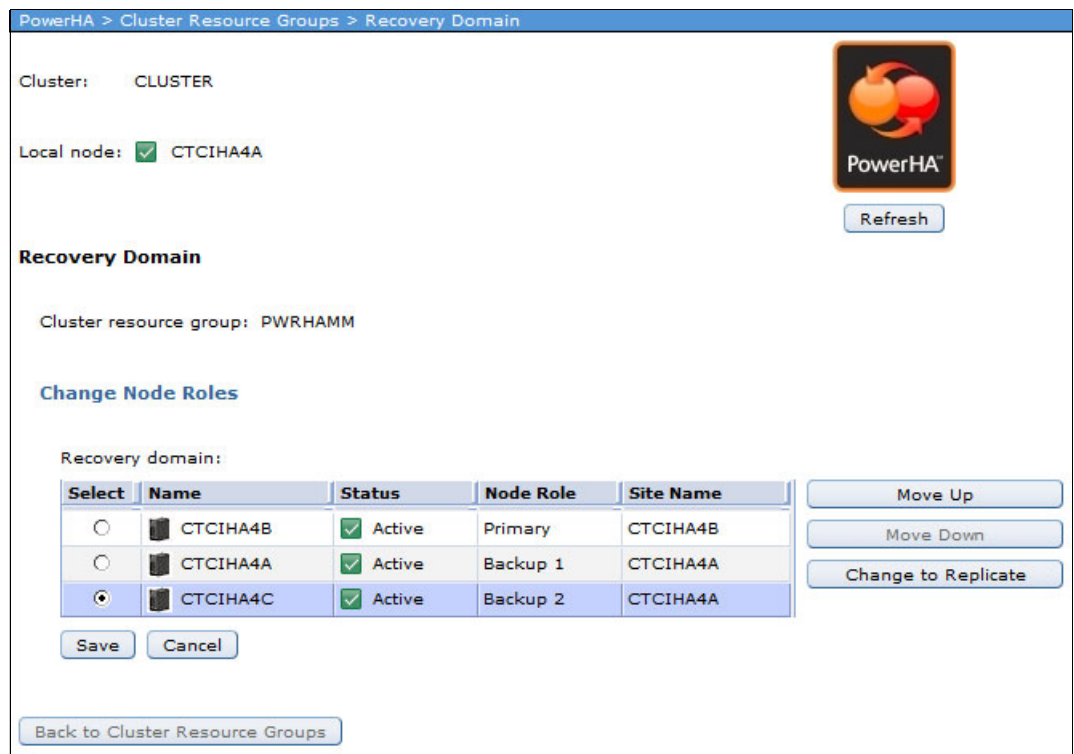

Figure 5-10   Changing a node's role

The switch process performs the following functions:

1. The process varies off the IASP on the current production cluster node if the IASP is not varied off already.

2. If a server takeover Internet Protocol (IP) address is defined in the CRG, this IP interface is ended on the current production cluster node.

3. The recovery domain information in the CRG is updated to show the current production node and current backup node.

4. If switchover reverse replication is configured with the default value of *YES in the auxiliary storage pool (ASP) session, the remote copy direction is switched on the DS8000 system.

5. If switchover reverse replication is configured as *NO in the ASP session, the remote copy is detached on the DS8000 system. When you want to restart the replication, you issue the following Change Auxiliary Storage Pool Session (`CHGASPSSN`) command:

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

6. On the current production node (Preferred Target location), the IASP is varied on if vary on is configured in the CRG device entry.

7. The current production node starts the server takeover IP address if a server takeover IP address is defined in the CRG.

8. If the CRG is set up to automatically vary on the IASP, the IASP is varied on.

You can check the status of the vary-on process with the Display ASP Vary Status
(**DSPASPSTS**) command on the current production system. During the vary-on process, several
steps are performed as shown in Figure 5-11.

```
                      Display ASP Vary Status

 ASP Device . . . . :    199 PWRHAMM      Current time . . . :   00:01:32
 ASP State  . . . . :    AVAILABLE        Previous time  . . :   00:00:00
 Step . . . . . . . :    35 / 35          Start date . . . . :   09/22/15


       Step                                              Elapsed time
       Cleaning up cross-reference files                   00:00:03
       Database access path recovery                       00:00:04
       Database cross-reference file merge                 00:00:46
       SPOOL initialization                                00:00:01
       Image catalog synchronization                       00:00:00
       Command analyzer recovery                           00:00:00
 >     Catalog validation                                  00:00:33

 Press Enter to continue


  F3=Exit    F5=Refresh    F6=Display UDFS    F7=Display STATFS    F12=Cancel
  F19=Automatic refresh
```

Figure 5-11   DSPASPSTS output

After the vary-on process finishes, the ASP state is AVAILABLE. If the IASP was not varied on
by the switch, you need to manually vary on the IASP and the IP addresses that are needed
for user access. You can then start your applications.

5.1.2 Changing the node to switch to in a three-site configuration with a command

The following steps show you how to switch in a three-site configuration by using control language (CL) commands:

1. In this example, if you want to switch back to CTCIHA4A, you must change the CRG by running the following Change Cluster Resource Group (**CHGCRG**) command:

   ```
   CHGCRG CLUSTER(CLUSTER) CRG(PWRHAMM) CRGTYPE(*DEV) RCYDMNACN(*CHGCUR)
   ```

   See Figure 5-12.

```
                    Change Cluster Resource Group (CHGCRG)

 Type choices, press Enter.

 Cluster  . . . . . . . . . . . . > CLUSTER       Name
 Cluster resource group . . . . . > PWRHAMM       Name
 Cluster resource group type  . . > *DEV          *DATA, *APP, *DEV, *PEER
 CRG exit program . . . . . . . .   *SAME         Name, *SAME, *NONE
   Library  . . . . . . . . . . .                 Name, *CURLIB
 Exit program format name . . . .   *SAME         *SAME, EXTP0100, EXTP0101...
 Exit program data  . . . . . . .   *SAME

 User profile . . . . . . . . . .   *SAME         Name, *SAME, *NONE
 Text description . . . . . . . .   *SAME

 Recovery domain action . . . . .   *CHGCUR       *SAME, *CHGPREFER, *CHGCUR

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

Figure 5-12   Changing the CRG

2. Press F9 to see additional parameters. Then, press Page Up and Page Down to see the recovery domain list. See Figure 5-13.

```
  Change Cluster Resource Group (CHGCRG)


  Type choices, press Enter.


