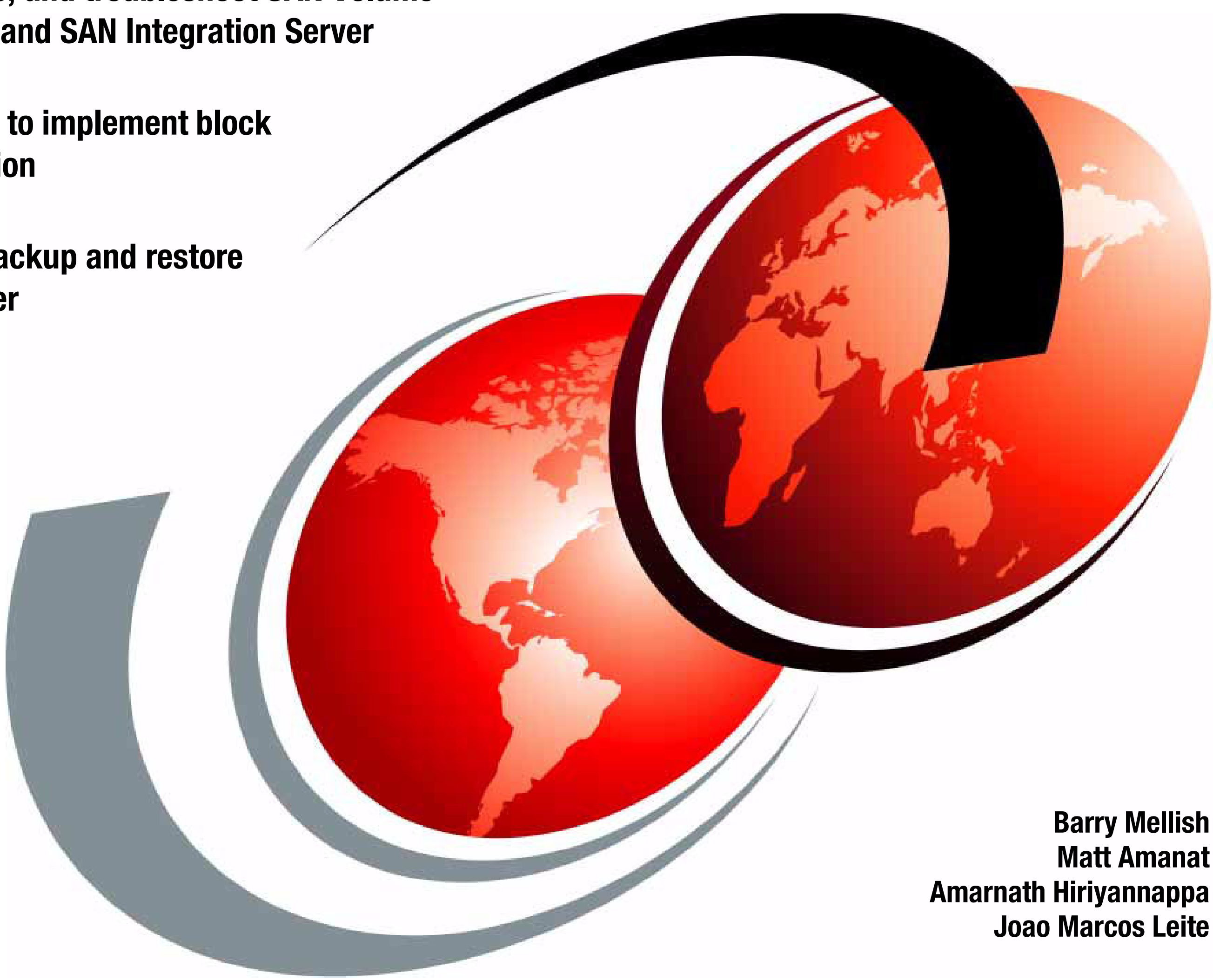
IBM

IBM TotalStorage SAN Volume Controller and SAN Integration Server

Install, use, and troubleshoot SAN Volume Controller and SAN Integration Server

Learn how to implement block virtualization

Perform backup and restore on a cluster



Barry Mellish
Matt Amanat
Amarnath Hiriyannappa
Joao Marcos Leite

Redbooks

ibm.com/redbooks

International Technical Support Organization

IBM TotalStorage SAN Volume Controller and SAN Integration Server

January 2004

SG24-6423-01

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

Second Edition (January 2004)

This edition applies to the San Volume Controller Version 1 Release 1.1 as announced on 14 October 2003. The testing carried out to produce this book was carried out with pre-GA code. Some of the window examples may differ slightly from the code that you use. However, the general principles of operation are consistent.

Contents

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 give 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 Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites 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.

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 illustrates 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. 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 IBM's application programming interfaces.

Trademarks

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

eserver®
eserver®
ibm.com®
xSeries®
AFS®
AIX 5L™
AIX®
developerWorks®
Enterprise Storage Server®
FlashCopy®
IBM®
Netfinity®
Redbooks (logo) ™
Redbooks™
RETAIN®
Sequent®
Storage Tank™
Tivoli®
TotalStorage®

The following terms are trademarks of other companies:

Intel, Intel Inside (logos), MMX, and Pentium are trademarks of Intel Corporation in the United States, other countries, or both.

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

Java and all Java-based trademarks and logos are trademarks or registered trademarks of Sun Microsystems, Inc. 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, and service names may be trademarks or service marks of others.

Preface

This IBM® Redbook is a detailed technical guide to the IBM TotalStorage® SAN Volume Controller (SVC) and IBM TotalStorage SAN Integration Server (SIS). It covers the November 2003 enhancements. Although much of the discussion in this redbook is based upon the SVC, it equally applies to the SIS. The SIS is a specially packaged SVC with the addition of storage, storage area network (SAN) switches, and an Ethernet switch.

The book covers the following areas:

- Storage virtualization: This is only a high-level overview since this topic is described in detail in other Redbooks.
- Architecture of the SVC and SIS
- Implementing and configuring the SVC and SIS
- Using virtualization and advanced functions such as FlashCopy® and Remote Copy
- Migrating existing storage to the SVC and SIS
- Backup and recovery

You can now combine the virtualization functions of the SAN Volume Controller storage software with the Cisco MDS 9000 Series Caching Services Module from the Cisco MDS 9000 Family. The virtualization functions of this new product are similar to those described in this book. For a full discussion of this new product, read the IBM Redbook, *Implementing the IBM TotalStorage SAN Volume Controller Storage Software on the Cisco MDS 9000,* SG24-7059.

The team that wrote this redbook

This book is the result of two residencies, the first at IBM Hursley, the second at IBM Raleigh. It was produced by a team of specialists from around the world.

Barry Mellish is a certified I/T Specialist and is currently working as a Regional Storage Virtualization Specialist in EMEA. Previous to this, he was a Project Leader at the International Technical Support Organization (ITSO), San Jose Center, for four years. He has co-authored 14 Redbooks and has taught many classes worldwide on storage subsystems. He joined IBM United Kingdom 20 years ago. Before joining the ITSO, he worked as a Senior Storage Specialist on the Disk Expert Team in EMEA.

Matt Amanat is an I/T Specialist with the Storage Technical Sales Support group in the Americas. He has taught storage networking concepts to specialists across the Americas and regularly acts as a storage networking and open systems storage subject matter expert (SME) for solution assurance reviews. Prior to joining the Storage Systems Group, Matt was a specialist with the Netfinity/IBM @server® xSeries® server group in Canada.

Amarnath Hiriyannappa is an I/T Specialist who has eight years experience with IBM in various technical positions in both marketing and development with servers, networking, and storage organization. He is in charge of defining comprehensive architectures and recommendations to clients evolving from traditional server centric to SAN data centric infrastructure. His responsibilities cover the complete range of technologies for designing and implementing optimized SAN information systems, such as IBM Tivoli® SAN and storage management, servers, business applications, networking, storage, and virtualization products.

Joao Marcos Leite is a certified I/T Specialist who has 24 years of experience as an I/T Specialist in Brazil. He joined the IBM Storage Subsystems Group in 2000 as a Field Technical Sales Support specialist. He has been appointed as the technical focal point for storage software products in Latin America, and as a Regional Designated Specialist for storage software and virtualization products for the IBM Americas Group – South Region (Latin America). He has taught storage virtualization for clients and in industry and IBM events, and acts as a storage SME in solution assurance reviews. Prior to joining IBM, he worked for Fujitsu as a pre- and post-sales support analyst for mainframe and storage products.

From left to right: Barry Mellish, Matt Amanat, Amarnath Hiriyannappa, and Joao Marcos Leite

The team that wrote the first edition is:

Craig McKenna is a Senior IT Specialist with 13 years of experience with IBM and is currently employed as a Field Technical Support Specialist within the Storage Systems Group (SSG) in Australia and New Zealand. He is involved throughout the pre-sales and post-sales process, from designing end-to-end storage solutions and competitive positioning to implementation and problem determination activities across the entire range of IBM TotalStorage offerings. He has taught storage subject matter to a variety of audiences and has spoken as a SME at industry events. He has most recently been identified as the technical focal point for storage software products in Australia and New Zealand.

Philippe Jachimczyk is a certified I/T Specialist who has 17 years experience with IBM in various technical positions in both marketing and development with servers, networking, and storage organization. He is in charge of defining comprehensive architectures and recommendations to clients evolving from traditional server centric to SAN data centric infrastructure. His responsibilities cover the complete range of technologies for designing and implementing optimized SAN information systems, such as Tivoli® SAN and storage management, servers, business applications, networking, storage, and virtualization products.

Robert Symons is a Senior IT Specialist working for ITS in Canada. He is the team lead of the Canadian disk support team in ITS and provides remote support to clients and system support representatives. He has over 10 years experience as a Country Support Representative. Rob's main focus is solving hardware disk storage problems.

Sean Crawford is an I/T Specialist with the Storage Technical Sales Support group in the Americas. He has taught storage networking concepts to specialists across the Americas and regularly acts as a storage networking and open systems storage SME for solution assurance reviews. Prior to joining the Storage Systems Group, Sean was a specialist with the Netfinity/xSeries server group in Canada.

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

There are many people that contributed to this book. In particular, we thank the development and PFE teams in Hursley, the SSG Advanced Technical Support team, and the Open Systems Lab in San Jose, where the practical work was carried out. Bill Wiegand of ATS and Jeff Irving of SSG Technical Marketing Support, although not authors, contributed many of the scripts and ideas that are documented in the book.

Eric Wong
IBM France

Reg Beer
Carlos Fuente
Ian Godwin
Peter Grandison
Marius Kalmantas
Steve Legg
Rob Nicholson
Nick O'Leary
Nick O'Rourke
Josh Sampays
Bill Scales
Clive Wells
IBM Hursley

Brian Sherman
IBM Canada

Ron Henkhaus
Bill Wiegand
IBM Advanced Technical Support

Malkit Hayun
IBM New York

Michael Dean
IBM Atlanta

Chuck Grimm
Dan O'Hare
IBM Rochester

Carl Bishop
Jim Blue
William Champion
Huang Chen

Tom Daniels
Mat Geiser
Tim Gilbert
Larry Gooch
Parker Grannis
Nathaniel Hatfield
Joao Molina
Jeff Ottman
Alan Petersburg
Jerry Poole
Joel Stapf
Claude Summers
Doug Vassello
David Watts
IBM Raleigh

Siebo Friesenborg
Bruce McNutt
John Ponder
IBM Tucson

Linda Duyanovich
Hemanth Kalluri
Kraig Klippel
Chris Saul
Brian Smith
Everend Thorne
John Zaborsky
IBM San Jose

Charlotte Brooks
Maritza M. Dubec
Jon Tate
IBM ITSO Center San Jose

Become a published author

Join us for a two- to six-week residency program! Help write an IBM Redbook dealing with specific products or solutions, while getting hands-on experience with leading-edge technologies. You'll team with IBM technical professionals, Business Partners and/or clients.

Your efforts will help increase product acceptance and client satisfaction. As a bonus, you'll develop a network of contacts in IBM development labs, and increase your productivity and marketability.

Find out more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

We want our Redbooks™ to be as helpful as possible. Send us your comments about this or other Redbooks in one of the following ways:

- Use the online **Contact us** review redbook form found at:

  `ibm.com/redbooks`

- Send your comments in an Internet note to:

  `redbook@us.ibm.com`

- Mail your comments to:

  IBM Corporation, International Technical Support Organization
  Dept. QXXE Building 80-E2
  650 Harry Road
  San Jose, California 95120-6099

1

Introduction to storage virtualization

This chapter describes the need for storage virtualization and the IBM approach to both symmetrical and asymmetrical storage virtualization. We articulate the fundamental differences between the two architectures to help you understand why IBM has chosen to use symmetrical virtualization for the IBM TotalStorage SAN Volume Controller (the topic of the remainder of this redbook).

1.1 The need for storage virtualization

At the business level, clients are faced with three major storage challenges:

- **Managing storage growth**: Storage needs continue to grow at over 50% per year. Managing storage is becoming more complex than ever. We now have to deal with multiple server platforms and different operating systems, which may be connected to a storage area network (SAN) with multiple and diverse storage platforms.
- **Increasing complexity**: Although the declining cost of storage per megabyte (MB) makes it attractive to add additional disks, the increasing complexity of managing this storage results in over-used staff and under-used information technology (IT) resources. Combining this with the shortage of skilled storage administrators, it is possible to add significant cost and introduce risk to storage management.
- **Maintaining availability**: The added complexity of 24x7 environments significantly reduces, for example, the efficiency of conducting routine maintenance, scheduling backups, data migration, and introducing new software and hardware. This problem is compounded by the fact that, as availability increases, so does the inherent costs.

Storage needs are rising, and the challenge of managing disparate storage systems is growing. IBM TotalStorage SAN Volume Controller brings your devices together in a *virtual pool* to make all your storage appear as:

- One “logical” device to centrally manage and to allocate capacity as needed
- One solution to help achieve the most effective use of your key storage resources on demand

Virtualization solutions can be implemented in the storage network, in the server, or in the storage device itself. The IBM storage virtualization solution is SAN-based, which helps allow for a more open virtualization implementation. Locating virtualization in the SAN, and therefore in the path of input/output (I/O) activity, helps to provide a solid basis for policy-based management. The focus of IBM on open standards means our virtualization solution supports freedom of choice in storage-device vendor selection.

The IBM TotalStorage SAN Volume Controller solution is to:

- Simplify storage management
- Reduce IT data storage complexity and costs while enhancing scalability
- Extend on-demand flexibility and resiliency to your IT infrastructure

1.2 Symmetrical virtualization

In a conventional SAN, the logical unit numbers (LUNs) that are defined within the storage subsystem are directly presented to the host or hosts. Symmetrical virtualization, otherwise known as *block aggregation* or *in-band virtualization*, essentially means having an appliance in the data path that can take physical storage from one or more storage subsystems and offer it to hosts in the form of a virtual disk.

The SNIA Block Aggregation Model (Figure 1-1) specifies that block aggregation can be performed within hosts (servers), in the storage network (storage routers, storage controllers), or in storage devices (intelligent disk arrays).

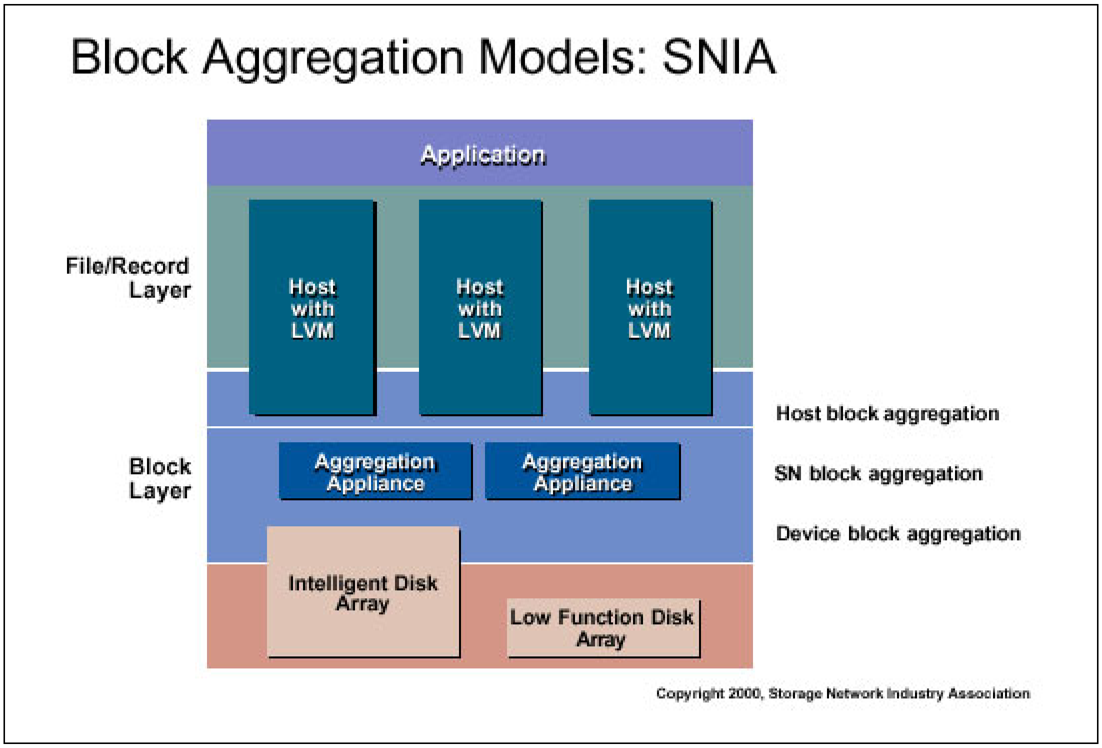


Figure 1-1 SNIA Block Aggregation Models

While each of these approaches has pros and cons and all are available in various forms from various vendors, we chose to develop our latest block aggregation product (IBM TotalStorage SAN Volume Controller) within the storage network.

Block aggregation within the storage network provides three significant benefits to clients:

- Increased storage administrator productivity

  Administrators can manage, add, and migrate physical disks non-disruptively from an application server point of view. This is accomplished by providing insulation between the server's view of the logical disks and the disks as presented by the storage subsystem. Productivity is improved by allowing administrators to perform management functions when convenient rather than waiting for ever decreasing maintenance windows. Downtime requirements are all but eliminated.

- Providing a common platform for advanced functions

  By providing a logical view of physical storage, advanced functions like disaster recovery can be done at a single point in the SAN in a consistent way regardless of the underlying physical storage. FlashCopy, Peer-to-Peer Remote Copy (PPRC), and data migration can also be performed in a consistent way. This common platform is used to provide other advanced functions over time such as advanced security and quality of service (QoS) capabilities.

- Improved capacity utilization

  Spare capacity on underlying physical disks can be reallocated non-disruptively from an application server point of view irrespective of the server operating system or platform type. Logical disks can be created from any of the physical disks being managed by the virtualization device (that is, vendor agnostic).

Figure 1-2 shows the IBM approach to block aggregation.

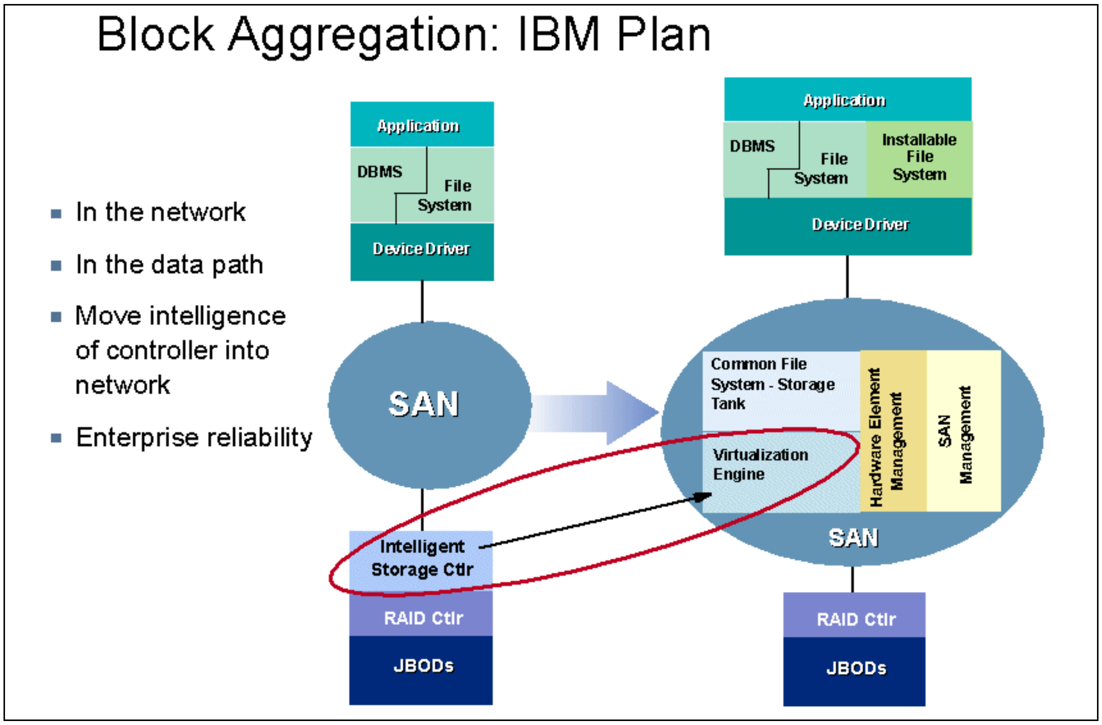


Figure 1-2 IBM plan for block aggregation

In addition to the three major benefits outlined above, abstracting the hosts from directly accessing the storage subsystem or subsystems has many other benefits over other methods of block aggregation including:

- It gives you the ability to add advanced functions and apply them to the entire storage infrastructure. The first release of the product offers:
  - Copy Services (PPRC and FlashCopy)
  - Data migration
  - Read and Write Caching
- Later releases of the product offer such functions as:
  - Quality of Service
  - Performance based data migration
  - Performance optimization in the data path
  - Advanced security
- It does not lock a client into a particular storage hardware vendor.
- It is not intrusive on the hosts.
- It can offload function from the hosts.
- It can support storage management from multiple ISVs.
- It offers superior scalability.

Our virtualization product provides redundant, modular, and scalable solutions. It is based on a clustered IBM SAN appliance running a Linux kernel to support high availability and performance. Additional nodes are capable of being added non-disruptively providing enterprise class scalability. Our long history of storage controller development has enabled us to develop systems where, in the exceptionally rare case that a failure occurs, the virtualization device can fail and recover gracefully. Figure 1-3 shows a representation of the IBM TotalStorage SAN Volume Controller.

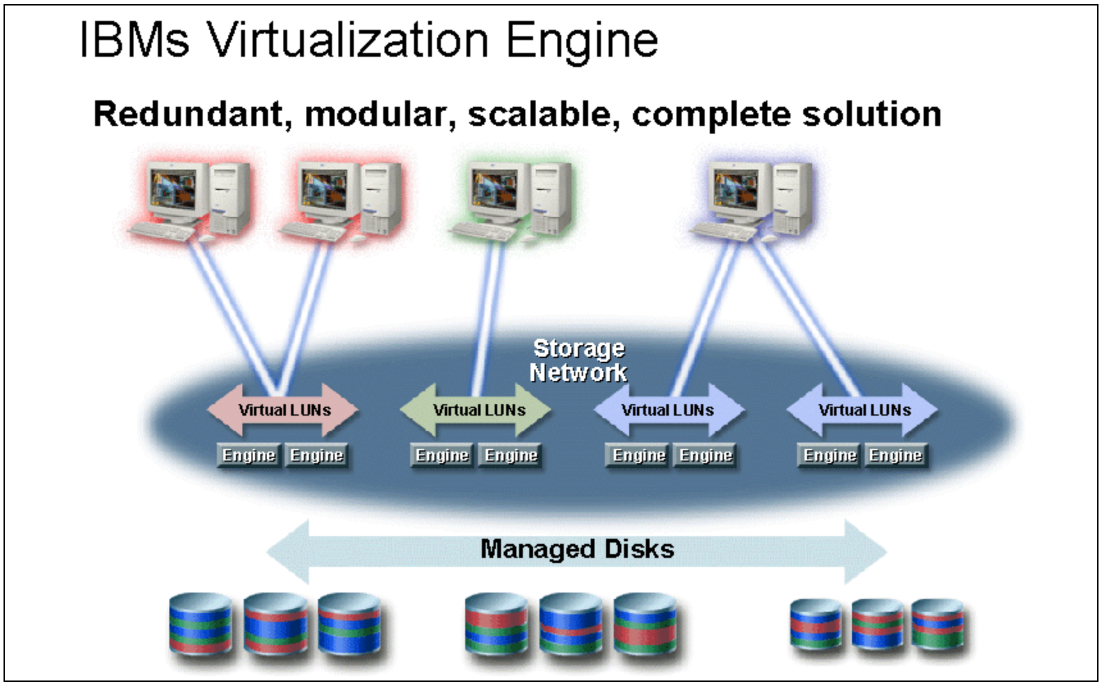


Figure 1-3 Conceptual diagram of the IBM SAN Volume Controller

In summary, we add enterprise class block aggregation functionality to the storage network. Our solution improves storage administrator productivity, provides a common base for advanced functions, and provides for more efficient use of storage. Our product is being designed to be delivered as a horizontally scalable, integrated solution based on IBM SAN appliance, and Linux using a fault tolerant clustered architecture.

1.3 Asymmetrical virtualization

Asymmetrical virtualization, otherwise known as *file aggregation* or *out-of-band virtualization*, is when the virtualization appliance is not in the data path. Typically, asymmetrical virtualization is more geared toward file sharing across the SAN. To this end, it typically involves a single file system in a single name space.

File aggregation is a similar technique as block aggregation. However, rather than dealing with blocks of data, file aggregation addresses the needs of accessing and sharing files in a storage network. In the SNIA model, hosts get file metadata from file system or Network Attached Storage (NAS) controllers, and then access the data directly. File aggregation can be used in conjunction with or independent from block aggregation. Figure 1-4 shows the SNIA file aggregation model.

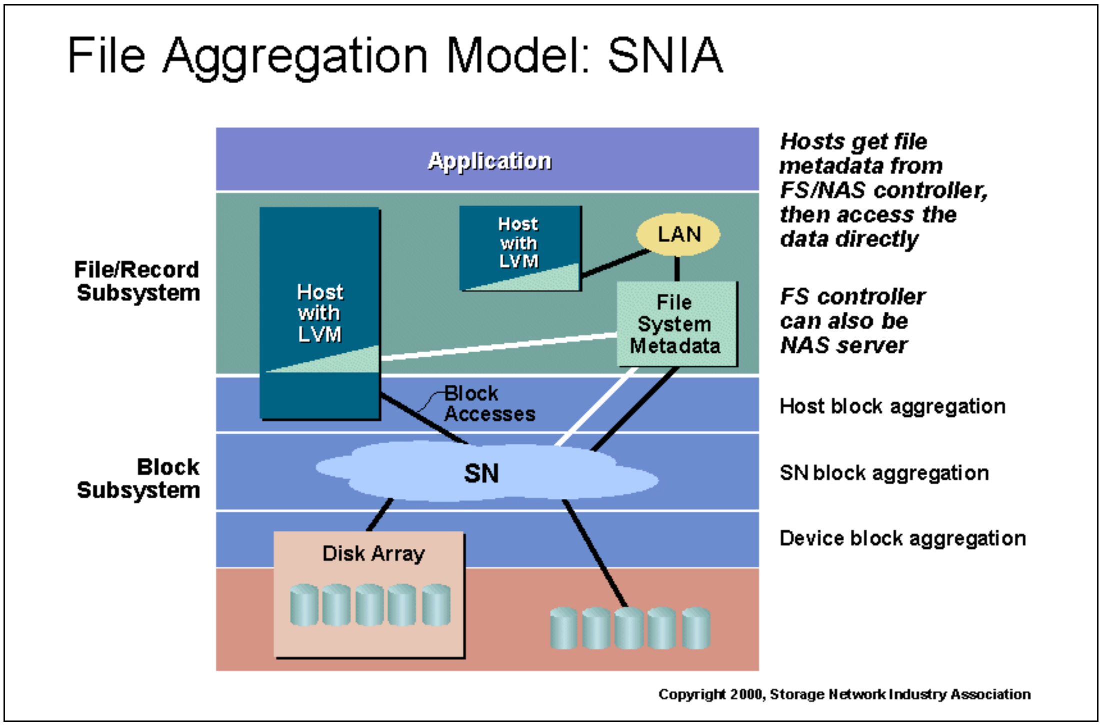


Figure 1-4 SNIA file aggregation model

Our approach is through the use of a common file system based on the IBM Storage Tank™ technology initiative. Initially, this file system covers all SAN-based files and is later expanded to cover all files in an enterprise. We provide a metadata server cluster for managing information about files and have designed our file system clients to access disks directly. For clients with both SANs and NAS, we provide a converged SAN and NAS solution based on the Storage Tank technology.

Our solution, the IBM TotalStorage SAN file system, is designed to provide a common file system specifically designed for storage networks. By managing file details (metadata) on the storage network instead of in individual servers, we can make a single file system available to all application servers on that network. Doing so provides immediate benefits: a single point of management and a single name space, and common management for all files in the network, eliminating management of files on a server by server basis.

The SAN file system's design automates routine and error prone tasks using policy based automation, initially to manage file placement and handle "out of space" conditions. The SAN file system's design also allows the first true heterogeneous file sharing, where the reader and writer of the exact same data can run different operating systems. Initially, the SAN file system provides a range of the most commonly used operating systems in the enterprise SANs (see Figure 1-5).

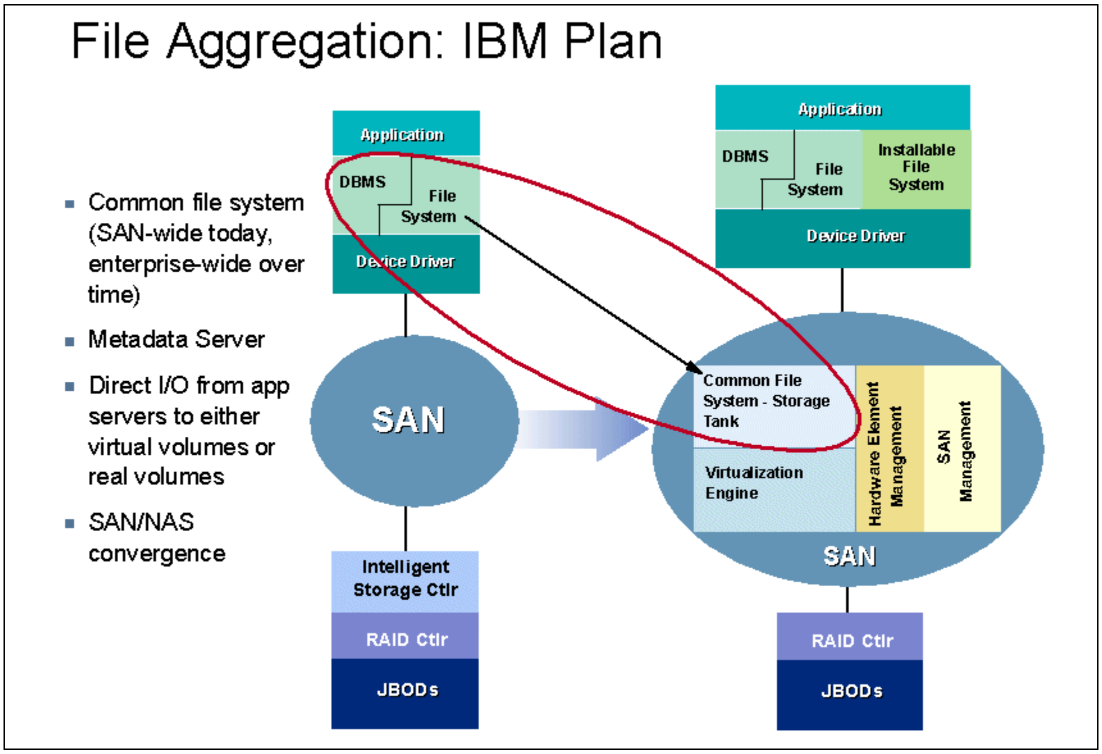


Figure 1-5 The IBM plan for file aggregation

The SAN file system metadata servers are based on clustered IBM SAN appliances running Linux to support high availability and performance. The metadata servers provide file locks and all other file information (such as location) to authorized application servers, which are running the SAN file system client code (no application changes necessary).

After file information is passed from the metadata server to the application server, the application server can access the blocks that comprise that file directly through the SAN (not through the metadata server). By providing direct access to data from the application server to the underlying storage (virtualized or not), our solution provides the benefits of heterogeneous file sharing with local file system performance.

Since the metadata servers have a complete understanding of all files on the SAN, including the essential metadata to make important decisions, it is a logical point to manage the storage in the network through policy-based controls. For example, when news files are created, the metadata server can decide where to place each file based on specified criteria such as file type.

The SAN file system metadata server provides the ability to group storage devices according to their characteristics, such as latency and throughput. These groupings, called *storage pools*, allow administrators to manage data according to the characteristics that matter to them. For example, an administrator can define a storage pool for mission critical applications using highly reliable storage arrays that are backed up nightly and have full disaster recovery capabilities. The administrator can also define a storage pool for less critical applications based on JBODs with weekly tape backups and minimal disaster recovery capabilities. Using this level of storage classification, an administrator can set up automated policies that determine which files are placed in which storage pools based on the required service levels.

Because the SAN file system metadata is separate from the application data, files can be manipulated while remaining active. For example, files being processed by a mission-critical application can be non-disruptively moved within or across storage pools without stopping the application. Similarly, data migration from one storage system to another can be handled non-disruptively by having the metadata server move the pools to new physical disks, and then disconnecting the old disks, all done without quiescing applications.

The SAN file system offers a logical extension to current NAS and SAN environments. While NAS has proven successful in the marketplace, it is does not take advantage of a SAN infrastructure. NAS Filers become the keepers of their own file metadata and must be managed separately from the files on SAN attached application servers. Our approach is to add NAS capabilities to the SAN file system thereby allowing storage administrators to manage the NAS file data with the same tools as for their application servers. This approach of SAN and NAS convergence helps lower TCO in these environments.

To facilitate the adoption of the SAN file system, the client code and the client reference implementation source code are licensed at no cost. In addition, the metadata server protocols are made publicly available. We will work with the industry to encourage convergence of the SAN file system protocols with other standards.

In summary, the IBM TotalStorage SAN file system is a common SAN-wide file system that permits centralization of management and improved storage utilization at the file level. The SAN file system is delivered in a highly available configuration based on an IBM SAN appliance with active-active failover and clustering for the metadata servers, providing high availability and fault tolerance. The SAN file system is also being designed to provide policy based storage automation capabilities for provisioning and data placement, nondisruptive data migration, and a single point of management for files on a storage network. The use of the SAN file system can greatly simplify the management of files on SANs and result in a significant reduction in TCO.

Figure 1-6 shows a diagram of the IBM SAN file system architecture.

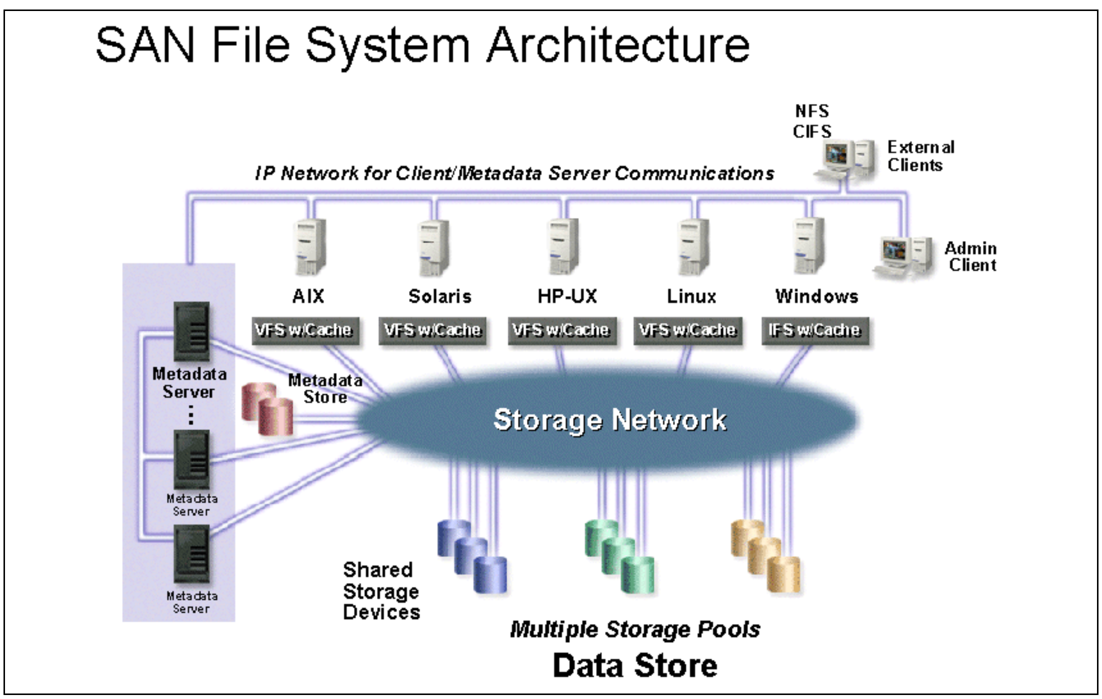


Figure 1-6 Conceptual diagram of the IBM SAN file system architecture

1.4 Conclusion

In conclusion, IBM TotalStorage SAN Volume Controller enables storage virtualization. This allows clients to reap the benefit of better application business responsiveness, maximized storage utilization, dynamic resource allocation, improved storage administration utilization and reduced storage outage.

Symmetrical and asymmetrical virtualization provide two very distinct yet complementary approaches to virtualization. IBM will extol the virtues of each in two separate products. Both products fulfill different requirements, and therefore, use different approaches to virtualization.

The rest of this redbook is dedicated to the IBM TotalStorage SAN Volume Controller and its method of symmetrical virtualization. For greater detail about the technology and implementation of the IBM TotalStorage SAN file system, see *IBM TotalStorage: Introducing the SAN File System*, SG24-7057.

2

IBM TotalStorage SAN Volume Controller overview

This chapter describes the major concepts behind the IBM TotalStorage SAN Volume Controller to provide the framework for discussion for the remainder of this redbook. We begin the chapter with a glossary of terms that are used.

2.1 Glossary of commonly used terms

We begin this chapter with a short glossary of terms commonly used throughout the remainder of this redbook.

Boss node

A single node acts as the boss node for overall management of the cluster. If the boss node fails, another node in the cluster will take over the responsibilities.

Configuration node

At any one time, a single node in the cluster is used to manage configuration activity. This configuration node manages a cache of the configuration information that describes the cluster configuration and provides a focal point for configuration commands. Similarly, at any one time, a single node acts as the boss node for overall management of the cluster.

Extent

An extent is a fixed size unit of data that is used to manage the mapping of data between mDisks and vDisks.

Front-end and back-end

SAN Volume Controller takes managed disks and represents these to application servers (hosts). The managed disks are looked after by the “back-end” application of the SAN Volume Controller. The virtual disks presented to hosts are looked after by the “front-end” application in the SAN Volume Controller.

Grain

A grain is the unit of data represented by a single bit in a FlashCopy bitmap, 256 K in SAN Volume Controller.

I/O group

An input/output (I/O) group contains two SAN Volume Controller nodes defined by the configuration process. Each SAN Volume Controller node is associated with exactly one I/O group. The nodes in the I/O group provide access to the vDisks in the I/O group.

LU and LUN

Strictly speaking, there is a difference between an logical unit (LU) and a logical unit number (LUN). A LUN is a unique identifier used on a SCSI bus that enables it to differentiate between up to eight separate devices (each of which is a logical unit). In practice, the two terms are used interchangeably. In this book, when we refer to a LUN, we refer to the unit of storage that is defined in a storage subsystem such as an IBM TotalStorage Enterprise Storage Server® (ESS) or IBM TotalStorage Fibre Array Storage Technology (FAStT) server.

Managed disk

Managed disk (mDisk) is a SCSI disk presented by a RAID controller and managed by the SAN Volume Controller. The mDisk is not visible to host systems on the SAN. In some of the code examples shown in this book, the mDisk is referred to as a VLUN.

Managed disk group

The managed disk group (MDG) is a collection of mDisks that jointly contain all the data for a specified set of vDisks.

Master console

The master console is the platform on which the software used to manage the SAN Volume Controller and SAN Integration Server (SIS) runs.

Node

Node is the name given to the individual servers in an SAN Volume Controller cluster on which the SAN Volume Controller software runs.

SAN Volume Controller

The SAN Volume Controller is a SAN appliance designed for attachment to a variety of host computer systems, which carries out block level virtualization of disk storage.

SAN Integration Server

The SIS is a pre-packaged system comprising an SAN Volume Controller cluster, back-end storage, SAN and Ethernet switches, and master console assembled and pre-configured in a rack. Expansion racks are available to accommodate additional back-end storage capacity.

Virtual disk

Virtual disk (vDisk) is a SAN Volume Controller device that appears to host systems attached to the SAN as a SCSI disk. Each vDisk is associated with exactly one I/O group.

2.2 Virtualization overview

The IBM TotalStorage SAN Volume Controller Storage nodes are the hardware elements of the IBM TotalStorage SAN Volume Controller, a member of the IBM TotalStorage virtualization family of solutions. The SAN Volume Controller combines servers into a high availability cluster. Each of the servers in the cluster is populated with 4 GB of high-speed memory, which serves as the cluster cache. A management card is installed in each server to monitor various parameters which the cluster uses to determine the optimum and continuous data path. The cluster is protected against data lose by uninterruptible power supplies. The SAN Volume Controller Storage nodes are always installed in pairs for high availability.

Storage virtualization addresses the increasing cost and complexity in data storage management. It addresses this increased complexity by shifting storage management intelligence from individual SAN controllers into the network via a virtualization cluster of nodes.

The SAN Volume Controller solution is designed to reduce both the complexity and costs of managing your SAN-based storage. With the SAN Volume Controller, you can:

- Simplify management and increase administrator productivity by consolidating storage management intelligence from disparate storage controllers into a single view.
- Improve application availability by enabling data migration between disparate disk storage devices non-disruptively.
- Improve disaster recovery and business continuance needs by applying and managing copy services across disparate disk storage devices within the Storage Area Network (SAN). These solutions include a Common Information Model (CIM) Agent, enabling unified storage management based on open standards for units that comply with CIM Agent standards.
- Provides advanced features and functions to the entire SAN, such as:
  - Large scalable cache
  - Copy Services

– Space management (later releases to include Policy Based Management)
– Mapping based on desired performance characteristics
– Quality of Service (QoS) metering and reporting

Note: The SAN Volume Controller is not a RAID controller. The disk subsystems attached to SANs that have the SAN Volume Controller provide the basic RAID setup. The SAN Volume Controller uses what is presented to it as a managed disk to create virtual disks.

2.3 Compass architecture

The IBM TotalStorage SAN Volume Controller is based on the COMmodity PArts Storage System (Compass) architecture developed at the IBM Almaden Research Center.

The overall goal of the Compass architecture is to create storage subsystem software applications that require minimal porting effort to leverage a new hardware platform. To meet this goal:

- Compass, although currently deployed on the Intel® hardware platform, can be ported to other hardware platforms.
- Compass, although currently deployed on a Linux kernel, can be ported to other Portable Operating System Interface (POSIX)-compliant operating systems.
- Compass uses commodity adapters and parts wherever possible. To the highest extent possible, it only uses functions in the commodity hardware that are commonly exercised by the other users of the parts. This is not to say that Compass software could not be ported to a platform with specialized adapters. However, the advantage in specialized function must be weighed against the disadvantage of future difficulty in porting and in linking special hardware development plans to the release plans for applications based on the Compass architecture.
- Compass is developed in such a way that it is as easy as possible to troubleshoot and correct software defects.
- Compass is designed as a scalable, distributed software application that can run in increasing sets of Compass nodes with near linear gain in performance while using a shared data model that provides a single pool of storage for all nodes.
- Compass is designed so that there is a single configuration and management view of the entire environment regardless of the number of Compass nodes in use.

The approach is to minimize the dependency on unique hardware, and to allow exploitation of or migration to new SAN interfaces simply by plugging in new commodity adapters. Performance growth over time is assured by the ability to port Compass to just about any platform and remain current with the latest processor and chipset technologies on each. The SAN Volume Controller implementation of the Compass architecture has exploited Linux as a convenient development platform to deploy this function. This has, and will continue to enhance the ability of IBM to deploy robust function in a timely way.

2.3.1 SAN Volume Controller clustering

In simple terms, a cluster is a collection of servers that, together, provide a set of resources to a client. The key point is that the client has no knowledge of the underlying physical hardware of the cluster. This means that the client is isolated and protected from changes to the physical hardware, which brings a number of benefits. Perhaps the most important of these benefits is high availability. Resources on clustered servers act as highly available versions of unclustered resources. If a node (an individual computer) in the cluster is unavailable, or too busy to respond to a request for a resource, the request is transparently passed to another node capable of processing it. Clients are, therefore, unaware of the exact locations of the resources they are using. For example, a client can request the use of an application without being concerned about either where the application resides or which physical server is processing the request. The user simply gains access to the application in a timely and reliable manner. Another benefit is scalability. If you need to add users or applications to your system and want performance to be maintained at existing levels, additional systems can be incorporated into the cluster.

The IBM TotalStorage SAN Volume Controller is a collection of up to four cluster nodes, added in pairs. In future releases, the cluster size will be increased to permit further performance scalability. These four nodes are managed as a set (cluster) and present a single point of control to the administrator for configuration and service activity.

> **Note:** Although the SAN Volume Controller code is based on a Linux kernel, the clustering feature is not based on Linux clustering code. The clustering failover and failback feature is part of the SAN Volume Controller application software.

Within each cluster, a node is defined as the configuration node. This node is assigned the cluster IP address and is responsible for transitioning additional nodes into the cluster.

During normal operation of the cluster the nodes keep in touch with each other. If a node is idle for a while (a few seconds), then a heartbeat signal is sent to assure connectivity with the cluster. Should a node fail for any reason, the workload intended for it is taken over by another node until the failed node has been restarted and re-admitted to the cluster (which happens automatically). In the event that the microcode on a node becomes corrupted, resulting in a failure, the workload is transferred to another node, the code on the failed node is repaired, and the node is re-admitted to the cluster (again, all automatically).

For I/O purposes, SAN Volume Controller nodes within the cluster are grouped into pairs, called *I/O groups*, with a single pair being responsible for serving I/O on a given vDisk. One node within the I/O group represents the preferred path for I/O to a given vDisk. The other node represents the non-preferred path. This preference alternates between nodes as each vDisk is created within an I/O group to balance the workload evenly between the two nodes.

> **Note:** The preferred node by no means signifies absolute ownership. The data is still accessed by the partner node in the I/O group in the event of a failure or if the preferred node workload becomes too high.

Beyond automatic configuration and cluster administration, the data transmitted from attached application servers is also treated in the most reliable manner. When data is written by the host, the preferred node within the I/O group stores a write in its own write cache and the write cache of its partner (non-preferred) node before sending an “I/O complete” status back to the host application. To ensure that data is written in the event of a node failure, the surviving node empties its write cache and proceeds in write-through mode until the cluster is returned to a fully operational state.

> **Note:** Write-through mode is where the data is not cached in the nodes, but written directly to the disk subsystem instead. While operating in this mode, performance is somewhat degraded. However more importantly, it ensures that the data makes it to its destination without the risk of data loss to which a single copy of data in cache would expose you.

As yet another data protection feature, the SAN Volume Controller is supplied with a pair of uninterruptible power supply units. In addition to voltage regulation to protect valuable electronic components within the SAN Volume Controller configuration, in the event of a main power outage, the uninterruptible power supply provides enough power to destage data to the SAN Volume Controller internal disk and shut down the nodes within the SAN Volume Controller cluster gracefully. This is a feature found in most high-end disk subsystems.

2.3.2 SAN Volume Controller virtualization

The SAN Volume Controller provides block aggregation and volume management for disk storage within the SAN. In simpler terms, this means that the SAN Volume Controller manages a number of back-end storage controllers and maps the physical storage within those controllers to logical disk images that can be seen by application servers and workstations in the SAN. The SAN is zoned in such a way that the application servers cannot see the back-end storage, preventing any possible conflict between SAN Volume Controller and the application servers both trying to manage the back-end storage.

As described earlier, when an application server performs I/O to a vDisk assigned to it by the SAN Volume Controller, it can access that vDisk via either of the nodes in the I/O group. Each node may only be in one I/O group and since each I/O group only has two nodes, the distributed redundant cache design in the SAN Volume Controller only needs to be two-way.

The SAN Volume Controller I/O groups are connected to the SAN in such a way that all back-end storage and all application servers are visible to all of the I/O groups. The SAN Volume Controller I/O groups see the storage presented to the SAN by the back-end controllers as a number of disks, known as *managed disks*. Because the SAN Volume Controller does not attempt to provide recovery from physical disk failures within the back-end controllers, mDisks are usually, but not necessarily, part of a RAID array. The application servers do not see the mDisks at all. Instead, they see a number of logical disks, known as virtual disks or vDisks, which are presented to the SAN by the SAN Volume Controller.

mDisks are collected into groups, known as *managed disk groups*. The mDisks that are used in the creation of a particular vDisk must all come from the same MDG. Each mDisk is divided up into a number of extents (default minimum size 16 MB, maximum size of 512 MB), which are numbered sequentially from the start to the end of each mDisk. Conceptually, this is represented as shown in Figure 2-1.

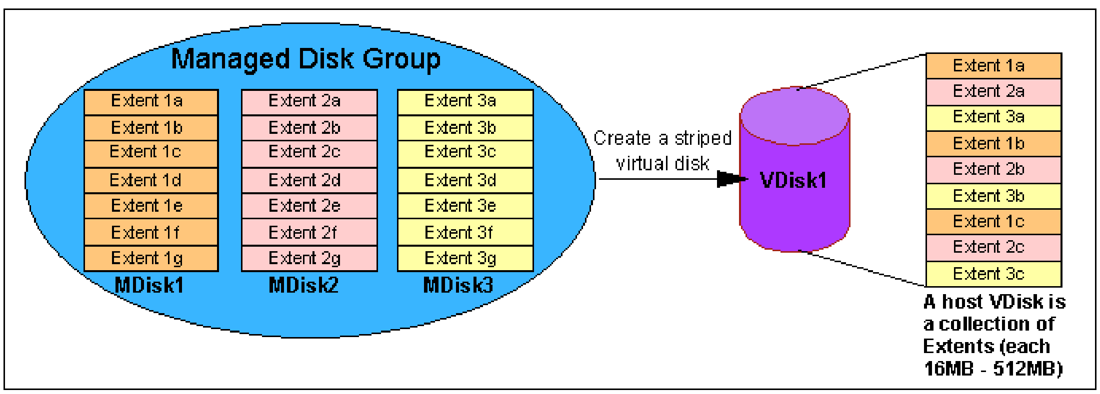


Figure 2-1 Extents being used to create a virtual disk

The virtualization function in the SAN Volume Controller maps the vDisks seen by the application servers on to the mDisks provided by the back-end controllers. I/O traffic for a particular vDisk is, at any one time, handled exclusively by the nodes in a single I/O group. Although a cluster can have many nodes within it, the nodes handle I/O in independent pairs. This means that the I/O capability of the SAN Volume Controller scales well (almost linearly), since additional throughput can be obtained by simply adding additional I/O groups.

Figure 2-2 summarizes the various relationships that bridge the physical disks through to the virtual disks within the SAN Volume Controller architecture.

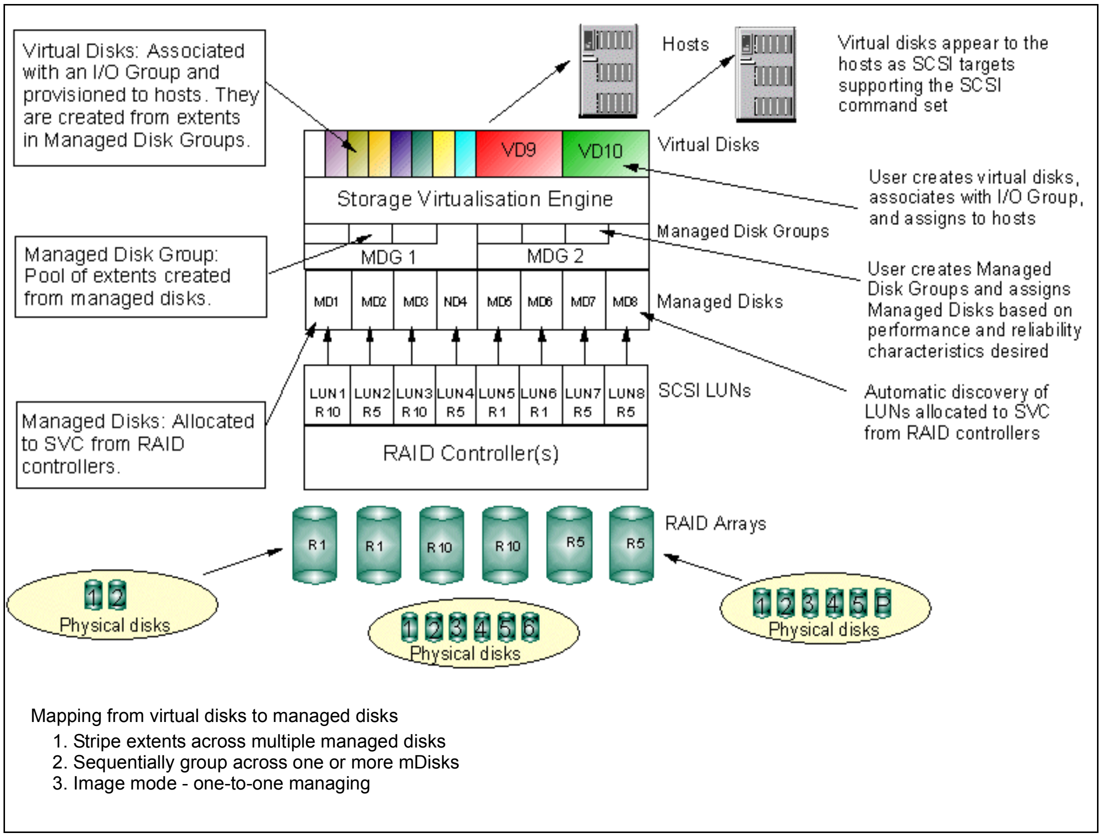


Figure 2-2 The relationship between physical and virtual disks

Virtualization mappings

Several different mapping functions are provided by the SAN Volume Controller:

- **Striped**: Here a vDisk is mapped to a number of mDisks in a MDG. The extents on the vDisk are striped over the mDisks. Therefore, if the vDisk is mapped to 5 mDisks, then the first, sixth, eleventh, etc. extent comes from the first mDisk; the second, seventh, twelfth (etc) extent will come from the second mDisk; and so on. This is the default mapping.
- **Sequential**: Here a vDisk is mapped to a single mDisk in a MDG. There is no guarantee that sequential extents on the mDisk map to sequential extents on the vDisk, although this may be the case when the vDisk is created.

> **Note:** There are no ordering requirements in the mDisk to vDisk extent mapping function for either striped or sequential vDisks. This means that if you examine the extents on an mDisk, it is quite possible for adjacent extents to be mapped to different vDisks. It is also quite possible for adjacent extents on the mDisk to be mapped to widely seperated extents on the same vDisk, or to adjacent extents on the vDisk. In addition, the position of the extents on the mDisks is not fixed by the initial mapping, and can be varied by the user performing data migration operations.

- **Image**: Image mode sets up a one-to-one mapping of extents on an mDisk to the extents on the vDisk. Because the vDisk has exactly the same extent mapping as the underlying mDisk, any data already on the disk is still accessible when migrated to an SAN Volume Controller environment. Within the SAN Volume Controller environment, the data can (optionally) be seamlessly migrated off the image mode vDisk to a striped or sequential vDisk within a MDG.

2.3.3 SAN Volume Controller multipathing

Each SAN Volume Controller node presents a vDisk to the SAN via multiple paths, usually four. In normal operation, two nodes provide redundant paths to the same storage. This means that, depending on zoning, a single host bus adapter (HBA) sees up to eight paths to each LUN presented by the SAN Volume Controller. Because most operating systems cannot resolve multiple paths back to a single physical device, IBM provides a multipathing device driver.

The multipathing driver supported by the SAN Volume Controller is the IBM Subsystem Device Driver (SDD). It manages the multiple paths from the host to the SAN Volume Controller making use of the preferred paths in a round robin manner before using any non-preferred path. SDD performs data path failover in the event of a failure within the SAN Volume Controller, or the host path while masking out the additional disks that is otherwise seen by the hosts due to the redundant paths through the SAN fabric.

FYI: The SDD code is updated to support both the SAN Volume Controller and the ESS. Provided that the latest version is used, IBM supports the concurrent connections of a host to both an SAN Volume Controller and native ESS environment.

2.4 SAN Volume Controller logical configuration

Figure 2-3 shows an example of a SAN Volume Controller configuration.

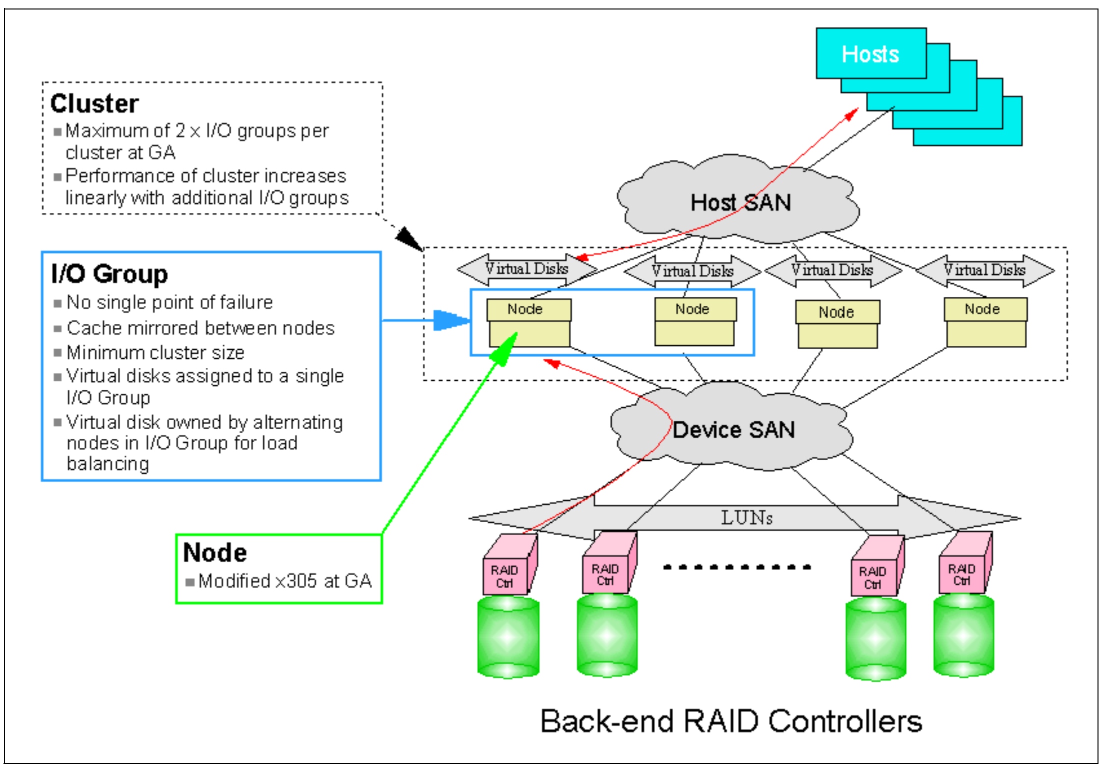


Figure 2-3 SAN Volume Controller logical view

Configuration notes

- The Fibre Channel SAN connections between the SAN Volume Controller and the switches are optical fibre running at 2 Gb/s. However, the SAN Volume Controller is also supported in 1 Gb/s Fibre Channel fabrics.
- To provide high availability, the SAN Volume Controller nodes should be configured in redundant SAN fabrics.
- The Fibre Channel switches need to be zoned to permit the hosts to see the SAN Volume Controller nodes and the SAN Volume Controller nodes to see the RAID Controllers. The SAN Volume Controller nodes within a cluster must be able to see each other and the master console must be able to see everything in the SAN Volume Controller environment. In addition, if there are two SAN Volume Controller clusters with Remote Copy services between them, zoning must be set so the two clusters see each other.
- In addition to a Fibre Channel connection or connections, each device has an Ethernet connection for configuration and error reporting, though only one of the nodes, the configuration node, binds an IP address to its Ethernet connection.

2.5 SAN Volume Controller compatibility

The SAN Volume Controller is capable of supporting Windows® NT® and 2000, AIX®, Red Hat Linux, Sun Solaris, and HP-UX hosts. The SAN switch support is also broad and includes the IBM TotalStorage SAN Switches and members of the McData, CNT, and Cisco families of SAN Switches and Directors.

Initial support of disk subsystems was limited to the IBM TotalStorage FAStT Storage Server and the IBM TotalStorage Enterprise Storage Server. It was further extended to support Hitachi Freedom Storage Thunder 9200, Hewlett-Packard StorageWorks Modular Array 8000 (single path), Hewlett-Packard StorageWorks Enterprise Modular Array 12000 (single path), and Hewlett-Packard StorageWorks Enterprise Modular Array 16000 (single path). Future releases of the SAN Volume Controller will include support for other third-party disk subsystems from manufacturers such as HDS, HPQ and EMC.

For the latest supported operating systems, hosts, HBAs, SAN switches and storage subsystems, see the supported hardware list for SAN Volume Controller on the Web at:

http://www.storage.ibm.com/support/2145

2.6 Software licensing

There are three parameters to consider when licensing the SAN Volume Controller software:

- The total amount of storage that is managed by the SAN Volume Controller cluster: This may be greater than the amount of storage that is virtualized by the cluster. See Figure 2-4.

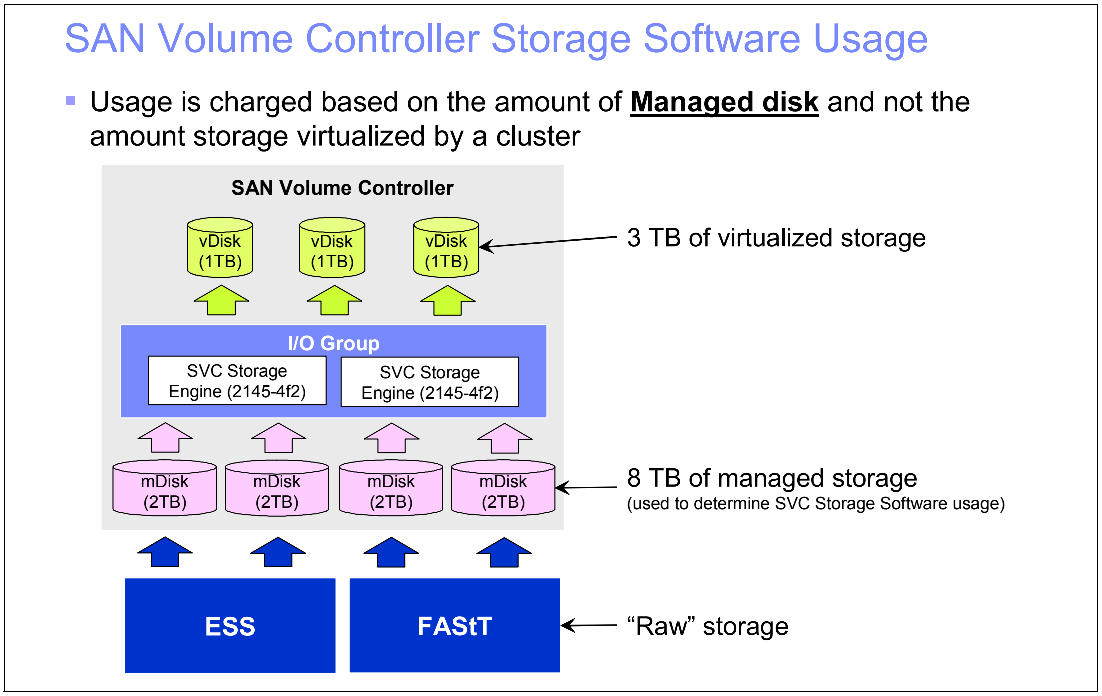


Figure 2-4 Base software license diagram

- The amount of virtualized storage that you want to FlashCopy: You may want to manage 8 TB of storage, but FlashCopy only 1 TB of storage. In this case, you need a 2 TB FlashCopy license since there is 1 TB of source and 1 TB of target FlashCopy volumes. See Figure 2-5.

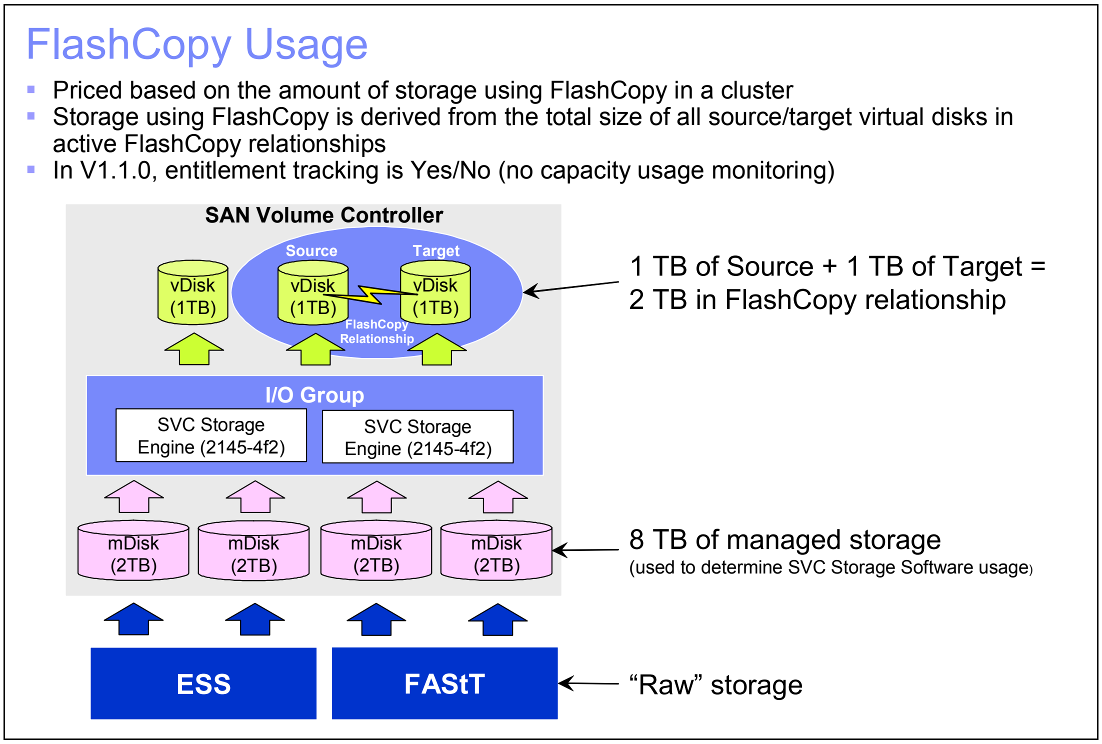


Figure 2-5 FlashCopy storage license

- The amount of virtualized storage that you want to Remote Copy to another system: You may be managing 8 TB of storage and FlashCopy copying 1 TB (2 TB Fibre Channel license required) and Remote Copy copying 1.5 TB. In this case, both clusters that are taking part in the Remote Copy relationship require a 1.5 TB Peer-to-Peer Remote Copy (PPRC) license. In the case of intracluster Remote Copy, where both primary and secondary volumes are in the same cluster, then the license must be large enough to cover both.

  See Figure 2-6 for the intracluster scenario. For intercluster PPRC, only the amount of virtualized storage that is in a PPRC relationship at that site needs to be covered by the PPRC license.

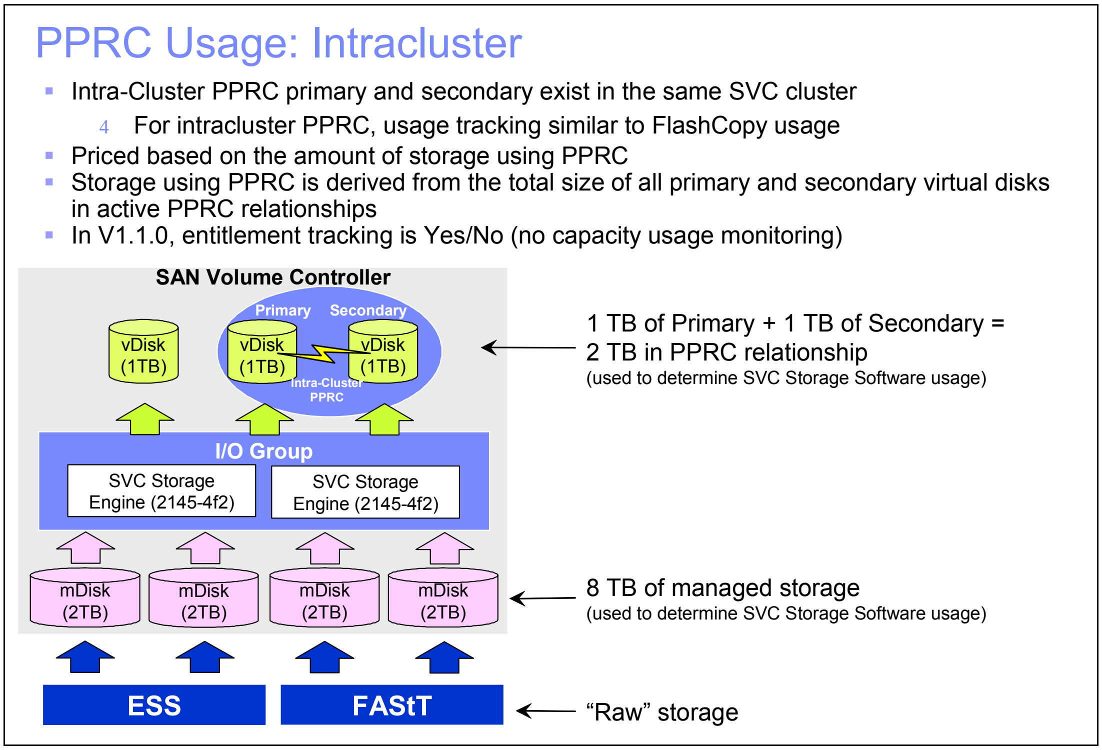


Figure 2-6 PPRC intracluster

Figure 2-7 shows the intercluster PPRC relationship and how this affects licensing.

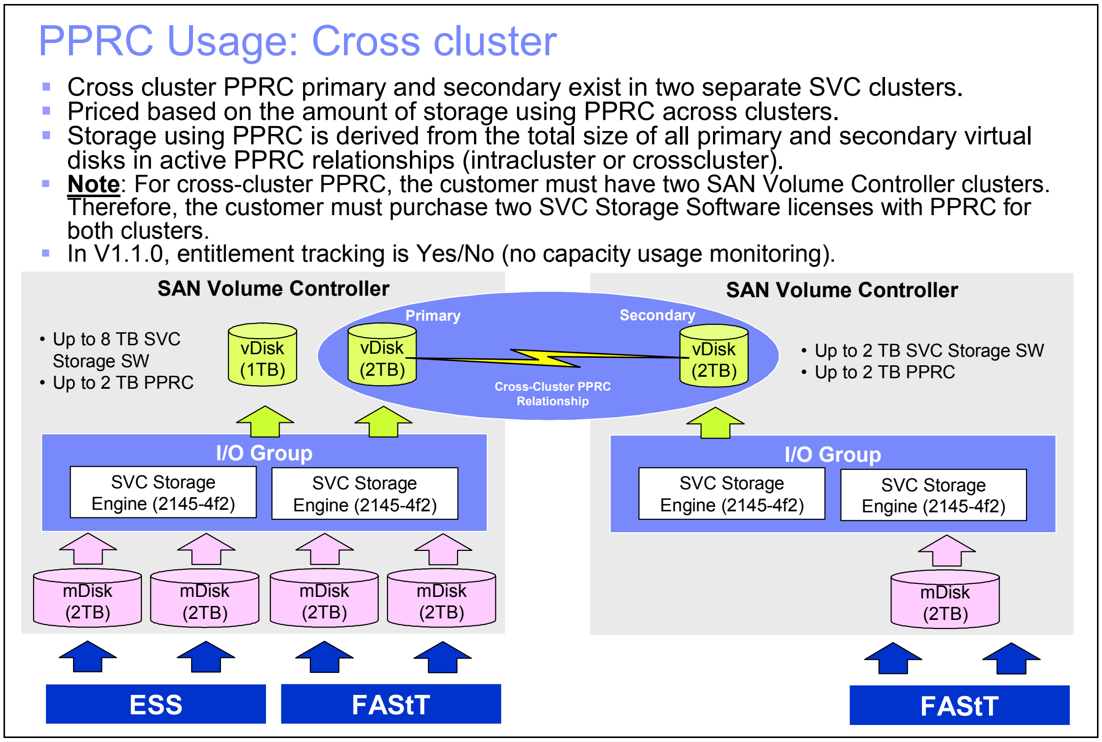


Figure 2-7 PPRC intercluster relationship

You can increase any one of these three licenses independently of the other. That is you can increase the total amount of managed storage without increasing the other licenses if the amounts of storage being copied remains unchanged. Similarly you can change the copy licenses independently of each other.

3

Planning and configuration

This chapter discusses the steps required when planning to install an IBM TotalStorage SAN Volume Controller (SVC) in your storage network. It looks at the implications for your storage network.

3.1 General planning rules

To achieve the most benefit from SVC, pre-installation planning should consider several important steps. These steps ensure that SVC provides the best possible performance, reliability, and ease of management for your application needs. Proper configuration also helps minimize downtime by avoiding changes to SVC and the storage area network (SAN) environment to meet future growth needs.

Planning the SVC requires the following steps:

1. Document the number of hosts (application servers), the traffic profile activity (read or write, sequential or random), the performance requirements (input/output (I/O) per seconds), the total storage capacity, the storage capacity per host, and the host logical unit number (LUN) sizes.
2. Define the local and remote SAN fabrics and clusters if a remote copy or a secondary site are needed.
3. Define the number of clusters and the number of pair of nodes (between 1 and 2) for each site. Each pair of nodes (an I/O group) is the container for the virtual disks. It depends on the overall performance needs.
4. Design the SAN according to the requirement for high availability. Consider the total number of ports and the bandwidth needed between the host to the SVC, the SVC to the back-end storage, between the SVC nodes, and for the ISL between the local and remote fabric.
5. Define the managed disks (mDisks) in the back-end storage (the RAID controller).
6. Define the managed disk groups (MDGs). It depends on the back-end storage in place and the needs for data migration.
7. Create and repartition the vDisks between the different I/O groups and the different MDGs in a way to optimize the I/O load between the hosts and the SVC. It can be an equal re-partition of all the vDisks between the different nodes or a re-partition, which takes into account the expected load from the different hosts.
8. Plan for the physical location of the equipment in the rack.

> **Note:** See *IBM TotalStorage Virtualization Family SAN Volume Controller: Planning Guide*, GA22-1052, for the hardware location and connection charts.

You need IP addresses for the SVC, the service IP address, master console, and switches. See the planning guide mentioned in the previous note box.

3.2 Physical planning

There are main factors to take into account when carrying out the physical planning of an SCVC/SIS installation. The physical site must have the following characteristics:

- Power, cooling, and location requirements for the SVC and uninterruptible power supplies.
- An SVC node is one EIA unit high.
- The uninterruptible power supply is two EIA units.
- The master console is two EIA units: one for the server and one for the keyboard and monitor.

- Other hardware devices may be in the rack such as IBM TotalStorage Fibre Array Storage Technology (FAStT), switches, Ethernet switch, and others.
- The maximum power rating of the rack and input power supply must not be exceeded.

3.2.1 Physical rules

The SVC must be installed in pairs to provide high availability. See Figure 3-1.

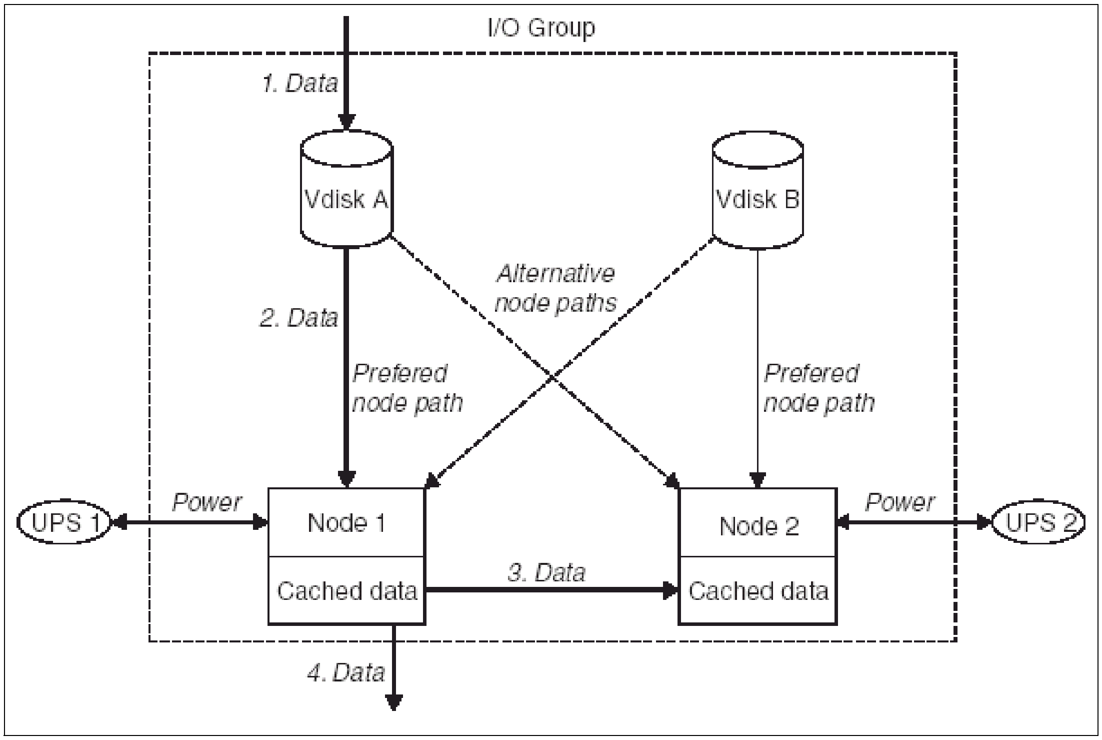


Figure 3-1 Node uninterruptible power supply setup

- Each SVC node of an I/O pair must be connected to a different uninterruptible power supply.
- Each uninterruptible power supply can support up to four SVC nodes.
- Each uninterruptible power supply that supports a pair of nodes must be connected to a different power domain (if possible) to reduce the chances of input power loss.
- The uninterruptible power supplies must be installed in the lowest available position in the rack. If necessary, move lighter units toward the top.
- A cluster can contain no more than four SVC nodes. (This increases in later releases of the product).

Figure 3-2 shows the layout within a rack.

EIA 36	Blank
EIA 35	Ethernet Hub 1
EIA 34	Blank
EIA 33	Blank
EIA 32	Blank
EIA 31	Blank
EIA 30	Blank
EIA 29	Blank
EIA 28	FC Switch 1
EIA 27	FC Switch 1
EIA 26	Blank
EIA 25	Blank
EIA 24	Blank
EIA 23	Blank
EIA 22	SAN Volume Controller 4
EIA 21	SAN Volume Controller 3
EIA 20	SAN Volume Controller 2
EIA 19	SAN Volume Controller 1
EIA 18	master console
EIA 17	Master console keyboard and monitor
EIA 16 EIA 15 EIA 14	RAID Controller 4
EIA 13 EIA 12 EIA 11	RAID Controller 3
EIA 10 EIA 9 EIA 8	RAID Controller 2
EIA 7 EIA 6 EIA 5	RAID Controller 1
EIA 4 EIA 3	Uninterruptible power supply 2
EIA 2 EIA 1	Uninterruptible power supply 1

Figure 3-2 Sample rack layout

3.2.2 Cable connections

Complete a cable connection table to document all of the connections required for the setup:

- Nodes
- Uninterruptible power supply
- Ethernet
- Fibre Channel ports
- Master console

A typical planning chart is shown in Figure 3-3.

Node number	Uninterruptible power supply	Power domain	Ethernet		FC Port-1	FC Port-2	FC Port-3	FC Port-4
			SAN Volume Controller port	Hub or Switch				

Master console	Ethernet		FC Port-1	FC Port-2
	Public network	VPN		

Figure 3-3 Cable connection table

3.3 SAN planning and configuration

SAN storage systems using SVC may be configured with two or four SVC nodes, arranged in a cluster. These are attached to the SAN fabric, along with RAID controllers, and host systems. The SAN fabric is zoned to allow the SVCs to “see” the RAID controllers, and for the hosts to “see” the SVCs. The hosts are not able to directly “see” or operate on the RAID controllers. The SVC nodes within a cluster must be able to see each other.

The zoning capabilities of the SAN switch are used to create these distinct zones. See Figure 3-4. The SVC in Release 1 supports 1 Gb/s or 2 Gb/s Fibre Channel fabric only.

A cluster of SVC nodes are connected to the same fabric and present virtual disks to the hosts. These virtual disks are created from managed disks presented by the RAID controllers. There are two distinct zones shown in the fabric:

- A host zone in which the hosts can see and address the SVC nodes
- A disk zone in which the SVC nodes can see and address the LUNs presented by the RAID controllers

Hosts are not permitted to operate on the RAID LUNs directly, and all data transfer happens through the SVC nodes. Under some circumstances, a RAID controller may present LUNs to both the SVC (as managed disks, which it then virtualizes to hosts) and to other hosts in the SAN. There are some configuration limitations as to how this may be achieved.

This is commonly described as *symmetric virtualization*. Figure 3-4 shows the data flow across the physical topology.

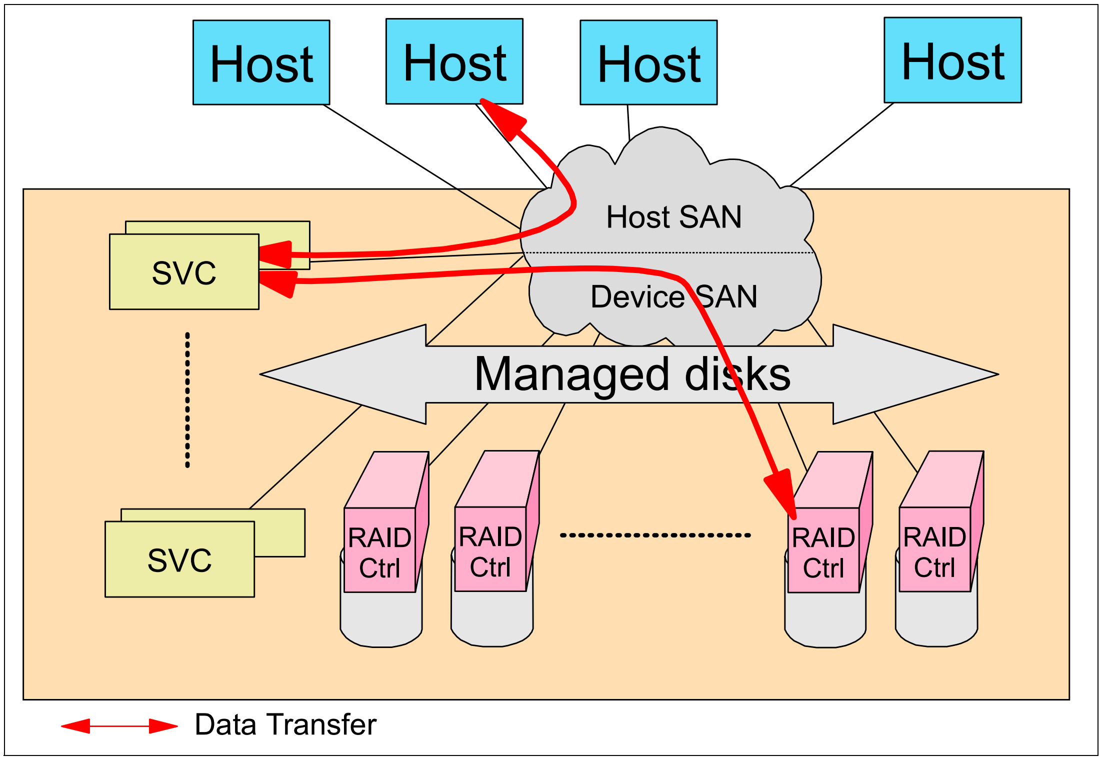


Figure 3-4 SVC physical topology

Logically, the two zones can be thought of as two separate logical SANs, leading to the diagram shown in Figure 3-5.

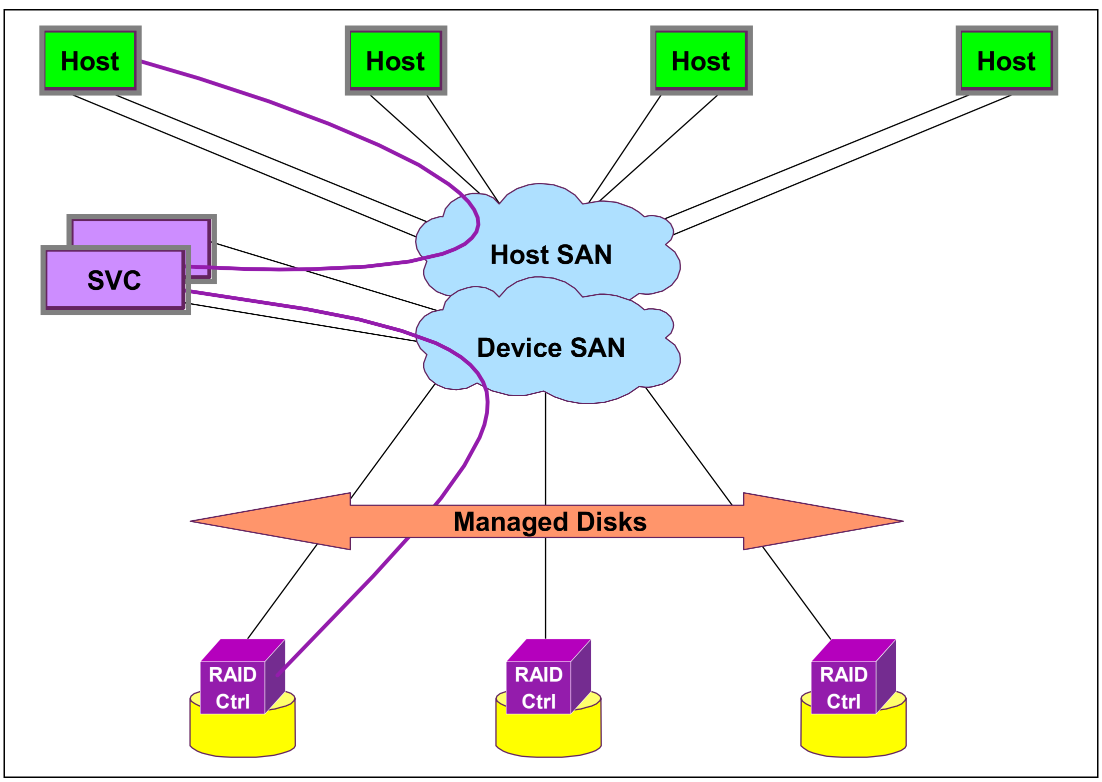


Figure 3-5 SVC logical topology

3.3.1 SAN definitions

The following definitions are used in this section.

ISL hop

This is a “hop” on an interswitch link (ISL). ISL hops are the number of ISLs traversed on the shortest route between the pair of nodes that are farthest apart. They consider all pairs of N-ports (end node) in a fabric. It measures distance only in terms of ISLs in the fabric.

Oversubscription

Oversubscription is the ratio of the sum of the traffic on the initiator N-port connection or connections to the traffic on the most heavily loaded ISL or ISLs where more than one is in parallel between these switches. This is assuming a symmetrical network and given a specific workload applied evenly from all initiators and directed evenly to all targets.

Symmetrical means that all the initiators are connected at the same level, and all the controllers are connected at the same level. The SVC node makes this calculation more interesting. It places its back-end traffic onto the same network. This back-end traffic varies by workload, so 100% read hit gives a different oversubscription to 100% write miss.

Redundant SAN

A redundant SAN is a SAN configuration in which any single component may fail. Connectivity between the devices within the SAN is maintained possibly with degraded performance. This is normally achieved by splitting the SAN into two independent counterpart SANs (two SAN fabrics).

Counterpart SAN

A counterpart SAN is a non-redundant portion of a redundant SAN. A counterpart SAN provides all the connectivity of the redundant SAN, but without the redundancy. An SVC node is typically connected to a redundant SAN made out of two counterpart SANs.

Local fabric

Since the SVC supports remote copy, there may be significant distances (up to 10 km in the first release of the product) between the components in the local cluster and those in the remote cluster. The local fabric comprises those SAN components (switches, cables, etc.), which connect the components (nodes, hosts, switches) of the local cluster together.

Remote fabric

Since the SVC supports remote copy, there may be significant distances (up to 10 km in the first release of the product) between the components in the local cluster and those in the remote cluster. The remote fabric comprises those SAN components (switches, cables, etc.) which connect the components (nodes, hosts, switches) of the remote cluster together.

Local and remote fabric interconnect

These are the SAN components that are used to connect the local and remote fabrics together. They may be simply be single mode optical fibers driven by high power GBICs as in SVC node release 1. Or they may be other more sophisticated components (channel extenders for example) in future releases of the SVC.

Fibre Channel port fan in

This is the number of hosts that can see any one SVC node port. Some controllers, such as the IBM TotalStorage Enterprise Storage Server (ESS) for example, recommend limiting the number of hosts that use each port, to prevent excessive queuing at that port. Clearly, if the port fails or the path to that port fails, the host may failover to another port and the fan in criteria may be exceeded in this degraded mode.

Channel extender

A channel extender is a device for long distance communication connecting other SAN fabric components. Generally, these may involve protocol conversion to asynchronous transfer mode (ATM) or Internet Protocol (IP) or some other long distance communication protocol.

3.3.2 General SAN design guidelines with SVC

> **Note:** The SVC is not a RAID controller.

To ensure high availability in SVC installations, the following rules apply when you design a SAN with the SVC.

For the SVC intracluster

The following general guidelines apply:

- An SVC node always contains two host bus adapters (HBAs), each of which has two Fibre Channel (FC) ports. If an HBA fails, this remains a valid configuration, and the node operates in degraded mode. If an HBA is physically removed from a SVC node, then the configuration is unsupported.
- To maintain application uptime in the unlikely event of an individual SVC node failing, SVC nodes are always deployed in pairs. If a node fails or is removed from the configuration, the remaining node operates in a degraded mode, but is still a valid configuration. The remaining mode operates in write through mode (the cache is disabled for the write).
- The uninterruptible power supply must be in the same rack as the nodes.
- The Fibre Channel SAN connections between the SVC node and the switches are optical fiber. The connections between the switches and the SVC nodes run at either 1 Gb/s or 2 Gb/s. However, all of the FC ports on SVC nodes in a single cluster must run at one speed. Operation with different speeds running on the node to switch connections in a single cluster is illegal.
- The two nodes within an I/O group shall be within 100 m of each other.

For the SVC intercluster

Two SVC clusters may not share the same back end storage array. The consequences of sharing the same array can result in data loss. If the same mDisk becomes visible on two different SVC, then this is a usage error that can cause data corruption.

For the SAN fabric

The following guidelines apply:

- On the FC SAN, the SVC nodes are always connected to SAN switches and nothing else. Each node must be connected to each of the counterpart SANs within the redundant SAN fabric. Operation with direct connections between host and node or storage controller and node is unsupported.
- The Fibre Channel switch must be zoned to permit the hosts to see the SVC nodes and the SVC nodes to see the RAID Controllers. The SVC nodes within a cluster must be able to see each other, the back-end storage, and the front-end host HBAs. However, the front-end host, HBAs, and the back-end storage must not be in the same zone.

- Mixed speeds are permitted within the fabric, but not for intracluster communication. The user may use lower speeds to extend distance or to make use of 1 Gb/s legacy components.
- The local or remote fabric should not contain more than three interswitch links within each fabric. Operation with more ISLs is unsupported. When a local and a remote fabric are connected together for remote copy purposes, then the ISL count between a local node and a remote node may not exceed seven. This means that some ISLs may be used in a cascaded switch link between local and remote clusters, provided that the local or remote cluster internal ISL count is less than three.
- The switch configuration in an SVC fabric must be legal with respect to the switch manufacturer's configuration rules. This may impose restrictions on the switch configuration. For example, it may be a switch manufacturer's requirement that no other manufacturer's switches are present in the SAN. Operation outside the switch manufacturer's rules is not supported.
- The SAN contains only supported switches as listed on the Web at:

  `http://www.ibm.com/storage/support/2145`

  Operation with other switches is unsupported.
- Host HBAs in dissimilar hosts or dissimilar HBAs in the same host are in separate zones. For example, if you have AIX and Microsoft® hosts, they need to be in separate zones. Here *dissimilar* means that the hosts are running different operating systems or are different hardware platforms. Therefore, different levels of the same operating system are regarded as *similar*. This is a SAN interoperability issue rather than an SVC requirement.
- If the SAN fabric does not have redundant components, it is viewed as an invalid configuration.

Back-end storage guidelines

The following guidelines apply:

- On the FC SAN, back-end storage is always connected to SAN switches and nothing else. Multiple connections are allowed from the redundant controllers in the back-end storage to improve data bandwidth performance. It is not mandatory to have a connection from each redundant controller in the back end storage to each counterpart SAN. For example, in a FAStT configuration in which the FAStT contains two redundant controllers, only two controller minihubs are normally used. This means that Controller A in the FAStT is connected to counterpart SAN A, and controller B in the FAStT is connected to counterpart SAN B. Operation with direct connections between host and controller is unsupported.
- The SVC is configured to manage LUNS exported only by RAID controllers as listed on the Web at:

  `http://www.ibm.com/storage/support/2145`

  Operation with other RAID controllers is unsupported.
- All SVC nodes in a SVC must be able to see the same set of back-end storage ports on each back-end controller. Operation in a mode where two nodes see a different set of ports on the same controller is degraded. The system logs errors requesting a repair action. This can occur if inappropriate zoning was applied to the fabric. It can also occur if inappropriate LUN masking is used. This has important implications for back-end storage, such as FAStT, which imposes exclusivity rules on which HBA world wide names (WWNs) a storage partition can be mapped to. It is up to the user to check that the planned configuration is supported. You can find the supported hardware list at the Web site listed in the previous point.

Host and application servers guidelines

The following guidelines apply:

- Each SVC presents a vDisk to the SAN through four paths. Since in normal operation two nodes are used to provide redundant paths to the same storage, this means that a host HBA can see eight paths to each LUN presented by the SVC. We use zoning to limit the pathing to fours paths. The hosts must run a multipathing device driver to resolve this back to a single device. The multipathing driver supported by SVC is the IBM subsystem device driver (SDD).
- The number of paths from the SVC nodes to a host must not exceed eight. The maximum number of host HBA ports must not exceed four. Without any switch zoning, note the following calculation:

  ```
  Up to four SVC ports x two nodes per I/O group x four HBA ports equals 32 paths
  ```

  To restrict the number of paths to a host, the switches should be zoned so that each host HBA port is zoned with a SVC node port of each node in the cluster.
- If a host has multiple HBA ports, then each port should be zoned to a different set of SVC ports to maximize high availability and performance.

For management

The following guidelines apply:

- In addition to a Fibre Channel connection, each device has an Ethernet connection for configuration and error reporting. These connections are aggregated together through an Ethernet switch.
- The master console must only be in the SVC storage zone.

3.3.3 Boot support

SVC is not supported to be used as the boot device for a host (boot from SAN). The reason for this is that any multipathing driver is loaded after boot (multipathing drivers are not currently part of the operating system). Therefore, access to the boot device is not multipath protected.

3.3.4 Configuration saving

The configuration data for the SVC cluster is hardened in each node by the advanced reliability, availability, and serviceability (RAS) design and integration with the uninterruptible power supply. There is no need to save the configuration externally to the cluster.

3.3.5 High availability SAN design and configuration rules with SVC

Figure 3-6 shows a basic two node configuration. To provide high availability, SVC should be configured in redundant SAN fabrics. Our configuration in the lab as shown in Figure 3-6 is a redundant fabric made up of two 16-port switches.

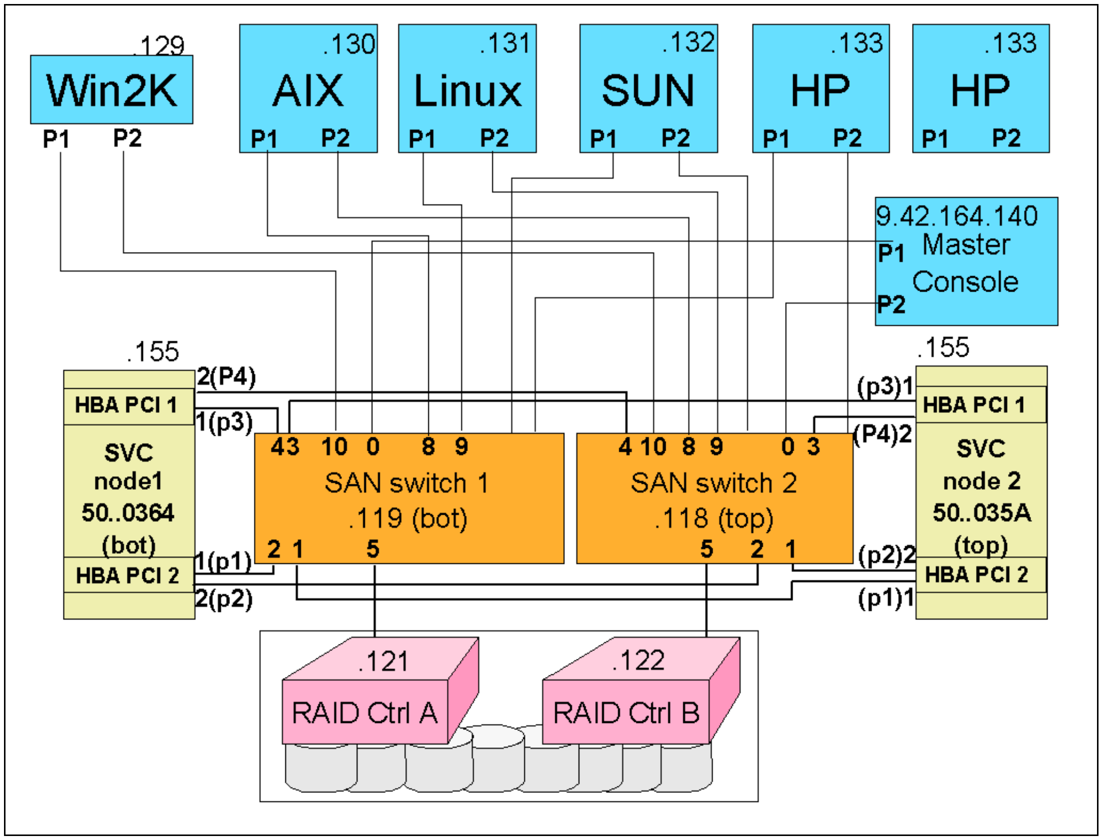


Figure 3-6 Simple two-node SVC high availability configuration

3.4 Zoning

To manage the zoning in the fabric, we recommend that you use *port zoning*. Port zoning provides protection from changing a component, such as a Fibre Channel adapter, in the fabric. There is a storage zone that contains the storage and all the SVC ports. The rest of the switch is zoned as half the SVC ports and one remaining port (set up for host connections). If additional storage is attached, the port used must be added to the storage zone. The master console is only in the storage zone.

Figure 3-7 shows the sample alias and zoning convention we used. The horizontal shaded areas make up a zone. The SVC ports, master console, and storage are in one zone. Each host port Host_FC0 is zoned to either of the ports 10, 20, 30, or 40 from both nodes. We took the port numbering on the nodes from their worldwide port names (WWPNs).

Node 1:

- 50:05:07:68:01:10:03:5a
- 50:05:07:68:01:20:03:5a
- 50:05:07:68:01:30:03:5a
- 50:05:07:68:01:40:03:5a

Node 2:

- 50:05:07:68:01:10:03:64
- 50:05:07:68:01:20:03:64

- 50:05:07:68:01:30:03:64
- 50:05:07:68:01:40:03:64

SVC1_Switch1										
Zone Name	0	1	2	3	4	5	7	8	9	10
SVC1_to_FAST1_A	SVC1_MC_P1	SVC1N2_10	SVC1N1_10	SVC1N2_30	SVC1N1_30	SVC1_FAST1_A				
RedHat1_FC0		SVC1N2_10	SVC1N1_10						Host2_FC0	
W2K_1_FC0				SVC1N2_30	SVC1N1_30					W2K_1_FC0
AIX1_FC0		SVC1N2_10	SVC1N1_10					AIX1_FC0		
AIX2_FC0				SVC1N2_30	SVC1N1_30		Host4_FC0			

SVC1_Switch2										
Zone Name	0	1	2	3	4	5	7	8	9	10
SVC1_to_FAST1_B	SVC1_MC_P2	SVC1N2_20	SVC1N1_20	SVC1N2_40	SVC1N1_40	SVC1_FAST1_B				
RedHat1_FC1		SVC1N2_20	SVC1N1_20						Host2_FC1	
W2K_1_FC1				SVC1N2_40	SVC1N1_40					W2K_1_FC1
AIX1_FC1		SVC1N2_20	SVC1N1_20					AIX1_FC1		
AIX2_FC1				SVC1N2_40	SVC1N1_40		Host4_FC1			

SVC1N1				10	20	PCI2
	PCI1	30	40			
		This is node 1 in SVC1				

SVC1N2				10	20	PCI2
	PCI1	30	40			
		This is node 2 in SVC1				

Figure 3-7 Sample zoning convention

Look at SVC1_Switch1 for the host AIX1 on port 8 of both switches. We created the host entries as AIX1_FC0 for adapter 1 and AIX2_FC1 for adapter 2. We continue this style of naming throughout our hosts.

Naming conventions become a real challenge. The challenges come from finding naming conventions that will continue to be steady as changes occur to the environment. We used the following naming convention to create the aliases for our two cluster, two fabric setup. We only used one interswitch link instead of two as shown in Figure 3-8.

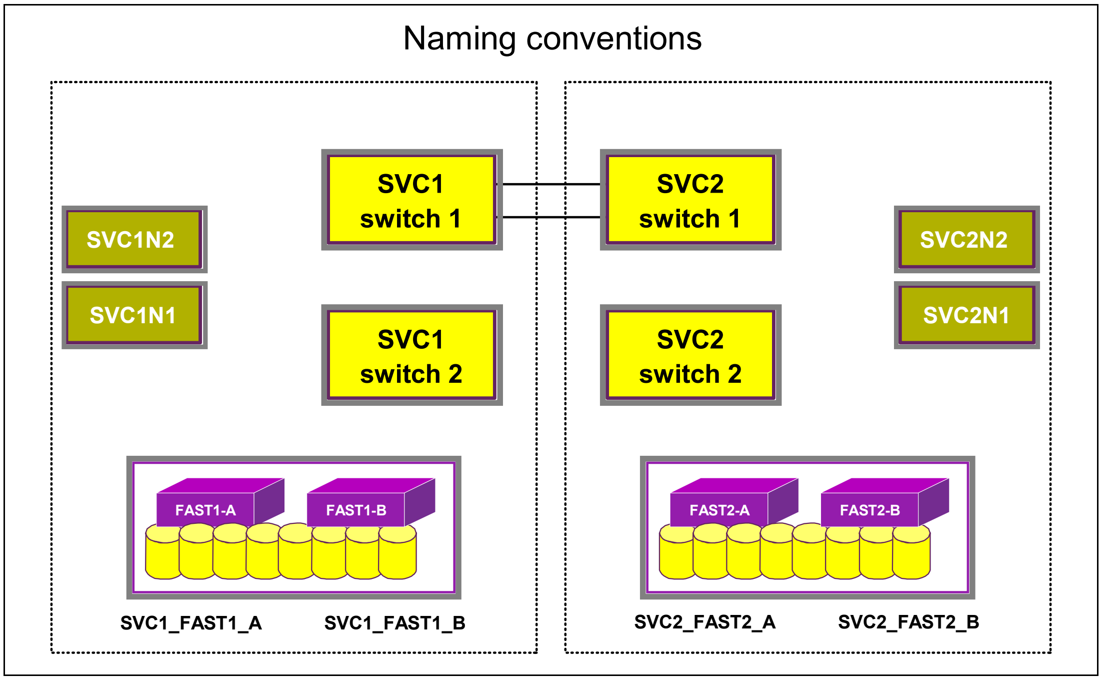


Figure 3-8 Dual SVC setup

We connected the fabrics together to use the Peer-to-Peer Remote Copy (PPRC) functions.

First SVC:

- SVC = SVC1

This is controller A and B on the first FAStT connected to the first SVC node:

- SVC1_FAST1_A
- SVC1_FAST1_B

For example, you can use SVC1_ESS1_Bay1A for ESS storage bay locations.

Node 1 ports10, 20, 30, and 40 in SVC1:

- SVC1N1_10
- SVC1N1_20
- SVC1N1_30
- SVC1N1_40

Node 2 ports10,20,30 and 40 in SVC1:

- SVC1N2_10
- SVC1N2_20
- SVC1N2_30
- SVC1N2_40

Master console ports P1 and P2:

- SVC1_MC_P1
- SVC1_MC_P2

Host Fibre Channel ports FC0 and FC1:

- HOSTNAME_FC0
- HOSTNAME_FC1

For example, AIX1_fc0, AIX_fc1

Storage zone names:

- SVC1_to_SVC1FAST1_A
- SVC1_to_SVC1FAST1_B

For example, you can use SVC1_to_ESS1_Bay1A for ESS storage bay locations.

Host zone names:

- HOSTNAME_FC0_to_SVC1
- HOSTNAME_FC1_to_SVC1

Master console zone name:

- SVC1_MC_P1_to_SVC1

We merged two SVC configurations together using the above naming conventions with little trouble. Remember to check the following items to make a fabric merge possible between two switches.

Check the switch domain IDs make sure they are unique:

```
SVC1_Switch1:admin> switchshow
switchName:     SVC1_Switch1
switchType:     9.2
```

```
switchState:    Online
switchMode:     Native
switchRole:     Subordinate
switchDomain:   11
```

Changing the IDs affects your zoning setup. Therefore, you must change them first before you change your zoning information. Part of the switch port address that is used by the operating system is determined by the core pid ID and domain ID. A change of the domain ID or core pid disrupts some UNIX® operating systems. It affects how the UNIX operating system addresses the device SCSI address. Make sure you check first before you attempt this when storage is already defined and in use. Here is an example port address:

```
xxypaa
```

Here, xx is the domain, y is the PID, p is the port, and aa is 00 for N ports or loop address for NL ports. Consider this example:

```
010100 domain=01, PID = 0, port=1
```

Some older Brocade/IBM switches have the pid set to 1 and switches do not merge with switches set to 0. This is from the early days of Fibre Channel. Now the switches need to use both bytes for port addressing. See the *IBM SAN Survival Guide*, SG24-6143, for more information regarding switches.

Note: A change of the domain ID or core pid disrupts some UNIX operating systems. Make sure you check first before you attempt this when storage is already defined and in use.

3.4.1 Lab configuration

Example 3-1 shows the configuration taken from our fabrics. We have two FAStT servers and an ESS for our back-end storage.

Example 3-1 Switch zoning

```
SVC1_Switch1:admin> switchshow
switchName:     SVC1_Switch1
switchType:     9.2
switchState:    Online
switchMode:     Native
switchRole:     Subordinate
switchDomain:   11
switchId:       fffc0b
switchWwn:      10:00:00:60:69:51:87:0a
switchBeacon:   OFF
Zoning:         ON (SVC1_SVC2)
port  0: id N2 Online       F-Port 21:01:00:e0:8b:28:af:d6
port  1: id N2 Online       F-Port 50:05:07:68:01:10:03:64
port  2: id N2 Online       F-Port 50:05:07:68:01:10:03:5a
port  3: id N2 Online       F-Port 50:05:07:68:01:30:03:64
port  4: id N2 Online       F-Port 50:05:07:68:01:30:03:5a
port  5: id N2 Online       F-Port 20:08:00:a0:b8:0f:bd:f1
port  6: id N2 No_Light
port  7: id N2 No_Light
port  8: id N1 Online       F-Port 10:00:00:00:c9:29:5a:9a
port  9: id N2 Online       F-Port 21:00:00:e0:8b:09:15:41
port 10: id N1 Online       F-Port 21:00:00:e0:8b:04:d7:51
port 11: id N2 No_Light
```

```
port 12: -- N2 No_Module
port 13: -- N2 No_Module
port 14: -- N2 No_Module
port 15: id N1 Online       E-Port 10:00:00:60:69:10:97:15 "SVC2_Switch1" (upstream)

SVC1_Switch2:admin> switchshow
switchName:     SVC1_Switch2
switchType:     9.2
switchState:    Online
switchMode:     Native
switchRole:     Subordinate
switchDomain:   12
switchId:       fffc0c
switchWwn:      10:00:00:60:69:51:87:2f
switchBeacon:   OFF
Zoning:         ON (SVC1_SVC2)
port  0: id N2 Online       F-Port 21:00:00:e0:8b:08:af:d6
port  1: id N2 Online       F-Port 50:05:07:68:01:20:03:64
port  2: id N2 Online       F-Port 50:05:07:68:01:20:03:5a
port  3: id N2 Online       F-Port 50:05:07:68:01:40:03:64
port  4: id N2 Online       F-Port 50:05:07:68:01:40:03:5a
port  5: id N2 Online       F-Port 20:09:00:a0:b8:0f:bd:f2
port  6: id N2 No_Light
port  7: id N2 No_Light
port  8: id N1 Online       F-Port 10:00:00:00:c9:26:6f:5b
port  9: id N2 Online       F-Port 21:01:00:e0:8b:29:15:41
port 10: id N1 Online       F-Port 21:01:00:e0:8b:24:d7:51
port 11: id N2 No_Light
port 12: -- N2 No_Module
port 13: -- N2 No_Module
port 14: -- N2 No_Module
port 15: id N1 Online       E-Port 10:00:00:60:69:10:02:50 "SVC2_Switch2" (upstream)
SVC1_Switch2:admin>

SVC2_Switch1:admin> switchshow
switchName:     SVC2_Switch1
switchType:     2.4
switchState:    Online
switchMode:     Native
switchRole:     Principal
switchDomain:   21
switchId:       fffc15
switchWwn:      10:00:00:60:69:10:97:15
switchBeacon:   OFF
Zoning:         ON (SVC1_SVC2)
port  0: id  Online       F-Port  21:01:00:e0:8b:25:bf:40
port  1: id  Online       F-Port  50:05:07:68:01:10:01:67
port  2: id  Online       F-Port  50:05:07:68:01:10:01:2c
port  3: id  Online       F-Port  50:05:07:68:01:40:01:67
port  4: sw  Online       F-Port  50:05:07:68:01:40:01:2c
port  5: id  Online       F-Port  20:04:00:a0:b8:0c:eb:7d
port  6: id  Online       F-Port  50:05:07:63:00:c4:0b:e8
port  7: id  No_Light
port  8: id  No_Light
port  9: id  No_Light
port 10: sw  Online       F-Port  10:00:00:00:c9:25:84:1d
port 11: id  No_Light
port 12: id  No_Light
port 13: id  No_Light
port 14: sw  No_Light
```

```
port 15: id  Online       E-Port  10:00:00:60:69:51:87:0a "SVC1_Switch1" (downstream)
SVC2_Switch1:admin>

SVC2_Switch2:admin> switchshow
switchName:     SVC2_Switch2
switchType:     2.3
switchState:    Online
switchMode:     Native
switchRole:     Principal
switchDomain:   22
switchId:       fffc16
switchWwn:      10:00:00:60:69:10:02:50
switchBeacon:   OFF
Zoning:         ON (SVC1_SVC2)
port  0: id  No_Light
port  1: id  Online       F-Port  50:05:07:68:01:20:01:67
port  2: id  Online       F-Port  50:05:07:68:01:20:01:2c
port  3: id  Online       F-Port  50:05:07:68:01:30:01:67
port  4: id  Online       F-Port  50:05:07:68:01:30:01:2c
port  5: id  Online       F-Port  20:05:00:a0:b8:0c:eb:7d
port  6: sw  Online       F-Port  50:05:07:63:00:c8:0b:e8
port  7: id  No_Light
port  8: id  No_Light
port  9: id  No_Light
port 10: id  No_Light
port 11: id  No_Light
port 12: id  No_Light
port 13: id  No_Light
port 14: --  No_Module
port 15: id  Online       E-Port  10:00:00:60:69:51:87:2f "SVC1_Switch2" (downstream)
SVC2_Switch2:admin>

SVC1_Switch1:admin> cfgshow
Defined configuration:
 cfg:   SVC1_SVC2
                AIX1_FC0_to_SVC1; AIX2_FC0_to_SVC2; RedHat1_FC0_to_SVC1;
                SVC1_MC_P1_to_SVC1; SVC1_to_SVC1FAST1_A; SVC2_to_SVC2FAST1_A;
                W2K_1_FC0_to_SVC1; AIX1_FC0_to_ESS1_B1A1;
                W2K_1_FC0_to_ESS1_B1A1; SVC1_to_SVC2
 cfg:   SVC1_SW1
                SVC1_MC_P1_to_SVC1; SVC1_to_SVC1FAST1_A; AIX1_FC0_to_SVC1;
                RedHat1_FC0_to_SVC1; W2K_1_FC0_to_SVC1
 cfg:   SVC2_SW1
                SVC2_to_SVC2FAST1_A; AIX2_FC0_to_SVC2
 zone:  AIX1_FC0_to_ESS1_B1A1
                AIX1_FC0; SVC2_ESS1_B1A1
 zone:  AIX1_FC0_to_SVC1
                AIX1_FC0; SVC1N1_30; SVC1N2_30
 zone:  AIX2_FC0_to_SVC2
                AIX2_FC0; SVC2N2_40; SVC2N1_40
 zone:  RedHat1_FC0_to_SVC1
                RedHat1_FC0; SVC1N1_30; SVC1N2_30
 zone:  SVC1_MC_P1_to_SVC1
                SVC1N1_10; SVC1N1_30; SVC1N2_10; SVC1N2_30; SVC1_MC_P1
 zone:  SVC1_to_SVC1FAST1_A
                SVC1_FAST1_A; SVC1N1_10; SVC1N1_30; SVC1N2_10; SVC1N2_30
 zone:  SVC1_to_SVC2
                SVC1N1_10; SVC1N1_30; SVC1N2_10; SVC1N2_30; SVC2N1_10;
                SVC2N1_40; SVC2N2_10; SVC2N2_40
 zone:  SVC2_MC_P1_to_SVC2
```

```
               W2K_2_FC0; SVC2N2_10; SVC2N1_10; SVC2N2_40; SVC2N1_40
 zone:  SVC2_to_ESS1_B1A1
               SVC2_ESS1_B1A1; SVC2N2_10; SVC2N1_10; SVC2N2_40; SVC2N1_40
 zone:  SVC2_to_SVC2FAST1_A
               SVC2N2_10; SVC2N1_10; SVC2N2_40; SVC2N1_40; SVC2_FAST1_A
 zone:  W2K_1_FC0_to_ESS1_B1A1
               W2K_1_FC0; SVC2_ESS1_B1A1
 zone:  W2K_1_FC0_to_SVC1
               SVC1N1_10; SVC1N2_10; W2K_1_FC0
 zone:  W2K_2_FC0_to_SVC2
               SVC2N2_10; SVC2N1_10; W2K_2_FC0
 alias: AIX1_FC0
               11,8
 alias: AIX2_FC0
               21,10
 alias: RedHat1_FC0
               11,9
 alias: SVC1N1_10
               11,2
 alias: SVC1N1_30
               11,4
 alias: SVC1N2_10
               11,1
 alias: SVC1N2_30
               11,3
 alias: SVC1_FAST1_A
               11,5
 alias: SVC1_MC_P1
               11,0
 alias: SVC2N1_10
               21,2
 alias: SVC2N1_40
               21,4
 alias: SVC2N2_10
               21,1
 alias: SVC2N2_40
               21,3
 alias: SVC2_ESS1_B1A1
               21,6
 alias: SVC2_FAST1_A
               21,5
 alias: SVC2_MC_P1
               21,0
 alias: W2K_1_FC0
               11,10
 alias: W2K_2_FC0
               21,9

Effective configuration:
 cfg:   SVC1_SVC2
 zone:  AIX1_FC0_to_ESS1_B1A1
               11,8
               21,6
 zone:  AIX1_FC0_to_SVC1
               11,8
               11,4
               11,3
 zone:  AIX2_FC0_to_SVC2
               21,10
               21,3
```

```
                21,4
 zone:  RedHat1_FC0_to_SVC1
                11,9
                11,4
                11,3
 zone:  SVC1_MC_P1_to_SVC1
                11,2
                11,4
                11,1
                11,3
                11,0
 zone:  SVC1_to_SVC1FAST1_A
                11,5
                11,2
                11,4
                11,1
                11,3
 zone:  SVC1_to_SVC2
                11,2
                11,4
                11,1
                11,3
                21,2
                21,4
                21,1
                21,3
 zone:  SVC2_to_SVC2FAST1_A
                21,1
                21,2
                21,3
                21,4
                21,5
 zone:  W2K_1_FC0_to_ESS1_B1A1
                11,10
                21,6
 zone:  W2K_1_FC0_to_SVC1
                11,2
                11,1
                11,10

SVC1_Switch2:admin> cfgshow
Defined configuration:
 cfg:   SVC1_SVC2
                AIX1_FC1_to_SVC1; AIX2_FC1_to_SVC2; SVC1_MC_P2_to_SVC1;
                SVC1_to_SVC1FAST1_B; SVC2_to_SVC2FAST1_B; W2K_1_FC1_to_SVC1;
                RedHat1_FC1_to_SVC1; AIX1_FC1_to_ESS1_B31A;
                W2K_1_FC1_to_ESS1_B31A; SVC1_to_SVC2
 cfg:   SVC1_SW2
                AIX1_FC1_to_SVC1; SVC1_MC_P2_to_SVC1; SVC1_to_SVC1FAST1_B
 cfg:   SVC2_SW2
                SVC2_to_SVC2FAST1_B; AIX2_FC1_to_SVC2
 zone:  AIX1_FC1_to_ESS1_B31A
                AIX1_FC1; SVC2_ESS1_B3A1
 zone:  AIX1_FC1_to_SVC1
                AIX1_FC1; SVC1N1_40; SVC1N2_40
 zone:  AIX2_FC1_to_SVC2
                AIX2_FC1; SVC2N1_30; SVC2N2_30
 zone:  RedHat1_FC1_to_SVC1
                RedHat1_FC1; SVC1N1_40; SVC1N2_40
 zone:  SVC1_MC_P2_to_SVC1
```

```
                SVC1N1_20; SVC1N1_40; SVC1N2_40; SVC1N2_20; SVC1_MC_P2
 zone:  SVC1_to_SVC1FAST1_B
                SVC1N1_20; SVC1N1_40; SVC1N2_20; SVC1N2_40; SVC1_FAST1_B
 zone:  SVC1_to_SVC2
                SVC1N1_20; SVC1N1_40; SVC1N2_20; SVC1N2_40; SVC2N1_20;
                SVC2N1_30; SVC2N2_20; SVC2N2_30
 zone:  SVC2_MC_P2_to_SVC2
                SVC2_MC_P2; SVC2N2_30; SVC2N2_20; SVC2N1_30; SVC2N1_20
 zone:  SVC2_to_ESS1_B3A1
                SVC2N1_20; SVC2N1_30; SVC2N2_20; SVC2N2_30; SVC2_ESS1_B3A1
 zone:  SVC2_to_SVC2FAST1_B
                SVC2N1_20; SVC2N1_30; SVC2N2_20; SVC2N2_30; SVC2_FAST1_B
 zone:  W2K_1_FC1_to_ESS1_B31A
                W2K_1_FC1; SVC2_ESS1_B3A1
 zone:  W2K_1_FC1_to_SVC1
                SVC1N2_20; SVC1N1_20; W2K_1_FC1
 zone:  W2K_2_FC1_to_SVC2
                W2K_2_FC1; SVC2N1_20; SVC2N2_20
 alias: AIX1_FC1
                12,8
 alias: AIX2_FC1
                22,10
 alias: RedHat1_FC1
                12,9
 alias: SVC1N1_20
                12,2
 alias: SVC1N1_40
                12,4
 alias: SVC1N2_20
                12,1
 alias: SVC1N2_40
                12,3
 alias: SVC1_FAST1_B
                12,5
 alias: SVC1_MC_P2
                12,0
 alias: SVC2N1_20
                22,2
 alias: SVC2N1_30
                22,4
 alias: SVC2N2_20
                22,1
 alias: SVC2N2_30
                22,3
 alias: SVC2_ESS1_B3A1
                22,6
 alias: SVC2_FAST1_B
                22,5
 alias: SVC2_MC_P2
                22,0
 alias: W2K_1_FC1
                12,10
 alias: W2K_2_FC1
                H{1,9}

Effective configuration:
 cfg:   SVC1_SVC2
 zone:  AIX1_FC1_to_ESS1_B31A
                12,8
                22,6
```

```
 zone:  AIX1_FC1_to_SVC1
                12,8
                12,4
                12,3
 zone:  AIX2_FC1_to_SVC2
                22,10
                22,4
                22,3
 zone:  RedHat1_FC1_to_SVC1
                12,9
                12,4
                12,3
 zone:  SVC1_MC_P2_to_SVC1
                12,2
                12,4
                12,3
                12,1
                12,0
 zone:  SVC1_to_SVC1FAST1_B
                12,2
                12,4
                12,1
                12,3
                12,5
 zone:  SVC1_to_SVC2
                12,2
                12,4
                12,1
                12,3
                22,2
                22,4
                22,1
                22,3
 zone:  SVC2_to_SVC2FAST1_B
                22,2
                22,4
                22,1
                22,3
                22,5
 zone:  W2K_1_FC1_to_ESS1_B31A
                12,10
                22,6
 zone:  W2K_1_FC1_to_SVC1
                12,1
                12,2
                12,10

SVC1_Switch2:admin>
```

3.4.2 Dual room high availability configuration with SVC

Figure 3-9 shows the high availability configuration of the SVC, when each node of the cluster is in a different room. This is the configuration that we recommend for maximum availability.

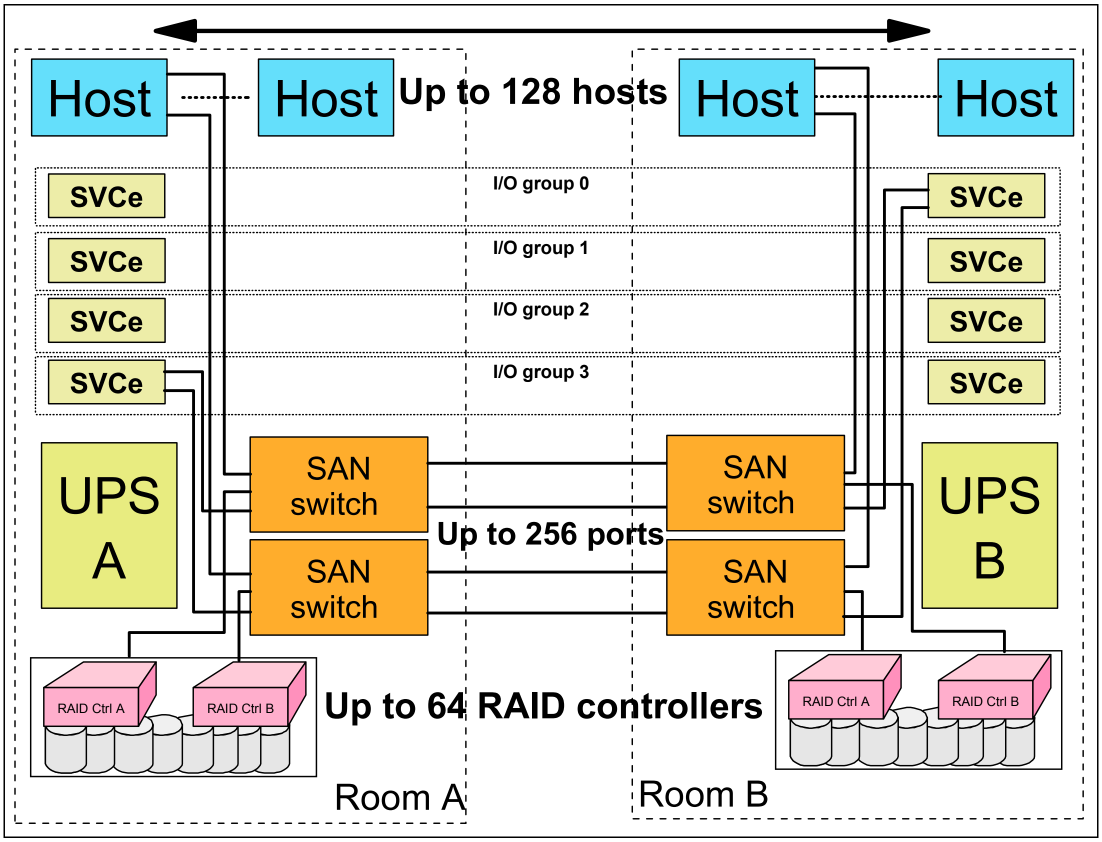


Figure 3-9 High availability SAN Volume Controller cluster in a two-room configuration

3.4.3 Local and remote SAN fabrics with SVC

SVC supports both intracluster and intercluster remote copy. From the intracluster point of view, any single cluster is a reasonable candidate for a remote copy operation. Intercluster operation needs a pair of clusters, separated by a number of moderately high bandwidth links. Such a configuration is shown in Figure 3-9:

- In this release of SVC, the local and remote fabric interconnect is a single ISL hop between a switch in the local fabric and a switch in the remote fabric. It is a single mode fibre, up to 10 km in length. Operation with other local and remote fabric interconnects is unsupported at present. Check the support Web site for any changes to this:

  http://www-1.ibm.com/support/docview.wss?rs=591&uid=ssg1S1001859

- For release 1 of SVC, operation with Fibre Channel extenders is not supported. If this is required, speak to your IBM representative about raising a Request for Price Quotation (RPQ).
- In Remote Copy configurations, additional zones are required that contain only the local nodes and the remote nodes. It is valid for the local hosts to see the remote nodes or for the remote hosts to see the local nodes. It is unsupported to create a zone that presents both the local and remote back-end storage and either local and remote nodes or both.

3.5 SVC back-end storage planning

This section describes the various types of disk groups and their relationships.

In the configuration shown in Figure 3-10, a RAID controller is presenting a LUN to the SVC and another host. SVC presents the vDisk or vDisks created from the mDisk to the other host. The two hosts are not required to have two different multipath drivers installed. The RAID controller being a FAStT server would have RDAC installed for the direct attached host and SDD installed for the host attached to SVC. This is a supported configuration.

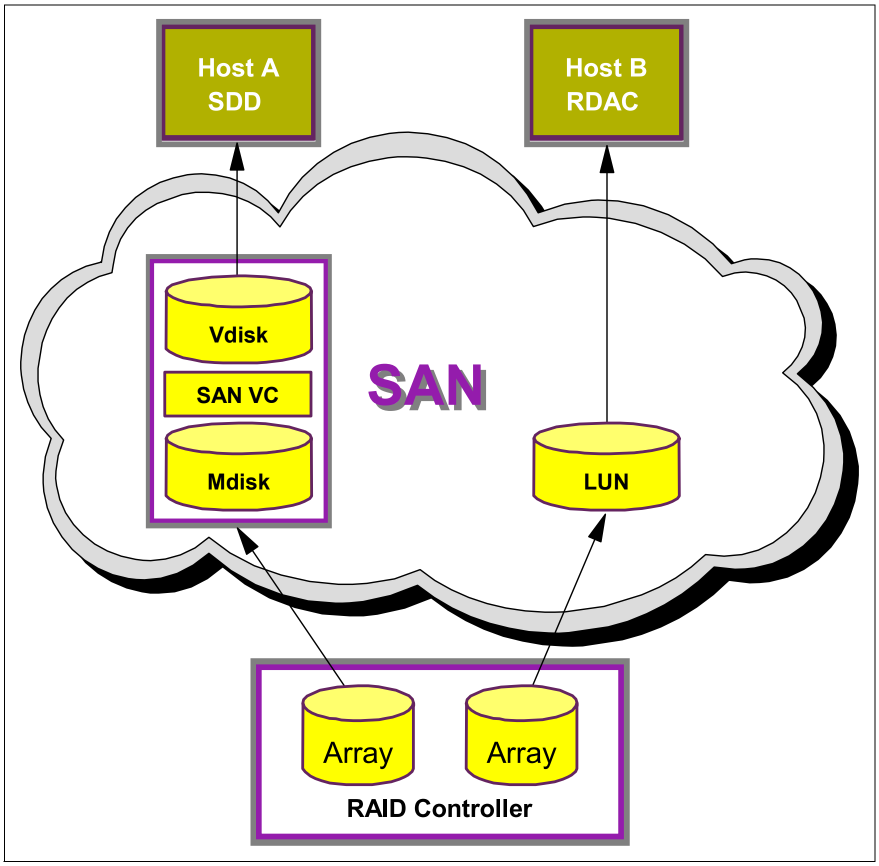


Figure 3-10 RAID controller shared

With the ESS, you can attach directly to an ESS LUN and to a vDisk from the SVC that came from the ESS. The host uses the subsystem device driver to access both parties. This is a supported configuration and is shown in Figure 3-11.

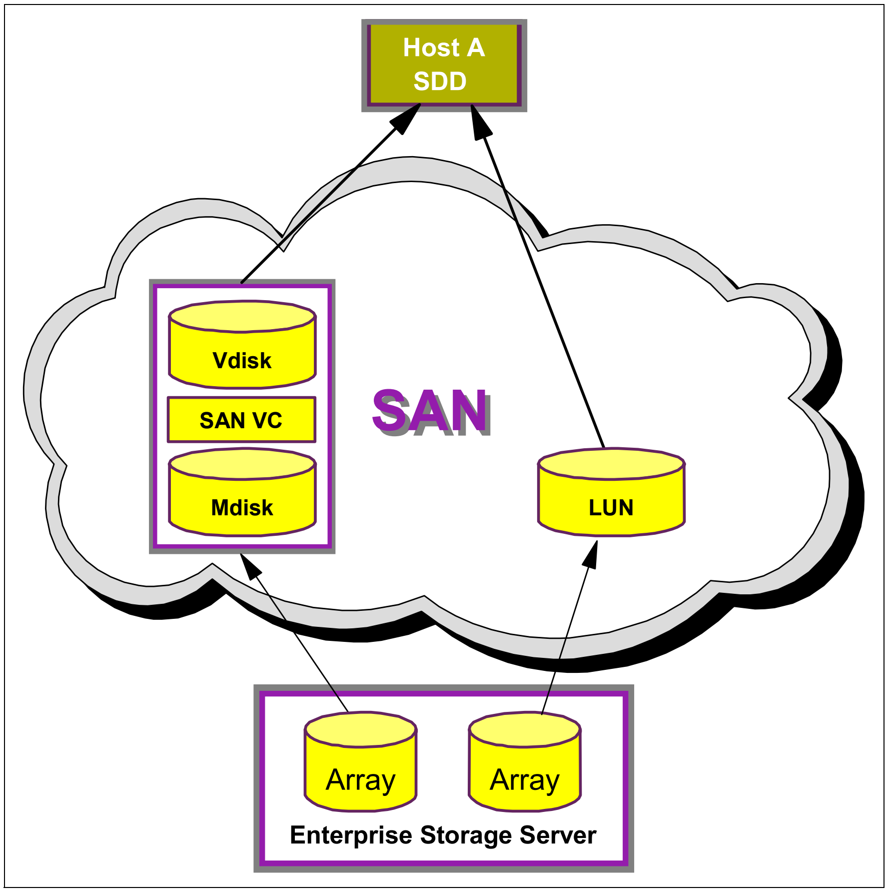


Figure 3-11 Host connected to ESS and SVC

In the configuration shown in Figure 3-12, the host does not need to have both the RDAC driver installed for access to the FAStT and the subsystem device driver installed to access the SVC. This configuration is *not supported*.

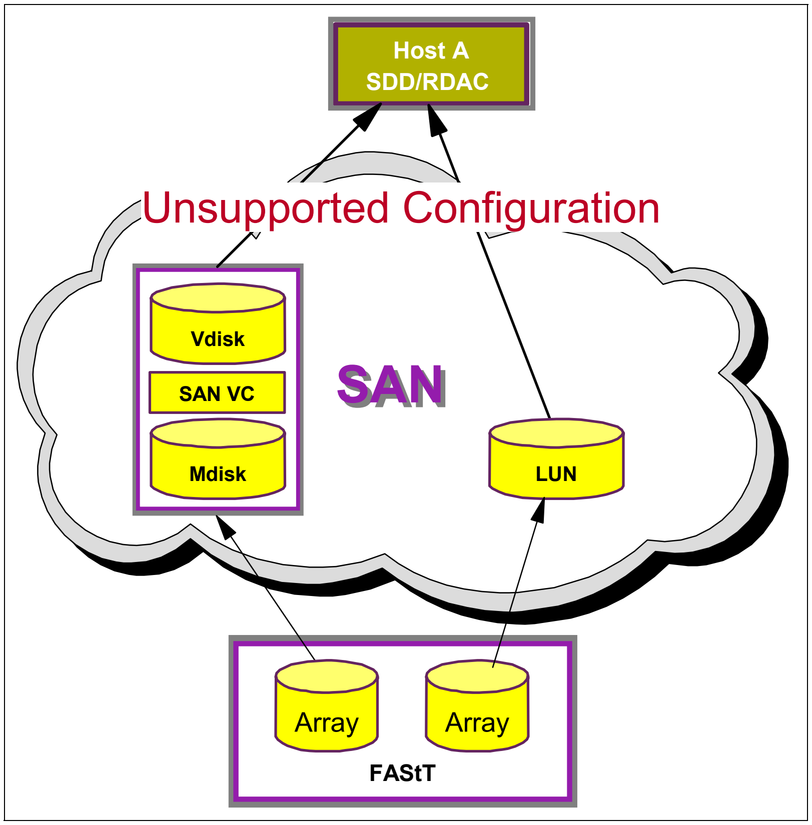


Figure 3-12 FAStT unsupported configuration

3.5.1 Block virtualization

The managed disk group is at the center of the many-to-many relationship between managed disks and virtual disks. It acts as a container into which managed disks contribute chunks of disk blocks, known as *extents*, and from which virtual disks consume these extents of storage.

- MDGs are collections of managed disks. A managed disk is contained within exactly one MDG.
- A SVC supports up to 128 MDGs.
- There is no limit to the number of virtual disks that can be in a MDG other than the per cluster limit.
- MDGs are also (normally) collections of virtual disks. Under normal circumstances, a virtual disk is associated with exactly one MDG. The exception to this is when a virtual disk is migrated between MDGs.

3.5.2 MDGs, I/O groups, virtual disks, and managed disk

Figure 3-13 shows three RAID controllers that were configured to provide a number of SCSI logical units. In SVC terminology, each of these logical units or LUNS is a managed disk. RAID controller A contains two managed disks, known as M1 and M2. RAID Controller B contains managed disks M3 and M4. And RAID controller C contains managed disks M5 and M6.

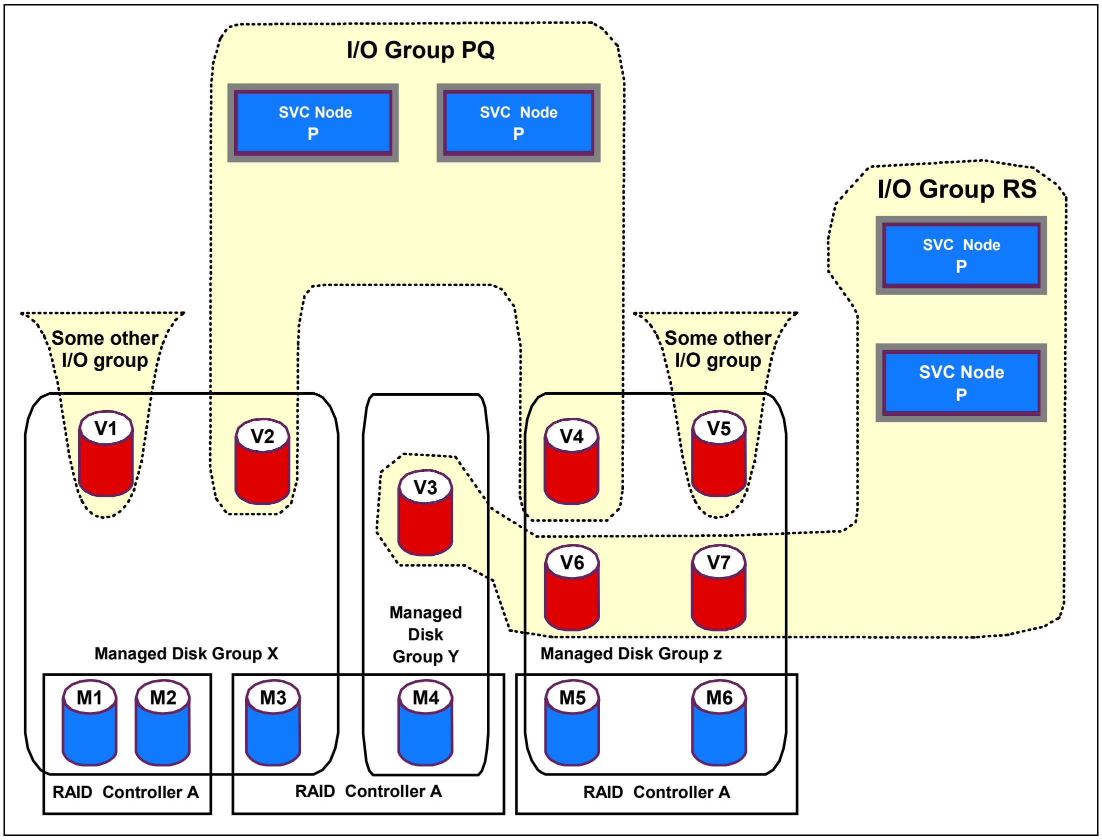


Figure 3-13 Disk relationships

Figure 3-13 also shows three MDGs, X, Y, and Z. Each managed disk is contained within a single MDG and that one MDG can span controllers. SVC supports an arbitrary relationship between RAID controllers and MDGs. The MDG simply contains a collection of managed disks from the set of available controllers.

Note: We recommend that only LUNs from one back-end controller form part of a MDG.

Plus, Figure 3-13 shows virtual disks numbered V1 to V7. Each virtual disk is contained entirely within a MDG. This is the normal situation. The only exception to this is during a migration.

Virtual disks are also members of another collection, namely I/O groups. Figure 3-13 shows two I/O groups named PQ and RS. An I/O group contains an arbitrary set of virtual disks and exactly two SVC nodes (unless one has failed). The I/O group defines which nodes support I/O access from hosts.

There is no fixed relationship between I/O groups and MDGs. An individual virtual disk is normally a member of one MDG and one I/O group:

- The MDG defines which managed disks provide back-end storage that makes up the virtual disk.
- The I/O group defines which SVC nodes provide I/O access to the virtual disk.

3.5.3 Extents

A virtual disk occupies an integer number of extents. It length does not need to be an integer multiple of the extent size, but must be an integer multiple of the block size. Any space left over between the last logical block in the virtual disk and the end of the last extent in the virtual disk is unused.

You can define a vDisk with the smallest granularity of 512 bytes (a block). However, an entire extent is reserved even if it is partially used. An extent is never shared between virtual disks. Each extent on a managed disk is contained with at most one virtual disk. Free extents are associated with no virtual disk.

SVC supports extent sizes of 16 MB, 32 MB, 64 MB, 128 MB, 256 MB, and 512 MB. The extent size is a property of the MDG, which is set when the MDG is created. It cannot be changed and all managed disks, which are contained in the MDG, have the same extent size. It follows that all virtual disks associated with the MDG must also have the same extent size. Table 3-1 shows the relationship between the extent size and the maximum capacity of the cluster.

Table 3-1 Extent size and maximum cluster capacities

Extent size	Maximum cluster capacity
16 MB	64 TB
32 MB	128 TB
64 MB	256 TB
128 MB	512 TB
256 MB	1 PB
512 MB	2 PB

3.5.4 Image mode virtual disk

Image mode provides a direct block-for-block translation from the managed disk to the virtual disk with no virtualization. This mode is intended to allow virtualization of managed disks, which already contain data that was written directly, not through a SVC node from a pre-virtualized subsystem. When an image mode virtual disk is created, it directly corresponds to the managed disk it is created from.

It allows a client to insert a SVC into the data path of an existing storage configuration with minimal downtime. After the SVC is inserted into the data path using image mode, you can use the migration facilities to migrate the data to managed mode and rearrange the data while an application is accessing the data.

