

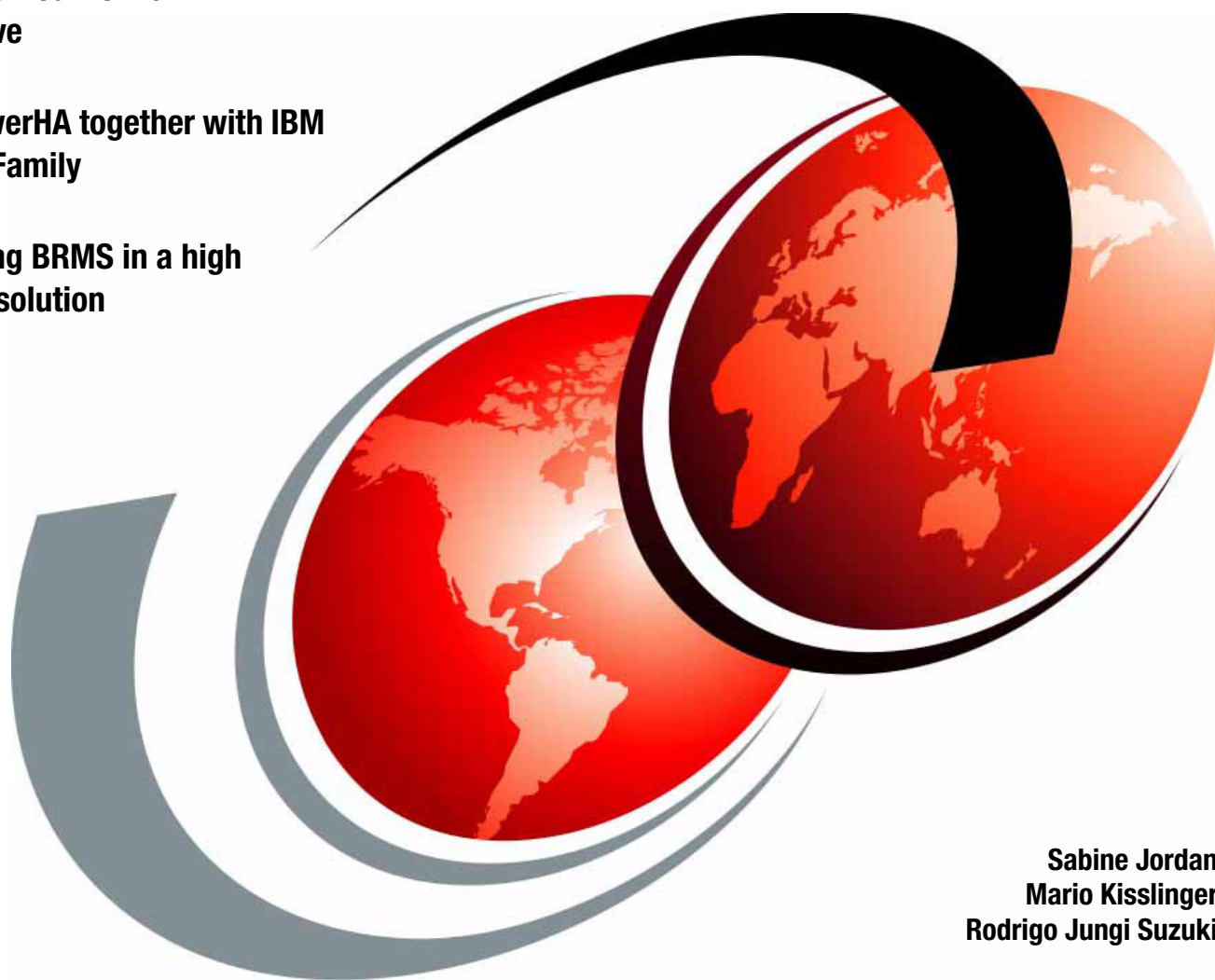
IBM i and IBM Storwize Family

A Practical Guide to Usage Scenarios

VIOS explained from an IBM i perspective

Using PowerHA together with IBM Storwize Family

Configuring BRMS in a high available solution



Sabine Jordan
Mario Kisslinger
Rodrigo Jungi Suzuki

Redbooks



International Technical Support Organization

**IBM i and IBM Storwize Family: A Practical
Guide to Usage Scenarios**

April 2014

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

First Edition (April 2014)

This edition applies to IBM i 7.1 Technology Refresh (TR) 7.

© Copyright International Business Machines Corporation 2014. All rights reserved.

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

Contents

Notices	vii
Trademarks	viii
Preface	1
Authors	1
Now you can become a published author, too!	3
Comments welcome	3
Stay connected to IBM Redbooks	3
Chapter 1. Introduction to IBM Storwize Family for IBM i	5
1.1 Decision support for internal versus external storage	6
1.2 Attachment options	7
1.2.1 Direct attached	7
1.2.2 Native attached	8
1.2.3 vSCSI attached	8
1.2.4 N-Port ID Virtualization (NPIV) attached	9
1.3 Basics of IBM Storwize family	10
1.4 IBM Storwize V7000	14
1.4.1 Control enclosure models	14
1.4.2 Expansion enclosure models	15
Chapter 2. Configuring IBM Storwize V7000 with IBM i	17
2.1 Initial setup	18
2.2 Initial configuration	19
2.3 Preparing the IBM Storwize V7000	25
2.3.1 Step 1: Creating the storage pool	25
2.3.2 Step 2: Creating a generic volume	29
2.3.3 Step 3: Creating a compressed volume	31
2.3.4 Step 4: Creating a host	34
2.3.5 Step 5: Mapping the LUN to the host	36
2.4 IBM Storwize V7000 and hardware compression	37
2.4.1 Using the IBM Comprestimator Utility	38
2.4.2 Creating a volume mirrored copy on a compressed volume	39
2.4.3 Comparing the performance of compressed and non-compressed volumes	43
2.5 Tips for implementing IBM Storwize V7000 in a IBM i environment	50
2.5.1 Storwize V7000 configuration options	50
2.5.2 Host attachment	51
2.5.3 Multipath	52
2.5.4 Zoning SAN switches	55
2.5.5 Data layout	57
2.5.6 Solid-State Drives	59
2.5.7 Booting from SAN	60
2.5.8 Sizing for performance	60
2.5.9 Performance monitoring	61
2.5.10 Copy services considerations	61
2.5.11 Translation from 520-byte blocks to 512-byte blocks	62
Chapter 3. Setting up VIOS	63
3.1 Before you start	64

3.2	VIOS setup by using HMC GUI	65
3.3	VIOS setup by using the CLI	69
3.4	Mapping storage through VIOS to IBM i	72
3.4.1	Setting up vSCSI connections	72
3.4.2	Mapping storage by using NPIV	76
3.5	Network setup	79
3.5.1	Optional step: LACP setup for adapter redundancy in each VIOS	80
3.5.2	Creating VLAN-capable trunk Ethernet adapters	80
3.5.3	Shared Ethernet Adapter setup	83
3.5.4	Transporting VLANs to the external network	85
3.6	Customizing	85
3.6.1	Changing ksh to bash style	86
3.6.2	Configuring monitoring by using SNMP	86
3.6.3	Creating an image repository for client D-IPL	87
3.7	Troubleshooting	92
3.7.1	Measuring performance	92
3.7.2	Restarting the RMC connection	95
3.7.3	Collecting troubleshooting data for support	96
Chapter 4. PowerHA and the IBM Storwize Family		97
4.1	Usage scenarios for IBM PowerHA SystemMirror for i	98
4.1.1	Metro Mirror	99
4.1.2	Global Mirror	99
4.1.3	LUN level switching	100
4.1.4	FlashCopy	102
4.2	Setting up IBM PowerHA SystemMirror for i	103
4.2.1	Requirements for PowerHA SystemMirror for i with SAN Volume Controller or Storwize	103
4.2.2	SAN Volume Controller split cluster and LUN level switching	105
4.2.3	LUN level switching in combination with Global Mirror	137
4.2.4	Metro Mirror setup	150
4.2.5	Adding LUNs to an existing IASP	150
Chapter 5. Live Partition Mobility		157
5.1	Usage scenarios for Live Partition Mobility	158
5.2	License considerations with Live Partition Mobility	162
5.3	Setting up Live Partition Mobility	163
5.3.1	Requirements	163
5.3.2	Preparation	164
5.3.3	Live Partition Mobility migration	168
5.3.4	Hints and tips for the validation process	170
Chapter 6. BRMS and FlashCopy		173
6.1	Introduction to PowerHA, BRMS, and FlashCopy scenario	174
6.2	PowerHA and FlashCopy	174
6.2.1	Preparation for FlashCopy on the Storwize side	174
6.2.2	Changes in PowerHA to add the third node into the cluster	175
6.3	Backup Recovery, and Media Services (BRMS)	176
6.3.1	Setting up the BRMS network group	176
6.3.2	Configuring BRMS media classes and media policies	178
6.3.3	Configuring the BRMS control group	179
Appendix A. PowerHA Tools IBM i		183
	PowerHA Tools for IBM i	184

IBM Lab Services Offerings for PowerHA for IBM i	185
Related publications	187
IBM Redbooks	187
Other publications	187
Online resources	187
Help from IBM	188

Notices

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

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

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

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

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

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

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

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

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


COPYRIGHT LICENSE:

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

Trademarks

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

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

AIX®	Informix®	Redbooks®
DB2®	Power Systems™	Redbooks (logo)  ®
DS8000®	POWER7®	Storwize®
Easy Tier®	PowerHA®	System Storage®
FlashCopy®	PowerVM®	SystemMirror®
IBM®	Real-time Compression™	Tivoli®

The following terms are trademarks of other companies:

ITIL is a registered trademark, and a registered community trademark of The Minister for the Cabinet Office, and is registered in the U.S. Patent and Trademark Office.

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

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

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

Preface

The use of external storage and the benefits of virtualization became a topic of discussion in the IBM® i area during the last several years. The question tends to be, what are the advantages of the use of external storage that is attached to an IBM i environment as opposed to the use of internal storage. The use of IBM PowerVM® virtualization technology to virtualize Power server processors and memory also became common in IBM i environments. However, virtualized access to external storage and network resources by using a VIO server is still not widely used.

This IBM Redbooks® publication gives a broad overview of the IBM Storwize® family products and their features and functions. It describes the setup that is required on the storage side and describes and positions the different options for attaching IBM Storwize family products to an IBM i environment. Basic setup and configuration of a VIO server specifically for the needs of an IBM i environment is also described.

In addition, different configuration options for a combined setup of IBM PowerHA® SystemMirror® for i and the Storwize family products are described and positioned against each other. Detailed examples are provided for the setup process that is required for these environments.

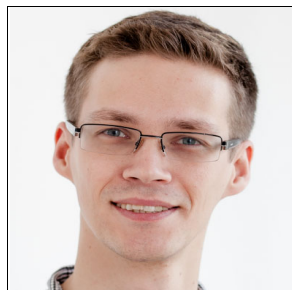
The information that is provided in this book is useful for clients, IBM Business Partners, and IBM service professionals who need to understand how to install and configure their IBM i environment with attachment to the Storwize family products.

Authors

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



Sabine Jordan is a Consulting IT Specialist in IBM Germany. She has more than 20 years experience as a Technical Specialist in the IBM i, specializing in high availability. Currently she works on IBM PowerHA SystemMirror for i implementations for SAP and non-SAP environments that use geographic mirroring, IBM DS8000®, and SAN Volume Controller remote copy services. She creates concepts for the design and implements the entire project (cluster setup, application changes), in addition to performing customer education and testing. Sabine also presents and delivers workshops (internally and externally) on IBM PowerHA SystemMirror for i, high availability, and disaster recovery.



Mario Kisslinger is a Solution Architect for UBL Informationssysteme GmbH in Germany. He has six years of experience with designing infrastructure solutions for IBM i customers. In complex project situations, he accompanies the whole project lifecycle. As a Pre-Sales consultant, he designs and configures the solution and supports the proposal. In the implementation phase, he installs the solution. He also helps with the transition into the operation phase. His areas of expertise include backup and recovery, migration scenarios, and serving IBM i out of the datacenter.



Rodrigo Jungi Suzuki is SAN Storage specialist at IBM Brazil Global Delivery Center in Hortolandia. Currently, Rodrigo is a Subject Matter Expert (SME) account focal point, and works with projects, implementations, and support for international clients. He has 20 years of IT Industry experience with the last five years in the SAN Storage area. He has a background in UNIX and Informix® database. He holds a bachelor degree in Computer Science from Universidade Paulista in Sao Paulo, Brazil, and is an IBM Certified IT Specialist. Rodrigo also is certified in NetApp NCPA, IBM Storwize V7000 Technical Solutions V2, and ITIL.

This edition of this IBM Redbooks project was led by Debbie Landon of the International Technical Support Organization, Rochester Center.

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

David Bhaskaran
Jenny Dervin
Steven Finnes
Brian Nordland
Marilyn Rodriguez
Kiswanto Thayib
Kristopher Whitney
Keith Zblewski
IBM Rochester, IBM i Development team

Ingo Dimmer
IBM Germany, Advanced Technical Skills (ATS) team

David Painter
IBM United Kingdom, STG Lab Services

Mike Schambureck
IBM US, IBM i Technology Center (iTC)

Alison Pate
IBM US, Advanced Technical Skills (ATS) team

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 this website:

<http://www.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 that is found at this website:

<http://www.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 Storwize Family for IBM i

With the entrance of the IBM Storwize Family into the market, many options to attach external storage to an IBM i host system became available. This chapter describes the pros and cons of external versus internal storage. This chapter also describes the available options to connect storage of the IBM Storwize Family to IBM Power Systems and explains the differences between the available Storwize models.

The following topics are covered in this chapter:

- ▶ Decision support for internal versus external storage
- ▶ Attachment options
- ▶ Basics of IBM Storwize family
- ▶ IBM Storwize V7000

1.1 Decision support for internal versus external storage

IBM i clients focus on a high number of fast disk arms to handle the small I/O block sizes that often are on IBM i. Disk performance is always mission critical.

It is reasonable to use internal storage for the following reasons:

- ▶ Easy to distribute specific physical disk arms between the production, test, and development partitions.
- ▶ Everything is managed from one box.
- ▶ In case of errors, there is only one contact person in support.
- ▶ Easy to set up within one department.
- ▶ There are no unwanted side effects if other partitions produce a high workload.

However, internal disks can be annoying for the following reasons:

- ▶ Low flexibility to expand or reduce the number of disks if needed.
- ▶ Complete disks or expansion units must be provisioned to partitions.
- ▶ No load balancing between disks with high read/write usage and those with low read/write usage.
- ▶ The IBM i administrator must worry about disk protection and the number of hot spare disks.

External storage helps solve the annoying points of internal storage and can provide solutions for most of the positive points. External storage features the following advantages:

- ▶ Storage area network (SAN) storage can be placed in different pools for production and test workloads. It can also be implemented in different storage tier; for example, with different reliability and performance demands.
- ▶ Availability and disk protection is managed by the storage administrator.
- ▶ It is much easier to move or replicate the storage into other areas to minimize the risk of disaster.
- ▶ Storage can be segmented into more and smaller parts, with the flexibility to move storage between partitions.
- ▶ Smaller footprint in the data center that leads to reduced power and cooling costs.
- ▶ Investment cycle for storage and server can be split apart, which allows longer use of the external storage when the IBM Power System is replaced.

As a rule, the advantages to use external storage are growing, as more partitions must be deployed. Every logical disk unit is distributed to many physical disks and automatically protected against failure. It is possible to set up new workloads when they are needed, without the demand for more disk adapters or storage boxes.

With the support for external storage in IBM PowerHA SystemMirror for IBM i or the built-in replication functions of the IBM Storwize Family, a broad range of disaster recovery and high availability options are available.

IBM PowerVM Live Partition Mobility also depends on external storage. With this product, logical partitions can be moved to another IBM Power System, without the need for downtimes. This can be used for planned outages like maintenance windows or migrations.

1.2 Attachment options

There are several different options to connect external storage to the IBM i operating system. This section explains the differences between these attachment options and provide some hints about when to use which option.

For more information about possible combinations of Host Bus Adapters, specific IBM Power Systems, and firmware levels, see the IBM System Storage® Interoperation Center (SSIC) at this website:

<http://www-03.ibm.com/systems/support/storage/ssic/interoperability.wss>

1.2.1 Direct attached

Figure 1-1 shows that direct attached means a connection from a member of the IBM Storwize Family to a logical partition that is running IBM i without a switch. It is important to connect each Host Bus Adapter to each Node of the storage system for redundancy.

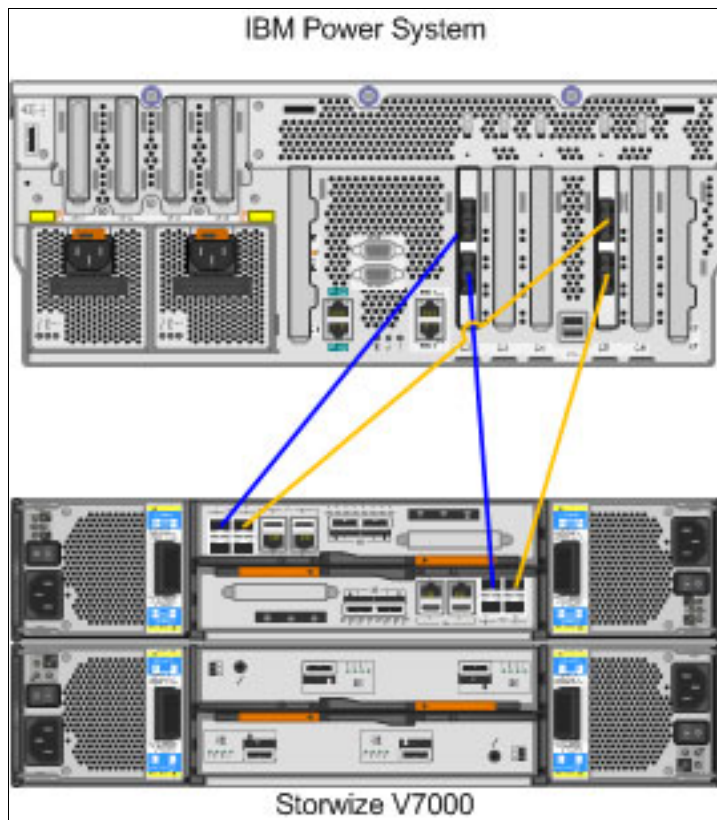


Figure 1-1 Direct attachment of IBM Storwize V7000 to IBM Power System

Important: Direct attached is restricted to the 4 GB adapters, 5774 or the low profile adapter 5276, and requires IBM i 7.1 TR6 with the latest Program Temporary Fixes (PTFs).

Direct attached storage is a good choice for systems with up to two logical partitions. The number of partitions is limited by the Fibre Channel Ports in the storage system. It is possible to use fewer Host Bus Adapters at the client partition, but this limits the number of active paths for the disks to one, which is not considered a best practice and can be error prone.

1.2.2 Native attached

As shown in Figure 1-2, in a native attached setup, the storage is connected to IBM i by using a SAN switch. IBM i 7.1 TR6 with the latest PTFs is the minimum requirement for this attachment option.

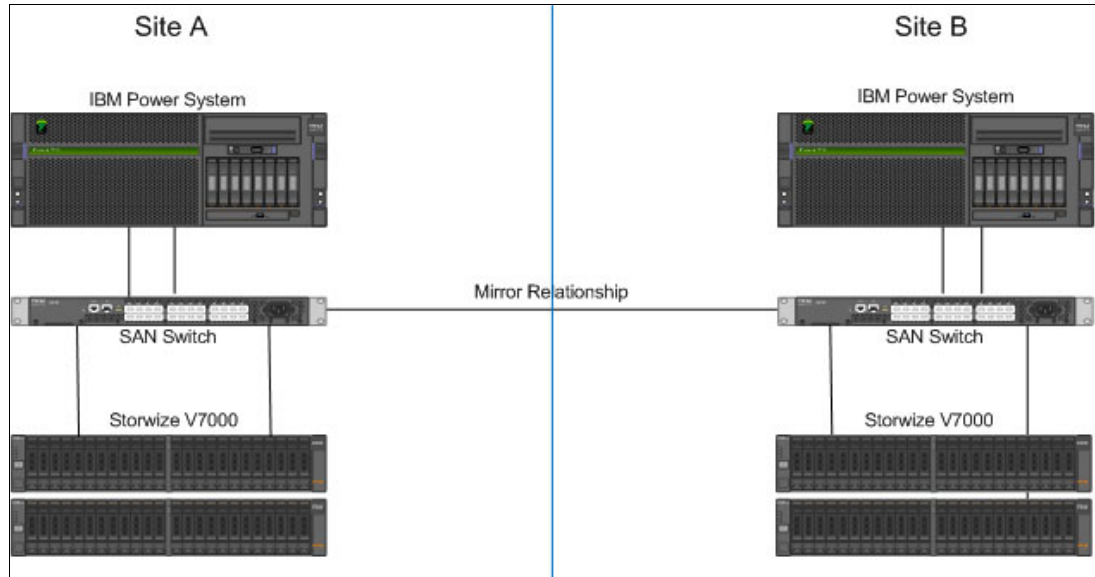


Figure 1-2 Two native attached IBM Storwize V7000 to two IBM Power Systems

The native attach option allows for a greater number of IBM i partitions (compared to the direct attach option) because all Host Bus Adapters are connected to the switch. The amount of the maximum partitions is limited by the number of SAN switch ports and slots in the IBM Power System.

1.2.3 vSCSI attached

For a storage attachment that uses vSCSI, an IBM Virtual I/O Server (VIOS) must be installed (see Figure 1-3 on page 9). The logical disks from shared SAN storage are assigned to the VIOS partitions. The VIOS provides them to the IBM i client by using virtual SCSI connections. For more information about the necessary steps for this configuration, see 3.4.1, “Setting up vSCSI connections” on page 72.

The use of VIOS as virtualization layer provides a highly flexible and customized solution where it is possible to use fewer physical adapters as compared to direct attached and native attached configurations.

Disks can be attached directly to the VIOS partitions (without the use of a switch) or they can be provided through the SAN fabric. In either case, the disks are mapped directly to the VIOS where they are shown as hdiskX devices. Afterward, a mapping from VIOS to the IBM i client must be configured.

Note: It is also possible to provide internal disks through VIOS to the IBM i partitions by using vSCSI and mirroring these disks through the IBM i operating system for resilience.

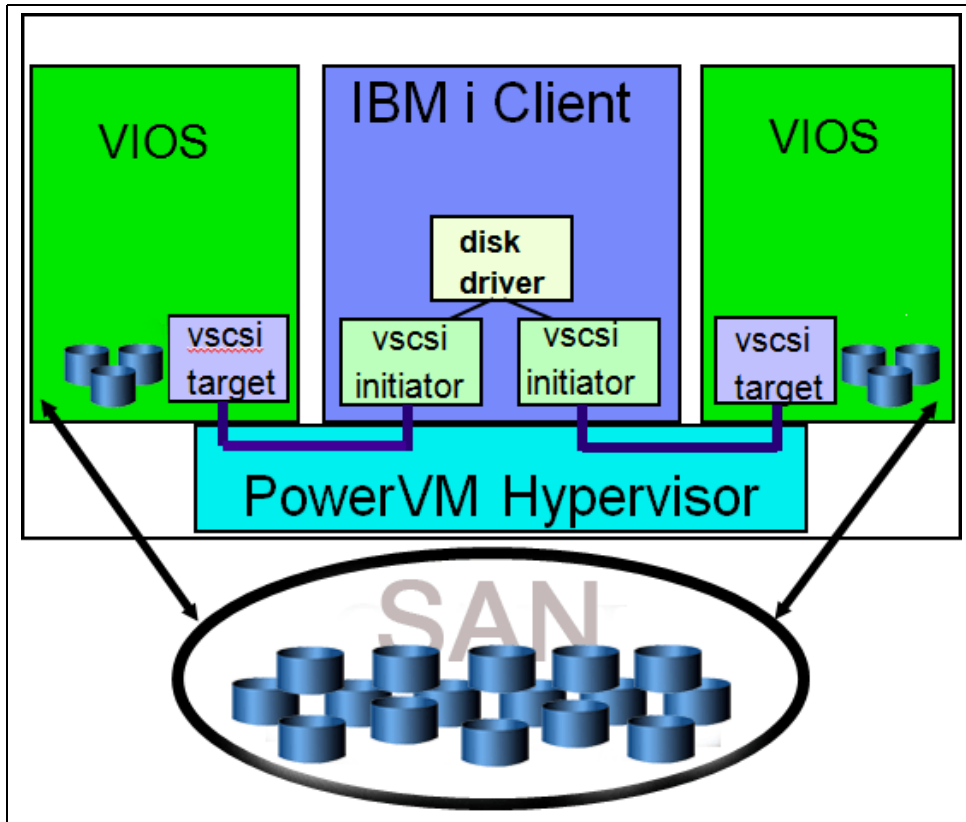


Figure 1-3 Dual VIOS setup for vSCSI disks

1.2.4 N-Port ID Virtualization (NPIV) attached

Figure 1-4 on page 10 shows the difference between vSCSI and N-Port ID Virtualization (NPIV) attached. Instead of mapping the disks to the VIOS, the disks are transported transparently through VIOS and mapped to the IBM i client. For more information about the setup process for this configuration, see 3.4.2, “Mapping storage by using NPIV” on page 76.

Virtual Fibre Channel client adapters are assigned to the IBM i partitions. The IBM Power server hypervisor automatically generates a pair of virtual World Wide Port Numbers (WWPNs) for this adapter that is used for SAN zoning and the host attachment within the SAN storage.

Note: An NPIV capable Host Bus Adapter and an NPIV enabled SAN switch are required.

The transparent connection of NPIV allows for better performance than virtual SCSI, simplified SAN setup, and multipath connections to the storage by adding them to different zones, as described in 2.5.3, “Multipath” on page 52.

The NPIV setup is the most flexible configuration for a large number of partitions on a single IBM Power System.

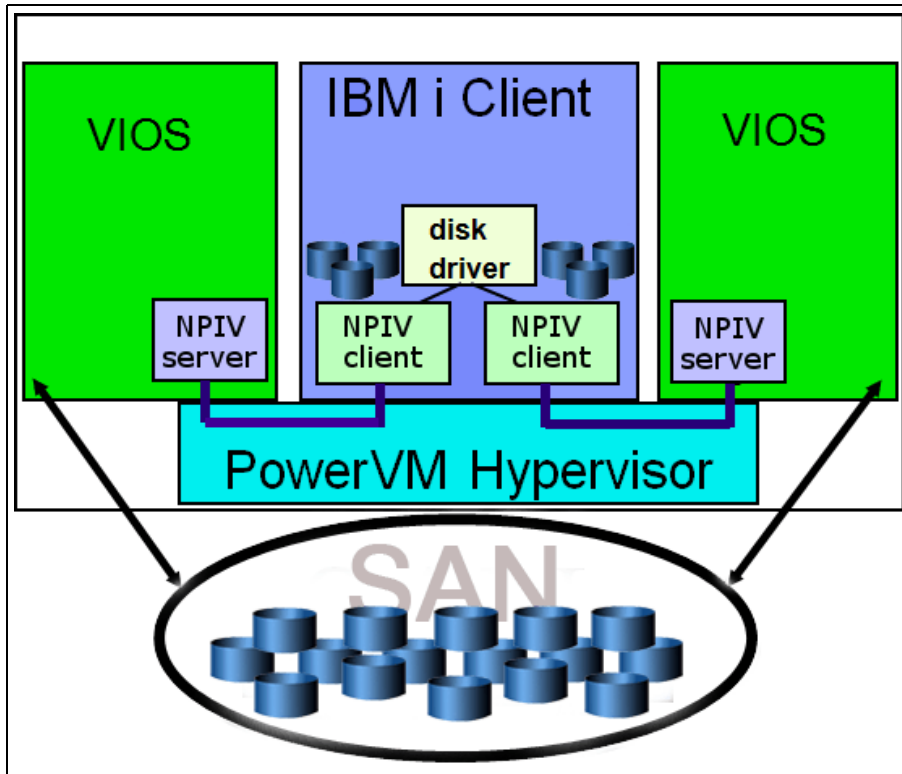


Figure 1-4 Dual VIOS setup for NPIV disks

1.3 Basics of IBM Storwize family

The SAN Volume Controller is a storage virtualization system that facilitates storage management, providing efficiency. It helps to improve storage space usage and SAN administrator productivity.

The SAN Volume Controller can be a single point of control to provision LUNs to your entire SAN environment. It can virtualize a variety of storage systems. For more information, see the following IBM System Storage SAN Volume Controller website:

<http://www-03.ibm.com/systems/storage/software/virtualization/svc/>

You can use IBM SAN Volume Controller with two or more nodes, which are arranged in a cluster. A *cluster* is a pair of nodes and can grow up to four SAN Volume Controller node pairs.

Figure 1-5 shows the front view of an IBM SAN Volume Controller.



Figure 1-5 SAN Volume Controller front view

The IBM SAN Volume Controller main function is to virtualize the disk drives that are provided from different storage systems that are connected at the back end. The main function of the IBM Storwize V3700 and IBM Storwize V7000 is to provide disks from their expansion enclosures. Although these are the most usual Storwize family usage situations, the IBM Storwize V3700 and Storwize V7000 can also virtualize external storages.

One of the main differences between the IBM Storwize V3700 and the IBM Storwize V7000 is capacity. The Storwize V3700 supports up to four expansion units and can have up to 240 TB of capacity. The Storwize V7000 can support up to nine expansion units, reaching up to 480 TB of capacity.

Figure 1-6 shows the front view of Storwize V3700.



Figure 1-6 Storwize V3700 front view

The IBM Storwize V3700 is an entry-level storage system that is targeted to meet the requirements of the small business market and the IBM Storwize V7000 is targeted toward more complex environments. There are differences with the cache per controller, maximum number of expansions, quantity of host interface, and so on.

If you need Real-time Compression™ function, that is supported on the IBM Storwize V7000. For more information, see 2.4, “IBM Storwize V7000 and hardware compression” on page 37.

Figure 1-7 shows the front view of the IBM Storwize V7000. For more information about configuring an IBM Storwize V7000 in an IBM i environment, see Chapter 2, “Configuring IBM Storwize V7000 with IBM i” on page 17.



Figure 1-7 Storwize V7000 front view

There also is an IBM Storwize V3500 storage system. It consists of a one control enclosure. The control enclosure contains disk drives and two nodes, which are attached to the SAN fabric and reaches up to 36 TB of storage capacity.

Important: The IBM Storwize V3500 is available only in the following countries and regions:

- ▶ People's Republic of China
- ▶ Hong Kong S.A.R. of the PRC
- ▶ Macao S.A.R. of the PRC
- ▶ Taiwan

The user interface for the IBM SAN Volume Controller is the same for all of the Storwize family, as shown in Figure 1-8.



Figure 1-8 IBM SAN Volume Controller Overview panel

For more information about the features, benefits, and specifications of the IBM Storwize family, see the following websites:

- ▶ IBM System Storage SAN Volume Controller:
<http://www-03.ibm.com/systems/storage/software/virtualization/svc/>
- ▶ IBM Storwize V7000 and Storwize V7000 Unified Disk Systems:
http://www-03.ibm.com/systems/storage/disk/storwize_v7000/index.html
- ▶ IBM Storwize V3700:
http://www-03.ibm.com/systems/storage/disk/storwize_v3700/index.html
- ▶ IBM Storwize V3500:
http://www-03.ibm.com/systems/hk/storage/disk/storwize_v3500/index.html

Table 1-1 list a series of videos that are available for the IBM Storwize family.

Table 1-1 IBM Storwize videos

Video Description	URL
Storwize Family	http://www.youtube.com/watch?v=M0hIfLKUk1U
Storwize V3700	http://www.youtube.com/watch?v=IdxnH9YWF9s
Storwize V7000	http://www.youtube.com/watch?v=S2p4IQxIiec&feature=c4-overview-v1&list=PL48FE37276F6014E8

Video Description	URL
IBM Real-time Compression and Storwize V7000 and SAN Volume Controller demonstration	http://www.youtube.com/watch?v=rgKj75kn2J0

1.4 IBM Storwize V7000

In this section, we provide more details about the IBM Storwize V7000.

An IBM Storwize V7000 system consists of machine type 2076 rack-mounted enclosures. The control enclosure contains the main processing units that control the entire system and can be connected to expansion enclosures. The expansion enclosures can hold up to 12 3.5-inch drives or up to 24 2.5-inch drives.

1.4.1 Control enclosure models

For the 2076-112 and 2076-124 models, there are two power supply slots numbered as 1 and two canister slots numbered as 2 in Figure 1-9. The canister slots have Fibre Channel ports, USB ports, Ethernet ports, and Serial-attached SCSI ports.

Machine type and model 2076-112 can hold up to 12 3.5-inch drives and the 2076-124 can hold up to 24 2.5-inch drives.

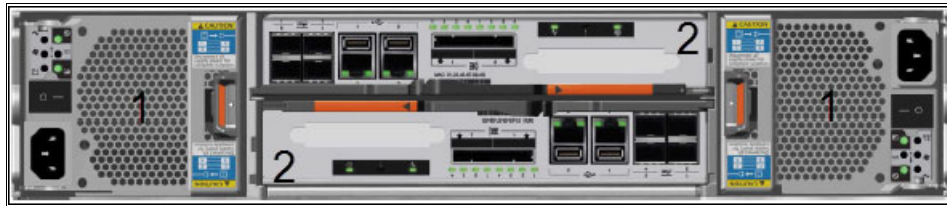


Figure 1-9 Storwize V7000 control enclosure rear view, models 112 and 124

Figure 1-10 shows the rear view of the control enclosure for the machine type and models 2076-312 and 324, which includes 10 Gbps Ethernet ports capabilities.

Machine type and model 2076-312 can hold up to 12 3.5-inch drives and the 2076-324 can hold up to 24 2.5-inch drives.

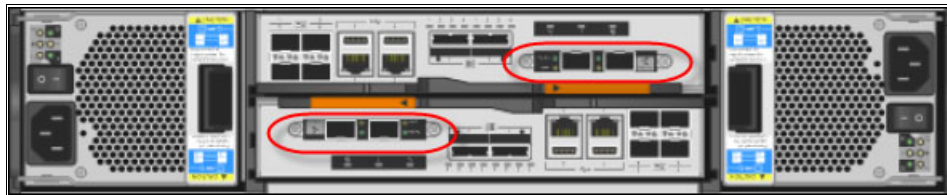


Figure 1-10 Storwize V7000 control enclosure rear view, models 312 and 324

1.4.2 Expansion enclosure models

Expansion enclosure model 2076-212 can hold up to 12 3.5-inch drives, as shown Figure 1-11.



Figure 1-11 Storwize V7000 expansion enclosure front view, models 2076-112, 212, and 312

Expansion enclosure model 2076-224 can hold up to 24 2.5-inch drives, as shown Figure 1-12.



Figure 1-12 Storwize V7000 expansion enclosure front view, models 2076-124, 224, and 324

The expansion enclosure has two power supply slots numbered as 1 and two canister slots numbered as 2 as shown in Figure 1-13.



Figure 1-13 Storwize V7000 expansion enclosure rear view

Each canister has two SAS ports that must be connected if you are adding an expansion enclosure.

Note: Control enclosures have Ethernet ports, Fibre Channel ports, and USB ports. Expansion enclosures do not have these ports.



Configuring IBM Storwize V7000 with IBM i

This chapter describes configuring an IBM Storwize V7000 with an IBM i system.

The following topics are covered in this chapter:

- ▶ Initial setup
- ▶ Initial configuration
- ▶ Preparing the IBM Storwize V7000
- ▶ IBM Storwize V7000 and hardware compression
- ▶ Tips for implementing IBM Storwize V7000 in a IBM i environment

2.1 Initial setup

This section describes how to perform the initial configuration of an IBM Storwize V7000.

The Storwize V7000 uses a USB key to start the initial configuration. The USB key comes with the Storwize V7000 and contains a program that is called `InitTool.exe` to start the initialization.

Tip: If you do not have the USB key that is shipped with the Storwize V7000, you can use another USB key. You can also download the `InitTool.exe` file from this website:

[http://www-933.ibm.com/support/fixcentral/swg/selectFixes?parent=Mid-range+disk+systems&product=ibm/Storage_Disk/IBM+Storwize+V7000+\(2076\)&release=All&platform=All&function=all](http://www-933.ibm.com/support/fixcentral/swg/selectFixes?parent=Mid-range+disk+systems&product=ibm/Storage_Disk/IBM+Storwize+V7000+(2076)&release=All&platform=All&function=all)

Complete the following steps to start the initial configuration of an IBM Storwize V7000:

1. Insert the USB key into a personal computer that is running the Microsoft Windows operating system. The initialization program is started automatically if the system is configured to autorun for USB keys or open the USB key from Windows system and click **InitTool.exe**.
2. In the IBM Storwize V7000 Initialization Tool window (see Figure 2-1), select **Initialize a new system using the USB key** and click **Next**.

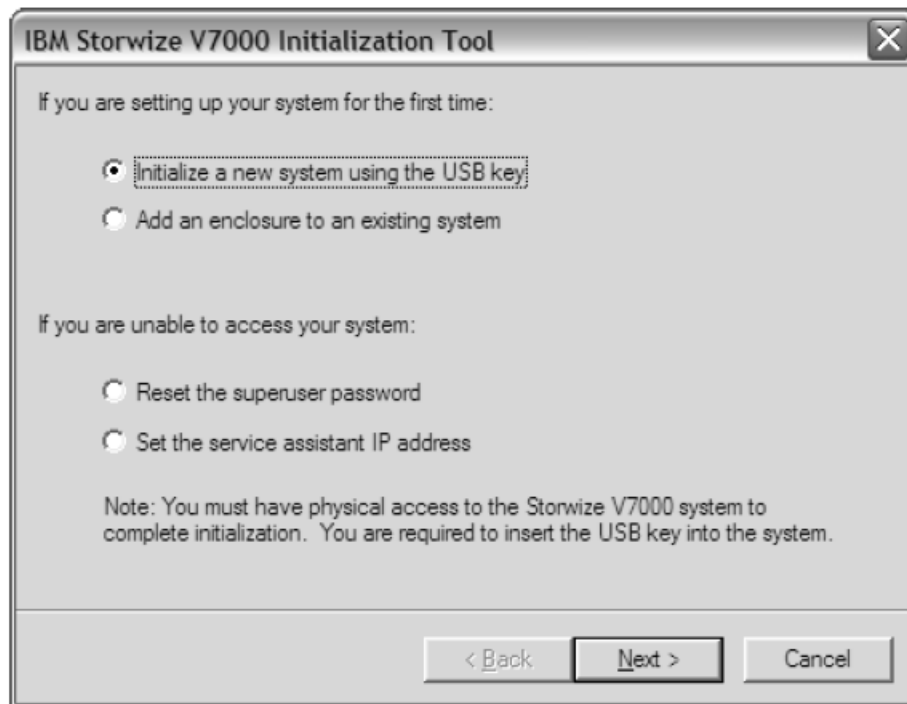


Figure 2-1 IBM Storwize V7000 initialization tool

3. The initialization panel guides you through the following steps:
 - a. Enter the IP address, subnet mask, and network gateway address.
 - b. After the initialization process finishes, eject the USB key and insert it into any USB port on the control enclosure of your Storwize V7000 system. After the USB key indicator lights stops flashing, remove the USB key from the system.

- c. Reinsert the USB key in your personal computer to continue setting up your system. On the Microsoft Windows operating system, you see the results of the operation.
4. From a personal computer that has an Ethernet connection, open a supported web browser and point to the system management IP address to start the graphical user interface.
5. Log in as superuser and use passw0rd for the password. It is considered a best practice to change the default password for security reasons.

Tip: Normally, the Storwize V7000 contains sufficient power for the system to start. If the batteries do not have sufficient charge, the system cannot start. You must wait until the system became available.

2.2 Initial configuration

Complete the following steps to complete the initial Storwize V7000 configuration:

1. From the Microsoft Windows operating system, open a web browser and point to the system management IP address. You see the IBM Storwize V7000 GUI, as shown in Figure 2-2.

Log in as superuser and use passw0rd for the password. Click **Continue**.



Figure 2-2 IBM Storwize V7000 initial configuration

2. Read and agree to the license agreement, as shown in Figure 2-3. Click **Next**.

License Agreement (Step 1 of 6)

Read the license agreement carefully.

License | IBM Notices | Java Notices | Non-IBM Licenses | Additional Licenses and Notices

International Program License Agreement

Part 1 - General Terms

BY DOWNLOADING, INSTALLING, COPYING, ACCESSING, CLICKING ON AN "ACCEPT" BUTTON, OR OTHERWISE USING THE PROGRAM, LICENSEE AGREES TO THE TERMS OF THIS AGREEMENT. IF YOU ARE ACCEPTING THESE TERMS ON BEHALF OF LICENSEE, YOU REPRESENT AND WARRANT THAT YOU HAVE FULL AUTHORITY TO BIND LICENSEE TO THESE TERMS. IF YOU DO NOT AGREE TO THESE TERMS,

- DO NOT DOWNLOAD, INSTALL, COPY, ACCESS, CLICK ON AN "ACCEPT" BUTTON, OR USE THE PROGRAM, AND
- PROMPTLY RETURN THE UNUSED MEDIA, DOCUMENTATION, AND PROOF OF ENTITLEMENT TO THE PARTY FROM WHOM IT WAS OBTAINED FOR A REFUND OF THE AMOUNT PAID. IF THE PROGRAM WAS DOWNLOADED, DESTROY ALL COPIES OF THE PROGRAM.

1. Definitions

"Authorized Use" - the specified level at which licensee is authorized to execute or run the Program. That level may be measured by number of users, millions of service units ("MSUs"), Processor Value Units ("PVUs"), or other level of use specified by IBM.

"IBM" - International Business Machines Corporation or one of its subsidiaries.

I agree with the terms in the license agreement.

I do not agree with the terms in the license agreement.

Next >

Figure 2-3 IBM Storwize V7000 license agreement

3. Enter the system name, time zone, date, and time, as shown in Figure 2-4. Click **Next**.

Name, Date, and Time (Step 2 of 6)

System Name

Time Zone

Set Date and Time

Date

12-Hour 24-Hour

Time

Set NTP Server IP Address

Next >

Figure 2-4 IBM Storwize V7000 system name, time zone, and date setup

4. You can add a license for external storage, remote copy, and real-time compression, as shown in Figure 2-5. Click **Next**.

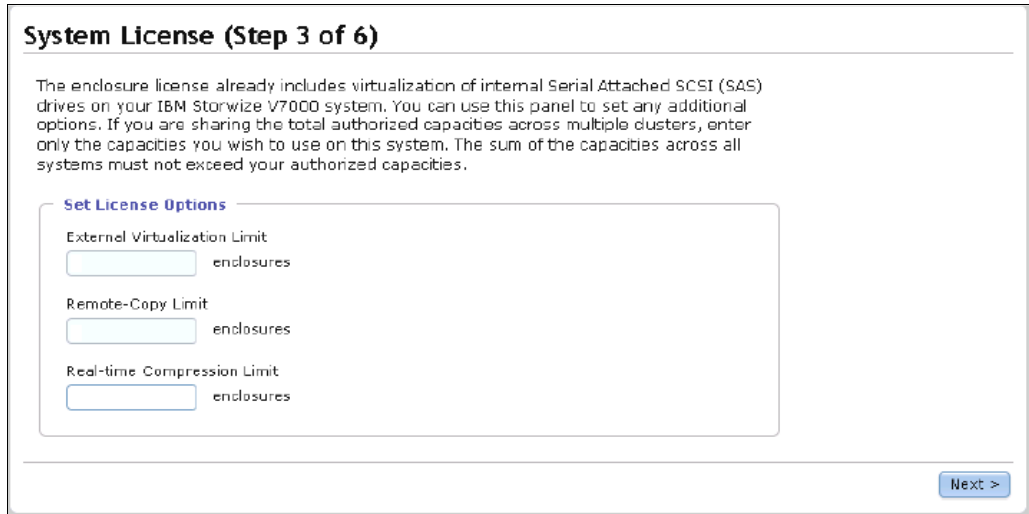


Figure 2-5 IBM Storwize V7000 system license

5. Configure call home and email alert.

You can configure your system to send an email report to IBM support if a hardware issue is detected. This function is called *Call Home*. To configure the Call Home feature, complete the following steps:

- a. Click **Configure Email Event Notifications**, as shown in Figure 2-6.

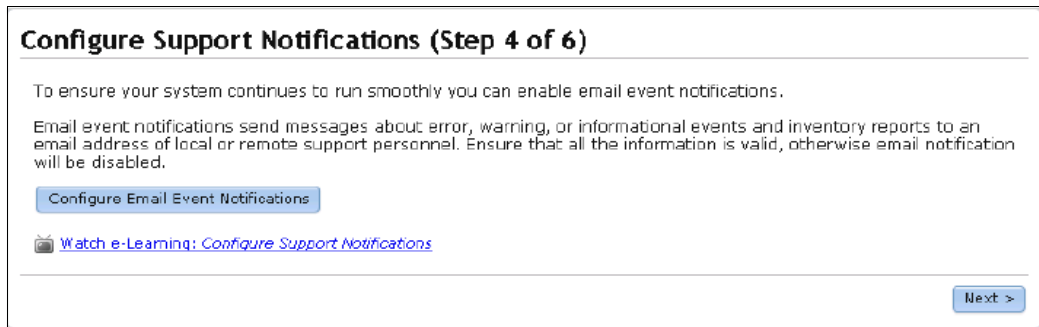


Figure 2-6 Email event notifications setup

- b. Configure the email account information as shown in Figure 2-7. Click **Next**.

Figure 2-7 Email account information

- c. Configure the email servers, as shown in Figure 2-8. Click **Next**.

Figure 2-8 Configuring the email servers

- d. Enter the IBM support email address to receive the notifications, as shown in Figure 2-9. Click **Next**.

Figure 2-9 Support notifications email address

- e. Verify the Summary information as shown in Figure 2-10 to ensure the information that is provided is correct. Click **Finish**.

Figure 2-10 Support notification summary

- f. After the support notification is configured, you can edit the configuration by clicking **Edit** or you can continue with the setup by clicking **Next**, as shown in Figure 2-11.

Figure 2-11 Support notification information

6. The next window displays the hardware configuration, as shown in Figure 2-12. Click **Next**.

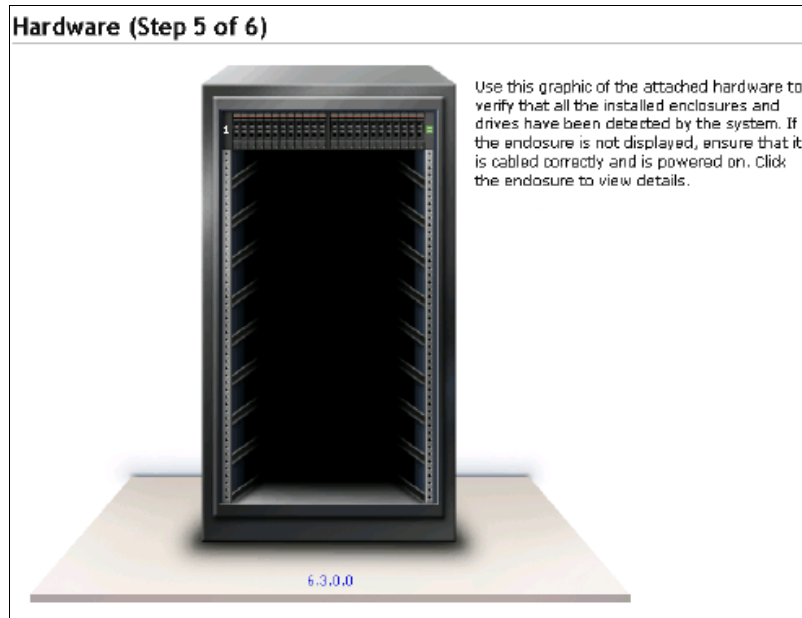


Figure 2-12 IBM Storwize V7000 hardware

7. An Add Enclosure pop-up window opens, as shown in Figure 2-13. Wait until the task is finished and then click **Close**.

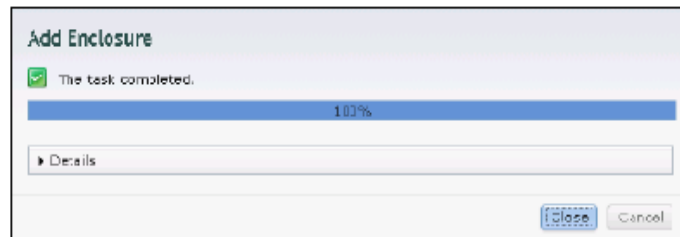


Figure 2-13 Adding an enclosure

8. The configuration wizard detects the internal disk drives, as shown in Figure 2-14. The following choices are available:
 - Accept the configuration that is proposed by selecting the check box to automatically configure the internal storage and click **Finish**.

Note: All available disks are configured into RAID sets and hot spare drives that are based on the number of drives that are found. However, this might not be what you intended to configure. Check the intended configuration before you click **Finish**.

- Leave the check box cleared, click **Finish**, and continue with the storage configuration.

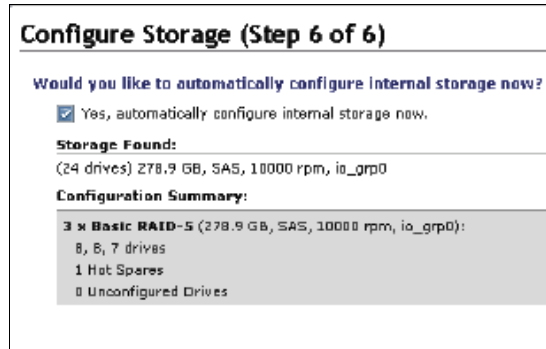


Figure 2-14 Internal disk drives configuration proposed

9. After you click **Finish**, wait until the setup Wizard is complete, as shown in Figure 2-15.

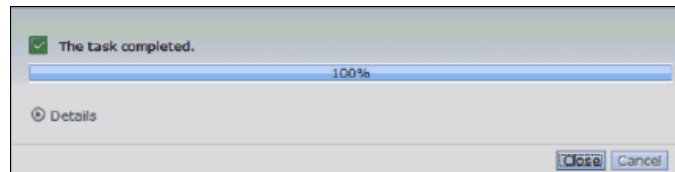


Figure 2-15 Wizard setup complete

2.3 Preparing the IBM Storwize V7000

This section describes the steps to set up the IBM Storwize V7000, which consists of the following tasks:

- ▶ 2.3.1, “Step 1: Creating the storage pool” on page 25
- ▶ 2.3.2, “Step 2: Creating a generic volume” on page 29
- ▶ 2.3.3, “Step 3: Creating a compressed volume” on page 31
- ▶ 2.3.4, “Step 4: Creating a host” on page 34
- ▶ 2.3.5, “Step 5: Mapping the LUN to the host” on page 36

2.3.1 Step 1: Creating the storage pool

Complete the following steps to create the storage pool:

1. Open the GUI by entering the management IP address into a web browser’s URL field and log in to the Storwize V7000, as shown Figure 2-16 on page 26.



Figure 2-16 IBM Storwize V7000 login panel

Figure 2-17 shows the Storwize V7000 overview window.

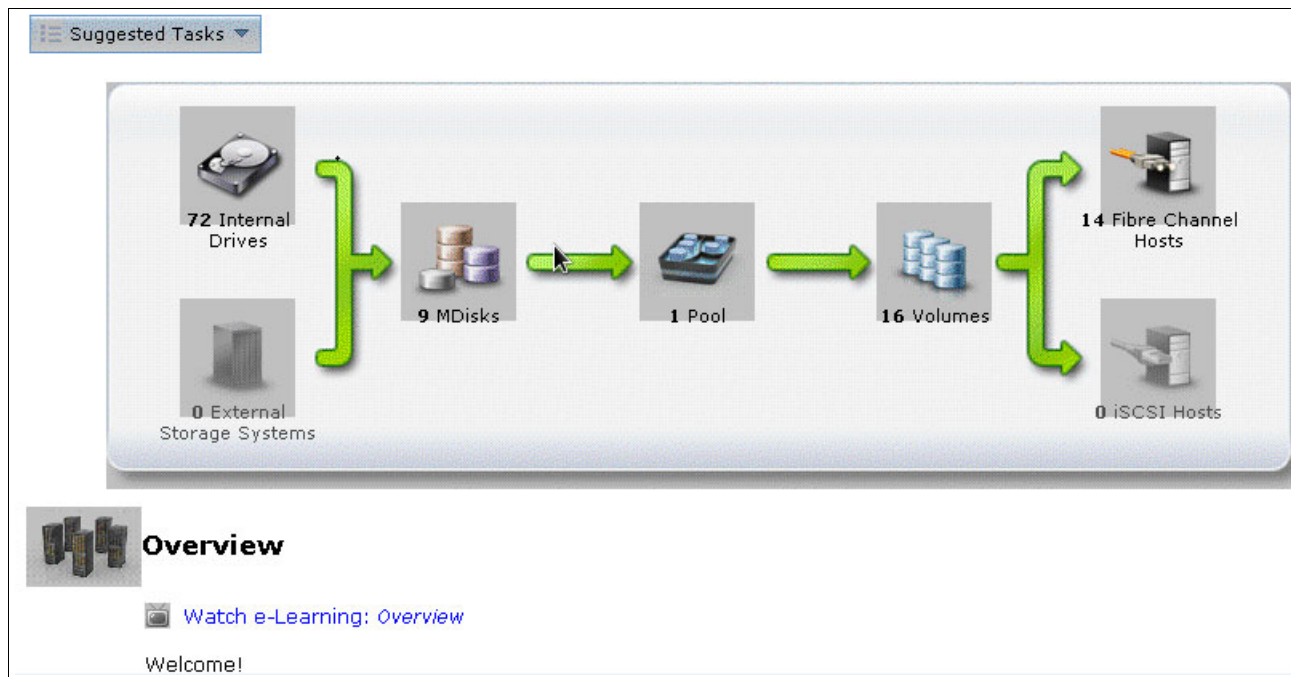


Figure 2-17 Storwize V7000 Overview panel

- Go to the storage pool panel to create a storage pool. Click the **Pools** image button, as shown Figure 2-18.

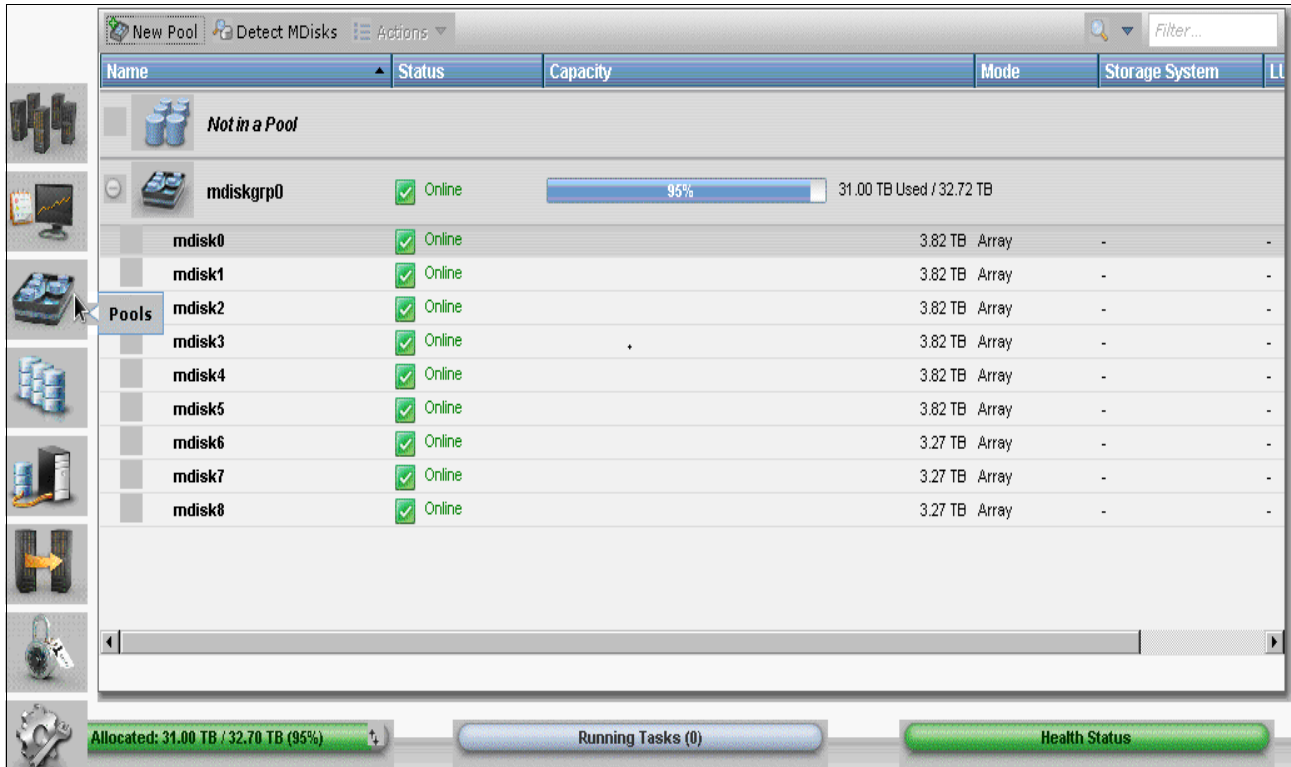


Figure 2-18 Selecting the IBM Storwize V7000 Pools panel

- To create a storage pool, click the **New Pool** button. In the Create Pool window, enter the pool name and click **Next**, as shown Figure 2-19.



Figure 2-19 Creating a pool

- Click **Detect MDisks** to show you the MDisks that are to be added to the storage pool, as shown in Figure 2-20.