  Recovery domain node list:
    Node identifier  . . . . . . .    CTCIHA4A       Name, *SAME
    Node role  . . . . . . . . . .    *PRIMARY       *SAME, *BACKUP, *PRIMARY...
    Backup sequence number . . . .    *SAME          1-127, *SAME, *LAST
    Site name  . . . . . . . . . .    CTCIHA4A       Name, *SAME, *NONE
    Data port IP address action  .    *SAME          *SAME, *ADD, *REMOVE
    Data port IP address . . . . .    *NONE


                   + for more values




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

Figure 5-13   Recovery domain list

You can change the roles in the Node role field. For the *BACKUP role, set the sequence number to 1 - first, 2 - second, and so on. Press Enter to save your changes. The next switch or failover will use these settings.

Note: Ensure that the current production node is specified as *PRIMARY and that all of the nodes in the recovery domain are specified with the correct values.

5.1.3 Special considerations when you switch in a Global Mirror setup

After you end your applications and vary off the IASP, ensure that all changes were synchronized between the source volumes and the target volumes.

You can ensure that all changes were synchronized between the source volumes and the target volumes by running the following steps in the PowerHA GUI:

1. From the left pane, select **PowerHA**. Then, on the main window, select **Independent ASPs**. Select **Details** for the IASP that you want to switch to the current backup system as shown in Figure 5-14.

Figure 5-14   IASP details option

2. From the Advanced actions drop-down menu (Figure 5-15), select **ASP Sessions and Copy Descriptions**.

Figure 5-15   ASP session and copy descriptions

3. For the selected IASP session, select **Properties** and check whether all data was synchronized between the Global Mirror source and the Global Mirror target. Alternatively, you can use the Display ASP Session (`DSPASPSSN`) command to see this information.

5.2 Failover after a production system outage

A *failover* occurs when the current production node failed and the backup node takes over. The default failover action/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 depends on previous testing and verification that business applications can run on the backup node and users can access the business applications on the backup node.

You can distinguish between the following failure scenarios:

▶ A primary node failure triggers an automatic failover. This scenario can occur either by a panic message that is sent to the current backup node or cluster monitors in place.

▶ A primary node or cluster communications failure results in node partition status.

Each scenario requires different failover and recovery actions, which are described in more detail in the following sections.

Primary node failure that triggers an automatic failover event

An unplanned automatic failover event is triggered for a switchable IASP in a CRG for a primary node failure event that is detected by cluster resource services in the following situations:

▶ By a panic message that was sent by the failing primary node, for example, for an action of ending a cluster node or powering down the system.

▶ By a power state change event that was sent by the Hardware Management Console (HMC) server for a partition failure to the registered IBM i cluster nodes when you use advanced node failure detection. For more information about advanced node failure detection, see the IBM Redbooks publication *IBM PowerHA SystemMirror for i: Preparation (Volume 1 of 4)*, SG24-8400.

For an unplanned automatic failover event, a CPABB02 inquiry message "`Cluster resource groups are failing over to node <nodename> (C / G)`" 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, the failover starts immediately without posting any message.

With the default cluster parameters FLVDFTACT=*PROCEED and FLVWAITTIM=*NOWAIT, an automatic failover is triggered. By setting a *failover wait time* with the FLVWAITTIM cluster parameter, which is either specified with a duration in minutes or *NOMAX, the user can respond to the CPABB02 inquiry message to either proceed with or cancel the failover. The *failover default action* (FLVDFTACT) parameter setting determines whether PowerHA either automatically proceeds or cancels the failover processing after the specified failover wait time expired without receiving a response for the inquiry message from the user.

When you use the default PowerHA session parameter setting FLVRVSREPL=*NO for a DS8000 ASP session, after a failover event, the session is only detached. Also, the remote copy relationships are not reversed so that the user can preserve the primary node IASP data for possible further failure analysis.

After the recovery of the preferred primary node, including the restart of its cluster services, the detached ASP session is reattached to resume remote copy replication and make the IASP highly available again.

> **Important:** Ensure that you reattach in the correct direction.

Run the following steps:

1. From the PowerHA GUI, click **Independent ASPs**. On the Independent ASPs window, from the Select Action pull-down menu, select **ASP Sessions and Copy Descriptions** as shown in Figure 5-16.
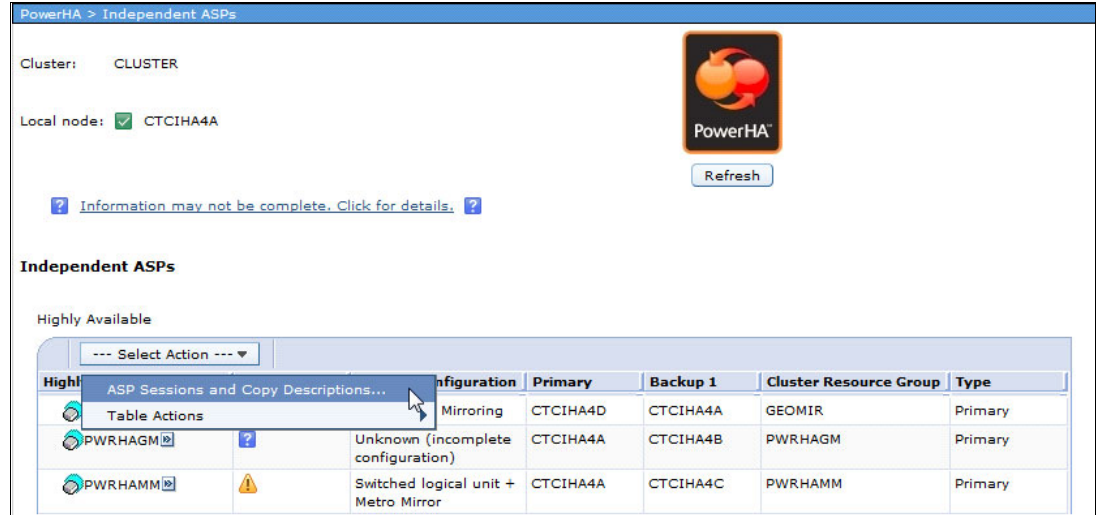

Figure 5-16   Working with ASP Sessions and Copy Descriptions

2. On the ASP Sessions and Copy Descriptions window, click the arrow next to the ASP session name and select **Properties** as shown in Figure 5-17.