When you create a image mode virtual disk, the managed disk specified must not be a member of a MDG. The managed disk is made a member of the specified MDG as a result of the creation of the image mode virtual disk.

Image mode provides direct mapping from managed disk to virtual disk. You can think of it as a property of both virtual disks and managed disks. Image mode is supported only for virtual disk creation. It is not possible to migrate a managed mode disk to an image mode disk.

The capacity specified must be less than or equal to the size of the managed disk. If it is less than the size of the managed disk, then the unused space in the managed disk is *not* available for use in any other virtual disk. There is no facility to specify an offset. Therefore, logical block number (LBA) “N” on the resulting image mode virtual disk maps directly to LBA

"N" on the image mode managed disk. Image mode virtual disks have a minimum size of one block (512 bytes) and always occupy at least one extent.

Image mode managed disks are members of a MDG, but do not contribute free extents to the pool of free extents. Therefore, an image mode managed disk can have at most one virtual disk associated with it.

3.5.5 Managed mode virtual disk

Disks operating in managed mode provide a full set of virtualization functions.

Within a MDG, SVC supports an arbitrary relationship between extents on (managed mode) virtual disks and extents on managed disks. Subject to the constraint that each managed disk extent is contained in at most one virtual disk, each virtual disk extent maps to exactly one managed disk extent (except when in the progress of migrating).

Figure 3-14 shows virtual disk V, which is made up of a number of extents. Each of these extents is mapped to an extent on one of the managed disks A,B or C. The mapping table stores the details of this indirection. You see that some of the managed disk extents are unused. That is to say there is no virtual disk extent which maps to them. These unused extents are available for use in creating new virtual disks, migration, expansion, etc.

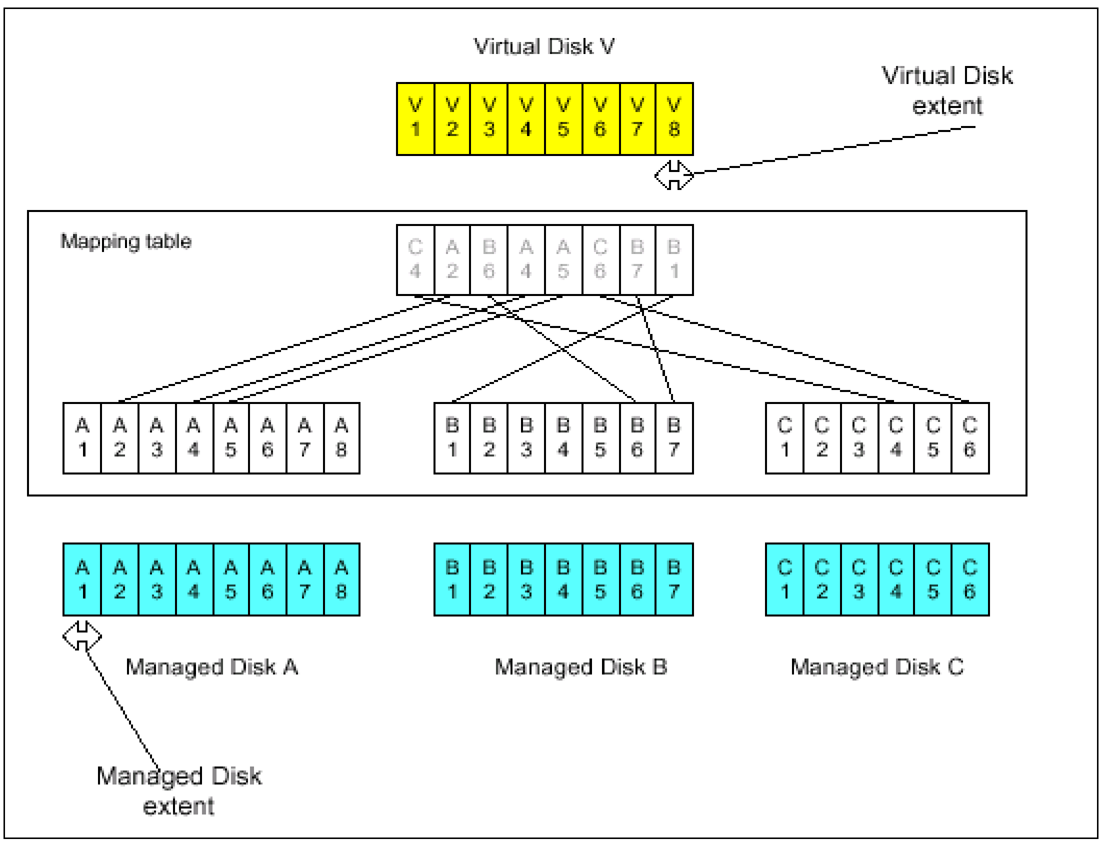


Figure 3-14 Simple view of block virtualization

Creating managed mode virtual disk

When a virtual disk is created, the SVC needs to know the policy to apply to create the initial assignment of managed disk extents to virtual disk extents. The supported policies are listed in the following sections. These policies are only used for the creation of a new virtual disk.

After the virtual disk is created, the policy has no effect and is not considered when making decisions during migration operations.

Striped

When a virtual disk is created using a striped policy, its extents are allocated from the specified ordered list of managed disks. The allocation algorithm starts with the first managed disk in the ordered list and attempts to allocate an extent from it then it moves to the next disk so on for each managed disk in turn. If the specified managed disk has no free extents, then it misses its turn and the turn passes to the next managed disk in the list. When the end of the list is reached, the algorithm loops back to the first disk in the list. Allocation proceeds until all extents required have been allocated.

A specific managed disk can appear more than once in the list. This causes two extents to be allocated from the disk on each pass of the list. This may be useful when striping across managed disks of different sizes.

When selecting which extent to allocate from the chosen managed disk, the policy followed is as described in the “Allocation of free extents” section that follows. This allocation policy leads to a coarse grained striping. The granularity of the striping is at the extent level. This coarse grained striping is unlikely to result in large bandwidth for sequential transfers but is likely to spread the workload caused by random small transactions across the managed disks from which the extents are allocated.

Wide striping increases the probability that the data on the virtual disk will be lost due to the failure of one of the managed disk across which the virtual disk is striped. It is acceptable for the list to contain only one disk. In this case, extents are allocated from a single disk as described in the “Allocation of free extents” section. Contrast this with the allocation scheme for the sequential policy.

Sequential

When a virtual disk is created using a sequential policy, its extents are allocated from a single specified managed disk. The SVC searches for regions of the target managed disk that contain free extents, which are sequential so the region is large enough to allocate the virtual disk from completely sequential extents. If it finds more than one such region, it chooses the smallest region which satisfies this condition. If it doesn’t find such regions, creation of the virtual disk fails.

Allocation of free extents

The migration operations and some of the virtualization operations require the allocation of a specific number of extents from a specific set of managed disks. The algorithm used to achieve this is described here.

Choosing the managed disk to allocate from:

Where the set of managed disks to allocate extents from contains more than one disk, extents are allocated from managed disks in a *round robin* fashion. If a managed disk has no free extents when its turn arrives, then its turn is missed and the round robin moves to the next managed disk in the set which has a free extent.

As the algorithm progresses, disks with no free extents on the previous pass of the round robin are queried for free extents on each turn of the algorithm in case extents become free.

Choosing the extent to allocate from a specific managed disk:

When an extent is to be allocated from a specific managed disk, the allocation policy is to allocate the next free extent from a list of free extents held by the SVC cluster for the specific managed disk.

3.5.6 Selecting MDGs

There are two key questions relating to the selection of MDGs:

- How many MDGs should I use?
- From which mDisk group should I create my vDisk?

The first question is discussed in 3.1, “General planning rules” on page 26. To summarize this, we make the following recommendations:

- Create at least one separate MDG for all the image mode virtual disks.
- Create one separate MDG for all the managed mode virtual disks with the sequential policy.
- Create MDGs for back-end controllers that provide the same level of performance, reliability, or both. You can group all the managed disks that are RAID 10 in one MDG, all the managed disks that are RAID 5 in another group and all the managed disks that are RAID 0 (no protection) in a third group.

  Note: It may be prudent to keep each back-end storage subsystem in a separate mDisk group. This prevents a failure in storage subsystem A from affecting storage subsystem B. If a vDisk is composed of mDisks from both A and B, then a failure in either A or B causes the vDisk to be unavailable until the fault on A or B is rectified.

- You should name them in way that it will be easier for you when you create a virtual disk to associate it with the MDG with the level of performance and reliability needed. For example SVC1_F1_array (FAStT number one), SVC1_E1_Volume Serial (ESS number one), etc.

To answer the second question, which mDisk group to use, keep in mind that an individual virtual disk is a member of one MDG and one I/O group:

- The MDG defines which managed disks provide the back-end storage which makes up the virtual disk.
- The I/O group defines which SVC Nodes provide I/O access to the virtual disk.

Note: There is no fixed relationship between I/O groups and MDGs.

Therefore, you define all the virtual disks to:

- Optimize the performance between the hosts and SVC by repartitioning the vDisks between the different nodes of the SVC cluster.
- Get the level of performance, reliability, and capacity you require by using the MDG that corresponds to your needs (you can access any MDG from any nodes).
- Each write is copied into the cache of the two nodes before any acknowledgment is sent to the host.

3.5.7 I/O handling and offline conditions

For a virtual disk to be online, *all* managed disks in the MDG or MDGs must be associated with the virtual disk to be online. This applies to image mode virtual disks and managed mode virtual disks. A virtual disk is offline if any managed disk in the MDG is offline even if that managed disk does not contribute any extents to the virtual disk in question or the managed disk has no allocated extents.

> **Note:** Normally a virtual disk is associated with just one MDG. However, for the duration of a migration between MDGs, the virtual disk is associated with two MDGs. In this case, the offline rules apply to both MDGs for the duration of the migration only.

Referring back to Figure 3-13 on page 49, this means that if managed disk M1 is taken offline by RAID controller A, virtual disks V1 and V2 are taken offline by SVC.

This notion of offline and online is managed on a node basis. Therefore, if a condition arises that causes one SVC node to see a managed disk offline, then the affected virtual disks are taken offline on that node only.

For example, again refer to Figure 3-13 on page 49. If the SAN connection between RAID Controller B and SVC Node P were to fail, then Node P would lose contact with managed disks M3 and M4. Since M3 is in MDG X and M4 is in MDG Y, this takes offline on Node P all virtual disks in MDGs X and Y. Therefore, hosts accessing Node P see virtual disk V2 go offline. Hosts accessing V2 via Node Q continue to see the virtual disk as online.

3.5.8 Quorum disks

The cluster must contain at least half of its nodes to function. A tie break situation can occur if exactly half the nodes in the cluster fail at the same time or if the cluster is divided so that exactly half the nodes in the cluster cannot communicate with the other half.

For example, in a cluster of four nodes, if any two nodes fail at the same time, or any two cannot communicate with the other two, a tie-break condition exists and must be resolved. To resolve the tie-break condition a quorum disk is used. The cluster automatically chooses three managed disks to be quorum disks. One of these disks is used to settle a tie-break condition. If a tie-break condition occurs, the first half of the cluster to access one of the quorum disk after the split has occurred locks the disk and continues to operate. The other side stops. This action prevents both sides from becoming inconsistent with each other.

3.5.9 Virtualization operations on virtual disks

You can perform several operations on virtual disks as explained in the following sections.

Expanding a virtual disk

A virtual disk can be expanded. The granularity of expansion is one block (512 bytes). If the expansion requires the allocation of additional extents then these are allocated to the virtual disk from the managed disks specified using the algorithm described in “Allocation of free extents” on page 52. Expanding a virtual disk using the sequential policy forces the virtualization policy to be changed to striped.

A virtual disk can also be expanded by a single extent. This gives the user the ability to expand the virtual disk by selecting Individual managed disk extents, allowing any desired mapping from virtual to managed extents to be created.

A security clear feature is provided to allow the resulting additional space to be set to all zeros.

Image mode virtual disks *cannot* be expanded. They must first be migrated to managed mode.

Note: Not all operating systems can tolerate the expansion of a virtual disk. A reboot or remount of the disk may be needed to use the additional space.

Reducing a virtual disk

A virtual disk can be shrunk. The granularity of shrinking is one block (512 bytes). If the shrink operation allows extents to be freed, then these are returned to the pool of free extents for allocation by later virtualization and migration operations.

Image mode virtual disks *cannot* be reduced in size. They must first be migrated to managed mode.

Warning: Not all operating systems can tolerate a virtual disk being reduced in size. You *must* be cautious and know where data resides or data loss can occur.

Deleting a virtual disk

A virtual disk can be deleted. When a virtual disk is deleted, all host mappings are deleted and any cached read or write data is discarded. Also, any FlashCopy mappings or remote copy relationships in which the disk is participating are also deleted.

If the virtual disk was operating in managed mode, then the extents are returned to the pool of free extents for allocation by later virtualization operations. If the virtual disk was an image mode virtual disk, deleting the virtual disk causes the managed disk to be ejected from the MDG. The mode of the managed disk is returned to “unmanaged” mode. This makes the delete operation the inverse of the create operation for image mode disks.

3.5.10 Creating an mDisk group (extent size rules)

There are several guidelines or rules that you must follow when creating an mDisk group.

Number of MDGs

The number of mDisk groups depends on the following factors:

- The need for image mode virtual disks (data migration)
- The need for managed mode virtual disks with sequential policy
- The models of the back-end controller (RAID controller with cache or without, Fast, ESS, Others) that have different properties on performance, availability, response time, etc.

It is possible to have a common MDG for all the SVC cluster. However, a virtual disk (vDisk) is offline if any managed disk in the MDG is offline, even if that managed disk does not contribute any extents to the virtual disk in question or the managed Disk has no allocated extents. The more managed disks there are in a MDG, the more your vDisk (host LUN) is stripped and the better the performance is.

We recommend that you:

- Create at least one separate MDG for all the image mode virtual disks.
- Create one separate MDG for each array (or RAID) type presented from a RAID controller.

  Note: It may be prudent to keep each back-end storage subsystem in a separate mDisk group. This prevents a failure in storage subsystem A from affecting storage subsystem B. If a vDisk is composed of mDisks from both A and B, then a failure in either A or B causes the vDisk to be unavailable.

- You can name them in way that it is easier for you when you create a virtual disk to associate it with the MDG with the level of performance and reliability needed. Consider these examples: pool1_high_perf_high_rela, pool2_grp_low_perf_low_rela, mDisk_grp_ESS1, mDisk_grp_FAStT2, and mDisk_grp_raid10.

Size of extent

If you want to migrate a vDisk from one mDisk group to another mDisk group, the extent size must be the same between the two mDisk groups. Because of this, it can be useful to set a common extent size for all the mDisk groups. A value of 32 MB (128 TB) or 64 MB (256 TB) can be a best trade-off between performance and capacity.

Configuration steps

The parameters needed to create a MDG are:

- The name you want to assign to the MDG
- List of the managed disk you want to include
- The extent size you want to use

The operations are:

1. Create a MDG.
2. Add a managed disk to the MDG.

3.5.11 Creating a managed disk

First you need to create the logical disks (LUNs) in your back-end storage. We recommend that you use the maximum LUN size to be presented as an mDisk. The discovery of the managed disk is automatically done by the SVC. The managed disk is in unmanaged mode until you include it in a MDG.

You need at least one managed disk for the support of the quorum disk used in the cluster. All SVC nodes must have access at any time to all the managed disks.

The size of the managed disk can be up to 2 TB. You can use some common sizes for all the managed disks (16 GB, 32 GB). That helps for simplicity and to ensure as much as possible that all the mDisk are used in the striping process for a managed disk with striped policy. If you have three managed disks, two of 4 GB and one of 210 GB, very quickly only the disk of 210 GB is used in the striping process.

For image mode and managed mode with sequential policy virtual disks, you must create managed disks (LUN in the back-end controller), at a minimum, of the same size that the origin disk you want to migrate (for the image mode) or you want to copy (for the managed mode with sequential policy). If no extents are available, the creation of the virtual disks fail.

Configuration steps

Define the logical or physical disks (logical units) in the back-end storage.

When you include a managed disk into a MDG, change the mode from unmanaged to *managed*.

3.5.12 Creating a virtual disk

An individual virtual disk is a member of one MDG and one I/O group. The MDG defines which managed disks provide the back-end storage that makes up the virtual disk. The I/O group defines which SVC Nodes provide I/O access to the virtual disk.

Note: There is no fixed relationship between I/O groups and MDGs.

Define all the virtual disks to:

- Optimize the performance between the hosts and SVC by repartitioning the vDisks between the different nodes of the SVC cluster.
- Get the level of performance, reliability, and capacity that you require by using the MDG that corresponds to your needs. You can access any MDG from any node.

When you create a vDisk, it is associated to one node of an I/O group. By default, every time you create a new vDisk, it is associated to the next node using a round robin algorithm. For example, you may have four hosts (host1, host2, host3, host4) with 100 vDisks for each host of the same size with the same level of I/O activity and a four node (two I/O groups) cluster. The result is 100 vDisks on each node (25 vDisks from host1, 25 vDisks from host2, etc.).

You can specify a preferred access node. This is the node through which to send I/O to the vDisk instead of using the round robin algorithm. Consider the example of one host with four vDisks (VD1, VD2, VD3, and VD4). VD1 and VD3 have a high level of I/O activity. VD2 and VD4 have a low level of I/O activity. If you use the round robin algorithm, VD1 and VD3 are on the same node 1 of the I/O group, and VD2 and VD4 are on the same node 2 of the I/O group. To avoid this, use the preferred node feature to specify VD1 on node 1 and VD3 on node 2. We recommend that you use the *preferred node* for I/O option when you create your vDisks. See Figure 6-36 on page 129.

A virtual disk is defined for an I/O group which provides the following benefits:

- The vDisk is “exported” by the two nodes of the I/O group to the host via eight paths (four paths for each node). We use zoning to limit it to two paths from each node.
- Each write is copied into the cache of the two nodes before acknowledgment is sent to the host.

Even if you have eight paths for each virtual disk, all I/O traffic flows only toward one node (the preferred node). Therefore, only four paths are really used by SDD. The other four are used only in case of a failure of the preferred node.

Before you create a virtual disk, you can check the amount of space that is available in the back-end storage. You can determine the free capacity for an mDisk or an mDisk group as shown in Example 3-2.

Example 3-2 lsmdiskgrp command

```
IBM_2145:admin>svcinfo lsmdiskgrp ARR36P5N
id 0
name ARR36P5N
status online
mDisk_count 2
vDisk_count 0
capacity 333.9GB
extent_size 32
free_capacity 333.9GB
```

Creating image mode virtual disks

Use image mode virtual disks when a managed disk already has data on it, from a previrtualized subsystem. When an image mode virtual disk is created, it directly corresponds to the managed disk from which it is created. Therefore, virtual disk LBA x = managed disk.

When you create an image mode, the managed disk must have a mode of *unmanaged* and, therefore, does not belong to any MDG. A capacity of 0 is not allowed. Image mode virtual disks may be created in sizes with a minimum granularity of 512 bytes, and must be at least one block (512 bytes) in size.

The SVC may reserve an integer number of extents on which to hold the image mode disk. It effectively rounds up its size to the nearest whole number of extents.

Creating managed mode virtual disks with sequential policy

When you create a managed mode virtual disk with sequential policy, you must use a complete new empty managed disk with a size equal or greater than the size of the virtual disk you want to create. This is due to the fact that, even if you know the number of free extents in a managed disk, you do not know if they are sequential.

3.5.13 Maximum number of virtual disks and total capacity per node

This section discusses the relationship between the number of virtual disks and the total capacity per node.

Modifying a virtual disk

You can change the I/O group with which a virtual disk is associated. This requires a flush of the cache within the nodes in the current I/O group to ensure that all data is written to disk. I/O should be suspended at the host level before you perform this operation.

You can set the I/O governing rate, a cap on the amount of I/O that is accepted for this virtual disk. You can set it in terms of I/O per second or MBs per second. By default, no I/O governing is set when a virtual disk is created.

Configuration steps

The parameters that you need to create a virtual disk are:

- The name you want to assign to the virtual disk
- The I/O group you want to use
- The MDG you want to use
- The capacity you want to define

The operations are:

1. Create a managed mode vDisk.
2. Create an image mode vDisk.
3. Modify the vDisk.

3.5.14 Creating a host group (LUN masking)

The maximum number of FC HBAs is four per host for dual HBAs or eight FC ports.

Configuration steps

The parameters needed to create a host and virtual disk to host mapping are:

- The name you want to assign to the host

- The list of the WWPN of the FC HBAs of the host
- The name of the vDisks you want to assign to the host

The operations are:

1. Create a host.
2. Create a vDisk to host mapping.

3.5.15 Expanding a SVC cluster configuration

You can expand a SVC cluster configuration as explained in the following sections.

Adding a node to a cluster

If you defined a two-node SVC cluster, you can easily add new nodes to add new hosts and back-end controller or to redistribute workload.

Adding a new disk controller to an mDisk group

mDisk groups can span controllers. We recommend that you *do not* do this. Each mDisk group should, in normal circumstance, comprise disks from one controller. See 3.5.2, “MDGs, I/O groups, virtual disks, and managed disk” on page 48.

vDisk size increase and decrease

The SVC allows to increase and decrease the size of vDisks. Not all operating systems allow this. See “Expanding a virtual disk” on page 54.

3.5.16 Migration

This facility allows the mapping of virtual disk extents to managed disk extents to be changed, without interrupting a host’s access to that virtual disk. You can perform this for any virtual disk managed by the SVC. You may use this for:

- Redistributing workload within a cluster across back-end storage
- Moving workload onto newly installed storage
- Moving workload off old or failing storage, ahead of decommissioning it
- Moving workload to rebalance a changed workload
- Migrating in data from legacy back-end storage to SVC managed storage

For further details about the migration facility, see Chapter 12, “Migration to the SAN Volume Controller” on page 407.

3.6 SVC supported capabilities

Table 3-2 lists the maximum values of parameters to use when you plan an installation.

Table 3-2 IBM TotalStorage SAN Volume Controller maximum values

Element	Maximum	Notes
Cluster properties		
Nodes	4	Arranged as two pairs; always two nodes per SVC I/O groups
I/O groups	4	Fixed number; currently only two are usable

Element	Maximum	Notes
Virtual disks (vDisks)	vDisks/Cluster 1024 Size 2 TB SCSI/mappings/ cluster 20 K vDisk paths/host 8	Maximum number of vDisk per (Linux and AIX: 1 TB) Separately zoned host port
mDisk groups or cluster	128	
mDisks	4096	Represents an average of 64 per controller
mDisks per mDisk group	128	
mDisk size	2 TB	Defined by a 32-bit LBA limit
Host properties		
Host LUNs	4096	This includes virtualized LUNs and image LUNs.
Host ports	128	Up to 128 distinct host WWPNs are recognized per cluster.
SCSI LUNs presented to a given host ID	4096	This is the maximum number of LUNs mapped to a single host ID, and therefore, any host port. SCSI LUNs must be in range 0 to 4095.
Virtual disks to SCSI LUN/host ID mappings	20000	This is the total number of map table entries. On average, each host LUN can be mapped approximately 2.5 times to distinct host IDs.
Host LUN size	2 TB	Defined by a 32-bit LBA limit
RAID controller properties		
RAID controllers	64	
Worldwide node name (WWNN) per controller WWPN per controller LUs per controller	2 16 4096	
SAN properties		
SAN ports	256	Maximum SAN ports per fabric
Maximum clusters per fabric	2	
RAID controller	64	A maximum of 128 ports
ISL hops in a fabric	3	Local or remote fabric should not contain more than three ISL hops
ISL hops from host to SVC node	3	
ISL hops from RAID controller to SVC node	3	
ISL hops from local cluster to remote cluster	7	

3.6.1 Adding ESS storage to the SVC

Perform the following steps:

1. Go to the Web site and check the prerequisites:

   `http://www.ibm.com/storage/support/2145`

2. You may need to upgrade the microcode level of the ESS to support this attachment.
3. Before the mDisks can be presented to the SVC, define the ESS ports to the storage zone to all SVC node ports. See Example 3-1 on page 38.
4. Sign on the ESS specialist. On the left side of the window that opens, select **Storage Allocation**. Figure 3-15 shows the Storage Allocation window.
5. On the Storage Allocation -- Graphical View panel, click **Open System Storage**.

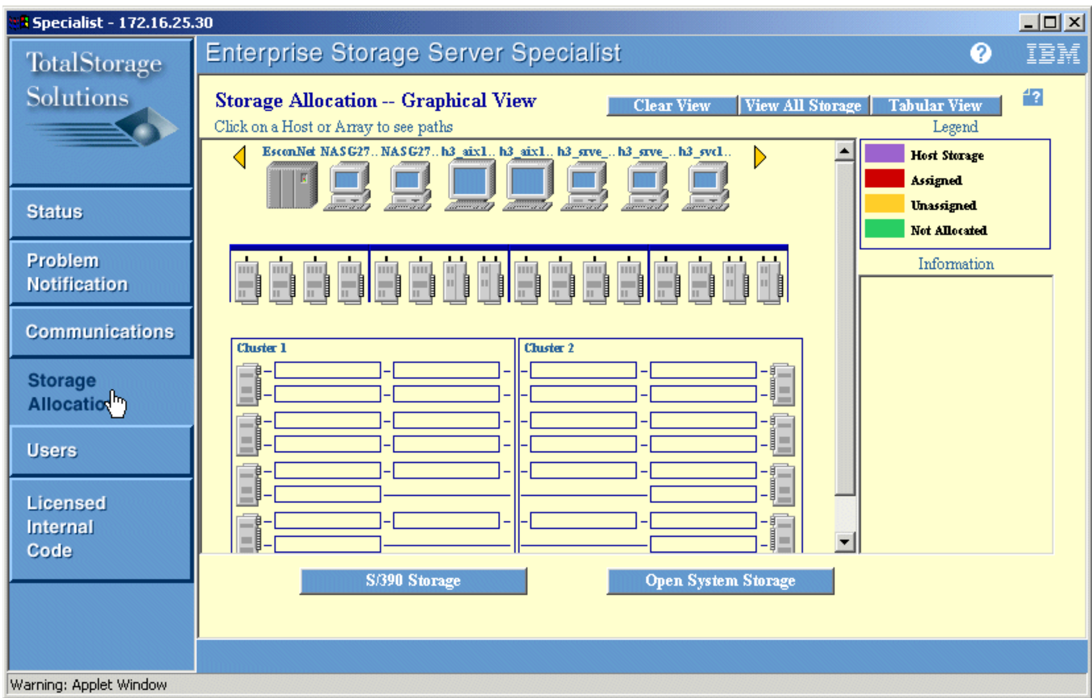


Figure 3-15 Storage Allocation window

6. Go to the Configure Host Adapter Ports window (Figure 3-16). Configure the host bay ports in the ESS if you haven't already. Select the port or ports. For Fibre Channel Topology, select **Point to Point (Switched Fabric)**. For Fibre Channel Protocol, select **FCP (Open Systems)**.

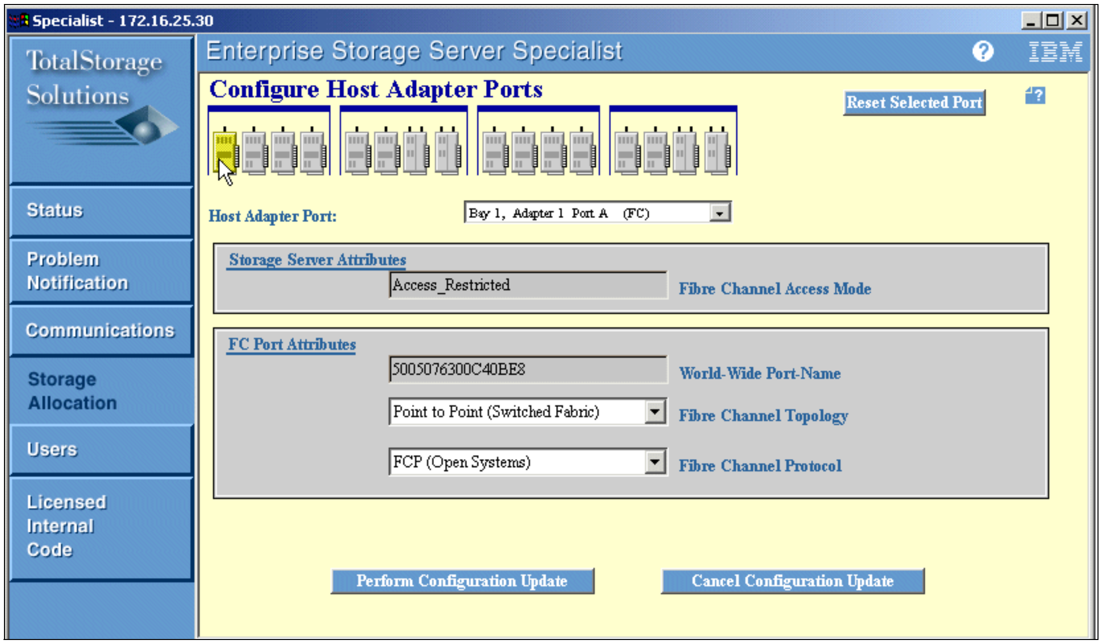


Figure 3-16 Configure Host Adapter Ports window

7. Go to the Modify Host Systems window (Figure 3-17). Follow these steps:

   a. Enter the host system nickname, which is case `h3_svclnode1_a`, in our case.
   b. Select its WWPN from the list of WWPNs or type it in manually.
   c. Select the Fibre Channel ports that you want to see the node.

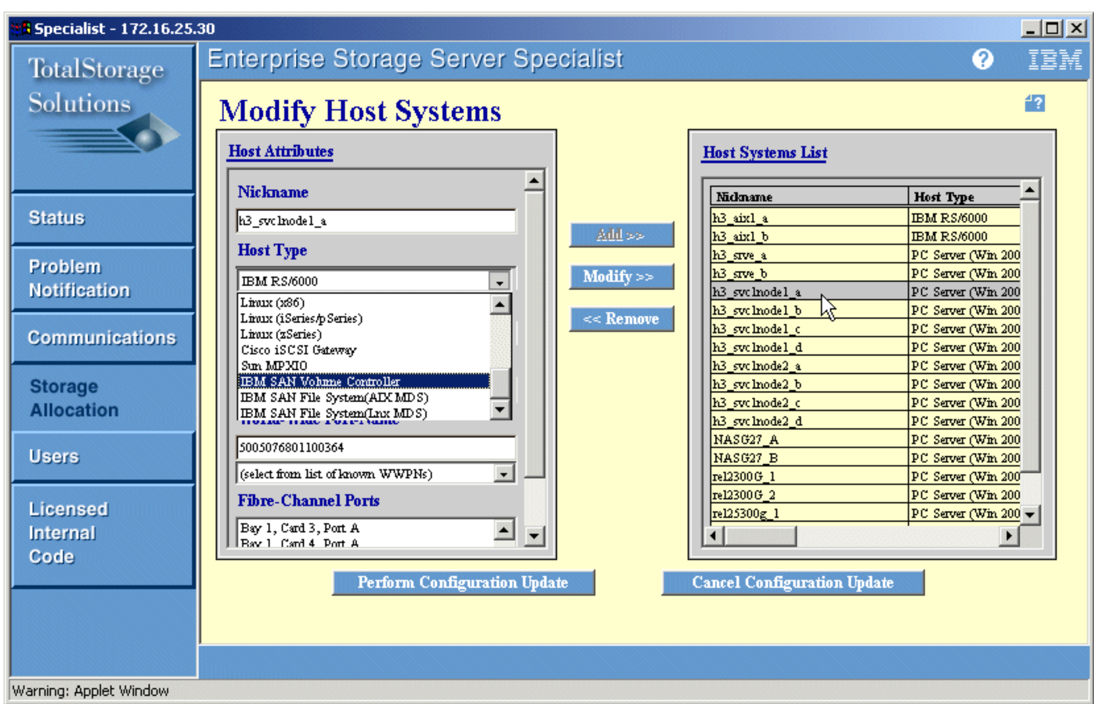


Figure 3-17 ESS Modify Host System window

We defined all four node ports for both nodes to bay1, card 1 port A, and bay 3, card 1 port A. This distributes access across two bays in the ESS for redundancy. See Figure 3-18.

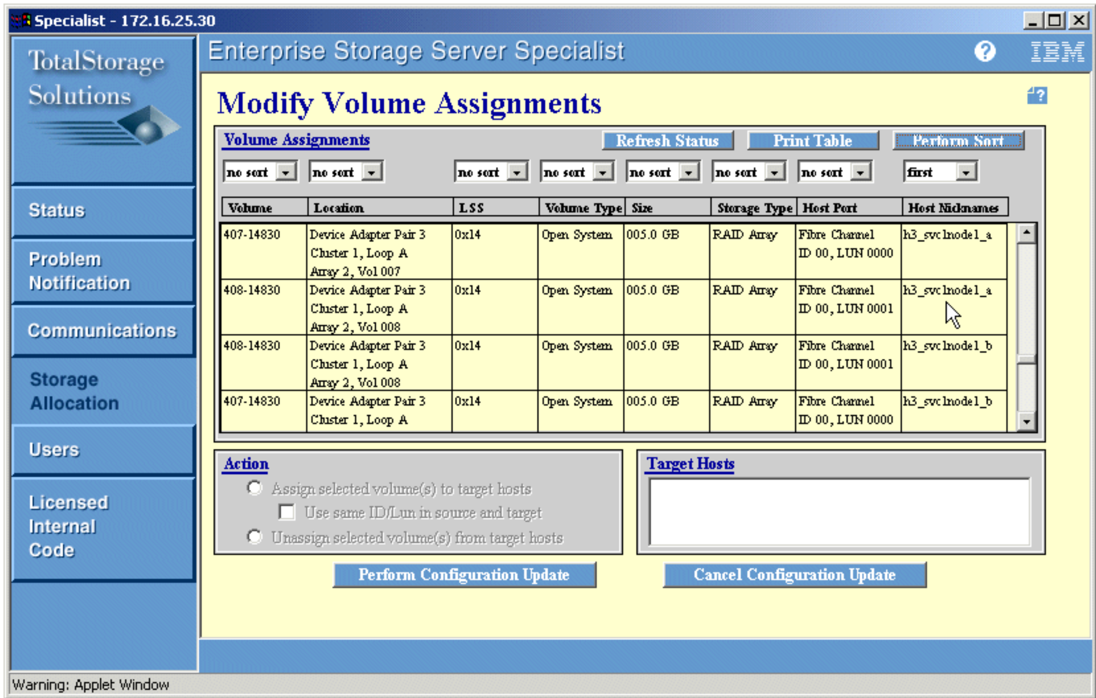


Figure 3-18 Volume assignments

8. Share the volumes across both ports in the ESS. Now the volumes are presented to the SVC on both ports. You can find the mDisks on the SVC and rename them to a unique name identifying the origin. Consider the example of E1_407_14830 - ESS number 1,_volume serial number in ESS 407-14830. You can add the mDisk to a MDG called ESS_14830, for example. See Figure 3-19.

> **Note:** The LUN size is normally a full array size that is presented to the SVC as an mDisk.

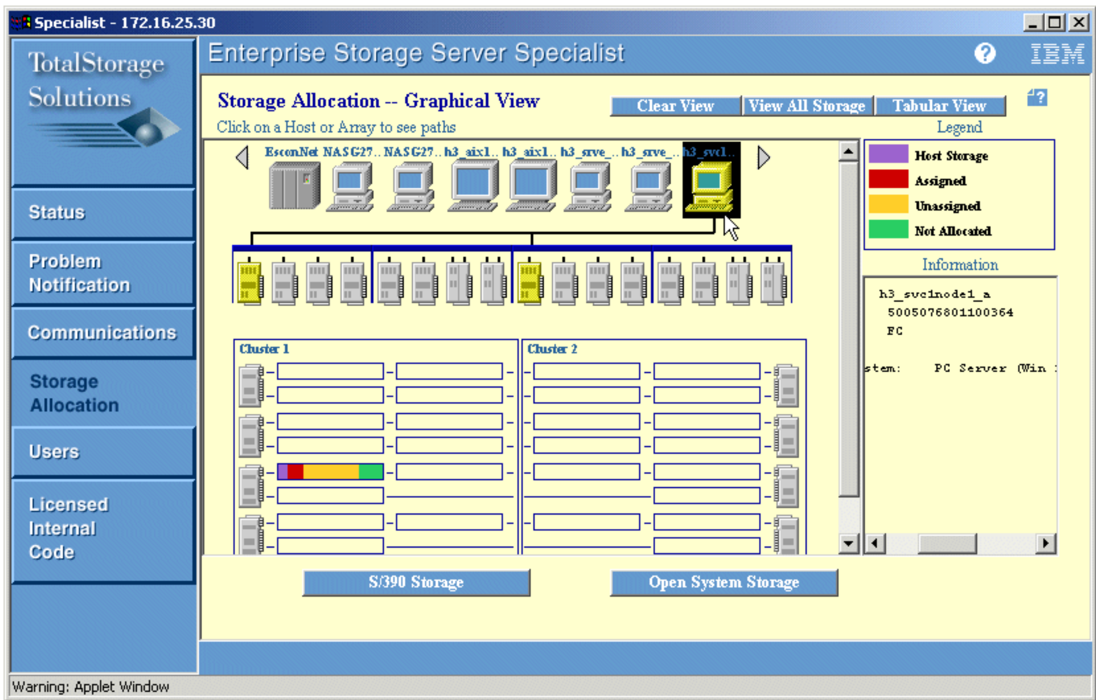


Figure 3-19 Viewing the two paths

3.6.2 Adding FAStT storage to the SVC

To add FAStT storage to the SVC, follow these steps:

1. Check the prerequisites on the Web at:

   `http://www.ibm.com/storage/support/2145`

2. Check the supported firmware levels and configurations before you connect to the SVC. See Figure 3-20.

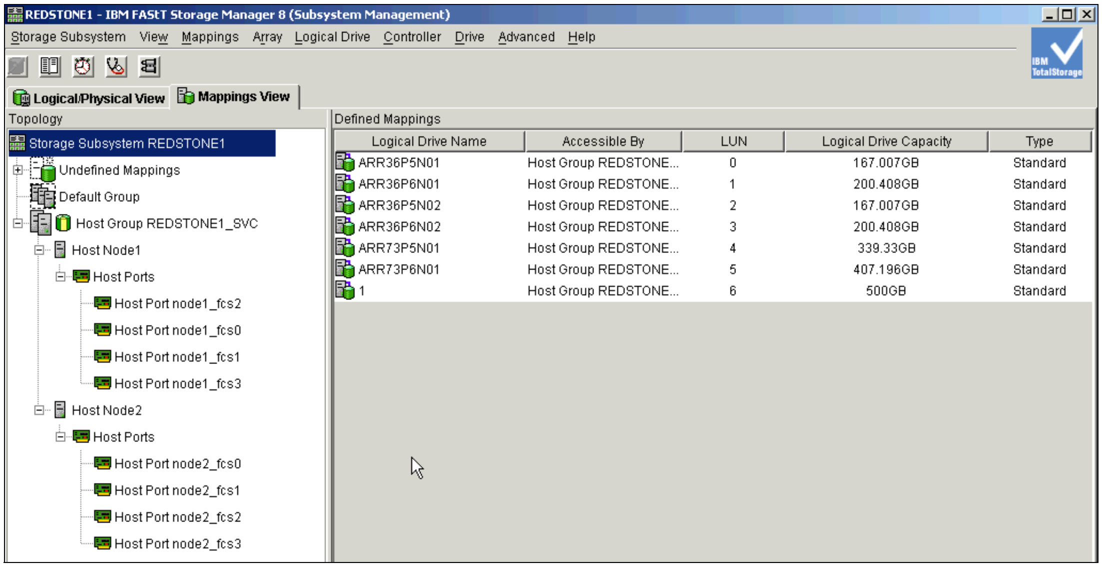


Figure 3-20 FAStT mappings view

3. We defined one storage partition in the FAStT with all of the SVC node ports defined, with one host partition with eight ports. See Figure 3-21.

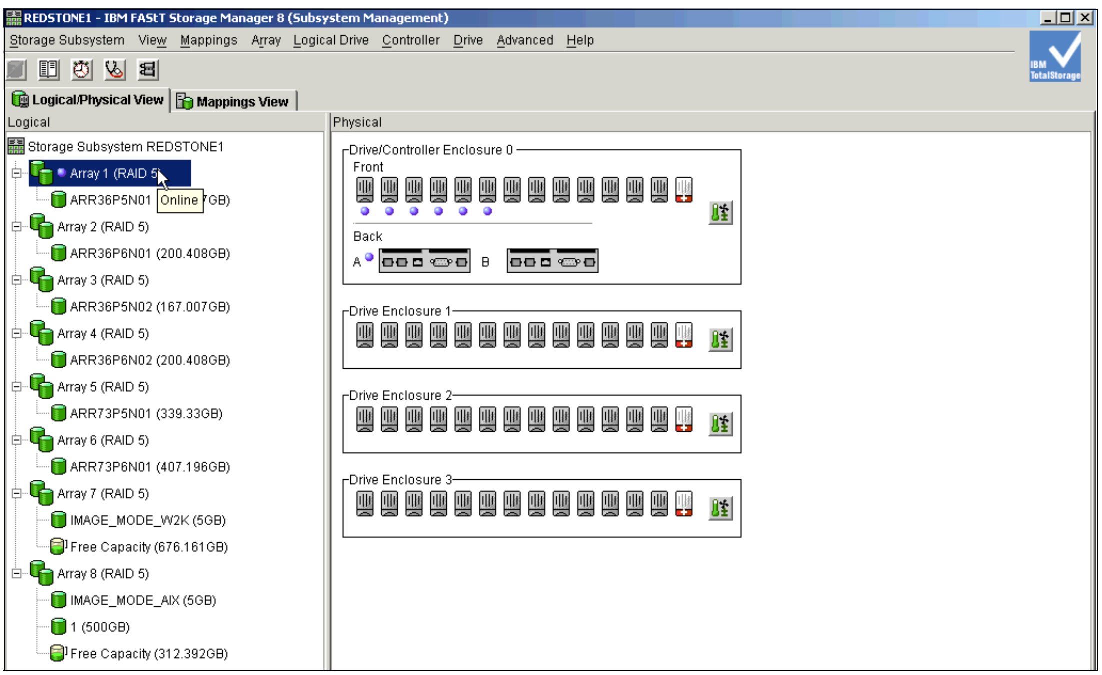


Figure 3-21 Array 1

These arrays were created with names that reflect the quantity and size of the physical drives. This may be important to reflect performance and size when you create mDisk groups. Our opinion is that it is more important to map back the mDisk names to the physical RAID control. Therefore, you can use F1_Array_number_LUNname. The arrays should alternate between controller A and controller B for the preferred path. See Figure 3-22.

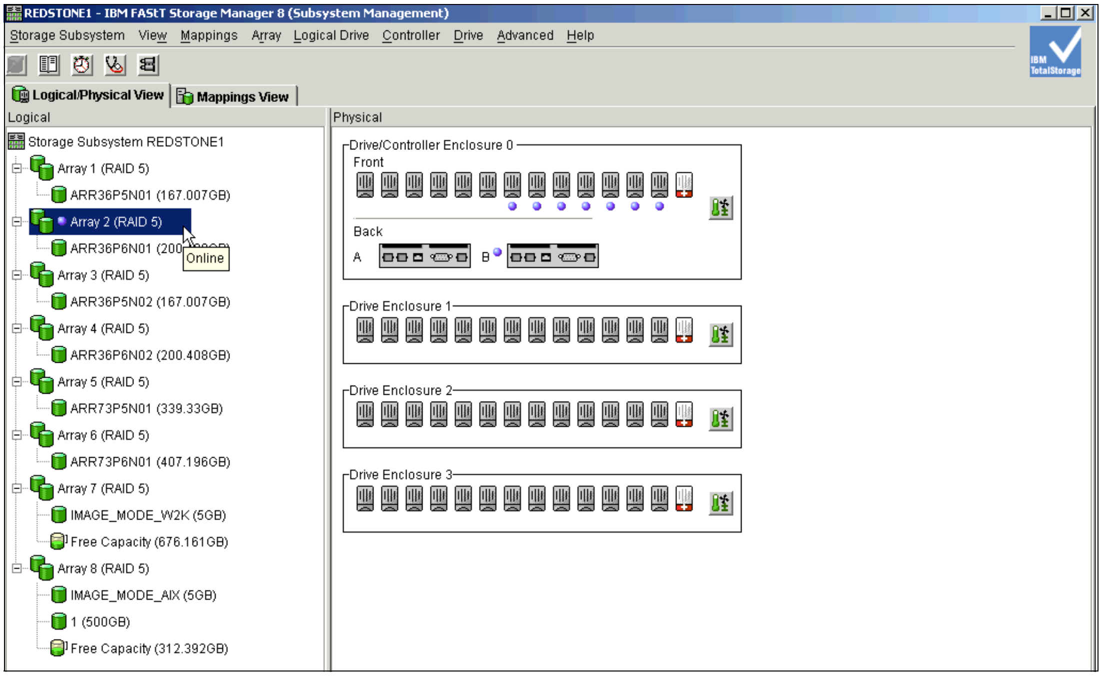


Figure 3-22 Array 2

Figure 3-23 shows the host type for SVC.

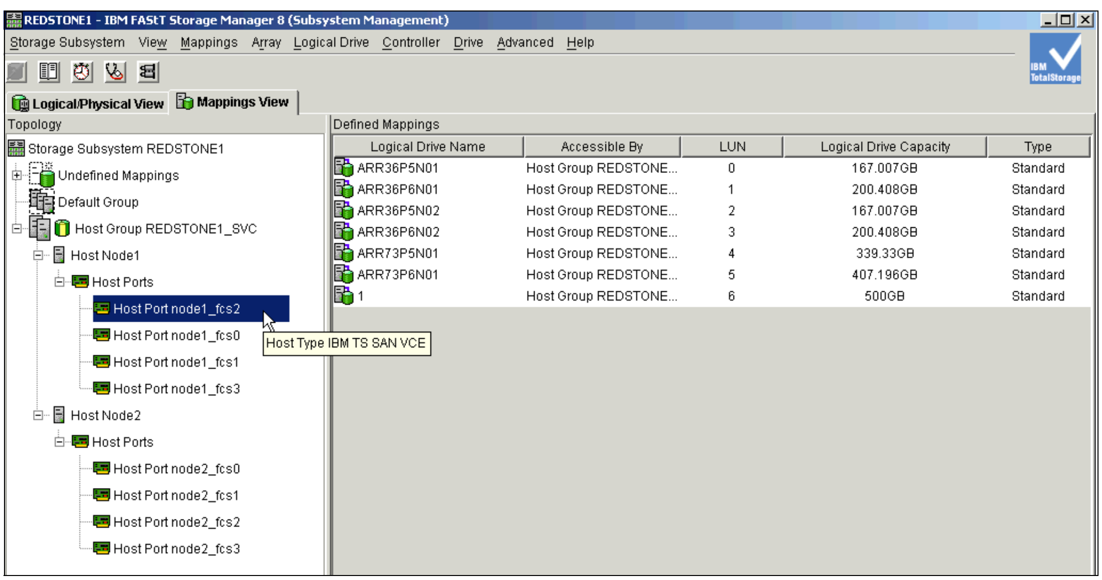


Figure 3-23 Host type for storage partition

Figure 3-24 shows the port mapping. Now the volumes are presented to the SVC on both ports. You can find the mDisks on the SVC and rename them to a unique name identifying the origin. Consider the example of F1_Array_LUN - FAStT number 1,_array number_LUN number. You can add the mDisk to an mDisk group called FAStT_14830, for example.

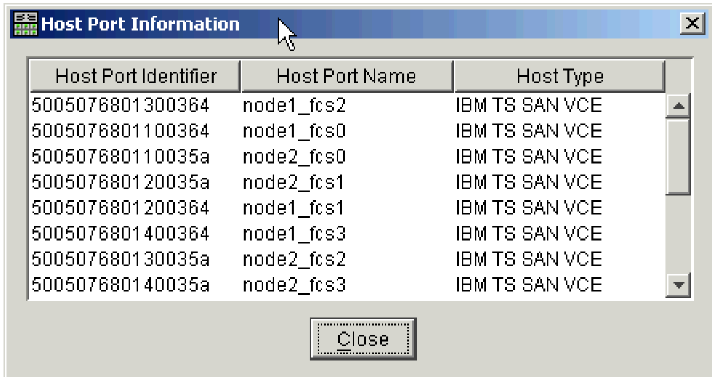


Figure 3-24 Port mapping

4

Initial installation and configuration of the SVC

This chapter describes the initial installation and configuration procedures for the IBM TotalStorage SAN Volume Controller (SVC) using the service panel and the cluster Web interface.

Note: The service panel consists of the display window and buttons on the front of each SVC node.

4.1 Preparing for installation

See Chapter 3, “Planning and configuration” on page 25, for information pertaining to physical connectivity, storage area network (SAN) zoning, and assigning disk to the SVC.

4.2 Secure Shell overview

Secure Shell (SSH) is used to secure data flow between SVC cluster (SSH server) and a client (either a command line client (command line interface (CLI) or the CIMOM)). The connection is secured by the means of a private or public key pair:

- A public key is uploaded to the SSH server
- A private key identifies the client and is checked against the public key during the connection. The private key must be protected.
- The two keys are generated together.
- The SSH server must also identify itself with a specific host key.
- If the client does not have that key yet, it is added to a list of known hosts.

Secure Shell is the communication vehicle between the management system (usually the master console) and the SVC cluster.

SSH is a client-server network application. The SVC cluster acts as the SSH server in this relationship. The SSH client provides a secure environment on which to connect to a remote machine. It uses the principles of public and private keys for authentication.

When an SSH client (A) attempts to connect to a server (B), a key is needed to authenticate the connection. The key consists of two halves: the public and private keys. The public key is put onto (B). When (A) tries to connect, the private key on (A) can authenticate with its public half on (B).

SSH keys are generated by the SSH client software. This includes a public key, which is uploaded and maintained by the cluster, and a private key that is kept private to the workstation that is running the SSH client. These keys authorize specific users to access the administration and service functions on the cluster. Each key is associated with a user-defined ID string that can consist of up to 40 characters. Up to 100 keys can be stored on the cluster. You can also add new IDs and keys or delete unwanted IDs and keys.

To use the CLI or SVC graphical user interface (GUI), you must have an SSH client installed on that system, generate the SSH key pair on the client system and store the client’s SSH public key on the SVC cluster or clusters.

The SVC master console has the free implementation of SSH-2 for Windows called PuTTY pre-installed. This software provides the SSH client function for users logged into the master console who want to invoke the CLI or GUI the managed the SVC cluster.

4.2.1 Generating public and private SSH key pair using PuTTY

Perform the following steps to generate SSH keys on the SSH client system (master console):

> **Note:** We use these keys when we reach 4.4.1, “Uploading the SSH public key to the SVC cluster” on page 82.

1. Start the PuTTY Key Generator to generate public and private SSH keys. From your desktop, select **Start -> Programs -> PuTTY -> PuTTYgen**.
2. On the PuTTY Key Generator GUI window (Figure 4-1), generate the keys:
   a. Select the **SSH2 RSA** radio button.
   b. Leave the number of bits in a generated key value at `1024`.
   c. Click **Generate**.

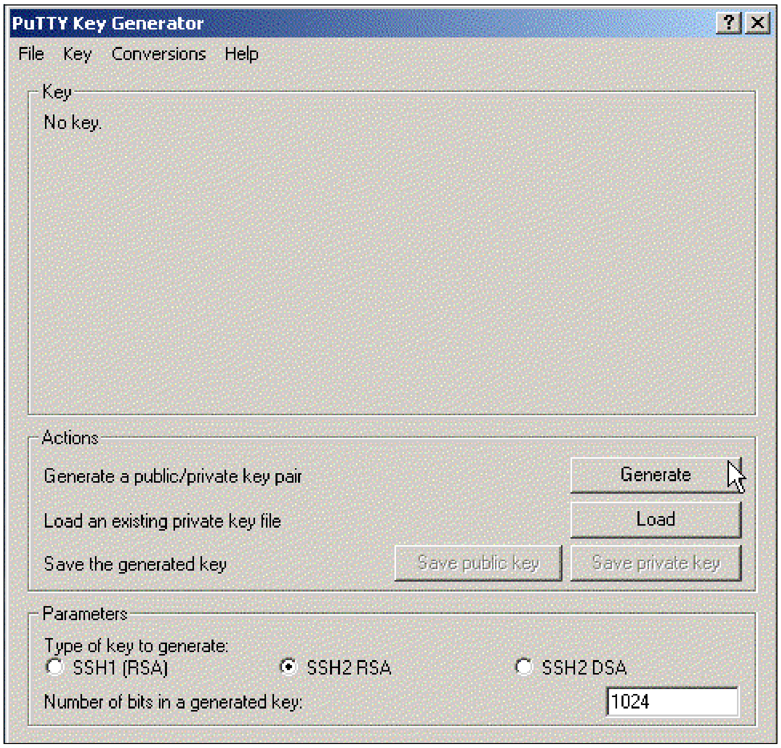


Figure 4-1 PuTTY key generator GUI

3. The message in the Key section of the window changes. Figure 4-2 shows this message.

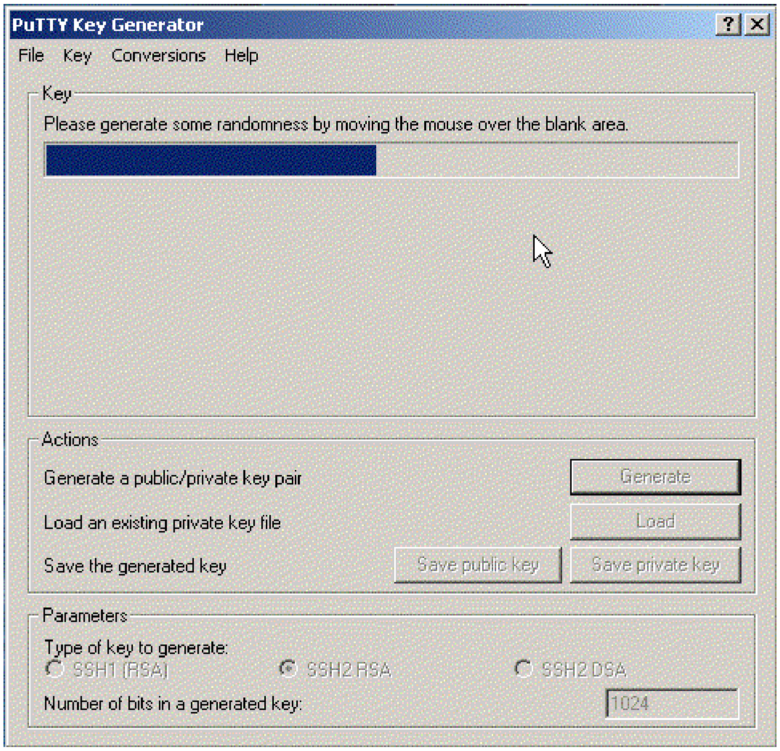


Figure 4-2 PuTTY random key generation

> **Note:** The blank area indicated by the message is the large blank rectangle on the GUI inside the section of the GUI labelled *Key*. Continue to move the mouse pointer over the blank area until the progress bar reaches the far right. This generates random characters to create a unique key pair.

4. After the keys are generated, save them for later use as follows:

   a. Click **Save public key** as shown in Figure 4-3.

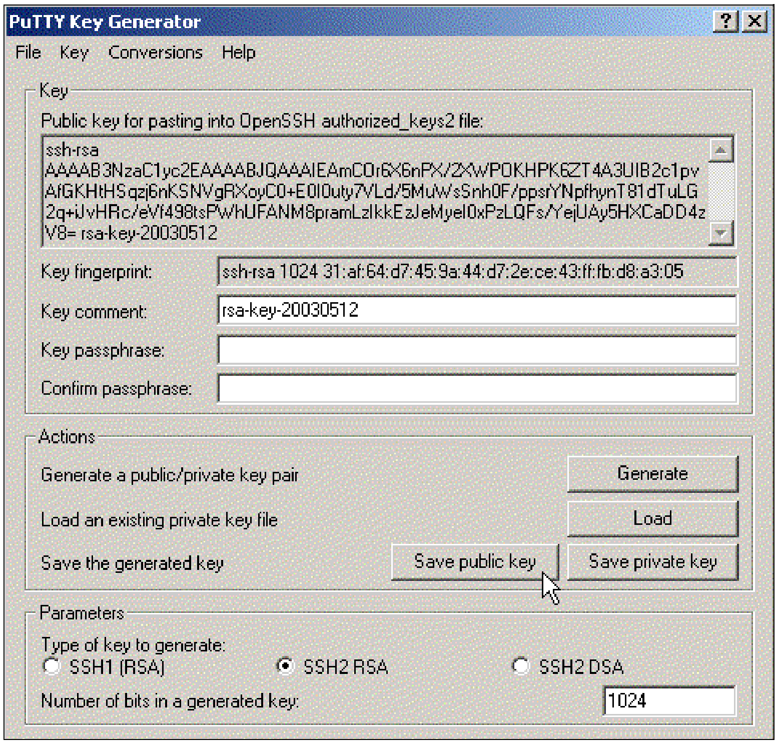


Figure 4-3 Saving the public key

   b. You are prompted for a name (for example, pubkey) and a location for the public key (for example, C:\Support Utils\PuTTY). Click **Save**.

      If you choose another name or location, ensure you have a record of them because you must identify the name and location of the SSH public key when you reach 4.4.1, “Uploading the SSH public key to the SVC cluster” on page 82.

      Note: The PuTTY Key Generator saves the public key with no extension by default.

   c. In the PuTTY Key Generator window, click **Save private key**.

   d. You are prompted with a warning message as shown in Figure 4-4. Click **Yes** to save the private key without a passphrase.

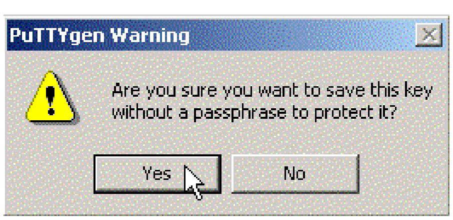


Figure 4-4 Saving the private key without passphrase

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

If you chose another name or location, ensure you have a record of it because you must identify the name and location of the SSH private key when you configure the PuTTY session as in Chapter 5, “Quickstart configuration using the CLI” on page 85.

> **Note:** The PuTTY Key Generator saves the private key with the PPK extension.

5. Close the PuTTY Key Generator GUI.
6. Using Windows Explorer on your master console, navigate to the directory where you saved your private key (for example, C:\Support Utils\PuTTY).
7. Copy the private key file (for example, icat.ppk) to the C:\Program Files\IBM\svcconsole\cimom directory.

> **Important:** If you named your private key something other than icat.ppk, make sure you rename it to `icat.ppk` in C:\Program Files\IBM\svcconsole\cimom folder. The GUI (which we use later) expects the file to be called icat.ppk and for it to be in this location.

4.3 Basic installation

This section provides step-by-step instructions for building your SVC cluster initially.

4.3.1 Creating the cluster (first time) using the service panel

This section provides the step-by-step instructions you need to create the cluster for the first time using the service panel. Use Figure 4-5 as a reference for the buttons you need to push in the steps that follow.

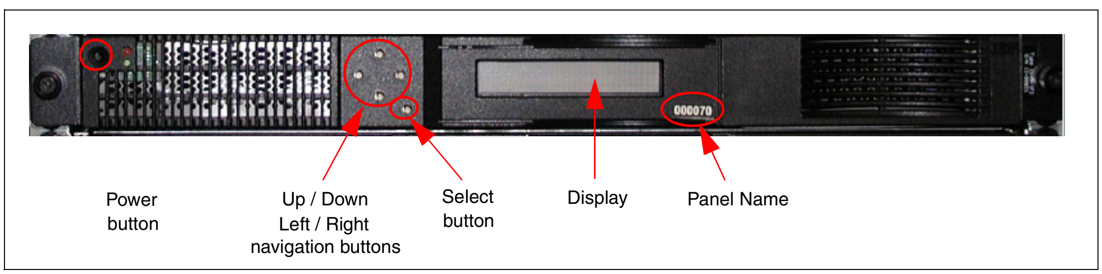


Figure 4-5 SVC service front panel

Prerequisites

Ensure that the SVC nodes are physically installed. Prior to configuring the cluster, ensure that you have the following information:

- License: The license indicates whether you are permitted to use FlashCopy, PPRC, or both. It also indicates how much capacity you are licensed to virtualize.
- Cluster IP addresses (one for the cluster and other for service access)
- Subnet IP mask
- Gateway IP address

Process

After the hardware is physically installed into racks, complete the following steps to initially configure the cluster through the service panel:

1. Choose any node that is to become a member of the cluster that you are creating.
2. At the service panel of that node, click and release the **Up** or **Down** navigation button continuously until *Node:* is displayed.

> **Important:** If a time-out occurs when making the input for the fields during these steps, you must begin from step 2. All the changes are lost, so be sure to have all the information in hand before you begin.

3. Click and release the **Left** or **Right** navigation button continuously until *Create Cluster?* is displayed.
4. Click the **Select** button. If *IP Address:* is displayed on line 1 of the screen, go to step 5. If *Delete Cluster?* is displayed in line 1 of the service display, this node is already a member of a cluster. Either you selected the wrong node, or this node was already used in a previous cluster. The ID of this existing cluster is displayed in line 2 of the service display.
   a. If you selected the wrong node, you can exit this procedure by clicking the **Left**, **Right**, **Up**, or **Down** button (it cancels automatically after 60 seconds).
   b. If you are sure that the existing cluster is not required, follow these steps:
      i. Click and hold the **Up** button.
      ii. Click and release the **Select** button. This deletes the cluster information from the node. Go back to step 1 and start again.

> **Attention:** When a cluster is deleted, all client data that is contained in that cluster is lost.

5. Click the **Select** button.
6. Use the **Up** or **Down** navigation button to change the value of the first field of the IP Address to the value that you have chosen.

> **Note:** Pressing and holding the **Up** or **Down** buttons will increment or decrease the IP address field by 10s. The field value rotates from 0 to 255 with **Down** button, and from 255 to 0 with **Up** button.

7. Use the **Right** navigation button to move to the next field. Use the **Up** or **Down** navigation buttons to change the value of this field.
8. Repeat step 7 for each of the remaining fields of the IP address.
9. When you changed the last field of the IP address, click the **Select** button.
10. Click the **Right** button. Subnet Mask: is displayed.
11. Click the **Select** button.
12. Change the fields for Subnet Mask in the same way that you changed the IP address fields.
13. When you changed the last field of Subnet Mask, click the **Select** button.
14. Click the **Right** navigation button. Gateway: is displayed.
15. Click the **Select** button.
16. Change the fields for Gateway in the same way that you changed the IP address fields.
17. When you completed the changes to all Gateway fields, click the **Select** button.

18. Click the **Right** navigation button until Create Now? is displayed.
19. If you are satisfied with your settings, click the **Select** navigation button.

    If you want to review your settings before you create the cluster, use the **Right** and **Left** buttons to review those settings. Make any necessary changes, return to Create Now?, and click the **Select** button.

    If the cluster is created successfully, *Password:* is displayed in line 1 of the service display panel. Line 2 contains a randomly generated password, which we use to complete the cluster configuration in the next section.

    > **Attention:** Make a *note* of this *password now*. It is *case sensitive*. The password is displayed only for approximately 60 seconds. If you do not record the password, you must start the cluster configuration procedure again.

20. After the Password: display times out, if the cluster is created successfully *Cluster:* is displayed in line 1 of the service display panel. Also, the cluster IP address is displayed on line 2 and you have successfully completed the initial creation of the cluster.

    If the cluster is not created, *Create Failed:* is displayed in line 1 of the service display. Line 2 contains an error code. Refer to the error codes that are documented in *IBM TotalStorage Virtualization Family SAN Volume Controller: Service Guide*, SC26-7542, to find the reason why the cluster creation failed and what corrective action to take.

> **Important:** At this time, do not repeat this procedure to add other nodes to the cluster. Adding nodes to the cluster is accomplished in 5.3, “Adding nodes to the cluster” on page 90, or in 6.2, “Adding nodes to the cluster” on page 105.

4.4 Completing the initial cluster setup using the Web interface

Perform the following steps to connect to the cluster:

1. Using a Web browser, access the SVC cluster. Typing the following line in the Web browser’s address field:

   `https://svcclusterip/create`

   Here *svcclusterip* is the SVC cluster IP address you configured in the service panel earlier (for example, 9.42.164.155).

   > **Important:** Because the SVC cluster interface is secure, it is required to access it as Secure HTTP (HTTPS).

2. When you see the security alert regarding certificates (Figure 4-6), select **Yes**.

Figure 4-6 Certificate Security Alert

3. The Enter Network Password window (Figure 4-7) opens. For User Name, type `admin`. For password, type the password that was shown on the service panel when you created the cluster in the previous section.

> **Note:** The password is case sensitive.

Figure 4-7 User ID and password validation

4. The browser accesses the SVC and displays the Create New Cluster wizard window as shown in Figure 4-8. Click **Continue**.

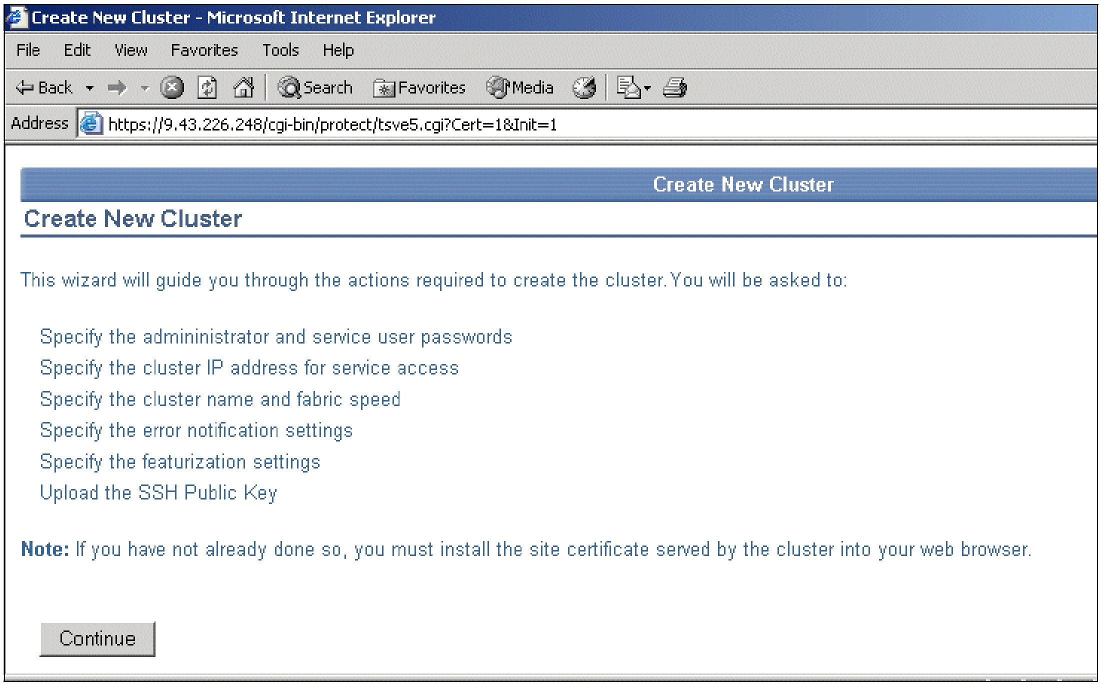


Figure 4-8 Create New Cluster wizard

5. The Create New Cluster page (Figure 4-9) opens. Fill in the following details:
   - A new admin password to replace the random one that the cluster generated: The password is case sensitive and can consist of A to Z, a to z, 0 to 9, and the underscore. It cannot start with a number (minimum of one character, maximum of 15 characters).
   - A service password to access the cluster for service operation: The password is case sensitive and can consist of A to Z, a to z, 0 to 9, and the underscore. It cannot start with a number (minimum of one character and a maximum of 15 characters).
   - A cluster name: The cluster name is case sensitive and can consist of A to Z, a to z, 0 to 9, and the underscore. It cannot start with a number (minimum of one character and a maximum of 15 characters).
   - A service IP address to access the cluster for service operations.

   > **Note:** The service IP address is different than the cluster IP address. However, because the service IP address is configured for the cluster, it must be on the same IP subnet.

   - The fabric speed of your Fibre Channel network (either 1 Gb/s or 2 Gb/s).
   - The Administrator Password Policy check box, if selected, enables you to reset the password from the service panel if you forget it. This check box is optional.

Note: Your SVC should be in a secure room if you enable this function since anyone who knows the correct key sequence can reset the admin password.

The key sequence is:

- From the Cluster: menu item displayed on the service panel, click the **Left** or **Right** button until you see Recover Cluster?.
- Click the **Select** button and you should see Service Access?.
- Click and hold the **Up** button and then click and release the **Select** button. This generates a new random password. Write it down.

Be careful because clicking and holding the Down button and clicking and releasing the Select button places the node in service mode.

Click the **Create New Cluster** button.

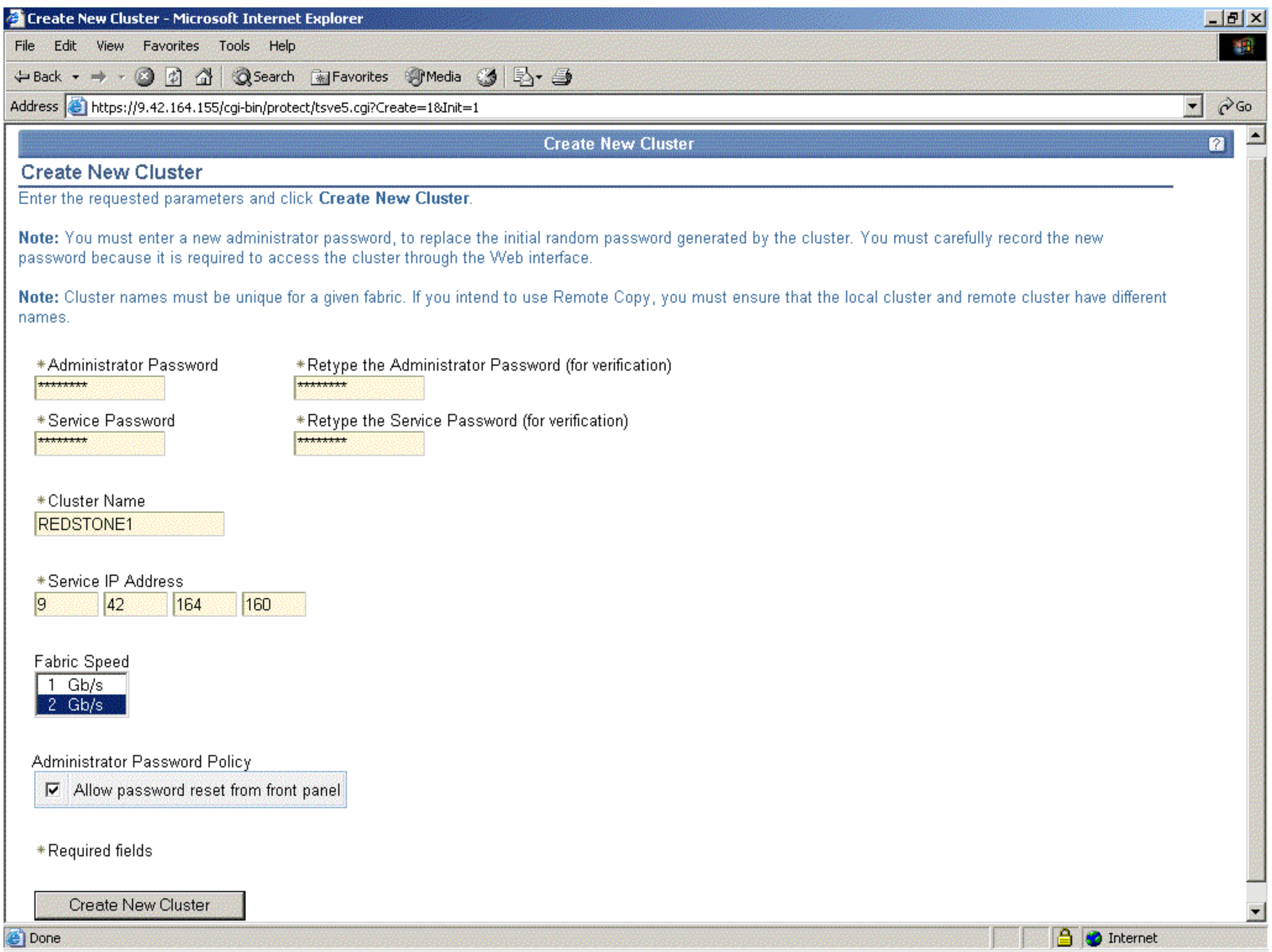


Figure 4-9 Create New Cluster Web interface

6. A number of progress windows appear as shown in Figure 4-10. Click **Continue** each time you are prompted.

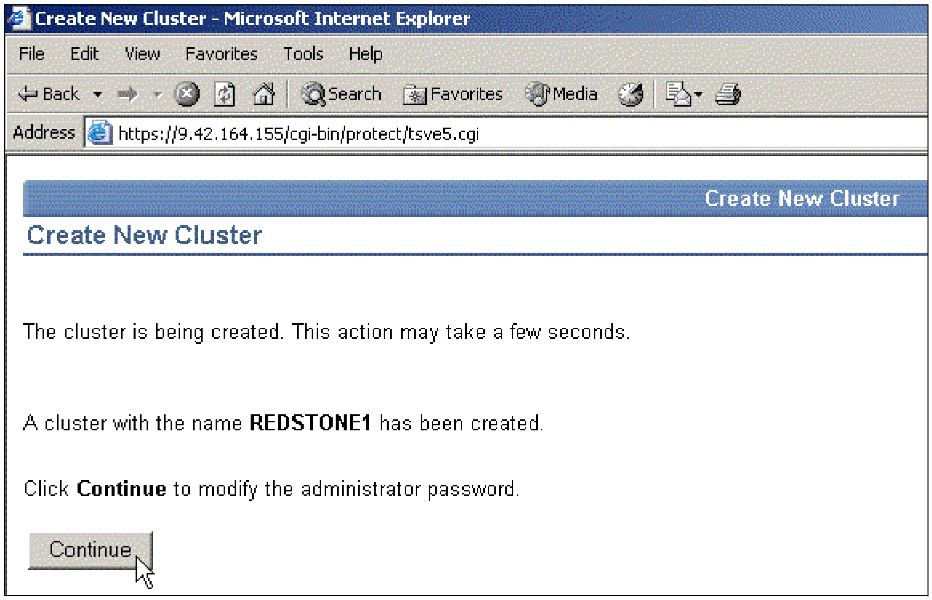


Figure 4-10 Create New Cluster Progress page

7. When the changes are accepted, the cluster displays the Enter Network Password window again. Type the User Name `admin` and the new admin password.

> **Note:** By this time, you should see the service panel display on the front of the configured node reflect the cluster name you set above (for example, REDSTONE1).

8. After you log back in, you see the first of several configuration pages beginning with the Error Notification Settings page as shown in Figure 4-11. We cover what these setting do in a later section. For now, click **Update Settings** and then on the next page, click **Continue** when prompted.

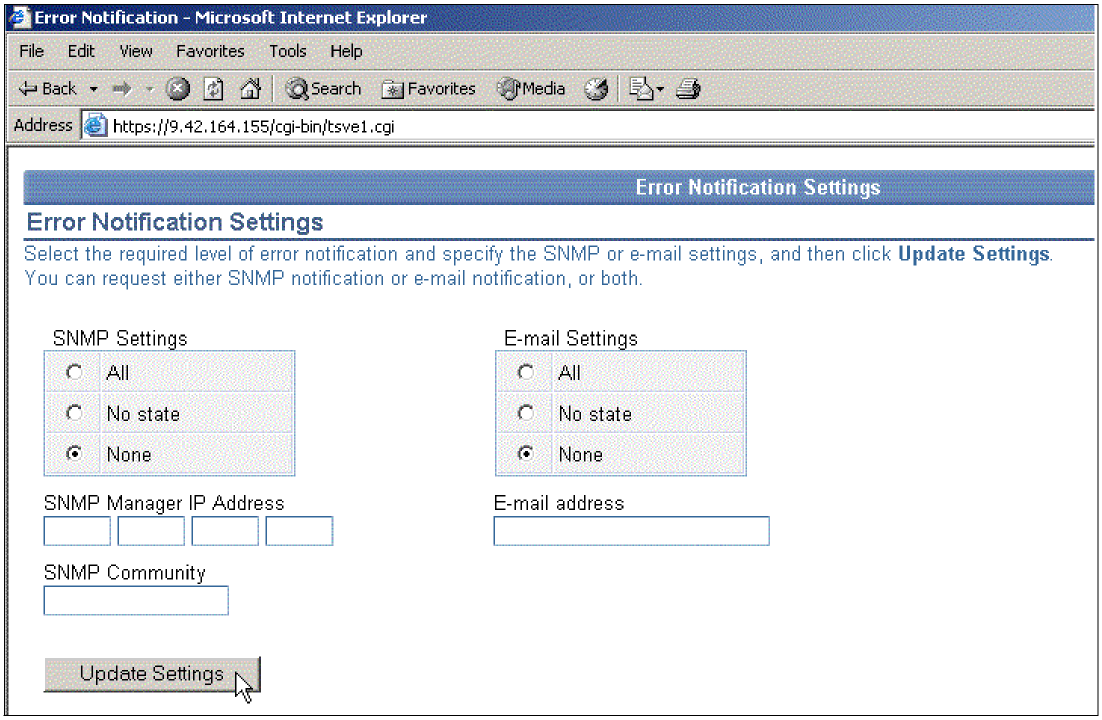


Figure 4-11 Error Notification Settings configuration page

You see the Featurization Settings window (Figure 4-12). To continue, you must at least fill in the Virtualization Limit (Gigabytes) field. If you are licensed for FlashCopy and PPRC, you may also select the **Enabled** radio buttons here. These features are discussed in more detail in later sections. Click the **Set Features** button.

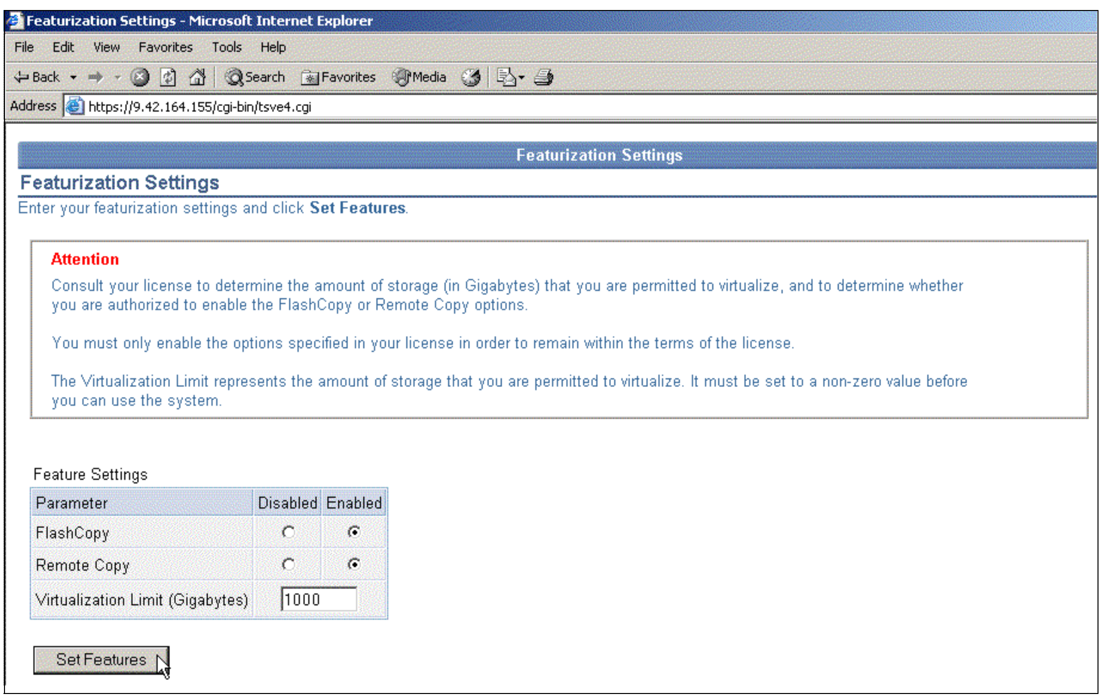


Figure 4-12 Featurization Settings Configuration page

> **Note:** The SVC uses the standard of 1 GB = 1024 MB. Therefore, typing 10 GB in the Featurization Settings page provides us with 10240 MB rather than 10000 MB as with other disk subsystems.

4.4.1 Uploading the SSH public key to the SVC cluster

After you click the Set Features button on the featurization page above, you see the Featurization Settings confirmation page (Figure 4-13). Follow these steps:

1. On this page you may either click the **Continue** button (to immediately upload an admin SSH key pair) or the **Bypass** button (to upload the SSH key pair at a later time). We recommend that you click **Continue** because we need to upload an SSH key pair to administer the box from either the CLI or GUI in the following chapters.

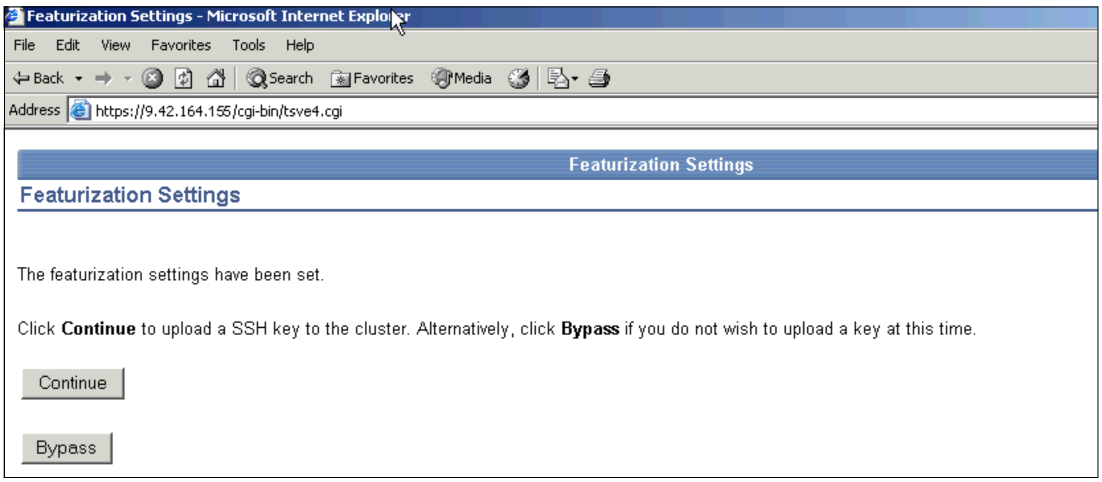


Figure 4-13 Featurization Settings confirmation page

2. The Add SSH Public Key page (Figure 4-14) opens. Browse or type the fully qualified directory path and file name of the public key you created and saved in 4.2.1, "Generating public and private SSH key pair using PuTTY" on page 70. Then type the name of the user you want to associate with this admin key pair (for example, admin) and click **Add Key**.

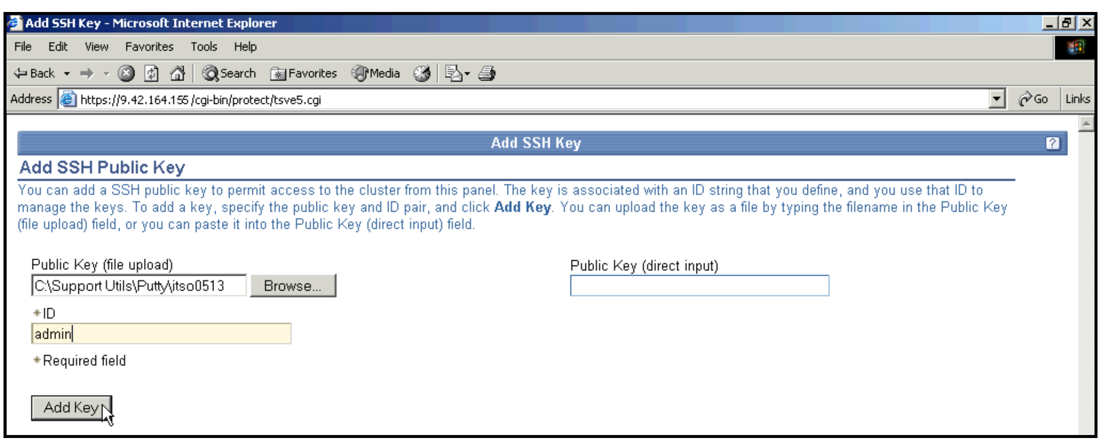


Figure 4-14 Add SSH Public Key

3. On the next page (Figure 4-15), a message is displayed indicating that a new SSH administrator key associated with the ID *admin* was added. Click **Continue**.

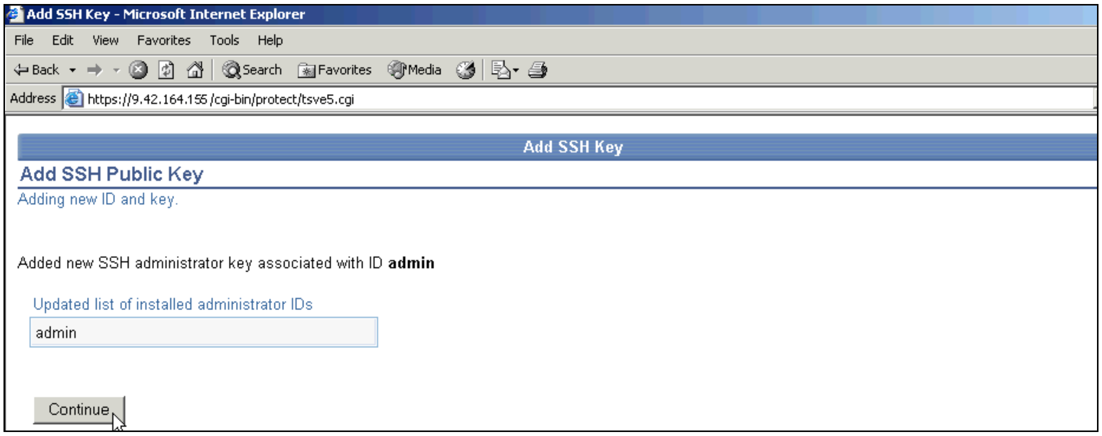


Figure 4-15 Adding SSH admin key successful

4. The cluster Web Interface administration home page (Figure 4-16) opens. Add a service SSH key pair. From the left navigation bar, select the **Maintain SSH Keys** link.

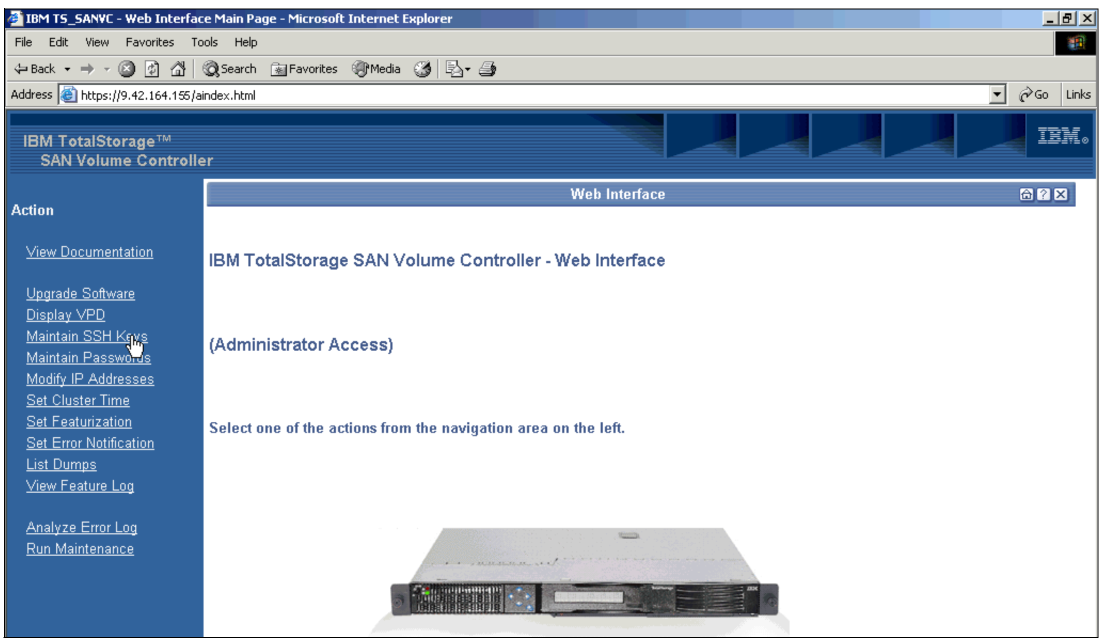


Figure 4-16 SVC Web Interface administration home page

5. On the SSH Key Maintenance panel (Figure 4-17), follow these steps:

   a. Type or browse to the location of the Public Key you created earlier.

   > **Note:** You can use the same key pair as for the admin account, or you can create a fresh key pair for the service account. We use the same key pair.

   b. Type the ID to be used.

   c. Under Access Level, select the **service** radio button to associate the key pair with the service user ID.

d. Click the **Add Key** button.

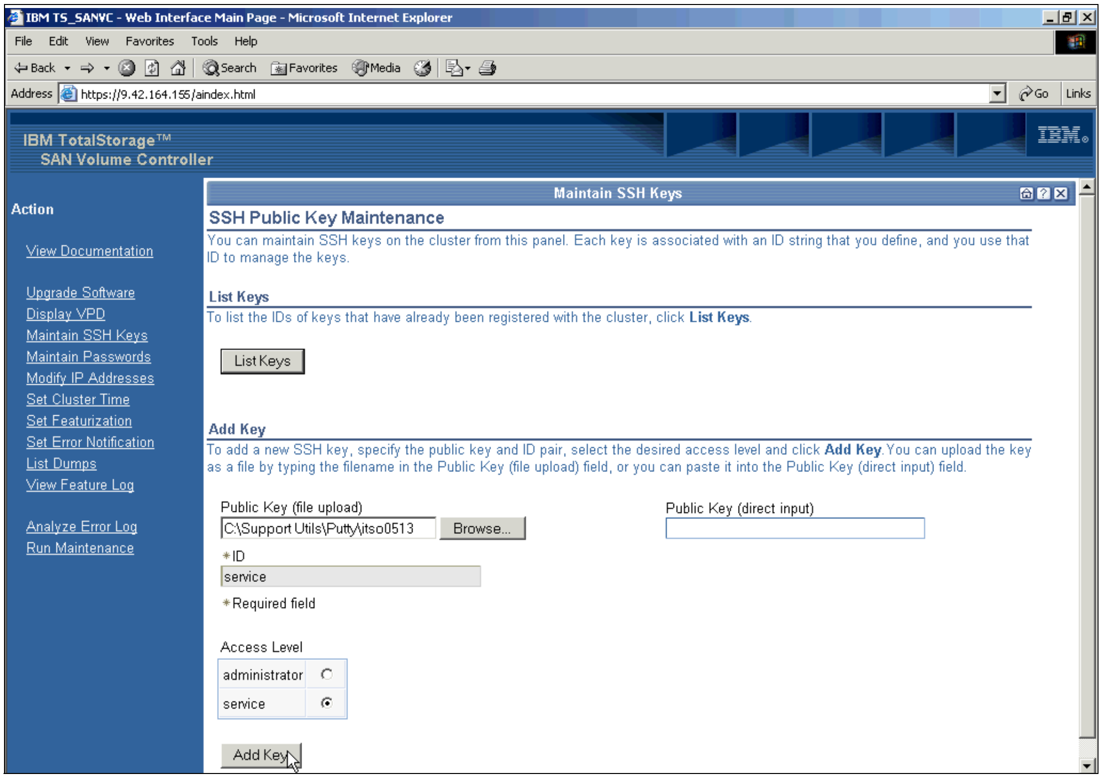


Figure 4-17 SSH Key Maintenance panel

6. On the next panel (Figure 4-18), a message appears indicating that a new SSH administrator key associated with the ID *service* was added. Close the Web interface.

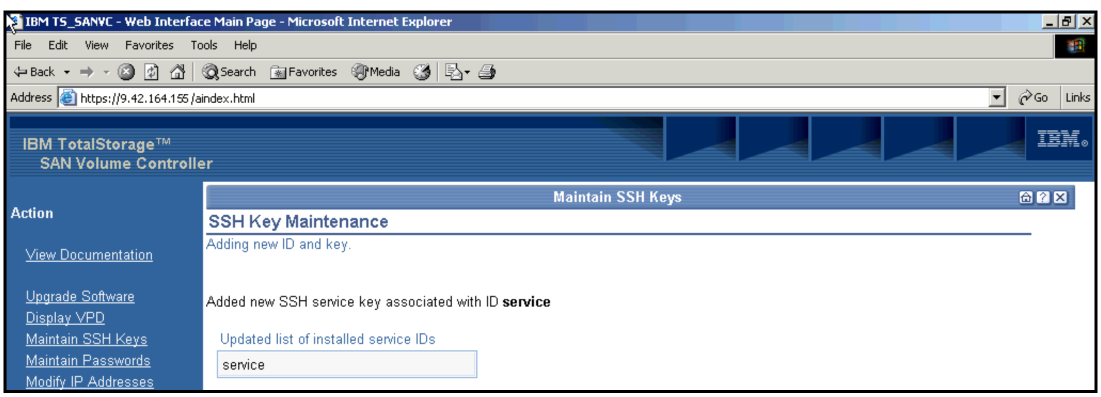


Figure 4-18 Adding SSH service Key Successful

We have now completed the basic setup requirements for the SVC cluster using the service panel and cluster Web interface. We are ready to complete the installation and configuration of the SVC cluster using either the CLI or CIM Agent and Console for SVC GUI.

See Chapter 5, “Quickstart configuration using the CLI” on page 85, or Chapter 6, “Quickstart configuration using the GUI” on page 101, for instructions about how to complete the initial configuration of your SVC cluster.

5

Quickstart configuration using the CLI

This chapter describes the basic configuration procedures required to get your IBM TotalStorage SAN Volume Controller (SVC) environment up and running as quickly as possible using the command line interface (CLI). See Chapter 8, “SVC configuration and administration using the CLI” on page 161, for more information about these and other configuration and administration procedures.

Important: The CLI is case sensitive.

5.1 Configuring the PuTTY session for the CLI

Before we can use the CLI, we must set up the PuTTY program using the SSH keys we generated earlier in 4.2.1, "Generating public and private SSH key pair using PuTTY" on page 70.

Perform these steps to configure the PuTTY session on the SSH client system:

1. From your Windows desktop, select **Start -> Programs -> PuTTY -> PuTTY** to open the PuTTY Configuration GUI window.
2. On the PuTTY Configuration window (Figure 5-1), from the Category panel on the left, click **Session**.

> **Note:** The items that you select in the Category panel affect the content that appears in the right panel.

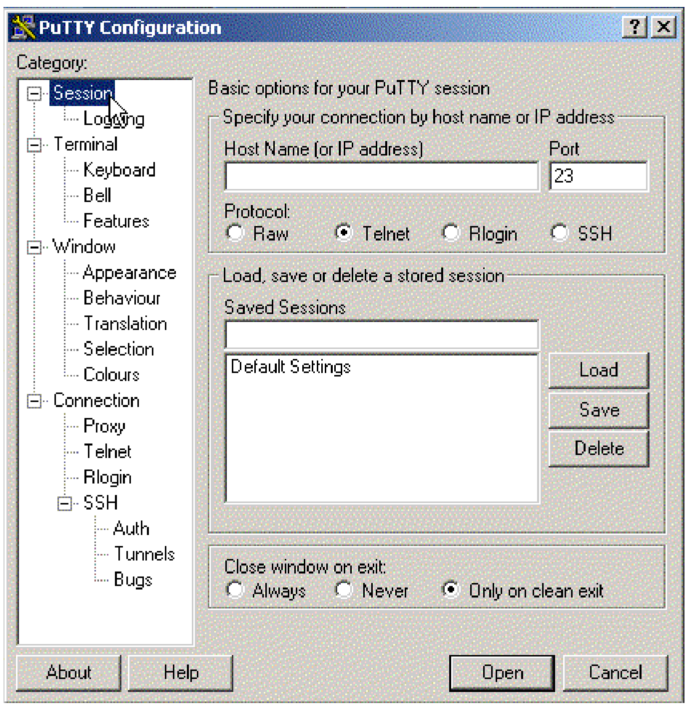


Figure 5-1 PuTTY Configuration window

3. In the right panel, under the Specify your connection by host name or IP address section, select the **SSH** radio button.
4. From the Category panel on the left side of the PuTTY Configuration window, click **Connection -> SSH**.
5. In the right panel, in the Preferred SSH protocol version section, select radio button **2**.
6. From the Category panel on the left side of the PuTTY Configuration window, click **Connection -> SSH -> Auth**.

7. In the right panel, in the Private key file for authentication: field under the Authentication Parameters section, type the fully qualified directory path and file name of the SSH client private key file you created earlier. See Figure 5-2.

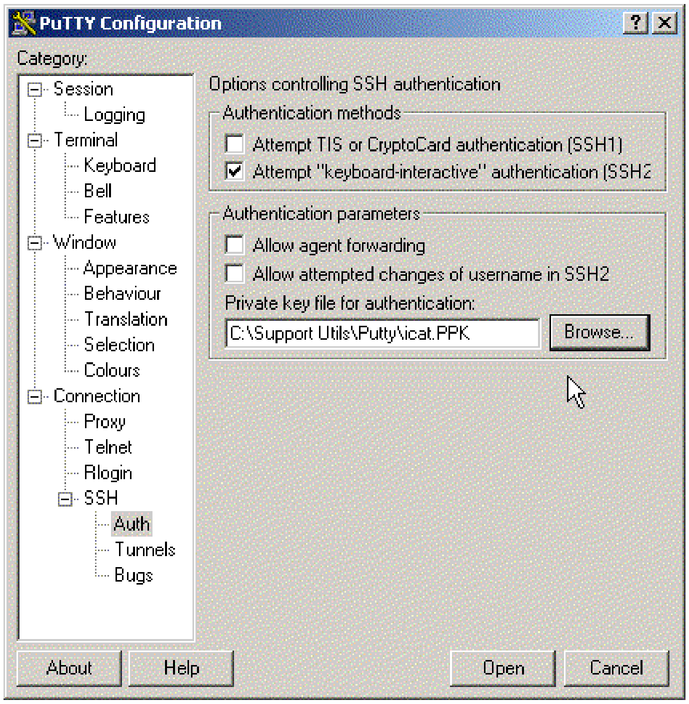


Figure 5-2 PuTTY Configuration: Private key location

> **Tip:** You can click **Browse** to select the file name from the system directory, or alternatively, type the fully qualified file name (for example, `C:\Support Utils\PuTTY\icat.PPK`).

8. From the Category panel on the left side of the PuTTY Configuration window, click **Session**.
9. In the right panel, follow these steps as shown in Figure 5-3:
   a. Under the Load, save or delete a stored session section, select **Default Settings** and click **Save**.
   b. For Host Name (or IP address), type the IP address of your SVC cluster.
   c. In the Save Sessions field, type a name (for example, SVC) to associate with this session.
   d. Click **Save**.

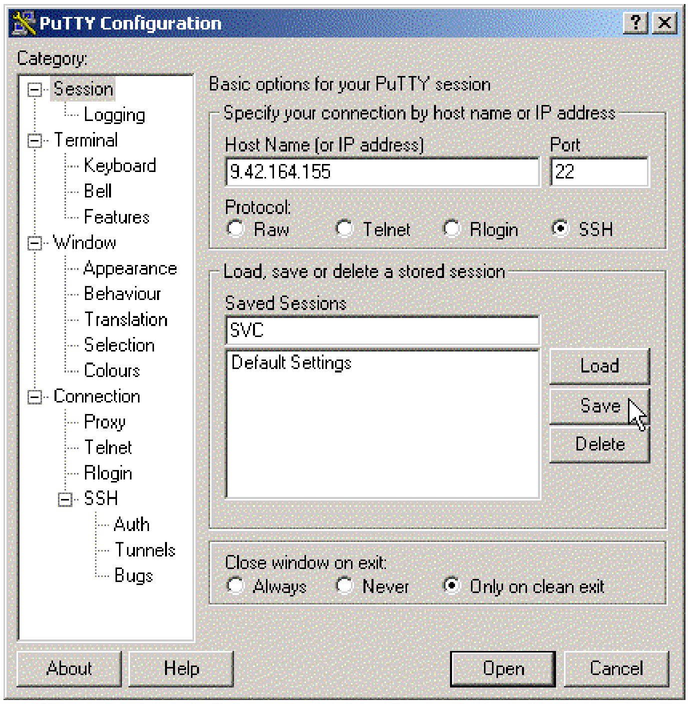


Figure 5-3 PuTTY Configuration: Saving a session

You can close the PuTTY Configuration window or leave it open to continue.

5.2 Starting the PuTTY CLI session

We need the PuTTY application for all of the following CLI tasks. If you closed it for any reason, restart the session as explained here:

1. From your master console desktop, open the PuTTY application by selecting **Start -> Programs -> PuTTY**.
2. On the PuTTY Configuration window (Figure 5-4), select the session you saved earlier (for example, SVC) and click **Load**.
3. Click **Open**.

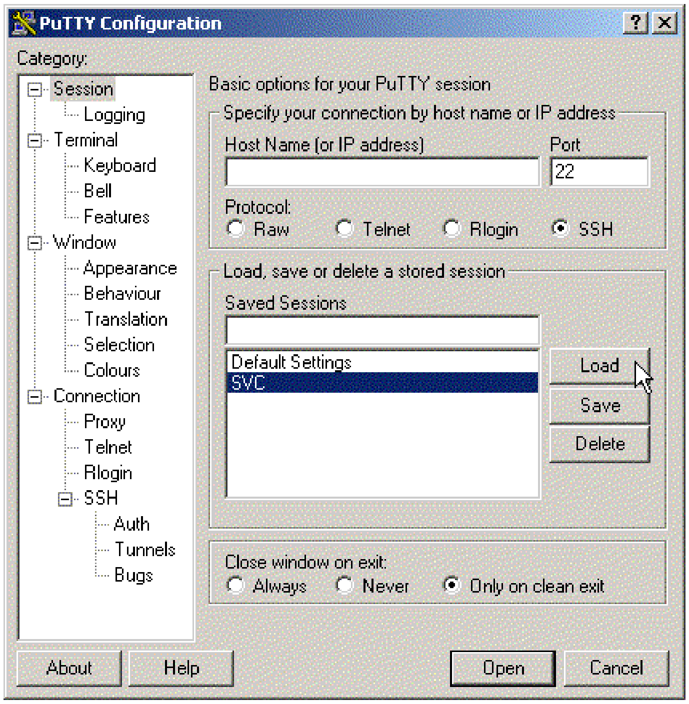


Figure 5-4 Open PuTTY command line session

4. If this is the first time you used the PuTTY application since generating and uploading the SSH key pair, you are prompted with a PuTTY Security Alert window stating a mismatch between private and public keys as shown in Figure 5-5. Click **Yes**, which invokes the CLI.

Figure 5-5 PuTTY Security Alert

5. At the Login as: prompt, type `admin` and press Enter (the user ID is case sensitive). As shown in Figure 5-6, the private key used in this PuTTY session is now authenticated against the public key that we uploaded to the SVC cluster in 4.4.1, “Uploading the SSH public key to the SVC cluster” on page 82.

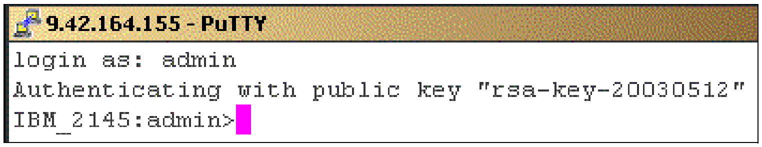


Figure 5-6 User authentication using SSH key pairs

5.3 Adding nodes to the cluster

After cluster creation is completed through the service panel and cluster Web interface, only one node (the configuration node) is set up. To be a fully functional SVC cluster, you must add a second node to the configuration.

To add a node to a cluster, gather the necessary information as explained in the following steps:

1. We need some information from the existing node. Gather this information by using the **`svcinfo lsnode node1`** command as shown in Example 5-1.

> **Note:** The name of the nodes already in the cluster are shown on the service panel displays on each node. By default, the first node is called *node1*. We show how to change this in a later topic.

Example 5-1 svcinfo lsnode command

```
IBM_2145:admin>svcinfo lsnode node1
id 1
name node1
UPS_serial_number YM100032B422
WWNN 5005076801000364
status online
IO_group_id 0
IO_group_name io_grp0
partner_node_id
partner_node_name
config_node yes
UPS_unique_id 20400000C2484082
port_id 5005076801100364
port_status active
port_id 5005076801200364
port_status active
port_id 5005076801300364
port_status active
port_id 5005076801400364
port_status
active
```

The most important information to look for here is the IO_group_name because we use this when adding our second node to the SVC cluster configuration.

> **Note:** You can see in Example 5-1 that no partner_node_id or partner_node_name exists for node1 as yet.

2. See what nodes are available for inclusion in the SVC cluster configuration. Enter the **`svcinfo lsnodecandidate`** command:

```
IBM_2145:admin>svcinfo lsnodecandidate
id               node_cover_name   UPS_serial_number     UPS_unique_id
500507680100035A 000683            YM100032B425          20400000C2484085
```

   If this command returns no information and your second node is powered on and zones are correctly defined, pre-existing cluster configuration data may be stored in it. If you are sure this node is not part of another active SVC cluster, you can use the service panel to delete the existing cluster information. After this is complete, re-issue the **`svcinfo lsnodecandidate`** command and you should see it listed.

   For information about how to delete an existing cluster configuration using the service panel, see 4.3.1, "Creating the cluster (first time) using the service panel" on page 74. For information about storage area network (SAN) zoning, see Chapter 3, "Planning and configuration" on page 25.

3. Using a combination of the information obtained in the previous steps, add the second node to the SVC cluster configuration. Enter the **`svctask addnode`** command. The full syntax of the command is:

```
addnode {-panelname panel_name | -wwnodename wwnn_arg} [-name new_name]
       -iogrp iogrp_name_or_id
```

   Note the following explanation:

   – **panelname**: Name of the node as it appears on the panel
   – **wwnodename**: Worldwide node name (WWNN) of the node
   – **name**: Name to be allocated to the node
   – **iogrp**: I/O group to which the node is added

   Here is an example of this command:

```
IBM_2145:admin>svctask addnode -wwnodename 500507680100035a -iogrp io_grp0
Node, id [2], successfully added
```

   In this example:

   – **500507680100035a** is the ID found using the **`svcinfo lsnodecandidate`** command.

   > **Note:** The wwnodname is one of the few things in the CLI that is not case sensitive.

   – **io_grp0** is the name of the I/O group to which node1 belonged as found using the **`svcinfo lsnode node1`** command.

> **Note:** Because we did not provide the -name parameter, the SVC automatically generates the name *nodeX* (where *X* is the ID sequence number assigned by the SVC internally). In our case, this is *node2*.
>
> If you want to provide a name, you can use A to Z, a to z, 0 to 9, and the underscore. It can be between one and 15 characters in length. However, it cannot start with a number or the word *node* because this prefix is reserved for SVC assignment only.

4. If we display the node information for node1 again, as shown in Example 5-2, node1 now has a *partner_node_id* of 2 and a *partner_node_name* of node2.

Example 5-2 svcinfo lsnode command

```
IBM_2145:admin>svcinfo lsnode node1
id 1
name node1
UPS_serial_number YM100032B422
WWNN 5005076801000364
status online
IO_group_id 0
IO_group_name io_grp0
partner_node_id 2
partner_node_name node2
config_node yes
UPS_unique_id 20400000C2484082
port_id 5005076801100364
port_status active
port_id 5005076801200364
port_status active
port_id 5005076801300364
port_status active
port_id 5005076801400364
port_status active
```

You have now completed the cluster configuration and you have a fully redundant SVC environment.

5.4 Setting the cluster time zone and time

Perform the following steps to set the cluster time zone and time:

1. Find out for which time zone your cluster is currently configured. Enter the **svcinfo showtimezone** command as shown here:

   ```
   IBM_2145:admin>svcinfo showtimezone
   id              timezone
   522 UTC
   ```

   If this setting is correct (for example, 522 UTC), you can skip to Step 4. If not, continue with Step 2.

2. To find what time zone code is associated with your time zone, enter the **svcinfo lstimezones** command as shown in Example 5-3. A truncated list is provided for this example.

Example 5-3 svcinfo lstimezones command

```
IBM_2145:admin>svcinfo lstimezones
. . .
508 UCT
509 Universal
510 US/Alaska
511 US/Aleutian
512 US/Arizona
513 US/Central
514 US/Eastern
515 US/East-Indiana
516 US/Hawaii
517 US/Indiana-Starke
518 US/Michigan
```

```
519 US/Mountain
520 US/Pacific
521 US/Samoa
. . .
```

3. Now that you know which time zone code is correct for you (in our case 514), set the time zone by issuing the **svctask settimezone** command:

   ```
   IBM_2145:admin>svctask settimezone -timezone 514
   ```

4. Set the cluster time by issuing the **svctask setclustertime** command:

   ```
   IBM_2145:admin>svctask setclustertime -time 0512150403
   ```

   The format of the time is MMDDHHmmYY.

You have now completed the tasks necessary to set the cluster time zone and time.

5.5 Creating host definitions

Perform the following steps to create host definitions within the SVC:

1. To determine which hosts ports are eligible for definition, issue the **svcinfo lshbaportcandidate** command as shown in Figure 5-4.

Example 5-4 svcinfo lshbaportcandidate command

```
IBM_2145:admin>svcinfo lshbaportcandidate
id
210000E08B08AFD6
210100E08B28AFD6
210000E08B04D751
210100E08B24D751
10000000C9295A9A
10000000C9266F5B
```

   This command shows all WWNs that are visible to the SVC which were not already defined to a host. If your WWN does not appear, verify that the host logged into the switch and that zoning is updated to allow SVC and host ports to see each other as explained in Chapter 3, “Planning and configuration” on page 25.

2. The output from this command shows that we have four QLogic ports (21xxx) and two Emulex ports (10xxx). By checking the hosts and confirming with the switch nameserver, you determine that the 10xxx WWNs belong to the AIX host. Therefore, you have everything necessary to create a host definition. We can do this in one of two ways:

   – You can add WWN port definitions to a host one at a time using the **mkhost** and **addhostport** commands as shown in Figure 5-5.

Example 5-5 svctask mkhost and addhostport commands

```
IBM_2145:admin>svctask mkhost -name aix1 -hbawwpn 10000000C9295A9A
Host id [0] successfully created

IBM_2145:admin>svctask addhostport -hbawwpn 10000000c9266f5b aix1
```

> **Note:** The -name and -hbawwpn parameters are optional. If you do not specify a -name, the default is host*X*, where X is the ID sequence number assigned by the SVC internally. If you do not specify the -hbawwpn parameter, an "empty" host is created.
>
> If you prefer to provide a name for your host (as we have), you can use letters A to Z, a to z, numbers 0 to 9, and the underscore. It can be between one and 15 characters in length. However, it cannot start with a number or the word *host* because this prefix is reserved for SVC assignment only.

The -name parameter is used to name the host (in our case `aix1`) and the -hbawwpn parameter is filled in using data retrieved from the **`lshbaportcandidate`** command.

– Add all ports at the same time using a modification of the **`mkhost`** command as shown here:

```
IBM_2145:admin>svctask mkhost -name aix1 -hbawwpn 10000000c9295a9a:10000000c9266f5b
Host id [0] successfully created
```

3. Check that the host definitions were correctly created using the **`svcinfo lshost`** command as shown in Example 5-6.

Example 5-6 svcinfo lshost commands

```
IBM_2145:admin>svcinfo lshost
id              name             port_count
0               aix1             2

IBM_2145:admin>svcinfo lshost aix1
id 0
name aix1
port_count 2
WWPN 10000000C9295A9A
port_logged_in_count 2
WWPN 10000000C9266F5B
port_logged_in_count 2
```

You have now completed the tasks required to add host definitions to your SVC configuration.

5.6 Displaying managed disks

Perform the following steps to display mDisks:

1. See which mDisks are available. Enter the **`svcinfo lsmdiskcandidate`** command as shown in Example 5-7. This displays all detected mDisks that are not currently part of a managed disk group (MDG).

Example 5-7 svcinfo lsmdiskcandidate command

```
IBM_2145:admin>svcinfo lsmdiskcandidate
id
0
1
2
```

```
3
4
5
6
7
```

Alternatively, you can list all mDisks (managed or unmanaged) by issuing the **svcinfo lsmdisk** command as shown in Example 5-8.

Example 5-8 svcinfo lsmdisk command

```
IBM_2145:admin>svcinfo lsmdisk -delim :
id:name:status:mode:mDisk_grp_id:mDisk_grp_name:capacity:ctrl_LUN_#:controller_name
0:mDisk0:online:unmanaged:::167.0GB:0000000000000000:controller0
1:mDisk1:online:unmanaged:::200.4GB:0000000000000001:controller0
2:mDisk2:online:unmanaged:::167.0GB:0000000000000002:controller0
3:mDisk3:online:unmanaged:::200.4GB:0000000000000003:controller0
4:mDisk4:online:unmanaged:::339.3GB:0000000000000004:controller0
5:mDisk5:online:unmanaged:::407.2GB:0000000000000005:controller0
6:mDisk6:online:unmanaged:::681.2GB:0000000000000006:controller0
7:mDisk7:online:unmanaged:::817.4GB:0000000000000007:controller0
```

From this output, you can see additional information about each mDisk (such as current status). For the purpose of our current task, we are only interested in the unmanaged disks because they are candidates for MDGs (all mDisk in our case).

> **Tip:** The -delim : parameter collapses output instead of wrapping text over multiple lines.

2. If not all mDisks that you expect are visible, rescan the Fibre Channel network available by entering the **svctask detectmdisk** command:

   ```
   IBM_2145:admin>svctask detectmdisk
   ```

3. If you run the **svcinfo lsmdiskcandidate** command again and your mDisk or mDisks are still not visible, check that the logical unit numbers (LUNs) from your subsystem have been properly assigned to the SVC and that appropriate zoning is in place (for example, SVC can see the disk subsystem). See Chapter 3, “Planning and configuration” on page 25, for details about how to set up your SAN fabric.

5.7 Creating MDGs

Perform the following steps to create a MDG:

1. From the information obtained in the previous section, add mDisks to MDGs using one of the following ways:

   – Issue the **svctask mkmDiskgrp** command as shown here where you add multiple mDisk to the MDG at the same time:

     ```
     IBM_2145:admin>svctask mkmdiskgrp -name ARR36P5N -ext 32 -mdisk mDisk0:mDisk2
     mDisk Group, id [0], successfully created
     ```

     This command creates an mDisk group called ARR36P5N. The extent size used within this group is 32 MB. Two mDisks (mDisk0 and mDisk2) are added to the group.

> **Note:** The -name and -mdisk parameters are optional. If you do not enter a -name, the default is mDiskgrp*X*, where *X* is the ID sequence number assigned by the SVC internally. If you do not enter the -mdisk parameter, an empty MDG is created.
>
> If you want to provide a name, you can use letters A to Z, a to z, numbers 0 to 9, and the underscore. It can be between one and 15 characters in length, but it cannot start with a number or the word *mDiskgrp* because this prefix is reserved for SVC assignment only.

By running the **svcinfo lsmdisk** command again, you should now see the mDisks (mDisk0 and mDisk2) as "managed" and part of the MDG ARR36P5N as shown in Example 5-9.

Example 5-9 svcinfo lsmdisk command

```
IBM_2145:admin>svcinfo lsmdisk -delim :
id:name:status:mode:mDisk_grp_id:mDisk_grp_name:capacity:ctrl_LUN_#:controller_name
0:mDisk0:online:managed:0:ARR36P5N:167.0GB:0000000000000000:controller0
1:mDisk1:online:unmanaged:::200.4GB:0000000000000001:controller0
2:mDisk2:online:managed:0:ARR36P5N:167.0GB:0000000000000002:controller0
3:mDisk3:online:unmanaged:::200.4GB:0000000000000003:controller0
4:mDisk4:online:unmanaged:::339.3GB:0000000000000004:controller0
5:mDisk5:online:unmanaged:::407.2GB:0000000000000005:controller0
6:mDisk6:online:unmanaged:::681.2GB:0000000000000006:controller0
7:mDisk7:online:unmanaged:::817.4GB:0000000000000007:controller0
```

– If you want to add an mDisk to an existing MDG or want to add mDisks one at a time, combine the **mkmDiskgrp** command to create the initial MDG and then use the **addmDisk** command, as shown in Example 5-10, to add other mDisks to it.

Example 5-10 svctask mkmDiskgrp and addmDisk commands

```
IBM_2145:admin>svctask mkmDiskgrp -name ARR36P6N -ext 32 -mDisk mDisk1
mDisk Group, id [1], successfully created

IBM_2145:admin>svctask addmDisk -mDisk mDisk3 ARR36P6N
```

The first command in this example creates an mDisk group called ARR36P6N. The extent size used within this group is 32 MB. One mDisk (mDisk1) is added to the group. The second command adds a second mDisk (mDisk3) to the same MDG.

By running the **svcinfo lsmdisk** command again, you now see the mDisks (mDisk1 and mDisk3) as "managed" and part of the MDG ARR36P6N (see Example 5-11).

Example 5-11 svcinfo lsmdisk command

```
IBM_2145:admin>svcinfo lsmdisk -delim :
id:name:status:mode:mDisk_grp_id:mDisk_grp_name:capacity:ctrl_LUN_#:controller_name
0:mDisk0:online:managed:0:ARR36P5N:167.0GB:0000000000000000:controller0
1:mDisk1:online:managed:1:ARR36P6N:200.4GB:0000000000000001:controller0
2:mDisk2:online:managed:0:ARR36P5N:167.0GB:0000000000000002:controller0
3:mDisk3:online:managed:1:ARR36P6N:200.4GB:0000000000000003:controller0
4:mDisk4:online:unmanaged:::339.3GB:0000000000000004:controller0
5:mDisk5:online:unmanaged:::407.2GB:0000000000000005:controller0
```

```
6:mDisk6:online:unmanaged:::681.2GB:0000000000000006:controller0
7:mDisk7:online:unmanaged:::817.4GB:0000000000000007:controller0
```

For information about other tasks, such as adding mDisks to MDGs, renaming MDGs or deleting MDGs, see Chapter 8, “SVC configuration and administration using the CLI” on page 161.

You have now completed the tasks required to create a MDG.

5.8 Creating a vDisk

When making a vDisk, you must enter several parameters (some mandatory, some optional) on the CLI. The full command syntax is:

```
mkvdisk -mDiskgrp name|id -iogrp name|id -size size [-fmtdisk]
      [-vtype seq|striped|image] [-node name|id] [-unit b|kb|mb|gb|tb|pb]
      [-mDisk name|id_list] [-name name]
```

The parameters are defined as:

- **-mDiskgrp**: Name or ID of the MDG in which to create the vDisk.
- **-iogrp**: Name or ID of I/O group which is to own the vDisk.
- **-size**: Capacity (numerical), not necessary for image mode vDisks.
- **-fmtdisk**: Optional parameter to force a format of the new vDisk.
- **-vtype**: Optional parameter to specify the type of vDisk (sequential, striped or image mode). Default (if nothing is specified) is striped.
- **-node**: Optional parameter to specify the name or ID of the preferred node. Default (if nothing is specified) is to alternate between nodes in the I/O group.
- **-unit**: Optional parameter to specify the data units for capacity parameter. Default (if nothing is specified) is MB.
- **-mDisk**: Optional parameter to specify the name or ID of the mDisk or mDisks to be used for the vDisk. This is only required for sequential and image mode vDisks since striped vDisks use all mDisks that are available in the MDG by default.

  Note: You can use this parameter for striped vDisks, for example, if you want to specify that the vDisk only uses a subset of the mDisks available within a MDG.

- **-name**: Optional parameter to assign a name to the new vDisk. Default (if nothing is specified) is to assign the name *vDiskX*, where *X* is the ID sequence number assigned by the SVC internally.

  Note: If you want to provide a name, you can use A to Z, a to z, 0 to 9, and the underscore. It can be between one and 15 characters in length. However, it cannot start with a number or the word *vDisk* since this prefix is reserved for SVC assignment only.

Now perform the following steps to create vDisks:

1. Create a striped vDisk using the **`svctask mkvdisk`** command (we cover sequential and image mode vDisks in a later section). See Example 5-12. This command creates a 10 GB, striped vDisk called *aix_vDisk0* within the MDG ARR36P5N and assigned it to the I/O group iogrp_0.

Example 5-12 svctask mkvdisk commands

```
IBM_2145:admin>svctask mkvdisk -mDiskgrp ARR36P5N -iogrp io_grp0 -size 10 -vtype striped
-unit gb -name aix_vDisk0
Host LUN, id [0], successfully created
```

2. Create the vDisks (four in this example) using the previous methods which can be displayed using the **svcinfo lsvdisk** command as shown in Example 5-13.

Example 5-13 svcinfo lsvdisk command

```
IBM_2145:admin>svcinfo lsvdisk -delim :
id:name:IO_group_id:IO_group_name:status:mDisk_grp_id:mDisk_grp_name:capacity:type:FC_id:FC
_name:RC_id:RC_name
0:aix_vDisk0:0:io_grp0:online:0:ARR36P5N:10.0GB:striped::::
1:aix_vDisk1:0:io_grp0:online:0:ARR36P5N:10.0GB:striped::::
2:aix_vDisk2:0:io_grp0:online:0:ARR36P5N:10.0GB:striped::::
3:aix_vDisk3:0:io_grp0:online:0:ARR36P5N:10.0GB:striped::::
```

To display more information about a specific vDisk, enter a variant of the **svcinfo lsvdisk** command as shown in Example 5-14.

Example 5-14 svcinfo lsvdisk command

```
IBM_2145:admin>svcinfo lsvdisk aix_vDisk0
id 0
name aix_vDisk0
IO_group_id 0
IO_group_name io_grp0
status online
mDisk_grp_id 0
mDisk_grp_name ARR36P5N
capacity 10.0GB
type striped
formatted no
mDisk_id
mDisk_name
FC_id
FC_name
RC_id
RC_name
throttling 0
preferred_node_id 1
```

For information about other tasks, such as deleting a vDisk, renaming a vDisk, or expanding a vDisk, see Chapter 8, “SVC configuration and administration using the CLI” on page 161.

You have now completed the tasks required to create a vDisk.

5.9 Assigning a vDisk to a host

Using the vDisk and host definition created in previous sections, assign vDisks to hosts ready for their use. To do this, use the **svctask mkvdiskhostmap** command:

```
IBM_2145:admin>svctask mkvdiskhostmap -host aix1 aix_vDisk0
Host LUN to Host map, id [0], successfully created
```

This command assigns vDisk aix_vDisk0 to host aix1.

> **Note:** The optional parameter -scsi scsi_num can help assign a specific LUN ID to a vDisk that is to be associated with a given host. The default (if nothing is specified) is to increment based on what is already assigned to the host.

For information about other tasks, such as deleting a vDisk to host mapping, see Chapter 8, “SVC configuration and administration using the CLI” on page 161.

You have now completed all the tasks required to assign a vDisk to an attached host. You are ready to proceed to Chapter 7, “Host configuration” on page 133, to begin to use the assigned vDisks.

6

Quickstart configuration using the GUI

This chapter describes the basic configuration procedures required to get your IBM TotalStorage SAN Volume Controller (SVC) environment up and running as quickly as possible using the CIM Agent and Console for SAN Volume Controller graphical user interface (GUI).

See Chapter 9, “SVC configuration and administration using the GUI” on page 217, for more information about these and other configuration and administration procedures.

Important: Data entries made through the GUI are case sensitive.

6.1 Configuring the GUI

If this is the first time you have used the central administration GUI, you must configure it as explained here:

1. Open the GUI using one of the following methods:

   – Double-click the icon marked **SAN Volume Controller Console** on your master console's desktop.

   – Open a Web browser on the master console and point to the address:

     `http://`*`localhost`*`:9080/ica`

   – Open a Web browser on a separate workstation and point to the address:

     `http://`*`masteconsoleipaddress`*`:9080/ica`

     In our case, the master console IP address is 9.42.164.140.

2. On the Signon page (Figure 6-7), type the user ID `superuser` and the default password of `password`. Click **OK**.

> **Note:** Passwords for the central administration GUI are separate from the passwords set for individual SVC clusters.

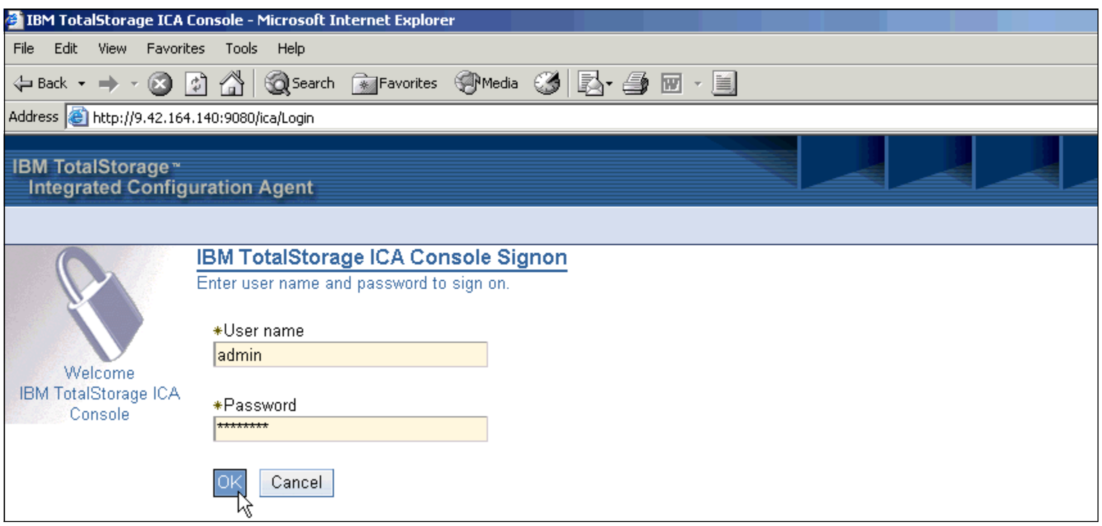


Figure 6-1 GUI signon

3. Change the default password.

> **Note:** Like all passwords, this is case sensitive.

4. On the GUI Welcome panel (Figure 6-2), click the **Add SAN Volume Controller Cluster** button in the center of the panel.

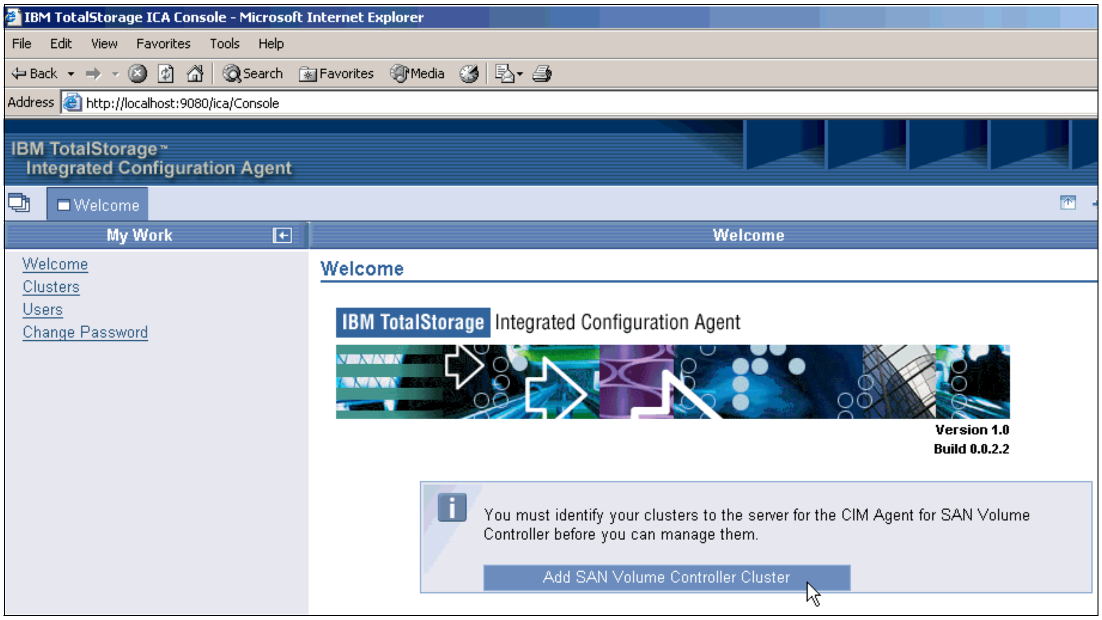


Figure 6-2 Adding the SVC cluster for management

5. On the Adding Clusters panel (Figure 6-3), type the IP address of your SVC cluster and click **OK**.

> **Important:** Do not select the Create (initialize) Cluster box. Doing so leads you through the initial cluster installation process, which is explained in Chapter 4, “Initial installation and configuration of the SVC” on page 69. If you select it, your cluster is re-initialized and any configuration settings you already made are lost.

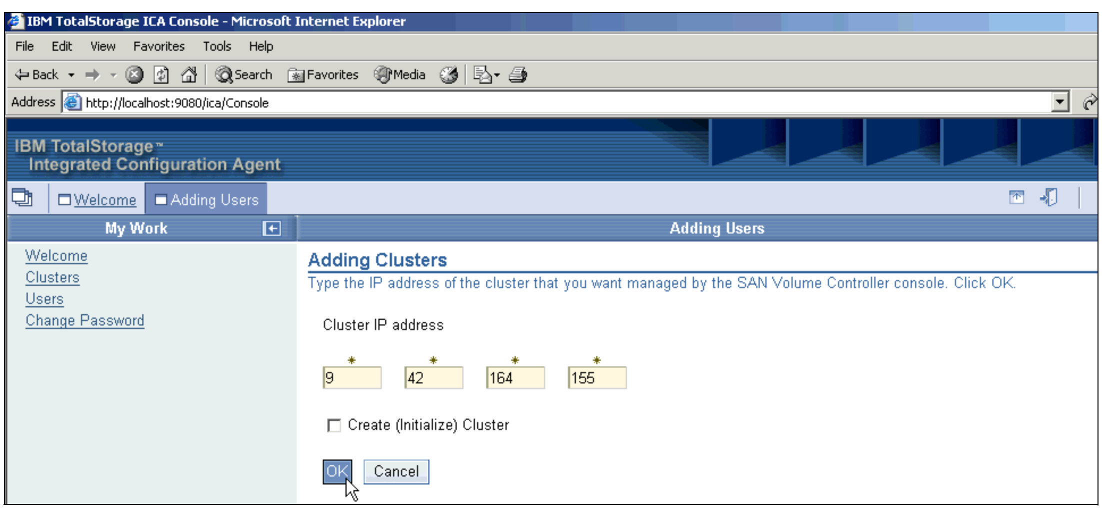


Figure 6-3 Adding Clusters panel

6. You are prompted for the user ID and password of the SVC cluster, as shown in Figure 6-4. Enter the user ID of `admin` and the password you set earlier. Click **OK**.

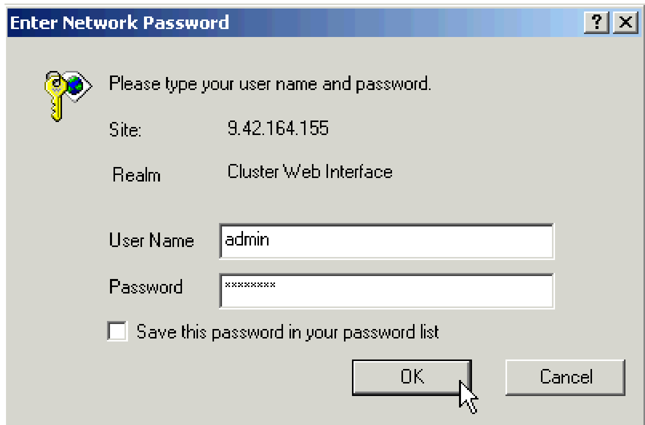


Figure 6-4 SVC cluster user ID and password signon window

7. The Maintaining SSH Keys panel (Figure 6-5) opens. We already uploaded our keys. To verify this, click the **List Keys** button.

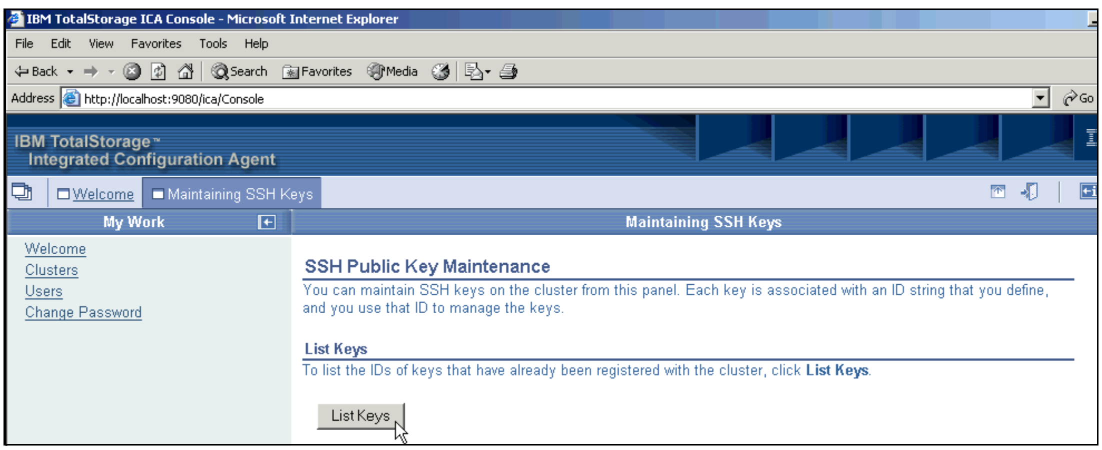


Figure 6-5 SSH Public Key Maintenance panel

8. You see the keys that already uploaded to the cluster as shown in Figure 6-6. In the upper right corner of this panel, click the **X** to close it.

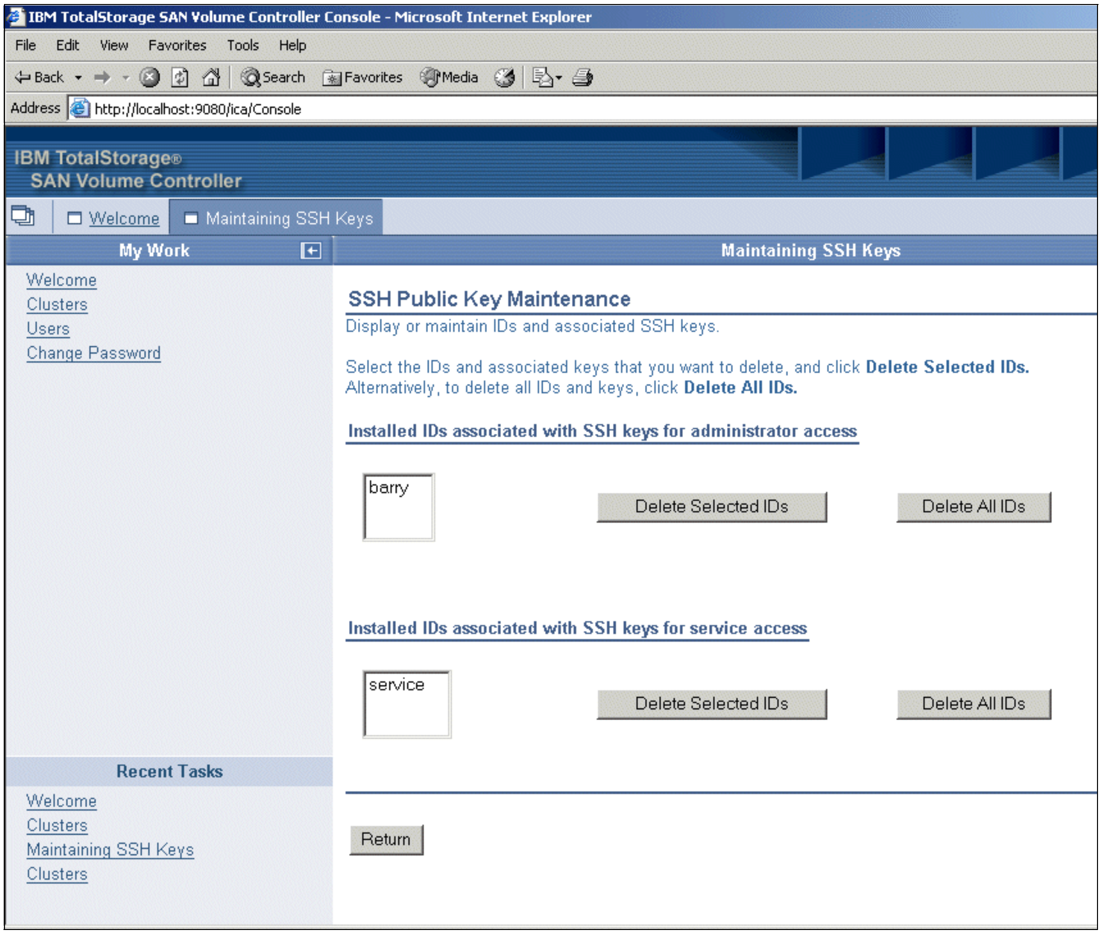


Figure 6-6 List SSH keys panel

9. You can now close the browser session completely or leave it open on the Welcome panel and continue with the next section.

You have completed the tasks required to configure the GUI for SVC administration. Continue with the next section and add the second node to the cluster.

6.2 Adding nodes to the cluster

After cluster creation is completed through the service panel and cluster Web interface, only one node (the configuration node) is set up. To be a fully functional SVC cluster, a second node must be added to the configuration.

Perform the following steps to add nodes to the cluster:

> **Note:** If you closed your browser at the end of the last section, start with step 1. If you left it open, skip to Step 3 on page 107.

1. Open the GUI using one of the following methods:
   - Double-click the **SAN Volume Controller Console** icon on your master console's desktop.
   - Open a Web browser on the master console and point to the address:

     `http://localhost:9080/ica`
   - Open a Web browser on a separate workstation and point to the address:

     `http://masteconsoleipaddress:9080/ica`

     In our case, the master console IP address is 9.42.164.140.
2. On the Signon page (Figure 6-7), type the user ID `superuser` and the password you set in the previous section. Click **OK**.

Figure 6-7 GUI signon page

3. You see the GUI Welcome page as shown in Figure 6-8. This page has several links: My Work (top left), a Recent Tasks list (bottom left), the GUI version and build level information (right, under the graphic), and a hypertext link to the SVC download page:

   http://www.ibm.com/storage/support/2145

   Under My Work on the left, click the **Clusters** link.

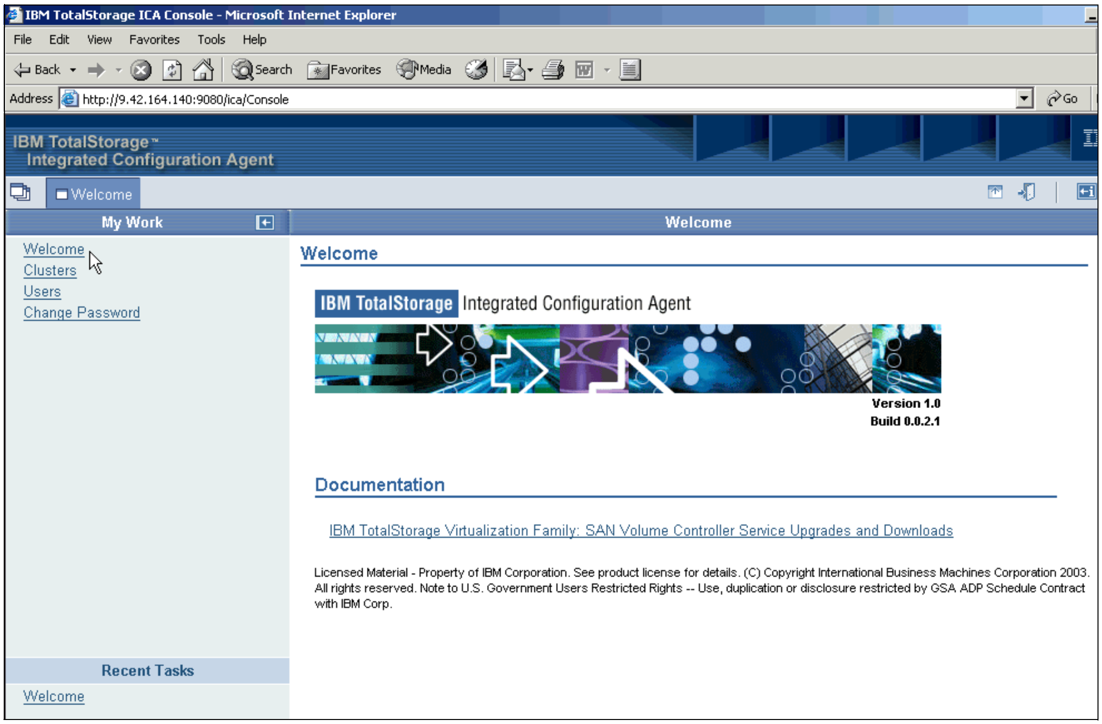


Figure 6-8 GUI Welcome page

4. On the Viewing Clusters panel (Figure 6-9), select the check box next to the cluster on which you want to perform actions (in our case **REDSTONE1**). Select **Launch the SAN Volume Controller application** from the list and click **Go**.