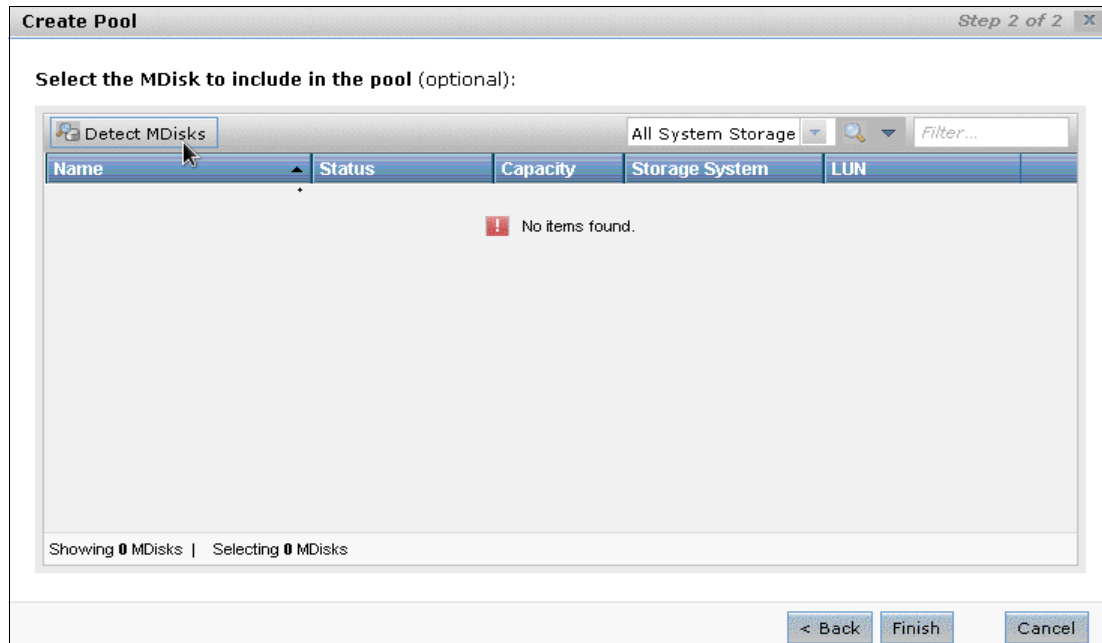


Figure 2-20 Detecting MDisks

- When the Discover Devices task completes, click **Close**, as shown in Figure 2-21.

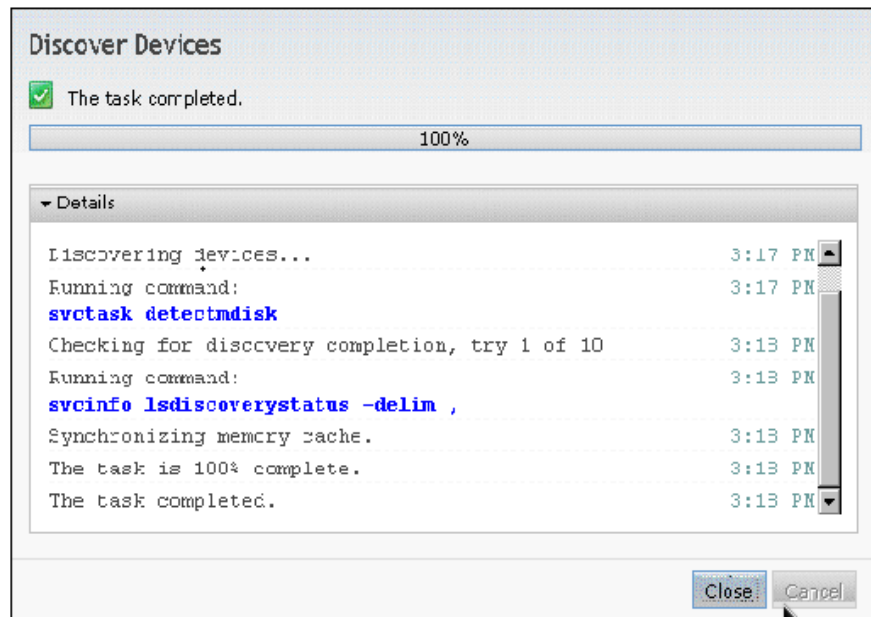


Figure 2-21 Discovering devices

- Select the MDisks and click **Finish** to create the storage pool.

The storage pool is created. The next steps describe how to create the volumes to be used by the systems.

2.3.2 Step 2: Creating a generic volume

Complete the following steps to create a generic volume:

1. Click the Volumes image and then select **Volumes**, as shown in Figure 2-22.



Figure 2-22 Volumes options

2. In the Volumes window, click **New Volume**. The New Volume window is displayed, as shown in Figure 2-23.

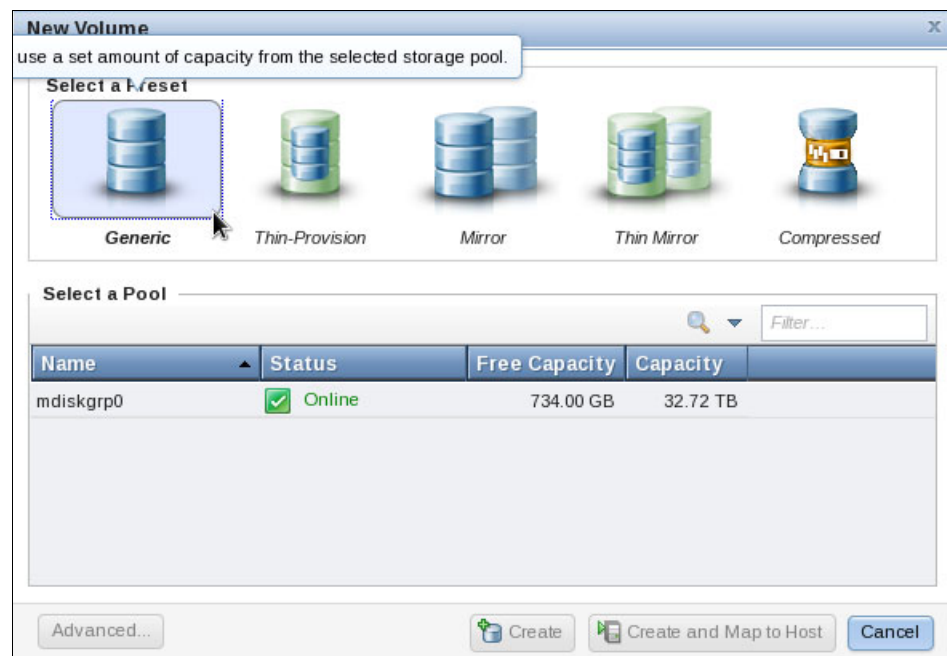


Figure 2-23 Creating a volume

- Click **Generic** and select the pool name. Enter the volume name and size, then click **Create**, as shown in Figure 2-24.

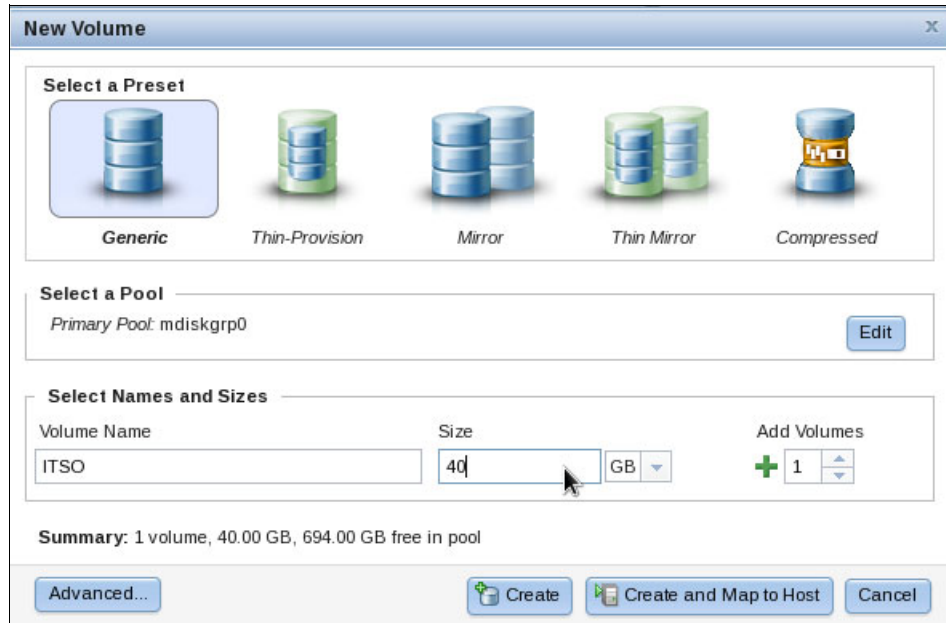


Figure 2-24 Creating a generic volume

- Figure 2-25 shows the output of the volume creation task. Click **Close**.

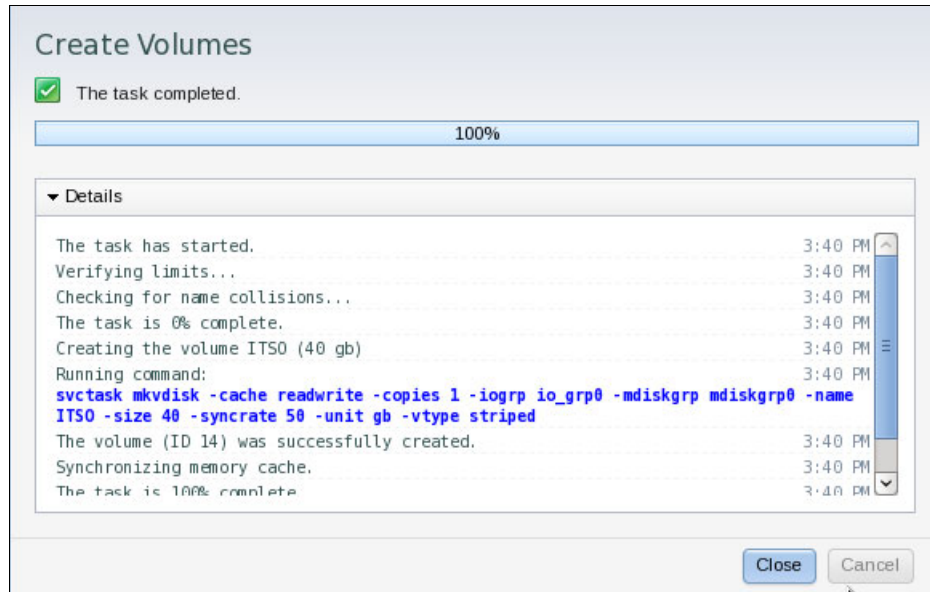


Figure 2-25 Volume creation task output

- In the Volumes window, you can see the new volume, as shown in Figure 2-26.

Name	Status	Capacity	Storage Pool	UID
ITSO	Online	40.00 GB	mdiskgrp0	60050768028100E7B80

Figure 2-26 Volume information

6. Right-click the volume to see the volume details, as shown in Figure 2-27.

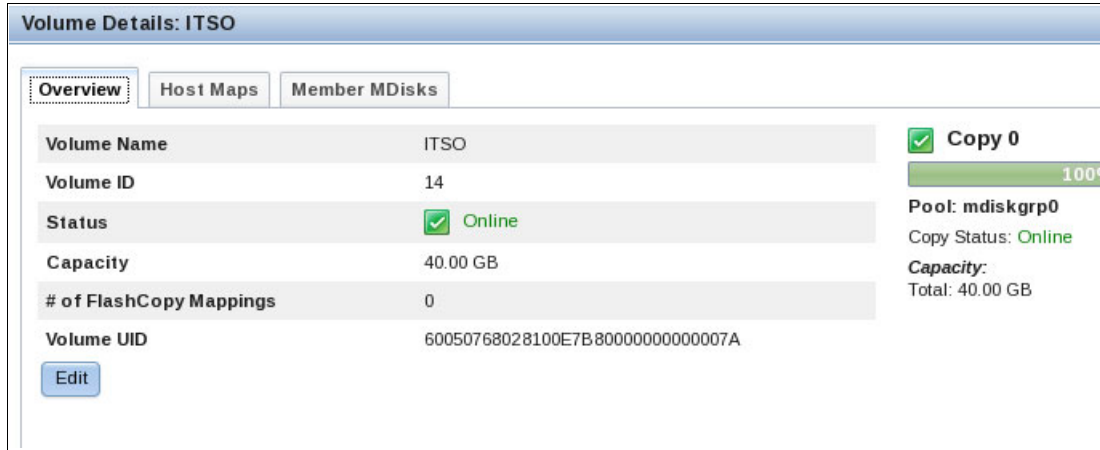


Figure 2-27 Displaying volume details

2.3.3 Step 3: Creating a compressed volume

This step is optional. IBM Storwize V7000 is designed to improve storage efficiency by compressing data through supported real-time compression for block storage. With real-time compression, you save a considerable amount of space. IBM Real-time Compression operates as data is written to disk, which avoids the need to store uncompressed data when waiting for compression. On the Storwize V7000, the IBM Real-time Compression feature is licensed by enclosures; on the IBM SAN Volume Controller, it is licensed by storage space (in terabytes).

For more information about hardware compression, see 2.4, “IBM Storwize V7000 and hardware compression” on page 37.

Complete the following steps if you want to create a compressed volume:

1. Click the Volumes image and select the **Volumes** option, as shown Figure 2-28.



Figure 2-28 IBM Storwize V7000 volumes options

2. At the top of the Volumes window, click **New Volume**.

3. In the New Volume window, click the **Compressed** icon to create a Compressed volume, as shown in Figure 2-29.

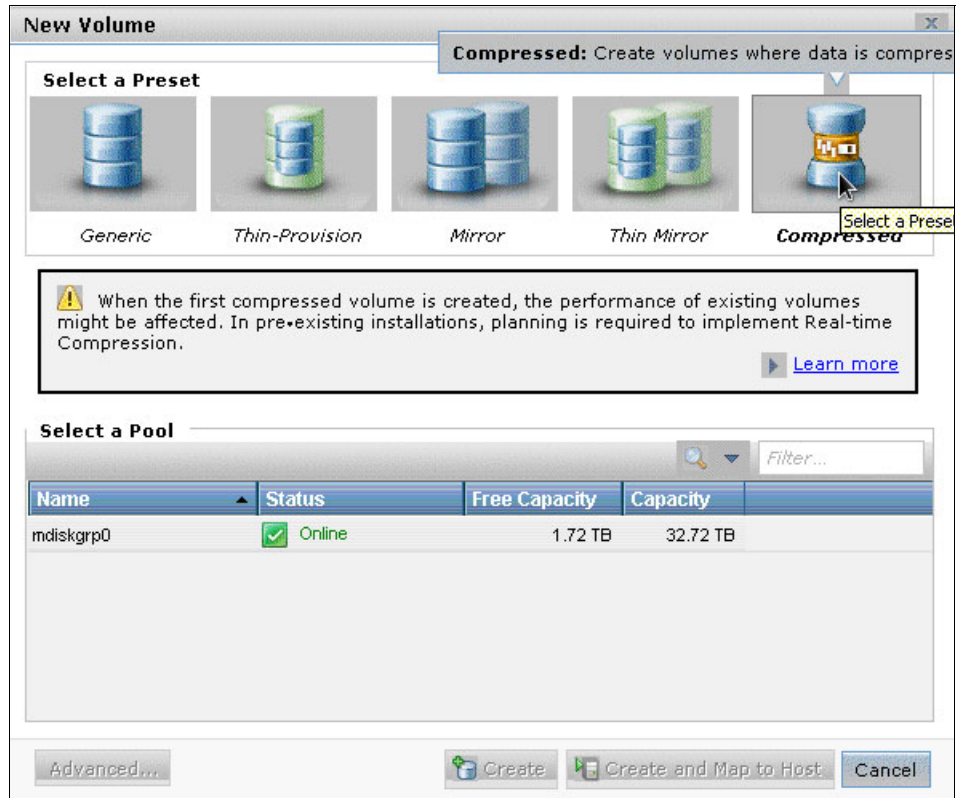


Figure 2-29 Creating a volume

- As shown in Figure 2-30, select the pool and enter the volume name and size. Click **Create**.

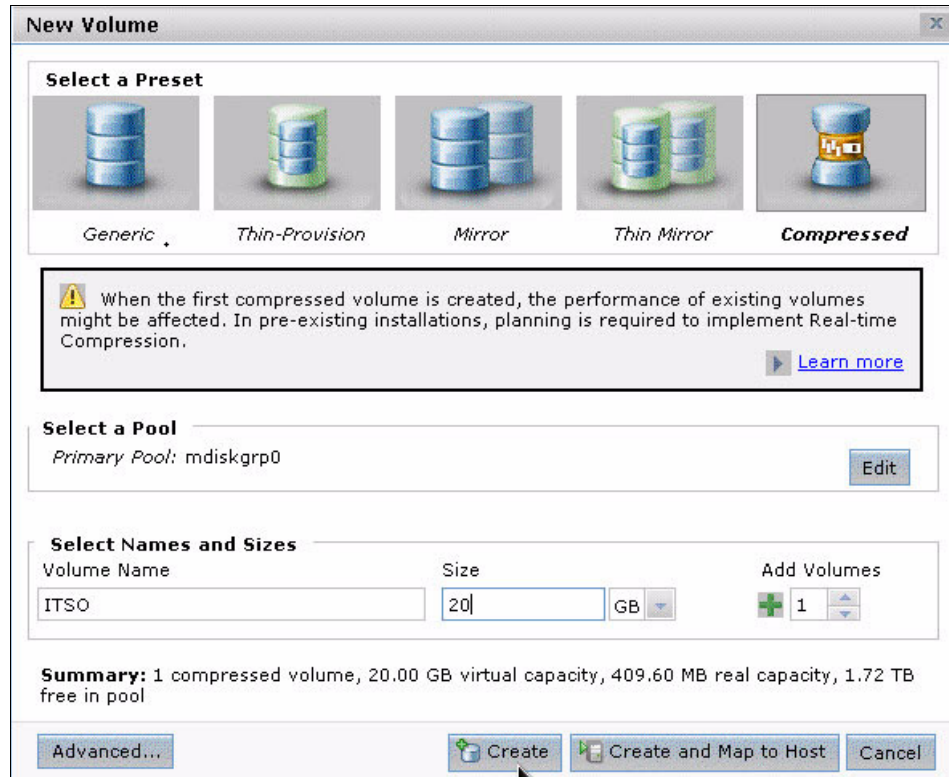


Figure 2-30 Creating a compressed volume

- After the volume is created, click **Close**, as shown in Figure 2-31.

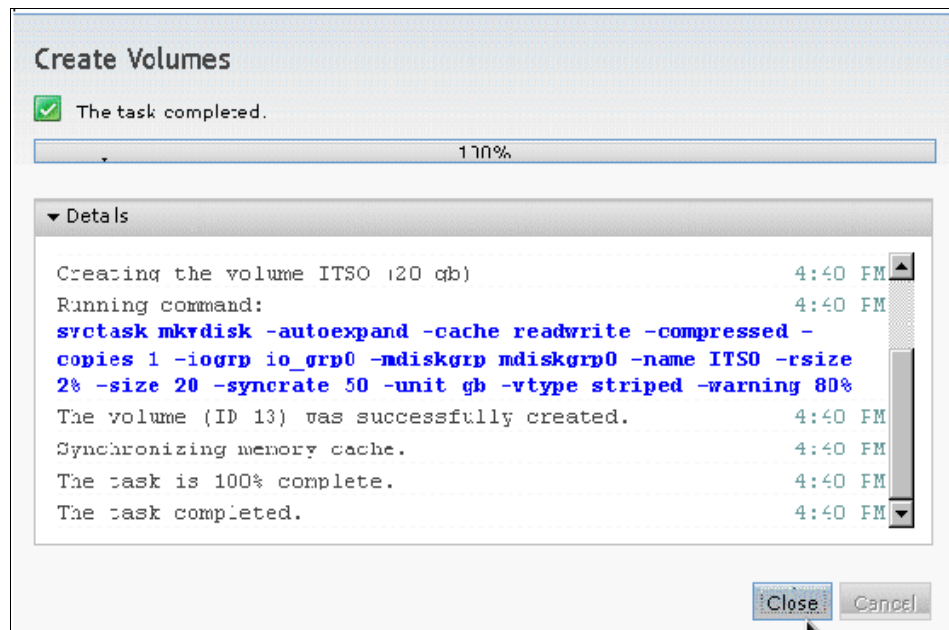


Figure 2-31 IBM Storwize V7000 Create Volumes task panel

6. In the Volume window, you see the new compressed volume, as shown in Figure 2-32.

Name	Status	Capacity	Storage Pool	UID	Host Mappings
ITSO	Online	20.00 GB	mdiskgrp0	60050768028100E7B800000000000078	No

Figure 2-32 IBM Storwize V7000 compressed volume

7. Right-click the new volume and select **Properties** to see the volume details, as shown in Figure 2-33.

Volume Details: ITSO

Overview | Host Maps | Member MDisks

Volume Name	ITSO	<input checked="" type="checkbox"/> Copy 0
Volume ID	13	<div style="width: 100%; height: 10px; background-color: #ccc; border: 1px solid #ccc;"></div>
Status	<input checked="" type="checkbox"/> Online	Pool: mdiskgrp0
Capacity	20.00 GB	Compressed, Striped
# of FlashCopy Mappings	0	Copy Status: Online
Volume UID	60050768028100E7B800000000000078	Easy Tier Status: Inactive
Caching I/O Group	io_grp0	Capacity:
Accessible I/O Groups	io_grp0	Used: 160.00 KB
Preferred Node	node2	Before Compression: 0 bytes
I/O Throttling	Disabled	Compression Savings: 0.00%
Mirror Sync Rate	50	Real: 425.60 MB
Cache Mode	Enabled	SSD Tier: 0 bytes
Cache State	Empty	HDD Tier: 425.60 MB
UDID (OpenVMS)		(Automatically Expand)
<input type="button" value="Edit"/>		Total: 20.00 GB
		Warning Threshold: 80 %

Show Details

Figure 2-33 IBM Storwize V7000 compressed volume details

2.3.4 Step 4: Creating a host

Complete the following steps to create a host:

1. Click the Hosts image and select **Hosts**, as shown in Figure 2-34.

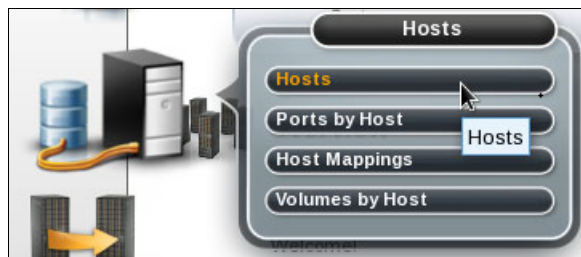


Figure 2-34 Hosts options

2. In the Hosts window, click **New Host**.

3. In the Create Host window, select **Fibre Channel Host** or **iSCSI Host**, as shown Figure 2-35.



Figure 2-35 Create Host options

4. In this example, **Fibre Channel Host** was selected. Enter the host name and add the WWPN. In the Advanced Settings section, select the settings that are related to your Storwize V7000 I/O Group configuration and then select the Host Type. Click **Create Host**, as shown in Figure 2-36.

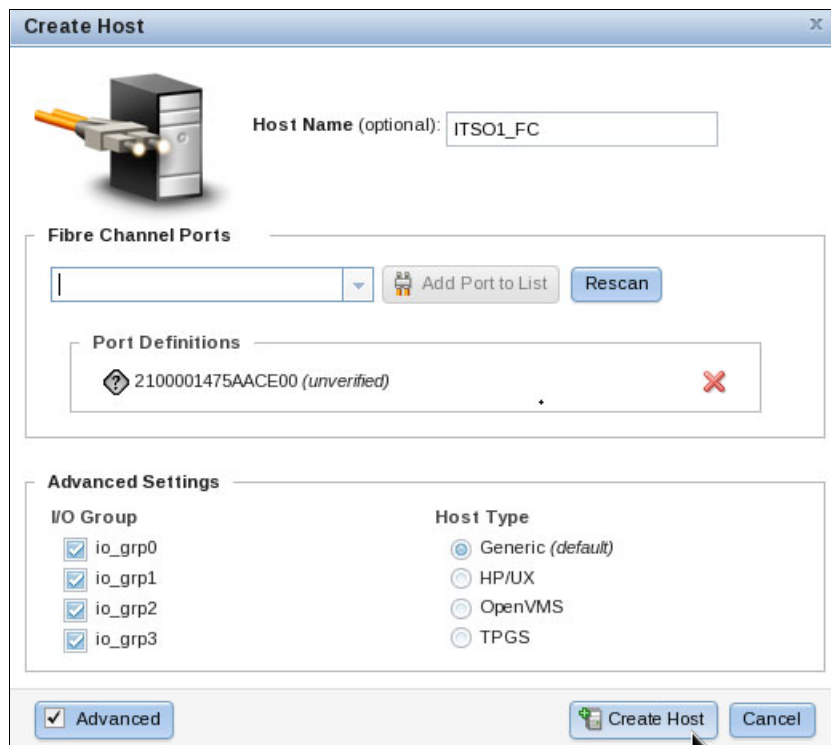


Figure 2-36 Creating a Fibre Channel host

- After the host is created, click **Close**, as shown in Figure 2-37.

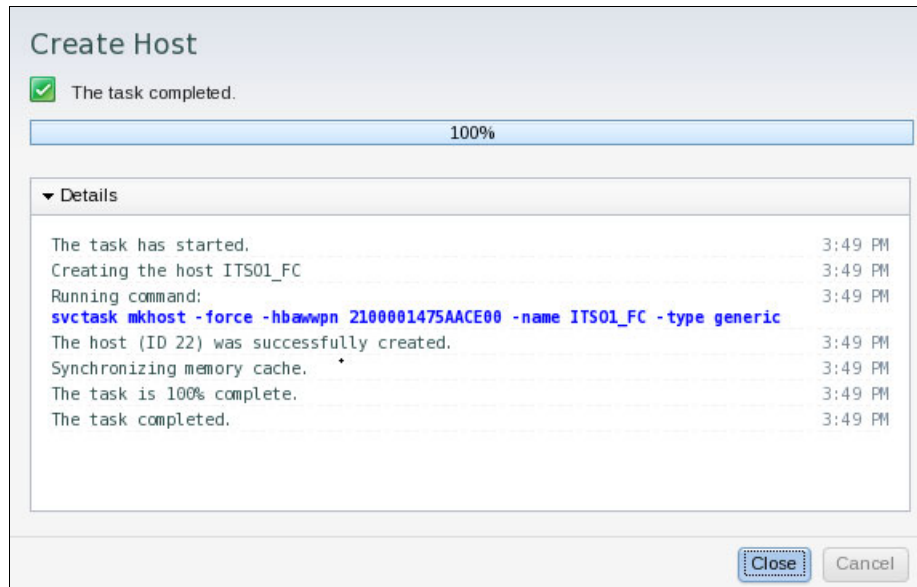


Figure 2-37 Create host task output

2.3.5 Step 5: Mapping the LUN to the host

Complete the following steps to map the LUN to the host:

- Click the Volumes image and select **Volumes**, as shown in Figure 2-38.



Figure 2-38 Volumes options

- In the Volumes window, right-click the volume and select **Map to Host**, as shown in Figure 2-39.

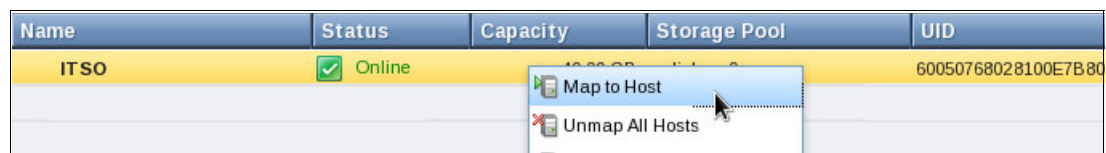


Figure 2-39 Volume mapping

- In the Modify Host Mappings window, select the host, select the volumes to be mapped, and click the right arrow to move the volume from the Unmapped Volumes column to the Volumes Mapped to the Host column. Click **Map Volumes**, as shown in Figure 2-40.

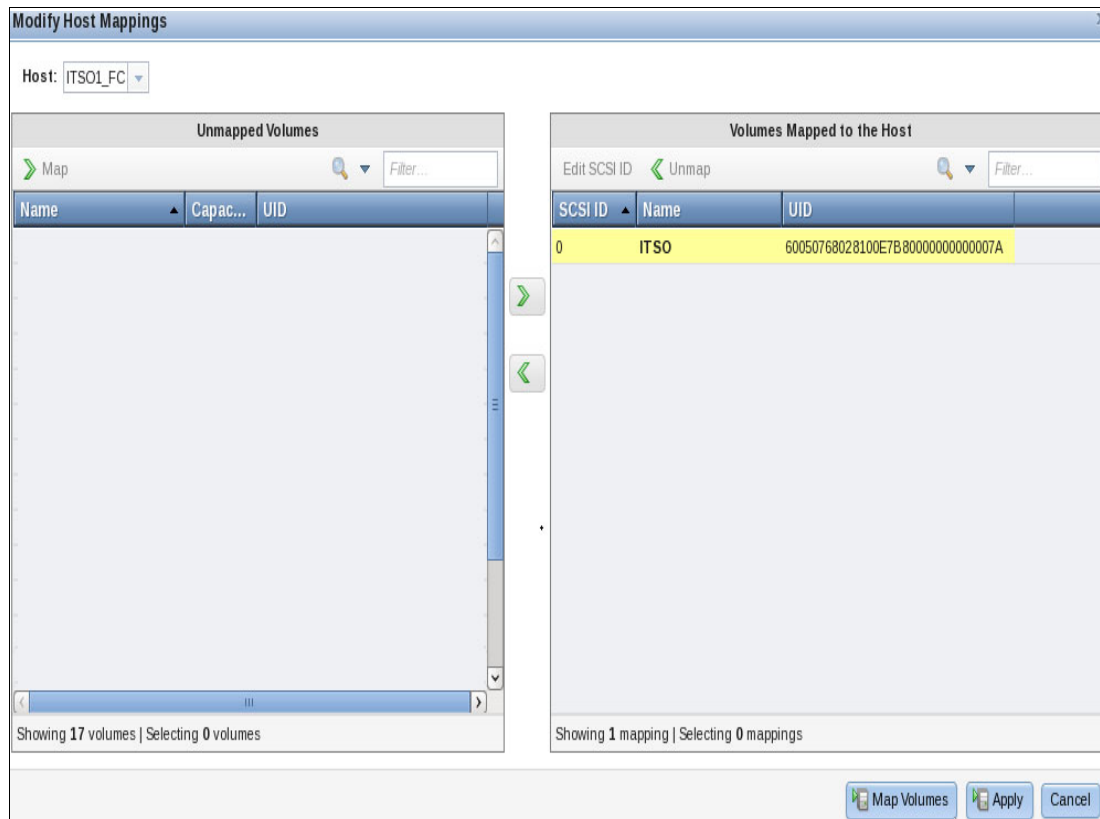


Figure 2-40 Mapping volumes

The volume is now mapped to the server. Ask the server administrator to verify that the new disks are now recognized.

2.4 IBM Storwize V7000 and hardware compression

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

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

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

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

<http://www-01.ibm.com/support/docview.wss?uid=ssg1S7003988&myns=s028&mynp=0CST3FR7&mync=E>

On average, an IBM i 7.1 system can achieve a data compression ratio of approximately 1:3, which is about a 66% of compression savings. For an example, see 2.4.2, “Creating a volume mirrored copy on a compressed volume” on page 39.

2.4.1 Using the IBM Comprestimator Utility

The IBM Comprestimator Utility is a host-based utility that can be used to provide an estimate for the achievable capacity savings by using Storwize V7000 compression and thin-provisioning. This utility is delivered as a binary executable file and is run from a host with access to the Storwize V7000 volumes to be analyzed. It performs read operations only.

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

<http://www-304.ibm.com/support/customer/sas/f/comprestimator/home.html>

Important: The IBM Comprestimator Utility does not run on IBM i; therefore, it can be used only with VIOS attached storage and not direct attached storage.

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

Example 2-1 Comprestimator output

```
# ./comprestimator -d /dev/hdisk1
Version: 1.2.01 (Build u0004)
Start time: 30/07/2012 03:50:58
Device name: /dev/hdisk1
Device size: 20.0 GB
Number of processes: 10
Exhaustive: no
```

Sample#	Device Name	Size(GB)	Compressed Size(GB)	Total Savings(GB)	Total Savings(%)	Thin Provisioning Savings(%)	Compression Savings(%)	Compression Accuracy Range(%)
3526	/dev/hdisk1	20.0	3.9	16.1	80.6%	22.9%	74.9%	4.7%

In this example of an IBM i load source of a 20 GB volume on the Storwize V7000, the utility estimates a thin-provisioning savings of 22.9% and a compression savings of 74.9%. This results in a compressed size of 3.9 GB, which is a total capacity savings of 80.6%. This result is the required physical storage capacity that is predicted to be reduced by 80.6% as a result of the combined savings from thin provisioning and compression.

Note: A Storwize V7000 compressed volume is always a thinly provisioned volume from an architectural perspective when it was created with a real size of 100%. Therefore, the reported compression savings are in addition to the savings that are provided by thin provisioning, which stores areas of zeros with minimal capacity.

Thin provisioning

IBM i can use thin provisioning because it is not apparent to the server. However, you must provide adequate hard disk drives (HDDs) to sustain the required performance regardless of whether the capacity is required. Also, IBM i 7.1 does not pre-format LUNs so that initial allocations can be thin-provisioned; therefore, there is no space reclamation and the effectiveness of the thin provisioning can decline over time.

You also must ensure that you have sufficient disks that are configured to maintain performance of the IBM i workload.

Thin provisioning might be more applicable to test or development environments.

2.4.2 Creating a volume mirrored copy on a compressed volume

IBM Storwize V7000 volume mirroring is required to migrate from a standard volume to a compressed volume. The following steps show how to create a volume (VDisk) copy from an existing non-compressed IBM i volume to a compressed volume for evaluating the new Storwize V7000 real-time compression feature; for example, in terms of compression ratio and performance:

1. From the V7000 GUI, right-click an existing non-compressed IBM i volume and select **Volume Copy Actions** → **Add Mirrored Copy** to create a VDisk mirrored copy on a compressed volume, as shown in Figure 2-41.

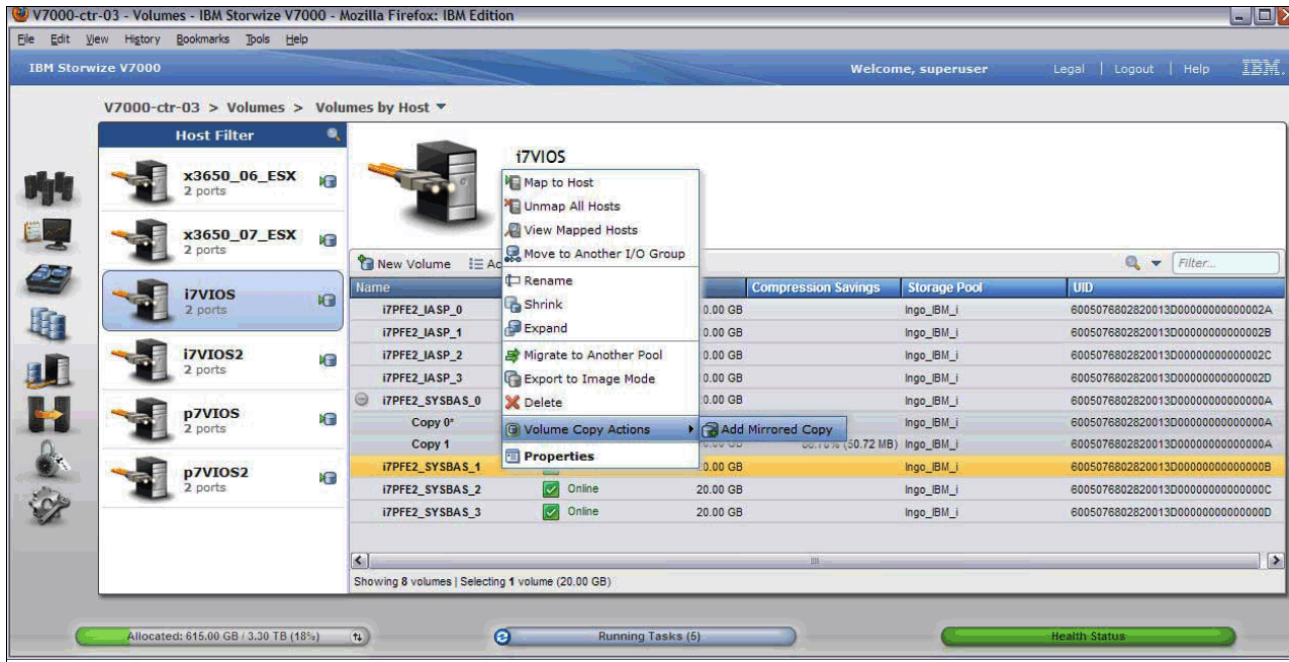


Figure 2-41 IBM Storwize V7000 adding a mirrored copy

- In the Add Volume Copy window, select **Compressed** as the volume type and a pool for the new VDisk copy. Click **Add Copy** to create a VDisk copy on a compressed volume, as shown in Figure 2-42.

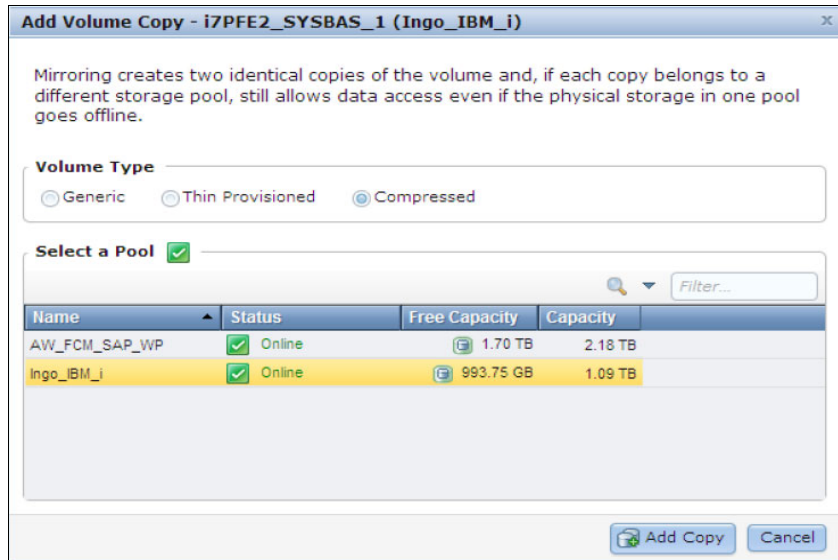


Figure 2-42 IBM Storwize V7000 adding a volume copy

As an alternative to the GUI procedure, the **addvdiskcopy** CLI command can be used for creating a compressed mirrored VDisk from an existing VDisk. The following CLI command shows an example of creating a compressed mirrored VDisk from an existing VDisk ID 10 in mdiskgroup 1, which initially allocates 2% real capacity and uses the thin-provisioning auto expand option:

```
svctask addvdiskcopy -mdiskgrp 1 -rsize 2% -autoexpand -compressed 10
```

The VDisk copy creation synchronization progress can be monitored from the Running Tasks information in the Storwize V7000 GUI or with the CLI command that is shown in Example 2-2, which also shows the estimated completion time.

Example 2-2 Compression progress

```
IBM_2076:V7000-ctr-03:admin>svcinfolsvdisksyncprogress 10
vdisk_id vdisk_name      copy_id progress estimated_completion_time
10      i7PFE2_SYSBAS_0 0      100
10      i7PFE2_SYSBAS_0 1      51      120625171149
```

In this example, VDisk 10 is a 20 GB LUN that is used as an IBM i 7.1 load source. From its 18,204 MB net capacity that is available to IBM i from the virtual SCSI disk unit, approximately 59.3% is used, as shown in Figure 2-43.

Work with Disk Status										17PFE2
Elapsed time: 00:00:00										10/16/13 12:02:19
Unit	Type	Size (M)	% Used	I/O Rqs	Request Size (K)	Read Rqs	Write Rqs	Read (K)	Write (K)	% Busy
1	6B22	19088	59.3	.0	.0	.0	.0	.0	.0	0
2	6B22	19088	51.2	.0	.0	.0	.0	.0	.0	0
3	6B22	19088	51.2	.0	.0	.0	.0	.0	.0	0
4	6B22	19088	51.2	.0	.0	.0	.0	.0	.0	0

Command
 ===>
 F3=Exit F5=Refresh F12=Cancel F24=More keys

Figure 2-43 IBM i Work with Disk Status display

From the audit log entry that is shown in Example 2-3, you can see that the mirrored copy creation was started at 14:21, with an estimated completion time of 17:11. This process takes about 2:50 hours to finish the 20 GB LUN that uses the default volume synchronization rate of 50%, which provides a synchronization data rate of approximately 2 MB.

Example 2-3 Audit log entry

```

IBM_2076:V7000-ctr-03:admin>svcinfol catauditlog -first 1
audit_seq_no timestamp cluster_user ssh_ip_address result res_obj_id action_cmd
819 120625142114 superuser 9.155.44.66 0 svctask addvdiskcopy
-mdiskgrp 1 -rsize 2% -warning 80 -autoexpand -compressed 10
  
```


You can review the VDisk volume properties by right-clicking the volume and selecting **Properties** from the context menu. In the Volume Details window, select **Show Details**. A compression savings for the IBM i load source disk unit of 65.5% is reported, as shown in Figure 2-44. Therefore, the total savings (including thin-provisioning) are $100\% \cdot 3.64 \text{ GB} / 20 \text{ GB} = 81.8\%$.

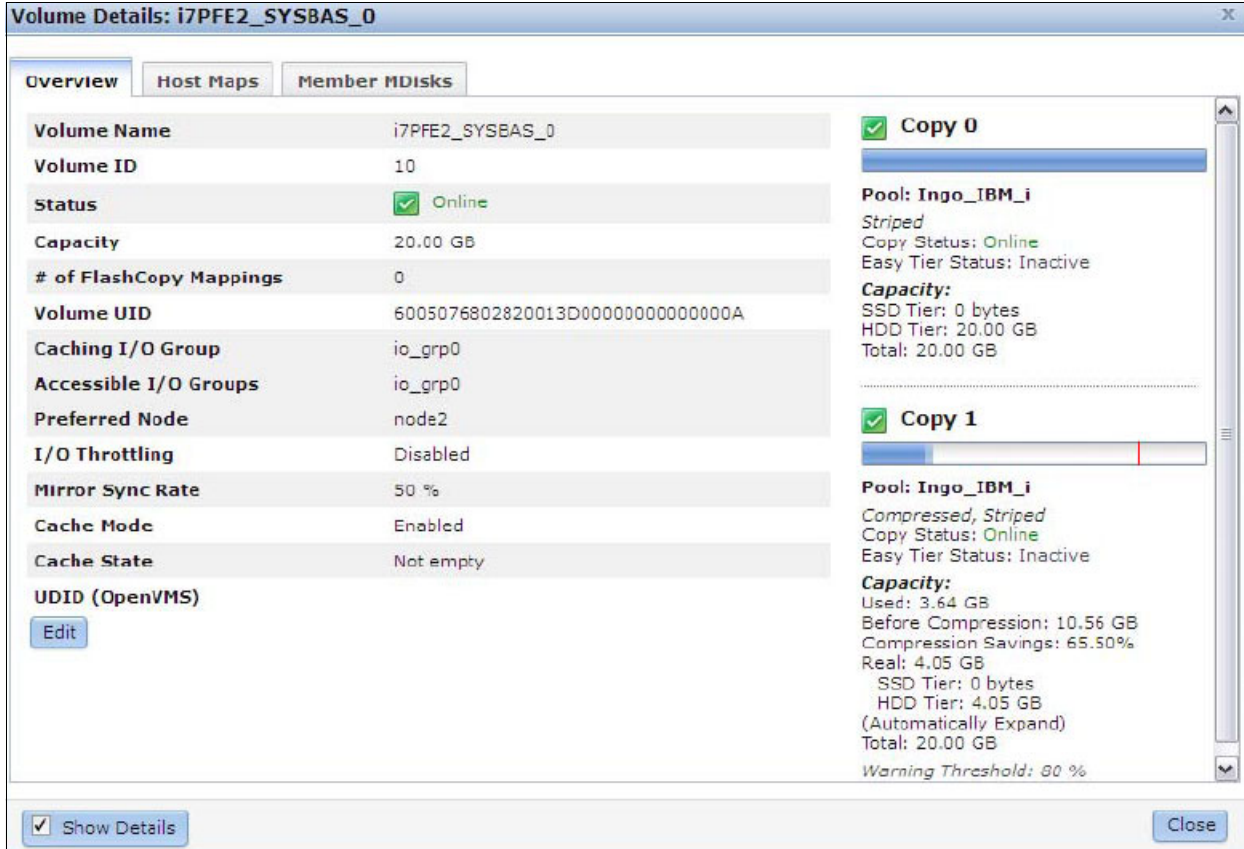


Figure 2-44 IBM Storwize V7000 volume details

Change the synchronization rate to 100% by clicking **Edit** or by using the `svctask chvdisk -syncrate 100 <vdisk_id>` command for the remaining three 20 GB LUNs from ASP1, which are on a V7000 RAID 10 array with 8x 10k RPM SAS drives that are each used by IBM i by 51.2%. This results in the following VDisk mirror creation times:

- ▶ 10 minutes for a single compressed volume with up to 16% V7000 RTC CPU usage
- ▶ 10 minutes for two compressed volumes with up to 32% V7000 RTC CPU usage

On average, approximately 66% of compression savings was achieved, which is a data compression ratio of approximately 1:3 for the IBM i 7.1 system as shown by the volume compression savings information that is highlighted in Figure 2-45 on page 43.

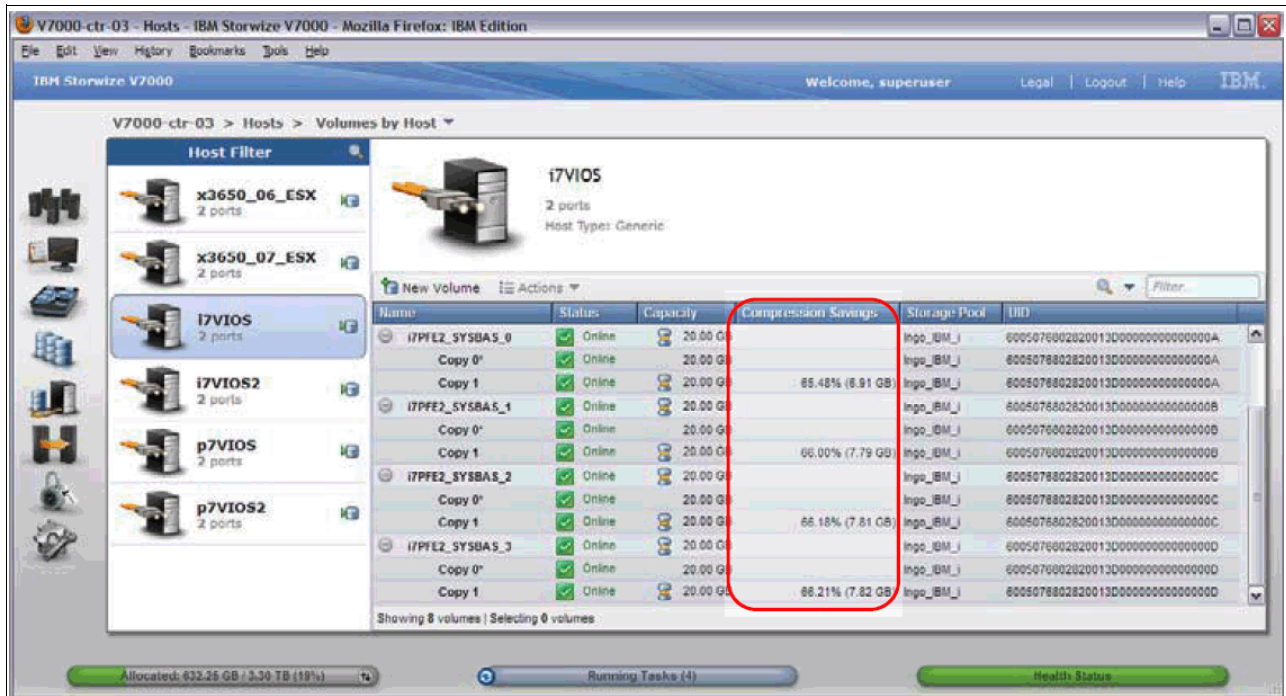


Figure 2-45 IBM Storwize V7000 Volumes by host panel

When the first compressed volume is created, some of the Storwize V7000 and SAN Volume Controller CPU cores and memory within the I/O group are reserved for use by the real-time compression engine. On a four-core system with Storwize V7000 or SAN Volume Controller CF8 nodes, three cores are reserved for the compression engine. On a six-core SAN Volume Controller CG8 nodes system, four cores are reserved.

These resources are returned for general system use again when the last compressed volume is deleted. It is considered a best practice to enable compression on four- or six-node systems only if the current system CPU usage is less than 25% for a four-node system and less than 50% for a six-node system. This can have significant performance implications to all hosts that are served by the storage unit.

2.4.3 Comparing the performance of compressed and non-compressed volumes

For comparing the performance of compressed and non-compressed volumes, an IBM i database test workload was used, which performs read and write I/Os on a 32 GB large DB2® database.

In the first scenario of using Storwize V7000 VDisk mirroring, both volume copies (that is, the compressed and non-compressed copies), were on the same RAID 10 rank that uses 8x 300 GB 10k RPM SAS drives. For the second scenario with a split VDisk mirroring, each copy was on its own dedicated RAID rank configuration.

Note: The small Storwize V7000 disk configuration that is used in this example was saturated by the IBM i workload. Therefore, it is not representative of a properly sized IBM i SAN storage configuration. However, this example still shows a compression performance evaluation.

Scenario 1: Volume mirroring

This scenario shows the achieved disk I/O performance results for the IBM i database workload with the mirrored VDisk configuration that was created as described earlier in this section. With the non-compressed volume flagged as the primary copy, approximately 899 read I/Os per second and 1165 write I/Os per second were achieved with a corresponding service time of 9 ms and 11 ms, as shown in Figure 2-46.

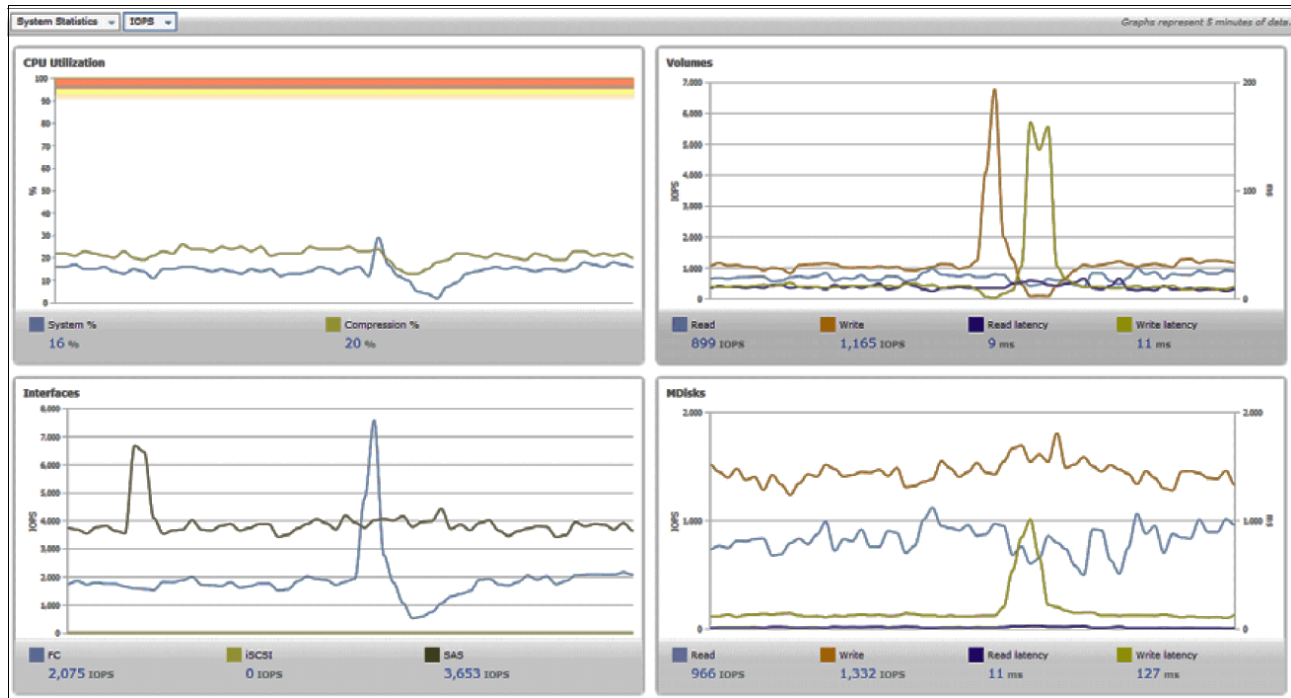


Figure 2-46 IBM Storwize V7000 performance for non-compressed primary mirrored VDisks

After the primary copy was switched, which designates the preferred copy of the mirrored VDisk to read from, from the non-compressed volume to the compressed volume by using the `svctask chvdisk ·primary <copy_ID> <vdisk_ID>` command, you can see a significant increase in read I/O performance with 4681 read I/Os per second and 1062 write I/Os per second with a service time of 1 ms and 12 ms, as shown in Figure 2-47 on page 45.

Note: Depending on the host workload characteristics, a significant read I/O performance improvement when compressed volumes are used can be achieved. Because of the compression for cache read misses and write de-stages from cache to disk, less data must be transferred.

This result can be verified by comparing the disk backend read I/O rate for the MDisks, which dropped from 966 read I/Os per second for the non-compressed primary (see Figure 2-46) to only 181 read I/Os per second for the compressed primary, as shown in Figure 2-47 on page 45.

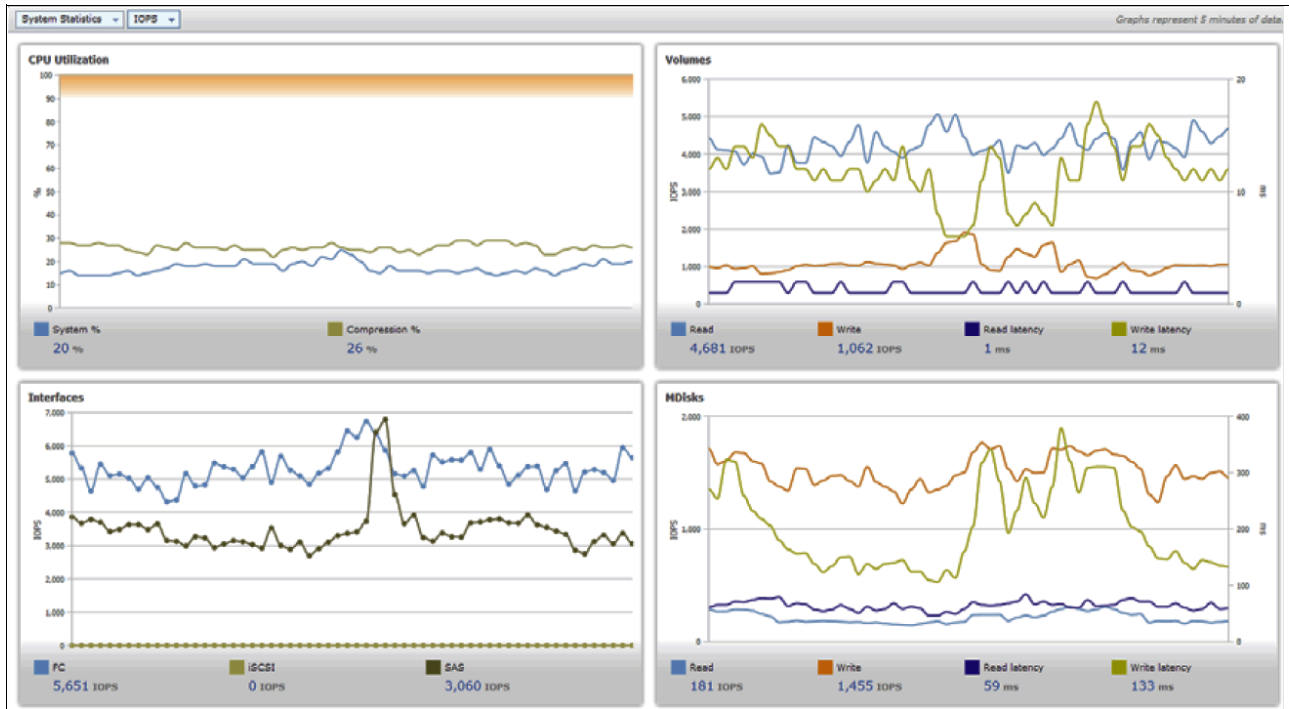


Figure 2-47 IBM Storwize V7000 performance for compressed primary mirrored VDisks

Scenario 2: Split volume mirroring

As described in “Scenario 1: Volume mirroring” on page 44, which used VDisk mirroring to create a compressed copy of a non-compressed volume, this can be seen as an interim solution for testing the efficiency of the Storwize V7000 compression engine before deciding to stay with independent (that is non-mirrored) compressed or non-compressed volumes.

In this section, a second scenario of comparing the Storwize V7000 disk I/O performance with IBM i database workload for non-mirrored non-compressed versus non-mirrored compressed volumes is reviewed. The previously mirrored volumes were split into two independent copies by migrating one mirrored copy to another equally configured RAID rank (MDisk) in another storage pool (mdiskgrp) by using the `svctask migratevdisk` command and by splitting the mirror by using the `svctask splitvdiskcopy` command.