Figure 5-17   ASP session properties

3. On the Properties window, from the Select Action pull-down menu, click **Reattach** as shown in Figure 5-18. Confirm your choice by clicking **YES** in the confirmation window that opens.
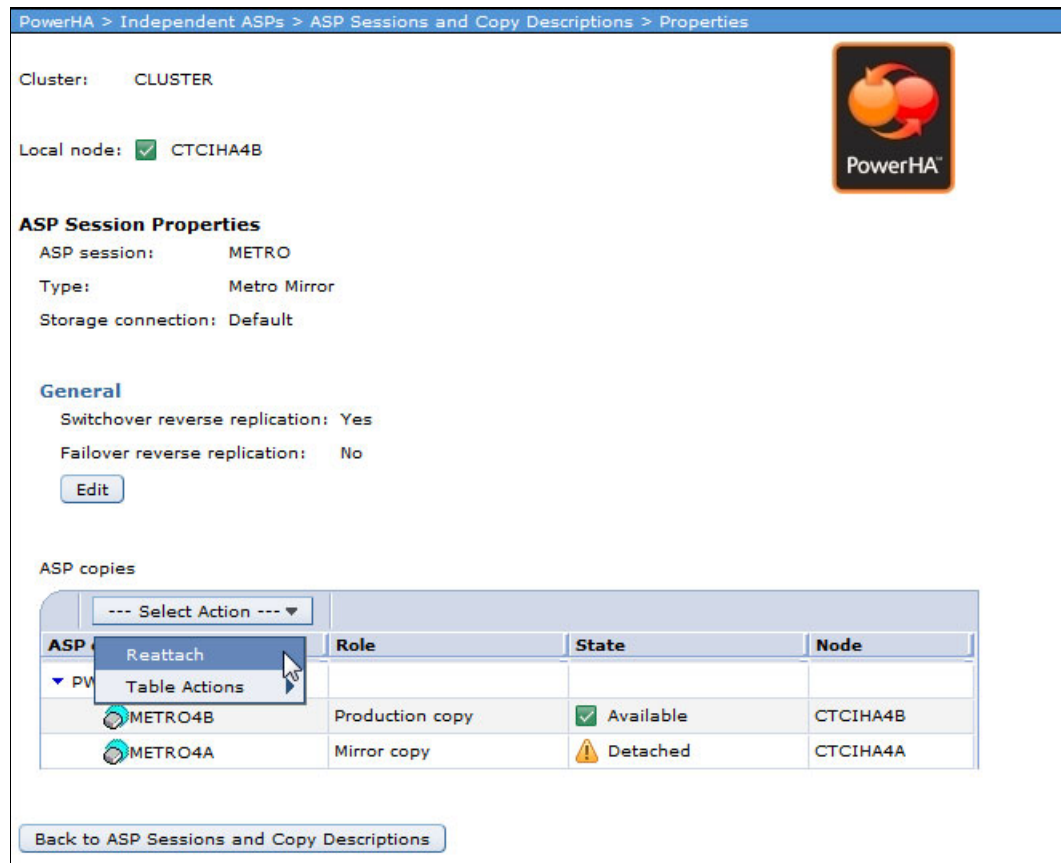

Figure 5-18   Reattaching replication

You can also reattach replication by using the following Change ASP Session (`CHGASPSSN`) command:

`CHGASPSSN SSN(SS_Name) OPTION(*REATTACH)`

After resynchronization finishes, a planned switch to the preferred primary node is then possible. It can be performed as described in 5.1, "Planned switching" on page 122 at the next planned role swap or at a convenient time.

Primary node failure without an automatic failover event

Even with advanced node failure detection, certain outages do not cause an automatic failure. In this case, the current backup system will put the nodes into a partitioned status. The cluster is not able to communicate between the partitions any more. The cluster cannot determine whether the inability to communicate between the partitions 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 when the current production system is still up and running.

If a partition status occurs, the message CPFBB47 "`Automatic failover not started for cluster resource group <crgname> in cluster <clustername>`" is sent to the QSYSOPR message queue on the current backup node. If this partition status occurs and you see this message, check whether the production partition experienced a failure.

If the production partition failed, you can change the status of the current primary node from PARTITIONED to FAILED by using the following Change Cluster Node Entry (`CHGCLUNODE`) command:

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

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

After the recovery of the preferred primary node, including a restart of its cluster services and of the device CRG, the ASP session is reattached on the primary node, which is the current backup node to resume replication.

Cluster node partition status

A primary node partition status can be triggered by a sudden primary node failure or cluster communication error. The use of advanced node failure detection, as described in the IBM Redbooks publication *IBM PowerHA SystemMirror for i: Preparation (Volume 1 of 4)*, SG24-8400, can help you to avoid cluster partition conditions for many cases of sudden node failures. The HMC or Virtual I/O Server (VIOS) can detect the failure and notify the registered IBM i cluster nodes to invoke an automatic failover.

However, certain cases, such as cluster heartbeat communication errors or sudden system power outages, can lead to a cluster node partition condition. The user needs to determine the reason for the partition condition and decide whether to fix it or change the production node to a failed status to invoke a cluster failover to a backup node.

A cluster partition condition with a failed cluster node, which is indicated by the message ID CPFBB20, requires manual actions to invoke a cluster failover. For example, you can set the primary cluster node from the *partition* status to the *failed* status by using the following Change Cluster Node Entry (`CHGCLUNODE`) command:

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

After this `CHGCLUNODE` command runs, the cluster node roles are changed and remote copy relationships are stopped so that the IASP can be varied on manually at the new primary node on the backup site.

After the recovery of the preferred primary node, including restarting its cluster services and the device CRG, you can perform a switchback to the preferred primary node by using the Change Cluster Resource Group Primary (`CHGCRGPRI`) command.

5.3 Scheduled or unscheduled switching with IASP Copy Services Manager

The IASP Copy Services Manager (ICSM) process for a role swap differs from the PowerHA process. ICSM is always a switchover, and it is run from the current backup node.