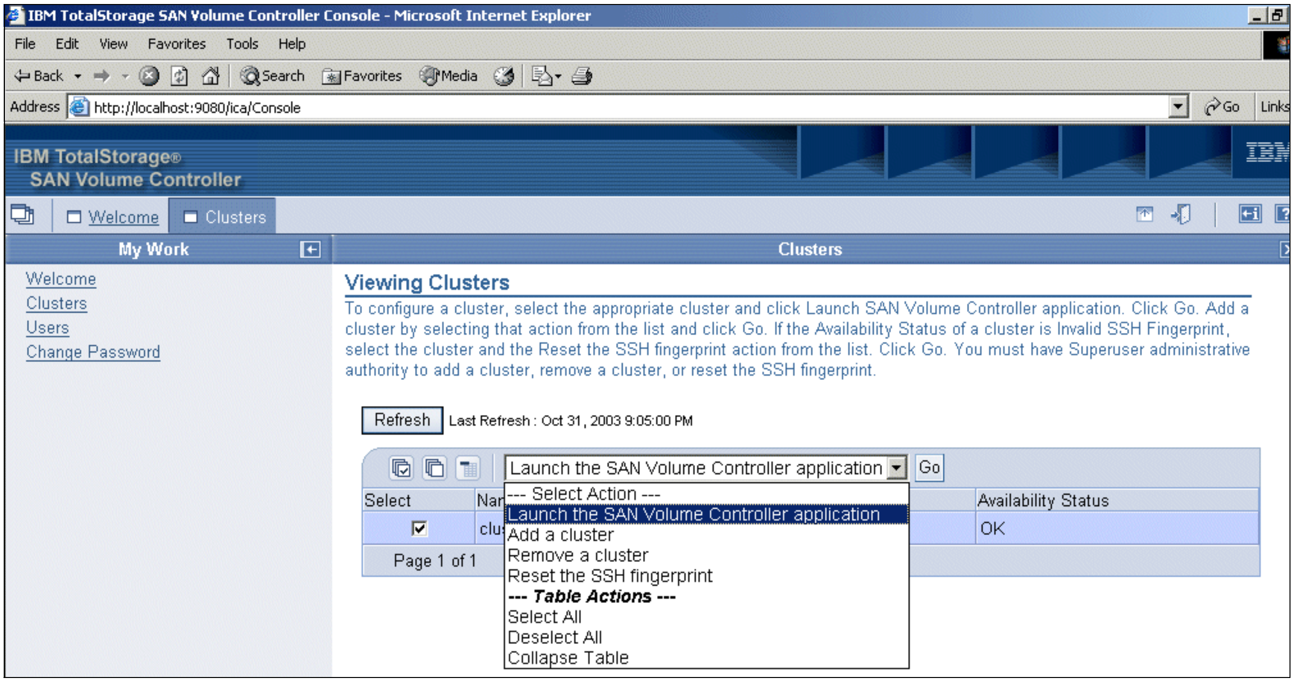


Figure 6-9 Selecting to launch the SAN Volume Controller application

5. The SAN Volume Controller Console Application launches in a separate browser window (Figure 6-10). On this page, as with the Welcome page, you can see several links under My Work (top left), a Recent Tasks list (bottom left), the SVC Console version and build level information (right, under graphic), and a hypertext link to the SVC download page:

   `http://www.ibm.com/storage/support/2145`

   Under My Work, click the **Work with Nodes** option and then the **Nodes** link.

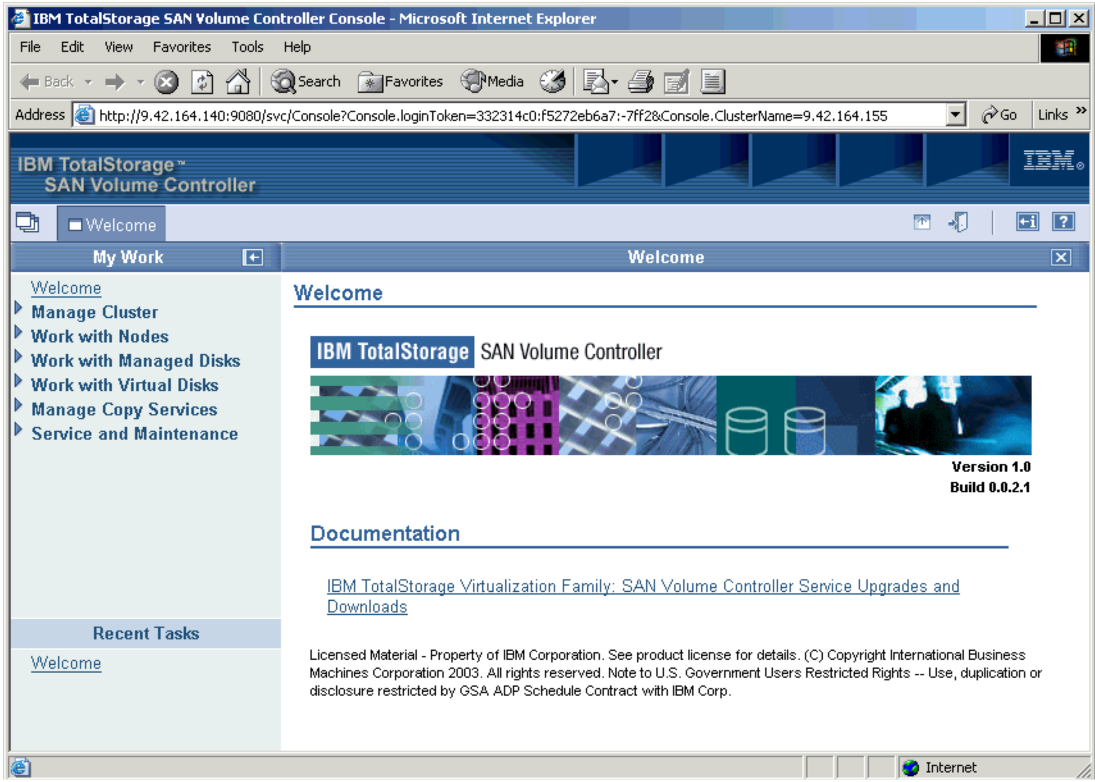


Figure 6-10 SVC Console Welcome page

6. The Viewing Nodes panel (Figure 6-11) opens. Note the input/output (I/O) group name (for example, `ITSO_grp0`). Select the node you want to add. Ensure that **Add a node** is selected from the list and click **Go**.

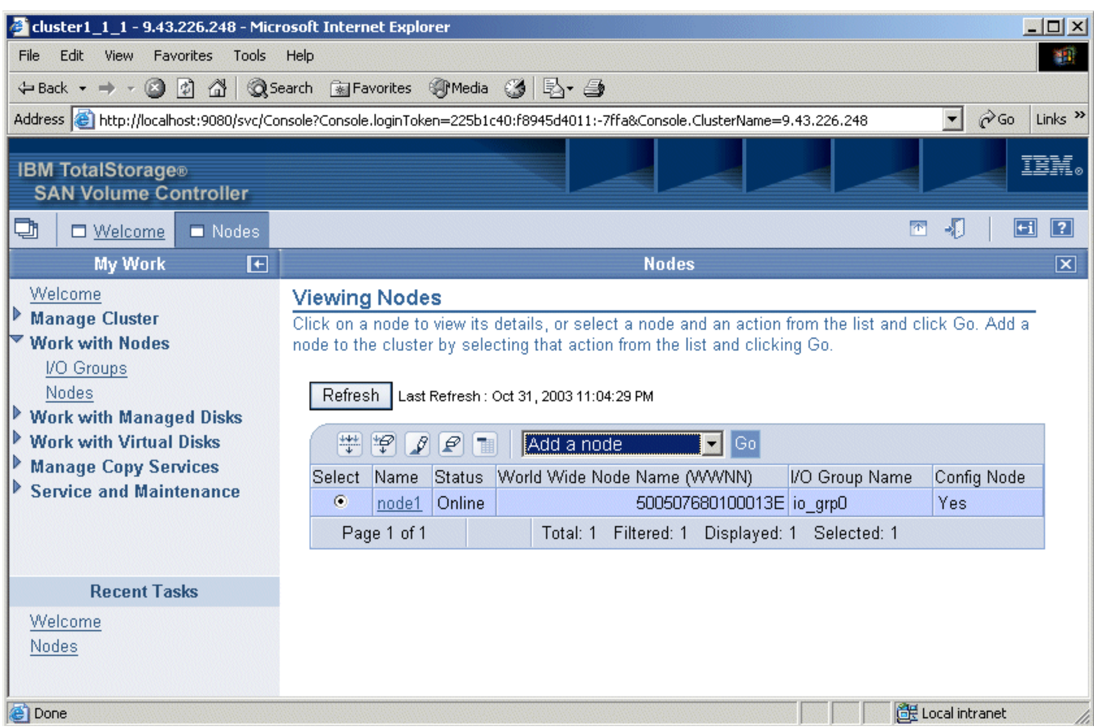


Figure 6-11 Viewing Nodes panel

Note: We renamed the existing node to `ITSO_node1` (we show how to do this later). On your panel, it should appear as `node1` by default.

7. The next panel (Figure 6-12) displays the available nodes. Select the node. Associate it with an I/O group and provide a name (for example, `ITSO_node2`). Click **OK**.

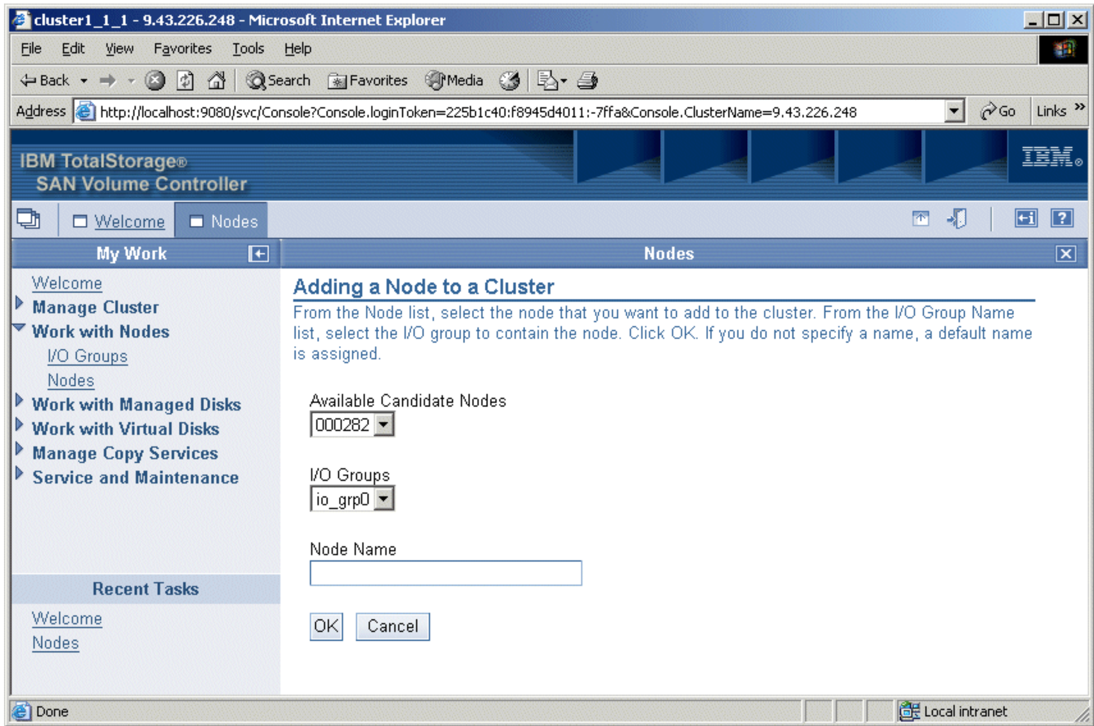


Figure 6-12 Adding a Node to a Cluster panel

Note: If you do not provide a name, the SVC automatically generates the name *nodeX*, where *X* is the ID sequence number assigned by the SVC internally. In our case, this is `node2`.

If you want to provide a name (as we have), you can use letters A to Z, a to z, numbers 0 to 9, and the underscore. It can be between one and 15 characters in length, but cannot start with a number or the word *node* since this prefix is reserved for SVC assignment only.

In our case, we only have enough nodes to complete the formation of one I/O group. Therefore, we added our new node to the I/O group that ITSO_node1 was already using, namely `ITSO_grp0` (which we renamed from the default of `iogrp0`; we explain this later).

If this panel does not display any available nodes, your second node is powered on, and zones are appropriately configured, pre-existing cluster configuration data may be stored on it. If you are sure this node is not part of another active SVC cluster, use the service panel to delete the existing cluster information. When this is complete, return to this panel and you should see the node listed.

For information about zoning requirements, see Chapter 3, “Planning and configuration” on page 25. For information about how to delete an existing cluster configuration using the service panel, see 4.3, “Basic installation” on page 74.

8. You return to the Viewing Nodes panel (Figure 6-13). It shows the status change of the node from Adding to *Online*.

> **Note:** This panel does not automatically refresh. Therefore, you continue to see the *adding* status only until you click the **Refresh** button.

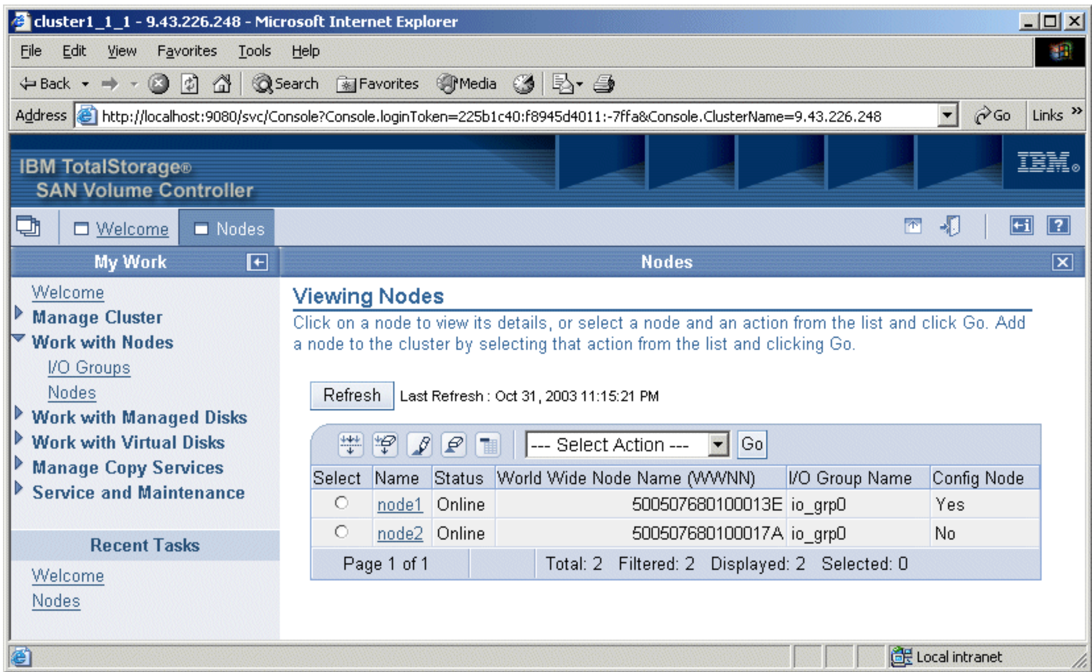


Figure 6-13 Node added successfully

You have completed the cluster configuration and have a fully redundant SVC environment.

6.3 Installing certificates

As we continue with setting up the SVC cluster, you come across many instances where you are prompted with security warnings regarding unrecognized certificates. The security warning panel (Figure 6-14) shows three options.

Figure 6-14 Security Alert window

These options are:

- **Yes**: Clicking yes accepts the certificate for this task. This option allows you to proceed using the unrecognized certificate. Each time you select a task which transmits secure information, you are prompted to accept another certificate. In most cases, you are prompted multiple times due to the two-way data exchange, which occurs between the management workstation and the SVC cluster. In some cases, this can cause your browser to crash.
- **No** (default): Clicking this option rejects the certificate for this task and does not allow you to proceed.
- **View Certificate**: Clicking this option launches the Certificate window (Figure 6-15), from where you can install the certificate. If you do not want to be prompted repeatedly to accept or reject certificates, we recommend you choose this option.

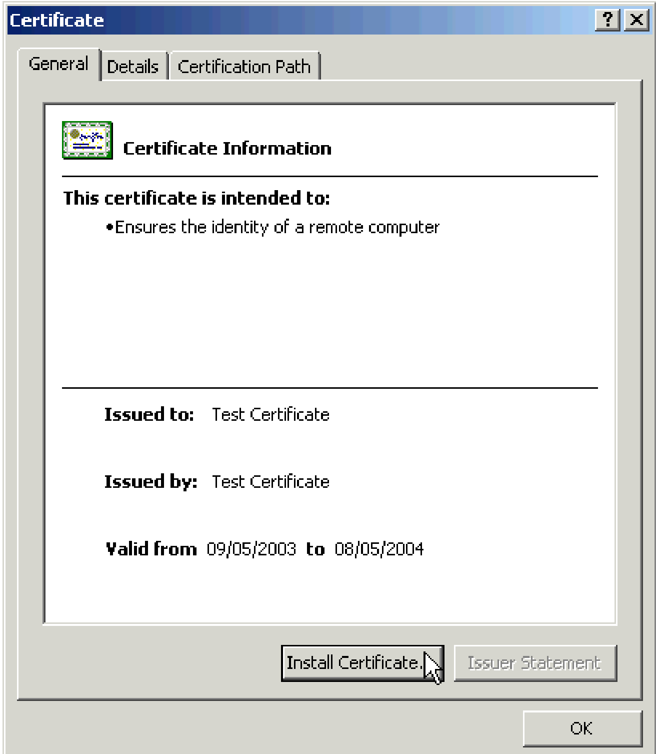


Figure 6-15 Certificate Information

Follow these steps to install a certificate:

1. From the Security Alert window (Figure 6-14), select **View Certificate**.
2. The Certificate window (Figure 6-15) opens. Click **Install Certificate**.
3. The Welcome to the Certificate Import Wizard information panel (Figure 6-16) opens. Click **Next**.

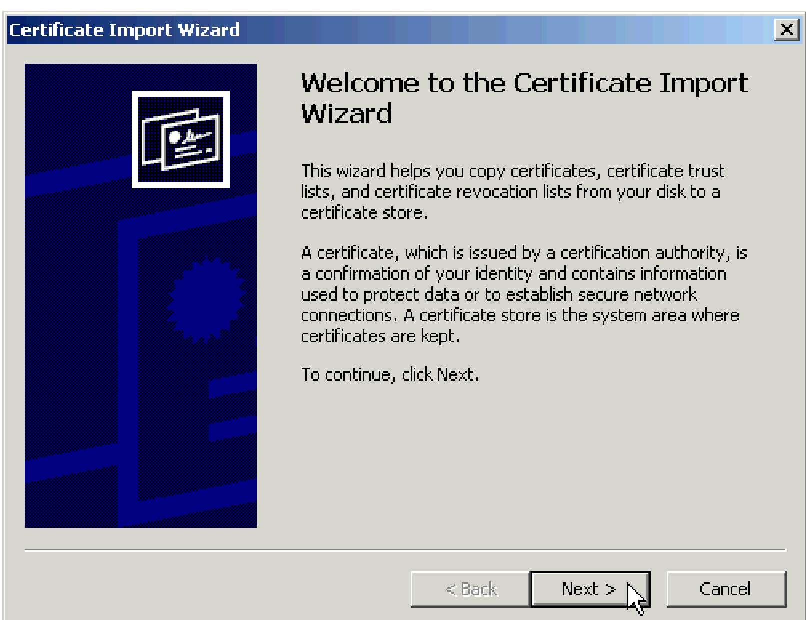


Figure 6-16 Certificate Import Wizard

4. On the Certificate Store panel (Figure 6-17), click **Next**.

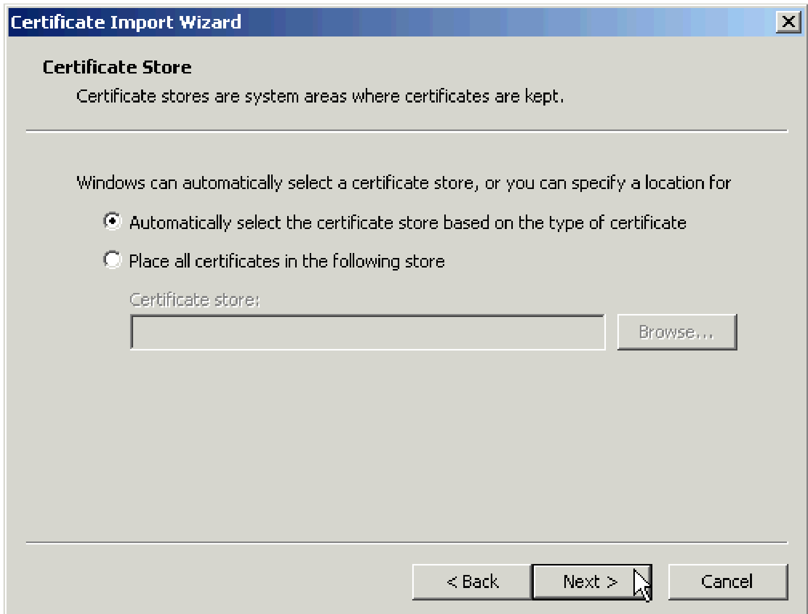


Figure 6-17 Certificate Store panel

5. You may be prompted with the Root Certificate Store confirmation window (Figure 6-18). If you are, click **Yes**.

Figure 6-18 Root Certificate Store

6. You should see a message stating that the import was successful (Figure 6-19). Click **OK**.

Figure 6-19 Certificate Import successful

7. You return to the Certificate Information window (Figure 6-15) you saw earlier. Click **OK**.
8. Provide the `admin` user ID and password when prompted.

From this point, you should no longer be asked to accept or reject certificates from the SVC cluster.

Note: Future code upgrades may result in new certificate IDs, so you may have to go through this process again.

6.4 Setting the cluster time zone and time

Perform the following steps to set the cluster time zone and time:

1. From the SVC Welcome page (Figure 6-10 on page 109), select the **Manage Cluster** option and the **Set Cluster Time** link.
2. The Cluster Date and Time Settings panel (Figure 6-20) opens. At the top of the panel, you see the current settings. If necessary, make adjustments and ensure that the **Update cluster data and time** and **Update cluster time zone** check boxes are selected. Click **Update**.

Note: You may be prompted for the cluster user ID and password. If you are, enter `admin` and the password you set earlier.

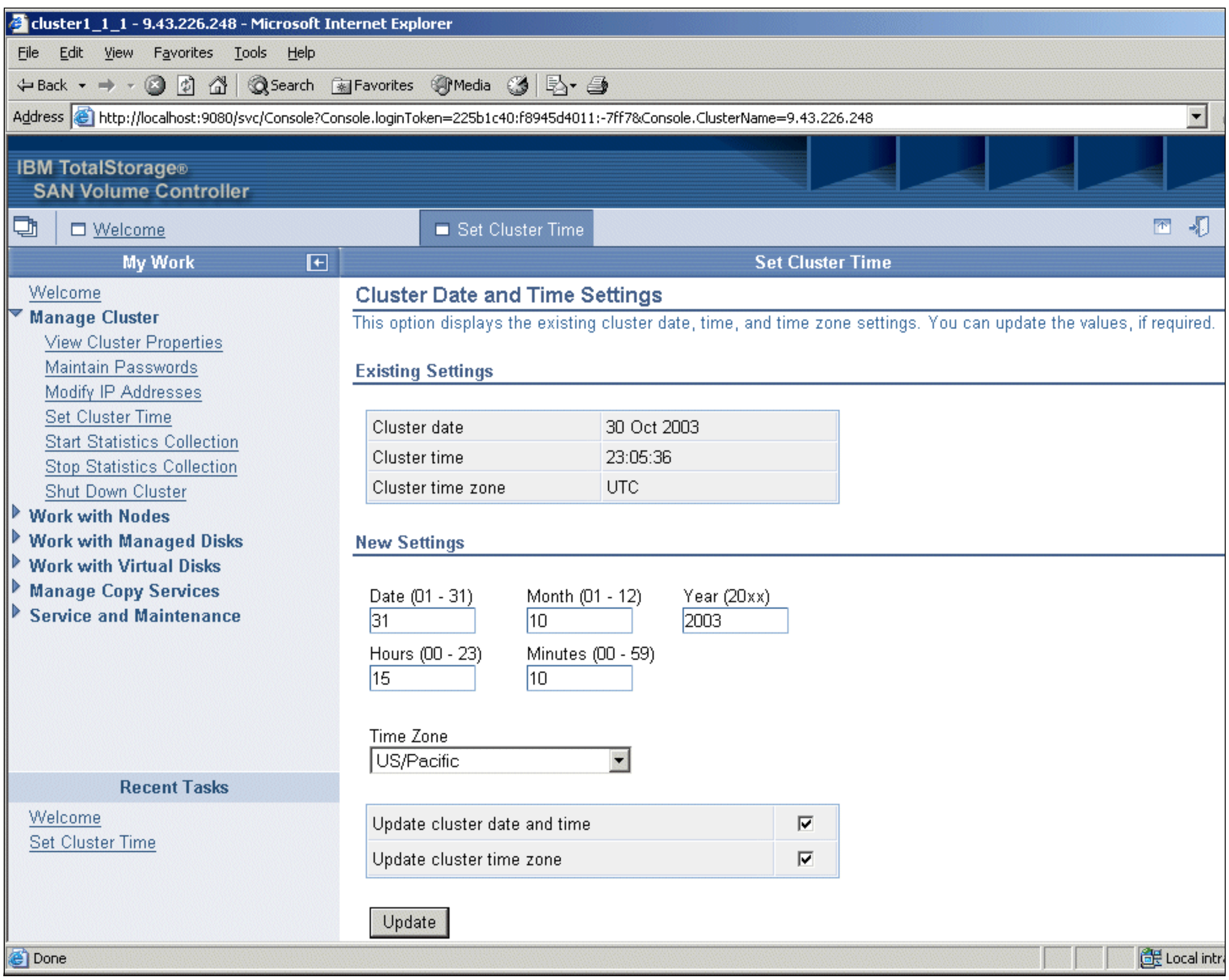


Figure 6-20 Cluster Date and Time Settings panel

3. You see the message “`The ClusterTime Zone Setting Has Been Updated`”.

You have now completed the tasks necessary to set the cluster time zone and time.

6.5 Creating host definitions

Perform the following steps to create host objects within the SVC:

1. From the SVC Welcome page (Figure 6-10 on page 109), select the **Working with Virtual Disks** option and then the **Hosts** link.
2. You see the Filtering Hosts panel (not shown), which we explain later. Click the **Bypass filter** button at the top of this panel.
3. The Viewing Hosts panel (Figure 6-21) opens. Select **Create a host** from the list and click **Go**.

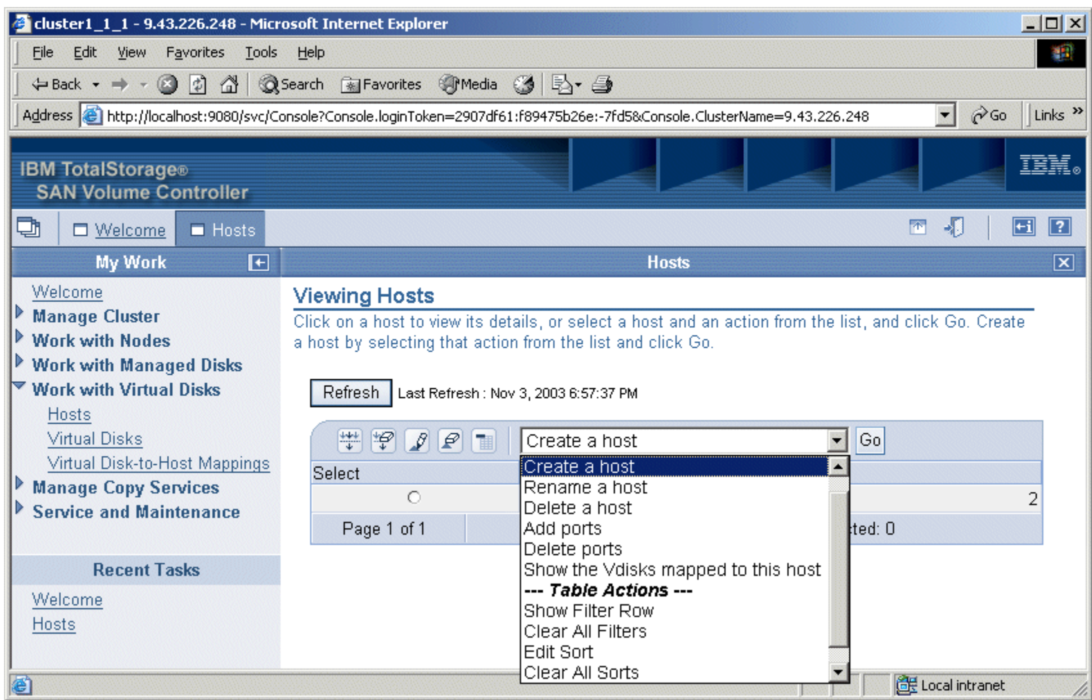


Figure 6-21 Viewing Hosts panel

4. On the Creating Hosts panel (Figure 6-22), follow these steps:
   a. Type a name for your host (for example, `Win2K_2`).

   > **Note:** If you do not provide a name, the SVC automatically generates the name *hostX*, where *X* is the ID sequence number assigned by the SVC internally.
   >
   > If you want to provide a name (as we have), you can use the letters A to Z, a to z, numbers 0 to 9, and the underscore. It can be between one and 15 characters in length. However, it cannot start with a number or the word *host* because this prefix is reserved for SVC assignment only.

   b. From the Available Port list, select the WWN or WWNs, one at a time, and click the **Add** button.

c. When you are done adding the WWNs, click **OK**.

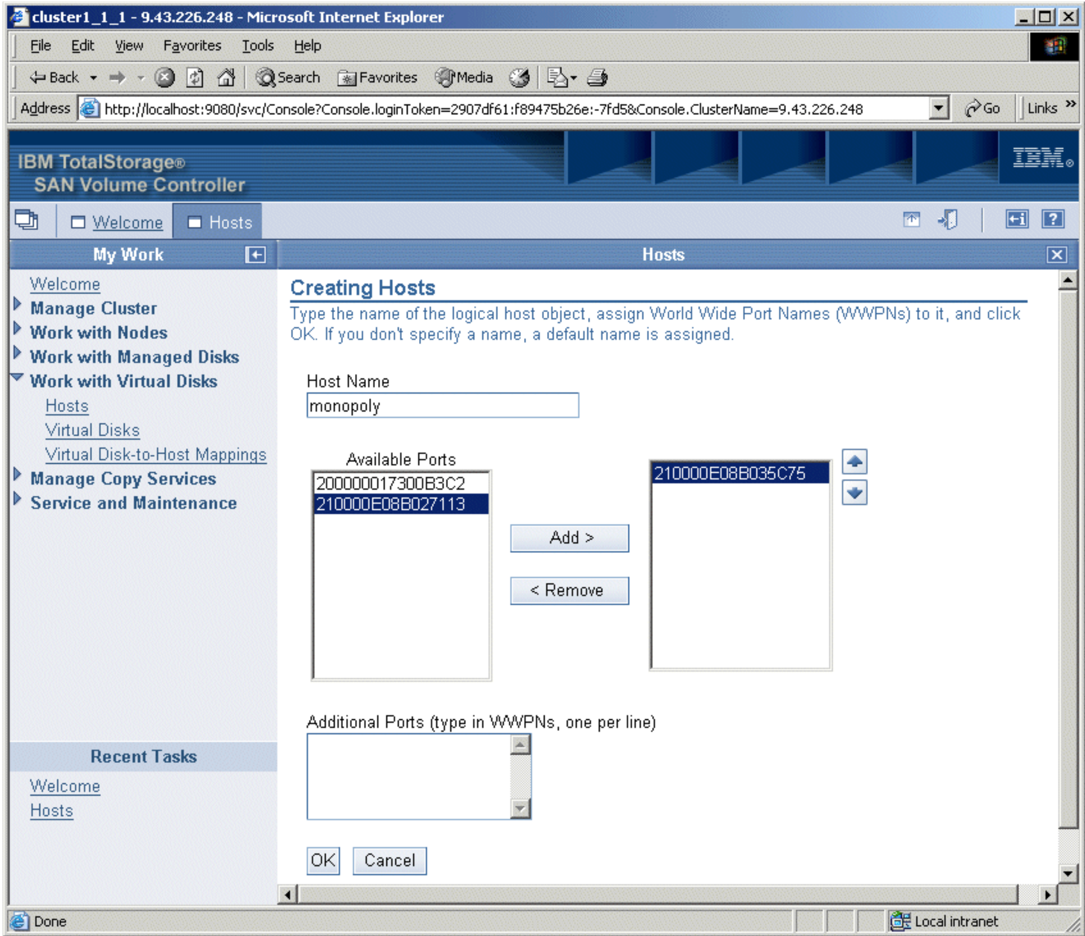


Figure 6-22 Creating Hosts panel

Note: This panel shows all WWNs that are visible to the SVC and that have not already been defined to a host. If your WWN does not appear, check that the host has logged into the switch and that zoning is updated to allow SVC and host ports to see each other. This is described in Chapter 3, “Planning and configuration” on page 25.

5. You return to the Viewing Hosts panel (Figure 6-23) where you should see your newly created host.

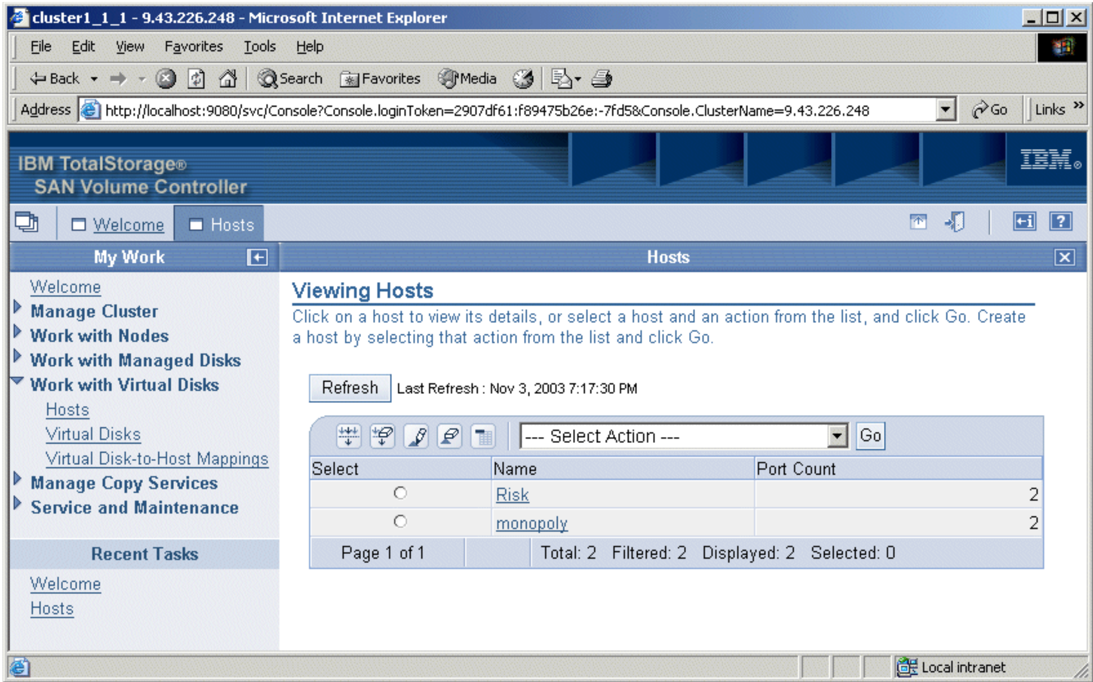


Figure 6-23 Host added successfully

For information about other tasks, such as adding host ports, deleting host ports, or deleting hosts, see Chapter 9, “SVC configuration and administration using the GUI” on page 217.

You have now completed the tasks required to add host definitions to your SVC configuration.

6.6 Displaying managed disks

Perform the following steps to display mDisks:

1. From the SVC Welcome page (Figure 6-10 on page 109), select the **Work with Managed Disks** option and then the **Managed Disks** link.
2. When the Filtering Managed Disks (mDisk) panel opens, click **Bypass filter** to open the Viewing Managed Disks panel.
3. On the Viewing Managed Disks panel (Figure 6-25), if your mDisks are not displayed, rescan the Fibre Channel network. Select **Discover mDisks** from the list and click **Go**.

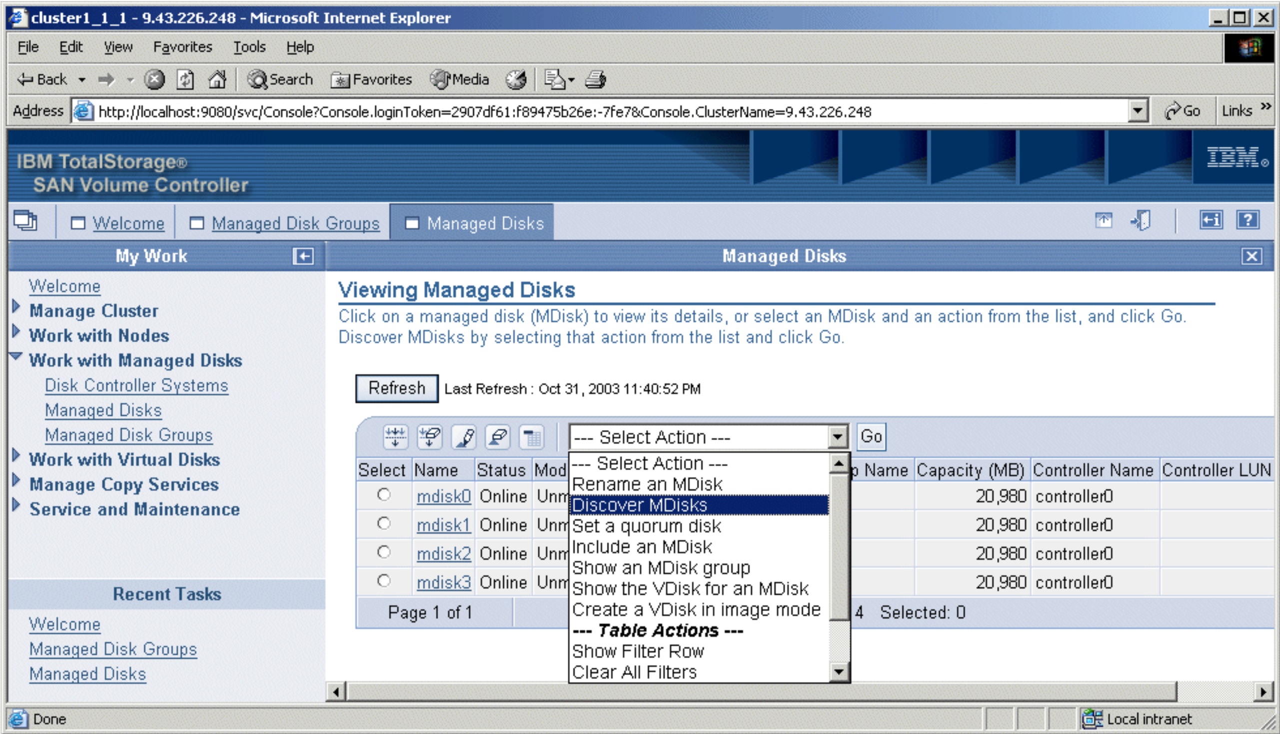


Figure 6-24 Discover mDisks

> **Note:** If your mDisks are still not visible, check that the logical unit numbers (LUNs) from your subsystem are properly assigned to the SVC and that appropriate zoning is in place (for example, SVC can see the disk subsystem). See Chapter 3, "Planning and configuration" on page 25, for more details about how to setup your storage area network (SAN) fabric.

6.7 Creating MDGs

Perform the following steps to create a managed disk group (MDG):

1. From the SVC Welcome page (Figure 6-10 on page 109), select the **Work with Managed Disks** option and then the **Managed Disks Groups** link.
2. When the Filtering Managed Disks (mDisk) Groups panel opens, click **Bypass filter**.
3. The Viewing Managed Disks Groups panel (Figure 6-25) opens. Select **Create an mDisk Group** from the list and click **Go**.

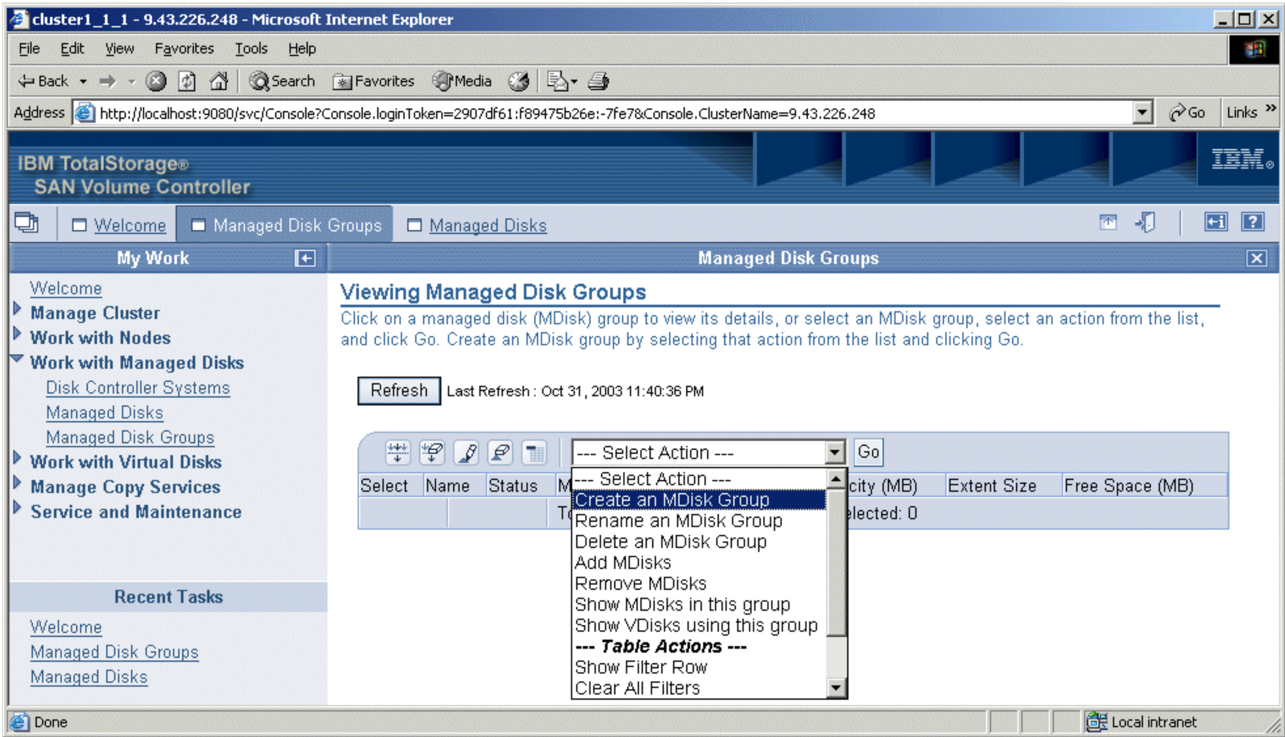


Figure 6-25 Selecting the option to create an mDisk group

4. On the Create Managed Disk Group panel (Figure 6-26), click **Next**.

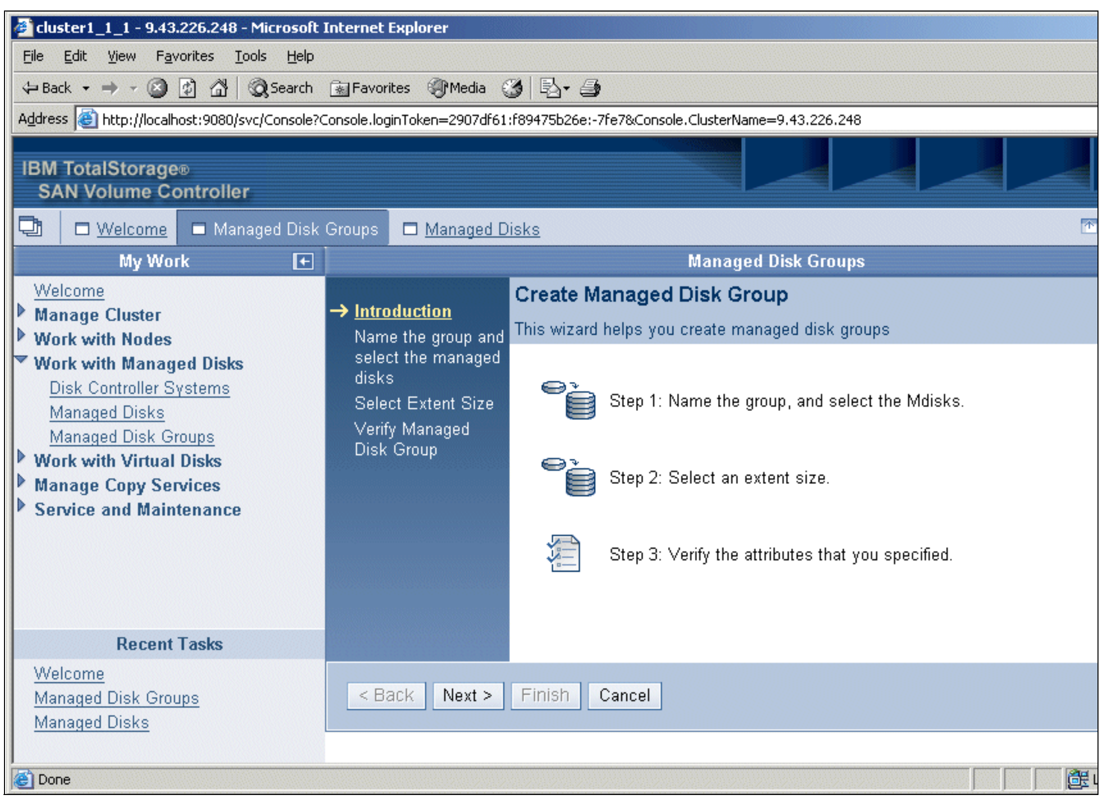


Figure 6-26 Create Managed Disk Group panel

5. On the Name the group and select the managed disks panel (Figure 6-27), follow these steps:

   a. Type a name for your MDG.

> **Note:** If you do not provide a name, the SVC automatically generates the name *mDiskgrpX*, where *X* is the ID sequence number assigned by the SVC internally.
>
> If you want to provide a name (as we have), you can use the letters A to Z, a to z, numbers 0 to 9, and the underscore. It can be between one and 15 characters in length, but cannot start with a number or the word *mDiskgrp* because this prefix is reserved for SVC assignment only.

   b. From the mDisk Candidates box, one at a time, select the mDisks you want to put into the MDG. Click **Add** to move them to the Selected mDisks box.

   c. Click **Next**.

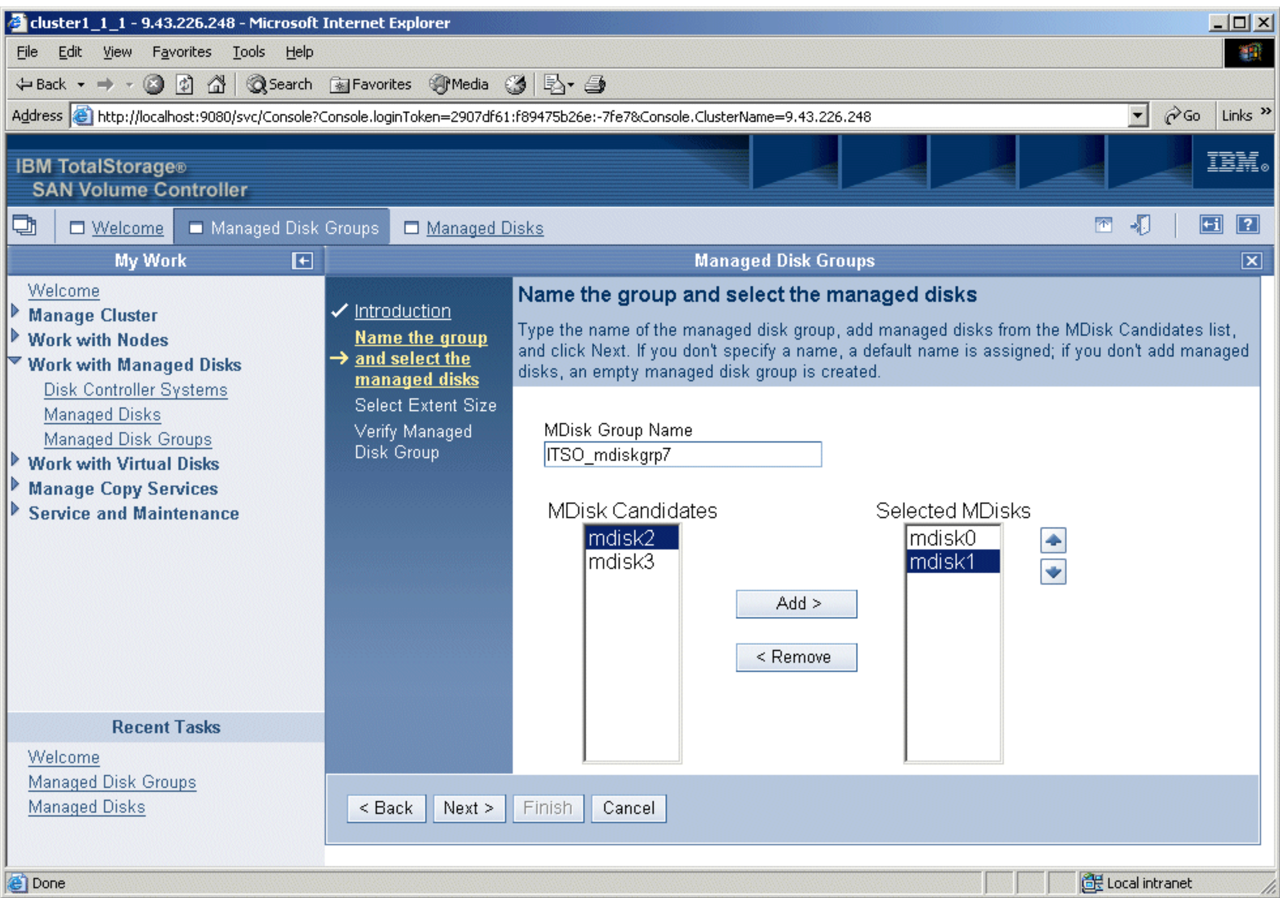


Figure 6-27 Name the group and select the managed disks panel

6. From the list shown in Figure 6-28, select the extent size you want to use. Then click **Next**.

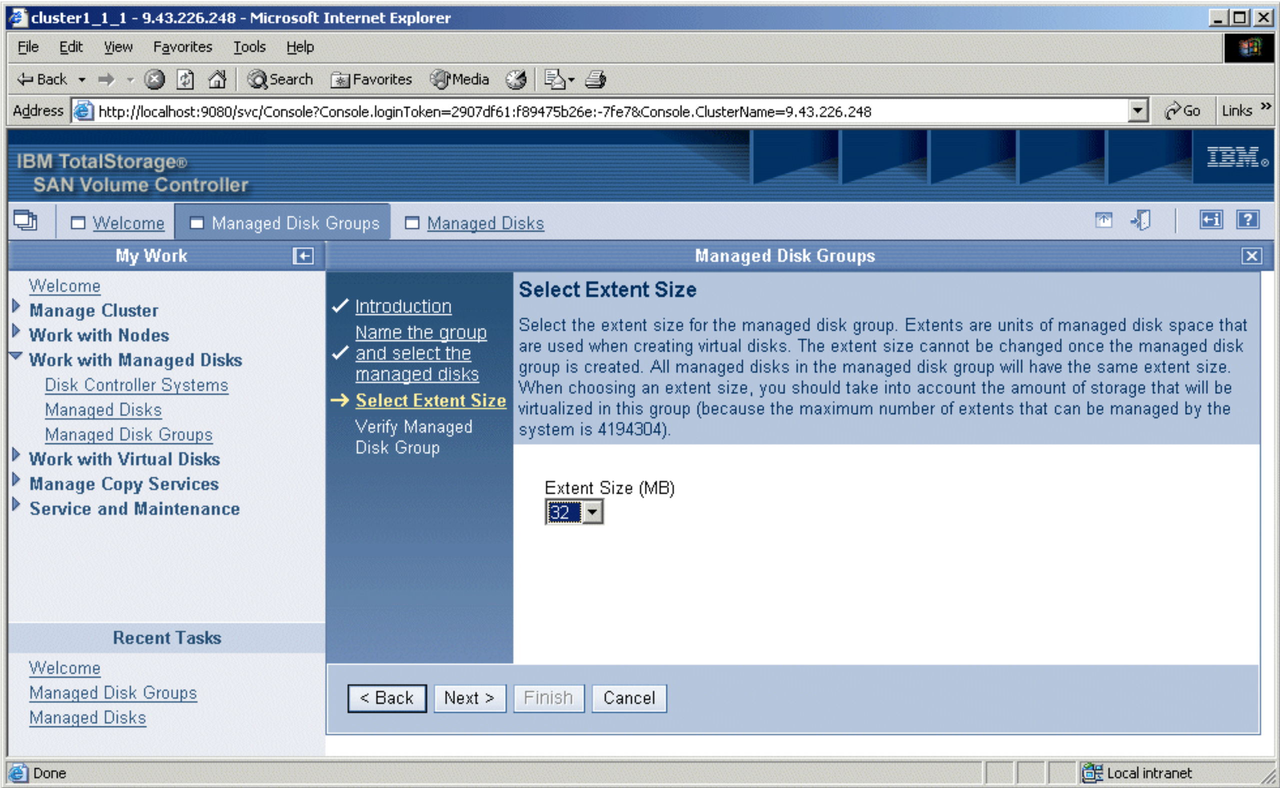


Figure 6-28 Select Extent Size panel

7. On the Verify Managed Disk Group panel (Figure 6-29), verify that the information you specified is correct. Then click **Finish**.

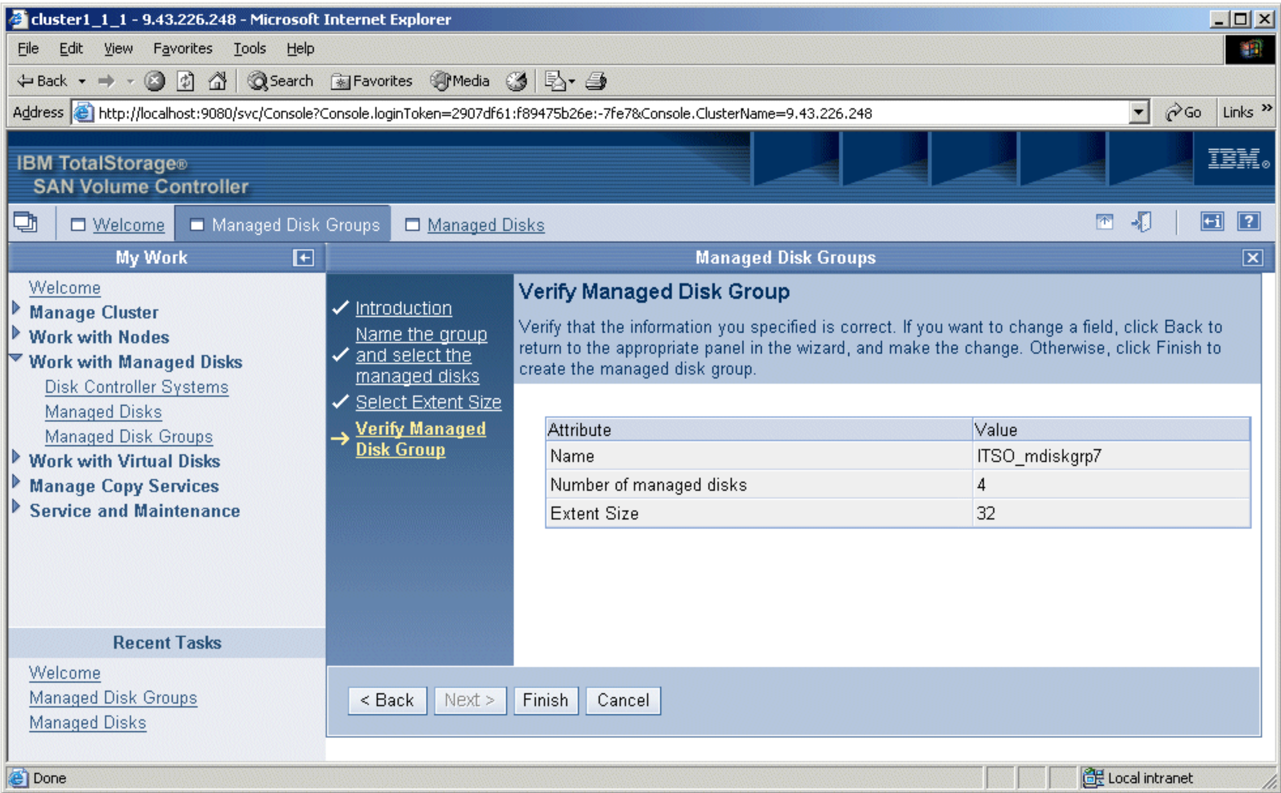


Figure 6-29 Verify MDG wizard

8. You should return to the Viewing Managed Disk Groups panel (Figure 6-30) where your new MDG (in our case ITSO_mDiskgrp7) is displayed.

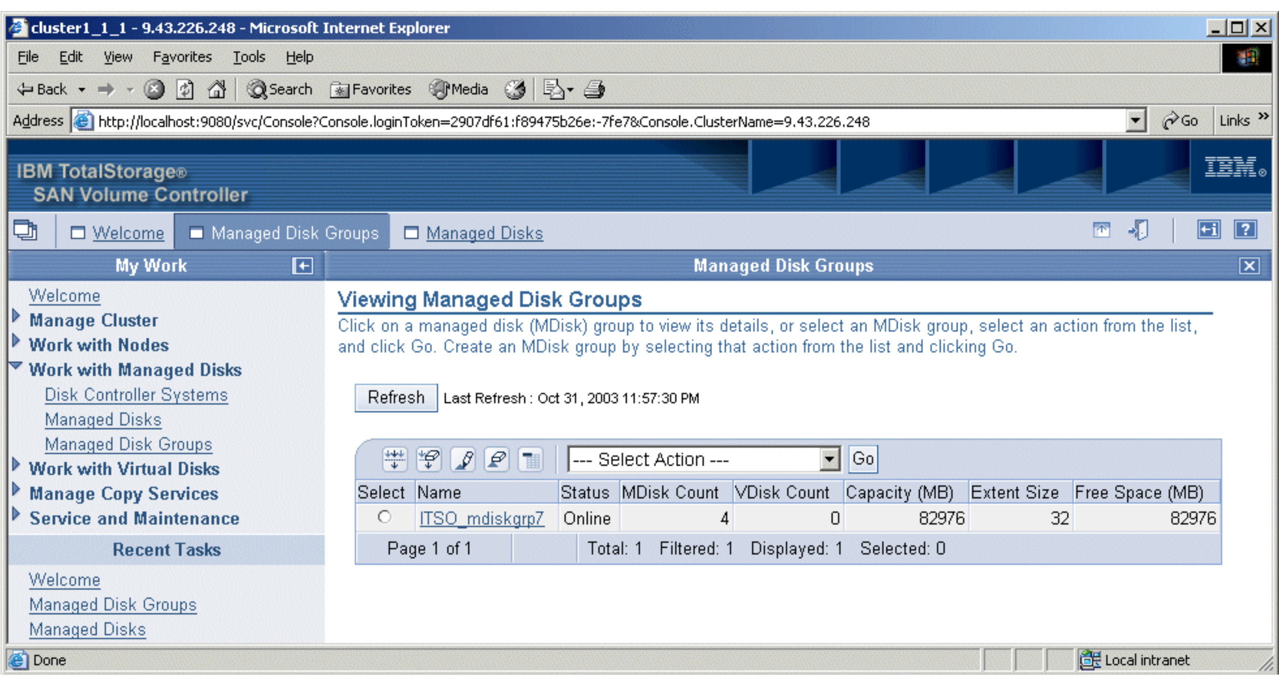


Figure 6-30 MDG added successfully

Note: As you can see, the mDisk Count in our MDG is 0 because we did not select any mDisk Candidates in the previous step that you should have.

For information about other tasks, such as adding mDisks to MDGs and renaming MDGs or deleting MDGs, see Chapter 9, "SVC configuration and administration using the GUI" on page 217.

You have completed the tasks required to create a MDG.

6.8 Creating a vDisk

Perform the following steps to create vDisks:

1. From the SVC Welcome page (Figure 6-10 on page 109), select the **Work with Virtual Disks** option and then the **Virtual Disks** link.
2. When the Filtering Virtual Disks (vDisks) panel opens, click **Bypass filter**.
3. The Viewing Virtual Disks panel (Figure 6-31) opens. Select **Create vDisk** from the list. Click **Go**.

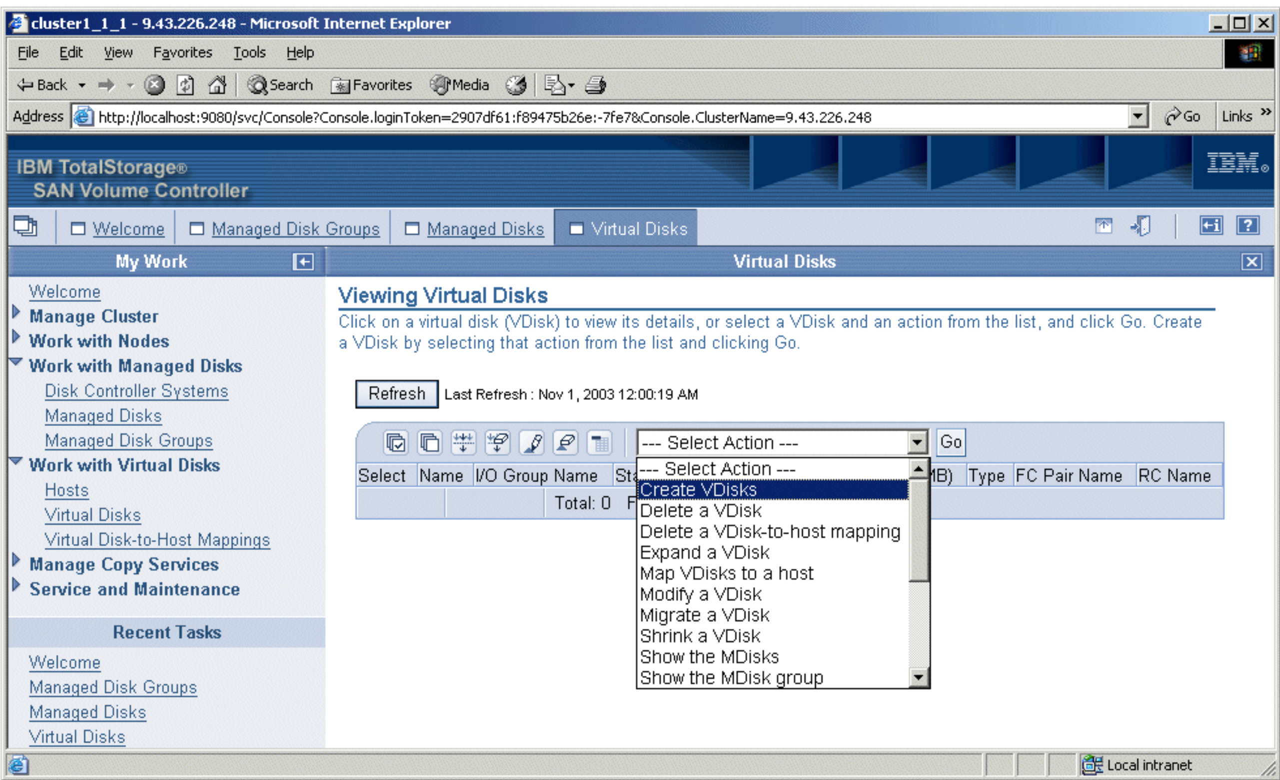


Figure 6-31 Viewing Virtual Disks panel

4. The Create Virtual Disks wizard (Figure 6-32) opens. Click the **Next** button.

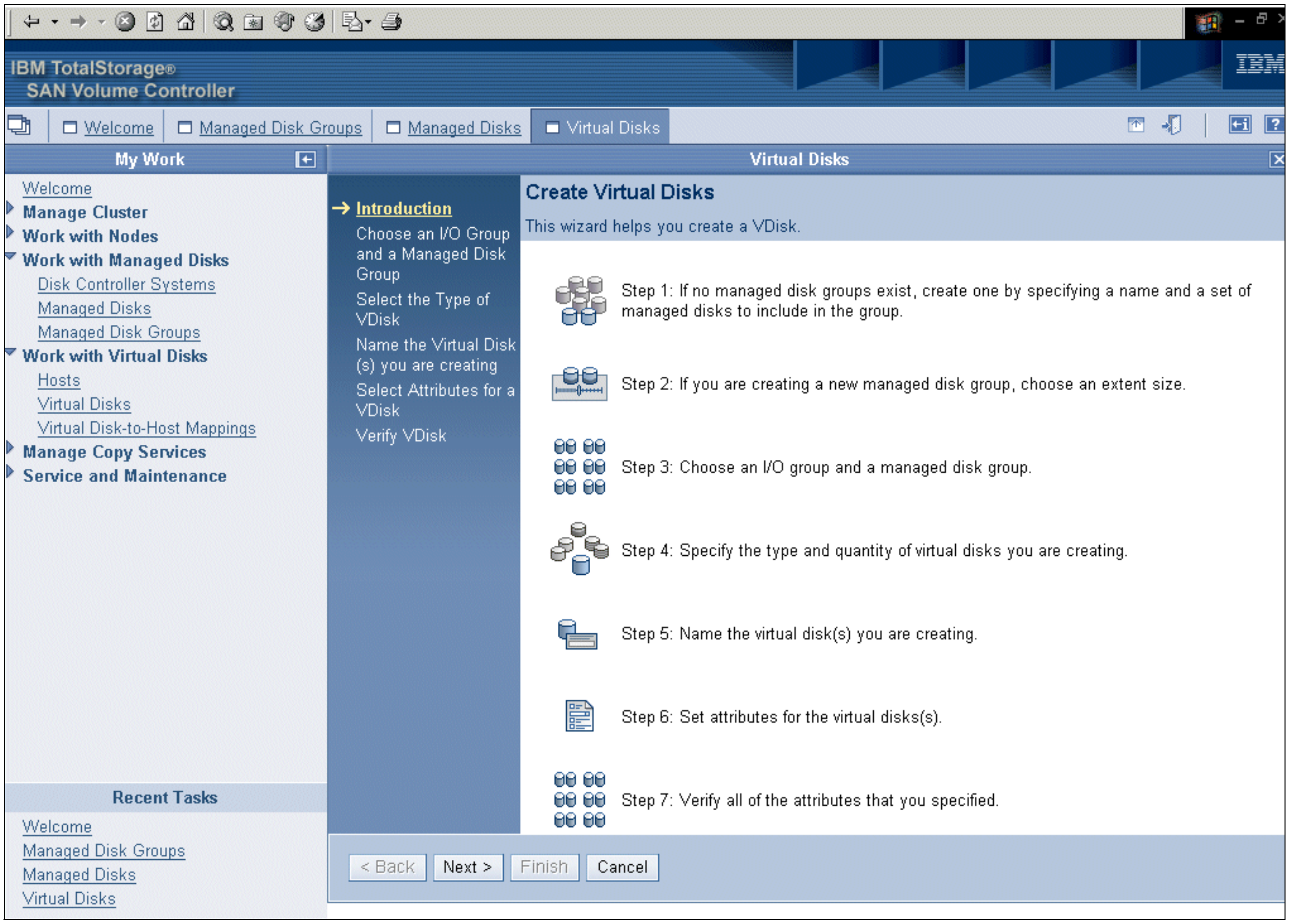


Figure 6-32 Create Virtual Disks wizard

5. On the Choose an I/O Group and a Managed Disk Group panel (Figure 6-33), follow these steps:

   a. Select the I/O group you want to associate the vDisk with from the list. In our case, we only have one, ITSO_grp0, so we must select it from the list.

   > **Note:** You can let the system choose the preferred node and I/O group.

   b. Optionally, choose a preferred node. The default (if nothing is selected) is to alternate between nodes in the I/O group.

   c. Select the mDisk group in which to create the vDisk from the list.

   d. Click **Next**.

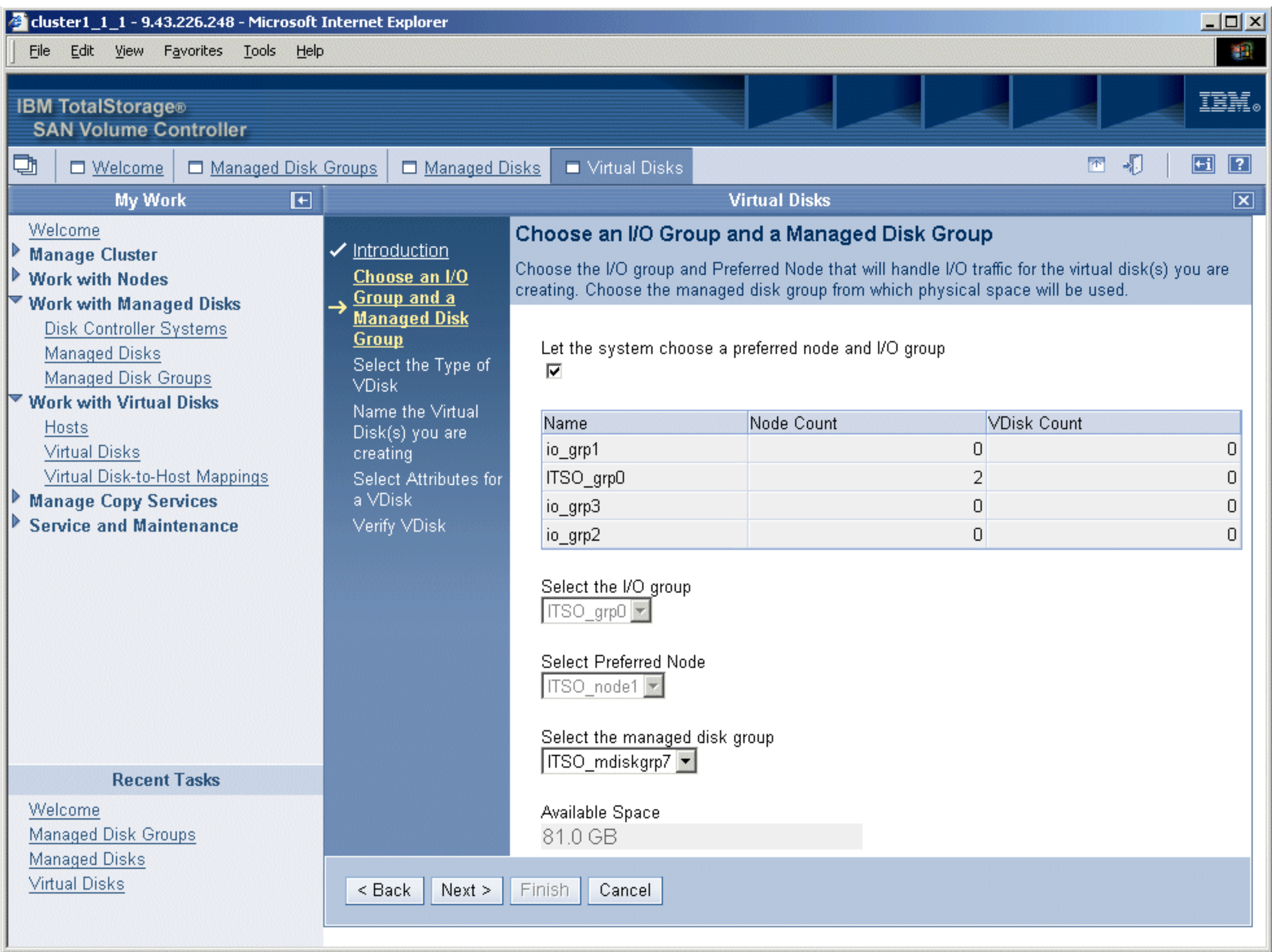


Figure 6-33 Choosing an I/O group and a MDG panel

6. On the Select the Type of vDisk panel (Figure 6-34), follow these steps:
   a. Select the type of vDisk you want to create (striped, sequential or image) from the list.
   b. Select the number of vDisks you want to create.
   c. Click **Next**.

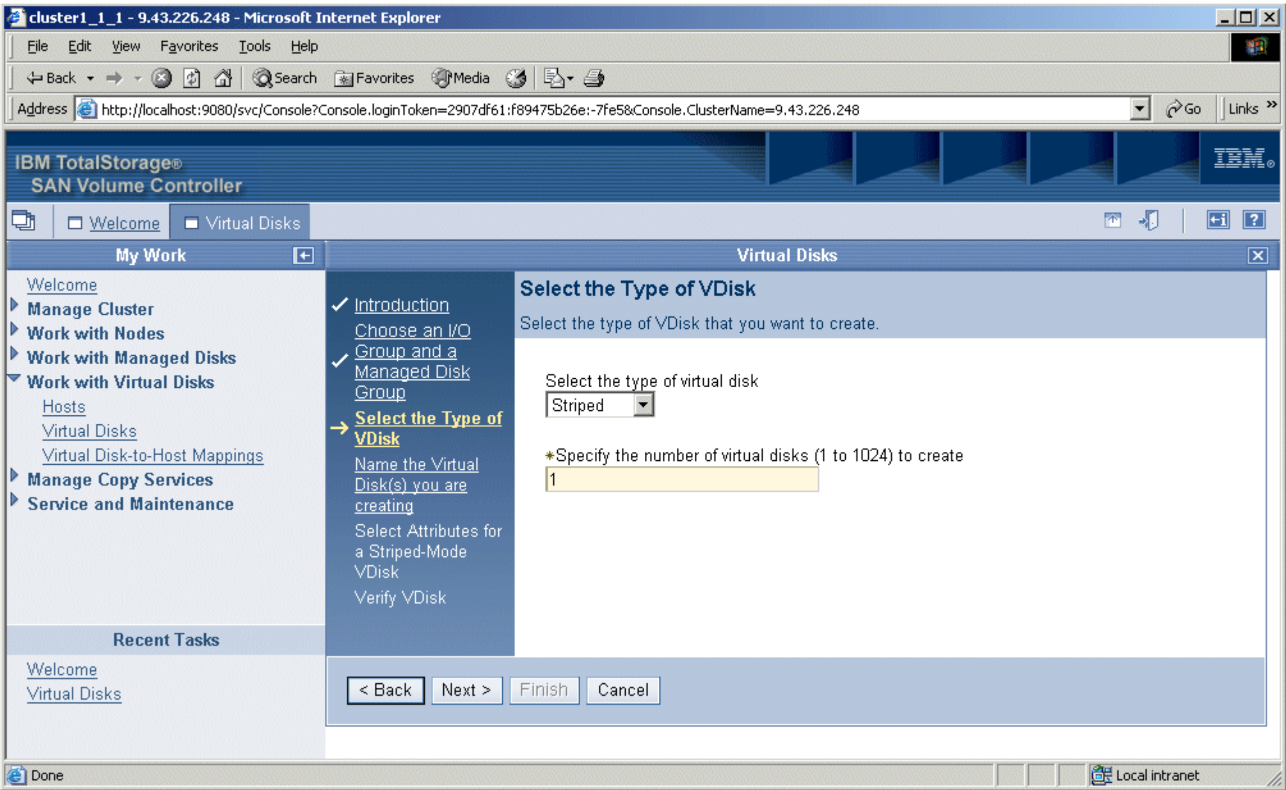


Figure 6-34 Select the Type of vDisk panel

7. On the Name the Virtual Disk(s) you are creating panel (Figure 6-35), type a name for your vDisk. Click **Next**.

> **Note:** If you do not provide a name, the SVC automatically generates the name *vDiskX*, where *X* is the ID sequence number assigned by the SVC internally.
>
> If you want to provide a name (as we have), you can use letters A to Z, a to z, numbers 0 to 9, and the underscore. It can be between one and 15 characters in length. However, it cannot start with a number or the word *vDisk* because this prefix is reserved for SVC assignment only.

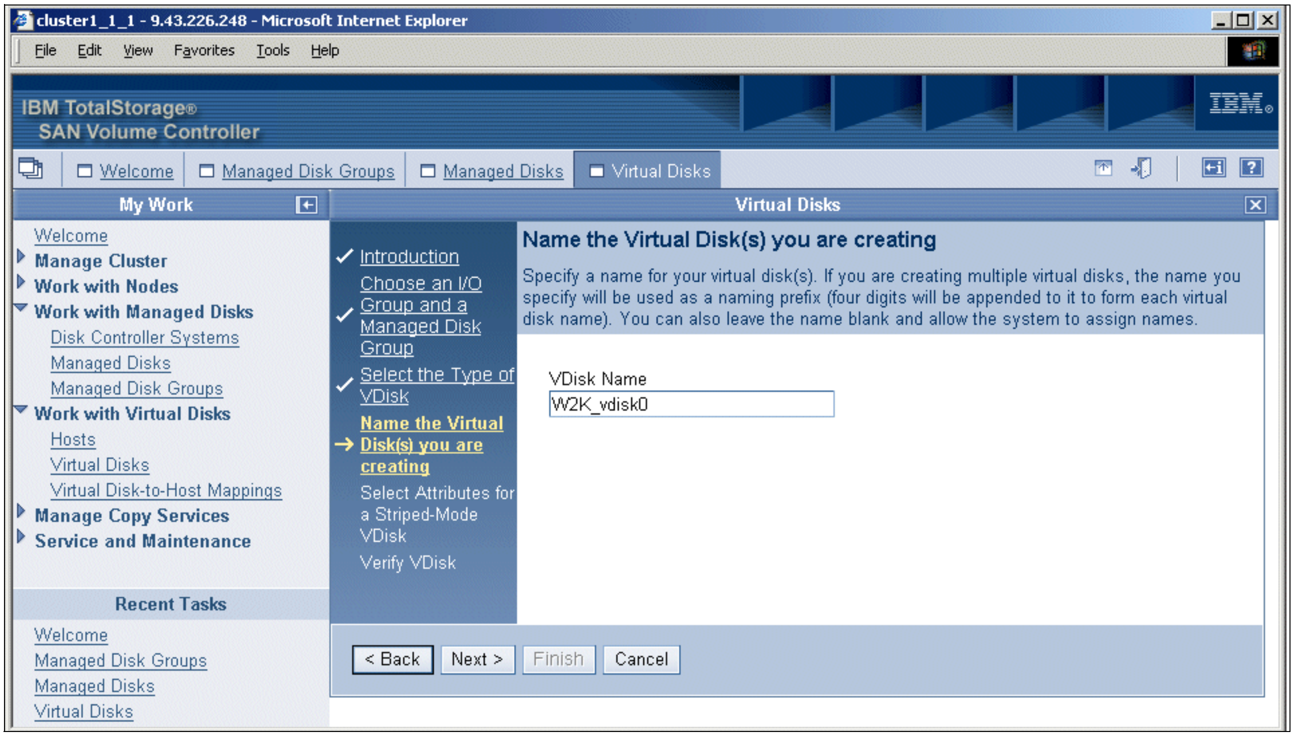


Figure 6-35 Name the Virtual Disk(s) you are creating panel

8. On the Select Attributes for *modetype*-mode vDisk panel (*modetype* is the type of vDisk you selected in the previous step) as shown in Figure 6-36, follow these steps:

   a. Optionally, choose the Managed Disk Candidates upon which to create the vDisk. Click **Add** to move them to the Managed Disks Striped in this Order box.

      Striped vDisks, by default, use all mDisks within a MDG. Therefore, it is not necessary to select anything here. However, you may want to select from the list, for example, if you want to specify that the vDisk only uses a subset of the mDisks available within a MDG.

      For image and sequential vDisks, you do not see the managed disk candidates or managed disks striped in the Order box. Instead you see the Managed Disk Used to Create vDisk and the top one in the list selected by default.

   b. Type the capacity of the vDisk. Select the unit of capacity from the list.

      Remember, capacity is calculated based on 1 GB = 1024 MB. Therefore, an entry of 10 GB actually provides 10240 MB instead of 10000 MB as with other disk subsystems.

   c. Optionally, choose to format the vDisk by selecting the check box.

      Note: Formatting destroys any existing data on the vDisk.

   d. After you complete all the necessary entry fields, click **Next**.

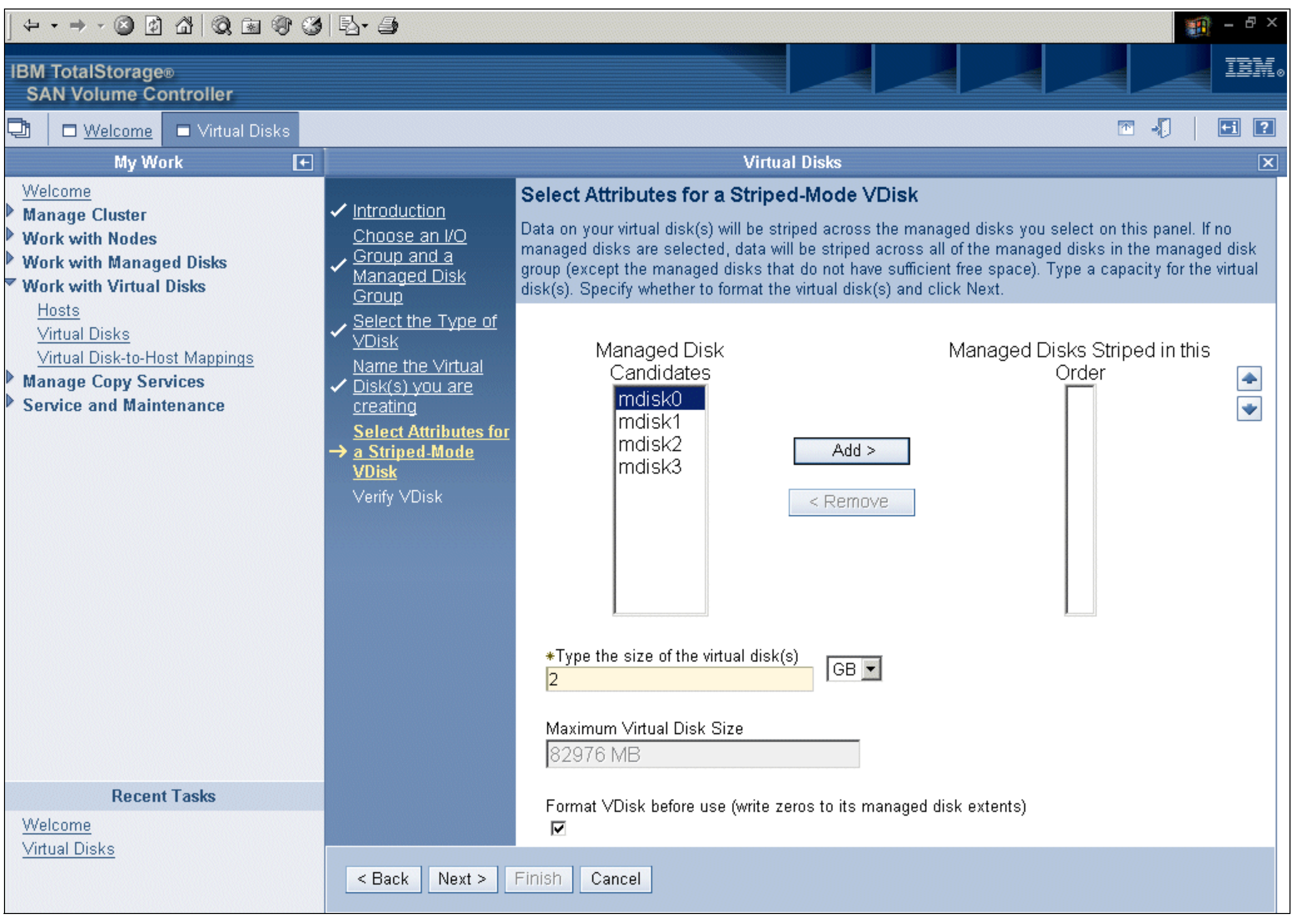


Figure 6-36 Select Attributes for a vDisk panel

9. On the Verify vDisk panel (Figure 6-37), verify your selections. You can select the **Back** button at any time to make changes.

   If you are satisfied with your selections, click **Finish**. You return to the Viewing Virtual Disks panel where you should now see your newly created vDisk.

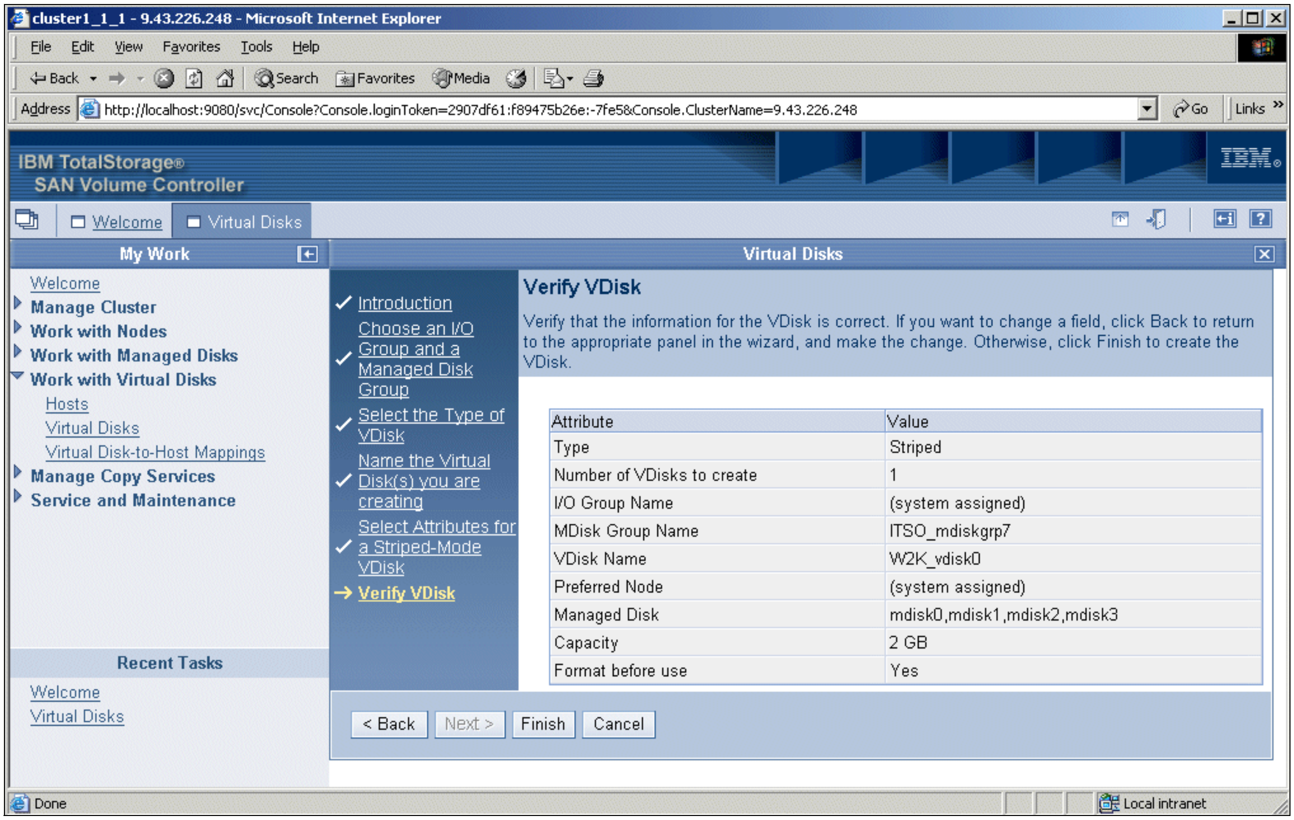


Figure 6-37 Verify vDisk

For information about other tasks, such as deleting a vDisk, renaming a vDisk or expanding a vDisk, see Chapter 9, “SVC configuration and administration using the GUI” on page 217.

You have now completed the tasks required to create a vDisk.

6.9 Assigning a vDisk to a host

Perform the following steps to map a vDisk to a host:

1. From the SVC Welcome page (Figure 6-10 on page 109), select the **Work with Virtual Disks** option and then the **Virtual Disks** link.
2. When the Filtering Virtual Disks (vDisks) panel opens, click **Bypass filter**.
3. On the Viewing Virtual Disks panel (Figure 6-38), select the radio button next to the vDisk you want to assign. Select **Map a vDisk to a host** from the list and click **Go**.

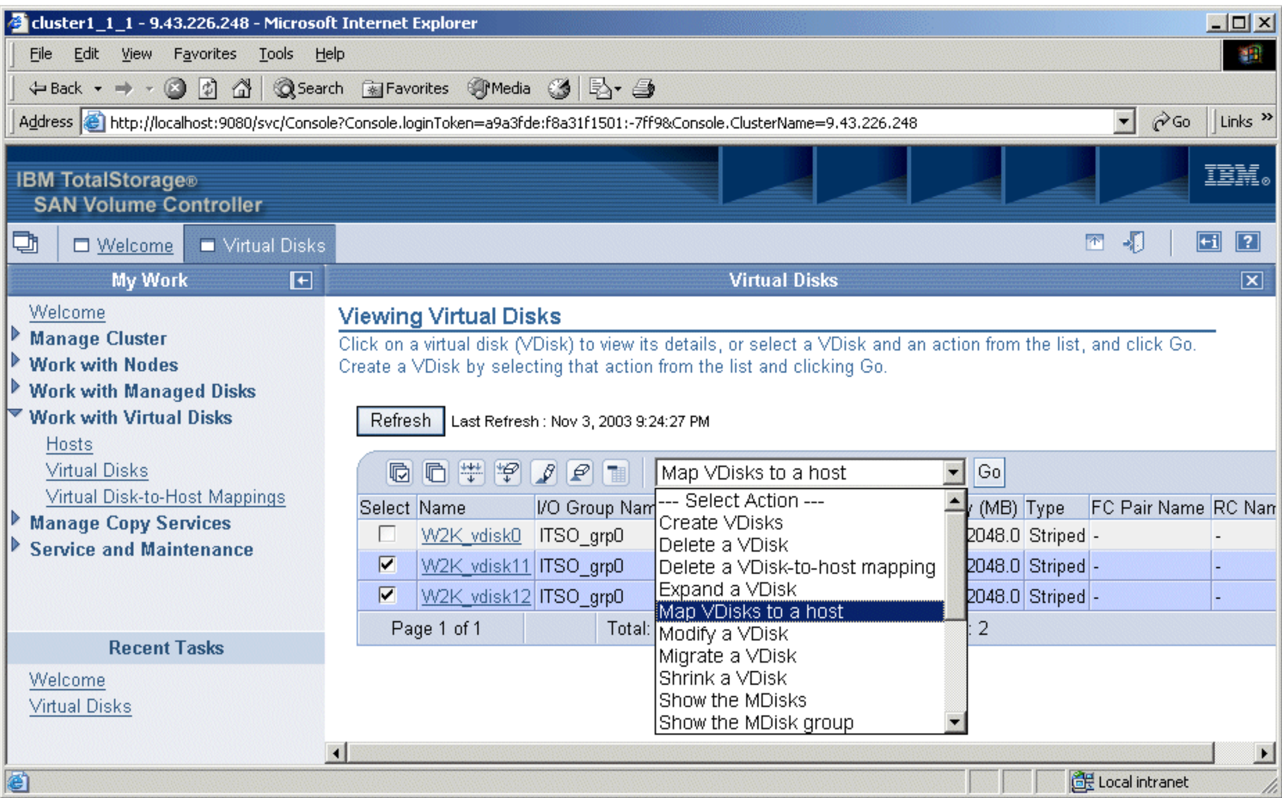


Figure 6-38 Assign a vDisk to a host

4. On the Creating a vDisk-to-Host Mappings panel (Figure 6-39), select the target host. Optionally assign a specific SCSI LUN ID, and click **OK**.

> **Note:** The SCSI LUN ID option allows you to assign a specific LUN ID to a vDisk that you are mapping to a host. The default, if nothing is selected, is to increment the SCSI LUN ID by one based on what is already assigned to the host.

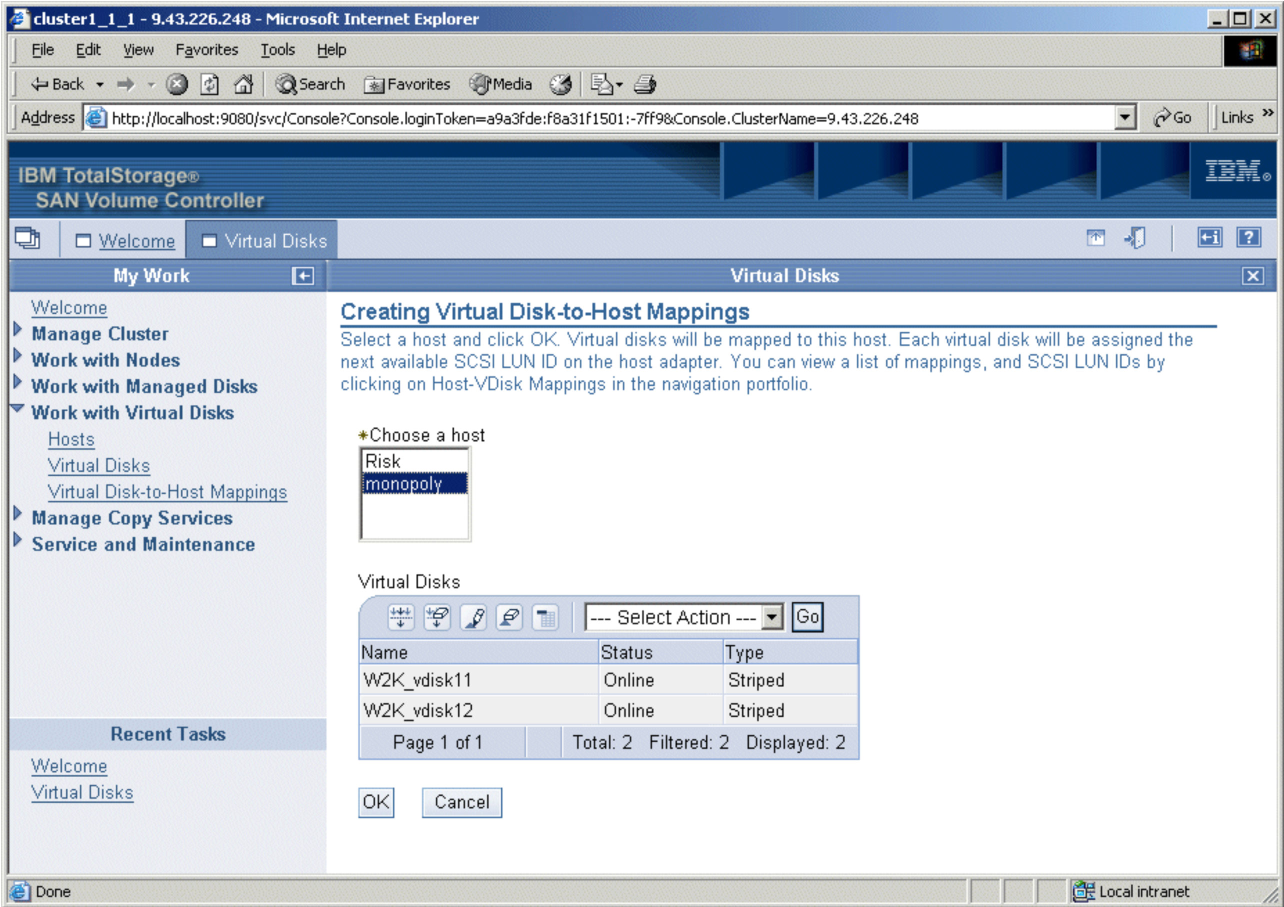


Figure 6-39 Creating vDisk-to-Host Mappings panel

5. You now return to the Viewing Virtual Disks panel.

For information about other tasks such as deleting a vDisk to host mapping, see Chapter 9, “SVC configuration and administration using the GUI” on page 217.

You have now completed all the tasks required to assign a vDisk to an attached host. You are ready to proceed to Chapter 7, “Host configuration” on page 133, to begin using the assigned vDisks.

7

Host configuration

This chapter describes the basic host configuration procedures required to connect a supported host to the IBM TotalStorage SAN Volume Controller (SVC).

7.1 SAN configuration

You must use great care when setting up the storage area network (SAN) to ensure that the zoning rules discussed in 3.4, “Zoning” on page 35, are followed. The SVC is configured so that all its ports are in an *SVC zone* together with psorts from the master console. There is a *storage zone* that comprises all SVC ports and all ports on the back-end storage system. Finally, there are multiple *host zones*, that each consist of one HBA port and one port from each of the two SVC nodes.

7.2 SVC setup

Figure 7-1 shows a basic configuration with multiple heterogeneous host connected to the SVC nodes through two switches.

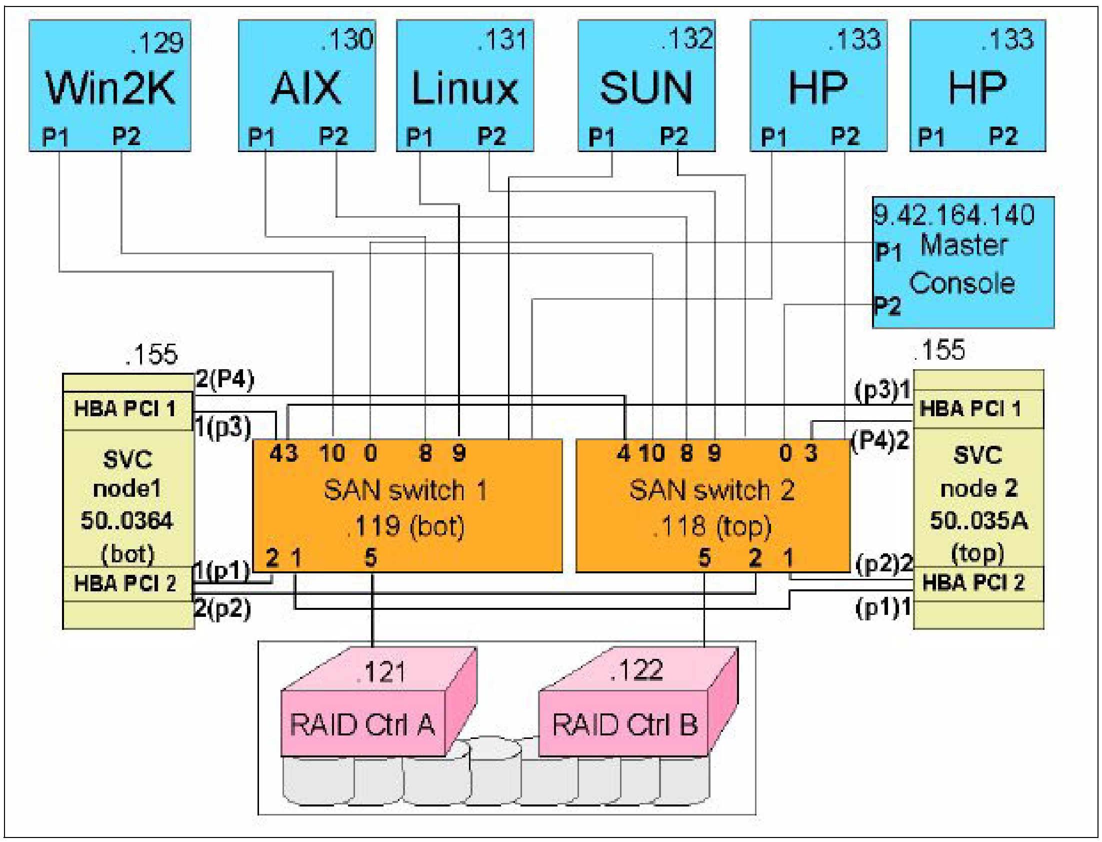


Figure 7-1 SAN Volume controller setup

7.3 Switch and zoning configuration

Example 7-1 lists the details of the zoning used on the switches in Figure 7-1.

Example 7-1 Switch 119 configuration SWITCH_1

```
2109_Switch_119:admin> switchshow
switchName:     2109_Switch_119
switchType:     9.2
switchState:    Online
switchMode:     Native
switchRole:     Principal
switchDomain:   1
switchId:       fffc01
switchWwn:      10:00:00:60:69:51:87:0a
switchBeacon:   OFF
Zoning:         ON (RedstoneConfig)
port  0: id N2 Online      F-Port 21:01:00:e0:8b:28:af:d6 --> Master Console
port  1: id N2 Online      F-Port 50:05:07:68:01:10:03:64 --> SVC node 2 port 1
port  2: id N2 Online      F-Port 50:05:07:68:01:10:03:5a --> SVC node 1 port 1
port  3: id N2 Online      F-Port 50:05:07:68:01:30:03:64 --> SVC node 2 port 3
port  4: id N2 Online      F-Port 50:05:07:68:01:30:03:5a --> SVC node 1 port 3
port  5: id N2 Online      F-Port 20:08:00:a0:b8:0f:bd:f1 --> FastT 600
port  6: id N2 No_Light
port  7: id N2 No_Light
port  8: id N1 Online      F-Port 10:00:00:00:c9:29:5a:9a --> AIX1
port  9: id N2 Online      F-Port 21:00:00:e0:8b:09:15:41 --> Linux RedHat1
port 10: id N1 Online      F-Port 21:00:00:e0:8b:04:d7:51 --> Win2k_1
port 11: id N2 No_Light
port 12: -- N2 No_Module
port 13: -- N2 No_Module
port 14: -- N2 No_Module
port 15: -- N2 No_Module

2109_Switch_119:admin> cfgshow
Defined configuration:
 cfg:   RedstoneConfig
                zone1; zone2; zone3; zone4; zone5; zone6; zone7; zone8; zone9; zone10;
zone11
 zone:  zone1   1,0; 1,1; 1,2; 1,5; 1,3; 1,4
 zone:  zone10  1,0; 1,1; 1,2; 1,10
 zone:  zone11  1,0; 1,3; 1,4; 1,9
 zone:  zone2   1,0; 1,3; 1,4; 1,6
 zone:  zone3   1,0; 1,1; 1,2; 1,7
 zone:  zone4   1,0; 1,3; 1,4; 1,8
 zone:  zone5   1,0; 1,3; 1,4; 1,15
 zone:  zone6   1,0; 1,1; 1,2; 1,14
 zone:  zone7   1,0; 1,3; 1,4; 1,13
 zone:  zone8   1,0; 1,1; 1,2; 1,12
 zone:  zone9   1,0; 1,3; 1,4; 1,11
```

Example 7-2 shows the IBM 2109 switch number two configuration.

Example 7-2 Switch 118 configuration SWITCH _2

```
IBM_2109_118:admin> switchshow
switchName:     IBM_2109_118
switchType:     9.2
switchState:    Online
switchMode:     Native
switchRole:     Principal
switchDomain:   1
```

```
switchId:       fffc01
switchWwn:      10:00:00:60:69:51:87:2f
switchBeacon:   OFF
Zoning:         ON (RedstoneConfig)
port  0: id N2 Online      F-Port 21:00:00:e0:8b:08:af:d6 --> Master console
port  1: id N2 Online      F-Port 50:05:07:68:01:20:03:64 --> svc node 2 port 2
port  2: id N2 Online      F-Port 50:05:07:68:01:20:03:5a --> SVC node 1 port 2
port  3: id N2 Online      F-Port 50:05:07:68:01:40:03:64 --> SVC node 2 port 4
port  4: id N2 Online      F-Port 50:05:07:68:01:40:03:5a --> SVC node 1 port 4
port  5: id N2 Online      F-Port 20:09:00:a0:b8:0f:bd:f2 --> FastT 600
port  6: id N2 No_Light
port  7: id N1 Online      F-Port 10:00:00:00:c9:25:84:1d --> AIX2
port  8: id N1 Online      F-Port 10:00:00:00:c9:26:6f:5b --> AIX1
port  9: id N2 Online      F-Port 21:01:00:e0:8b:29:15:41 --> Linux RedHat1
port 10: id N1 Online      F-Port 21:01:00:e0:8b:24:d7:51 --> Win2K_1
port 11: id N2 No_Light
port 12: -- N2 No_Module
port 13: -- N2 No_Module
port 14: -- N2 No_Module
port 15: -- N2 No_Module


IBM_2109_118:admin> cfgshow
Defined configuration:
 cfg:   RedstoneConfig
                zone1; zone2; zone3; zone4; zone5; zone6; zone7; zone8; zone9; zone10;
zone11
 zone:  zone1   1,0; 1,1; 1,2; 1,5; 1,3; 1,4
 zone:  zone10  1,0; 1,1; 1,2; 1,10
 zone:  zone11  1,0; 1,3; 1,4; 1,9
 zone:  zone2   1,0; 1,3; 1,4; 1,6
 zone:  zone3   1,0; 1,1; 1,2; 1,7
 zone:  zone4   1,0; 1,3; 1,4; 1,8
 zone:  zone5   1,0; 1,3; 1,4; 1,15
 zone:  zone6   1,0; 1,1; 1,2; 1,14
 zone:  zone7   1,0; 1,3; 1,4; 1,13
 zone:  zone8   1,0; 1,1; 1,2; 1,12
 zone:  zone9   1,0; 1,3; 1,4; 1,11
```

Even if there are 16 possible paths (two server ports X 8 SVC node ports), only four paths exist because of the switch zoning. A zone consists in one port of the server and one port of each node.

Figure 7-2 shows an example for the *zone10* used with the host system “Win2k_1”. There are two paths from the FC Host adapter 1 to port1 of SVC node1, and port 1 of SVC node 2. Two other paths are from the FC Host adapter 2 to port2 of SVC node1, and port 2 of SVC node 2. That gives the host system “Win2k_1” a maximum of four paths.

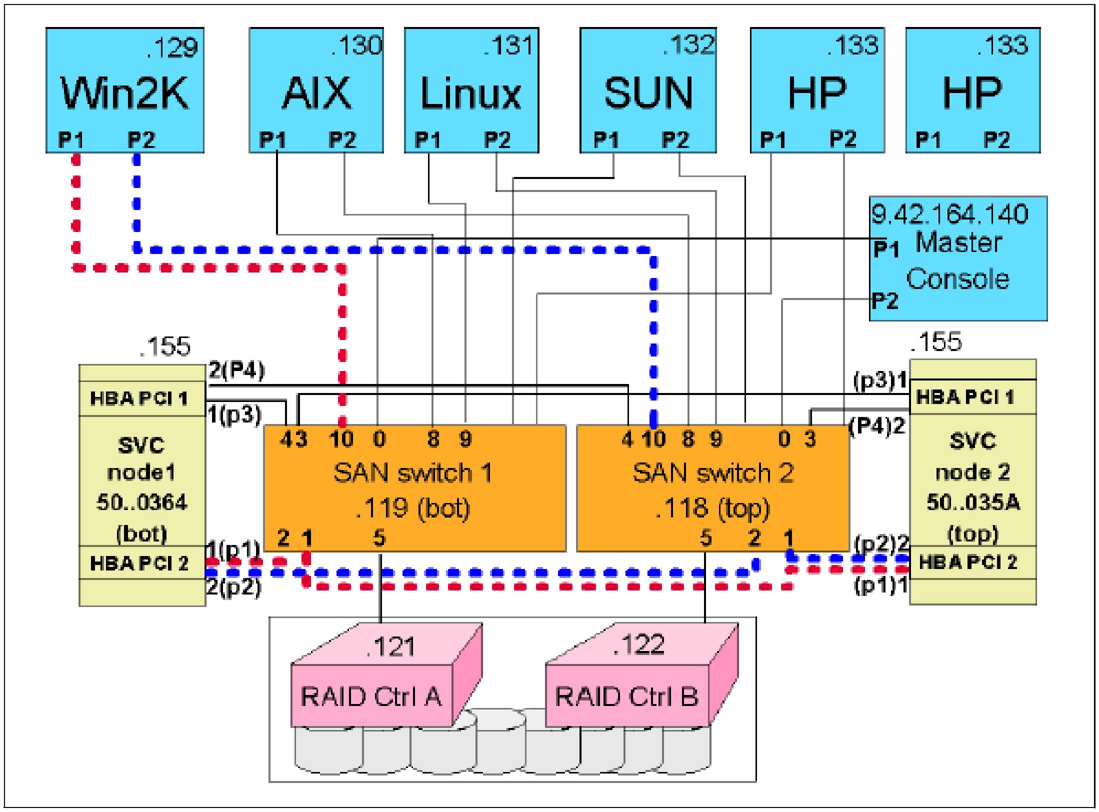


Figure 7-2 Zoning for zone10 in switch 1 (red) and switch 2 (blue)

7.4 AIX-specific information

The following section details specific information that pertains to the connection of AIX-based hosts to the SVC environment.

7.4.1 Configuring the AIX host

To configure the AIX host, follow these steps:

1. Install the HBA or HBAs on your AIX host system.
2. Install and configure the 2145 and SDD drivers.
3. Configure the switches (zoning) if needed.
4. Connect your AIX host system to the switches.
5. Reboot your AIX host system.
6. Configure your host, mDisks, and host mapping in the SAN Volume Controller.
7. Run the **cfgmgr** command to discover the mDisks created on SVC.

7.4.2 Support information

This section details the current support information. You must check the Web site that are listed for the latest details.

Operating system versions and maintenance levels

At GA, the versions of AIX listed in Table 7-1 were supported. See the following Web site:

http://www.storage.ibm.com/support/2145

Table 7-1 Versions of AIX supported with SVC at GA

Operating system level	Machine level
AIX 4.3.3	Recommended Maintenance Level 11
AIX 5.1	Recommended Maintenance Level 3
AIX 5.2	

SDD package version

At GA, the following version of SDD for AIX is supported. See the following Web site for the latest information:

http://www.storage.ibm.com/support/2145

- SDD for AIX version 1.4.0.1
- SDD for HACMP (non-concurrent)
- SDD for HACMP (concurrent)

Example 7-3 shows checking the installation of device driver for San Volume Controller 2145 and SDD.

Example 7-3 Checking of 2145 SAN Volume Controller and SDD device driver

```
# lslpp -l *sdd*
  Fileset                      Level  State      Description
  ----------------------------------------------------------------------------
Path: /usr/lib/objrepos
  devices.sdd.51.rte         1.4.0.0  COMMITTED  IBM Subsystem Device Driver
                                                 for AIX V51

Path: /etc/objrepos
  devices.sdd.51.rte         1.4.0.0  COMMITTED  IBM Subsystem Device Driver
                                                 for AIX V51

# lslpp -l *2145*
  Fileset                      Level  State      Description
  ----------------------------------------------------------------------------
Path: /usr/lib/objrepos
  ibm2145.rte          4.3.2002.1111  COMMITTED  IBM 2145 TotalStorage SAN
                                                 Volume Controller

Path: /etc/objrepos
  ibm2145.rte          4.3.2002.1111  COMMITTED  IBM 2145 TotalStorage SAN
                                                 Volume Controller
```

You can also check that the SDD server is operational as shown in Example 7-4.

Example 7-4 SDD server is operational

```
# lssrc -s sddsrv
Subsystem         Group            PID     Status
 sddsrv                            5828    active
# ps -eaf |grep sdd
    root  5828  7744   0 09:27:13      -  0:22 /usr/sbin/sddsrv
    root 18666 17554   1 18:00:55  pts/1  0:00 grep sdd
```

Supported host adapters

At GA, the HBAs listed in Table 7-2 are supported for AIX. See the following Web site for the latest information:

http://www.storage.ibm.com/support/2145

Table 7-2 Supported FC Host Adapter for AIX at GA

FC Host Adapter	AIX 4.3.3	AIX 5.1
IBM FC PCI adapter (FC 6227)	D/D: devices.pci. df1000f7.rte 4.3.3.76 Firmware 3.30X1	D/D: devices.pci. df1000f7.rte 5.1.0.35 Firmware 3.30X1
IBM FC 64-bit PCI adapter (FC 6228)	D/D: devices.pci. df1000f9.rte 4.3.3.76 Firmware 3.82A1	D/D: devices.pci. df1000f9.rte 5.1.0.35 Firmware 3.82A1

7.4.3 Host adapter configuration settings

You can check the availability of the FC Host Adapters by using the command shown in Example 7-5.

Example 7-5 FC Host Adapter availability

```
# lsdev -Cc adapter |grep fcs
fcs0    Available 20-58    FC Adapter
fcs1    Available 20-60    FC Adapter
```

You can also find the worldwide port name (WWPN) of your FC Host Adapter and check the firmware level as shown in Example 7-6.

Example 7-6 FC Host Adapter settings and WWPN

```
# lscfg -vl fcs0
  DEVICE            LOCATION          DESCRIPTION
  fcs0              20-58             FC Adapter


        Part Number.................09P4038
        EC Level....................A
        Serial Number...............1A1360381D
        Manufacturer................001A
        FRU Number..................09P4039
```

```
        Network Address.............10000000C9266F5B
        ROS Level and ID............02903331
        Device Specific.(Z0)........4002206D
        Device Specific.(Z1)........10020193
        Device Specific.(Z2)........3001506D
        Device Specific.(Z3)........03000909
        Device Specific.(Z4)........FF101493
        Device Specific.(Z5)........02903331
        Device Specific.(Z6)........06113331
        Device Specific.(Z7)........07113331
        Device Specific.(Z8)........20000000C9266F5B
        Device Specific.(Z9)........SS3.30X1
        Device Specific.(ZA)........S1F3.30X1
        Device Specific.(ZB)........S2F3.30X1
        Device Specific.(YL)........P2-I1/Q1


# lscfg -vl fcs1

  DEVICE            LOCATION          DESCRIPTION
  fcs1              20-60             FC Adapter

        Part Number.................09P4038
        EC Level....................A
        Serial Number...............1A12500B4A
        Manufacturer................001A
        FRU Number..................09P4039
        Network Address.............10000000C9295A9A
        ROS Level and ID............02903331
        Device Specific.(Z0)........4002206D
        Device Specific.(Z1)........10020193
        Device Specific.(Z2)........3001506D
        Device Specific.(Z3)........03000909
        Device Specific.(Z4)........FF101493
        Device Specific.(Z5)........02903331
        Device Specific.(Z6)........06113331
        Device Specific.(Z7)........07113331
        Device Specific.(Z8)........20000000C9295A9A
        Device Specific.(Z9)........SS3.30X1
        Device Specific.(ZA)........S1F3.30X1
        Device Specific.(ZB)........S2F3.30X1
        Device Specific.(YL)........P2-I2/Q1
```

7.4.4 Discovering the assigned vDisk

Before you add a new volume from the SAN Volume Controller, the AIX host system “aix1” had the configuration as shown in Example 7-7.

Example 7-7 Status of AIX host system ‘aix1’

```
# lspv
hdisk0          0007905873b5c2fb                    rootvg
hdisk1          0007905873b5c364                    rootvg

# lsvg
rootvg
```

The configuration of the host aix1, the vDisks aix_vDisk1 to aix_vDisk4, and the mapping between the host and the vDisks are defined in the SAN Volume Controller as shown in Example 7-8.

You can check that the WWPN as listed in Example 7-6 are logged into the SAN Volume Controller for the host "aix1" by entering:

```
svcinfo lshost aix1
```

You can also find the serial numbers of the vDisks with the following command:

```
svcinfo lshostvdiskmap
```

Then verify them by entering the following command as shown in Example 7-9:

```
lsattr -El vpath0
```

Example 7-8 SVC definitions for host system 'aix1'

```
IBM_2145:admin>svcinfo lshost aix1
id 0
name aix1
port_count 2
WWPN 10000000C9295A9A
port_logged_in_count 2
WWPN 10000000C9266F5B
port_logged_in_count 2

IBM_2145:admin>svcinfo lshostvdiskmap
id/name/SCSI_id/vDisk_id/vDisk_name / wwpn / vDisk_UID
0/aix1/ 0   / 0   / aix_vDisk0 / 10000000C9295A9A / 600507680182001B2000000000000000
0/aix1 / 1 / 1    / aix_vDisk1 / 10000000C9295A9A / 600507680182001B2000000000000001
0/ aix1/ 2  / 2   / aix_vDisk2 / 10000000C9295A9A / 600507680182001B2000000000000002
0/ aix1/ 3  / 3   / aix_vDisk3 / 10000000C9295A9A / 600507680182001B2000000000000003
0/ aix1/ 0  / 0   / aix_vDisk0 / 10000000C9266F5B / 600507680182001B2000000000000000
0/ aix1/ 1  / 1   / aix_vDisk1 / 10000000C9266F5B / 600507680182001B2000000000000001
0/ aix1/ 2  / 2   / aix_vDisk2 / 10000000C9266F5B / 600507680182001B2000000000000001
0/ aix1/ 3  / 3   / aix_vDisk3 / 10000000C9266F5B / 600507680182001B2000000000000003

IBM_2145:admin>svcinfo lsvdisk aix_vDisk0
id 0
name aix_vDisk0
IO_group_id 0
IO_group_name ITSO_grp0
status online
mDisk_grp_id 0
mDisk_grp_name ARR36P5N
capacity 10.0GB
type striped
formatted no
mDisk_id
mDisk_name
FC_id
FC_name
RC_id
RC_name
throttling 0
preferred_node_id 1
```

```
IBM_2145:admin>svcinfo lsvdiskhostmap aix_vDisk0
id  / name / SCSI_id / host_id / host_name / wwpn          / vDisk_UID
0  / aix_vDisk0 / 0  / 0 / aix1 / 10000000C9295A9A /600507680182001B2000000000000000
0  / aix_vDisk0 / 0  / 0 / aix1 / 10000000C9266F5B /600507680182001B2000000000000000
```

You have to run **cfgmgr** one time per adapter:

```
cfgmgr -l fcs0
cfgmgr -l fcs1
```

The disk configuration of the AIX host system appears now as shown in Example 7-9. The vDisk “aix_vDisk0” is seen as vpath0. The vpath0 is made of hdisk2, hdisk6, hdisk10, and hdisk14. They correspond to the four paths between the host “aix1” and the mDisk “aix_vDisk0”. Even if there are 16 possible paths (two server ports X eight SVC node ports), only four exist because of the switch zoning. A zone consists in one port of the server and one port of each node.

You can check the serial number of the vDisks by entering the following command (see Example 7-9):

```
lsattr -El vpath0
```

Example 7-9 Four vDisks from SVC added with four different paths for each vDisk

```
# lsdev -Cc disk
hdisk0  Available 10-60-00-8,0 16 Bit SCSI Disk Drive
hdisk1  Available 10-60-00-9,0 16 Bit SCSI Disk Drive
hdisk2  Available 20-58-01     IBM TotalStorage SAN Volume Controller device
hdisk3  Available 20-58-01     IBM TotalStorage SAN Volume Controller device
hdisk4  Available 20-58-01     IBM TotalStorage SAN Volume Controller device
hdisk5  Available 20-58-01     IBM TotalStorage SAN Volume Controller device
hdisk6  Available 20-58-01     IBM TotalStorage SAN Volume Controller device
hdisk7  Available 20-58-01     IBM TotalStorage SAN Volume Controller device
hdisk8  Available 20-58-01     IBM TotalStorage SAN Volume Controller device
hdisk9  Available 20-58-01     IBM TotalStorage SAN Volume Controller device
hdisk10 Available 20-60-01     IBM TotalStorage SAN Volume Controller device
hdisk11 Available 20-60-01     IBM TotalStorage SAN Volume Controller device
hdisk12 Available 20-60-01     IBM TotalStorage SAN Volume Controller device
hdisk13 Available 20-60-01     IBM TotalStorage SAN Volume Controller device
hdisk14 Available 20-60-01     IBM TotalStorage SAN Volume Controller device
hdisk15 Available 20-60-01     IBM TotalStorage SAN Volume Controller device
hdisk16 Available 20-60-01     IBM TotalStorage SAN Volume Controller device
hdisk17 Available 20-60-01     IBM TotalStorage SAN Volume Controller device
vpath0  Available              Data Path Optimizer Pseudo Device Driver
vpath1  Available              Data Path Optimizer Pseudo Device Driver
vpath2  Available              Data Path Optimizer Pseudo Device Driver
vpath3  Available              Data Path Optimizer Pseudo Device Driver

# lspv
hdisk0          0007905873b5c2fb                    rootvg
hdisk1          0007905873b5c364                    rootvg
hdisk2          none                                None
hdisk3          none                                None
hdisk4          none                                None
hdisk5          none                                None
hdisk6          none                                None
hdisk7          none                                None
hdisk8          none                                None
hdisk9          none                                None
hdisk10         none                                None
```

```
hdisk11         none                                None
hdisk12         none                                None
hdisk13         none                                None
hdisk14         none                                None
hdisk15         none                                None
hdisk16         none                                None
hdisk17         none                                None
vpath0          000790582dd7c20c                    itsoaix1vg
vpath1          none                                None
vpath2          none                                None
vpath3          none                                None

# lsattr -El vpath0
pvid          000790582dd7c20c0000000000000000         Data Path Optimizer Parent False
policy        df                                       Scheduling Policy          True
active_hdisk  hdisk2/600507680182001B2000000000000000  Active hdisk               False
active_hdisk  hdisk6/600507680182001B2000000000000000
active_hdisk  hdisk10/600507680182001B2000000000000000
active_hdisk  hdisk14/600507680182001B2000000000000000
serial_number 600507680182001B2000000000000000         N/A                        False

# lsattr -El vpath1
pvid          none                                     Data Path Optimizer Parent False
policy        df                                       Scheduling Policy          True
active_hdisk  hdisk3/600507680182001B2000000000000001  Active hdisk               False
active_hdisk  hdisk7/600507680182001B2000000000000001
active_hdisk  hdisk11/600507680182001B2000000000000001
active_hdisk  hdisk15/600507680182001B2000000000000001
serial_number 600507680182001B2000000000000001         N/A                        False

# lsattr -El vpath2
pvid          none                                     Data Path Optimizer Parent False
policy        df                                       Scheduling Policy          True
active_hdisk  hdisk4/600507680182001B2000000000000002  Active hdisk               False
active_hdisk  hdisk8/600507680182001B2000000000000002
active_hdisk  hdisk12/600507680182001B2000000000000002
active_hdisk  hdisk16/600507680182001B2000000000000002
serial_number 600507680182001B2000000000000002         N/A                        False

# lsattr -El vpath3
pvid          none                                     Data Path Optimizer Parent False
policy        df                                       Scheduling Policy          True
active_hdisk  hdisk5/600507680182001B2000000000000003  Active hdisk               False
active_hdisk  hdisk9/600507680182001B2000000000000003
active_hdisk  hdisk13/600507680182001B2000000000000003
active_hdisk  hdisk17/600507680182001B2000000000000003
serial_number 600507680182001B2000000000000003         N/A                        False

# lspv vpath0
PHYSICAL VOLUME:    vpath0                   VOLUME GROUP:     itsoaix1vg
PV IDENTIFIER:      000790582dd7c20c VG IDENTIFIER     0007905800004c00000000f52dd7d38a
PV STATE:           active
STALE PARTITIONS:   0                        ALLOCATABLE:      yes
PP SIZE:            16 megabyte(s)           LOGICAL VOLUMES:  2
TOTAL PPs:          639 (10224 megabytes)    VG DESCRIPTORS:   2
FREE PPs:           576 (9216 megabytes)     HOT SPARE:        no
USED PPs:           63 (1008 megabytes)
FREE DISTRIBUTION:  128..65..127..128..128
USED DISTRIBUTION:  00..63..00..00..00
```

7.4.5 Using SDD

You can also check that device 3 (vpath0/aix_vDisk1) is in the “open” state because it is used by a volume group. Only two paths of the four paths have activity. They are the two paths connected the preferred node of the input/output (I/O) group of the SVC cluster. The two other paths are used as alternate paths in case of a failure of the preferred path (see Example 7-10).

Example 7-10 SDD commands used to check the availability of the devices

```
# datapath query adapter
Active Adapters :2
Adpt#     Adapter Name   State     Mode     Select     Errors  Paths  Active
    0           fscsi1  NORMAL   ACTIVE      24638          0      8       2
    1           fscsi0  NORMAL   ACTIVE      24655          0      8       2
Total Devices : 4

# datapath query device
DEV#:   0  DEVICE NAME: vpath1  TYPE: 2145           POLICY:    Optimized
SERIAL: 600507680182001B2000000000000001
==========================================================================
Path#      Adapter/Hard Disk          State     Mode     Select     Errors
    0          fscsi1/hdisk3          CLOSE   NORMAL          0          0
    1          fscsi1/hdisk7          CLOSE   NORMAL          0          0
    2         fscsi0/hdisk11          CLOSE   NORMAL          0          0
    3         fscsi0/hdisk15          CLOSE   NORMAL          0          0

DEV#:   1  DEVICE NAME: vpath2  TYPE: 2145           POLICY:    Optimized
SERIAL: 600507680182001B2000000000000002
==========================================================================
Path#      Adapter/Hard Disk          State     Mode     Select     Errors
    0          fscsi1/hdisk4          CLOSE   NORMAL          0          0
    1          fscsi1/hdisk8          CLOSE   NORMAL          0          0
    2         fscsi0/hdisk12          CLOSE   NORMAL          0          0
    3         fscsi0/hdisk16          CLOSE   NORMAL          0          0

DEV#:   2  DEVICE NAME: vpath3  TYPE: 2145           POLICY:    Optimized
SERIAL: 600507680182001B2000000000000003
==========================================================================
Path#      Adapter/Hard Disk          State     Mode     Select     Errors
    0          fscsi1/hdisk5          CLOSE   NORMAL          0          0
    1          fscsi1/hdisk9          CLOSE   NORMAL          0          0
    2         fscsi0/hdisk13          CLOSE   NORMAL          0          0
    3         fscsi0/hdisk17          CLOSE   NORMAL          0          0

DEV#:   3  DEVICE NAME: vpath0  TYPE: 2145           POLICY:    Optimized
SERIAL: 600507680182001B2000000000000000
==========================================================================
Path#      Adapter/Hard Disk          State     Mode     Select     Errors
    0          fscsi1/hdisk2           OPEN   NORMAL      24637          0
    1          fscsi1/hdisk6           OPEN   NORMAL          0          0
    2         fscsi0/hdisk10           OPEN   NORMAL      24655          0
    3         fscsi0/hdisk14           OPEN   NORMAL          0          0

# lsvpcfg
vpath0 (Avail pv itsoaix1vg) 600507680182001B2000000000000000 = hdisk2 (Avail ) hdisk6
(Avail ) hdisk10 (Avail ) hdisk14 (Avail )
```

```
vpath1 (Avail ) 600507680182001B2000000000000001 = hdisk3 (Avail ) hdisk7 (Avail ) hdisk11
(Avail ) hdisk15 (Avail )
vpath2 (Avail ) 600507680182001B2000000000000002 = hdisk4 (Avail ) hdisk8 (Avail ) hdisk12
(Avail ) hdisk16 (Avail )
vpath3 (Avail ) 600507680182001B2000000000000003 = hdisk5 (Avail ) hdisk9 (Avail ) hdisk13
(Avail ) hdisk17 (Avail )
```

7.4.6 Creating and preparing volumes for use

A volume group *itsoaix1vg* is created on vpath0 (aix_vDisk1), and the file system *itsofs1* is mounted on *itsoaix1vg*. See Example 7-11.

Example 7-11 'aix1' host system new volume group and file system configuration

```
# lsvg -o
itsoaix1vg
rootvg

# lsvg -l itsoaix1vg
itsoaix1vg:
LV NAME             TYPE       LPs    PPs    PVs  LV STATE      MOUNT POINT
loglv00             jfslog     1      1      1    open/syncd    N/A
lv00                jfs        62     62     1    open/syncd    /itsofs1

# df -k
Filesystem    1024-blocks      Free %Used    Iused %Iused Mounted on
/dev/hd4            16384      7960   52%     1306    16% /
/dev/hd2          1867776   1265440   33%    20505     5% /usr
/dev/hd9var         16384     11020   33%      433    11% /var
/dev/hd3            32768     29196   11%       57     1% /tmp
/dev/hd1            16384     15820    4%       18     1% /home
/proc                   -         -    -         -     -  /proc
/dev/hd10opt        32768     26592   19%      294     4% /opt
/dev/lv00         1015808    403224   61%       19     1% /itsofs1
```

7.4.7 Running a command line (CLI) on an AIX host system

To issue CLI commands you must install and prepare the SSH client system on the AIX host system. For AIX 5L™ Power 5.1 and 5.2, you can get OpenSSH from the Bonus Packs. You also need its prerequisite, OpenSSL, from the AIX toolbox for Linux applications for Power Systems. For AIX 4.3.3, the software is available from the AIX toolbox for Linux applications.

The AIX installation images from IBM developerWorks® are available at the following Web site:

http://oss.software.ibm.com/developerworks/projects/openssh

Refer to Chapter 4 or Chapter 5 "Basic Installation" for the SSH client system installation and preparation.

7.5 Windows NT and 2000 specific information

The following sections detail specific information about the connection of Windows 2000 based hosts to the SVC environment.

7.5.1 Configuring the Windows 2000 host

Follow these steps:

1. Install the HBA or HBAs on your Windows 2000 server.
2. Install and configure SDD.
3. Shut down your Windows 2000 host system.
4. Configure the switches (zoning) if needed.
5. Connect your Windows 2000 server FC Host adapters to the switches.
6. Restart your Windows 2000 host system.
7. Configure your host, mDisks and host mapping in the SAN Volume Controller.
8. Use Rescan disk in Computer Management of your Windows 2000 server to discover the mDisks created on the SAN Volume Controller.

7.5.2 Support information

This section tells where you can obtain various types of support information.

Operating system versions and maintenance levels

At the time of this writing, the versions of Windows listed in Table 7-3 are supported. See the following Web site for the latest information:

`http://www.storage.ibm.com/support/2145`

Table 7-3 Versions of Windows supported with SVC at the time of this writing

Operating system level	Machine level
Windows 2000 Server and Advanced Server	Service Pack 3
Windows NT4 Enterprise Server	Service Pack 6a

Supported host adapters

See the supported hardware list for the latest information about the supported HBAs and driver levels for Windows:

`http://www.storage.ibm.com/support/2145`

7.5.3 Host adapter installation and configuration

Refer to manufacturer's instructions for installation and configuration of the HBAs. You can check that the FC adapter is installed as shown in Figure 7-3.

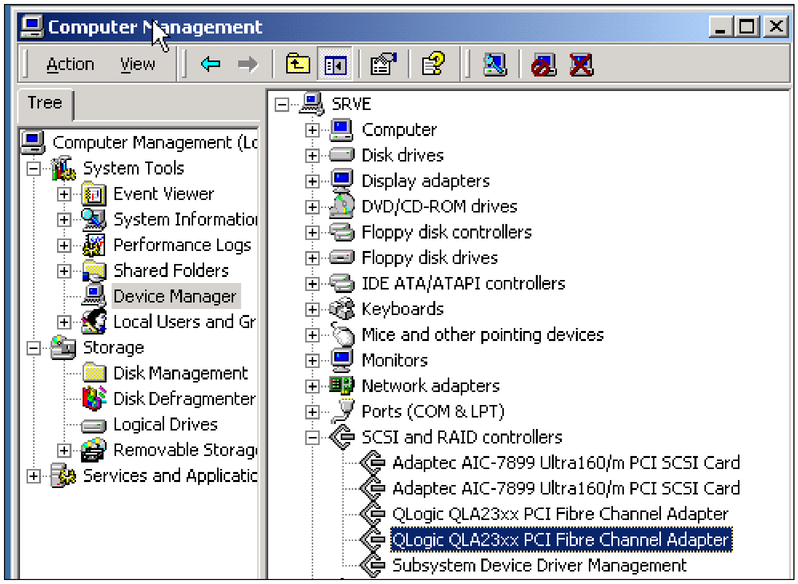


Figure 7-3 QLogic FC Host Adapter

7.5.4 SDD installation

At the time of this writing, Version 1.4.0.1 for Windows NT 4 and 1.4.0.2 for Windows 2000 of SDD are supported. See the following Web site for the latest information about SDD for Windows:

`http://www.storage.ibm.com/support/2145`

After you download the appropriate version of the SDD from this Web site, run setup to install SDD. Answer *Yes* to Windows Digital Signature prompt. Answer *Yes* to reboot the system.

You can check that the installation of SDD is complete. From your Windows desktop, click **Start-> Programs-> Subsystem Device Driver-> readme**.

7.5.5 Discovering assigned vDisk

Before you add a new volume from the SAN Volume Controller, the Windows 2000 host system “Win2K_1” had the configuration as shown in Figure 7-4, with only local disks.

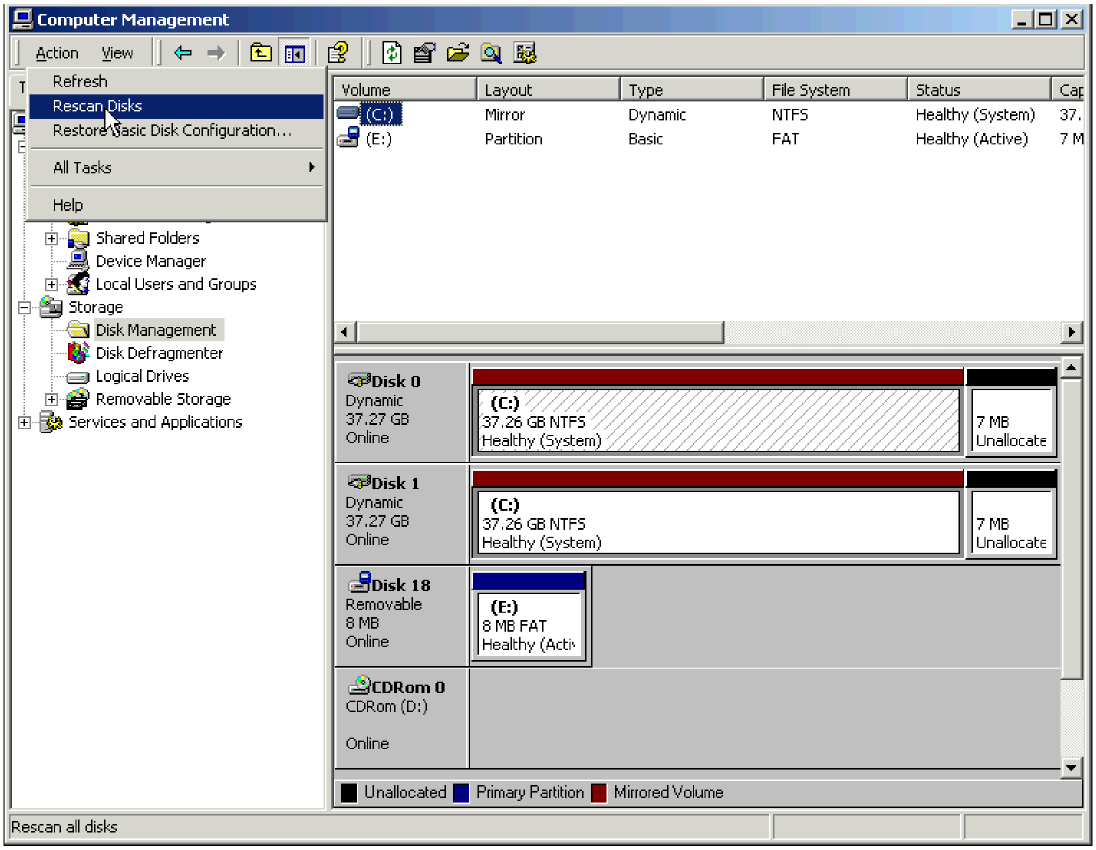


Figure 7-4 Windows 2000 host system before adding a new volume from SVC

The configuration of the host “Win2K_1”, the vDisk “windisk1”, and the mapping between the host and the vDisk are defined in the SAN Volume Controller as described in Example 7-12.

You can check that the WWPN is logged into the SAN Volume Controller for the host “Win2K_1” by entering the following command:

```
svcinfo lshost Win2K_1
```

You can also find the serial number of the vDisks by entering the following command:

```
svcinfo lshostvdiskmap
```

Example 7-12 SVC configuration for Windows 2000 host system

```
IBM_2145:admin>svcinfo lshost Win2K_1
id 1
name Win2K_1
port_count 2
WWPN 210000E08B04D751
port_logged_in_count 2
WWPN 210100E08B24D751
port_logged_in_count 2
```

```
IBM_2145:admin>svcinfo lshostvdiskmap Win2K_1
id / name    / SCSI_id  / vDisk_id/ vDisk_name  / wwpn  / vDisk_UID
1 / Win2K_1  / 0 / 4      / windisk1 / 210000E08B04D751 /600507680182001B2000000000000004
1 / Win2K_1  / 0 / 4      / windisk1 / 210100E08B24D751 /600507680182001B2000000000000004

IBM_2145:admin>svcinfo lsvdisk windisk1
id 4
name windisk1
IO_group_id 0
IO_group_name ITSO_grp0
status online
mDisk_grp_id 0
mDisk_grp_name ARR36P5N
capacity 200.0MB
type striped
formatted no
mDisk_id
mDisk_name
FC_id
FC_name
RC_id
RC_name
throttling 0
preferred_node_id 3

IBM_2145:admin>svcinfo lsvdiskhostmap windisk1
id / name     / SCSI_id / host_id / host_name       / wwpn         / vDisk_UID
4 / windisk1  / 0   / 1   / Win2K_1   / 210000E08B04D751 /600507680182001B2000000000000004
4 / windisk1  / 0   / 1   / Win2K_1   / 210100E08B24D751 /600507680182001B2000000000000004
```

After the rescan disks is complete from the Computer Management window, the new volume is found and assigned to the X: letter as shown in Figure 7-5.

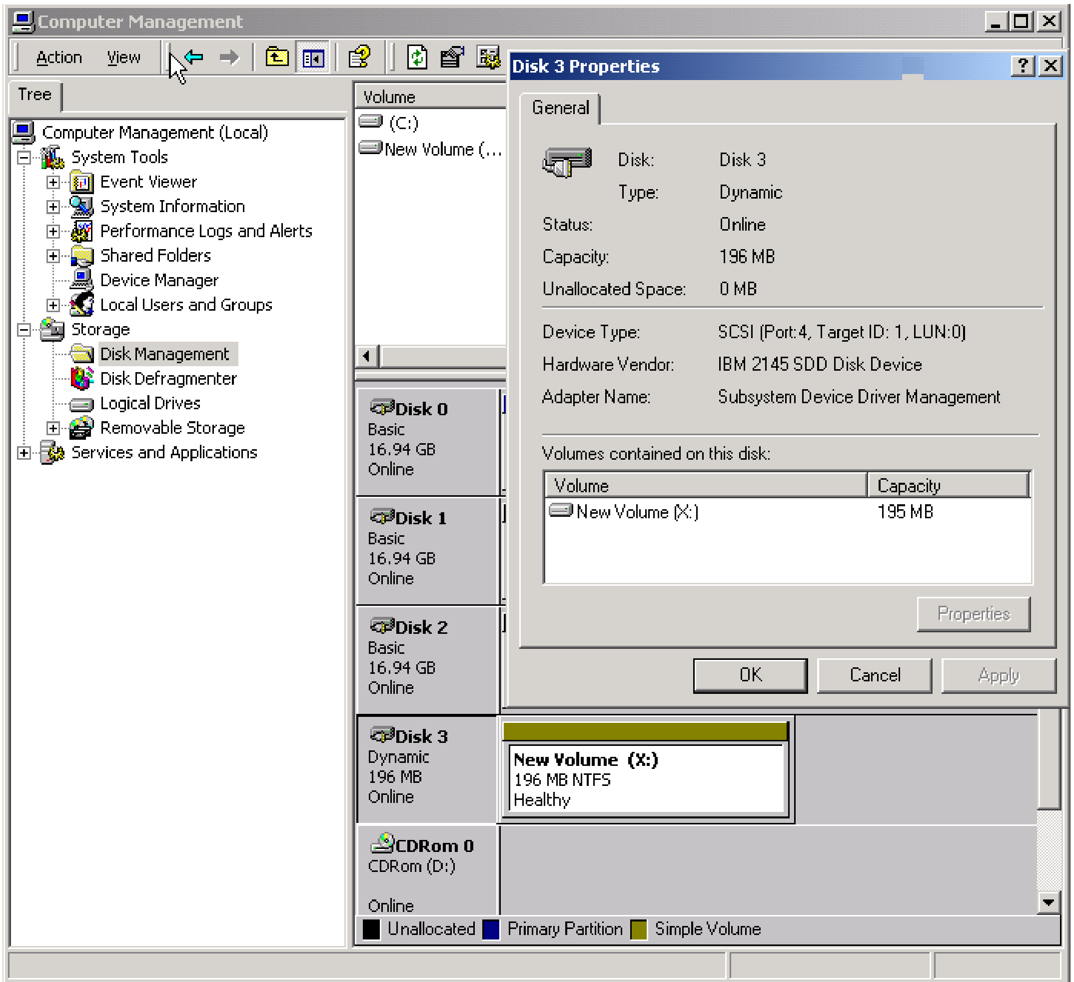


Figure 7-5 Windows 2000 host system with a new volume from SVC

The volume is identified as an IBM 2145 SCSI Disk Device. The number of IBM 2145 SCSI Disk Devices that you see is equal to:

```
(# of vDisks) x (# of Nodes) x (4 Ports per Node) x (# of HBAs)
```

This is shown in Figure 7-6. This corresponds to the number of paths between the Windows 2000 host system and the SAN Volume Controller. However, you see one IBM 2145 SDD Disk Device per vDisk.