IBM i storage management uses I/O bundling to limit the amount of physical disk I/O. That is, multiple logical I/Os can be coalesced into a single physical disk I/O with a larger transfer size. The degree of IBM i I/O bundling also depends on the disk I/O service times. From this perspective, the IBM i disk I/O transfer sizes can vary for different storage configurations such that it is more accurate here to compare the achieved data rates (MBps) than the I/O throughput (I/O per second).

Figure 2-48 shows the IBM i Work with Disk Status display for the case of the use of non-compressed volumes.

Work with Disk Status										17PFE2
Elapsed time: 00:07:46										10/16/13 14:08:39
Unit	Type	Size (M)	% Used	I/O Rqs	Request Size (K)	Read Rqs	Write Rqs	Read (K)	Write (K)	% Busy
1	6B22	19088	78.6	1564	26.2	1097	466.9	4.6	77.0	79
2	6B22	19088	74.0	1857	26.0	1302	555.3	4.5	76.4	80
3	6B22	19088	74.0	1851	26.0	1303	548.1	4.5	77.2	76
4	6B22	19088	74.0	1869	25.8	1302	566.7	4.5	74.7	78

Command
 ===>
 F3=Exit F5=Refresh F12=Cancel F24=More keys

Figure 2-48 Work with Disk Status display, non-compressed volumes

Figure 2-49 shows the Work with Disk Status display for compressed volumes, which shows the IBM i I/O bundling effect with slightly smaller write transfers being used for the lower performing non-compressed volumes.

Work with Disk Status										17PFE2
Elapsed time: 00:07:46										10/16/13 14:08:39
Unit	Type	Size (M)	% Used	I/O Rqs	Request Size (K)	Read Rqs	Write Rqs	Read (K)	Write (K)	% Busy
1	6B22	19088	78.7	2310	23.3	1744	566.5	4.5	81.3	66
2	6B22	19088	74.1	2714	23.3	2047	667.6	4.5	81.1	69
3	6B22	19088	74.1	2710	23.4	2050	659.2	4.5	82.3	68
4	6B22	19088	74.1	2744	23.2	2048	695.7	4.5	78.3	68

Command
 ===>
 F3=Exit F5=Refresh F12=Cancel F24=More keys

Figure 2-49 Work with Disk Status display, compressed volumes

By using non-compressed volumes for the IBM i database workload, looking at the 7-minute averages from Figure 2-48 on page 46, 5004 read I/Os per second (22 MBps) and 2137 write I/Os per second (159 MBps) with services times of 2 ms and 5 ms were achieved, as shown in Figure 2-50.

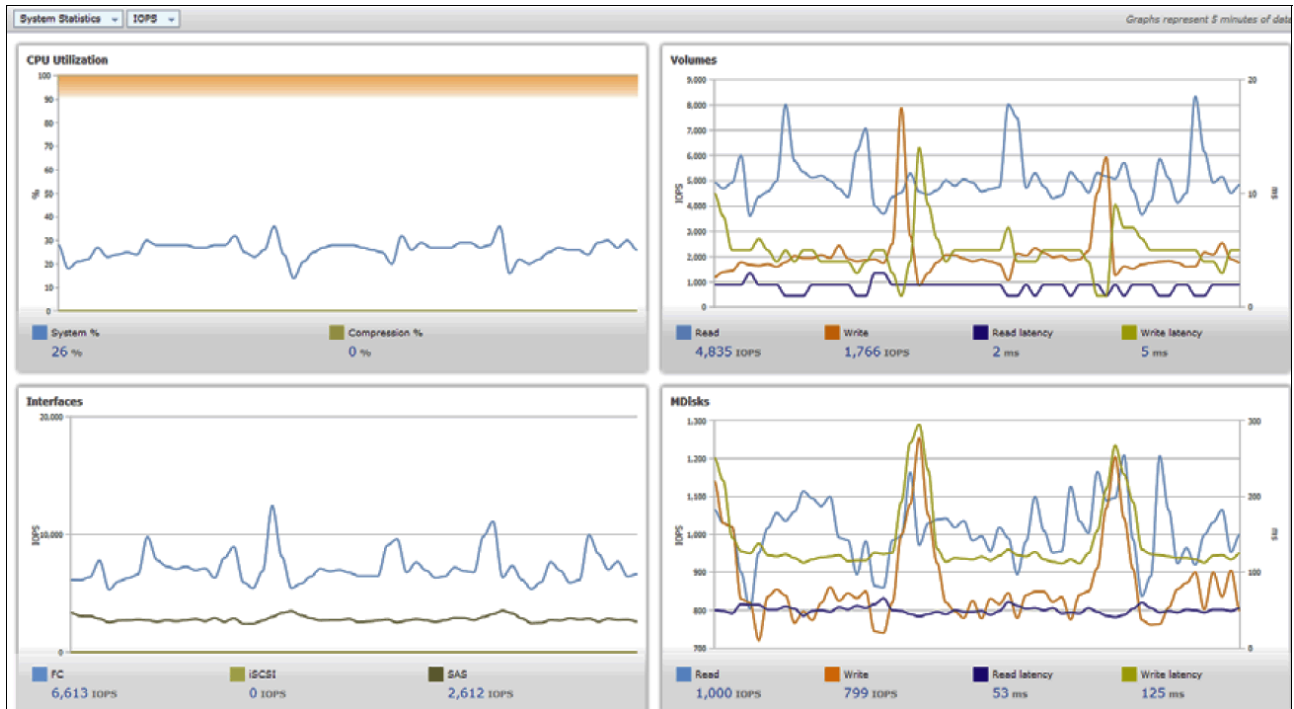


Figure 2-50 IBM Storwize V7000 performance for split mirror uncompressed volumes (IOPS)

This is compared to looking at the 5-minute averages from Figure 2-49 on page 46, in which 7889 read I/Os per second (35 MBps) and 2589 write I/Os per second (204 MBps) with 0 ms and 4 ms was achieved with the compressed volumes, as shown in Figure 2-51.

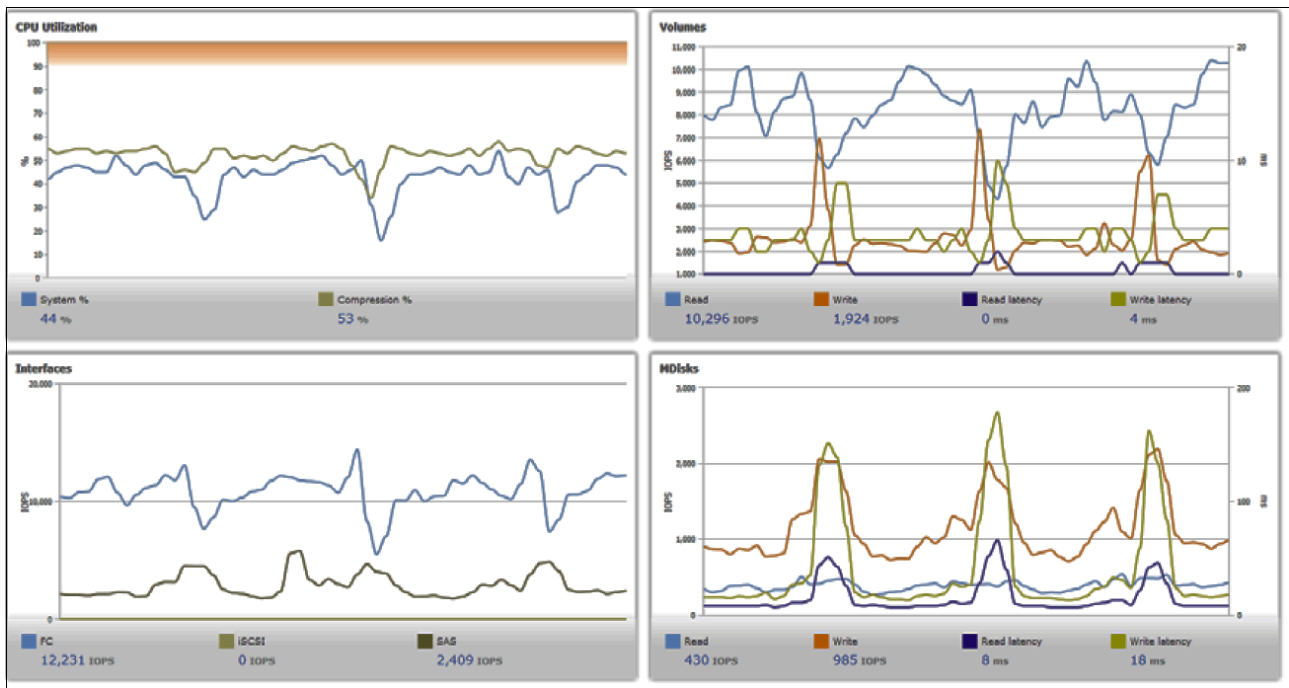


Figure 2-51 IBM Storwize V7000 performance for split mirror compressed volumes (IOPS)

As described in “Scenario 1: Volume mirroring” on page 44, significant disk I/O performance improvements are possible by using compression when the V7000/SAN Volume Controller disk backend is already highly used. This is due to the compression in which less data must be read and written from and to the disk backend, which causes a lower disk usage rate and, therefore, better disk response times.

Looking at the V7000 disk backend data rates, you can see a significant reduction of the backend workload from 78 MBps reads and 172 MB/ps for the case of non-compressed volumes that are shown in Figure 2-52 on page 49. This is down to only 14 MBps reads and 16 ps writes for compressed volumes, as shown in Figure 2-53 on page 49.

However, as can also be seen from Figure 2-53 on page 49, with the system CPU utilization reaching around 50% and the real-time compression CPU utilization peaking at 56%, a proper sizing of the Storwize V7000 system with regards to CPU utilization as described in 2.4.2, “Creating a volume mirrored copy on a compressed volume” on page 39 becomes more important when the compression feature is used.



Figure 2-52 IBM Storwise V7000 performance for split mirror non-compressed volumes (MBps)

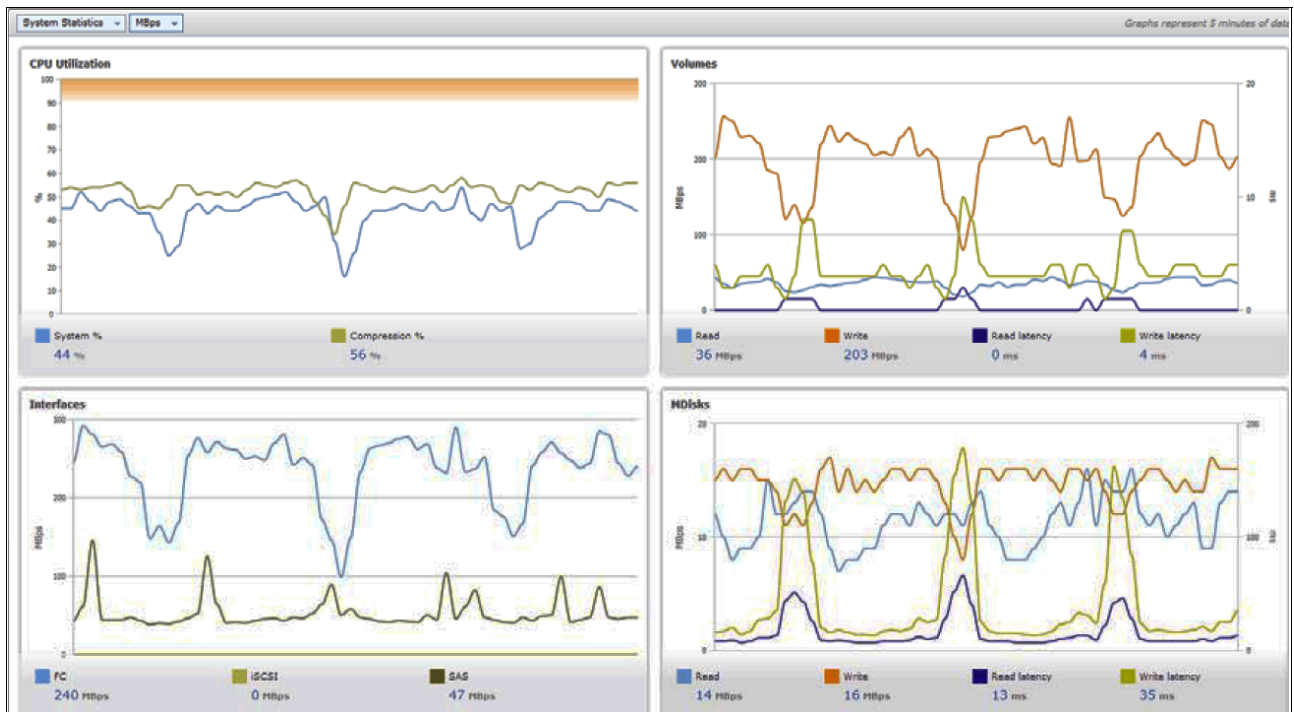


Figure 2-53 IBM Storwise V7000 performance for split mirror compressed volumes (MBps)

Conclusion

In this evaluation of the Storwize V7000 V6.4 real-time compression feature with IBM i database workload, an average compression savings of 66% was achieved. This compression ratio of 1:3 can be considered representative because it also aligns well with the typical IBM i tape drive compaction ratio.

From a performance perspective for the tested configuration and workload, a 23% increase in IBM i database disk throughput with even lower response times was observed when V7000 compressed volumes were used versus non-compressed volumes. This improvement was gained for a V7000/SAN Volume Controller disk backend, which was already highly utilized when the non-compressed volumes were used.

A significant performance improvement by using real-time compression can be expected only for already highly utilized disk backends and is achieved because of the compression, in which less data must be read from and written to the disk backend. However, the reduction of the disk backend utilization comes with a higher V7000/SAN Volume Controller CPU utilization. This is due to the processing power that is required for the software real-time compression engine so that for a proper V7000/SAN Volume Controller, sizing the aspect of CPU utilization is not neglected.

2.5 Tips for implementing IBM Storwize V7000 in a IBM i environment

This section describes the following useful tips for implementing the IBM Storwize V7000 in a IBM i environment:

- ▶ 2.5.1, “Storwize V7000 configuration options” on page 50
- ▶ 2.5.2, “Host attachment” on page 51
- ▶ 2.5.3, “Multipath” on page 52
- ▶ 2.5.4, “Zoning SAN switches” on page 55
- ▶ 2.5.5, “Data layout” on page 57
- ▶ 2.5.6, “Solid-State Drives” on page 59
- ▶ 2.5.7, “Bootting from SAN” on page 60
- ▶ 2.5.8, “Sizing for performance” on page 60
- ▶ 2.5.9, “Performance monitoring” on page 61
- ▶ 2.5.10, “Copy services considerations” on page 61
- ▶ 2.5.11, “Translation from 520-byte blocks to 512-byte blocks” on page 62

2.5.1 Storwize V7000 configuration options

Different hardware and RAID options are available for the Storwize V7000 and can be validated in terms of performance by using the Disk Magic sizing tool. You can configure the RAID level and array width according to the solution that you modeled in Disk Magic tool. It is considered a best practice to follow the default configuration options as outlined in the Storwize V7000 GUI configuration wizard.

The default configuration option in the Storwize V7000 GUI is RAID 5 with a default array width of 7+P for SAS HDDs, RAID 6 for Nearline HDDs with a default array width of 10+P+Q, and RAID 1 with a default array width of 2 for SSDs.

It is considered a best practice to create a dedicated storage pool for IBM i with enough managed disks that are backed by enough spindles to handle the expected IBM i workload. Modeling with the Disk Magic tool by using actual client performance data can be performed to size the storage system properly.

Note: Disk Magic is a tool that is used to size IBM disk subsystems that are based on the performance and objectives of the workloads that are hosted on the disk subsystem. IBM Disk Magic is a product to be used only by IBM personnel or an IBM Business Partner.

For more information about Disk Magic sizing for IBM i, see Chapter 5, Sizing for Midrange Storage, in the IBM Redbooks publication *IBM i and Midrange External Storage*, SG24-7668.

2.5.2 Host attachment

Figure 2-54 shows that an IBM i system only logs in to a Storwize V7000 node a single time from an I/O port adapter on the IBM i LPAR.

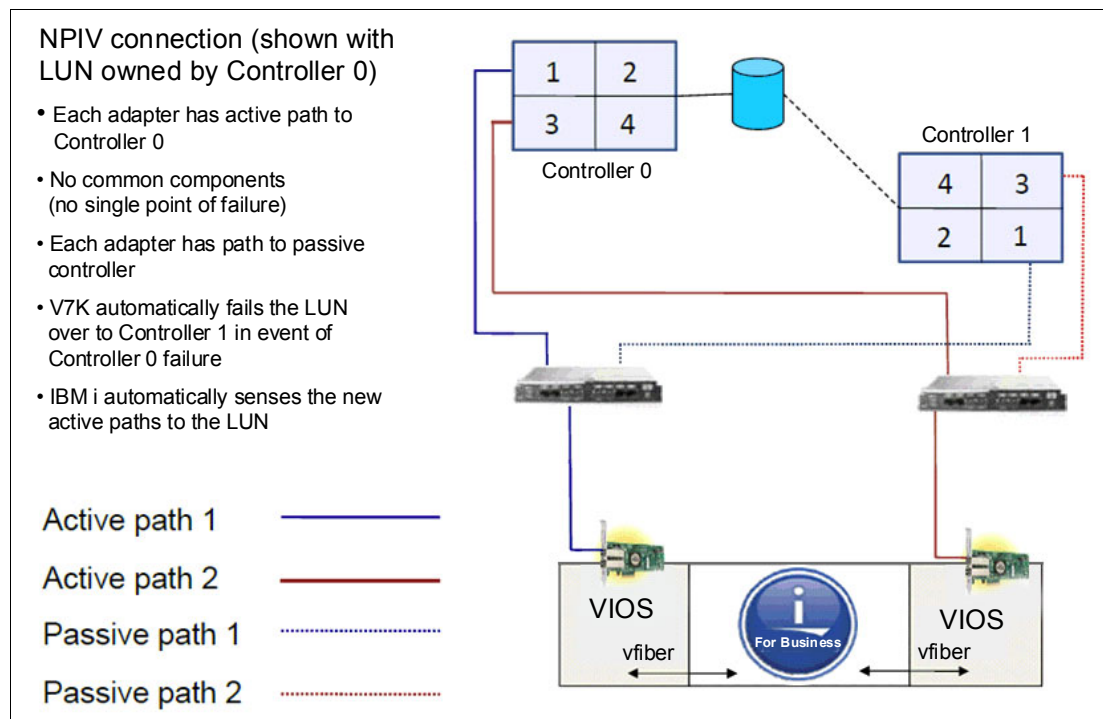


Figure 2-54 NPIV connection

The same connection considerations apply when you are connecting by using the native connection option without VIOS. The following best practices guidelines are useful:

- ▶ Isolate host connections from remote copy connections (Metro Mirror or Global Mirror) where possible.
- ▶ Isolate other host connections from IBM i host connections on a host port basis.
- ▶ Always have symmetric pathing by connection type. That is, use the same number of paths on all host adapters that are used by each connection type.
- ▶ Size the number of host adapters that are needed based on expected aggregate maximum bandwidth and maximum IOPS (use the Disk Magic tool or other common sizing methods that are based on actual or expected workload).

2.5.3 Multipath

Multipath provides greater resiliency for SAN-attached storage. IBM i supports up to eight paths to each LUN. In addition to the availability considerations, laboratory performance testing shows that two or more paths provide performance improvements when compared to a single path. Typically, two paths to a LUN is the ideal balance of price and performance.

Disk Magic tool: The Disk Magic tool supports multipathing only over two paths.

You might want to consider more than two paths for workloads where there is a high wait time or where high I/O rates are expected to LUNs.

Multipath for a LUN is achieved by connecting the LUN to two or more ports that belong to different adapters in an IBM i partition, as shown in the following examples:

- ▶ With native connection, although it is not a requirement, it is considered a recommended practice for availability reasons to have the ports for multipath be on different physical adapters in IBM i.
- ▶ With VIOS NPIV connection, the virtual Fibre Channel adapters for multipath should be evenly spread across the different VIOS.
- ▶ With VIOS vSCSI connection, the virtual SCSI adapters for multipath should be evenly spread across the different VIOS.

Figure 2-55 shows a high-level view of multipath in different types of connections.

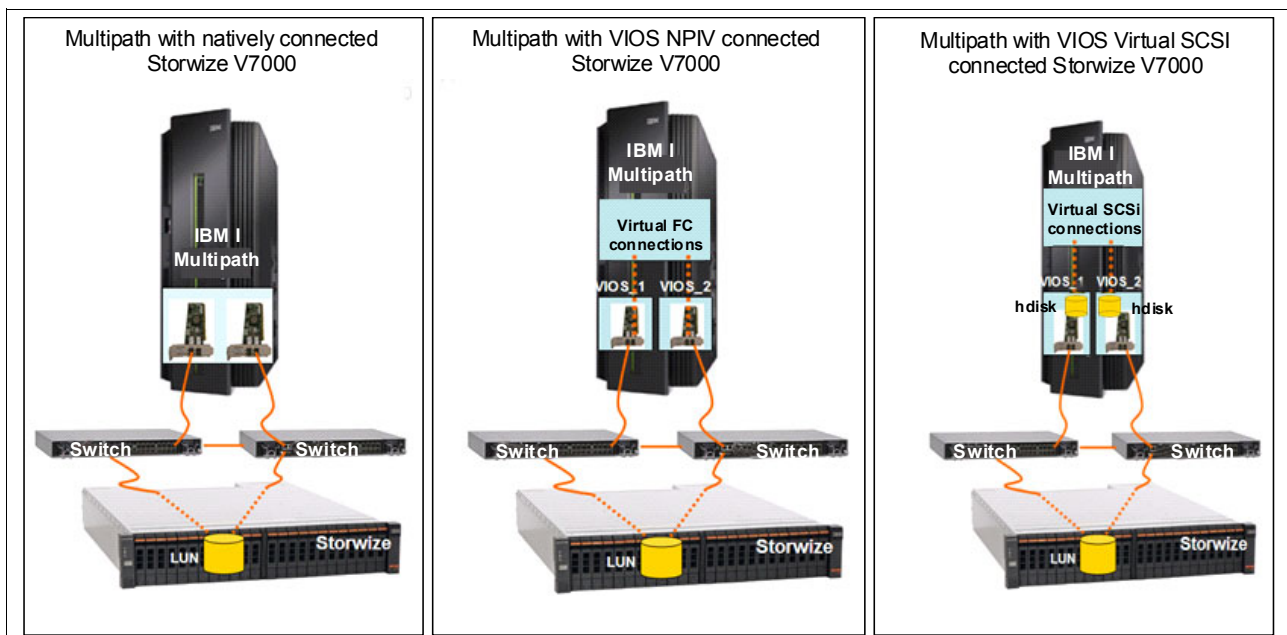


Figure 2-55 IBM i multipath options

IBM i multipath provides resiliency in case the hardware for one of the paths fails. It also provides performance improvement as multipath uses I/O load balancing in a round-robin mode among the paths.

Every LUN in Storwize V7000 uses one V7000 node as a preferred node. The I/O traffic to or from the particular LUN normally goes through the preferred node. If that node fails, the I/O is transferred to the remaining node.

With IBM i multipath, all the paths to a LUN through the preferred node are active and the path through the non-preferred node is passive. Multipath employs the load balancing among the paths to a LUN that go through the node, which is preferred for that LUN.

Multipath with native and VIOS NPIV connections

In native and VIOS NPIV connections, multipath is achieved by assigning the same LUN to multiple physical or virtual FC adapters in IBM i. The same LUN is assigned to multiple WWPNs, each from one port in a physical or virtual FC adapter, with each virtual FC adapter assigned to different VIOS.

For clarity, this description is limited to a multipath with two WWPNs. By using the preferred practice of switch zoning, you can achieve the situation where four paths are established from a LUN to the IBM i where two of the paths going through adapter 1 (in NPIV also through VIOS 1) and two of the paths going through adapter 2 (in NPIV also through VIOS 2). From the two paths that go through each adapter, one goes through the preferred node and other goes through the non-preferred node. Therefore, two of the four paths are active, each of them going through different adapter, and different VIOS if NPIV is used. Two of the paths are passive, each of them going through different adapter, and different VIOS if NPIV is used. IBM i multipathing uses a round-robin algorithm to balance the I/O among the paths that are active.

Figure 2-56 shows the detailed view of paths in multipath with native or VIOS NPIV connection to the Storwize V7000. The solid lines refer to active paths, and the dotted lines refer to the passive paths. The red lines present one switch zone and green lines present the other switch zone.

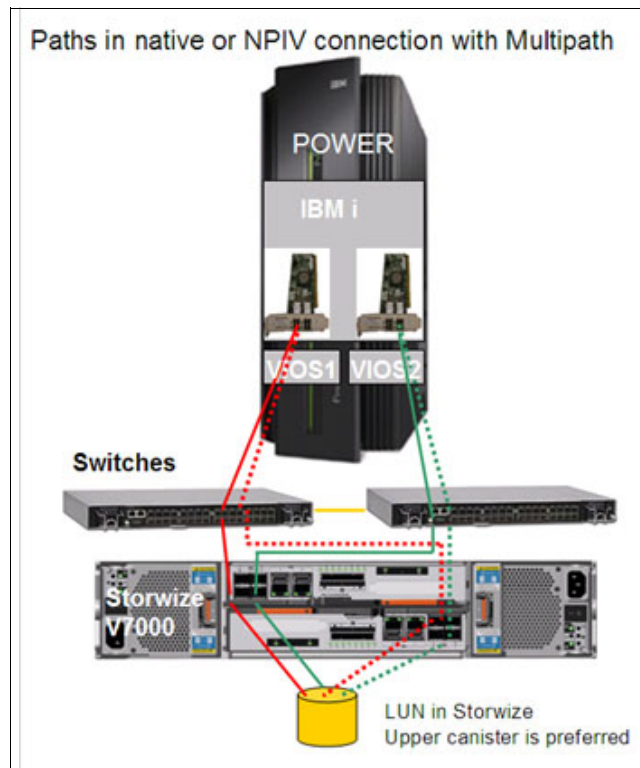


Figure 2-56 Native or NPIV connection multipath view

Figure 2-57 show the IBM i view of paths to the LUN that is connected in VIOS NPIV. You can see two active and two passive paths are established to each LUN.

Display Disk Path Status						
ASP	Unit	Serial Number	Type	Model	Resource Name	Path Status
1	1	YF033F00001D	2145	050	DMP001	Active
					DMP007	Passive
					DMP008	Active
1	2	YF033F00001E	2145	050	DMP002	Passive
					DMP003	Passive
					DMP010	Active
1	3	YF033F00001F	2145	050	DMP005	Active
					DMP009	Passive
					DMP004	Active
					DMP012	Passive
					DMP006	Passive
					DMP011	Active

Press Enter to continue.

F3=Exit F5=Refresh F9=Display disk unit details
F11=Display encryption status F12=Cancel

Figure 2-57 Displaying disk path status

Multipath with a VIOS vSCSI connection

In this type of connection, the LUN in Storwize V7000 is assigned to multiple VIOS. IBM i establishes one path to the LUN through each VIOS. For clarity, this description is limited to a multipath with two VIOS.

The LUN reports as device hdisk in each VIOS. The I/O rate from VIOS (device hdisk) to the LUN uses all the paths that are established from VIOS to Storwize V7000. Multipath across these paths and load balancing and I/O through a preferred node are handled by the VIOS multipath driver. The two hdisks that represent the LUN in each VIOS are mapped to IBM i through different virtual SCSI adapters. Each of the hdisks reports in to IBM i as a different path to the same LUN (disk unit). IBM i establishes multipath to the LUN by using both paths. Both paths are active and the load balancing in a round-robin algorithm is used for the I/O traffic.

Figure 2-58 shows the detailed view of paths in multipath with VIOS vSCSI connected to IBM Storwize V7000.

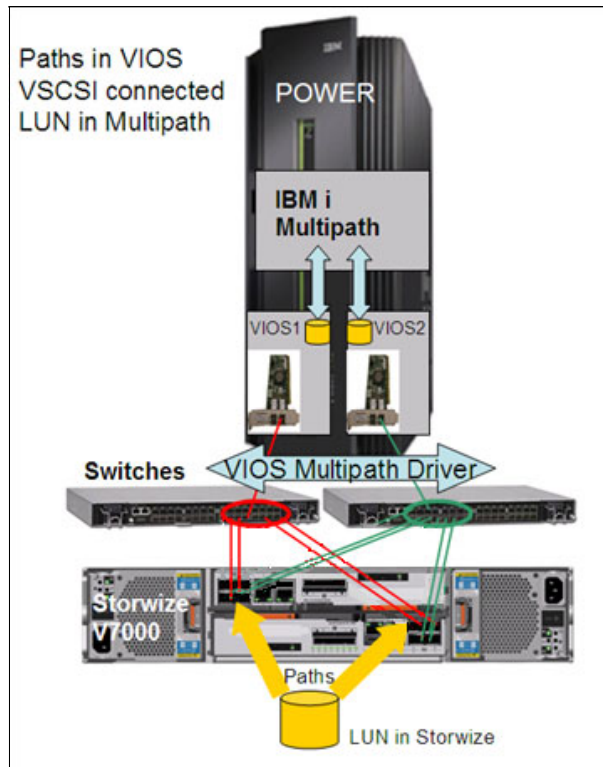


Figure 2-58 VIOS vSCSI multipath

IBM i uses two paths to the same LUN, each path through one VIOS to the relevant hdisk connected with a virtual SCSI adapter. Both paths are active and the IBM i load balancing algorithm is used for I/O traffic. Each VIOS has eight connections to the Storwize V7000; therefore, eight paths are established from each VIOS to the LUN. The I/O traffic through these paths is handled by the VIOS multipath driver.

Figure 2-59 shows both paths for the LUN in IBM i.

Display Disk Path Status						
ASP	Unit	Serial Number	Type	Model	Resource Name	Path Status
1	1	YTXBA98C5SSL	6B22	050	DMP274	Active
					DMP275	Active

Figure 2-59 Disk paths from IBM i client

2.5.4 Zoning SAN switches

With native connection and the connection with VIOS NPIV, it is considered a best practice to zone the switches so that one WWPN of one IBM i port is in zone with two ports of Storwize V7000, with each port from one node canister. This way, you ensure resiliency for the I/O to or from a LUN assigned to that WWPN. If the preferred node for that LUN fails, the I/O rate continues by using the non-preferred node.

Note: A port in a virtual FC adapter in IBM i has two WWPNs. For connecting external storage, use the first WWPN, the second WWPN is used for Live Partition Mobility (LPM). Therefore, it is a good idea to zone both WWPNs if you plan to use LPM, otherwise zone just the first WWPN.

When connection with VIOS virtual SCSI is used, it is recommended to zone one physical port in VIOS with all available ports in Storwize V7000, or with as many ports as possible to allow load balancing. There are a maximum of eight paths that are available from VIOS to the Storwize V7000. Storwize V7000 ports zoned with one VIOS port can be evenly spread between Storwize V7000 node canisters.

Examples of zoning for VIOS vSCSI connection

Figure 2-60 shows the use of one physical FC port in VIOS with a mapped IBM i virtual Fibre Channel client adapter, which was zoned with 2 * ports in Storwize V7000, 1 * ports from each node canister. In this way, the IBM i client LPAR sees one active and one passive path for each configured IBM i volume on the Storwize V7000.

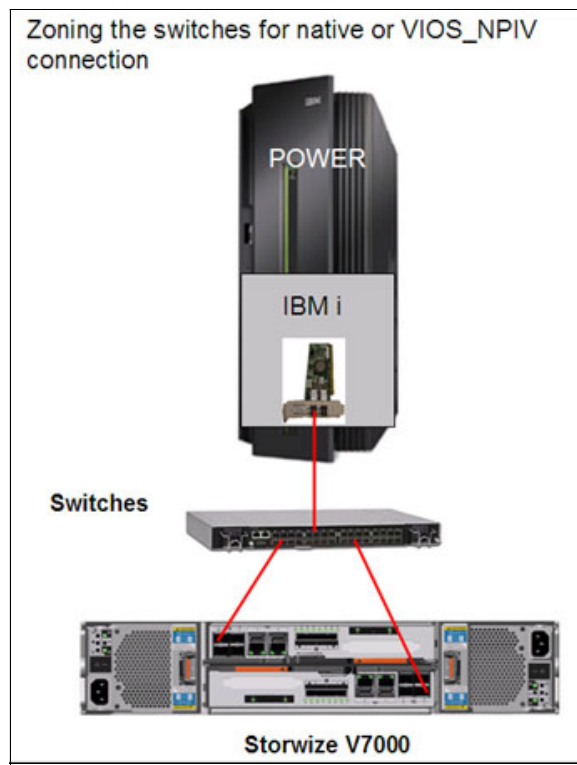


Figure 2-60 Zoning with one VIOS adapter

Figure 2-61 shows the use of two adapters in VIOS, zone one port from each adapter with 4 * V7000 ports, with two ports from each canister. In this way, the Storwize V7000 is balanced between VIOS ports and between the Storwize V7000 node canisters and do not exceed the eight paths from VIOS to Storwize V7000.

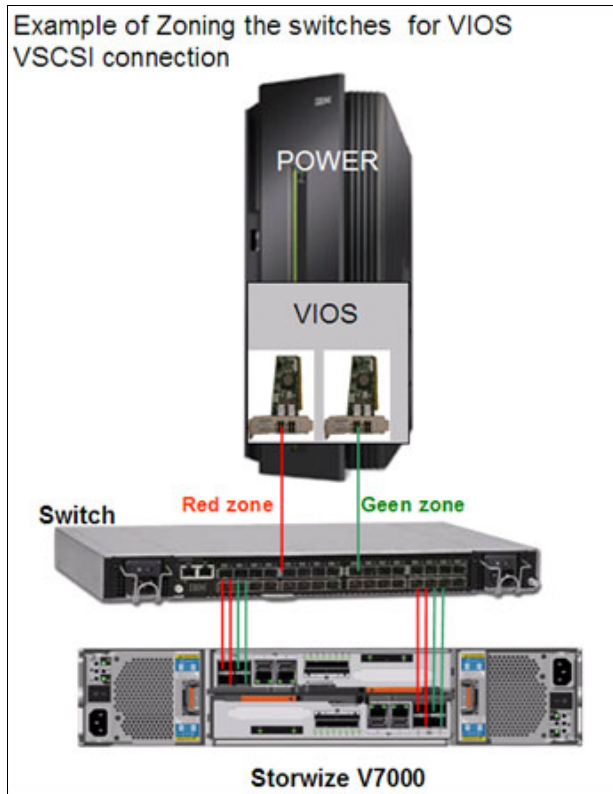


Figure 2-61 Zoning with two VIOS adapters

2.5.5 Data layout

Selecting an appropriate data layout strategy depends on your primary objectives.

Spreading workloads across all components maximizes the utilization of the hardware components. This includes spreading workloads across all the available resources. However, it is always possible when sharing resources that performance problems can arise because of contention on these resources.

To protect critical workloads, you can isolate them to minimize the chance that non-critical workloads can affect the performance of critical workloads.

A *storage pool* is a collection of managed disks from which volumes are created and presented to the IBM i system as LUNs. The primary property of a storage pool is the extent size, which is 1 GB by default with Storwize V7000 release 7.1 and 256 MB in earlier versions. This extent size is the smallest unit of allocation from the pool.

When you add managed disks to a pool, they can have the following performance characteristics:

- ▶ Same RAID level.
- ▶ Roughly the same number of drives per array.

- ▶ Same drive type (SAS, NL_SAS, and SSD unless Easy Tier® is used) because data from each volume is spread across all MDisks in the pool so the volume performs approximately at the speed of the slowest MDisk in the pool.
- ▶ The exception to this rule is that if Easy Tier is used, you can have two different tiers of storage in the same pool, but the MDisks within the tiers can still have the same performance characteristics.

Workload isolation is most easily accomplished where each ASP or LPAR has its own managed storage pool. This configuration ensures that you can place data where you intend with I/O activity balanced between the two nodes or controllers on the Storwize V7000.

Important: Make sure that you isolate critical workloads. A best practice is to have only IBM i LUNs on any storage pool (rather than mixed with non-IBM i). If you mix production and development workloads in storage pools, make sure that you understand this might affect production performance.

LUN compared to disk arm

The Storwize V7000 LUN that is connected to an IBM i reports in to the IBM i as a disk unit. IBM i storage management employs the management and performance functions as though the LUN was a disk arm. In fact, the LUN often is spread across multiple physical disk arms in the Storwize V7000 disk pool, where the LUN uses some capacity from each disk arm.

All of the disk arms in the disk pools are shared among all the LUNs that are defined in that disk pool. Figure 2-62 shows an example of a Storwize V7000 disk pool with three disk arrays of V7000 internal disk arms (MDisks) and a LUN that was created in the disk pool, with the LUN using an extent from each disk array in turn.

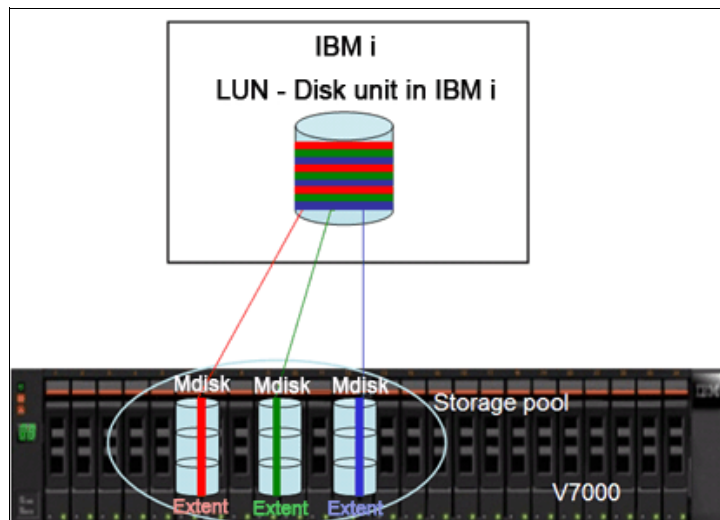


Figure 2-62 Storwize V7000 disk pool

LUN size

The maximum supported LUN size by IBM i is 2 TB - 1 Byte; that is, 2 TB LUNs are not supported by IBM i.

The number of LUNs that are defined often is related to the wait time component of the response time. If there are insufficient LUNs, wait time often increases. The sizing process determines the correct number of LUNs that are required to access the needed capacity, which meets performance objectives.

The number of LUNs drives the requirement for more FC adapters on the IBM i because of the addressing restrictions of IBM i if you are using native attachment. Each path to a LUN counts towards the maximum addressable LUNs on each IBM i IOA.

For any ASP, define all the LUNs to be the same size where 80 GB is the recommended minimum LUN size. A minimum of six LUNs for each ASP or LPAR is considered a best practice.

To support future product enhancements, it is considered a best practice that load source devices be created with at least 80 GB. A smaller number of larger LUNs reduces the number of I/O ports that are required on the IBM i and the Storwize V7000. In an IASP environment, you might use larger LUNs in the IASPs, but SYSBAS might require more, smaller LUNs to maintain performance.

The Disk Magic tool does not always accurately predict the effective capacity of the ranks depending on the DDM size that is selected and the number of spares that are assigned. The IBM Capacity Magic tool can be used to verify capacity and space utilization plans.

To calculate the number of LUNs that you require, you can use a tool, such as the IBM Systems Workload Estimator, which is found at the following website:

<http://www-912.ibm.com/wle/EstimatorServlet>

Adding LUNs to ASP

Adding a LUN to an ASP generates I/O activity on the rank as the LUN is formatted. If there is production work that is sharing the same rank, you might see a performance effect. For this reason, it is considered a best practice that you schedule adding LUNs to ASPs outside peak intervals.

2.5.6 Solid-State Drives

One of the most exiting innovations to happen to enterprise storage is solid-state drives (SSDs). Solid-state storage means using a memory-type device for mass storage rather than the use of a spinning disk or tape.

Solid-state storage technology has the following benefits:

- ▶ Significantly improved performance for hard-to-tune, I/O bound applications. No code changes are required.
- ▶ Reduced floor space. Can be filled near 100% without performance degradation.
- ▶ Greater IOPS.
- ▶ Faster access times.
- ▶ Reduced energy use.

The use of SSDs with the Storwize V7000 is through Easy Tier. Even if you do not plan to install SSDs, you can still use Easy Tier to evaluate your workload and provide information about the benefit you might gain by adding SSDs. Easy Tier is included with the Storwize V7000. However, it does require you to purchase a license and obtain a license key on a Storwize V3700.

When Easy Tier automated management is used, it is important to allow Easy Tier some space to move data. Do not allocate 100% of the pool capacity, but leave some capacity deallocated to allow for Easy Tier migrations.

At a minimum, leave one extent available per tier in each storage pool. However, for optimum use of future functions, plan to leave 10 extents available total per pool.

There also is an option to create a disk pool of SSDs in Storwize V7000 and to create an IBM i ASP that uses disk capacity from an SSD pool. The applications that are running in that ASP can experience a performance boost.

Note: IBM i data relocation methods (such as ASP balancing and media preference) are not available to use with SSD in Storwize V7000.

2.5.7 Booting from SAN

All connection options of native, VIOS NPIV, and VIOS virtual SCSI support boot from SAN. The LoadSource is on a Storwize V7000 LUN, which is connected the same way as the other LUNs. There are no special requirements for LoadSource connection.

When you are installing the IBM i operating system with disk storage on a Storwize V7000, you can select from one of the available Storwize V7000 LUNs when the installation prompts you to select the load source disk.

When you are migrating from internal disk drives or from another storage system to the Storwize V7000, you can use IBM i ASP balancing to migrate all of the disk capacity, except the LoadSource. After the non-LoadSource data is migrated to the Storwize V7000 with load balancing, you can then migrate the LoadSource by using the DST function to copy disk unit data to copy from the previous disk unit to the LUN in Storwize V7000. The Storwize V7000 LUN must be of equal or greater size than the disk unit that was previously used for the LoadSource. This type of migration can be done with the Storwize V7000 connections of native, VIOS NPIV, and VIOS vSCSI.

2.5.8 Sizing for performance

It is important to size a storage subsystem that is based on I/O activity rather than capacity requirements alone. This is true of an IBM i environment because of the sensitivity to I/O performance. IBM has excellent tools for modeling the expected performance of your workload and configuration. This section provides some guidelines for sizing for performance. However, these guidelines provide a starting point only for sizing with the appropriate tools.

LUN size is flexible. Choose the LUN size that gives you enough LUNs for good performance according to the Disk Magic tool. A best practice is to start modeling at 80 GB.

It is equally important to ensure that the sizing requirements for your SAN configuration also take into account the other resources that are required when Copy Services is enabled. Use the Disk Magic tool to model the overheads of replication (Global Mirror and Metro Mirror), particularly if you are planning to enable Metro Mirror.

It is also a good idea to perform a bandwidth sizing for Global Mirror and Metro Mirror.

Note: The Disk Magic tool does not support modeling FlashCopy® (which is also known as Point-in-Time Copy functions). Make sure that you do not size the system to maximum recommended utilizations if you want to also use FlashCopy snapshots for backups.

You must collect IBM i performance data to be analyzed. Generally, you collect performance data for a week's worth of performance for each system or LPAR and analyze the result reports for sizing.

Each set of reports includes print files for the following reports:

- ▶ System report: Disk utilization (required)
- ▶ Component report: Disk activity (required)
- ▶ Resource Interval report: Disk utilization detail (required)
- ▶ System report: Storage pool utilization (optional)

2.5.9 Performance monitoring

After your storage subsystem is installed, it is essential that you continue to monitor the performance of the subsystem. IBM i Performance Tools reports provide information about I/O rates and response times to the server. This allows you to track trends in increased workload and changes in response time. Review these trends to ensure that your storage subsystem continues to meet your performance and capacity requirements. Make sure that you are current on fixes to ensure that the Performance Tools reports are reporting your external storage correctly.

If you have multiple servers that are attached to a storage subsystem, particularly if you have other platforms that are attached in addition to IBM i, it is essential that you have a performance tool that enables you to monitor the performance from the storage subsystem perspective.

The IBM Tivoli® Storage Productivity Center license product provides a comprehensive tool for managing the performance of Storwize V7000. Collect data from all attached storage subsystems in 15-minute intervals. If there is a performance problem, IBM asks for this data.

There is a simple performance management reporting interface that is available through the IBM Storwize V7000 GUI. This provides a subset of the performance metrics that are available from the IBM Tivoli Storage Productivity Center tool.

For more information about the IBM Tivoli Storage Productivity Center tool, see this website:

<http://www-03.ibm.com/software/products/us/en/tivostorprodcent/>

2.5.10 Copy services considerations

The Storwize V7000 has two options for Global Mirror: the classic Global Mirror and the Change Volumes enhancement, which allows for a flexible and configurable RPO with which Global Mirror can be maintained during peak periods of bandwidth constraint.

Change Volumes are not currently supported by PowerHA, so it is essential that you size the bandwidth to accommodate the peaks or otherwise risk an effect to production performance. There is a limit of 256 Global Mirror with change volume relationships per system.

The current zoning guidelines for mirroring installations advise that a maximum of two ports on each Storwize V7000 node canister be used for mirroring, with the remaining two ports on the node not having any visibility to any other cluster. If you are experiencing performance issues when mirroring is operating, implementing zoning in this fashion might help to alleviate this situation.

Consulting services are available from the IBM STG Lab Services to assist in the planning and implementation of Storwize V7000 Copy Services in an IBM i environment. For more information, see this website:

http://www-03.ibm.com/systems/services/labservices/platforms/labservices_i.html

2.5.11 Translation from 520-byte blocks to 512-byte blocks

IBM i disks have a block size of 520 bytes. Most fixed block (FB) storage devices are formatted with a block size of 512 bytes, so a translation or mapping is required to attach these to IBM i.

Note: The IBM DS8000 Storage system supports IBM i with a native disk format of 520 bytes.

IBM i performs the following change of the data layout to support 512-byte blocks (sectors) in external storage: for every page (8 * 520-byte sectors) it uses another ninth sector; it stores the 8-byte headers of the 520-byte sectors in the ninth sector; therefore, it changes the previous 8* 520-byte blocks to 9* 512-byte blocks. The data that was previously stored in 8 * sectors is now spread across 9 * sectors, so the required disk capacity on Storwize V7000 is 9/8 of the IBM i usable capacity. Conversely, the usable capacity in IBM i is 8/9 of the allocated capacity in Storwize V7000.

Therefore, when an Storwize V7000 is attached to IBM i (whether through vSCSI, NPIV, or native attachment), this mapping of 520:512 byte blocks means that you have a capacity overhead and lose around 11% of the configured SAN storage capacity, which is not available for IBM i user data.

The effect of this translation to IBM i disk performance is negligible.



Setting up VIOS

This chapter describes how to set up an IBM PowerVM Virtual I/O Server (VIOS) by using the Hardware Management Console (HMC) graphical user interface (GUI), which was introduced in HMC Version 7.7.7.0 Service Pack 1. This chapter also explains how to set up VIOS by using a command-line interface (CLI).

The configuration of vSCSI and N-Port ID Virtualization (NPIV) host attachment are also explained by using both GUI and CLI interfaces, as the VIOS handles I/O requests to the storage.

Finally, the network setup between the client partitions and the external network is described. This explanation starts with a simple setup that uses one virtual network for the communication between the partitions and the external network and ends with multiple virtual networks to isolate traffic between the partitions, forwarding these VLAN tags to the external network, and balancing the VIOS partitions for network load.

The following topics are covered in this chapter:

- ▶ Before you start
- ▶ VIOS setup by using HMC GUI
- ▶ VIOS setup by using the CLI
- ▶ Mapping storage through VIOS to IBM i
- ▶ Network setup
- ▶ Customizing
- ▶ Troubleshooting

3.1 Before you start

The two DVDs that are called “PowerVM Virtual I/O Server Installation Media” are required, as physical media or as image files.

If your setup depends on multipath disk drive connections and vSCSI, the Subsystem Device Driver Path Control Module (SDDPCM) driver is required. The SDDPCM driver provides for specific SAN storage features, such as load balancing, that is not included in the native MPIO multipath driver.

The SDDPCM driver can be downloaded from this website:

<http://www-01.ibm.com/support/docview.wss?uid=ssg1S4000201>

The following steps are not explained in detail. For more information about how to create a logical partition and assign resources to it, see the IBM Redbooks publication *IBM Power Systems HMC Implementation and Usage Guide*, SG24-7491:

1. Create the logical partition from the HMC. Select the IBM Power server that hosts the VIOS from the HMC GUI on the left.
2. Click **Configuration** → **Create Partition** → **VIO Server** and assign the necessary resources for storage and network.
3. Assign a minimum set of physical adapters to the VIOS partitions for presenting the storage to the client partitions as needed. Add Ethernet adapters for VIOS management and Ethernet virtualization.

The following minimum partition requirements must be met:

- ▶ One physical Ethernet adapter for external network access
- ▶ One physical storage adapter (for VIOS disk and client virtualization)
- ▶ Two virtual server serial adapters for console connections
- ▶ One or two virtual server SCSI adapters (one for virtual optical device; one for vSCSI disks)
- ▶ One or two virtual server Fibre Channel adapters (one for NPIV disks; one for external NPIV Fibre Channel tape device)

Table 3-1 lists the VIOS design limits for IBM i.

Table 3-1 VIOS design limits for IBM i

Design Item	Maximum limit
Number of virtual target SCSI devices per virtual SCSI connection	16
Number of virtual devices per virtual NPIV port	64
Number of storage volumes per virtual Fibre Channel connection (NPIV)	64
Number of client LPARs (hosts) per physical Fibre Channel port (NPIV)	64
Number of NPIV virtual WWPN pairs per Power System	32000
Number of physical volumes (LUNs) per volume group (storage pool)	1024
Number of logical volumes per volume group (storage pool)	1024
Number of virtual adapters per LPAR	65535
Number of virtual Ethernet adapters per LPAR	32767
Number of 802.1Q VLAN tags per Shared Ethernet Adapter (SEA)	20

3.2 VIOS setup by using HMC GUI

Complete the following steps to set up your VIOS partition by using the HMC GUI:

Important: When it comes to network setup, make sure that communication with the HMC is possible by using the specified IP addresses. This is used for the setup of the partitions and is necessary to communicate with the partitions after the setup to add resources dynamically.

1. On the HMC GUI, activate the VIOS partition by selecting the partition and clicking **Activate** → **Profile**, as shown in Figure 3-1.

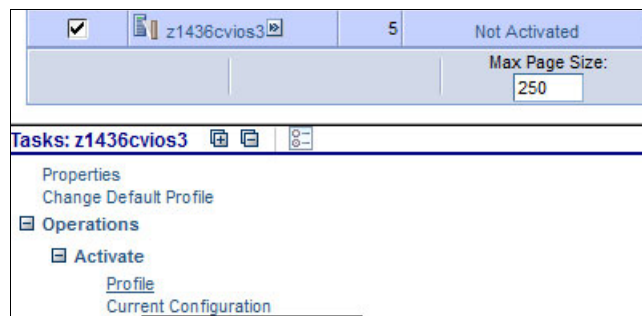


Figure 3-1 Activating the VIOS partition

2. In the Activate Logical Partition window (see Figure 3-2), select **Yes** to install the Virtual I/O Server as part of activation process. Click **OK** to start the installation wizard by using the HMC GUI.

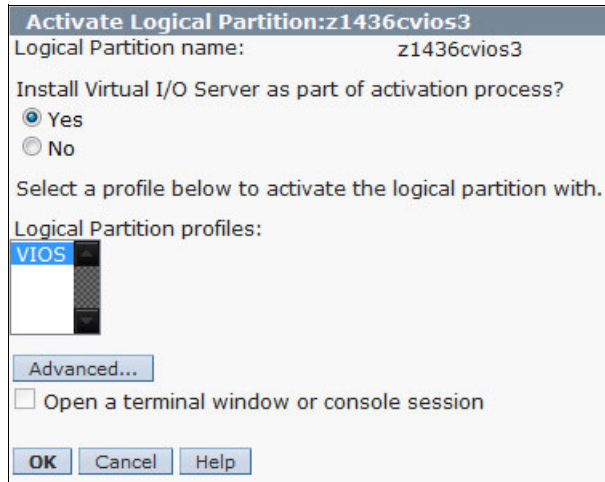


Figure 3-2 Starting the VIOS installation wizard

The setup wizard starts the partition and scans for network hardware.

Important: Make sure that no console is connected to this partition because this causes an error in hardware detection.

3. There are three possible ways to select a source for the installation media. A good approach for IBM i clients is to use the physical media, which must be inserted into the HMC optical drive. As an alternative, image files can be imported, which are on a remote FTP site; for example, on an existing IBM i partition.

If physical media is used for the installation, continue to step 8 on page 68.

- To import the image file into the local repository, select the Virtual I/O Server installation source **Local Repository** and click **Import**, as shown in Figure 3-3.

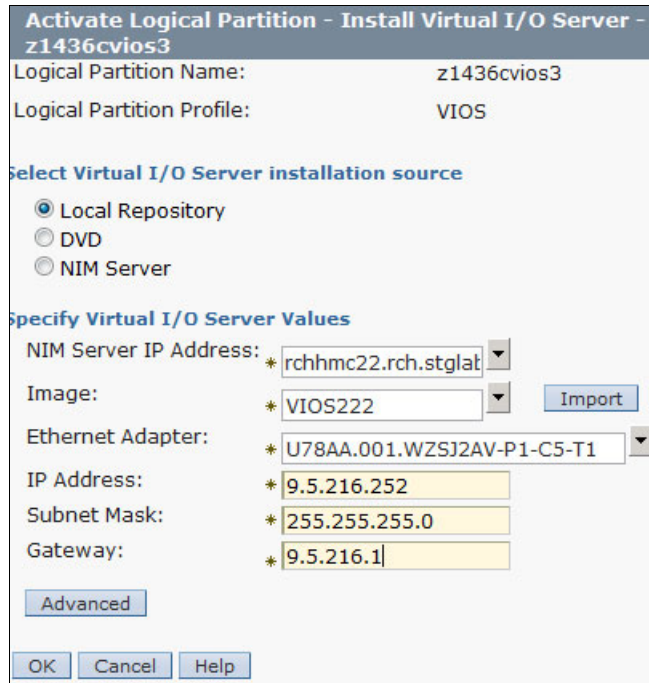


Figure 3-3 Installing VIOS from a local repository

- In the Import Virtual I/O Server Image window (see Figure 3-4), enter all of the necessary information and click **OK**.

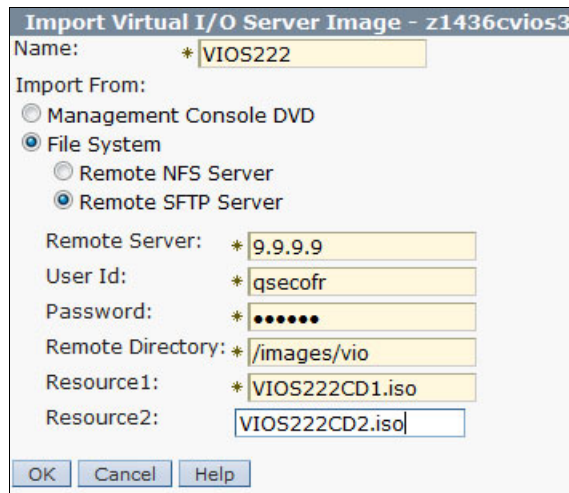


Figure 3-4 Importing the VIOS images

- Figure 3-5 shows the loading panel. Wait until the files are uploaded to the HMC and then continue with the installation.

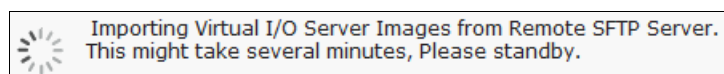


Figure 3-5 Importing process

Attention: Do not use blanks or special characters for the repository name.

Tip: It is possible to manage the Virtual I/O Server Image Repository from the left side of the HMC in the HMC Management menu.

7. Select the newly created repository to begin the installation process. Verify that the network information is correct and then continue with step 10 on page 68.
8. As shown in Figure 3-6, all of the necessary information to install the VIOS partition from physical media in the HMC optical drive was entered. Select the Virtual I/O Server installation source of **DVD** and specify the correct network adapter and all the network details for your VIOS partition. Click **OK**.

Activate Logical Partition - Install Virtual I/O Server - z1436cvios3

Logical Partition Name: z1436cvios3
Logical Partition Profile: VIOS

Select Virtual I/O Server installation source

Local Repository
 DVD
 NIM Server

Specify Virtual I/O Server Values

NIM Server IP Address: * rchhmc22.rch.stglal
Ethernet Adapter: * U78AA.001.WZSJ2AV-P1-C5-T1
IP Address: * 9.5.216.252
Subnet Mask: * 255.255.255.0
Gateway: * 9.5.216.1

Advanced
OK Cancel Help

Figure 3-6 Installing VIOS from DVD

9. After all of the information is copied from the first installation media, a window opens, in which you are prompted for the second media to be inserted into the HMC optical drive. Exchange the DVD and click **OK** to continue the installation.
10. You see the message that is shown in Figure 3-7 after the installation is complete.

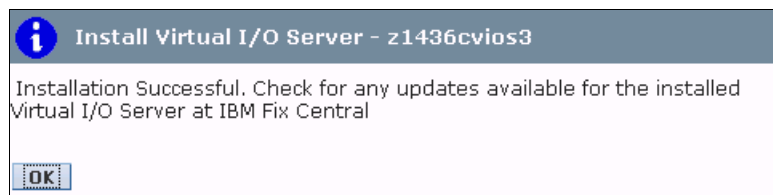


Figure 3-7 Installation success message

3.3 VIOS setup by using the CLI

Before you set up the VIOS by using the CLI, ensure that you created the VIOS partition profile as described in 3.1, “Before you start” on page 64.