The Switch PPRC (`SWPPRC`) command is shown in Figure 5-19.

```
                       Switch PPRC (SWPPRC)

 Type choices, press Enter.

 Environment name . . . . . . . .   iaspname     Name
 Switch type  . . . . . . . . . .   *SCHEDULED   *SCHEDULED, *UNSCHEDULED...
 Type . . . . . . . . . . . . . .   *MMIR        *, *GMIR, *LUN, *MMIR...
 Auto Vary On . . . . . . . . . .   *YES         *YES, *NO
 Auto replicate . . . . . . . . .   *DFT         *DFT, *YES, *NO
 Switch paused MMIR . . . . . . .   *NO          *YES, *NO




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

Figure 5-19   SWPPRC command

Figure 5-19 shows a parameter for switch type. This parameter can be *SCHEDULED, which is similar to a PowerHA planned switch. This value requires that the whole environment is operational. Alternatively, this value can be *UNSCHEDULED, which is similar to a PowerHA failover. In either case, a message is sent to the system operator for permission to continue the switch.

Other parameters can be set if you want to vary on the IASP after the switch, start replication after the switch, and switch even if the replication was paused at the time.

> **Important:** If you switch while the replication is suspended for any reason, you will lose data unless the IASP was varied off before the replication was suspended and it was not used before you performed the switch.

5.4 Working with PowerHA FlashCopy

When you work with FlashCopy in a PowerHA environment, you must understand what is included in the copy and what is not. FlashCopy creates a point-in-time copy of logical disk units. Therefore, unless you take the correct steps, any data that still resides in main memory and that is not committed to disk is not copied to the FlashCopy target logical unit numbers (LUNs).

IBM i provides a *quiesce* function that can suspend database transactions and database and integrated file system (IFS) file change operations for the system and configured basic auxiliary storage pools (ASPs) or independent ASPs (IASPs). Transactions that were started are allowed to continue until they reach a commit boundary, subject to a time limit. The initiation of new transactions or new file (database or IFS) operations is temporarily halted. In addition, changed pages of memory are committed to disk.

From a user application perspective, the system seems to respond more slowly than expected during the suspended time. Application transactions will not break unless the application uses timeouts and that time is exceeded.

The Change ASP Activity (`CHGASPACT`) command is shown in Figure 5-20 with the following parameters:

► For the ASP device (ASPDEV) parameter, specify the name of the IASP.

► The Option parameter *SUSPEND starts the suspend process. A CPCB717 message (see Figure 5-21 on page 141) is sent to the QSYSOPR message queue when the suspend is achieved. This message also contains data about how long it took to achieve the quiesce point. If it is difficult to achieve a suspend point in your environment, you can use the *FRCWRT parameter instead. The *FRCWRT parameter does not attempt to reach a quiesce point. The final option is *RESUME, which is run after you take the FlashCopy to release the quiesce and allow normal operations to continue. A CPCB718 message (see Figure 5-22 on page 142) is sent to the QSYSOPR queue when normal operations resume.

► The Suspend timeout (SSPTIMO) parameter specifies the amount of time in seconds that the command is allowed to take before it completes.

► The Suspend timeout action (SSPTIMOACN) parameter defines the action that needs to be taken if a full suspend of the ASP was not reached in the specified amount of time. *CONT allows the suspend to keep going in the background. *END cancels the suspend and resumes normal operations.

```
                       Change ASP Activity (CHGASPACT)

 Type choices, press Enter.

 ASP device . . . . . . . . . . .  ASPDEV       > PWRHAMM
 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 5-20   CHGASPACT command

Figure 5-21 shows a CPCB717 message.

```
                    Additional Message Information

 Message ID . . . . . . :   CPCB717      Severity . . . . . . . :   00
 Message type . . . . . :   Completion
 Date sent  . . . . . . :   09/25/15     Time sent  . . . . . . :   11:32:56


 Message . . . . :   Access to ASP GEOMIR is suspended.
 Cause . . . . . :   Access to ASP GEOMIR is suspended.  New transactions will
   not be allowed to start until access is resumed.  Transaction quiescing used
   0 of 60 available seconds.  The reason code is 0.
     Possible reason codes are:
     0 -- Transaction quiescing completed successfully within the specified
   time.
     -1 -- Transaction quiescing did not complete in the specified time.  A
   larger timeout value is required to allow existing transactions to complete.
 Recovery  . . . :   None.


                                                                      Bottom
 Press Enter to continue.


 F3=Exit   F6=Print   F9=Display message details   F12=Cancel
 F21=Select assistance level
```

*Figure 5-21   Message that is sent to QSYSOPR when CHGASPACT *SUSPEND completes*

If you want to know the total amount of time that your applications were suspended during your FlashCopy process, follow these steps:

1. Obtain the time stamp of the CPCB718 message on QSYSOPR (11:33:02 in Figure 5-22 on page 142).

2. Obtain the time stamp of the CPCB717 message on QSYSOPR (11:32:56 in Figure 5-21).

3. Obtain the number of seconds that were used for transaction quiescing in the CPCB717 message (0 in Figure 5-21).

4. Subtract the two time stamps from each other (6 seconds in this example).

5. Add the number of seconds to achieve the quiesce (0 in this example, which is a total of 6 seconds of suspension of service).

See the CPCB718 message in Figure 5-22.

```
                        Additional Message Information

 Message ID . . . . . . :   CPCB718        Severity . . . . . . . :   00
 Message type . . . . . :   Completion
 Date sent  . . . . . . :   09/25/15       Time sent  . . . . . . :   11:33:02


 Message . . . . :    Access to ASP GEOMIR successfully resumed.
 Cause . . . . . :    Access to ASP GEOMIR is successfully resumed.  Users may
   access data within the ASP.



Press Enter to continue.


 F3=Exit    F6=Print    F9=Display message details    F12=Cancel
 F21=Select assistance level
```

*Figure 5-22   Message that is sent when CHGASPACT *RESUME completes*

You also need to understand the consequences of taking a FlashCopy from an IASP that was not quiesced. The FlashCopy operation finishes successfully, which it will if you do not use the **CHGASPACT** command at all because that FlashCopy occurs on the storage system, and the storage system is unaware of the interdependencies of activities 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 did not reach a commit boundary are identified, 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.

> **Important:** Restoring this data and then running an Apply Journal Changes (**APYJRNCHG**) command by using journal entries from your production system to reach a current data state involve significant manual work and might be difficult. You must identify the files for which a rollback took place during the database recovery. Then, you must find the individual starting point (last commit point) for each of these files manually from the journal entries. Then, start the **APYJRNCHG** command for each file with a different starting point.
>
> The chances are that you will end up with inconsistent data when you try to restore data this way.

To make a FlashCopy of your IASP, you need to run the following steps:

1. On the production node, either vary off the IASP or use the **CHGASPACT** command to quiesce the database.

2. Run the Start ASP Session (**STRASPSSN**) command (Figure 5-23 on page 143) to start the FlashCopy process. Ensure that you specify the correct source and target copy descriptions. You can also decide whether the IASP will be fully copied or only changes will be copied and whether the relationship will persist after a copy completes.