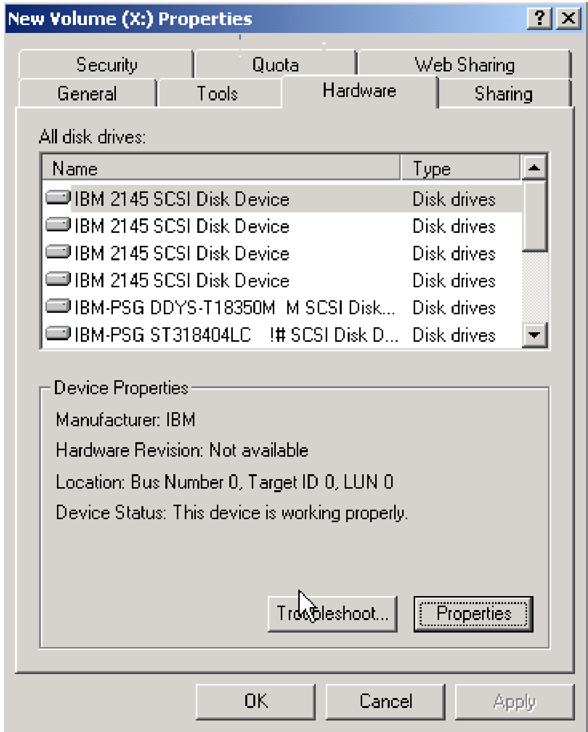
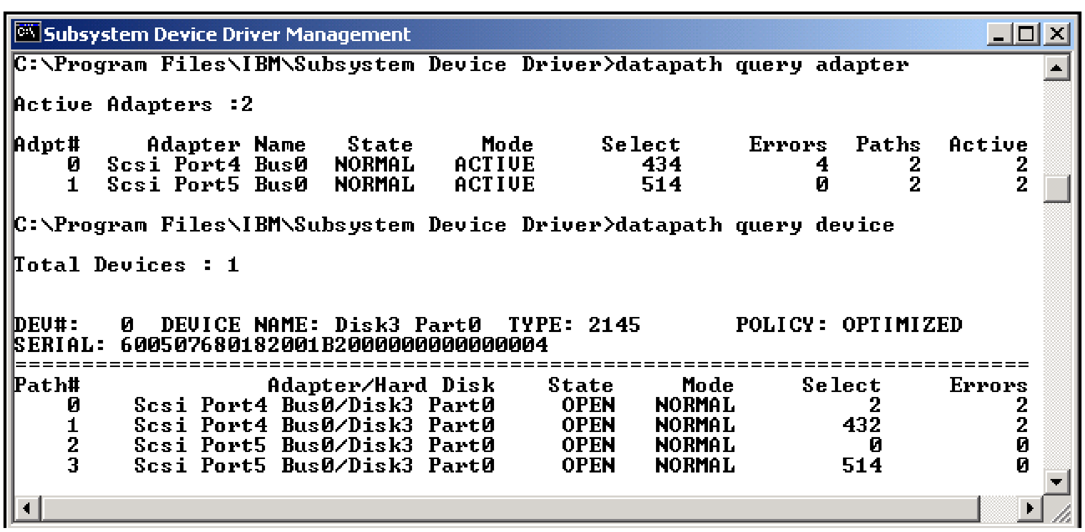


Figure 7-6 Number of devices found related to the number of paths

7.5.6 Using SDD

To open a command window for SDD, from your desktop, click **Start-> Programs-> Subsystem Device Driver-> Subsystem Device Driver Management**.

You can use the SDD-specific commands explained in *Subsystem Device Driver User's Guide for the IBM TotalStorage Enterprise Storage Server and the IBM TotalStorage SAN Volume Controller*, SC26-7540, as shown in Figure 7-7.



```
Subsystem Device Driver Management
C:\Program Files\IBM\Subsystem Device Driver>datapath query adapter

Active Adapters :2

Adpt#     Adapter Name  State     Mode      Select     Errors  Paths  Active
    0   Scsi Port4 Bus0  NORMAL   ACTIVE       434          4      2       2
    1   Scsi Port5 Bus0  NORMAL   ACTIVE       514          0      2       2

C:\Program Files\IBM\Subsystem Device Driver>datapath query device

Total Devices : 1


DEV#:   0  DEVICE NAME: Disk3 Part0  TYPE: 2145       POLICY: OPTIMIZED
SERIAL: 600507680182001B2000000000000004
============================================================================
Path#              Adapter/Hard Disk      State     Mode      Select     Errors
    0     Scsi Port4 Bus0/Disk3 Part0     OPEN    NORMAL          2          2
    1     Scsi Port4 Bus0/Disk3 Part0     OPEN    NORMAL        432          2
    2     Scsi Port5 Bus0/Disk3 Part0     OPEN    NORMAL          0          0
    3     Scsi Port5 Bus0/Disk3 Part0     OPEN    NORMAL        514          0
```

Figure 7-7 Datapath query commands

Or you can open your browser with the host IP address as shown in Figure 7-8.

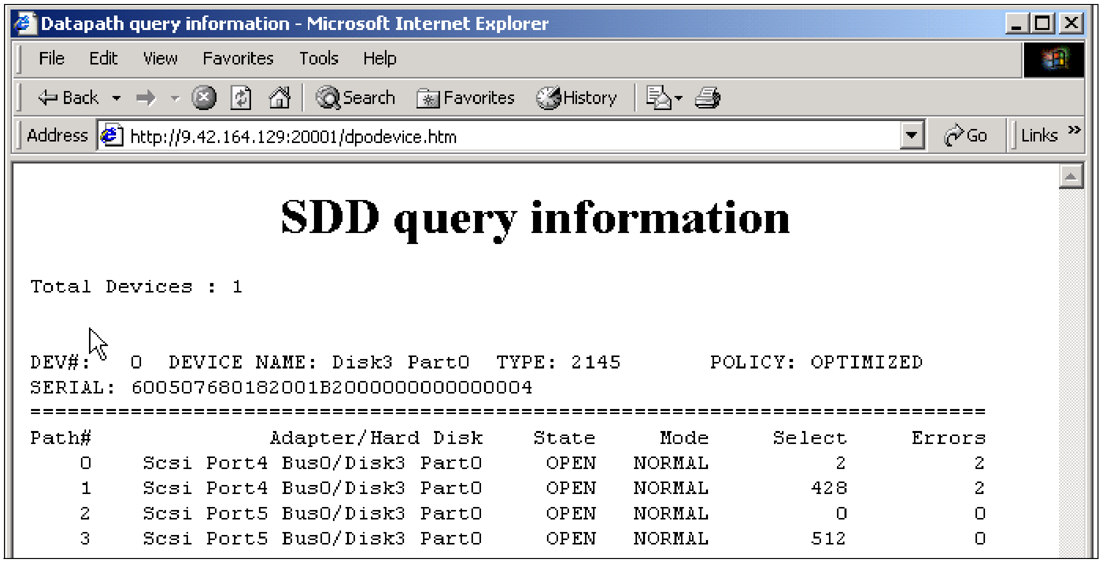


Figure 7-8 SDD query information using Web browser at <Win2k_1 ip add>:20001

7.5.7 Running a command line (CLI) on a Windows host system

To issue CLI commands, you must install and prepare the SSH client system on the Windows host system.

You can install the PuTTY SSH client software on a Windows host using the PuTTY Installation program putty-0.53b-installer.exe. This is in the SSHClient\PuTTY directory of the SAN Volume Controller Console CD-ROM. Or, you can download PuTTY from the following Web site:

http://www.chiark.greenend.org.uk/~sgtatham/putty/

The following Web site offers SSH client alternatives for Windows:

http://www.openssh.com/windows.html

Cygwin software has an option to install an OpenSSH client. You can download Cygwin from the following Web site:

http://www.cygwin.com/

Refer to Chapter 4 or Chapter 5 “Basic Installation” for the SSH client system installation and preparation.

7.6 Linux (on Intel) specific information

The following section details specific information pertaining to the connection of Linux on Intel-based hosts to the SVC environment.

7.6.1 Configuring the Linux host

Follow these steps to configure the Linux host:

1. Install the HBA or HBAs on your Linux server.
2. Install Kernel 2.4.9e16.

3. Configure the switches (zoning) if needed.
4. Connect your Linux server FC Host adapters to the switches.
5. Install SDD for Linux.
6. Configure your host, vDisks, and host mapping in the SAN Volume Controller.
7. Reboot your Linux server to discover the vDisks created on SVC.

7.6.2 Support information

For the following information, consult the IBM TotalStorage SAN Volume Controller (2145) Web site at:

`http://www.storage.ibm.com/support/2145`

- Operating system versions and maintenance levels: At GA, Red Hat Version Advanced Server 2.1 with Kernel 2.4.9e16 was supported.
- SDD: At GA, Version 1.4.0.0-8 of SDD for Linux was supported
- Supported host adapters: At announcement, the QLA2310FL, QLA2340FL, and QLA2342FL adapters were supported.

7.6.3 Host adapter configuration settings

See *IBM TotalStorage Virtualization Family SAN Volume Controller: Host Attachment Guide*, SC26-7563.

7.6.4 Discovering assigned vDisk

The **cat /proc/scsi/scsi** command shows the devices that the SCSI driver has probed. In our configuration, we can see two vDisks per SCSI adapter. See Figure 7-9.

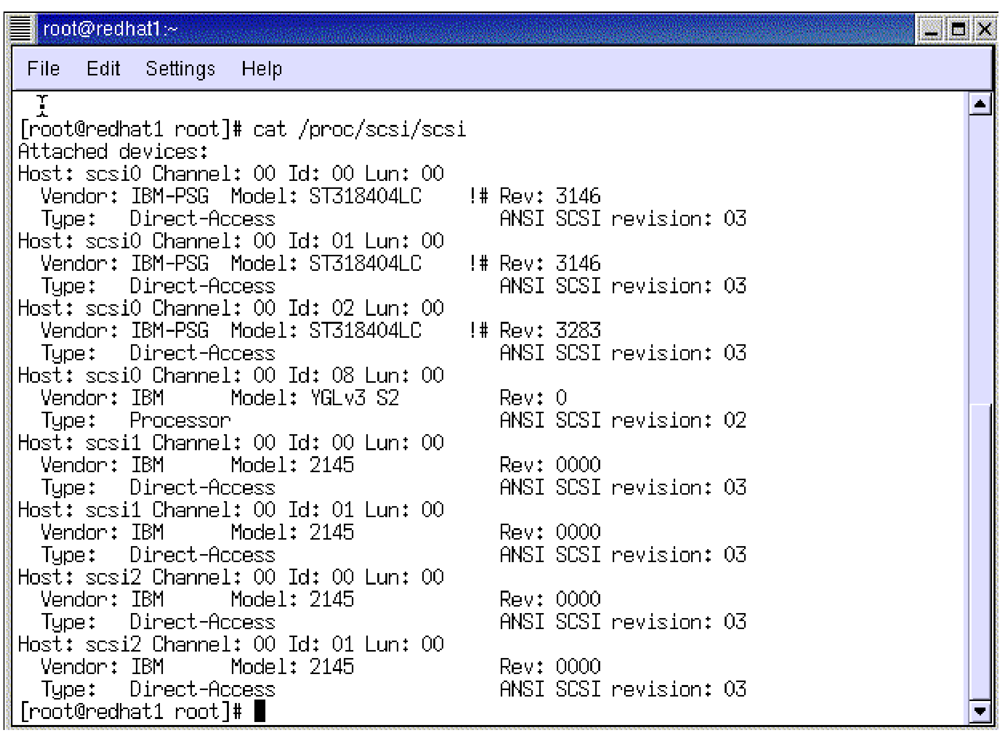


Figure 7-9 proc/scsi/scsi

7.6.5 Using SDD

The **`rpm -iv IBMsdd-1.4.0.0-4.i686.rh7.rpm`** command installs the package. To install the subsystem device driver, make sure the RPM is executable.

Issue the following command (see Figure 7-10):

```
chmod 755 IBMsdd-1.4.0.0-4.i686.rh7.rpm
```

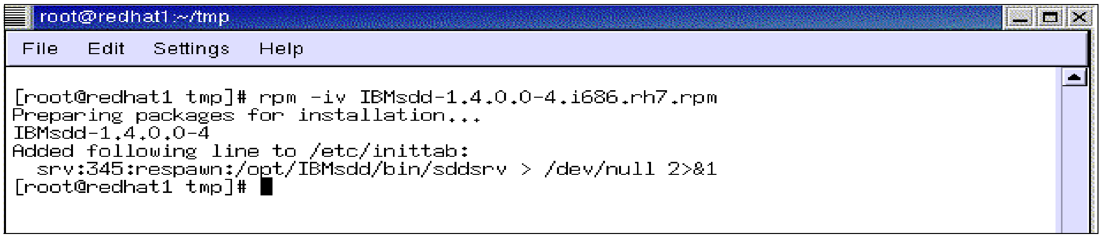


Figure 7-10 install SDD rpm

SDD is installed to the /opt/IBMSDD/bin directory as shown in Figure 7-11.

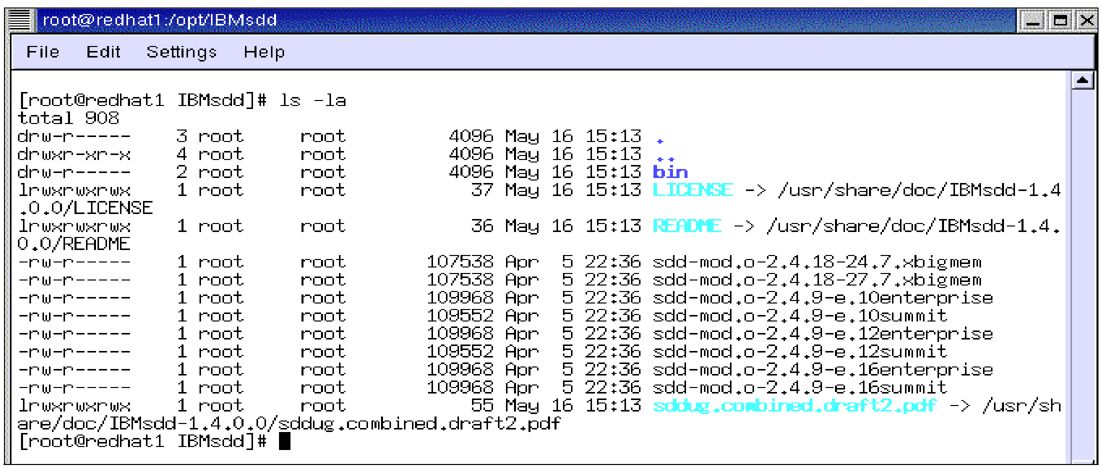


Figure 7-11 opt SDD directory

Configure SDD by matching the level of the Kernel to the SDD module. Enter the following command as shown in Figure 7-12:

```
insmod ./sdd-mod.o-2.4.9-e.16enterprise
```

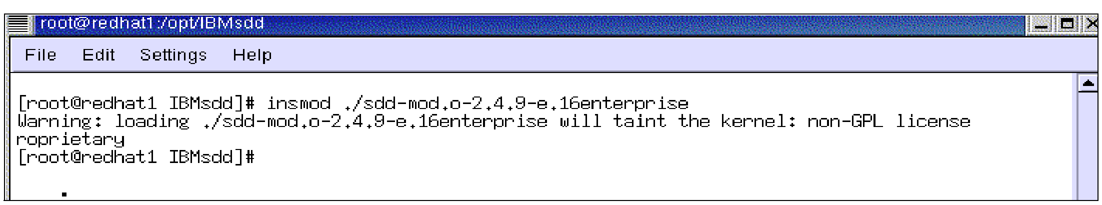


Figure 7-12 insmod

Issue the **cfgvpath** query to view the name and serial number of the vDisk configured in the SAN Volume Controller (see Figure 7-13).

```
root@redhat1:/opt/IBMsdd
File  Edit  Settings  Help

[root@redhat1 IBMsdd]# cfgvpath query
/dev/sda  ( 8,   0) host=0  ch=0  id=0  lun=0   vid=IBM-PSG  pid=ST318404LC   serial=xxxxxxxxxxxx
 df_ctlr=0 X
/dev/sdb  ( 8,  16) host=0  ch=0  id=1  lun=0   vid=IBM-PSG  pid=ST318404LC   serial=xxxxxxxxxxxx
 df_ctlr=0 X
/dev/sdc  ( 8,  32) host=0  ch=0  id=2  lun=0   vid=IBM-PSG  pid=ST318404LC   serial=xxxxxxxxxxxx
 df_ctlr=0 X
/dev/sdd  ( 8,  48) host=1  ch=0  id=0  lun=0   vid=IBM      pid=2145         serial=600507680182001b2000000000000007
 df_ctlr=0
/dev/sde  ( 8,  64) host=1  ch=0  id=1  lun=0   vid=IBM      pid=2145         serial=600507680182001b2000000000000007
 df_ctlr=0
/dev/sdf  ( 8,  80) host=2  ch=0  id=0  lun=0   vid=IBM      pid=2145         serial=600507680182001b2000000000000007
 df_ctlr=0
/dev/sdg  ( 8,  96) host=2  ch=0  id=1  lun=0   vid=IBM      pid=2145         serial=600507680182001b2000000000000007
 df_ctlr=0
[root@redhat1 IBMsdd]#
```

Figure 7-13 cfgvpath query

The **cfgvpath** command configures the SDD vpath devices. The configuration information is saved by default to the file /etc/vpath.conf. You can save the configuration information to a specified file name by entering the following command (see Figure 7-14):

```
cfgvpath -f file.name
```

```
root@redhat1:/opt/IBMsdd
File  Edit  Settings  Help

[root@redhat1 IBMsdd]# cfgvpath
crw-r--r--    1 root     root     252,   0 May 16 15:28 /dev/IBMsdd
vpatha has already been configured.
[root@redhat1 IBMsdd]# more /etc/vpath
/etc/vpath: No such file or directory
[root@redhat1 IBMsdd]# more /etc/vpath.conf
vpatha 600507680182001b2000000000000007
[root@redhat1 IBMsdd]#
```

Figure 7-14 cfgvpath

Issue the **chkconfig --** level command to enable SDD to run at system startup:

```
chkconfig -level 5 sdd on
```

To verify the setting, enter the following command:

```
chkconfig -list sdd
```

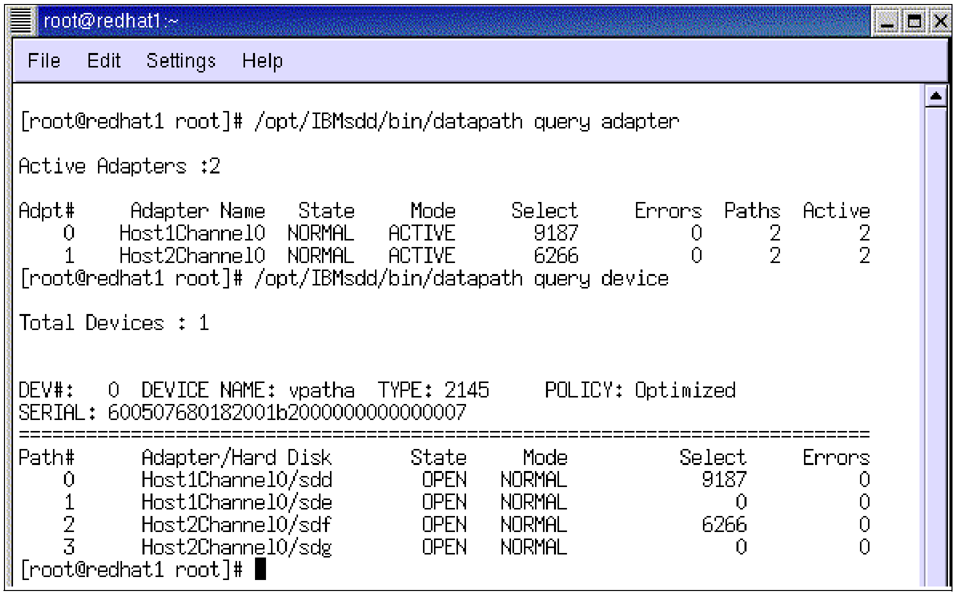
This is shown in Figure 7-15.

```
root@redhat1:/dev
File  Edit  Settings  Help

[root@redhat1 dev]# chkconfig --level 5 sdd on
[root@redhat1 dev]# chkconfig --list sdd
sdd             0:off   1:off   2:off   3:off   4:off   5:on    6:off
[root@redhat1 dev]#
```

Figure 7-15 SDD run level

Run the datapath query commands to display the online adapters, paths to adapters. Notice that the preferred paths are used from one of the nodes, that is path 0 and 2. Paths 1 and 3 connect to the other node and are used as alternate or backup paths for high availability. See Figure 7-16.



```
root@redhat1:~
File  Edit  Settings  Help

[root@redhat1 root]# /opt/IBMsdd/bin/datapath query adapter

Active Adapters :2

Adpt#     Adapter Name   State     Mode      Select     Errors  Paths  Active
    0     Host1Channel0  NORMAL    ACTIVE      9187          0      2       2
    1     Host2Channel0  NORMAL    ACTIVE      6266          0      2       2
[root@redhat1 root]# /opt/IBMsdd/bin/datapath query device

Total Devices : 1


DEV#:   0  DEVICE NAME: vpatha  TYPE: 2145     POLICY: Optimized
SERIAL: 600507680182001b2000000000000007
============================================================================
Path#      Adapter/Hard Disk       State     Mode           Select     Errors
    0      Host1Channel0/sdd        OPEN    NORMAL            9187          0
    1      Host1Channel0/sde        OPEN    NORMAL               0          0
    2      Host2Channel0/sdf        OPEN    NORMAL            6266          0
    3      Host2Channel0/sdg        OPEN    NORMAL               0          0
[root@redhat1 root]#
```

Figure 7-16 The datapath query

You can see the preferred paths in Example 7-13. it shows the vDisk information from the SVC command line.

Example 7-13 svcinfo redhat1

```
IBM_2145:admin>svcinfo lshost redhat1
id 2
name redhat1
port_count 2
WWPN 210000E08B091541
port_logged_in_count 2
WWPN 210100E08B291541
port_logged_in_count 2

IBM_2145:admin>svcinfo lshostvdiskmap -delim : redhat1
id:name:SCSI_id:vDisk_id:vDisk_name:wwpn:vDisk_UID
2:redhat1:0:5:redhatdisk1:210000E08B091541:600507680182001B2000000000000007
2:redhat1:0:5:redhatdisk1:210100E08B291541:600507680182001B2000000000000007

IBM_2145:admin>svcinfo lsvdisk redhatdisk1
id 5
name redhatdisk1
IO_group_id 0
IO_group_name ITSO_grp0
status online
mDisk_grp_id 0
mDisk_grp_name ARR36P5N
capacity 300.0MB
type striped
formatted no
```

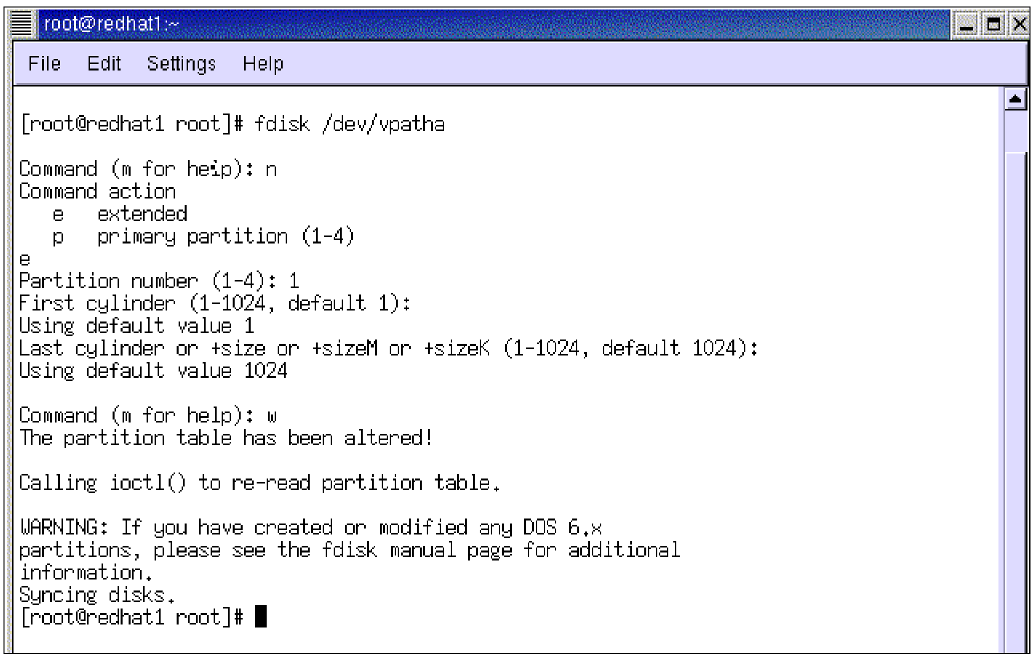
```
mDisk_id
mDisk_name
FC_id
FC_name
RC_id
RC_name
throttling 0
preferred_node_id 0

IBM_2145:admin>svcinfo lsvdiskhostmap redhatdisk1 -delim :
id:name:SCSI_id:host_id:host_name:wwpn:vDisk_UID
0:aix_vDisk0:0:0:aix1:10000000C9295A9A:600507680182001B2000000000000000
0:aix_vDisk0:0:0:aix1:10000000C9266F5B:600507680182001B2000000000000000
```

7.6.6 Creating and preparing volumes for use

Follow these steps:

1. Create a partition on the vpath device. See Figure 7-17.



```
root@redhat1:~
File  Edit  Settings  Help

[root@redhat1 root]# fdisk /dev/vpatha

Command (m for help): n
Command action
   e   extended
   p   primary partition (1-4)
e
Partition number (1-4): 1
First cylinder (1-1024, default 1):
Using default value 1
Last cylinder or +size or +sizeM or +sizeK (1-1024, default 1024):
Using default value 1024

Command (m for help): w
The partition table has been altered!

Calling ioctl() to re-read partition table.

WARNING: If you have created or modified any DOS 6.x
partitions, please see the fdisk manual page for additional
information.
Syncing disks.
[root@redhat1 root]#
```

Figure 7-17   fdisk

2. Create a file system on the vpath as shown in Figure 7-18.

```
[root@redhat1 root]# mkfs -t ext2 /dev/vpatha
mke2fs 1.26 (3-Feb-2002)
Filesystem label=
OS type: Linux
Block size=1024 (log=0)
Fragment size=1024 (log=0)
76912 inodes, 307200 blocks
15360 blocks (5.00%) reserved for the super user
First data block=1
38 block groups
8192 blocks per group, 8192 fragments per group
2024 inodes per group
Superblock backups stored on blocks:
        8193, 24577, 40961, 57345, 73729, 204801, 221185

Writing inode tables: done
Writing superblocks and filesystem accounting information: done

This filesystem will be automatically checked every 27 mounts or
180 days, whichever comes first.  Use tune2fs -c or -i to override.
[root@redhat1 root]#
```

Figure 7-18 mkfs

3. Create the mount point and mount the vpath drive as shown in Figure 7-19.

```
root@redhat1:/
File   Edit   Settings   Help

[root@redhat1 /]# mkdir /ITSOlv1mp
[root@redhat1 /]# mount -t ext2 /dev/vpatha /ITSOlv1mp
[root@redhat1 /]# df
Filesystem           1k-blocks      Used Available Use% Mounted on
/dev/sdc1             17488652   3232480  13367780  20% /
/dev/sda1                46636     23770     20458  54% /boot
none                    449376         0    449376   0% /dev/shm
/dev/vpatha             297485        13    282112   1% /ITSOlv1mp
[root@redhat1 /]# datapath query device

Total Devices : 1


DEV#:   0  DEVICE NAME: vpatha  TYPE: 2145     POLICY: Optimized
SERIAL: 600507680182001b2000000000000007
============================================================================
Path#      Adapter/Hard Disk       State     Mode         Select     Errors
    0      Host1Channel0/sdd        OPEN    NORMAL        131871          0
    1      Host1Channel0/sde        OPEN    NORMAL             0          0
    2      Host2Channel0/sdf        OPEN    NORMAL        132163          0
    3      Host2Channel0/sdg        OPEN    NORMAL             0          0
[root@redhat1 /]#
```

Figure 7-19 mountpoint

7.7 SUN Solaris and HP-UX support information

For the latest information about SUN Solaris and UP-UX support, see:

http://www.storage.ibm.com/support/2145

Operating system versions and maintenance levels

At GA, Sun Solaris Release 8 Version 7 or later is supported. HP-UX V11i was supported at GA.

SDD package version

At GA, SDD for both Solaris and HP-UX was supported.

8

SVC configuration and administration using the CLI

This chapter describes how to use the command line interface (CLI) to perform additional and advanced configuration and administration tasks that were not covered in Chapter 5, “Quickstart configuration using the CLI” on page 85. It also discusses the backup and recovery function that was announced in October 2003.

8.1 Managing the cluster

This section details the various configuration and administration tasks that you can perform on the cluster.

You must issue all of the following commands from a secure SSH command line. To launch your PuTTY command line, follow these steps:

1. Open the PuTTY application. From your master console desktop, select **Start -> Programs -> PuTTY**.
2. On the main screen (Figure 8-1), select the session you created and saved in 5.1, "Configuring the PuTTY session for the CLI" on page 86 (for example, `SVC`), and click **Load**. Then click **Open** to begin your session.
3. At the Login as: prompt, type `admin` and press Enter.

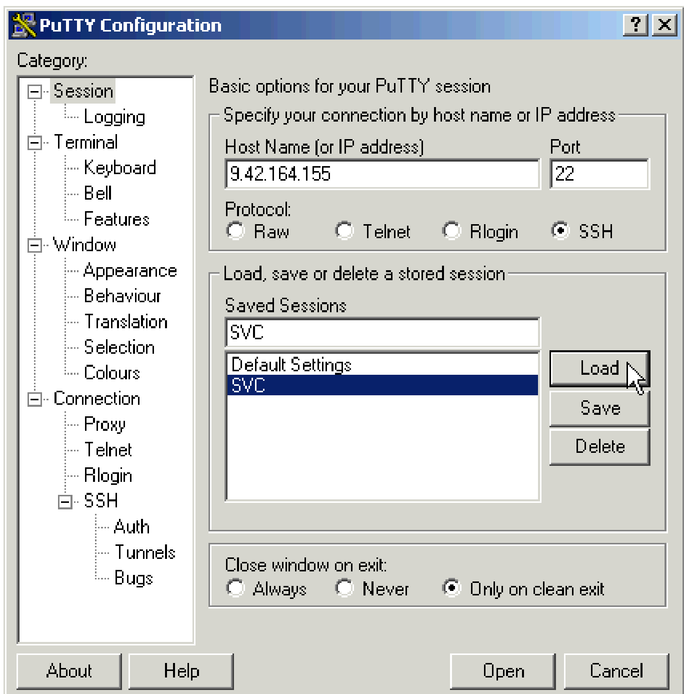


Figure 8-1 Starting PuTTY

Command syntax

Two major command sets are available to you. The **`svcinfo`** command set allows you to query the various components within the IBM TotalStorage SAN Volume Controller (SVC) environment. The **`svctask`** command set allows you to make changes to the various components within the SVC.

When the command syntax is shown, you see some parameters in square brackets, for example, `[parameter]`. This indicates that the parameter is optional in most, if not all instances. Anything that is not in square brackets is required information. You can view the syntax of a command by entering one of the following commands:

- **`svcinfo -?`**: Show a complete list of information commands
- **`svctask -?`**: Shows a complete list of task commands
- **`svcinfo commandname -?`**: Shows the syntax of information commands
- **`svctask commandname -?`**: Shows the syntax of task commands

- **`svcinfo commandname -filtervalue?`**: Shows what filters you can use to reduce output of the information commands

If you look at the syntax of the command by typing **`svcinfo command name -?`**, you often see -filter listed as a parameter. Be *aware* that the correct parameter is -filtervalue as stated above.

> **Tip:** You can use the up and down keys on your keyboard to recall commands recently issued. Then, you can use the left and right, backspace, and delete keys to edit commands before you resubmit them.

8.1.1 Organizing on-screen content

Sometimes the output of a command can be long and difficult to read on screen. In cases where you need information about a subset of the total number of available items, you can use filtering to reduce the output to a more manageable size.

Filtering

To reduce the output that is displayed by an **`svcinfo`** command, you can specify a number of filters depending on which **`svcinfo`** command you are running. To see which filters are available to you, type the command followed by the -filtervalue? flag to see such output as that which is shown in Example 8-1.

Example 8-1 svcinfo lsvdisk -filtervalue? command

```
IBM_2145:admin>svcinfo lsvdisk -filtervalue?

Filters for this view are :
     vDisk_name
     vDisk_id
     IO_group_id
     IO_group_name
     status
     mDisk_grp_name
     mDisk_grp_id
     capacity
     type
     FC_id
     FC_name
     RC_id
     RC_name
     is_mapped
```

When you know the filters, you can be more selective in generating output. For instance, if you issue the **`svcinfo lsvdisk`** command with no filters, you see the output as shown in Example 8-2.

Example 8-2 svcinfo lsvdisk command: No filters

```
IBM_2145:admin>svcinfo lsvdisk -delim :
id:name:IO_group_id:IO_group_name:status:mDisk_grp_id:mDisk_grp_name:capacity:type:FC_id:FC
_name:RC_id:RC_name
0:aix_vDisk0:0:ITSO_grp0:online:0:ARR36P5N:10.0GB:striped::::
```

```
1:aix_vDisk1:0:ITSO_grp0:online:0:ARR36P5N:10.0GB:seq::::
2:aix_vDisk2:0:ITSO_grp0:online:0:ARR36P5N:10.0GB:seq::::
3:aix_vDisk3:0:ITSO_grp0:online:0:ARR36P5N:10.0GB:striped::::
4:windisk1:0:ITSO_grp0:online:0:ARR36P5N:200.0MB:striped::::
5:redhatdisk1:0:ITSO_grp0:online:0:ARR36P5N:300.0MB:striped::::
6:aix_vDisk4:0:ITSO_grp0:online:0:ARR36P5N:1.0GB:striped::::
7:rhat2_vDisk0:0:ITSO_grp0:online:1:ARR36P6N:500.0MB:striped::::
8:aix_vDisk1_src:0:ITSO_grp0:online:1:ARR36P6N:500.0MB:striped:0:flashmapaix1::
9:aix_vDisk1_tgt:0:ITSO_grp0:online:2:ARR73P5N:500.0MB:striped:0:flashmapaix1::
10:TSM_POOL:0:ITSO_grp0:online:0:ARR36P5N:1.0GB:striped::::
11:SAP_DATA03:0:ITSO_grp0:online:2:ARR73P5N:500.0MB:seq::::
12:c1_aix_log_src:0:ITSO_grp0:online:1:ARR36P6N:500.0MB:striped:::12:pprcmapaixlog
13:c1_aix_data_src:0:ITSO_grp0:online:2:ARR73P5N:1.0GB:striped:::13:pprcmapaixdata
```

Tip: The -delim : parameter truncates the on-screen content and separates data fields with colons as opposed to wrapping text over multiple lines.

If you now add a filter to our **svcinfo** command (such as type), you can reduce the output dramatically as shown in Example 8-3.

Example 8-3 svcinfo lsvdisk command: With filter

```
IBM_2145:admin>svcinfo lsvdisk -filtervalue 'type=seq' -delim :
id:name:IO_group_id:IO_group_name:status:mDisk_grp_id:mDisk_grp_name:capacity:type:FC_id:FC
_name:RC_id:RC_name
1:aix_vDisk1:0:ITSO_grp0:online:0:ARR36P5N:10.0GB:seq::::
2:aix_vDisk2:0:ITSO_grp0:online:0:ARR36P5N:10.0GB:seq::::
11:SAP_DATA03:0:ITSO_grp0:online:2:ARR73P5N:500.0MB:seq::::
```

This command shows all sequential mode virtual disks (vDisks) in the SVC environment.

Note: Some filter values did not appear to work at the time of writing this redbook.

You are now ready to continue with the rest of this chapter.

8.1.2 Viewing cluster properties

Use the **svcinfo lscluster** command to display summary information about all clusters visible to the SVC. To display more detailed information about a specific cluster, run the command again and append the cluster name parameter (for example, REDSTONE1). Both of these commands are shown in Example 8-4.

Example 8-4 svcinfo lscluster command

```
IBM_2145:admin>svcinfo lscluster -delim :
id:name:location:partnership:bandwidth:cluster_IP_address:cluster_service_IP_address
0000010030400364:REDSTONE1:local:::9.42.164.155:9.42.164.160

IBM_2145:admin>svcinfo lscluster REDSTONE1
id 0000010030400364
name REDSTONE1
location local
partnership
```

```
bandwidth
cluster_IP_address 9.42.164.155
cluster_service_IP_address 9.42.164.160
total_mDisk_capacity 3479.9GB
space_in_mDisk_grps 1981.4GB
space_allocated_to_vDisks 449.7GB
total_free_space 3030.2GB
statistics_status on
statistics_frequency 15
required_memory 4096
cluster_locale en_US
SNMP_setting all
SNMP_community public
SNMP_server_IP_address 9.42.164.140
subnet_mask 255.255.255.0
default_gateway 9.42.164.1
time_zone 522 UTC
email_setting none
email_id storageadmin@ibm.com
code_level 00000000
FC_port_speed 2Gb
```

8.1.3 Maintaining passwords

Use the **svctask chcluster** command to change the admin and services passwords. The full syntax of the **svctask chcluster** command is:

```
svctask chcluster [-clusterip ip_address] [-serviceip ip_address]
          [-admpwd [password]] [-servicepwd [password]] [-gw gateway]
          [-mask subnet_mask] [-speed speed]
```

Note the following explanation:

- **clusterip**: IP address to access the cluster
- **serviceip**: IP address used if a node has been expelled from the cluster
- **admpwd**: Administrator's password
- **servicepwd**: Service user's password
- **gw**: cluster's gateway IP address
- **mask**: Subnet mask
- **speed**: Fabric speed

The command to change the admin password is:

```
IBM_2145:admin>svctask chcluster -admpwd admin -servicepwd service
```

This command changes the current admin password to `admin` and the current service password to `service`.

> **Note:** You can use letters A to Z, a to z, numbers 0 to 9, and the underscore in a password. The password can be between one and 15 characters in length.
>
> Since you are using an SSH key pair in your PuTTY command line session, these password changes only affect graphical user interface (GUI) users who must use passwords for initial authentication. Also, as you can see, these password are not encrypted on screen, so make sure no one is looking over your shoulder.

You have now completed the tasks required to change the admin and service passwords for your SVC cluster.

8.1.4 Modifying IP addresses

Using the **svctask chcluster** command again, you can change the cluster IP address as shown here:

```
IBM_2145:admin>svctask chcluster -clusterip 9.42.164.156 -serviceip 9.42.164.161
```

This command changes the current IP address of the cluster to 9.42.164.156 and the current service IP address to 9.42.164.161.

Important: If you specify a new cluster IP address, the existing communication with the cluster through the CLI is broken and the PuTTY application automatically closes. You must relaunch the PuTTY application and point to the new IP address.

Modifying the IP address of the cluster, although quite simple, means some reconfiguration for other items within the SVC environments (such reconfiguring our PuTTY application and the central administration GUI).

The -clusterip and -serviceip parameters can be used in isolation (as shown in the last two examples) or in combination with other **chcluster** command parameters.

You have now completed the tasks required to change the IP addresses (cluster and service) of your SVC environment.

8.1.5 Setting the cluster time zone and time

Perform the following steps to set the cluster time zone and time:

1. Determine what time zone your cluster is currently configured for by issuing the **svcinfo showtimezone** command as shown here:

   ```
   IBM_2145:admin>svcinfo showtimezone
   id              timezone
   522 UTC
   ```

   If this setting is correct (for example, 522 UTC), skip to Step 4. If not, continue with Step 2.

2. Determine the time zone code that is associated with your time zone. To find this, enter the **svcinfo lstimezones** command shown in Example 8-5. The list was truncated for purposes of the example.

Example 8-5 svcinfo lstimezones command

```
IBM_2145:admin>svcinfo lstimezones
. . .
508 UCT
509 Universal
510 US/Alaska
511 US/Aleutian
512 US/Arizona
513 US/Central
514 US/Eastern
515 US/East-Indiana
516 US/Hawaii
517 US/Indiana-Starke
518 US/Michigan
519 US/Mountain
```

```
520 US/Pacific
521 US/Samoa
. . .
```

In this example, the correct time zone code is 514.

3. Set the time zone by issuing the **svctask settimezone** command:

   ```
   IBM_2145:admin>svctask settimezone -timezone 514
   ```

4. With the correct time zone, set the cluster time by issuing the **svctask setclustertime** command:

   ```
   IBM_2145:admin>svctask setclustertime -time 0512150403
   ```

   The format of the time is MMDDHHmmYY.

You have now completed the tasks necessary to set the cluster time zone and time.

8.1.6 Starting a statistics collection

Use the **svctask startstats** command to start the collection of statistics within your cluster:

```
IBM_2145:admin>svctask startstats -interval 15
```

The interval you specify (minimum 15, maximum 60) is in minutes. This command starts statistics collection and gathers data at 15 minute intervals.

Note: To verify that statistics collection is set, display the cluster properties again, as shown in Example 8-6, and look about half way down.

Example 8-6 Statistics collection status and frequency

```
IBM_2145:admin>svcinfo lscluster REDSTONE1
id 0000010030400364
name REDSTONE1
location local
partnership
bandwidth
cluster_IP_address 9.42.164.155
cluster_service_IP_address 9.42.164.160
total_mDisk_capacity 3479.9GB
space_in_mDisk_grps 1981.4GB
space_allocated_to_vDisks 449.7GB
total_free_space 3030.2GB
statistics_status on
statistics_frequency 15
required_memory 4096
cluster_locale en_US
SNMP_setting all
SNMP_community public
SNMP_server_IP_address 9.42.164.140
subnet_mask 255.255.255.0
default_gateway 9.42.164.1
time_zone 522 UTC
email_setting none
email_id storageadmin@ibm.com
code_level 00000000
FC_port_speed 2Gb
```

You have now completed the tasks required to start statistics collection on your cluster.

8.1.7 Stopping a statistics collection

Use the **svctask stopstats** command to start the collection of statistics within your cluster:

```
IBM_2145:admin>svctask stopstats
```

This command stop statistics collection.

> **Note:** To verify that statistics collection is stopped, display the cluster properties again, as shown in Example 8-7, and look about half way down.

Example 8-7 Statistics collection status and frequency

```
IBM_2145:admin>svcinfo lscluster REDSTONE1
id 0000010030400364
name REDSTONE1
location local
partnership
bandwidth
cluster_IP_address 9.42.164.155
cluster_service_IP_address 9.42.164.160
total_mDisk_capacity 3479.9GB
space_in_mDisk_grps 1981.4GB
space_allocated_to_vDisks 449.7GB
total_free_space 3030.2GB
statistics_status off
statistics_frequency 15
required_memory 4096
cluster_locale en_US
SNMP_setting all
SNMP_community public
SNMP_server_IP_address 9.42.164.140
subnet_mask 255.255.255.0
default_gateway 9.42.164.1
time_zone 522 UTC
email_setting none
email_id storageadmin@ibm.com
code_level 00000000
FC_port_speed 2Gb
```

Notice that the interval parameter is not changed but the status is *off*. You have now completed the tasks required to stop statistics collection on your cluster.

8.1.8 Shutting down a cluster

If all input power to a SVC cluster is to be removed for more than a few minutes (for example, if the machine room power is to be shutdown for maintenance), it is important that you shut down the cluster before you remove the power. The reason for this is that if the input power is removed from the uninterruptible power supply units without first shutting down the cluster and the uninterruptible power supplies themselves, the uninterruptible power supply units remain operational and eventually become drained of power.

When input power is restored to the uninterruptible power supplies, they start to recharge. However the SVC does not permit any input/output (I/O) activity to be performed to the vDisks until the uninterruptible power supplies are charged enough to enable all the data on the SVC nodes to be destaged in the event of a subsequent unexpected power loss. Recharging the uninterruptible power supply may take as long as three hours.

Shutting down the cluster prior to removing input power to the uninterruptible power supply units prevents the battery power from being drained. It also makes it possible for I/O activity to be resumed as soon as input power is restored.

The following procedure shuts down the cluster.

> **Important:** Make sure all connected hosts are shutdown prior to performing this task. Failure to do so can cause data corruption or loss.

Use the **`svctask stopcluster`** command to shut down your SVC cluster:

```
IBM_2145:admin>svctask stopcluster
```

This command shuts down the SVC cluster. At this point you lose administrative contact with your cluster, and the PuTTY application automatically closes.

You have now completed the tasks required to shut down the cluster. You can now shut down the uninterruptible power supplies by pressing the power button on their front panels.

> **Note:** To restart the cluster, you must first restart the uninterruptible power supply units by pressing the power button on their front panels. Then, go to the service panel of one of the nodes within your cluster and push the power on button. After it is fully booted up (for example, displaying `Cluster:` on line 1 and the cluster name on line 2 of the display panel), you can start the other nodes in the same way.
>
> As soon as all nodes are fully booted and you re-establish administrative contact using PuTTY, your cluster is fully operational again.

8.2 Working with nodes

This section explains the various configuration and administration tasks that you can perform on the nodes within an SVC cluster.

8.2.1 I/O groups

This section explains the tasks that you can perform on an I/O group level.

Viewing I/O group details

Use the **`svcinfo lsiogrp`** command, as shown in Example 8-8, to view information about I/O groups defined within your SVC environment.

Example 8-8 I/O group details

```
IBM_2145:admin>svcinfo lsiogrp
id              name            node_count      vDisk_count
0               ITSO_grp0       2               10
1               ITSO_grp1       0               0
```

```
2               ITSO_grp2       0             0
3               ITSO_grp3       0             0
4               recovery_io_grp 0             0
```

As you can see, the SVC predefines five I/O groups. In a two-node cluster (like ours) only one I/O group is actually in use. In a four -node cluster, you would have two I/O groups in use. The other “standard” I/O groups (for example, `ITSO_grp2` and `ITSO_grp3`) are reserved for future expansion.

The recovery I/O group is a temporary home for vDisks when both nodes in the I/O group that normally owns them have suffered multiple failures. This allows you to move the vDisks to the recovery I/O group and then into a working I/O group. Of course, while temporarily assigned to the recovery I/O group, I/O access is not possible.

Renaming an I/O group

Use the **`svctask chiogrp`** command to rename an I/O group:

```
IBM_2145:admin>svctask chiogrp -name grp3_ITSO ITSO_grp3
```

This command renames the I/O group ITSO_grp3 to `grp3_ITSO`.

> **Note:** The **`chiogrp`** command specifies the new name first.
>
> If you want to provide a name, you can use letters A to Z, a to z, numbers 0 to 9, and the underscore. It can be between one and 15 characters in length. However, it cannot start with a number or the word *iogrp* since this prefix is reserved for SVC assignment only.

To see the whether your renaming was successful, issue the **`svcinfo lsiogrp`** command again and you should see the change reflected.

You have now completed the tasks required to rename an I/O group.

8.2.2 Nodes

This section details the tasks which can be performed on an individual node level.

Viewing node details

Use the **`svcinfo lsnode`** command to view summary information about nodes defined within your SVC environment. To view more details about a specific node, append the node name (for example, `ITSO_node1`) to the command.

Both of these commands are shown in Example 8-9.

> **Tip:** The -delim : parameter truncates the onscreen content and separates data fields with colons as opposed to wrapping text over multiple lines.

Example 8-9 svcinfo lsnode command

```
IBM_2145:admin>svcinfo lsnode -delim :
id:name:UPS_serial_number:WWNN:status:IO_group_id:IO_group_name:config_node:UPS_unique_id
0:ITSO_node1:YM100032B422:5005076801000364:online:0:ITSO_grp0:yes:20400000C2484082
1:ITSO_node2:YM100032B425:500507680100035A:online:0:ITSO_grp0:no:20400000C2484085
```

```
IBM_2145:admin>svcinfo lsnode ITSO_node1
id 0
name ITSO_node1
UPS_serial_number YM100032B422
WWNN 5005076801000364
status online
IO_group_id 0
IO_group_name ITSO_grp0
partner_node_id 9
partner_node_name ITSO_node2
config_node yes
UPS_unique_id 20400000C2484082
port_id 5005076801100364
port_status active
port_id 5005076801200364
port_status active
port_id 5005076801300364
port_status active
port_id 5005076801400364
port_status active
```

Adding a node

Before you can add a node, you must know which unconfigured nodes you have as "candidates". You can find this out by issuing the **svcinfo lsnodecandidate** command:

```
IBM_2145:admin>svcinfo lsnodecandidate
id               node_cover_name   UPS_serial_number     UPS_unique_id
5005076801000364 000667            YM100032B422          20400000C2484082
```

Now that you know the available nodes, you can use the **svctask addnode** command to add the node to the SVC cluster configuration. The complete syntax of the **addnode** command is:

```
addnode {-panelname panel_name | -wwnodename wwnn_arg} [-name new_name]
        -iogrp iogrp_name_or_id
```

Note the following explanation:

- **panelname**: Name of the node as it appears on the panel
- **wwnodename**: Worldwide node name (WWNN) of the node
- **name**: Name to be allocated to the node
- **iogrp**: I/O group to which the node is added

The command to add a node to your SVC cluster is:

```
IBM_2145:admin>svctask addnode -panelname 000667 -name ITSO_node1 -iogrp ITSO_grp0
Node, id [10], successfully added
```

This command adds the candidate node with the panelname of `000667` (actually displayed as `node_cover_name` in the **svcinfo lsnodecandidate** output) to the I/O group ITSO_grp0 and name it `ITSO_node1`.

We used the -panelname parameter (`000667`), but could have used the -wwnodename parameter (`5005076801000364`) instead, for example:

```
svctask addnode -wwnodename 5005076801000364 -name ITSO_node1 -iogrp ITSO_grp0
```

We also used the optional -name parameter (`ITSO_node1`). If you do not provide the -name parameter, the SVC automatically generates the name *nodeX* (where *X* is the ID sequence number assigned by the SVC internally).

Note: If you want to provide a name (as we did), you can use letters A to Z, a to z, numbers 0 to 9, and the underscore. It can be between one and 15 characters in length. However, it cannot start with a number or the word *node* because this prefix is reserved for SVC assignment only.

Deleting a node

Use the **`svctask rmnode`** command to remove a node from the SVC cluster configuration:

```
IBM_2145:admin>svctask rmnode ITSO_node1
```

This command removes node ITSO_node1 from the SVC cluster.

Since ITSO_node1 was also the configuration node, the SVC transfers the configuration node responsibilities to a surviving node (in our case ITSO_node2). Unfortunately the PuTTY session cannot be dynamically passed to the surviving node. Therefore the PuTTY application loses communication and closes automatically.

You *must* restart the PuTTY application to establish a secure session with the new configuration node.

Note: You are not allowed to delete the last node in an I/O group when vDisks are still associated with it.

Renaming a node

Use the **`svctask chnode`** command to rename a node within the SVC cluster configuration:

```
IBM_2145:admin>svctask chnode -name ITSO_node0 ITSO_node1
```

This command renames node ITSO_node1 to `ITSO_node0`.

Note: The **`chnode`** command specifies the new name first. The name can consist of letters A to Z, a to z, numbers 0 to 9, and the underscore. It can be between one and 15 characters in length. However, it cannot start with a number or the word *node* because this prefix is reserved for SVC assignment only.

Shutting down a node

Earlier you learned how to shut down the complete SVC cluster in a controlled manner. On occasion, it may be necessary to shut down a single node within the cluster, to perform such tasks as scheduled maintenance, while leaving the SVC environment up and running.

Use the **`svctask stopcluster -node`** command as shown in Example 8-10 to shut down a node.

Example 8-10 svctask stopcluster -node command

```
IBM_2145:admin>svctask stopcluster -node ITSO_node1
Are you sure that you want to continue with the shut down? Ensure that you have stopped all
FlashCopy mappings, Remote Copy relationships, data migration operations and forced
deletions before continuing.
yes <ENTER>
```

This command shuts down ITSO_node1 in an orderly manner.

Note: Since this is not the configuration node in this case, the PuTTY session remains intact.

By re-issuing the **svcinfo lsnode** commands you saw earlier (as shown in Example 8-11), you see that the node is now offline.

Example 8-11 svcinfo lsnode

```
IBM_2145:admin>svcinfo lsnode -delim :
id:name:UPS_serial_number:WWNN:status:IO_group_id:IO_group_name:config_node:UPS_unique_id
0:ITSO_node1:YM100032B422:0000000000000000:offline:0:ITSO_grp0:no:20400000C2484082
1:ITSO_node2:YM100032B425:500507680100035A:online:0:ITSO_grp0:yes:20400000C2484085

IBM_2145:admin>svcinfo lsnode ITSO_node1
CMMVC5782E The object specified is offline
```

To restart the node, you simply go to the service panel of the node and push the power on button.

You have now completed the tasks required to view, add, delete, rename and shut down a node within an SVC environment.

8.3 Working with managed disks

This section details the various configuration and administration tasks that you can perform on the managed disks (mDisks) within the SVC environment.

8.3.1 Disk controller systems

This section details the tasks that you can perform on a disk controller level.

Viewing disk controller details

Use the **svcinfo lscontroller** command to display summary information about all available back-end storage systems. To display more detailed information about a specific controller, run the command again and append the controller name parameter (for example, `controller0`). Both of these commands are shown in Example 8-12.

Tip: The -delim : parameter truncates the onscreen content and separates data fields with colons as opposed to wrapping text over multiple lines.

Example 8-12 svcinfo lscontroller command

```
IBM_2145:admin>svcinfo lscontroller -delim :
id:controller_name:ctrl_s/n:vendor_id:product_id_low:product_id_high
0:controller0::IBM     :1722-600:

IBM_2145:admin>svcinfo lscontroller controller0
id 0
controller_name controller0
WWNN 200800A0B80FBDF0
```

```
mDisk_link_count 9
max_mDisk_link_count 9
degraded no
vendor_id IBM
product_id_low 1722-600
product_id_high
product_revision 0520
ctrl_s/n
WWPN 200900A0B80FBDF2
path_count 6
max_path_count 14
WWPN 200800A0B80FBDF1
path_count 8
max_path_count 16
```

Renaming a controller

Use the **svctask chcontroller** command to change the name of a storage controller. To verify the change, run the **svcinfo lscontroller** command. Both of these commands are shown in Example 8-13.

Example 8-13 svctask chcontroller command

```
IBM_2145:admin>svctask chcontroller -name fastt600 controller0

IBM_2145:admin>svcinfo lscontroller -delim :
id:controller_name:ctrl_s/n:vendor_id:product_id_low:product_id_high
0:fastt600::IBM     :1722-600:
```

As you can see, this command renamed the controller named controller0 to `fastt600`.

Note: The **chcontroller** command specifies the new name first. The name can consist of letters A to Z, a to z, numbers 0 to 9, and the underscore. It can be between one and 15 characters in length. However, it cannot start with a number or the word *controller* because this prefix is reserved for SVC assignment only.

8.3.2 Managed disks

This section details the tasks that you can perform on an mDisk level.

mDisk information

Use the **svcinfo lsmdisk** command to display summary information about all available managed disks. To display more detailed information about a specific mDisk, run the command again and append the mDisk name parameter (for example, `ITSO_mDisk0`). Both of these commands are shown in Example 8-14.

Example 8-14 svcinfo lsmdisk command

```
IBM_2145:admin>svcinfo lsmdisk -delim :
id:name:status:mode:mDisk_grp_id:mDisk_grp_name:capacity:ctrl_LUN_#:controller_name
0:ITSO_mDisk0:online:managed:0:ARR36P5N:167.0GB:0000000000000000:fastt600
1:ITSO_mDisk1:online:managed:1:ARR36P6N:200.4GB:0000000000000001:fastt600
2:ITSO_mDisk2:online:managed:0:ARR36P5N:167.0GB:0000000000000002:fastt600
```

```
3:ITSO_mDisk3:online:managed:1:ARR36P6N:200.4GB:0000000000000003:fastt600
4:ITSO_mDisk4:online:managed:2:ARR73P5N:339.3GB:0000000000000004:fastt600
5:ITSO_mDisk5:online:managed:2:ARR73P5N:407.2GB:0000000000000005:fastt600
6:mDisk6:excluded:unmanaged:::681.2GB:0000000000000006:fastt600
7:mDisk7:excluded:unmanaged:::817.4GB:0000000000000007:fastt600
8:mDisk8:online:unmanaged:1:ARR36P6N:500.0GB:0000000000000006:fastt600

IBM_2145:admin>svcinfo lsmdisk ITSO_mDisk0
id 0
name ITSO_mDisk0
status online
mode managed
mDisk_grp_id 0
mDisk_grp_name ARR36P5N
capacity 167.0GB
quorum_index 0
block_size 512
controller_name fastt600
ctrl_type 4
ctrl_WWNN 200800A0B80FBDF0
controller_id 0
path_count 2
max_path_count 2
ctrl_LUN_# 0000000000000000
```

Renaming an mDisk

Use the **svctask chmDisk** command to change the name of an mDisk. To verify the change, run the **svcinfo lsmdisk** command. Both of these commands are shown in Example 8-15.

Example 8-15 svctask chmDisk command

```
IBM_2145:admin>svctask chmDisk -name ITSO_mDisk8 mDisk8

IBM_2145:admin>svcinfo lsmdisk -delim :
id:name:status:mode:mDisk_grp_id:mDisk_grp_name:capacity:ctrl_LUN_#:controller_name
0:ITSO_mDisk0:online:managed:0:ARR36P5N:167.0GB:0000000000000000:fastt600
1:ITSO_mDisk1:online:managed:1:ARR36P6N:200.4GB:0000000000000001:fastt600
2:ITSO_mDisk2:online:managed:0:ARR36P5N:167.0GB:0000000000000002:fastt600
3:ITSO_mDisk3:online:managed:1:ARR36P6N:200.4GB:0000000000000003:fastt600
4:ITSO_mDisk4:online:managed:2:ARR73P5N:339.3GB:0000000000000004:fastt600
5:ITSO_mDisk5:online:managed:2:ARR73P5N:407.2GB:0000000000000005:fastt600
6:mDisk6:excluded:unmanaged:::681.2GB:0000000000000006:fastt600
7:mDisk7:excluded:unmanaged:::817.4GB:0000000000000007:fastt600
8:ITSO_mDisk8:online:unmanaged:1:ARR36P6N:500.0GB:0000000000000006:fastt600
```

As you can see, this command renamed the mDisk named mDisk8 to `ITSO_mDisk8`.

> **Note:** The **chmDisk** command specifies the new name first. The name can consist of letters A to Z, a to z, numbers 0 to 9, and the underscore. It can be between one and 15 characters in length. However, it cannot start with a number or the word *mDisk* because this prefix is reserved for SVC assignment only.

Discovering mDisks

Use the **svctask detectmdisk** command to scan for newly added mDisks:

```
IBM_2145:admin>svctask detectmdisk
```

To check whether any newly added mDisk was successfully detected, run the **svcinfo lsmdisk** command as before. If your disks do not appear, check that the disk is appropriately assigned to SVC in your disk subsystem and that your zones are properly set up as explained in Chapter 3, “Planning and configuration” on page 25.

Setting a quorum disk

The SVC cluster, after the process of node discovery, automatically chooses three mDisks as quorum disks. Each disk is assigned an index number of 0, 1, or 2.

The quorum disks are only created once when at least one managed mDisk with an available extent is placed in managed mode.

In the event that half the nodes in a cluster are missing for any reason, the other half cannot simply assume that the nodes are “dead”. It can simply mean that the cluster state information is not being successfully passed between nodes for some reason (network failure for example). For this reason, if half the cluster disappears from the view of the other, each surviving half attempts to lock the first quorum disk (`index 0`). In the event of quorum disk `index 0` not being available, the next disk (`index 1`) becomes the quorum, and so on.

The half of the cluster that is successful in locking the quorum disk becomes the exclusive processor of I/O activity. It attempts to reform the cluster with any nodes it can still see.

If for any reason you want to set your own quorum disks, you can use the **svctask setquorum** command, as shown in Example 8-16, to reassign the quorum indexes.

Example 8-16 svctask setquorum command

```
IBM_2145:admin>svctask setquorum -quorum 0 ITSO_mDisk8

IBM_2145:admin>svcinfo lsmdisk ITSO_mDisk8
id 8
name ITSO_mDisk8
status online
mode managed
mDisk_grp_id 1
mDisk_grp_name ARR36P6N
capacity 500.0GB
quorum_index 0
block_size 512
controller_name fastt600
ctrl_type 4
ctrl_WWNN 200800A0B80FBDF0
controller_id 0
path_count 2
max_path_count 2
ctrl_LUN_# 0000000000000006
```

As you can see, this command has set `ITSO_mDisk8` as a quorum disk using quorum index `0`.

Including an mDisk

If a significant number of errors occur on an mDisk, the SVC automatically excludes it. These errors can be from a hardware problem, a storage area network (SAN) zoning problem, or the result of poorly planned maintenance. If it was a hardware fault, you should receive Simple Network Management Protocol (SNMP) alerts as the state of the disk subsystem (before the disk was excluded) and undertaken preventative maintenance. If not, the hosts that were using vDisks, which used the excluded mDisk, now have I/O errors.

By running the **svcinfo lsmdisk** command, you can see that mDisk5 is excluded in Example 8-17.

Example 8-17 svcinfo lsmdisk command: Excluded mDisk

```
IBM_2145:admin>svcinfo lsmdisk -delim :
id:name:status:mode:mDisk_grp_id:mDisk_grp_name:capacity:ctrl_LUN_#:controller_name
0:mDisk0:online:unmanaged:::167.0GB:0000000000000000:F600_23C1436
1:mDisk1:online:unmanaged:::200.4GB:0000000000000001:F600_23C1436
2:mDisk2:online:unmanaged:::167.0GB:0000000000000002:F600_23C1436
3:mDisk3:online:unmanaged:::200.4GB:0000000000000003:F600_23C1436
4:mDisk4:online:managed:2:Migrated_vDisks:339.3GB:0000000000000004:F600_23C1436
5:mDisk5:excluded:managed:2:Migrated_vDisks:407.2GB:0000000000000005:F600_23C1436
6:mDisk6:online:unmanaged:::500.0GB:0000000000000006:F600_23C1436
7:mDisk7:online:managed:1:ess_mDiskgrp0:4.7GB:0000000000000000:ESS
8:mDisk8:online:managed:1:ess_mDiskgrp0:4.7GB:0000000000000001:ESS
9:mDisk9:online:managed:1:ess_mDiskgrp0:4.7GB:0000000000000002:ESS
10:mDisk10:online:managed:1:ess_mDiskgrp0:4.7GB:0000000000000003:ESS
```

After taking the necessary corrective action to repair the mDisk (for example, replace failed disk, repair SAN zones, etc.), you can tell the SVC to include the mDisk again by issuing the **svctask includemDisk** command:

```
IBM_2145:admin>svctask includemDisk mDisk5
```

Running the **svcinfo lsmdisk** command again should show mDisk5 online again, as shown in Example 8-18.

Example 8-18 svcinfo lsmdisk command: Verifying that mDisk is included

```
IBM_2145:admin>svcinfo lsmdisk -delim :
id:name:status:mode:mDisk_grp_id:mDisk_grp_name:capacity:ctrl_LUN_#:controller_name
0:mDisk0:online:unmanaged:::167.0GB:0000000000000000:F600_23C1436
1:mDisk1:online:unmanaged:::200.4GB:0000000000000001:F600_23C1436
2:mDisk2:online:unmanaged:::167.0GB:0000000000000002:F600_23C1436
3:mDisk3:online:unmanaged:::200.4GB:0000000000000003:F600_23C1436
4:mDisk4:online:managed:2:Migrated_vDisks:339.3GB:0000000000000004:F600_23C1436
5:mDisk5:online:managed:2:Migrated_vDisks:407.2GB:0000000000000005:F600_23C1436
6:mDisk6:online:unmanaged:::500.0GB:0000000000000006:F600_23C1436
7:mDisk7:online:managed:1:ess_mDiskgrp0:4.7GB:0000000000000000:ESS
8:mDisk8:online:managed:1:ess_mDiskgrp0:4.7GB:0000000000000001:ESS
9:mDisk9:online:managed:1:ess_mDiskgrp0:4.7GB:0000000000000002:ESS
10:mDisk10:online:managed:1:ess_mDiskgrp0:4.7GB:0000000000000003:ESS
```

Showing the mDisk group

Use the **svcinfo lsmdisk** command as before to display information about the managed disk group (MDG) to which an mDisk belongs, as shown in Example 8-19.

Example 8-19 svcinfo lsmdisk command

```
IBM_2145:admin>svcinfo lsmdisk -delim :
id:name:status:mode:mDisk_grp_id:mDisk_grp_name:capacity:ctrl_LUN_#:controller_name
0:ITSO_mDisk0:online:managed:0:ARR36P5N:167.0GB:0000000000000000:fastt600
1:ITSO_mDisk1:online:managed:1:ARR36P6N:200.4GB:0000000000000001:fastt600
2:ITSO_mDisk2:online:managed:0:ARR36P5N:167.0GB:0000000000000002:fastt600
3:ITSO_mDisk3:online:managed:1:ARR36P6N:200.4GB:0000000000000003:fastt600
4:ITSO_mDisk4:online:managed:2:ARR73P5N:339.3GB:0000000000000004:fastt600
5:ITSO_mDisk5:online:managed:2:ARR73P5N:407.2GB:0000000000000005:fastt600
6:mDisk6:excluded:unmanaged:::681.2GB:0000000000000006:fastt600
7:mDisk7:excluded:unmanaged:::817.4GB:0000000000000007:fastt600
8:ITSO_mDisk8:online:unmanaged:1:ARR36P6N:500.0GB:0000000000000006:fastt600
```

See 8.3.3, “MDGs” on page 180, for more details about MDGs.

Showing the vDisk for mDisk

Use the **svcinfo lsmdiskmember** command to display information about the vDisks that use space on a specific mDisk, as shown in Example 8-20.

Example 8-20 svcinfo lsmdiskmember command

```
IBM_2145:admin>svcinfo lsmdiskmember ITSO_mDisk0
id
0
1
3
4
5
6
```

This command shows you that the vDisks with IDs 0 to 6 are all using space on ITSO_mDisk0.

To correlate the IDs displayed in this output to vDisk names, you can run the **svcinfo lsvdisk** command. We discuss this command in more detail in 8.4, “Working with virtual disks” on page 183.

Creating a vDisk in image mode

An image mode disk is a vDisk that has an exact one-to-one (1:1) mapping of vDisk extents with the underlying mDisk. For example, extent 0 on the vDisk contains the same data as extent 1 on the mDisk and so on. Without this one-to-one mapping (for example, if extent 0 on the vDisk mapped to extent 3 on the mDisk), there is little chance that the data on a newly introduced mDisk is still readable.

Image mode is intended for the purpose of migrating data from an environment outside the SVC, to an environment within the SVC. A logical unit number (LUN) that was previously directly assigned to a SAN attached host can now be reassigned to the SVC (possible short outage) and given back to the same host as an image mode vDisk. During the same outage, the host and zones can be reconfigured to access the disk via the SVC.

After access is re-established, the host workload can resume while the SVC manages the transparent migration of the data to other SVC managed vDisks on the same or another disk subsystem.

We recommend that, during the migration phase of the SVC implementation, you add one mDisk at a time to the SVC environment. This reduces the possibility of error. It also means that the short outages required to reassign the LUNs from the subsystem or subsystems and reconfigure the SAN and host can be staggered over a period of time to minimize the business impact.

> **Important:** Creating an image mode vDisk can be done only using an unmanaged disk (for example, before you added it to a MDG). We recommend that you create an *empty* MDG, called image_mode, or similar since you need to add your newly created image mode vDisk to a MDG. See 8.3.3, “MDGs” on page 180, for information about creating a MDG.
>
> If you noticed earlier, ITSO_mDisk8 is currently unmanaged. For example, it does not belong to a MDG. Therefore, we use it to demonstrate this process.

Use the **`svctask mkvdisk`** command to create an image mode vDisk. The full syntax of this command is:

```
svctask mkvdisk -mDiskgrp name|id -iogrp name|id -size size [-fmtdisk]
        [-vtype seq|striped|image] [-node name|id] [-unit b|kb|mb|gb|tb|pb]
        [-mDisk name|id_list] [-name name]
```

Here, the parameters are defined as follows:

- **mDiskgrp**: Name or ID of the MDG in which to create the vDisk.
- **iogrp**: Name or ID of I/O group which is to own the vDisk.
- **Size**: Capacity (numerical); not necessary for image mode vDisks.
- **fmtdisk**: Optional parameter to force a format of the new vDisk.
- **vtype**: Optional parameter to specify the type of vDisk (sequential, striped, or image mode). Default (if nothing is specified) is striped.
- **node**: Optional parameter to specify the name or ID of the preferred node. Default (if nothing is specified) is to alternate between nodes in the I/O group.
- **unit**: Optional parameter to specify the data units for capacity parameter. Default (if nothing is specified) is megabytes (MB).
- **mDisk**: Optional parameter to specify the name or ID of the mDisk or mDisks to be used for the vDisk. This is only required for sequential and image mode vDisks because striped vDisks use all mDisks that are available in the MDG by default.

  > **Note:** You can use this parameter for striped vDisks, for example, if you want to specify that the vDisk only uses a subset of the mDisks available within a MDG.

- **name**: Optional parameter to assign a name to the new vDisk. Default (if nothing is specified) is to assign the name *vDiskX*, where *X* is the ID sequence number assigned by the SVC internally.

> **Note:** If you do not provide the -name parameter, the SVC automatically generates the name *vDiskX* (where *X* is the ID sequence number assigned by the SVC internally).
>
> If you want to provide a name, you can use letters A to Z, a to z, numbers 0 to 9, and the underscore. It can be between one and 15 characters in length. However, it cannot start with a number or the word *vDisk* because this prefix is reserved for SVC assignment only.

The command to create an image mode vDisk is:

```
IBM_2145:admin>svctask mkvdisk -mDiskgrp image_mode -iogrp ITSO_grp0 -vtype image -mDisk
ITSO_mDisk8 -name SAP_DATA01
Host LUN, id [7], successfully created
```

This command creates an image mode vDisk called `SAP_DATA01` using mDisk ITSO_mDisk8. The vDisk belongs to the MDG image_mode and is owned by the I/O group ITSO_grp0 (ITSO_grp0 is our only I/O group).

If you run the **svcinfo lsmdisk** command again, you notice that ITSO_mDisk8 now has a status of *image* as shown in Example 8-21.

Example 8-21 svcinfo lsmdisk command: mDisk status

```
IBM_2145:admin>svcinfo lsmdisk -delim :
id:name:status:mode:mDisk_grp_id:mDisk_grp_name:capacity:ctrl_LUN_#:controller_name
0:ITSO_mDisk0:online:managed:0:ARR36P5N:167.0GB:0000000000000000:fastt600
1:ITSO_mDisk1:online:managed:1:ARR36P6N:200.4GB:0000000000000001:fastt600
2:ITSO_mDisk2:online:managed:0:ARR36P5N:167.0GB:0000000000000002:fastt600
3:ITSO_mDisk3:online:managed:1:ARR36P6N:200.4GB:0000000000000003:fastt600
4:ITSO_mDisk4:online:managed:2:ARR73P5N:339.3GB:0000000000000004:fastt600
5:ITSO_mDisk5:online:managed:2:ARR73P5N:407.2GB:0000000000000005:fastt600
6:mDisk6:excluded:unmanaged:::681.2GB:0000000000000006:fastt600
7:mDisk7:excluded:unmanaged:::817.4GB:0000000000000007:fastt600
8:ITSO_mDisk8:online:image:3:image_mode:500.0GB:0000000000000006:fastt600
```

8.3.3 MDGs

This section explains the tasks that you can perform on a MDG level.

Viewing mDisk group information

Use the **svcinfo lsmdiskgrp** command, as shown in Example 8-22, to display information about the MDGs defined in the SVC.

Example 8-22 svcinfo lsmdiskgrp command

```
IBM_2145:admin>svcinfo lsmdiskgrp -delim :
id:name:status:mDisk_count:vDisk_count:capacity:extent_size:free_capacity
0:ARR36P5N:online:2:7:333.9GB:32:292.4GB
1:ARR36P6N:online:2:1:400.8GB:32:400.3GB
2:ARR73P5N:online:2:1:746.5GB:32:746.0GB
3:image_mode:online:1:1:500.0GB:32:0
```

Creating an mDisk group

Use the **svctask mkmdiskgrp** command to create an MDG. The full syntax of this command is:

```
svctask mkmdiskgrp [-name name] [-mDisk name|id_list] -ext size
```

Note the following explanation:

- **name**: Name to assign to new group
- **mdisk**: List of names or IDs of mDisks to assign to group
- **extent**: Size of extents in this group

The command to create an MDG is:

```
IBM_2145:admin>svctask mkmDiskgrp -name image_mode -ext 32
mDisk Group, id [3], successfully created
```

This command creates an MDG called `image_mode` with an extent size of 32 MB. Since we did not specify any mDisks to add to the group with the -mDisk parameter, this is an empty MDG.

If you run the **svcinfo lsmdiskgrp** command, you should see the MDG you created as shown in Example 8-23.

Example 8-23 svcinfo lsmdiskgrp command

```
IBM_2145:admin>svcinfo lsmdiskgrp -delim :
id:name:status:mDisk_count:vDisk_count:capacity:extent_size:free_capacity
0:ARR36P5N:online:2:7:333.9GB:32:292.4GB
1:ARR36P6N:online:2:1:400.8GB:32:400.3GB
2:ARR73P5N:online:2:1:746.5GB:32:746.0GB
3:image_mode:online:0:0:0:32:0
```

Renaming an mDisk group

Use the **svctask chmDiskgrp** command to change the name of an MDG. To verify change, run the **svcinfo lsmdiskgrp** command. Both of these commands are shown in Example 8-24.

Example 8-24 svctask chmDiskgrp command

```
IBM_2145:admin>svctask chmDiskgrp -name IMAGE image_mode

IBM_2145:admin>svcinfo lsmdiskgrp -delim :
id:name:status:mDisk_count:vDisk_count:capacity:extent_size:free_capacity
0:ARR36P5N:online:2:7:333.9GB:32:292.4GB
1:ARR36P6N:online:2:1:400.8GB:32:400.3GB
2:ARR73P5N:online:2:1:746.5GB:32:746.0GB
3:IMAGE:online:0:0:0:32:0
```

As you can see, this command renamed the MDG from image_mode to `IMAGE`.

> **Note:** The **chmDiskgrp** command specifies the new name first. The name can consist of letters A to Z, a to z, numbers 0 to 9, and the underscore. It can be between one and 15 characters in length. However, it cannot start with a number or the word *mDiskgrp* because this prefix is reserved for SVC assignment only.

Deleting an mDisk group

Use the **svctask rmmdiskgrp** command to remove a MDG from the SVC cluster configuration:

```
IBM_2145:admin>svctask rmmdiskgrp IMAGE
```

This command removes the MDG `IMAGE` from the SVC configuration.

> **Note:** If there are mDisks within the MDG, you must use the -force flag, for example:
>
> ```
> svctask -force rmmdiskgrp IMAGE
> ```

Adding mDisks

If you created an empty MDG as we did, or you simply assign additional mDisks to your SVC environment at a later date, you can use the **svctask addmDisk** command to populate the MDG:

```
IBM_2145:admin>svctask addmDisk -mDisk ITSO_mDisk8 image_mode
```

You can *only* add unmanaged mDisks to a MDG. This command adds mDisk ITSO_mDisk8 to the MDG named image_mode.

> **Attention:** Do not do this if you want to create an image mode vDisk from the mDisk you are adding. As soon as you add an mDisk to a MDG, it becomes managed and extent mapping is not necessarily 1:1 anymore.

Removing mDisks

Use the **svctask rmmdisk** command to remove an mDisk from a MDG:

```
IBM_2145:admin>svctask rmmdisk -mDisk ITSO_mDisk8 IMAGE
```

This command removes the mDisk called ITSO_mDisk8 from the MDG named IMAGE.

> **Note:** If vDisks are using the mDisks you are removing from the MDG, you *must* use the -force flag:
>
> ```
> svctask rmmdisk -force -mDisk ITSO_mDisk8 IMAGE
> ```
>
> Even then, the removal only takes place if there is sufficient space to migrate the vDisk data to other extents on other mDisks which remain in the MDG. After you remove the mDisk group, it takes some time to change the mode from managed to unmanaged.

Showing mDisks in this group

Use the **svcinfo lsmdisk -filtervalue** command, as shown in Example 8-25, to see which mDisks are part of a specific MDG. This command shows all mDisks that are part of the MDG ess_mDiskgrp0.

Example 8-25 svcinfo lsmdisk -filtervalue: mDisks in MDG

```
IBM_2145:admin>svcinfo lsmdisk -filtervalue 'mDisk_grp_name=ess_mDiskgrp0' -delim :
id:name:status:mode:mDisk_grp_id:mDisk_grp_name:capacity:ctrl_LUN_#:controller_name
7:mDisk7:online:managed:1:ess_mDiskgrp0:4.7GB:0000000000000000:ESS
8:mDisk8:online:managed:1:ess_mDiskgrp0:4.7GB:0000000000000001:ESS
9:mDisk9:online:managed:1:ess_mDiskgrp0:4.7GB:0000000000000002:ESS
10:mDisk10:online:managed:1:ess_mDiskgrp0:4.7GB:0000000000000003:ESS
```

Showing vDisks using this group

Use the **svcinfo lsvdisk -filtervalue** command, as shown in Example 8-26, to see which vDisks are part of a specific MDG. This command shows all vDisks that are part of the MDG ess_mDiskgrp0.

Example 8-26 svcinfo lsvdisk -filtervalue: vDisks in MDG

```
IBM_2145:admin>svcinfo lsvdisk -filtervalue 'mDisk_grp_name=ess_mDiskgrp0' -delim:
id:name:IO_group_id:IO_group_name:status:mDisk_grp_id:mDisk_grp_name:capacity:type:FC_id:FC
_name:RC_id:RC_name
0:vDisk0:0:io_grp0:online:1:ess_mDiskgrp0:4.7GB:striped::::
```

You have now completed the tasks required to manage the disk controller systems, managed disks, and MDGs within an SVC environment.

8.4 Working with virtual disks

This section details the various configuration and administration tasks which can be performed on the vDisks within the SVC environment.

8.4.1 Hosts

This section explains the tasks that you can perform on a host level.

Host information

Use the **svcinfo lshost** command to display summary information about all hosts defined within the SVC environment. To display more detailed information about a specific host, run the command again and append the host name parameter (for example, `aix1`). Both of these commands are shown in Example 8-27.

> **Tip:** The -delim: parameter truncates the on-screen content and separates data fields with colons as opposed to wrapping text over multiple lines.

Example 8-27 svcinfo lshost command

```
IBM_2145:admin>svcinfo lshost
id              name            port_count
0               aix1            2
1               Win2K_1         2
2               redhat1         2
3               aix2            1
4               essai           1
5               essai2          1
6               essai1          1
7               ITSO_W2K        0
8               redhat2         0

IBM_2145:admin>svcinfo lshost aix1
id 0
name aix1
port_count 2
```

```
WWPN 10000000C9295A9A
node_logged_in_count 1
WWPN 10000000C9266F5B
node_logged_in_count 1
```

Creating a host

Before you create a host, you need to know that its host bus adapter (HBA) worldwide port names (WWPNs) are visible to the SVC. To do this, you issue the **svcinfo lshbaportcandidate** command as shown in Example 8-28.

Example 8-28 svcinfo lshbaportcandidate command

```
IBM_2145:admin>svcinfo lshbaportcandidate
id
210000E08B08AFD6
210100E08B28AFD6
```

After you know the WWPNs that are displayed match our host (use host or SAN switch utilities to verify), use the **svctask mkhost** command to create an image mode vDisk. The full syntax of this command is:

```
svctask mkhost [-name name] -hbawwpn wwpn_list [-force]
```

Note the following explanation:

- **name**: Name to be assigned to the host
- **hbawwpn**: List of HBA WWPNs to be added to host
- **force**: Force the creation using the user entered WWPNs

Note: If you do not provide the -name parameter, the SVC automatically generates the name *hostX* (where *X* is the ID sequence number assigned by the SVC internally).

If you want to provide a name, you can use letters A to Z, a to z, numbers 0 to 9, and the underscore. It can be between one and 15 characters in length. However, it cannot start with a number or the word *host* because this prefix is reserved for SVC assignment only.

The command to create a host is shown here:

```
IBM_2145:admin>svctask mkhost -name ITSO_W2K -hbawwpn 210000e08b08afd6
Host id [7] successfully created
```

This command creates a host called ITSO_W2K using WWPN `210000e08b08bafd6`.

Note: You can define a host with multiple ports by using the separator (:) between WWPNs:

```
svctask mkhost -name ITSO_W2K -hbawwpn 210000e08b08afd6:210100e08b28afd6
```

Or you can use the **addport** command, which we show later.

Perhaps your WWPN or WWPNs did not display when you issued the **svcinfo lshbaportcandidate** command, but you are *sure* your adapter is functioning (for example, you see WWPN in the switch name server) and your zones are correctly setup. In this case, you can type the WWPN of your HBA or HBAs and use the -force flag to create the host regardless, as shown here:

```
IBM_2145:admin>svctask mkhost -name ITSO_FORCE -hbawwpn 10000000c935b472 -force
Host id [8] successfully created
```

This command forces the creation of a host called ITSO_FORCE using WWPN `10000000C935b472`.

> **Note:** WWPNs are one of the few things within the CLI that are not case sensitive.

If you run the **`svcinfo lshost`** command again, you should now see your host.

Renaming a host

Use the **`svctask chhost`** command to change the name of a host. To verify change, run the **`svcinfo lshost`** command. Both of these commands are shown in Example 8-29.

Example 8-29 svctask chhost command

```
IBM_2145:admin>svctask chhost -name redhat2 ITSO_FORCE

IBM_2145:admin>svcinfo lshost
id              name            port_count
0               aix1            2
1               Win2K_1         2
2               redhat1         2
3               aix2            1
4               essai           1
5               essai2          1
6               essai1          1
7               ITSO_W2K        2
8               redhat2         1
```

As you can see, this command renamed the host from ITSO_FORCE to `redhat2`.

> **Note:** The **`chhost`** command specifies the new name first. The name can consist of letters A to Z, a to z, 0 to 9, and the underscore. It can be between one and 15 characters in length. However, it cannot start with a number or the word *host* because this prefix is reserved for SVC assignment only.

Deleting a host

Use the **`svctask rmhost`** command to delete a host from the SVC configuration. This command deletes the host called `redhat2` from the SVC configuration.

```
IBM_2145:admin>svctask rmhost redhat2
```

> **Note:** If no vDisks are assigned to the host, you must use the -force flag, for example:
>
> `svctask rmhost -force redhat2`

Adding ports

If you add a HBA to a server that is already defined within the SVC, you can use the **`svctask addhostport`** command to add WWPN definitions to it.

Before you add the new WWPN, you need to know that it is visible to the SVC. To do this, you issue the **`svcinfo lshbaportcandidate`** command as shown here:

```
IBM_2145:admin>svcinfo lshbaportcandidate
id
210100E08B28AFD6
```

After you know the WWPNs that are displayed match our host (use host or SAN switch utilities to verify), use the **svctask addhostport** command to add the port or ports to the host. The full syntax of this command is:

```
svctask addhostport -hbawwpn wwpn_list [-force] name|id
```

Note the following explanation:

- **hbawwpn**: The list of HBA WWPNs
- **force**: Indicates to force the system to use the provided WWPNs
- **namelid**: The name or ID of the host to which the ports are added

The command to add a host port is:

```
IBM_2145:admin>svctask addhostport -hbawwpn 210100e08b28afd6 ITSO_W2K
```

This command adds the WWPN of `210000e08b08bafd6` to the host `ITSO_W2K`.

> **Note:** You can add multiple ports at a time by using the separator (:) between WWPNs, for example:
>
> ```
> svctask addhostport -hbawwpn 210100e08b28afd6:210100e08c08afd6 ITSO_W2K
> ```

Perhaps your WWPN or WWPNs did not display when you issued the **svcinfo lshbaportcandidate** command, but you are **sure** your adapter is functioning (for example, you see WWN in the switch name server) and your zones are correctly setup. In this case, you can manually type the WWPN of your HBA or HBAs and use the -force flag to create the host regardless as shown here:

```
IBM_2145:admin>svctask addhostport -hbawwpn 10000000c935b623 -force redhat2
```

This command forces the addition of the WWPN `1000000c935b623` to the host called redhat2.

> **Note:** WWPNs are one of the few things within the CLI that are not case sensitive.

If you run the **svcinfo lshost** command again, you should see your host with an updated port count (2 in our example).

Example 8-30 svcinfo lshost command: port count

```
IBM_2145:admin>svcinfo lshost
id              name            port_count
0               aix1            2
1               Win2K_1         2
2               redhat1         2
3               aix2            1
4               essai           1
5               essai2          1
6               essai1          1
7               ITSO_W2K        2
8               redhat2         2
```

Deleting ports

If you make a mistake when adding or if you remove a HBA from a server that is already defined within the SVC, you can use the **svctask rmhostport** command to remove WWPN definitions from an existing host.

Before you remove the WWPN, be sure that it is the right one. To find this out, you issue the **svcinfo lshost HOST** command (our host is redhat2) as shown in Example 8-31.

Example 8-31 svcinfo lshost command

```
IBM_2145:admin>svcinfo lshost redhat2
id 8
name redhat2
port_count 2
WWPN 10000000C935B472
node_logged_in_count 0
WWPN 10000000C935B623
node_logged_in_count 0
```

When you know the WWPN, use the **svctask rmhostport** command to delete a host port. The full syntax of this command is:

```
svctask rmhostport -hbawwpn wwpn_list [-force] name|id
```

Note the following explanation:

- **hbawwpn**: List of HBA WWPNs
- **force**: Forces the system to use the provided WWPNS
- **namelid**: Name or ID of the host from which the ports are removed

The command to remove a Host port is:

```
IBM_2145:admin>svctask rmhostport -hbawwpn 210100e08b28afd6 ITSO_W2K
```

This command removes the WWPN of 210000e08b08bafd6 from host ITSO_W2K.

Note: You can remove multiple ports at a time by using the separator (:) between WWPNs, for example:

```
svctask rmhostport -hbawwpn 210100e08b28afd6:210100e08c08afd6 ITSO_W2K
```

Showing the vDisks mapped to this host

Use the **svcinfo lshostvdiskmap** command to show which vDisks are assigned to a specific host:

```
IBM_2145:admin>svcinfo lshostvdiskmap -delim : aix2
id:name:SCSI_id:vDisk_id:vDisk_name:wwpn:vDisk_UID
3:aix2:0:9:aix_vDisk1_tgt:10000000C925841D:600507680182001B200000000000000D
```

From this command, you can see that the host aix2 has only one vDisk called aix_vDisk1_tgt assigned.

Note: Although the -delim: flag normally comes at the end of the command string, in this case, you must specify this flag *before* the host name. Otherwise, it returns information for the host with the ID of 0 (aix1 in our case).

8.4.2 Virtual disks

This section explains the tasks that you can perform on a virtual disk level.

vDisk information

Use the **svcinfo lsvdisk** command to display summary information about all vDisks defined within the SVC environment. To display more detailed information about a specific vDisk, run the command again and append the vDisk name parameter (for example, aix_vDisk0). Both of these commands are shown in Example 8-32.

Example 8-32 svcinfo lsvdisk command

```
IBM_2145:admin>svcinfo lsvdisk -delim :
id:name:IO_group_id:IO_group_name:status:mDisk_grp_id:mDisk_grp_name:capacity:type:FC_id:FC
_name:RC_id:RC_name
0:aix_vDisk0:0:ITSO_grp0:online:0:ARR36P5N:10.0GB:striped::::
1:aix_vDisk1:0:ITSO_grp0:online:0:ARR36P5N:10.0GB:seq::::
2:aix_vDisk2:0:ITSO_grp0:online:0:ARR36P5N:10.0GB:seq::::
3:aix_vDisk3:0:ITSO_grp0:online:0:ARR36P5N:10.0GB:striped::::
4:windisk1:0:ITSO_grp0:online:0:ARR36P5N:200.0MB:striped::::
5:redhatdisk1:0:ITSO_grp0:online:0:ARR36P5N:300.0MB:striped::::
6:aix_vDisk4:0:ITSO_grp0:online:0:ARR36P5N:1.0GB:striped::::
7:rhat2_vDisk0:0:ITSO_grp0:online:1:ARR36P6N:500.0MB:striped::::
8:aix_vDisk1_src:0:ITSO_grp0:online:1:ARR36P6N:500.0MB:striped:0:flashmapaix1::
9:aix_vDisk1_tgt:0:ITSO_grp0:online:2:ARR73P5N:500.0MB:striped:0:flashmapaix1::

IBM_2145:admin>svcinfo lsvdisk aix_vDisk0
id 0
name aix_vDisk0
IO_group_id 0
IO_group_name ITSO_grp0
status online
mDisk_grp_id 0
mDisk_grp_name ARR36P5N
capacity 10.0GB
type striped
formatted no
mDisk_id
mDisk_name
FC_id
FC_name
RC_id
RC_name
throttling 0
preferred_node_id 12
```

Creating a vDisk

Use the **svctask mkvdisk** command to create an image mode vDisk. The full syntax of the **mkvdisk** command is:

```
mkvdisk -mDiskgrp name|id -iogrp name|id -size size [-fmtdisk]
        [-vtype seq|striped|image] [-node name|id] [-unit b|kb|mb|gb|tb|pb]
        [-mDisk name|id_list] [-name name]
```

Here the parameters are defined as follows:

- **-mDiskgrp**: Name or ID of the MDG in which to create the vDisk.

- **-iogrp** Name or ID of I/O group that is to own the vDisk.
- **-size**: Capacity (numerical); not needed for image mode vDisks.
- **-fmtdisk**: Optional parameter to force a format of the new vDisk.
- **-vtype**: Optional parameter to specify the type of vDisk (sequential, striped or image mode). Default (if nothing is specified) is striped.
- **-node**: Optional parameter to specify the name or ID of the preferred node. Default (if nothing is specified) is to alternate between nodes in the I/O group.
- **-unit**: Optional parameter to specify the data units for capacity parameter. Default (if nothing is specified) is MB.
- **-mDisk**: Optional parameter to specify the name or ID of the mDisk or mDisks to be used for the vDisk. This is only required for sequential and image mode vDisks because striped vDisks use all mDisks that are available in the MDG by default.

  Note: You can use this parameter for striped vDisks. For example, you may want to specify that the vDisk only uses a subset of the mDisks available within a MDG.

- **-name**: Optional parameter to assign a name to the new vDisk. Default (if nothing is specified) is to assign the name *vDiskX*, where *X* is the ID sequence number assigned by the SVC internally.

  Note: If you do not provide the -name parameter, the SVC automatically generates the name *vDiskX* (where *X* is the ID sequence number assigned by the SVC internally).

  If you want to provide a name, you can use letters A to Z, a to z, numbers 0 to 9, and the underscore. It can be between one and 15 characters in length. However, it cannot start with a number or the word *vDisk* because this prefix is reserved for SVC assignment only.

The command to create an vDisk is shown here:

```
IBM_2145:admin>svctask mkvdisk -mDiskgrp ARR73P5N -iogrp ITSO_grp0 -size 500 -name
SAP_DATA02
Host LUN, id [10], successfully created
```

This command creates a striped (default) vDisk called SAP_DATA02 of 500 MB (MB default). The vDisk belongs to the MDG ARR73P5N and is owned by the I/O group ITSO_grp0 (ITSO_grp0 is our only I/O group).

To create a sequential vDisk, simply insert the -vtype and -mDisk parameters as shown here:

```
IBM_2145:admin>svctask mkvdisk -mDiskgrp ARR73P5N -iogrp ITSO_grp0 -size 500 -vtype seq
-mDisk ITSO_mDisk5 -name SAP_DATA03
Host LUN, id [11], successfully created
```

This command creates a sequential vDisk called SAP_DATA03 of 500 MB (MB default). The vDisk is created on mDisk ITSO_mDisk5 within the MDG ARR73P5N and is owned by the I/O group ITSO_grp0 (ITSO_grp0 is our only I/O group).

Note: An entry of 1 GB uses 1024 MB.

Deleting a vDisk

Use the **svctask rmvDisk** command to delete a vDisk from your SVC configuration:

```
IBM_2145:admin>svctask rmvDisk SAP_DATA02
```

This command deletes vDisk SAP_DATA02 from the SVC configuration. If the vDisk is assigned to a host, you need to use the -force flag to delete the vDisk, for example:

```
svctask rmvDisk -force SAP_DATA02
```

Deleting a vDisk-to-host mapping

If you mapped a vDisk to a host by mistake, or you simply want to reassign the vDisk to another host, use the **svctask rmvDiskhostmap** command to unmap a vDisk from a host:

```
IBM_2145:admin>svctask rmvDiskhostmap -host redhat2 SAP_DATA03
```

This command unmaps the vDisk called SAP_DATA03 from the host `redhat2`.

Expanding a vDisk

Expanding a vDisk presents a larger capacity disk to your operating system. Although easily done using the SVC, you must ensure your operating systems supports expansion before using this function.

Assuming your operating system supports it, you can use the **svctask expandvdisksize** command to increase the capacity of a given vDisk. The full syntax of the **expandvdisksize** command is:

```
expandvdisksize -size size [-mDisk name|id_list] [-fmtdisk]
                [-unit b|kb|mb|gb|pb|tb] name|id
```

Note the following explanation:

- **size**: Capacity by which to expand
- **mDisk**: Disks to use as stripe set (optional for striped)
- **fmtdisk**: Format disk before use
- **unit**: Unit for capacity
- **nameid**: Name or ID of disk to be expanded

Important: Be very careful here. The format option formats the entire vDisk, not only the new extents.

A sample of this command is:

```
IBM_2145:admin>svctask expandvdisksize -size 1 -unit gb SAP_DATA02
```

This command expands the 500 MB SAP_DATA02 vDisk by 1 GB for a total of 1.5 GB.

Note: With sequential vDisks, you must specify the mDisk from which you want to obtain space. We do *not* support the expansion of image mode vDisks.

If there are not enough extents to expand your vDisk to the specified size, you receive an error message.

Mapping a vDisk to a host

To map a vDisk, use the **svctask mkvdiskhostmap** command:

```
IBM_2145:admin>svctask mkvdiskhostmap -host redhat2 SAP_DATA02
Host LUN to Host map, id [1], successfully created
```

This command maps the vDisk called SAP_DATA02 to the host called `redhat2`.

Modifying a vDisk

Use the **`svctask chvdisk`** command to change the name of a vDisk. The full syntax of the **`svctask chvdisk`** command is:

```
svctask chvdisk [-iogrp name|id] [-rate rate [-unitmb]] [-name name] [-force] name|id
```

Note the following explanation:

- **iogrp**: Name or ID of new I/O group
- **rate**: Throttling rate
- **unitmb**: Specify throttle rate in mb (default is ios)
- **name**: New name for disk
- **force**: Force the removal of the disk
- **namelid**: Existing name, or ID, of disk being changed

Changing the name of a vDisk is quite an obvious task. However, the I/O governing parameter is a new concept.

I/O governing

I/O governing effectively throttles the amount of I/Os per second (or MBs per second) that can be achieved to and from a specific vDisk. You may want to do this if you have a vDisk that has an access pattern, which adversely affects the performance of other vDisks on the same set of mDisks. For example, it uses most of available bandwidth.

Of course, if this application is highly important, then migrating the vDisk to another set of mDisks may be advisable. However, in some cases, it is an issue with the I/O profile of the application rather than a measure of its use or importance.

The choice between I/O and MB as the I/O governing throttle should be based on the disk access profile of the application. Database applications generally issue large amounts of I/O but only transfer a relatively small amount of data. In this case, setting an I/O governing throttle based on MBs per second does not achieve much. It is better to use an I/O per second throttle.

On the other extreme, a streaming video application generally issues a small amount of I/O, but transfers large amounts of data. In contrast to the database example, setting an I/O governing throttle based on I/Os per second does not achieve much, it is better to use a MBs per second throttle.

> **Note:** An I/O governing rate of 0 (displayed as *throttling* in CLI output of **`svcinfo lsvdisk`** command) does not mean that zero I/O per second (or MBs per second) can be achieved. It means that no throttle is set.

An example of the **`chvdisk`** command is shown here:

```
IBM_2145:admin>svctask chvdisk -name TSM_POOL SAP_DATA02

IBM_2145:admin>svctask chvdisk -rate 20 -unitmb TSM_POOL
```

> **Note:** The **`chvdisk`** command specifies the new name first. The name can consist of letters A to Z, a to z, numbers 0 to 9, and the underscore. It can be between one and 15 characters in length. However, it cannot start with a number or the word *vDisk* because this prefix is reserved for SVC assignment only.

The first command changes the name of the vDisk from SAP_DATA02 to `TSM_POOL` while the second command changes the vDisk throttling to 20 MB/sec.

Note: At the time of writing this redbook, these two commands cannot be run concurrently, for example:

```
svctask chvdisk -rate 20 -unit mb -name TSM_POOL SAP_DATA02
```

If you want to verify the changes, issue the **`svcinfo lsvdisk`** command, as shown in Example 8-33.

Example 8-33 svcinfo lsvdisk command: Verifying throttling

```
IBM_2145:admin>svcinfo lsvdisk TSM_POOL
id 10
name TSM_POOL
IO_group_id 0
IO_group_name ITSO_grp0
status online
mDisk_grp_id 0
mDisk_grp_name ARR36P5N
capacity 1.5GB
type striped
formatted no
mDisk_id
mDisk_name
FC_id
FC_name
RC_id
RC_name
virtual_disk_throttling (MB) 20
preferred_node_id 1
```

Migrating a vDisk

From time to time, you may want to migrate vDisks from one set of mDisks to another to retire an old disk subsystem, to have better balance performance across your virtualized environment, or simply to migrate data into the SVC environment transparently using image mode. To do so, use the **`svctask migratevdisk`** command. The full syntax of the command is:

```
migratevdisk -mDiskgrp name|id [-threads threads] -vDisk name|id
```

Note the following explanation:

- **mDiskgrp**: mDisk group name or ID
- **threads**: Number of threads
- **vDisk**: vDisk name or ID

Important: After migration is started, it continues to completion unless it is stopped or suspended by an error condition or the vDisk being migrated is deleted.

As you can see from the above parameters, before you can migrate your vDisk, you must know the name of our mDisk and the name of the MDG to which you want to migrate. To find the name, simply run the **`svcinfo lsvdisk`** and **`svcinfo lsmdiskgrp`** commands you learned earlier.

Restriction: Before you can migrate an image mode vDisk with a partial extent, you *must* change it into a managed disk.

When you know these details, you can issue the **migratevdisk** command as shown here:

```
IBM_2145:admin>svctask migratevdisk -mDiskgrp ARR73P5N -vDisk TSM_POOL
```

This command moves the vDisk TSM_POOL to the MDG ARR73P5N.

Note: If insufficient extents are available within your target MDG, you receive an error message. Make sure the source and target mDisk group have the same extent size.

The optional threads parameter allows you to assign a priority to the migration process. The default is 4, which is the highest priority setting. However, if you want the process to take a lower priority over other types of I/O, you can specify 3, 2, or 1.

You can run the **svcinfo lsmigrate** command at any time to see the status of the migration process. This is shown in Example 8-34.

Example 8-34 svcinfo lsmigrate command

```
IBM_2145:admin>svcinfo lsmigrate
migrate_type mDisk_Group_Migration
progress 0
migrate_source_vDisk_index 10
migrate_target_mDisk_grp 0
max_thread_count 4

IBM_2145:admin>svcinfo lsmigrate
migrate_type mDisk_Group_Migration
progress 36
migrate_source_vDisk_index 10
migrate_target_mDisk_grp 0
max_thread_count 4

IBM_2145:admin>svcinfo lsmigrate
migrate_type mDisk_Group_Migration
progress 86
migrate_source_vDisk_index 10
migrate_target_mDisk_grp 0
max_thread_count 4
```

Note: The progress is a percent complete.

Shrinking a vDisk

The method that the SVC uses to shrink a vDisk is to remove the required number of extents from the end of the vDisk. Depending on where your data actually resides on the vDisk, this can be quite destructive.

For example, you may have a vDisk that consists of 128 extents (0 to 127) of 16 MB (2 GB capacity) and you want to decrease the capacity to 64 extents (1 GB capacity). In this case, the SVC simply removes extents 64 to 127. Depending on the operating system, there is no easy way to ensure that your data resides entirely on extents 0 to 63. Therefore, you *may* lose data. Although easily done using the SVC, you must ensure your operating system supports shrinking, either natively or by using third-party tools, before using this function. In addition, we recommend that you *always* have a good up-to-date backup before you execute this task.

Assuming your operating system supports it, you can use the **svctask shrinkvdisksize** command to decrease the capacity of a given vDisk. The full syntax of the **svctask shrinkvdisksize** command is:

```
svctask shrinkvdisksize -size size [-unit b|kb|mb|gb|pb|tb] name|id
```

Note the following explanation:

- **size**: Capacity by which to shrink
- **unit**: Units applicable to capacity
- **name|id**: The name or ID of the disk to be shrunk

Here is an example of this command:

```
IBM_2145:admin>svctask shrinkvdisksize -size 500 TSM_POOL.
```

This command shrinks the vDisk TSM_POOL by 500 MB to a new total size of 1 GB.

Showing the mDisks

Use the **svcinfo lsvdiskmember** command to show which mDisks are used by a specific vDisk.

Example 8-35 svcinfo lsvdiskmember command

```
IBM_2145:admin>svcinfo lsvdiskmember TSM_POOL
id
0
2
```

If you want to know more about these mDisks, you can run the **svcinfo lsmdisk** command as explained in 8.3.2, “Managed disks” on page 174 (using the ID displayed above rather than the name).

Showing the mDisk group

Use the **svcinfo lsvdisk** command, as shown in Example 8-36, to show to which MDG a specific vDisk belongs.

Example 8-36 svcinfo lsvdisk command: MDG name

```
IBM_2145:admin>svcinfo lsvdisk TSM_POOL
id 10
name TSM_POOL
IO_group_id 0
IO_group_name ITSO_grp0
status online
mDisk_grp_id 0
mDisk_grp_name ARR36P5N
capacity 1.0GB
type striped
formatted no
mDisk_id
mDisk_name
FC_id
```

```
FC_name
RC_id
RC_name
virtual_disk_throttling (MB) 20
preferred_node_id 1
```

If you want to know more about these MDGs, you can run the **svcinfo lsmdiskgrp** command as explained in 8.3.3, "MDGs" on page 180.

Showing the hosts

To show the hosts to which a specific vDisk has been assigned, run the **svcinfo lsvdiskhostmap** command as shown in Example 8-37.

Example 8-37 svcinfo lsvdiskhostmap command

```
IBM_2145:admin>svcinfo lsvdiskhostmap -delim : TSM_POOL
id:name:SCSI_id:host_id:host_name:wwpn:vDisk_UID
10:TSM_POOL:1:8:redhat2:00100000000092546D:600507680182001B2000000000000001A
10:TSM_POOL:1:8:redhat2:10000000C923546D:600507680182001B2000000000000001A
```

This command shows the host or hosts to which the vDisk TSM_POOL was mapped.

> **Note:** Although the optional -delim : flag normally comes at the end of the command string, in this case you must specify this flag *before* the vDisk name. Otherwise, the command does not return any data.

You have now completed the tasks required to manage the hosts and vDisks within an SVC environment.

8.5 Managing Copy Services

See Chapter 10, "Copy Services: FlashCopy" on page 297, and Chapter 11, "Copy Services: Peer-to-Peer Remote Copy" on page 337, for more information.

8.6 Service and maintenance

This section details the various Service and Maintenance tasks that you can execute within the SVC environment.

8.6.1 PuTTY Secure Copy

Many of the tasks explained in this section require you to know how to use the PuTTY Secure Copy command to transfer data to and from your SVC cluster. PuTTY Secure Copy provides a file transfer mechanism for Secure Shell (SSH) to copy files either between two directories on the SVC configuration node, or between the configuration node and another host.

You must have appropriate permissions on the source and destination directories on your respective hosts to use the **pscp** command. Secure copy is available to you when you install an SSH client on your management workstation. It is already installed on the master console.

The PuTTY Secure Copy interface deliberately limits the permissions to the file systems inside the SVC. If you log on as `admin`, the only writable file system is /home/admin and similarly /home/service for the service login.

The PuTTY Secure Copy executable is command line driven. Therefore, to execute any of the following commands, you must either run the commands from the directory where the pscp.exe file is located or add the directory where it is located to the PATH variable. To modify the PATH variable, enter the following command from a command prompt:

```
C:\>set PATH=C:\Support Utils\PuTTY;%PATH%
```

Here C:\Support Utils\PuTTY is the directory location of the pscp.exe file.

Copying files from a remote to local system using PuTTY Secure Copy

Assume you want to copy a file called svcinfo.trc from the /dumps directory of your cluster. You want to copy this file from the cluster called `testsvc` to your local directory, d:\svc_dumps, where you name the file `svcinfo.txt`. Enter the following command:

```
C:\>pscp admin@testsvc:/dumps/svcinfo.trc d:\svc_dumps\svcinfo.txt
Authenticating with public key "rsa-key-20030514"
svcinfo.txt               |          3 kB |   3.1 kB/s | ETA: 00:00:00 | 100%
```

Copying files from a local to remote system using PuTTY Secure Copy

Assume you want to copy a file called software_upgrade.pkg from the local directory d:\svc_code (on your management workstation) to the upgrade directory on the machine called `testsvc`. Enter the following command:

```
C:\pscp d:\svc_code\software_upgrade.pkg admin@testsvc:/upgrade
Authenticating with public key "rsa-key-20030514"
software_upgrade.pkg |          1 kB |   1.0 kB/s | ETA: 00:00:00 | 100%
```

8.6.2 Upgrading software

This section explains how to upgrade the SVC software.

Package numbering and version

The format for software upgrade packages is three positive integers separated by dots. For example, a software upgrade package contains something similar to `1.15.0`.

Each software package is given a unique number. Two software packages are fully compatible if the first two numbers are identical. For example, 1.2.15 and 1.2.16 are fully compatible. However, 1.2.15 and 1.3.0 are not compatible.

If two software packages are fully compatible, then you can upgrade or downgrade between these packages. If two software packages are not fully compatible, then the new software upgrade package contains code to upgrade from a list of previous software levels. In this case, you can apply the software upgrade only if the cluster is currently running one of the previous software levels. After you apply it, you cannot remove it.

Precaution before upgrade

Before you upgrade the SVC software, ensure that all I/O paths between all hosts and SANs are working. Otherwise, the applications may have I/O failures during the software upgrade.

You must also ensure that all I/O paths are working for each host that is running I/O operations to the SANs during the software upgrade. You can check the I/O paths by using **`datapath query`** commands. *See IBM TotalStorage ESS, SAN Volume Controller, SAN Volume Controller for Cisco*, SC26-7608, for more information about **`datapath query`** commands. You do not need to check for hosts that have no active I/O operations to the SANs during the software upgrade.

Procedure

To upgrade the SVC cluster software, perform the following steps:

1. Upload the new software package using PuTTY Secure Copy (pscp). Enter the following command:

```
C:\pscp d:\svc_code\030516_full_tgz.gpg admin@REDSTONE1:/upgrade
Authenticating with public key "rsa-key-20030514"
030516_full_tgz.gpg      |     80394 kB |   2063.7 kB/s | ETA: 00:00:12 | 92%
```

2. Check the package was successfully delivered through the PuTTY command line application by entering the **svcinfo lssoftwaredumps** command:

```
IBM_2145:admin>svcinfo lssoftwaredumps
id              software_filename
0               030516_full.tgz.gpg
```

3. Now that the package is uploaded, use the hidden **svcservicetask** command set to apply the software upgrade as shown here:

```
IBM_2145:admin>svcservicetask applysoftware -file 030516_full.tgz.gpg
```

4. The new code is distributed and applied to each node in the SVC cluster. After installation, each node is automatically restarted in turn. If a node does not restart automatically during the upgrade, repair or manually delete that node from the cluster to complete the back out process.

5. Eventually both nodes should display `Cluster:` on line one on the SVC front panel and the name of your cluster on line 2. Be prepared for a long wait (we waited up to 40 minutes in the upgrade).

> **Note:** During this process, both your CLI and GUI vary from sluggish (very slow) to unresponsive. The important thing is that I/O to hosts can continue.

6. To verify that the upgrade was successful, you can perform either of the following options:
   - Run the **svcinfo lscluster** and **svcinfo lsnodevpd** commands as shown in Example 8-38. The **nodevpd** information is truncated for purposes of this example. The real output is much longer.

Example 8-38 svcinfo lscluster and lsnodevpd commands

```
IBM_2145:admin>svcinfo lscluster REDSTONE1
id 0000020060C006B4
name REDSTONE1
location local
partnership
bandwidth
cluster_IP_address 9.42.164.155
cluster_service_IP_address 9.42.164.160
total_mDisk_capacity 2000.0GB
space_in_mDisk_grps 765.2GB
space_allocated_to_vDisks 34.3GB
total_free_space 1965.7GB
statistics_status off
statistics_frequency 15
required_memory 4096
cluster_locale en_US
```

```
SNMP_setting none
SNMP_community
SNMP_server_IP_address 0.0.0.0
subnet_mask 255.255.255.0
default_gateway 9.42.164.1
time_zone 522 UTC
email_setting none
email_id
code_level 0.6.03051600
FC_port_speed 2Gb

IBM_2145:admin>svcinfo lsnodevpd REDSTONE1_node1
software: 5 fields
code_level 0.6.03051600
node_name REDSTONE1_node1
ethernet_status 1
WWNN 0x500507680100035a
id 1

IBM_2145:admin>svcinfo lsnodevpd REDSTONE1_node2
software: 5 fields
code_level 0.6.03051600
node_name REDSTONE1_node2
ethernet_status 1
WWNN 0x5005076801000364
id 2
```

– Copy the error log to your management workstation as explained in 8.6.3, “Running maintenance procedures” on page 198. Open it in WordPad and search for `Software Install completed`.

You have now completed the tasks required to upgrade the SVC software.

8.6.3 Running maintenance procedures

Use the **svctask finderr** command to generate a list of any unfixed errors in the system. This command analyzes the last generated log that resides in the `/dumps/elogs/` directory on the cluster.

If you want to generate a new log before analyzing for unfixed errors, run the **svctask dumperrlog** command:

```
IBM_2145:admin>svctask dumperrlog
```

This generates a file called errlog_timestamp, such as errlog_000667_030527_182350, where:

- **errlog** part of the default prefix for all error log files
- **000667** is the panel name of the current configuration node
- **030527** is the date (YYMMDD)
- **182350** is the time (HHMMSS)

If you want, you can add the -prefix parameter to your command to change the default prefix of errlog to something else, for example:

```
svctask dumperrlog -prefix svcerrlog
```

This command creates a file called svcerrlog_timestamp.

After you generate your error log, you can issue the **svcinfo finderr** command to scan it for any unfixed errors as shown here:

```
IBM_2145:admin>svctask finderr
Highest priority unfixed error code is [2030]
```

As you can see, we have one unfixed error on our system.

To know more about this unfixed error, you need to look at the error log in more detail. Use the PuTTY Secure Copy process to copy the file from the cluster to your local management workstation as shown in Example 8-39.

Example 8-39 pscp command: Copy error logs off SVC

```
C:\Support Utils\PuTTY>pscp admin@REDSTONE1:/dumps/elogs/errlog_000667_030527_182350
c:/SVC_Dumps/errlog.txt
Authenticating with public key "rsa-key-20030514"
errlog.txt                |        367 kB | 367.1 kB/s | ETA: 00:00:00 | 100%
```

This command copies the file called errlog_000667_030527_182350 to the C:\SVC_Dumps directory on our local workstation and call the file errlog.txt.

Open the file in WordPad (Notepad does not format the screen as well). You should see the information similar to Example 8-40. The list was truncated for purposes of this example.

Example 8-40 errlog in WordPad

```
//-------------------
// Error Log Entries
//-------------------


Error Log Entry 0
  Node Identifier       : ITSO_node2
  Object Type           : cluster
  Object ID             : 0
  Sequence Number       : 408
  Root Sequence Number  : 408
  First Error Timestamp : Mon May 26 17:11:44 2003
                        : Epoch + 1053983504
  Last Error Timestamp  : Mon May 26 17:11:44 2003
                        : Epoch + 1053983504
  Error Count           : 1
  Error ID              : 981001 : Cluster Fabric View updated by fabric discovery
  Error Code            :
  Status Flag           : SNMP trap raised
  Type Flag             : INFORMATION

CA 17 00 00 63 02 00 00 02 00 00 00 02 00 00 00
05 00 00 00 03 00 00 00 0B 00 00 00 EE 03 02 00
00 00 00 00 12 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
```

Scrolling through, or searching for the term “unfixed” (without quotation marks), you should find more detail about the problem.

After you take the necessary steps to rectify the problem, you can mark the error as fixed in the log by issuing the **svctask cherrstate** command against its sequence number:

```
IBM_2145:admin>svctask cherrstate -sequencenumber 591
```

If you accidentally mark the wrong error as `fixed`, you can mark it as `unfixed` again by entering the same command and appending the -unfix flag to the end, for example:

```
svctask cherrstate -sequencenumber 591 -unfix
```

> **Note:** The **-unfix** command did not work at the time of writing this redbook.

8.6.4 Maintaining SSH keys

To maintain SSH keys, perform the following steps:

1. Obtain a list of keys stored on the SVC cluster by issuing the **svcinfo lssshkeys** command:

   ```
   IBM_2145:admin>svcinfo lssshkeys -user all
   id              userid_key_identifier
   1 admin admin
   2 service service
   3 admin itso0513
   ```

   This command shows *all* SSH keys that are currently stored in the cluster.

2. You can either delete the keys or add keys.

   – To delete keys use the **svctask rmsshkey** command as shown here:

     ```
     IBM_2145:admin>svctask rmsshkey -user admin -key itso0513
     ```

     This command removes the itso0513 key associated with the admin user.

     Alternatively you can enter the **svctask rmallsshkeys** command to remove all keys as shown here:

     ```
     svctask rmallsshkeys -user admin|service|all
     ```

     Here, *user* indicates to whom the key applies. An example of this command is shown here:

     ```
     IBM_2145:admin>svctask rmallsshkeys -user all
     ```

     This command removes *all* SSH keys from the SVC environment.

     > **Important:** Normally, deleting all IDs is considered if you believe security has been compromised and that unauthorized personnel have access to admin or service IDs. If you delete all IDs, you *must* generate a fresh pair and upload the pair into your cluster to administer the SVC environment again. If you did not enable the *Password reset via service panel* option, the only way to gain access to your cluster (to upload the new key) is to rebuild the cluster using the service panel. This destroys *all* data.
     >
     > For this reason, if you find yourself in a situation where security has been compromised, we recommend that you generate a fresh pair of keys and upload the public half before you *selectively* delete all other IDs.

– To add new keys, copy the new key to the /tmp directory on the server using PuTTY Secure Copy as shown here:

```
C:\>pscp c:\Support Utils\PuTTY\itsopub admin@REDSTONE1:/tmp
Authenticating with public key "rsa-key-20030514"
itsopub                  |          0 kB |   0.3 kB/s | ETA: 00:00:00 | 100%
```

For information about how to generate a SSH key pair, see 4.2.1, “Generating public and private SSH key pair using PuTTY” on page 70. For information about how to use PuTTY Secure Copy, see 8.6.1, “PuTTY Secure Copy” on page 195.

After you copy the file over to your cluster, you can use the **svctask addsshkey** command as shown here:

```
IBM_2145:admin>svctask addsshkey -label itsoadmin -file /tmp/itsopub -user admin
```

This command adds the key file *itsopub* from the /tmp directory and associates it with the admin user called *itsoadmin*.

8.6.5 Setting up error notification

To set up error notification, use the **svctask setevent** command. The full syntax of the **setevent** command is:

```
setevent [-snmptrap all|no_state|none] [-snmpip ip_address]
        [-community community] [-alert all|no_state|none] [-email email_address]
```

Note the following explanation:

- **snmptrap**: When to raise a trap
- **snmpip**: IP address of host running SNMP
- **community**: SNMP community
- **alert**: When to raise an e-mail notification
- **email**: E-mail address for notifications

An example of the **setevent** command is shown here:

```
IBM_2145:admin>svctask setevent -snmptrap all -snmpip 9.42.164.160 -community SVC -alert
no_state -email support@ibm.com
```

This command sends *all* events (errors and changes in state) to the SVC community on the SNMP manager with the IP address 9.42.164.160. This command also sets e-mail alerting to forward error events (no state changes) to the e-mail address support@ibm.com (fictitious) by setting the -alert parameter to no_state.

Note: At the time of writing this redbook, the no_state parameter was invalid through the CLI.

8.6.6 Analyzing the error log

The following types of events and errors are logged in the error log:

- **Events**: State changes that are detected by the cluster software and that are logged for informational purposes. Events are recorded in the cluster error log.
- **Errors**: Hardware or software problems that are detected by the cluster software and that require some repair. Errors are recorded in the cluster error log.
- **Unfixed errors**: Errors that were detected and recorded in the cluster error log and that have not yet been corrected or repaired.

- **Fixed errors**: Errors that were detected and recorded in the cluster error log and that have subsequently been corrected or repaired.

To display the error log, use the **svcinfo lserrlog** or **svcinfo caterrlog** commands as shown in Example 8-41 (output is the same).

Example 8-41 svcinfo caterrlog command

```
IBM_2145:admin>svcinfo caterrlog -delim :
id:type:fixed:SNMP_trap_raised:error_type:node_name:sequence_number:root_sequence_number:fi
rst_timestamp:last_timestamp:number_of_errors:error_code
0:cluster:no:yes:6:ITSO_node2:406:406:030526195944:030526195944:1:00981001
0:cluster:no:yes:6:ITSO_node1:407:406:030526195959:030526195959:1:00981001
0:unknown:no:no:5:ITSO_node2:0:0:030526195940:030526195940:0:00990221
0:unknown:no:no:5:ITSO_node2:0:0:030526195944:030526195944:0:00990219
9:node:no:no:5:ITSO_node2:0:0:030526201513:030526201513:0:00990387
9:node:no:no:5:ITSO_node2:0:0:030526202635:030526202635:0:00990386
9:node:no:no:5:ITSO_node2:0:0:030526202729:030526202729:0:00990387
9:node:no:no:5:ITSO_node2:0:0:030526202748:030526202748:0:00990386
254:rc_const_grp:no:no:5:ITSO_node2:0:0:030526203516:030526203516:0:00990240
0:unknown:no:no:5:ITSO_node2:0:0:030526210137:030526210137:0:00990220
```

This command views the error log that was last generated. It shows that nine events are logged. Use the method described in 8.6.3, “Running maintenance procedures” on page 198, to upload and analyze the error log in more detail.

To clear the error log, you can issue the **svctask clearerrlog** command as shown here:

```
IBM_2145:admin>svctask clearerrlog
Do you really want to clear the log?  y
```

This clears existing errors. However, if you have any unfixed errors, you must force the clearing of the log using the -force flag.

8.6.7 Setting features

To change the licensing feature settings, use the **svctask chlicense** command. The full syntax of the **svctask chlicense** command is:

```
svctask chlicense [-flash on|off] [-remote on|off] [-size capacity]
```

Note the following explanation:

- **flash**: Enable/disable FlashCopy
- **remote**: Enable/disable Peer-to-Peer Remote Copy (PPRC)
- **size**: Set licensed capacity (in GBs)

Before you change the licensing, see what license you already have by issuing the **svcinfo lslicense** command as shown in Example 8-42.

Example 8-42 svcinfo lslicense command

```
IBM_2145:admin>svcinfo lslicense
feature_flash on
feature_remote off
feature_num_gb 2000
```

Consider, for example, that you have purchased an additional 4 TB of licensing and the PPRC premium feature. The command you need to enter is shown here:

```
IBM_2145:admin>svctask chlicense -remote on

IBM_2145:admin>svctask chlicense -size 6000
```

The first command turns the remote copy feature on. The second command changes the licensed capacity to 6000 GB (6 TB) which is 4 TB more than what it was.

> **Note:** At the time of writing this redbook, these two commands cannot be entered concurrently, for example:
>
> `svctask chlicense -remote on -size 6000`

To verify that the changes you made are reflected in your SVC configuration, you can issue the **svcinfo lslicense** command as before. See Example 8-43.

Example 8-43 svcinfo lslicense command: Verifying changes

```
IBM_2145:admin>svcinfo lslicense
feature_flash on
feature_remote on
feature_num_gb 6000
```

8.6.8 Viewing the feature log

To view the feature log using the CLI, you must first create a feature log dump. Then copy the feature log to your management workstation using PuTTY Secure Copy. Finally, open the file in WordPad.

To create the feature log dump, enter the **svctask dumpinternallog** command as shown here:

```
IBM_2145:admin>svctask dumpinternallog
```

This creates a file called feature.txt in the /dumps/feature directory on the cluster. To see whether creation was successful, you can enter the **svcinfo lsfeaturedumps** command, as shown here to see that the file was created:

```
IBM_2145:admin>svcinfo lsfeaturedumps
id              feature_filename
0               feature.txt
```

> **Note:** Only one of these files exists. Therefore, each time you run the **dumpinternallog** command, it overwrites any existing feature.txt file.

Now that you created the file, you must copy it to your management workstation using PuTTY Secure Copy as shown here:

```
C:\PuTTY>pscp admin@REDSTONE1:/dumps/feature/feature.txt c:\svc_dumps\feature.txt
Authenticating with public key "rsa-key-20030514"
feature.txt               |         18 kB |  18.5 kB/s | ETA: 00:00:00 | 100%
```

Now open the file in WordPad (Notepad does a poor job formatting) to view the output. It should look similar to Example 8-44. The output list was truncated for purposes of this example.

Example 8-44 Feature dump in WordPad

```
//---------------------
//---------------------
// Feature Log Entries
//---------------------
time     type     value0   value1   value2   value3   value4   value5
3ebfe92f 00000001 00000000 00000000 00000000 00000000 00000000 00000000
3ebfe92f 00000003 00000000 00000000 00000000 00000000 00000000 00000000
3ebfe92f 00000005 00000000 00000000 00000400 00000000 00000000 00000000
3ecb9ddf 00000005 00000000 00000400 000007d0 0000015c 00000000 00000000
3ed286b8 00000003 00000000 00000000 00000000 00000000 00000000 00000000
3ed286bd 00000005 00000000 000007d0 00001770 0000002e 00000000 00000000
00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000
00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000
00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000
00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000
........
00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000
00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000
30400364 00000100
aa226003 f9e4c5f1 829a130c 6d12be12
```

8.6.9 Dumping the configuration

To dump configuration information, use the **svctask dumpconfig** command as shown here:

```
IBM_2145:admin>svctask dumpconfig -prefix REDSTONE1
The configuration data has been written to REDSTONE1_000667_030527_184138
```

This command creates a configuration dump with the prefix of REDSTONE1 and places it in the /dumps/configs directory on the cluster. The SVC automatically appends a time stamp to the file name specified as shown where:

- 000667 is the panel name of the current configuration node.
- 030527 is the date (YYMMDD).
- 184138 is the time (HHMMSS).

Now that the configuration is dumped, you can copy it to your management workstation using PuTTY Secure Copy as shown in Example 8-45.

Example 8-45 pscp: Copying configuration dumps

```
C:\PuTTY>pscp admin@REDSTONE1:/dumps/configs/REDSTONE1_000667_030527_184138
c:\svc_dumps\redstone1_030527.txt
Authenticating with public key "rsa-key-20030514"
redstone1_030527.txt          |         17 kB |  17.6 kB/s | ETA: 00:00:00 | 100%
```

Open the file in WordPad (Notepad does not format well) to view the output. It should look similar to Example 8-46. The output was truncated for purposes of this example.

Example 8-46 config dump in WordPad

```
cluster_view_array[0] =
{
  name       : REDSTONE1
  unique_id  : 10030400364
  state      : ICOC_OBJ_STATE_UNINITIALISED
  used       : TRUE
}



controller count = 1
device_view_array[0] =
{
  name                        : fastt600
  product_serial_number_array :
  vendor_id                   : 5278866513604911136
  product_id_low              : 3546358422665834544
  product_id_high             : 2314885530818453536
  used                        : TRUE
}



host count = 9
  name             : aix1
  port_count       : 2
  state            : ICOC_OBJ_STATE_OFFLINE
  used             : TRUE
}
  name             : Win2K_1
  port_count       : 2
  state            : ICOC_OBJ_STATE_OFFLINE
  used             : TRUE
```

8.6.10 Listing dumps

Several commands are available for you to list the dumps that were generated over a period of time. You can use the **lsxxxxdumps** command, where *xxxx* is the object dumps, to return a list of dumps in the appropriate directory.

The dumps in the SVC are contained in the following directory structure:

- /dumps
- /dumps/configs
- /dumps/elogs
- /dumps/feature
- /dumps/iostats
- /dumps/iotrace

Software upgrade packages are contained in the /home/admin/upgrade directory. These directories exist on every node in the cluster.

Configuration dump

Dumps contained in the /dumps/configs directory are dumps of the cluster configuration data. A configuration dump is created by using the **svctask dumpconfig** command. This dumps the configuration of the cluster, including all object details, to the /dumps/configs directory. If no filename prefix is supplied, the default config_ is used. The full, default file name is config_*NNNNNN*_YYMMDD_HHMMSS. Here *NNNNNN* is the node front panel name. If the command is used with the -prefix option, then the value entered for the -prefix is used instead of *config*.

The command to list all dumps in the /dumps/configs directory is **svcinfo lsconfigdumps**.

Error or event dump

Dumps contained in the /dumps/elogs directory are dumps of the contents of the error and event log at the time that the dump was taken. An error or event log dump is created by using the **svctask dumperrlog** command. This dumps the contents of the error or event log to the /dumps/elogs directory. If no filename prefix is supplied, the default errlog_ is used. The full, default file name is errlog_*NNNNNN*_YYMMDD_HHMMSS. Here *NNNNNN* is the node front panel name. If the command is used with the -prefix option, then the value entered for the -prefix is used instead of `errlog`.

The command to list all dumps in the /dumps/elogs directory is **svcinfo lserrlogdumps**.

Featurization log dump

Dumps contained in the /dumps/feature directory are dumps of the featurization log. A featurization log dump is created by using the **svctask dumpinternalllog** command. This dumps the contents of the featurization log to the /dumps/feature directory to a file called feature.txt. Only one of these files exists, so every time the **svctask dumpinternallog** command is run, this file is overwritten.

The command to list all dumps in the /dumps/feature directory is **svcinfo lsfeaturedumps**.

I/O statistics dump

Dumps contained in the /dumps/iostats directory are dumps of the I/O statistics for disks on the cluster. An I/O statistics dump is created by using the **svctask startstats** command. As part of this command, you can specify a time interval at which you want the statistics to be written to the file (the default is 15 minutes). Every time the time interval is encountered, the I/O statistics that are collected up to this point are written to a file in the /dumps/iostats directory. The file names used for storing I/O statistics dumps are m_stats_*NNNNNN*_YYMMDD_HHMMSS, or v_stats_*NNNNNN*_YYMMDD_HHMMSS, depending on whether the statistics are for mDisks or vDisks. Here *NNNNNN* is the node front panel name.

The command to list all dumps in the /dumps/iostats directory is **svcinfo lsiostatsdumps**.

I/O trace dump

Dumps contained in the /dumps/iotrace directory are dumps of I/O trace data. The type of data that is traced depends on the options specified by the **svctask settrace** command. The collection of the I/O trace data is started by using the **svctask starttrace** command. The I/O trace data collection is stopped when the **svctask stoptrace** command is used. When the trace is stopped, the data is written to the file. The file name is *prefix*_*NNNNNN*_YYMMDD_HHMMSS. Here *NNNNNN* is the node front panel name, and *prefix* is the value entered by the user for the -filename parameter in the **svctask settrace** command.

The command to list all dumps in the /dumps/iotrace directory is **svcinfo lsiotracedumps**.

Application abends dump

Dumps contained in the /dumps directory are dumps resulting from application abends. Such dumps are written to the /dumps directory. The default file names are dump.*NNNNNN*.YYMMDD.HHMMSS. Here *NNNNNN* is the node front panel name. In addition to the dump file, it is possible that there may be some trace files written to this directory. These are named NNNNNN.trc.

The command to list all dumps in the /dumps directory is **`svcinfo ls2145dumps`**.

Software dump

The final option available in the **`svcinfo lsxxxxdumps`** command series is the **`svcinfo lssoftwaredumps`** command. This command lists the contents of the /home/admin/upgrade directory. Any files in this directory are copied there at the time that you want to perform a software upgrade.

Example 8-47 shows these commands.

Example 8-47 Listing dumps

```
IBM_2145:admin>svcinfo lsconfigdumps
id              config_filename
0               REDSTONE1_000683_030526_173553
1               REDSTONE1_000683_030526_173949

IBM_2145:admin>svcinfo lserrlogdumps
id              filename
0               errlog_000683_030523_120658
1               errlog_000683_030526_170137
2               errlog_000683_030526_171618
3               errlog_000683_030526_171725

IBM_2145:admin>svcinfo lsfeaturedumps
id              feature_filename
0               feature.txt

IBM_2145:admin>svcinfo lsiostatsdumps
id              iostat_filename
0               m_stats_000683_030522_145411
1               v_stats_000683_030522_145412
2               m_stats_000683_030522_150911
3               v_stats_000683_030522_150912
4               v_stats_000683_030522_152412
5               m_stats_000683_030522_152412

IBM_2145:admin>svcinfo ls2145dumps
id              2145_filename
0               dump.000683.030526.173553
1               dump.000683.030526.173949
2               000683.trc

IBM_2145:admin>svcinfo lsiotracedumps
id              iotrace_filename
0               vDisktrace_000683_030527_162005

IBM_2145:admin>svcinfo lssoftwaredumps
id              software_filename
0               030515_full.tgz.gpg
```

Other node dumps

All of the **`svcinfo lsxxxxdumps`** commands can accept a node identifier as input (for example, append the node name to the end of any of the above commands). If this identifier is not specified, then the list of files on the current configuration node (in our case ITSO_node2) is displayed. If the node identifier is specified, then the list of files on that node is displayed.

However, files can only be copied from the current configuration node (using PuTTY Secure Copy). Therefore, you must issue the **svctask cpdumps** command to copy the files from a non-configuration node to the current configuration node. Subsequenty you can copy them to the management workstation using PuTTY Secure Copy.

For example, you issue the **svcinfo lsconfigdumps ITSO_node1** command. You discover a dump file and want to copy it to your management workstation for further analysis. In this case, you must first copy the file to your current configuration node.

To copy dumps from other nodes to the configuration node, use the **svctask cpdumps** command. The full syntax of the **svctask cpdumps** command is:

```
svctask cpdumps -prefix prefix name|id
```

Note the following explanation:

- **prefix**: Directory and or files to retrieve
- **name|id**: Name or ID of node from which the dumps must be retrieved

The prefix you enter depends on which dumps you hope to retrieve from the remote host. The valid -prefix directories are:

- /dumps
- /dumps/iostats
- /dumps/iotrace
- /dumps/feature
- /dumps/configs
- /dumps/elogs

An sample of the command is shown here:

```
IBM_2145:admin>svctask cpdumps -prefix /dumps/configs ITSO_node1
```

Now that you copied the configuration dump file from ITSO_node1 to your configuration node, you can use PuTTY Secure Copy to copy the file to your management workstation for further analysis as described earlier.

To clear the dumps, you can run the **svctask cleardumps** command. Again, you can append the node name if you want to clear dumps off a node other than the current configuration node (the default for the **svctask cleardumps** command). The full syntax of the command is:

```
svctask cleardumps -prefix prefix [name|id]
```

Note the following explanation:

- **prefix**: Directory or file filter
- **name|id**: Name or ID of the node

Here, -prefix must be one of these directories:

- /dumps
- /dumps/iostats
- /dumps/iotrace
- /dumps/feature
- /dumps/configs
- /dumps/elogs

The commands in Example 8-48 clear all logs or dumps from the ITSO_node1 node.

Example 8-48 svctask cleardumps command

```
IBM_2145:admin>svctask cleardumps -prefix /dumps ITSO_node1
IBM_2145:admin>svctask cleardumps -prefix /dumps/iostats ITSO_node1
IBM_2145:admin>svctask cleardumps -prefix /dumps/iotrace ITSO_node1
IBM_2145:admin>svctask cleardumps -prefix /dumps/feature ITSO_node1
IBM_2145:admin>svctask cleardumps -prefix /dumps/configs ITSO_node1
IBM_2145:admin>svctask cleardumps -prefix /dumps/elogs ITSO_node1
```

8.7 Tracing a host disk back to its source physical disk

You may find these steps useful:

1. On your host, run the **datapath query device** command from your host. You see a long disk serial number for each vpath device as shown in Example 8-49.

Example 8-49 datapath query device

```
C:\Program Files\IBM\Subsystem Device Driver>datapath query device

Total Devices : 1


DEV#:   0  DEVICE NAME: Disk3 Part0  TYPE: 2145       POLICY: OPTIMIZED
SERIAL: 600507680183001AD00000000000000B
============================================================================
Path#              Adapter/Hard Disk    State     Mode     Select     Errors
    0    Scsi Port4 Bus0/Disk3 Part0     OPEN   NORMAL          4          0
    1    Scsi Port5 Bus0/Disk3 Part0     OPEN   NORMAL          3          0
    2    Scsi Port5 Bus0/Disk3 Part0     OPEN   NORMAL          0          0
    3    Scsi Port4 Bus0/Disk3 Part0     OPEN   NORMAL          0          0
```

2. Run the **svcinfo lshostvdiskmap** command to return a list of all assigned vDisks.

   ```
   IBM_2145:admin>svcinfo lshostvdiskmap -delim : W2K
   id:name:SCSI_id:vDisk_id:vDisk_name:wwpn:vDisk_UID
   0:W2K:0:0:vDisk0:21000000B3474536:600507680183001AD00000000000000B
   ```

   Look for the disk serial number that matches your **datapath query device** output. This host was defined in our SVC as W2K.

3. Run the **svcinfo lsvdiskmember vDiskname** command for a list of the mDisk or mDisks that make up the specified vDisk:

   ```
   IBM_2145:admin>svcinfo lsvdiskmember vDisk0
   id
   0
   ```

4. Query the mDisks with the **svcinfo lsmdisk mDiskID** to find their controller and LUN number information as shown in Example 8-50. The output displays the controller name and the controller LUN ID, which should be enough (provided you named your controller something unique such as a serial number) to track back to a LUN within the disk subsystem.

Example 8-50 svcinfo lsmdisk command

```
IBM_2145:admin>svcinfo lsmdisk 0
id 0
name mDisk0
status online
mode managed
mDisk_grp_id 0
mDisk_grp_name mDiskgrp0
capacity 100.0GB
quorum_index 0
block_size 512
controller_name F600_231436
ctrl_type 4
ctrl_WWNN 200400A0B80CEB7C
controller_id 0
path_count 2
max_path_count 2
ctrl_LUN_# 0
```

8.8 SVC cluster configuration backup and recovery

The SVC configuration data is stored on all the nodes in the cluster. It is specially hardened so that, in normal circumstances, the SVC should never lose its configuration settings. However in exceptional circumstances, such as a rogue fire sprinkler soaking the SVC cluster or a multiple hardware failure, this data may become corrupted or lost.

This section details the tasks that you can perform to save the configuration data from a SVC configuration node and restore it. The following configuration information is backed up:

- SVC cluster
- Storage controllers
- Hosts
- I/O groups
- Software licenses
- Managed disks
- MDGs
- SVC nodes
- SSH keys
- Virtual disks
- vDisk-to-host mappings

> **Important:** Before you begin the restore process, consult IBM Support to determine the cause as to why you cannot access your original configuration data. After the restore process starts, the original data on the vDisks is destroyed. Therefore, you must ensure that you have a backup of all user data on the vDisks.

The SVCCONFIG command line tool is a script, used under CLI provided, to save and restore configuration data. It uses secure communications to communicate with a configuration node. The tool is designed to work if the hardware configuration for restoration is identical to that during saving.

The prerequisites for having a successful backup are:

- All nodes in the cluster must be online.
- No object name may begin with an underscore (_).

- Do not run any independent operations that could change the cluster configuration while the backup command runs.
- Do not make any changes to the fabric or the cluster between backup and restore. If changes are made, back up your configuration again or you may not be able to restore it later.

> **Note:** We recommend that you make a backup of the SVC configuration data after each major change in the environment, such as defining or changing a vDisks, vDisk-to-host mappings, etc. In addition, you can make a backup after each change. Be aware that only *two versions* of the backup file are maintained for each cluster (the previous one has .bak appended), unless you copy the XML or XML BAK files to another folder.

8.8.1 Backing up the SVC cluster configuration

You can backup your cluster configuration by using the Backing Up a Cluster Configuration panel or the CLI **`svcconfig`** command. This section describes the overall procedure for backing up your cluster configuration and the conditions that must be satisfied to perform a successful backup.

The backup command extracts configuration data from the cluster and saves it to svc.config.backup.xmlin in /tmp. A file svc.config.backup.sh is also produced. You can study this file to see what other commands were issued to extract information.

A log svc.config.backup.logis also produced. You can study this log for details in regard to what was done and when. This log also includes information about the other commands issued.

Any pre-existing file svc.config.backup.xml is archived as svc.config.backup.bak. Only one such archive is kept. We recommend that you move immediately the XML file and related KEY files (see limitations below) *off the cluster* for archiving. Then erase the files from /tmp using the **`clear`** command. We also recommend that you change all objects having default names to *non-default names*. Otherwise, a warning is produced for objects with default names. Also the object with the default name is restored with its original name with an "_r" appended. The prefix _(underscore) is reserved for backup and restore command usage, and should not be used in any object names.

> **Attention:** The tool backs up logical configuration data only, not client data. It does not replace a traditional data backup and restore tool, but supplements such a tool with a way to backup and restore the client's configuration. To provide a complete backup and disaster recovery solution, the user must backup both user (non-configuration) data and configuration (non-user) data. After restoration of the SVC configuration, the client is expected to fully restore user (non-configuration) data to the cluster's disks.

The full syntax of the **`svcconfig backup`** command is:

```
svcconfig backup [-quite] [-v on|off]
```

Note the following explanation:

- **-quiet**: Suppresses standard output (STDOUT) messages from the console.
- **-v on | off**: *On* means verbose messages are displayed. *Off* means normal messages (the default) are displayed.

Prerequisites

You must have the following prerequisites in place:

- All nodes must be online.
- No object name may begin with an underscore.
- All objects should have non-default names, that is, names that are not assigned by the SAN Volume Controller.

Although we recommend that objects have non-default names at the time the backup is taken, this is not mandatory. Objects with default names are renamed when they are restored.

Example 8-51 shows an example of the **svcconfig backup** command.

Example 8-51 svcconfig backup command

```
IBM_2145:admin>svcconfig backup
............
CMMVC6112W io_grp io_grp1 has a default name
.
CMMVC6112W io_grp io_grp2 has a default name
.
CMMVC6112W io_grp io_grp3 has a default name
.
CMMVC6112W io_grp recovery_io_grp has a default name
.................
CMMVC6136W No SSH key file svc.config.barry.admin.key
CMMVC6136W No SSH key file svc.config.service.service.key
...................
IBM_2145:admin>svcconfig clear -all
IBM_2145:admin>
```

Example 8-52 shows the **pscp** command.

Example 8-52 pscp command

```
C:\Support Utils\Putty>pscp admin@9.43.226.248:/tmp/svc.config.backup.xml c:\cli
backup.xml
Authenticating with public key "rsa-key-20031031"
clibackup.xml             |         22 kB |  22.2 kB/s | ETA: 00:00:00 | 100%

C:\Support Utils\Putty>
```

Context

The following scenario illustrates the value of configuration backup:

1. Use the **svcconfig** command to create a backup file on the cluster that contains details about the current cluster configuration.
2. Store the backup configuration on some form of tertiary storage. You must copy the backup file from the cluster or it becomes lost if the cluster crashes.
3. If a severe failure occurs, it causes the cluster to become lost. Both configuration data (for example, the cluster definitions of hosts, I/O groups, MDGs, mDisks) and the application data on the virtualized disks are lost. In this scenario, it is assumed that the application data can be restored from normal client backup procedures. However, before you can

carry this out, you must reinstate the cluster, as configured at the time of the failure. This means you restore the same MDGs, I/O groups, host definitions, and the vDisks that existed prior to the failure. You can copy the application data back onto these vDisks and resume operations.