Important: Make sure that the VIOS installation media is inserted into the IBM Power System optical drive and that the adapter that is connected to the drive is assigned to the VIOS partition.

Complete the following steps:

1. On the HMC GUI, activate the VIOS partition by selecting the partition and clicking **Activate** → **Profile**, as shown in Figure 3-8.

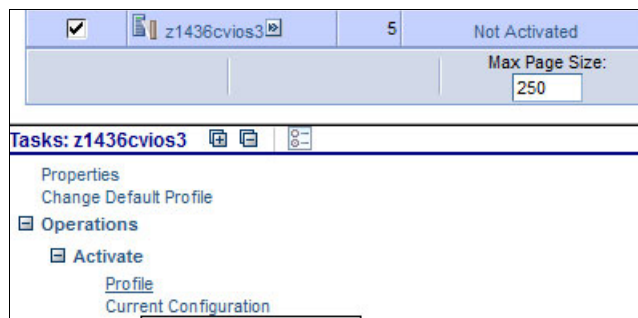


Figure 3-8 Activating the VIOS partition

2. In the Activate Logical Partition window (see Figure 3-9), click **Advanced**. Select **SMS** and click **OK** to change the boot mode. Activate the VIOS partition by clicking **OK**.

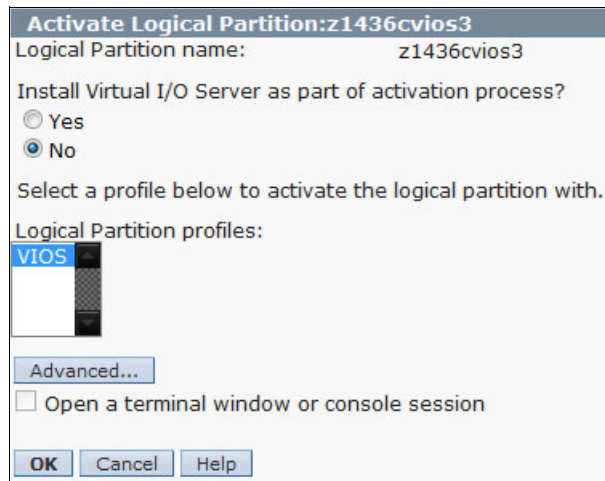


Figure 3-9 Starting the VIOS partition without starting the installation wizard

3. Complete the following steps to open a console connection to the VIOS:
 - a. Connect by using Secure Shell (SSH) to the HMC.
 - b. Log on with `hscroot` or a similar account.
 - c. Run the `vtmenu` command to open a server or partition selection.

Tip: The console window can be closed by entering ~.

- d. Browse to the VIOS partition and select it, as shown in Example 3-1.

Example 3-1 Opening a console window

```
hscroot@rchhmc22.rch.stglabs.ibm.com:~> vtmenu

Retrieving name of managed system(s) . . . z1436be:
-----
Partitions On Managed System: fz1436be:
OS/400 Partitions not listed
-----
1) "z1436cvios3:" Running

Enter Number of Running Partition (q to quit): 1

Opening Virtual Terminal On Partition z1436cvios3. . .

Open in progress..
```

4. Complete the following steps to select the boot device:
 - a. Select **Select Boot Options** and press Enter.
 - b. Select **Select Install/Boot Device** and press Enter.
 - c. Select **Select 1st Boot Device** and press Enter.
 - d. Select **CD/DVD** and press Enter.
 - e. Select the media type that corresponds to the optical device and press Enter.
 - f. Select the device number that corresponds to the optical device and press Enter.
 - g. Set the boot sequence to configure the first boot device. The optical device is now the first device in the Current Boot Sequence list.
 - h. Exit the SMS menu by pressing X and confirming that you want to exit SMS.
5. Complete the following steps to install the VIOS:
 - a. Select the console that you want and press Enter.
 - b. Select a language for the Base Operating System (BOS) menus and press Enter.
 - c. Select **Start Install Now with Default Settings** and press Enter.
 - d. Select **Change/Show Installation Settings and Install** to change the installation and system settings. This is necessary if more than one disk for VIOS is mapped to the system or for more advanced virtualization features are used, such as LPM.Figure 3-10 on page 71 shows the restrictions on each PowerVM edition.

Features	Express	Standard	Enterprise
Maximum VMs	3 / Server	1000 / Server	1000 / Server
Management	VMControl, IVM	VMControl, IVM*, HMC	VMControl, IVM*, HMC
Virtual I/O Server	✓	✓ (Dual)	✓ (Dual)
Suspend/Resume		✓	✓
NPIV	✓	✓	✓
Shared Processor Pools		✓	✓
Shared Storage Pools		✓	✓
Thin Provisioning		✓	✓
Active Memory Sharing			✓
Live Partition Mobility			✓

Figure 3-10 IBM PowerVM editions

- e. Select **Continue with Install** and press Enter. The system reboots after the installation is complete.
6. Finish the installation by configuring an IP address to the VIOS partition, as shown in Example 3-2.

Example 3-2 Configuring an IP address

```
$ lsdev en*
$ lsdev -dev ent0 -vpd
ent0          U78AA.001.WZSJG7K-P1-C7-T1 4-Port Gigabit Ethernet
PCI-Express Adapter (e414571614102004)
```

```
PCIe2 4-port 1GbE Adapter:
FRU Number.....00E2873
EC Level.....D77470
Customer Card ID Number....576F
Part Number.....00E2872
Feature Code/Marketing ID...5899
Serial Number.....YL502032N09E
Manufacture ID.....6CAE8B01EF14
Network Address.....6CAE8B01EF14
ROM Level.(alterable).....10050160
Hardware Location Code.....U78AA.001.WZSJG7K-P1-C7-T1
```

PLATFORM SPECIFIC

```
Name: ethernet
Node: ethernet@0
Device Type: network
Physical Location: U78AA.001.WZSJG7K-P1-C7-T1
```

```
$ cfgassist
```

VIOS TCP/IP Configuration

```
* Hostname [z1436cvios3]
* Internet ADDRESS (dotted decimal) [9.5.216.252]
```

Network MASK (dotted decimal)	[255.255.255.0]	
* Network INTERFACE	en0	
Default Gateway (dotted decimal)	[9.5.216.1]	
NAMESERVER		
Internet ADDRESS (dotted decimal)	[9.10.255.200]	
DOMAIN Name	[rch.stglabs.ibm.com]	
CableType	bnc	+

3.4 Mapping storage through VIOS to IBM i

The VIOS partitions are the base to provide storage, whether internal or external to the IBM i client partitions. If internal storage is used, only vSCSI connections are possible.

The limitations of vSCSI and NPIV are listed in Table 3-1 on page 65. If the focus of the setup is for performance, it is considered a best practice to use NPIV instead of vSCSI.

3.4.1 Setting up vSCSI connections

The management of vSCSI disks is built into the HMC GUI. It is also possible to configure everything from the VIOS command line. For multipath environments, the installation of the IBM SDDPCM provides load balancing between the available Fibre Channel paths.

Downloading and installing the SDDPCM multipath driver

Complete the following steps to download and install the SDDPCM multipath driver:

1. Check which level of AIX® is the base for the installed VIOS. This can be done by using the `oem_platform_level` command, as shown in Example 3-3.

The SDDPCM driver can be downloaded from this website:

<http://www-01.ibm.com/support/docview.wss?uid=ssg1S4000201>

Example 3-3 SDDPCM installation

```
$ oem_platform_level
AIX Version:
6.1.0.0
$ tar xf devices.sddpcm.61.rte.tar
$ cfgassist install_update
```

Install Software

```
* INPUT device / directory for software      .
* SOFTWARE to install                        [_all_latest]      +
PREVIEW only? (install operation will NOT occur)  no                +
COMMIT software updates?                     yes               +
SAVE replaced files?                          no                +
AUTOMATICALLY install requisite software?     yes               +
EXTEND file systems if space needed?          yes               +
OVERWRITE same or newer versions?            no                +
VERIFY install and check file sizes?         no                +
Include corresponding LANGUAGE filesets?     yes               +
DETAILED output?                              no                +
Process multiple volumes?                     yes               +
```

```

ACCEPT new license agreements?          no          +
[MORE...8]

$ shutdown -restart

```

- Transfer the downloaded files to the VIOS partition by using FTP. Run the commands that are shown in Example 3-3. They include a change of environment from VIOS commands to a broadened selection of AIX commands to finding the AIX version. Then, return to the VIOS environment, extract the downloaded archive, open the wizard to install the SDDPCM drivers, and restart the VIOS.

Attention: If IBM i partitions are running and no secondary VIOS partition is configured for failover, a restart of the VIOS partition does affect them all.

- Test if the SDDPCM driver is working by switching to the OEM setup environment by using the `oem_setup_env` and `pcmpath query` device commands.

Mapping disks from VIOS to IBM i partition by using HMC GUI

To map disks from VIOS to an IBM i partition, complete the following steps:

- From the HMC GUI, click **Systems Management** → **Servers** and select the Power System.
- From the Tasks menu, click **Configuration** → **Virtual Resources** → **Virtual Storage Management**. Select the VIOS partition that you want to work with from the drop-down menu and then click **Query**.
- The result is displayed as shown in Figure 3-11. In the first tab (Virtual Disks), virtual disks can be created and mapped to the client partitions. This creates files in the VIOS that serve as disk units, which are comparable to mounting storage by using a Network Server Description (NWSA).

In the third tab (Physical Volumes), physical volumes can be mapped from the VIOS to the client partition. With the use of external storage, such as the IBM Storwize Family, the disk sizing can be done in the SAN before they are mapped to the VIOS. Select the disk to be connected and click **Modify assignment**.

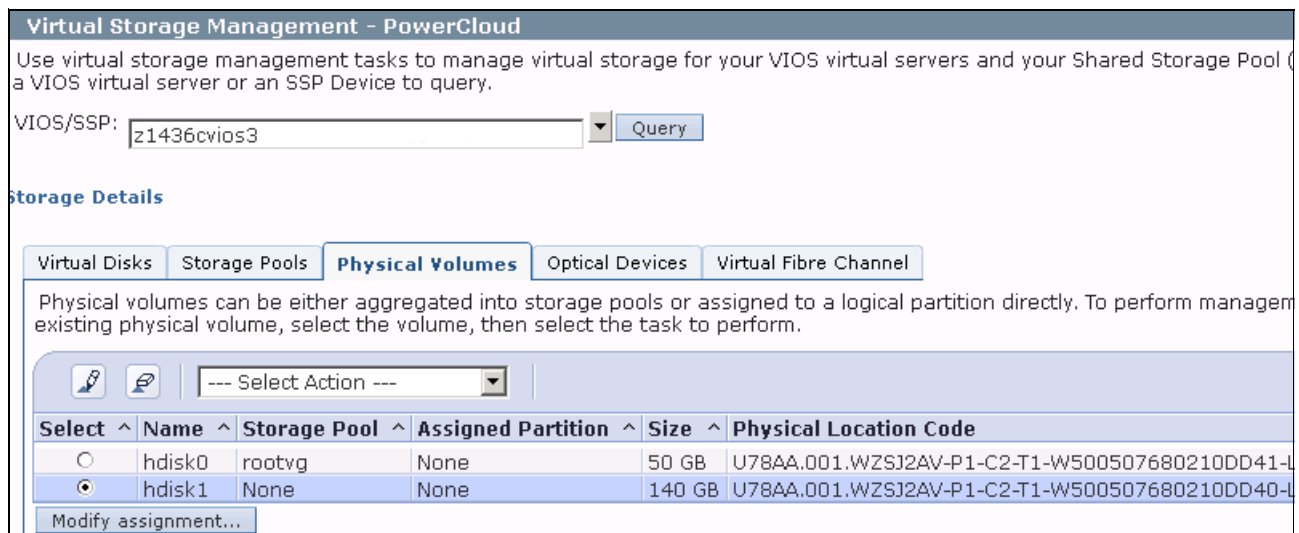


Figure 3-11 VIOS virtual storage management

- In the Modify Physical Disk Partition Assignment window, select the partition and adapter ID, as shown in Figure 3-12. Click **OK**. Disks can be added to partitions when they are active, but can be removed only from partitions that are inactive. The Assigned Partition column in Figure 3-11 changes to the partition name and slot ID that is selected.

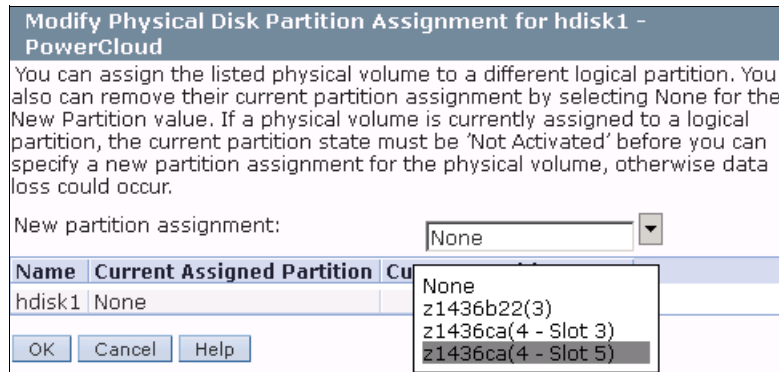


Figure 3-12 Selecting partition and slot ID for disk

The disk is now visible within IBM i and can be added to the ASP as normal.

Mapping disks from VIOS to IBM i partition using CLI

To map disks by using the VIOS CLI, complete the following steps:

- Connect to the VIOS partition by using an SSH utility. Log on as `padmin` user.
- Use the `lsmdev`, `lsmmap`, and `mkvdev` or `rmvdev` commands to display your disk devices, list the active mappings, and add or remove disks to or from partitions.

Example 3-4 shows how to display all available disks by using the `hdisk` command.

Tip: The location code V1-C6 that is shown in Example 3-5 references the VIOS partition ID 1 (V1) and virtual adapter ID 6 (C6).

Example 3-4 Displaying all available disks

```

$ lsmmap | grep hdisk
hdisk0          Available  MPIIO FC 2145
hdisk1          Available  MPIIO FC 2145

```


In Example 3-5 the active mappings are displayed and the hdisk1 device is mapped to the IBM i partition.

Example 3-5 Displaying disk mappings and adding a disk to the host

```
$ lsmmap -all
```

SVSA ID	Physloc	Client Partition
vhost2	U8205.E6C.06F759R-V1-C6	0x00000004

```
VTD NO VIRTUAL TARGET DEVICE FOUND
```

```
$ mkvdev -vdev hdisk1 -vadapter vhost2  
vtscsi0 Available
```

```
$ lsmmap -all
```

SVSA ID	Physloc	Client Partition
vhost2	U8205.E6C.06F759R-V1-C6	0x00000004

```
VTD vtscsi0  
Status Available  
LUN 0x8100000000000000  
Backing device hdisk1  
Physloc U78AA.001.WZSJ2AV-P1-C2-T1-W500507680210DD40-L1000000000000  
Mirrored false
```

- In the IBM i service tools, the LUN ID from VIOS is visible as a physical location for the logical resource, as shown in Figure 3-13.

You can select this display by clicking **STRSST** → **Start a service tool** → **Hardware service manager** → **Logical hardware resources** → **System bus resources** → **Resources associated with IOP** before the correct 290A adapter → **Display detail** at the disk unit.

Auxiliary Storage Hardware Resource Detail	
Description	Disk Unit
Type-model	6B22-050
Status	Operational
Serial number	YWAD6K386TL6
Part number	
Resource name	DPH001
Physical location	U8205.E6C.06F759R-V4-C3-T1-L8100000000000000

Figure 3-13 Location of logical hardware resource

3.4.2 Mapping storage by using NPIV

Figure 3-14 shows four IBM i partitions that are hosted by a dual VIOS configuration. Each IBM i partition uses two virtual Fibre Channel (FC) adapters to get redundant storage paths through each VIOS partition.

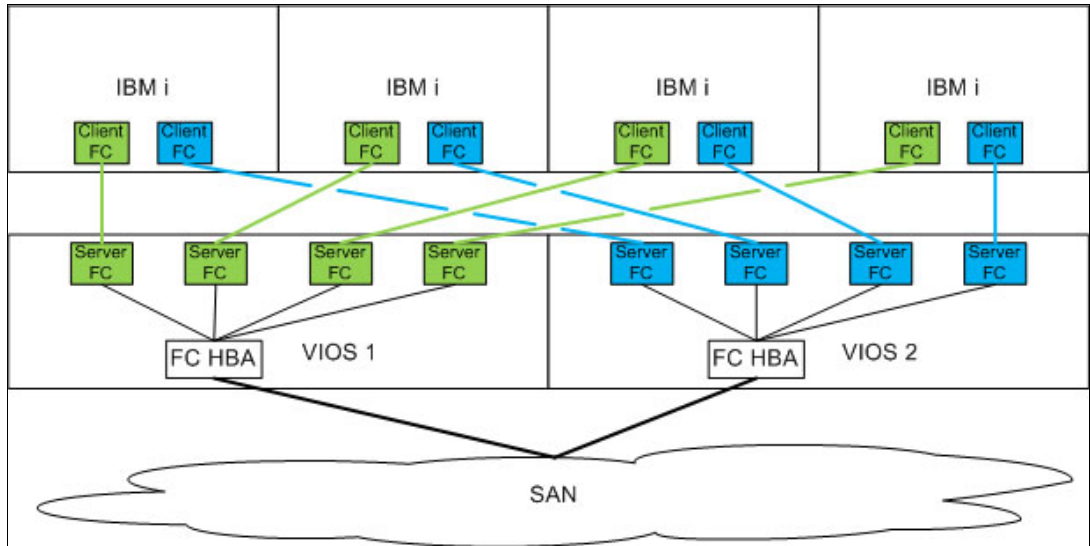


Figure 3-14 N-Port ID virtualization

For the virtual client adapters, the HMC generates two World Wide Port Numbers (WWPNs). The first WWPN is used for the SAN zoning. The second WWPN is important for Live Partition Mobility (LPM) and is used to create the partition profile at the remote system.

The server-client pairs can be mapped to the same physical HBA and are visible to the SAN and to the physical HBA WWPNs. Often, only the virtual WWPNs are used for further SAN zoning operations because they are connected to the IBM i client partition.

To map disks by using the HMC GUI, complete the following steps:

1. From the HMC GUI, click **Systems Management** → **Servers** and then select the Power Systems that you want to use.
2. From the Tasks menu, click **Configuration** → **Virtual Resources** → **Virtual Storage Management**. Select the VIOS partition that you want to work with from the drop-down menu and click **Query**.
3. The result is shown in Figure 3-15. In the last tab (Virtual Fibre Channel) of this window, virtual server Fibre Channel adapters can be mapped to physical ports. The location of the physical ports is shown in the Physical Location Code column. Distribute the client partitions across the physical adapters according to the I/O load.
Select the Fibre Channel port to work with and then click **Modify partition connections**.

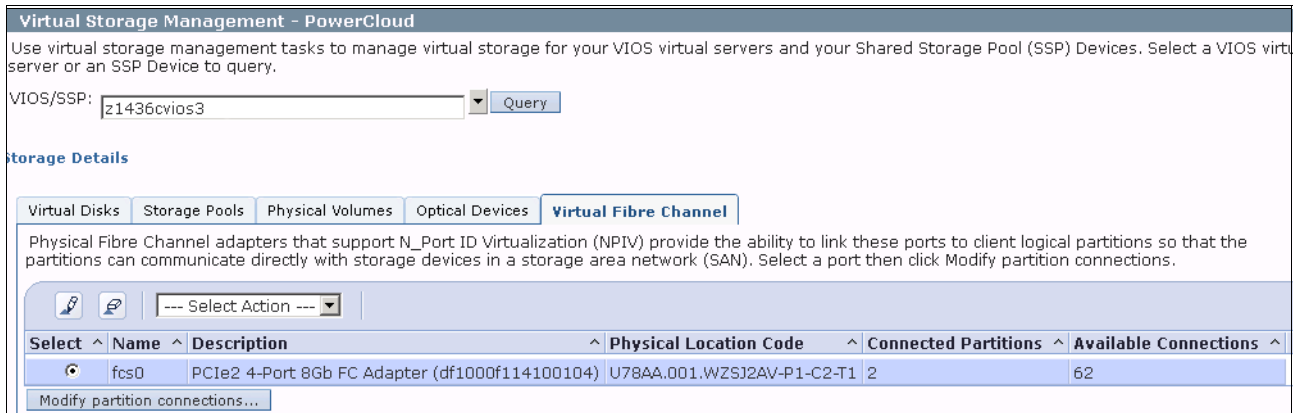


Figure 3-15 VIOS virtual Fibre Channel management

4. Figure 3-16 shows the active connections for the host bus adapter port fcs0. The connections can be modified by selecting and clearing the check boxes. Click **OK** to save the changes.

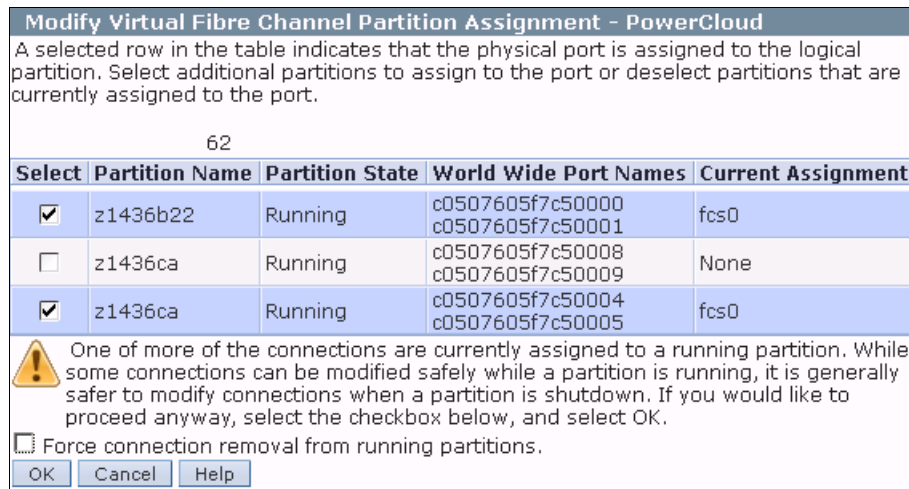


Figure 3-16 Modifying virtual Fibre Channel assignments

All disks that are zoned to the IBM i partition and connected to the IBM i virtual WWPN from the IBM Storwize Family SAN storage are now visible within IBM i and can be added to the ASP as normal. If disks are added or removed from the SAN, the changes are immediately visible to the IBM i partition without further configuration of the VIOS.

Mapping disks from VIOS to IBM i partition by using the CLI

To map disks by using the VIOS CLI, complete the following steps:

1. Connect to the VIOS partition by using an SSH utility and log on as the padmin user.
2. Use the `lsnports`, `lsmmap`, and `vfcmmap` commands to display your host bus adapter devices, list the active mappings, and add or remove virtual connections to or from the partitions.

Example 3-6 shows displaying all available adapters and the location codes. The output contains the following fields:

- Physical port name (name)
- Physical port location code (physloc)
- Fabric support (fabric)
- Total number of NPIV ports (tports)
- Available number of NPIV ports (aports)
- Supported number of WWPNs (swwpns)
- Available number of WWPNs (awwpns)

Example 3-6 Displaying host bus adapters and location codes

```
$ lsnports
name                physloc                fabric tports aports swwpns awwpns
fcs0                U78AA.001.WZSJ2AV-P1-C2-T1      1     64     62  2048   2040
```

In Example 3-7, the active mappings are displayed and the virtual server host bus adapter is mapped to the IBM i partition. The physical location code of the adapter is set to V1-C7 because it is a virtual adapter with the adapter ID 7 on the VIOS partition.

Example 3-7 Displaying disk mappings and adding a virtual server port to the host bus adapter

```
$lsmmap -all -npiv

Name                Physloc                CIntID CIntName          CIntOS
-----
vfchost2           U8205.E6C.06F759R-V1-C7      4

Status:NOT_LOGGED_IN
FC name:                FC loc code:
Ports logged in:0
Flags:1<NOT_MAPPED,NOT_CONNECTED>
VFC client name:        VFC client DRC:

$ vfcmmap -vadapter vfchost2 -fcp fcs0
$ lsmmap -all -npiv
Name                Physloc                CIntID CIntName          CIntOS
-----
vfchost2           U8205.E6C.06F759R-V1-C7      4 z1436ca          IBM i

Status:LOGGED_IN
FC name:fcs0                FC loc code:U78AA.001.WZSJ2AV-P1-C2-T1
Ports logged in:1
Flags:a<LOGGED_IN,STRIP_MERGE>
VFC client name:DC02        VFC client DRC:U8205.E6C.06F759R-V4-C7
```

3. In the IBM i service tools, the LUN ID from VIOS is visible as a physical location for the logical resource, as shown in Figure 3-17.

You can select this display by clicking **STRSST** → **Start a service tool** → **Hardware service manager** → **Logical hardware resources** → **System bus resources** → **Resources associated with IOP** before the correct 6B25 adapter → **Display detail** at the disk unit.

The disk type equals the type of the storage. The serial number of the disk contains the LUN ID, which is visible in the IBM Storwize Family web-based interface.

Auxiliary Storage Hardware Resource Detail	
Description	Disk Unit
Type-model	2145-050
Status	Operational
Serial number	YF033F000066
Part number	
Resource name	DMP274
Physical location	U8205.E6C.06F759R-V4-C12-T1- W500507680210DD40-L0

Figure 3-17 Location of logical hardware resource

3.5 Network setup

Network setup in VIOS can be a complex topic. It is considered a best practice to set up at least two VIOS instances for failover reasons. This way, VIOS updates can be done concurrently and the network infrastructure can be used redundantly.

The network setup in VIOS consists of the following steps:

- ▶ 3.5.1, “Optional step: LACP setup for adapter redundancy in each VIOS” on page 80
- ▶ 3.5.2, “Creating VLAN-capable trunk Ethernet adapters” on page 80
- ▶ 3.5.3, “Shared Ethernet Adapter setup” on page 83 for external network access and failure resiliency
- ▶ 3.5.4, “Transporting VLANs to the external network” on page 85 to Transport VLAN tags out of the IBM Power System

Figure 3-18 on page 80 displays the complete VIOS networking setup, including all of these steps.

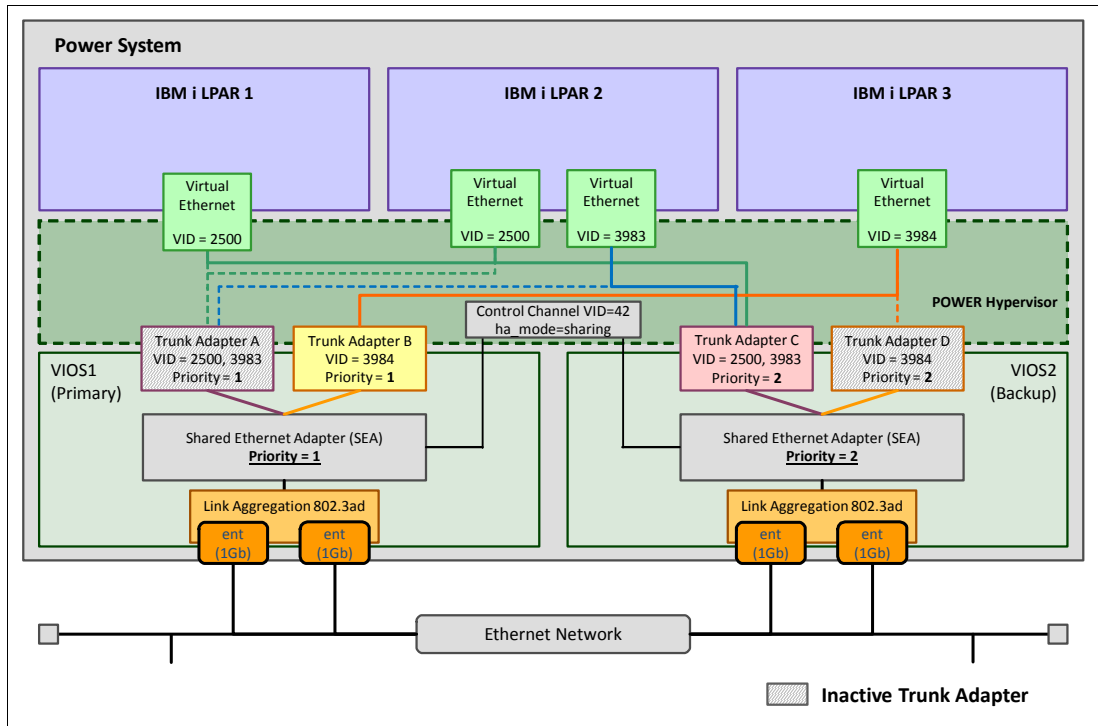


Figure 3-18 VIOS networking

3.5.1 Optional step: LACP setup for adapter redundancy in each VIOS

The link aggregation control protocol (LACP) allows for Ethernet port redundancy. This depends on the correct switch configuration. It is also a good idea to coordinate this configuration with the network department.

A wizard guides you through the setup of LACP. Example 3-8 shows the LACP configuration for ent0 and ent1.

Example 3-8 LACP configuration for ent0 and ent1

```
$ mkvdev -lnagg ent0,ent1
ent14 Available
en14
et14
```

3.5.2 Creating VLAN-capable trunk Ethernet adapters

As shown in Figure 3-18, a virtual trunk Ethernet adapter is required on the VIOS partitions to become part of the SEA configuration that was described in 3.5.3, “Shared Ethernet Adapter setup” on page 83.

Complete the following steps to set up the Ethernet adapter by using the HMC GUI:

1. From the HMC GUI, select **Systems Management** → **Servers** → IBM Power System.
2. Select the primary VIOS partition.
3. In the Tasks area, click **Configuration** → **Manage Profiles** → profile name.

- Click the **Virtual Adapters** tab to create the virtual adapters for this partition, as shown in Figure 3-19. Click **Actions** → **Create Virtual Adapter** → **Ethernet Adapter**.

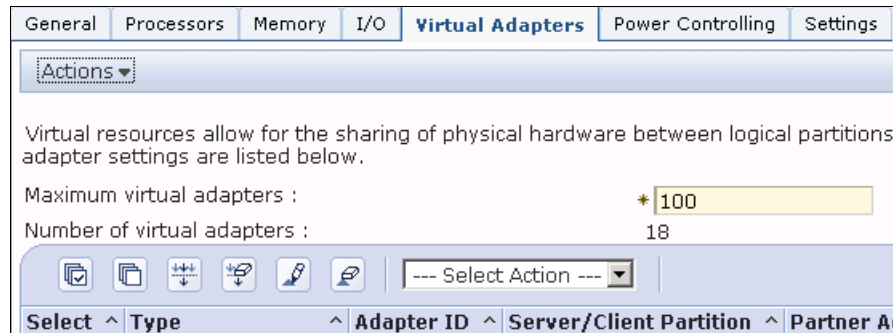


Figure 3-19 VIOS virtual adapter configuration

- In this step, a virtual trunk Ethernet adapter is created that can bridge traffic to the external network and allow for multiple VLANs, as shown in Figure 3-20 on page 82.

If you do not plan on having multiple virtual networks, do not select the IEEE 802.1q compatible adapter option, which adds the VLAN tagging options.

Tip: It is considered a best practice to use only one default VLAN for client traffic and another internal VLAN for the shared Ethernet adapter high availability.

Tip: VLAN IDs can range 0 - 4095. The IDs 0 and 4095 are reserved. VLAN ID 1 is the default VLAN ID for most hardware providers. Do not use these IDs for VLAN setups to avoid unexpected behavior.

It is important that the VLAN ID that is added to the packets that are leaving the client partitions (the default VLAN ID number) is not removed on entering the Virtual I/O Server. This is what happens when a default configuration is used.

It is for this reason that the default VLAN ID on the Virtual I/O Servers virtual Ethernet adapter (as shown in Figure 3-20 on page 82) must be set to an unused VLAN ID (in this example, 42 is used) that is not used outside the system.

If a packet arrives with this VLAN ID (42), the VLAN ID tag is stripped off (untagged) and cannot be forwarded to the correct external VLAN. If this was sent through the shared Ethernet adapter to the external physical network, it arrives at the Ethernet switch as untagged.

Click **OK** to save the settings and close the window.

General | Advanced

Virtual ethernet adapter
 Adapter ID : * 2
 VSwitch : [Dropdown] [Get VSwitch ID List...]
 Port Virtual Ethernet (VLAN ID): 42 [Dropdown] [View Virtual Network...]

IEEE Settings
 Select this option to allow additional virtual LAN IDs for the adapter.
 IEEE 802.1q compatible adapter
 Maximum number of VLANs: 20
 Add VLAN ID: [Input] [Add]
 Additional VLAN IDs: [List: 2500, 3983] [Remove]

Shared Ethernet Settings
 Select Ethernet bridging to link (bridge) the virtual Ethernet to a physical network
 Use this adapter for Ethernet bridging
 Priority: 1 [Spin] (1 or 2)

OK Cancel Help

Figure 3-20 Creating a virtual Ethernet trunk adapter for external network access

6. Repeat steps 3 on page 80 through step 5 on page 81 for the secondary VIOS. Use a priority value of two for the virtual trunk Ethernet adapter.
7. Add another virtual Ethernet adapter (as shown in Figure 3-21) for the control channel of the SEA with as high a VLAN ID as possible. Make sure to coordinate with the network department. When the SEA is created, it checks for a virtual Ethernet adapter, with no access to the external network, and starts at the highest VLAN ID.

Tip: With an active RMC connection between VIOS and HMC, all changes to an active partition can be done by using DLPAR operations. Take care to save all changes to the LPAR profile afterward.

General | Advanced

Virtual ethernet adapter
 Adapter ID : * 3
 VSwitch : [Dropdown] [Get VSwitch ID List...]
 Port Virtual Ethernet (VLAN ID): 4094 [Dropdown] [View Virtual Network...]

This adapter is required for virtual server activation.

IEEE Settings
 Select this option to allow additional virtual LAN IDs for the adapter.
 IEEE 802.1q compatible adapter

Shared Ethernet Settings
 Select Ethernet bridging to link (bridge) the virtual Ethernet to a physical network
 Use this adapter for Ethernet bridging

OK Cancel Help

Figure 3-21 Creating a control channel virtual Ethernet adapter

3.5.3 Shared Ethernet Adapter setup

The Shared Ethernet Adapter (SEA) bridges the traffic between the virtual trunk Ethernet adapter and the physical adapter interface. It is also used for high availability reasons. Complete the following steps to create an SEA adapter:

1. From the HMC GUI, click **Systems Management** → **Servers** → IBM Power System.
2. From the Tasks area, click **Configuration** → **Virtual Network Management**.
3. As shown in Figure 3-22, select the default VSwitch and select the VLAN ID of 42, as this was the defined virtual trunk Ethernet adapter that was created in the example in step 5 on page 81. Click **Create SEA**.

Tip: You can select each VLAN ID to get a list of the associated adapter resource name in every LPAR.

VSwitch

This panel allows you to manage Virtual Switch (VSwitch) and Virtual Local Area Network (VLAN) configuration. VSwitches can be created from this panel. VLANs will be created on the activation of partitions using the VLANs.

Select	VSwitch	Switch Mode
<input checked="" type="radio"/>	ETHERNET0(Default)	VEB

Virtual LANs

Use Virtual VLANs to view the VLANs defined for the managed system. You may also view VLANs by their partition participation by changing the "View by" selection to Partitions.

View by:

Select a virtual local area network (VLAN) to manage. You then can view configuration details for the VLAN and select management tasks for the VLAN.

Select	VLAN ID	Bridge
<input type="radio"/>	23	
<input checked="" type="radio"/>	42	
<input type="radio"/>	2500	
<input type="radio"/>	3983	
<input type="radio"/>	3984	
<input type="radio"/>	4093	
<input type="radio"/>	4094	

Details

Partitions		Shared Ethernet Adapters		
Partition	Virtual Adapter	Shared Adapter	Priority	VIOS
P740_VIOS1	ent2(Slot 2)			
P740_VIOS2	ent2(Slot 2)			

Figure 3-22 HMC virtual network editor

4. Select your primary and secondary VIOS partitions and the physical network interface, as shown in Figure 3-23.

Create Shared Ethernet Adapter - PowerCloud

This panel allows you to create a Shared Ethernet Adapter (SEA) for this Virtual Local Area Network (VLAN) using Virtual I/O Servers. You can create a single SEA and an optional Failover SEA for this VLAN. The failover can be created at a later time if you wish to only create a single SEA; However no other configuration are allowed.

VLAN ID: 42

VIOS: P740_VIOS1(1)

Physical adapter: ent0 - 2-Port 10/100/1000 Base-TX PCI-Express Adapter (14104003) (U78AA.001.WZSJ2AV-P1-C4)

Configure virtual network failover

Failover VIOS: P740_VIOS2(2)

Failover Physical adapter: ent0 - 2-Port 10/100/1000 Base-TX PCI-Express Adapter (14104003) (U78AA.001.WZSJ2AV-P1-C4)

OK Cancel Help

Figure 3-23 SEA creation

5. The wizard returns to the Virtual Network Management window. If you select the VLAN ID of 4094, you can see that this VLAN acts now as the control channel VLAN for the newly created SEA, as shown in Figure 3-24.

Select	VSwitch	Switch Mode
<input checked="" type="radio"/>	ETHERNET0(Default)	VEB

Virtual LANs

Use Virtual VLANs to view the VLANs defined for the managed system. You may also view VLANs by their partition participation by changing the "View by" selection to Partitions.

View by: VLANs

Select a virtual local area network (VLAN) to manage. You then can view configuration details for the VLAN and select management tasks for the VLAN.

Select	VLAN ID	Bridge
<input type="radio"/>	23	
<input type="radio"/>	42	P740_VIOS1(ent0), P740_VIOS2(ent0)
<input type="radio"/>	2500	P740_VIOS1(ent0), P740_VIOS2(ent0)
<input type="radio"/>	3983	P740_VIOS1(ent0), P740_VIOS2(ent0)
<input type="radio"/>	3984	
<input type="radio"/>	4093	
<input checked="" type="radio"/>	4094	

Details

Partitions		Shared Ethernet Adapters		
Partition	Virtual Adapter	Shared Adapter	Priority	VIOS
P740_VIOS1	ent3(Slot 3) - VLAN 42 Control Channel			
P740_VIOS2	ent3(Slot 3) - VLAN 42 Control Channel			

Figure 3-24 HMC virtual network editor with control channel VLAN

The SEA configuration is now finished and client partitions with virtual Ethernet adapters on VLAN 42 can access the external network. VLAN tags are not yet exported to the external network. If you decide to use other VLAN tags in VLAN 2500 and 3983, continue with 3.5.4, "Transporting VLANs to the external network" on page 85.

6. This optional step adds a virtual trunk Ethernet adapter to the SEA and reconfigures it to use load sharing, as shown in Figure 3-18 on page 80. For more information, see Example 3-9.

Important: The two steps that are shown in Example 3-9 must be run in each VIOS partition, starting with the primary site.

Example 3-9 Adding a trunk adapter to the SEA and change ha_mode attribute

```
ent6: SEA
ent2: trunk adapter which is already part of the SEA
ent5: new trunk adapter adding to the SEA

$ chdev -dev ent6 -attr virt_adapters=ent2,ent5
ent6 changed
Changing the SEA to Load Sharing mode.
$ chdev -dev ent6 -attr ha_mode=sharing
ent6 changed
```

3.5.4 Transporting VLANs to the external network

Complete the following steps to transport the VLAN IDs (that can be used only within the vSwitch until now) to the external network. The network hardware that is attached must be configured for VLAN tagging:

1. Connect to the VIOS partitions by using SSH or open a console connection.
2. Use the `mkvdev` command to add VLAN devices to the shared Ethernet adapter, as shown in Example 3-10.

Example 3-10 Exporting the VLAN IDs using the SEA adapter ent6

```
$ mkvdev -vlan ent6 -tagid 2500
ent7 Available
en7
et7
$ mkvdev -vlan ent6 -tagid 3983
ent8 Available
en8
et8
$ mkvdev -vlan ent6 -tagid 3984
ent9 Available
en9
et9
```

3.6 Customizing

This section describes how to ease navigation within the VIOS shell. This section also shows how to set up the SNMP daemon to monitor VIOS performance remotely. The following topics are covered:

- ▶ 3.6.1, “Changing ksh to bash style” on page 86
- ▶ 3.6.2, “Configuring monitoring by using SNMP” on page 86
- ▶ 3.6.3, “Creating an image repository for client D-IPL” on page 87

3.6.1 Changing ksh to bash style

To permanently change the VIOS shell to a more intuitive way of operation, follow the commands that are shown in Example 3-11. This example uses the editor `vi` that is preinstalled with VIOS.

Use the `o` key to insert a new line and change to editing mode. Use the `Esc` key to exit editing mode.

Outside the editing mode, you can use commands to browse `vi`. Enter `:d` to delete a line and `:x` to exit and save the file.

Example 3-11 Changing ksh to bash style

```
$ oem_setup_env
# chmod 644 /usr/ios/cli/.profile
# vi /usr/ios/cli/.profile
```

add the following lines to the file:

```
export ENV=/usr/ios/cli/.kshrc
export PS1="$(ioscli hostname)$ "
alias __A=`echo "\020"~` #ctrl-p => up
alias __B=`echo "\016"~` #ctrl-n => down
alias __C=`echo "\006"~` #ctrl-f => right
alias __D=`echo "\002"~` #ctrl-b => left
alias __H=`echo "\001"~` # Home
alias __F=`echo "\005"~` # End
set -o emacs
```

exit the editor via **ESC** and `:x`

```
# exit
```

3.6.2 Configuring monitoring by using SNMP

Complete the following steps to enable the Simple Network Management Protocol (SNMP) daemon:

1. Edit the file `snmpdv3.conf` with the `vi` editor, as described in 3.6.1, “Changing ksh to bash style” on page 86.
2. Edit the file `rc.tcpip` to enable the SNMP daemon auto start.
3. Restart the SNMP daemons.

For more information, see Example 3-12.

Example 3-12 Configuring SNMP daemons

```
$ oem_setup_env
# vi /etc/snmpdv3.conf
```

Put a # before the following line to comment it out:

```
VACM_VIEW defaultView          1.3.6.1.4.1.2.6.191          - excluded -
```

Add the next four lines to the end of the file, save, and exit:

```
VACM_GROUP group1 SNMPV2c public -
VACM_ACCESS group1 - - noAuthNoPriv SNMPV2c defaultView - defaultView -
COMMUNITY public public noAuthNoPriv 0.0.0.0 0.0.0.0 -
VACM_VIEW defaultView          1                      - included -
```

```
# vi /etc/rc.tcpip
```

Uncomment the following lines, save, and exit:

```
start /usr/sbin/snmpd "$src_running"
start /usr/sbin/hostmibd "$src_running"
start /usr/sbin/snmpmibd "$src_running"
start /usr/sbin/aixmibd "$src_running"
```

Restart the services:

```
# for s in snmpd snmpmibd hostmibd aixmibd; do stopsrc -s $s; sleep 2; done
# for s in snmpd snmpmibd hostmibd aixmibd; do startsrc -s $s; sleep 2; done
# exit
```

3.6.3 Creating an image repository for client D-IPL

If no physical DVD drive is available to the client partition, create an image repository that is on the VIOS partitions disk. This can be done by using the HMC GUI or from the VIOS CLI.

Attention: Unlike image catalogs in IBM i, the images in the repository are not autoloaded.

Before you start, a virtual DVD drive must be created on the VIOS partition by using the VIOS command line. Complete the following steps:

1. Create a virtual server SCSI adapter on the VIOS partition profile and a matching virtual client SCSI adapter on the client partition that is used as alternative Load Source device.
2. Run the `mkvdev` command from the VIOS CLI for the created device (as shown in Example 3-13 on page 88) to create a virtual optical device.

Tip: Use the `-dev` parameter with the `mkvdev` command to set a device name, such as `vtopt_iLPAR1` for later reference.

Example 3-13 Creating a virtual optical device

```
$ lsmmap -all
SVSA          Physloc          Client Partition
ID
-----
vhost0        U8205.E6C.06F759R-V1-C4  0x00000003

VTD          NO VIRTUAL TARGET DEVICE FOUND

$ mkvdev -fbo -vadapter vhost0
vtopt0 Available
$ lsmmap -all
SVSA          Physloc          Client Partition
ID
-----
vhost0        U8205.E6C.06F759R-V1-C4  0x00000003

VTD          vtopt0
Status       Available
LUN          0x8100000000000000
Backing device
Physloc
Mirrored     N/A
```

Creating a repository and adding images by using the HMC GUI

Complete the following steps to create the repository, add the I_BASE_01 image, and assign it to the IBM i client partition:

1. Select your IBM Power Server and open the Virtual Storage Management from the Tasks area by clicking **Configuration** → **Virtual Resources** → **Virtual Storage Management**.
2. Select the VIOS partition and click **Query**.
3. As shown in Figure 3-25 on page 89, click **Create Library** in the Optical Devices tab to add a new library.

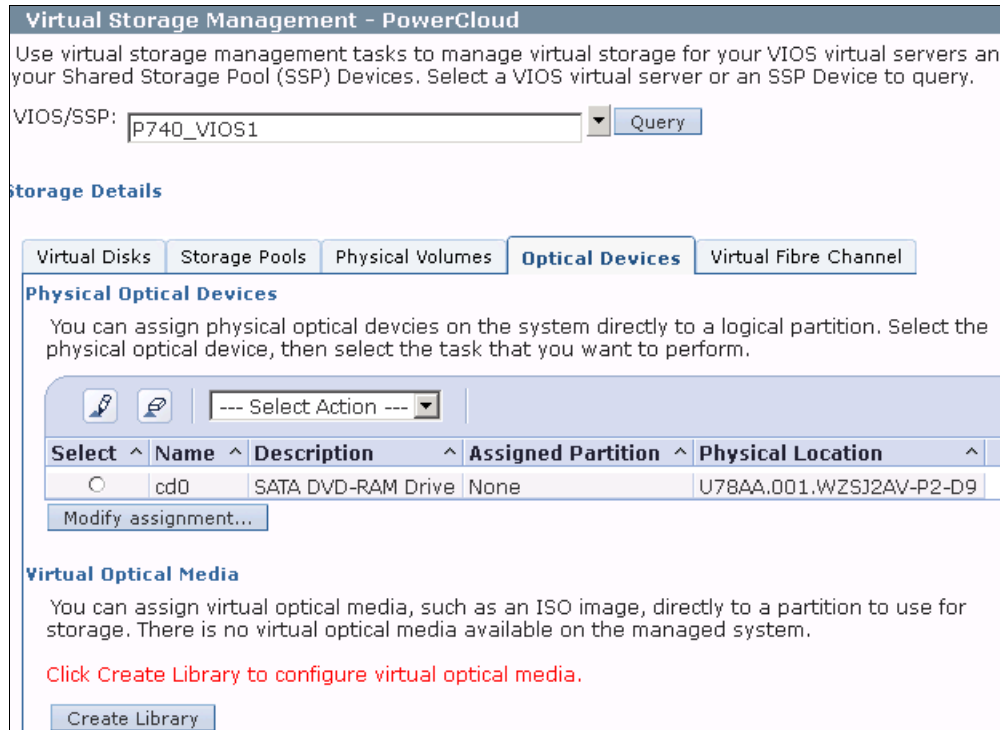


Figure 3-25 Managing optical devices

- Figure 3-26 shows the Create Media Library window. Enter the media library size. The maximum size depends on the VIOS disk that is used. Ensure that it is large enough to store the image files. Click **OK**.

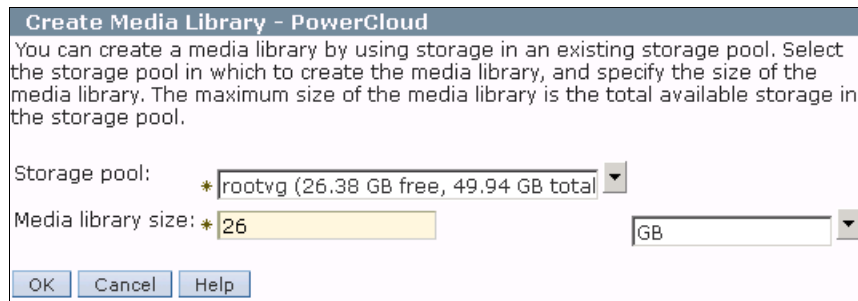


Figure 3-26 Creating the media library

- Transfer the image file by using FTP or an SCP client to the VIOS server.

6. Add the transferred image file to the repository, as shown in Figure 3-27. Enter a media name by which the image can be identified and the full path to the image file. Click **OK** to create a Virtual Media object.

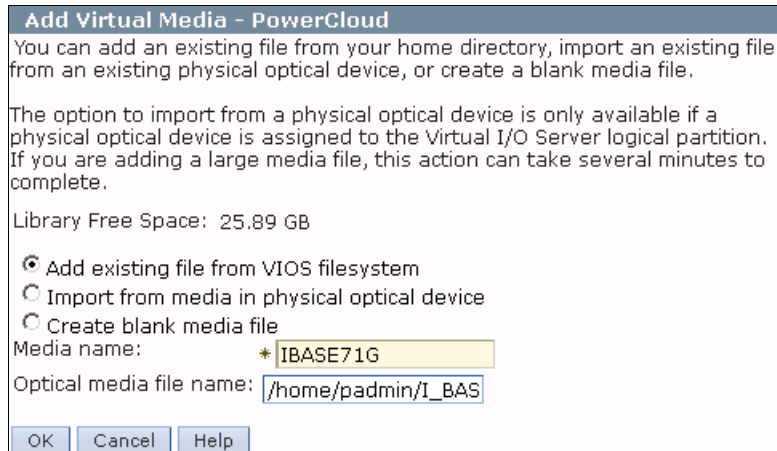


Figure 3-27 Adding virtual media

Figure 3-28 shows the upload progress.

Tip: You can delete the image file from the VIOS after it is copied from its location to the repository.

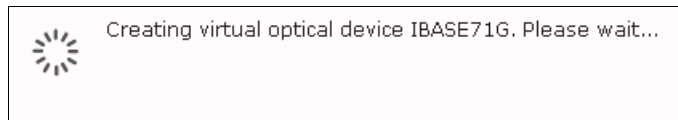


Figure 3-28 Creating the virtual optical media

7. Select the image file and assign it to the virtual optical device, as shown in Figure 3-29.

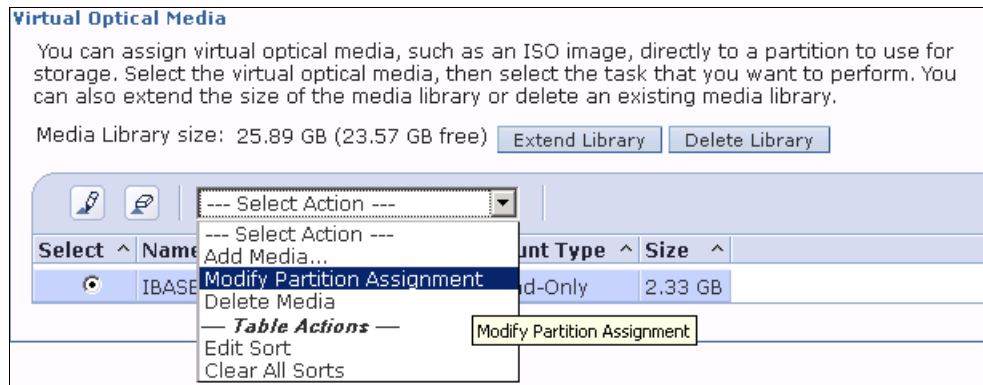


Figure 3-29 Assigning the virtual optical media

8. Select the partition and virtual optical drive. Select the **Force reassignment on running partitions** option and click **OK**, as shown in Figure 3-30. Use read-only as the access mode to load the image into multiple virtual drives.

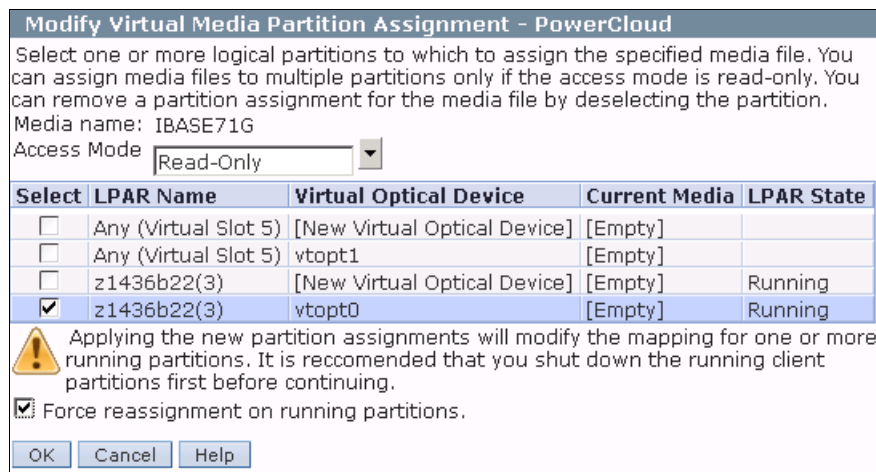


Figure 3-30 Assigning the virtual media

Creating a repository and adding images by using the VIOS CLI

Example 3-14 shows how to create an image repository, add an image to the repository, and load the image for the client partition by using the VIOS CLI. Use read-only as the access mode to load the image into multiple virtual drives.

Example 3-14 Creating a repository, adding an image, and loading the image

```
$ lssp
Pool                Size(mb)  Free(mb)  Alloc Size(mb)  BDs Type
rootvg              51136    27008    64              0 LVPPOOL
$ mkrep -sp rootvg -size 26G
Virtual Media Repository Created
Repository created within "VMLibrary" logical volume
$ mkvopt -name IBASE71G -file I_BASE_01_710G.iso -ro
$ loadopt -disk IBASE71G -vtd vtopt0
$ lsrep
Size(mb) Free(mb) Parent Pool          Parent Size      Parent Free
 26515   24134 rootvg                51136            384

Name                                File Size Optical
Access
IBASE71G                             2381 vtopt0      ro
$ unloadopt -vtd vtopt0
```

3.7 Troubleshooting

This section describes how to improve VIOS performance and find bottlenecks. This section also shows how to restart the Resource Monitoring and Control (RMC) connection and which information is necessary for a support call and how to gather that information.

The following topics are covered:

- ▶ 3.7.1, “Measuring performance” on page 92
- ▶ 3.7.2, “Restarting the RMC connection” on page 95
- ▶ 3.7.3, “Collecting troubleshooting data for support” on page 96

3.7.1 Measuring performance

Starting with VIOS 2.2.2, a tool to measure the performance and check the VIOS configuration is available. Complete the following steps:

1. Run the **part** command (as shown in Example 3-15) to collect the performance information.

Example 3-15 Gathering performance statistics

```
$ part
usage: part {-i INTERVAL | -f FILENAME} [-t LEVEL] [--help|-?]
      -i <minutes>      interval can range between 10-60
      -f <file>         any nmon recording
      -t <level>        1 - Basic logging, 2 - Detailed logging
      -?                usage message
$ part -i 30
part: Reports are successfully generated in z1436cvios3_131011_20_26_48.tar
```

2. Download the archive from the VIOS and extract it at your workstation.
3. Open the `vios_advisor_report.xml` file with a web browser.

The resulting output is split into different sections. In Figure 3-31, the overview of the system is displayed. Click the icons to see an explanation for each item.










Monitoring Start Time : 2013-10-11T20:23:18		
Monitoring Stop Time : 2013-10-11T21:23:10 Duration : 59 min		
IBM Systems Workload Estimator link : http://ibm.com/systems/support/tools/estimator (VIOS Sizings)		
SYSTEM - CONFIGURATION		
	Name	Value
	Processor Family	Architecture PowerPC Implementation POWER7_COMPAT_mode 64 bit
	Server Model	IBM 8202-E4D
	Server Frequency	3612.0 MHz
	Server - Online CPUs	10.0 cores
	Server - Maximum Supported CPUs	10.0 cores
	VIOS Level	2.2.2.2
	AIX Version	6.1.8.15
	AIX Build	08
	VIOS Advisor Release	0.1

Figure 3-31 System configuration

Figure 3-32 shows information about average and maximum I/O activity within the measured interval.







VIOS - I/O ACTIVITY							
	Name	Value					
	Disk I/O Activity	avg: 2057 iops @ 14.20 KB peak: 5823 iops @ 41KB					
	Network I/O Activity	[avgSend: 0 iops 0.0 MBps , avgRcv: 0 iops 0.0MBps] [peakSend: 0 iops 0.0MBps , peakRcv: 0 iops 0.0MBps]					
VIOS - DISK ADAPTERS							
	Name	Measured Value	Suggested Value	First Observed	Last Observed	Risk 1=lowest 5=highest	Impact 1=lowest 5=highest
	FC Adapter Count	2	-	2013-10-21T11:01:16	-	-	-
	FC Avg IOps	avg: 2057 iops @ 13KB	-	2013-10-21T11:01:16	2013-10-21T12:01:10	-	-
	FC Adapter Utilization	optimal	-	-	-	-	-
	FC Port Speeds	running at speed	-	-	-	-	-

Figure 3-32 VIOS I/O and network activity

Figure 3-33 shows the details about the physical disk drives that are visible to the VIOS. This example uses NPIV attached disks. Only the disk that contains the VIOS is shown.




VIOS - DISK DRIVES							
	Name	Measured Value	Suggested Value	First Observed	Last Observed	Risk 1=lowest 5=highest	Impact 1=lowest 5=highest
	Physical Drive Count	1	-	2013-10-11T20:23:18	-	-	-
	I/Os Blocked	pass	-	-	-	-	-
	Long I/O Latency	pass	-	-	-	-	-

Figure 3-33 VIOS disk drives

Figure 3-34 displays the average CPU usage. The VIOS footprint is small if you use NPIV. The CPU load increases with a higher virtualized network load.