3. As soon as the **STRASPSSN** command completes, you must issue the **CHGASPACT *RESUME** command, if needed.

4. On the FlashCopy node, now vary on the IASP and perform the backup or other processing that you want to perform.

See Figure 5-23 for the **STRASPSSN** command.

```
                        Start ASP Session (STRASPSSN)

 Type choices, press Enter.

 Session  . . . . . . . . . . . . > FLASHCOPY     Name
 Session type . . . . . . . . . . > *FLASHCOPY    *GEOMIR, *METROMIR...
 ASP copy:
   Preferred source . . . . . . . > SOURCECOPY    Name
   Preferred target . . . . . . . > TARGETCOPY    Name
               + for more values
 Device domain  . . . . . . . . .   *             Name, *
 FlashCopy type . . . . . . . . .   *NOCOPY       *COPY, *NOCOPY
 Persistent relationship  . . . .   *NO           *YES, *NO




                                                                      Bottom
 F3=Exit   F4=Prompt   F5=Refresh   F10=Additional parameters   F12=Cancel
 F13=How to use this display       F24=More keys
```

Figure 5-23   STRASPSSN command

5.5 Integration between PowerHA and BRMS

When you perform a save from a FlashCopy, the original data and the original system do not know that a save occurred. Therefore, for example, the Backup, Recovery, and Media Services (BRMS) recovery reports will not show the correct instructions for restoring data.

However, BRMS was enhanced to support a FlashCopy environment. By using the correct setup, the original system "thinks" that it performed the save operations itself and the recovery reports on the original system reflect that a save was performed from the FlashCopy. By using the Advanced feature of BRMS, you can even set the save timestamp to the time that the FlashCopy was taken instead of the time that the actual save occurred.

5.5.1 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. You are required to install the BRMS network feature. Run the following steps:

1. Log on to the IBM i partition and open the BRMS policy menu by using the following command:

   WRKPCYBRM *SYS

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

3. On the Change Network Group panel (Figure 5-24), change the communication method to
   *IP and add your IBM i partitions as nodes to the BRMS network.

   Note: The host name must be reachable in the network. Add a host table entry, if
   necessary.

```
                           Change Network Group                    CTCIHA4A TEST

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     CTCIHA4C     TEST


                                                                         Bottom

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

Figure 5-24   Adding node CTCIHA4C to the BRMS network

4. Import the BRMS database information from the leading BRMS system to the secondary
   partitions by using the following command (ignore all messages):

   INZBRM OPTION(*NETSYS) FROMSYS(TEST.CTCIHA4C)

5. After you add the system to the BRMS network, you can share the information about the
   backup between all partitions in the network.

User-defined IASP time stamps

When you use FlashCopy as the basis for your save operations, you must be aware that the
actual save time differs from the time that the FlashCopy was taken (and that represents the
"real" save time from a data perspective). For example, if the FlashCopy occurred at 1:00 PM
but the backups were not performed until 3:00 PM, BRMS by default uses 3:00 PM as the
time of the backup and the reference point for future incremental backups. However, the
FlashCopy time reflects when the IASP data was last changed.

The following interfaces allow users to define their own FlashCopy timestamps for specific IASPs. These commands are issued on the system where the actual save operation takes place:

► To add or update a timestamp:

```
CALL QBRM/Q1AOLD PARM('FLASHTIME' '*ENABLE' 'IASPNAME' 'FILESYSTEM TYPE'
'1071213143302')
```

IASPNAME is the name of the IASP. FILESYSTEM TYPE is either *QSYS or *IFS, and 1071213143302 is the timestamp in *CYYMMDDHHMMSS* format.

For example, if you want a timestamp for IASP ASP45 for both *IFS and *QSYS types, you need to use *ENABLE twice: one time for *QSYS and one time for *IFS. These timestamps 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, issue the following command:

```
CALL QBRM/Q1AOLD PARM('FLASHTIME' '*DISPLAY')
```

When the configurations are displayed, two timestamps are shown. The first is the user-specified timestamp, and the second timestamp is used by BRMS internally.

► To remove a configuration for a specific IASP, issue the following command:

```
CALL QBRM/Q1AOLD PARM('FLASHTIME' '*DISABLE' 'IASPNAME' 'FILESYSTEM TYPE')
```

► To remove all of the user-defined timestamp configurations, issue the following command:

```
CALL QBRM/Q1AOLD PARM('FLASHTIME' '*DISABLE' '*ALL')
```

Internally, BRMS requires a unique timestamp in certain places. In those cases, BRMS generates a unique timestamp that is incremented from the user-defined timestamp. Save While Active sync points and incremental reference timestamps always use the value that is defined by the user. BRMS maintains a list of the timestamps that are used and clears this list when the Start Maintenance for BRMS (**STRMNTBRM**) command with the parameter RUNCLNUP(*YES) is specified.

Important: The BRMS Advanced Feature (57xxBR1 option 1 (Feature Code (FC) 5102)) is required on the system where the FLASHTIME interface is used:

► User-defined timestamps are not used for archival backups.

► User-defined timestamps are not used for *LINK from IASP *ALLAVL or *CURASPGRP.

► Any SAVBRM or *LINK control group entry, which saves from multiple IASPs, does not use user-defined timestamps.

► 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, we suggest that different user-defined timestamps are used for each IASP entry of the *QSYS type.

► Do not use timestamps that are later than the current system date and time.

5.5.2 Considerations when you use FlashCopy on a Global Mirror target

A FlashCopy of a Global Mirror with Change Volumes target can be up to two cycle periods behind the status of the source copy, depending on when the FlashCopy or detach is performed in relationship to the cycle time.

The calculation is the same calculation to calculate the maximum recovery point objective (RPO). Even if a vary off is performed before the detach, to ensure an up-to-date detached copy, wait two cycle periods after the vary off before you complete the detach. This approach assumes that your cycle time is set to a value that corresponds to your data change rate and your available bandwidth.

5.6 Using FlashCopy when you change your application

When you perform larger changes to data in your IASP, the use of FlashCopy can provide an attractive way to create a copy of the status before you perform the changes. This approach provides a quick way to get 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 you perform major changes, you want to ensure 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 5-25. The session type must be *FLASHCOPY. Ensure that you specify the ASP copy description that describes the current production copy of your IASP as the preferred source. The preferred target is the ASP copy description of the FlashCopy ASP.

```
                    Start ASP Session (STRASPSSN)

  Type choices, press Enter.