4. Recover the hardware. This includes hosts, SVCs, disk controller systems, disks, and SAN fabric. The hardware and SAN fabric must physically be the same as those used before the failure.
5. Re-initialize the cluster.
6. Restore your cluster configuration using the backup configuration file generated.
7. Restore the data on your virtual disks (vDisks) using your preferred restore solution or with help from IBM Service.
8. Resume normal operations.

8.8.2 Restoring the SVC cluster configuration

> **Important:** After the recovery process is started, the user data on the vDisks is destroyed and must be recovered from your usual tape backup process. Before you begin this process, consult IBM Support to determine the cause of the loss of your cluster configuration.

You can use the **`restore`** command to take information from configuration files in the /tmp directory of SVC and restore the cluster to that configuration. You can enter this command immediately only after a cluster is created. This command restores the target cluster configuration from the svc.config.backup.xml file, and associated .key files (if present) in the configuration files directory.

The full syntax of the **`svcconfig restore`** command is:

```
svcconfig restore [-f |-force] [-q |-quite] [-prepare [-fmt|-fmtdisk]| -execute
[-fmt|-fmtdisk] ] [-v on|off]
```

Note the following explanation:

- **-prepare**: Checks the current configuration against the information held in svc.config.backup.xml on the configuration to be restored. It prepares commands for execution in svc.config.restore.sh, and produces a log of events in svc.config.restore.prepare.log.
- **-fmt | fmtdisk**: Includes the -fmtdisk option and all **`mkvdisk`** commands to be issued.
- **-execute**: Executes the command script ?svc.config.restore.sh. It produces a log of events in svc.config.restore.execute.log.
- **-f | force**: Forces continued execution where possible.
- **-q | quiet**: Suppresses console output (STDOUT).
- **-v on | off**: Produces verbose output (on). The default is regular output (off).

The **`pscp`** command shown in Example 8-53 enables files to be transferred using PuTTY secure copy. See 8.6.1, “PuTTY Secure Copy” on page 195, for full details.

Example 8-53 pscp command

```
C:\Support Utils\Putty>pscp C:\svc.config.backup.xml admin@9.43.226.248:/tmp
Authenticating with public key "rsa-key-20031031"
svc.config.backup.xml     |         22 kB |  22.2 kB/s | ETA: 00:00:00 | 100%
C:\Support Utils\Putty>
```

The **svcconfig restore** command shown in Example 8-54 enables the previously saved configuration that was transferred using OSCP to be restored.

Example 8-54 svcconfig restore command

```
IBM_2145:admin>svcconfig restore -prepare
............................
IBM_2145:admin>svcconfig restore -execute
.................................
```

8.9 Scripting and its usage under CLI for SVC task automation

Usage of scripting constructs is better for automation of regular operational jobs. You can use available shells to develop it. And to run a SVC console where operating system is Windows 2000, you can either purchase licensed shell emulation software or download Cygwin from:

http://www.cygwin.com

Scripting enhances the productivity of SVC administrators and integration of their storage virtualization environment. Example 8-55 shows script used to reset the SVC MDG and hosts to default before restoring the SVC configuration backup.

Example 8-55 SVC reset script

```
#########################################################
#
#  Program: reset_demo
#
#  Purpose:   Remove MDG and all hosts defined.
#
##########################################################

progname=`basename $0`

. /usr/local/bin/democonfig

# Verify the SVCSSH was updated

if [ "${SVCSSH}" = "CHANGEME" ]
then
   printf "${progname}: Fatal Error: Script must be customized before use\n"
   exit 1
fi

#  First ensure were starting with a clean slate

# Catch & report a problem if the group already exists
```

```
mdisklist=`${SVCSSH} svcinfo lsmdiskgrp -delim : -nohdr | cut -f2 -d:`

for mdgrp in  ${mdisklist}
do
    echo "svctask rmmdiskgrp -force ${mdgrp}"
    ${SVCSSH} "svctask rmmdiskgrp -force ${mdgrp}"
done

hlist=`${SVCSSH} svcinfo lshost -delim : -nohdr | cut -f2 -d:`
for hname in  ${hlist}
do
   echo "svctask rmhost ${hname}"
   ${SVCSSH} "svctask rmhost ${hname}"
done
```

You can create your own customized scripts to automate a large number of tasks for completion at a variety of times and run them through the CLI.

9

SVC configuration and administration using the GUI

This chapter describes how to use the CIM Agent and Console for the IBM TotalStorage SAN Volume Controller graphical user interface (GUI). They allow you to perform additional and advanced configuration and administration tasks, which are not covered in Chapter 6, “Quickstart configuration using the GUI” on page 101.

9.1 Managing the cluster using the GUI

This section explains the various configuration and administration tasks that you can perform on the cluster.

9.1.1 Installing certificates

You may have already accepted certificates, as suggested in Chapter 6, “Quickstart configuration using the GUI” on page 101. If you did not, you may notice many instances as we continue where your are prompted with security warnings regarding unrecognized certificates. Return to 6.3, “Installing certificates” on page 112, and complete the steps.

9.1.2 Organizing onscreen content

In the following sections, there are several panels within the SVC GUI for you to perform filtering (to minimize the amount of data shown on screen) and sorting (to organize the content on the screen). Since we do not cover these functions elsewhere, this section provides a brief overview of these functions.

For the purpose of illustration, on the SVC Welcome page, we click the **Work with Virtual Disks** option and click the **Virtual Disks** link.

Filtering

When you select an item, such as the Virtual Disks link, you are first prompted with a filtering panel (Figure 9-1). On this panel, you can provide the criteria from which the SVC generates the next panel, a subset of all items available for display.

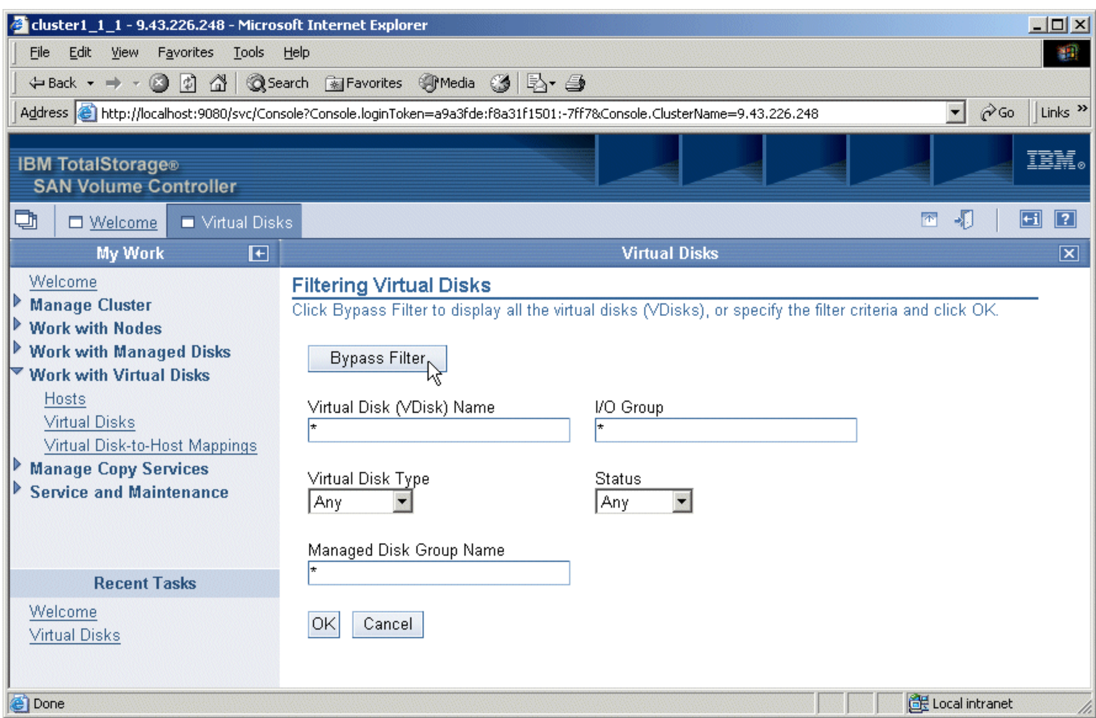


Figure 9-1 Filtering

> **Note:** The asterisk (*) acts as a wildcard.

After you enter the criteria (in our case, `*W2K*`), click the **OK** button. Then the Viewing Virtual Disks panel (Figure 9-2) is displayed. It shows only those virtual disks (vDisks) that meet your filtering criteria.

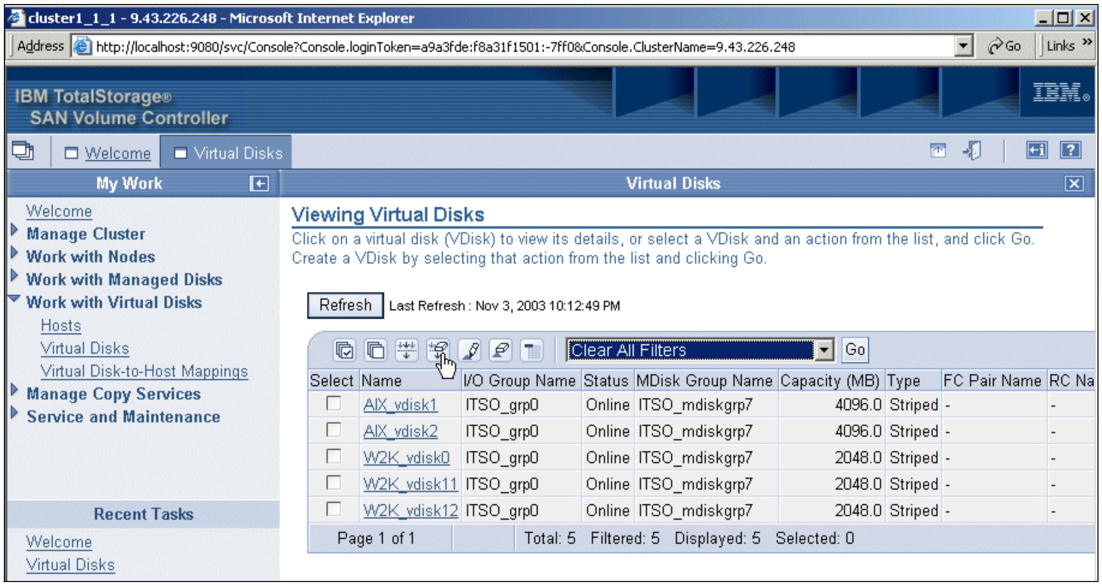


Figure 9-2 Viewing Virtual Disks: Filtered view

If you want to clear the filter so you can see all vDisks, simply select **Clear All Filters** from the list and click **Go**. Or click the small icon highlighted by the mouse pointer in Figure 9-2.

> **Note:** At the time of writing this redbook, the Clear All Filters function did not work. You must close the Viewing vDisks panel and re-open it by clicking the **Bypass Filter** button.

Sorting

Regardless of whether you filter, when you are on the Viewing Virtual Disks panel, you can sort the displayed data by selecting **Edit Sort** from the list and clicking **Go**. Or you can click the small icon highlighted by the mouse pointer in Figure 9-3.

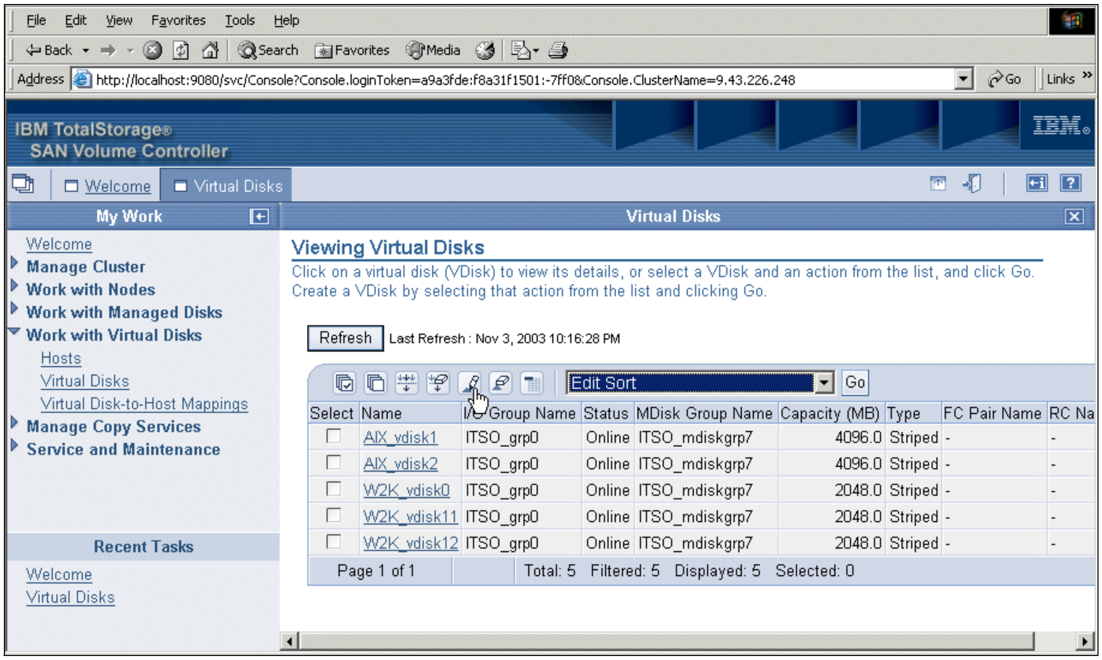


Figure 9-3 Selecting Edit Sort

As shown in Figure 9-4, you can sort based on up to three criteria including: name, I/O group name, status, mDisk group name, capacity (MB), type, FC pair name, and RC name.

Note: The actual sort criteria differs based on the information that you are sorting.

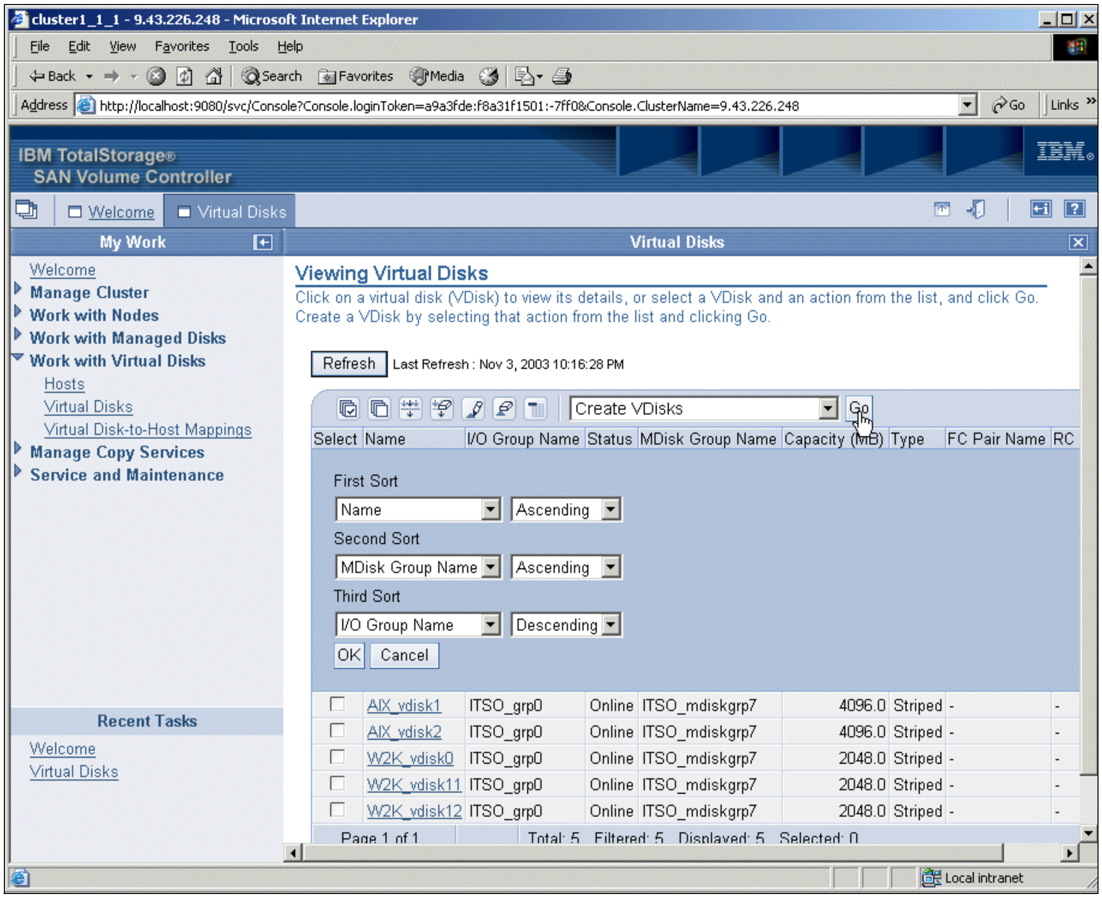


Figure 9-4 Sorting criteria

When you finish making your choices, click **OK** to regenerate the display based on your sorting criteria.

If you want to clear the sort, simply select **Clear All Sorts** from the list and click **Go**. Or click the small icon highlighted by the mouse pointer in Figure 9-5.

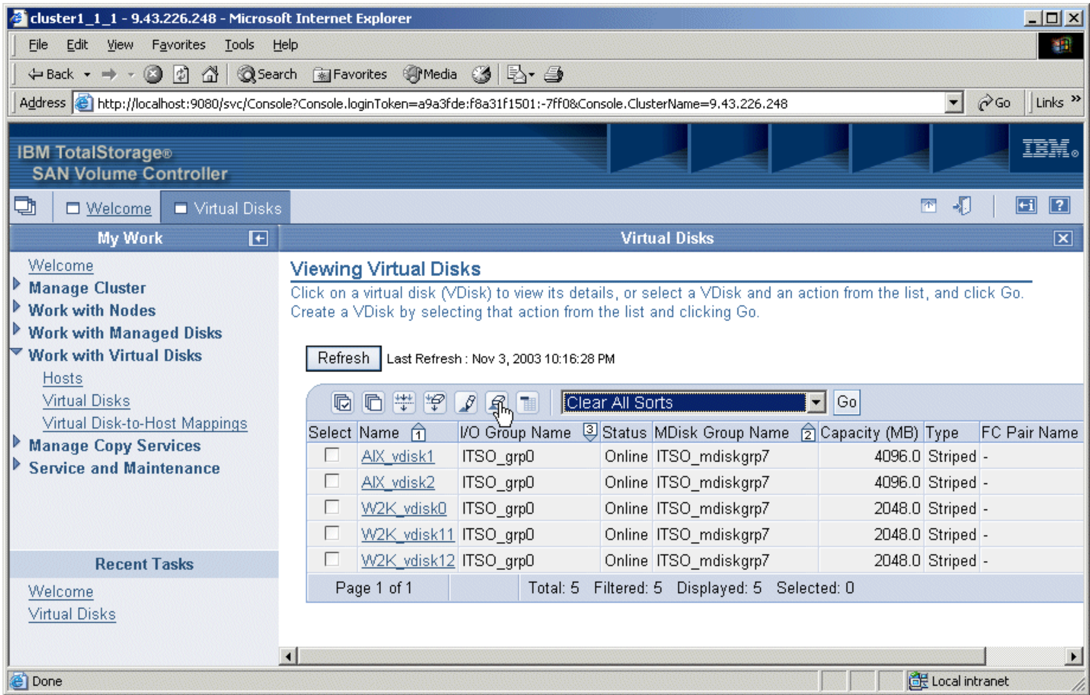


Figure 9-5 Selecting to clear all sorts

Documentation

If you need to access to online documentation, in the upper right corner of the panel, click the icon.

Help

If you need to access to online help, in the upper right corner of the panel, click the icon. This displays a context-sensitive help window.

General housekeeping

If at any time the content in the right side of the frame is “cut off”, you can collapse the My Work column by clicking the icon at the top of the My Work column. When collapsed, the small arrow changes from pointing to the left to pointing to the right ( ). Clicking the small arrow that points right expands the My Work column back to its original size.

In addition, each time you open a configuration or administration window using the GUI in the following sections, it creates a link for that panel along the top of your Web browser beneath the main banner graphic. As a general house keeping task, we recommend that you close each window when you finish with it by clicking the icon to the right of the panel name, but below the icon. Be careful not to close the entire browser.

You are now ready to continue.

9.1.3 Viewing cluster properties

Perform the following steps to display the cluster properties:

1. From the SVC Welcome page, select the **Manage Cluster** option and then the **View Cluster Properties** link.
2. The Viewing General Properties panel (Figure 9-6) opens. Click the **IP Addresses**, **Space**, **SNMP**, or **Statistics** links and you see additional information that pertains to your cluster.

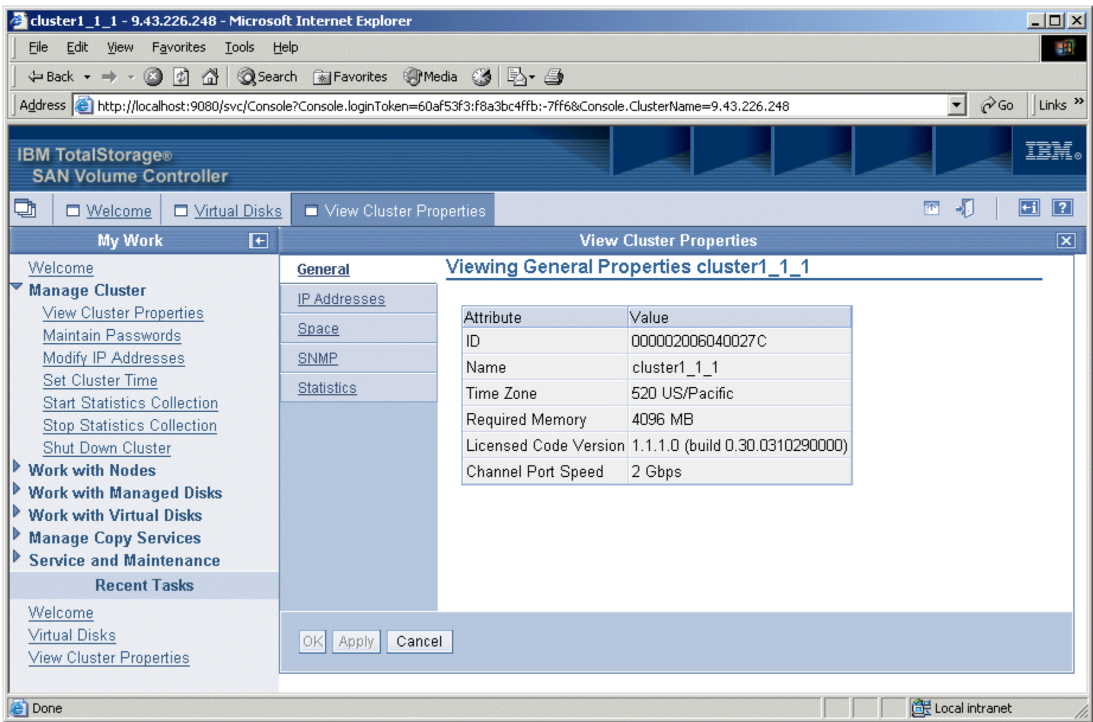


Figure 9-6 View Cluster Properties: General properties

9.1.4 Maintaining passwords

Perform the following steps to maintain passwords:

1. From the SVC Welcome page, select the **Manage Cluster** option and then the **Maintain Passwords** link.
2. Before you can access the Maintain Passwords panel, enter the existing SVC administration user ID and password when prompted. Click **OK**.
3. The Maintain Passwords panel (Figure 9-7) opens. Enter the new passwords for the administrator account, the service account, or both. Click **Modify Password**.

> **Note:** The password can consist of letters A to Z, a to z, numbers 0 to 9, and the underscore. It can be between one and 15 characters in length and is case sensitive.

4. Before the next panel is displayed, enter the new user ID and password combination when prompted.

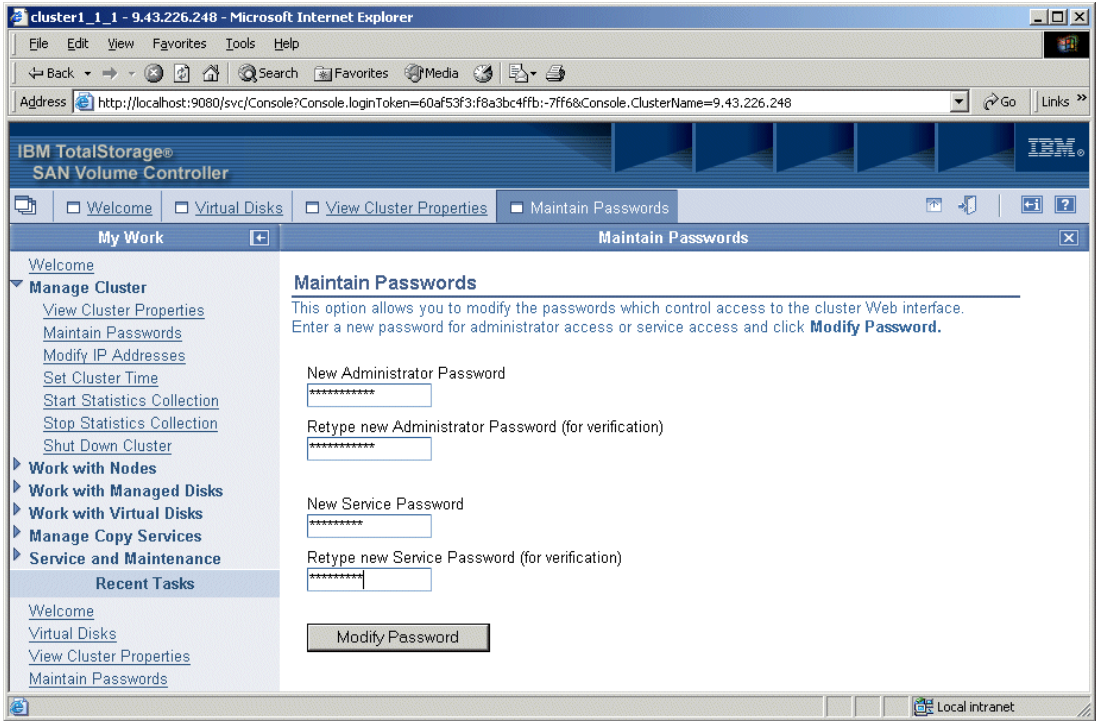


Figure 9-7 Maintain Passwords panel

When complete, you see the successful update messages as shown in Figure 9-8.

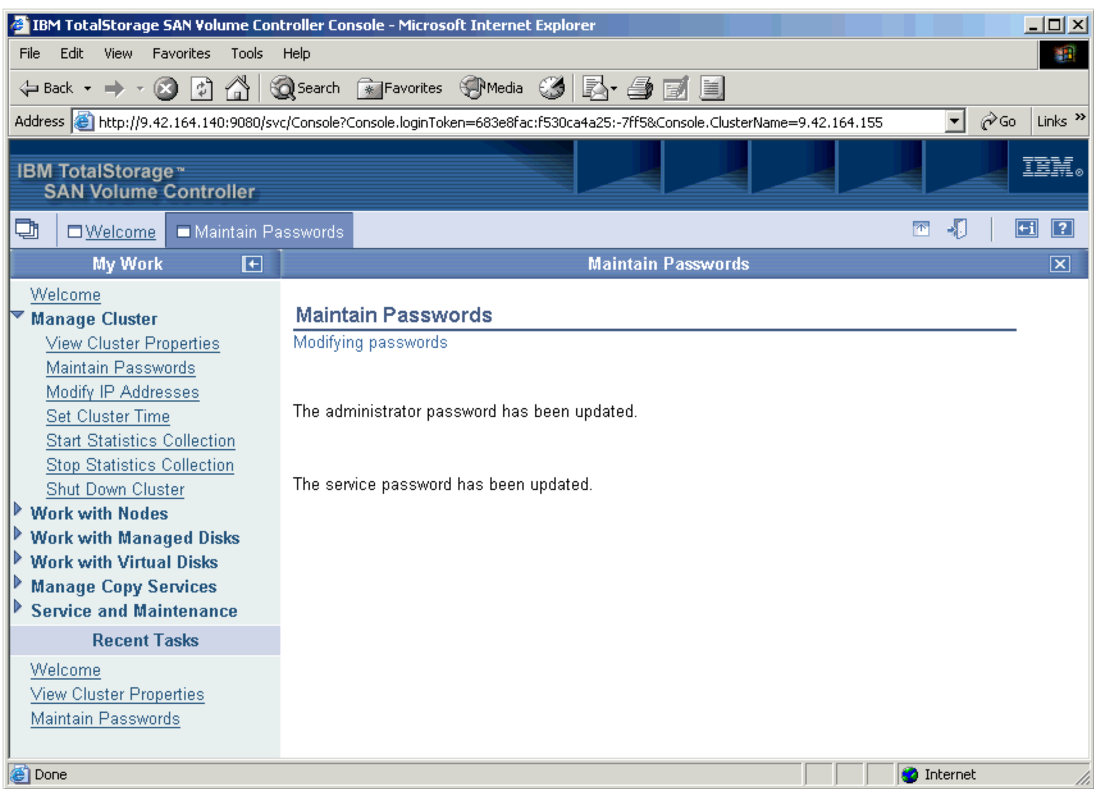


Figure 9-8 Modifying passwords successful update messages

You have now completed the tasks required to change the admin and service passwords for your SVC cluster.

9.1.5 Modifying IP addresses

> **Important:** If you specify a new cluster IP address, the existing communication with the cluster through the GUI is broken. You need to relaunch the SAN Volume Controller Application from the GUI Welcome page.
>
> Modifying the IP address of the cluster, although quite simple, requires some reconfiguration for other items within the SVC environments. This may include, for example, reconfiguring the PuTTY application and the central administration GUI.

Perform the following steps to modify the cluster and service IP addresses of our SVC configuration:

1. From the SVC Welcome page, select the **Manage Cluster** option and the **Modify IP Addresses** link.
2. The Modify IP Addresses panel (Figure 9-9) opens. Make any necessary changes. Then click **Modify Settings**.

   > **Note:** You should use your mouse or the tab key to move between fields.

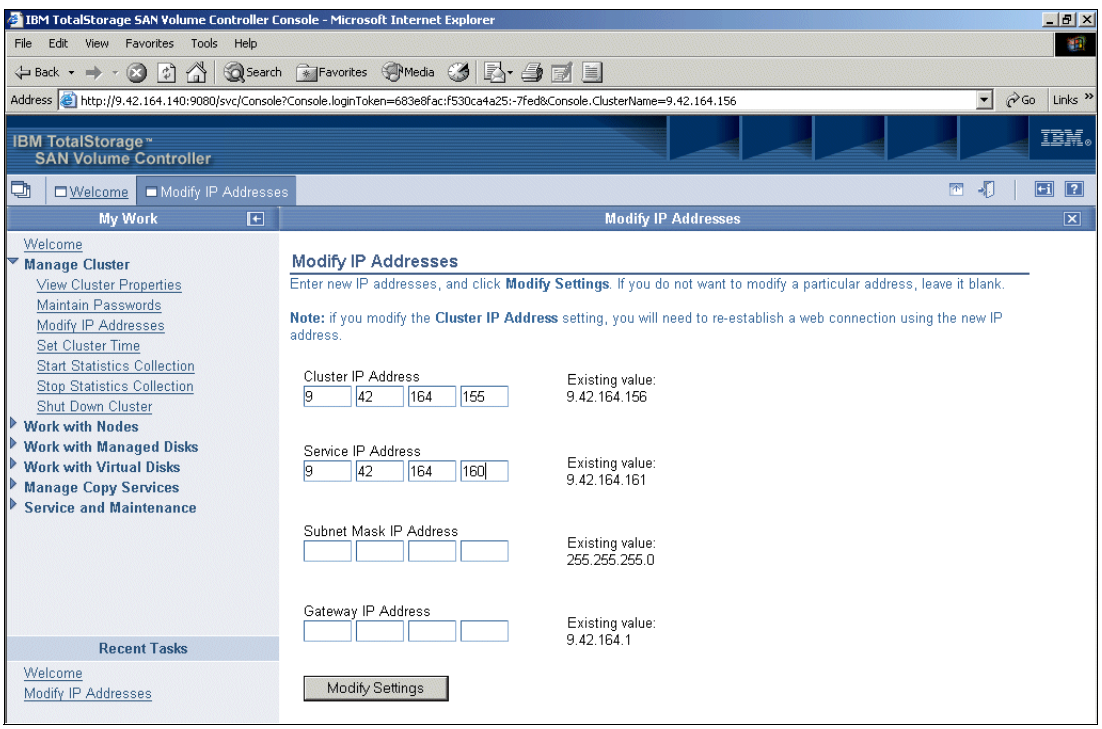


Figure 9-9 Modify IP Addresses

3. You advance to the next panel which shows a message indicating that the IP addresses were updated. See Figure 9-10.

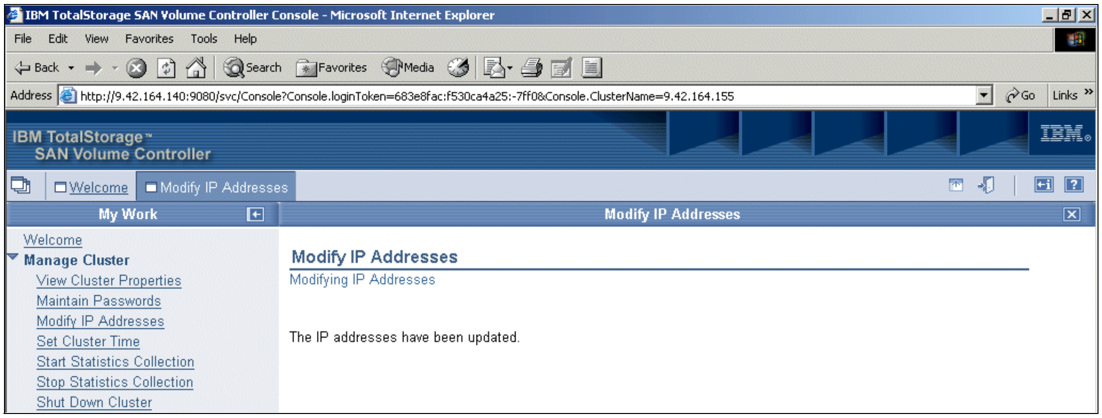


Figure 9-10 Modify IP Addresses update successful message

You have now completed the tasks required to change the IP addresses (cluster and service) for your SVC environment.

9.1.6 Setting the cluster time zone and time

Perform the following steps to set the cluster time zone and time:

1. From the SVC Welcome page, select the **Working with Clusters** options and the **Set Cluster Time** link.
2. The Cluster Date and Time Settings panel (Figure 9-11) opens. At the top of panel, you see the current settings. If necessary, make adjustments and ensure that the **Update cluster data and time** and **Update cluster time zone** check boxes are selected. Click **Update**.

> **Note:** You may be prompted for the cluster user ID and password. If you are, type `admin` and the password you set earlier.

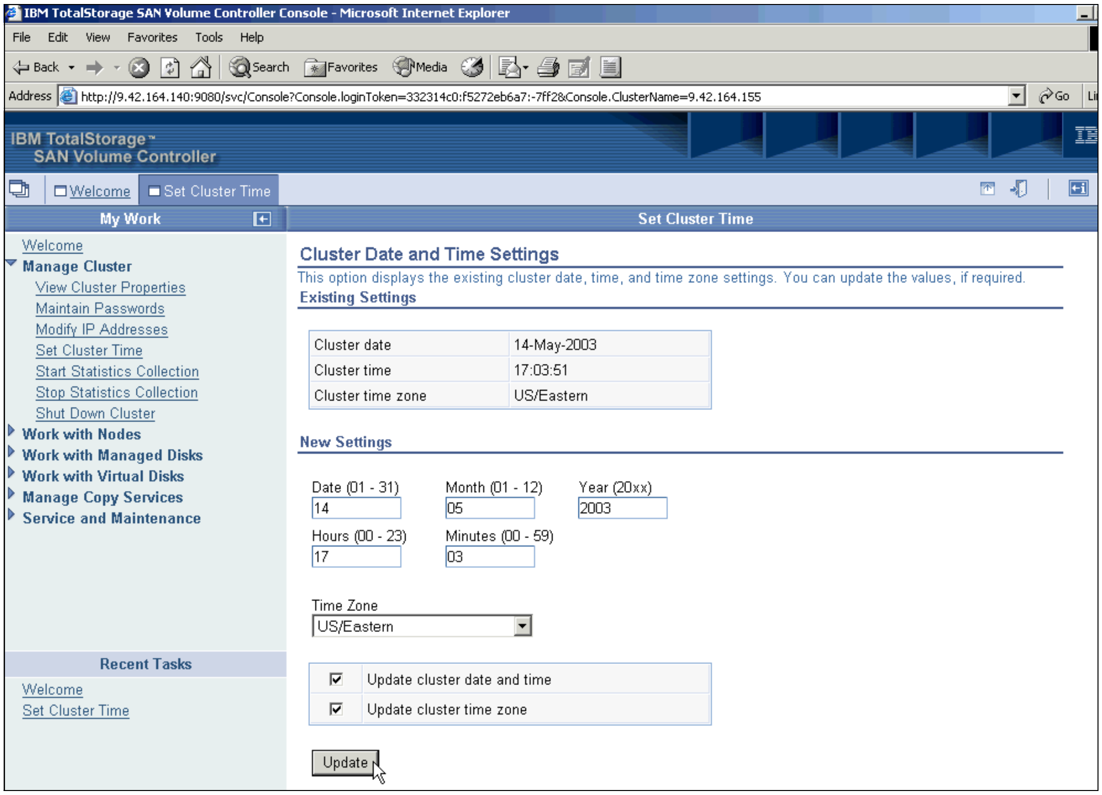


Figure 9-11 Cluster Date and Time Settings panel

3. You return to the Cluster Date and Time Settings panel (Figure 9-12), which shows the new settings.

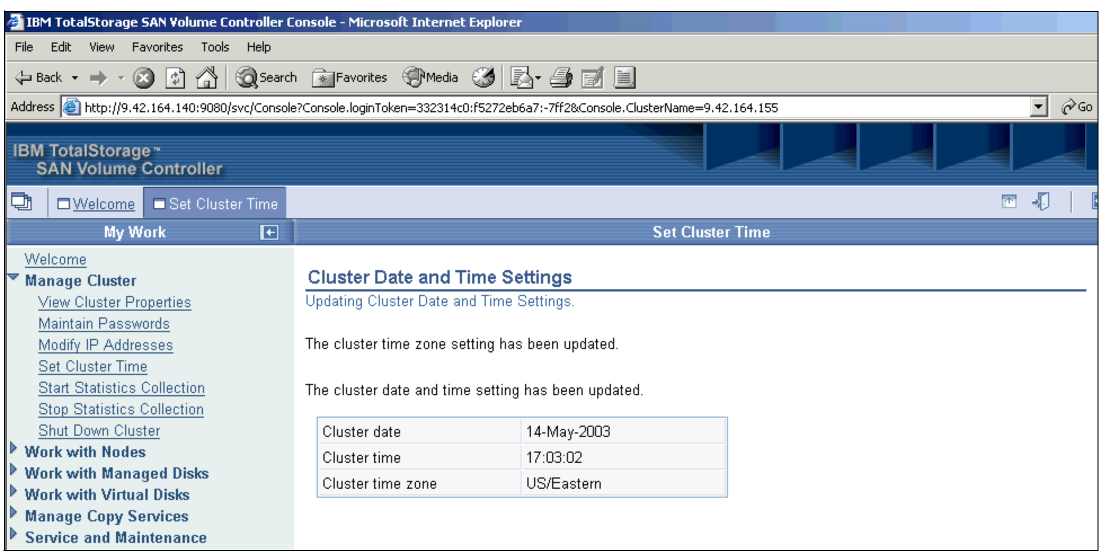


Figure 9-12 Cluster Date and Time Settings update confirmation

You have now completed the tasks necessary to set the cluster time zone and time.

9.1.7 Starting the statistics collection

Perform the following steps to start statistics collection on your cluster:

1. From the SVC Welcome page, select the **Manage Cluster** option and the **Start Statistics Collection** link.
2. The Starting the Collection of Statistics panel (Figure 9-13) opens. Make an interval change, if desired. The interval you specify (minimum 15, maximum 60) is in minutes. Click **OK**.

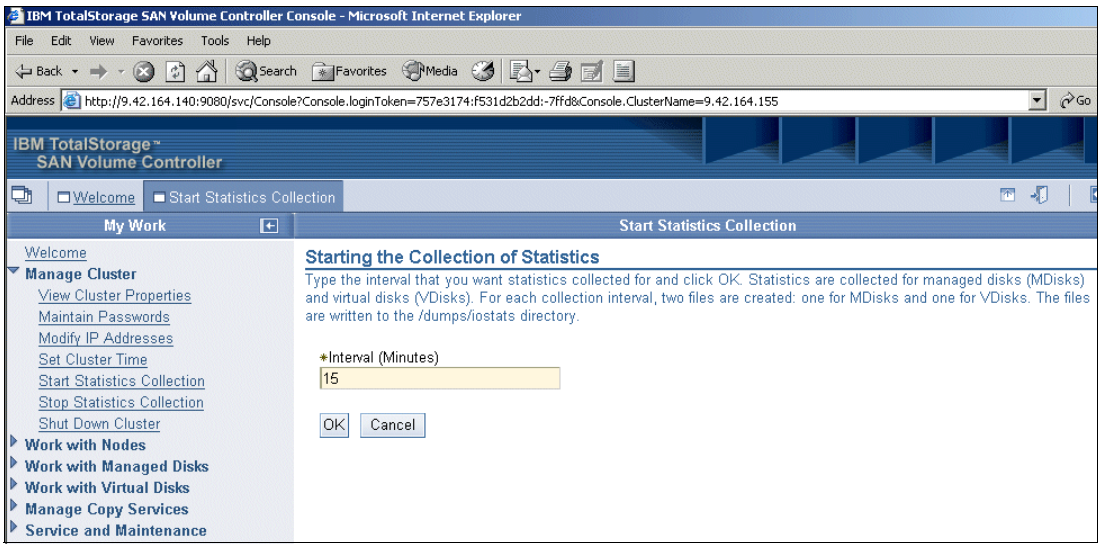


Figure 9-13 Starting collection of statistics

3. Although it does not state the current status and it is not obvious, clicking OK turns on the statistics collection. To verify, click the **Cluster Properties** link as you did in 9.1.3, "Viewing cluster properties" on page 223. Then, click the **Statistics** link. You see the interval as specified in Step 2 and the status of *On* as shown in Figure 9-14.

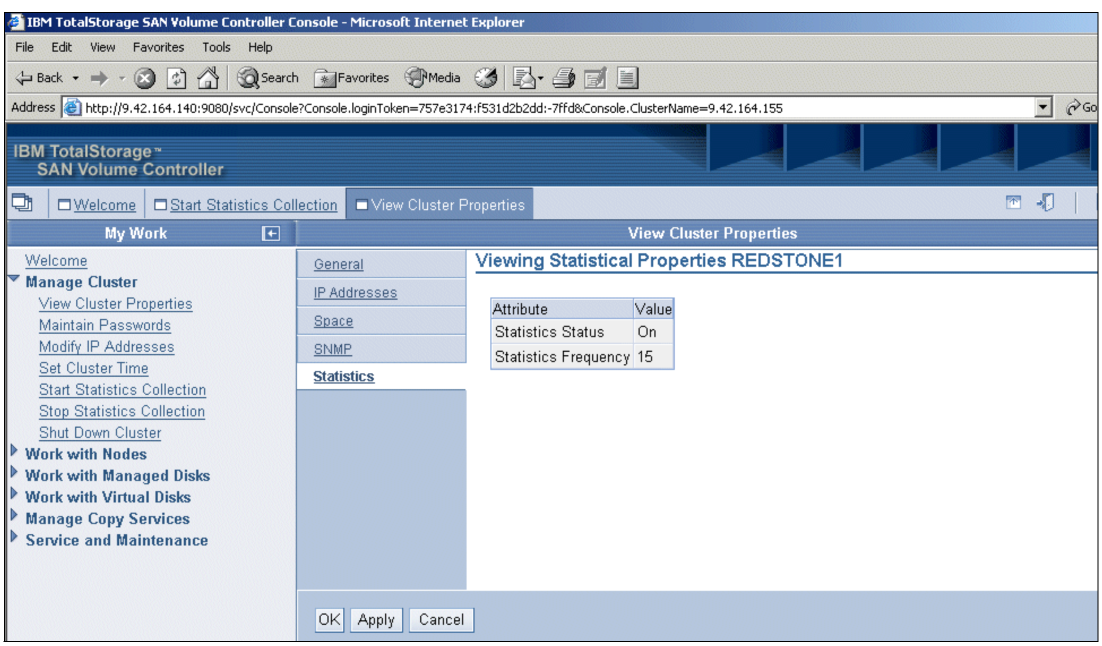


Figure 9-14 Verifying that statistics collection is on

You have now completed the tasks required to start statistics collection on your cluster.

9.1.8 Stopping the statistics collection

Perform the following steps to stop statistics collection on your cluster:

1. From the SVC Welcome page, select the **Manage Cluster** option and the **Stop Statistics Collection** link.
2. The Stopping the Collection of Statistics panel (Figure 9-15) opens. You see a message asking whether you are sure that you want to stop the statistics collection. Click **Yes** to stop the ongoing task.

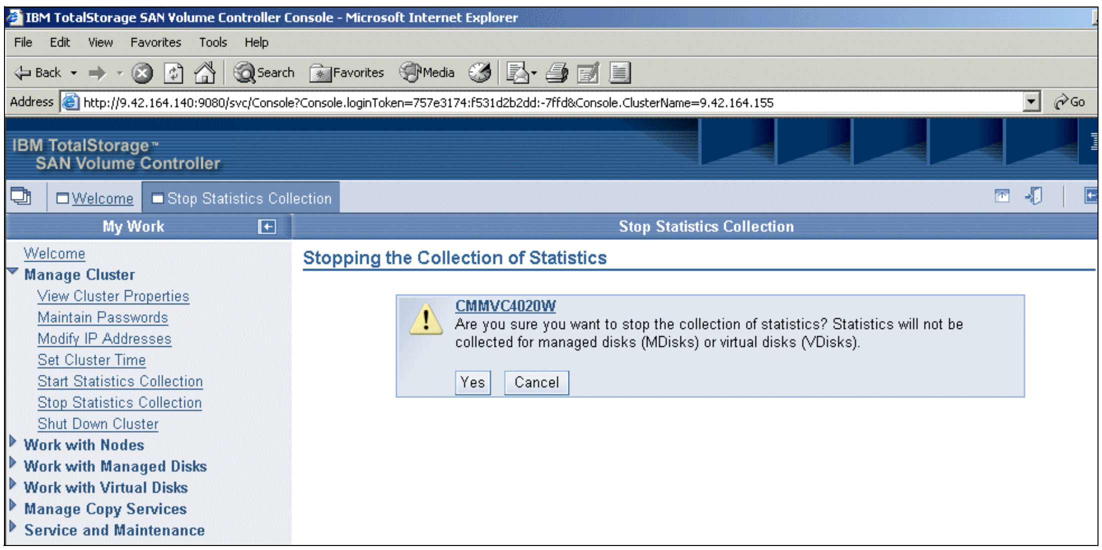


Figure 9-15 Stopping the collection of statistics

3. The window closes. To verify that the collection has stopped, click the **Cluster Properties** link as you did in 9.1.3, “Viewing cluster properties” on page 223. Then, click the **Statistics** link. Now you see the status has changed to *Off* as shown in Figure 9-16.

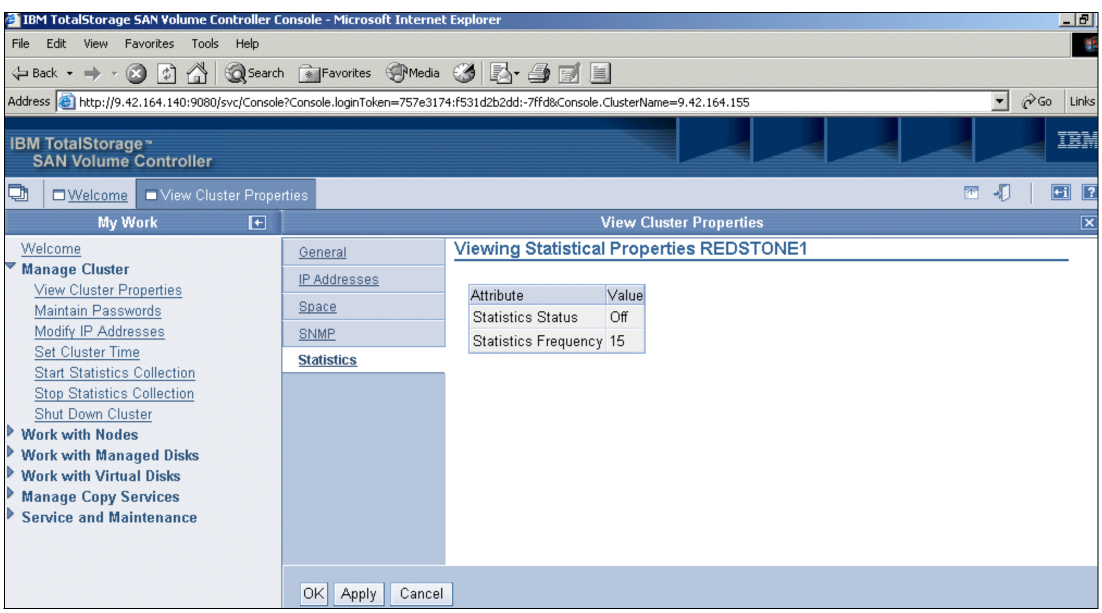


Figure 9-16 Verifying that statistics collection is off

You have now completed the tasks required to stop statistics collection on your cluster.

9.1.9 Shutting down the cluster

If all input power to a SAN Volume Controller cluster is to be removed for more than a few minutes (for example, if the machine room power is to be shutdown for maintenance), it is important that you shut down the cluster before you remove the power. The reason for this is that if the input power is removed from the uninterruptible power supply units without shutting down the cluster and the uninterruptible power supply units themselves first, the uninterruptible power supply units remain operational and eventually become drained of power.

When input power is restored to the uninterruptible power supply units, they start to recharge. However, the SAN Volume Controller does not permit any input/output (I/O) activity to be performed to the vDisks until the uninterruptible power supply units are charged enough to enable all the data on the SAN Volume Controller nodes to be de-staged in the event of an unexpected power loss. This may take as long as three hours. Shutting down the cluster prior to removing input power to the uninterruptible power supply units prevents the battery power from being drained and makes it possible for I/O activity to be resumed as soon as input power is restored.

Perform the following steps to shut down your cluster:

Important: Make sure that all connected hosts are shut down prior to undertaking this task. Failure to do so can cause data corruption or loss.

1. From the SVC Welcome page, select the **Manage Cluster** option and the **Shut Down Cluster** link.
2. The Shutting down cluster panel (Figure 9-17) opens. You see a message asking you to confirm whether you want to shut down the cluster. Click **Yes** to begin the shutdown process.

   Note: At this point, you lose administrative contact with your cluster.

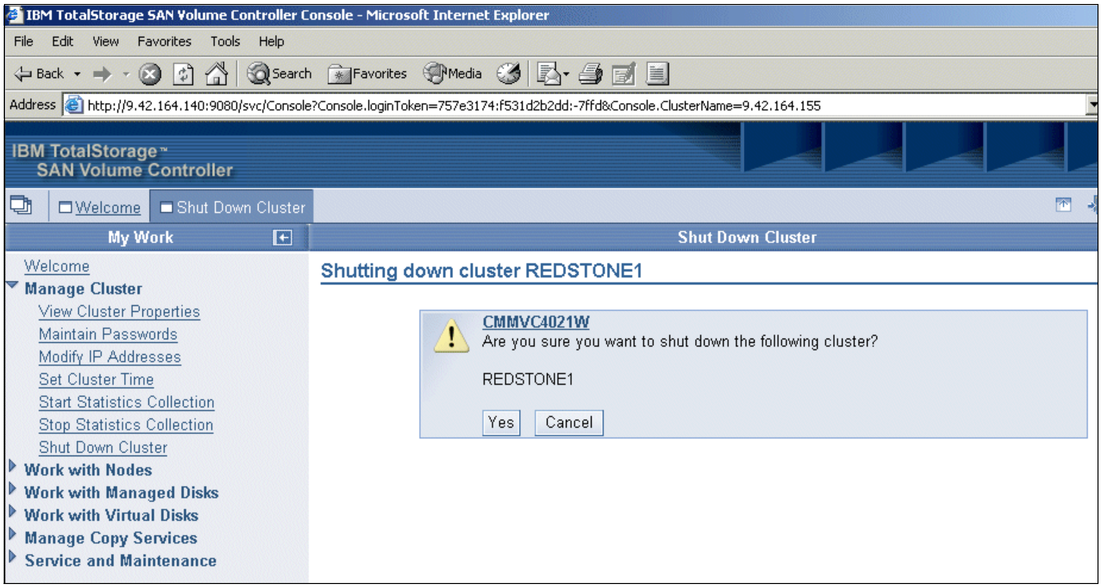


Figure 9-17 Shutting down the cluster

You have now completed the tasks required to shut down the cluster. Now you can shut down the uninterruptible power supplies by pressing the power button on their front panels.

Note: To restart the SVC cluster, you must first restart the uninterruptible power supply units by pressing the power button on their front panels. After they are on, go to the service panel of one of the nodes within your SVC cluster and press the *power on* button, releasing it quickly. After it is fully booted (for example, displaying `Cluster:` on line 1 and the cluster name on line 2 of the SVC front panel), you can start the other nodes in the same way.

As soon as all nodes are fully booted and you re-established administrative contact using the GUI, your cluster is fully operational again.

9.2 Working with nodes using the GUI

This section discusses the various configuration and administration tasks that you can perform on the nodes within an SVC cluster.

9.2.1 I/O groups

This section details the tasks that can be performed on an I/O group level.

Renaming an I/O group

Perform the following steps to rename an I/O group:

1. From the SVC Welcome page, select the **Work with Nodes** option and the **I/O Groups** link.
2. The Viewing Input/Output Groups panel (Figure 9-18) opens. Select the radio button to the left of the I/O group you want to rename. In our case, we select **io_grp0**. Ensure that **Rename an I/O Group** is selected from the list. Click **Go**.

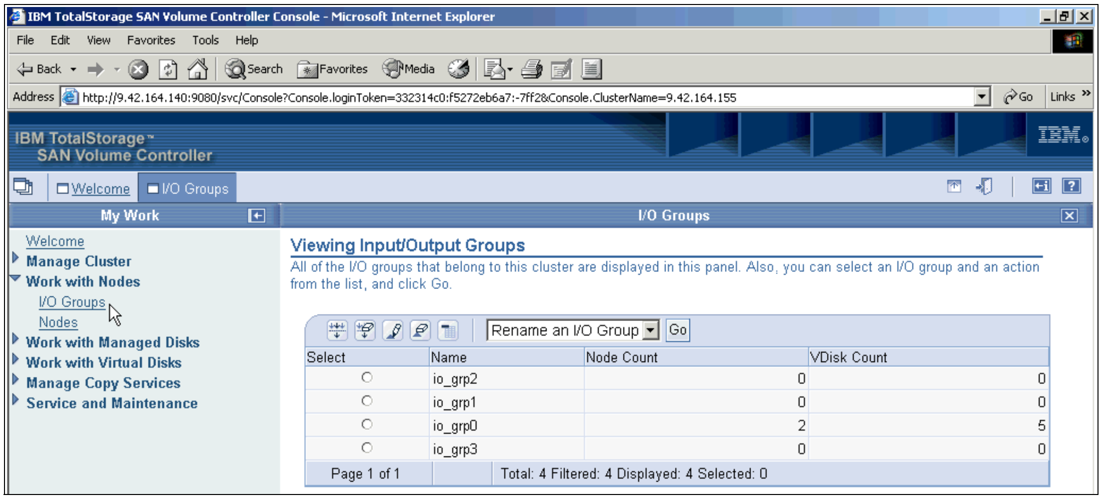


Figure 9-18 Viewing Input/Output Groups

3. On the Renaming I/O Groups *I/Ogroupname* panel (*I/Ogroupname* is the I/O group you selected in the previous step), type the new name you want to assign to the I/O group. Click **OK**. See Figure 9-19.

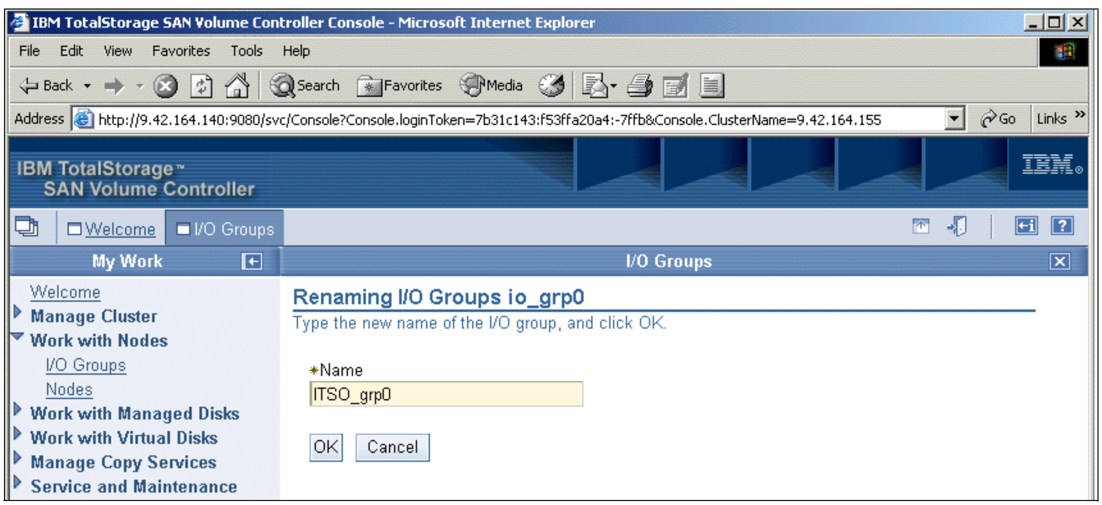


Figure 9-19 Renaming the I/O group

Note: The name can consist of letters A to Z, a to z, numbers 0 to 9, and the underscore. It can be between one and 15 characters in length, but cannot start with a number or the word *iogrp* because this prefix is reserved for SVC assignment only.

You have now completed the tasks required to rename an I/O group.

9.2.2 Nodes

This section discusses the tasks that you can perform on a node level. You perform each task from the Viewing Nodes panel (Figure 9-20). To access this panel, from the SVC Welcome page, select the **Work with Nodes** options and the **Nodes** link.

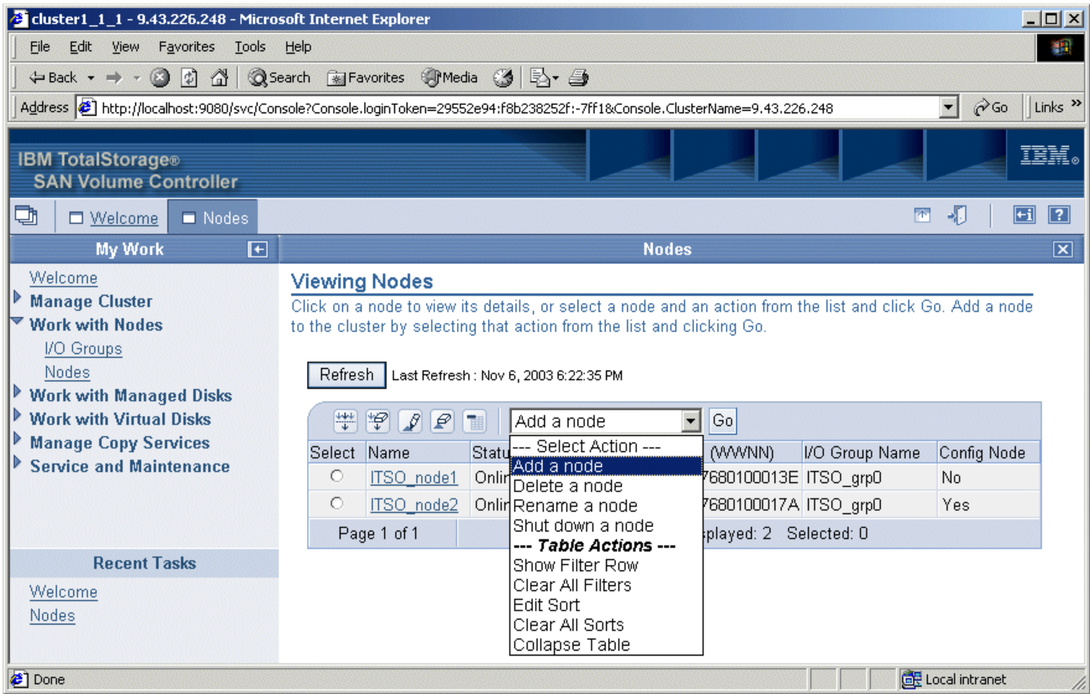


Figure 9-20 Viewing Nodes

Viewing the node details

Perform the following steps to view information about a node within the SVC cluster:

1. From the Viewing Nodes panel (Figure 9-20), click the highlighted name of the node.
2. The Viewing General Details *nodename* panel (where *nodename* is the node you chose) opens as shown in Figure 9-21. Click the **Ports** and **Vital Product Data** links to view additional information about your selected node.

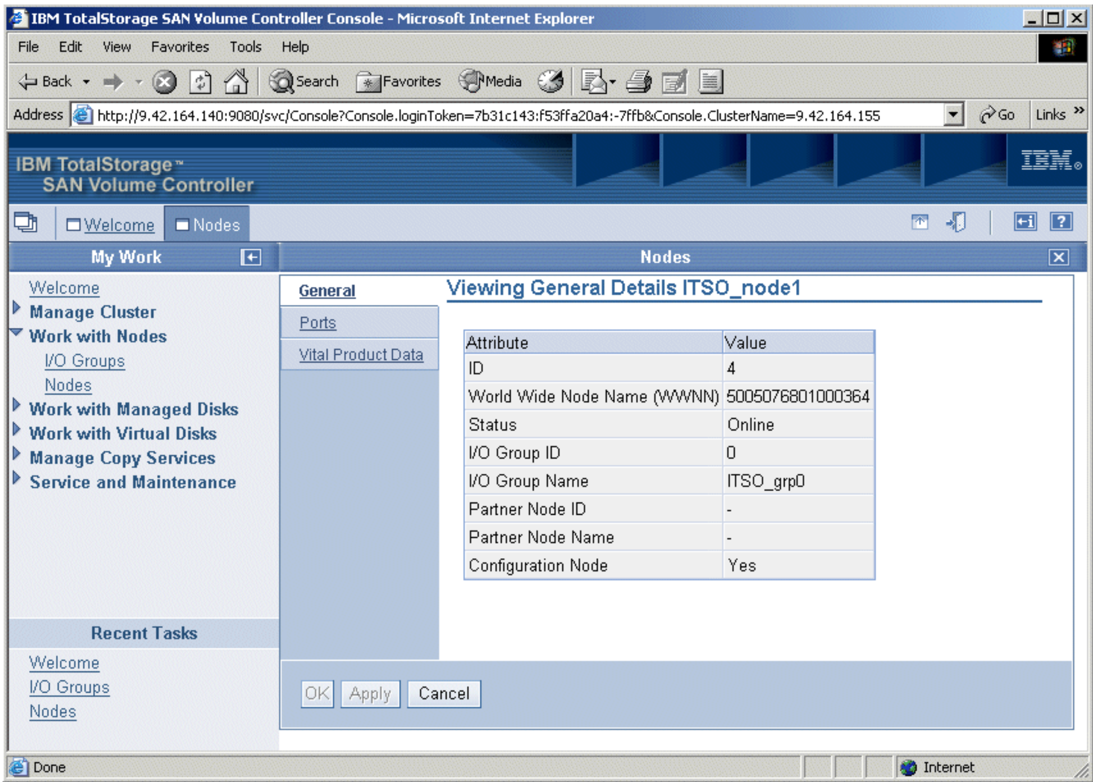


Figure 9-21 General node details

Adding a node

Perform the following steps to add a node to the SVC cluster:

1. From the Viewing Nodes panel (Figure 9-20), select **Add a node** and click **Go**.
2. On the Adding a Node to a Cluster panel (Figure 9-22), select a node from the list of available nodes. Select the I/O group to which you want to assign the new node. Enter a suitable name for the new node. Click **OK**.

> **Note:** If you do not provide the name, the SVC automatically generates the name *nodeX* (where *X* is the ID sequence number assigned by the SVC internally). The name can consist of letters A to Z, a to z, numbers 0 to 9, and the underscore. It can be between one and 15 characters in length, but cannot start with a number or the word *node* because this prefix is reserved for SVC assignment only.

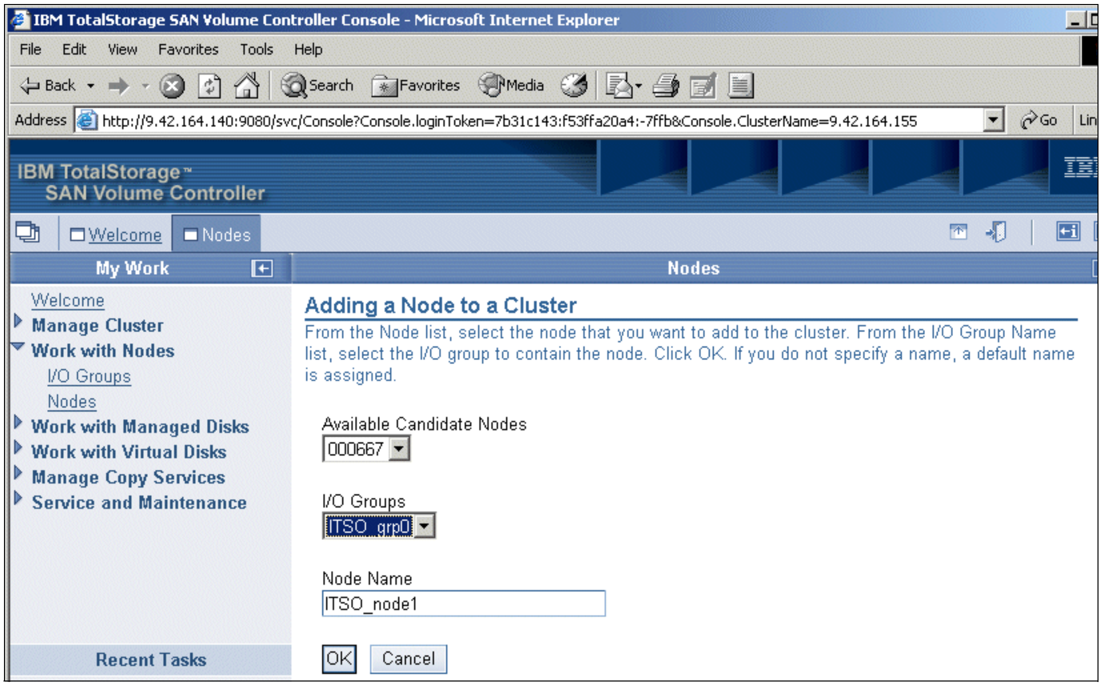


Figure 9-22 Adding a node

Deleting a node

Perform the following steps to delete a node from the SVC cluster:

1. From the Viewing Nodes panel (Figure 9-20 on page 233), select the radio button to the left of the node you want to delete. Select **Delete a node** from the list and click **Go**.
2. On the Deleting a Node from Cluster *nodename* panel (where *nodename* is the name of the node you selected in the previous step), confirm your decision by selecting **Yes**. See Figure 9-23.

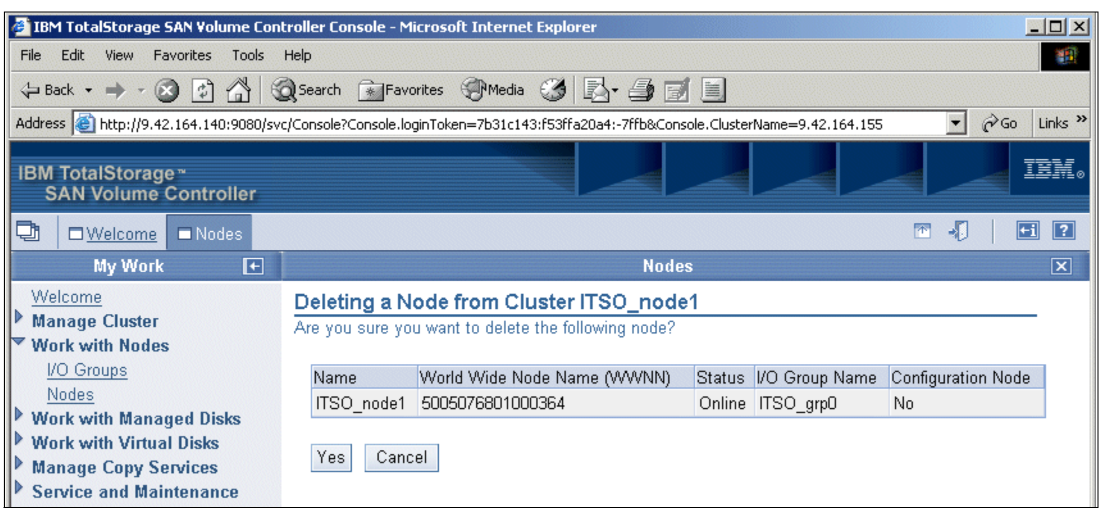


Figure 9-23 Deleting a node

Note: If the node you are deleting is the configuration node, the SVC transfers the configuration node responsibilities to a surviving node prior to deletion.

If the node you are deleting is the last node in an I/O group, and vDisks are still associated with it, you are not permitted to delete it.

Renaming a node

Perform the following steps to rename a node in SVC cluster:

1. From the Viewing Nodes panel (Figure 9-20 on page 233), select the radio button to the left of the node you want to rename. Select **Rename a node** from the list, and click **Go**.
2. On the Renaming Node *nodename* panel (where *nodename* is the node you selected in the previous step), type the new name you want to assign to the node. Click **OK**. See Figure 9-24.

   Note: The name can consist of letters A to Z, a to z, numbers 0 to 9, and the underscore. It can be between one and 15 characters in length, but cannot start with a number or the word *node* because this prefix is reserved for SVC assignment only.

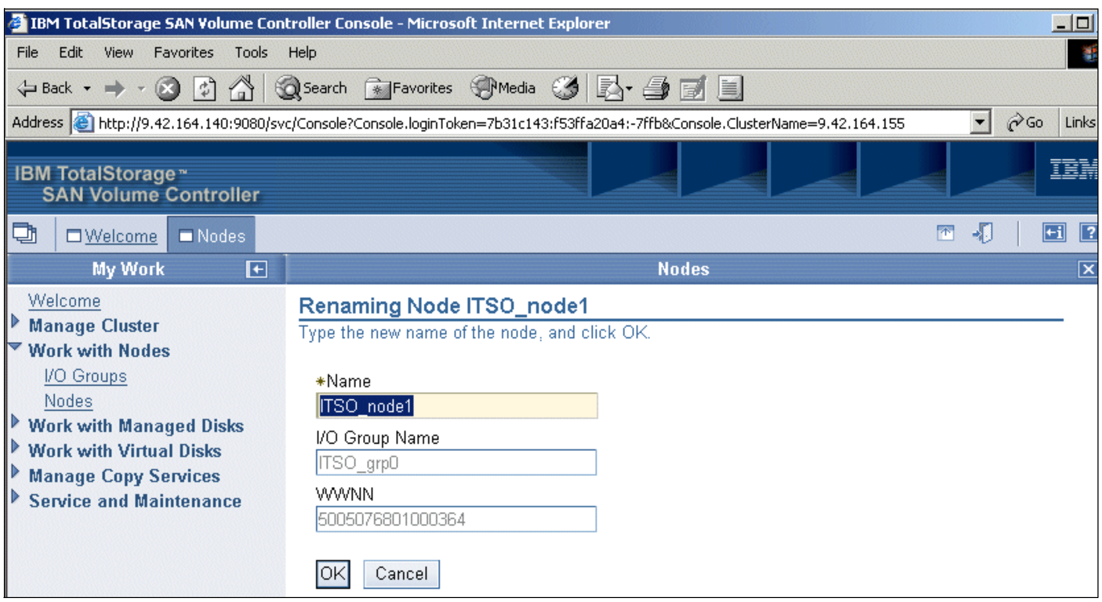


Figure 9-24 Renaming a node

Shutting down a node

Earlier you learned how to shut down the complete SVC cluster in an controlled manner. On occasion, it may be necessary to shut down a single node within the cluster to perform such tasks as scheduled maintenance, while leaving the SVC environment up and running.

To shut down a single node in an SVC cluster, perform the following steps:

1. From the Viewing Nodes panel (Figure 9-20 on page 233), select the radio button to the left of the node you want to shut down. Select **Shut down a node** from the list. Click **Go**.
2. On the confirmation panel (Figure 9-25) that appears next, select **Yes** to continue with the shutdown process.

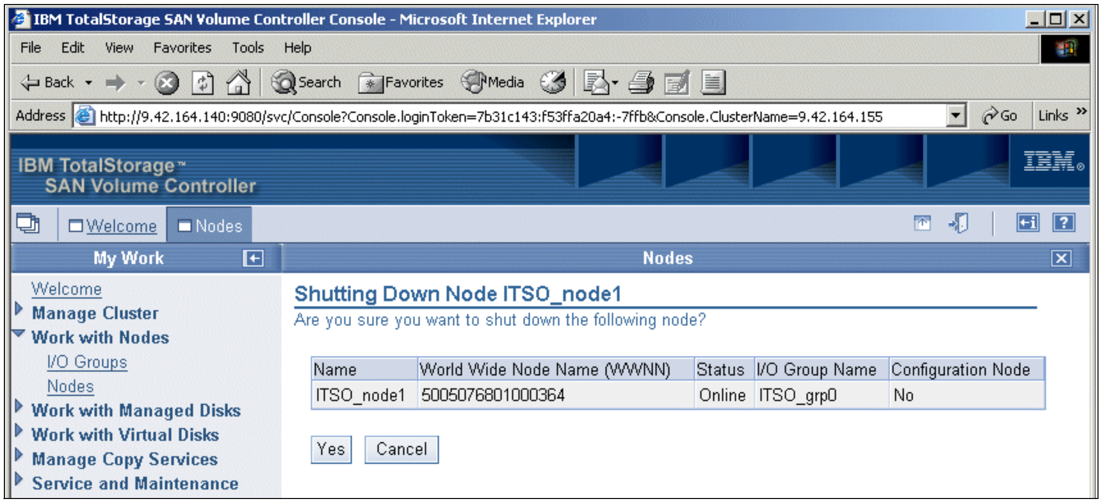


Figure 9-25 Shutting down a node

To restart the SVC node, simply go to the front panel of that node and push the power on button.

You have now completed the tasks that are required to view, add, delete, rename, and shut down a node within a SVC environment.

9.3 Working with managed disks

This section details the various configuration and administration tasks that you can perform on the managed disks (mDisks) within the SVC environment.

9.3.1 Disk controller systems

This section details the tasks that you can perform on a disk controller level.

Viewing disk controller details

Perform the following steps to view information about a back-end disk controller in use by the SVC environment:

1. Select the **Work with Managed Disks** option and then the **Disk Controller Systems** link.
2. The Viewing Disk Controller Systems panel (Figure 9-26) opens. For more detailed information about a specific controller, click its **ID** (highlighted by the mouse cursor in Figure 9-26).

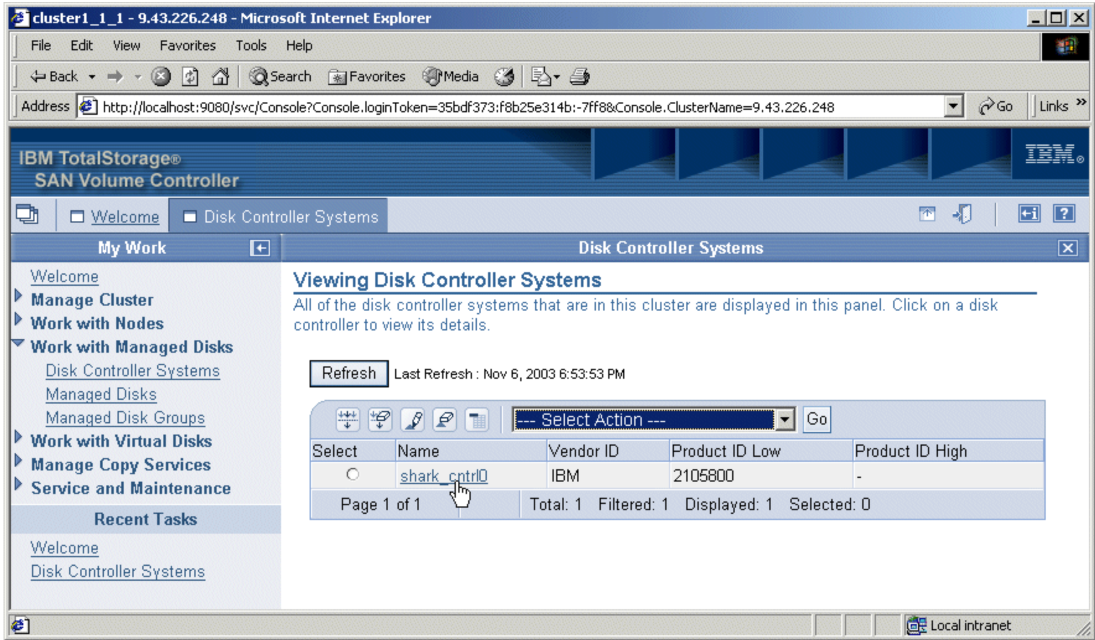


Figure 9-26 Disk controller systems

3. The Viewing General Details for *IDname* panel (Figure 9-27) opens (where *IDname* is the ID you selected). Review the details and click **Cancel** to return to the previous panel.

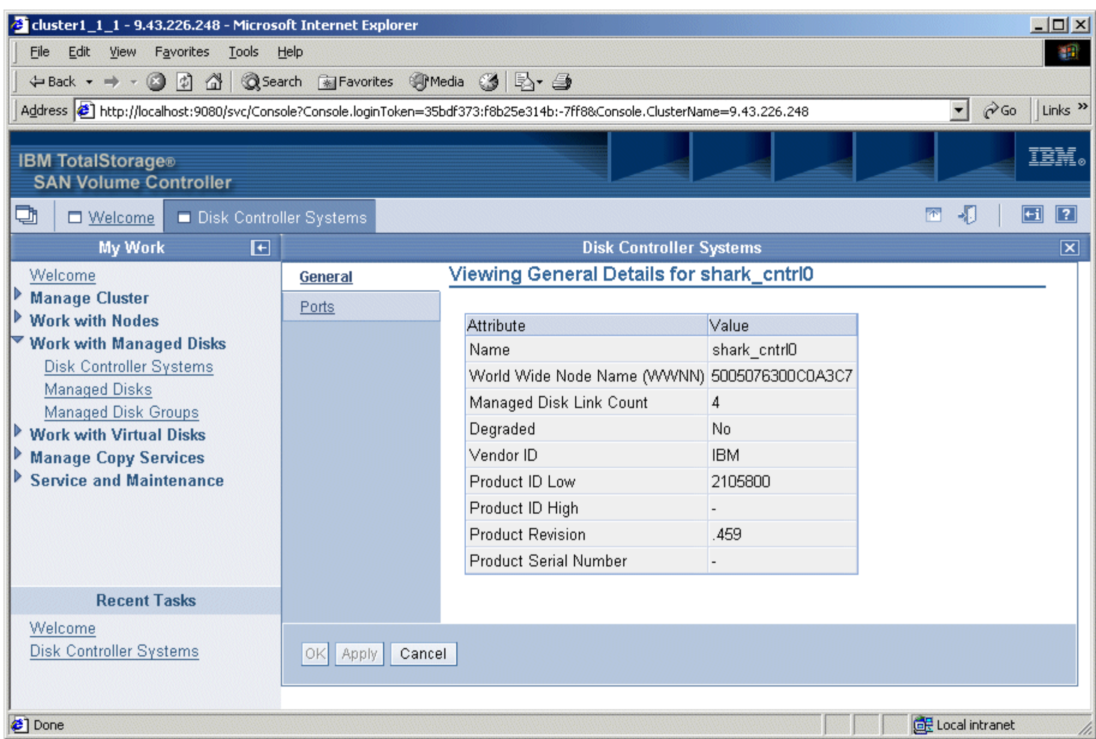


Figure 9-27 Viewing general details about a disk controller

Renaming a disk controller

Perform the following steps to rename a disk controller used by the SVC cluster:

1. Select the radio button to the left of the controller you want to rename. Then select **Rename disk controller system** from the list and click **Go**.
2. On the Renaming Disk Controller System *controllername* panel (where *controllername* is the controller you selected in the previous step), type the new name you want to assign to the controller and click **OK**. See Figure 9-28.

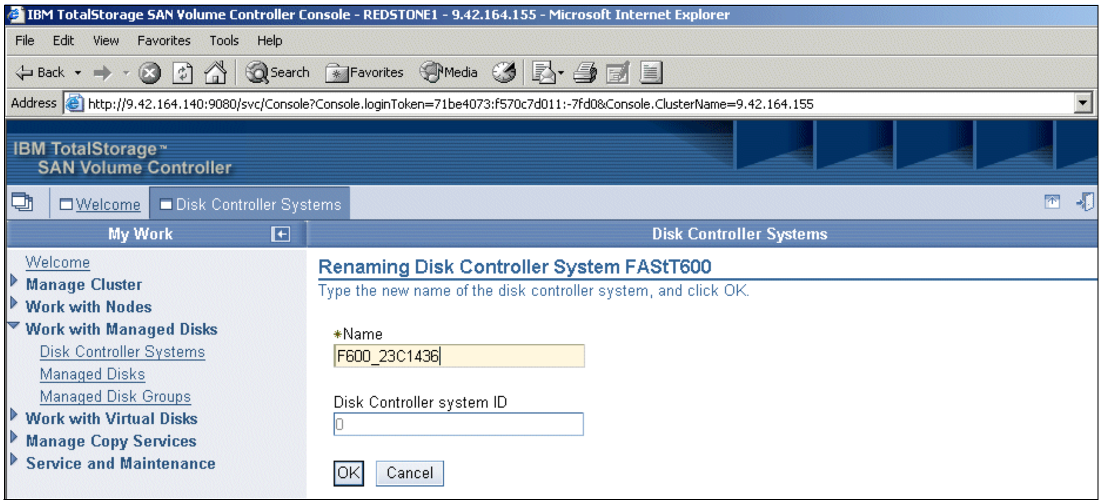


Figure 9-28 Renaming a controller

3. You return to the Disk Controller Systems panel. You should now see the new name of your controller displayed.

> **Note:** The name can consist of letters A to Z, a to z, numbers 0 to 9, and the underscore. It can be between one and 15 characters in length. However, it cannot start with a number or the word *controller* because this prefix is reserved for SVC assignment only.

9.3.2 Managed disks

This section details the tasks which can be performed on an mDisk level. You perform each of the following tasks from the Managed Disk*s* panel (Figure 9-29). To access this panel, from the SVC Welcome page, click the **Work with Managed Disks** option and then the **Managed Disks** link.

> **Note:** At the Filtering Managed Disks (mDisks) panel, click **Bypass filter**.

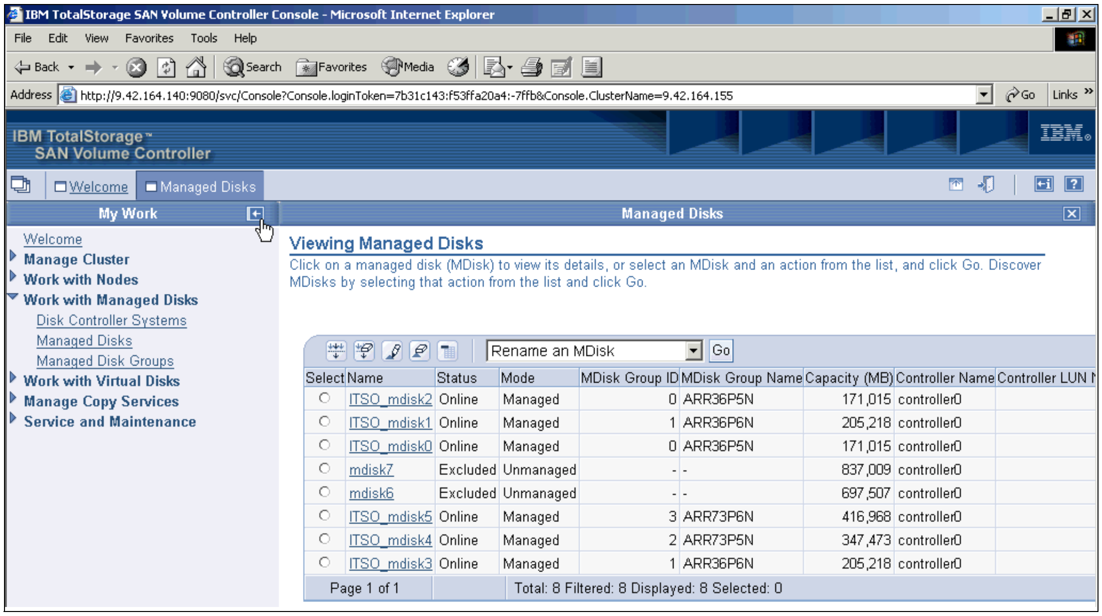


Figure 9-29 Viewing Managed Disks panel

mDisk information

To retrieve information about a specific mDisk, perform the following steps:

1. On the Viewing Managed Disks panel (Figure 9-30), click the name of any mDisk in the list to reveal more detailed information about the specified mDisk.

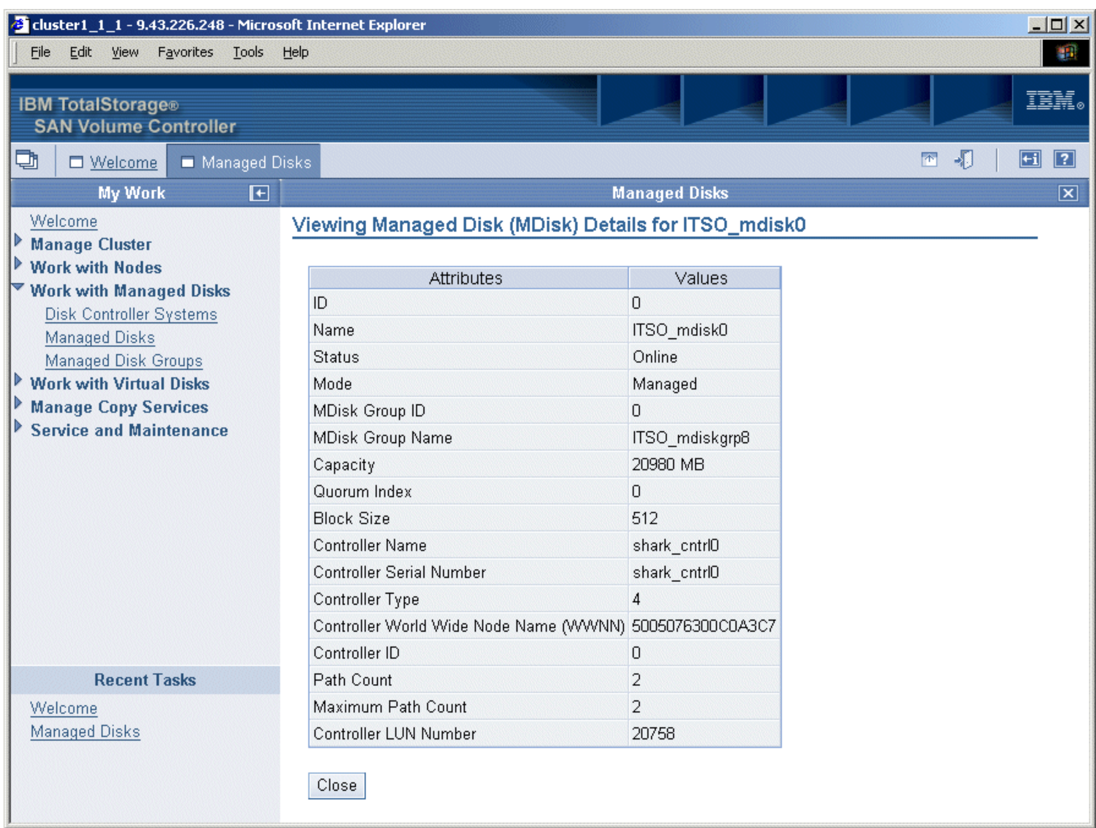


Figure 9-30 Managed disk details

> **Tip:** If at any time, the content in the right side of frame is “cut off”, you can minimize the My Work column by clicking the arrow to the right of the My Work heading at the top right of column (highlighted with mouse pointer in Figure 9-29).
>
> After you minimize the column, you see an arrow in the far left position in the same location where the My Work column formerly appeared. See Figure 9-31.

2. Review the details and then click **Close** to return to the previous panel.

Renaming an mDisk

Perform the following steps to rename an mDisk controlled by the SVC cluster:

1. Select the radio button to the left of the mDisk that you want to rename. Select **Rename an mDisk** from the list and click **Go**.
2. On the Renaming Managed Disk *mDiskname* panel (where *mDiskname* is the mDisk you selected in the previous step), type the new name you want to assign to the mDisk and click **OK**. See Figure 9-31.

> **Note:** The name can consist of letters A to Z, a to z, numbers 0 to 9, and the underscore. It can be between one and 15 characters in length. However, it cannot start with a number or the word *mDisk* because this prefix is reserved for SVC assignment only.

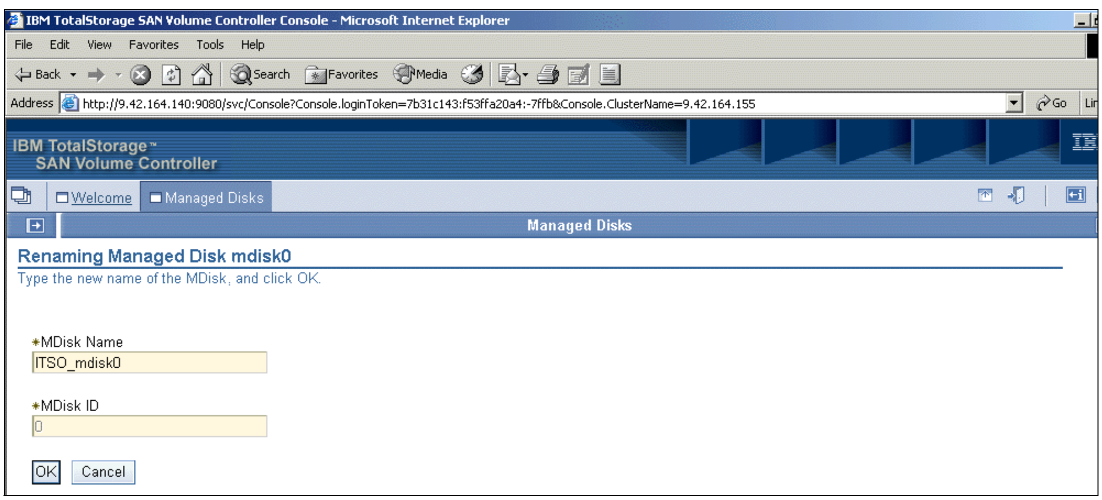


Figure 9-31 Renaming an mDisk

Discovering mDisks

Perform the following steps to discover newly assigned mDisks:

1. Select **Discover mDisks** from the list and click **Go**.
2. Any newly assigned mDisks are now displayed.

Setting a quorum disk

The SVC cluster, after the process of node discovery, automatically chooses three mDisks as quorum disks. Each disk is assigned an index number of either 0, 1, or 2.

In the event that half the nodes in a cluster are missing for any reason, the other half cannot simply assume that the nodes are “dead”. It can simply mean that the cluster state information is not being successfully passed between nodes for some reason (for example, network failure). For this reason, if half the cluster disappears from the view of the other, each surviving half attempts to lock the first quorum disk (index 0). In the event of quorum disk index 0 not being available, the next disk (index 1) becomes the quorum, and so on.

The half of the cluster that is successful in locking the quorum disk becomes the exclusive processor of I/O activity. It attempts to reform the cluster with any nodes it can still see.

If for any reason you want to set your own quorum disks, complete the following tasks:

1. Select the radio button to the left of the mDisk that you want to designate as a quorum. Then select **Set a quorum disk** from the list and click **Go**.
2. On the Setting a Quorum Disk: *mDiskname* panel (where *mDiskname* is the mDisk you selected in the previous step), assign a quorum index of 0, 1, or 2 and click **OK**.

Quorum disks are only created if at least one mDisk is in managed mode (that is, it was formatted by SVC with extents in it). Otherwise, a 1330 cluster error message is displayed in the SVC front panel. You can correct it only when you place mDisks in managed mode.

Including an mDisk

If a significant number of errors occurs on an mDisk, the SVC automatically excludes it. These errors can be from a hardware problem, a storage area network (SAN) zoning problem or the result of poorly planned maintenance. If it is a hardware fault, you should receive SNMP alerts in regard to the state of the hardware (before the disk were excluded) and undertaken preventative maintenance. If not, the hosts that were using vDisks, which used the excluded mDisk, now have I/O errors.

From the Viewing Managed Disks panel (Figure 9-32), you can see that mdisk5 is excluded.

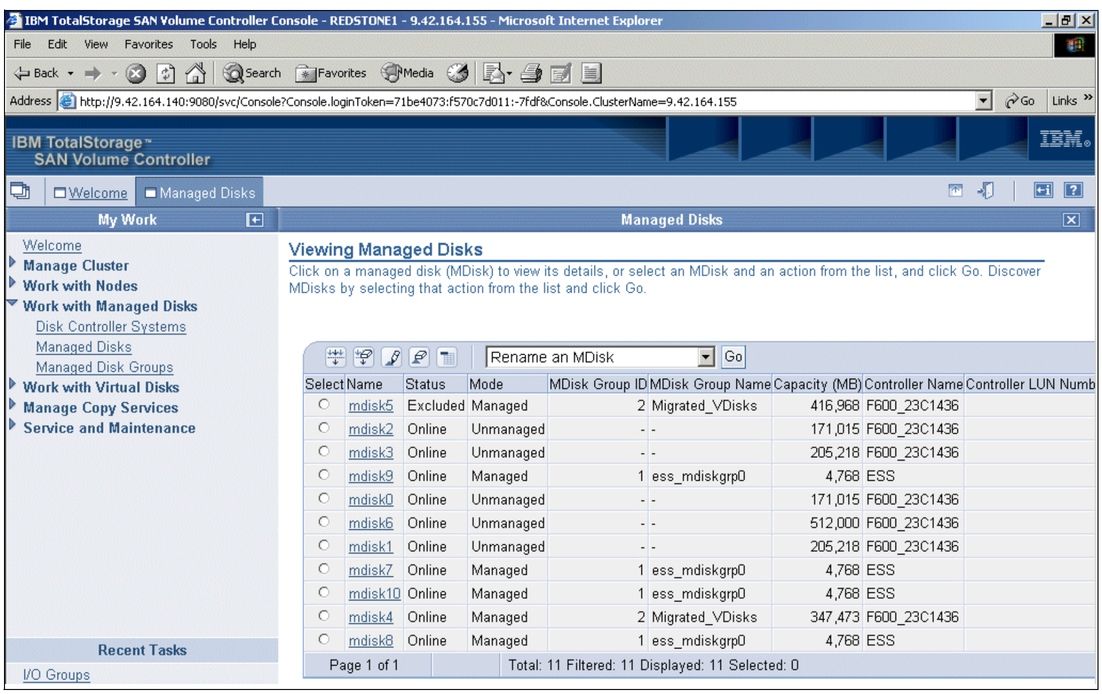


Figure 9-32 Viewing Managed Disks: Excluding an mDisk

After you take the necessary corrective action to repair the mDisk (for example, replace the failed disk, repair SAN zones), you can tell the SVC to include the mDisk again. Select the radio button to the left of the excluded mDisk. Then select **Include an mDisk** from the list and click **Go**. See Figure 9-33.

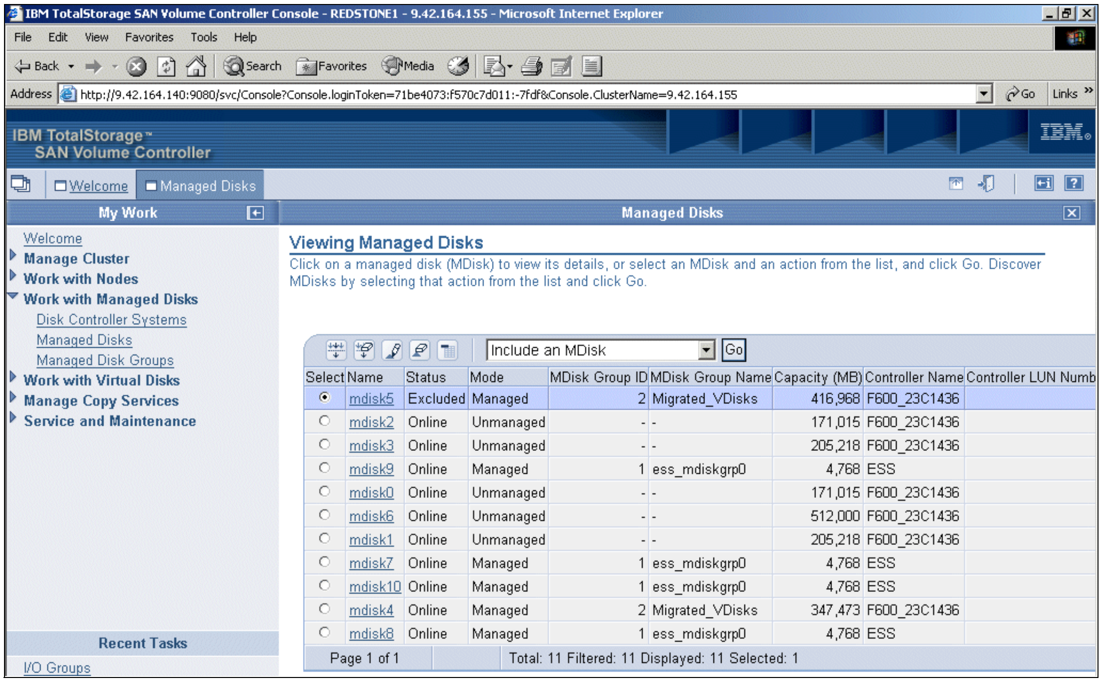


Figure 9-33 Including an mDisk

When you return to the Viewing Managed Disks panel (Figure 9-34), you see that mdisk5 is now back in an online state.

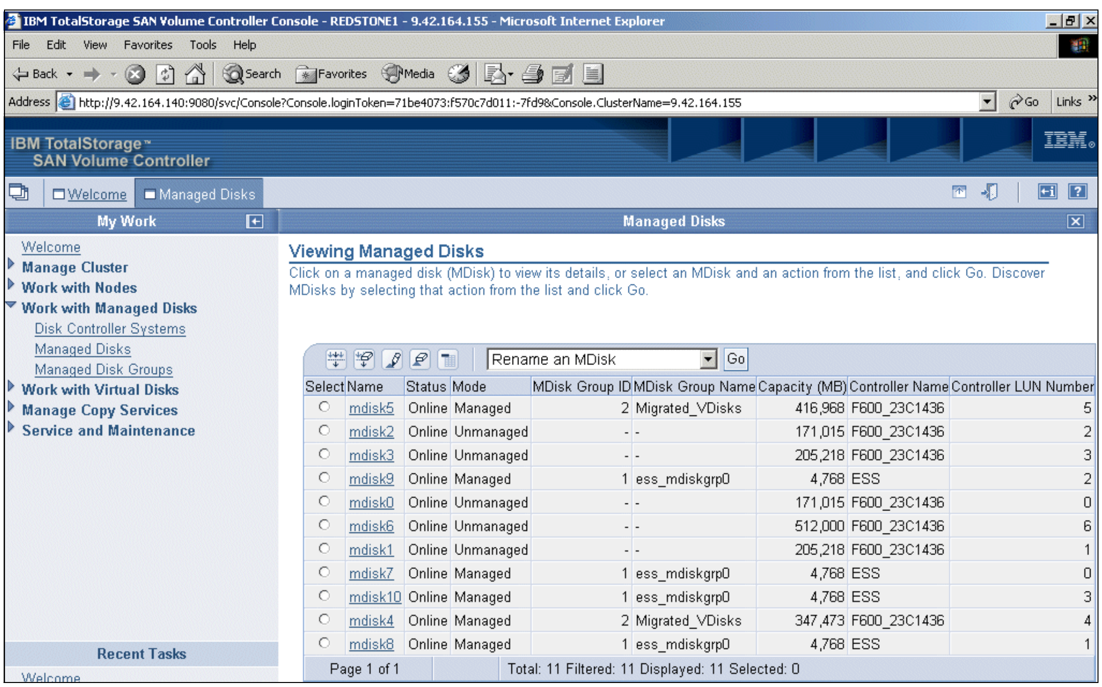


Figure 9-34 Viewing Managed Disks: Verifying the included mDisk

Note: At the time of writing this redbook, the Include an mDisk function did not work in the GUI. However, you can include the mDisk using the CLI.

Showing an mDisk group

To display information about the managed disk group (MDG) to which an mDisk belongs, perform the following steps:

1. Select the radio button to the left of the mDisk you want to obtain MDG information about. Select **Show mDisk Group** from the list and click **Go**.
2. You now see a subset (specific to the mDisk you chose in the previous step) of the View MDGs panel. We cover the View MDGs panel in more detail in 9.3.3, “MDGs” on page 247.

Showing a vDisk for an mDisk

To display information about vDisks that reside on an mDisk, perform the following steps:

1. Select the radio button to the left of the mDisk you want to obtain vDisk information about. Select **Show vDisk for mDisk** from the list and click **Go**.
2. You now see a subset (specific to the mDisk you chose in the previous step) of the View Virtual Disks panel. We cover the View Virtual Disks panel in more detail in 9.4, “Working with virtual disks” on page 255.

Creating a vDisk in image mode

An image mode disk is a vDisk that has an exact one-to-one (1:1) mapping of vDisk extents with the underlying mDisk. For example, extent 0 on the vDisk contains the same data as extent 1 on the mDisk and so on. Without this 1:1 mapping (for example, if extent 0 on the vDisk mapped to extent 3 on the mDisk), there is little chance that the data on a newly introduced mDisk is still readable.

Image mode is intended for the purpose of migrating data from an environment without the SVC to an environment with the SVC. A LUN that was previously directly assigned to a SAN-attached host can now be reassigned to the SVC (during a short outage) and returned to the same host as an image mode vDisk, with the user’s data intact. During the same outage the host, cables and zones can be reconfigured to access the disk, now via the SVC.

After access is re-established, the host workload can resume while the SVC manages the transparent migration of the data to other SVC managed mDisks on the same or another disk subsystem.

We recommend that, during the migration phase of the SVC implementation, you add one mDisk at a time to the SVC environment. This reduces the possibility of error. It also means that the short outages required to reassign the LUNs from the subsystem or subsystems and reconfigure the SAN and host can be staggered over a period of time to minimize the business impact.

Important: You can create an image mode vDisk only by using an unmanaged disk. That is, you must do this before you add the mDisk that corresponds to your original logical volume to a MDG.

To create an image mode vDisk, perform the following steps:

1. Select the radio button to the left of the unmanaged mDisk on which you want to create an image mode vDisk. Select **Create vDisk in image mode** from the list and click **Go**.

2. The Image Mode vDisk wizard (Figure 9-35) opens. Type the name you want to assign to the vDisk (`SAP_DATA_01` in our example), the I/O group from the list of available I/O groups (`ITSO_grp0` is the only choice in our example), and the MDG to which you want to assign the vDisk. Click **Next**.

> **Note:** If you do not provide a name, the SVC automatically generates the name *vdiskX* (where *X* is the ID sequence number assigned by the SVC internally).
>
> If you want to provide a name, you can use letters A to Z, a to z, numbers 0 to 9, and the underscore. It can be between one and 15 characters in length, but cannot start with a number or the word *vdisk* because this prefix is reserved for SVC assignment only.

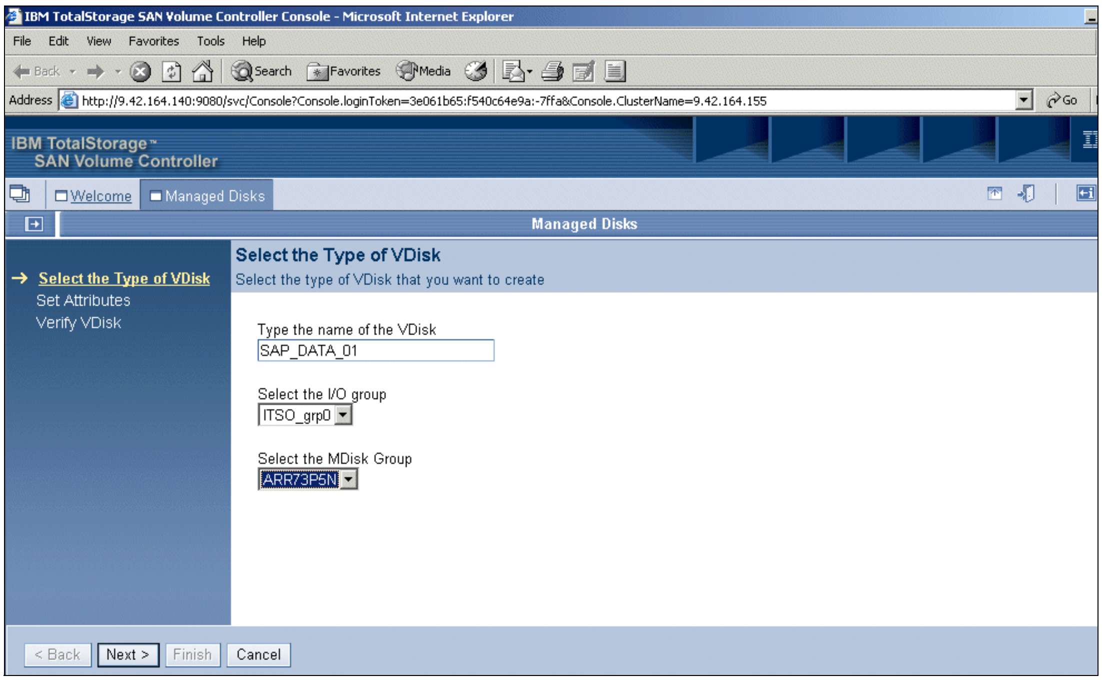


Figure 9-35 Type of vDisk for creating an image mode vDisk

3. On the next panel (Figure 9-36), select the node within the I/O group that you want to be the preferred owner of your vDisk. Click **Next**.

> **Note:** Not selecting a node (leaving the selection blank) allows the SVC to decide which node should be preferred based on the quantity of vDisks already assigned to each node within the I/O group.

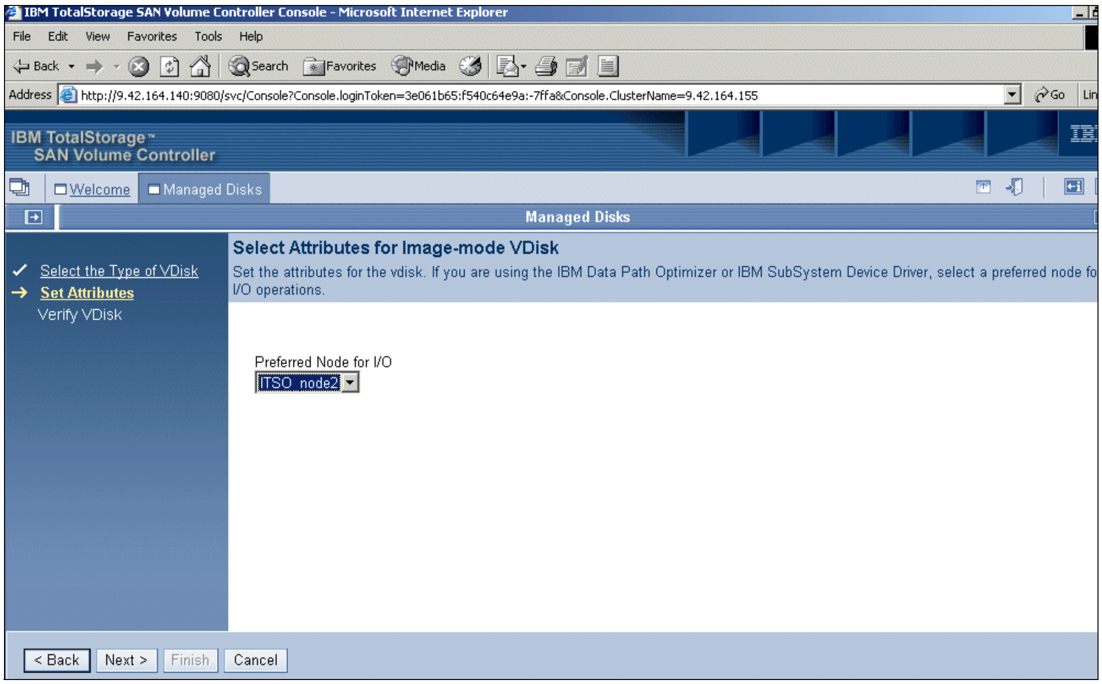


Figure 9-36 Preferred node for creating an image mode vDisk

4. You now see the final confirmation panel (Figure 9-37). Review and revise any of the specified details. When you are done, click **Finish**.

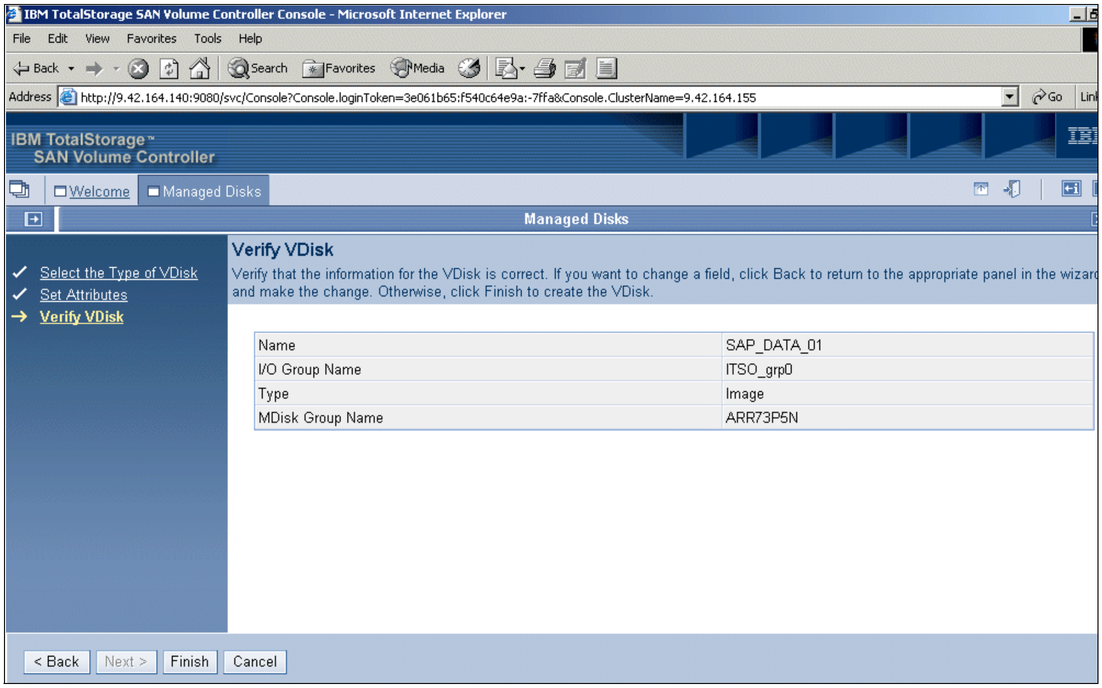


Figure 9-37 Finishing the creation of an image mode vDisk

9.3.3 MDGs

This section details the tasks that can be performed on a MDG level. Each of the following tasks are performed from the View Managed Disk Groups panel (Figure 9-38). To access this panel, from the SVC Welcome page, click the **Work with Managed Disks** option and then the **Managed Disk Groups** link.

> **Note:** At the Filtering Managed Disk (mDisk) Groups panel, click **Bypass filter**.

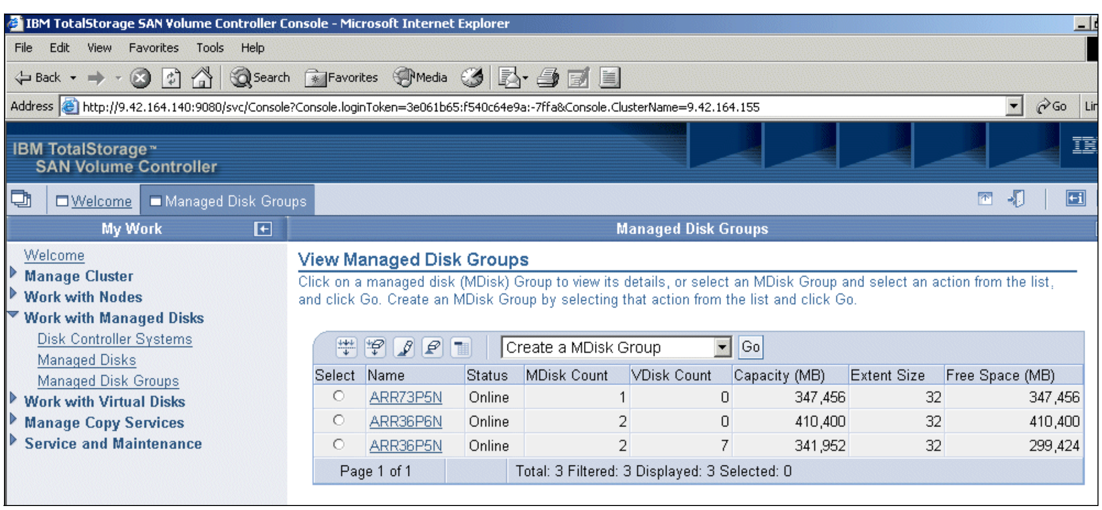


Figure 9-38 Viewing MDGs

Viewing mDisk group information

To retrieve information about a specific MDG, perform the following steps:

1. On the Viewing Managed Disk Groups panel (Figure 9-38), click the name of any MDG in the list.
2. On the View mDisk Group Details panel (Figure 9-39), you see more detailed information about the specified mDisk. Here you see information pertaining to the number of mDisks and vDisks as well as the capacity (both total and free space) within the MDG. When you finish viewing the details, click **Close** to return to the previous panel.

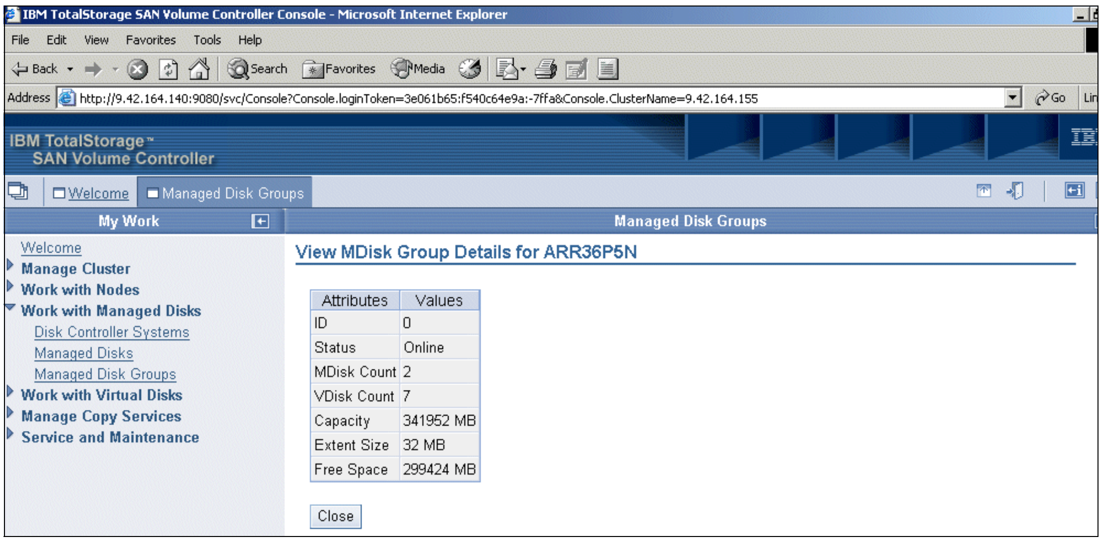


Figure 9-39   MDG details

Creating an mDisk group

To create a MDG, perform the following steps:

1. Select **Create an mDisk group** from the list and click **Go**.
2. On the Create Managed Disk Group wizard panel (Figure 9-40), click **Next**.

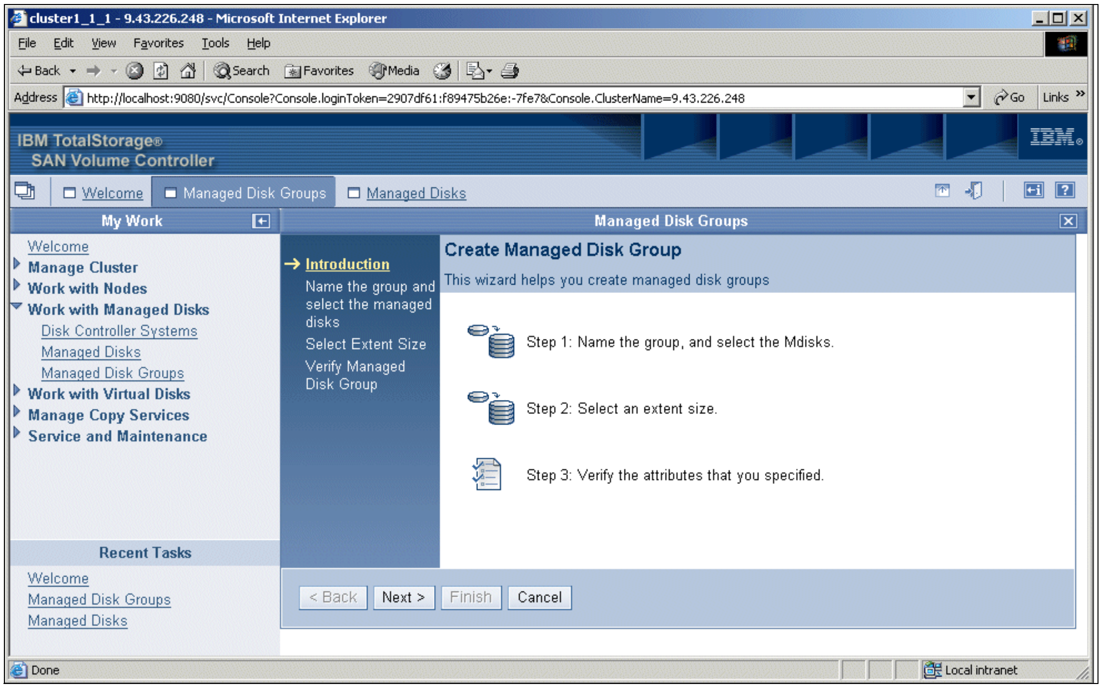


Figure 9-40 Create MDG wizard

3. On the Name the group and select the managed disks panel (Figure 9-41), give your MDG a name. Optionally select the mDisk Candidates and add them to the Selected mDisks list (one at a time) in the desired order.

   Selecting **No mDisk Candidates** creates an “empty” MDG. You can add mDisks to an “empty” MDG at a later time.

   Click **Next**.

> **Note:** If you do not provide a name, the SVC automatically generates the name *mdiskgrpX* (where *X* is the ID sequence number assigned by the SVC internally).
>
> If you want to provide a name, you can use letters A to Z, a to z, numbers 0 to 9, and the underscore. It can be between one and 15 characters in length. However, it cannot start with a number or the word *mdiskgrp* because this prefix is reserved for SVC assignment only.

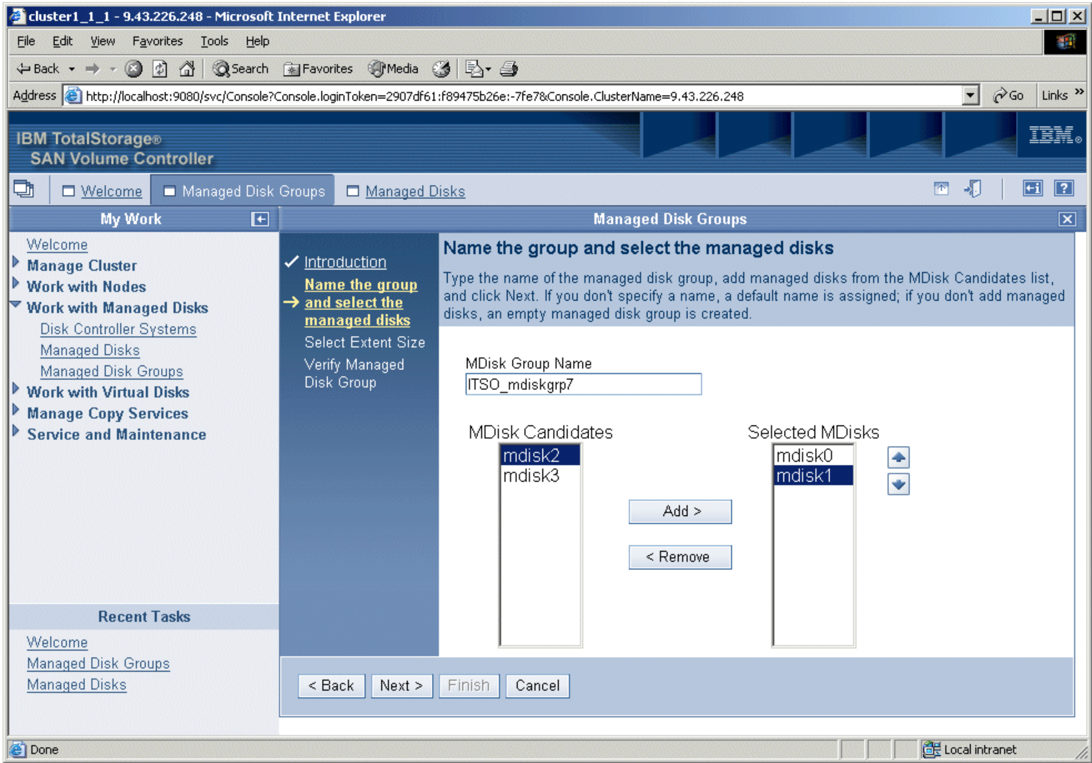


Figure 9-41 Name the group and select the managed disks

4. On the Select Extent Size panel (Figure 9-42), select from the list the size of the extents you want to format your MDG. Click **Next**.

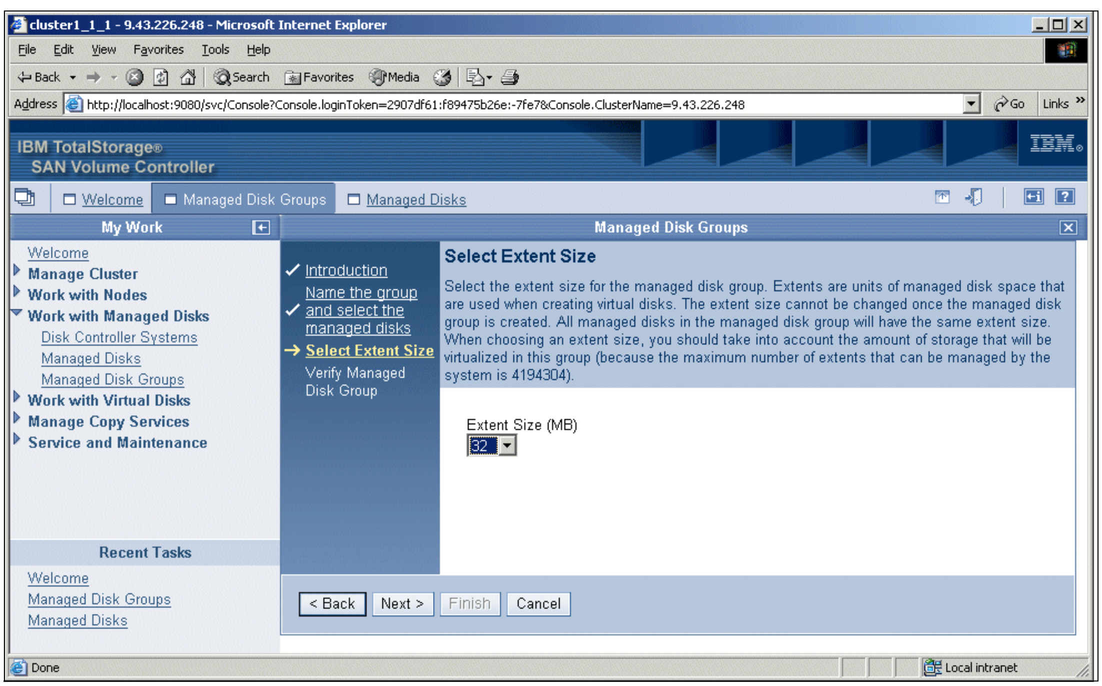


Figure 9-42 Select Extent Size for the MDG

5. Verify the information that you specified in the previous panels, as shown in Figure 9-43. If it is correct, click **Finish**. If you need to correct something, click **Back**.

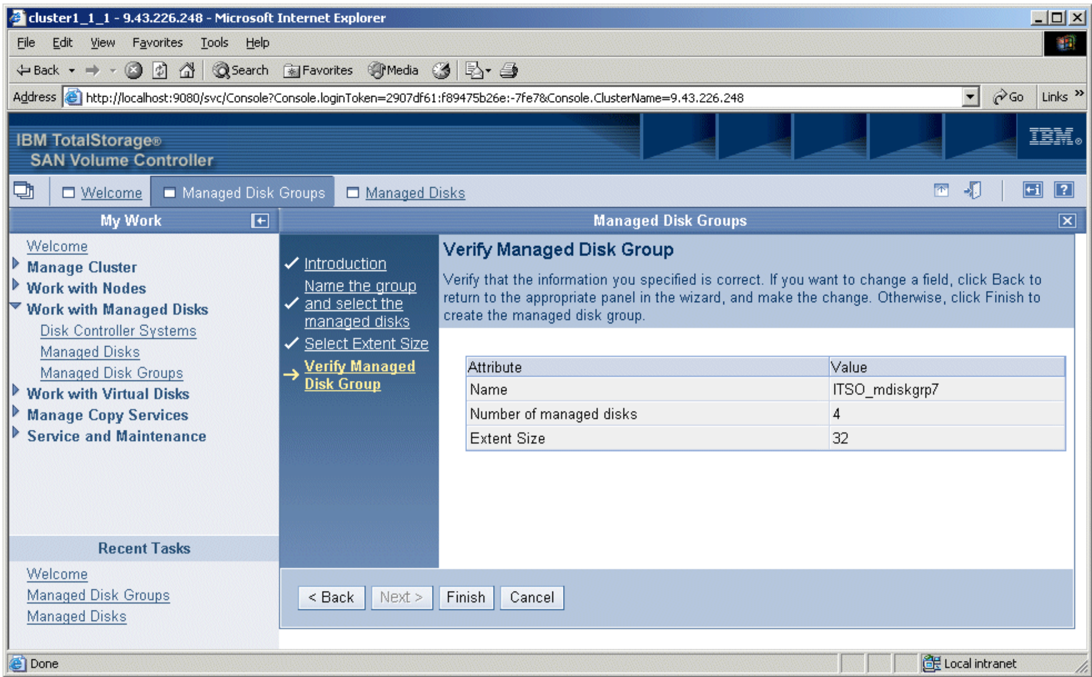


Figure 9-43 Verifying the information about the MDG

Renaming an mDisk group

To rename a MDG, perform the following steps:

1. Select the radio button to the left of the MDG you want to rename. Select **Rename an mDisk Group** from the list and click **Go**.
2. From the Renaming MDG *MDGname* panel (where *MDGname* is the MDG you selected in the previous step), type the new name you want to assign and click **OK**. See Figure 9-44.

> **Note:** The name can consist of letters A to Z, a to z, numbers 0 to 9, and the underscore. It can be between one and 15 characters in length, but cannot start with a number or the word *mdiskgrp* because this prefix is reserved for SVC assignment only.

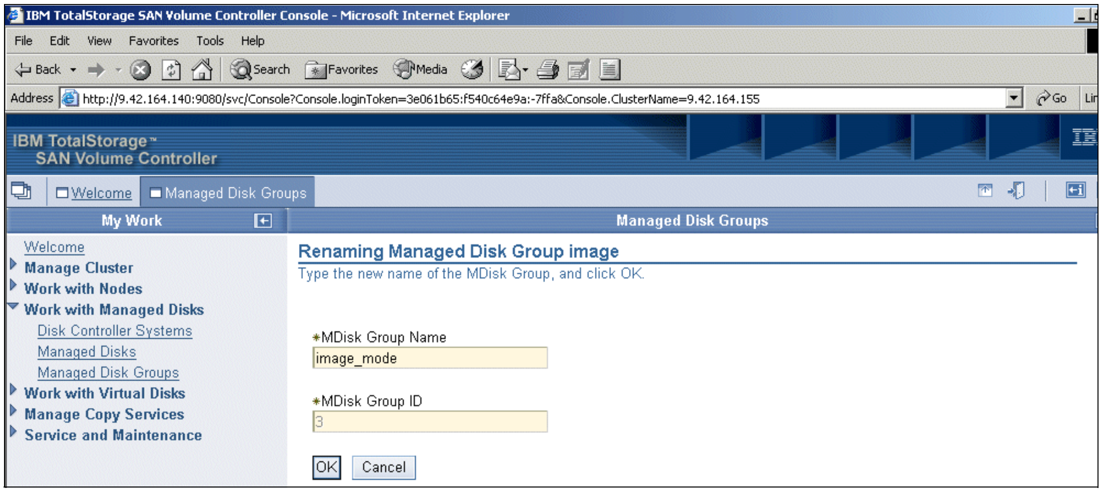


Figure 9-44 Renaming a MDG

Deleting an mDisk group

To delete a MDG, perform the following steps:

1. Select the radio button to the left of the MDG you want to delete. Select **Delete an mDisk Group** from the list and click **Go**.
2. On the Deleting a Managed Disk (mDisk) Group: *MDGname* panel (where *MDGname* is the MDG you selected in the previous step), click **OK** to confirm that you want to delete the MDG. See Figure 9-45.

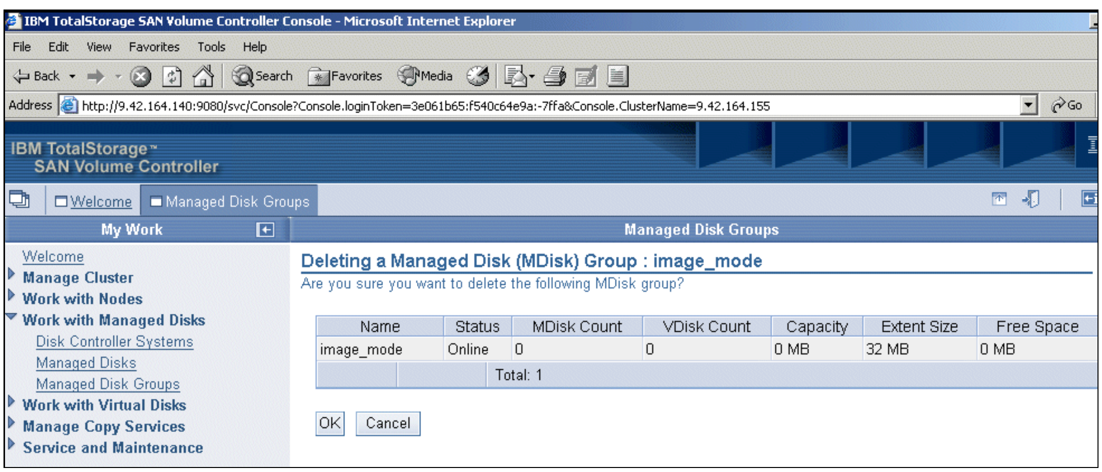


Figure 9-45 Deleting a MDG

If there are mDisks and vDisks within the MDG you are deleting, you are required to select **Forced deletion** for the MDG as shown in Figure 9-46.

> **Attention:** If you delete a MDG with the Forced deletion option, and vDisks were associated with that mDisk group, you lose the data on your vDisks, since they it is deleted before the mDisk Group. If you want to save your data, migrate the vDisks to another mDisk group before you delete the mDisk group previously assigned to it.

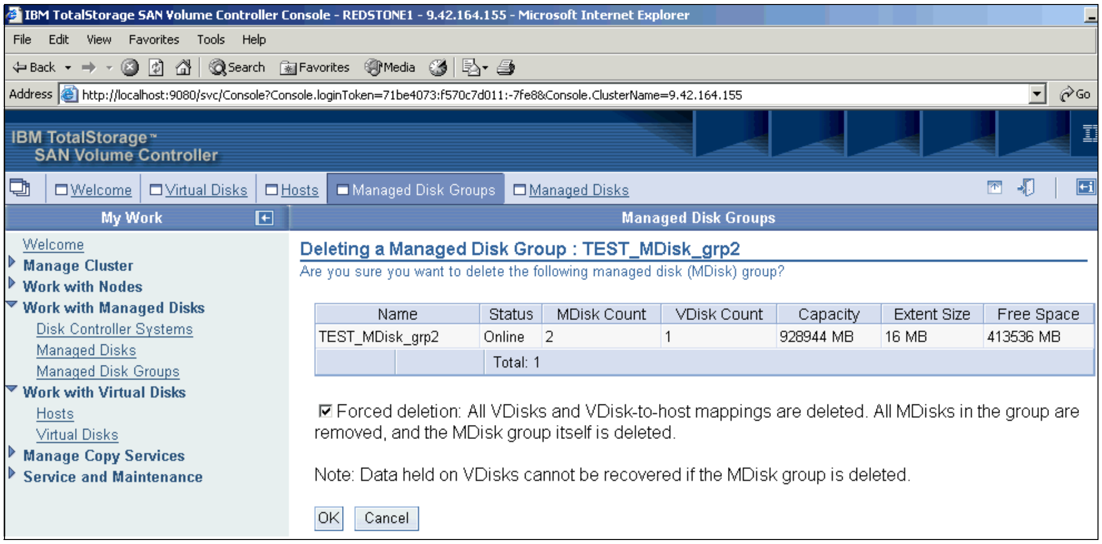


Figure 9-46 Confirming forced deletion of a MDG

Adding mDisks

If you created an "empty" MDG as we did or you simply assign additional mDisks to your SVC environment later, you can add mDisks to existing MDGs by performing the following steps:

> **Note:** You can only add unmanaged mDisks to a MDG.

1. Select the radio button to the left of the MDG to which you want to add mDisks. Select **Add mDisks** from the list and click **Go**.
2. From the Adding Managed Disks to an mDisk Group *mDiskname* panel (where *mDiskname* is the MDG you selected in the previous step), select the desired mDisk from the mDisk Candidates list (one at a time) and click **Add** to move it to the Chosen mDisks list. After you select all the desired mDisks, click **OK**. See Figure 9-47.

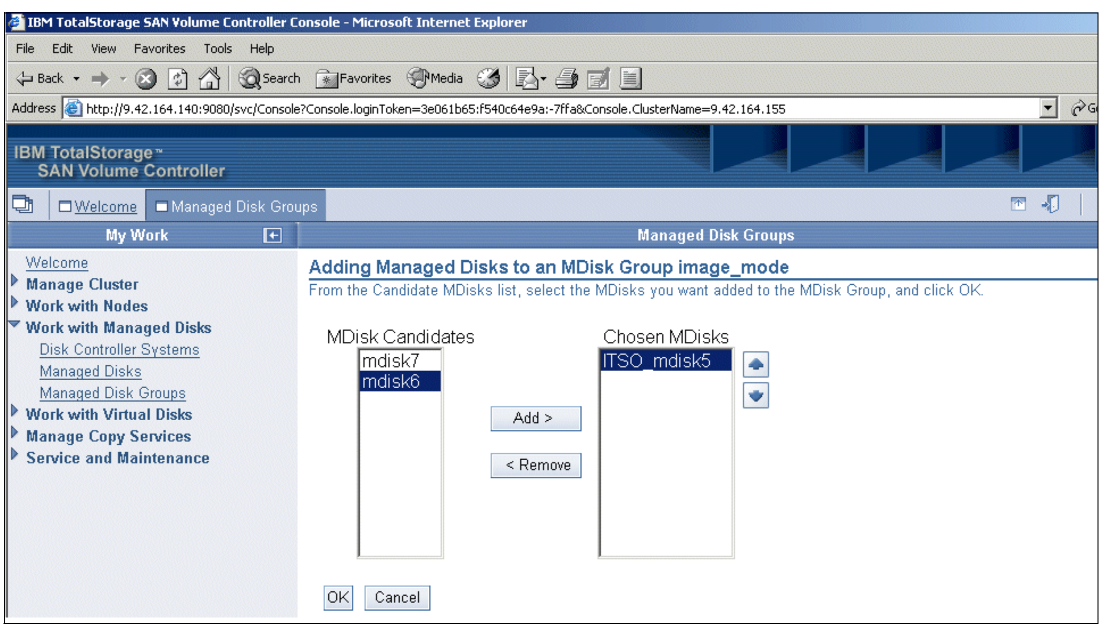


Figure 9-47 Adding mDisks to a MDG

Removing mDisks

To remove an mDisk from a MDG, perform the following steps:

1. Select the radio button to the left of the MDG from which you want to remove an mDisk. Select **Remove mDisks** from the list and click **Go**.
2. From the Deleting Managed Disks from MDG *MDGname* panel (where *MDGname* is the MDG you selected in the previous step), select the desired mDisk from the mDisk Candidates list (one at a time). Click **Add** to move the mDisk to the mDisks to Delete list. After you select all the desired mDisks, click **OK**. See Figure 9-48.

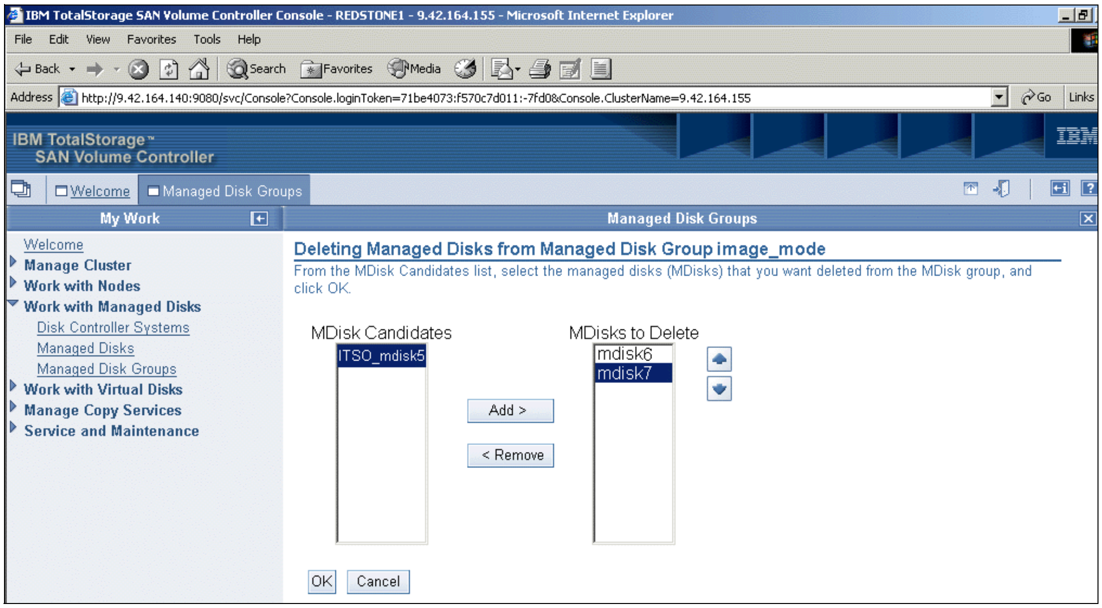


Figure 9-48 Removing mDisks from a MDG

If vDisks are using the mDisks that you are removing from the MDG, you are required to click the **Forced Delete** button to confirm the removal of the mDisk, as shown in Figure 9-49. Even then, the removal only takes place if there is sufficient space to migrate the vDisk data to other extents on other mDisks that remain in the MDG.