VIOS - CPU				
	Name	Measured Value	Suggested Value	First Observed
	CPU Capacity	1.0 ent	-	2013-10-21T11:01:16
	CPU consumption	avg:3.5% (cores:0.1) high:12.0% (cores:0.2)	-	-
	Processing Mode	Shared CPU (Capped)	-	2013-10-21T11:01:16
	Virtual Processors	10	-	2013-10-21T11:01:16
	SMT Mode	SMT4	-	2013-10-21T11:01:16

Figure 3-34 VIOS CPU usage

Figure 3-35 shows the memory usage. If the VIOS starts to page, increase the size of the memory by using a DLPAR operation without the need for downtime.







VIOS - MEMORY							
	Name	Measured Value	Suggested Value	First Observed	Last Observed	Risk 1=lowest 5=highest	Impact 1=lowest 5=highest
	Real Memory	4.000 GB	4.000 GB	2013-10-21T11:01:16	-	1	2
	Available Memory	0.764 GB	1.5 GB Avail.	2013-10-21T11:01:09	2013-10-21T12:01:10	-	-
	Paging Rate	0.0 MB/s Paging Rate	-	-	-	-	-
	Paging Space Size	1.500 GB	-	2013-10-21T11:01:16	-	-	-
	Free Paging Space	1.484 GBfree	-	-	-	-	-
	Pinned Memory	1.716 GB pinned	-	-	-	-	-

Figure 3-35 VIOS memory usage

As shown in Figure 3-35 on page 94, there is a warning about the memory usage. Click the warning icon to see a detailed explanation, as shown in Figure 3-36. As this system still has about 20% of its memory unused, it is not necessary to take action, but the amount of available memory can decrease soon.

<p>What Is This? Real Memory: This is the amount of memory (in GB) assigned to the VIOS partition through the HMC.</p> <p>Why Important? Increase the amount of memory allocated to the VIOS partition or reduce memory consumption. Failing to do so can lead to memory page-in/outs, which can impact overall performance. The VIOS may be limited in its capabilities to support additional I/O devices.</p> <p>How To Modify? DLPAR add memory to the VIOS partition using the HMC or change the VIOS partition's profile using the HMC and power off/on the VIOS partition. Memory consumption can be reduced by lowering the amount of memory allocated for J2 inode cache-ing. Run these commands with root authority: <code>ioo -p -o j2_inodeCacheSize=200</code> and <code>ioo -p -o j2_metadataCacheSize=200</code></p>

Figure 3-36 Memory warning explanation

Alternatively, the VIOS Advisor tool can be downloaded from this website:

<https://www.ibm.com/developerworks/mydeveloperworks/wikis/home/wiki/Power%20Systems/page/VIOS%20Advisor>

The tool allows for a longer collection period of up to 1440 minutes and provides greater detail on I/O activity and network configuration.

3.7.2 Restarting the RMC connection

When the connection between HMC and VIOS is interrupted, the RMC connection can be lost. Without an active RMC connection, dynamic LPAR operations are no longer possible. Use the commands that are shown in Example 3-16 to check the connection state and restart the connection.

Example 3-16 Checking and restarting the RMC connection

```
$ oem_setup_env
# /usr/sbin/rsct/bin/rmcdomainstatus -s ctrmc

# /usr/sbin/rsct/install/bin/recfgct
0513-071 The ctcas Subsystem has been added.
0513-071 The ctrmc Subsystem has been added.
0513-059 The ctrmc Subsystem has been started. Subsystem PID is 9502788.

# /usr/sbin/rsct/bin/rmcdomainstatus -s ctrmc
Management Domain Status: Management Control Points
  I A 0x69e3763d06415474 0002 9.5.103.98
  I A 0x48864326026f84b4 0001 9.5.104.205

# exit
```

The `rmcdomainstatus` command now shows a connection to both attached HMCs.

3.7.3 Collecting troubleshooting data for support

This section describes the information you need to collect for support.

Gather information about the type, model, and serial number of the IBM Power System.

Use the commands that are shown in Example 3-17 to collect information about the installed operating system levels.

Example 3-17 Collecting operating system information

```
RZMB02V10S1$ ioslevel
2.2.2.1
RZMB02V10S1$ oem_platform_level
AIX Version:
6.1.0.0
```

To collect information about the virtual adapter configuration (including the error logs), run the **snap** command and download the resulting `snap.pax.Z` file by using FTP to your workstation. The process resembles what is shown in Example 3-18 and runs for several minutes.

Example 3-18 Running the snap command

```
$ snap
    Gathering data.  This may take 5 to 10 minutes.  Please wait.

Checking space requirement for general information..... done.
Checking space requirement for tcpip information..... done.
Checking space requirement for kernel information..... done.
Checking space requirement for printer information..... done.
Checking space requirement for dump information.....done.
Checking space requirement for sna information..... done.
Checking space requirement for filesys information..... done.
Checking space requirement for async information..... done.
Checking space requirement for lang information..... done.
Checking space requirement for XS25 information.....done.
Log directory set to /tmp/ibmsupt/ctsupt
Gathering information.....
Starting tar/compress process.....
*****done*****
Checking space requirement for install information..... done.
Checking space requirement for ssa information..... done.
...
...
...
Gathering pcixscsi system information..... done.
Gathering sissas system information..... done.
Gathering IB system information.....
Creating compressed pax file...
Starting pax/compress process... Please wait... done.

-rw-----  1 0      0      43131685 Oct 11 11:31 snap.pax.Z
    Finished gathering data.  Check your home directory for snap.pax.Z
$
```



PowerHA and the IBM Storwize Family

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

This chapter describes various usage scenarios for IBM PowerHA SystemMirror for i with the SAN Volume Controller and IBM Storwize storage servers. This chapter also describes the possibilities, position them against each other, and gives you an overview of how to configure them.

The following topics are covered in this chapter:

- ▶ Usage scenarios for IBM PowerHA SystemMirror for i
- ▶ Setting up IBM PowerHA SystemMirror for i

4.1 Usage scenarios for IBM PowerHA SystemMirror for i

The IBM PowerHA SystemMirror for i solution offers a complete end-to-end integrated clustering solution for high availability (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 high availability through planned and unplanned outages.

Because data resiliency is managed within the IBM i storage management architecture, there is no operator involvement. Just as there is no operator involvement with RAID 5 or disk mirroring.

As shown in Figure 4-1, the architecture of PowerHA SystemMirror for i provides several different implementation possibilities.

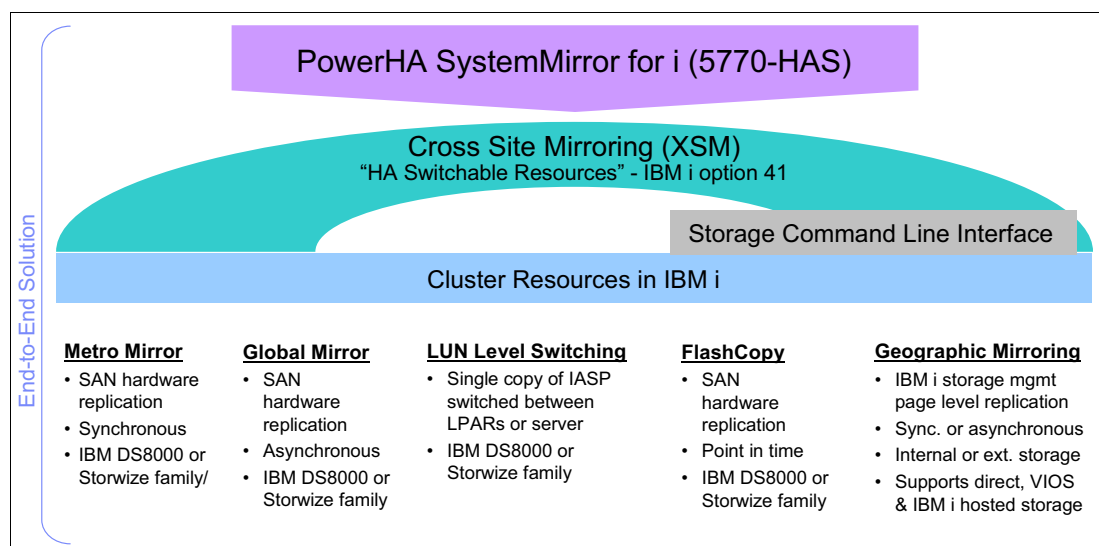


Figure 4-1 Overview of IBM PowerHA SystemMirror for i

Geographic mirroring offers IBM i clients an IBM i-based page-level replication solution for implementing high availability and disaster recovery with any kind of IBM i-supported internal or external storage solution. With IBM System Storage DS8000, SAN Volume Controller, or IBM Storwize family, storage servers clients can use storage-based remote replication functions for high availability and disaster recovery, LUN-level switching for local high availability, and FlashCopy for reducing save window outages by enabling the creation of a copy that is attached to a separate partition for offline backup to tape.

Which of these solutions make most sense in your environment depends largely on the storage infrastructure in place.

The following solutions are described in this section:

- ▶ 4.1.1, “Metro Mirror” on page 99
- ▶ 4.1.2, “Global Mirror” on page 99
- ▶ 4.1.3, “LUN level switching” on page 100
- ▶ 4.1.4, “FlashCopy” on page 102

For more information about the Storwize Copy Services functions, see the IBM Redbooks publication, *IBM System Storage SAN Volume Controller and Storwize V7000 Replication Family Services*, SG24-7574.

4.1.1 Metro Mirror

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

Figure 4-2 shows a typical setup with PowerHA and Metro Mirror. Each system has its own system ASP that contains the operating system, license programs, and some other objects. The IASP that contains application programs and data is mirrored from the primary storage system to the secondary storage system. Data is kept synchronous on a page level.

Metro Mirror is commonly used with PowerHA to achieve a high availability and disaster recovery solution for metropolitan distances up to 300 km when Storwize V3700, V5000, V7000, or SAN Volume Controller is used.

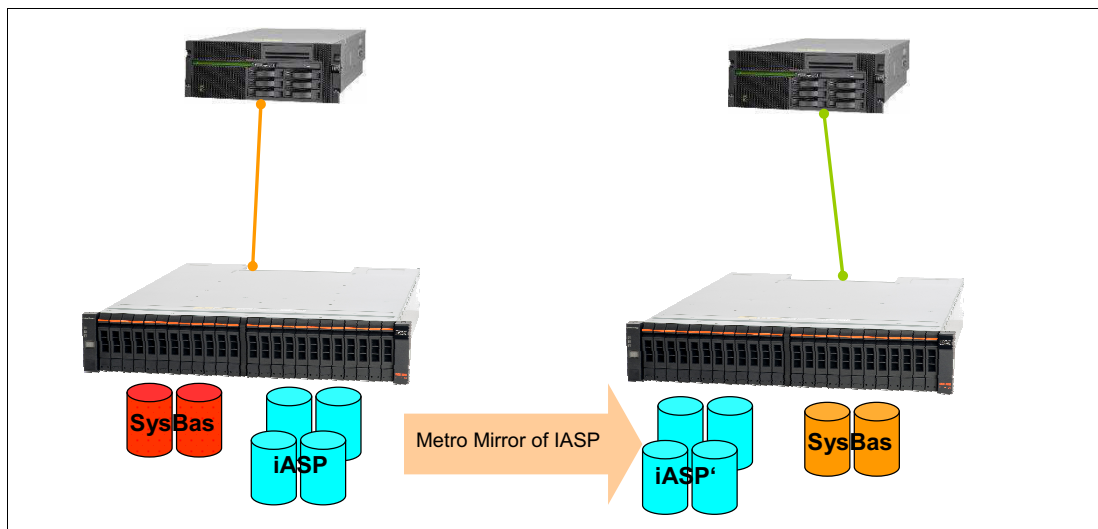


Figure 4-2 Metro Mirror

4.1.2 Global Mirror

Global Mirror is an asynchronous remote copy relationship between two volumes (VDisks) of equal (virtual) size. When the remote copy relationship is established, the preferred primary volume is designated as the master volume and the preferred secondary volume as the auxiliary volume. Every host write I/O sent to the Global Mirror primary volume is acknowledged back to the host after it is committed to the write cache of both nodes for the corresponding I/O group of the primary Storwize family system. At some later time (that is, asynchronously), this write update is sent by the primary Storwize family system to the secondary Storwize family system. Therefore, Global Mirror provides the capability to perform remote copy over long distances, up to the maximum supported round-trip latency of 80 ms, which exceeds the performance-related limitations of synchronous remote copy without host write I/O performance affects that are caused by remote replication delays.

Though the data is sent asynchronously from the primary to the secondary Storwize family system, the write ordering is maintained by sequence numbers that are assigned to acknowledged host write I/Os and with the secondary applying writes in order by their sequence number.

Consistency of the data at the remote site is maintained always. However, during a failure condition, the data at the remote site might be missing recent updates that were not sent or that were in-flight when a replication failure occurred. Therefore, the use of journaling to allow for proper crash consistent data recovery is of key importance. It is also considered a best practice to ensure crash consistent data recovery even when remote replication is not used. This allows you to recover from a loss of transient modified data in IBM i main storage after a potential sudden server crash due to a severe failure condition.

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

Global Mirror is commonly used with PowerHA to achieve disaster recovery where you want to put some distance between your sites. You want to combine it with a local high availability solution because the asynchronous implementation of Global Mirror does not ensure that the last transaction reached the backup system.

4.1.3 LUN level switching

LUN level switching, as shown in Figure 4-3 on page 101, provides support for an automated IBM i high availability solution by using a single copy of an independent ASP that can be switched (or ASPs), which can be switched between two IBM i cluster nodes for a local high availability solution.

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

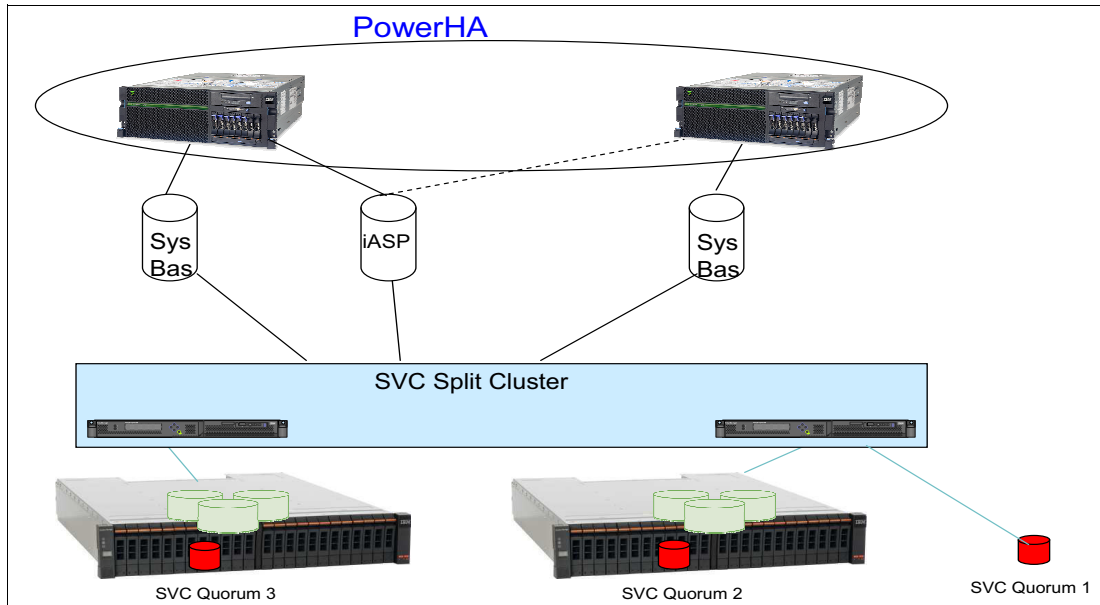


Figure 4-3 LUN level switching

SAN Volume Controller split cluster

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

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

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

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

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

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

4.1.4 FlashCopy

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

FlashCopy also is supported across heterogeneous storage systems that are attached to the IBM Storwize family, but only within the same Storwize family system. Up to 4096 FlashCopy relationships are supported per IBM Storwize family system, and up to 256 copies are supported per FlashCopy source volume. With IBM Storwize family version 6.2 and later, a FlashCopy target volume can also be a non-active remote copy primary volume, which eases restores from a previous FlashCopy in a remote copy environment by using the FlashCopy reverse function.

FlashCopy is commonly used to minimize downtimes for backup operations to tape. Instead of saving data from the production system, you create a FlashCopy of that data and attach it to another system where the actual save to tape is done, as shown in Figure 4-4 on page 103.

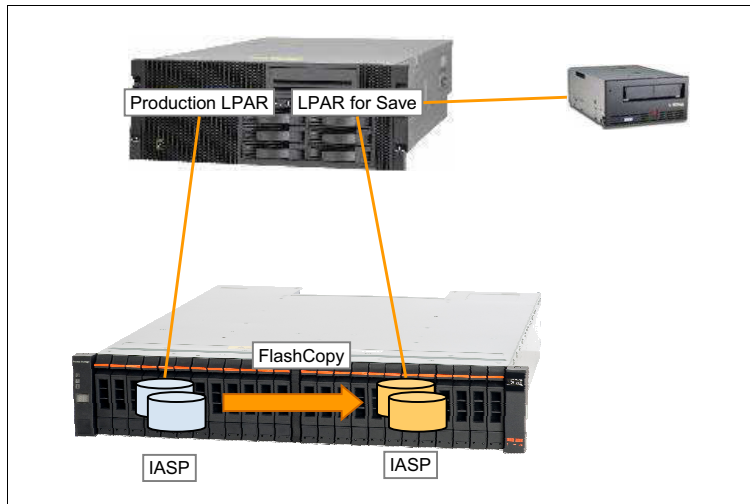


Figure 4-4 FlashCopy

In addition, FlashCopy can be used to provide a kind of disaster recovery option when large changes are made to your application. Create the FlashCopy before you start the application change. When you find that something does not work as expected, you can revert to the state before the application changes without waiting for a lengthy restore operation.

For more information about how an IASP is suspended for a brief period, the FlashCopy is taken, and the BRMS setup is done, see Chapter 6, “BRMS and FlashCopy” on page 173.

4.2 Setting up IBM PowerHA SystemMirror for i

This section describes the step-by-step setup instructions for configuring IBM PowerHA SystemMirror for i with the IBM Storwize family. The section includes the following topics:

- ▶ 4.2.1, “Requirements for PowerHA SystemMirror for i with SAN Volume Controller or Storwize” on page 103
- ▶ 4.2.2, “SAN Volume Controller split cluster and LUN level switching” on page 105
- ▶ 4.2.3, “LUN level switching in combination with Global Mirror” on page 137
- ▶ 4.2.4, “Metro Mirror setup” on page 150
- ▶ 4.2.5, “Adding LUNs to an existing IASP” on page 150

4.2.1 Requirements for PowerHA SystemMirror for i with SAN Volume Controller or Storwize

To work properly, PowerHA SystemMirror for i must communicate with the IBM Storwize Family system. This is done by using Secure Shell (SSH). You must create an SSH key pair on IBM i and attach the SSH public key to a user profile on the IBM Storwize Family system. The corresponding private key file to be used is specified in the creation of the ASP copy descriptions. It must be distributed to the same directory on all nodes in the cluster.

Generating the SSH key requires the following IBM i products to be installed:

- ▶ 5770-SS1 option 33 (Portable Application Solution Environment)
- ▶ 5733-SC1 base and option 1 (IBM Portable Utilities for i and OpenSSH, OpenSSL, zlib)

Note: The Enterprise edition of PowerHA is required for two-site Copy Services solutions.

Generating the key file is done on IBM i from Qshell. Example 4-1 shows the commands in Qshell to generate the key and the files that are created. Ensure that the directory you specify in the **keygen** command exists before the command is issued. Also, the user profile you are using when you are generating the key can have a maximum length of eight characters only.

Example 4-1 Generation of the SSH key pair in Qshell

```
> cd /QIBM/UserData/HASM/hads/.ssh/  
$  
> ssh-keygen -t rsa -f id_rsa -N ''  
Generating public/private rsa key pair.  
Your identification has been saved in id_rsa.  
Your public key has been saved in id_rsa.pub.  
The key fingerprint is:  
76:49:e1:9e:04:0a:c5:e2:68:3a:d0:6b:0b:4b:e2:2e  
powerha@Z1436B23.rch.stglabs.ibm.com  
$  
> ls -la  
total 64  
drwxrwsrwx 2 powerha 0 8192 Oct 07 17:02 .  
drwx--Sr-x 3 qsys 0 8192 Oct 07 15:19 ..  
-rw----- 1 powerha 0 1679 Oct 07 17:02 id_rsa  
-rw-r--r-- 1 powerha 0 414 Oct 07 21 17:02 id_rsa.pub
```

You then must import the `id_rsa.pub` file into a user on the Storwize family system. This user must have a minimum role of Copy Operator to perform the functions that are needed for PowerHA to work properly. If you want to use LUN level switching, make sure that the user has a minimum role of Administrator because PowerHA must change host attachments when switching the IASP to the secondary system.

Transfer the file to your PC and import it to the Storwize user, as shown in Figure 4-5.

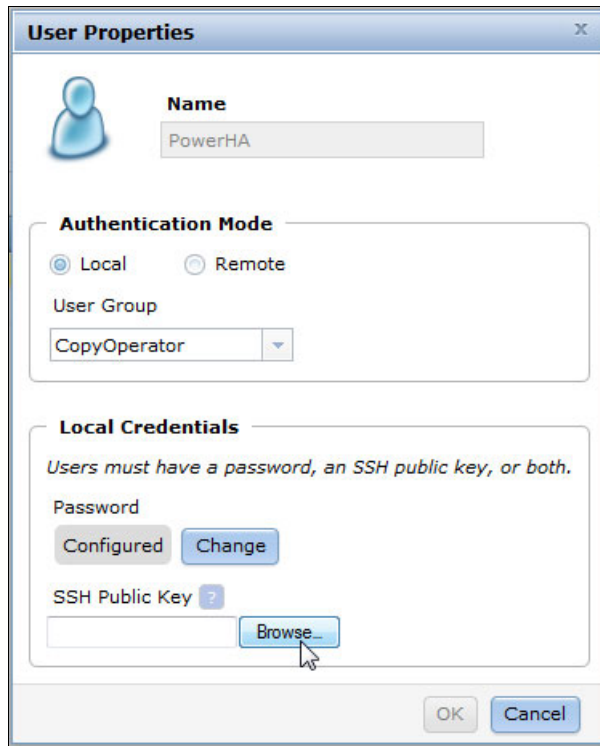


Figure 4-5 Importing SSH public key file to SAN Volume Controller user

Make sure to distribute the `id_rsa` file to all nodes in your cluster to the same directory. The user profile QHAUSRPRF must have at least *R data authority to the key file on each system.

4.2.2 SAN Volume Controller split cluster and LUN level switching

This section provides an example of configuring SAN Volume Controller split cluster and LUN level switching. In this example, it is assumed that SAN Volume Controller split cluster was set up correctly. Also, ensure that SAN zoning is set up so that the IASP can be reached from the primary cluster node and from the secondary cluster node.

Complete the following steps by using the GUI. These steps cover creating the cluster, creating the cluster resource group, and creating the ASP copy descriptions.

To complete these steps by using a CLI, see “Using CLI commands to setup SAN Volume Controller split cluster and LUN level switching” on page 136:

1. Log in to the IBM Systems Director Navigator for i by using the following URL:
`http://your_server_IP_address:2001`

- From the list of possible tasks that is shown in Figure 4-6, select **PowerHA** to access the PowerHA configuration and management GUI.

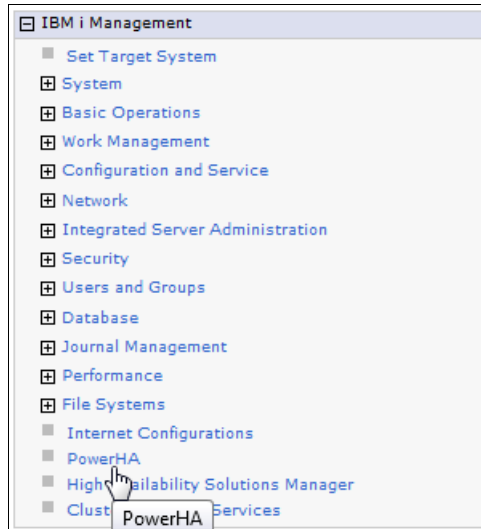


Figure 4-6 Cluster creation from IBM Systems Director Navigator

- As shown in Figure 4-7, there is no configured cluster. Click **Create a new Cluster** to start the cluster creation wizard.

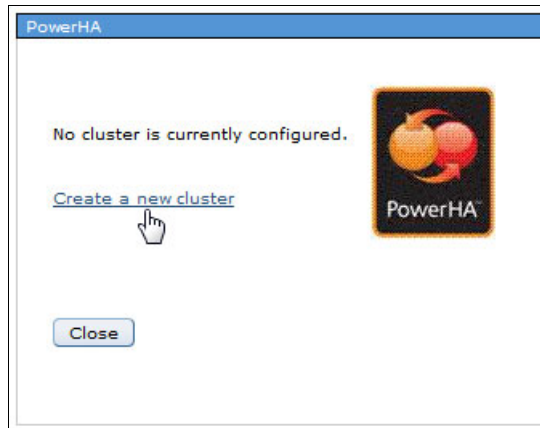


Figure 4-7 Creating a cluster

- On the Create Cluster welcome panel, click **Next** to continue to the cluster setup wizard.

5. As shown in Figure 4-8, on the Create Cluster Name and Version panel, you must provide a name for the cluster. With current fixes installed, the PowerHA version is 2.2. Click **Next**.

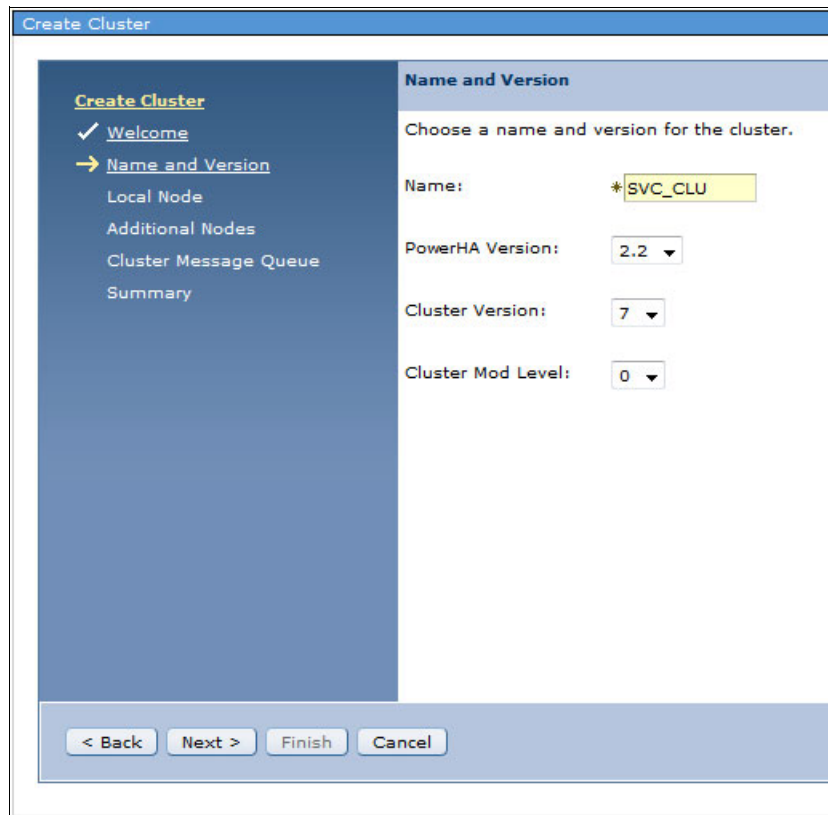


Figure 4-8 Cluster creation, specifying the cluster name and version

- In the Create Cluster Local Node panel, provide information about the local cluster node, as shown in Figure 4-9. Specify a node name and one or two IP addresses that are used for cluster communication. You can enter these addresses manually or click the **Use entry from below** drop-down menu and select the address from the list that is shown.

Tip: It is a common practice to use two different IP connections for the cluster heartbeat so that if there is a partial network failure, cluster communication is still available.

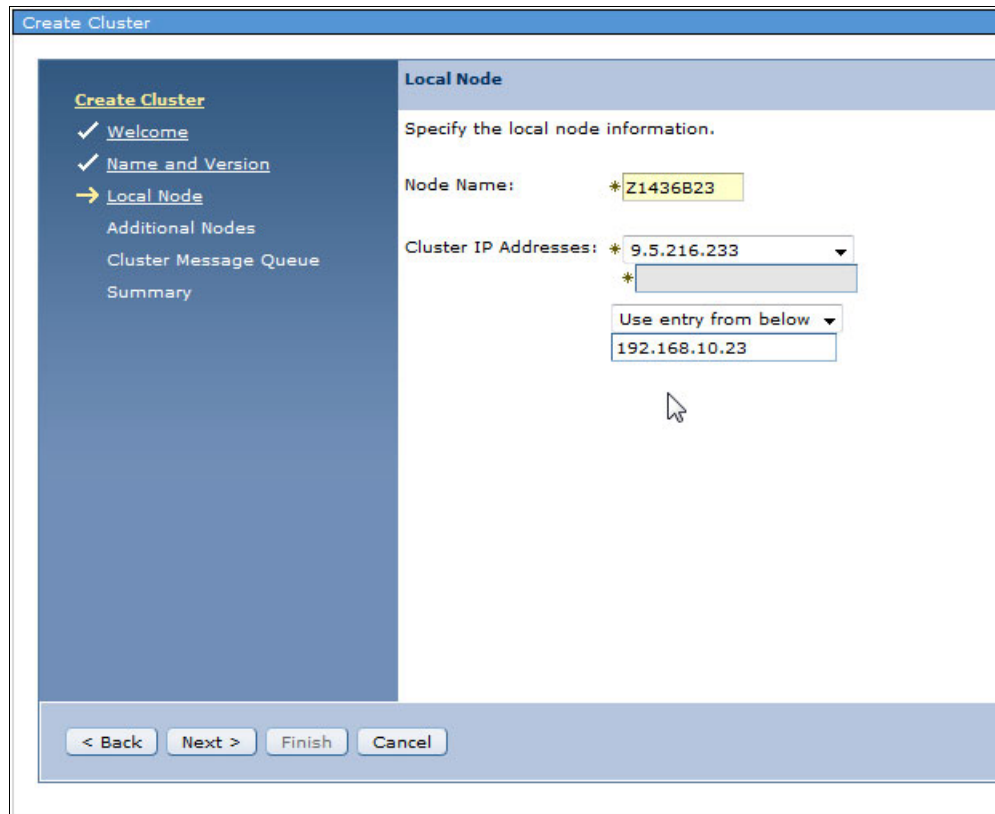


Figure 4-9 Cluster creation, local node specifications

7. In the Create Cluster Additional Nodes panel, specify the information for the second node in the cluster. Click **Add**, as shown in Figure 4-10, and specify the node name and the IP addresses for the second cluster node. You can also specify whether clustering is started after the creation of the cluster is finished.

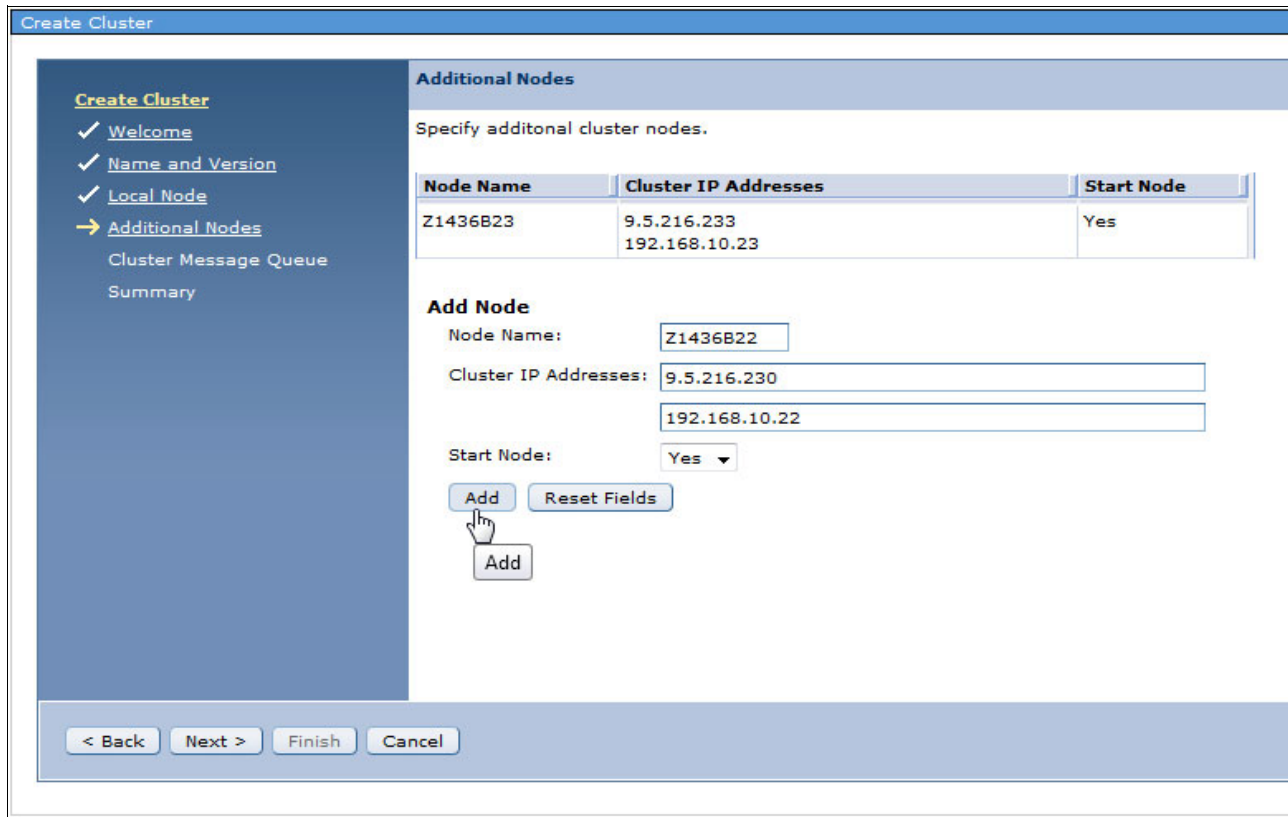


Figure 4-10 Cluster creation, adding the second node

8. When all of the nodes in your cluster are specified, click **Next**, as shown in Figure 4-11.

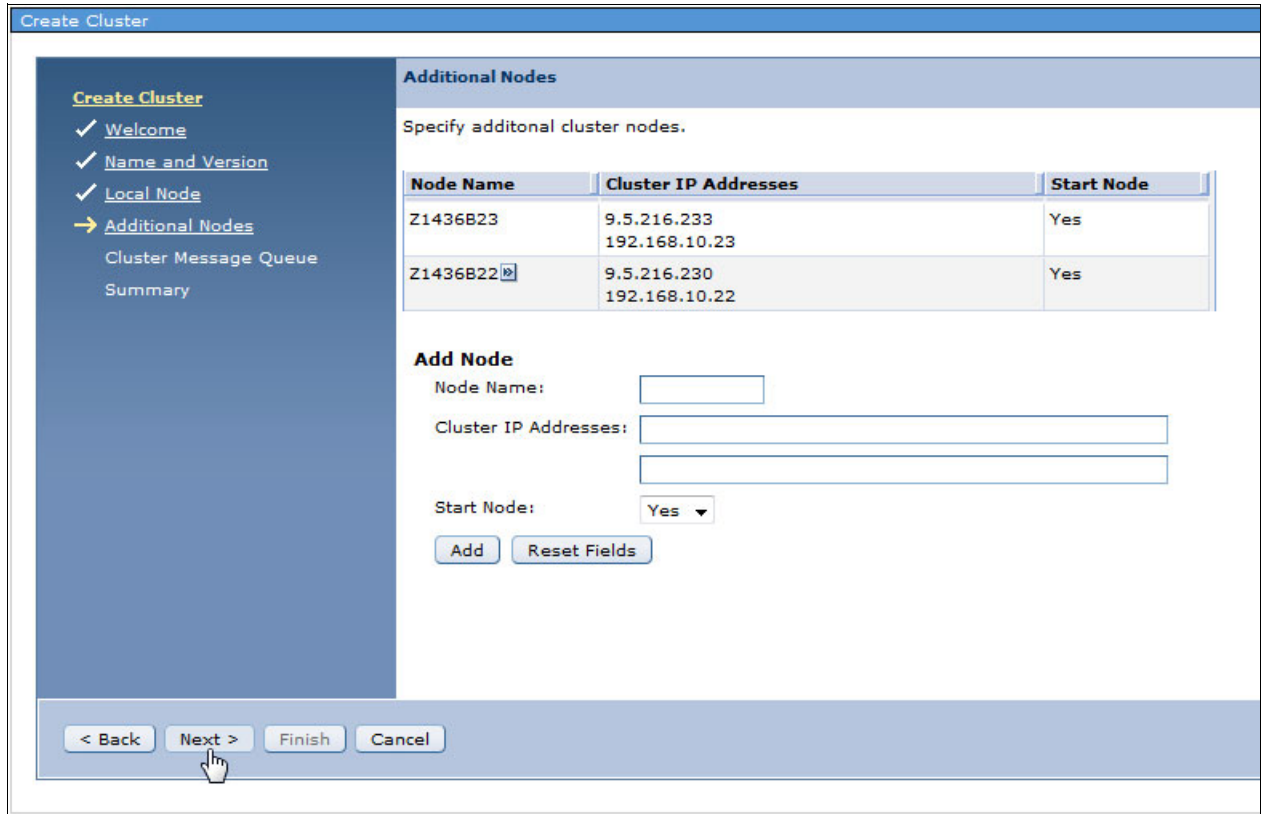


Figure 4-11 Cluster creation, all nodes specified

9. In the Create Cluster Message Queue panel, you must decide whether you want to use a cluster message queue, as shown in Figure 4-12. When a message queue is specified, a message is sent to that message queue on the secondary system when there is a failure on the primary system. You can also specify how long that message waits for an answer and what the standard action is after that time passes. Click **Next**.

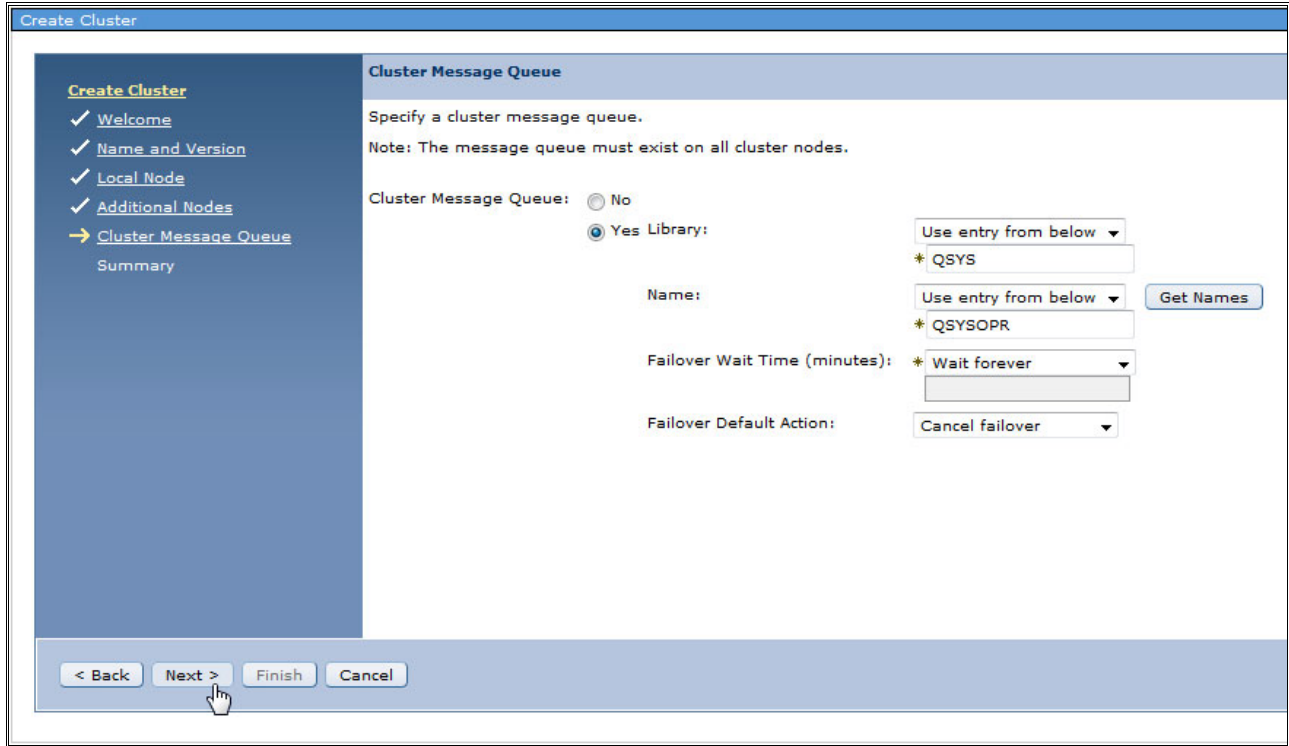


Figure 4-12 Cluster creation, cluster message queue

10. As shown in Figure 4-13, you see an overview of the configuration options that you chose. Click **Finish** to start the cluster creation.

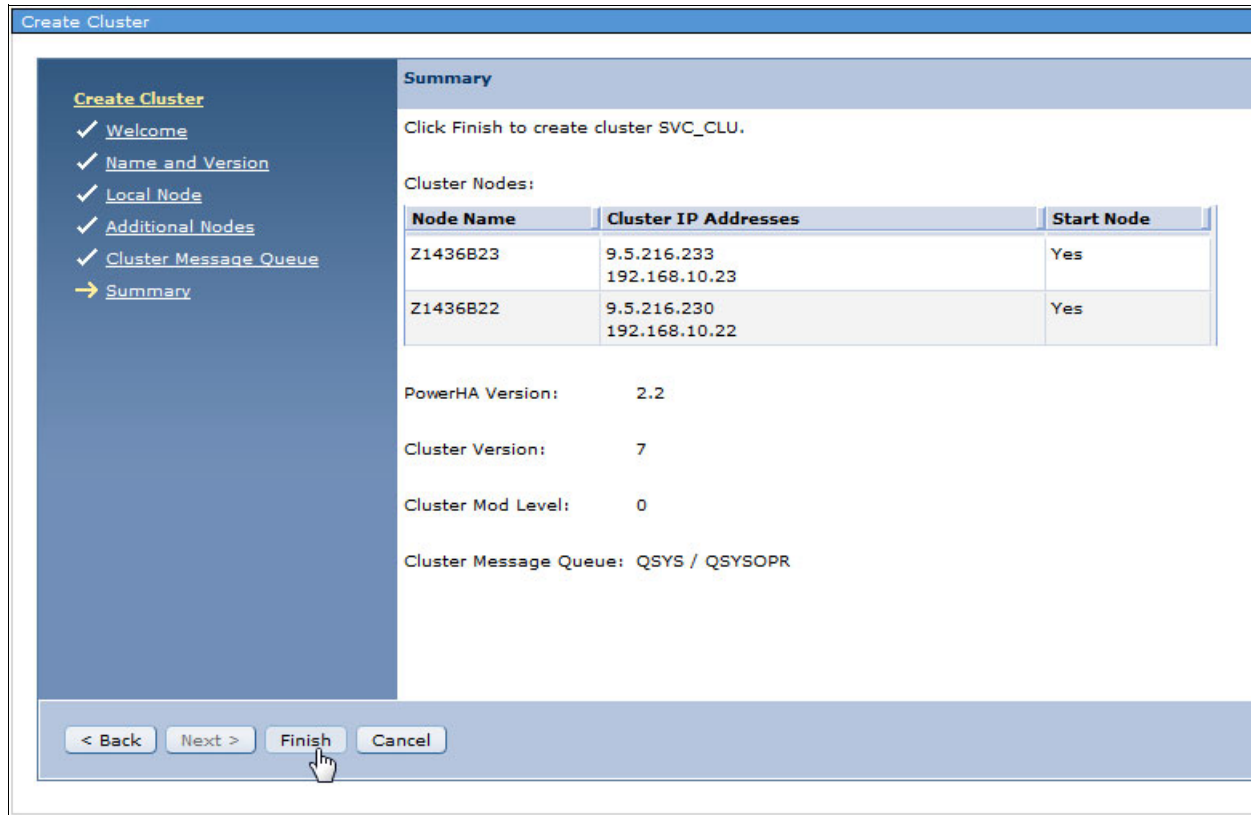


Figure 4-13 Cluster creation summary

11. Status messages appear as the cluster is created, as shown in Figure 4-14. Click **Close** to return to the PowerHA main menu.

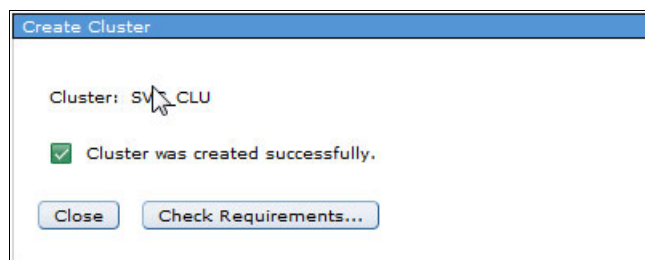


Figure 4-14 Cluster creation successful message

12. From the PowerHA main menu (as shown in Figure 4-15), select **Cluster Nodes** to edit the cluster nodes that were created.

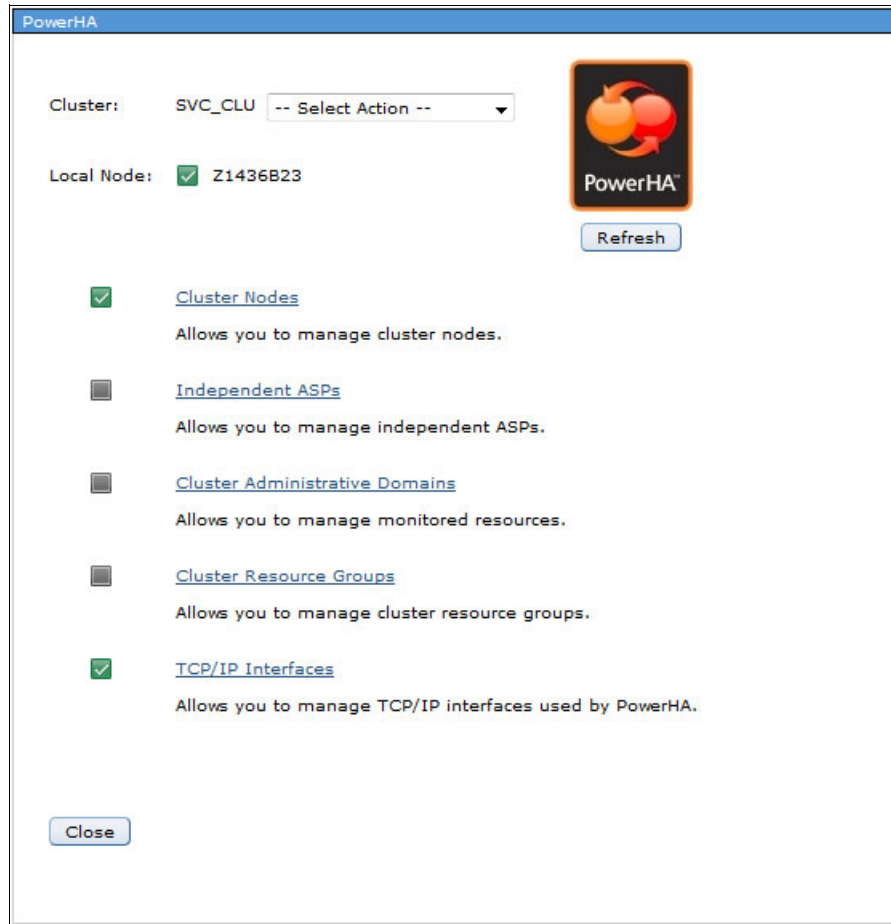


Figure 4-15 PowerHA main menu

13. Figure 4-16 shows the two nodes of the cluster that was created in this example. Currently, there is no device domain that is specified for the two nodes. The device domain is needed so that you can switch an IASP between the nodes.

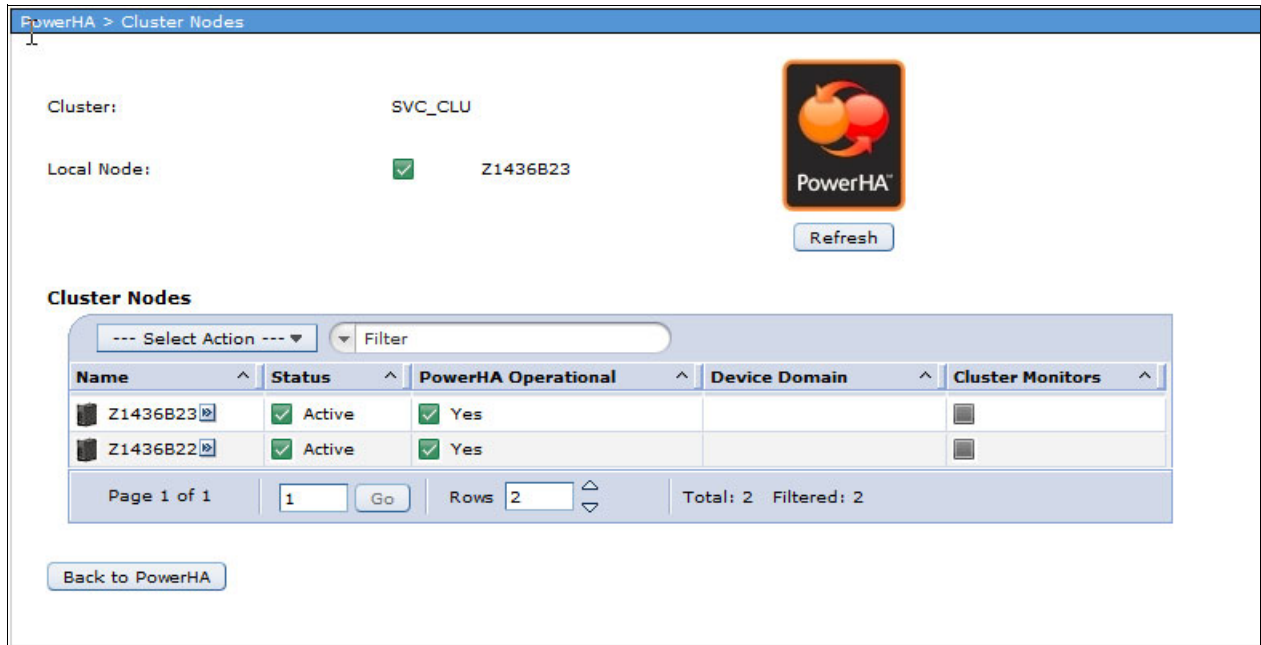


Figure 4-16 Nodes in the cluster

14. Right-click each node in the cluster and select **Properties**, as shown in Figure 4-17.

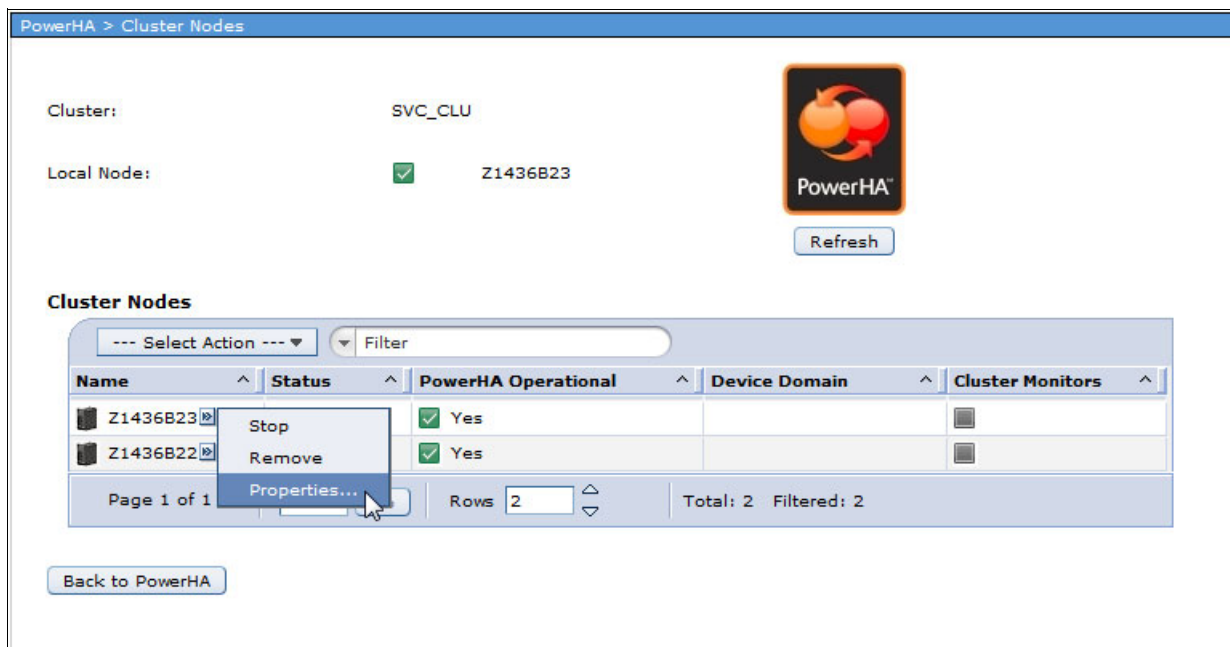


Figure 4-17 Node properties selection

15. Figure 4-18 shows the properties for cluster node Z1436B23 from our example. There is no entry for the device domain. Click **Edit**.

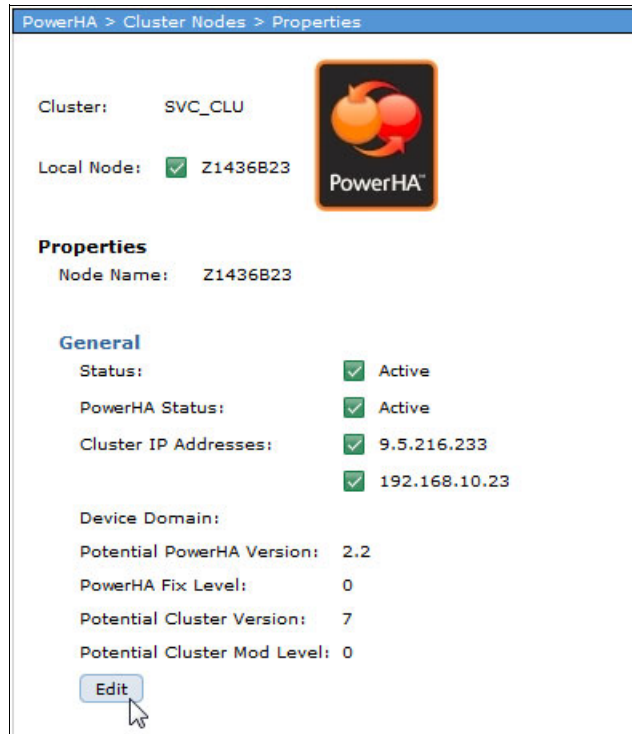


Figure 4-18 Cluster node properties

16. Define the name of the device domain, as shown in Figure 4-19. If there is already a device domain that is created within the cluster, you can choose the device domain from the drop-down menu.

Note: If an IASP exists on one of the cluster nodes, that node must be added first into the device domain. Only the first node in a device domain can have an existing IASP. Nodes with an existing IASP cannot be added to an existing device domain.

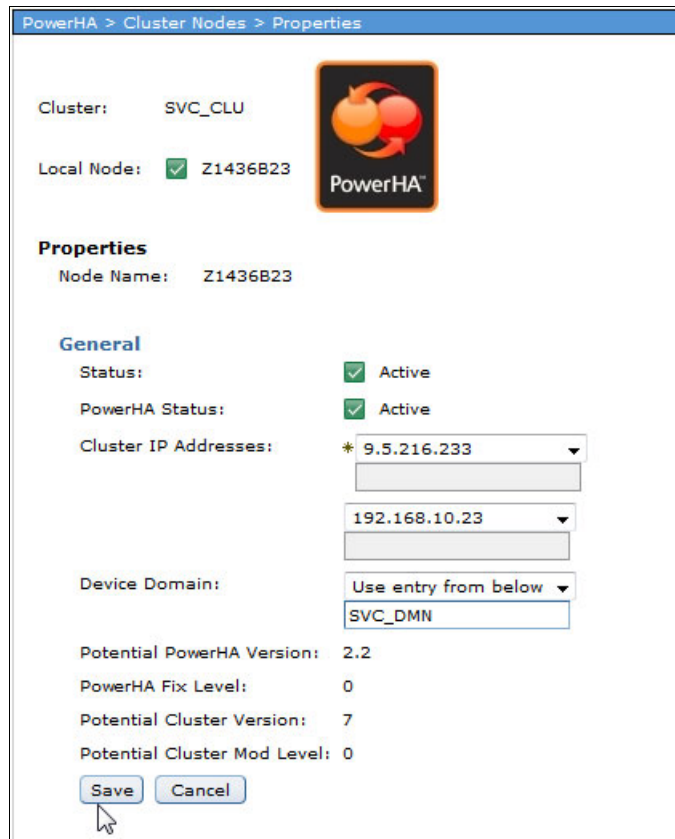



Figure 4-19 Specifying the device domain





17. Device domains must be defined for each node in the cluster. As shown in Figure 4-20, the cluster node overview shows the active cluster nodes with their device domains.