  Session  . . . . . . . . . . .    FLASH        Name
  Session type . . . . . . . . . .               *GEOMIR, *METROMIR...
  ASP copy:
    Preferred source . . . . . . .  METRO4A      Name
    Preferred target . . . . . . .  FLASH        Name
              + for more values
  Device domain  . . . . . . . . .  *            Name, *



 F3=Exit   F4=Prompt   F5=Refresh   F10=Additional parameters   F12=Cancel
  F13=How to use this display        F24=More keys
```

Figure 5-25   Start FlashCopy

3. Vary on the IASP on your FlashCopy partition and ensure that the data is usable.

4. Vary on the production copy of your IASP again and perform the changes. If you discover that for any reason you need to revert to the old status, perform these steps:

> **Important:** These steps reverse *all* changes that you made to the production IASP data after you created the FlashCopy. This reversal includes any normal production work that occurred after you performed the FlashCopy.

   a. Vary off the production IASP.

   b. Suspend the ASP session between the current production and the current backup:

   ```
   CHGASPSSN SSN(FLASH) OPTION(*SUSPEND)
   ```

   c. Reverse the FlashCopy. This reversal copies the data from the FlashCopy point in time over the current data of the IASP, eliminating *all* changes that were made to the data in the IASP after the FlashCopy was taken:

   ```
   CHGASPSSN SSN(FLASH) OPTION(*REVERSE)
   ```

   d. Vary on the production IASP and resume the ASP session between the current production and the current backup:

   ```
   CHGASPSSN SSN(FLASH) OPTION(*RESUME)
   ```

5.7 Working with Full System FlashCopy

When you use Full System FlashCopy, you need to understand the restrictions in this setup. The only way to ensure a 100% save and a consistent copy of a full system is to power off the source system. Performing a quiesce of the source system by using the Change ASP Copy Description (`CHGASPACT`) command suspends only database activities and file and IFS object change operations.

While you take a warm FlashCopy with the full system online, even when the full system is 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.

Another consideration of Full System FlashCopy is the chance that when you take a warm FlashCopy, certain IBM i internal objects might be corrupted at IPL time. These objects might be corrupted even when you successfully quiesce the system ASP by using the Change ASP Activity (`CHGASPACT`) command. This corruption might lead to a situation where the backup partition cannot successfully finish the IPL. You might need a Serial Line Internet Protocol (SLIP) installation, or another FlashCopy might be needed for a recovery.

5.8 Full System Copy Services Manager (FSCSM)

You can use Full System Copy Services Manager (FSCSM) to make a FlashCopy of a whole system, and it does not require an IASP. It also brings additional automation to run your backup. If you use BRMS for your backups, FSCSM has integration to bring the BRMS information back to your production system.

The FlashCopy is started from a controlling system by using the Make Full System Copy (`MKSYSCPY`) command as shown in Figure 5-26.

```
                      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 5-26   Full System FlashCopy

The `MKSYSCPY` command starts the Full System FlashCopy and performs the backup commands that are configured for the environment. The `MKSYSCPY` command performs the following steps:

1. Prepare the production partition for the FlashCopy process.
2. Power off or quiesce the specified partition.
3. Perform the FlashCopy of the partition LUNs.
4. IPL or release the quiesce of the production partition.
5. Activate the target partition.
6. Perform the backups.
7. If BRMS is used for backups, transfer the backup information to the production partition.
8. Optional: Power off the FlashCopy partition.

5.9 HyperSwap

HyperSwap operations are described. The following topics are covered:

5.9.1 User-initiated swap

A HyperSwap swap can be initiated manually. Figure 5-27 shows a HyperSwap environment.
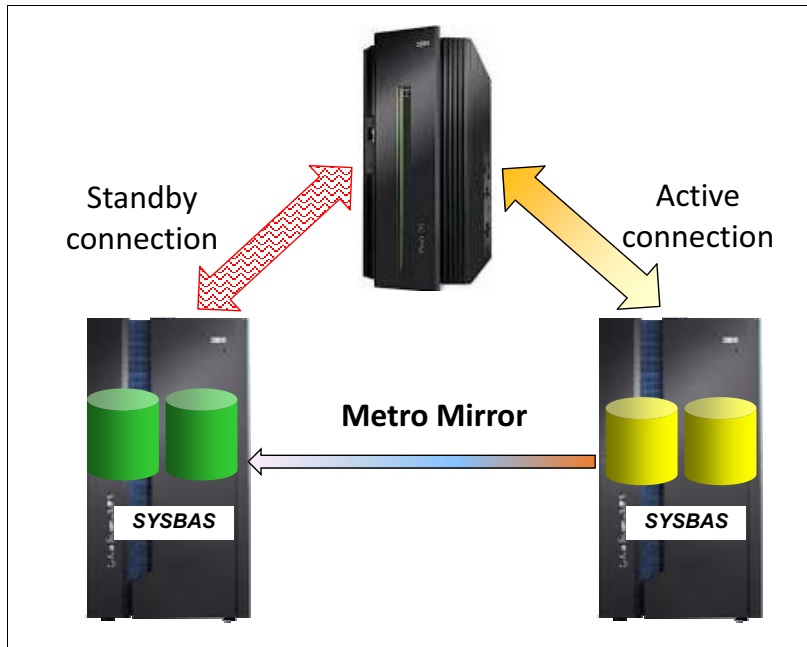

Figure 5-27   HyperSwap environment

The HyperSwap operation is performed by issuing the Change HyperSwap Status (**CHGHYSSTS**) command:

```
CHGHYSSTS OPTION(*SWAP)
```

The switch operation is fast (less than 1 second), and the replication is reversed automatically.

If you look at the HyperSwap status by using the Display HyperSwap Status (`DSPHYSSTS`) command before the swap (see Figure 5-28), you can see the replication direction before the switch.

```
                       Display HyperSwap Status                    CTCHYPER
                                                          09/25/15  14:27:50
                Volume                    Copy              Volume
 Resource       Identifier                Status  Resource  Identifier
 DMP083         IBM.2107-75XA511/8800     --->    DMP113    IBM.2107-75ZF751/8800
 DMP085         IBM.2107-75XA511/8801     --->    DMP114    IBM.2107-75ZF751/8801
 DMP086         IBM.2107-75XA511/8802     --->    DMP115    IBM.2107-75ZF751/8802
 DMP088         IBM.2107-75XA511/8900     --->    DMP116    IBM.2107-75ZF751/8900
 DMP090         IBM.2107-75XA511/8901     --->    DMP117    IBM.2107-75ZF751/8901
 DMP092         IBM.2107-75XA511/8902     --->    DMP118    IBM.2107-75ZF751/8902




                                                                     Bottom
 Press Enter to continue


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

Figure 5-28   HyperSwap copy status before a switch

If you run the `DSPHYSSTS` command again after the swap, you see that the direction is reversed now (Figure 5-29).