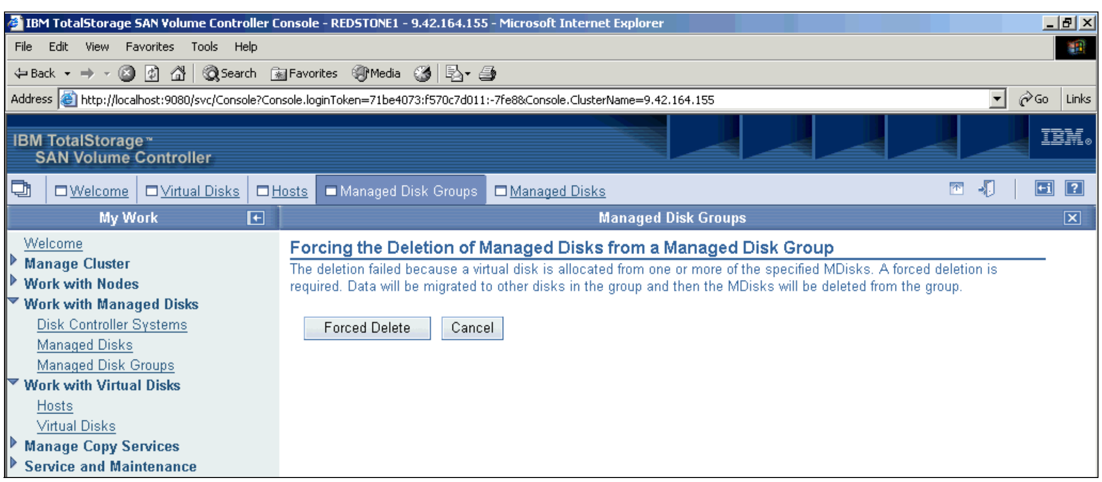


Figure 9-49 Confirming forced deletion of mDisk from MDG

Showing mDisks in this group

To show a list of mDisks within a MDG, perform the following steps:

1. Select the **radio button** to the left of the MDG from which you want to retrieve mDisk information. Select **Show mDisks in this group** from the list and click **Go**.
2. You now see a subset (specific to the MDG you chose in the previous step) of the Viewing Managed Disk panel (Figure 9-50) from 9.3.2, "Managed disks" on page 239.

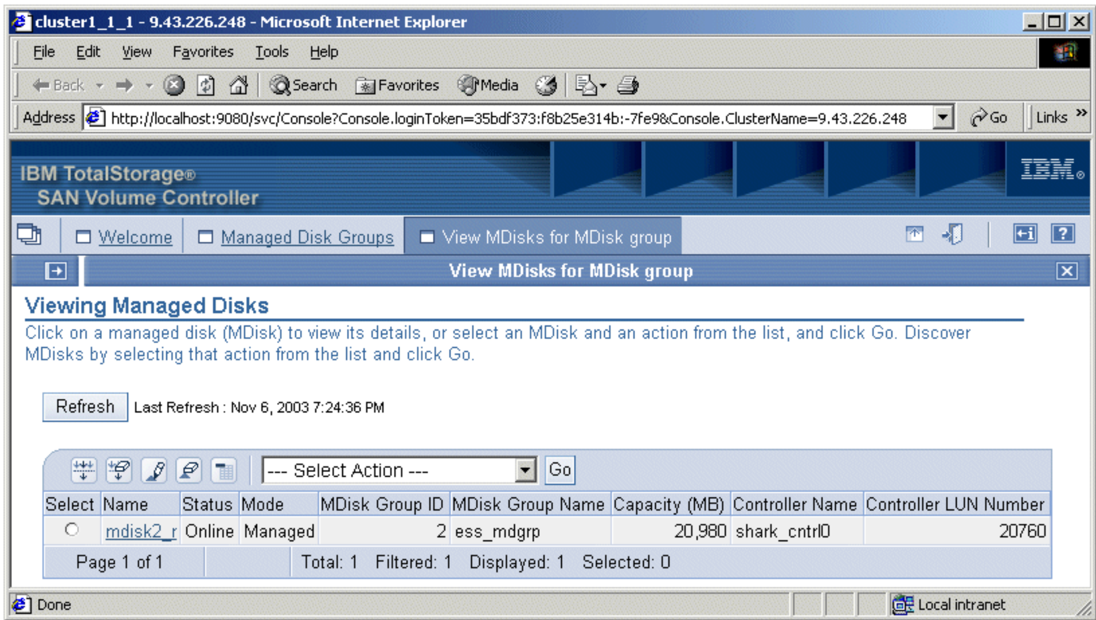


Figure 9-50 Viewing mDisks in a MDG

> **Note:** Remember, you can collapse the column titled My Work at any time by clicking the arrow to the right of the My Work column heading.

Showing vDisks using this group

To show a list of vDisks associated with mDisks within a MDG, perform the following steps:

1. Select the radio button to the left of the MDG from which you want to retrieve vDisk information. Select **Show vDisks using this group** from the list and click **Go**.
2. You see a subset (specific to the MDG you chose in the previous step) of the Viewing Virtual Disks panel. We cover the Viewing Virtual Disks panel in more detail in "vDisk information" on page 262.

You have now completed the tasks required to manage the disk controller systems, managed disks, and MDGs within the SVC environment.

9.4 Working with virtual disks

This section details the various configuration and administration tasks that you can perform on the vDisks within the SVC environment.

9.4.1 Hosts

This section details the tasks that you can perform on a host level. Each of the following tasks are performed from the Viewing Hosts panel (Figure 9-51). To access this panel, from the SVC Welcome page, click the **Work with Virtual Disks** option and then the **Hosts** link.

> **Note:** At the Filtering Hosts panel, click **Bypass filter**.

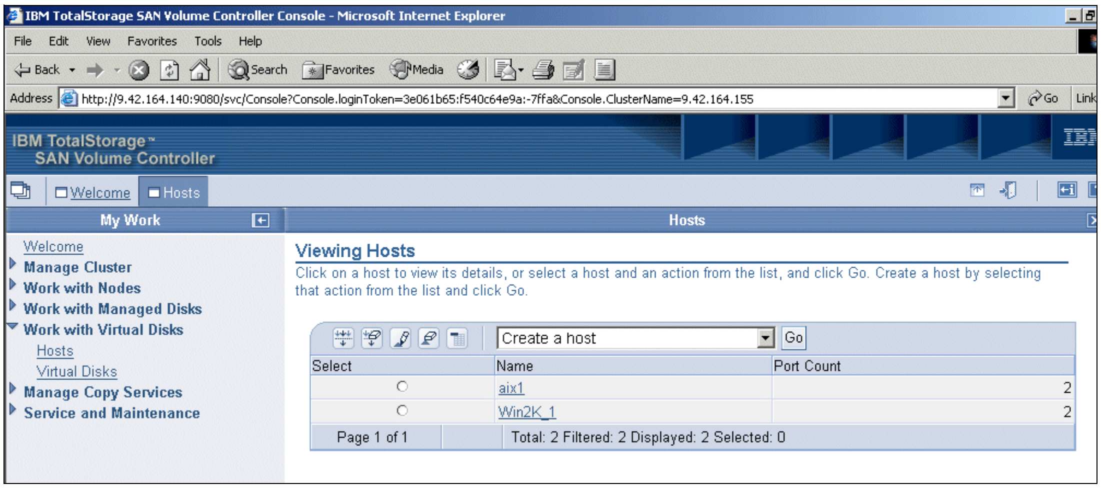


Figure 9-51 Viewing hosts

Host information

To retrieve information about a specific host, perform the following steps:

1. On the Viewing Hosts panel, click the name of any host in the list displayed.
2. On the Viewing General Details panel (Figure 9-52), you can see more detailed information about the specified host. You can click the **Port Details** link to see information about the Fibre Channel Host Bus Adapters (HBAs) that were defined within the host.

   When you are finished viewing the details, click **Cancel** to return to the previous panel.

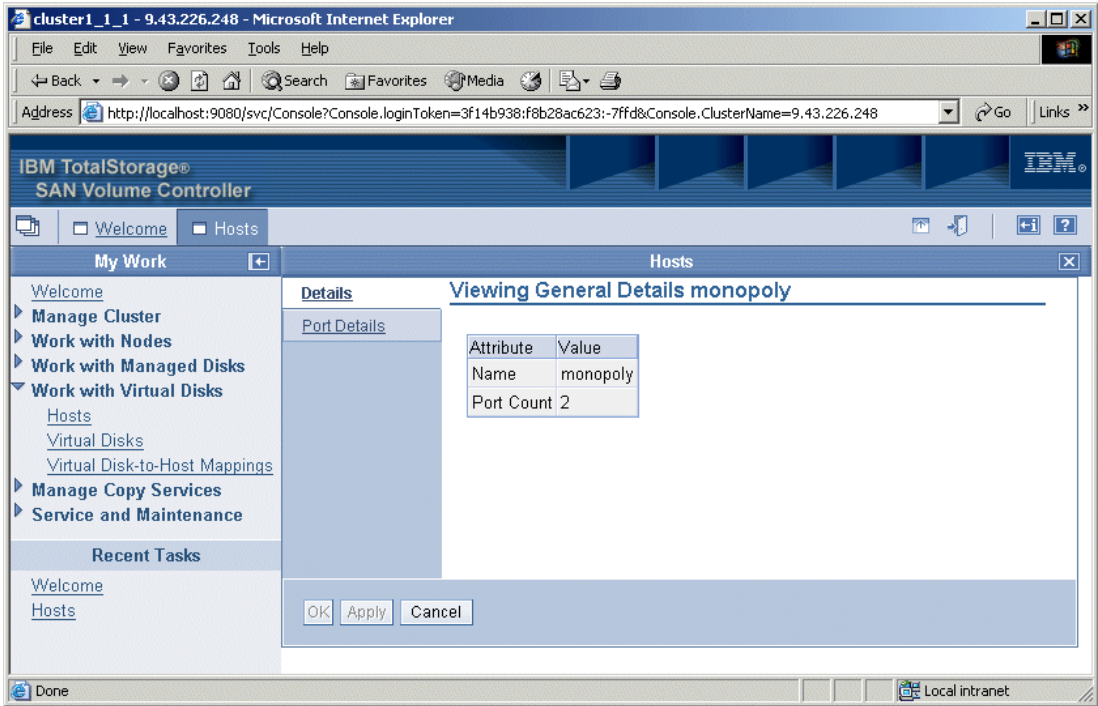


Figure 9-52 Host details

Creating a host

To create a new host, perform the following steps:

1. Select **Create a host** from the list and click **Go**.
2. On the Creating Hosts panel (Figure 9-53), type a name for your host.

> **Note:** If you do not provide a name, the SVC automatically generates the name *hostX* (where *X* is the ID sequence number assigned by the SVC internally).
>
> If you want to provide a name, you can use letters A to Z, a to z, numbers 0 to 9, and the underscore. It can be between one and 15 characters in length. However, it cannot start with a number or the word *host* because this prefix is reserved for SVC assignment only.

Select and add the worldwide port names (WWPNs) that correspond to your HBA or HBAs. Click **OK**.

Your WWPN or WWPNs may not display, although you are *sure* your adapter is functioning (for example, you see WWPN in the switch name server) and your zones are correctly setup. In this case, you can manually type the WWPN of your HBA or HBAs into the Additional Ports field (type in WWPNs, one per line) at the bottom of the panel before you click OK.

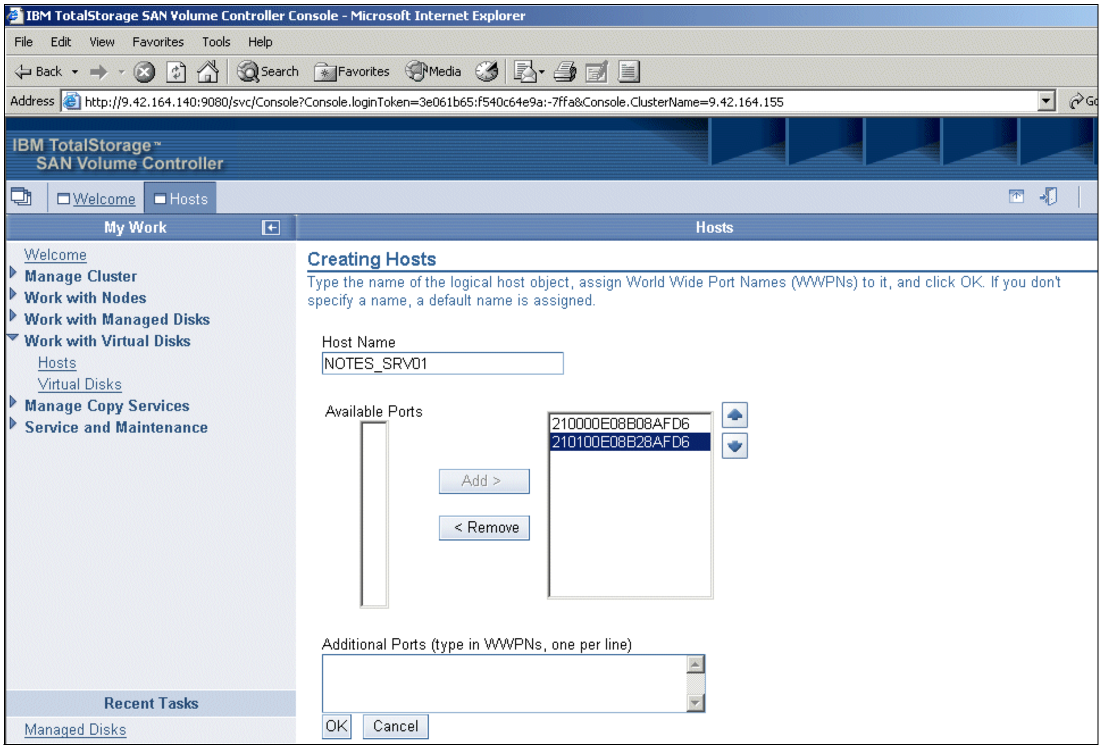


Figure 9-53 Creating a new host

Renaming a host

To rename a host, perform the following steps:

1. Select the radio button to the left of the host you want to rename. Select **Rename a host** from the list and click **Go**.
2. From the Renaming Host panel (Figure 9-54), type the new name you want to assign and click **OK**.

> **Note:** The name can consist of letters A to Z, a to z, numbers 0 to 9, and the underscore. It can be between one and 15 characters in length. However, it cannot start with a number or the word *host* because this prefix is reserved for SVC assignment only.

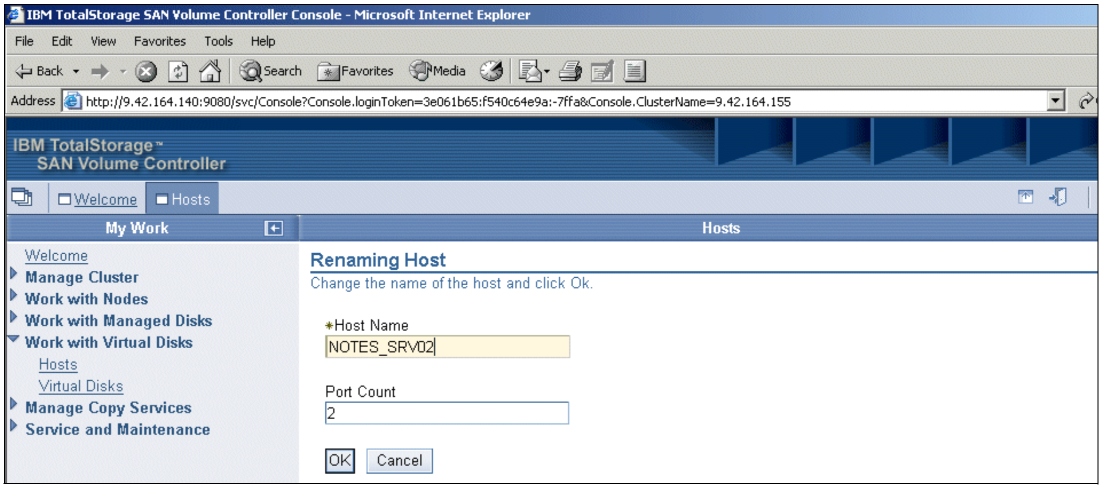


Figure 9-54 Renaming a host

Deleting a host

To delete a Host, perform the following steps:

1. Select the radio button to the left of the host you want to delete. Select **Delete a host** from the list and click **Go**.
2. On the Deleting Host *hostname* panel (where *hostname* is the host you selected in the previous step), click **OK** if you are sure you want to delete the host. See Figure 9-55.

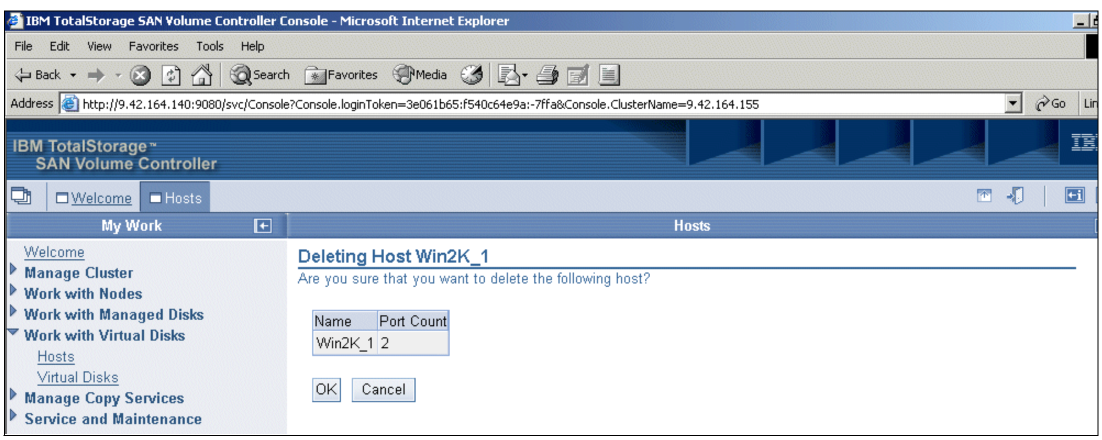


Figure 9-55 Deleting a host

> **Note:** If you still have vDisks associated with the host, you see a panel requesting confirmation for the forced deletion of the host. Click **OK** and all the mappings between this host and its vDisks are deleted before the host is deleted.

Adding ports

If you add a HBA to a server that is already defined within the SVC, you can simply add additional ports to your host definition by performing the following steps:

1. Select the radio button to the left of the host to which you want to add WWPNs. Select **Add ports** from the list and click **Go**.

2. From the Adding ports to *hostname* panel (where *hostname* is the host you selected in the previous step), select the desired **WWPN** from the Available Ports list (one at a time) and click **Add**. After you select all the desired WWPNs, click **OK**. See Figure 9-56.

   Your WWPN or WWPNs may not display, although you are *sure* your adapter is functioning (for example, you see WWPN in the switch name server) and your zones are correctly set up. In this case, you can manually type the WWPN of your HBA or HBAs into the Add Additional Ports field at the bottom of the panel before you click **OK**.

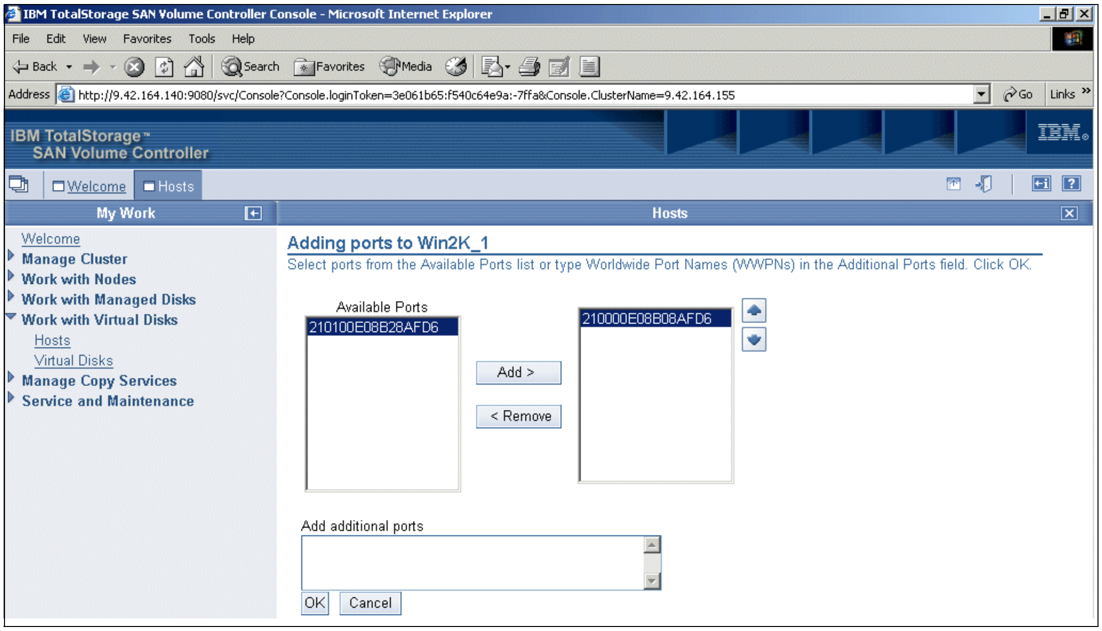


Figure 9-56 Adding ports to a host

Deleting ports

To delete a port from a host, perform the following steps:

1. Select the radio button to the left of the host from which you want to delete a port. Select **Delete ports** from the list and click **Go**.
2. On the Deleting Ports From *hostname* panel (where *hostname* is the host you selected in the previous step), select the ports you want to delete from the Available Ports list and click **Add**. When you select all the ports you want to delete from your host to the column to the right, click **OK**. See Figure 9-57.

> **Note:** If you have vDisks that are associated with the host, you receive a warning about deleting a host port. You need to confirm your action when prompted.

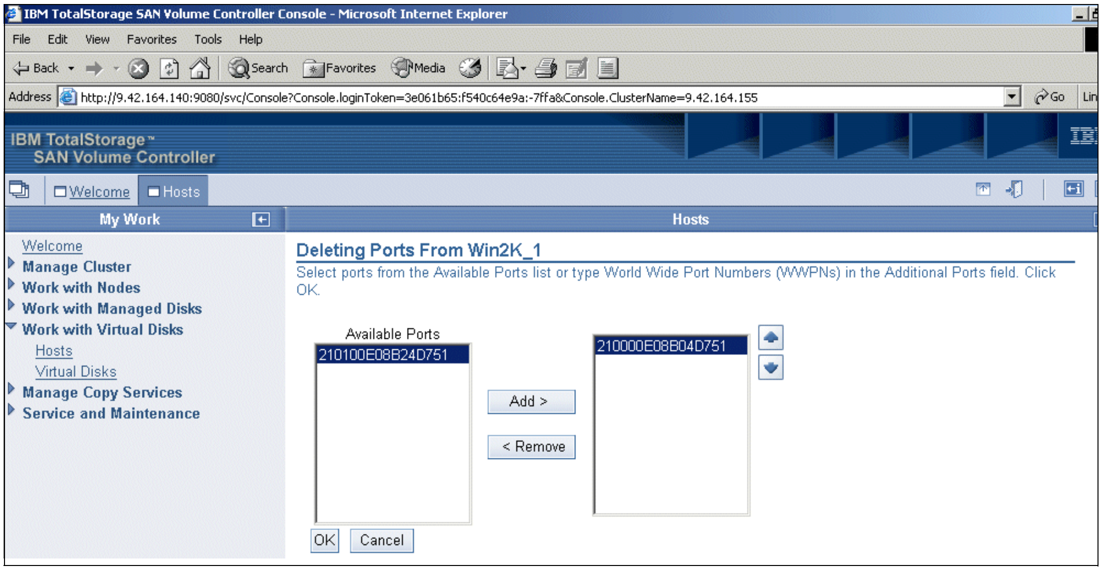


Figure 9-57 Deleting a port from a host

Showing vDisks mapped to a host

To show the vDisks assigned to a specific host, perform the following steps:

1. Select the radio button to the left of the host you want to view vDisk information about. Select **Show the vDisks mapped to this host** from the list and click **Go**.
2. You see a subset (specific to the host you chose in the previous step) of the Viewing Virtual Disks panel (Figure 9-58). You see that the vDisks, when accessed in this manner, are listed twice (once for each defined host port).

   We discuss the Viewing Virtual Disks panel in more detail in “vDisk information” on page 262.

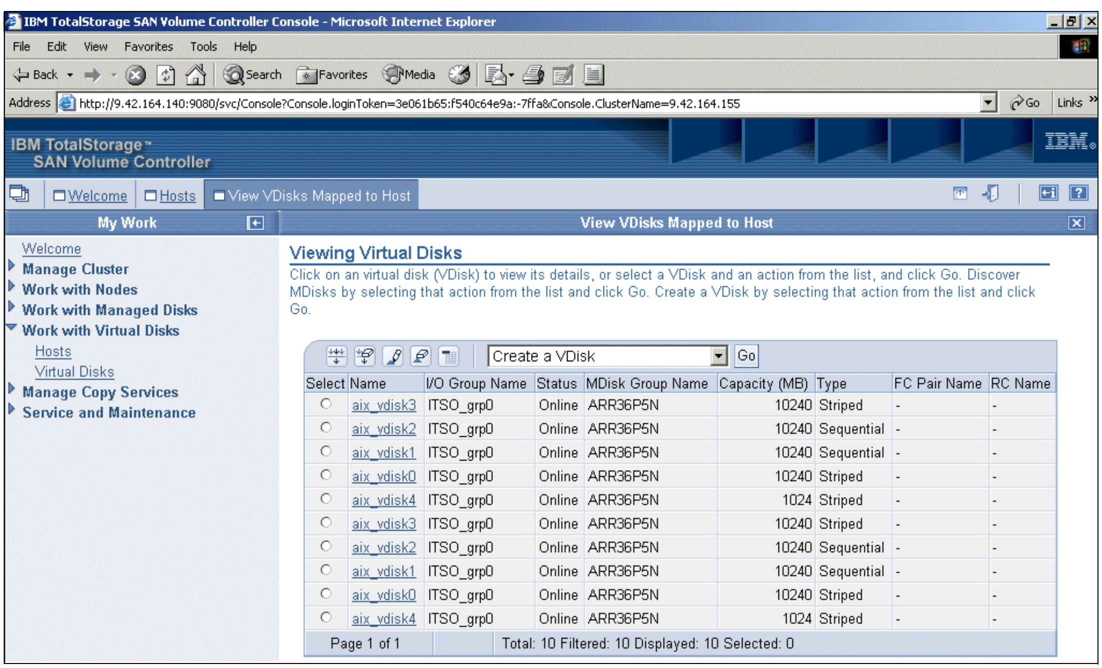


Figure 9-58 Showing the vDisk assignment for a host

9.4.2 Virtual disks

This section discusses the tasks that you can perform on a vDisk level. Each of the following tasks are performed from the Viewing Virtual Disks panel (Figure 9-59). To access this panel, from the SVC Welcome page, click the **Work with Virtual Disks** option and then the **Virtual Disks** link.

Note: At the Filtering Virtual Disks (vDisks) panel, click **Bypass filter**.

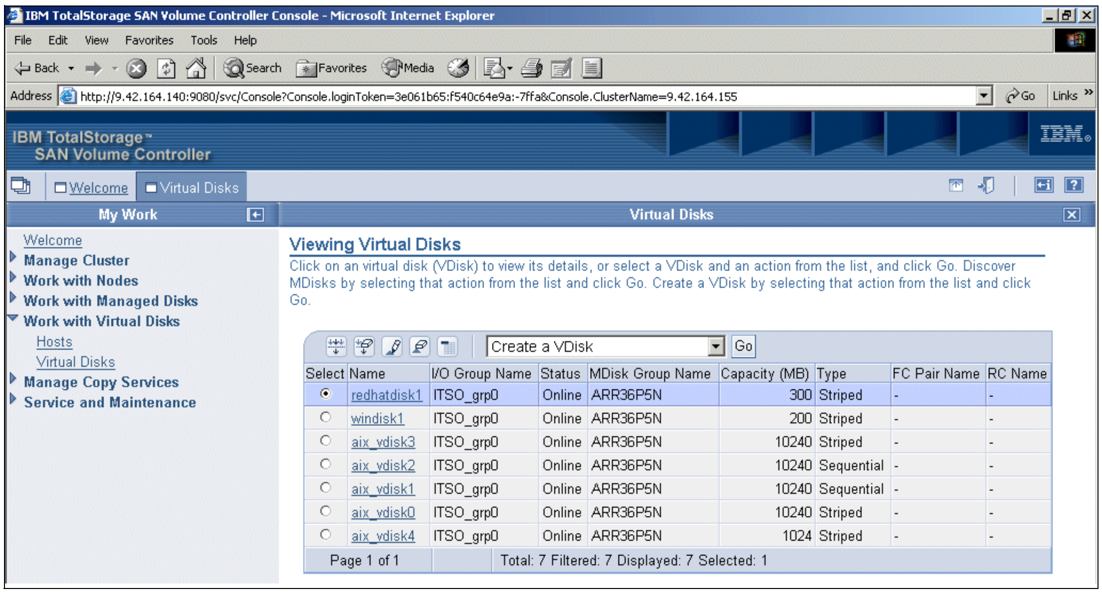


Figure 9-59 Viewing Virtual Disks

vDisk information

To retrieve information about a specific vDisk, perform the following steps:

1. On the Viewing Virtual Disks panel, click the name of the desired vDisk in the list.
2. The next panel (Figure 9-60) that opens shows you detailed information. Review the information. When you are done, click **Close** to return to the Viewing Virtual Disks panel.

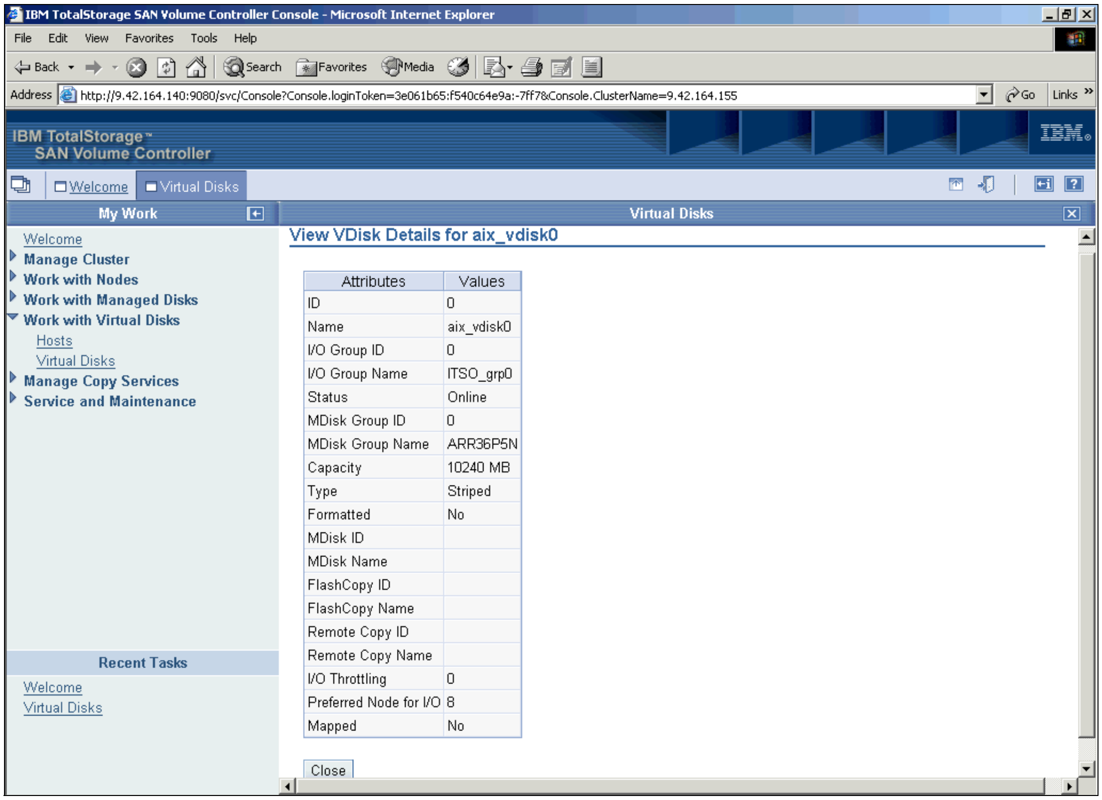


Figure 9-60 vDisk details

Creating a vDisk

To create a new vDisk, perform the following steps:

1. Select **Create a vDisk** from the list and click **Go**.
2. The Create Virtual Disks wizard launches (Figure 9-61). Click **Next**.

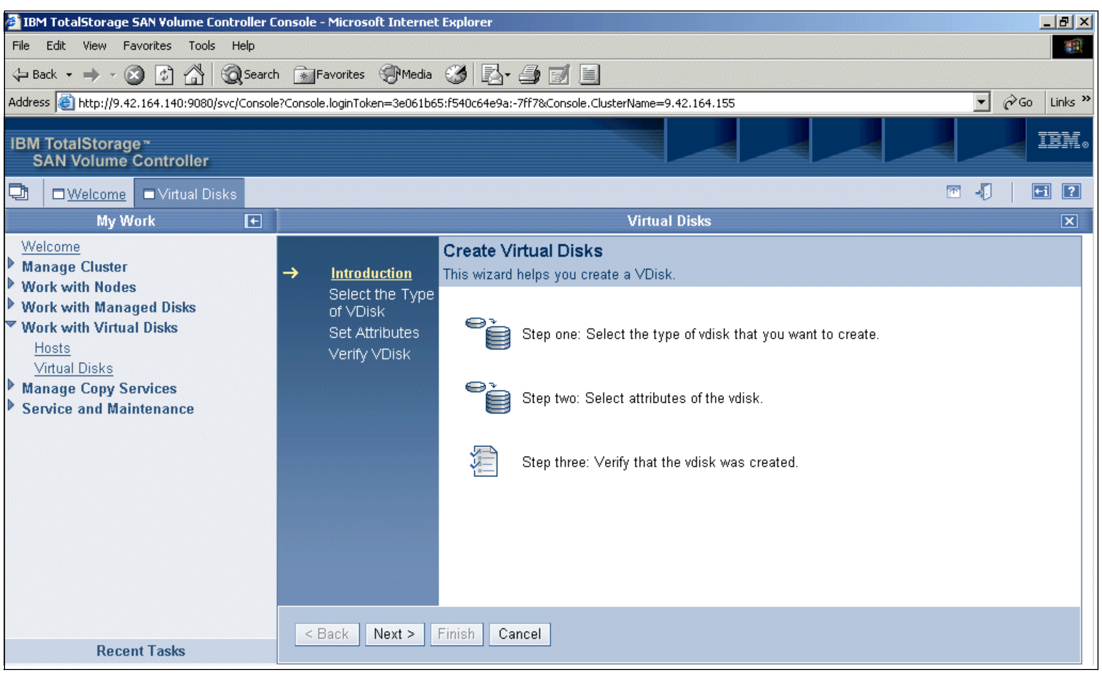


Figure 9-61 Creating a vDisk wizard: Introduction

3. The Select the Type of vDisk panel (Figure 9-62) opens.Type a name for your new vDisk, select the I/O group to which you want to assign the vDisk (in our example, `ITSO_grp0` is the only option). Select the **MDG** within which you want to create the vDisk. Select the type of **vDisk** you want to create (striped, sequential, or image mode). Click **Next**.

> **Note:** If you do not provide a name, the SVC automatically generates the name *vdiskX* (where *X* is the ID sequence number assigned by the SVC internally).
>
> If you want to provide a name, you can use letters A to Z, a to z, numbers 0 to 9, and the underscore. It can be between one and 15 characters in length, but cannot start with a number or the word *vdisk* because this prefix is reserved for SVC assignment only.

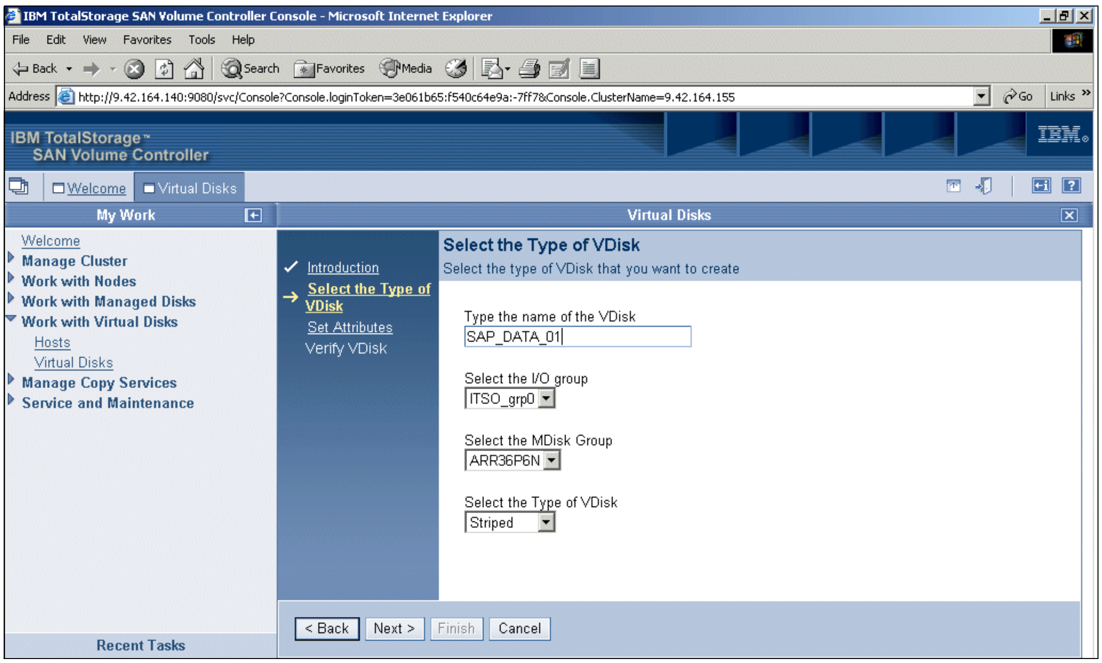


Figure 9-62 Creating a vDisk wizard: Select the Type of vDisk

4. The Select Attributes for *vDisktype* mode vDisk panel (where *vDisktype* is the type of vDisk you chose in the previous step) opens. Choose from the following options. After you select the appropriate options, depending on the type of the vDisk being created, click **Next**.

   – As shown in Figure 9-63, with a *striped vDisk*, you can:

     - Optionally, specify the preferred node to process the I/O requests for this vDisk. The default is to alternate between the nodes in the I/O group based on the quantities that are already assigned.
     - Optionally, from the Managed Disk Candidates list, select one by one a subset of the mDisks available to place your vDisk. Click **Add** to add them to the Managed Disks Striped in This Order list. When you are done, you can change the stripe order. Select an mDisk and click the up or down arrows to the right of the list. The default is to use all mDisks in the MDG.
     - Enter the size of the vDisk you want to create and select the capacity measurement (MB or GB) from the list.

Note: An entry of 1 GB uses 1024 MB.

- Optionally, format the new vDisk by selecting the **Format virtual disk (write zeros to its managed disk extents)** check box at the bottom of the panel.

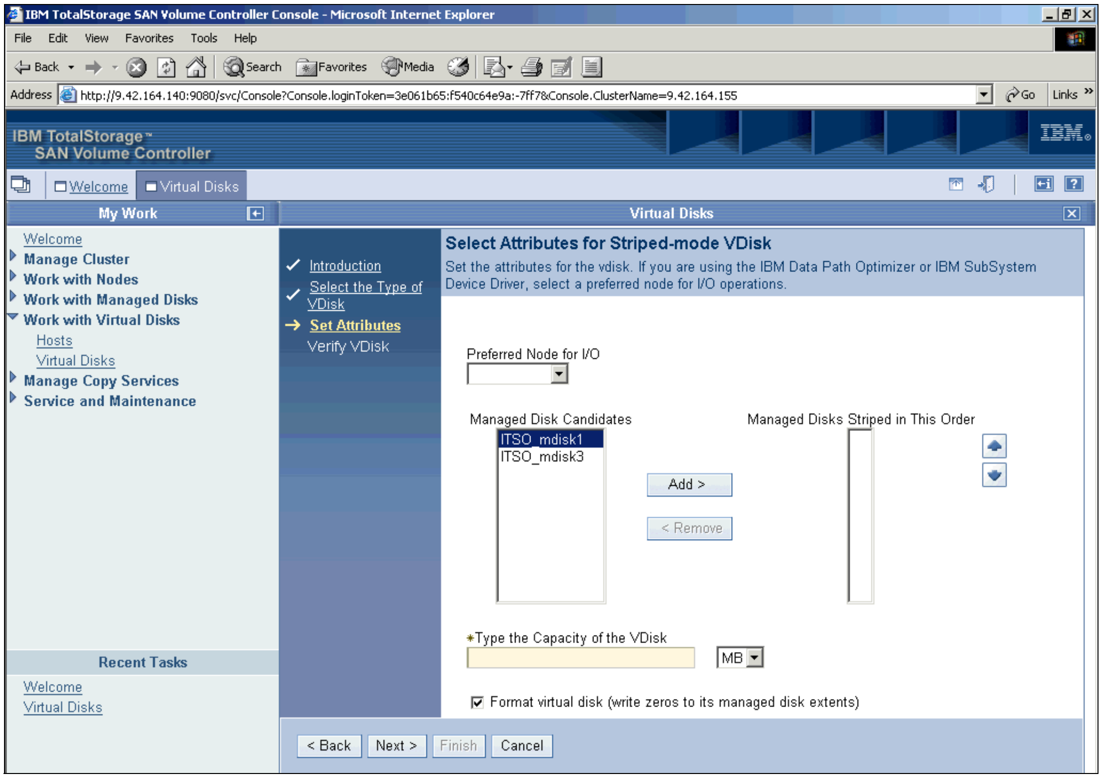


Figure 9-63 Creating a vDisk wizard: Select Attributes for Striped-mode vDisk

– As shown in Figure 9-64, with a *sequential vDisk*, you can:

- Optionally, specify the preferred node to process the I/O requests for this vDisk. The default it to alternate between the nodes in the I/O group based on the quantities that are already assigned.
- From the Managed Disk used to Create vDisk list, select the mDisk on which to place your vDisk.
- In the Type the Capacity of the vDisk field, enter the size of the vDisk you want to create. Select the capacity measurement (MB or GB) from the list.
- Optionally, format the new vDisk by selecting the **Format virtual disk (write zeros to its managed disk extensions)** check box at the bottom of the panel.

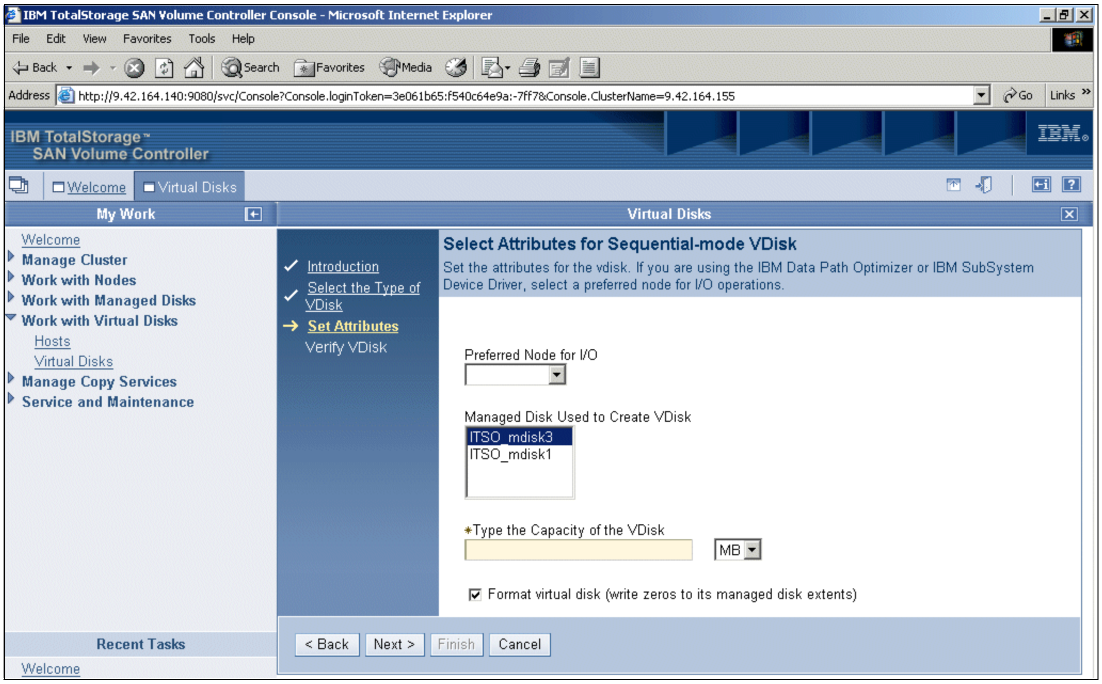


Figure 9-64 Creating a vDisk wizard: Select Attributes for Sequential-mode vDisk

– As shown in Figure 9-65, with an *image-mode vDisk*, you can:

- Optionally, specify the preferred node to process the I/O requests for this vDisk. The default it to alternate between the nodes in the I/O group based on quantities already assigned.
- From the Managed Disk Used to Create vDisk list, select the mDisk from which to create your image-mode vDisk.

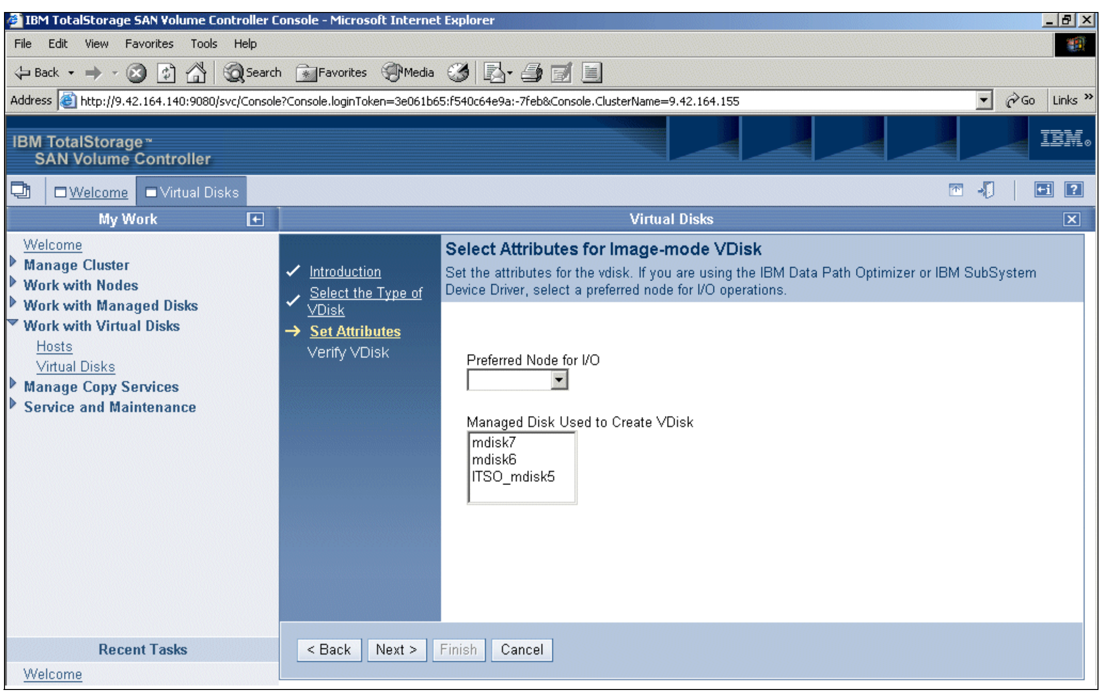


Figure 9-65 Creating a vDisk wizard: Select Attributes for Image-mode vDisk

5. On the Verify vDisk panel (Figure 9-66), if you are satisfied with the information shown, click **Finish** to complete the task. Otherwise, click **Back** to return and make any corrections.

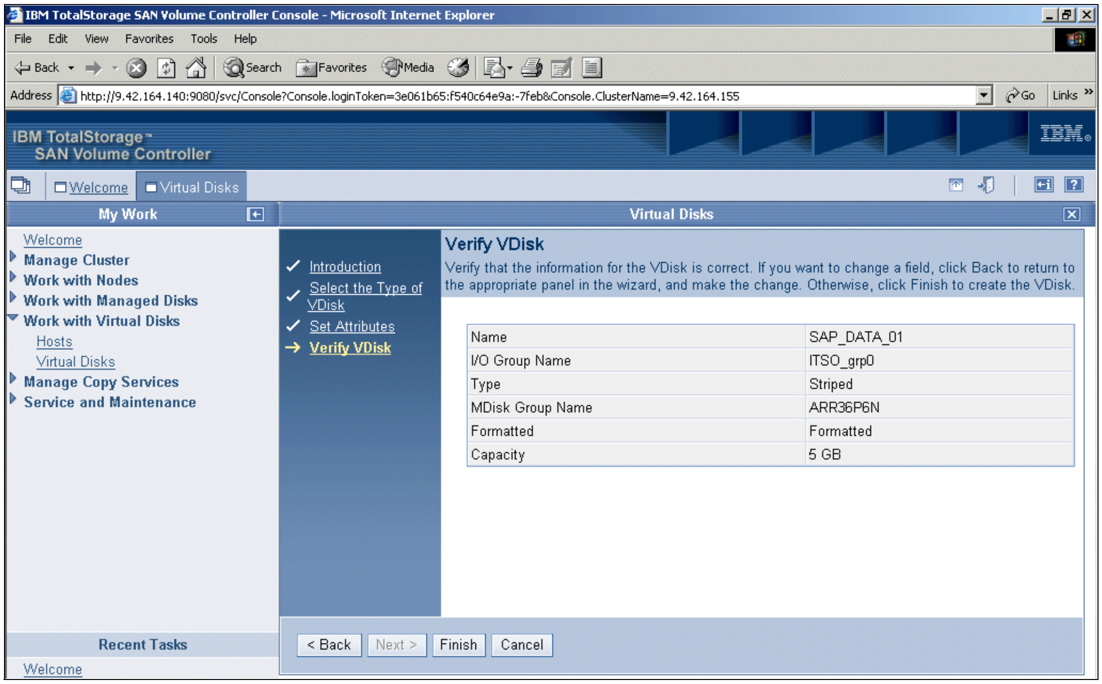


Figure 9-66 Creating a vDisk wizard: Verifying the information

Deleting a vDisk

To delete a vDisk, perform the following steps:

1. Select the radio button to the left of the vDisk you want to delete. Select **Delete a vDisk** from the list and click **Go**.
2. On the Deleting Virtual Disk *vDiskname* panel (where *vDiskname* is the vDisk you selected in the previous step), click **OK** to confirm your desire to delete the vDisk. See Figure 9-67.

   Remember, you can collapse the My Work column by clicking the arrow to the right of the My Work column heading.

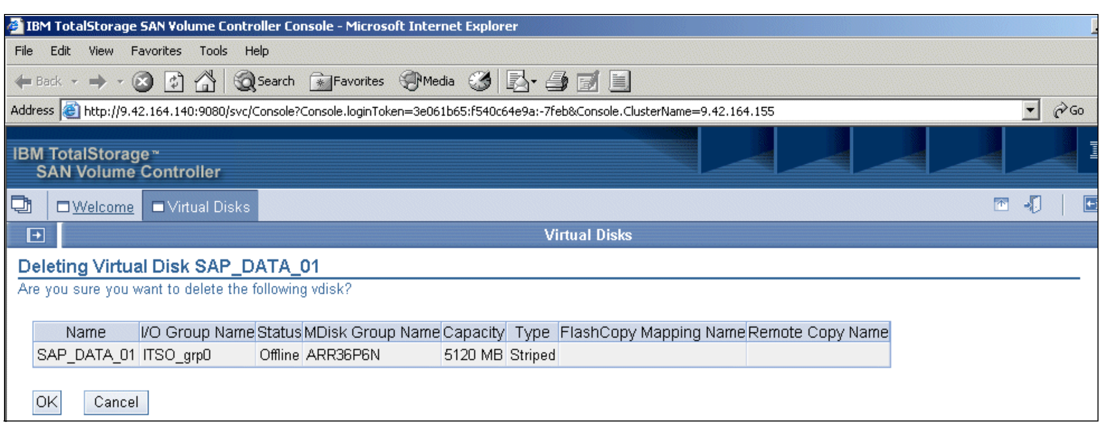


Figure 9-67 Deleting a vDisk

If the vDisk is currently assigned to a host, you receive a secondary message where you must click **Forced Delete** to confirm your decision. See Figure 9-68. This deletes the vDisk-to-host mapping before deleting the vDisk.

Attention: Deleting a vDisk is a destructive action related to user's data residing in that vDisk.

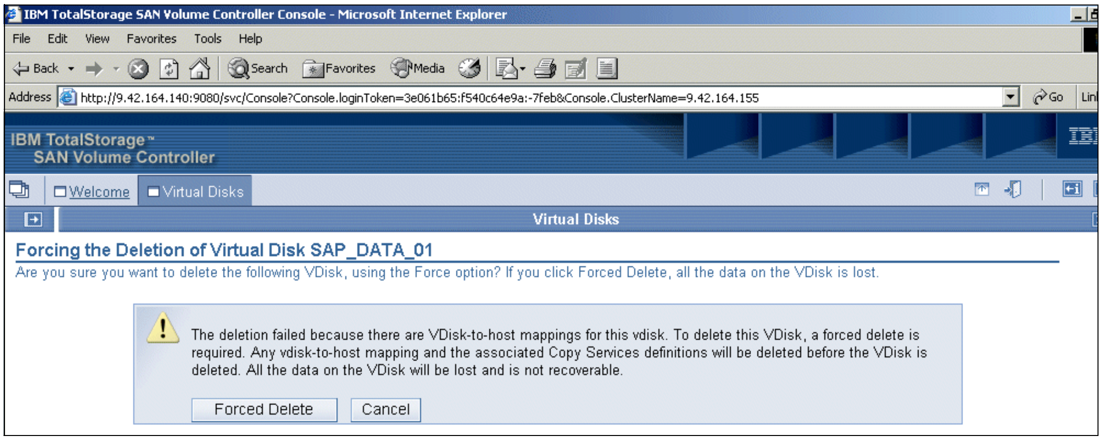


Figure 9-68 Deleting a vDisk: Forcing a deletion

Deleting a vDisk-to-host mapping

To unmap (unassign) a vDisk from a host, perform the following steps:

1. Select the radio button to the left of the vDisk you want to unmap. Select **Delete a vDisk-to-host mapping** from the list and click **Go**.
2. On the Deleting a vDisk-to-host mapping panel (Figure 9-69), from the Host Name list, select the host from which to unassign the vDisk. Click **OK**.

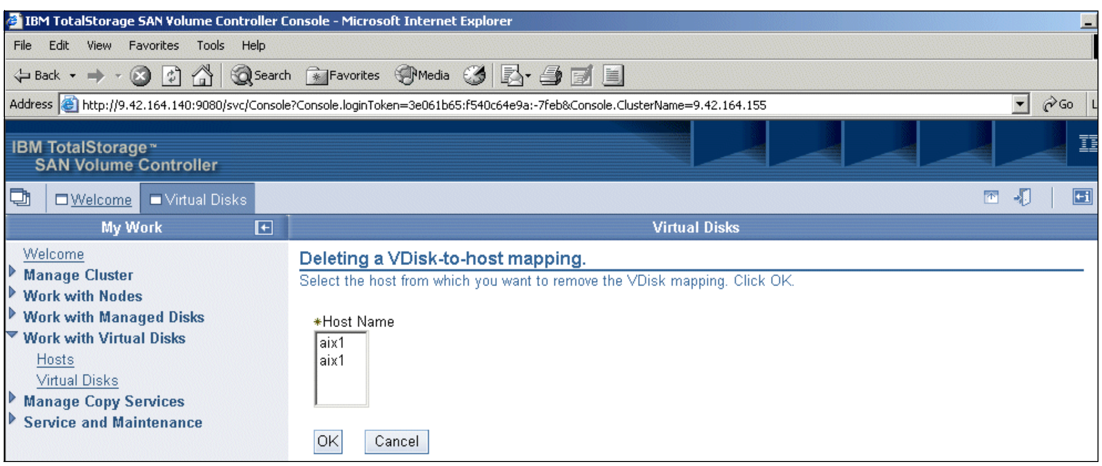


Figure 9-69 Deleting a vDisk-to-host mapping

Note: You see a host instance for each port defined (two in our case). You only need to unassign one of them. Then after you click OK, it may take 10 to 20 seconds to refresh the panel.

Expanding a vDisk

Expanding a vDisk presents a larger capacity disk to your operating system. Although you can do this easily using the SVC, you must ensure that your operating system is prepared for and supports the volume expansion before you use this function.

Assuming your operating system supports it, to expand a vDisk, perform the following steps:

1. Select the radio button to the left of the vDisk you want to expand. Select **Expand a vDisk** from the list and click **Go**.
2. The Expanding Virtual Disks *vDiskname* panel (where *vDiskname* is the vDisk you selected in the previous step) opens. See Figure 9-70. Follow these steps:
   a. Select the new size of the vDisk. This is the new size, not the increment to add. For example, if you have a 5 GB disk and you want it to become 10 GB, you specify 10 GB in this field.

   b. Optionally, select the managed disk candidates from which to obtain the additional capacity. The default for striped vDisk is to use equal capacity from each mDisk in the MDG.

   > **Notes:**
   > - With sequential vDisks, you must specify the mDisk from which you want to obtain space.
   > - There is no support for the expansion of image mode vDisks.
   > - If there are not enough extents to expand your vDisk to the specified size, you receive an error message.

   c. Optionally, format the full vDisk with zeros by selecting the **Format virtual disk (write zeros to its managed disk extents)** check box at the bottom of the panel.

   > **Important:** The **Format Virtual Disk** check box is selected by default. This formats the *entire* vDisk, not just the new extents, so be *very* careful.

   When you are done, click **OK**.

   If there are not enough extents to expand your vDisk to the specified size, you receive an error message.

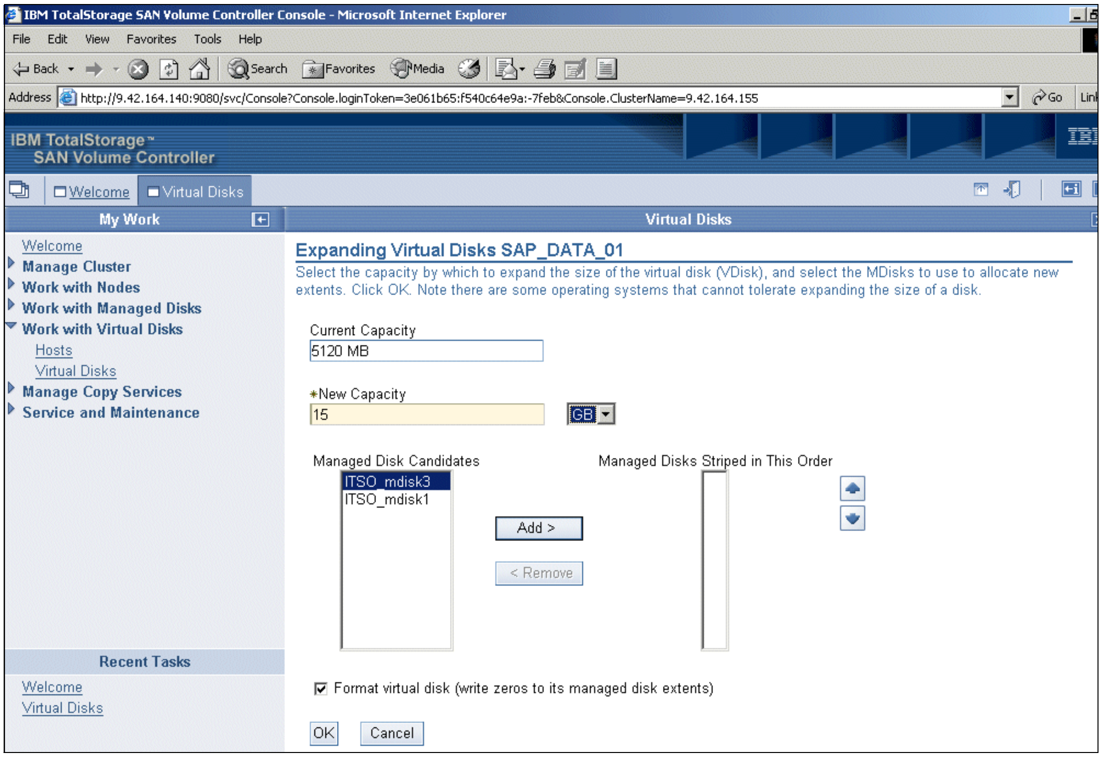


Figure 9-70 Expanding a vDisk

3. Go to your host and perform necessary operations to discover the additional space and expand your volumes into it. This procedure differs depending on the operating system.

Mapping a vDisk to a host

To map (assign) a virtual disk to a host, perform the following steps:

1. Select the radio button to the left of the vDisk you want to assign to a host. Select **Map a vDisk to a host** from the list and click **Go**.
2. On the Creating a vDisk-to-Host mapping *vDiskname* panel (where *vDiskname* is the vDisk you selected in the previous step), from the Target Host list, select the desired host. Optionally specify a SCSI LUN ID. The default is to increment based on what is already assigned to the host. Click **OK**. See Figure 9-71.

> **Note:** After you click OK, it can take 10 to 20 seconds to refresh the page.

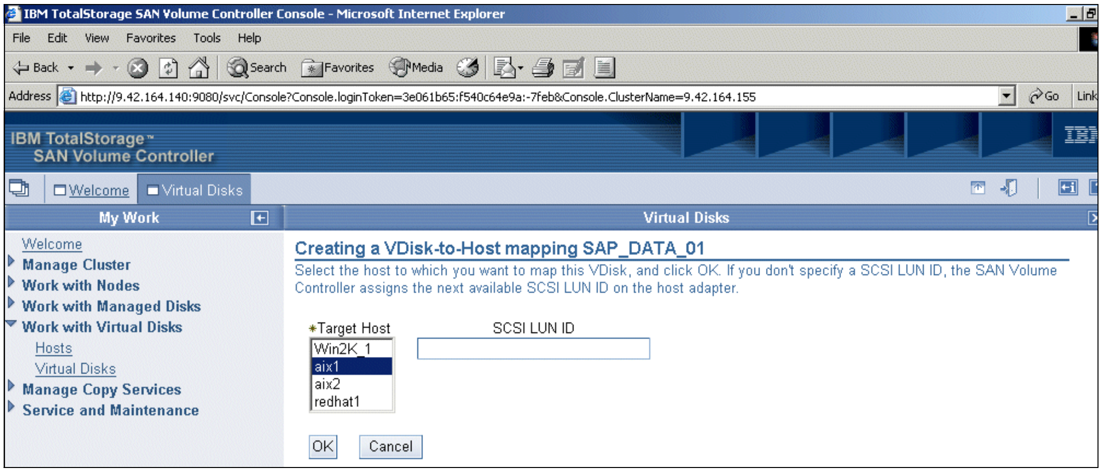


Figure 9-71   Mapping a vDisk to a host

Modifying a vDisk

The Modifying a vDisk menu item allows you to rename the vDisk, reassign the vDisk to another I/O group, and set throttling parameters.

To modify a vDisk, perform the following steps:

1. Select the radio button to the left of the vDisk you want to modify. Select **Modify a vDisk** from the list and click **Go**.
2. The Modifying virtual disk *vDiskname* panel (where *vDiskname* is the vDisk you selected in the previous step) opens. See Figure 9-72. You can perform the following steps separately or in combination:
   a. Type a new name for you vDisk.

   > **Note:** The name can consist of letters A to Z, a to z, numbers 0 to 9, and the underscore. It can be between one and 15 characters in length. However, it cannot start with a number or the word *vdisk* because this prefix is reserved for SVC assignment only.

   b. Select an alternate I/O group from the list to alter the I/O group to which it is assigned.
   c. Set performance throttling for a specific vDisk. In the I/O Governing field, type a number and select either **I/O** or **MB** from the list.

      I/O governing effectively throttles the amount of I/Os per second (or MBs per second) that can be achieved to and from a specific vDisk. You may want to do this if you have a vDisk that has an access pattern that adversely affects the performance of other vDisks on the same set of mDisks. For example, it uses most of the available bandwidth.

      If this application is highly important, then migrating the vDisk to another set of mDisks might be advisable. However, in some cases, it is an issue with the I/O profile of the application rather than a measure of its use or importance.

      The choice between I/O and MB as the I/O governing throttle should be based on the disk access profile of the application. Database applications generally issue large amounts of I/O but only transfer a relatively small amount of data. In this case, setting an I/O governing throttle based on MBs per second does not achieve much. It is better for you to use an I/O per second throttle. On the other extreme, a streaming video application generally issues a small amount of I/O, but transfers large amounts of data.

In contrast to the database example, setting an I/O governing throttle based on I/Os per second does not achieve much. Therefore, you should use an MBs per second throttle.

Click **OK** when you are done making changes.

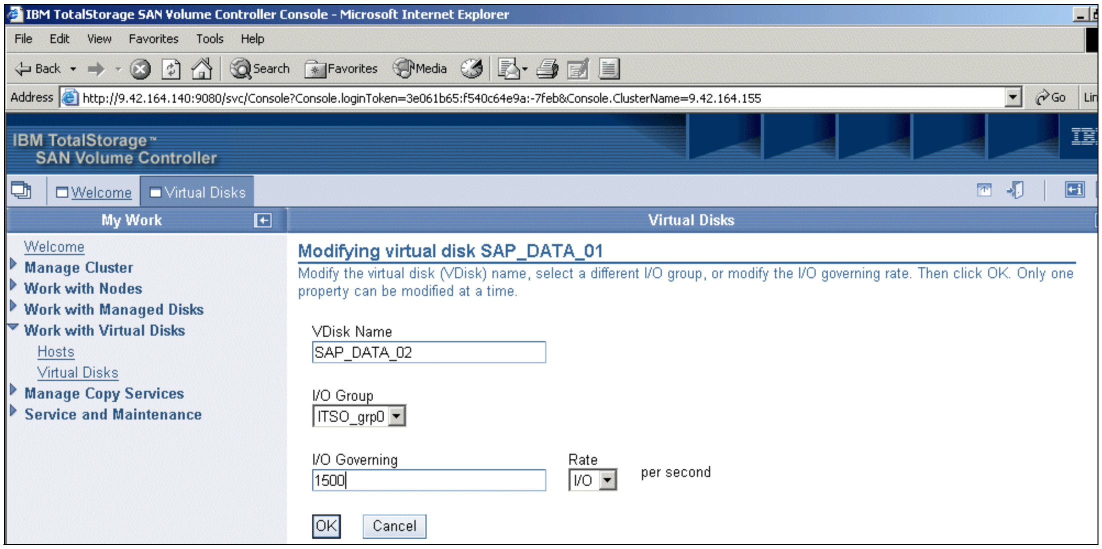


Figure 9-72 Modifying a vDisk

Migrating a vDisk

To migrate a vDisk, perform the following steps:

1. Select the radio button to the left of the vDisk you want to migrate, select **Migrate a vDisk** from the list and click **Go**.
2. The Migrating Virtual Disks-*vDiskname* panel (where *vDiskname* is the vDisk you selected in the previous step) opens as shown in Figure 9-73. From the mDisk Group Name list, select the MDG to which you want to reassign the vDisk. Specify the number of threads to devote to this process (a value from 1 to 4).

   The optional threads parameter allows you to assign a priority to the migration process. A setting of 4 is the highest priority setting. If you want the process to take a lower priority over other types of I/O, you can specify 3, 2, or 1.

   > **Important:** After a migration is started, there is no way to stop it. Migration continues until it is complete unless it is stopped or suspended by an error condition or the vDisk being migrated is deleted.

   When you are done making your selections, click **OK** to begin the migration process.

3. You need to manually refresh your browser or close it and return to the Viewing Virtual Disks panel periodically to see the mDisk Group Name column in the Viewing Virtual Disks panel update to reflect the new MDG name.

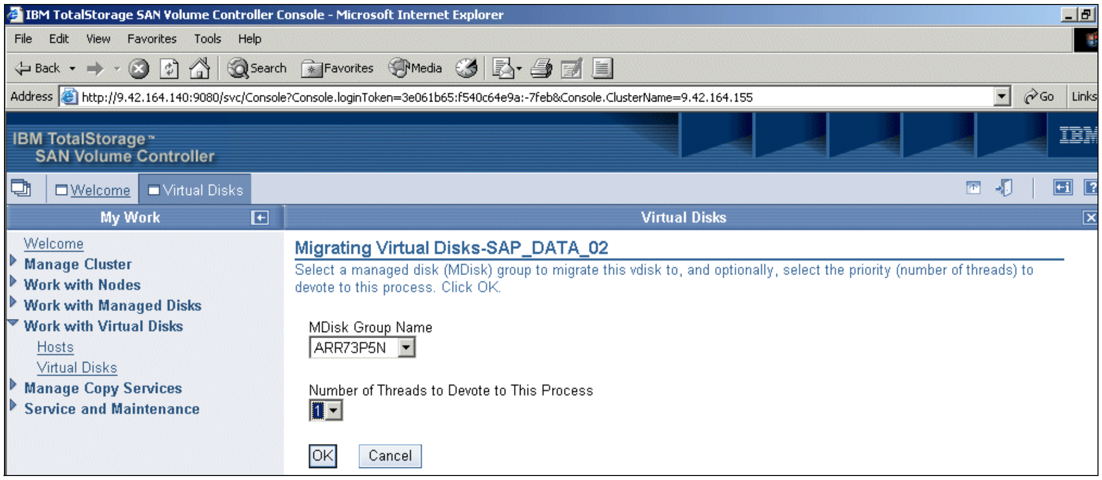


Figure 9-73 Migrating a vDisk

Shrinking a vDisk

The method the that the SVC uses to shrink a vDisk is to remove the required number of extents from the end of the vDisk. Depending on where your data actually resides on the vDisk, this can be quite destructive. For example, you may have a vDisk that consists of 128 extents (0 to 127) of 16 MB (2 GB capacity) and you want to decrease the capacity to 64 extents (1 GB capacity). In this case, the SVC simply removes extents 64 to 127. Depending on the operating system, there is no easy way to ensure that your data resides entirely on extents 0 through 63 so be aware that you *may* lose data.

Although easily done using the SVC, you must ensure your operating system supports shrinking, either natively or by using third-party tools, before using this function. In addition, we recommend that you *always* have a good current backup before you execute this task.

Shrinking a vDisk is useful in certain circumstances such as:

- Reducing the size of a candidate target vDisk of a PPRC relationship to turn it the same size as the source
- Releasing space from vDisks to have free extents in the MDG, considering you do not use that space anymore and take precautions with the remaining data, as explained earlier

Assuming your operating system supports it, perform the following steps to shrink a vDisk:

1. Perform any necessary steps on your host to ensure that you are not using the space you are about to remove.
2. Select the radio button to the left of the vDisk you want to shrink. Select **Shrink a vDisk** from the list and click **Go**.
3. The Shrinking Virtual Disks *vDiskname* panel (where *vDiskname* is the vDisk you selected in the previous step) opens as shown in Figure 9-74. In the Reduce Capacity By field, enter the capacity you want to reduce. Select MB or GB accordingly. The final capacity of the vDisk is the Current Capacity minus the capacity that you specify.

   > **Note:** Be careful with the capacity information. The Current Capacity field shows it in MBs, while you can specify a capacity to reduce in GBs. SVC calculates 1 GB as being 1024 MB.

   When you are done, click **OK**. The changes should become apparent on your host.

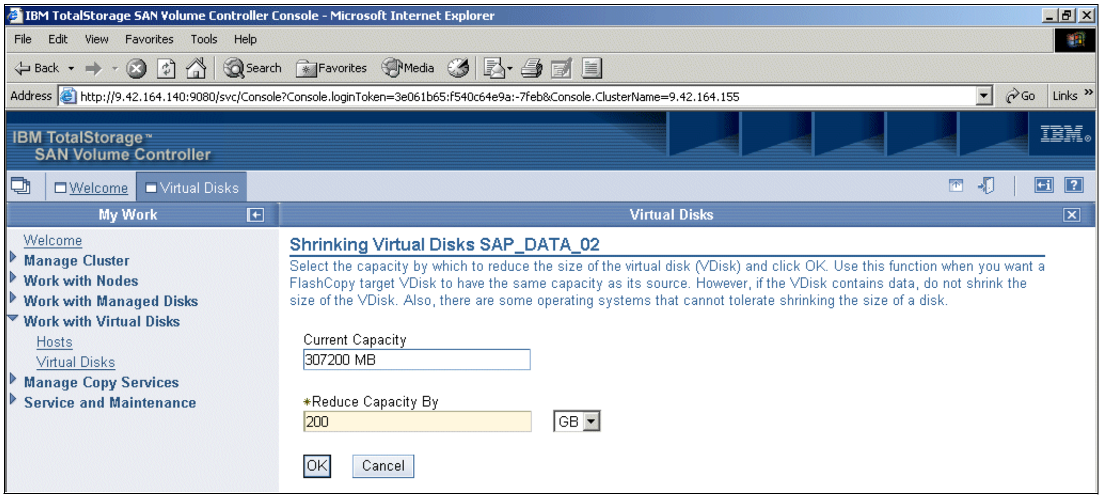


Figure 9-74 Shrinking a vDisk

Showing the mDisks

To show the mDisks that are used by a specific vDisk, perform the following steps:

1. Select the radio button to the left of the vDisk you want to view mDisk information about. Select **Show the mDisks** from the list and click **Go**.
2. You see a subset (specific to the vDisk you chose in the previous step) of the Viewing Managed Disks panel (Figure 9-75).

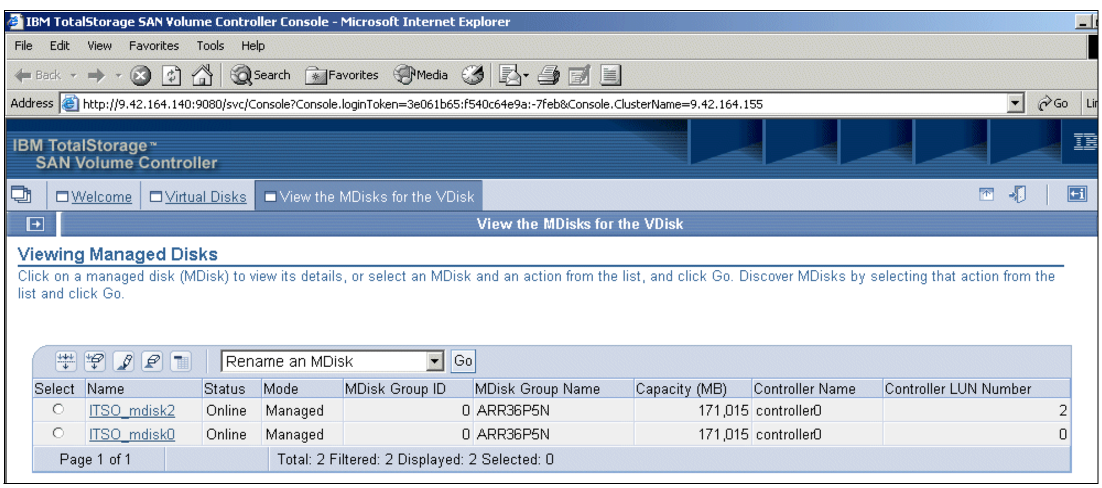


Figure 9-75 Showing mDisks used by a vDisk

For information about what you can do on this panel, see 9.3.2, “Managed disks” on page 239.

Showing the mDisk group

To show the MDG to which a specific vDisk belongs, perform the following steps:

1. Select the radio button to the left of the vDisk you want to view MDG information about. Select **Show the mDisk group** from the list and click **Go**.
2. You see a subset (specific to the vDisk you chose in the previous step) of the Viewing MDGs panel (Figure 9-76).

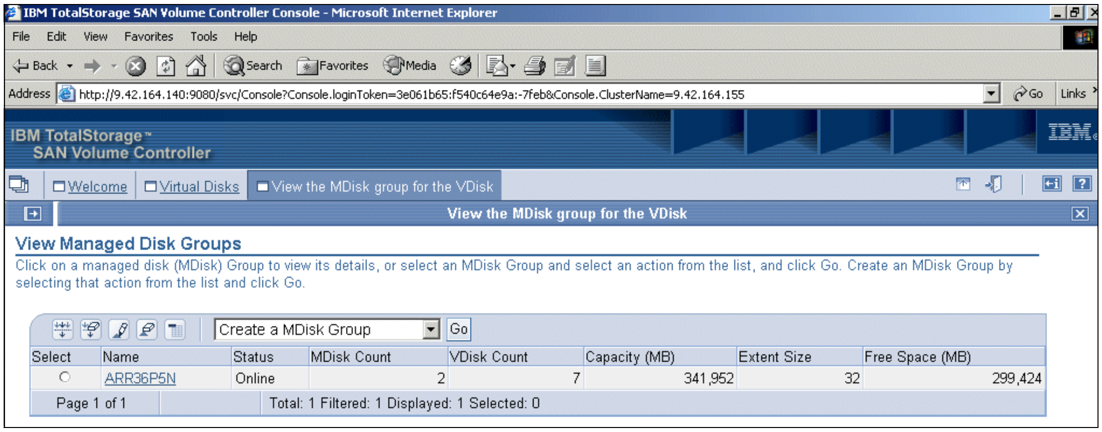


Figure 9-76 Showing a MDG for a vDisk

For information about what you can do on this panel, see 9.3.3, “MDGs” on page 247.

You have now completed the tasks required to manage the hosts and virtual disks within a SVC environment.

9.5 Managing Copy Services

See Chapter 10, “Copy Services: FlashCopy” on page 297, and Chapter 11, “Copy Services: Peer-to-Peer Remote Copy” on page 337, for more information about the tasks related to the management of the Copy Services in the SVC environment.

9.6 Service and maintenance using the GUI

This section discusses the various Service and Maintenance tasks that you can perform within the SVC environment. To perform all of the following activities, on the SVC Welcome page (Figure 9-77), select the **Service and Maintenance** option.

> **Note:** You are prompted for a user ID and password for some of the following tasks.

Figure 9-77 Service and Maintenance functions

9.6.1 Upgrading software

This section explains how to upgrade the SVC software.

Package numbering and version

The format for software upgrade packages is three positive integers separated by dots. For example, a software upgrade package contains something similar to `1.15.0`.

Each software package is given a unique number. Two software packages are fully compatible if the first two numbers are identical. For example, 1.2.15 and 1.2.16 are fully compatible. However, 1.2.15 and 1.3.0 are not compatible.

If two software packages are fully compatible, then you can upgrade or downgrade between these packages. If two software packages are not fully compatible, then the new software upgrade package contains code to upgrade from a list of previous software levels. In this case, you can apply the software upgrade only if the cluster is currently running one of the previous software levels. After you apply it, you cannot remove it.

Precaution before upgrade

Before you upgrade the SVC software, ensure that all I/O paths between all hosts and SANs are working. Otherwise, the applications may have I/O failures during the software upgrade.

You must also ensure that all I/O paths are working for each host that is running I/O operations to the SANs during the software upgrade. You can check the I/O paths by using **`datapath query`** commands. See the *Subsystem Device Driver User's Guide for the IBM TotalStorage Enterprise Storage Server and the IBM TotalStorage SAN Volume Controller*, SC26-7540 for more information about **`datapath query`** commands. You do not need to check for hosts that have no active I/O operations to the SANs during the software upgrade.

Procedure

To upgrade the SVC cluster software, perform the following steps:

1. From the SVC Welcome page, click the **Service and Maintenance** option and then the **Upgrade Software** link.
2. When prompted, enter the `admin` user ID and password, and click **Yes** if prompted with security alerts concerning certificates.
3. On the Upgrade Software panel shown in Figure 9-78, you can either upload a new software upgrade file or list the upgrade files. Click the **Upload** button to upload the latest SVC cluster code.

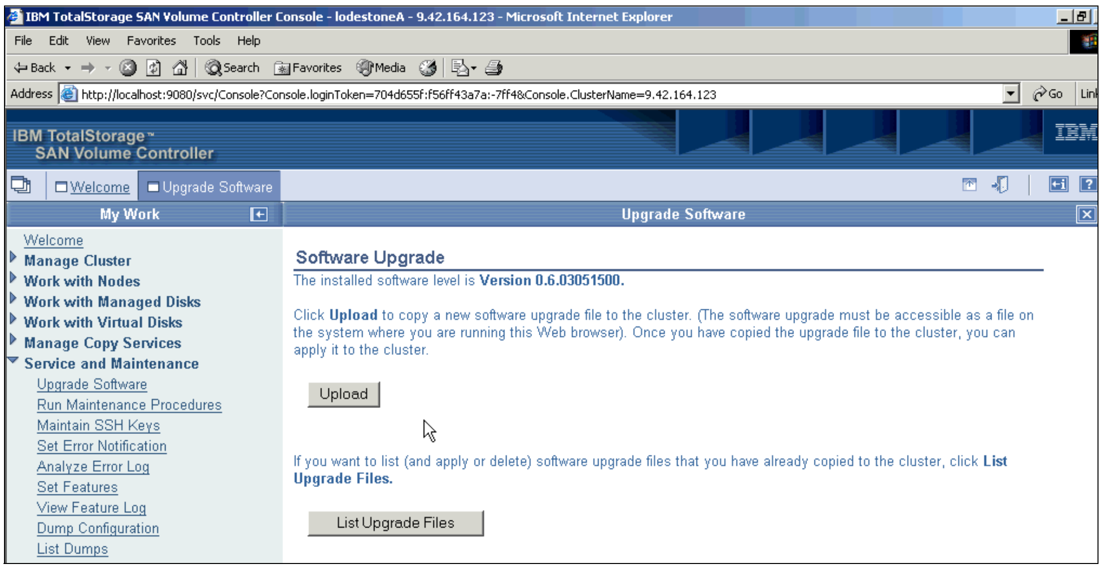


Figure 9-78 Update Software panel

4. On the Software Upgrade (file upload) panel (Figure 9-79), type or browse to the directory on your management workstation (for example, master console) where you stored the latest code level and click **Upload**.

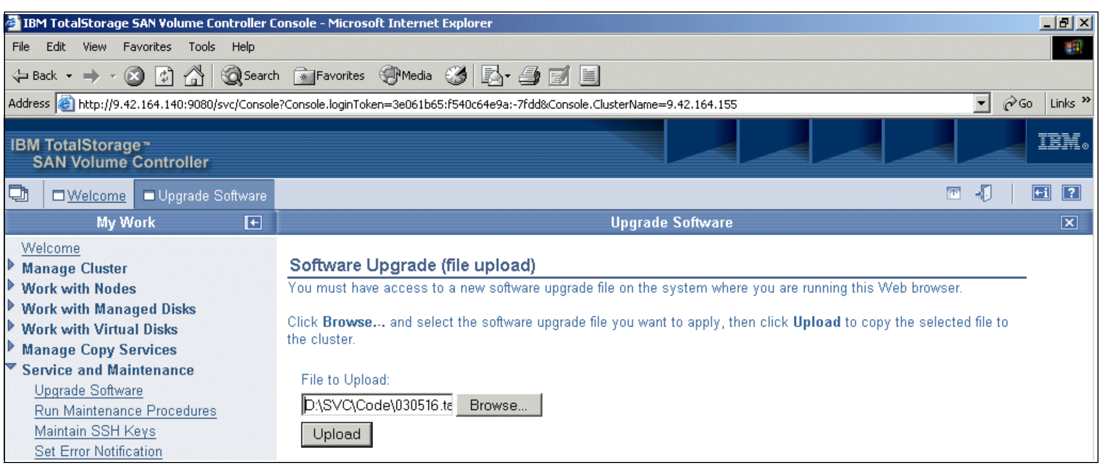


Figure 9-79 Software Upgrade (file upload)

5. The Software Upgrade panel (Figure 9-80) lists the available software packages. Make sure the radio button next to the package you want to apply is selected. Click the **Apply** button.

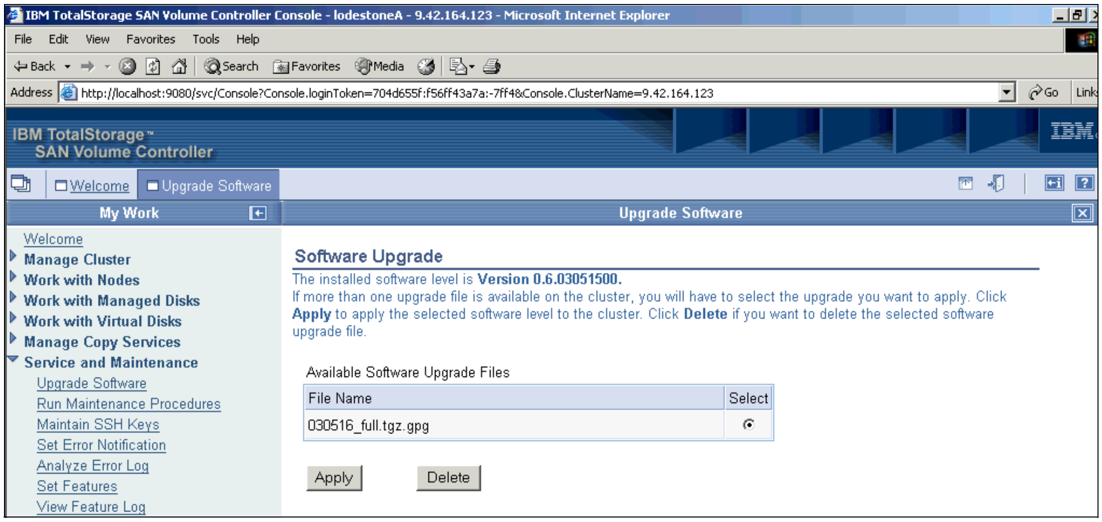


Figure 9-80 Software Upgrade

6. On the next panel, click the **Confirm** button to begin the upgrade process.
7. The Software Upgrade Status panel (Figure 9-81) opens. Click the **Check Upgrade Status** button periodically. However be prepared for a long wait. We waited up to 40 minutes for our upgrade.

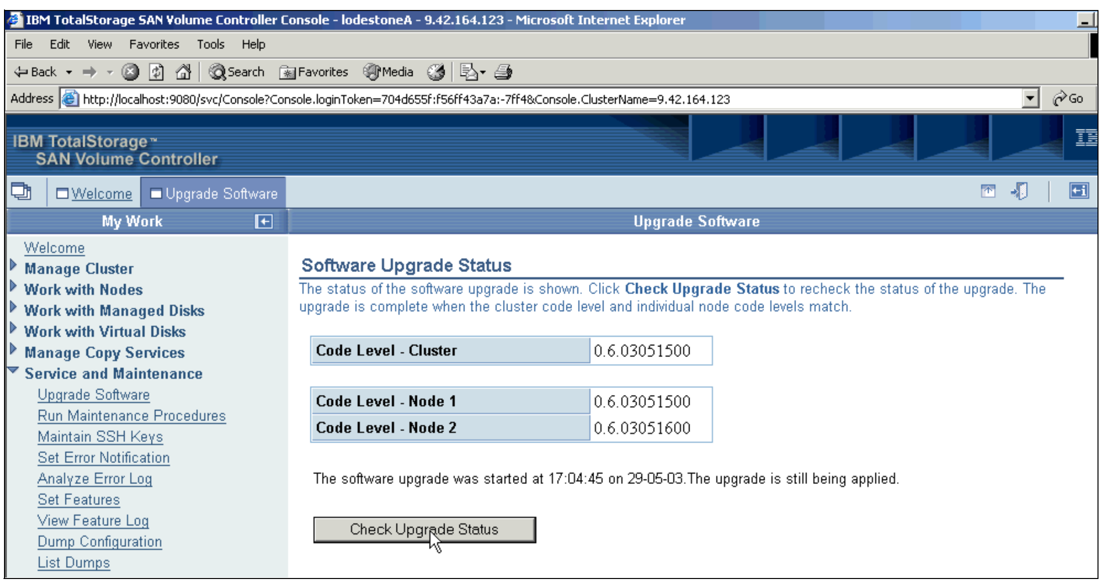


Figure 9-81 Software Upgrade Status

8. The new code is distributed and applied to each node in the SVC cluster. After installation, each node is automatically restarted in turn. If a node does not restart automatically during the upgrade, repair or manually delete that node from the cluster to complete the backout process.
9. Eventually you see the Software Upgrade Status panel (Figure 9-82) display the same level of code on all nodes and the cluster. Verify that the upgrade was successful. Choose either of these actions:
   – Return to the Upgrade Software panel (Figure 9-78), to see the version of the software that is displayed at the top of the panel. It should reflect the new level.

– Go to the Analyze Error Log panel. Search for `Software Install completed`. Select the **Sort by date with the newest first** radio button and then click **Perform**. This should list the software near the top. For more information about how to work with the Analyze Error Log panel, see 9.6.5, “Analyzing the error log” on page 284.

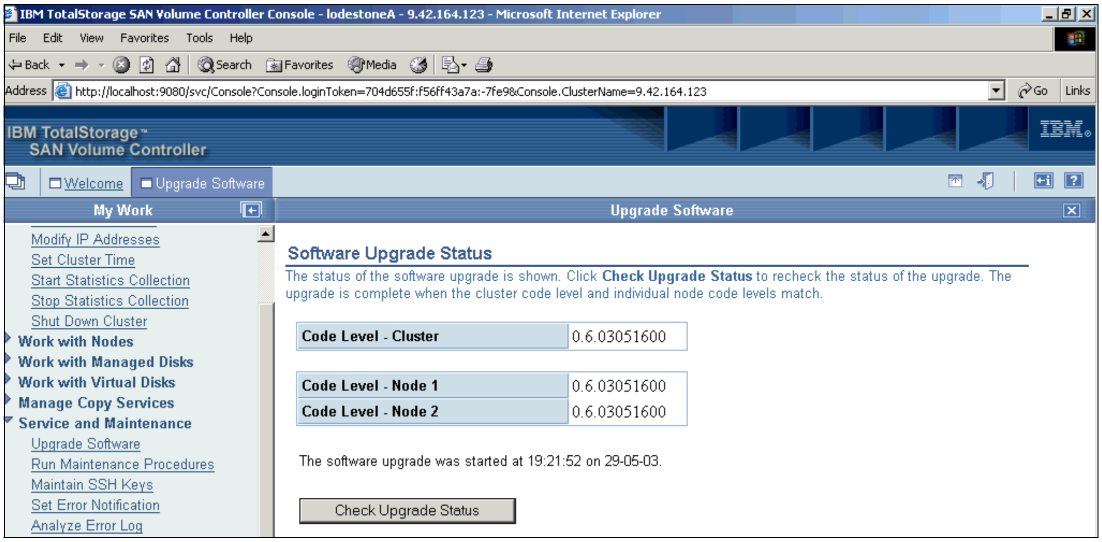


Figure 9-82 Software Upgrade Success

10. You have now completed the tasks required to upgrade the SVC software. Click the **X** icon in the upper right corner of the display area to close the Upgrade Software panel. Do *not* close the browser by mistake.

9.6.2 Running maintenance procedures

To run the maintenance procedures on the SVC cluster, perform the following steps:

1. From the SVC Welcome page, click the **Service and Maintenance** option and then the **Run Maintenance Procedures** link.
2. As shown in Figure 9-83, you can select whether you want to use the existing cluster error log or generate a fresh one using the radio buttons. Using the existing log file displays the unfixed errors that exist in the log file that was last generated. If this is your first time using this option, no error log exists.

   If you want to obtain the latest status of your cluster, or it is the first time you are using this option, select the **Generate a new error log file** option to create an error log file named errlog_000667_030521_160750 in the /dumps/elogs/ directory.

   – **errlog** part of the file name is generic for all error log files.
   – **000667** is the panel name of the current configuration node.
   – **030521** is the date (YYMMDD).
   – **160750** is the time.

   Click **Start Analysis**.

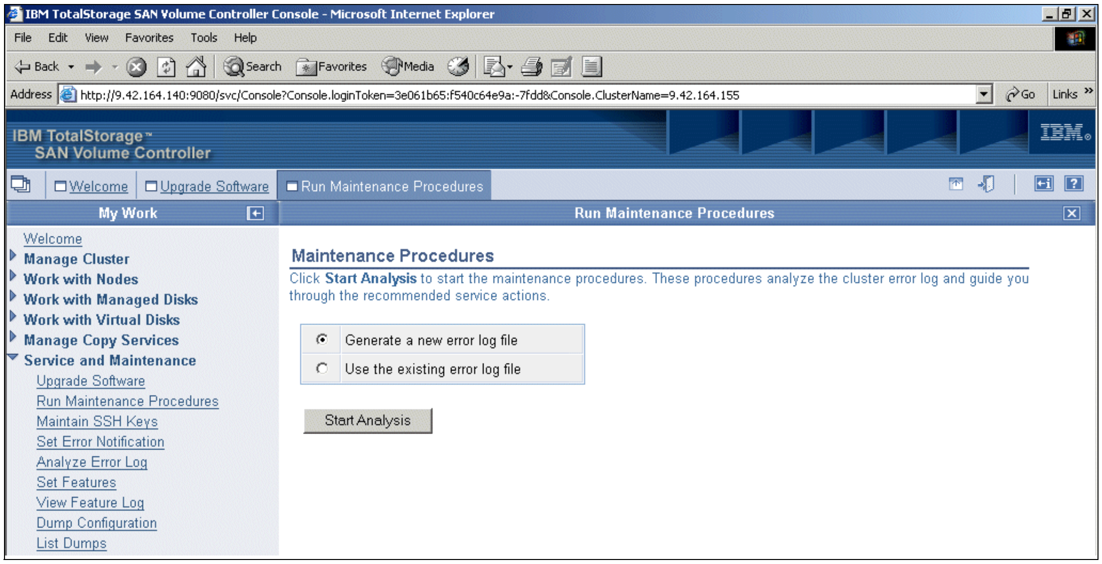


Figure 9-83 Maintenance Procedures

3. The log file (existing or newly generated) is scanned for unfixed errors. This task may return a list of unfixed errors (for example, things that require administrative intervention). Or, as in our case, it shows a message indicating that there are no unfixed errors in the cluster. See Figure 9-84.

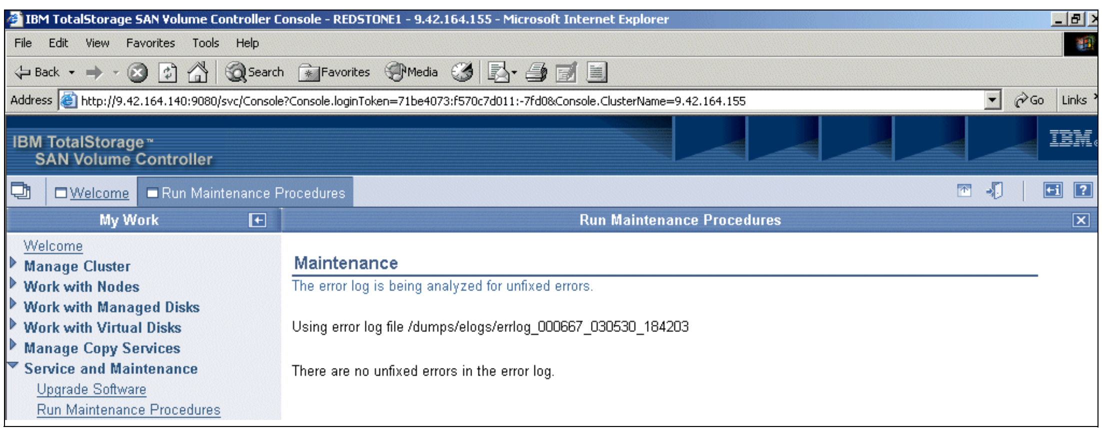


Figure 9-84 Maintenance procedures: Starting the analysis

4. Click the **X** icon in the upper right corner of the display area to close the Run Maintenance Procedures panel. Do not close the browser by mistake.

9.6.3 Maintaining SSH keys

To maintain the SSH keys, perform the following steps:

1. From the SVC Welcome page, click the **Service and Maintenance** options and then the **Maintain SSH Keys** link.
2. As shown in the SSH Key Maintenance panel (Figure 9-85), you can either list keys or add a key. To list existing keys, click the **List Keys** button. To add additional keys, complete the appropriate fields and click **Add Key**.

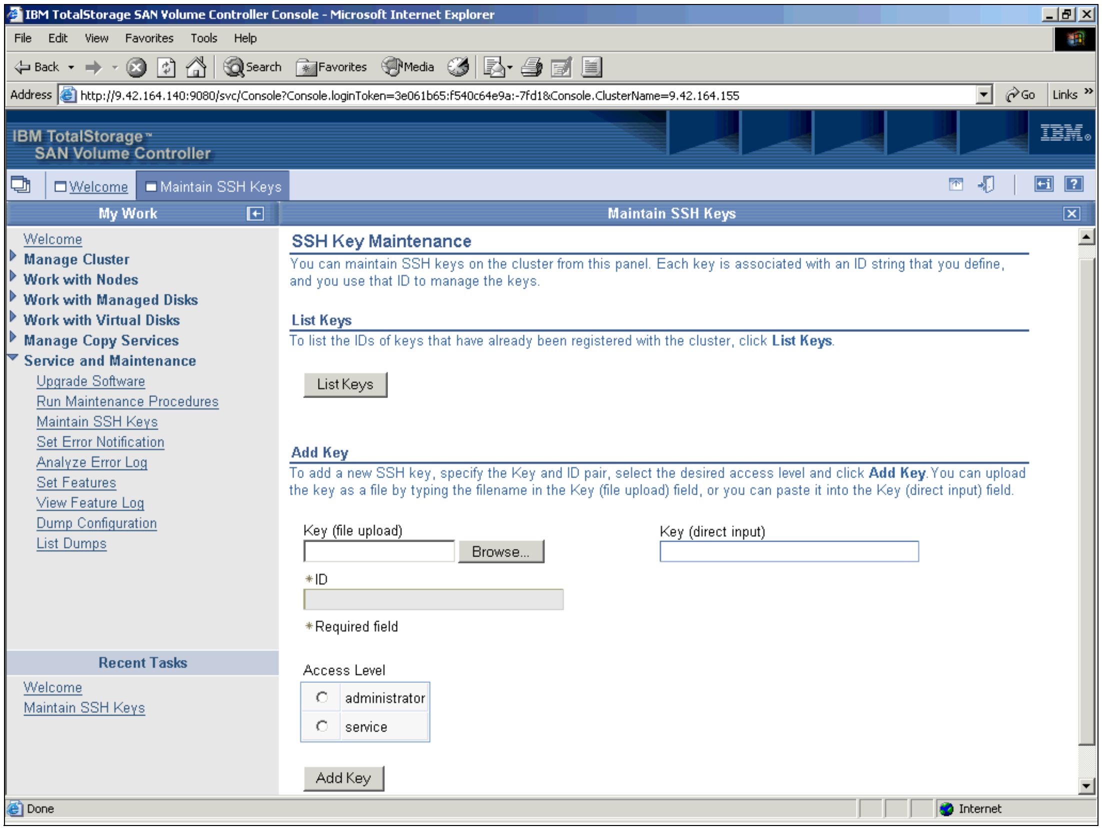


Figure 9-85 Maintain SSH Keys

– When you click List Keys, you see a panel similar to the one in Figure 9-86. It shows which keys already exist on the SVC cluster. You can delete IDs that are no longer required or authorized. You can either select the IDs you want to delete (use CTRL key to select multiple entries) and click the **Delete Selected IDs** button, or delete all IDs by clicking the **Delete All IDs** button.

> **Important:** Deleting all IDs is normally considered only if you believe that security has been compromised and that unauthorized personnel have access to admin or service IDs. If you delete all IDs, you *must* generate a fresh pair and upload into your cluster in order to administer the SVC environment again. If you did not enable the *Password reset via service panel* option, the only way for you to gain access to your cluster (to upload the new key) is to rebuild the cluster using the service panel. This *destroys all* data.
>
> For this reason, if you find yourself in a situation where security has been compromised, we recommend that you generate a fresh pair of keys and upload the public half before you selectively delete all other IDs.
>
> For information about how to generate SSH Keys, see 4.2.1, “Generating public and private SSH key pair using PuTTY” on page 70.

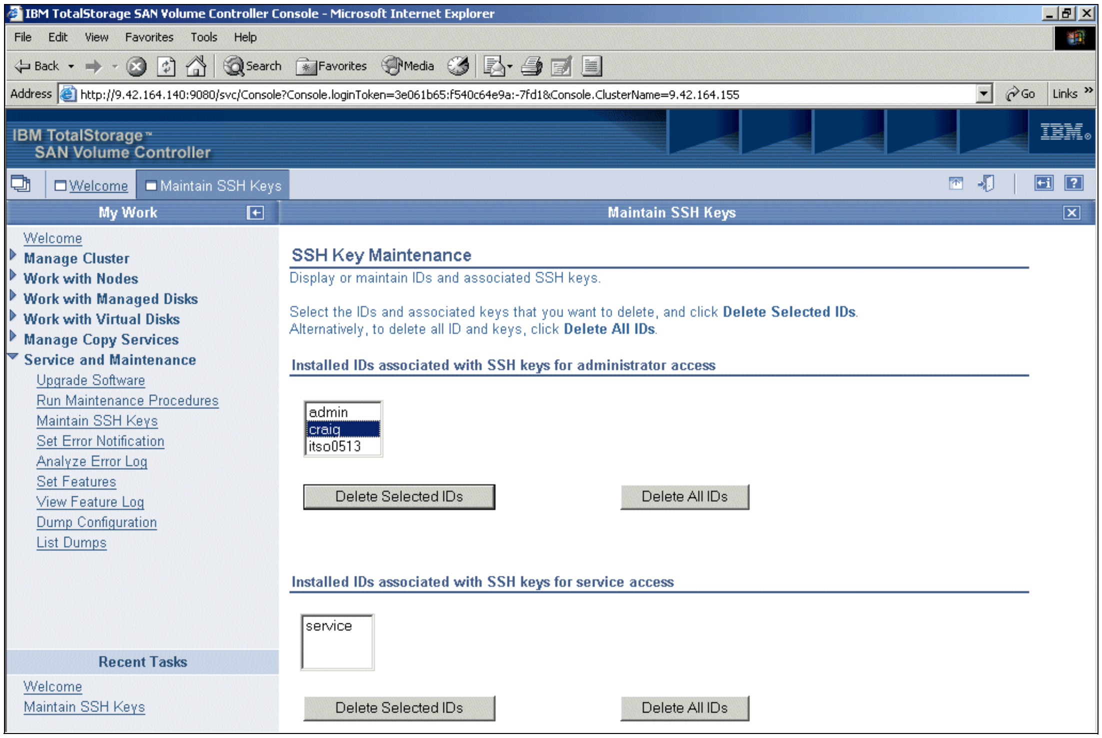


Figure 9-86 List Keys

– If you want to add a key, complete the appropriate information as shown in Figure 9-87, and click **Add Key**.

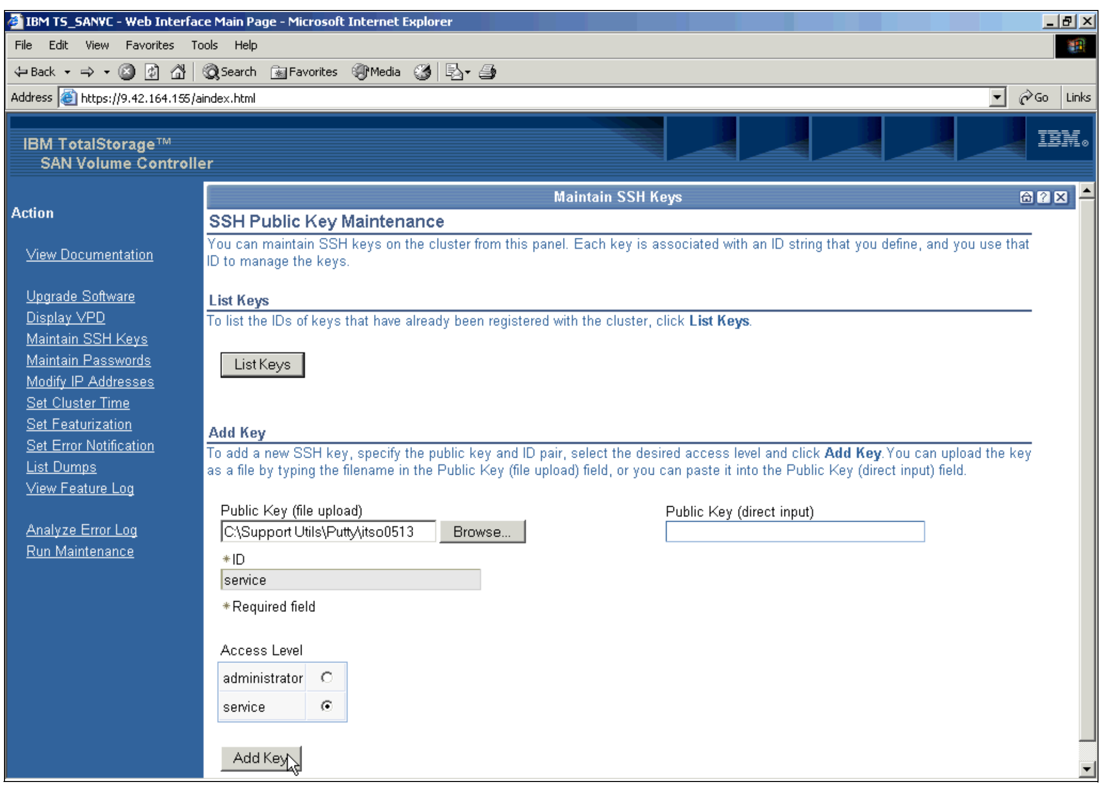


Figure 9-87 SSH Public Key Maintenance

3. A new panel opens and displays the message `Added new SSH service key associated with ID userid` as shown in Figure 9-88 (where *userid* is the user ID you specified).

> **Note:** If we selected the **admin** radio button, the panel that opens would read `Added new SSH admin key associated with ID IDname`.

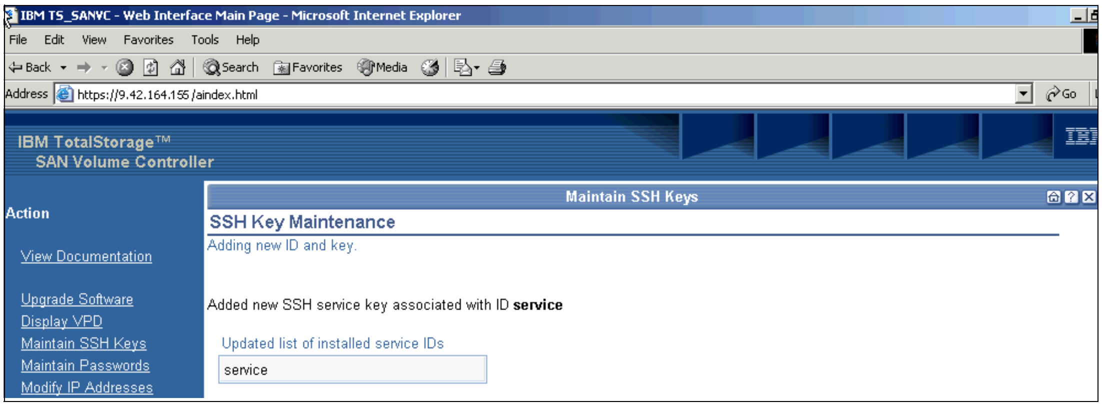


Figure 9-88 Adding SSH service key successful

4. Click the **X** icon in the upper right corner of the display area to close the SSH Keys Maintenance panel. Do not close the browser by mistake.

9.6.4 Setting error notification

To set up error notification, perform the following steps:

1. From the SVC Welcome page, click the **Service and Maintenance** option and then the **Set Error Notifications** link.
2. On the Modify Error Notification Settings panel (Figure 9-89), select the level of notification (default is None) to apply to both SNMP and e-mail alerting. Type the IP address of your SNMP Manager and the e-mail address to which to forward alerts. Click **Modify Settings**.

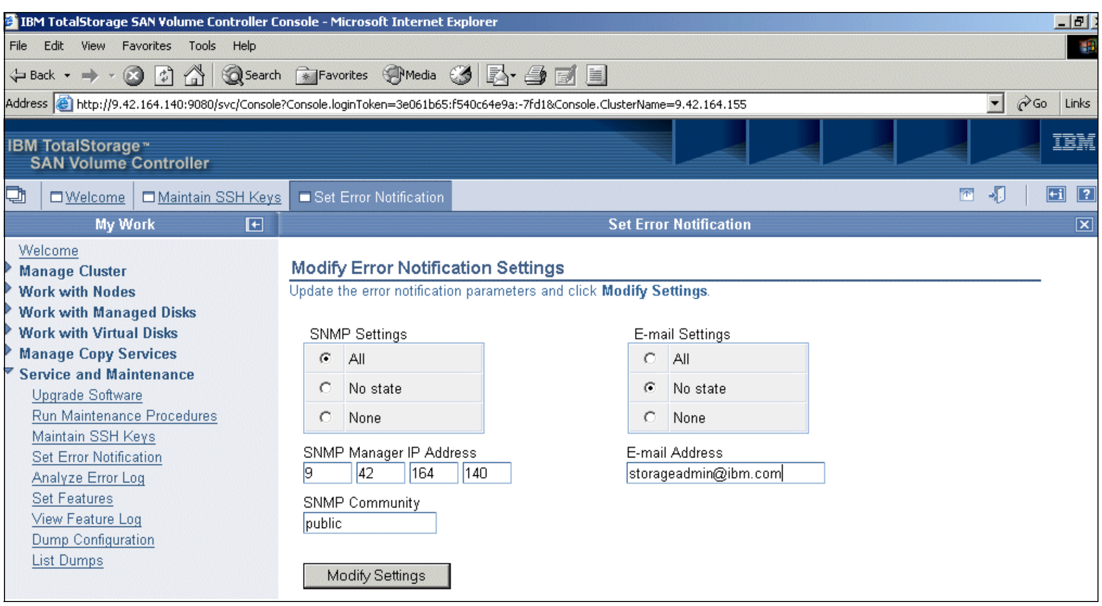


Figure 9-89 Setting error notification

3. The Modifying Error Notification Settings panel now shows these message:

```
The SNMP notification settings have been updated.
The e-mail notification settings have been updated.
```

4. Click the **X** icon in the upper right corner of the display area to close the Set Error Notification panel. Do not close the browser by mistake.

9.6.5 Analyzing the error log

The following types of events and errors are logged in the error log:

- **Events**: State changes that are detected by the cluster software and that are logged for informational purposes. Events are recorded in the cluster error log.
- **Errors**: Hardware or software problems that are detected by the cluster software and that require some sort of repair. Errors are recorded in the cluster error log.
- **Unfixed errors**: Errors that were detected and recorded in the cluster error log and that were not yet corrected or repaired.
- **Fixed errors**: Errors that were detected and recorded in the cluster error log and that were subsequently corrected or repaired.

To display the error log for analysis, perform the following steps:

1. From the SVC Welcome page, click the **Service and Maintenance** options and then the **Analyze Error Log** link.
2. From the Error Log Analysis panel (Figure 9-90), you can choose either the Process or Clear Log button.

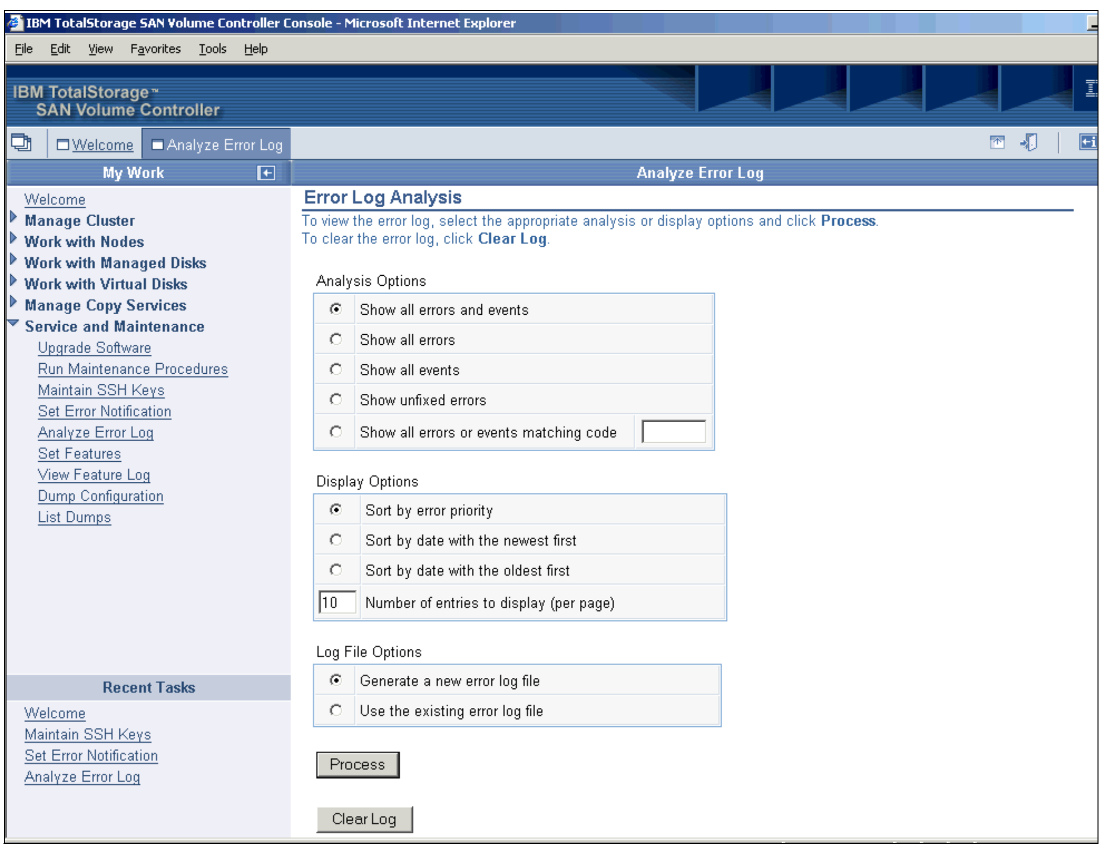


Figure 9-90 Analyzing the error log

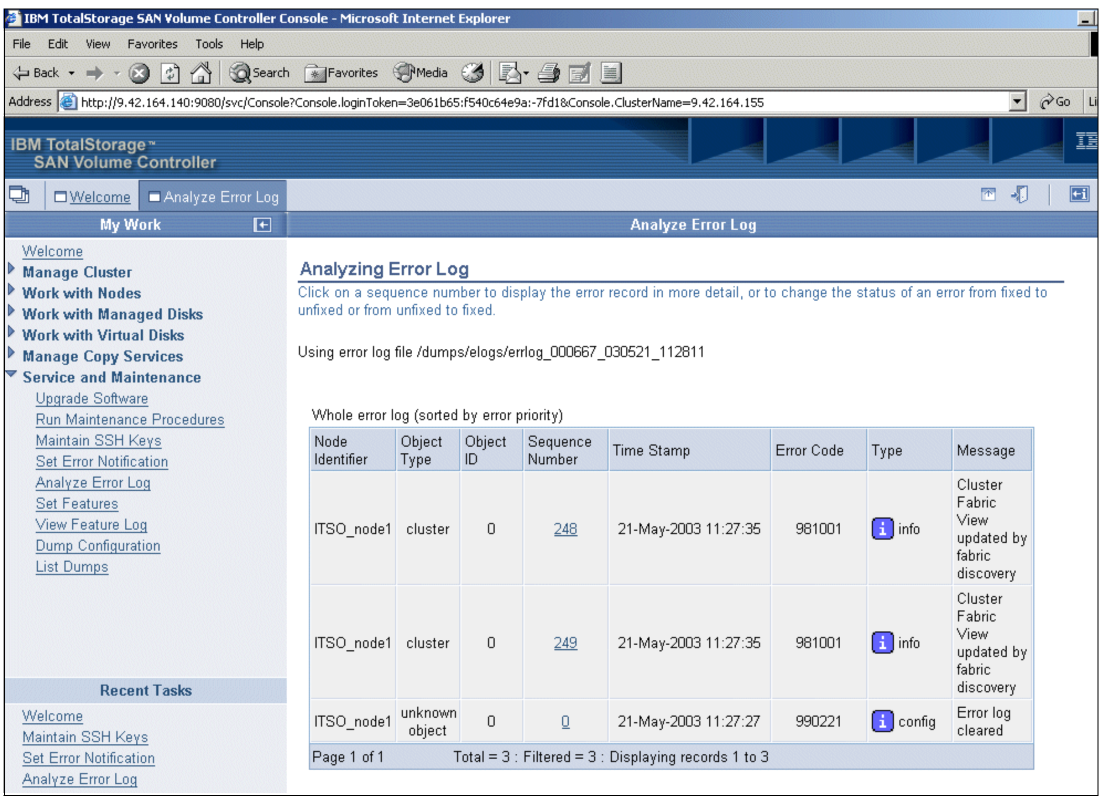
a. Select the appropriate radio buttons and click the **Process** button to display the log for analysis. The Analysis Options and Display Options radio button boxes allow you to filter the results of your log enquiry to reduce the output.

You can display the whole log, or you can filter the log so that only errors, events, or unfixed errors are displayed. You can also sort the results by selecting the appropriate display options. For example, you can sort the errors by error priority (lowest number = most serious error) or by date. If you sort by date, you can specify whether the newest or oldest error is to display at the top of the table. You can also specify the number of entries you want to display on each page of the table.

Click the Log File Options radio button to use the existing log file or to generate a fresh one. Using the existing log file displays entries that exist in the log file that was last generated. If this is the first time you are using this option, no error log exists. To obtain the latest status of your cluster, or if it is the first time you are using this option, select the **Generate a new error log file** option. The errlog_000667_030521_112811 error log file is created in the /dumps/elogs/ directory and is ready for analysis.

- **errlog** part of the file name is generic for all error log files.
- **000667** is the panel name of the current configuration node.
- **030521** is the date (YYMMDD).
- **112811** is the time.

In our example (Figure 9-91), we recently cleared our log (see the third message listed) and have only a few information events logged.



Node Identifier	Object Type	Object ID	Sequence Number	Time Stamp	Error Code	Type	Message
ITSO_node1	cluster	0	248	21-May-2003 11:27:35	981001	info	Cluster Fabric View updated by fabric discovery
ITSO_node1	cluster	0	249	21-May-2003 11:27:35	981001	info	Cluster Fabric View updated by fabric discovery
ITSO_node1	unknown object	0	0	21-May-2003 11:27:27	990221	config	Error log cleared

Page 1 of 1 Total = 3 : Filtered = 3 : Displaying records 1 to 3

Figure 9-91 Analyzing Error Log: Process

b. Click the **Clear Log** button at the bottom of the panel to clear the log. If the error log contains unfixed errors, a warning message is displayed when you click Clear Log.

3. Click the **X** icon in the upper right corner of the display area to close the Analyze Error Log panel. Do not close the browser by mistake.

9.6.6 Setting features

To change licensing feature settings, perform the following steps:

1. From the SVC Welcome page, click the **Service and Maintenance** options and then the **Set Features** link.
2. On the Featurization Settings panel (Figure 9-92), consult your license before you make changes in this panel. If you purchased additional features (for example, FlashCopy or PPRC) or increase the capacity of your license, make the appropriate changes. Then click the **Update Feature Settings** button.

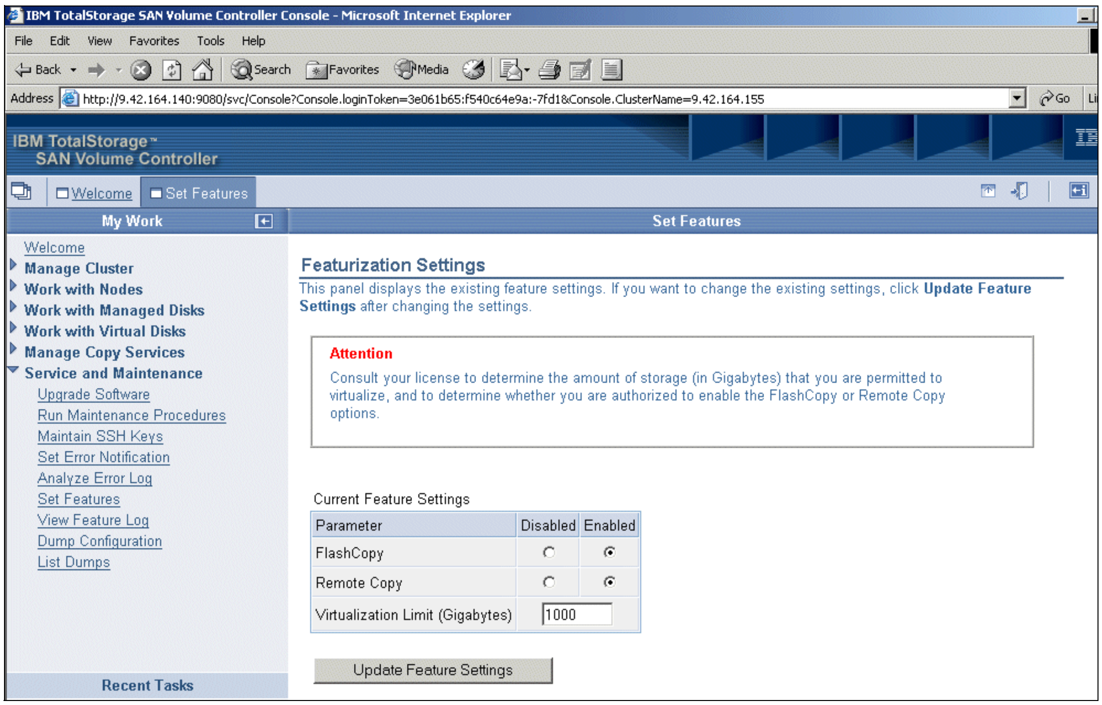


Figure 9-92 Setting features

3. You now see a license confirmation panel as shown in Figure 9-93. Review this panel and ensure that you are in compliance. If you are in compliance, click **I Agree** to make the requested changes take effect.

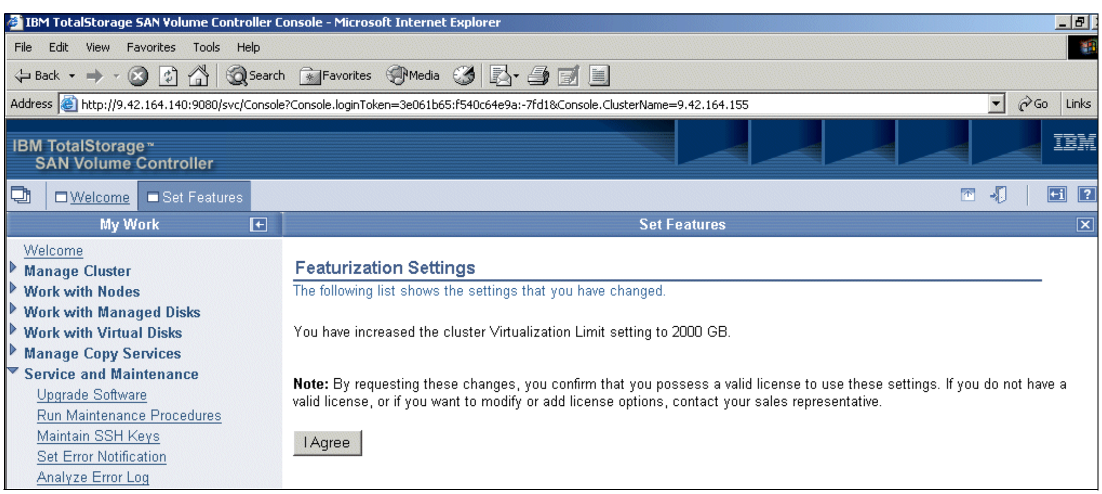


Figure 9-93 License agreement

4. You return to the Set Features panel where your changes should be reflected.
5. Click the **X** icon in the upper right corner of the display area to close the Set Features panel. Do not close the browser by mistake.

9.6.7 Viewing the feature log

To view the feature log, which registers the events related to the SVC licensed features, perform the following steps:

1. From the SVC Welcome page, click the **Service and Maintenance** option and then the **View Feature Log** link.
2. The Feature Log panel (Figure 9-94) opens. It displays the current feature settings and a log of when changes were made.

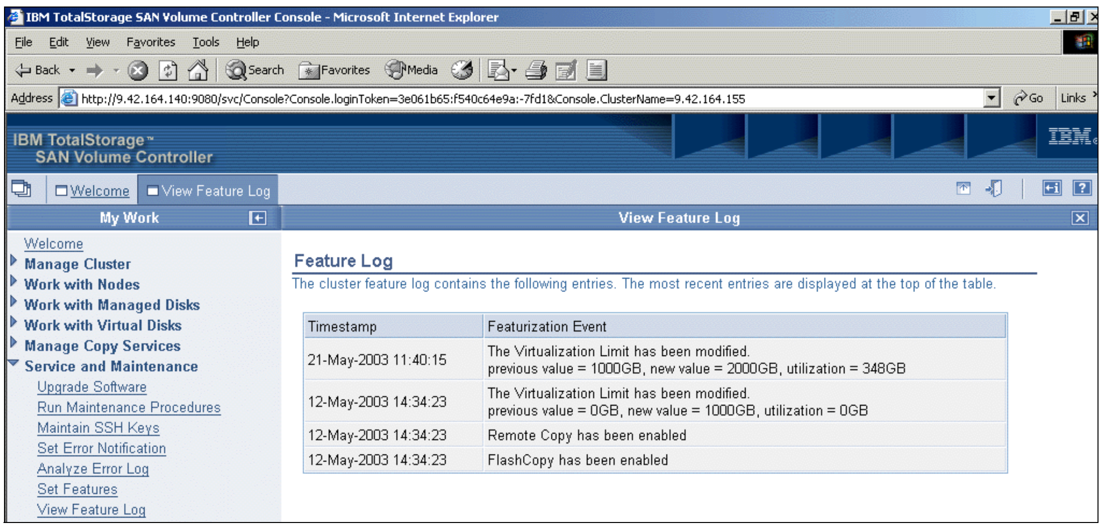


Figure 9-94 Feature Log

3. Click the **X** icon in the upper right corner of the display area to close the View Feature Log panel. Do not close the browser by mistake.

9.6.8 Dump configuration

To dump configuration information, perform the following steps:

1. From the SVC Welcome page, click the **Service and Maintenance** options and then the **Dump Configuration** link.

> **Note:** The purpose of performing a dump of the configuration data is only for problem diagnosis. It is not intended for configuration backup. To back up the SVC configuration data, see the 9.7, “Backing up the SVC configuration” on page 290.

2. On the Dumping Cluster Configuration panel (Figure 9-95), type a unique prefix for the dump file name. Click **OK** to create the dump file.

> **Note:** If you don’t specify a file name prefix, SVC uses the word *config*. The node front panel name and a time stamp are automatically appended to the file name prefix.

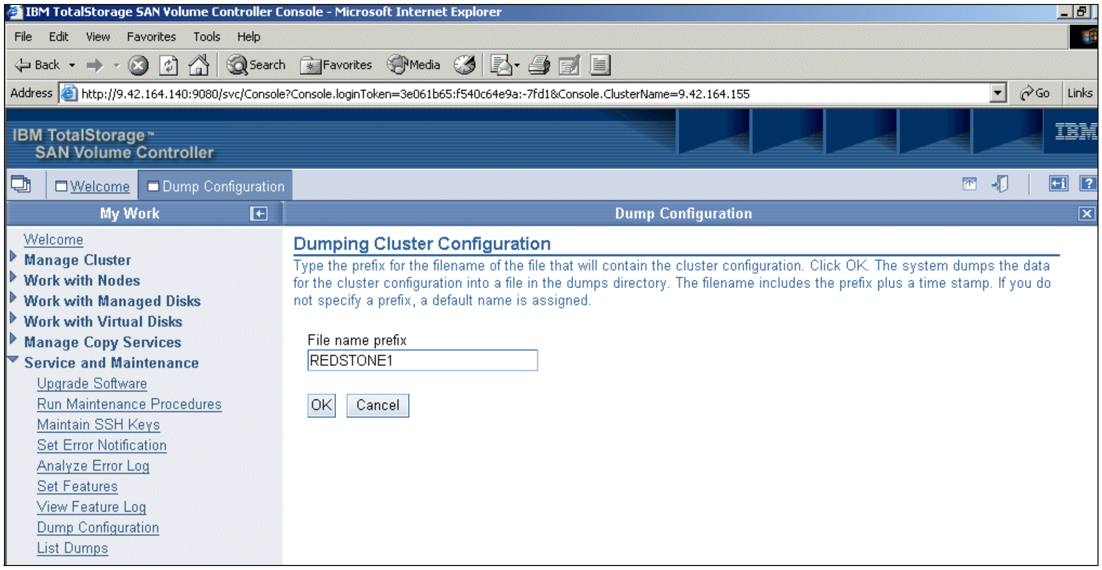


Figure 9-95 Dump configuration

3. Click the **X** icon in the upper right corner of the display area to close the Dump Configuration panel. Do not close the browser by mistake.

9.6.9 Listing dumps

To list the dumps that were generated, perform the following steps:

1. From the SVC Welcome page, click the **Service and Maintenance** option and then the **List Dumps** link.
2. On the List Dumps panel (Figure 9-96), you see several dumps and log files that were generated over time on our cluster. They include the configuration dump we generated in the previous section. Click any of the available links (the underlined text in the table under the List Dumps heading) to go to another panel that displays the available dumps.

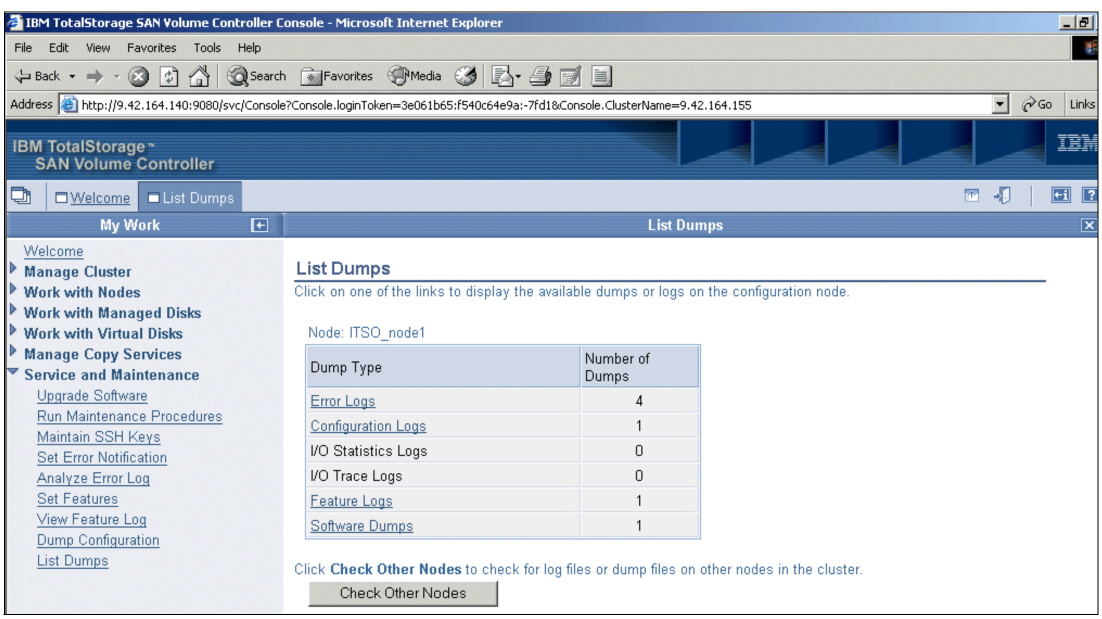


Figure 9-96 List Dumps

For example, if you click the **Error Logs** link, you see information like the example shown in Figure 9-97.

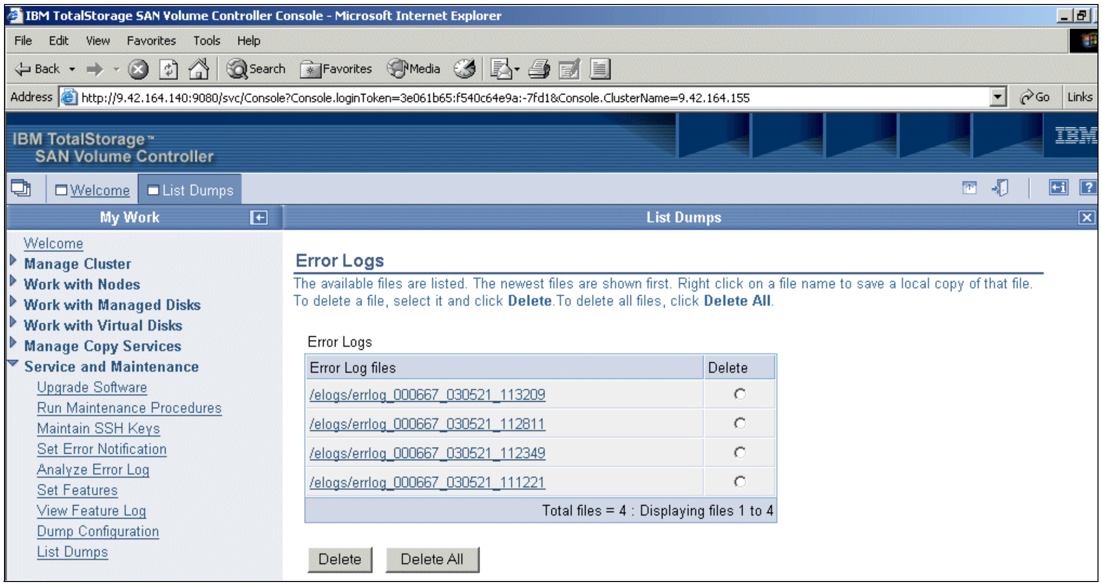


Figure 9-97 List Dumps: Error Logs

3. From this panel, you can perform either of the following tasks:
   - Click any of the available log file links (indicated by the underlined text) to display the log in complete detail as shown in Figure 9-98.
   - Delete one or all of the dump or log files. To delete all, click the **Delete All** button. To delete one, select the radio button to the right of the file and click the **Delete** button.

> **Note:** By default, the dump and log information that is displayed is available from the configuration node (in our example ITSO_node1). In addition to these files, each node in the SVC cluster keeps a local software dump file. Occasionally, other dumps are stored on them. Click the **Check Other Nodes** button at the bottom of the List Dumps panel to see which dumps or logs exist on other nodes in your cluster.

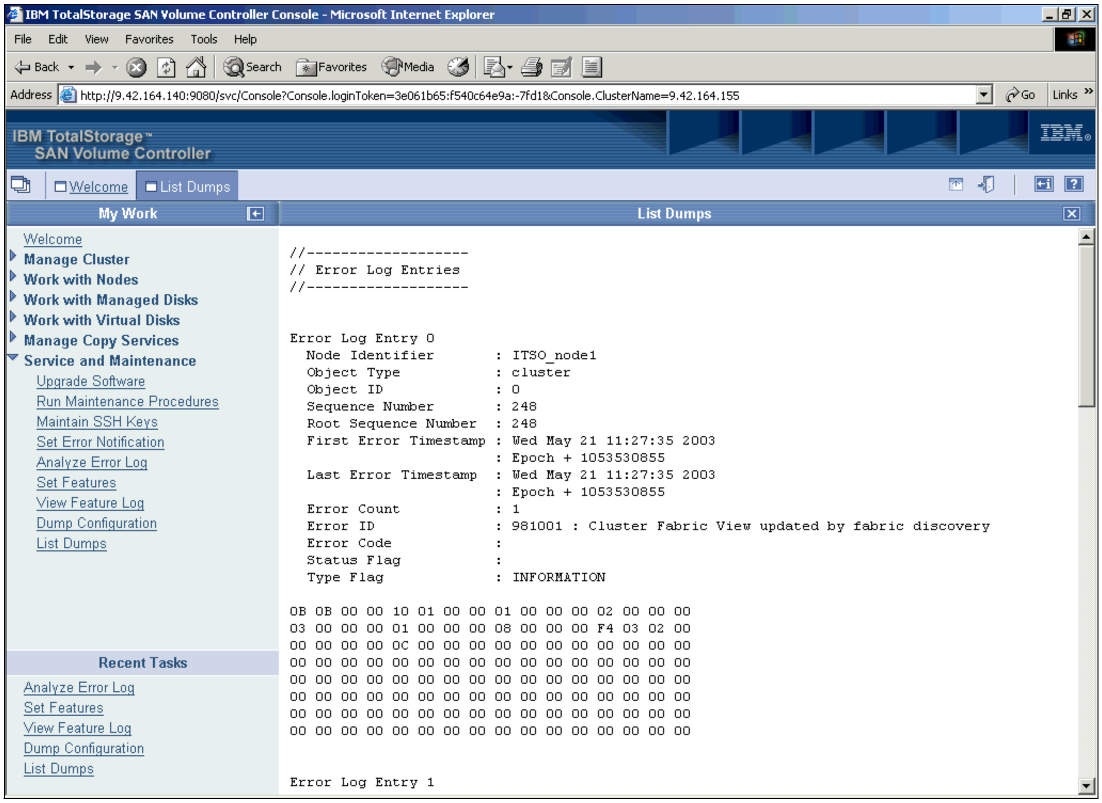


Figure 9-98 List Dumps: Error log detail

4. Click the **X** icon in the upper right corner of the display area to close the List Dumps panel. Do not close the browser by mistake.

9.7 Backing up the SVC configuration

The SVC configuration data is stored on all the nodes in the cluster. It is specially hardened so that, in normal circumstances, the SVC should never lose its configuration settings. However in exceptional circumstances, such as a rogue fire sprinkler soaking the SVC cluster or a multiple hardware failure, this data may become corrupted or lost.

This section details the tasks that you can perform to save the configuration data from a SVC configuration node and restore it. The following configuration information is backed up:

- SVC cluster
- Storage controllers
- Hosts
- I/O groups
- Software licenses
- Managed disks
- MDGs
- SVC nodes
- SSH keys
- Virtual disks
- vDisk-to-host mappings

Before you begin the restore process, consult IBM Support to determine the cause as to why you cannot access your original configuration data. After the restore process starts, the original data on the vDisks is destroyed. Therefore, you must ensure that you have a backup of all user data on the vDisks.

The prerequisites for having a successful backup are:

- All nodes in the cluster must be online.
- No object name may begin with an underscore (_).
- Do not run any independent operations that could change the cluster configuration while the backup command runs.
- Do not make any changes to the fabric or cluster between backup and restore. If changes are made, back up your configuration again or you may not be able to restore it later.

Note: We recommend that you make a backup of the SVC configuration data after each major change in the environment, such as defining or changing a vDisks, vDisk-to-host mappings, etc.

The output of the SVC configuration backup is a file with the name svc.config.backup.xml that is stored in the C:\Program Files\IBM\svcconsole\cimom\backup*SVCclustername* folder in the SVC master console (where *SVCclustername* is the SVC cluster name of the configuration from which you backed up). This differs from backing up the configuration using CLI. The svc.config.backup.xml file is stored in the /tmp folder on the configuration node and must be copied to an external and secure place for backup purposes.

Important: We strongly recommend that you change the default names of all objects to non-default names. For objects with a default name, a warning is produced and the object is restored with its original name and “_r” appended to it.

To back up the SVC configuration data, perform the following steps:

1. From the SVC Welcome page, click the **Service and Maintenance** option and then the **Backup Configuration** link.
2. On the Backing up a Cluster Configuration panel (Figure 9-99), click the **Backup** button.

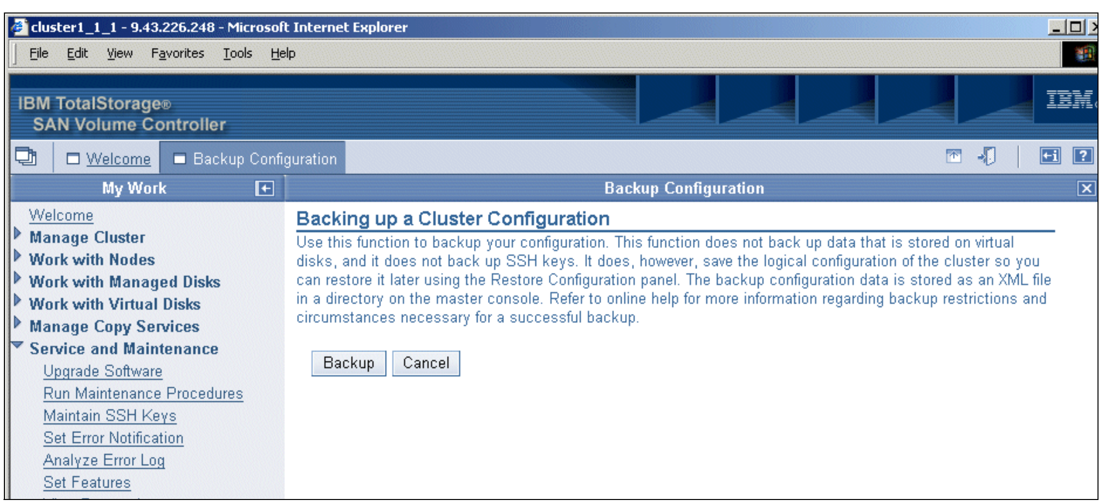


Figure 9-99 Backing up a Cluster Configuration data

3. After the configuration backup is successfully done, you see the message as shown in Figure 9-100. Make sure you that read, understand, and document the warning messages, since they can influence the restore procedure.