Cluster: SVC_CLU

Local Node: Z1436B23

 Refresh

Cluster Nodes

Name	Status	PowerHA Operational	Device Domain	Cluster Monitors
 Z1436B23	<input checked="" type="checkbox"/> Active	<input checked="" type="checkbox"/> Yes	SVC_DMN	
 Z1436B22	<input checked="" type="checkbox"/> Active	<input checked="" type="checkbox"/> Yes	SVC_DMN	

Page 1 of 1 | 1 | Go | Rows 2 | Total: 2 Filtered: 2

Back to PowerHA

Figure 4-20 Device domains added to each node in the cluster

18. Define the cluster resource group that is used to switch the IASP between the cluster nodes. From the main PowerHA window, click **Cluster Resource Group**, as shown in Figure 4-21.

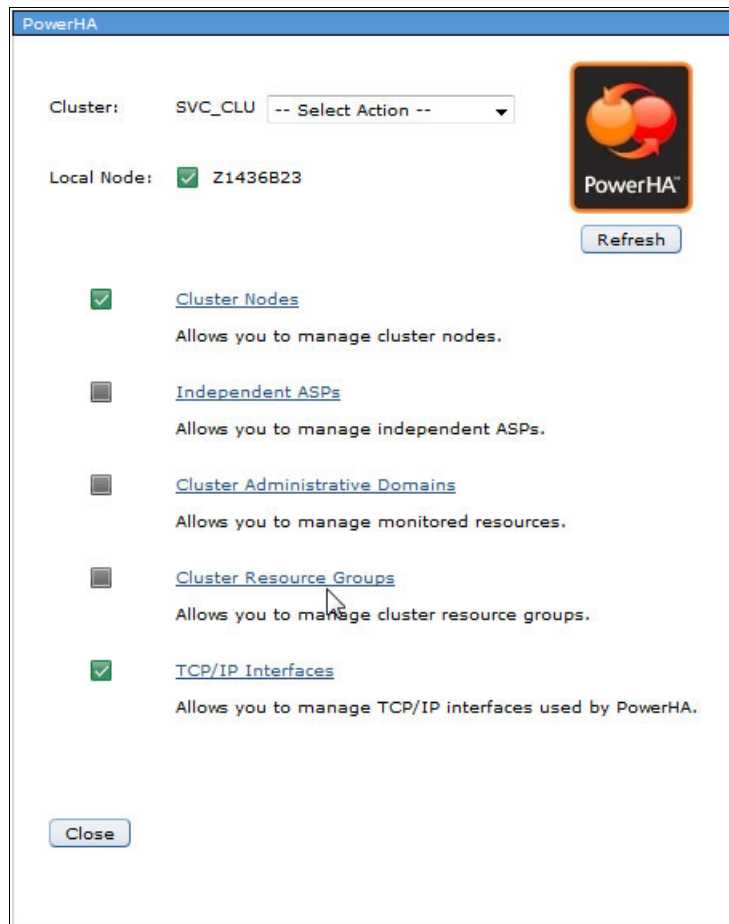


Figure 4-21 PowerHA main menu, selecting Cluster Resource Groups

19. In the PowerHA Cluster Resource Groups window, select **Create Cluster Resource Group** from the Select Action drop-down menu, as shown in Figure 4-22.

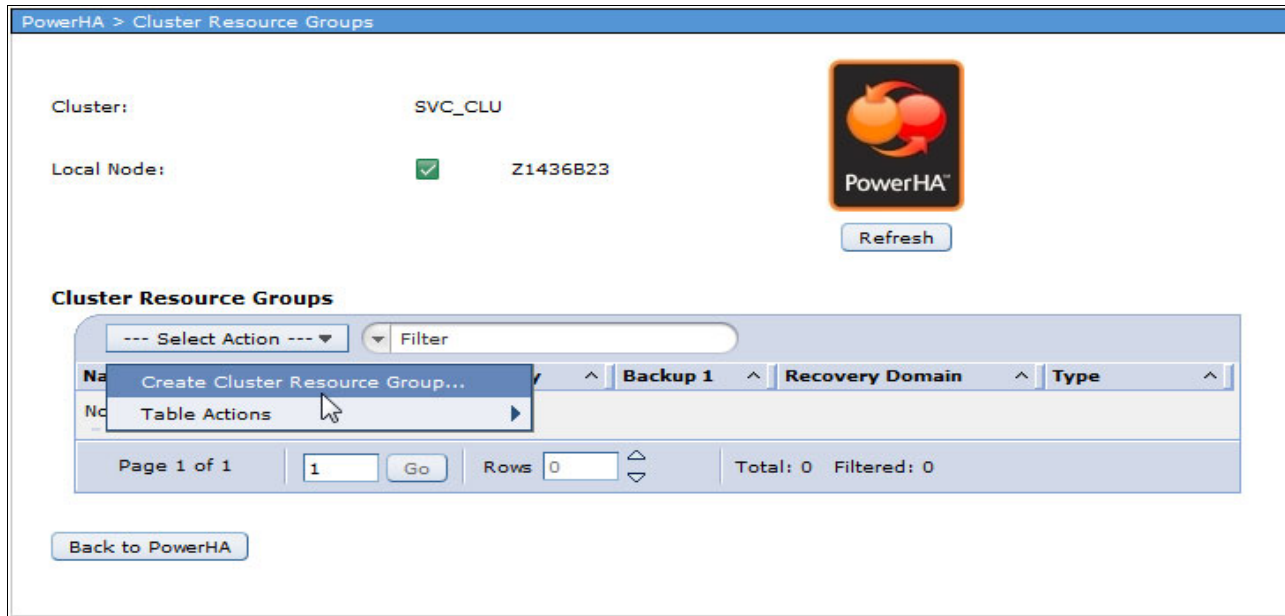


Figure 4-22 Specifying to create a cluster resource group

20. In the Create Cluster Resource Group Name and Type panel, specify a name for the cluster resource group. The type must be **Device**, as shown in Figure 4-23, because you want to switch an IASP device that uses this cluster resource group. Click **Next**.

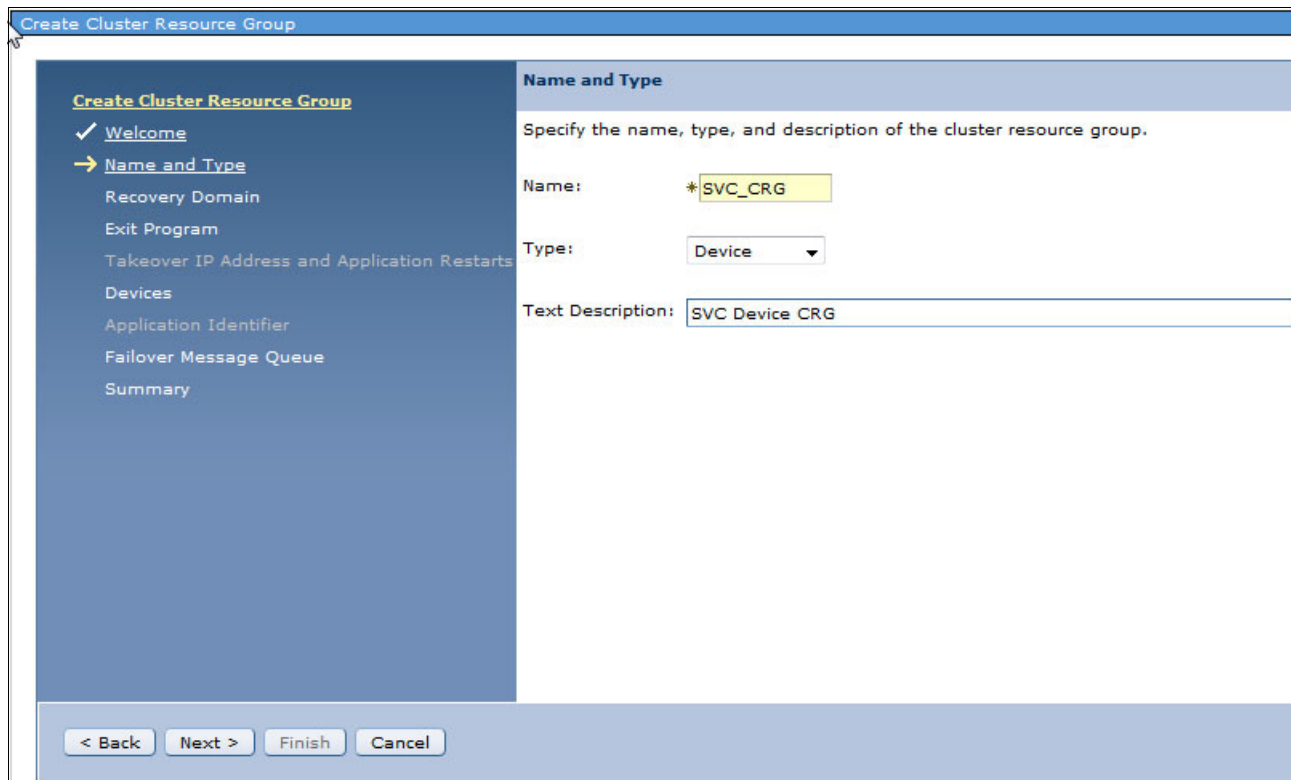


Figure 4-23 Create Cluster Resource Group, specifying the name and type

21. In the Create Cluster Resource Group Recovery Domain panel, define the recovery domain. The recovery domain defines which node in the cluster is the primary node and the order of backup nodes in the cluster.

Add the primary node into the recovery domain as shown in Figure 4-24. A site name also must be specified. Click **Add** to add the primary node into the recovery domain.

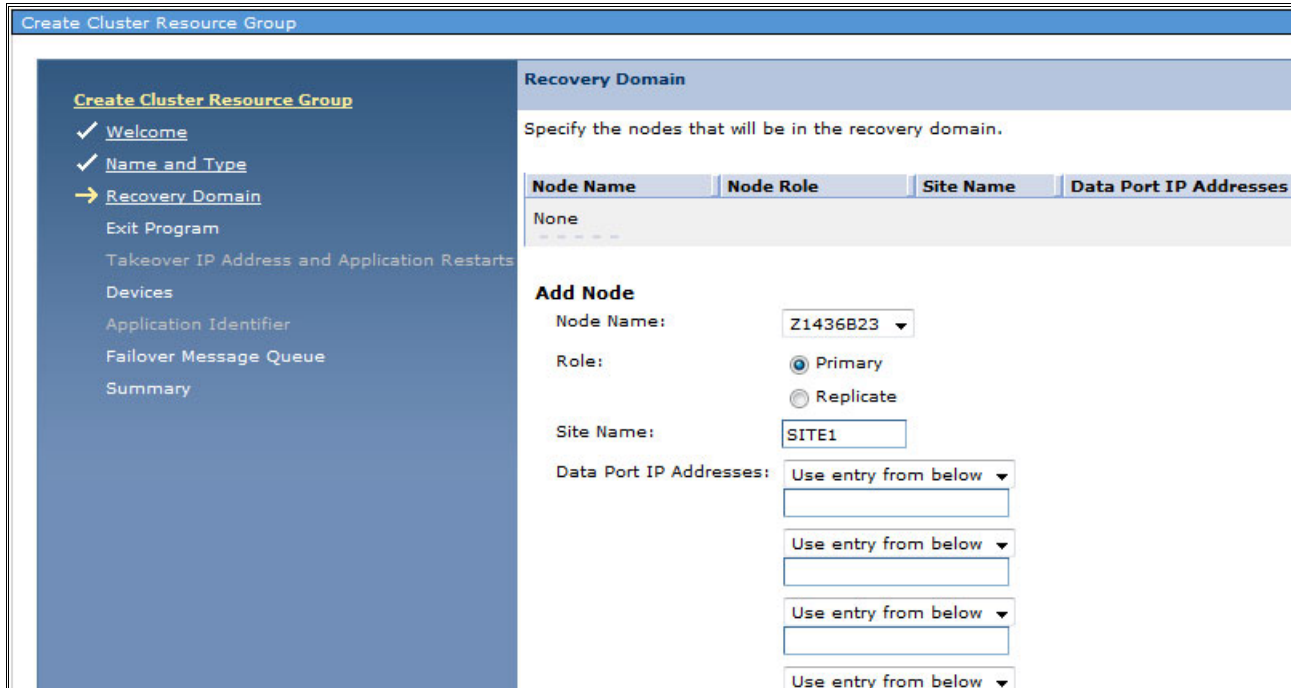


Figure 4-24 Create Cluster Resource Group, recovery domain for the primary node

22. Add the secondary node into the recovery domain. Specify Backup as the role. Ensure that the site name you provide here is identical to the site name that you provided for the primary node, as shown in Figure 4-25 on page 121. Click **Add** to add the node into the recovery domain.

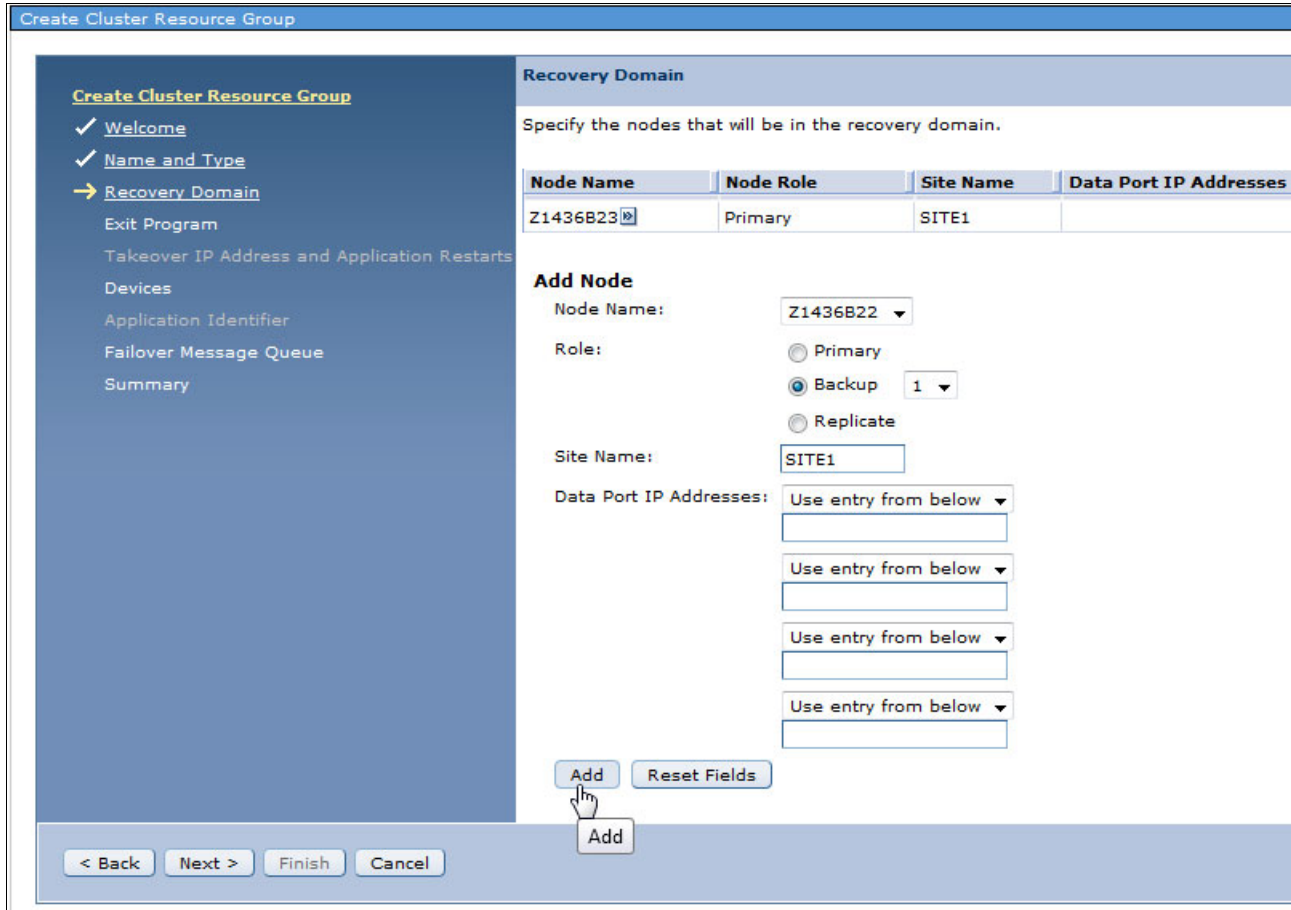


Figure 4-25 Create Cluster Resource Group, recovery domain for the backup node

23. The recovery domain overview shows the two cluster nodes with their node roles and site names, as shown in Figure 4-26. Click **Next**.

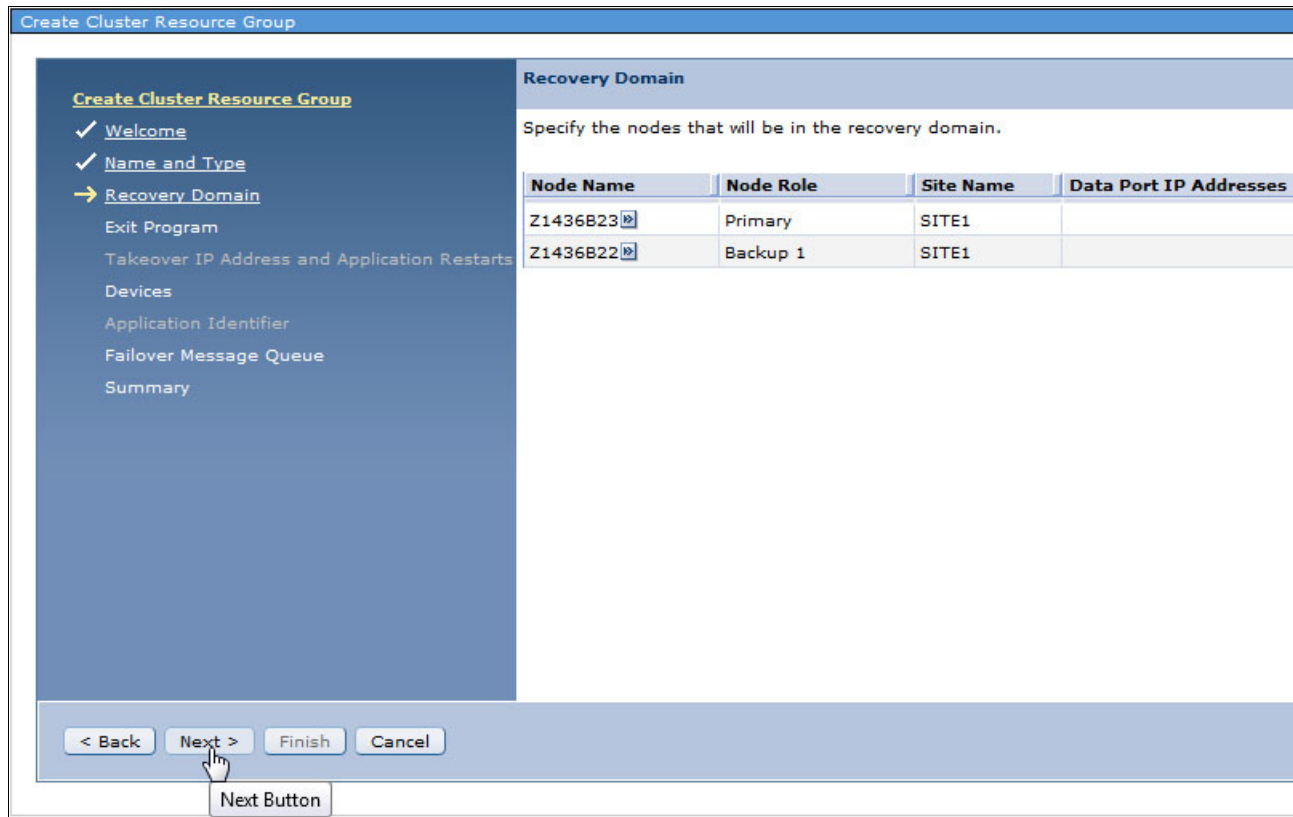


Figure 4-26 Create Cluster Resource Group, recovery domains created

24. In the Create Cluster Resource Group Exit Program panel, you can define a cluster resource group exit program. This program can be used to start functions that are specific to your environment when failover or switchover is done; for example, to start a certain application. For this example, an exit program is not used, as shown in Figure 4-27. Click **Next**.

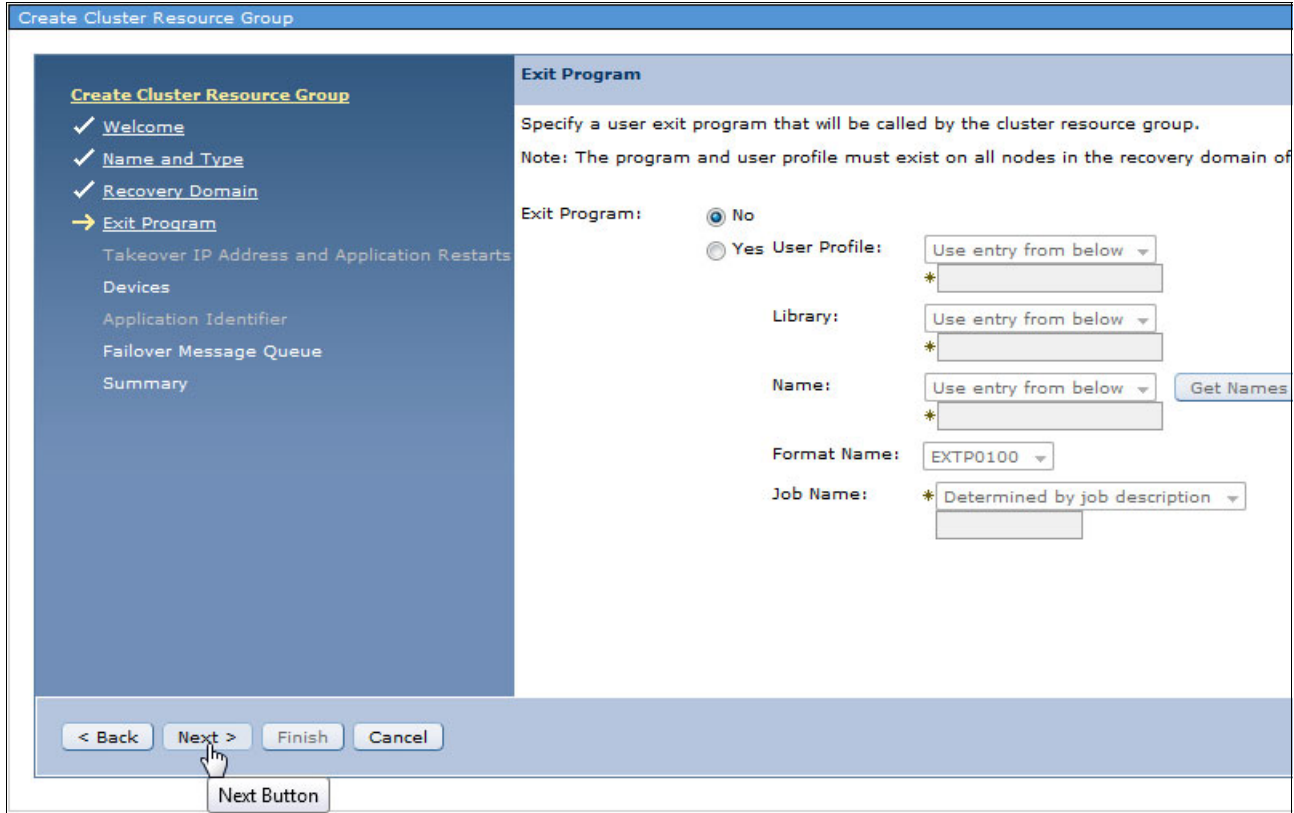


Figure 4-27 Create Cluster Resource Group, specifying an exit program

25. In the Create Cluster Resource Group Devices panel, define which device is controlled by the cluster resource group. Specify the name of the IASP, as shown in Figure 4-28.

You also can specify whether the IASP is varied on at the secondary system after a failover or switchover occurs. The server takeover IP address can be used to move one IP address to the current primary node. Click **Add** to add the device into the cluster resource group.

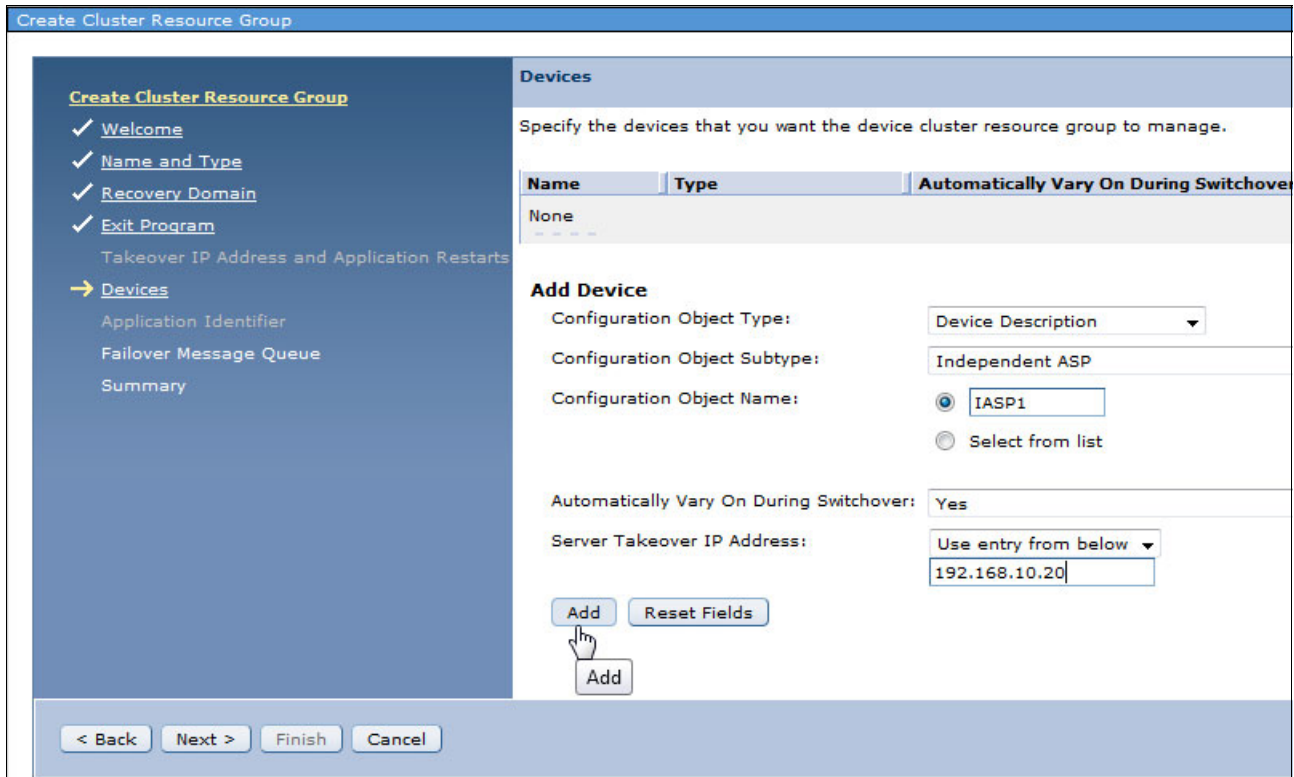


Figure 4-28 Create Cluster Resource Group, defining devices that are controlled by the cluster

26. You can add multiple devices into one cluster resource group. Click **Next** when the devices you need are part of the cluster resource group, as shown in Figure 4-29.

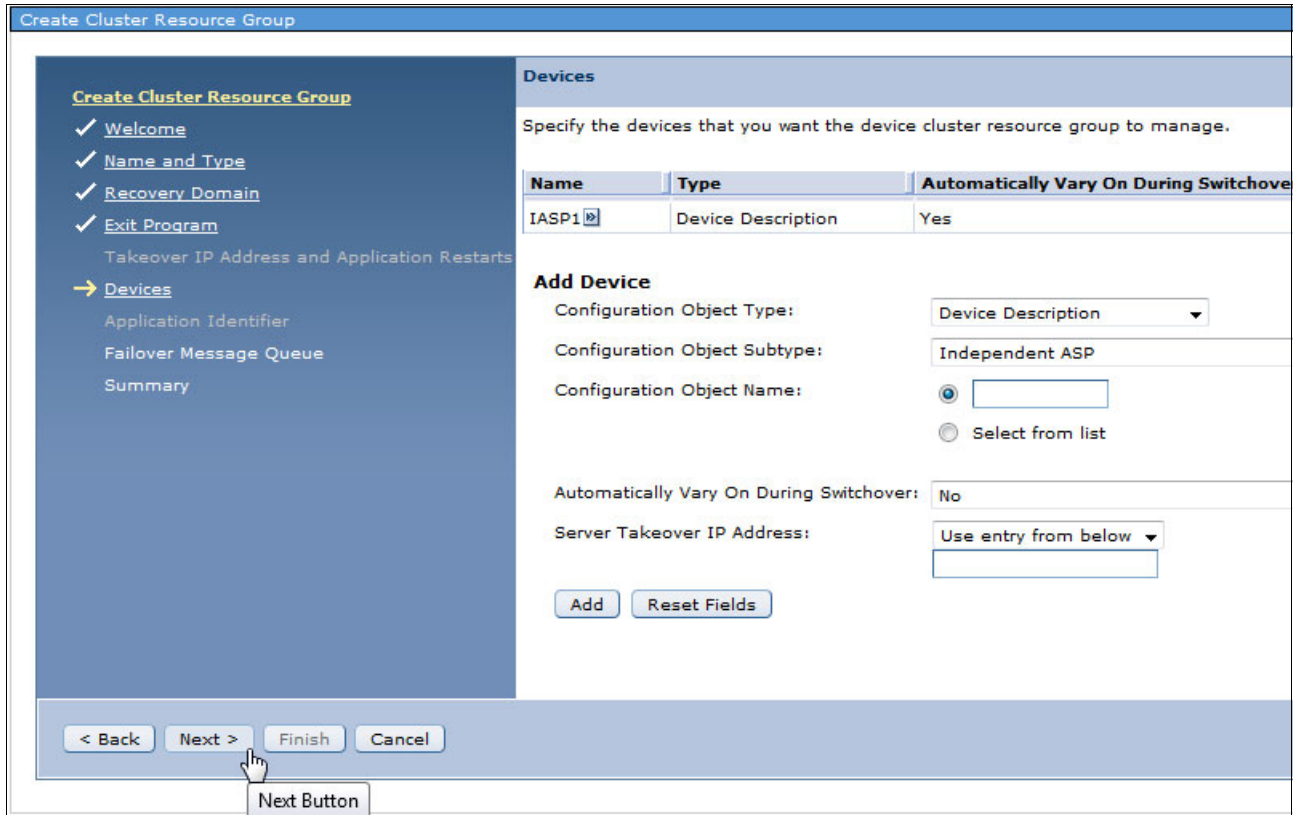


Figure 4-29 Create Cluster Resource Group, device added

27. In the Create Cluster Resource Group Failover Message Queue panel, you can define a failover message queue for the cluster resource group. As our example already defined a cluster-wide failover message queue, it does not define another one at the cluster resource level, as shown in Figure 4-30. Click **Next**.

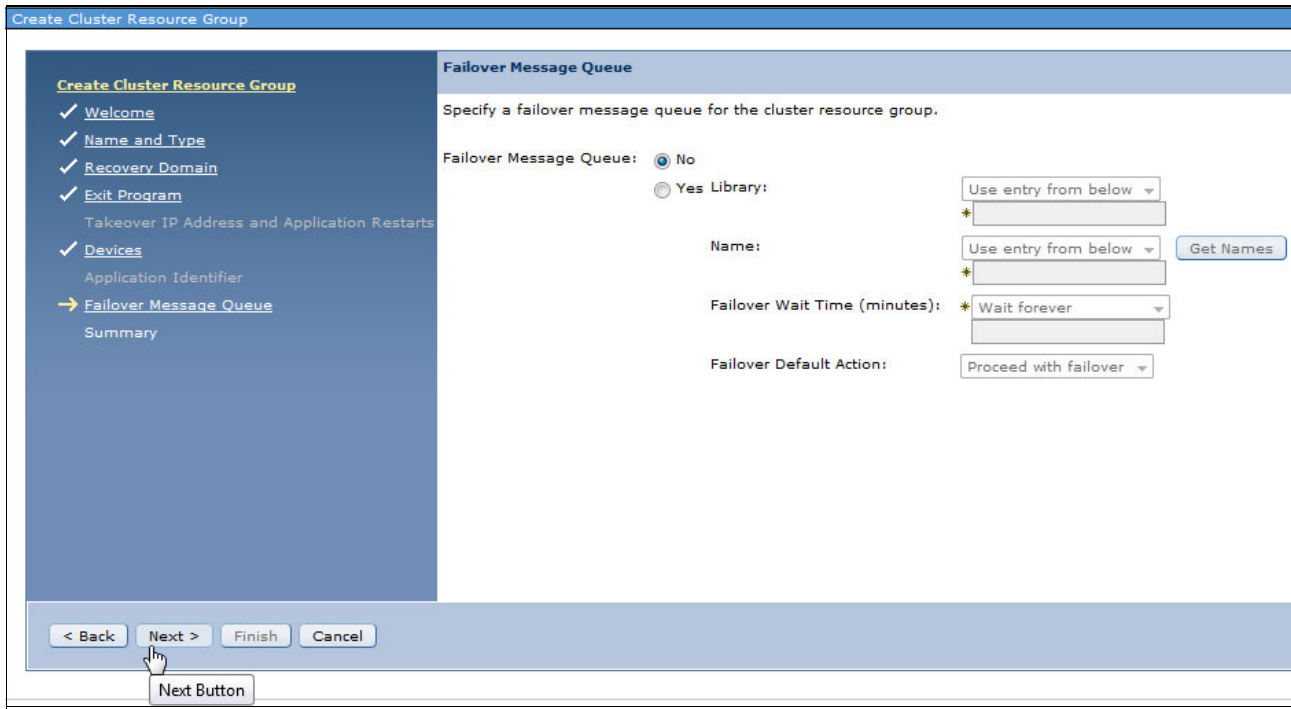


Figure 4-30 Create Cluster Resource Group, specifying a failover message queue

28. The Create Cluster Resource Group Summary panel that is shown in Figure 4-31 displays the configuration options that were entered. Click **Finish** to create the cluster resource group.

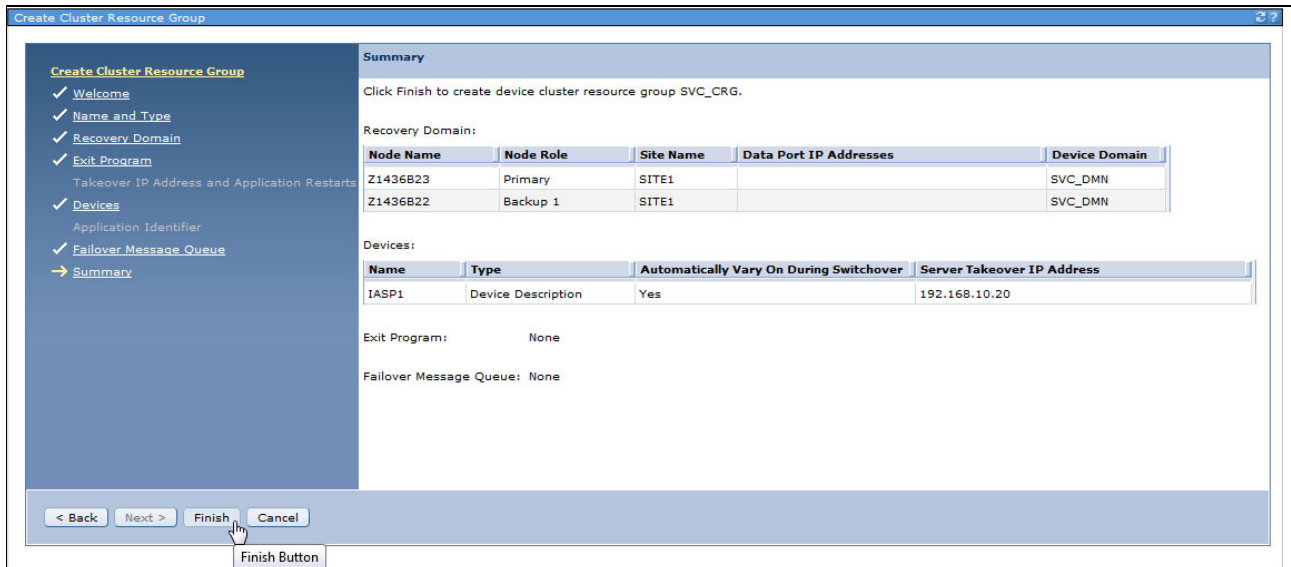



Figure 4-31 Create Cluster Resource Group, summary

The cluster resource group is created and is in an inactive status, as shown in Figure 4-32.

PowerHA > Cluster Resource Groups



Cluster: SVC_CLU

Local Node: Z1436B23

 Refresh

Cluster Resource Groups

--- Select Action --- Filter

Name	Status	Primary	Backup 1	Recovery Domain	Type
 SVC_CRG	 Inactive	Z1436B23	Z1436B22	<input checked="" type="checkbox"/>	Device

Page 1 of 1 1 Go Rows 1 Total: 1 Filtered: 1

[Back to PowerHA](#)

Figure 4-32 Cluster resource group created

29. The last step to set up LUN level switching with PowerHA SystemMirror for i is to create an ASP Copy description. From the PowerHA main menu, click **Independent ASPs**, as shown in Figure 4-33.

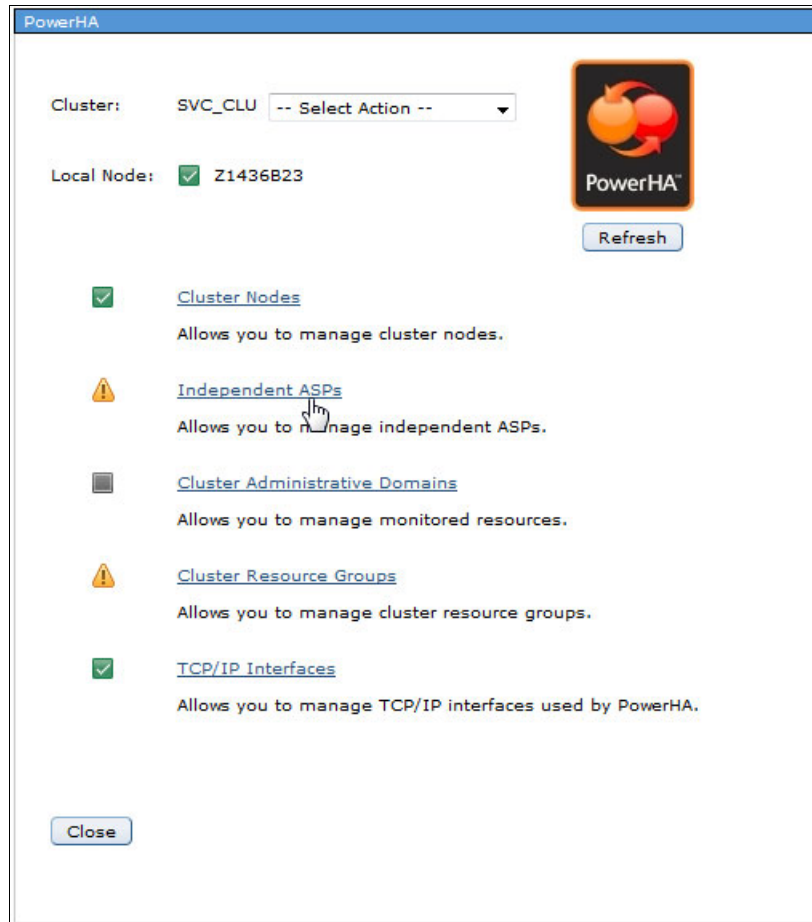


Figure 4-33 PowerHA main menu, selecting Independent ASPs

30. In the PowerHA Independent ASPs window, select **ASP Sessions and Copy Descriptions** from the Select Action drop-down menu, as shown in Figure 4-34.



Figure 4-34 Selecting ASPs and Copy descriptions

31. In the ASP Sessions and Copy Descriptions window, select **Add SVC ASP Copy Description** from the Select Action pull-down menu in the ASP Copy Descriptions section, as shown in Figure 4-35. This option is used when the storage attached is a member of the IBM Storwize family.

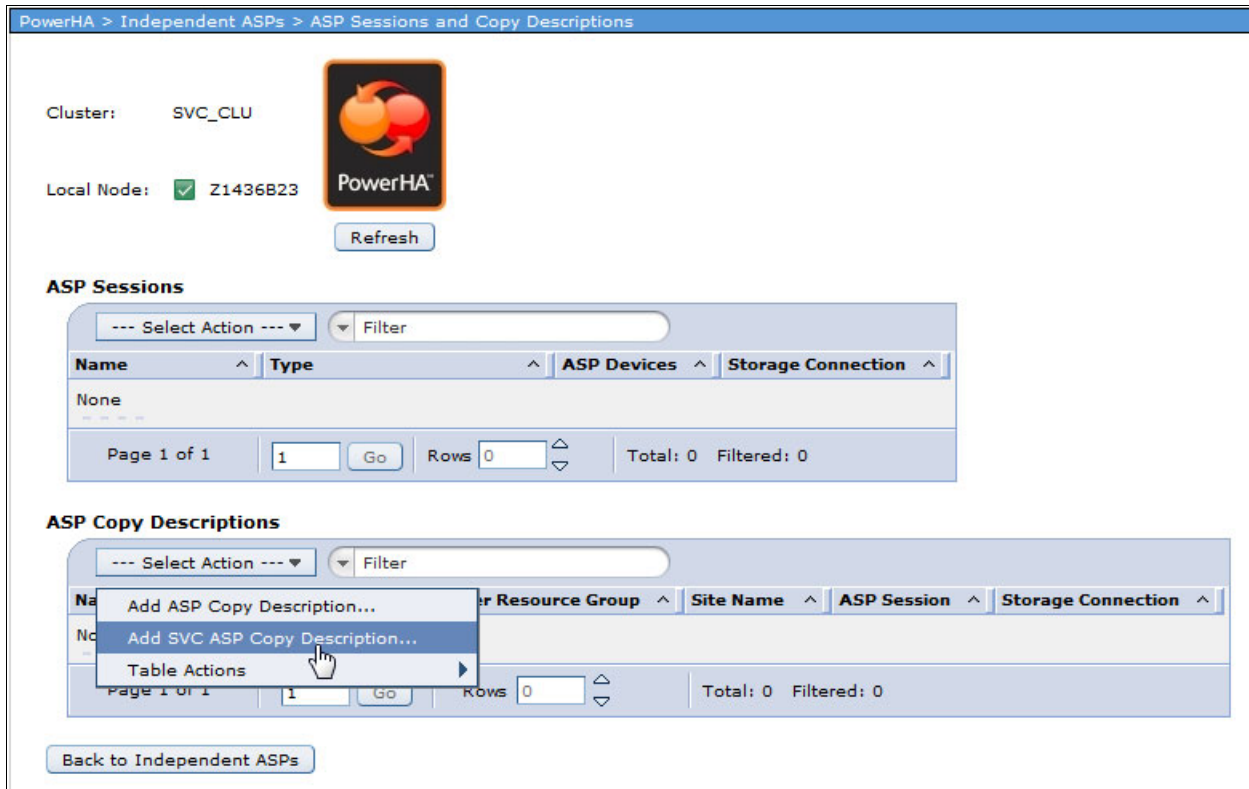


Figure 4-35 Adding an SAN Volume Controller ASP Copy description

32. In the Add SVC ASP Copy Description window, you must define the connection to the SAN Volume Controller and the disk range and host connections from the SAN Volume Controller setup for the LUNs of the IASP. As shown in Figure 4-36, enter the following information:

- Enter a name for the ASP Copy Description.
- The ASP Device, the Cluster Resource Group, and the Site name must be the names that were specified previously in our example.
- User Id is the user ID on the SAN Volume Controller that you imported the SSH key into. This field is case-sensitive. Make sure that you enter the user ID exactly as you specified it in the SAN Volume Controller.
- The Secure Shell Key File points to the SSH private key that is stored in the same directory for all nodes in the cluster
- The IP Address is the network address of the SAN Volume Controller.
- Enter the disk range that is used for the IASP by clicking **Add** to add the information to your configuration setup. This information can be found in the SAN Volume Controller GUI. You can also add several disk ranges here. Make sure that the **Use All Host Identifiers** option is selected.

The screenshot shows the 'Add SVC ASP Copy Description' window with the following fields and values:

- *ASP Copy Description: SVC_Source
- ASP Device: IASP1
- Cluster Resource Group: SVC_CRG
- Site Name: SITE1
- Storage Host:
 - *User Id: PowerHA
 - *Secure Shell Key File: /QIBM/UserData/HASM/hads/.ssh/id_rsa
 - *IP Address: 9.5.216.80
- Virtual Disk Ranges:

Range Start	Range End	Host Identifiers
None	---	---

Add Virtual Disk Range

Range Start: 2 (0-8191) Range End: 2 (0-8191) Use All Host Identifiers

Buttons: Add, Cancel
- Recovery Domain:

Node N	Host Identifiers
None	---

Buttons: Add, Add Recovery Domain Entry

Bottom buttons: OK, Cancel

Figure 4-36 SAN Volume Controller ASP Copy Description creation, adding a virtual disk range

33. You also need to enter the host identifier for all the nodes in the recovery domain. Click **Add Host Identifier** for each node, enter the Host Identifier, and click **Add** to add the information to your configuration, as shown in Figure 4-37. The host identifier can be found in the SAN Volume Controller GUI.

The screenshot displays the configuration interface for creating an ASP Copy description. The 'Storage Host' section includes fields for *User Id (PowerHA), *Secure Shell Key File (/QIBM/UserData/HASM/hads/.ssh/id_rsa), and *IP Address (9.5.216.80). The 'Virtual Disk Ranges' table shows a single entry with Range Start 2, Range End 2, and Host Identifiers All. The 'Recovery Domain' section is currently empty, showing 'None' for the Node Name. The 'Add Recovery Domain Entry' section is active, showing Node: Z1436B23, Host Identifiers: [Empty], and Host Identifier: 3. A mouse cursor is clicking the 'Add Host Identifier' button.

Figure 4-37 SAN Volume Controller ASP Copy description creation, adding a host identifier for the primary node

34. Click **Add** to add the Recovery Domain Entry, as shown in Figure 4-38. Make sure that you add the host identifier information for each node in the cluster.

ASP Device: IASP1

Cluster Resource Group: SVC_CRG

Site Name: SITE1

Storage Host

*User Id: PowerHA

*Secure Shell Key File: /QIBM/UserData/HASM/hads/.ssh/id_rsa

*IP Address: 9.5.216.80

Range Start	Range End	Host Identifiers
2	2	All

Add Virtual Disk Range

Recovery Domain:

Node Name	Host Identifiers
None	

Add Recovery Domain Entry

Node: Z1436B23

Host Identifiers: 3

Remove Host Identifier

Host Identifier:

Add Host Identifier

Reset Field

(0-512)

Add Cancel

OK Cancel

Figure 4-38 SAN Volume Controller ASP Copy description creation, adding the primary node

35. Figure 4-39 shows the finished configuration with the disk range for the IASP and two cluster nodes in the recovery domain and their corresponding host identifiers. Click **OK** to finish the configuration.

Add SVC ASP Copy Description

*ASP Copy Description:

ASP Device:

Cluster Resource Group:

Site Name:

Storage Host

*User Id:

*Secure Shell Key File:

*IP Address:

Virtual Disk Ranges:

Range Start	Range End	Host Identifiers
2	2	All


Recovery Domain:

Node Name	Host Identifiers
Z1436B23	3
Z1436B22	4

Figure 4-39 SAN Volume Controller ASP Copy description creation, recovery domain completed

Figure 4-40 shows the completed setup. Click **Back to Independent ASPs**.

PowerHA > Independent ASPs > ASP Sessions and Copy Descriptions

Cluster: SVC_CLU 

Local Node: Z1436B23

ASP Sessions

--- Select Action --- Filter

Name	Type	ASP Devices	Storage Connection
None			

Page 1 of 1 1 Go Rows 0 Total: 0 Filtered: 0

ASP Copy Descriptions

--- Select Action --- Filter

Name	ASP Device	Cluster Resource Group	Site Name	ASP Session	Storage Connection
SVC_SOURCE	IASP1	SVC_CRG	SITE1		SVC

Page 1 of 1 1 Go Rows 1 Total: 1 Filtered: 1

Figure 4-40 SAN Volume Controller ASP Copy description creation finished

36. Start the cluster resource group. From the main PowerHA menu, click **Cluster Resource Groups**.

37. In the PowerHA Cluster Resource Groups window, right-click the cluster resource name and select **Start**, as shown in Figure 4-41. The SAN Volume Controller with LUN level switching by using PowerHA SystemMirror for i setup is now complete.

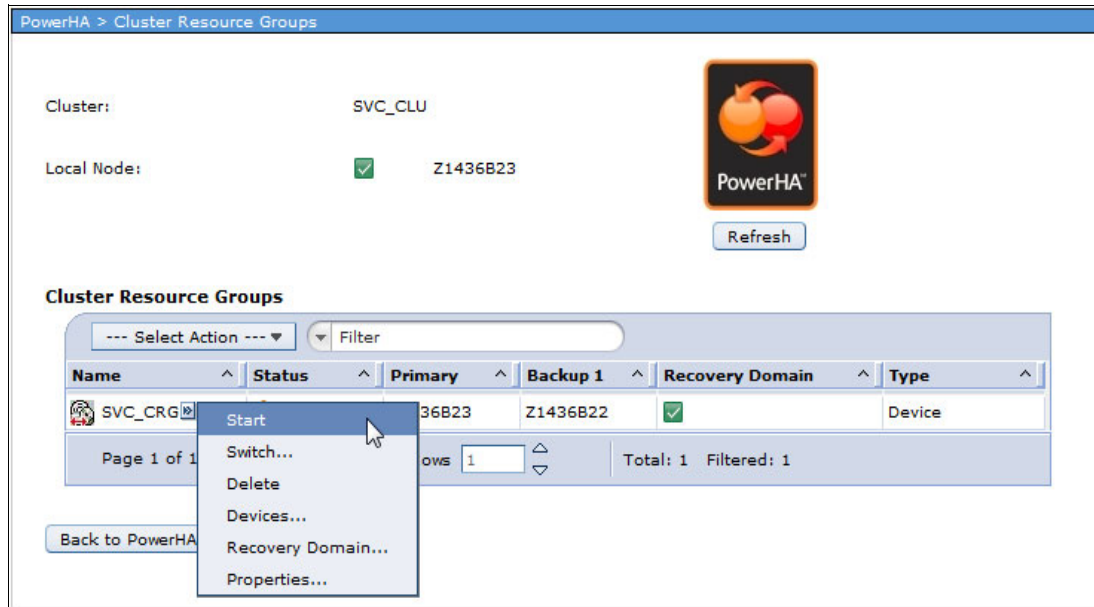


Figure 4-41 Starting the cluster resource group

Using CLI commands to setup SAN Volume Controller split cluster and LUN level switching

You can use the CLI commands that are shown in Example 4-2 to complete this setup.

Example 4-2 CLI commands to set up the SAN Volume Controller split cluster and LUN level switching

On the backup system:

```
CRTDEVASP DEVD(IASP1) RSRNAME(IASP1)
```

On the production system:

```
CRTCLU CLUSTER(SVC_CLU) NODE((Z1436B23 ('9.5.216.233' '192.168.10.23'))
(Z1436B22 ('9.5.216.230' '192.168.10.22')))
```

```
CLUMSGQ(QSYS/QSYSOPR) FLVWAITTIM(*NOMAX) FLVDFTACN(*CANCEL)
```

```
ADDDEVMNE CLUSTER(SVC_CLU) DEVDMN(SVC_DMN) NODE(Z1436B23)
```

```
ADDDEVMNE CLUSTER(SVC_CLU) DEVDMN(SVC_DMN) NODE(Z1435B22)
```

```
CRTCRG CLUSTER(SVC_CLU) CRG(SVC_CRG) CRGTYPE(*DEV)
```

```
EXITPGM(*NONE) USRPRF(*NONE)
```

```
RCYDMN((Z1436B23 *PRIMARY *LAST SITE1 *NONE)
```

```
(Z1436B22 *BACKUP *LAST SITE1 *NONE))
```

```
CFGGOBJ((IASP1 *DEVD *ONLINE '10.10.10.10'))
```

```
ADDSVCCPYD ASPCPY(PROD) ASPDEV(IASP1) CRG(SVC_CRG) SITE(SITE1) NODE(*CRG)
```

```
SVCHOST(PowerHA '/QIBM/UserData/HASM/hads/.ssh/id_rsa' '9.5.216.80')
```

```
VRTDSKRNG((1 4 *ALL)) RCYDMN((Z1436B23 (3)) (Z1436B22 (4)))
```

4.2.3 LUN level switching in combination with Global Mirror

LUN level switching can also be used with Global Mirror in a three-node cluster. In this setup, LUN level switching provides local high availability. In addition, if the whole local site goes down, Global Mirror provides a disaster recovery option to a remote site. Depending on the distance between the two sites, you can also consider the use of Metro Mirror instead of Global Mirror.

For more information about Metro Mirror, Global Mirror, and LUN level switching, see 4.1.1, “Metro Mirror” on page 99, 4.1.2, “Global Mirror” on page 99, and 4.1.3, “LUN level switching” on page 100.

It is assumed that the following steps were completed:

- ▶ The PowerHA setup for the local LUN level switching was completed, as described in 4.2.2, “SAN Volume Controller split cluster and LUN level switching” on page 105.
- ▶ Global Mirror for all volumes of the IASP were configured.
- ▶ The private SSH key that is used for communication was distributed between PowerHA and the SAN Volume Controller to the third cluster node and the public SSH key was assigned to a user on the secondary SAN Volume Controller. For more information, see 4.2.1, “Requirements for PowerHA SystemMirror for i with SAN Volume Controller or Storwize” on page 103.

To complete these steps by using a CLI interface, see “Using CLI commands to setup LUN level switching with Global Mirror in a three-node cluster” on page 149.

Complete the following steps by using the GUI interface:

1. Add the third cluster node into the existing cluster. To do this, select **Cluster Nodes** from the PowerHA main menu in IBM Navigator for i.
2. In the PowerHA Cluster Nodes window, select **Add Cluster Node** from the Select Action pull-down menu, as shown in Figure 4-42 on page 138.

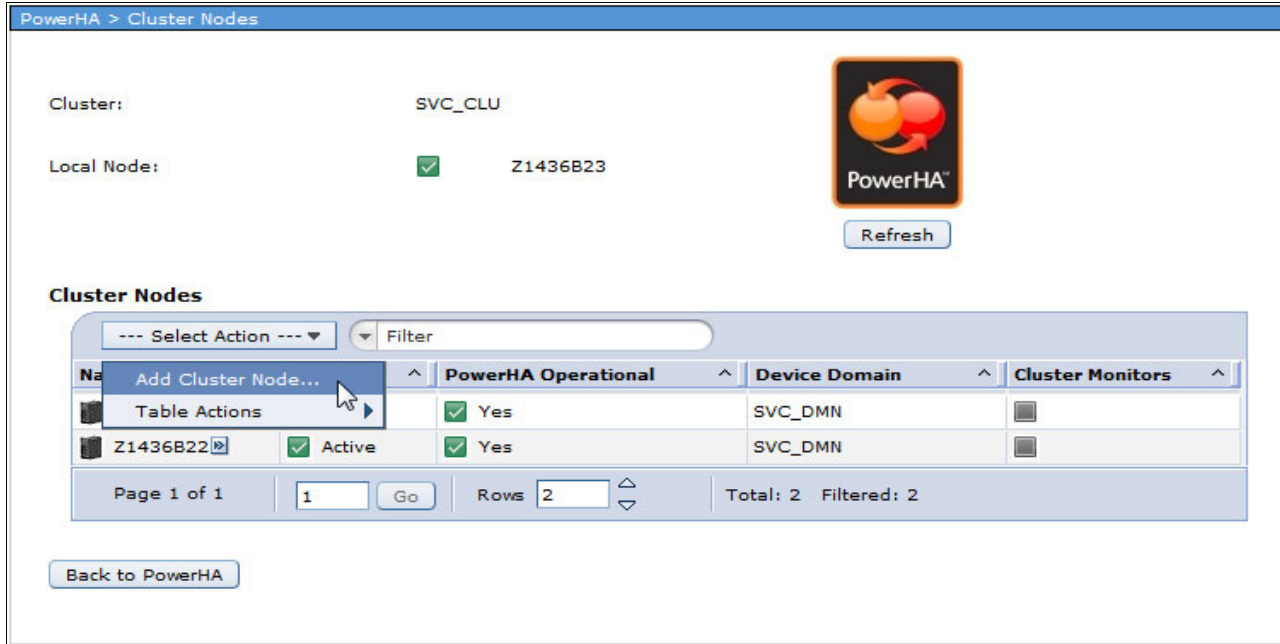


Figure 4-42 Adding a cluster node to create a three-node cluster

3. In the Add Cluster Node window, enter the node name and the cluster IP addresses that are used for cluster communication, as shown in Figure 4-43. Click **OK**.