```
                       Display HyperSwap Status                    CTCHYPER
                                                          09/25/15  14:29:41
                Volume                    Copy              Volume
 Resource       Identifier                Status  Resource  Identifier
 DMP083         IBM.2107-75XA511/8800     <---    DMP113    IBM.2107-75ZF751/8800
 DMP085         IBM.2107-75XA511/8801     <---    DMP114    IBM.2107-75ZF751/8801
 DMP086         IBM.2107-75XA511/8802     <---    DMP115    IBM.2107-75ZF751/8802
 DMP088         IBM.2107-75XA511/8900     <---    DMP116    IBM.2107-75ZF751/8900
 DMP090         IBM.2107-75XA511/8901     <---    DMP117    IBM.2107-75ZF751/8901
 DMP092         IBM.2107-75XA511/8902     <---    DMP118    IBM.2107-75ZF751/8902




                                                                     Bottom
 Press Enter to continue


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

Figure 5-29   HyperSwap status after a switch

HyperSwap does not retain a preferred direction. If you IPL the system in the state that is shown in Figure 5-29 on page 150, after the IPL, you will see that the storage unit that is listed on the left of the panel changed to become the last used storage unit as shown in Figure 5-30. Figure 5-30 shows this same environment after an IPL.

> **Important:** Always pay attention to which storage unit is the current source and ensure that you know the storage unit serial numbers.

```
                     Display HyperSwap Status                      CTCHYPER
                                                        09/25/15  15:34:10
                Volume                  Copy             Volume
  Resource      Identifier              Status  Resource Identifier
  DMP083        IBM.2107-75ZF751/8800   --->    DMP113   IBM.2107-75XA511/8800
  DMP085        IBM.2107-75ZF751/8801   --->    DMP114   IBM.2107-75XA511/8801
  DMP086        IBM.2107-75ZF751/8802   --->    DMP115   IBM.2107-75XA511/8802
  DMP088        IBM.2107-75ZF751/8900   --->    DMP116   IBM.2107-75XA511/8900
  DMP090        IBM.2107-75ZF751/8901   --->    DMP117   IBM.2107-75XA511/8901
  DMP092        IBM.2107-75ZF751/8902   --->    DMP118   IBM.2107-75XA511/8902




                                                                    Bottom
  Press Enter to continue


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

Figure 5-30   HyperSwap status after an IPL

5.9.2 Automatic swap

Any of the following situations can cause an automatic swap:

► Failure of the DS8000 storage unit
► Host connection failure to DS8000
► Configuration issue or mistake

See Figure 5-31.
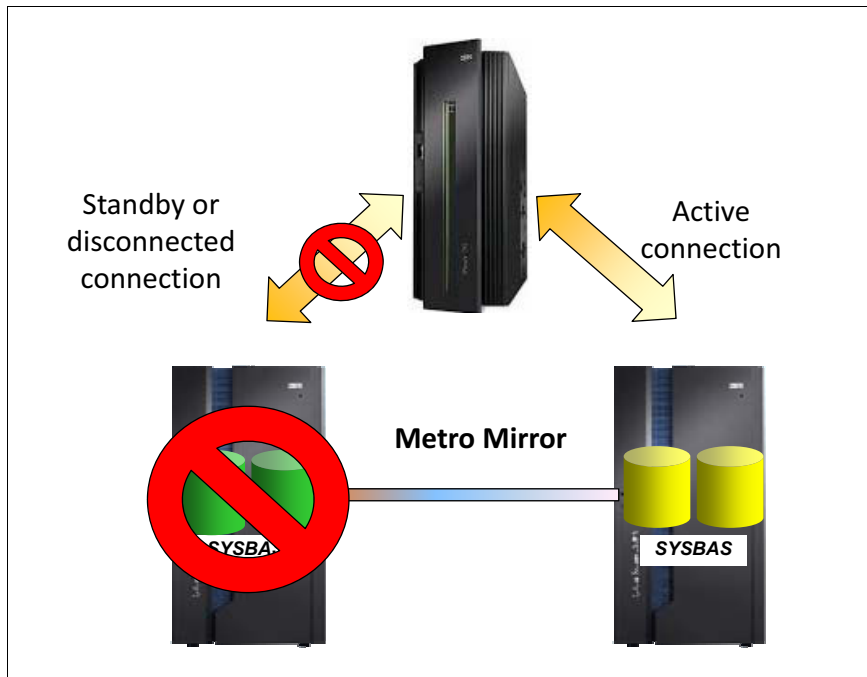
Figure 5-31   Failed connections in a HyperSwap environment

The failover process is automatic and transparent to the operating system (performed within milliseconds).

If the storage configuration is changed, such as removing a disk, the failover process can take longer. IBM i knows when configuration changes are performed in the DS8000. If replication is suspended, it must be restarted manually. QSYSOPR reports when a HyperSwap swap occurs with message CPI09B0.

When Metro Mirror is not active, the target disk will be in an Ineligible status. See Figure 5-32.

```
                        Display Disk Path Status

        Serial                          Resource   Path
ASP Unit    Number       Type Model Name          Status
 1   1      50-880024A    2107 A07  DMP007    Active
            50-880024A    2107 A07  DMP008    Active
            50-8800622    2107 A07  DMP087    Ineligible
            50-8800622    2107 A07  DMP001    Ineligible
 1   2      50-890024A    2107 A07  DMP013    Active
            50-890024A    2107 A07  DMP014    Active
            50-8900622    2107 A07  DMP090    Ineligible
            50-8900622    2107 A07  DMP004    Ineligible
 1   3      50-890124A    2107 A07  DMP015    Active
            50-890124A    2107 A07  DMP016    Active
            50-8901622    2107 A07  DMP091    Ineligible
            50-8901622    2107 A07  DMP005    Ineligible
 1   4      50-890124A    2107 A07  DMP016    Active
            50-890124A    2107 A07  DMP017    Active
                                                              More...
```

Figure 5-32   HyperSwap, working with disks in system service tools (SST) after a failure

If all disk connections are lost, the lost connections are displayed in the `DSPHYSSTS` command as shown in Figure 5-33.