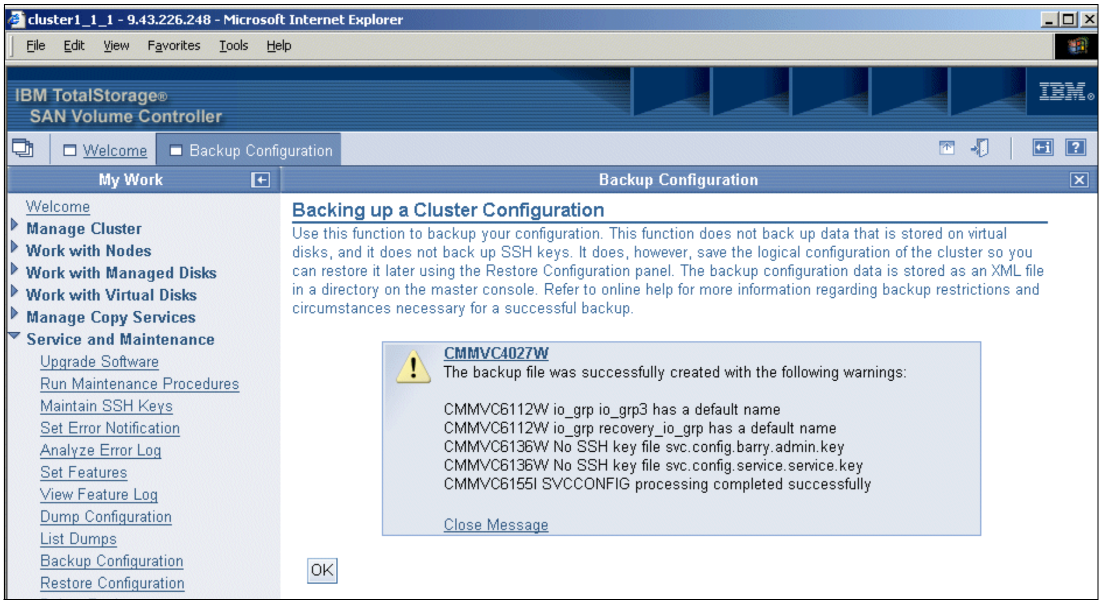


Figure 9-100 Configuration backup successful message and warnings

4. Click the **X** icon in the upper right corner of the display area to close the Backing up a Cluster Configuration panel. Do not close the browser by mistake.

9.7.1 Restoring the SVC configuration

Important: You can carry out the restore procedure only under the direction of IBM third-level support. If the process is carried out incorrectly, *data corruption* can occur.

This section discusses the tasks that you can perform to restore the configuration data to a SVC configuration node. This offers you a way to quickly recover from a disaster situation.

Restoring the SVC configuration data *does not* restore a user's data (vDisk contents) nor the metadata to find the physical extents in the mDisks related to the original vDisks. Its purpose is to have minimum downtime when recovering in a disaster situation, restoring the configuration information backed up as shown in 9.7, "Backing up the SVC configuration" on page 290, without manually defining everything again (mDisks, vDisks, Hosts, vDisk-to-host mappings, etc.). After that, you can restore client data using the normal restore routines and resume operations.

To successfully restore a configuration, these rules apply:

- Do not make any changes to the fabric or the cluster between backup and restore. If you make changes, you should perform another backup. Make sure that:
  – The cluster name is the same.
  – Enough I/O groups exist.
  – Each required mDisk in the backup cluster configuration is present.
  – All required disk controller systems are available.
- There should be only one node in the cluster to which you are restoring the configuration, that is, the default configuration node.

- There should be no MDGs, vDisks, hosts, FlashCopy consistency groups, FlashCopy mappings, Remote Copy consistency groups, or Remote Copy relationships defined.
- The hosts, SVC cluster, disk controller, disks, and the SAN fabric should be operational. To make the SVC cluster operational, you must perform the initial setup as explained in 4.3, “Basic installation” on page 74, and 4.4, “Completing the initial cluster setup using the Web interface” on page 76. During the initial setup, the SSH keys you backed up are uploaded to SVC.
- Do not make changes to the fabric or to the cluster between the two phases of the restore process (prepare and execute).
- The backup configuration should be in the C:\Program Files\IBM\svcconsole\cimom \backup*SVCclustername* folder on the SVC master console (where *SVCclustername* is the name to which you are restoring the configuration).

To restore the SVC Configuration data, perform the following steps:

1. From the SVC Welcome page, click the **Service and Maintenance** options and then the **Restore Configuration** link.
2. On the Restoring a Cluster Configuration panel (Figure 9-101), click the **OK** button. Do not select the Format virtual disks (write zeros to managed disk extents during restoration) check box.

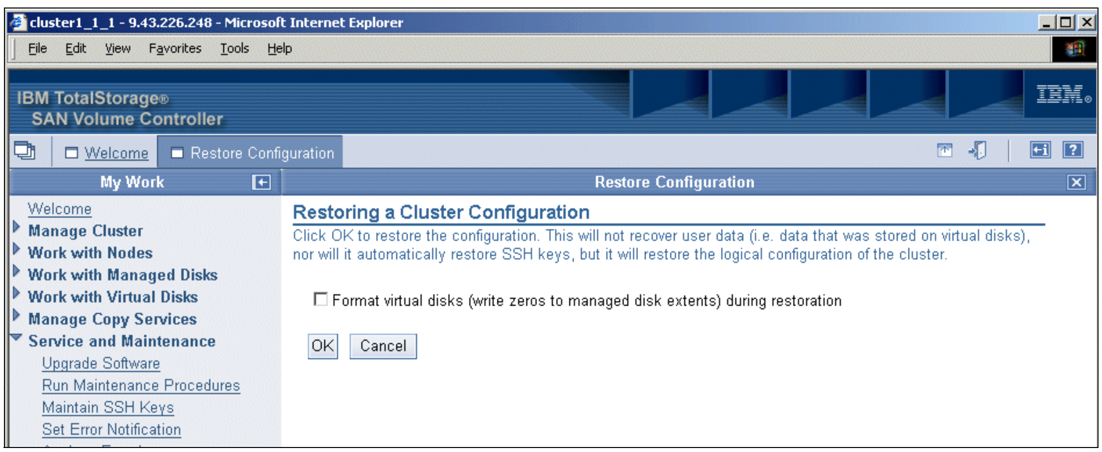


Figure 9-101 Restoring a Cluster Configuration panel

3. Click the **Restore Configuration** button to confirm the restore operation. See Figure 9-102.

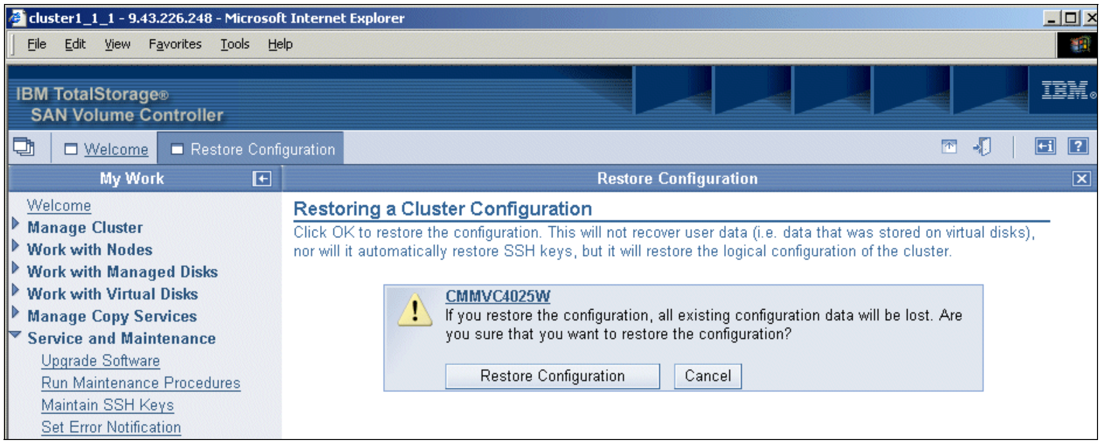


Figure 9-102 Confirming the configuration restore operation

4. If you did not meet the prerequisites for a configuration restore, as explained earlier in this section, the Prepare phase fails. You receive CMMVCxxxxE error messages as shown in Figure 9-103. Correct the errors and try to restore the configuration again.

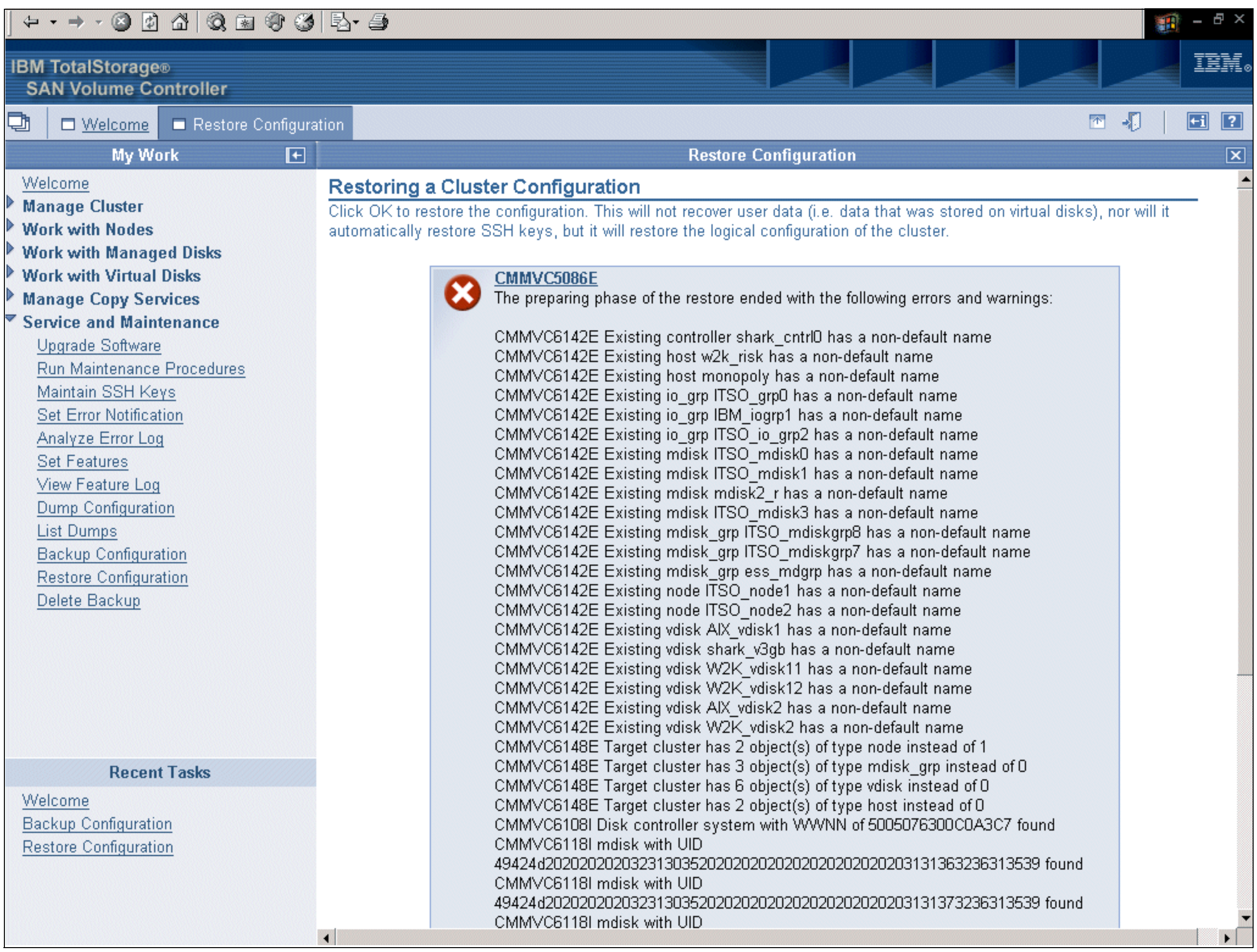


Figure 9-103 Restoring a configuration error messages

5. If you meet all the prerequisites, the Prepare phase should pass. Then you receive only warning and informational messages as shown in Figure 9-104. If you are sure these messages are expected, click the **Restore Configuration** button to start the Execute phase of the configuration restore.

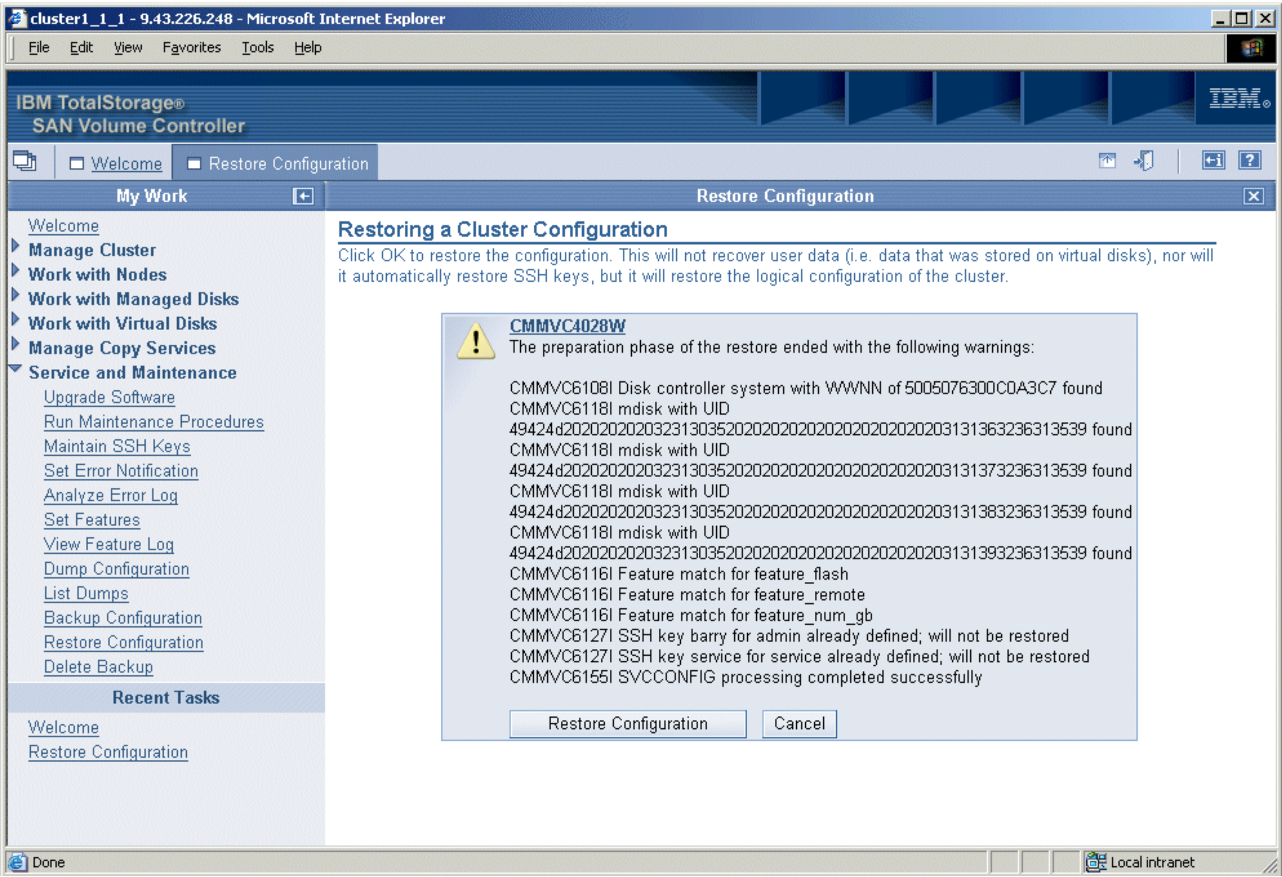


Figure 9-104 Restoring the configuration is successful

6. You are now ready to rediscover the vDisks from the hosts, restore the application data from backups to the corresponding vDisks using your backup and restore routines, and resume operations. Click the **X** icon in the upper right corner of the display area to close the Restoring a Cluster Configuration panel. Do not close the browser by mistake.

9.7.2 Deleting configuration backup

This section details the tasks that you can perform to delete the configuration backup files from the default folder in the SVC master console. You can do this if you already copied them to other external and secure place.

To delete the SVC Configuration backup files, perform the following steps:

1. From the SVC Welcome page, click the **Service and Maintenance** options and then the **Delete Configuration** link.

2. On the Deleting a Cluster Configuration panel (Figure 9-105), click the **OK** button to confirm the deletion. This deletes the C:\Program Files\IBM\svcconsole\cimom \backup*SVCclustername* folder (where *SVCclustername* is the SVC cluster name on which you are working) on the SVC master console and all its contents.

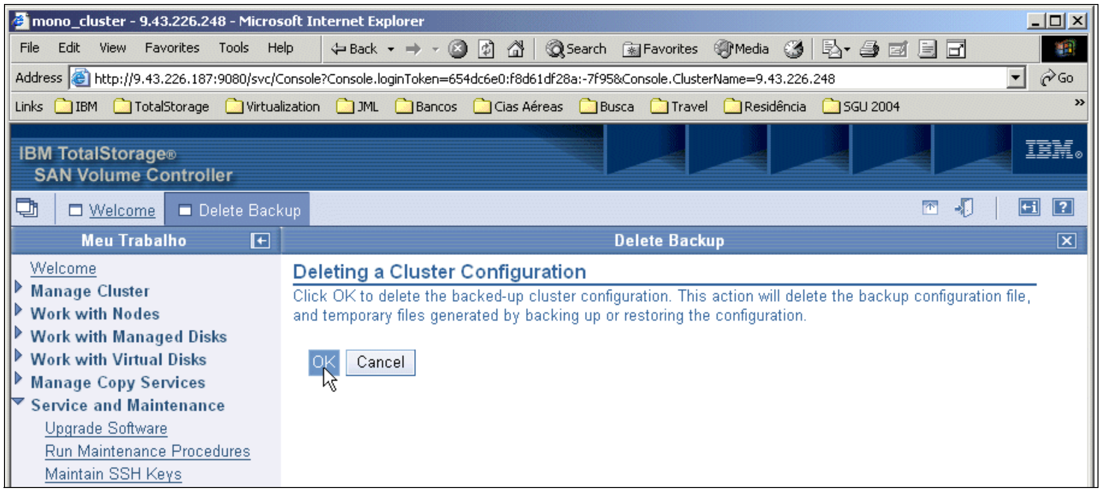


Figure 9-105 Deleting a cluster configuration

3. Click **Delete** to confirm the deletion of the configuration backup data. See Figure 9-106.

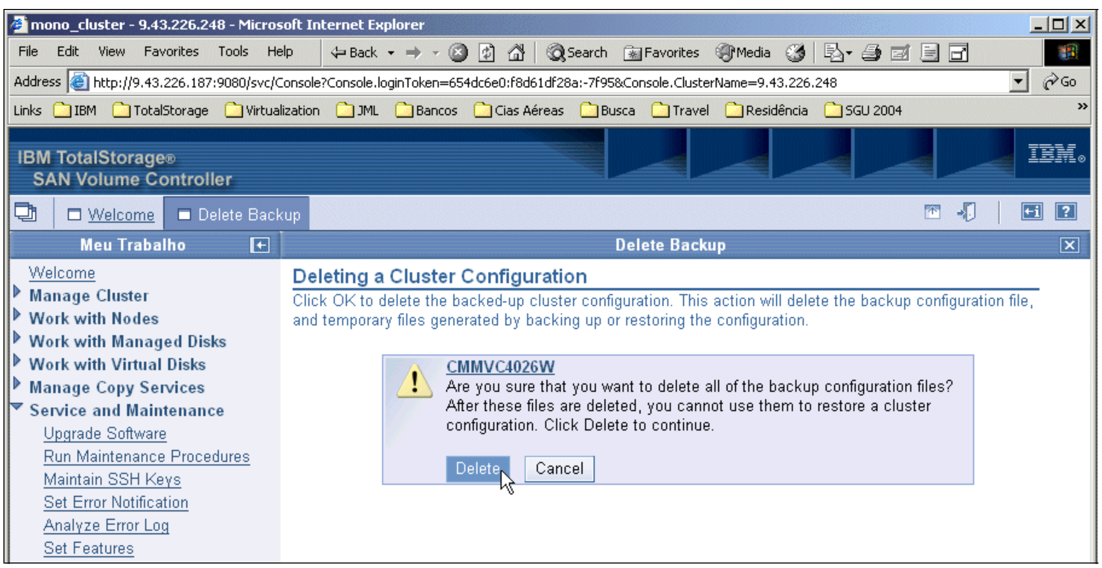


Figure 9-106 Deleting a Cluster Configuration confirmation message

4. Click the **X** icon in the upper right corner of the display area to close the Deleting a Cluster Configuration panel. Do not close the browser by mistake.

10

Copy Services: FlashCopy

This chapter introduces and describes the FlashCopy capability of the IBM TotalStorage SAN Volume Controller (SVC).

10.1 FlashCopy

This section looks at the fundamental principles of the operation of FlashCopy.

10.1.1 How it works

FlashCopy makes a copy of a set of source virtual disks to a set of target virtual disks. The original contents of the target virtual disks are lost. After the copy operation occurs, the target virtual disks have the contents of the source virtual disks as they existed at a single point in time (PIT), known as a *T(0) copy*. That is to say that, although the copy operation takes finite time, the resulting data at the target appears as though the copy was made instantaneously. This time is several orders of magnitude less than the time which is required to copy the data using conventional techniques.

As shown in Figure 10-1, both the source and target virtual disks are available for read and write operation, although not all the data is copied across from the source to target volumes.

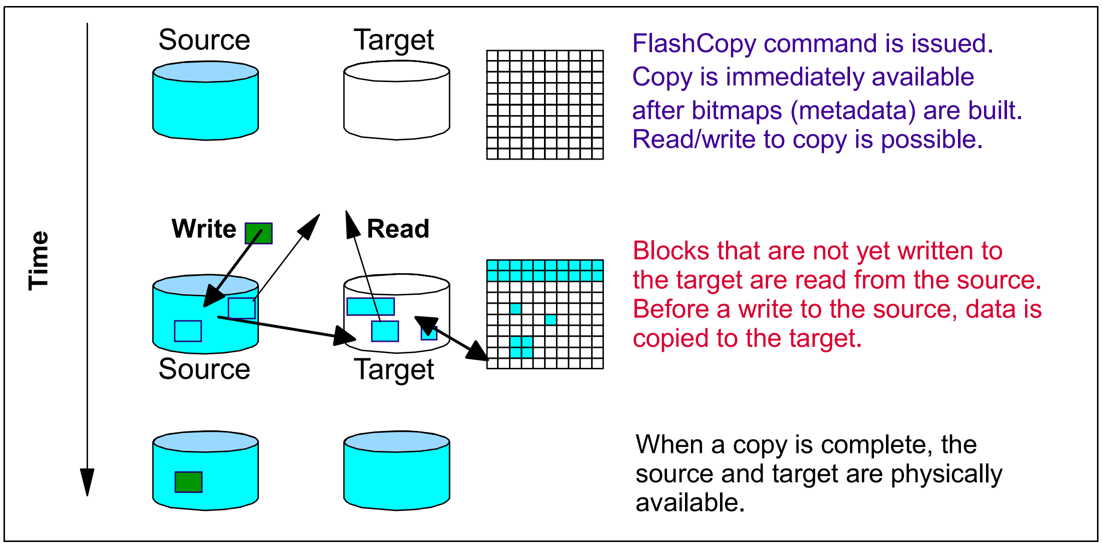


Figure 10-1 Implementation of SVC FlashCopy

Rationale

Point-in-time copy techniques are used to help solve the problem that it is difficult to make a consistent copy of a data set, which is being constantly updated. If a copy of a data set is taken using a technology that did not provide point-in-time semantics and the data set changes during the copy operation, then the resulting copy may contain data that is not self consistent.

For example, a reference to an object may be copied earlier than the object itself and the object is moved before it is itself copied. In this case, the copy contains the referenced object at its new location, but the reference points to the old location.

FlashCopy applications

The business applications for FlashCopy are many and various. An important use is for taking consistent backups of changing data. In this application, a FlashCopy is created to capture a point in time. The resulting image is backed up to tertiary storage such as tape. After the copied data is on tape, the FlashCopy target is redundant.

Practical uses for FlashCopy include those that are explained in the following sections.

Moving and migrating data
Anytime you need to move data from one server to another, FlashCopy can be useful. Do not forget to quiesce disk access before you make a FlashCopy and verify the consistency of the data before you attach it to the target server.

Moving workload
In the same way that you can move data from one server to another, you can move workload between servers.

Backup
FlashCopy does not usually speed up your backup, but it allows you to run your application while you back up. Therefore, the backup speed becomes less important for you. You may need fewer tape drives and a lower performance backup server. You only need to shut down your application for the time the FlashCopy task is started and can restart it almost immediately.

You can keep FlashCopy targets online after you back them up for some time. Therefore, you can copy the files that need to be restored from the FlashCopy target rather than restoring from tape. FlashCopy also enables you to perform backups whenever you want (not only during the off-shift period) because you do not need to wait for a lighter server load to do the backup.

Application testing
You can test new applications and new operating system releases against a FlashCopy of your production data. The risk of data corruption is eliminated, and your application does not need to be taken offline for an extended period of time to perform the copy of the data.

Other examples
Data mining is a good example of an area where FlashCopy can help you. Data mining can now extract data without affecting your application.

Moving from a single host to a cluster
In a clustered environment, two or more servers are accessing the same resources, such as disk drives. Different cluster models are available, which provide the following benefits. A cluster increases the availability of the data, as control of the resources from a server that is not available anymore is transferred to the remaining server or servers. The failover, failback, and access to the resources are controlled by special cluster software running on all servers within a cluster. Depending on the cluster model, the workload may be shared between the nodes within a cluster. In addition, if you are currently running your applications in a non-clustered environment and you plan to move to a clustered environment, you can use SVC FlashCopy functionality to create a test environment for the cluster quickly and easily with minimal impact on production.

Testing a clustered environment
If there is already a clustered environment, you can use FlashCopy to duplicate the shared data on the target volume or volumes.

Data backup with minimal impact on production
SVC FlashCopy functionality can be used to perform online backups of production data. When the data on the FlashCopy target is immediately available after the FlashCopy relationship is established, you need to ensure a consistent state of the data on the source volume for a few seconds only. After the short time of FlashCopy establishment, the data on the target can be written to backup media immediately. This decreases the time of backup windows of production data significantly.

Overview of FlashCopy

FlashCopy supports these features:

- The target is the time-zero copy of the source (known as *FlashCopy mappings*).
- The source virtual disk and target virtual disk are available (almost) immediately.
- Consistency groups are supported to enable FlashCopy across multiple logical unit numbers (LUNs) (up to 128 groups per cluster).
- The target may be updated independently of the source.
- Bitmaps are maintained in both nodes of the input/output (I/O) group.
- It is useful for test, backup, and improved availability.

10.1.2 FlashCopy mappings

In the SVC, FlashCopy occurs between a source virtual disk and a target virtual disk. The virtual disks must be the same size. The minimum granularity that SVC supports for FlashCopy is an entire virtual disk. The FlashCopy part of a virtual disk is not supported.

The source and target virtual disks must both be managed by the same SVC Cluster, but may be in different I/O groups within that Cluster. SVC FlashCopy associates a source virtual disk and a target virtual disk together in a FlashCopy mapping. Each virtual disk may be a member of only one FlashCopy mapping and a FlashCopy mapping always has exactly one source and one target virtual disk. Therefore, it is not possible for a virtual disk to simultaneously be the source for one FlashCopy mapping and the target for another.

Virtual disks (vDisk), which are members of a FlashCopy mapping, cannot have their size increased or decreased while they are members of the FlashCopy mapping. A SVC supports the creation of enough FlashCopy mappings to allow every virtual disk to be a member of a FlashCopy mapping.

A *mapping* is the act of creating a relationship between a source vDisk and a target vDisk. Mappings can be stand-alone or combined in consistency groups. You can perform the act of preparing, starting, or stopping on either the stand-alone mapping or the consistency group. Figure 10-2 illustrates the concept of FlashCopy mapping.

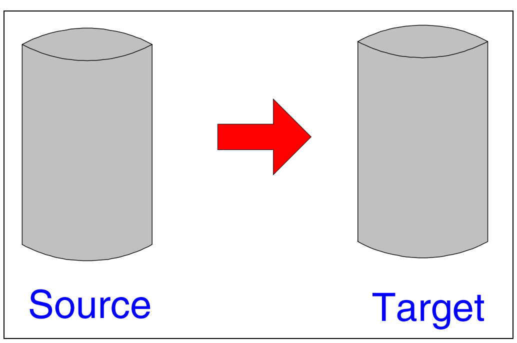


Figure 10-2 FlashCopy mapping

10.1.3 Consistency groups

Consistency groups address the issue that the using application may have related data which spans multiple virtual disks. FlashCopy must be performed in a way which preserves data integrity across multiple virtual disks. A requirement for preserving the integrity of data being written is to ensure that "dependent writes" are executed in the application's intended sequence.

FlashCopy mappings must be part of a consistency group, either part of their own group or part of default group 0. FlashCopy commands can be issued to the consistency group, the FlashCopy mapping name, or the FlashCopy mapping ID.

Figure 10-3 illustrates the concept of FlashCopy consistency groups.

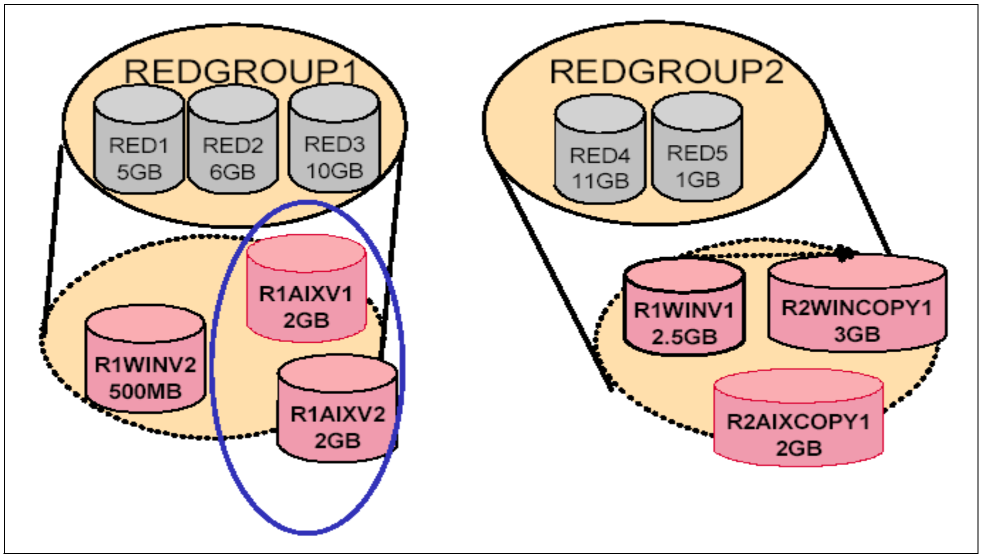


Figure 10-3 FlashCopy consistency groups

Dependent writes

Consider the following typical sequence of writes for a database update transaction:

1. Execute a write to update the database log indicating that a database update is about to take place.
2. Execute a second write to update the database.
3. Execute a third write to update the database log indicating that the database update has completed successfully.

The database ensures the correct ordering of these writes by waiting for each step to complete before starting the next. However if the database log (updates 1 and 3) and the database itself (update 2) are on different virtual disks and a FlashCopy mapping is started during this update, then you need to exclude the possibility that the database itself is copied slightly before the database log resulting in the target virtual disks seeing writes (1) and (3) but not (2).

In this case, if the database were restarted from a backup made from the FlashCopy target disks, the database log would indicate that the transaction had completed successfully when, in fact, that is not the case. The transaction is lost and the integrity of the database is in question.

Therefore, it may be the case that, to create a consistent image of the client data, it is necessary to perform a FlashCopy operation on multiple virtual disks as an atomic operation. To meet this need, SVC supports the concept of a consistency group. A consistency group contains a number of FlashCopy mappings. A FlashCopy mapping must either be a member of a consistency group or in the special *pseudo consistency group* 0 which is discussed later. A consistency group can contain an arbitrary number of FlashCopy mappings up to the maximum number of FlashCopy mappings supported by a SVC Cluster. SVC allows the Start command which causes the point in time copy to occur to be directed at a consistency group.

In this case, all of the FlashCopy mappings in the consistency group are started at the same time, resulting in a point in time copy which is consistent across all of the FlashCopy mappings which are contained in the consistency group.

Consistency group zero

For clients who do not need the complexity of consistency groups, SVC allows a FlashCopy mapping to be treated as an independent entity. In this case, the FlashCopy mapping reports that it is a member of a pseudo consistency group 0.

For FlashCopy mappings that are configured in this way, the **prepare** and **start** commands are directed at the FlashCopy mapping name or FlashCopy mapping ID rather than the consistency group ID. A **prepare** or **start** command directed toward a FlashCopy mapping, which is a member of any other consistency group, is illegal and fails because pseudo consistency group zero cannot be started or prepared.

Limits

SVC supports 128 consistency groups per cluster including consistency group 0. These are numbered 0 to 127.

10.1.4 FlashCopy indirection layer

FlashCopy requests a copy from the source virtual disk to the target virtual disk. It does this by using a Web interface, command line interface (CLI), or scripts. Then it creates a FlashCopy relationship between the source and target virtual disks. It establishes an algorithm in which it:

- Flushes write data in cache on a source virtual disk or disks
- Places cache into write-through on the source
- Discards cache on the target
- Establishes a sync point on all virtual disks
- Enables cache on both the source and target

FlashCopy provides the semantics of a point in time copy by using an indirection layer which intercepts I/Os targeted at both the source and target virtual disks. The act of starting a FlashCopy mapping causes this indirection layer to become active in the I/O path. This occurs as an atomic command across all FlashCopy mappings in the consistency group. The indirection layer makes a decision about each I/O. This decision is based upon:

- The virtual disk and logical block number (LBA) to which the I/O is addressed
- Its direction (read or write)
- The state of an internal data structure, the FlashCopy bitmap

The Indirection layer either allows the I/O through to the underlying storage, redirects the I/O from the target virtual disk to the source virtual disk, or stalls the I/O while it arranges for data to be copied from the source virtual disk to the target virtual disk.

Grains and the FlashCopy bitmap

You can tune the background copy rate, that is the speed at which data is copied from source to target volumes, as explained here:

1. The volumes are copied from the source to the target.
2. The backup copy rate establishes a back-end copy rate goal. A value of 0 has the same affect as the ESS NOCOPY option. The data rate is 128 KB/sec. to 64 MB/sec.
3. Attempts to read and write data that is already copied proceed as normal.

4. Attempts to process data not yet copied are intercepted. Use copy on demand or read from source as necessary.
5. The creates the affect of an instant copy.

When data is copied from the source virtual disk to the target virtual disk, it is copied in units of address space known as *grains*. In the SVC, the grain size is 256 K. The FlashCopy bitmap contains one bit for each grain. The bit records whether the associated grain has yet been split by copying the grain from the source to the target. The rate at which the grains are copied across from source to target volumes can be altered.

Source reads

Reads of the source are always passed through to the underlying source disk.

Target reads

In order for FlashCopy to process a read from the target disk, it must consult its bitmap. If the data being read is already copied to the target, then the read is sent to the target disk. If it has not, then the read is sent to the source disk. Clearly, this algorithm requires that while this read is outstanding, no writes are allowed to execute which would change the data being read from the source. The SVC satisfies this requirement using by a cluster wide locking scheme.

> **Note:** The current implementation of FlashCopy limits the number of concurrent reads to an unsplit target grain to one. If more than one concurrent read to an unsplit target grain is received by the FlashCopy mapping layer, then they are serialized.

Writes to the source or target

Where writes occur to source or target to an area (grain) which has not yet been copied, these will usually be stalled while a copy operation is performed to copy data from the source to the target, to maintain the illusion that the target contains its own copy. A specific optimization is performed where an entire grain is written to the target virtual disk. In this case the new grain contents are written to the target virtual disk, and if this succeeds, then the grain is marked as split in the FlashCopy bitmap without a copy from the source to the target having been performed. If the write fails, then the grain is not marked as split.

FlashCopy follows the algorithm of the *indirection layer*. You can think of the FlashCopy indirection layer as the I/O traffic cop when a FlashCopy relationship is underway. The I/O is intercepted and may be handled differently depending on the state and completion of the FlashCopy. Table 10-1 summarizes this algorithm that is followed by FlashCopy.

Table 10-1 Summary of FlashCopy I/O path actions

Virtual disk being accessed	Has the grain been copied yet?	Host I/O operations	
		Read	**Write**
Source	No	Read from source	Copy grain to target then write source
	Yes	Read from source	Write to source
Target	No	Read from source	Copy grain to target then write target
	Yes	Read from target	Write to target

This can also be expressed as shown in Figure 10-4.

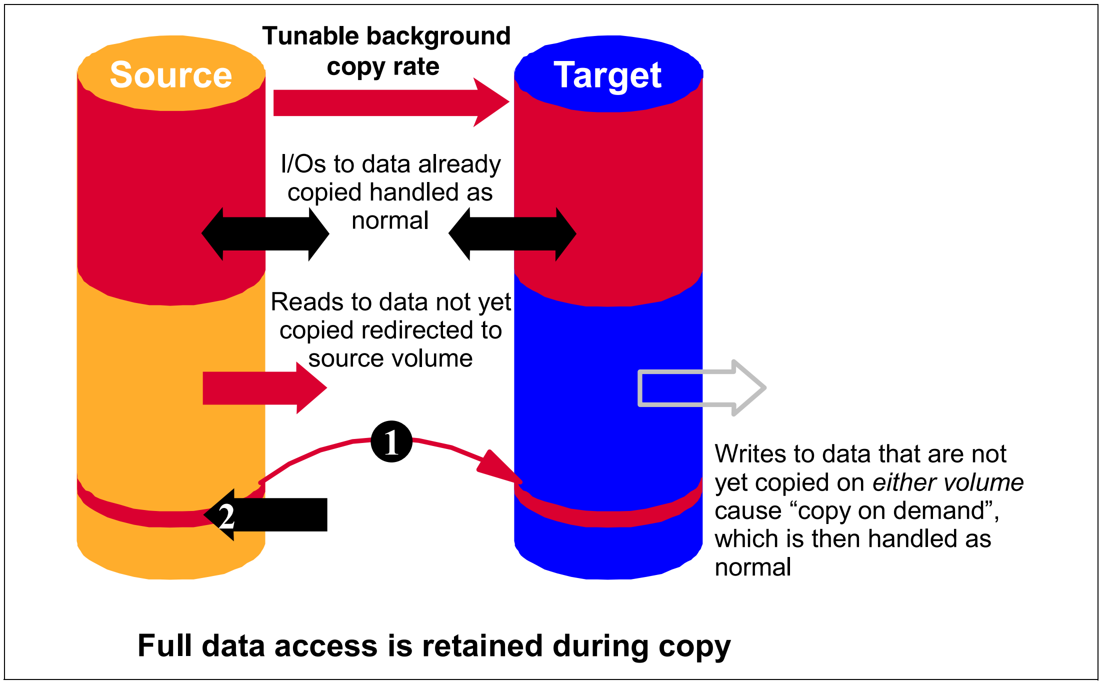


Figure 10-4 I/O processing with FlashCopy

Interaction with the cache

This copy-on-write process can introduce significant latency into write operations. To isolate the using application from this latency, the FlashCopy indirection layer is placed logically below the cache. This means that the copy latency is typically only seen on destage from the cache rather than for write operations from a using application which otherwise may be blocked waiting for the write to complete Figure 10-5.

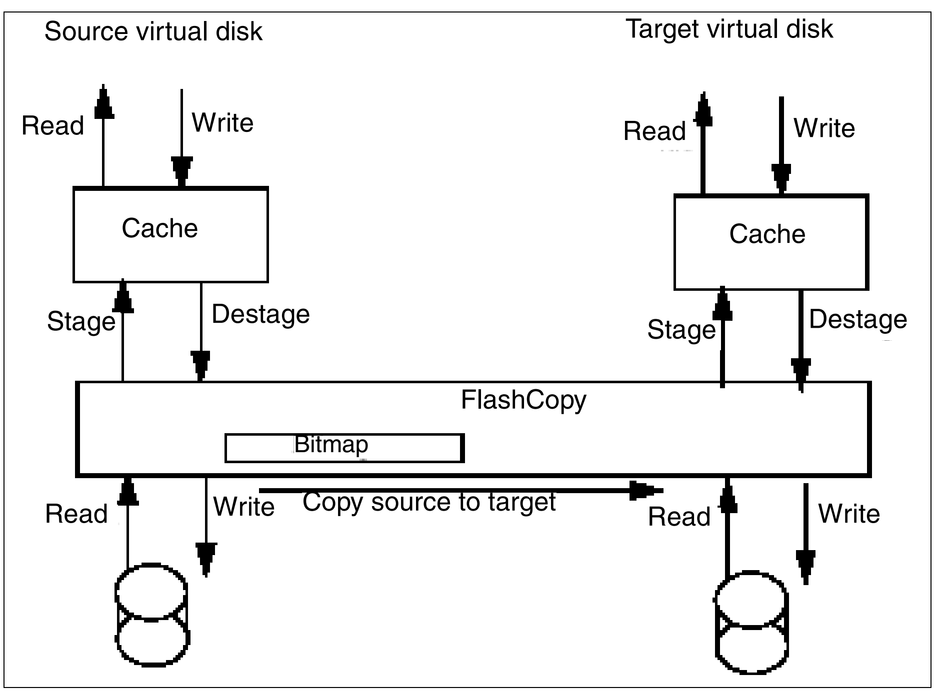


Figure 10-5 Conceptual block diagram showing SVC data flows

10.1.5 FlashCopy rules

There is a per I/O group limit of 16 TB on the quantity of source virtual disk address space, which may participate in FlashCopy mappings. This address space is allocated in units of 8 GB. That is to say that creating a FlashCopy mapping between a pair of virtual disks whose size is less than 8 GB consumes 8 GB of FlashCopy mapping address space.

The maximum number of FlashCopy mappings that are supported is 512, which is half the maximum number of vDisks, 1024, that are currently supported. This means that there is a maximum of 512 source and 512 target vDisks.

Here are some rules to consider:

- There is one-to-one mapping of the source to the target virtual disk.
- The source and target must be within the same cluster but can be across I/O groups.
- The complete source virtual disk is copied to the target virtual disk.
- Minimum granularity is the entire vDisk.
- The source and target must be the same size.
- There is no multiple relationship support, therefore, one FlashCopy per volume.
- There is no incremental support.
- An existing FlashCopy relationship can be stopped.
- A FlashCopy relationship is persistent until it is explicitly unmapped.
- The size of source and target virtual volumes cannot be altered (increased or decreased) during the mapping.
- There is no limit on the number of virtual disks that are supported. However, the maximum quantity of source virtual disk per I/O group is 16 TB.

FlashCopy and image mode disks

You can use FlashCopy with an image mode vDisk. The target vDisk must be exactly the same size as the source. When the target vDisk is created, the size should be specified to be the same as the size of the image mode vDisk. The vDisk is presented with the size specified, regardless of the number of extents used to create it.

Extents are not used in verifying equal size. The comparison is based on sizes as presented to the user. The following steps explain how to verify the size in bytes. The number of bytes is always rounded up to fit 512 bytes (disk sector size).

1. To get the vDisk size in bytes, run the following command:

   ```
   svcinfo lsvdisk -bytes (name of vdisk)
   ```

2. Run the **`svctask mkvdisk`** command with the -size and -units options to create the vDisk the same number of bytes in size as you get from Step 1.
3. You can use the expand and shrink vDisk commands to make the size right since they also support the -units option.

10.1.6 FlashCopy mapping events

There is a series of events that take place during the FlashCopy mapping process:

- **Create**: A new FlashCopy mapping is created between the specified source virtual disk and the specified target virtual disk. The operation fails if either the source or target virtual disks is already a member of a FlashCopy mapping. The operation also fails if the source and target virtual disks are difference sizes.

- **Prepare**: The prepare command is directed to either a consistency group for FlashCopy mappings which are members of a normal consistency group or to the mapping name for FlashCopy mappings which are members of the special consistency group 0. The prepare command places the FlashCopy mapping in the preparing state. It is important to note that the act of preparing for start may corrupt any data which previously resided on the target virtual disk since cached writes are discarded. Even if the FlashCopy mapping is never started, the data from the target may be logically changed by the act of preparing for start. See Figure 10-6.

1. Associate the source dataset with a target location (one or more virtual disks).
2. Create FlashCopy mapping between each source virtual disk to its corresponding target virtual disk. The target is equal in size to the source.
3. Discontinue access to the target (application dependent).
4. Prepare (pre-trigger) the FlashCopy:
   - Flush cache for the source.
   - Discard cache for the target.
5. Start (trigger) the FlashCopy:
   - Pause I/O (briefly) on the source.
   - Resume I/O on the source.
   - Start I/O on the target.

Figure 10-6 FlashCopy: Sequence of events

- **Flush done**: The FlashCopy relationship moves from the preparing state to the prepared state automatically after all cached data for the source is flushed and all cached data for the target is invalidated.
- **Start**: After all of the FlashCopy mapping or mappings in a consistency group are in the prepared state, the FlashCopy relationships can be started. Some other FlashCopy products refer to this event as *triggering* the FlashCopy.

  During the start command, the following actions occur:

  – New reads and writes to all source virtual disks in the consistency group are paused in the cache layer until all ongoing reads and writes below the cache layer are completed.

  – After all FlashCopy mappings in the consistency group are paused, internal metadata is set to allow FlashCopy operation.

  – After all FlashCopy mappings in the consistency group have their metadata set, read and write operations are unpaused on the source virtual disks.

  – The target virtual disks are brought online.

  As part of the start command, read, and write caching is enabled for both the source and target virtual disks.
- **Modify**: A FlashCopy mapping has two properties which can be modified. These are the background copy rate and the consistency group. The background copy rate can be modified in any state. However attempting to modify the consistency group in any state other than idle/copied fails.
- **Stop**: There are two mechanisms by which a FlashCopy mapping can be stopped. It can be stopped by a user command or by an I/O error.
- **Delete**: This event requests that the specified FlashCopy mapping be deleted. If the FlashCopy mapping is in the stopped state, the Force flag must be used. Deleting a FlashCopy mapping in the stopped state may allow unflushed write data from the cache to

be destaged to what was the target virtual disk. This does not affect the data integrity of the system because following a forced delete, nothing can be said about the contents of the target virtual disk. The data contained in the target virtual disk could be anything. The destaging of old data to what was the target virtual disk does not affect the future use of the virtual disk. Any new data is written over this old data, in the cache or on disk.

- **Flush failed**: If the flush of data from the cache cannot be completed, then FlashCopy mapping enters the *stopped* state.
- **Copy complete**: After every grain of the source and target is copied, the source and target are independent and the state machine enters the *copied* state.

10.1.7 FlashCopy mapping states

The FlashCopy mapping states are explained in the following sections.

Idling/copied

Read and write caching is enabled for both the source and the target. A FlashCopy mapping exists between the source and target but they behave as independent virtual disks in this state.

Copying

The target is "flashed" from the source. Reads and writes are executed on the target as though the contents of the source were instantaneously copied to the target during the start command The source and target can be independently updated. Internally, the target depends on the source for some tracks. Read and write caching is enabled on the source and the target.

Stopped

The FlashCopy was stopped either by user command or by an I/O error. When a FlashCopy mapping is stopped, any useful data in the target virtual disk is lost. Because of this, while the FlashCopy mapping is in this state, the target virtual disk is in the *offline* state. To regain access to the target, the mapping must be started again (the previous FlashCopy is lost) or the FlashCopy mapping must be deleted. While in the *stopped* state, any data that was written to the target virtual disk and was not flushed to disk before the mapping was stopped is pinned in the cache. It cannot be accessed but does consume resource. This data is destaged after a subsequent `delete` command or discarded during a subsequent `prepare` command. The source virtual disk is accessible and read and write caching is enabled for the source.

Suspended

The target was *point-in-time copied* from the source, and was in the copying state. Access to the metadata is lost. As a consequence, both the source and target virtual disks are offline. The background copy process is halted. When the metadata becomes available again, the FlashCopy mapping returns to the *copying* state. Then access to the source and target virtual disks are restored, and the background copy process is resumed. Unflushed data that was written to the source or target before the FlashCopy was suspended is pinned in the cache, consuming resources, until the FlashCopy mapping leaves the suspended state.

Preparing

Placing the FlashCopy function below the cache addresses, the write latency problem introduces a requirement. That is, in order for the resulting copy to be consistent, there must be no read or write data for the target and no write data for the source in the cache at the time that the FlashCopy operation is started.

Performing the necessary flush and invalidate operations as part of the start unnecessarily delays the start command and introduces a large latency for I/Os received after the start command. These I/Os need to wait for the start command to complete the cache flush. For this reason, SVC FlashCopy supports the **`prepare`** command, which prepares for a FlashCopy start while I/O continues to the source virtual disk.

In the preparing state, the FlashCopy mapping is prepared for starting by:

- Flushing any modified write data associated with the source virtual disk from the cache. Read data for the source is left in the cache.
- Placing the cache for the source virtual disk into write through mode, so that subsequent writes wait until data has been written to disk before completing the write command received from the using host.
- Discarding any read or write data associated with the target virtual disk from the cache.

While in this state, writes to the source virtual disk experience additional latency because the cache is operating in write through mode. While the FlashCopy mapping is in this state, the target virtual disk is in the offline state.

Before starting the FlashCopy mapping, it is important that any caches above SVC, perhaps in the using operating system, are also instructed to perform the same flush of source write data and invalidate of target read and write data that the SVC cache performs. A description of how this is arranged is outside the scope of this document.

Prepared

The FlashCopy mapping is ready to perform a start. While the FlashCopy mapping is in this state, the target virtual disk is in the *offline* state.

While in this state, writes to the source virtual disk experience additional latency because the cache is operating in write through mode.

Table 10-2 shows the various mapping states that are possible.

Table 10-2 FlashCopy mapping state summary (a) due to FlashCopy

State	Source		Target	
	Online/Offline	Cache state	Online/Offline	Cache state
Idling/Copied	Online[a]	Write-back	Online[a]	Write-back
Copying	Online[a]	Write-back	Online[a]	Write-back
Stopped	Online[a]	Write-back	Offline[a]	-
Suspended	Offline	Write-back	Offline[a]	-
Preparing	Online[a]	Write-through	Offline[a]	-
Prepared	Online[a]	Write-through	Offline[a]	-

10.1.8 Background copy rate

A FlashCopy mapping has a property Background Copy Rate. This is expressed as a percentage and can take values between 0 and 100. The background copy rate can be changed when the FlashCopy mapping is in any state. If a value of 0 is specified, then background copy is disabled. This is equivalent to the ESS NOCOPY option. One expected use of this is for short lived FlashCopy mapping, which is to be used for backup purposes only. Since the source data set is not expected to change much during the lifetime of the

FlashCopy mapping, it is more efficient in terms of managed disk I/Os not to perform a background copy. The relationship of the background copy rate value to the attempted number of grains to be split per second is shown in Table 10-3.

Table 10-3 Point-in-time copy rate

User percentage	KB per second	Grains per second
1-10	128	0.5
11-20	256	1
21-30	512	2
31-50	2048	8
91-100	64 MB	256

The grains/sec numbers represent goals that the code tries to achieve. SVC is unable to achieve these goals if insufficient bandwidth is available from the SVC nodes to the physical disks making up the managed disks after taking into account the requirements of foreground I/O. If this situation arises, then background copy I/O contends for resources on an equal basis with I/O arriving from hosts. Both tend to see an increase in latency and consequential reduction in throughput with respect to the situation had the bandwidth not been limited. Degradation is graceful.

Both background copy and foreground I/O continue to make forward progress, and do not stop, hang, or cause the node to fail. The background copy is performed by one of the nodes belonging to the I/O group in which the source virtual disk resides. This responsibility is failed over to the other node in the I/O group in the event of the failure of the node performing the background copy. The background copy is performed “backwards” that is to say it starts with the grain containing the highest LBAs and works backward toward the grain containing LBA 0. This is done to avoid any unwanted interactions with sequential write streams from the using application.

10.1.9 Synthesis

The FlashCopy functionality in SVC simply creates copy virtual disks. All the data in the source virtual disk is copied to the destination virtual disk. This includes operating system control information as well as application data and metadata.

Some operating systems are unable to use FlashCopy without an additional step which is termed *synthesis*. In general, synthesis performs some transformation on the operating system metadata in the target virtual disk so that the operating system can use the disk. Operating system specifics are discussed in Appendix B, “Open system specifics” on page 495.

10.1.10 Metadata management

A bitmap is maintained, with a bit for each grain in the copy FlashCopy mapping.

The bitmap is maintained in nonvolatile storage on the two SVC nodes making up the I/O group for the source virtual disk. While both of these nodes are functioning members of the Cluster the two copies of the bitmaps are updated and kept consistent with one another.

Other nodes, which are not members of the I/O group for the source virtual disk maintain a volatile *pessimistic bitmap* of the FlashCopy mapping. Pessimistic means that the bitmap may hold a bit which indicates that the grain must be copied, where in fact it does not need to be copied. Access to grains which have not yet been copied are coordinated by the source extent owner only. The actual reads and writes are performed on the node that wants to access the grain.

No provision is made for storing bitmaps within the back-end storage.

10.1.11 I/O handling

To understand how the cluster behaves in the presence of errors, it is necessary to describe how the FlashCopy I/O path operates. It is important to note that the binding of virtual disks to I/O groups affects only the cache and layers above the cache. Below the cache in the SVC software stack, virtual disks are available for I/O on all nodes.

As mentioned earlier, the background copy is performed by one of the nodes belonging to the source virtual disk's I/O group. It should be clear that any of the nodes in a SVC cluster may want to submit an I/O from the FlashCopy layer to a virtual disk.

It is assumed that the nodes and managed disks in the cluster have complete connectivity. If the nodes comprising the source I/O group do not have access to the managed disk extents comprising the target virtual disk, then an I/O error occur. And FlashCopy mapping is probably stopped. Similarly, the nodes comprising the target I/O group must have access to the managed disk extents comprising the source virtual disk.

10.1.12 Serialization of I/O by FlashCopy

In general, the FlashCopy function in SVC introduces no explicit serialization into the I/O path. Therefore, many concurrent I/Os are allowed to the source and target virtual disks. However, there is a lock for each grain. The lock can be taken shared or exclusive. The lock is taken in the following modes under the following conditions:

- The lock is taken shared for the duration of a read from the target virtual disk which touches a grain which is yet to be split.
- The lock is taken exclusive during a grain split. This happens prior to FlashCopy actioning any destage (or write through) from the cache to a grain which is yet to be split (the destage waits for the grain to be split). The lock is held during the grain split and released before the destage is processed.

If the lock is held shared and another process wants to take the lock shared then this request is granted unless a process is already waiting to take the lock exclusive. If the lock is held shared and it is requested exclusive, then the requesting process must wait until all holders of the shared lock free it. Similarly, if the lock is held exclusive then a process wanting to take the lock in either shared or exclusive mode, must wait for it to be freed.

10.1.13 Error handling

When a FlashCopy mapping is *not* copying, the FlashCopy function does not affect the error handling or reporting of errors in the I/O path. *Only* when a FlashCopy mapping is copying, error handling and reporting are affected by FlashCopy. These scenarios are described in the following sections.

Node failure

To explain what the product does, it is necessary to explain a little of how it does it. Normally, two copies of the FlashCopy bitmaps are maintained, one on each of the two nodes making up the I/O group of the source virtual disk. When a node fails, one copy of the bitmaps for all FlashCopy mappings whose source virtual disk is a member of the failing node's I/O group become inaccessible. FlashCopy continues with a single copy of the FlashCopy bitmap being stored nonvolatile in the remaining node in the source I/O group. The cluster metadata is updated to indicate that the missing node no longer holds up to date bitmap information.

When the failing node recovers or a replacement node is added to the I/O group, up-to-date bitmaps are re-established on the new node. Once again, it provides a redundant location for the bitmaps.

If access to both nodes in an I/O group is lost, or if access to the single remaining valid copy of a FlashCopy bitmap is lost, then any FlashCopy mappings, which were in the copying state and for which the source was in the lost I/O group, enter the suspended state. As stated access to both the source and target virtual disks in the FlashCopy mapping is suspended.

If both nodes in the I/O group to which the target virtual disk belongs become unavailable, then host access to the target virtual disk is not possible. This is standard SVC behavior and is unaffected by FlashCopy. The FlashCopy state is unaffected by this, and any background copy continues.

Path failure (path offline state)

In a fully functioning cluster, all nodes have a software representation of every virtual disk in the cluster within their application hierarchy. Since the storage area network (SAN) which links the SVC nodes to each other and to the managed disks is made up of many independent links, it is possible for some subset of the nodes to be temporarily isolated from some of the managed disks. When this happens, the managed disks are said to be *path offline* on some nodes.

> **Note:** Other nodes may see the managed disks as online, because their connection to the managed disks is still functioning.

When a managed disk enters the path offline state on a SVC node, all the virtual disks, which have any extent on the managed disk, also become path offline. Again, this happens only on the affected nodes. When a virtual disk is path offline on a particular SVC node, this means that host access to that virtual disk through the node will fail with SCSI sense indicating offline.

Path offline for the source vDisk

If a FlashCopy mapping is in the copying state and the source virtual disk goes path offline, then this path offline state is propagated to both source and target virtual disks. Again note that path offline is a state which exists on a per-node basis. Other nodes may not be affected. If the source virtual disk again becomes online then both the target and source virtual disks are brought back online

Path offline for the target vDisk

If the target virtual disk goes path offline but the source virtual disk is still online then above FlashCopy in the application stack, only the target virtual disk is in the path offline state. The source virtual disk remains online.

I/O errors caused by path failures

Special handling is invoked where:

- A FlashCopy mapping is copying.
- The source has *not* suffered a path failure (online).
- The target *has* suffered a path failure (path offline).
- A copy operation is requested, either by:

  – A background copy operation
  – A host I/O to the source (but specifically not the target)

- The copy operation fails when the write is attempted to the target.

In this case, the FlashCopy mapping is placed into the stopped state. I/O to the source is allowed to continue unaffected by FlashCopy. The target virtual disk is held OFFLINE and all I/O to it fail.

Although this situation arises on just one node, the FlashCopy mapping state is held on a cluster-wide basis and is propagated to all nodes. In all other cases, I/Os failing due to path failures simply result in the original I/O being failed because of the path failure.

10.1.14 Asynchronous notifications

FlashCopy raises informational error logs when mappings or consistency groups make certain state transitions as detailed below. These state transitions occur as a result of configuration events which complete asynchronously and the informational errors can be used to generate Simple Network Management Protocol (SNMP) traps to notify the user. Other configuration events complete synchronously, and no informational errors are logged as a result of these events.

- **PREPARE_COMPLETED**: This is logged when the FlashCopy mapping or consistency group enters the PREPARED state as a result of a user request to prepare. The user can now start (or stop) the mapping or group.
- **COPY_COMPLETED**: This is logged when the FlashCopy mapping or consistency group enters the IDLE_COPIED state when it was previously in the *copying* state. This indicates that the target disk now contains a complete copy and no longer depends on the source.
- **STOP_COMPLETED**: This is logged when the FlashCopy mapping or consistency group enters the STOPPED state as a result of a user request to Stop. It is distinct from the error that is logged when a mapping or group enters the *stopped* state as a result of an I/O error.

10.2 FlashCopy commands

This section lists and explains the commands used to create, modify, and delete FlashCopy copies. For complete details about the FlashCopy commands, see *IBM TotalStorage Virtualization Family SAN Volume Controller: Command-Line Interface User's Guide*, SC26-7544.

10.2.1 Create mapping (svctask mkfcmap)

The **`mkfcmap`** command enables you to create a new FlashCopy mapping, which maps a source virtual disk to a target virtual disk ready for subsequent copying. When executed, this command creates a new FlashCopy mapping logical object. This mapping persists until it is deleted. The mapping specified the source and destination virtual disks. The destination must be identical in size to the source, or the create will fail. The source and destination cannot be in an existing mapping. That is, a virtual disk can be either a source or destination disk in one and only one mapping. A mapping is triggered at the point in time the copy is required.

The mapping can optionally be given a name and assigned to a consistency group. These are groups of mappings that can be triggered at the same point in time. This allows multiple virtual disks to be copied at the same time, creating a consistent copy of multiple disks. This is required by some database products where the database and log files reside on different disks.

The background copy rate specifies the priority that should be given to completing the copy. If 0% is specified, the copy does not proceed in the background. The default is 50%

10.2.2 Modify mapping (svctask chfcmap)

This command allows you to modify certain attributes of an existing mapping. When executed, it allows you to modify the specified attributes of the mapping supplied. When modifying the name of a mapping, you cannot modify any of the other attributes at the same time.

Modifying the consistency group the mapping belongs to can only be done when the mapping is inactive. That is it has not been triggered, or if it has been triggered the copy has run to completion. Similarly, if the target consistency group is active, the mapping cannot be moved.

10.2.3 Delete mapping (svctask rmfcmap)

This command allows you to delete an existing mapping. When you execute this command, it attempts to delete the mapping specified. If the mapping is active, the command fails unless the force flag is specified.

Deleting a mapping only deletes the logical relationship between the two virtual disks. It does not effect the virtual disks themselves. However, forcing the delete renders the data on the destination virtual disk as inconsistent.

10.2.4 Prepare (pre-trigger) FlashCopy mappings (svctask prestartfcmap)

This command prepares a mapping for triggering. It flushes the cache of any data destined for the source virtual disk, and force the cache into write through until the mapping is triggered.

When executed, this command prepares either a single mapping or a group of mappings (on a consistency group basis) for subsequent triggering. The prepare step ensures that any data that resides in the cache for the source virtual disk is first flushed to disk. This ensures that the copy made is consistent with what the operating system thinks is on disk.

The mappings or consistency group enters the preparing state. When the prepare is complete, they change to the prepared state. At this point, the mapping or group is ready for triggering.

Preparing and the subsequent triggering is usually performed on a consistency group basis. Only mappings belonging to consistency group 0 can be prepared on their own.

10.2.5 Start (trigger) FlashCopy mappings (svctask startfcmap)

This command starts (triggers) a FlashCopy mapping or a group of mappings. This makes a point-in-time copy of the source virtual disk at the moment the command is executed.

When executed, this command triggers either a single mapping or a group of mappings (on a consistency group basis). Triggering means to take a point in time copy of the source virtual disk(s).

The mapping or group must first be prepared for triggering. See **`svctask prestartfcmap`** and **`svctask prestartfcconsistgrp`**. However, you can run this command with the optional prepare flag. This prepares the mapping or group and trigger as soon as the preparation is complete. Note that this means it is under the systems control when the trigger happens. That is, the prepare step may take some time to complete before the trigger is executed. If you want to control the triggering you should use the prepare command first.

The mappings or consistency group enter the copying state. The way the copy proceeds depends on the background copy rate attribute of the mapping. If the mapping is set to 0%, then only data that is subsequently updated on the source is copied to the destination. This means that the destination can only be used as a backup copy while the mapping exist in the copying state. If the copy is stopped, the destination is not be usable. If you want to end up with a duplicate copy of the source at the destination, you should ensure that the background copy rate is greater than 0. This means that the system copies all the data (even unchanged data) to the destination and eventually reaches the idle/copied state. At this time, you can delete the mapping and have a usable point in time copy of the source at the destination.

Triggering is usually performed on a consistency group basis. Only mappings that belong to consistency group 0 can be triggered on their own.

10.2.6 Stop FlashCopy mappings (svctask stopfcconsistgrp)

This command allows you to stop an active (copying) or suspended mapping or consistency group. When executed, this command stops either a single mapping or a group of mappings (on a consistency group basis). If the copy is stopped, the destination is not usable. The mapping or group needs to be reprepared and retriggered.

Stopping is usually performed on a consistency group basis. Only mappings belonging to consistency group 0 can be stopped on their own.

10.2.7 Create consistency group (svctask mkfcconsistgrp)

This command allows you to create a new FlashCopy consistency group. When executed, this command creates a new consistency group. The ID of the new group is returned.

10.2.8 Modify consistency group (svctask chfcconsistgrp)

This command allows you to modify the name of an existing consistency group. When executed, this command changes the name of the consistency group that is specified.

10.2.9 Delete consistency group (svctask rmfcconsistgrp)

This command allows you to delete an existing consistency group. When executed, this command deletes the consistency group specified. If there are mappings that are members of the group, the command fails unless the force flag is specified.

If you want to delete all the mappings in the consistency group as well, you must first delete the mappings, and then delete the consistency group.

10.3 Scripting

FlashCopy may be invoked from the SVC graphical user interface (GUI), but this may not make much sense if you plan to create a large number of FlashCopy copies at varying times. In this case, you can use the CLI and scripting. You can create your own customized scripts to automate a large number of tasks for completion at a variety of times and run them through the CLI.

10.4 Step-by-step guide to FlashCopy

The following example was carried out using two AIX (aix1 and aix2) host systems attached to the SAN Volume Controller. See Figure 10-7.

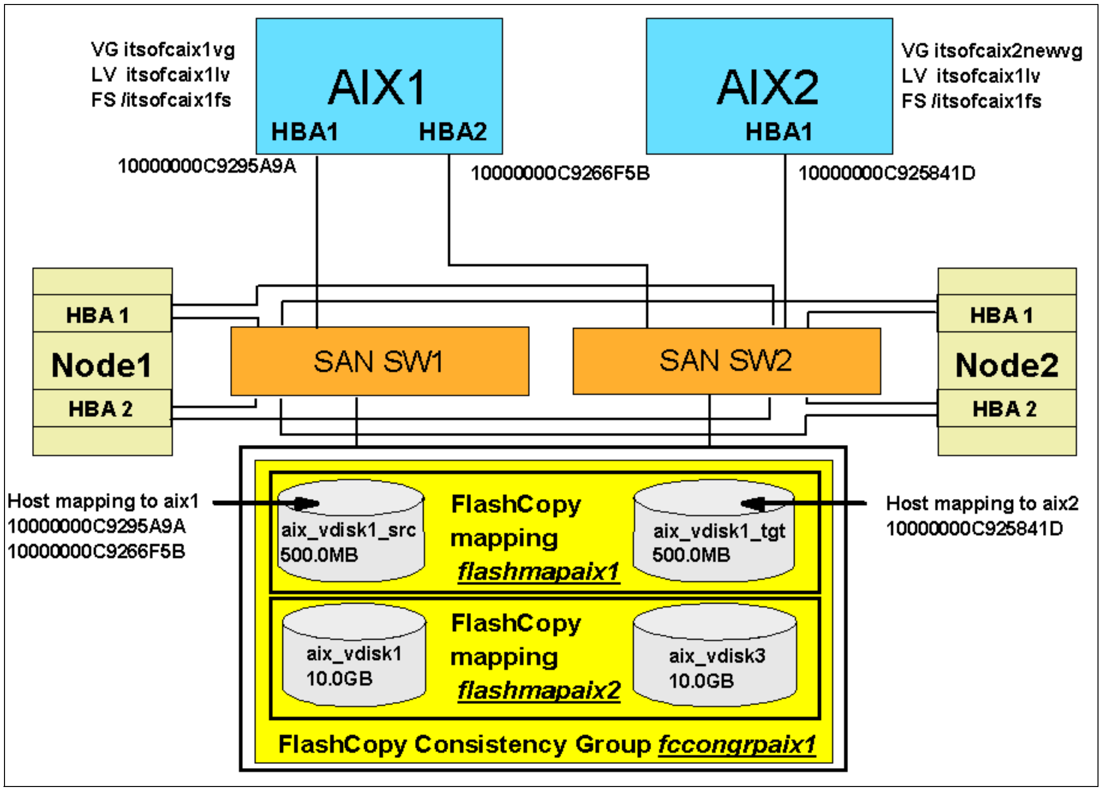


Figure 10-7 FlashCopy scenario

The vDisk aix_vDisk1_src is mapped and in production in host system aix1.

In this FlashCopy scenario, the following actions occur:

1. Create a FlashCopy consistency group fccongrpaix1.
2. Create the FlashCopy mapping between the source vDisk aix_vDisk1_src and the target vDisk aix_vDisk1_tgt in the FlashCopy consistency group fccongrpaix1.
3. Prepare the FlashCopy consistency group fccongrpaix1.
4. Start the FlashCopy consistency group fccongrpaix1.

The target vDisk aix_vDisk1_tgt is ready to be mapped and used by the host system aix2. Example 10-1 shows the storage definition on the firt system before FlashCopy occurs.

Example 10-1 Host system aix1 storage definition before FlashCopy

```
AIX1#lspv
hdisk0          0007905873b5c2fb                    rootvg
hdisk1          0007905873b5c364                    rootvg
hdisk2          none                                None
hdisk3          none                                None
hdisk4          none                                None
hdisk5          none                                None
```

```
vpath0          00079058479c31c5                    itsofcaix1vg

AIX1#lsattr -El vpath0
pvid          00079058479c31c50000000000000000        Data Path Optimizer Parent False
policy        df                                      Scheduling Policy          True
active_hdisk  hdisk2/600507680182001B200000000000000C Active hdisk               False
active_hdisk  hdisk3/600507680182001B200000000000000C
active_hdisk  hdisk4/600507680182001B200000000000000C
active_hdisk  hdisk5/600507680182001B200000000000000C
serial_number 600507680182001B200000000000000C        N/A                        False

AIX1#lsvg -o
itsofcaix1vg
rootvg

AIX1#df -k
Filesystem    1024-blocks      Free %Used    Iused %Iused Mounted on
/dev/hd4            16384      7936   52%     1281    16% /
/dev/hd2          1867776   1265812   33%    20504     5% /usr
/dev/hd9var         16384     10936   34%      433    11% /var
/dev/hd3            32768     29188   11%       58     1% /tmp
/dev/hd1            16384     15820    4%       18     1% /home
/proc                   -         -    -         -     -  /proc
/dev/hd10opt        32768     26592   19%      294     4% /opt
/dev/itsofcaix1lv       40960     39636    4%       17     1% /itsofcaix1fs

AIX1# more /itsofcaix1fs/file1.txt
this file was created on AIX1

>>>>>>>> unmount the file system before the FlashCopy

AIX1#unmount /itsofcaix1fs

AIX1#df -k

Filesystem    1024-blocks      Free %Used    Iused %Iused Mounted on
/dev/hd4            16384      7936   52%     1281    16% /
/dev/hd2          1867776   1265812   33%    20504     5% /usr
/dev/hd9var         16384     10936   34%      433    11% /var
/dev/hd3            32768     29188   11%       58     1% /tmp
/dev/hd1            16384     15820    4%       18     1% /home
/proc                   -         -    -         -     -  /proc
/dev/hd10opt        32768     26592   19%      294     4% /opt

>>>>>>>   now ready for FlashCopy
```

Example 10-2 shows the storage definition on the second host system before FlashCopy.

Example 10-2 Host system aix2 storage definition before FlashCopy

```
AIX2#lsvg -o
rootvg

AIX2#lspv
hdisk0        000b21dd2cf9b83d    rootvg
```

```
AIX2#df -k
Filesystem    1024-blocks      Free %Used    Iused %Iused Mounted on
/dev/hd4            32768     26392   20%     1164     8% /
/dev/hd2           393216     79132   80%    14361    15% /usr
/dev/hd9var         32768     30332    8%      152     2% /var
/dev/hd3            32768     31056    6%       60     1% /tmp
/dev/hd1            32768     31684    4%       18     1% /home

AIX2#lsdev -Cc adapter |grep fcs
fcs0     Available 10-90     FC Adapter
```

10.4.1 Checking the status of the FlashCopy feature

Check that the feature FlashCopy is either *on* or *enabled*.

Using the CLI

This is shown in Example 10-3.

Example 10-3 Feature PPRC/Remote Copy is on

```
SVC1:admin>svcinfo lslicense
feature_flash on
feature_remote on
feature_num_gb 2000
```

Using the GUI

Using the GUI, follow these steps:

1. From the left navigation panel, select **Service and Maintenance**.
2. Under Action in the left panel, select **Set Featurization**.
3. In the panel on the right, under Current Feature Settings, verify the feature settings for FlashCopy. As shown in Figure 10-8, the setting indicates *Enabled*.

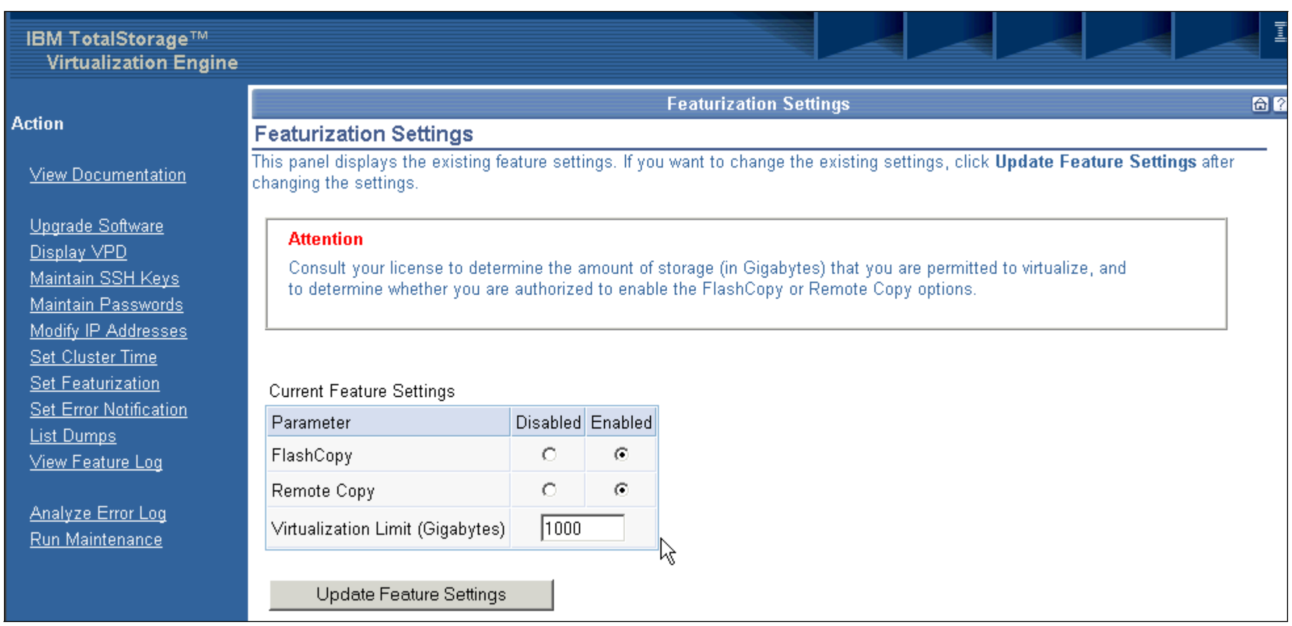


Figure 10-8 Checking that the FlashCopy feature is enabled

10.4.2 Checking the status of the FlashCopy target vDisk

Ensure that your FlashCopy target disk is not mapped to your host server.

Using the CLI

This is shown in Example 10-4.

Example 10-4 Host mapping definition for host system aix1 and aix2

```
IBM_2145:admin>svcinfo lshost
id              name            port_count
0               aix1            2
1               Win2K_1         2
2               redhat1         2
3               aix2            1

>>>>>  vDisk  aix_vDisk1_src and aix_vDisk1_tgt has been defined

IBM_2145:admin>svcinfo lsvdisk -delim /
id/name/IO_group_id/IO_group_name/status/mDisk_grp_id/mDisk_grp_name/capacity/type/FC_id/
FC_name/RC_id/RC_name
0/aix_vDisk0/0/ITSO_grp0/online/0/ARR36P5N/10.0GB/striped////
1/aix_vDisk1/0/ITSO_grp0/online/0/ARR36P5N/10.0GB/seq////
2/aix_vDisk2/0/ITSO_grp0/online/0/ARR36P5N/10.0GB/seq////
3/aix_vDisk3/0/ITSO_grp0/online/0/ARR36P5N/10.0GB/striped////
4/windisk1/0/ITSO_grp0/online/0/ARR36P5N/200.0MB/striped////
5/redhatdisk1/0/ITSO_grp0/online/0/ARR36P5N/300.0MB/striped////
6/aix_vDisk4/0/ITSO_grp0/online/0/ARR36P5N/1.0GB/striped////
7/SAP_DATA_02/0/ITSO_grp0/online/1/ARR36P6N/300.0GB/striped////
8/aix_vDisk1_src/0/ITSO_grp0/online/1/ARR36P6N/500.0MB/striped////
9/aix_vDisk1_tgt/0/ITSO_grp0/online/2/ARR73P5N/500.0MB/striped////

>>>>> vDisk aix_vDisk1_src mapped to host aix1

IBM_2145:admin>svcinfo lshostvdiskmap aix1 -delim /
id/name/SCSI_id/vDisk_id/vDisk_name/wwpn/vDisk_UID
0/aix1/5/8/aix_vDisk1_src/10000000C9295A9A/600507680182001B200000000000000C
0/aix1/5/8/aix_vDisk1_src/10000000C9266F5B/600507680182001B200000000000000C

IBM_2145:admin>svcinfo lsvdiskhostmap aix_vDisk1_src
id / name  / SCSI_id       host_id        host_name        wwpn           vDisk_UID
8 / aix_vDisk1_src / 5 / 0 / aix1 / 10000000C9295A9A /600507680182001B200000000000000C
8 / aix_vDisk1_src / 5 / 0 / aix1 / 10000000C9266F5B /600507680182001B200000000000000C

>>> no vDisk mapped to aix2

IBM_2145:admin>svcinfo lshostvdiskmap aix2
IBM_2145:admin>



>>>>>>>>>>  check that target disk aix_vDisk1_tgt is not mapped to any host

IBM_2145:admin>svcinfo lsvdiskhostmap aix_vDisk1_tgt
IBM_2145:admin>
```

Using the GUI

Using the GUI, follow these steps as shown in Figure 10-9:

1. Under My Work on the left, select **Work with Virtual Disks** and select **Hosts**.
2. In the Viewing Hosts panel on the right, under the Select column, select the host you want.
3. Select **Show the vDisks mapped to this host** from the list and click **Go**.

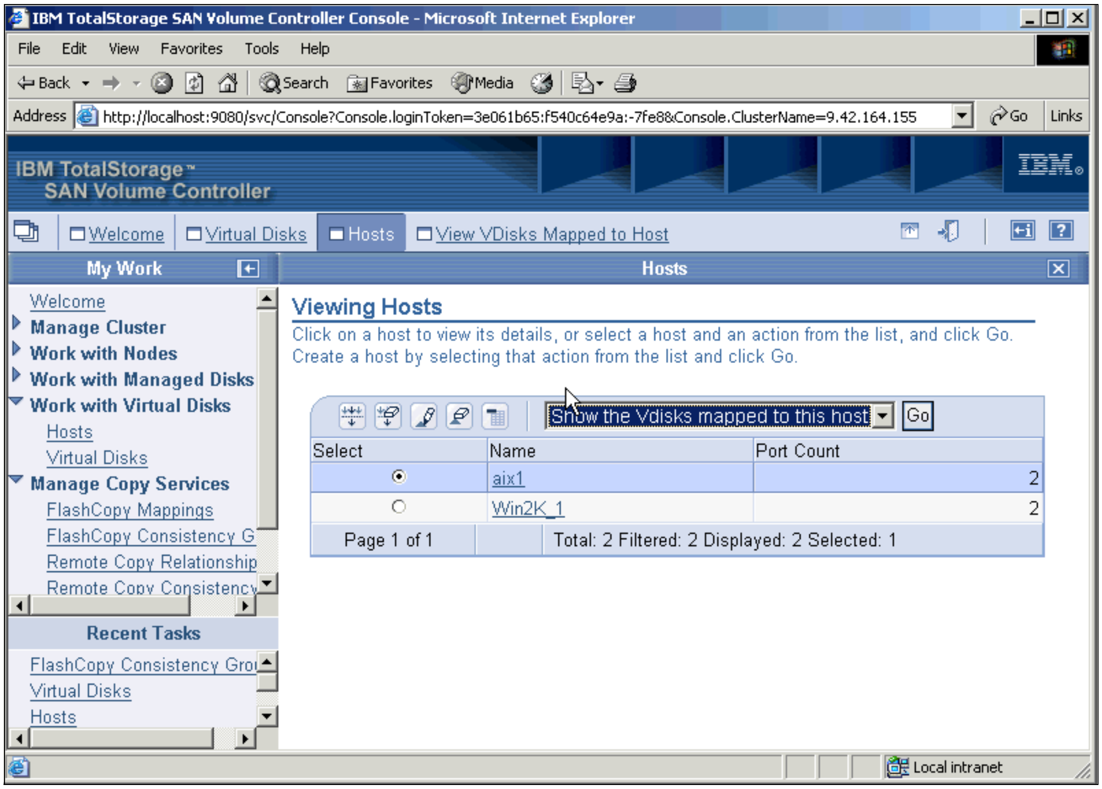


Figure 10-9 vDisk mapped to host

You can see that the vDisk aix_vDisk1_src is mapped to host system aix1 as shown in Figure 10-10.

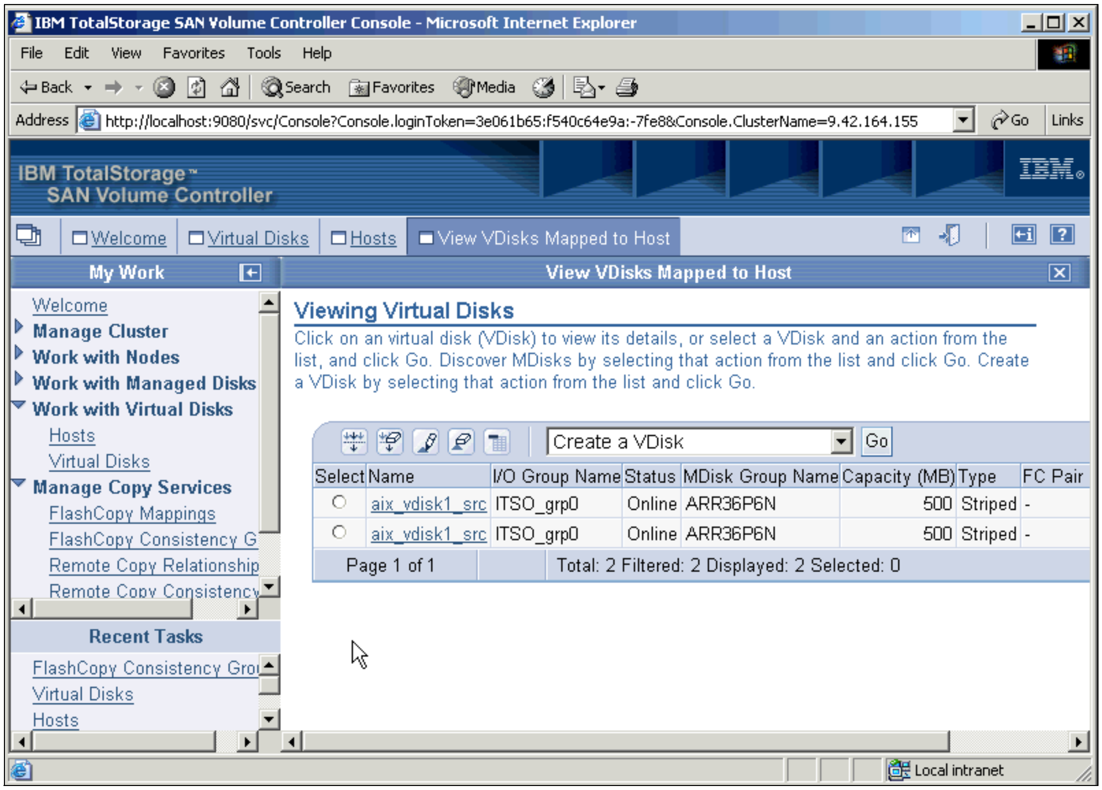


Figure 10-10 vDisk-to-host mapping

10.4.3 Creating a consistency group

As a practice, always use consistency groups for FlashCopy mappings. Although not required, if you add a second FlashCopy mapping, it is helpful to create a consistency group. If you do not use a consistency group, the FlashCopy mappings that are created are members of the pseudo consistency group zero.

Using the CLI

The command to create consistency group fccongrpaix1 is:

```
svctask mkfcconsistgrp -name fccongrpaix1
```

See Example 10-5.

Example 10-5 Creating the consistency group fccongrpaix1

```
IBM_2145:admin>svctask mkfcconsistgrp -name fccongrpaix1

IBM_2145:admin>svcinfo lsfcconsistgrp
id                name              status
1                 fccongrpaix1      idle_copied
```

Using the GUI

Complete the following steps as shown in Figure 10-11:

1. Under My Work on the left, select **Manage Copy Services** and **FlashCopy Consistency Groups**.

2. On the FlashCopy Consistency Groups panel, select **Create a consistency group** from the list and click **Go**.

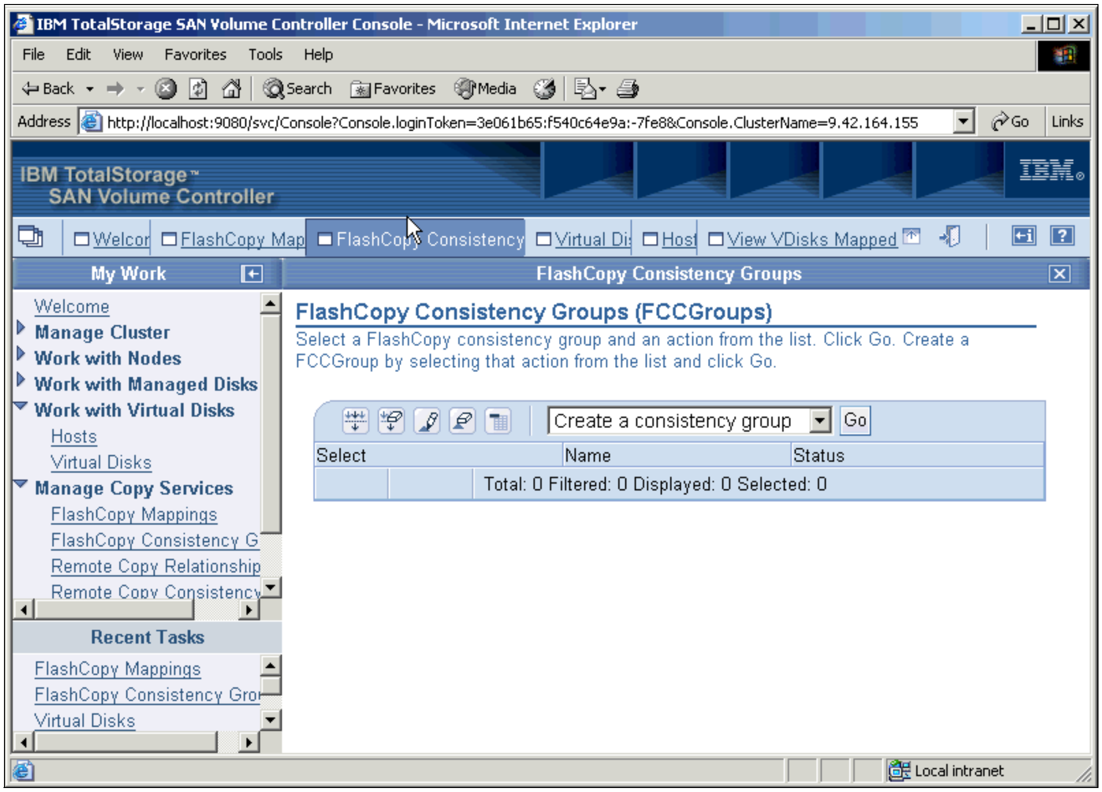


Figure 10-11 Creating a consistency group

3. On the Creating FlashCopy Consistency Groups panel (Figure 10-12), enter the name for your FlashCopy consistency group and click **OK** to complete.

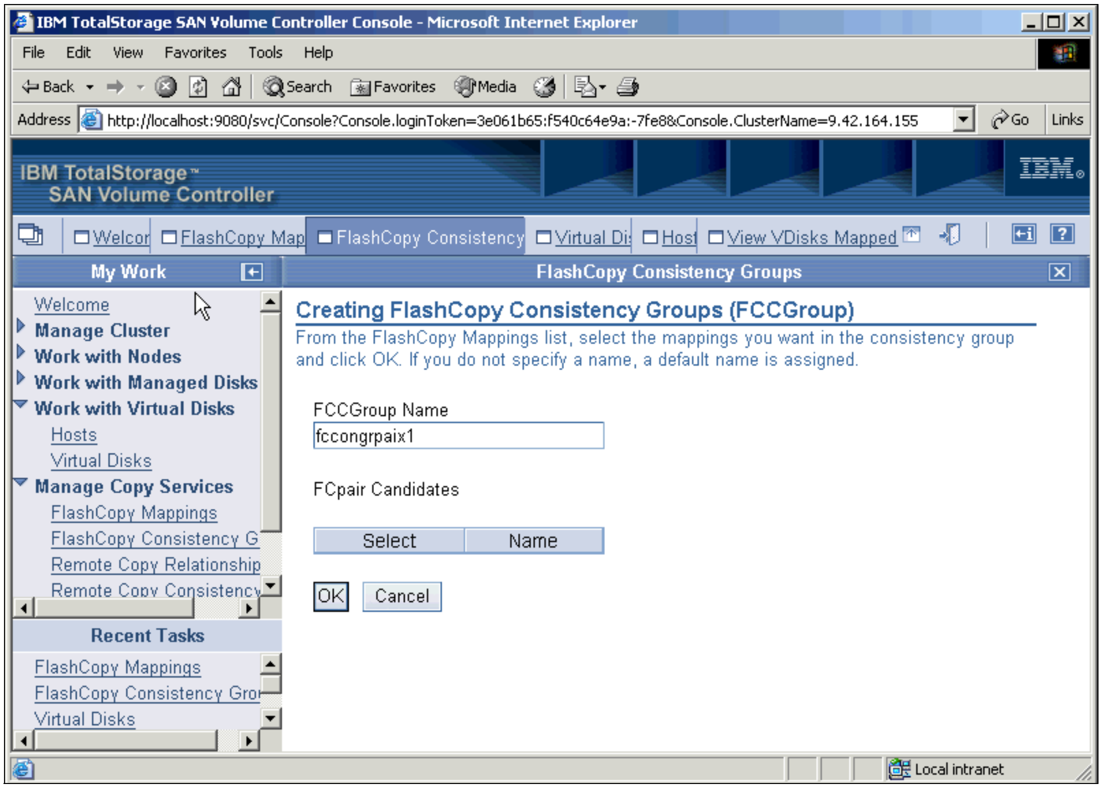


Figure 10-12 Creating FlashCopy consistency group fccongrpaix1

You can see the results in Figure 10-13.

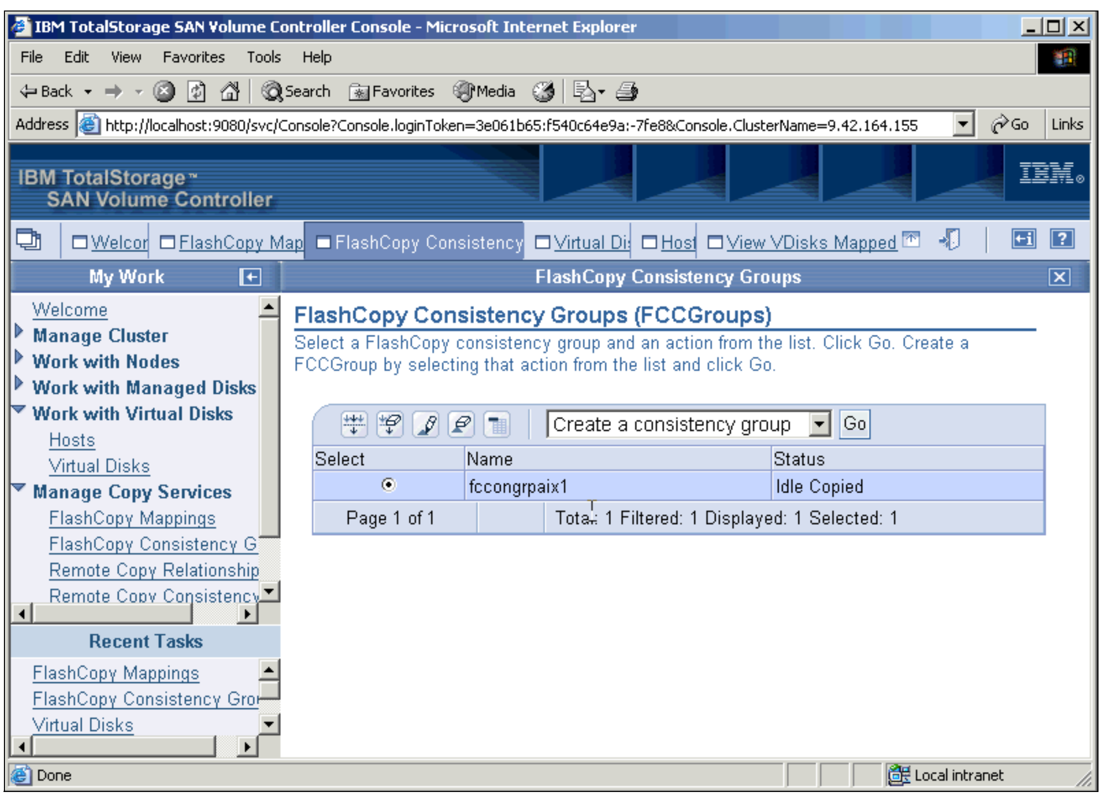


Figure 10-13 Consistency group fccongrpaix1 created

10.4.4 Creating the FlashCopy mapping

The virtual disks source and target must be the same size to create a FlashCopy mapping. The FlashCopy mapping is between aix_vDisk1_src as the source, and aix_vDisk1_tgt as the target. The background copy rate must be specified. Zero does not perform a full copy (NOCOPY). Any other rate processes the copy in the background. The default is 50% (approximately 2048 KB). A 100% value uses a rate of 64 MB and can impact other processing on the node.

Using the CLI

The command to create a FlashCopy mapping in the consistency group fccongrpaix1 between aix_vDisk1_src as the source and aix_vDisk1_tgt as the target with a background copy of 50% is:

```
svctask mkfcmap -source aix_vDisk1_src -target aix_vDisk1_tgt -name flashmapaix1
-consistgrp fccongraix1 -copyrate 50
```

See Example 10-6.

Example 10-6 Creating a FlashCopy mapping flashmapaix1

```
IBM_2145:admin>svctask mkfcmap -source aix_vDisk1_src -target aix_vDisk1_tgt -name
flashmapaix1 -consistgrp fccongraix1 -copyrate 50
Flash Copy Mapping, id [0], successfully created

>>>>>>> vDisk info have been changed for aix_vDisk1_src and aix_vDisk1_tgt
IBM_2145:admin>svcinfo lsvdisk -delim /
id/name/IO_group_id/IO_group_name/status/mDisk_grp_id/mDisk_grp_name/capacity/type/FC_id/FC
_name/RC_id/RC_name
0/aix_vDisk0/0/ITSO_grp0/online/0/ARR36P5N/10.0GB/striped////
1/aix_vDisk1/0/ITSO_grp0/online/0/ARR36P5N/10.0GB/seq////
2/aix_vDisk2/0/ITSO_grp0/online/0/ARR36P5N/10.0GB/seq////
3/aix_vDisk3/0/ITSO_grp0/online/0/ARR36P5N/10.0GB/striped////
4/windisk1/0/ITSO_grp0/online/0/ARR36P5N/200.0MB/striped////
5/redhatdisk1/0/ITSO_grp0/online/0/ARR36P5N/300.0MB/striped////
6/aix_vDisk4/0/ITSO_grp0/online/0/ARR36P5N/1.0GB/striped////
7/SAP_DATA_02/0/ITSO_grp0/online/2/ARR73P5N/5.0GB/striped////
8/aix_vDisk1_src/0/ITSO_grp0/online/1/ARR36P6N/500.0MB/striped/0/flashmapaix1//
9/aix_vDisk1_tgt/0/ITSO_grp0/online/2/ARR73P5N/500.0MB/striped/0/flashmapaix1//

>>>>>>>> fc mapping "flashmapaix1" is included in consistency group "fccongrpaix1"

IBM_2145:admin>svcinfo lsfcconsistgrp
id               name              status
1                fccongrpaix1      idle_copied

IBM_2145:admin>svcinfo lsfcconsistgrp fccongrpaix1
id 1
name fccongrpaix1
status idle_copied
FC_mapping_id 0
FC_mapping_name flashmapaix1

>>>>>>>>>> detailed information about the fc mapping flashmapaix1

IBM_2145:admin>svcinfo lsfcmap -delim /
id/name/source_vDisk_id/source_vDisk_name/target_vDisk_id/target_vDisk_name/group_id/group_
name/status/progress/copy_rate
```

```
0/flashmapaix1/8/aix_vDisk1_src/9/aix_vDisk1_tgt/1/fccongrpaix1/idle_copied//50


IBM_2145:admin>svcinfo lsfcmap flashmapaix1
id 0
name flashmapaix1
source_vDisk_id 8
source_vDisk_name aix_vDisk1_src
target_vDisk_id 9
target_vDisk_name aix_vDisk1_tgt
group_id 1
group_name fccongrpaix1
status idle_copied
progress
copy_rate 50

IBM_2145:admin>svcinfo lsfcmapprogress flashmapaix1
id              progress
0               0
```

Using the GUI

Complete the following steps as shown in Figure 10-14:

1. Under My Work, select **Manage Copy Services** and **FlashCopy Mappings**.
2. On the FlashCopy Mappings panel, select **Create a mapping** from the list and click **Go**.

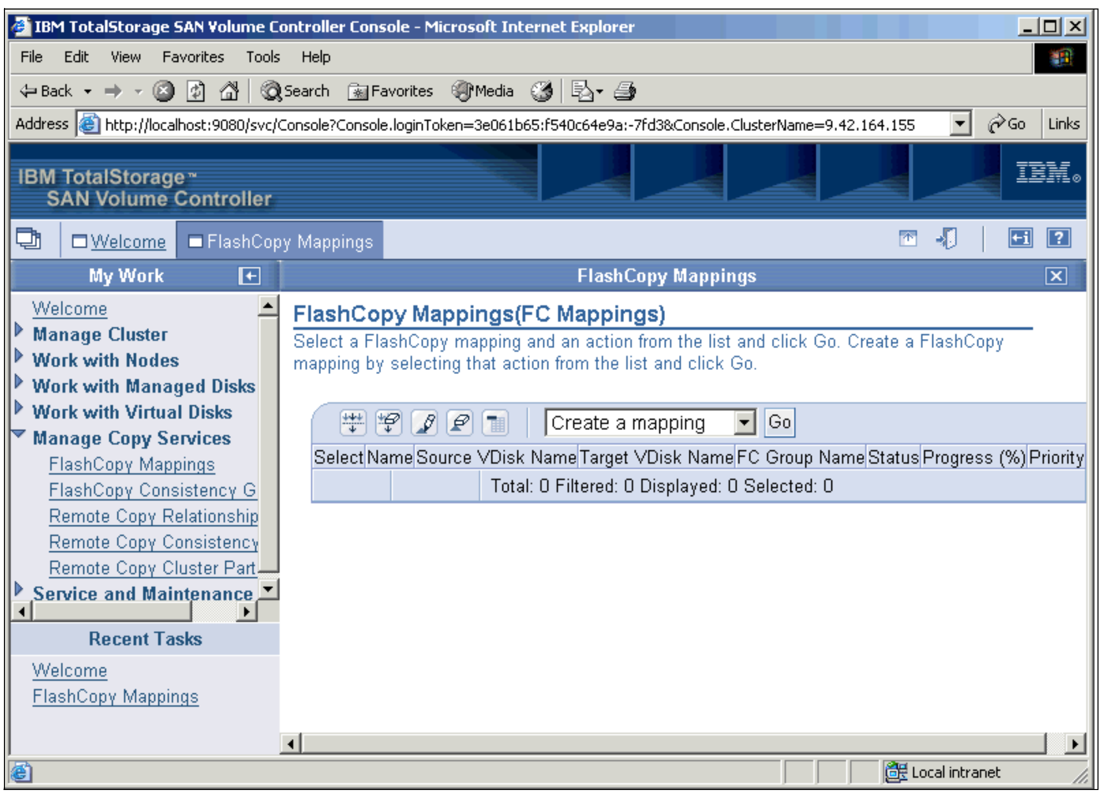


Figure 10-14 FlashCopy Mappings panel

3. As shown in Figure 10-15, follow these steps:

   a. Type the name of your FlashCopy mapping.
   b. Select the source and target virtual disks.

c. Select the FlashCopy consistency group with which this mapping is associated.
d. Select the background copy rate.
e. Click **OK**.

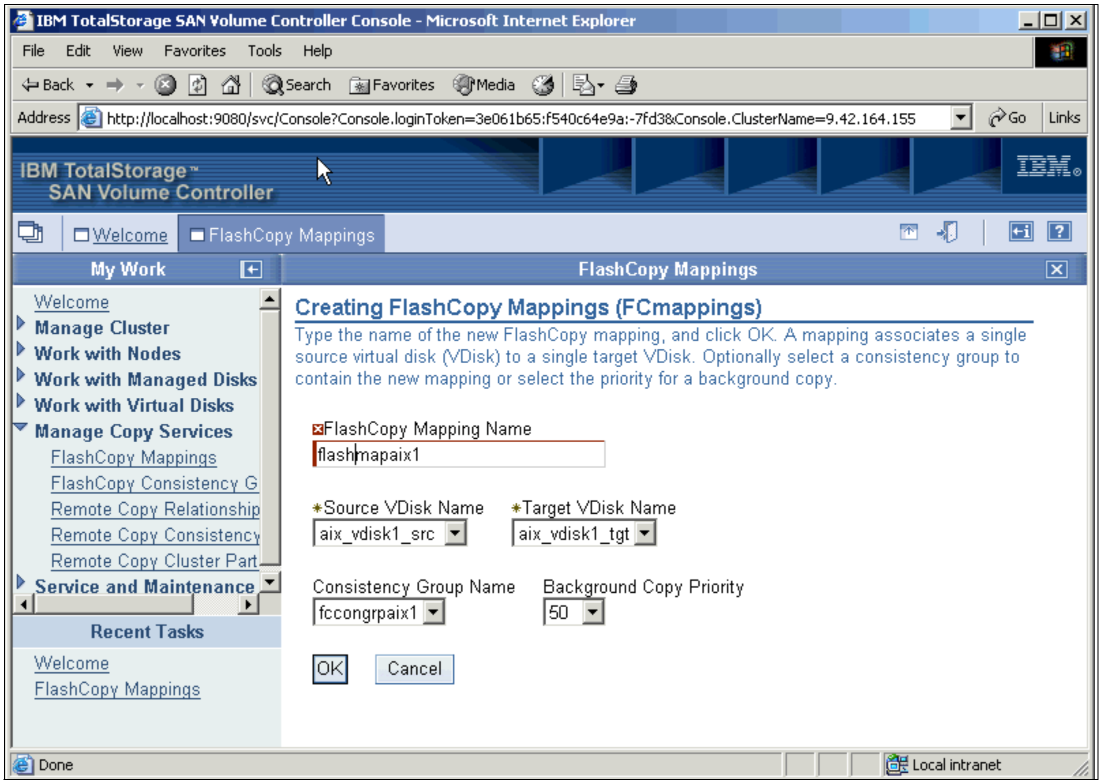


Figure 10-15 Creating mapping "flashmapaix1"

The result is shown in Figure 10-16.

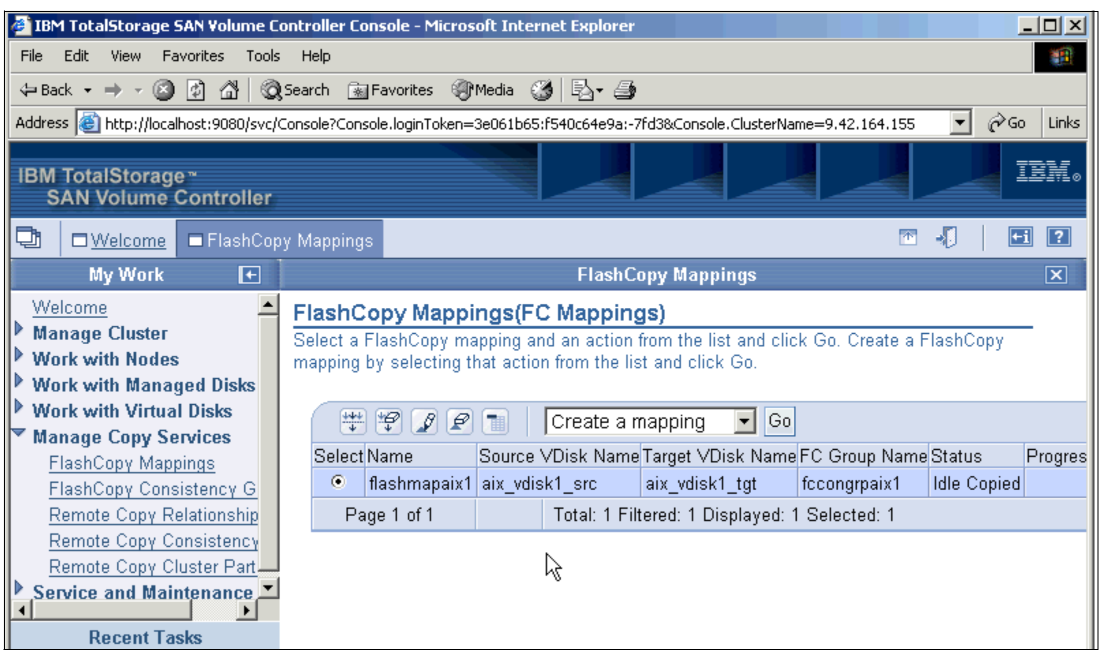


Figure 10-16 FC mapping created

10.4.5 Preparing the consistency group

Now that the FlashCopy and consistency groups are created, it is possible to start the FlashCopy operations. First, the source and target must be prepared. The FlashCopy **prepare** command flushes the cache on the source (and remains in write-through mode) and discards any data in cache destined for the target. By specifying the consistency group name, all disks in the group are prepared for FlashCopy.

> **Note:** Such operations as start, stop, and prepare can be done on a consistency group or a single mapping.

Using the CLI

The command to prepare the consistency group fccongrpaix1 is:

```
svctask prestartfcconsistgrp fccongrpaix1
```

To check the progress of the preparing, enter the following command using the CLI:

```
svcinfo lsfcconsistgrp
```

The status column says *Preparing* until it is prepared as shown in Example 10-7.

Example 10-7 Preparing the consistency group fccongrpaix1 for FlashCopy

```
IBM_2145:admin>svctask prestartfcconsistgrp fccongrpaix1

>>>>>>>>> Consistency group “fccongrpaix1” is prepared and ready for FlashCopy

IBM_2145:admin>svcinfo lsfcconsistgrp
id              name              status
1               fccongrpaix1      prepared

IBM_2145:admin>svcinfo lsfcconsistgrp fccongrpaix1
id 1
name fccongrpaix1
status prepared
FC_mapping_id 0
FC_mapping_name flashmapaix1

IBM_2145:admin>svcinfo lsfcmap -delim /
id/name/source_vDisk_id/source_vDisk_name/target_vDisk_id/target_vDisk_name/group_id/group_
name/status/progress/copy_rate
0/flashmapaix1/8/aix_vDisk1_src/9/aix_vDisk1_tgt/1/fccongrpaix1/prepared/0/50

IBM_2145:admin>svcinfo lsfcmap flashmapaix1
id 0
name flashmapaix1
source_vDisk_id 8
source_vDisk_name aix_vDisk1_src
target_vDisk_id 9
target_vDisk_name aix_vDisk1_tgt
group_id 1
group_name fccongrpaix1
status prepared
progress 0
copy_rate 50
```

Using the GUI

Complete the following steps as shown in Figure 10-17:

1. Under My Work, select **Manage Copy Services** and **FlashCopy Consistency Groups**.
2. On the FlashCopy Consistency Groups panel, under Select, select the consistency group you want to prepare.
3. Select **Prepare a consistency group** from the list and click **Go**.

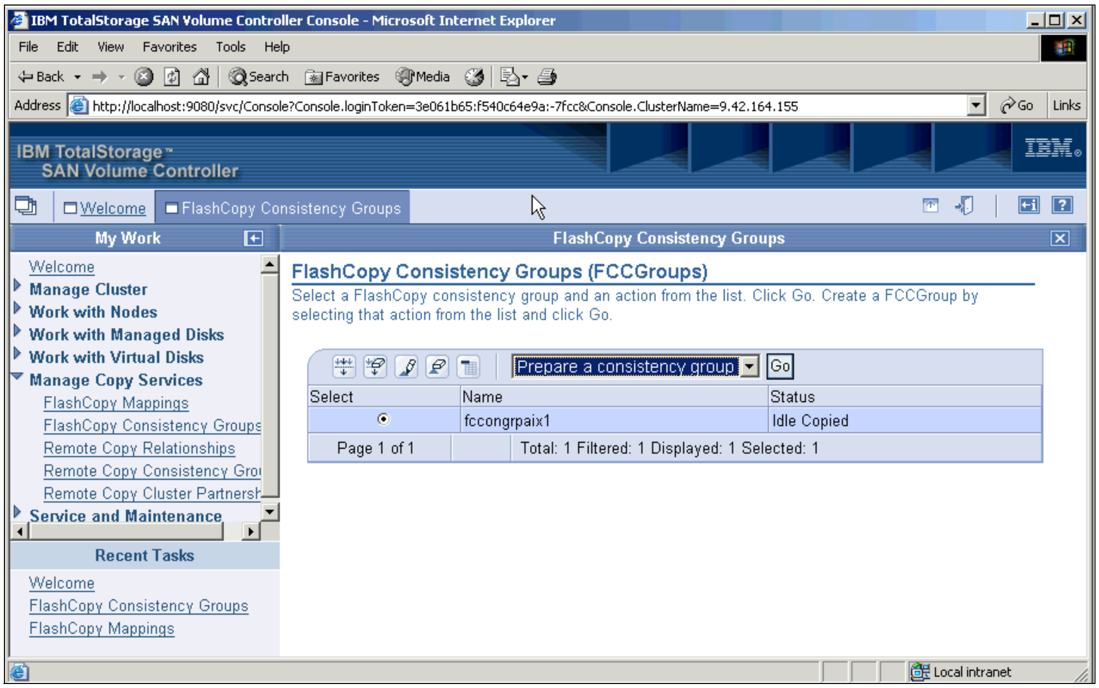


Figure 10-17 Preparing a consistency group

The status column now says *Preparing*. See Figure 10-18.

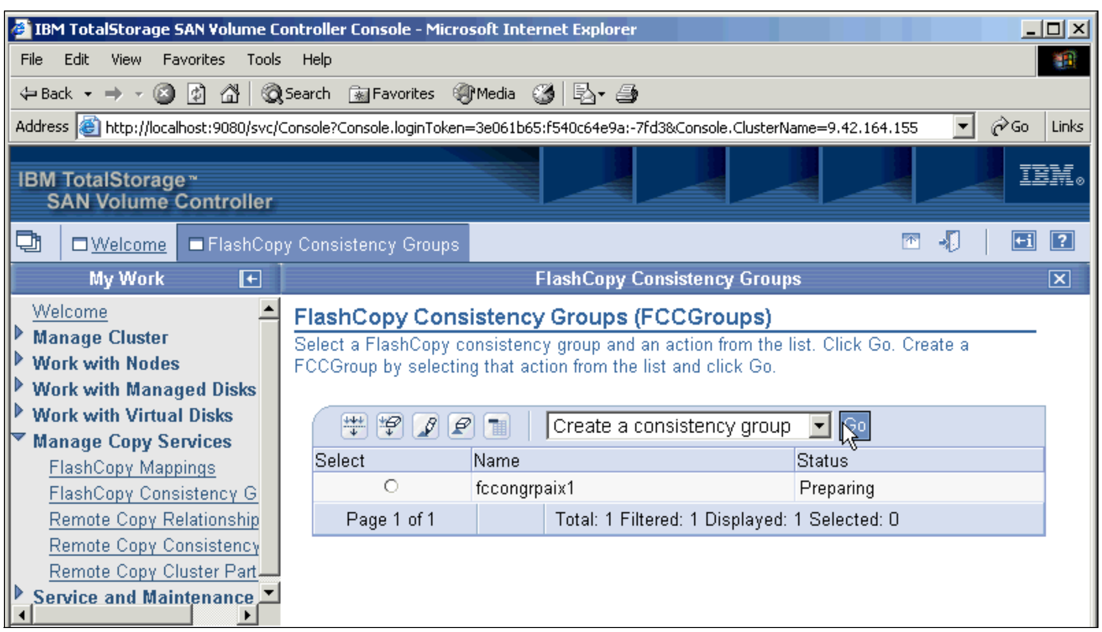


Figure 10-18 Preparing state for the consistency group FC

The status column says *Prepared* when the preparation is done. See Figure 10-19.

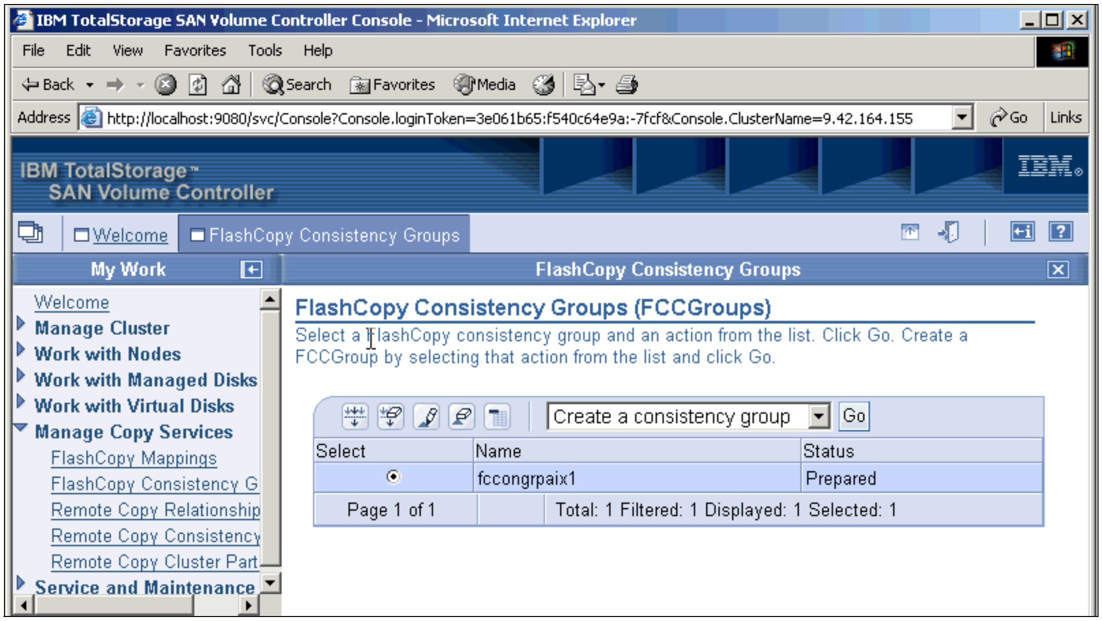


Figure 10-19 Consistency group fccongrpaix1 prepared

10.4.6 Starting the FlashCopy consistency group

After the FlashCopy consistency group is prepared, it is ready to be started.

Using the CLI

The command to start the FlashCopy consistency group fccongrpaix1 is:

```
svctask startfcconsistgrp fccongrpaix1
```

The command to monitor the progress of the FlashCopy mapping flashmapaix1 is:

```
svcinfo lsfcmapprogress flashmapaix1 and svcinfo lsfcmap flashmapaix1
```

The status starts from copying and progress 0 to complete with status idle_copied and progress 100. See Example 10-6.

Example 10-8 Starting and monitoring a FlashCopy consistency group

```
IBM_2145:admin>svctask startfcconsistgrp fccongrpaix1

IBM_2145:admin>svcinfo lsfcmapprogress flashmapaix1
id                progress
0                 0

IBM_2145:admin>svcinfo lsfcmap flashmapaix1
id 0
name flashmapaix1
source_vDisk_id 8
source_vDisk_name aix_vDisk1_src
target_vDisk_id 9
target_vDisk_name aix_vDisk1_tgt
group_id 1
group_name fccongrpaix1
```

```
status copying
progress 0
copy_rate 50

>>>>>>>> 24 percent of progress for the FlashCopy task

IBM_2145:admin>svcinfo lsfcmapprogress flashmapaix1
id              progress
0               24

IBM_2145:admin>svcinfo lsfcmap flashmapaix1
id 0
name flashmapaix1
source_vDisk_id 8
source_vDisk_name aix_vDisk1_src
target_vDisk_id 9
target_vDisk_name aix_vDisk1_tgt
group_id 1
group_name fccongrpaix1
status copying
progress 24
copy_rate 50

IBM_2145:admin>svcinfo lsfcmap -delim /
id/name/source_vDisk_id/source_vDisk_name/target_vDisk_id/target_vDisk_name/group_id/group_
name/status/progress/copy_rate
0/flashmapaix1/8/aix_vDisk1_src/9/aix_vDisk1_tgt/1/fccongrpaix1/copying/48/50


IBM_2145:admin>svcinfo lsfcconsistgrp
id              name                status
1               fccongrpaix1        copying

IBM_2145:admin>svcinfo lsfcconsistgrp fccongrpaix1
id 1
name fccongrpaix1
status copying
FC_mapping_id 0
FC_mapping_name flashmapaix1




>>>>> FlashCopy is complete

IBM_2145:admin>svcinfo lsfcmapprogress flashmapaix1
id              progress
0               100

IBM_2145:admin>svcinfo lsfcconsistgrp
id              name                status
1               fccongrpaix1        idle_copied

IBM_2145:admin>svcinfo lsfcconsistgrp fccongrpaix1
id 1
name fccongrpaix1
status idle_copied
FC_mapping_id 0
FC_mapping_name flashmapaix1
```

```
IBM_2145:admin>svcinfo lsfcmap -delim /
id/name/source_vDisk_id/source_vDisk_name/target_vDisk_id/target_vDisk_name/group_id/group_
name/status/progress/copy_rate
0/flashmapaix1/8/aix_vDisk1_src/9/aix_vDisk1_tgt/1/fccongrpaix1/idle_copied/100/50

IBM_2145:admin>svcinfo lsfcmap flashmapaix1
id 0
name flashmapaix1
source_vDisk_id 8
source_vDisk_name aix_vDisk1_src
target_vDisk_id 9
target_vDisk_name aix_vDisk1_tgt
group_id 1
group_name fccongrpaix1
status idle_copied
progress 100
copy_rate 50
```

Using the GUI

Complete the following steps as shown in Figure 10-20:

1. Under My Work, select **Manage Copy Services and FlashCopy Consistency Groups**.
2. On the FlashCopy Consistency Groups panel, under the Select column, select the **FlashCopy consistency group** that you want to start.
3. Select **Start a consistency group** from the list and click **Go**.

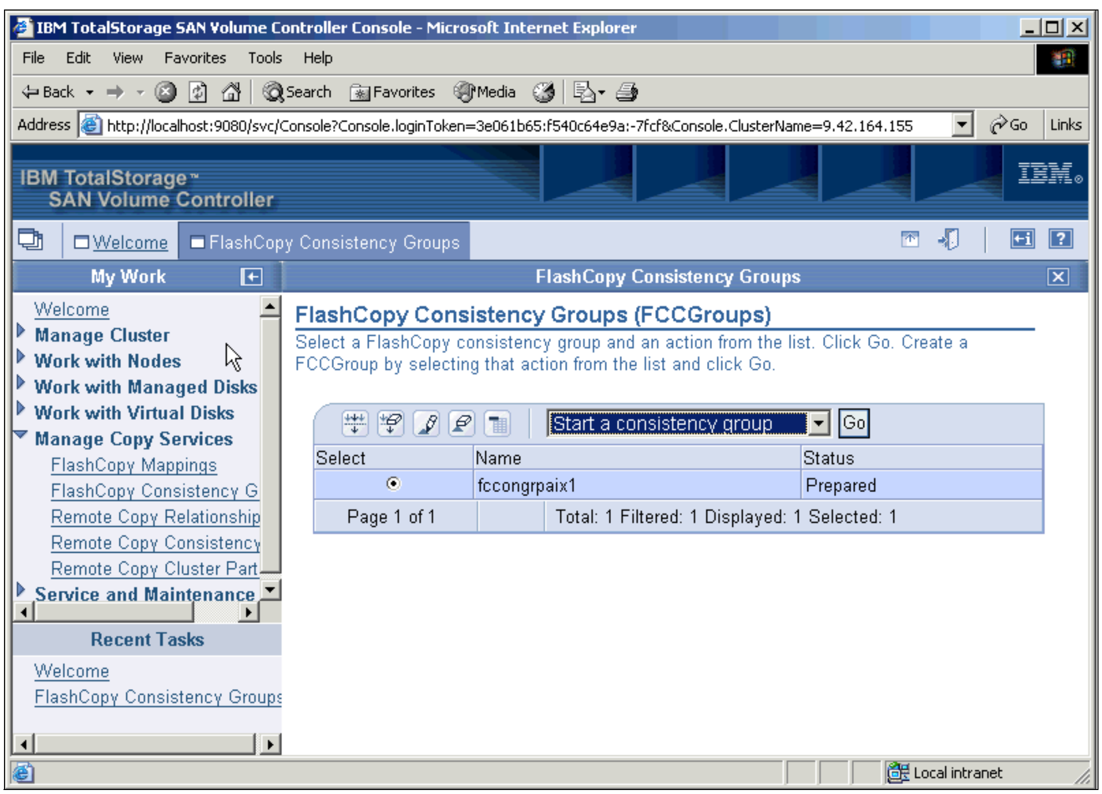


Figure 10-20 Starting FlashCopy consistency group fccongrpaix1

4. After confirming the FlashCopy consistency group creation, the GUI prompts "`Before FlashCopy started, cache will be flushed of any data destined for the source`

`virtual disk`". Click **OK** to complete creation of FlashCopy consistency group. The status column now says *Copying*.

You can see the status of the FlashCopy task with the windows shown in Figure 10-21, Figure 10-22, and Figure 10-23.

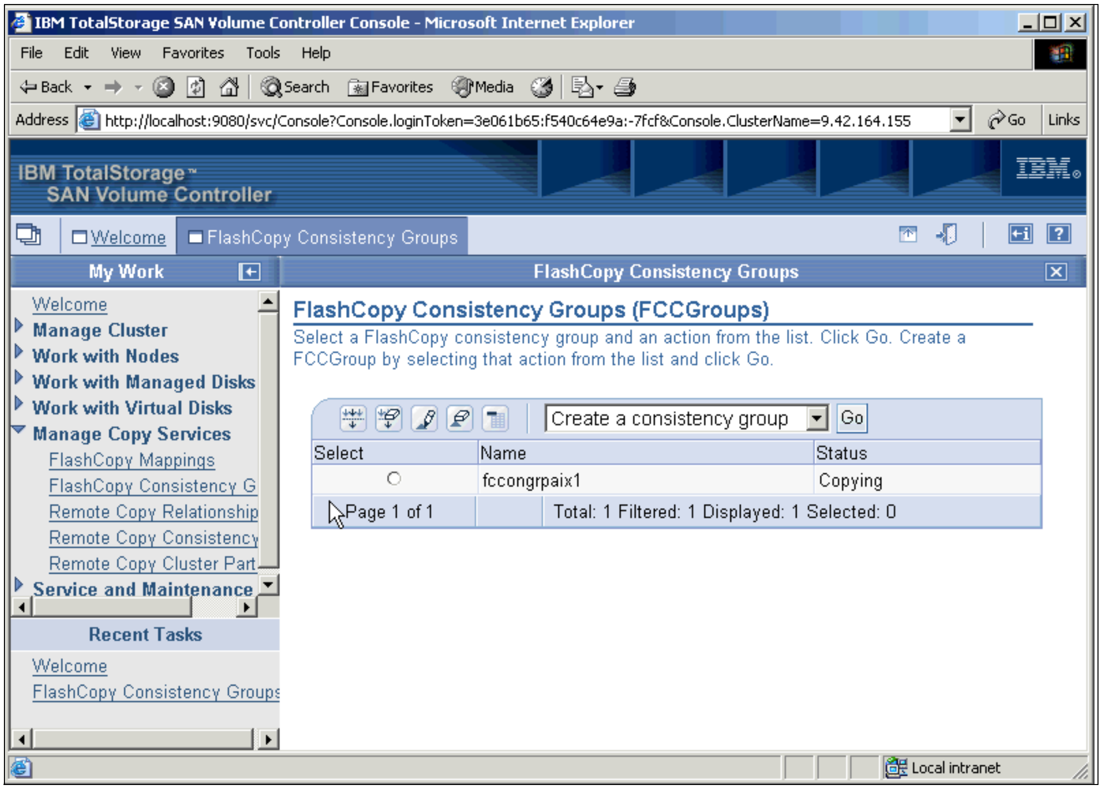


Figure 10-21 FlashCopy consistency group ccongrpaix1 in state Copying

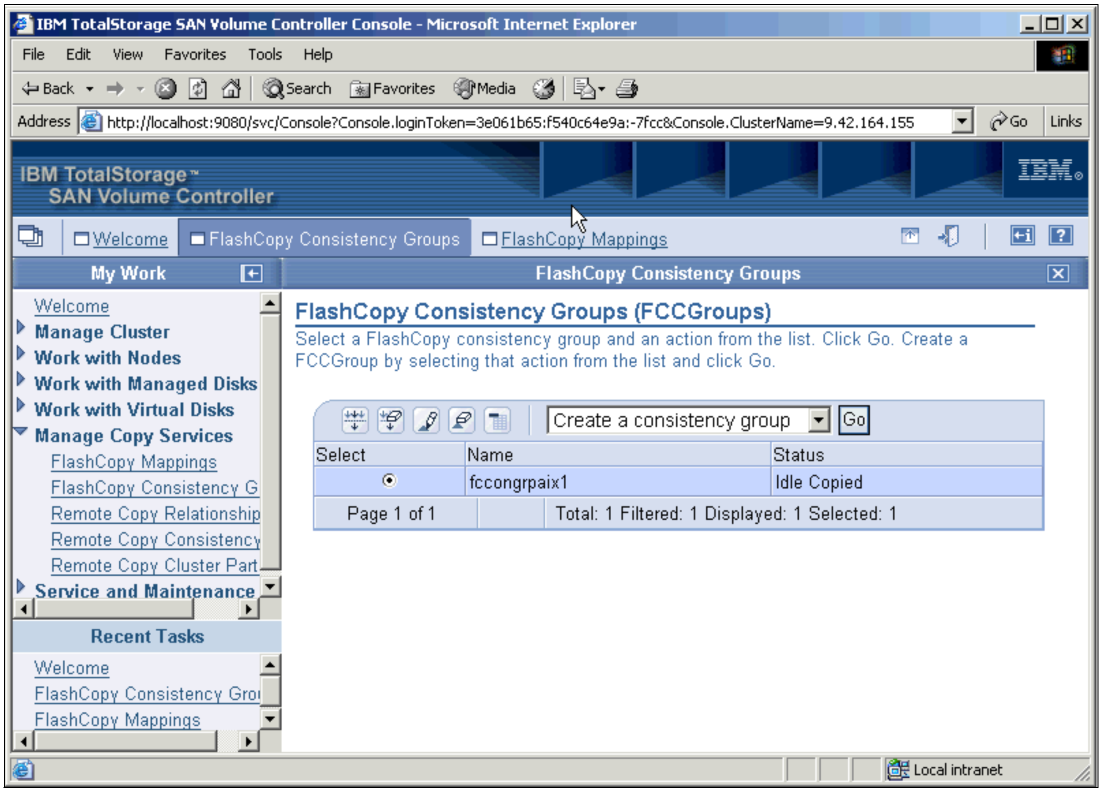


Figure 10-22 FlashCopy consistency group fccongrpaix1 in state Idle Copied

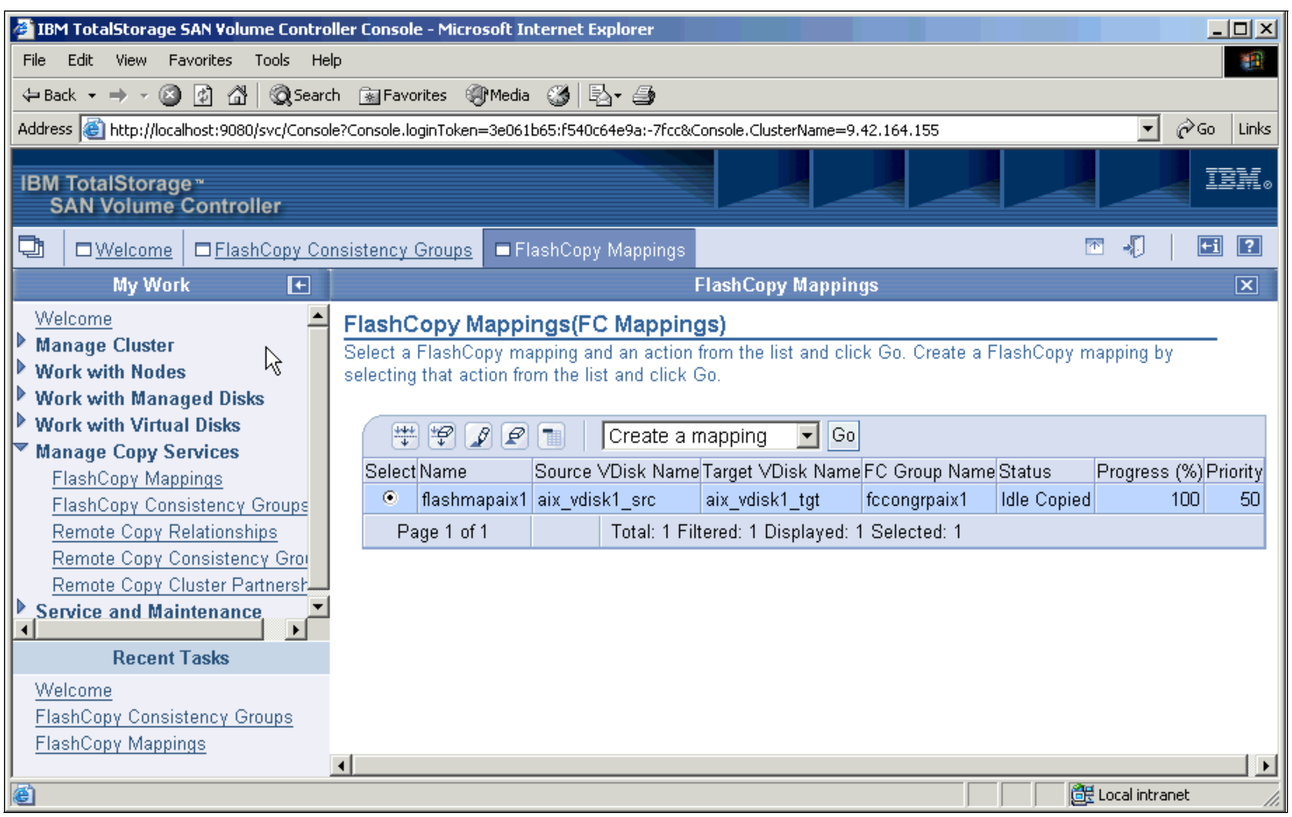


Figure 10-23 FlashCopy mapping flashmapaix1 in state Idle Copied with progress 100%

10.4.7 Using FlashCopy target vDisk on the host system aix2

After the FlashCopy is completed, map the target virtual disk to the host. See Example 10-9.

Example 10-9 Host system aix2 storage configuration after FlashCopy

```
AIX2#cfgmgr

AIX2#lspv
hdisk0         000b21dd2cf9b83d    rootvg
hdisk1         00079058479c31c5    None
hdisk2         00079058479c31c5    None
vpath0         none                None


AIX2#lsdev -Cc disk
hdisk0 Available 10-80-00-2,0 16 Bit SCSI Disk Drive
hdisk1 Available 10-90-01     IBM TotalStorage SAN Volume Controller device
hdisk2 Available 10-90-01     IBM TotalStorage SAN Volume Controller device
vpath0 Available              Data Path Optimizer Pseudo Device Driver

AIX2#datapath query device
Total Devices : 1

DEV#:   0  DEVICE NAME: vpath0  TYPE: 2145           POLICY:    Optimized
SERIAL: 600507680182001B200000000000000D
==================================================================
Path#      Adapter/Hard Disk          State     Mode     Select     Errors
    0          fscsi0/hdisk1          CLOSE   NORMAL          0          0
    1          fscsi0/hdisk2          CLOSE   NORMAL          0          0

AIX2#datapath query adapter
Active Adapters :1
Adpt#     Adapter Name   State     Mode     Select     Errors  Paths  Active
    0           fscsi0  NORMAL   ACTIVE          0          0      2       0

AIX2#lsvpcfg
vpath0 (Avail ) 600507680182001B200000000000000D = hdisk1 (Avail pv ) hdisk2 (Avail pv )

AIX2#importvg -y itsofcaix2newvg hdisk1
itsofcaix2newvg

AIX2#lspv
hdisk0         000b21dd2cf9b83d    rootvg
hdisk1         00079058479c31c5    itsofcaix2newvg
hdisk2         00079058479c31c5    itsofcaix2newvg
vpath0         none                None

AIX2#hd2vp itsofcaix2newvg
itsofcaix2newvg
itsofcaix2newvg contains these hdisks and pvids
   00079058479c31c5 hdisk1
hdisk1 is part of vpath0.
itsofcaix2newvg
Deleting pvid of hdisk1 from ODM database
Create vpath0 pvid of 00079058479c31c5 in ODM database.
lname=vpath0, pvid=00079058479c31c5
Deleting pvid of hdisk2 from ODM database
varyon itsofcaix2newvg was successful.
```

```
itsofcaix2newvg is converted to dpo device vpathes successfully!

AIX2#lspv
hdisk0        000b21dd2cf9b83d    rootvg
hdisk1        none                None
hdisk2        none                None
vpath0        00079058479c31c5    itsofcaix2newvg

AIX2#datapath query adapter
Active Adapters :1
Adpt#     Adapter Name   State     Mode     Select     Errors  Paths  Active
    0           fscsi0  NORMAL   ACTIVE         51          0      2       2

AIX2#datapath query device
Total Devices : 1
DEV#:   0  DEVICE NAME: vpath0  TYPE: 2145           POLICY:    Optimized
SERIAL: 600507680182001B200000000000000D
==========================================================================
Path#      Adapter/Hard Disk         State     Mode     Select     Errors
    0          fscsi0/hdisk1          OPEN   NORMAL          0          0
    1          fscsi0/hdisk2          OPEN   NORMAL         51          0


AIX2#lsvpcfg
vpath0 (Avail pv itsofcaix2newvg) 600507680182001B200000000000000D = hdisk1 (Avail ) hdisk2
(Avail )

AIX2#varyonvg itsofcaix2newvg

AIX2#lsvg -o
itsofcaix2newvg
rootvg

AIX2#lsvg itsofcaix2newvg
VOLUME GROUP:   itsofcaix2newvg         VG IDENTIFIER:  0007905800004c00000000f5479f2450
VG STATE:       active                  PP SIZE:        4 megabyte(s)
VG PERMISSION:  read/write              TOTAL PPs:      124 (496 megabytes)
MAX LVs:        256                     FREE PPs:       113 (452 megabytes)
LVs:            2                       USED PPs:       11 (44 megabytes)
OPEN LVs:       0                       QUORUM:         2
TOTAL PVs:      1                       VG DESCRIPTORS: 2
STALE PVs:      0                       STALE PPs:      0
ACTIVE PVs:     1                       AUTO ON:        yes
MAX PPs per PV: 1016                    MAX PVs:        32

IX2#fsck - /itsofcaix1fs
** Checking /dev/ritsofcaix1lv (/itsof)
** Phase 0 - Check Log
log redo processing for /dev/ritsofcaix1lv
** Phase 1 - Check Blocks and Sizes
** Phase 2 - Check Pathnames
** Phase 3 - Check Connectivity
** Phase 4 - Check Reference Counts
** Phase 5 - Check Inode Map
** Phase 6 - Check Block Map
9 files 2656 blocks 79264 free

AIX2#mount /itsofcaix1fs

AIX2#df -k
```

```
Filesystem    1024-blocks      Free %Used    Iused %Iused Mounted on
/dev/hd4            32768     26388   20%     1179     8% /
/dev/hd2           393216     79132   80%    14361    15% /usr
/dev/hd9var         32768     30288    8%      153     2% /var
/dev/hd3            32768     31052    6%       60     1% /tmp
/dev/hd1            32768     31684    4%       18     1% /home
/dev/itsofcaix1lv       40960     39632    4%       18     1% /itsofcaix1fs

AIX2#cd /itsofcaix1fs
AIX2#ls
file1.txt   lost+found
AIX2#more file1.txt
this file was created on AIX1

AIX2#ls
file1.txt                 file_new_after_flask.txt  lost+found
AIX2#more file_new_after_flask.txt
this file was created on AIX2 after FlashCopy
```

The storage configuration on host aix1 is shown in Example 10-10.

Example 10-10 Host system aix1 storage configuration after FlashCopy

```
>>>>>>>>  FlashCopy done

AIX1#fsck -y /itsofcaix1fs
** Checking /dev/ritsofcaix1lv (/itsof)
** Phase 0 - Check Log
log redo processing for /dev/ritsofcaix1lv
** Phase 1 - Check Blocks and Sizes
** Phase 2 - Check Pathnames
** Phase 3 - Check Connectivity
** Phase 4 - Check Reference Counts
** Phase 5 - Check Inode Map
** Phase 6 - Check Block Map
9 files 2656 blocks 79264 free

AIX1#mount /itsofcaix1fs

AIX1#df -k
Filesystem    1024-blocks      Free %Used    Iused %Iused Mounted on
/dev/hd4            16384      7936   52%     1281    16% /
/dev/hd2          1867776   1265812   33%    20504     5% /usr
/dev/hd9var         16384     10936   34%      433    11% /var
/dev/hd3            32768     29188   11%       58     1% /tmp
/dev/hd1            16384     15820    4%       18     1% /home
/proc                   -         -    -         -     -  /proc
/dev/hd10opt        32768     26592   19%      294     4% /opt
/dev/itsofcaix1lv       40960     39632    4%       18     1% /itsofcaix1fs

AIX1# more /itsofcaix1fs/file1.txt
this file was created on AIX1

AIX1#ls
file1.txt                 file_new_after_flask.txt  lost+found
AIX1#more file_new_after_flask.txt
this file was created on AIX1 after FlashCopy
```

11

Copy Services: Peer-to-Peer Remote Copy

This chapter describes the Peer-to-Peer Remote Copy (PPRC) function of Copy Services. PPRC in an IBM TotalStorage SAN Volume Controller (SVC) is similar in function to PPRC in the IBM TotalStorage Enterprise Storage Server (ESS) and IBM TotalStorage Fibre Array Storage Technology (FAStT) subsystems. This means that the SVC provides a single point of control to provide remote copy in your storage network.

11.1 Peer-to-Peer Remote Copy

This section looks at the principles of PPRC.

11.1.1 How it works

The general application of Peer-to-Peer Remote Copy seeks to maintain two copies of a data set. Often the two copies are separated by some distance, hence the term *remote*, but this is not required.

The SVC assumes that the Fibre Channel (FC) fabric to which it is attached contains hardware, which achieves the long distance requirement for the application. This hardware makes storage “at a distance” accessible as though it were local storage. Specifically, it enables two SVC clusters to connect (FC login) to each other and establishes communications in the same way as though they were located nearby on the same fabric. The only difference is in the expected latency of that communication, the bandwidth capability of the link, and the availability of the link as compared with the local fabric.

Refer to Chapter 3, “Planning and configuration” on page 25, to learn how a FC fabric must be configured to support SVC Peer-to-Peer Remote Copy and the distance limits on this support.

The relationship between the two copies is not symmetric. One copy of the data set is considered the primary copy (sometimes also known as the *source copy*). This copy provides the reference for normal run-time operation. Updates to this copy are shadowed to a secondary copy (sometimes known as the *destination copy* or *target copy*). The secondary copy is not normally referenced for performing I/O.

If the primary copy fails, the secondary copy can be enabled for I/O operation. A typical use of this function may involve two sites where the first provides service during normal running and the second is only activated when a failure of the first site is detected.

The secondary copy is not accessible for application I/O other than the I/Os that are performed for the remote copy process itself. SVC allows read-only access to the secondary storage when it contains a “consistent” image. This is only intended to allow boot time operating system discovery to complete without error, so that any hosts at the secondary site can be ready to start up the applications with minimum delay if required. For instance, many operating systems need to read LBA 0 to configure a logical unit. Although read access is allowed at the secondary in practice, the data on the secondary volumes cannot be read by a host. The reason for this is that most operating systems write a “dirty bit” to the file system when it is mounted. Because this write operation is not allowed on the secondary volume, the volume cannot be mounted.

This access is only provided where consistency can be guaranteed. However, there is no way in which coherency can be maintained between reads performed at the secondary and later write I/Os performed at the primary.