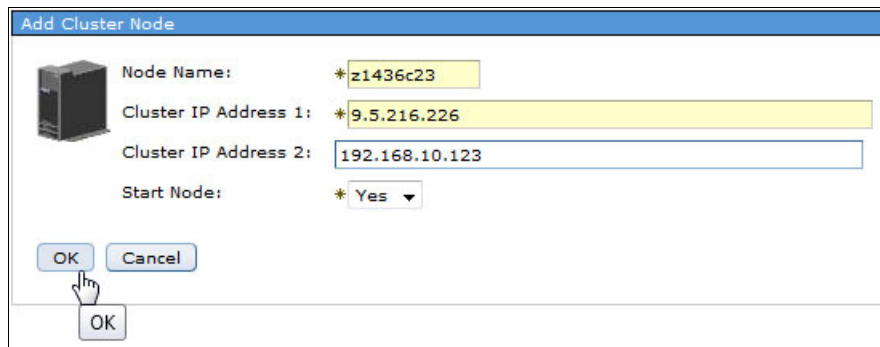


Figure 4-43 Three-node cluster, specifying the third cluster node details

4. As shown in Figure 4-44, the cluster node is added and started. Click **Close**.

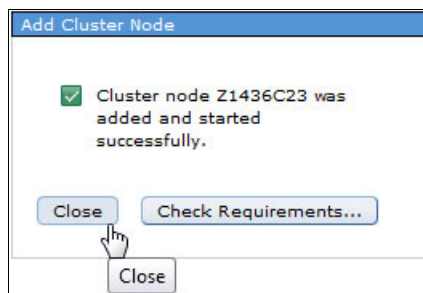


Figure 4-44 Three-node cluster: Third node added successfully

- As shown in Figure 4-45, our example now has a three-node cluster. The third node was not added to the device domain. You can add the node manually to the device domain or PowerHA can handle this for you when the cluster node is added into the cluster resource group. Click **Back to PowerHA** to return to the main PowerHA menu.

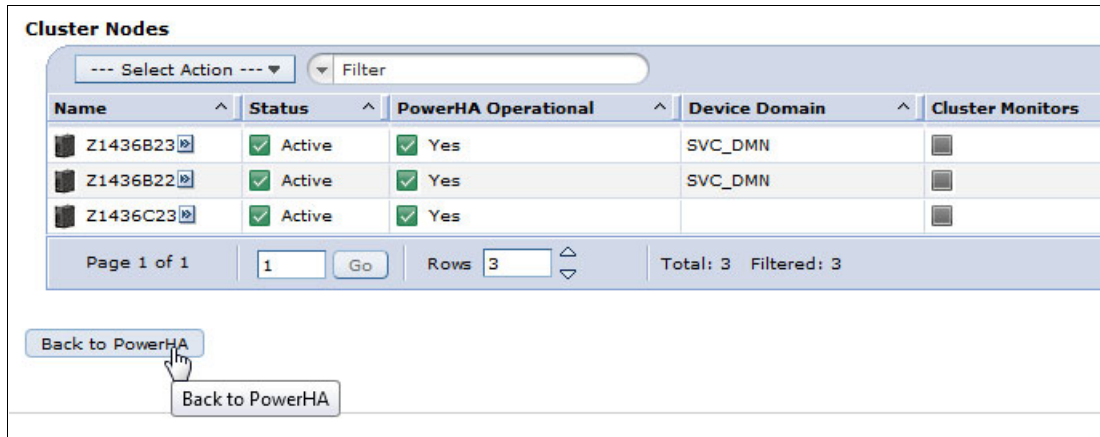


Figure 4-45 Three-node cluster: Node overview

- In the PowerHA main menu, select **Cluster Resource Groups**.
- In the PowerHA Cluster Resource Groups window, select **Recovery Domain** from the Select Action pull-down menu, as shown in Figure 4-46.

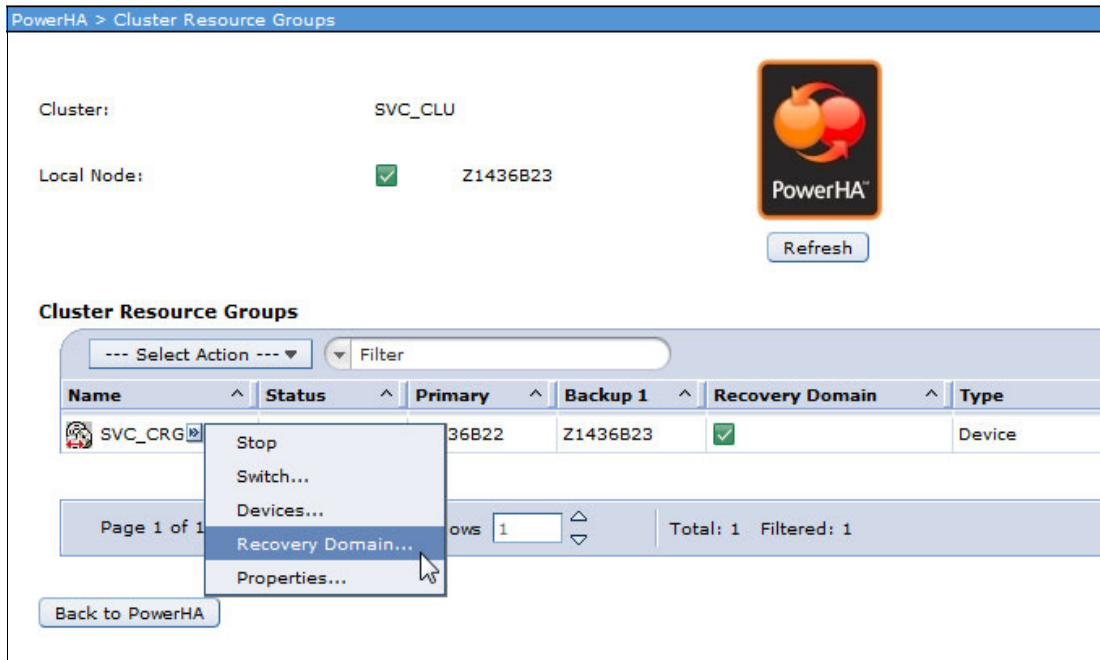


Figure 4-46 Three-node cluster: Creating a recovery domain

8. In the PowerHA Cluster Resource Groups Recovery Domain window, select **Add Recovery Domain Node** from the Select Action pull-down menu, as shown in Figure 4-47.



Figure 4-47 Three-node cluster: Adding a recovery domain node

9. In the Add Recovery Domain Node window (as shown in Figure 4-48), enter the following information:
 - Node name of the third node.
 - The node role is set to Backup with an order number of 2. In this example, there already is a primary node and a first backup node that is defined in the cluster setup.
 - Because this is a different site from the site that was used for LUN level switching, you must specify a different site name here. Think of the site name as a representation for each copy of the IASP in the recovery domain.

Click **OK**.

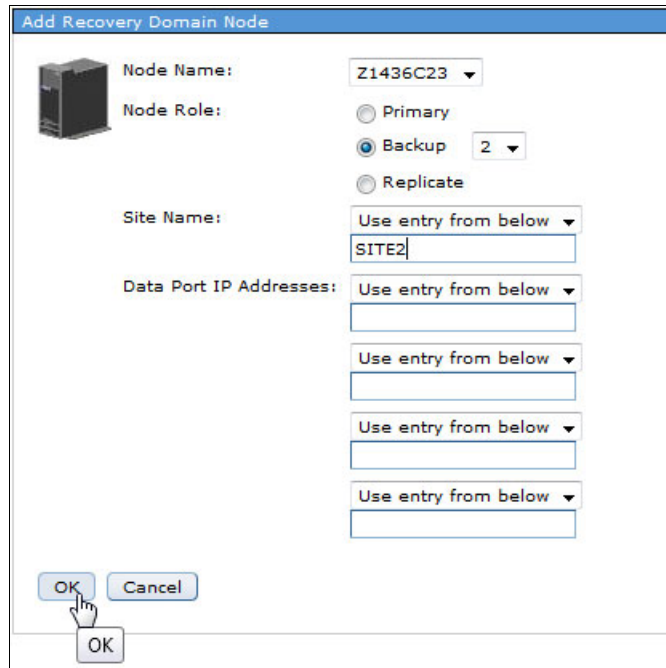


Figure 4-48 Three-node cluster: Specifying the recovery domain details

10. The new node is not yet added into the cluster device domain, as shown in the message that is shown in Figure 4-49. Click **OK** and the new node is automatically added to the same device domain that is used by the other nodes in the recovery domain.

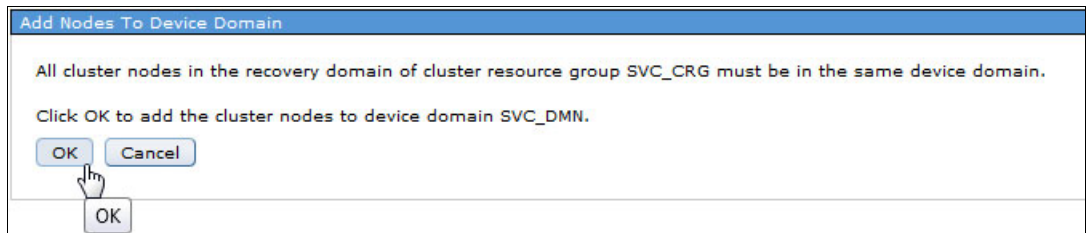


Figure 4-49 Three-node cluster: Warning message

11. A status message appears as the node is added to the device domain, as shown in Figure 4-50.

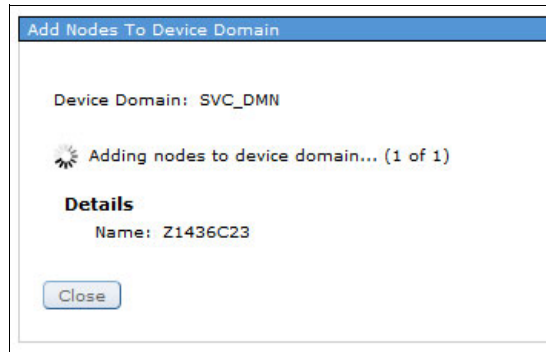


Figure 4-50 Three-node cluster: Status message

12. A status message is then displayed that shows that the third node was added to the recovery domain, as shown in Figure 4-51. Click **Close**.

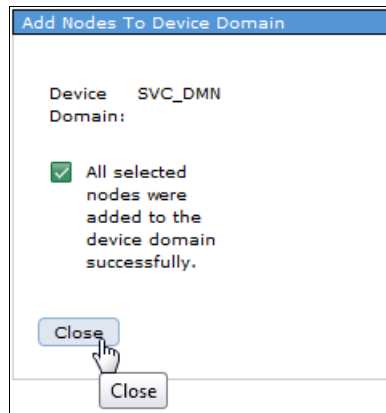


Figure 4-51 Three-node cluster: Node successfully added to recovery domain

13. The recovery domain now consists of three cluster nodes, as shown in Figure 4-52. Click **Back to Cluster Resource Groups** and return to the main PowerHA menu.

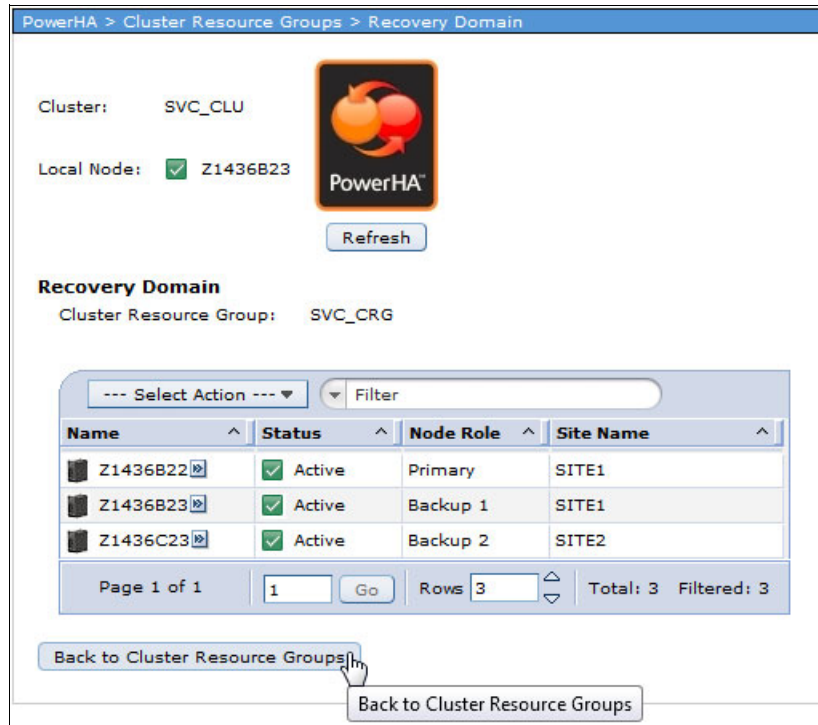


Figure 4-52 Three-node cluster: Recovery domain overview

14. In the main PowerHA menu, select **Independent ASP**.

15. In the PowerHA Independent ASPs window, add an ASP Copy Description for the IASP on the third node. Select **ASP Sessions and Copy Descriptions** from the Select Action pull-down menu, as shown in Figure 4-53.

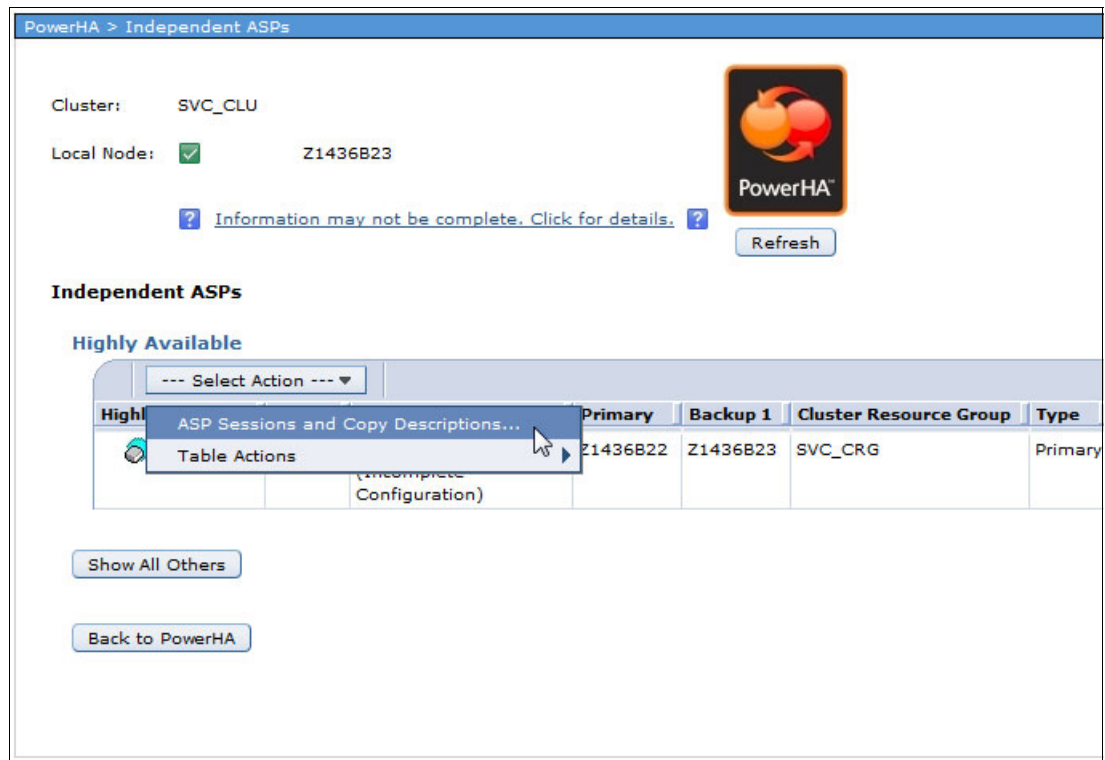


Figure 4-53 Three-node cluster: ASP session

16. In the PowerHA Independent ASPs, ASP Sessions and Copy Descriptions window, select **Add SVC ASP Copy Description** from the Select Action pull-down menu (as shown in Figure 4-54) because this example is working with a SAN Volume Controller/Storwize V7000 environment.

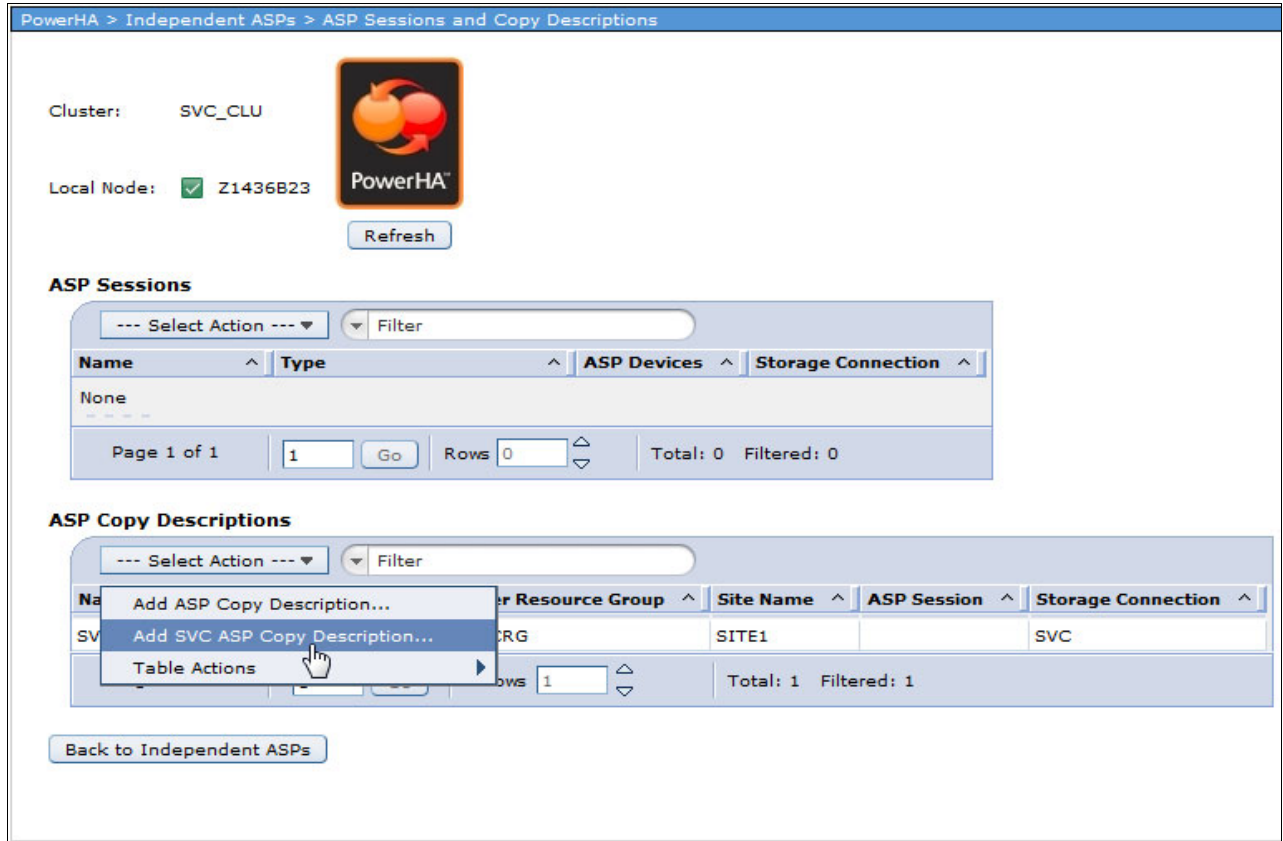


Figure 4-54 Three-node cluster: Adding a SAN Volume Controller copy description

17. In the Add SAN Volume Controller ASP Copy Description window, specify the following parameters for the ASP, as shown in Figure 4-55:

- ASP Copy Description name.
- ASP Device is the name of the IASP that is used.
- Name of the Cluster Resource Group.
- Enter the new Site Name that was created in step 9 on page 141.
- Define how PowerHA can talk to the storage system where the Global Mirror target LUNs are located. The User Id is case-sensitive.

Click **Add Virtual Disk Range**.

The screenshot shows the 'Add SVC ASP Copy Description' dialog box with the following configuration:

- *ASP Copy Description: IASP_GM
- ASP Device: IASP1
- Cluster Resource Group: SVC_CRG
- Site Name: SITE2
- Storage Host
 - *User Id: PowerHA
 - *Secure Shell Key File: /QIBM/UserData/HASM/hads/.ssh/id_rsa
 - *IP Address: 9.5.216.236
- Virtual Disk Ranges: A table with columns 'Range Start', 'Range End', and 'Host Identifiers'. The current entry is 'None'.
- Recovery Domain: A table with columns 'Node Name' and 'Host Identifiers'. The current entry is 'None'.

Buttons at the bottom include 'OK', 'Cancel', 'Add Virtual Disk Range', and 'Add Recovery Domain Entry'.

Figure 4-55 Three-node cluster: Specifying copy description details

18. Enter the numbers for your virtual disk range from the SAN Volume Controller, as shown in Figure 4-56. Click **Add** to add this information into the ASP copy description. You can specify multiple disk ranges here. Click **OK**.

Add SVC ASP Copy Description

*ASP Copy Description:

ASP Device:

Cluster Resource Group:

Site Name:

Storage Host

*User Id:

*Secure Shell Key File:

*IP Address:

Virtual Disk Ranges:

Range Start	Range End	Host Identifiers
None		

Add Virtual Disk Range

Range Start: Range End: Use All Host Identifiers

(0-8191) (0-8191)

Recovery Domain:

Node	Host Identifiers
None	

Figure 4-56 Three-node cluster: Adding the virtual disk range

19. Define an ASP session that tells PowerHA the relationship between ASP copy descriptions. Select **Start SVC ASP Session** from the Select Action pull-down menu in the ASP Sessions section, as shown in Figure 4-57.

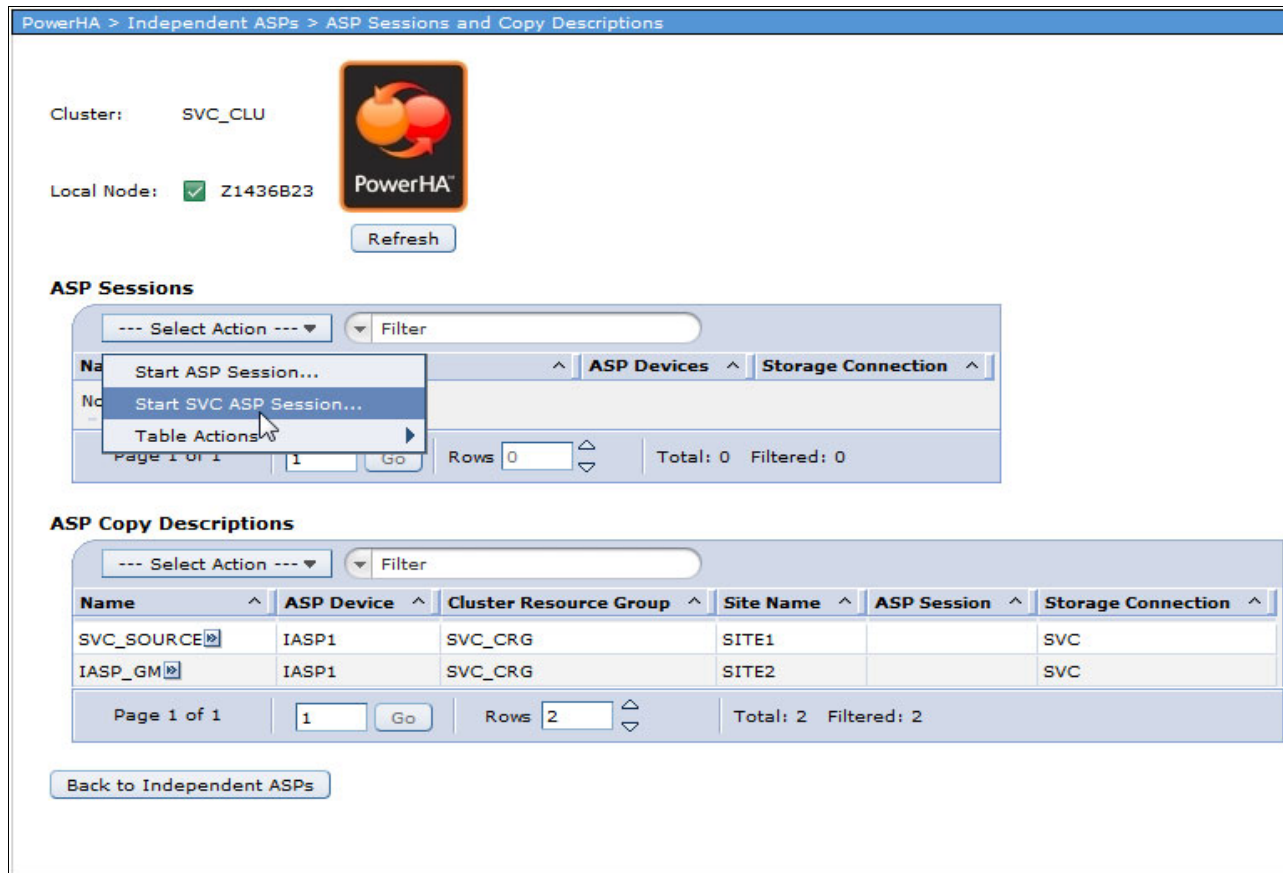


Figure 4-57 Three-node cluster: Defining a SAN Volume Controller ASP session

20. In the Start SVC ASP Session window, specify the following information, as shown in Figure 4-58:

- Type: Defines the relationship between the ASP copy descriptions. This example set up a Global Mirror connection between the primary and the secondary LUNs.
- Specify the Cluster Resource Group that is used.
- You can also decide whether replication between the LUNs is reversed after doing a switchover or a failover. By default, replication is not reversed after a failover so that the original data in the primary IASP is preserved to provide a chance for error analysis.

Click **OK**.

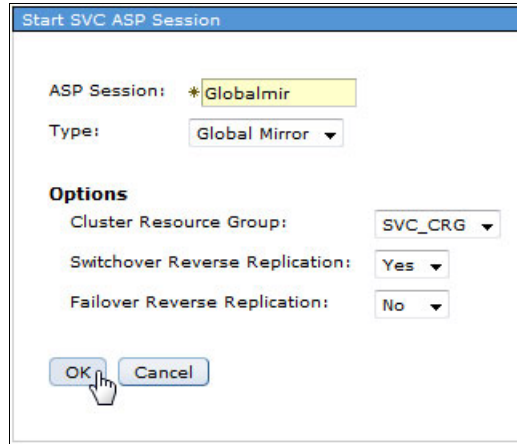


Figure 4-58 Three-node cluster: Starting an SAN Volume Controller ASP session

The three-node cluster that uses LUN level switching for local high availability and Global Mirror for remote disaster recovery is now ready for use.

Using CLI commands to setup LUN level switching with Global Mirror in a three-node cluster

If you prefer to set up this environment by using CL commands, use the commands that are shown in Example 4-3.

Example 4-3 Setup for adding Global Mirror by using CL commands

On the third cluster node:

```
CRTDEVASP DEVD(IASP1) RSRNAME(IASP1)
```

On the primary cluster node:

```
ADDCLUNODE CLUSTER(SVC_CLU) NODE(Z1436C23 ('9.5.216.226' '192.168.10.123'))
```

```
ADDDEVDMNE CLUSTER(SVC_CLU) DEVDMN(SVC_DMN) NODE(Z1436C23)
```

```
ADDCRGNODE CLUSTER(SVC_CLU) CRG(SVC_CRG)
RCYDMN(Z1436C23 *BACKUP *LAST SITE2 *NONE)
```

```
ADDSVCCPYD ASPCPY(IASP_GM) ASPDEV(IASP1) CRG(SVC_CRG) SITE(SITE2) NODE(*CRG)
SVCHOST(powerha '/QIBM/UserData/HASM/hads/.ssh/id_rsa' '9.5.216.236')
VRTDSKRNG((1 4 *ALL))
```

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

4.2.4 Metro Mirror setup

Setting up PowerHA SystemMirror for i in a Storwize family environment by using Metro Mirror is comparable to the setup that was described in 4.2.3, “LUN level switching in combination with Global Mirror” on page 137.

The main differences are the configuration on the Storwize system where you specify to use Metro Mirror instead of Global Mirror. You must setup a cluster with two nodes in it. You then create the cluster resource group. Make sure to specify different site names for the node identifiers in the recovery domain. Create two SAN Volume Controller Copy Descriptions that point to the LUNs that are used as Metro Mirror source volumes and Metro Mirror target volumes. When you are starting the SAN Volume Controller session, make sure to specify *METROMIR as the session type.

For more information about setup instructions for this environment, see *PowerHA SystemMirror for IBM i Cookbook*, SG24-7994.

4.2.5 Adding LUNs to an existing IASP

Complete the following steps to add LUNs to an existing IASP that is used with PowerHA:

1. Define the new LUNs on the storage system and map them to the correct host.
2. If you are using Metro Mirror or Global Mirror, define remote copy relationships between the LUNs on your storage system and start the copy relationship. When these copy relationships reach a consistent status, you add the copy relationships into the existing consistency group.
3. If you are using FlashCopy, make sure to add the new LUNs into the existing consistency group that you are using.
4. Add the LUNs into the IASP on IBM i from SST.

5. Add the ranges of the newly created LUN ranges to the corresponding ASP copy descriptions. You can use the Change SAN Volume Controller Copy Description (CHGSVCCPYD) command or use the Navigator GUI by completing the following steps:
 - a. From **Independent ASP** link in the PowerHA GUI, select **ASP Sessions and Copy Description** from the Select Action pull-down menu, as shown in Figure 4-59.

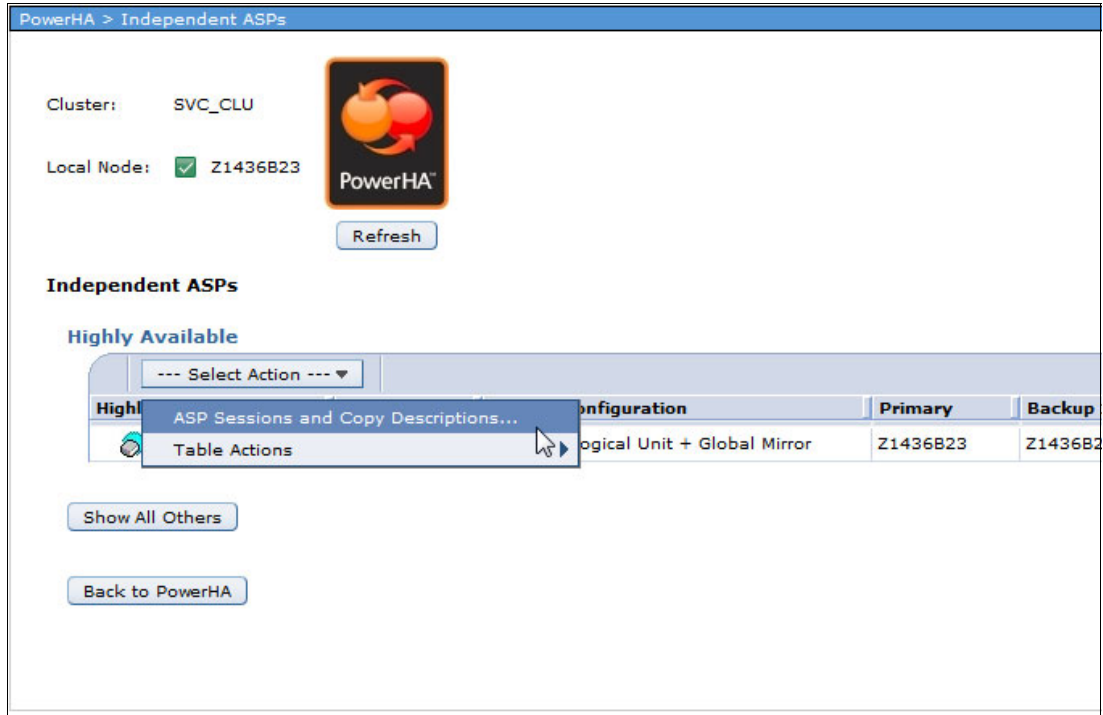


Figure 4-59 Adding LUNs to ASP session and copy description

- b. In the PowerHA Independent ASP > ASP Sessions and Copy Descriptions window, click the **Properties** for the IASP to which you added the LUNs, as shown in Figure 4-60.



Figure 4-60 Adding LUNs, selecting properties of the IASP

- c. In the Properties window, click **Edit** to change the ASP Copy Description, as shown in Figure 4-61.

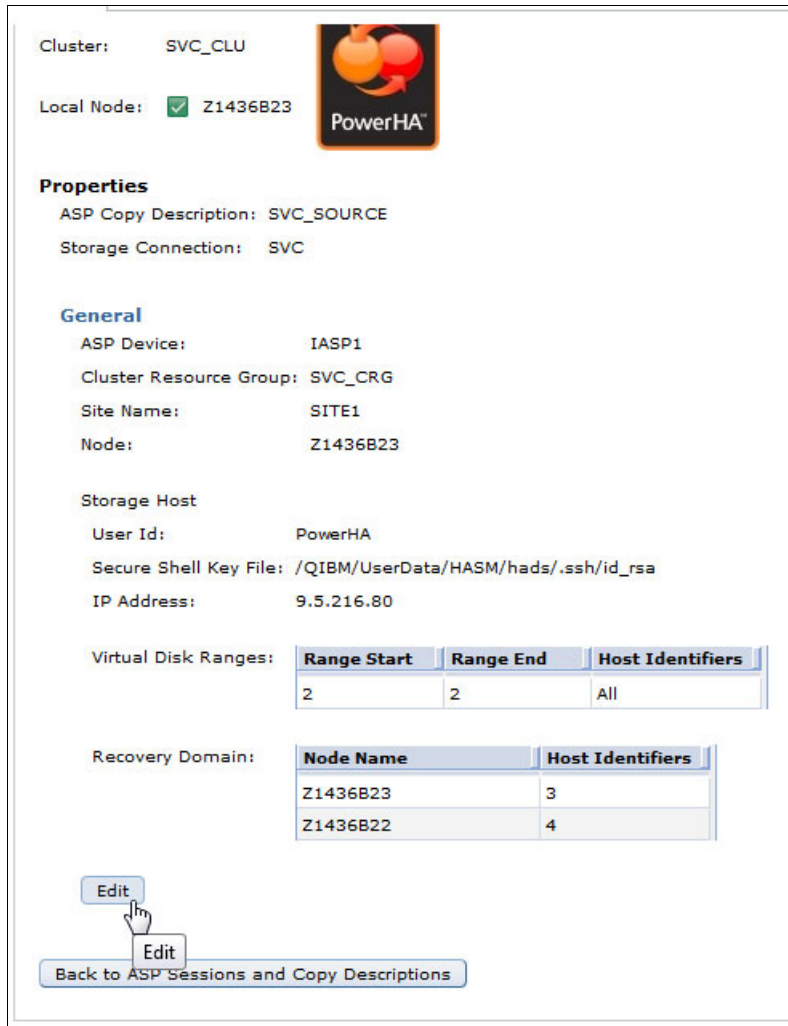


Figure 4-61 Adding LUNs, editing properties

- d. In the Properties window, you can see the virtual disk range that is assigned to the IASP. Click **Add Virtual Disk Range**, as shown in Figure 4-62.

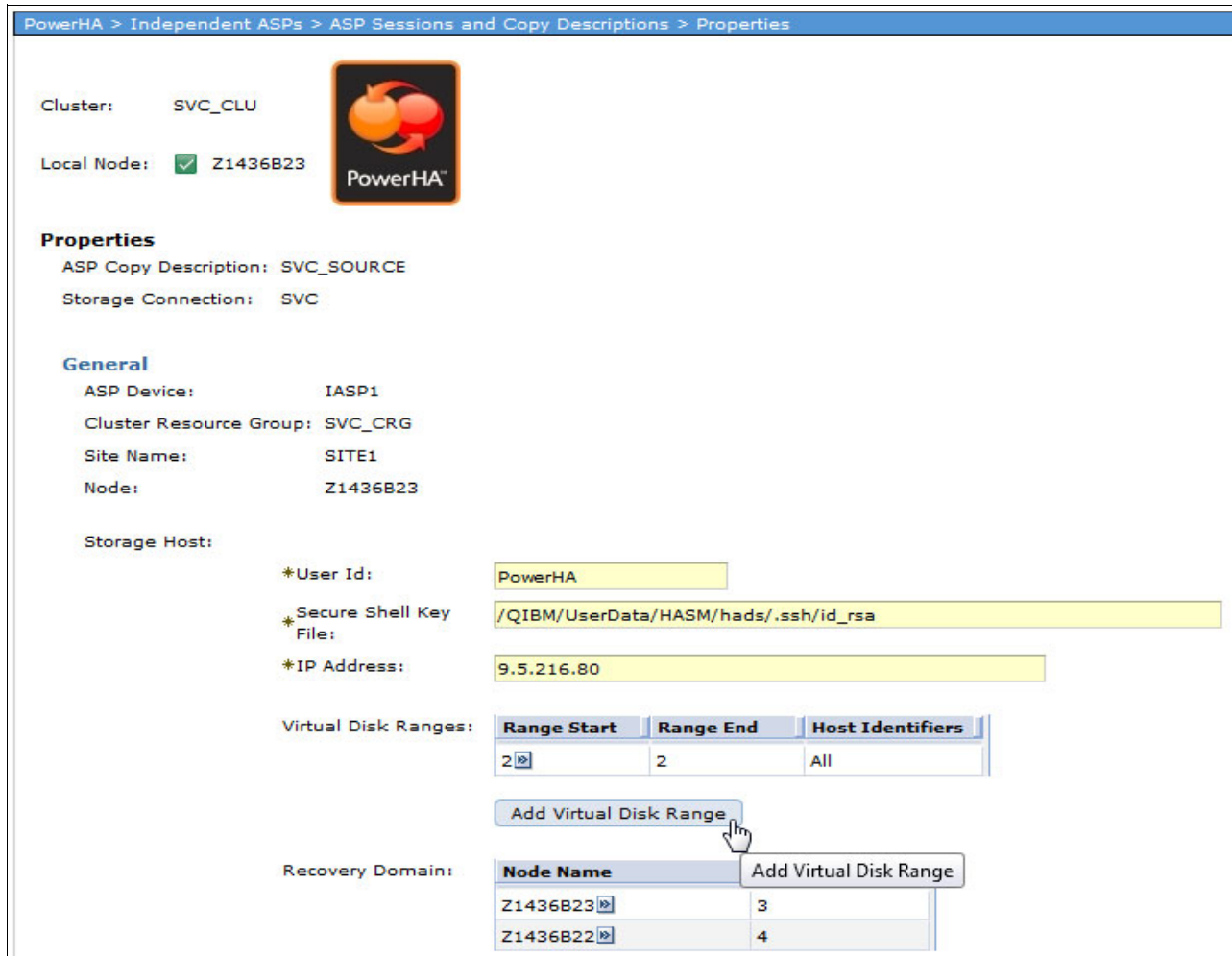


Figure 4-62 Adding LUNs, adding virtual disk range

e. Enter the disk range of the newly added LUNs and click **Add**, as shown in Figure 4-63.

ASP Device: IASP1
Cluster Resource Group: SVC_CRG
Site Name: SITE1
Node: Z1436B23

Storage Host:

*User Id: PowerHA
*Secure Shell Key File: /QIBM/UserData/HASM/hads/.ssh/id_rsa
*IP Address: 9.5.216.80

Virtual Disk Ranges:

Range Start	Range End	Host Identifiers
2	2	All

Add Virtual Disk Range

Range Start: 9 (0-8191) Range End: 11 (0-8191) Use All Host Identifiers

Add Cancel

Recovery Domain:

Node	Add	Host Identifiers
Z1436B23		3
Z1436B22		4

Add Recovery Domain Entry

Save Cancel

Figure 4-63 Adding LUNs, entering a new virtual disk range

- f. When all of the new LUN ranges are entered, click **Save** (as shown in Figure 4-64) to save the new configuration for the ASP copy description.

ASP Copy Description: SVC_SOURCE
 Storage Connection: SVC

General

ASP Device: IASP1
 Cluster Resource Group: SVC_CRG
 Site Name: SITE1
 Node: Z1436B23

Storage Host:

*User Id: PowerHA
 *Secure Shell Key File: /QIBM/UserData/HASM/hads/.ssh/id_rsa
 *IP Address: 9.5.216.80

Virtual Disk Ranges:

Range Start	Range End	Host Identifiers
2	2	All
9	11	All

Add Virtual Disk Range

Recovery Domain:

Node Name	Host Identifiers
Z1436B23	3
Z1436B22	4

Add Recovery Domain Entry

Save Cancel

Save

Back to ASP Sessions and Copy Descriptions

Figure 4-64 Adding LUNs, saving the new configuration



Live Partition Mobility

By using Live Partition Mobility, you can migrate partitions and their hosted applications from one physical server to another without disrupting the infrastructure services. The migration operation, which takes a few seconds, maintains complete system transactional integrity. The migration transfers the entire system environment, including processor state, memory, attached virtual devices, and connected users.

This chapter describes the general steps that are part of a Live Partition Mobility operation. It gives an overview of the prerequisites to use Live Partition Mobility and guides you through the setup process.

The following topics are covered in this chapter:

- ▶ Usage scenarios for Live Partition Mobility
- ▶ License considerations with Live Partition Mobility
- ▶ Setting up Live Partition Mobility

5.1 Usage scenarios for Live Partition Mobility

IBM Power System servers are designed to offer the highest stand-alone availability in the industry. Still, enterprises must occasionally restructure their infrastructure to meet new IT or business requirements. By letting you move your running production applications from one physical server to another, Live Partition Mobility allows for non-disruptive maintenance or modification to a system. This mitigates the effect on partitions and applications that were caused by the occasional need to shut down a system.

Because Live Partition Mobility works only when the current production system is up and running, it is not a replacement for another high availability solution. Instead, it is another feature that you want to use with your high availability solution.

Live Partition Mobility can help in reducing planned downtimes for hardware maintenance or replacement of systems and in moving logical partitions to another system to achieve a more balanced workload distribution. As the complete logical partition is moved to a new system and nothing remains on the original system, Live Partition Mobility cannot help in scenarios where you need access to one environment when you are performing productive work on another environment. This includes operating system maintenance or application maintenance.

The following steps provide an overview of the process that is involved in using Live Partition Mobility to move a running LPAR to a new system:

1. As a starting point in Figure 5-1, the IBM i client partition is running on System #1. On System #2, there is a Virtual I/O Server that is running that is connected to the storage system that System #1 is using. As shown in Figure 5-1, there is no configuration that is available on System #2 for the IBM i LPAR that we want to migrate to that system.

The environment is checked for required resources, such as main memory or processor resources.

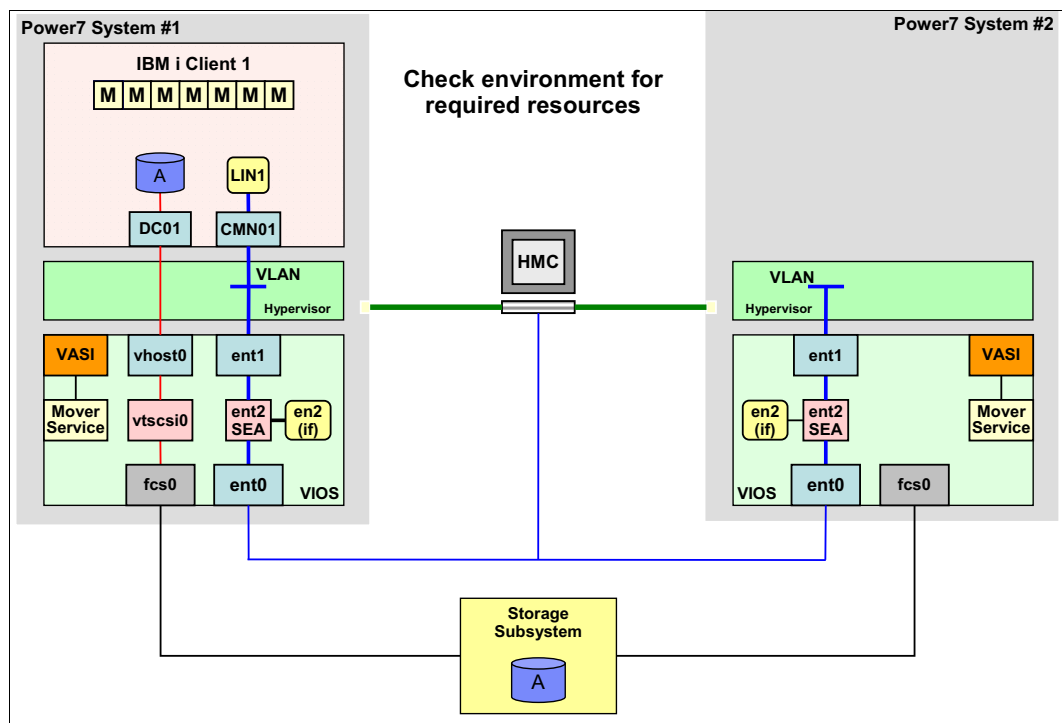


Figure 5-1 Live Partition Mobility, resource check

- If this check is successful, a shell LPAR is created on System #2, as shown in Figure 5-2.

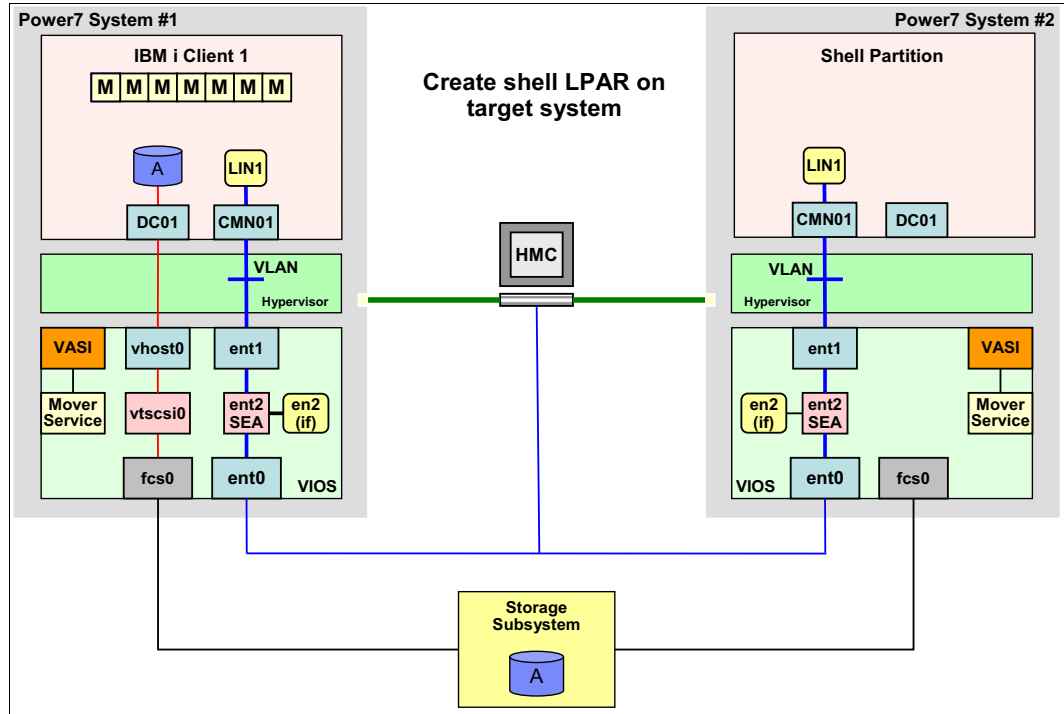


Figure 5-2 Live Partition Mobility, shell LPAR created

- Virtual SCSI or NPIV devices are created in the virtual target VIOS server and in the shell partition to give the partition access to the storage subsystem that is used by System #1, as shown in Figure 5-3.

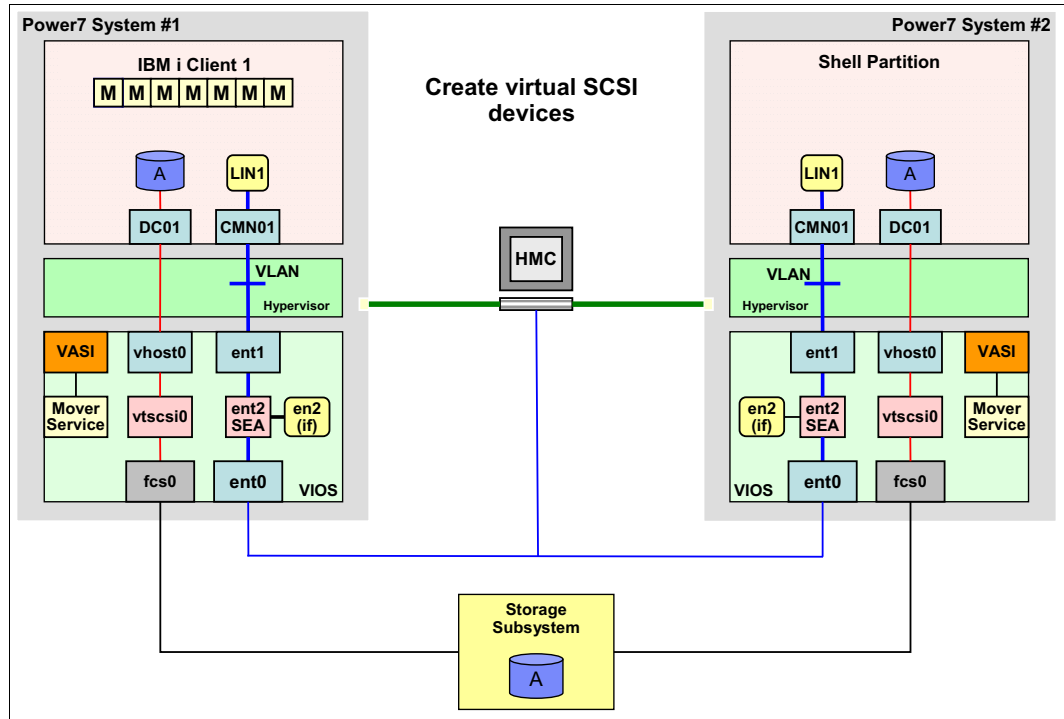


Figure 5-3 Live Partition Mobility, virtual SCSI devices created

- Main memory is migrated from System #1 to System #2, as shown in Figure 5-4. Users can still work on System #1 when this process is occurring. Whenever a page from main memory is changed on System #1 after it was moved to System #2, that page must be moved from System #1 to System #2 again. Therefore, it is considered a best practice to start a partition migration during times of low activity on the system.

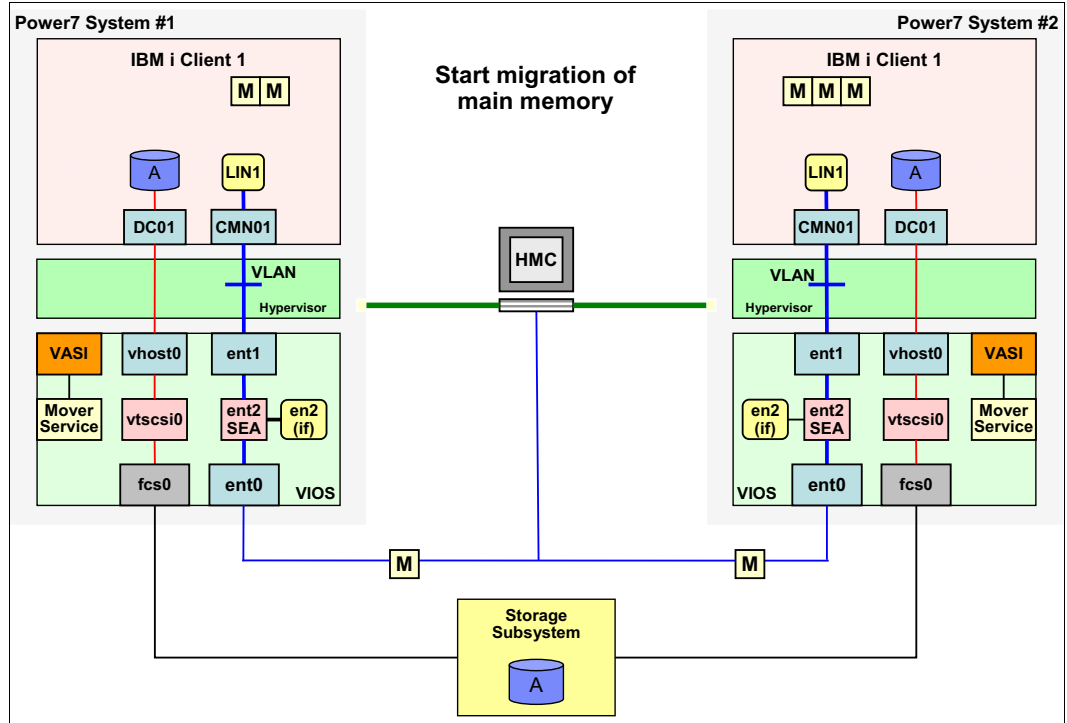


Figure 5-4 Live Partition Mobility, memory migration

- When the system reaches a state where only a small number of memory pages were not migrated, System #1 is suspended, as shown in Figure 5-5. From this point, no work is possible on System #1. Users notice that their transactions seem to be frozen. The remaining memory pages are then moved from System #1 to System #2.

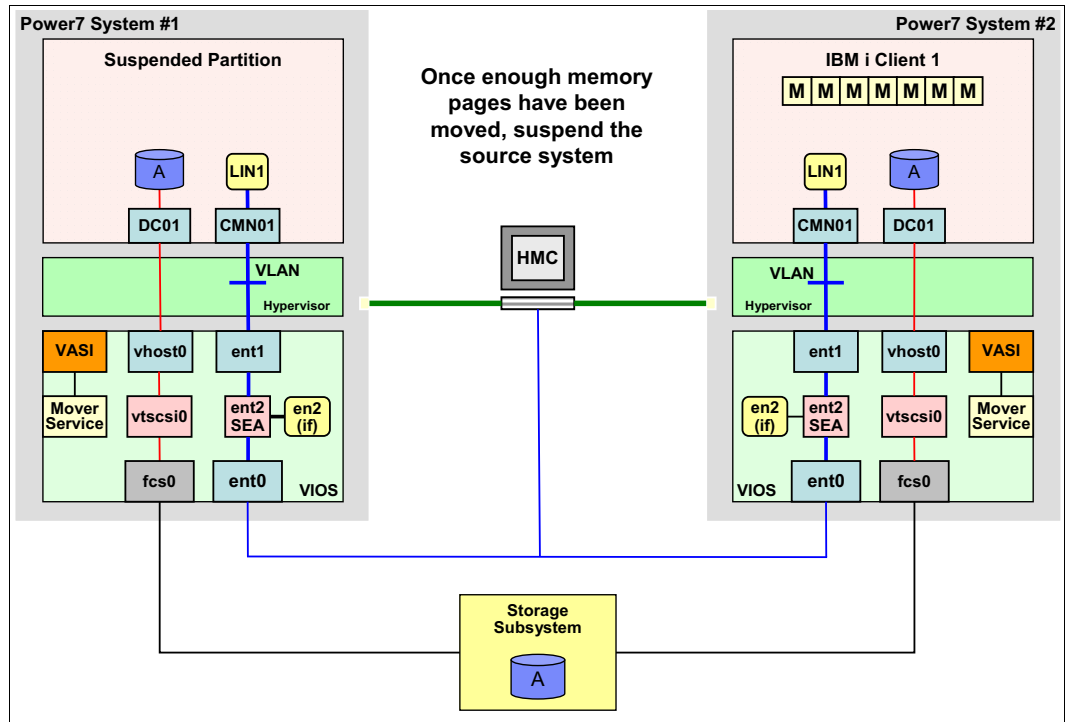


Figure 5-5 Live Partition Mobility, suspend of source system

6. After the memory pages migration is finished, the original LPAR definition on System #1 is deleted with the virtual I/O definitions in the Virtual I/O Server, as shown in Figure 5-6. The target system resumes and user transactions start to run again.

All IBM i jobs are moved over to the new systems. Users experience a delay in response time when the system is suspended. However, because jobs are still active on the system after the migration finishes, users can continue to work without having to log on to the system again.

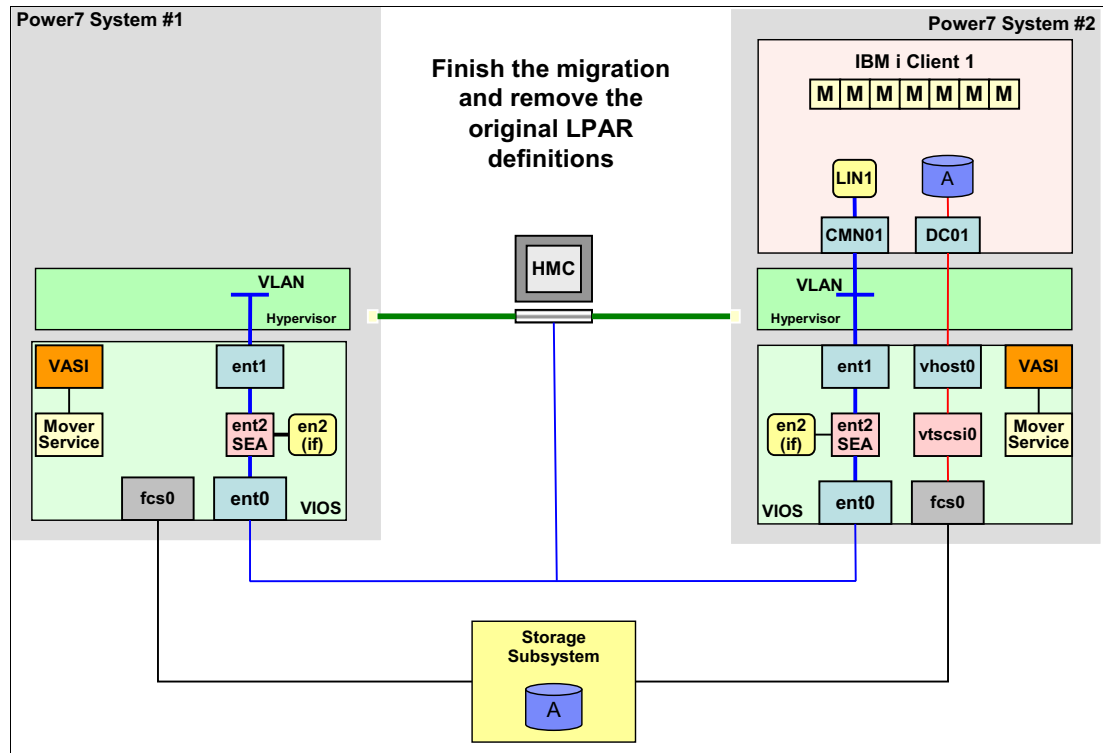


Figure 5-6 Live Partition Mobility, finished status

5.2 License considerations with Live Partition Mobility

When partitions are moved from one system to another, you must ensure that you still adhere to license requirements in your environment. For IBM i licenses, the following rules apply:

- ▶ When an LPAR is moved to another system permanently, considerations when using Live Partition Mobility are the same as when a manual move to another system is performed. IBM i entitlements are licensed to the hardware serial number. Therefore, you need a new license on the system the LPAR is migrated to or you can transfer entitlements from the old to the new system under certain circumstances. Licenses for IBM i licensed programs can be moved to the new system.
- ▶ When an LPAR is moved to another system temporarily, the following scenarios are possible:
 - All necessary licenses are available on both systems. No special considerations need to be taken.
 - The secondary system is a CBU edition. In this case, normal CBU regulations apply.