```
Display HyperSwap Status           CTCHYPER
                                   08/27/14 14:35:00
          Volume            Copy         Volume
Resource  Identifier        Status Resource    Identifier
DMP087*   IBM.2107-75ZF751/8800   <-\-  DMP007   IBM.2107-75XA511/8800
DMP088*   IBM.2107-75ZF751/8801   <-\-  DMP010   IBM.2107-75XA511/8801
DMP089*   IBM.2107-75ZF751/8802   <-\-  DMP011   IBM.2107-75XA511/8802
DMP090*   IBM.2107-75ZF751/8900   <-\-  DMP013   IBM.2107-75XA511/8900
DMP091*   IBM.2107-75ZF751/8901   <-\-  DMP015   IBM.2107-75XA511/8901
DMP092*   IBM.2107-75ZF751/8902   <-\-  DMP017   IBM.2107-75XA511/8902
```

Figure 5-33   HyperSwap status that shows the lost connection to all disks

Figure 5-34 shows the results that are displayed in `DSPHYSSTS` with a single lost connection to disk.

```
Display HyperSwap Status          CTCHYPER
                                  08/27/14  14:35:00
        Volume          Copy          Volume
Resource  Identifier       Status Resource   Identifier
DMP087*   IBM.2107-75ZF751/8800  <-\-  DMP007    IBM.2107-75XA511/8800
DMP088    IBM.2107-75ZF751/8801  XXXX  DMP010    IBM.2107-75XA511/8801
DMP089    IBM.2107-75ZF751/8802  XXXX  DMP011    IBM.2107-75XA511/8802
DMP090    IBM.2107-75ZF751/8900  XXXX  DMP013    IBM.2107-75XA511/8900
DMP091    IBM.2107-75ZF751/8901  XXXX  DMP015    IBM.2107-75XA511/8901
DMP092    IBM.2107-75ZF751/8902  XXXX  DMP017    IBM.2107-75XA511/8902
```

Figure 5-34   HyperSwap status that shows the lost connection to one disk

If you need to resume replication in a HyperSwap environment, for example after a failure, you can use the Change HyperSwap Status (`CHGHYSSTS`) command as shown in Figure 5-35. This command starts Peer-to-Peer Remote Copy (PPRC) replication transparently. The replication is started in the direction that is stored in the System Licensed Internal Code (SLIC) (that `DSPHYSSTS` does not show).

> **Important:** Use of the data storage command-line interface (DSCLI) to start the replication in the wrong direction will result in loss of access to the disk.

```
                    Change HyperSwap Status (CHGHYSSTS)

 Type choices, press Enter.

 Option . . . . . . . . . . . . . *START        *STOP, *START, *SWAP
```

Figure 5-35   Starting replication in a HyperSwap environment

5.10 Full System replication switch

Full System replication (FSR) is switched with the Switch Copy Services Environment (`SWCSE`) command that is shown in Figure 5-36.

```
                    Switch Copy Services Environment (SWCSE)

 Type choices, press Enter.

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




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

Figure 5-36   SWCSE command

The `SWCSE` command performs the following functions:

► If the production node is running, send an *INQ message (IAS0029) to the QSYSOPR message queue that requests permission to proceed.

► Call a client-specified power off command.

► Wait until the production node partition is deactivated.

► Switch the replication direction between the storage units.

► IPL the disaster recovery (DR) (new production) node.

► Call the client's startup program:

  – `FIXSTRPRSC` processes the information for the *CMN, *STG, *BRM, and *RTE differences that are entered with the `WRKSTRPRSC` command.

  – Change the license keys for other programs that depend on the system serial number.

  – Perform other client-specified processing.

► Start clustering.

Troubleshooting

For a complete description of IBM PowerHA for i troubleshooting procedures, see Chapter 6, "Troubleshooting and collecting data for problem determination," in the IBM Redbooks publication *IBM PowerHA SystemMirror for i: Preparation (Volume 1 of 4)*, SG24-8400.

Support contact: If you encounter a 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:

https://ibm.biz/BdRCLT

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 risks and improve resiliency for critical applications.
► Simplifies setup and automation of HA, DR, and backup solutions.
► Reduces the cost of maintaining and regularly testing an HA/DR environment.
► Facilitates flexible deployment options for single or multiple 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 a PowerHA installation and ongoing operations for independent auxiliary storage pool (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 IASP-level vary on and no system IPL)					
FlashCopy	Automates FlashCopy of the 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	
Logical unit number (LUN)-level switching	Simplifies the 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 the initial deployment and automates ongoing server and storage management of a two-site Metro Mirror or Global Mirror HA or DR solution. 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 functionality to provide a three-site server or storage replication solution that contains 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 (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 the backup window for online daily backups. Enables an entry solution while you plan 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. 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 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.

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 associated with this book is available in softcopy on the Internet from the IBM Redbooks web server. Point your web browser at this website:

`ftp://www.redbooks.ibm.com/redbooks/SG248403`

Alternatively, you can go to the IBM Redbooks website:

`ibm.com/redbooks`

Select **Additional materials** and open the directory that corresponds with the IBM Redbooks form number, SG248403.

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 cluster resource group status
MONIASPSTS.mbr	Sample CL program to verify the IASP utilization status

Related publications

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

IBM Redbooks

The following IBM Redbooks publications provide additional information about the topic in this document. Note that some publications referenced in this list might be available in softcopy only.

► *IBM PowerHA SystemMirror for i: Preparation (Volume 1 of 4)*, SG24-8400
► *IBM PowerHA SystemMirror for i: Using IBM Storwize (Volume 3 of 4)*, SG24-8402
► *IBM PowerHA SystemMirror for i: Using Geographic Mirroring (Volume 4 of 4)*, SG24-8401
► *DS8000 Copy Services for IBM i with VIOS*, REDP-4584
► *IBM i and IBM System Storage: A Guide to Implementing External Disks on IBM i*, SG24-7120
► *IBM System Storage DS8000: Architecture and Implementation*, SG24-8886

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

These publications are also relevant as further information sources:

► *Multipath Subsystem Device Driver User's Guide*:

https://ibm.biz/Bd4uNN

► *IBM System Storage DS Command-Line Interface User's Guide*:

https://ibm.biz/BdXYpn

Online resources

These websites are also relevant as further information sources:

- ▶ IBM DS8000 series:

  https://ibm.biz/Bd4ukG

- ▶ IBM DS8000 Series V7.2 Knowledge Center:

  http://www.ibm.com/support/knowledgecenter/HW213_7.2.0/com.ibm.storage.ssic.help.doc/f2c_ichomepage.htm?lang=en

- ▶ IBM Service and support best practices for Power Systems website:

  https://ibm.biz/BdRCLT

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

Redbooks

IBM PowerHA SystemMirror for i: Using DS8000

Printed in U.S.A.

Get connected

ibm.com/redbooks