- ▶ No licenses are available on the secondary system. In this case, the 70-day grace period applies. You can use the system for 70 consecutive days. When you are performing multiple Live Partition Mobility operations, this 70-day grace period restarts after each migration. To work in this way, licenses must be available on the source system. Source and target systems must belong to the same company or must be leased by the same company. In addition, the source system must be in a higher or equal processor group as the target system. To receive support when working on the target system, it is recommended to have at least one IBM i license with a valid software maintenance contract on that system.

5.3 Setting up Live Partition Mobility

The use of Live Partition Mobility requires several preparation steps that are described in this section. In addition, there are several error messages and warnings that might appear during a Live Partition Migration operation. The most common messages and warnings and their causes are explained in this section.

5.3.1 Requirements

The following minimum prerequisites for the use of Live Partition Mobility in an IBM i environment must be met on the source system and target system:

- ▶ POWER7® server with firmware level 740_40 or 730_51, or higher
- ▶ HMC V7R7.5 or later
- ▶ VIOS 2.2.1.4 (FP25-SP2) or later
- ▶ IBM i 7.1 Technology Refresh 4 or later
- ▶ PowerVM Enterprise Edition
- ▶ All VIOS virtual adapters assigned as required
- ▶ If NPIV is used, both logical WWPNs of a virtual FC server adapter must be zoned because the second WWPN is used on destination server.
- ▶ VLAN ID for IBM i client to be migrated must be identical on source and target system.

5.3.2 Preparation

Live Partition Mobility is enabled when PowerVM Enterprise Edition is installed on the server, as shown in the HMC server properties window in Figure 5-7. You can also see that the system is enabled for IBM i partition mobility.

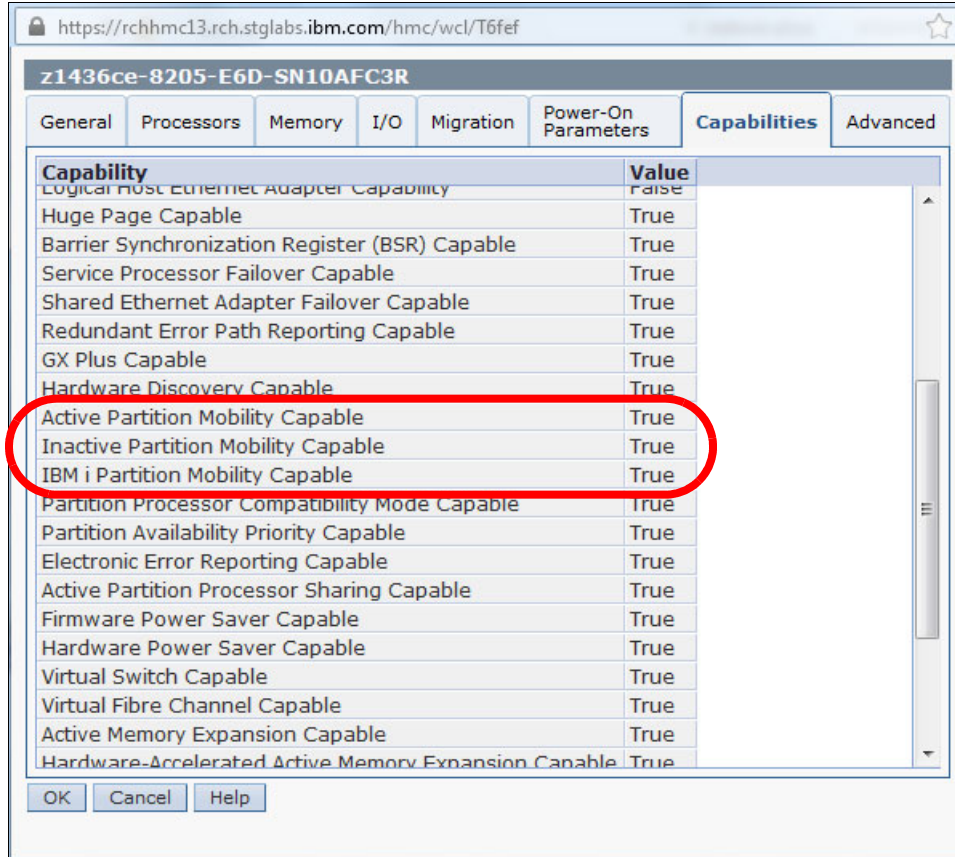


Figure 5-7 Live Partition Mobility, IBM i partition mobility capable

All I/O for the LPAR that you want to move to another system must be virtual and all disk storage for that LPAR must be attached externally by using VSCSI or NPIV. Access to that storage must be shared between VIO servers on the source and the target system.

No physical adapters can be assigned to the IBM i LPAR. The IBM i LPAR must be set to Restricted I/O, which is done in the Properties section of the LPAR, as shown in Figure 5-8. Setting this parameter requires an IPL of the LPAR. It prevents the addition of physical adapters to the LPAR.

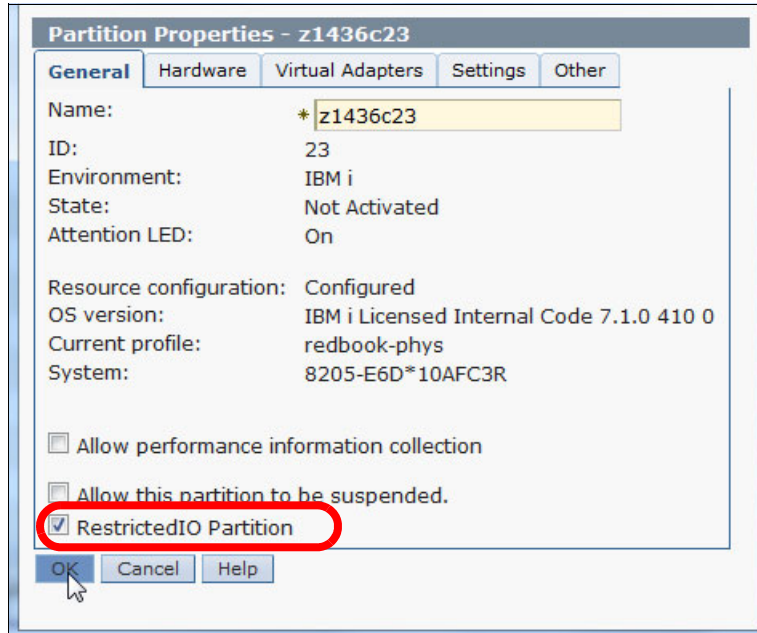


Figure 5-8 Live Partition Mobility, restricted I/O partition

You must specify which VIO Server on the source and on the target LPAR is designated as the mover service partition (MSP). This enables the specified Virtual I/O Server partition to asynchronously extract, transport, and install partition state information. Mover service partitions are not used for inactive migrations. If the MSP flag is not set for any VIO server on the system, the error message that is shown in Figure 5-9 opens during the Live Partition Mobility validation.

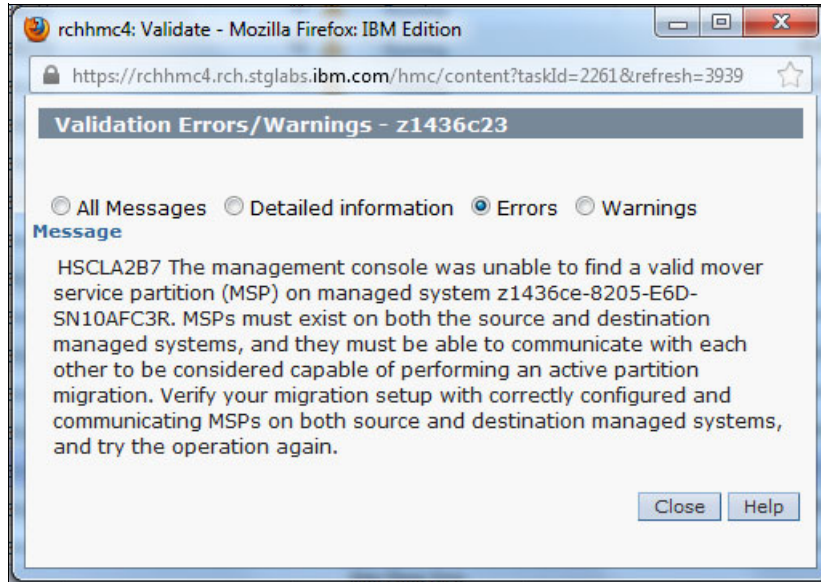


Figure 5-9 Live Partition Mobility, no mover service partition specified

The attribute for the mover service partition is set in the properties of the VIO server LPAR, as shown in Figure 5-10. The attribute becomes active immediately and does not require an IPL of the LPAR.

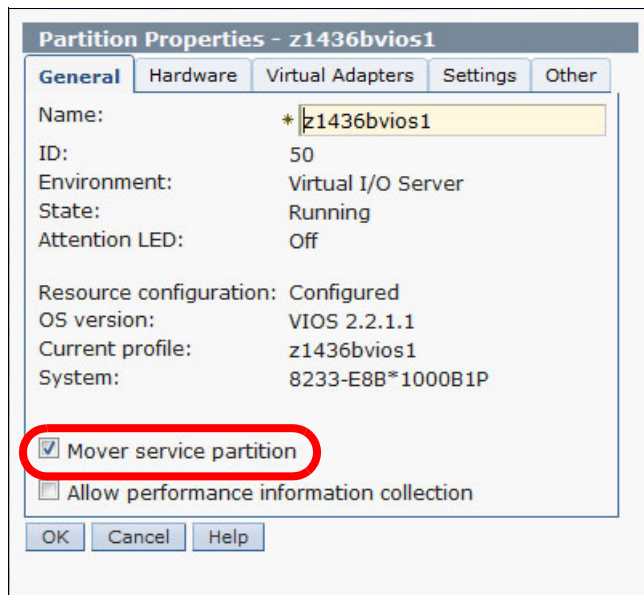


Figure 5-10 Live Partition Mobility, mover service partition flag

All VIO servers that are taking part in the Live Partition Mobility activity must have a working RMC connection. If VIO servers are cloned by using copy operations, this is not the case because their `ct_node_id` is not unique. In this case, the error message that is shown in Figure 5-11 is displayed. For more information about the steps that are required to set the `ct_node_id`, see 3.7.2, “Restarting the RMC connection” on page 95.

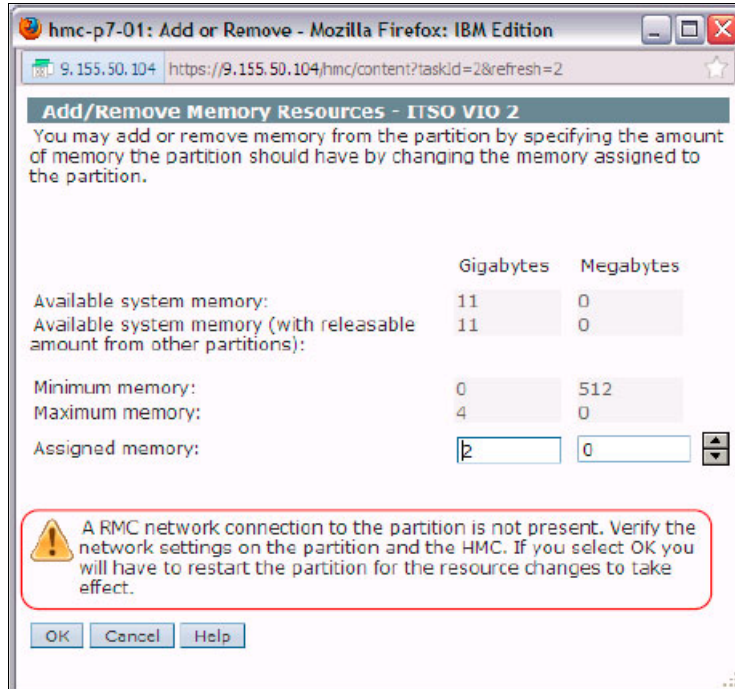


Figure 5-11 Live Partition Mobility, VIO server without RMC connection

Ensure that the memory region sizes of the source and the target system are identical. If this setting is different on the systems that are involved, the validation process displays the error message that is shown in Figure 5-12. If you must change this setting, open the ASMI interface for the system that you want to change. The value can be found by clicking **Performance Setup** → **Logical Memory Block Size**. Changing the memory region size requires a complete shutdown of the server.

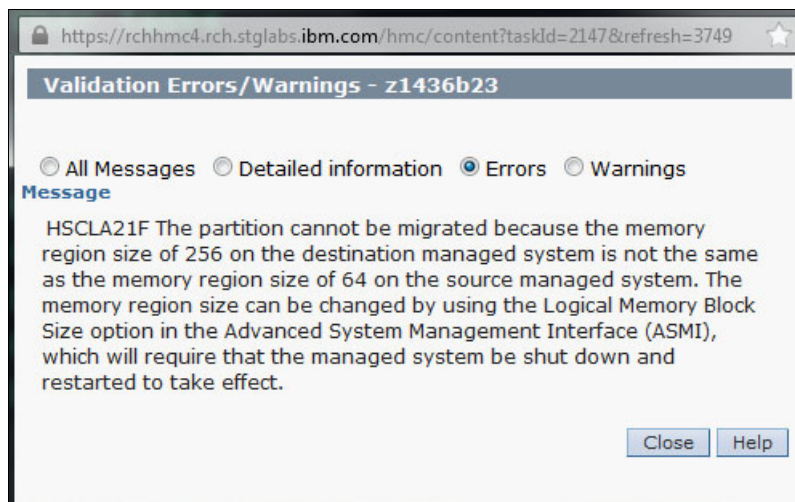


Figure 5-12 Live Partition Mobility, error for different memory region sizes

5.3.3 Live Partition Mobility migration

A Live Partition Mobility migration consists of the following general steps (the validation and the actual migration):

1. A Live Partition Mobility validation is started from the HMC GUI, as shown in Figure 5-13.

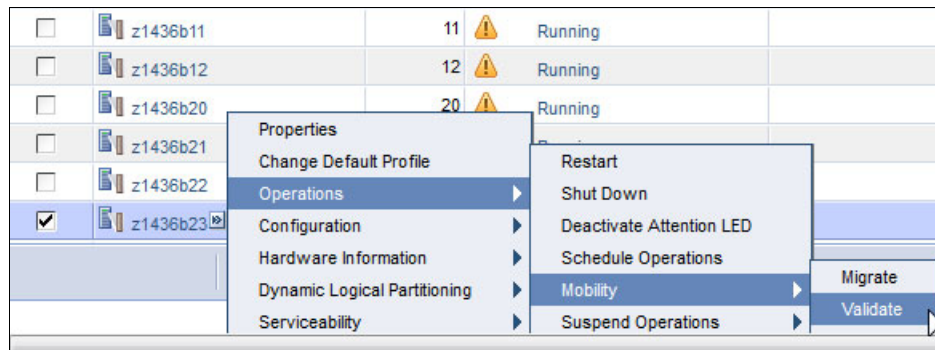


Figure 5-13 Live Partition Mobility, starting validation

2. Select the destination system to which you want to move the LPAR, as shown in Figure 5-14. Click **Validate**.

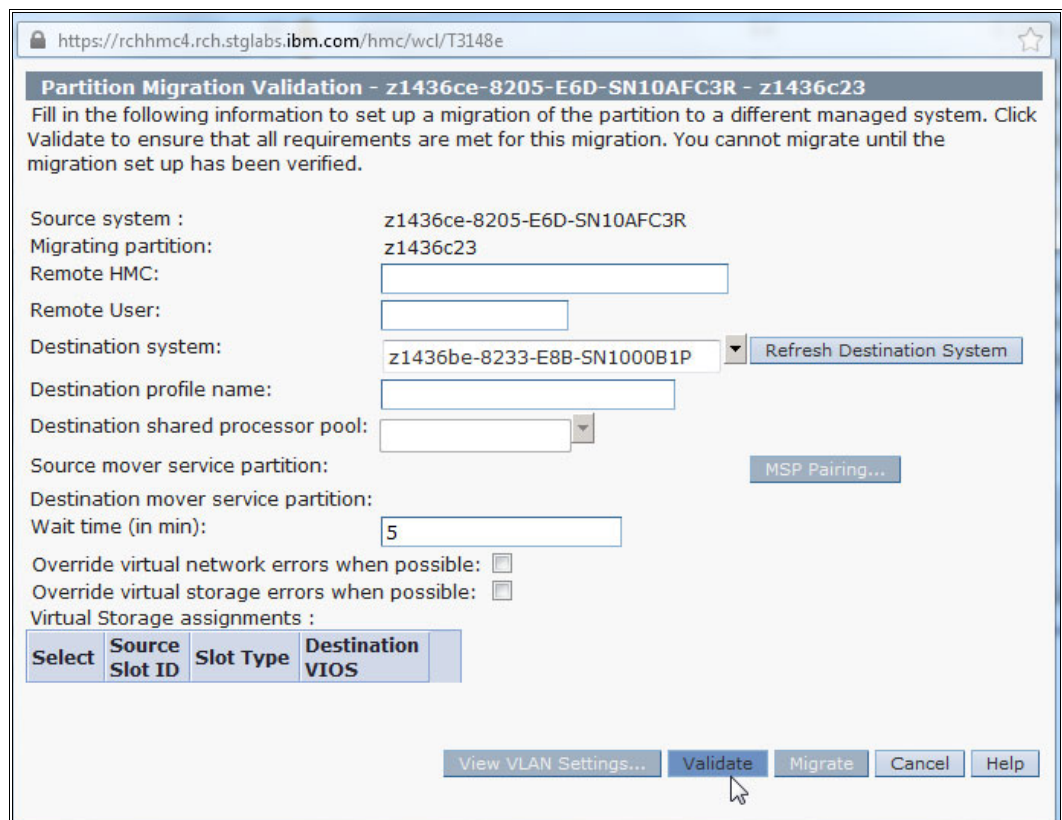


Figure 5-14 Live Partition Mobility, selecting the destination system

- After the validation finishes without errors, you can start the migration by clicking **Migrate** as shown in Figure 5-15. If there are multiple VIO servers available on the target system to provide I/O for the migrated LPAR, the Live Partition Mobility process provides a suggestion for the virtual storage assignment. This predefined setting can also be changed.

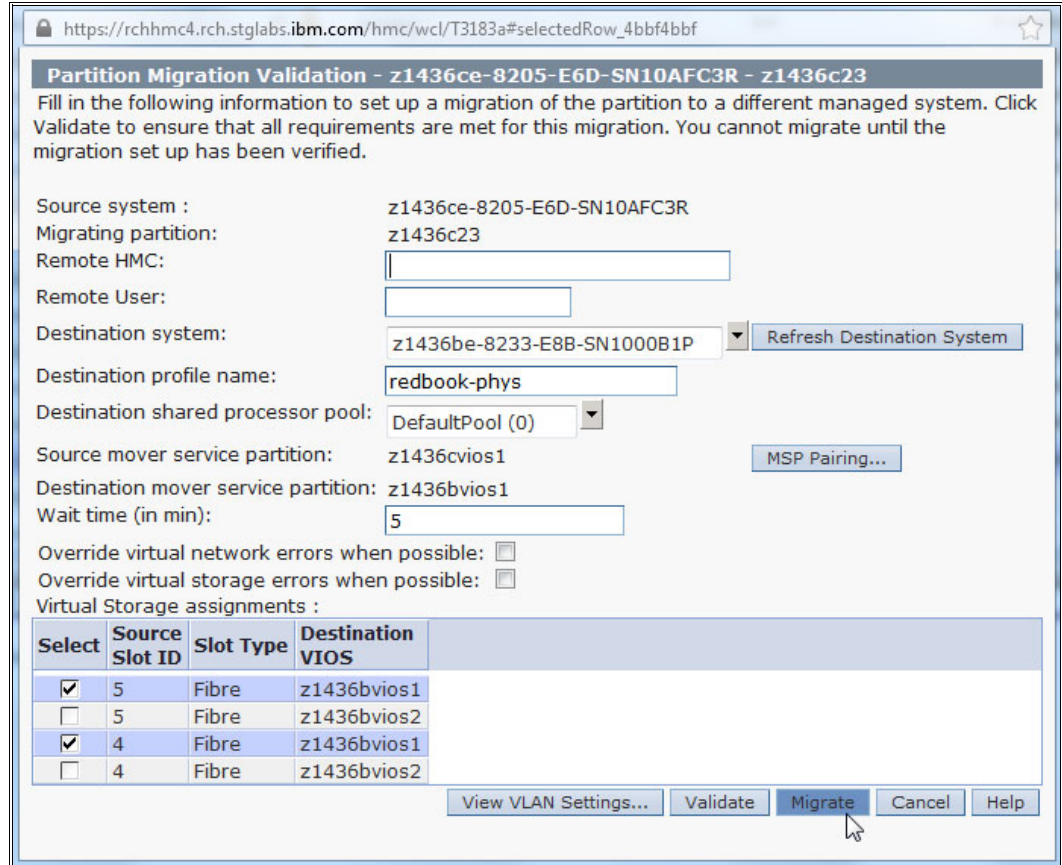


Figure 5-15 Live Partition Mobility, starting migration

- During the migration, a status window displays the progress, as shown in Figure 5-16.

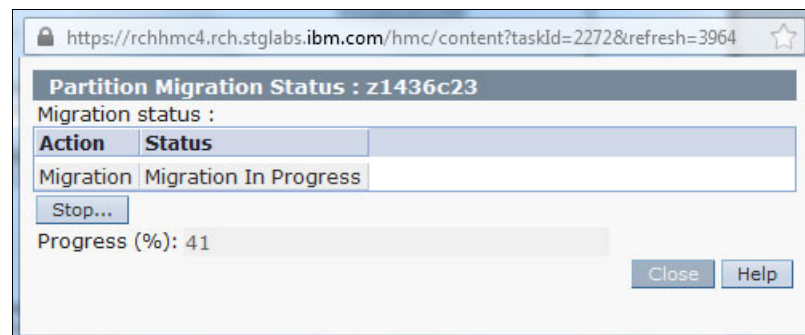


Figure 5-16 Live Partition Mobility, migration status

- On IBM i, the messages in the QSYSOPR message queue (as shown in Figure 5-17) indicate that a partition mobility action is in progress. The license information is displayed after a successful migration if a license key for the target system is not available.

```

                                Display Messages
                                System:      Z1436C23
Queue . . . . . : QSYSOPR                Program . . . . . : *DSPMSG
  Library . . . . : QSYS                    Library . . . . . :
Severity . . . . : 60                      Delivery . . . . . : *HOLD

Type reply (if required), press Enter.
  Partition suspend request in progress.
  Partition resumed after migration.
  License key will not be valid in 70 days on 01/01/14.
  License key will not be valid in 70 days on 01/01/14.

```

Figure 5-17 Live Partition Mobility QSYSOPR messages

5.3.4 Hints and tips for the validation process

During the validation, the following errors and warnings might appear:

- Virtual optical devices must be deallocated before a Live Partition Mobility operation can start. Reset the virtual IOP #290A from IBM i SST and unconfigure the virtual optical target device in VIOS by using the following command:

```
rmdev -dev vtopt0 -ucfg
```

Otherwise, the error message that is shown in Figure 5-18 is displayed.

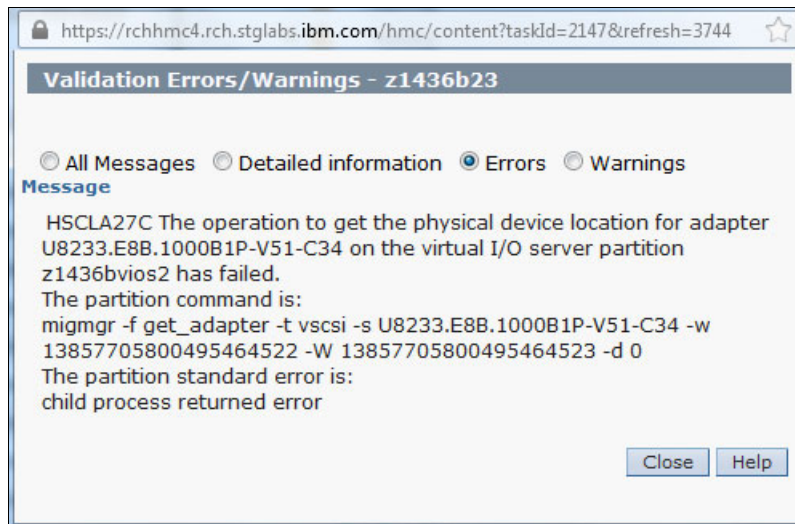


Figure 5-18 Live Partition Mobility error message if the virtual optical device is not deallocated

- ▶ Tape devices must be varied off from the IBM i side. If a tape is still varied on, this does not result in an error message during the validation phase. During the actual migration, the error message that is shown in Figure 5-19 is displayed and the Live Partition Mobility operation stops.



Figure 5-19 Live Partition Mobility error message if tape device is varied on

In addition, the information that is shown in Figure 5-20 can be found in the QSYSOPR message queue of the LPAR that you attempted to move.

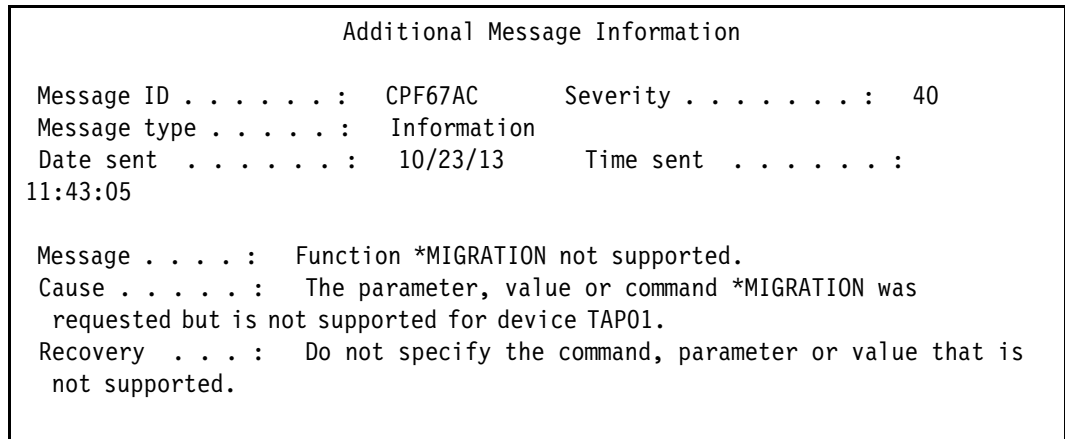


Figure 5-20 Live Partition Mobility QSYSOPR message if tape device not varied off

- ▶ A warning message is shown during validation if the slot numbers for server virtual Fibre Channel adapters or server virtual SCSI adapters cannot be retained, as shown in Figure 5-21. This is a warning only and it does not stop the migration process.

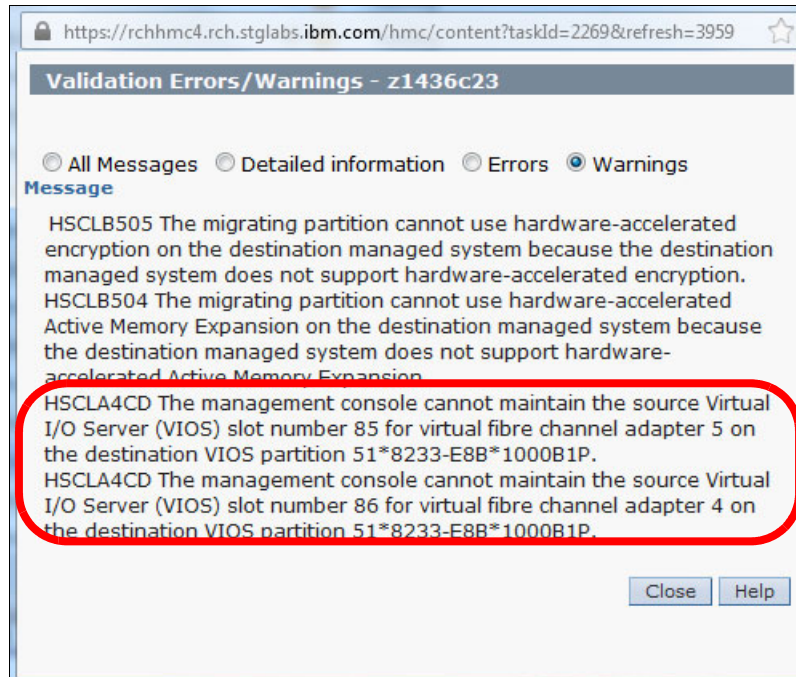


Figure 5-21 Live Partition Mobility warning message regarding slot numbers

- ▶ If you mapped two virtual Fibre Channel adapters to one physical port for the LPAR on the source system, the error message that is shown in Figure 5-22 is displayed. The same error message appears if the Live Partition Mobility operation results in a setup on the target system where two virtual Fibre Channel adapters are mapped to one physical port.

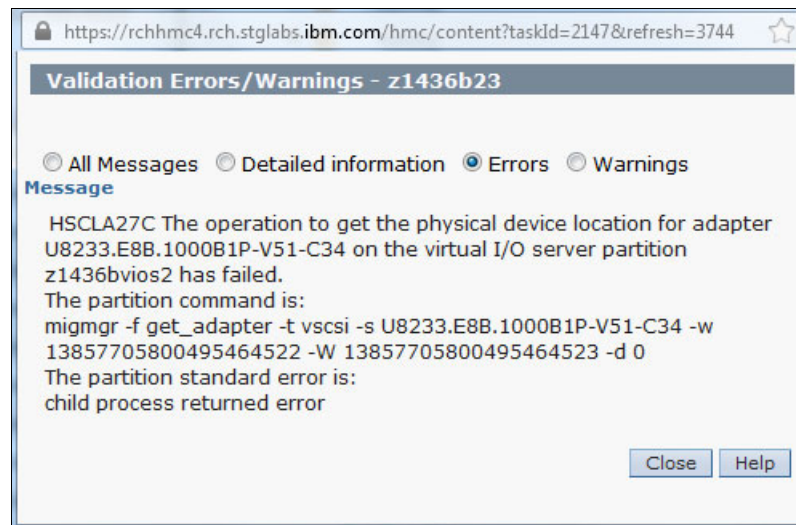


Figure 5-22 Live Partition Mobility error message for physical device location not found



BRMS and FlashCopy

In times where business demands for ever rising availability levels on systems with a high workload, backup windows are closing fast. With Backup, Recovery and Media Services (BRMS) and FlashCopy, an independent ASP (IASP) can be saved without a diminishing effect on applications and disk performance. With PowerHA, FlashCopy management is done from the IBM i partition for ease of use.

This chapter provides information about how an IASP is suspended for a brief period, the FlashCopy is taken, and the BRMS setup is done.

The following topics are covered in this chapter:

- ▶ Introduction to PowerHA, BRMS, and FlashCopy scenario
- ▶ PowerHA and FlashCopy
- ▶ Backup Recovery, and Media Services (BRMS)

6.1 Introduction to PowerHA, BRMS, and FlashCopy scenario

A common setup for this environment is a three-node cluster, as shown in Figure 6-1. High availability is provided between node A and node B by using Metro Mirror on the storage system. In addition, a FlashCopy point in time copy is taken from the LUNs that make up the IASP. This copy of the IASP is attached to the third node in the cluster. Data in that copy of the IASP is saved to tape. By using BRMS, these saves can then be restored to cluster nodes A and B.

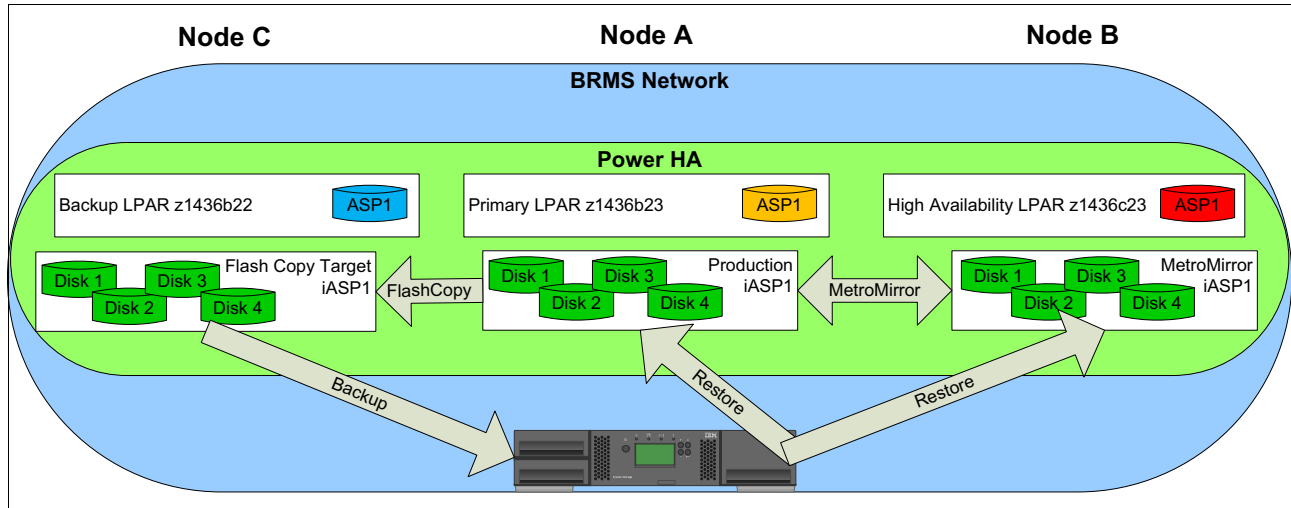


Figure 6-1 PowerHA, FlashCopy, and BRMS scenario

6.2 PowerHA and FlashCopy

The use of FlashCopy in an IBM i environment requires a tight integration between IBM i and the storage system that is running the FlashCopy operation. There are steps that must be taken on the IBM i side, such as quiescing the application to reach a consistent state of data on the disks. When this is accomplished, the FlashCopy is taken on the storage system. After this operation is finished, access to applications and data on the IBM i side must be started again.

PowerHA provides this integration. There are IBM i commands available with PowerHA that communicate with the storage system by using Secure Shell (SSH) communication to run the required operations.

6.2.1 Preparation for FlashCopy on the Storwize side

Preparation on the Storwize family side is easy because most of the configuration tasks are taken over by IBM PowerHA SystemMirror for i. You must create only the required LUNs and attach them to the host you want to use them on.

Preparation must be done on another LPAR. You cannot attach this copy of the IASP to your backup node. The FlashCopy consistency groups and FlashCopy mappings are automatically created when you start a FlashCopy session from the IBM i.

6.2.2 Changes in PowerHA to add the third node into the cluster

As shown in the example in Figure 6-1 on page 174, the two-node cluster for high availability is already set up with node z1436b23 as the primary system and node z1436c23 as the secondary system. To use FlashCopy and BRMS, complete the following steps to extend that configuration to include a third node z1436b22:

1. On the third node, create a device description for the IASP, as shown in the following example:

```
CRTDEVASP DEVD(IASP1) RSRNAME(IASP1)
```

2. On one of the existing cluster nodes, add the third node into the cluster, as shown in the following example:

```
ADDCLUNODE CLUSTER(SVC_CLU) NODE(Z1436B22 ('9.5.216.230' '192.168.10.22'))
```

3. Add the new cluster node into the existing device domain, as shown in the following example:

```
ADDDEVDMNE CLUSTER(SVC_CLU) DEVDMN(SVC_DMN) NODE(Z1436B22)
```

4. Add an SAN Volume Controller copy description that points to the newly created LUNs, as shown in the following example:

```
ADDSVCCPYD ASPCPY(FLASHSVCB) ASPDEV(IASP1) CRG(*NONE) SITE(*NONE)
NODE(Z1436B22) SVCHOST(powerha '/QIBM/UserData/HASM/hads/.ssh/id_rsa'
'9.5.216.236') VRTDSKRNG((8 11 *ALL))
```

Note: For the target copy description that is to be used 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 owned the target copy of the IASP.

A FlashCopy can now be taken by using the **STRSVCSSN** command, as shown in Figure 6-2.

```

                                Start SVC Session (STRSVCSSN)

Type choices, press Enter.

Session . . . . . > FLASH1      Name
Session type . . . . . > *FLASHCOPY *METROMIR, *GLOBALMIR...
Device domain . . . . . *        Name, *
ASP copy:
  Preferred source . . . . . IASP_MM Name
  Preferred target . . . . . FLASHSVCB Name
      + for more values
Incremental flash . . . . . *NO    *NO, *YES
Copy rate . . . . . 0            0-100
Cleaning rate . . . . . 0        0-100
Grain size . . . . . 256        256, 64
Consistency group . . . . . *GEN
Reverse consistency group . . . *GEN

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

```

Figure 6-2 Using the STRSVCSSN command to start FlashCopy

By using the **STRSVCSSN** command, the user can create an incremental FlashCopy relationship and specify a FlashCopy background copy rate and grain size. PowerHA requires the FlashCopy relationships to be included in a consistency group that is, by default, newly created by PowerHA on the SVC/V7000 when a FlashCopy session is started. Alternatively, the user can specify the name for an existing FlashCopy consistency group.

Issuing an End SVC ASP Session (**ENDSVCSSN**) command, as shown in Figure 6-3, ends the FlashCopy relationship. The consistency group is deleted by default.

```

                                End SVC ASP Session (ENDSVCSSN)

Type choices, press Enter.

Session . . . . . FLASH1      Name
Device domain . . . . . *      Name, *
Delete consistency group . . . . *YES    *YES, *NO

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

Figure 6-3 Using the **ENDSVCSSN** command to end FlashCopy

6.3 Backup Recovery, and Media Services (BRMS)

Backup, Recovery, and Media Services (BRMS) for i is the IBM strategic solution for planning and managing the backup of your IBM i server.

BRMS provides the IBM i server with support for policy-oriented setup and the running of backup, recovery, archive, and other removable-media-related operations. BRMS uses a consistent set of intuitive concepts and operations that can be used to develop and implement a backup strategy that is tailored to your business requirements.

For more information about BRMS, see *Backup Recovery and Media Services for OS/400: A Practical Approach*, SG24-4840.

6.3.1 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. Complete the following steps:

1. Log on to the IBM i LPAR 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.

- In the Change Network Group display (see Figure 6-4), change the communication method to *IP and add your IBM i LPARs as nodes to the BRMS network. The host name must be reachable in the network. Add a host table entry, if necessary.

Tip: If the add operation fails, change the DDM attributes to allow a connection without password by using the following command:

```
CHDDMTCPA PWDRQD(*USRID)
```

Change Network Group				Z1436B22 APPN
Network group :	*MEDINV	Position to :		
FlashCopy state . . . :	*ENDPRC			
Communication method :	*IP	*AVAIL, *IP, *SNA, *NONE		
Notify period :	30	30-99999 seconds		
Type options, press Enter.				
1=Add 4=Remove 8=Set time				
	Remote Local	Remote	System	Network
Opt	Location Name	Network ID	Status	Status
1	Z1436B23	APPN		
	Z1436C23	APPN	Online	Active
				Bottom
F3=Exit F5=Refresh F11=BRMS Media Information F12=Cancel				

Figure 6-4 Adding node z1436b23 to the BRMS network

- 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(APPN.Z1436B22)
```

- From the system that backs up the IASP, define which systems receive the save information and define that they also own the save history. If the option to change the save history owner is implemented, it appears as though the target system performed the backup. This is useful for the **PRTRCYRPT** and **STRRCYBRM** commands. Complete the following steps:

- To add a specific system sync, run the following command to change the system name to make it look as though the backup was done by this system and synchronize the reference date and time:

```
CALL QBRM/Q1AOLD PARM('HSTUPDSYNC' '*ADD' 'Z1436B23' 'APPN' 'IASP1' '*CHGSYSNAM')
```

- To add a specific system sync, run the following command to keep the name of who did the backup and synchronize the reference date and time:

```
CALL QBRM/Q1AOLD PARM('HSTUPDSYNC' '*ADD' 'Z1436C23' 'APPN' 'IASP1' '*NORMAL')
```

- To display what is setup, run the following command:

```
CALL QBRM/Q1AOLD PARM('HSTUPDSYNC' '*DISPLAY')
```

- To remove a specific system, use the following command:

```
CALL QBRM/Q1AOLD PARM('HSTUPDSYNC' '*REMOVE' 'Z1436C23' 'APPN' 'IASP1')
```

6.3.2 Configuring BRMS media classes and media policies

BRMS divides tape media into different classes. These classes must be added first because they are used later in the media policy. When the media policy is created, groups of tapes are divided by these classes. If there is a fixed set of tapes for daily, weekly, and monthly backups, create multiple policies.

Complete the following steps to create a media class and a media policy:

1. Run the Work with Media Classes command **WRKCLSBRM** to open the media classes menu.
2. Add the media class, as shown in Figure 6-5. Repeat this step as necessary for your normal BRMS save routines.

Add Media Class		
Type choices, press Enter.		
Media class	DAILYIASP	Name
Density	*ULTRIUM5	F4 for list
Media capacity	*DENSITY	*DENSITY, Number nnnnn.nn
Unit of measure		1=KB, 2=MB, 3=GB
Mark for label print	*NONE	*NONE, *MOVE, *WRITE
Label size	1	1=6 LPI, 2=8 LPI, 3=9 LPI
Label output queue	*SYSPCY	Name, *SYSPCY, *PRTF
Library		Name, *LIBL
Shared media	*YES	*YES, *NO
Write once media	*NO	*YES, *NO
Text		used for save of IASP1

Figure 6-5 Adding a media class for IASP save

- Run the **WRKPCYBRM TYPE(*MED)** command to add a new policy for the save operation, as shown in Figure 6-6. This example uses a retention period of one day. In most cases, you might prefer longer retention periods.

```

                                Create Media Policy

Type choices, press Enter.

Media policy . . . . . DAILYIASP      Name
Retention type . . . . . 2              1=Date, 2=Days,
                                         3=Versions, 4=Permanent
Retain media . . . . . 1                Date, Number
Deleted library retention. . . . . *NONE      Number, *NONE
Move policy . . . . . *NONE             Name, *NONE, *ADSM, F4
Media class . . . . . DAILYIASP        Name, *SYSPCY, *ADSM, F4
Storage location . . . . . *ANY         Name, *ANY, F4 for list
Save to save file . . . . . *NO         *YES, *NO
ASP for save files . . . . . *SYSTEM    Name, *SYSTEM, 1-32
Save file retention type . . . . . 4     1=Date, 2=Days,
                                         3=Permanent, 4=None
Retain save files . . . . . *NONE       Date, Number, *NONE
ASP storage limit . . . . . *SYS        *SYS, 1-99
Secure media . . . . . *NO              *YES, *NO, *ADSM

More...

Type choices, press Enter.

Required volumes . . . . . *NONE        *NONE, 1-9999
Mark volumes for duplication . . . . . *NO      *NO, *YES
Mark history for duplication . . . . . *NO      *NO, *YES
Text . . . . . used for save of IASP1

```

Figure 6-6 Add media policy for IASP save

6.3.3 Configuring the BRMS control group

Control groups define logical groups of libraries and objects that possess similar backup, retention, and recovery requirements. In addition to allowing you to define the order in which backup, archive, and recovery processing occurs, control groups provide for special related actions, such as tape loads, processing subsystems, and job queues. Control groups provide exits for user-defined processing during the backup cycle.

Figure 6-7 shows in which order the BRMS control groups that are created in this section must be run. This can be done with the job scheduler or by using the Run Remote Command (**RUNRMTCMD**) command.

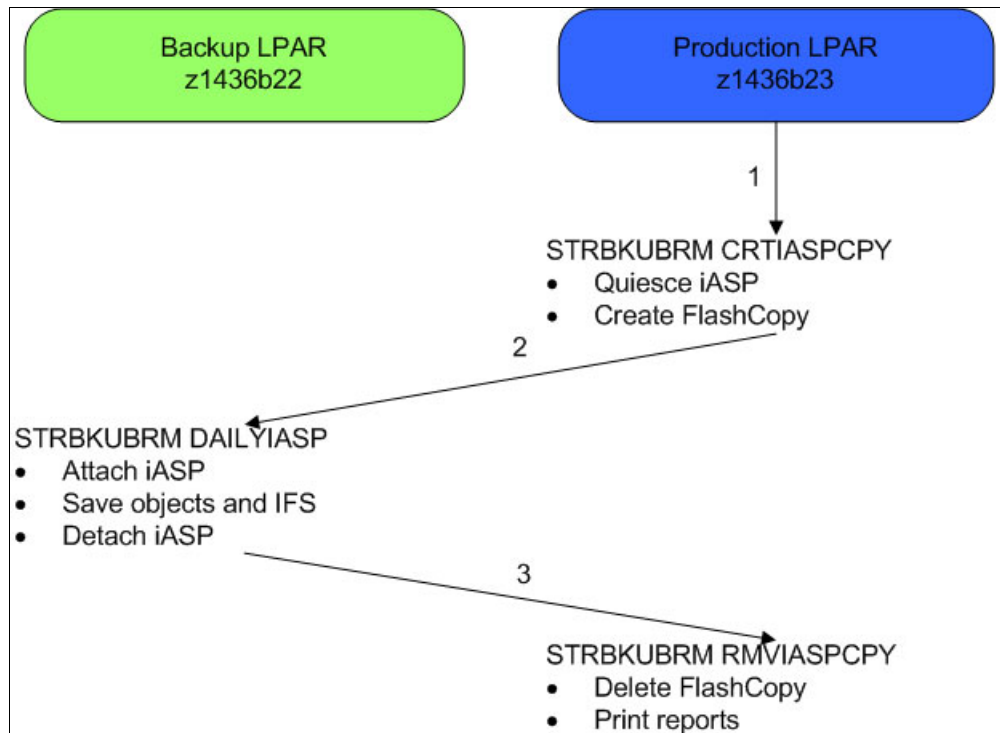


Figure 6-7 Backup run for FlashCopy IASP

Use the Work with Backup Control Groups command (**WRKCTLGBRM**) to create BRMS control groups. Complete the following steps:

1. As shown in Example 6-1, create control groups on the primary and high availability node, which is used to create and remove the FlashCopy of the IASP1. Also, the recovery reports are printed on the primary node.

Example 6-1 Creating two BRMS control groups with CRTIASPCPY and RMVIASPCPY

```

CRTIASPCPY *EXIT CHGASPACT ASPDEV(IASP1) OPTION(*SUSPEND) SSPTIMO(3)
           *EXIT STRSVCSSN SSN(FLASH1) TYPE(*FLASHCOPY) ASPCPY((IASP_MM FLASHVCB))
           *EXIT CHGASPACT ASPDEV(IASP1) OPTION(*RESUME)

RMVIASPCPY *EXIT ENDSVCSSN SSN(FLASH1)
           *EXIT SETASPGRP ASPGRP(IASP1)
           *EXIT STRMNTBRM RUNCLNUP(*NO) RTVVOLSTAT(*NO) PRTEXPMED(*NO) PRTBKUACT(*NO) PRTRCYRPT(*RCYANL)
  
```

2. On the backup node, create the control group that is used to back up the IASP copy, as shown in Example 6-2.

Example 6-2 Backup node control group

```

DAILYIASP *EXIT          DAILYIASP  DAILYIASP          VRYCFG CFGOBJ(IASP1) CFGTYPE(*DEV) STATUS(*ON)
           *ALLUSR   IASP1  DAILYIASP  DAILYIASP  *YES
           *LINK     IASP1  DAILYIASP  DAILYIASP  *YES
           *EXIT          DAILYIASP  DAILYIASP          VRYCFG CFGOBJ(IASP1) CFGTYPE(*DEV) STATUS(*OFF)
           *EXIT          DAILYIASP  DAILYIASP          SNDMSG MSG('BACKUPS ARE COMPLETE') TOUSR(*SYSOPR)
  
```

3. On the backup LPAR, the production LPAR was added to the list of owners for the saved object, as described in 6.3.1, "Setting up the BRMS network group" on page 176.
4. Run the BRMS control groups by using Start Backup by using the BRM (**STRBKUBRM**) command in the correct order to save the IASP objects.

Figure 6-8 shows that after running the following Start Recovery by using BRM (**STRRCYBRM**) command with ***LCL** as the originating system, all information about the saved IASP objects from the backup LPAR are available:

```
STRRCYBRM OPTION(*ALLUSR) ACTION(*RESTORE) FROMSYS(*LCL)
```

```

                                Select Recovery Items                                Z1436B23
                                                                                   Select action . . . . . : *ALL
                                                                                   Select volume . . . . . :
Type options, press Enter.
  1=Select  4=Remove  5=Display  7=Specify object

  Saved      Save      Save      Save      Parallel Volume      File      Expire
Opt Item     Date       Time      Type      Devices  Serial      Sequence  Date
  QRCL00033  10/16/13  15:29:15 *FULL    1FA390   3          10/17/13
  QSYS200033 10/16/13  15:29:15 *FULL    1FA390   2          10/17/13
  SYSIB00033 10/16/13  15:29:15 *FULL    1FA390   4          10/17/13
  TESTJORDAN 10/16/13  15:29:15 *FULL    1FA390   5          10/17/13
                                                                                   Bottom

F3=Exit  F5=Refresh  F9=Recovery defaults  F11=Object View
F12=Cancel  F14=Submit to batch  F16=Select

```

Figure 6-8 STRRCYBRM command on the production system



A

PowerHA Tools IBM i

This appendix describes the PowerHA Tools for IBM i offerings and services that are available from IBM Systems Lab Services.

The PowerHA Tools for IBM i complement and extend the PowerHA and IBM storage capabilities for high availability (HA) and disaster recover (DR).

The PowerHA Tools for IBM i provide the following benefits:

- ▶ Helps reduce business risk and improve resiliency for critical applications.
- ▶ Simplifies setup and automation of HA, DR, and backup solutions.
- ▶ Reduces the cost of maintaining and regular testing of an HA/DR environment.
- ▶ Facilitates flexible deployment options for single or multi-site protection.
- ▶ Assures consistent deployment by using best practices and experienced consultants.

For more information about PowerHA Tools for IBM i, see the following IBM Systems Lab Services and Training website:

<http://www.ibm.com/systems/services/labservices>

PowerHA Tools for IBM i

Table A-1 lists the PowerHA Tools for IBM i that are available from IBM Systems Lab Services.

Table A-1 PowerHA Tools for IBM i

PowerHA Tools for IBM i	Capability	Benefit	DS8000	Storwize	Internal
Smart Assist for PowerHA on IBM i	Provides operator commands and scripts to supplement PowerHA installation and ongoing operations for IASP enabled applications.	Simplifies deployment and ongoing management of HA for critical IBM i applications.	X	X	X
IASP Copy Services Manager (Automated recovery with faster IASP-level vary on, no system IPL)					
Flashcopy	Automates Flashcopy of IASP for daily offline backup with seamless BRMS integration.	Increases application availability by reducing or eliminating backup window for routine daily backups.	X	X	
LUN-level Switching	Simplifies deployment and automates switching an IASP between IBM i cluster nodes in one data center.	Enables a business continuity manager to provide a simple, single site HA solution.	X ¹	²	
Metro Mirror or Global Mirror	Simplifies initial deployment and automates ongoing server and storage management of two-site Metro Mirror or Global Mirror HA or DR solutions. Requires IASP enabled applications.	Enables a business continuity manager to provide seamless operation of integrated server and storage operations for two-site HA and DR.	X		
Metro Global Mirror (MGM)	Extends PowerHA functionality to provide three-site server or storage replication solution containing Metro Mirror for HA with Global Mirror for DR. Requires IASP enabled applications and IBM Tivoli Productivity Center for Replication.	Enables a business continuity manager to further lower business risk and maximize business resilience for critical applications that require three-site HA and DR protection.	X		
Full System Copy Services Manager (Automated recovery, requires full system IPL)					
Flashcopy	Automates full system Flashcopy for daily offline backup with integrated support for BRMS without IASP enabled applications.	Increases application availability by reducing or eliminating backup window for online daily backups. Enables an entry solution while planning IASP enablement.	X	X	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 solitons 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		

¹ DS8000 support available with PowerHA Tools for IBM i 6.1 or earlier; included in PowerHA SystemMirror 7.1

² V7000 support included with PowerHA 7.1 TR6

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 and 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.

Related publications

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

IBM Redbooks

The following IBM Redbooks publications provide more information about the topics in this document. Some of the publications that are referenced in this list might be available in softcopy only:

- ▶ *PowerHA SystemMirror for IBM i Cookbook*, SG24-7994
- ▶ *IBM Power Systems HMC Implementation and Usage Guide*, SG24-7491
- ▶ *Simple Configuration Example for Storwize V7000 FlashCopy and PowerHA SystemMirror for i*, REDP-4923
- ▶ *IBM i and Midrange External Storage*, SG24-7668
- ▶ *IBM System Storage SAN Volume Controller and Storwize V7000 Replication Family Services*, SG24-7574
- ▶ *Backup Recovery and Media Services for OS/400: A Practical Approach*, SG24-4840

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

<http://www.ibm.com/redbooks>

Other publications

Another relevant source of more information is *IBM Real-time Compression Evaluation User Guide*, S7003988, which is available at this website:

<http://www-01.ibm.com/support/docview.wss?uid=ssg1S7003988&mys=s028&mynp=0CST3FR7&mync=E>

Online resources

These websites are also relevant as further information sources:

- ▶ IBM System Storage Interoperation Center (SSIC):
<http://www-03.ibm.com/systems/support/storage/ssic/interoperability.wssDescription1>
- ▶ IBM System Storage SAN Volume Controller:
<http://www-03.ibm.com/systems/storage/software/virtualization/svc/>
- ▶ IBM Storwize V3500:
http://www-03.ibm.com/systems/hk/storage/disk/storwize_v3500/index.html

- ▶ IBM Storwize V3700:
http://www-03.ibm.com/systems/storage/disk/storwize_v3700/index.html
- ▶ IBM Storwize V7000 and Storwize V7000 Unified Disk Systems:
http://www-03.ibm.com/systems/storage/disk/storwize_v7000/index.html
- ▶ IBM Comprestimator Utility:
<http://www-304.ibm.com/support/customercare/sas/f/comprestimator/home.html>
- ▶ IBM Systems Workload Estimator:
<http://www-912.ibm.com/wle/EstimatorServlet>
- ▶ Subsystem Device Driver Path Control Module (SDDPCM) driver:
<http://www-01.ibm.com/support/docview.wss?uid=ssg1S4000201>
- ▶ Virtual I/O Server Advisor tool:
<https://www.ibm.com/developerworks/community/wikis/home?lang=en#!/wiki/Power%20Systems/page/VIOS%20Advisor>

Help from IBM

IBM Support and downloads:

<http://www.ibm.com/support>

IBM Global Services:

<http://www.ibm.com/services>



IBM i and IBM Storwize Family: A Practical Guide to Usage Scenarios

(0.2"spine)
0.17"->0.473"
90->249 pages



IBM i and IBM Storwize Family

A Practical Guide to Usage Scenarios



VIOS explained from an IBM i perspective

Using PowerHA together with IBM Storwize Family

Configuring BRMS in a high available solution

The use of external storage and the benefits of virtualization became a topic of discussion in the IBM i area during the last several years. The question tends to be, what are the advantages of the use of external storage that is attached to an IBM i environment as opposed to the use of internal storage. The use of IBM PowerVM virtualization technology to virtualize Power server processors and memory also became common in IBM i environments. However, virtualized access to external storage and network resources by using a VIO server is still not widely used.

This IBM Redbooks publication gives a broad overview of the IBM Storwize family products and their features and functions. It describes the setup that is required on the storage side and describes and positions the different options for attaching IBM Storwize family products to an IBM i environment. Basic setup and configuration of a VIO server specifically for the needs of an IBM i environment is also described.

In addition, different configuration options for a combined setup of IBM PowerHA SystemMirror for i and the Storwize family products are described and positioned against each other. Detailed examples are provided for the setup process that is required for these environments.

The information that is provided in this book is useful for clients, IBM Business Partners, and IBM service professionals who need to understand how to install and configure their IBM i environment with attachment to the Storwize family products.

INTERNATIONAL TECHNICAL SUPPORT ORGANIZATION

BUILDING TECHNICAL INFORMATION BASED ON PRACTICAL EXPERIENCE

IBM Redbooks are developed by the IBM International Technical Support Organization. Experts from IBM, Customers and Partners from around the world create timely technical information based on realistic scenarios. Specific recommendations are provided to help you implement IT solutions more effectively in your environment.

For more information:
ibm.com/redbooks

SG24-8197-00

ISBN 0738439223