Enabling the secondary copy for active operation requires some SVC, operating system, and possibly application specific work. This needs to be performed as part of the entire failover process. The SVC software at the secondary must be instructed to stop the relationship which has the affect of making the secondary logical unit accessible for normal I/O access. The operating system may need to mount file systems, or similar work, which can typically only happen when the logical unit is accessible for writes. The application may have some log of work to recover.

Note that this property of Peer-to-Peer Remote Copy, the requirement to “enable” the secondary copy, differentiates it from RAID 1 mirroring. The latter aims to emulate a single, reliable disk regardless of what system is accessing it. Peer-to-Peer Remote Copy retains the property that there are two volumes in existence but suppresses one while the copy is being maintained.

Using a secondary copy involves some conscious policy decision by a user that a failover is required. The application work involved in establishing operation on the secondary copy is substantial. The goal is to make this rapid (much faster compared to recovering from a backup copy) but not seamless. Most clients aim to automate this through failover management software. SVC provides Simple Network Management Protocol (SNMP) traps and interfaces to enable this automation.

The underlying storage at the primary or secondary of a remote copy is normally RAID storage, but may be any storage which can be managed by SVC.

Asynchronous versus synchronous remote copy

This section looks at the differences between synchronous and asynchronous remote copy. The remote copy can be maintained in one of two modes, synchronous or asynchronous. The definition of an asynchronous remote copy needs to be supplemented with some measure describing the maximum degree of asynchronicity.

Synchronous remote copy ensures that updates are committed at both primary and secondary before the application is given completion to an update. This ensures that the Secondary is fully up-to-date should it be needed in a failover. However, this means that the application is fully exposed to the latency and bandwidth limitations of the communication link to the secondary. Where this is truly remote, this can have a significant adverse effect on application performance.

In asynchronous remote copy, the application is given completion to an update before that update is necessarily committed at the secondary. Therefore, on a failover, some updates may be missing at the secondary. The application must have some external mechanism for recovering the missing updates and reapplying them. This mechanism may involve user intervention.

Asynchronous remote copy provides comparable functionality to a continuous backup process that is missing the last few updates. Recovery on the secondary site involves bringing up the application on the recent backup and then re-applying the most recent updates to bring the secondary up to date.

The asynchronous remote copy must present a secondary which is intelligible to the application. This means it cannot apply updates in an arbitrary order as may be permitted at the primary side. The reason is that, at the primary side, the application is enforcing an ordering implicitly by not scheduling an I/O until a previous dependent I/O has completed. We cannot hope to know the actual ordering constraints of the application.

The best that can be achieved is to choose an ordering which the application may see if I/O at the primary were stopped at a suitable point. One such example is to apply I/Os at the secondary in the order they were completed at the primary. Thus the secondary always reflects a state that could have been seen at the primary if we froze I/O there.

Note that a total ordering must be maintained across all data to which an application is sensitive. This total ordering may involve a single point of control assigning an order to an appropriate series of events at the primary. In applying global coordination to the asynchronous remote copy, it is important to maintain the performance of the system.

Future releases of SVC should incorporate asynchronous remote copy.

The importance of write ordering

Many applications that use block storage have a requirement to survive failures such as loss of power or software crash and not lose data that existed prior to the failure. Many such applications need to make non-atomic updates to that storage. An example is an update that involves more than one write operation.

For instance, consider the case of an update to a database involving a funds transfer from Account A to Account B. Two writes are required, one to each of the disk regions that contain the balances Account A and B. If these two writes were simply issued in parallel and power failed part-way through, then a number of outcomes may result:

- The accounts would be unchanged.
- The money might be deducted from A but not appear in B.
- The money might have appeared in B but not deducted from A.
- The transfer might be complete.

A certain degree of uncertainty is inevitable, since for many error scenarios, it is not possible to define at the point at which the error occurred whether writes that did not receive completion happened.

However, after the error is repaired it is important that any such uncertainty is removed and the system recover to a well-defined state. In the previous example, it is important to ensure that either the transfer did happen completely, or that it did not. If it did not, then a higher level recovery is required (for example, the user can re-request it). Any alternative states (such as the middle two half-complete cases) must be eliminated.

Many complex algorithms are used to eliminate such uncertainty, but the principles of operation are common. The following is a simple example of those principles, although real systems employ much more elaborate schemes.

A common name for this technique is *logging* since the technique involves a log file (or journal log file). The log contains a record of recent transactions that have been started, but not finished. It is used to keep track of all updates such that they can be recovered in the event of a disruption. The log must contain enough information to recover the transaction.

The sequence of operation may now become:

1. An entry is made in the log that a transfer is being performed. The entry includes the details of the Accounts (A and B) being updated and the final balance each will have when the transfer is complete. The entry also includes a unique number which distinguishes this log entry from all others. This write must complete before the next step.
2. Updates to each of the accounts are issued. These writes happen as above. *Both* of these writes must complete before the next step.
3. A second entry is made in the log stating that the transfer is complete. This references the first entry's unique ID.

Now what is the impact of a failure such that the application stops receiving completions at a certain point? The application must read the log and find any entries which do not have corresponding completion entries:

- If the first log entry is not visible, then the update did not happen. Both accounts remain at their initial balances. The transfer must be re-attempted if it is still needed.
- If both log entries are visible, then all the updates happened. Both accounts are at the desired and final balances with the transfer complete.

- If the first log entry is visible, but not the second, then the application simply restarts the sequence at Step 2. The writes to update the account must be reissued. Both of these must complete before the final completion log entry is made.

This simple application is tolerant of failures and interruptions. It uses two principles to ensure its operation:

- If it receives successful write completion for one I/O, it sees the data on a later read.
- It can withhold certain writes until completion is received for a preceding write and that way ensure that, when one write happens, other data is certain to have been written and is certain to appear on the disk.

Consider what happens if the writes appear in the wrong order and then power is lost:

- If the account update entries happen before the first log entry write occurs: A power failure before the first log entry means that there is no trace of information with which to perform a recovery. The two database entries may or may not occur successfully. The accounts may end up in any of the states identified in the first (unlogged) scenario.
- If the second log entry happens before the account update entries: A power failure immediately after the second log entry means the transaction is considered complete and the two account update entries may never be updated.

This scenario demonstrates that maintaining write ordering is key to ensuring the correct operation of applications following a disruption.

In this example, it is possible to locate the two account entries and the log on three separate disks, accessed through three different storage controllers. Nevertheless, the application is dependent on write ordering across all these systems to maintain correct operation following a power failure.

11.1.2 SVC PPRC functions

The SVC PPRC supports the following capabilities:

- SVC supports synchronous remote copy.
- SVC implements remote copy between vDisk pairs, each vDisk in the pair being managed by a SVC cluster.
- SVC supports intracluster remote copy, where both vDisks belong to the same cluster.
- SVC supports intercluster remote copy, where one vDisks comes from each of two clusters. A given SVC cluster can be configured to recognize a single other cluster. A cluster can communicate with only one other cluster. All intercluster remote copy takes place between those two clusters.
- Intercluster and intracluster remote copy can be used concurrently within a cluster for different relationships.
- SVC does *not* require a control network or fabric to be installed to manage remote copy. For intercluster remote copy SVC maintains a control link between the two clusters. This control link is used to control state and co-ordinate updates at either end. The control link is implemented on top of the same FC fabric connection as SVC uses for I/O.
- SVC implements a configuration model which maintains the remote copy configuration and state through major events such as failover, recovery, and resynchronization to minimize user configuration action through these events.
- SVC maintains and polices a strong concept of consistency and makes this available to guide configuration activity.

- SVC implements flexible resynchronization support enabling it to re-synchronize vDisk pairs which have suffered write I/O to both disks and to resynchronize only those regions which are known to have changed.

How PPRC works

There are several steps in the Remote Copy process:

1. A partnership is created between two SVC clusters.
2. A consistency group is created if you need multiple relationships to be treated as one entity.
3. A relationship is created between two vDisk of the same size.
4. The relationship is started, and synchronized.
5. Once synchronized, the relationship is ready for production data.
6. To access the auxiliary vDisk, the relationship is stopped and made inconsistent.
7. The remote host server is mapped to the auxiliary vDisk and the disk is available for I/O.

Intercluster communication

For intercluster remote copy, you must create a partnership between the two clusters. The clusters must have storage area network (SAN) interswitch links (ISLs) between them. The partnership must be created on both clusters. Figure 11-1 illustrates this concept.

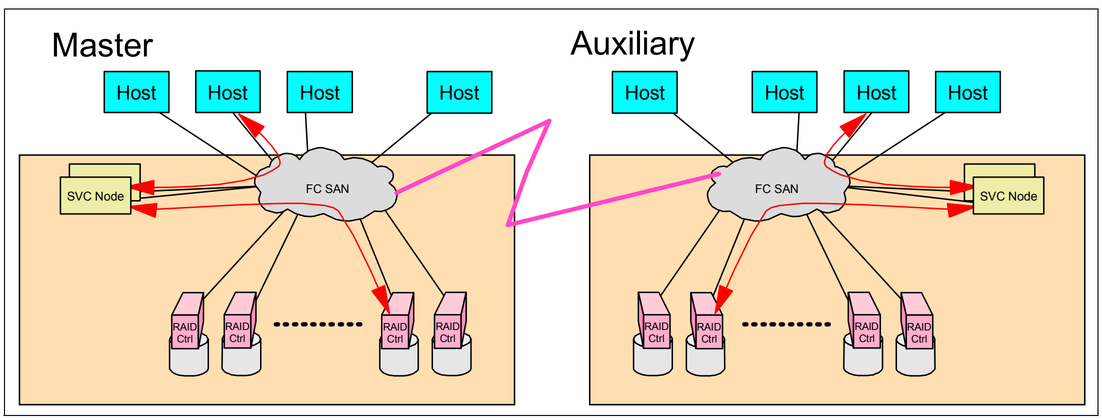


Figure 11-1 PPRC cluster partnership

Each SVC cluster can be configured to recognize a single other cluster as its partner (see 11.3.3, “Creating a PPRC intercluster partnership” on page 361). When both clusters are correctly configured and can reach each other over the fabric, they establish further communication facilities between the nodes in each of the cluster. This comprises:

- A single control channel which is used to exchange and coordinate configuration information
- I/O channels between each of the nodes in the clusters

These channels are maintained and updated as nodes appear and disappear and as links fail, and are repaired to maintain operation where possible. When communication between the clusters is lost an error log is created. This can be configured to lead to SNMP traps and drive the repair procedures.

Intercluster zoning

Figure 11-2 shows the schematic diagram of simple zoning required for PPRC. A PPRC zone can be thought as separate zone for copy service purposes.

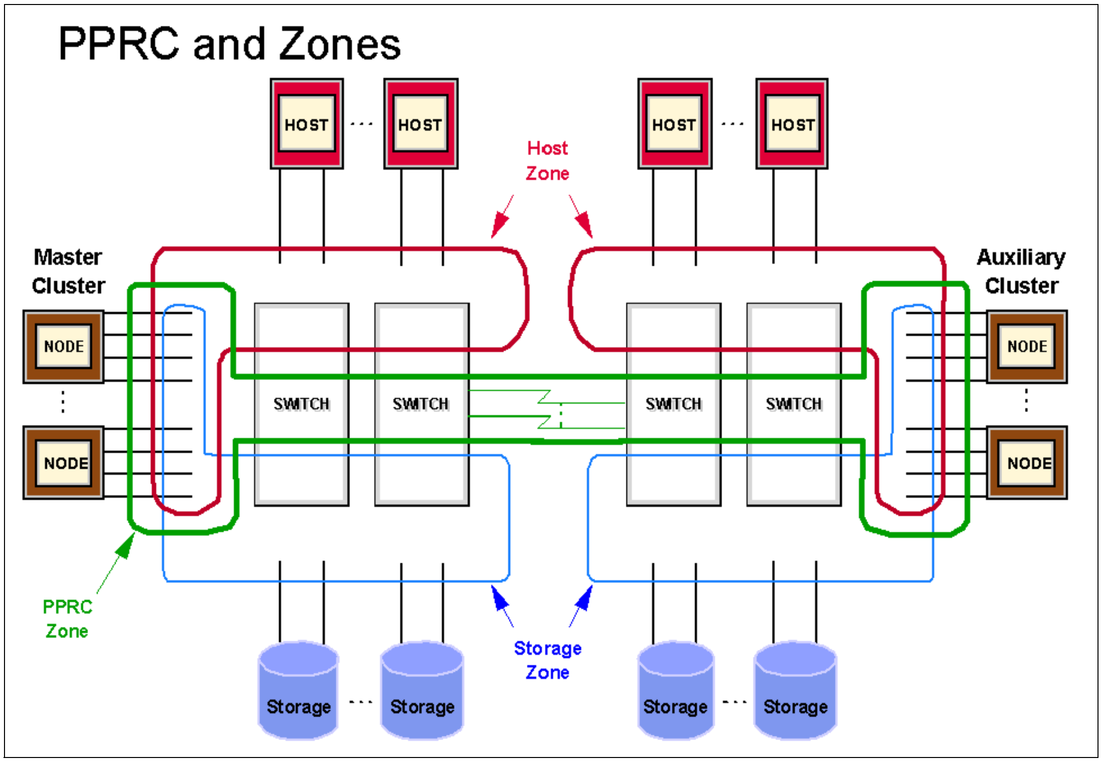


Figure 11-2 Schematic diagram of PPRC cluster zoning

Maintenance of the intercluster link

Every SVC node maintains a database of the other devices that are visible on the fabric. This is updated as devices appear and disappear. Devices that advertise themselves as SVC nodes are categorized according to the cluster to which they belong. Nodes that belong to the same cluster establish communication channels between themselves and begin to exchange messages to implement the clustering and functional protocols of SVC. Nodes that are in different clusters do not exchange messages after the initial discovery is complete unless they have been configured together to perform remote copy.

The intercluster link carries the control traffic to coordinate activity between the two clusters. It is formed between one node in each cluster which is termed the *focal point*. One of the FC logins between the two focal points is selected to carry all the communications traffic.

If this connection should fail because of a problem with the specific path chosen, a new login between the same two focal points is selected.

If the focal point node should fail (or all its logins to the remote cluster fail), then a new focal point is chosen to carry the control traffic. Changing the focal point causes I/O to pause but does not cause relationships to become ConsistentStopped.

SVC configuration model

There are several relationships to take into account.

PPRC relationship

Remote copy relationships are similar to FlashCopy mappings. They can be stand-alone or combined in consistency groups. Start and stop commands can be issued either against the stand-alone relationship or the consistency group. Figure 11-3 illustrates the PPRC relationship.

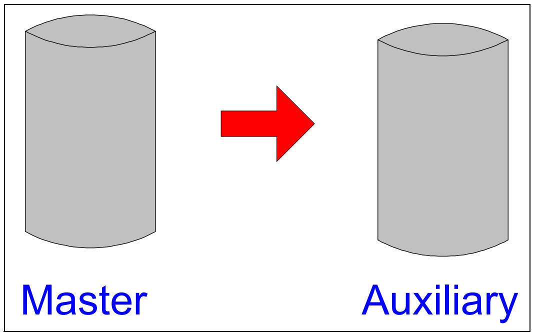


Figure 11-3 PPRC relationships

A remote copy relationship is composed of two virtual disks, a *master* virtual disk and an *auxiliary* virtual disk.

In the most common use of Remote Copy, the master virtual disk contains a production copy of the data and is the one “normally” used by an application. The auxiliary virtual disk contains a backup copy of the data and is used in disaster recovery scenarios. The terms master and auxiliary help support this use. In other uses, the terms master and auxiliary need to be interpreted appropriately.

The two virtual disks in a remote copy relationship must be of the same size.

The two virtual disks can both be in the same cluster in which case they must be in the same I/O group. Otherwise, the two virtual disks can come from each of the two clusters that have been configured to recognize each other and can be in any I/O group in each of those two clusters.

Each virtual disk in a remote copy relationship takes on a specific role, behaving as a primary or a secondary. A *primary virtual disk* contains a valid copy of application data and is accessible for application write I/O. A *secondary virtual disk* may contain a valid copy of application data, but is not available for application write I/O.

The two virtual disks in a remote copy relationship either both assume the primary role, or one is primary and one is secondary, depending on the state of the relationship. In the second case, with one primary and one secondary, Remote Copy makes the data on the secondary virtual disk match the data on the primary virtual disk by performing writes to the secondary virtual disk.

When a remote copy relationship is first created, the master virtual disk is always assigned the role of primary and the auxiliary virtual disk is assigned the role of secondary. Hence, the copying direction is from master to auxiliary.

Later, after an appropriate sequence of events and commands, the roles may be reversed and the copying direction is from auxiliary to master. The direction of copying is reflected in the state of the system.

In certain states, Remote Copy performs a background copy from primary to secondary. The rate of this copy is influenced by a CopyPriority parameter.

Consistency groups

Consistency groups allow you to combine similar tasks and complete them as one complete entity. Relationships can be combined into a consistency group. Figure 11-4 illustrates the concept of consistency groups.

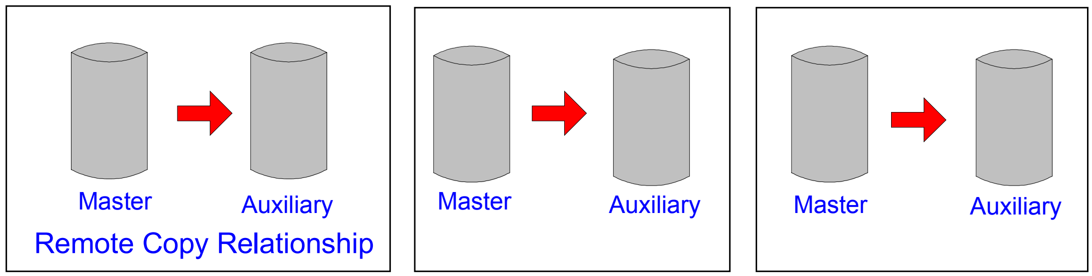


Figure 11-4 Remote copy consistency groups

Certain uses of Remote Copy require the manipulation of more than one relationship. Remote Copy provides the ability to group relationships, so that they are manipulated in unison.

For some uses, it may be that the relationships share some loose association and that the grouping simply provides a convenience for the administrator. But a more significant use arises where the relationships contain virtual disks that have a tighter association. One example is where an application's data is spread across more than one virtual disk. A more complex example is where multiple applications run on different host systems, where each application has data on different virtual disks, and these applications exchange data with each other. Both examples are cases where specific rules exist as to how the relationships must be manipulated, in unison, to ensure that the set of secondary virtual disks contain usable data. The key property is that these relationships be consistent. Hence the groups are called consistency groups:

- A relationship can be part of a single consistency group or not be part of a consistency group at all.
- Relationships which are not part of a consistency group are called *stand-alone relationships*.
- A consistency group can contain zero or more relationships. An *empty consistency group*, with zero relationships in it, has little purpose until it is assigned its first relationship except that it has a name.
- All the relationships in a consistency group must having matching master and auxiliary clusters.

Although it is possible that consistency groups can be used to manipulate sets of relationships that do not need to satisfy these strict rules, that manipulation can lead to some undesired side effects. The rules behind consistency mean that certain configuration commands are prohibited where this would not be the case if the relationship was not part of a consistency group.

For example, consider the case of two applications that are completely independent, yet they are placed into a single consistency group. In the event of an error, there is a loss of synchronization, and a background copy process is required to recover synchronization. While this process is in progress, Remote Copy rejects attempts to enable access to the Secondary vDisks of either application. If one application finishes its background copy much more quickly than the other, Remote Copy still refuses to grant access to its secondary, even though this is safe in this case, because the Remote Copy policy is to refuse access to the entire consistency group if any part of it is inconsistent.

Stand-alone relationships and consistency groups share a common configuration and state model. All the relationships in a non-empty consistency group have the same state as the consistency group.

State overview

A number of concepts of state are key to understanding the configuration concepts and are introduced here.

Connected versus disconnected

This distinction can arise when a Remote Copy relationship is created with the two virtual disks in different clusters. Under certain error scenarios, communications between the two clusters may be lost. For instance, power may fail causing one complete cluster to disappear. Alternatively, the fabric connection between the two clusters might fail, leaving the two clusters running but unable to communicate with each other.

When the two clusters can communicate, the clusters and the relationships spanning them are described as connected. When they cannot communicate, the clusters and the relationships spanning them are described as disconnected.

In this scenario, each cluster is left with half the relationship and has only a portion of the information that was available to it before. Some limited configuration activity is possible, a subset of what was possible before. The disconnected relationships are portrayed as having changed state. The new states describe what is known about the relationship and what configuration commands are permitted.

When the clusters can communicate again the relationships become connected once again. Remote Copy automatically reconciles the two state fragments, taking into account any configuration or other event that took place while the relationship was disconnected. As a result, the relationship can either return to the state it was in when it became disconnected or it can enter a different connected state.

Relationships that are configured between virtual disks in the *same* cluster will never be described as being in a disconnected state.

Consistent versus inconsistent

Relationships that contain vDisks operating as secondaries can be described as being consistent or inconsistent. Consistency groups that contain relationships can also be described as being consistent or inconsistent. The consistent or inconsistent property describes the relationship of the data on the secondary to that on the primary virtual disk. It can be considered a property of the secondary vDisk itself.

A secondary is described as consistent if it contains data that could have been read by a host system from the primary if power had failed at some imaginary point in time while I/O was in progress and power was later restored. This definition assumes that the primary maintains write ordering and therefore requires that the secondary maintains write ordering also.

Requirements of this definition are:

- The secondary vDisk contains the data from all writes for which the host has received good completion and that data has not been overwritten by a subsequent write.
- For writes for which the host did *not* receive good completion (that is it received bad completion or no completion at all) and the host subsequently performed a read from the primary of that data and that read returned good completion and no later write was sent, the secondary contains the same data as that returned by the read from the primary.

Consistent does not mean that the data is up-to-date. A copy can contain data that is frozen at some point in time in the past.

From the point of view of an application, consistency means that a secondary vDisk contains the same data as the primary vDisk at the point at which the imaginary power failure occurred.

If an application is designed to cope with unexpected power failure this guarantee of consistency means that the application can use the secondary and begin operation just as though it was restarted after the hypothetical power failure.

Again, the application is dependent on the key properties of consistency:

- Write ordering
- Read stability for correct operation at the secondary

If a relationship, or set of relationships, is inconsistent and an attempt is made to start an application using the data in the secondaries a number of outcomes are possible:

- The application may decide that the data is corrupt and crash or exit with an error code.
- The application may fail to detect the data is corrupt and return erroneous data.
- The application may work without a problem.

Because of the risk of data corruption, in particular undetected data corruption, Remote Copy strongly enforces the concept of consistency and prohibits access to inconsistent data. Consistency as a concept can be applied to a single relationship or a set of relationships in a consistency group. Write ordering is a concept that an application can maintain across a number of disks accessed through multiple systems and therefore consistency must operate across all those disks.

When deciding how to use consistency groups, the administrator must consider the scope of an application's data, taking into account all the interdependent systems which communicate and exchange information. If two programs or systems communicate and store details as a result of the information exchanged, then either of the following actions may occur:

- All the data accessed by the group of systems must be placed into a single consistency group.
- The systems must be recovered independently (each within its own consistency group). Then each system must perform recovery with the other applications to become consistent with them.

Consistent verses synchronized

A copy which is consistent *and* up-to-date is described as synchronized. In a synchronized relationship, the primary and secondary virtual disks are only different in regions where writes are outstanding from the host.

Consistency does *not* mean that the data is up-to-date. A copy can be consistent and yet contain data which was frozen at some point in time in the past. Write I/O may have continued to a primary and not have been copied to the secondary. This state arises when it becomes impossible to keep up-to-date and maintain consistency. An example is a loss of communication between clusters when writing to the secondary.

When communication is lost for an extended period of time, Remote Copy tracks the changes that happen at the primary but not the order of such changes nor the details of such changes (write data). When communication is restored, it is impossible to make the secondary synchronized without sending write data to the secondary out-of-order and therefore losing consistency.

Two policies can be used to cope with this:

- Take a point-in-time copy of the consistent secondary before allowing the secondary to become inconsistent. In the event of a disaster before consistency is achieved again, the point-in-time copy target provides a consistent, though out-of-date, image.
- Accept the loss of consistency, and loss of useful secondary, while making it synchronized.

11.1.3 PPRC configuration limits

The limits in Table 11-1 apply to configuration of Remote Copy.

Table 11-1 Peer-to-Peer Remote Copy configuration limits

Parameter	Value
Number of consistency groups	32
Number of relationships	1024 (512 for intracluster PPRC)
Total virtual disk size per I/O group	32 TB (up to 128 TB per cluster)
Maximum virtual disk size	2 TB

11.2 PPRC commands

For all the details about the Peer-to-Peer Remote Copy Commands, see *IBM TotalStorage Virtualization Family SAN Volume Controller: Command-Line Interface User's Guide*, SC26-7544.

The command set for Peer-to-Peer Remote Copy contains two broad groups:

- Commands to create, delete and manipulate relationships and consistency groups
- Commands to cause state changes

Where a configuration command affects more than one cluster, Remote Copy performs the work to coordinate configuration activity between the clusters. Some configuration commands can only be performed when the clusters are connected and fail with no effect when they are disconnected.

Other configuration commands are permitted even though the clusters are disconnected. The state is reconciled automatically by Remote Copy when the clusters become connected once more.

For any given command, with one exception, a single cluster actually receives the command from the administrator. This is significant for defining the context for a CreateRelationShip **mkrcrelationship** or CreateConsistencyGroup **mkrcconsistgrp** command in which case, the cluster receiving the command is called the *local cluster*.

This exception mentioned previously is the command that sets clusters into a Remote Copy partnership. The **mkpartnership** command must be issued to both the local and to the remote cluster.

The commands here are described as an abstract command set. These are implemented as:

- A command line interface (CLI) which can be used for scripting and automation
- A graphical user interface (GUI) which can be used for one-off tasks

11.2.1 Commands to manipulate relationships and consistency groups

Peer-to-Peer Remote Copy does not use the same concept as FlashCopy with respect to consistency group 0. With Remote Copy, relationships that do not belong to a consistency group are known as *stand-alone relationships* and have a null consistency group attribute. Consequently, consistency group 0 is a normal, valid consistency group

Show the candidate list virtual disk (svcinfo lsrcrelationshipcandidate)

Remote Copy provides a candidate list of virtual disks that are eligible to be made part of a Remote Copy. The candidate list provides a list of virtual disks that are of the correct size and are not disallowed by some other configuration state, such as being a FlashCopy target.

Create a PPRC relationship (svctask mkrcrelationship)

This command is used to create relationships with vDisks in the same cluster or in two different clusters. To create a relationship in two different clusters, the clusters must be connected at the time the create command is received.

- Master and auxiliary vDisks must be of the same size.
- They must not currently be part of a remote copy relationship.
- They must not be a target of a FlashCopy mapping
- The relationship name must be unique. For intracluster relationships, it must be unique within that cluster.
- For intercluster relationships, it must be unique across all relationship names in either of the two clusters involved.

Modify a PPRC relationship (svctask chrcrelationship)

This command can modify the specified attributes of the relationship supplied. You can:

- Change the name of a consistency group.
- Add a relationship to a group.
- Remove a relationship from a group.
- Move a relationship from one group to another.

Delete PPRC relationship (svctask rmrcrelationship)

This command is used to delete the relationship that is specified. Deleting a relationship only deletes the logical relationship between the two virtual disks. It does not affect the virtual disks themselves.

If the relationship is disconnected at the time that the command is issued, then the relationship is only deleted on the cluster on which the command is being run. When the clusters reconnect, then the relationship is automatically deleted on the other cluster. Alternatively, if the clusters are disconnected, and you still wish to remove the relationship on both clusters, you can issue the rmrcrelationship command independently on both of the clusters.

A relationship cannot be deleted if it is part of a consistency group. You must first remove the relationship from the consistency group.

If you delete an inconsistent relationship, the secondary virtual disk becomes accessible even though it is still inconsistent. This is the one case in which Remote Copy does not inhibit access to inconsistent data.

Create a PPRC consistency group (svctask mkrcconsistgrp)

This command creates a new consistency group. The name must be unique across all consistency groups known to the clusters owning this consistency group. If the consistency group involves two clusters, the clusters must be in communication throughout the create process.

The new consistency group does not contains any relationships and is in the empty state. Remote copy relationships can be added to the group using the **svctask mkrcrelationship** command.

Modify a PPRC consistency group (svctask chrcconsistgrp)

You can modify the name of an existing consistency group, remove a relationship, and add a stand-alone relationship.

Delete a PPRC consistency group (svctask rmrcconsistgrp)

This is used to delete a consistency group. This command deletes the specified consistency group. You can issue this command for any existing consistency group. If the consistency group is disconnected at the time that the command is issued, then the consistency group is only deleted on the cluster on which the command is being run. When the clusters reconnect, then the consistency group is automatically deleted on the other cluster.

Alternatively, if the clusters are disconnected, and you still want to remove the consistency group on both clusters, you can issue the **svctask rmrcconsistgrp** command separately on both of the clusters.

If the consistency group is not empty, then the relationships within it are removed from the consistency group before the group is deleted. These relationships then become stand-alone relationships. The state of these relationships is not changed by the action of removing them from the consistency group.

11.2.2 Commands to cause state changes in PPRC relationships

The remaining configuration commands manage the state of either:

- Stand-alone relationships
- Sets of relationships within a consistency group

To emphasize this, they can only be addressed to these objects. Commands addressed to relationships that are part of a consistency group are rejected.

A cluster is aware of all relationships and consistency group in which it participates, and provides state information, and accepts configuration commands, for any of those relationships and consistency group.

Start a PPRC relationship (svctask startrcrelationship)

This command is used to start the copy process, set the direction of copy if undefined, and optionally mark the secondary vDisk of the relationship as clean. It is used to start a stand-alone relationship. The command fails it if it is used to attempt to start a relationship that is part of a consistency group.

This command can only be issued to a relationship that is connected. For a relationship that is idling, this command assigns a copy direction (primary and secondary roles) and begins the copy process. Otherwise this command restarts a previous copy process that was stopped either by a stop command or by some I/O error.

If the resumption of the copy process leads to a period when the relationship is not consistent, then you must specify the force flag when restarting the relationship. This situation can arise if, for example, the relationship was stopped, and then further writes were performed on the original primary of the relationship. The use of the force flag here is a reminder that the data on the secondary is not useful for disaster recovery purposes while it is in an inconsistent state.

In the idling state, you must provide the primary argument. In other connected states, you can provide the primary argument, but it must match the existing setting.

Stop a PPRC relationship (svctask stoprcrelationship)

This command is used to stop the copy process for a relationship. It may also be used to enable write access to a consistent secondary vDisk.

This command applies to a stand-alone relationship. It is rejected if it is addressed to a relationship that is part of a consistency group. You can issue this command to stop a relationship that is copying from primary to secondary.

If the relationship is in an inconsistent state, any copy operation stops and does not resume until you issue a **`svctask startrcrelationship`** command. Write activity is no longer copied from the primary to the secondary virtual disk. For a relationship in the ConsistentSynchronized state, this command causes a *consistency freeze*.

When a relationship is in a consistent state (that is, in the ConsistentStopped, ConsistentSynchronized, or ConsistentDisconnected state) then the -access argument may be used with the **`stoprcrelationship`** command to enable write access to the secondary virtual disk.

Start a PPRC consistency group (svctask startrcconsistgrp)

This command starts a Remote Copy consistency group. This command can only be issued to a consistency group that is connected.

For a consistency group that is idling, this command assigns a copy direction (primary and secondary roles) and begins the copy process. Otherwise this command restarts a previous copy process that was stopped either by a stop command or by some I/O error.

Stop a PPRC consistency group (svctask stoprcconsistgrp)

This command is used to stop the copy process. It may also be used to enable write access to the secondary vDisks in the group if the group is in a consistent state.

This command applies to a consistency group. You can issue this command to stop a consistency group that is copying from primary to secondary.

If the consistency group is in an inconsistent state, any copy operation stops and does not resume until you issue the **`svctask startrcconsistgrp`** command. Write activity is no longer copied from the primary to the secondary virtual disks belonging to the relationships in the group. For a consistency group in the ConsistentSynchronized state, this command causes a consistency freeze.

When a consistency group is in a consistent state (for example, in the ConsistentStopped, ConsistentSynchronized, or ConsistentDisconnected state), then the -access argument may be used with the **`stoprcconsistgrp`** command to enable write access to the secondary virtual disks within that group.

Reverse a PPRC relationship (svctask switchrcrelationship)

This command is used to reverse the roles of primary and secondary virtual disk when a relationship is in a consistent state.

Reverse a PPRC consistency group (svctask switchrcconsistgrp)

This command is used to reverse the roles of primary and secondary virtual disks when a consistency group is in a consistent state. This change is applied to all the relationships in the consistency group.

11.2.3 Detailed states

The following sections detail the states which are portrayed to the user, for either consistency groups or relationships. It also details the extra information available in each state. The different major states are constructed to provide guidance as to the configuration commands that are available.

InconsistentStopped

This is a connected state. In this state, the primary is accessible for read and write I/O but the secondary is not accessible for either. A copy process needs to be started to make the secondary consistent.

This state is entered when the relationship or consistency group was InconsistentCopying and has either suffered a persistent error or received a Stop command which has caused the copy process to stop.

A Start command causes the relationship or consistency group to move to the InconsistentCopying state. A Stop command is accepted, but has no effect.

If the relationship or consistency group becomes disconnected, the secondary side transitions to InconsistentDisconnected. The primary side transitions to IdlingDisconnected.

InconsistentCopying

This is a connected state. In this state, the primary is accessible for read and write I/O but the secondary is not accessible for either.

This state is entered after a Start command is issued to an InconsistentStopped relationship or consistency group. It is also entered when a forced start is issued to an idling or ConsistentStopped relationship or consistency group.

A background copy process runs which copies data from the primary to the secondary virtual disk.

In the absence of errors, an InconsistentCopying relationship is *active*, and the CopyProgress increases until the copy process completes. In some error situations, the copy progress may freeze or even regress.

A persistent error or Stop command places the relationship or consistency group into InconsistentStopped state. A Start command is accepted, but has no effect.

If the background copy process completes on a stand-alone relationship, or on all relationships for a consistency group, the relationship or consistency group transitions to ConsistentSynchronized.

If the relationship or consistency group becomes disconnected, then the secondary side transitions to InconsistentDisconnected. The primary side transitions to IdlingDisconnected.

ConsistentStopped

This is a connected state. In this state, the secondary contains a consistent image, but it may be out-of-date with respect to the primary.

This state can arise when a relationship was in ConsistentSynchronized state and suffers an error which forces a consistency freeze. It can also arise when a relationship is created with a CreateConsistentFlag set to TRUE.

Normally, following an I/O error, subsequent write activity cause updates to the primary and the secondary is no longer synchronized (set to FALSE). In this case, to re-establish synchronization, consistency must be given up for a period. A Start command with the Force option must be used to acknowledge this, and the relationship or consistency group transitions to InconsistentCopying. Do this only after all outstanding errors are repaired.

In the unusual case where the primary and secondary are still synchronized (perhaps following a user stop, and no further write I/O was received), a Start command takes the relationship to ConsistentSynchronized. No Force option is required. Also in this unusual case, a Switch command is permitted which moves the relationship or consistency group to ConsistentSynchronized and reverses the roles of the primary and secondary.

If the relationship or consistency group becomes disconnected, then the secondary side transitions to ConsistentDisconnected. The primary side transitions to IdlingDisconnected.

An informational status log is generated every time a relationship or consistency group enters the *ConsistentStopped with a status of Online* state. This can be configured to enable an SNMP trap and provide a trigger to automation software to consider issuing a Start command following a loss of synchronization.

ConsistentSynchronized

This is a connected state. In this state, the primary vDisk is accessible for read and write I/O. The secondary vDisk is accessible for read-only I/O.

Writes that are sent to the primary vDisk are sent to both primary and secondary vDisks. Either good completion must be received for both writes, or the write must be failed to the host, or a state transition out of ConsistentSynchronized must take place before a write is completed to the host.

A Stop command takes the relationship to ConsistentStopped state. A Stop command with the -access argument takes the relationship to the Idling state.

A Switch command leaves the relationship in the ConsistentSynchronized state, but reverses the primary and secondary roles.

A Start command is accepted, but has no effect.

If the relationship or consistency group becomes disconnected, the same transitions are made as for ConsistentStopped.

Idling

This is a connected state. Both master and auxiliary disks are operating in the primary role. Consequently both are accessible for write I/O.

In this state, the relationship or consistency group accepts a Start command. Remote Copy maintains a record of regions on each disk which received write I/O while Idling. This is used to determine what areas need to be copied following a Start command.

The Start command must specify the new copy direction. A Start command can cause a loss of consistency if either virtual disk in any relationship has received write I/O. This is indicated by the synchronized status. If the Start command leads to loss of consistency, then a Forced flag must be specified.

Following a Start command, the relationship or consistency group transitions to ConsistentSynchronized if there is no loss of consistency, or to InconsistentCopying if there is.

Also, while in this state, the relationship or consistency group accepts a Clean option on the Start command.

If the relationship or consistency group becomes disconnected, then both sides change their state to IdlingDisconnected.

IdlingDisconnected

This is a disconnected state. The virtual disk or disks in this half of the relationship or consistency group are all in the primary role and accept read or write I/O.

The main priority in this state is to recover the link and make the relationship or consistency group connected once more.

No configuration activity is possible (except for deletes or stops) until the relationship becomes connected again. At that point, the relationship transition to a connected state. The exact connected state which is entered depends on the state of the other half of the relationship or consistency group, which depends on:

- The state when it became disconnected
- The write activity since it was disconnected
- The configuration activity since it was disconnected

If both halves are IdlingDisconnected, then the relationship becomes idling when reconnected.

While IdlingDisconnected, if a write I/O is received which causes loss of synchronization (synchronized attribute transitions from TRUE to FALSE) and the relationship was not already stopped (either through user stop or a persistent error), then an error log is raised to notify this. This error log is the same as that raised when the same situation arises when ConsistentSynchronized.

InconsistentDisconnected

This is a disconnected state. The virtual disks in this half of the relationship or consistency group are all in the secondary role and do not accept read or write I/O.

No configuration activity except for deletes is permitted until the relationship becomes connected again.

When the relationship or consistency group becomes connected again, the relationship becomes InconsistentCopying automatically unless either:

- The relationship was InconsistentStopped when it became disconnected
- The user issued a Stop while disconnected

In either case, the relationship or consistency group becomes InconsistentStopped.

ConsistentDisconnected

This is a disconnected state. The vDisks in this half of the relationship or consistency group are all in the secondary role and accept read I/O but not write I/O.

This state is entered from ConsistentSynchronized or ConsistentStopped when the secondary side of a relationship becomes disconnected.

In this state, the relationship or consistency group displays an attribute of FreezeTime which is the point in time that Consistency was frozen. When entered from ConsistentStopped, it retains the time it had in that state. When entered from ConsistentSynchronized, the FreezeTime shows the last time at which the relationship or consistency group was known to be consistent. This corresponds to the time of the last successful heartbeat to the other cluster.

A Stop with EnableAccessFlag set to TRUE transitions the relationship or consistency group to IdlingDisconnected state. This allows write I/O to be performed to the virtual disks and is used as part of a disaster recovery scenario.

When the relationship or consistency group becomes connected again, the relationship or consistency group becomes ConsistentSynchronized only if this does not lead to a loss of Consistency. This is the case provided:

- The relationship was ConsistentSynchronized when it became disconnected.
- No writes received successful completion at the primary while disconnected.

Otherwise the relationship become ConsistentStopped. The FreezeTime setting is retained.

Empty

This state only applies to consistency groups. It is the state of a consistency group which has no relationships and no other state information to show.

It is entered when a consistency group is first created. It is exited when the first relationship is added to the consistency group at which point the state of the relationship becomes the state of the consistency group.

11.2.4 Background copy

Remote Copy paces the rate at which background copy is performed by the appropriate relationships. Background copy takes place on relationships which are in InconsistentCopying state with a Status of Online.

The CopyPriority attribute governs the order in which relationships are allowed to perform background copy. Only those relationships that share the highest CopyPriority of any relationship in an eligible state (InconsitentCopying/Online) are permitted to perform any background copy. Relationships with lower CopyPriority do not perform any background copy until all relationships with a higher priority have either become ConsistentSynchronized or have become ineligible for background copy. A relationship becomes ineligible if it transitions to InconsistentStopped due to a User Stop or a persistent error, or if either primary or secondary virtual disks become offline.

The quota of background copy (configured on the intercluster link) is divided evenly between the nodes that are performing background copy for one of the eligible relationships. This allocation is made without regard for the number of disks that node is responsible for. Each node in turn divides its allocation evenly between the multiple relationships performing a background copy.

For intracluster relationships, each node is assigned a static quota of 50 MB/s.

11.3 Step-by-step guide to Peer-to-Peer Remote Copy

The following example was carried out using two AIX host systems (aix1 and aix2) attached to two SVCs in two different clusters. The back-end storage is made of a FAStT 600 for the cluster SVC1 (name REDSTONE1) and of a FAStT 700 for the cluster SVC2 (name lodestoneA). See Figure 11-5.

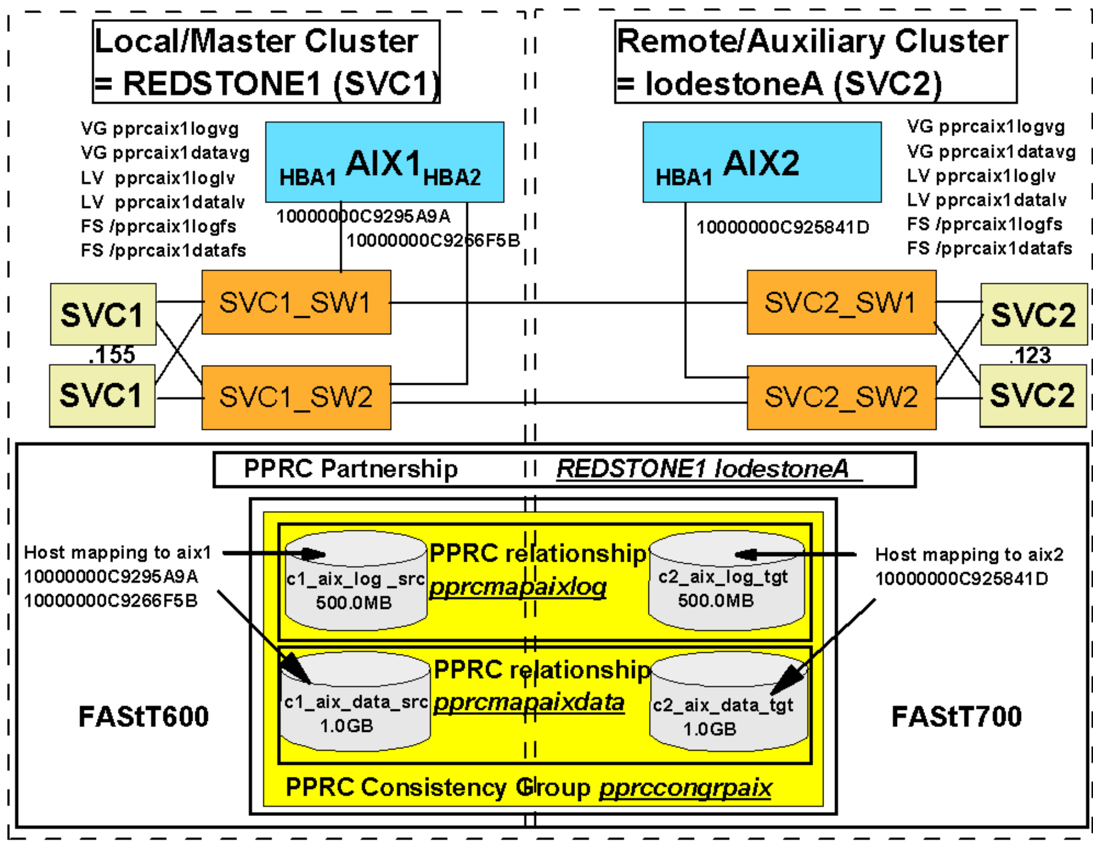


Figure 11-5 SVC PPRC scenario

PPRC/Remote Copy allows for the creation of a remote vDisk that is identical to the source. To ensure that both of the vDisks contain identical data they are maintained together in a synchronous relationship. The SVC supports remote copy between vDisks within a cluster or outside of a cluster the master and auxiliary disks MUST be the same capacity. The SVC also provides the ability to create consistency groups which allow for multiple disks to be grouped as part of a remote copy set. The auxiliary disk is read-only, while the remote copy relationship is consistent.

To create the SVC PPRC, you complete these steps:

1. Create a PPRC intercluster partnership between the local and master cluster REDSTONE1 and the remote and auxiliary cluster lodestoneA.
2. Create a PPRC/Remote Copy consistency group pprccongpraix.
3. Create two PPRC/Remote Copy relationships that will be mapped in the PPRC consistency group pprccongpraix:

   – pprcmapaixlog between the master vDisk c1_aix_log_src and the auxiliary vDisk c2_aix_log_tgt
   – pprcmapaixdata between the master vDisk c1_aix_data_src and the auxiliary vDisk c2_aix_data_tgt
4. Start the PPRC/Remote Copy consistency group pprccongpraix to synchronize the master and the auxiliary vDisks.
5. Stop the PPRC/Remote Copy consistency group pprccongpraix with option -access write access to the secondary vDisks to the aix2 host system.

The volume definition for aix1 and aix2 host systems is shown in Example 11-1 and Example 11-2.

Example 11-1 Volume definition for aix1 before PPRC

```
>>>>> 2 disks from ESS 2105 (vpath0 and vpath1) are already defined and configured

AIX1#lspv
hdisk0          0007905873b5c2fb                    rootvg
hdisk1          0007905873b5c364                    rootvg
hdisk2          none                                None
hdisk3          none                                None
hdisk4          none                                None
hdisk5          none                                None
vpath0          00079058573c2177                    essdisk1vg
vpath1          00079058573c65a9                    essdisk2vg

AIX1#lsdev -Cc disk
hdisk0 Available 10-60-00-8,0 16 Bit SCSI Disk Drive
hdisk1 Available 10-60-00-9,0 16 Bit SCSI Disk Drive
hdisk2 Available 20-58-01     IBM FC 2105F20
hdisk3 Available 20-58-01     IBM FC 2105F20
hdisk4 Available 20-60-01     IBM FC 2105F20
hdisk5 Available 20-60-01     IBM FC 2105F20
vpath0 Available              Data Path Optimizer Pseudo Device Driver
vpath1 Available              Data Path Optimizer Pseudo Device Driver

AIX1#datapath query adapter
Active Adapters :2
Adpt#     Adapter Name   State     Mode      Select     Errors  Paths  Active
    0           fscsi0  NORMAL   ACTIVE         94          0      2       2
    1           fscsi1  NORMAL   ACTIVE        108          0      2       2

AIX1#datapath query device
Total Devices : 2
DEV#:   0  DEVICE NAME: vpath0  TYPE: 2105F20         POLICY:    Optimized
SERIAL: 40514830
==========================================================================
Path#      Adapter/Hard Disk          State     Mode      Select     Errors
    0          fscsi0/hdisk2           OPEN   NORMAL          51          0
    1          fscsi1/hdisk4           OPEN   NORMAL          53          0

DEV#:   1  DEVICE NAME: vpath1  TYPE: 2105F20         POLICY:    Optimized
SERIAL: 40914830
==========================================================================
Path#      Adapter/Hard Disk          State     Mode      Select     Errors
    0          fscsi0/hdisk3           OPEN   NORMAL          41          0
    1          fscsi1/hdisk5           OPEN   NORMAL          55          0
```

```
AIX1#lsvg -o
essdisk2vg
essdisk1vg
rootvg

AIX1#df -k
Filesystem    1024-blocks      Free %Used    Iused %Iused Mounted on
/dev/hd4            16384      7620   54%     1292    16% /
/dev/hd2          1867776   1265400   33%    20526     5% /usr
/dev/hd9var         16384      9924   40%      436    11% /var
/dev/hd3            32768     29176   11%       62     1% /tmp
/dev/hd1            16384     15820    4%       18     1% /home
/proc                   -         -    -         -     -  /proc
/dev/hd10opt        32768     26592   19%      294     4% /opt
/dev/essdisk1lv       81920     79308    4%       17     1% /tmp/essdisk1fs
/dev/essdisk2lv       81920     79308    4%       17     1% /tmp/essdisk2fs
```

The volume definition on host aix2 is shown in Example 11-2.

Example 11-2 Volume definition for aix2 before PPRC

```
>>>>> 1 disk from SVC 2145 (vpath0) is already defined and configured

AIX2#lspv
hdisk0         000b21dd2cf9b83d    rootvg
hdisk1         none                None
hdisk2         none                None
vpath0         000b21dd56751ac9    testvg

AIX2#lsdev -Cc disk
hdisk0 Available 10-80-00-2,0 16 Bit SCSI Disk Drive
hdisk1 Available 10-90-01     IBM TotalStorage SAN Volume Controller device
hdisk2 Available 10-90-01     IBM TotalStorage SAN Volume Controller device
vpath0 Available              Data Path Optimizer Pseudo Device Driver

AIX2#datapath query adapter
Active Adapters :1
Adpt#     Adapter Name   State     Mode     Select     Errors  Paths  Active
    0           fscsi0  NORMAL   ACTIVE       2030          0      2       2

AIX2#datapath query device
Total Devices : 1
DEV#:   0  DEVICE NAME: vpath0  TYPE: 2145           POLICY:    Optimized
SERIAL: 600507680186000960000000000000000
==========================================================================
Path#      Adapter/Hard Disk          State     Mode     Select     Errors
    0          fscsi0/hdisk1           OPEN   NORMAL       2030          0
    1          fscsi0/hdisk2           OPEN   NORMAL          0          0

AIX2#lsvpcfg
vpath0 (Avail pv testvg) 600507680186000960000000000000000 = hdisk1 (Avail ) hdisk2 (Avail )

AIX2#lsvg -o
testvg
rootvg

AIX2#df -k
```

```
Filesystem    1024-blocks      Free %Used    Iused %Iused Mounted on
/dev/hd4            32768     26380   20%     1180     8% /
/dev/hd2           393216     79132   80%    14361    15% /usr
/dev/hd9var         32768     30268    8%      155     2% /var
/dev/hd3            32768     31052    6%       61     1% /tmp
/dev/hd1            32768     31684    4%       18     1% /home
/dev/testlv         40960     39632    4%       18     1% /testfs
```

11.3.1 Checking the status of the PPRC feature

Check that the feature PPRC or Remote Copy is either on or enabled.

Using the CLI

This is shown in Example 11-3.

Example 11-3 Feature PPRC/Remote Copy is on

```
SVC1:admin>svcinfo lslicense
feature_flash on
feature_remote on
feature_num_gb 2000
```

Using the GUI

Follow these steps as shown in Figure 11-6:

1. From the left pane, select **Service and Maintenance**.
2. Under Action on the left pane, select **Set Featurization**.
3. In the Featurization Settings panel, check the feature setting for Remote Copy. In our example, it is *Enabled*.

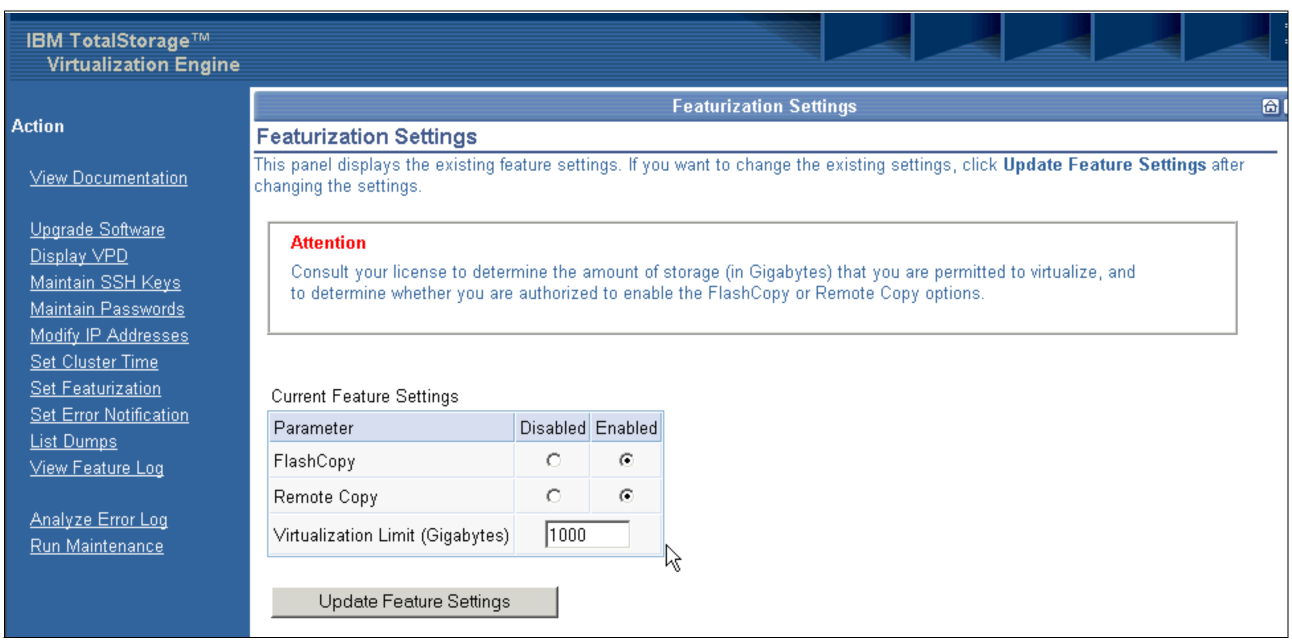


Figure 11-6 Checking that the feature PPRC/Remote Copy is enabled

11.3.2 Checking the status of the PPRC target vDisk

Ensure that your PPRC target disks are not mapped to your host server aix2.

Using the CLI

This procedure is shown using the CLI in Example 11-4 and Example 11-5.

Example 11-4 shows the configuration on SVC1 before PPRC.

Example 11-4 aix1 and aix2 host system configurations on SVC1 before PPRC

```
SVC1:admin>svcinfo lsvdisk -delim /
id/name/IO_group_id/IO_group_name/status/mDisk_grp_id/mDisk_grp_name/capacity/type/FC_id/FC
_name/RC_id/RC_name
12/c1_aix_log_src/0/ITSO_grp0/online/1/ARR36P6N/500.0MB/striped////
13/c1_aix_data_src/0/ITSO_grp0/online/2/ARR73P5N/1.0GB/striped////

-------------- vDisk c1_aix_log_src and c1_aix_data_src mapped to host aix1 ------------

SVC1:admin>svcinfo lshostvdiskmap aix1 -delim /
id/name/SCSI_id/vDisk_id/vDisk_name/wwpn/vDisk_UID
0/aix1/0/12/c1_aix_log_src/10000000C9295A9A/600507680182001B2000000000000018
0/aix1/1/13/c1_aix_data_src/10000000C9295A9A/600507680182001B2000000000000019
0/aix1/0/12/c1_aix_log_src/10000000C9266F5B/600507680182001B2000000000000018
0/aix1/1/13/c1_aix_data_src/10000000C9266F5B/600507680182001B2000000000000019

--------------- no vDisk mapped to aix2 -----------------------------------

SVC1:admin>svcinfo lshostvdiskmap aix2
SVC1:admin>
```

Example 11-5 shows the configuration on SVC2 before PPRC.

Example 11-5 aix1 and aix2 host system configurations on SVC2 before PPRC

```
SVC2:admin>svcinfo lsvdisk -delim /
id/name/IO_group_id/IO_group_name/status/mDisk_grp_id/mDisk_grp_name/capacity/type/FC_id/FC
_name/RC_id/RC_name
0/svc2aix2test/0/io_grp0/online/0/SVC2_F1_MDG1/100.0MB/striped////
1/c2_aix_log_tgt/0/io_grp0/online/0/SVC2_F1_MDG1/500.0MB/striped////
2/c2_aix_data_tgt/0/io_grp0/online/0/SVC2_F1_MDG1/1.0GB/striped////

--------- no host mapping to aix2 for c2_aix_log_tgt and c2_aix_data_tgt---------

SVC2:admin>svcinfo lshostvdiskmap aix2 -delim /
id/name/SCSI_id/vDisk_id/vDisk_name/wwpn/vDisk_UID
0/aix2/0/0/svc2aix2test/10000000C925841D/600507680186000960000000000000000
```

Using the GUI

As shown in Figure 11-7, follow these steps:

1. Under My Work, select **Work with Virtual Disks** and select **Hosts**.
2. In the Viewing Virtual Disks panel, under the Select column, select the host you want.
3. Select **Create a vDisk** from the list and click **Go**.

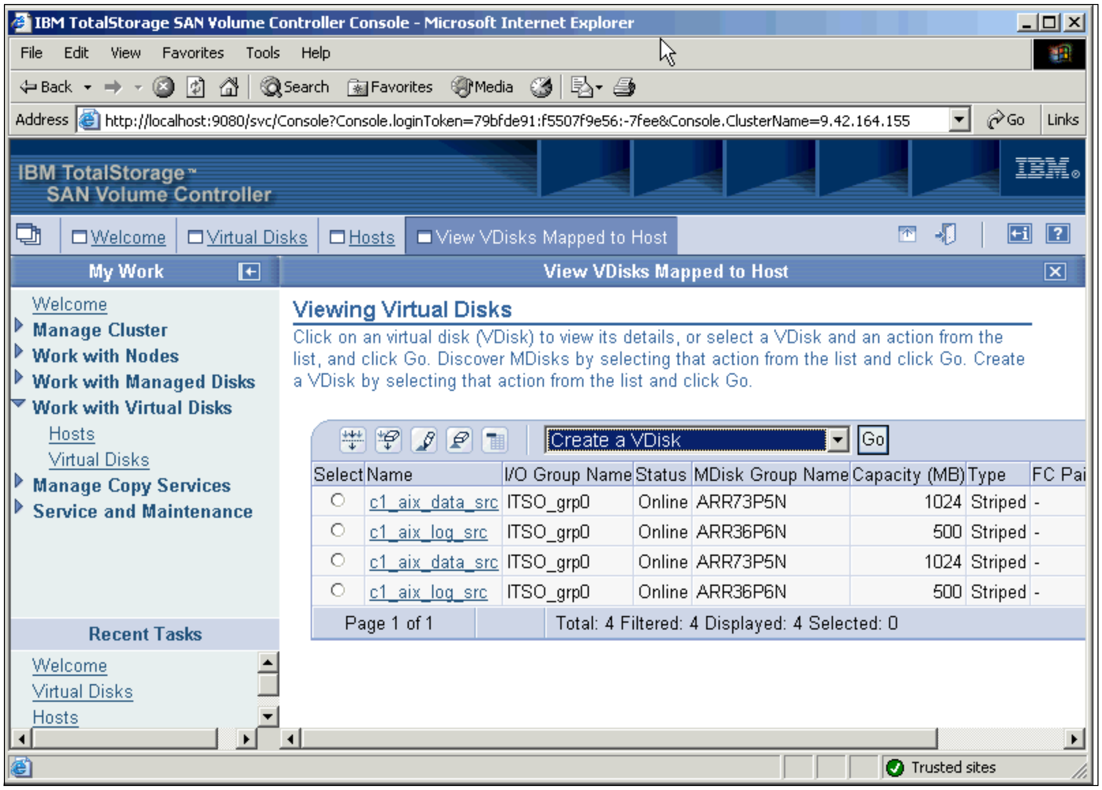


Figure 11-7 Check for vDisk c1_aix_data_src and c1_aix_log_src mapped to host aix1 in SVC1

On SVC1 and SVC2, no vDisk is mapped to aix2 host system.

11.3.3 Creating a PPRC intercluster partnership

To configure an intercluster Peer-to-Peer Remote Copy, you must create a partnership between the two SVC clusters. The creation of the partnership must be completed on both clusters for this to be successful.

Before creating the PPRC intercluster partnership, make sure that a correct zone is defined between the SVC nodes of the local cluster SVC1 (REDSTONE1) and the remote cluster SVC2 (lodestoneA) for all the FC ports of the nodes.

Using the CLI

To look for candidate clusters, enter the following command:

```
svcinfo lsclustercandidate
```

To create the partnership on REDSTONE1 (SVC1), enter:

```
svctask mkpartnership -bandwidth 50 lodestoneA
```

To create the partnership on lodestoneA (SVC2), enter:

```
svctask mkpartnership -bandwidth 50 REDSTONE1
```

Example 11-6 shows creating the partnership on REDSTONE1 (SVC1).

Example 11-6 Creating a PPRC intercluster partnership

```
SVC1:admin>svcinfo lscluster -delim :
id:name:location:partnership:bandwidth:cluster_IP_address:cluster_service_IP_address
0000010030400364:REDSTONE1:local:::9.42.164.155:9.42.164.160

SVC1:admin>svcinfo lsclustercandidate

>>>>>> no cluster candidate because no zoning is defined the two SVC clusters
>>>>>> after the creation of a SVC inter_cluster zone

SVC1:admin>svcinfo lsclustercandidate
id               configured     cluster_name
0000010030C0012C no             lodestoneA

>>>>>>>>>>>>>>> create partnership from SVC1 to SVC2
>>>>>>>>>>>>>>> Partnership from SVC2 to SVC1 already done

SVC1:admin>svctask mkpartnership -bandwidth 50 lodestoneA

SVC1:admin>svcinfo lscluster -delim /
id/name/location/partnership/bandwidth/cluster_IP_address/cluster_service_IP_address
0000010030400364/REDSTONE1/local///9.42.164.155/9.42.164.160
0000010030C0012C/lodestoneA/remote/fully_configured/50/9.42.164.123/9.42.164.120
```

Example 11-7 shows creating the partnership on lodestoneA (SVC2).

Example 11-7 PPRC intercluster partnership between REDSTONE1 and lodestoneA on SVC2

```
>>>>>>>>>>>>>>>>>>>>>>>  status of cluster lodestoneA (SVC2)

SVC2:admin>svcinfo lscluster -delim /
id/name/location/partnership/bandwidth/cluster_IP_address/cluster_service_IP_address
0000010030C0012C/lodestoneA/local///9.42.164.123/9.42.164.120

>>>>>>>>>> Remote cluster REDSTONE1 candidate for intercluster partnership

SVC2:admin>svcinfo lsclustercandidate
id               configured     cluster_name
0000010030400364 no             REDSTONE1

>>>>>>>>>>>>>>> create partnership from SVC2 to SVC1
>>>>>>>>>>>>>>> Partnership from SVC1 to SVC2 not yet done

SVC2:admin>svctask mkpartnership -bandwidth 50 REDSTONE1

SVC2:admin>svcinfo lscluster -delim /
id/name/location/partnership/bandwidth/cluster_IP_address/cluster_service_IP_address
0000010030C0012C/lodestoneA/local///9.42.164.123/9.42.164.120
0000010030400364/REDSTONE1/remote/partially_configured_local/50/9.42.164.155/9.42.164.160

>>>>>> after cluster partnership on SVC1 defined

SVC2:admin>svcinfo lscluster -delim /
id/name/location/partnership/bandwidth/cluster_IP_address/cluster_service_IP_address
0000010030C0012C/lodestoneA/local///9.42.164.123/9.42.164.120
0000010030400364/REDSTONE1/remote/fully_configured/50/9.42.164.155/9.42.164.160
```

Using the GUI

You see the candidate when you create the partnership. To create the PPRC intercluster partnership, follow these steps as shown in Figure 11-8:

1. Under My Work, select **Manage Copy Services** and select **Remote Copy Cluster Partnership**.
2. In the Remote Copy Cluster Partnership panel, click the **Create** button.

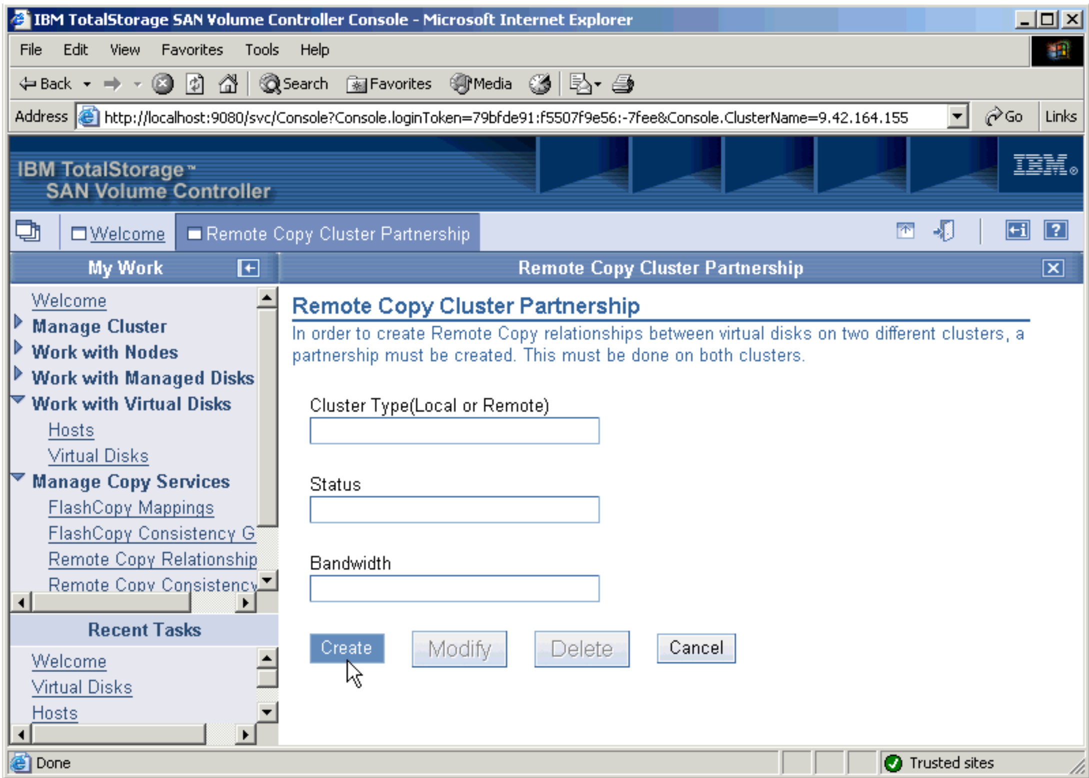


Figure 11-8 Remote Copy Cluster Partnership

3. On the Create Cluster Partnership panel (Figure 11-9), select the remote cluster. Enter the bandwidth (MB) for the connection between the clusters and click **OK**.

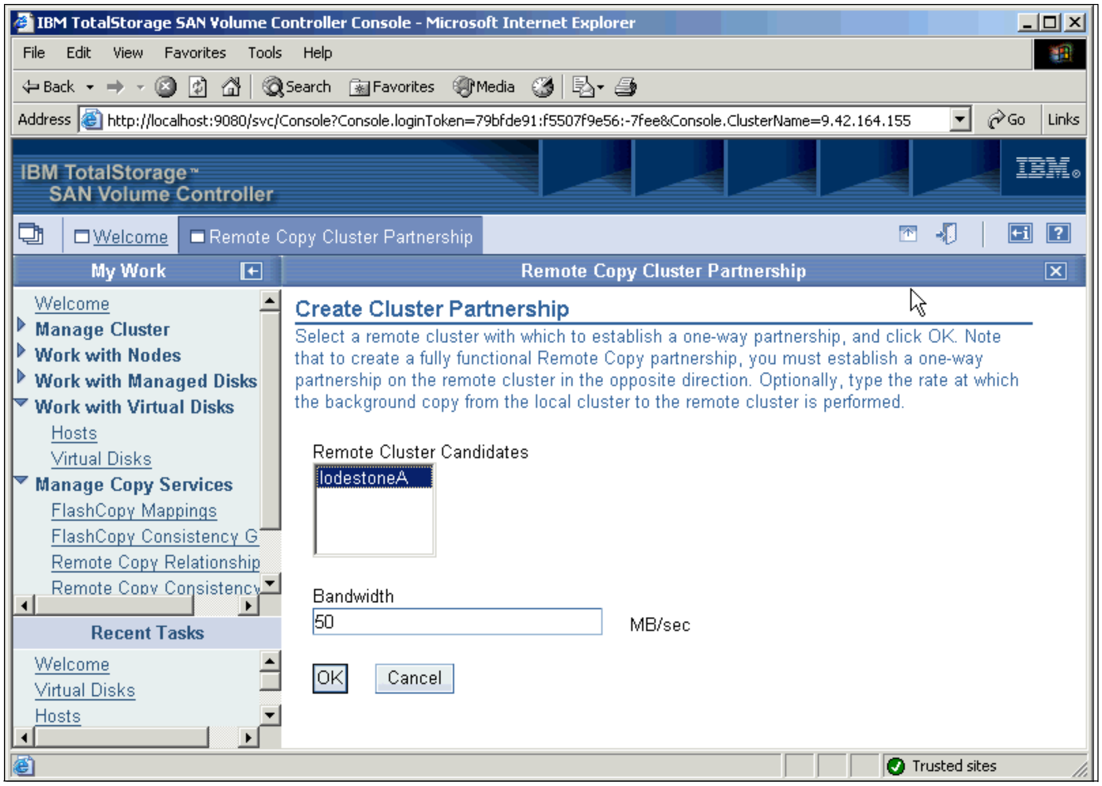


Figure 11-9 Create Cluster Partnership

On the Remote Copy Cluster Partnership panel (Figure 11-10), you see the status and bandwidth of the partnership.

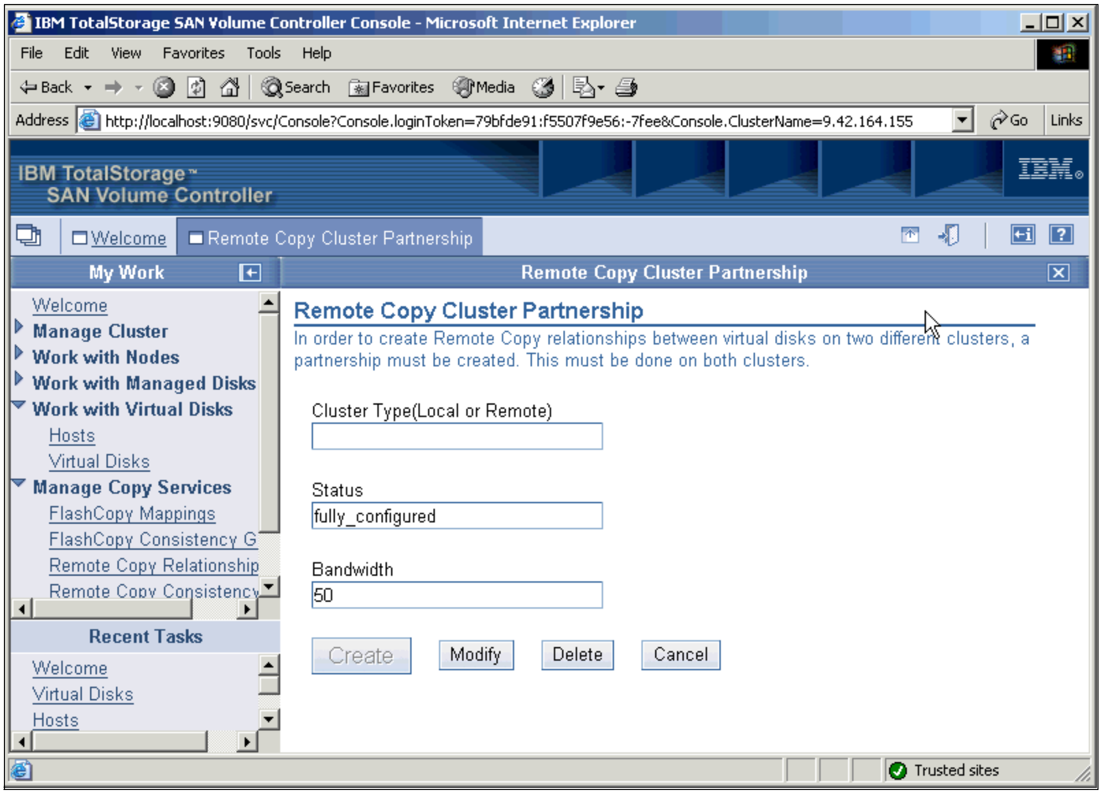


Figure 11-10 Creation of a PPRC intercluster partnership done

Attention: Remember to do this for *both* SVC clusters.

After the creation of the PPRC intercluster partnership, you do not need any more to work on the remote cluster SVC2 (lodestoneA). All the commands are issued from the master cluster SVC1 (REDSTONE1)

11.3.4 Creating a PPRC consistency group

After the partnership is established, you can continue. At this point, you can create a PPRC/Remote Copy consistency group and then create PPRC relationships. Or you can go straight to creating the PPRC/Remote Copy relationships.

As a practice, always use consistency groups for PPRC relationships. Although they are not required, if you add a PPRC relationship, it is helpful to have a consistency group created.

You can enter commands to start, stop, and reverse copy direction against either the PPRC relationship or the PPRC consistency group.

Using the CLI

To create a PPRC consistency group pprccongrpaix between the local cluster and the remote cluster lodestoneA (intercluster), enter the following command:

```
svctask mkrcconsistgrp -name pprccongrpaix -cluster lodestoneA
```

If you want to create a PPRC consistency group within the cluster (intracluster), you don't have to specify the remote cluster name:

```
svctask mkrcconsistgrp -name pprccongrpaix
```

Example 11-8 shows this.

Example 11-8 Creating PPRC consistency group pprccongrpaix on SVC1

```
>>>>>>>>>>>>>> create the consistency group with the option of remote cluster lodestoneA

SVC1:admin>svctask mkrcconsistgrp -name pprccongrpaix -cluster lodestoneA
RC Consistency Group, id [255], successfully created

>>>>> PPRC consistency group "pprccongrpaix" is created as empty (no PPRC relationship)

SVC1:admin>svcinfo lsrcconsistgrp -delim /
id/name/master_cluster_id/master_cluster_name/aux_cluster_id/aux_cluster_name/primary/state
/relationship_count
255/pprccongrpaix/0000010030400364/REDSTONE1/0000010030C0012C/lodestoneA//empty/0

SVC1:admin>svcinfo lsrcconsistgrp pprccongrpaix
id 255
name pprccongrpaix
master_cluster_id 0000010030400364
master_cluster_name REDSTONE1
aux_cluster_id 0000010030C0012C
aux_cluster_name lodestoneA
primary
state empty
relationship_count 0
freeze_time
status
sync
```

To monitor the state of the PPRC consistency groups, you can use:

```
svcinfo lsrcconsistgrp
svcinfo lsrcconsistgrp pprccongrpaix
```

You can see this in Example 11-9.

Example 11-9 PPRC consistency group pprccongrpaix on SVC2

```
>>>>>>>>>> Before the creation of consistency group "pprccongrpaix"

SVC2:admin>svcinfo lsrcconsistgrp

>>>>>>>>>> After the creation of consistency group "pprccongrpaix" on SVC1

SVC2:admin>svcinfo lsrcconsistgrp -delim /
id/name/master_cluster_id/master_cluster_name/aux_cluster_id/aux_cluster_name/primary/state
/relationship_count
255/pprccongrpaix/0000010030400364/REDSTONE1/0000010030C0012C/lodestoneA//empty/0
```

Using the GUI

To create a PPRC consistency group pprccongrpaix between the local cluster and the remote cluster lodestoneA (intercluster), follow these steps:

1. As shown in Figure 11-11, under My Work, select **Manage Copy Services** and select **Remote Copy Consistency Groups**.
2. In the View Remote Copy Consistency Groups panel, select **Create a Consistency Group** from the list and click **Go**.

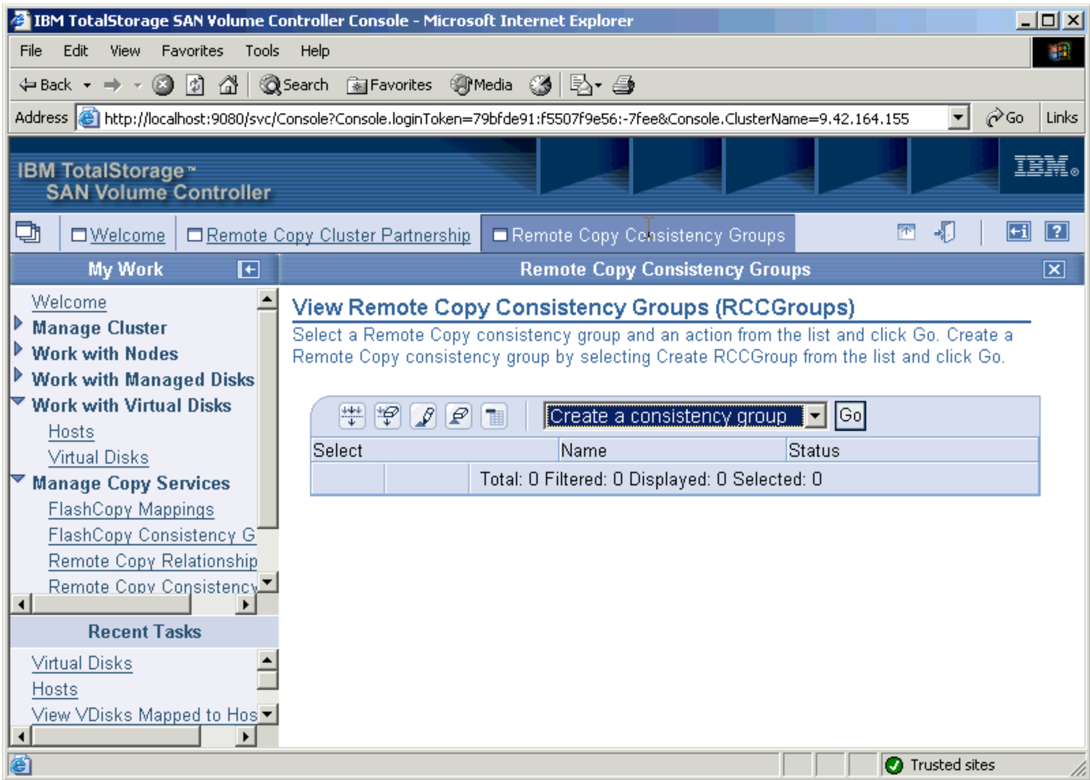


Figure 11-11 View Remote Copy Consistency Groups

3. On the first panel of the Remote Copy Consistency Groups wizard (Figure 11-12), click **Next**.
4. On the next panel (not shown), specify the name of your PPRC consistency group pprccongrpaix. Select the **remote cluster in RCpair candidate**. Click **OK**.

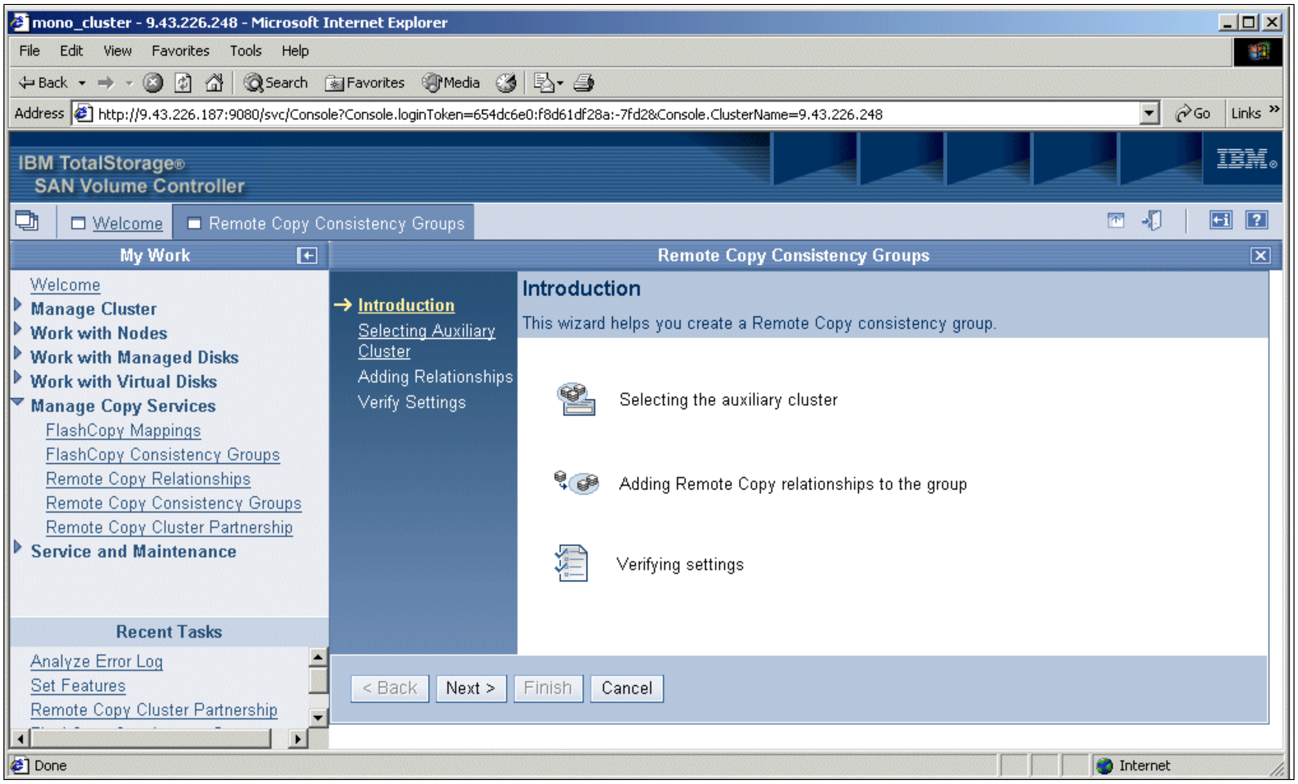


Figure 11-12 Remote Copy Consistency Groups wizard

5. On the next panel (Figure 11-13), determine whether the relationship is intracluster or intercluster. Click **Next**.

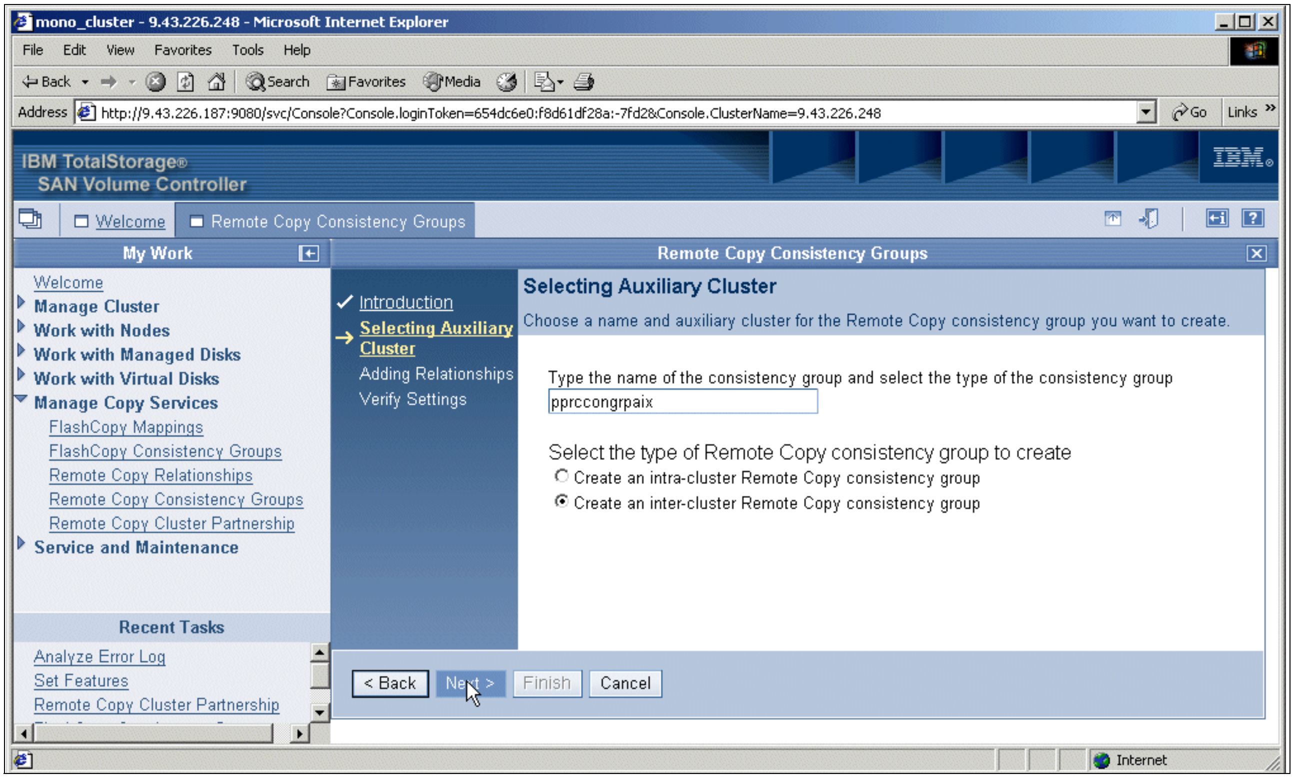


Figure 11-13 Selecting Auxiliary Cluster

6. On the Add Remote Copy Relationships panel (Figure 11-14), add remote copy relationships to the consistency group. When you are finished, click **Next**.

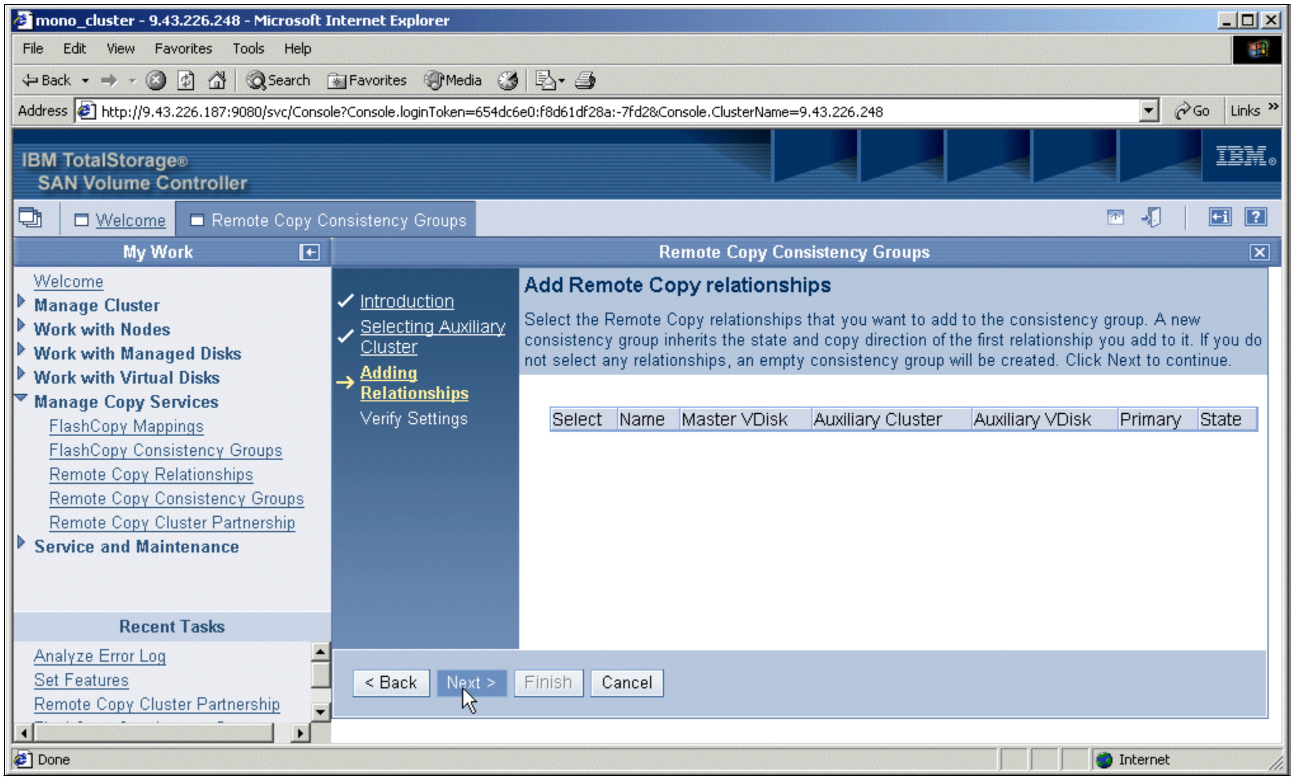


Figure 11-14 Add Remote Copy relationships

7. On the Verify Settings panel (Figure 11-15), verify your settings and click **Finish**.

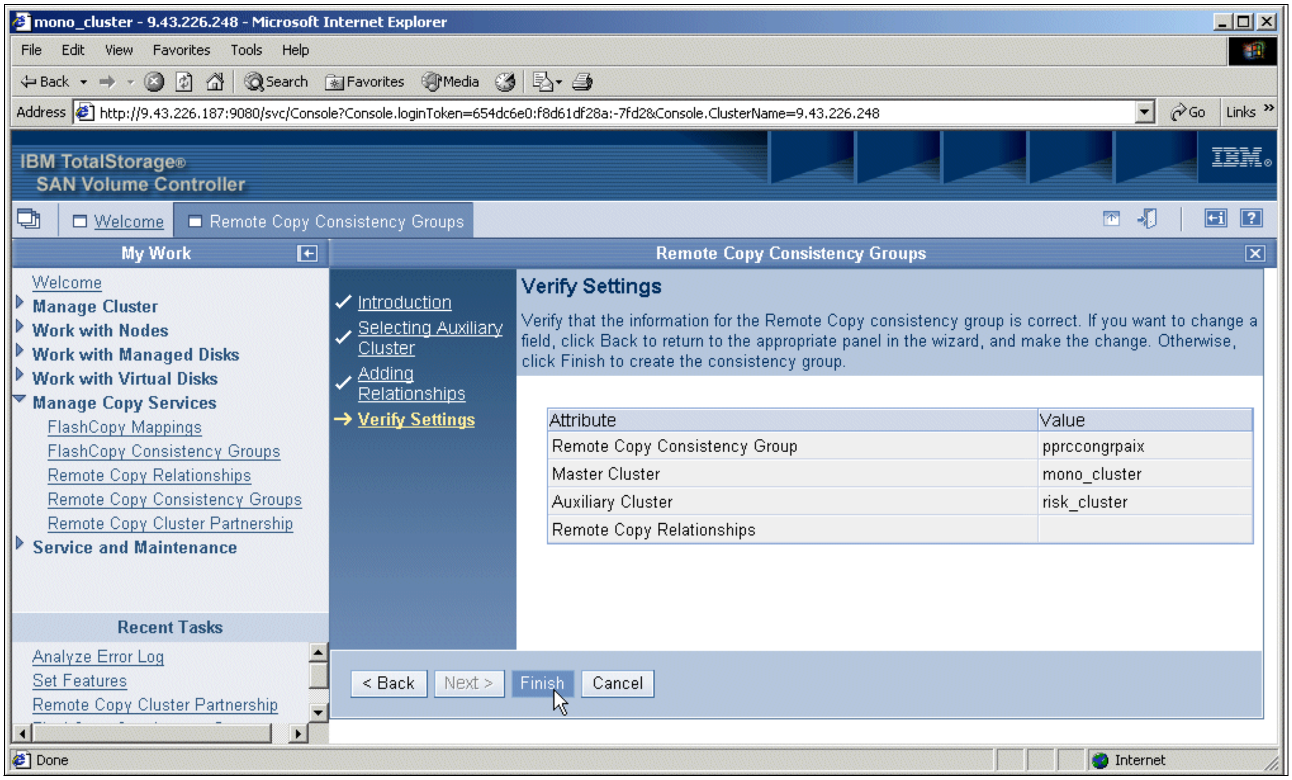


Figure 11-15 Verify Settings

8. On the View Remote Copy Consistency Groups panel (Figure 11-16), verify the PPRC consistency group.

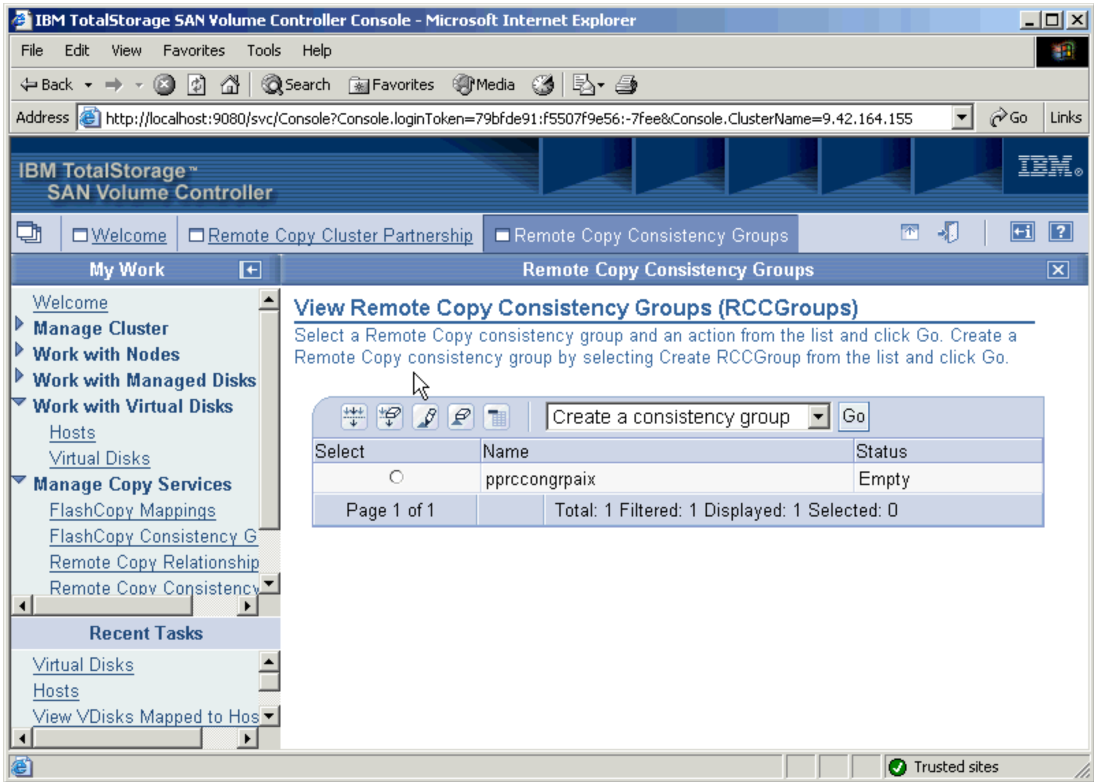


Figure 11-16 Creating a PPRC consistency group pprccongrpaix done

11.3.5 Creating the PPRC relationship

We create two PPRC/Remote Copy relationships between the following vDisks:

- A first PPRC relationship pprcmapaixlog between the master vDisk c1_aix_log_src in local cluster SVC1 and the auxiliary vDisk c2_aix_log_src in remote cluster SVC2.
- A second PPRC relationship pprcmapaixdata between the master vDisk c1_aix_data_src in local cluster SVC1 and the auxiliary vDisk c2_aix_data_src in remote cluster SVC2.

Both relationships are assigned to the PPRC consistency group pprccongrpaix.

Using the CLI

The following command is used to create a PPRC relationship pprcmapaixdata between a master vDisk c1_aix_data_src and an auxiliary vDisk c2_aix_data_tgt which is in the remote cluster lodestoneA:

```
svctask mkrcrelationship -master c1_aix_data_src -aux c2_aix_data_tgt -cluster lodestoneA
-name pprcmapaixdata -consistgrp pprccongrpaix -priority 50
```

The PPRC relationship is included in the PPRC consistency group pprccongrpaix. The background copy priority is set to `50`. When background Copy Services, such as PPRC, are running, copy activities can be given priorities. This way, copy services the with highest priority are synchronized before any attempt is made to synchronize a lower priority copy service.

In the local cluster or in the remote cluster, to find the vDisks that are available for PPRC relationships, you can use the command:

```
svcinfo lsrcrelationshipcandidate
```

To monitor the PPRC relationships, you can use either of the following commands:

```
svcinfo lsrcrelationship
svcinfo lsrcrelationship pprcmapaixdata
```

Example 11-10 shows this.

Example 11-10 Creating PPRC relationship pprcmapaixdata on SVC1

```
>>>>>> The following vDisks are available for the PPRC relationship

SVC1:admin>svcinfo lsrcrelationshipcandidate
id                vDisk_name
0                 aix_vDisk0
1                 aix_vDisk1
2                 aix_vDisk2
3                 aix_vDisk3
4                 windisk1
5                 redhatdisk1
6                 aix_vDisk4
7                 rhat2_vDisk0
8                 aix_vDisk1_src
10                SAP_DATA02
11                SAP_DATA03
12                c1_aix_log_src
13                c1_aix_data_src

>>>>>>>>>> The first PPRC relationship “pprcmapaixlog” has been created using the GUI

SVC1:admin>svcinfo lsrcrelationship -delim /
id/name/master_cluster_id/master_cluster_name/master_vDisk_id/master_vDisk_name/aux_cluster
_id/aux_cluster_name/aux_vDisk_id/aux_vDisk_name/primary/consistency_group_id/consistency_g
roup_name/state/bg_copy_priority/progress
12/pprcmapaixlog/0000010030400364/REDSTONE1/12/c1_aix_log_src/0000010030C0012C/lodestoneA/1
//master/255/pprccongrpaix/inconsistent_stopped/50/0

SVC1:admin>svcinfo lsrcrelationship pprcmapaixlog
id 12
name pprcmapaixlog
master_cluster_id 0000010030400364
master_cluster_name REDSTONE1
master_vDisk_id 12
master_vDisk_name c1_aix_log_src
aux_cluster_id 0000010030C0012C
aux_cluster_name lodestoneA
aux_vDisk_id 1
aux_vDisk_name
primary master
consistency_group_id 255
consistency_group_name pprccongrpaix
state inconsistent_stopped
bg_copy_priority 50
progress 0
freeze_time
status online
```

```
sync

>>>>> the PPRC relationship "pprcmapaixlog" has been added to PPRC consistency group
"pprccongrpaix"

SVC1:admin>svcinfo lsrcconsistgrp -delim /
id/name/master_cluster_id/master_cluster_name/aux_cluster_id/aux_cluster_name/primary/state
/relationship_count
255/pprccongrpaix/0000010030400364/REDSTONE1/0000010030C0012C/lodestoneA/master/inconsisten
t_stopped/1

>>>> The state of the PPRC consistency group changed from "empty" to "inconsistent_stopped"
>>>> when we added the first PPRC relationship which state was "inconsistent_stopped"

SVC1:admin>svcinfo lsrcconsistgrp pprccongrpaix
id 255
name pprccongrpaix
master_cluster_id 0000010030400364
master_cluster_name REDSTONE1
aux_cluster_id 0000010030C0012C
aux_cluster_name lodestoneA
primary master
state inconsistent_stopped
relationship_count 1
freeze_time
status online
sync
RC_rel_id 12
RC_rel_name pprcmapaixlog

>>>>> the vDisk candidates for PPRC relationship "pprcmapaixdata"

SVC1:admin>svcinfo lsrcrelationshipcandidate
id              vDisk_name
0               aix_vDisk0
1               aix_vDisk1
2               aix_vDisk2
3               aix_vDisk3
4               windisk1
5               redhatdisk1
6               aix_vDisk4
7               rhat2_vDisk0
8               aix_vDisk1_src
10              SAP_DATA02
11              SAP_DATA03
13              c1_aix_data_src

>>>>>>>>>>>>    Creation of the second relationship "pprcmapaixdata"

SVC1:admin>svctask mkrcrelationship -master c1_aix_data_src -aux c2_aix_data_tgt -cluster
lodestoneA -name pprcmapaixdata -consistgrp pprccongrpaix -priority 50
RC Relationship, id [13], successfully created

>>>>>>> now the PPRC consistency group "pprccongrpaix" has two PPRC relationships

SVC1:admin>svcinfo lsrcconsistgrp -delim /
id/name/master_cluster_id/master_cluster_name/aux_cluster_id/aux_cluster_name/primary/state
/relationship_count
255/pprccongrpaix/0000010030400364/REDSTONE1/0000010030C0012C/lodestoneA/master/inconsisten
t_stopped/2
```

```
SVC1:admin>svcinfo lsrcconsistgrp pprccongrpaix
id 255
name pprccongrpaix
master_cluster_id 0000010030400364
master_cluster_name REDSTONE1
aux_cluster_id 0000010030C0012C
aux_cluster_name lodestoneA
primary master
state inconsistent_stopped
relationship_count 2
freeze_time
status online
sync
RC_rel_id 12
RC_rel_name pprcmapaixlog
RC_rel_id 13
RC_rel_name pprcmapaixdata


SVC1:admin>svcinfo lsrcrelationship -delim /
id/name/master_cluster_id/master_cluster_name/master_vDisk_id/master_vDisk_name/aux_cluster
_id/aux_cluster_name/aux_vDisk_id/aux_vDisk_name/primary/consistency_group_id/consistency_g
roup_name/state/bg_copy_priority/progress
12/pprcmapaixlog/0000010030400364/REDSTONE1/12/c1_aix_log_src/0000010030C0012C/lodestoneA/1
//master/255/pprccongrpaix/inconsistent_stopped/50/0
13/pprcmapaixdata/0000010030400364/REDSTONE1/13/c1_aix_data_src/0000010030C0012C/lodestoneA
/2//master/255/pprccongrpaix/inconsistent_stopped/50/0

SVC1:admin>svcinfo lsrcrelationship pprcmapaixdata
id 13
name pprcmapaixdata
master_cluster_id 0000010030400364
master_cluster_name REDSTONE1
master_vDisk_id 13
master_vDisk_name c1_aix_data_src
aux_cluster_id 0000010030C0012C
aux_cluster_name lodestoneA
aux_vDisk_id 2
aux_vDisk_name
primary master
consistency_group_id 255
consistency_group_name pprccongrpaix
state inconsistent_stopped
bg_copy_priority 50
progress 0
freeze_time
status online
sync

>>>>>> the vDisks c1_aix_log_src and c1_aix_data_src are mapped to two PPRC relationships

SVC1:admin>svcinfo lsvdisk -delim /
id/name/IO_group_id/IO_group_name/status/mDisk_grp_id/mDisk_grp_name/capacity/type/FC_id/FC
_name/RC_id/RC_name
0/aix_vDisk0/0/ITSO_grp0/online/0/ARR36P5N/10.0GB/striped////
1/aix_vDisk1/0/ITSO_grp0/online/0/ARR36P5N/10.0GB/seq////
2/aix_vDisk2/0/ITSO_grp0/online/0/ARR36P5N/10.0GB/seq////
3/aix_vDisk3/0/ITSO_grp0/online/0/ARR36P5N/10.0GB/striped////
4/windisk1/0/ITSO_grp0/online/0/ARR36P5N/200.0MB/striped////
```

```
5/redhatdisk1/0/ITSO_grp0/online/0/ARR36P5N/300.0MB/striped////
6/aix_vDisk4/0/ITSO_grp0/online/0/ARR36P5N/1.0GB/striped////
7/rhat2_vDisk0/0/ITSO_grp0/online/1/ARR36P6N/500.0MB/striped////
8/aix_vDisk1_src/0/ITSO_grp0/online/1/ARR36P6N/500.0MB/striped/0/flashmapaix1//
9/aix_vDisk1_tgt/0/ITSO_grp0/online/2/ARR73P5N/500.0MB/striped/0/flashmapaix1//
10/SAP_DATA02/0/ITSO_grp0/online/2/ARR73P5N/500.0MB/striped////
11/SAP_DATA03/0/ITSO_grp0/online/2/ARR73P5N/500.0MB/seq////
12/c1_aix_log_src/0/ITSO_grp0/online/1/ARR36P6N/500.0MB/striped///12/pprcmapaixlog
13/c1_aix_data_src/0/ITSO_grp0/online/2/ARR73P5N/1.0GB/striped///13/pprcmapaixdata
```

Example 11-11 shows the CLI format to check PPRC relationships.

Example 11-11 Checking PPRC relationships pprcmapaixlog and pprcmapaixdata on SVC2

```
>>>>>>>>>> no PPRC relationship is created yet in the consistency group “pprccongrpaix”

SVC2:admin>svcinfo lsrcrelationship

>>>>>>>>>> The following vDisks are the vDisks candidate for a PPRC relationship on SVC2

SVC2:admin>svcinfo lsrcrelationshipcandidate
id               vDisk_name
0                svc2aix2test
1                c2_aix_log_tgt
2                c2_aix_data_tgt

>>> after the creation of PPRC relationships on SVC1

SVC2:admin>svcinfo lsrcconsistgrp -delim /
id/name/master_cluster_id/master_cluster_name/aux_cluster_id/aux_cluster_name/primary/state
/relationship_count
255/pprccongrpaix/0000010030400364/REDSTONE1/0000010030C0012C/lodestoneA/master/inconsisten
t_stopped/2

SVC2:admin>svcinfo lsrcconsistgrp pprccongrpaix
id 255
name pprccongrpaix
master_cluster_id 0000010030400364
master_cluster_name REDSTONE1
aux_cluster_id 0000010030C0012C
aux_cluster_name lodestoneA
primary master
state inconsistent_stopped
relationship_count 2
freeze_time
status online
sync
RC_rel_id 1
RC_rel_name pprcmapaixlog
RC_rel_id 2
RC_rel_name pprcmapaixdata

SVC2:admin>svcinfo lsrcrelationship -delim /
id/name/master_cluster_id/master_cluster_name/master_vDisk_id/master_vDisk_name/aux_cluster
_id/aux_cluster_name/aux_vDisk_id/aux_vDisk_name/primary/consistency_group_id/consistency_g
roup_name/state/bg_copy_priority/progress
```

```
1/pprcmapaixlog/0000010030400364/REDSTONE1/12//0000010030C0012C/lodestoneA/1/c2_aix_log_tgt
/master/255/pprccongrpaix/inconsistent_stopped/50/0
2/pprcmapaixdata/0000010030400364/REDSTONE1/13//0000010030C0012C/lodestoneA/2/c2_aix_data_t
gt/master/255/pprccongrpaix/inconsistent_stopped/50/0

SVC2:admin>svcinfo lsrcrelationship pprcmapaixlog
id 1
name pprcmapaixlog
master_cluster_id 0000010030400364
master_cluster_name REDSTONE1
master_vDisk_id 12
master_vDisk_name
aux_cluster_id 0000010030C0012C
aux_cluster_name lodestoneA
aux_vDisk_id 1
aux_vDisk_name c2_aix_log_tgt
primary master
consistency_group_id 255
consistency_group_name pprccongrpaix
state inconsistent_stopped
bg_copy_priority 50
progress 0
freeze_time
status online
sync


SVC2:admin>svcinfo lsrcrelationship pprcmapaixdata
id 2
name pprcmapaixdata
master_cluster_id 0000010030400364
master_cluster_name REDSTONE1
master_vDisk_id 13
master_vDisk_name
aux_cluster_id 0000010030C0012C
aux_cluster_name lodestoneA
aux_vDisk_id 2
aux_vDisk_name c2_aix_data_tgt
primary master
consistency_group_id 255
consistency_group_name pprccongrpaix
state inconsistent_stopped
bg_copy_priority 50
progress 0
freeze_time
status online
sync

SVC2:admin>svcinfo lsvdisk -delim /
id/name/IO_group_id/IO_group_name/status/mDisk_grp_id/mDisk_grp_name/capacity/type/FC_id/FC
_name/RC_id/RC_name
0/svc2aix2test/0/io_grp0/online/0/SVC2_F1_MDG1/100.0MB/striped////
1/c2_aix_log_tgt/0/io_grp0/offline/0/SVC2_F1_MDG1/500.0MB/striped///1/pprcmapaixlog
2/c2_aix_data_tgt/0/io_grp0/offline/0/SVC2_F1_MDG1/1.0GB/striped///2/pprcmapaixdata
```

Using the GUI

Create a PPRC relationship pprcmapaixlog between a master vDisk c1_aix_log_src and an auxiliary vDisk c2_aix_log_tgt which is in the remote cluster lodestoneA. The PPRC relationship is included in the PPRC consistency group pprccongrpaix. The background copy priority is set to 50.

Complete the following steps:

1. As shown in Figure 11-17, under My Work, select **Manage Copy Services** and select **Remote Copy Relationships**.
2. In the Remote Copy Relationship panel, select **Create a relationship** from the list and click **Go**.

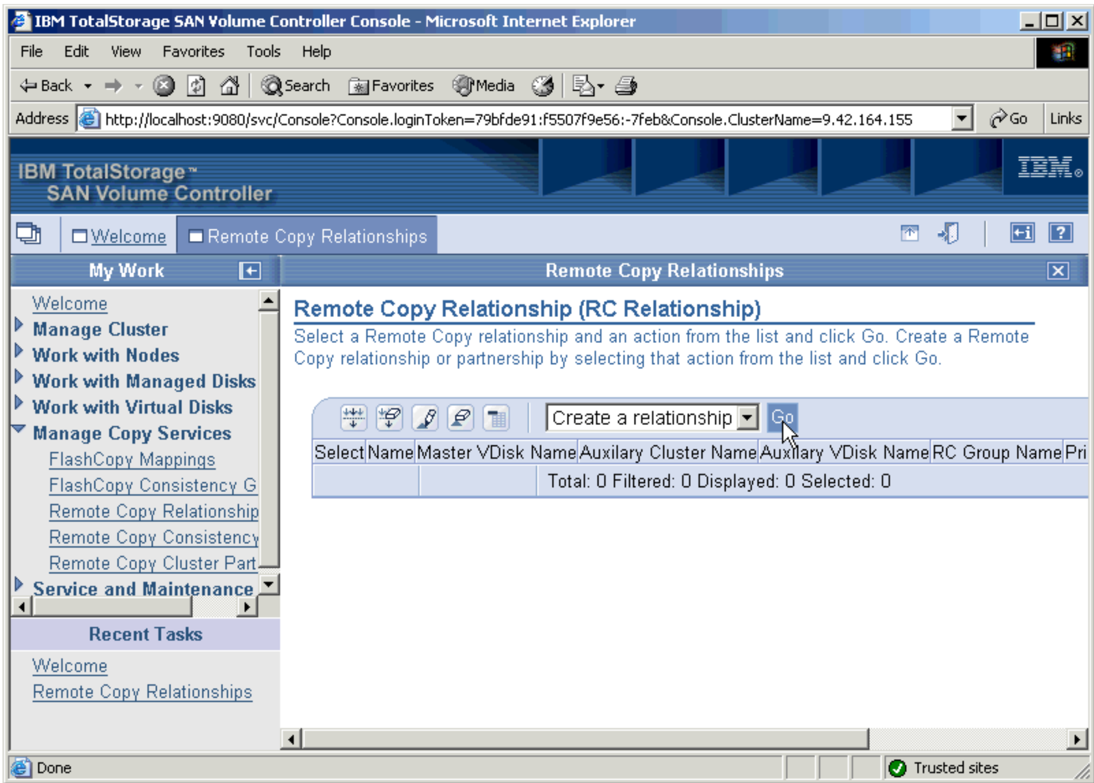


Figure 11-17 Remote Copy Relationship

3. The Remote Copy Relationships wizard (Figure 11-18) opens. Click **Next**.

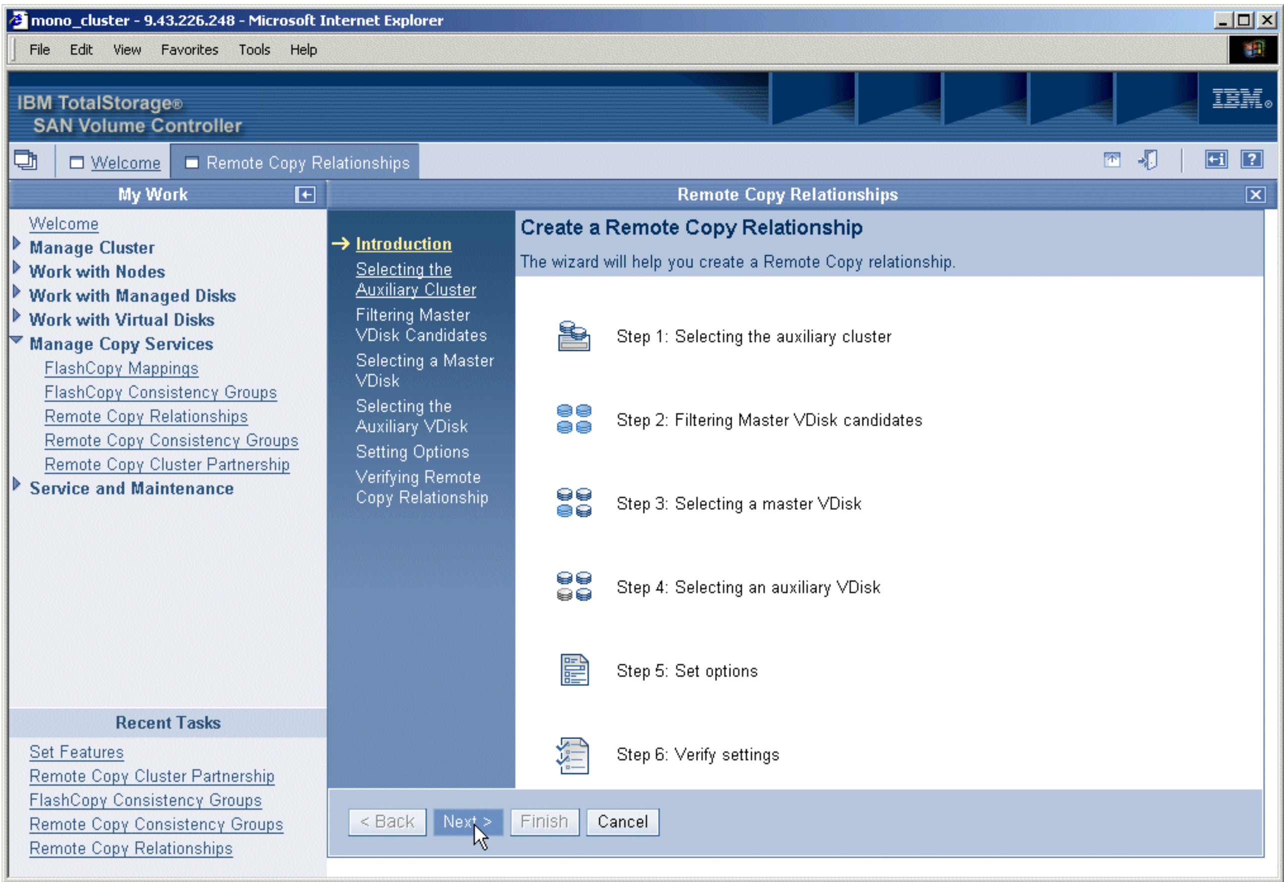


Figure 11-18 Create a Remote Copy Relationship wizard

4. On the Selecting the Auxiliary Cluster panel (Figure 11-9), follow these steps:
   a. Type the name of PPRC relationship, which in this example is `pprcmapaixlog`.
   b. Select the first option **Inter-cluster relationship between local cluster (mono_cluster) and remote cluster (risk_cluster), the local cluster will be the master cluster and the remote cluster will be the auxiliary cluster**.
   c. Click **Next**.

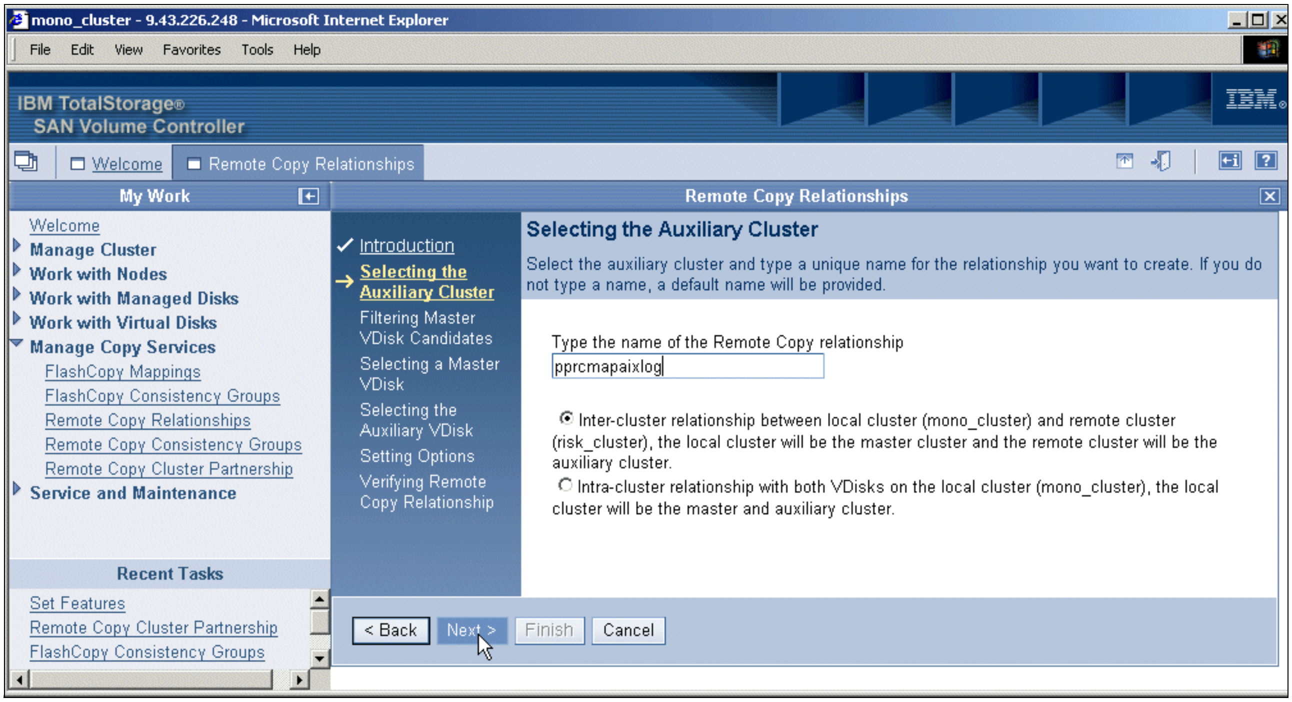


Figure 11-19 Selecting the Auxiliary Cluster

5. On the Filtering Master vDisk Candidates panel (Figure 11-20), filter the list of vDisk candidates. Click **Next**.

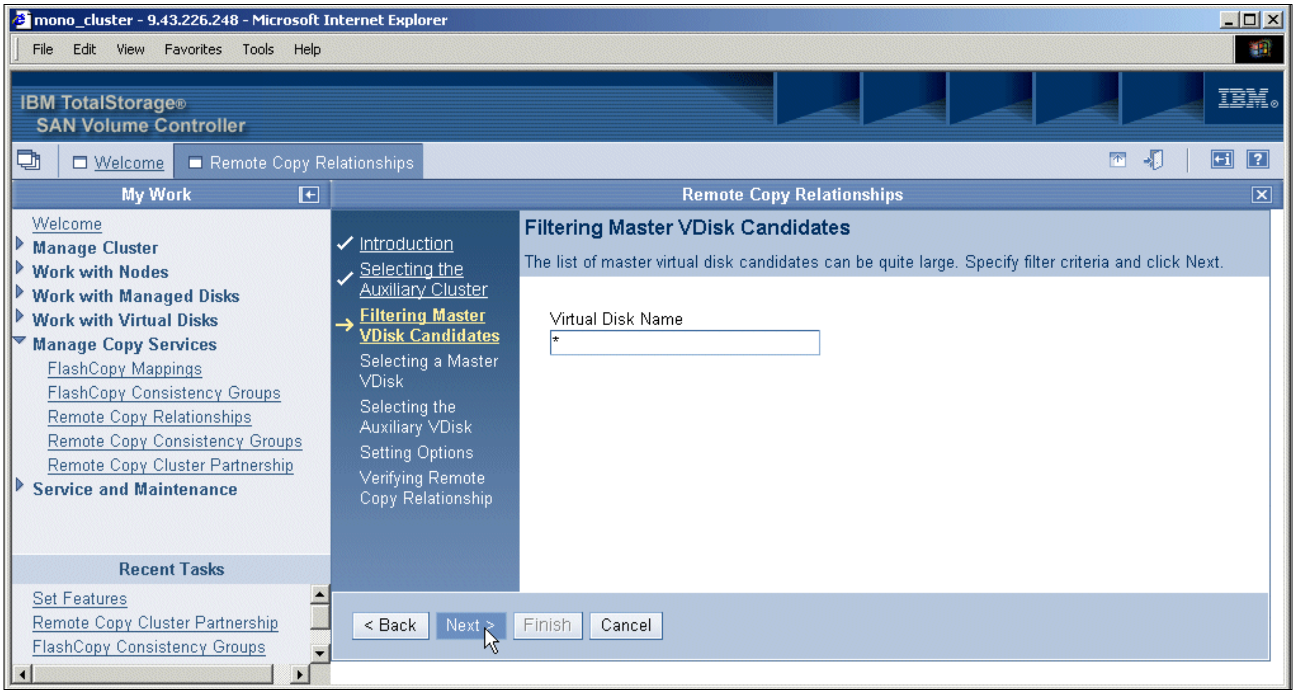


Figure 11-20 Filtering Master vDisk Candidates

6. Select the disk that you want to use from the list as shown in Figure 11-21. Click **Next**.

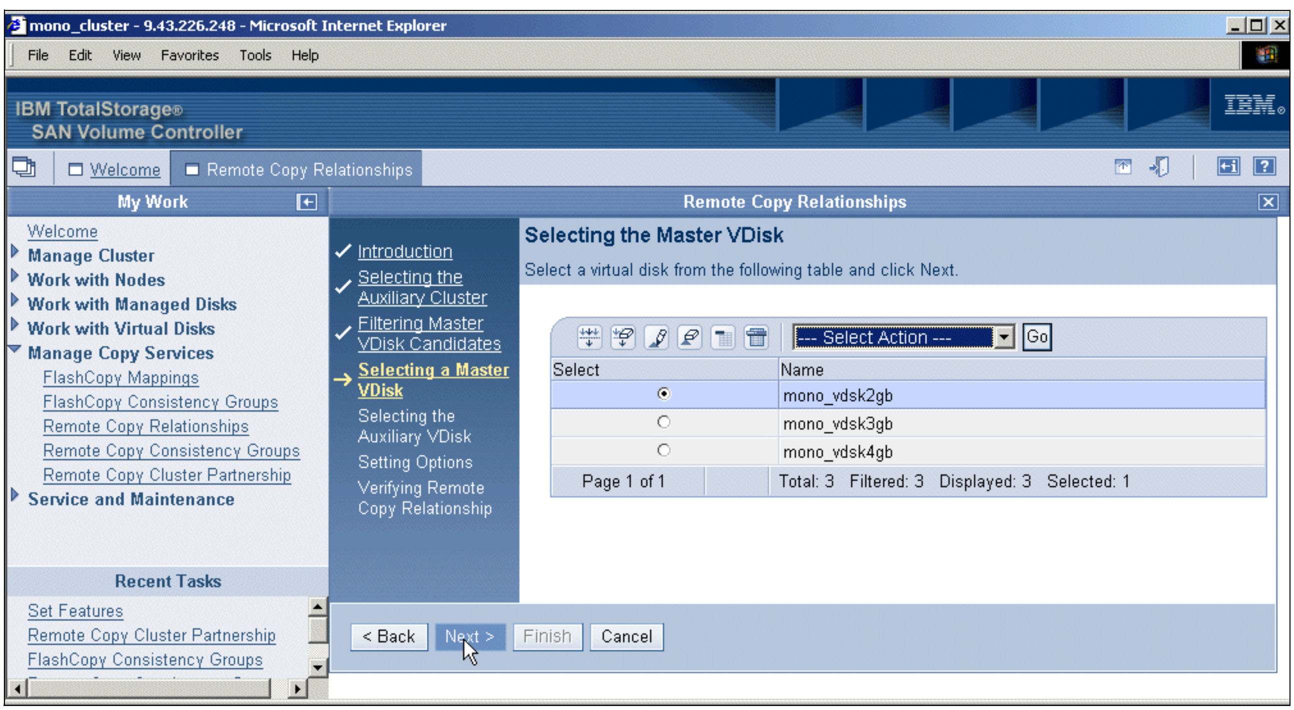


Figure 11-21 Setting the Master vDisk

7. Select the auxiliary disk as shown in Figure 11-22.

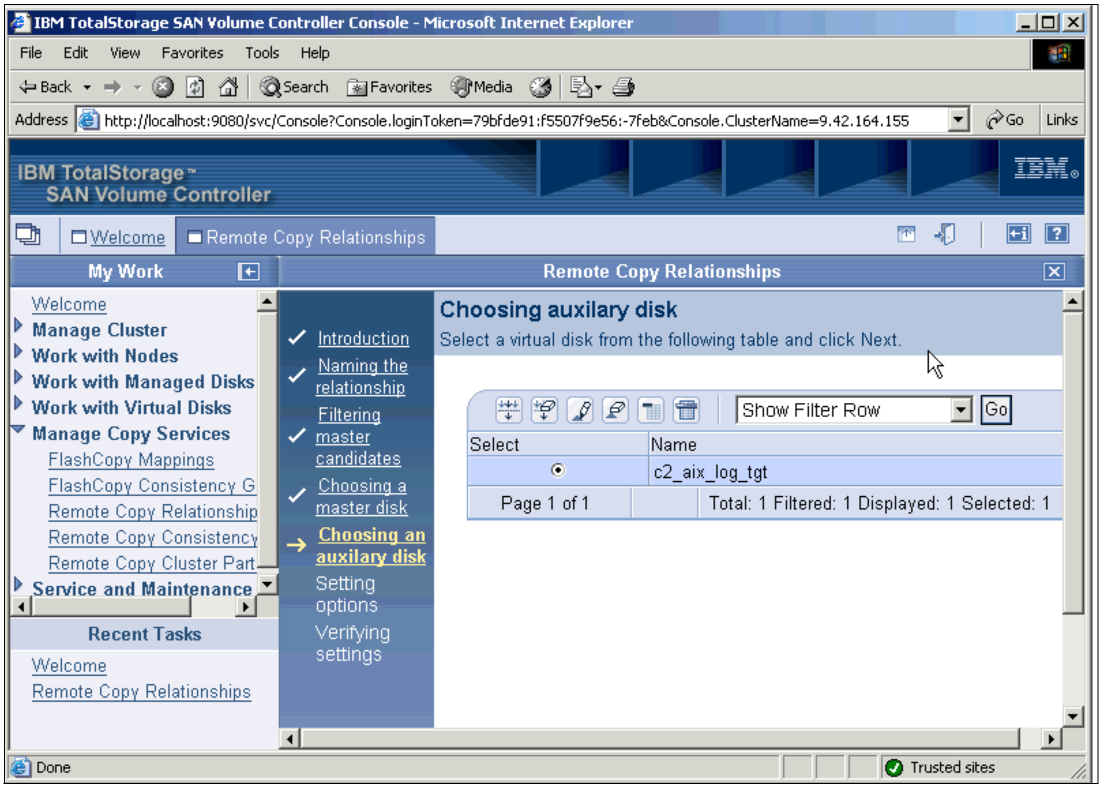


Figure 11-22 Choosing auxiliary disk

8. As shown in Figure 11-23, choose the consistency group and the priority. Select the **Synchronized** check box only if the master and auxiliary are already synchronized.

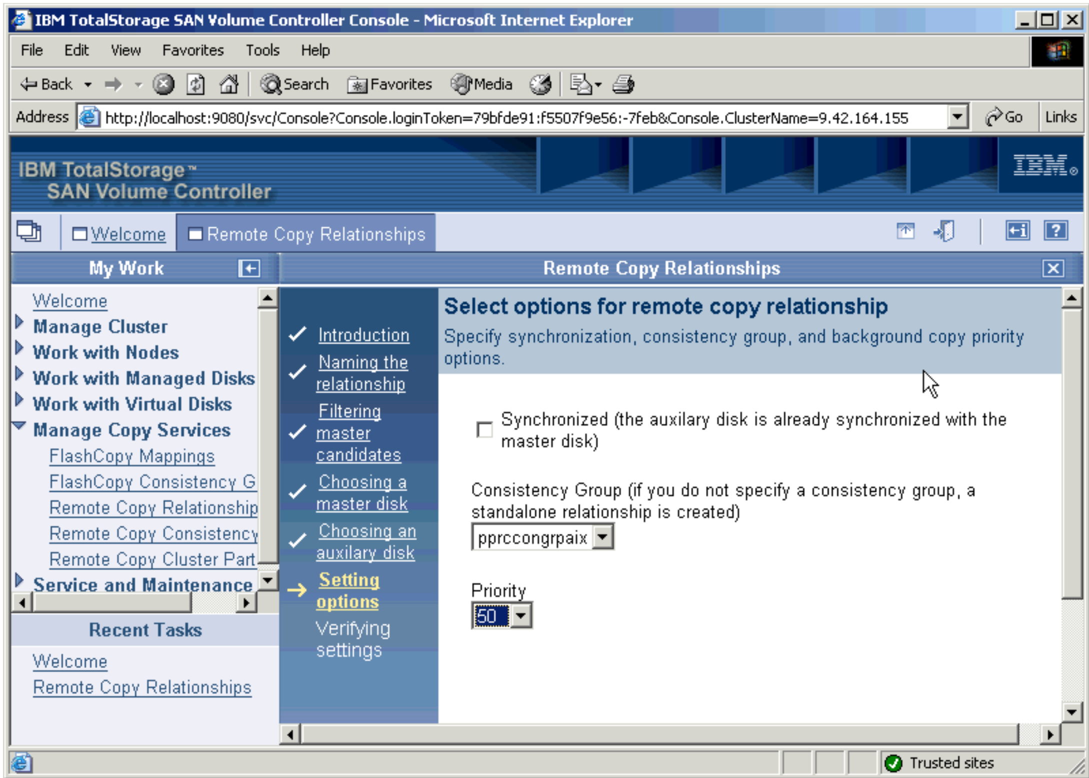


Figure 11-23 Select options for remote copy relationship

9. On the Verify remote copy relationship panel (Figure 11-24), verify all the settings that you have chosen.

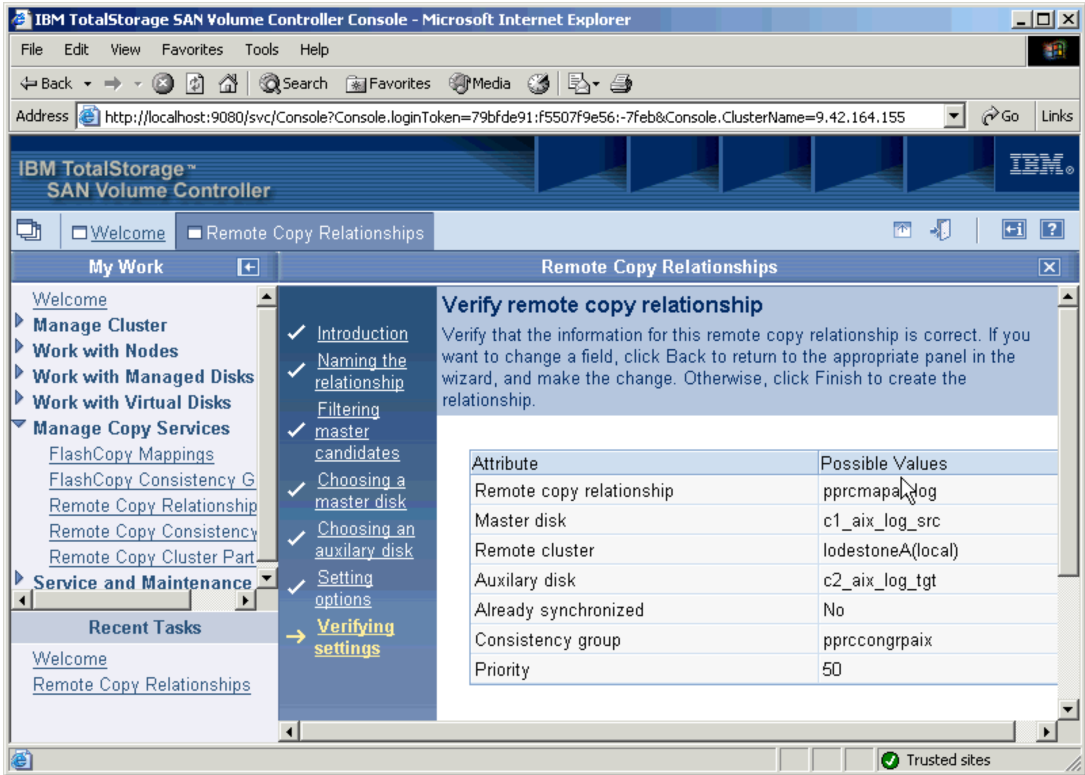


Figure 11-24 Verify remote copy relationship

The panel shown in Figure 11-25 is where the relationship is actually created.

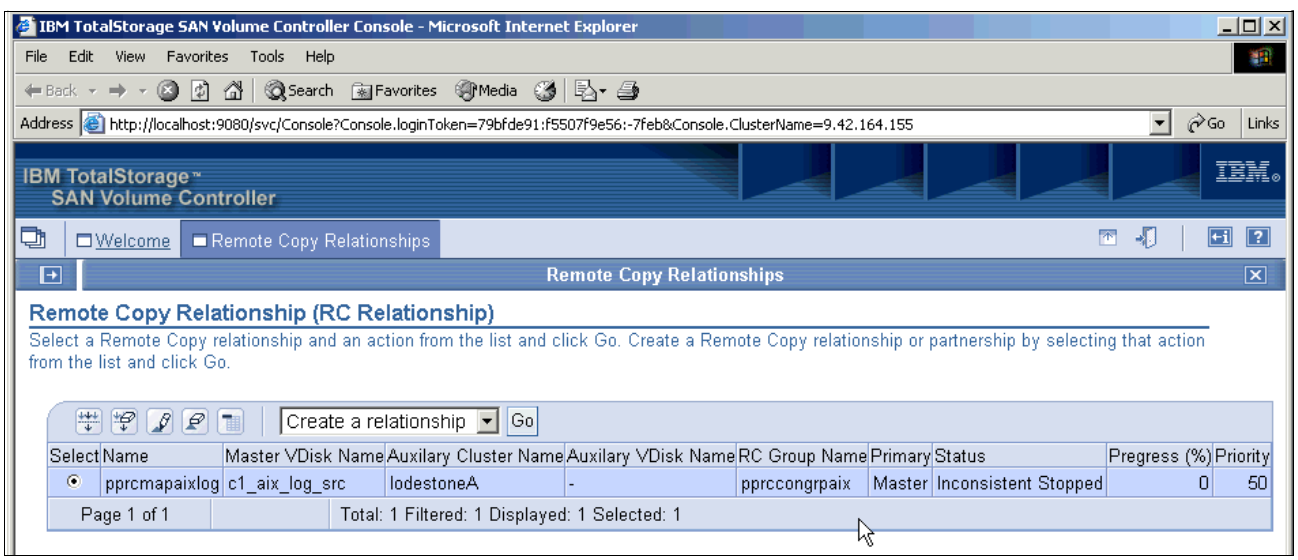


Figure 11-25 Creation of PPRC relationship pprcmapaixlog on SVC1

You can check that, on the remote cluster SVC2, the consistency group pprccongrpaix was also created and updated from the empty state to the inconsistent-stopped state. See Figure 11-26.

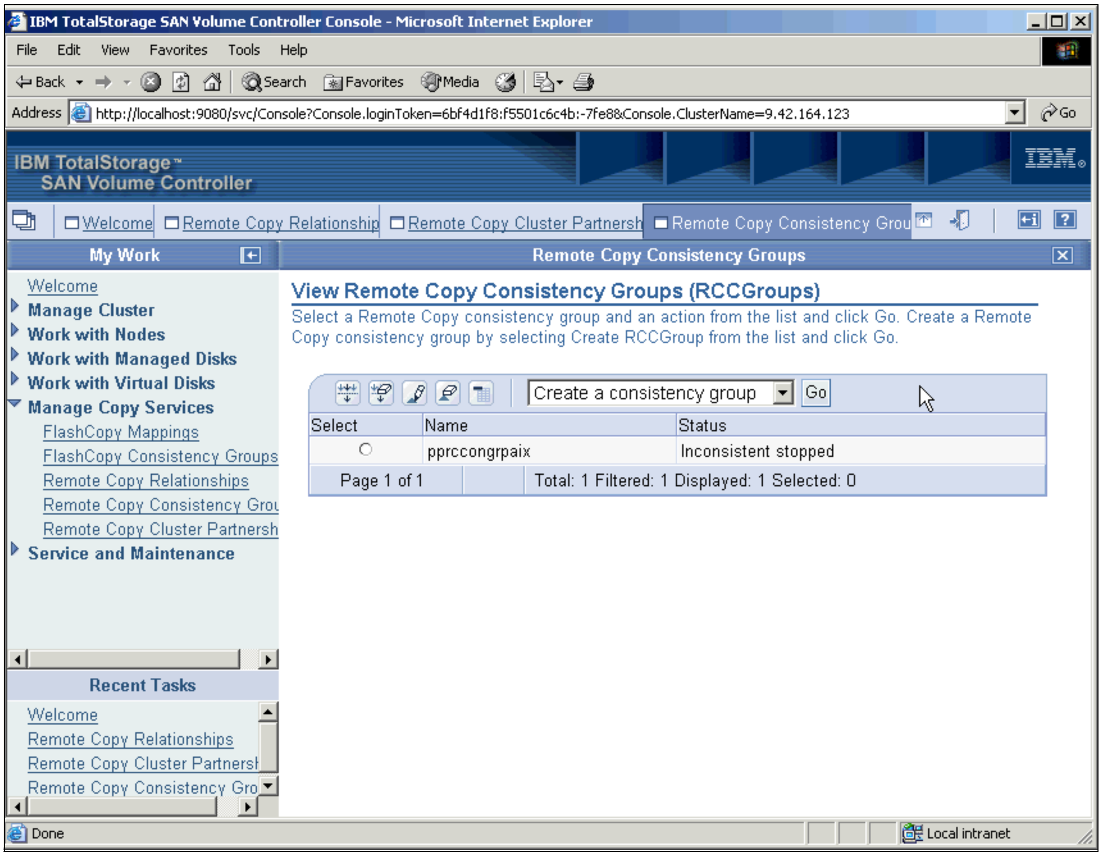


Figure 11-26 Checking PPRC consistency group pprccongrpaix on SVC2

On SVC2, you can check that two PPRC relationships were created and that they are both in the inconsistent-stopped state. This is shown in Figure 11-27.

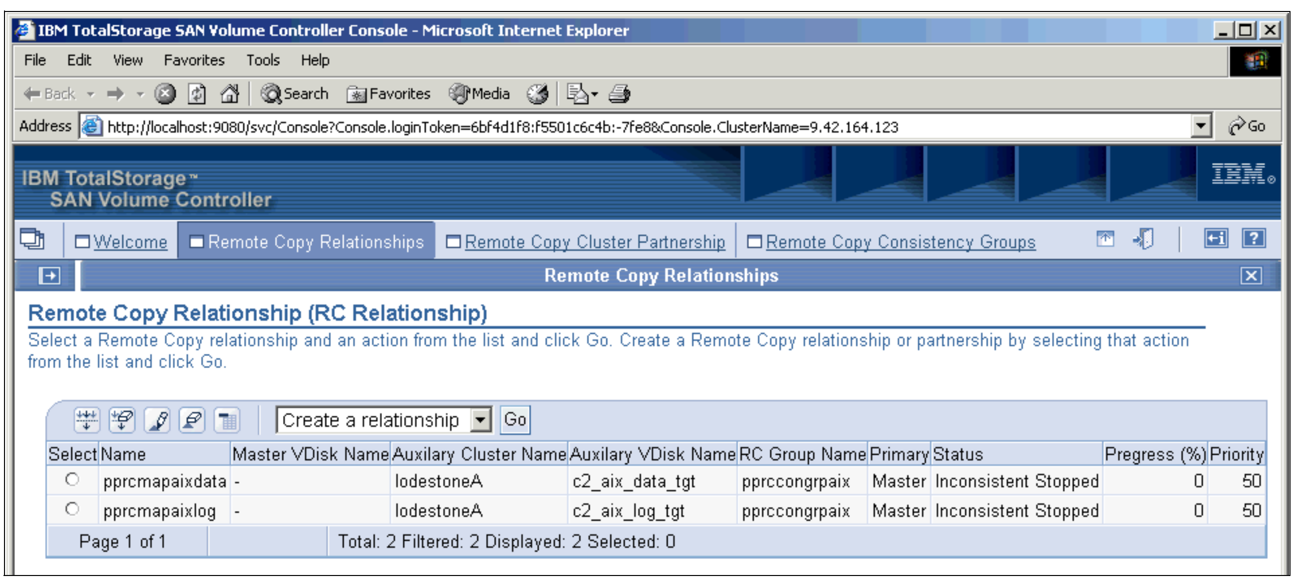


Figure 11-27 Checking of PPRC relationships pprcmapaixlog and pprcmapaixdata on SVC2

11.3.6 Starting the PPRC consistency group

We have the clusters joined by a partnership and the virtual disks joined by a relationship. Now we can start the PPRC consistency group.

Using the CLI

To start PPRC consistency group pprccongrpaix from the master cluster, which is also the primary vDisks for the synchronization, enter:

```
svctask startrcconsistgrp pprccongrpaix
```

To check on the progress of the remote copy synchronization process, you can look at the progress of the PPRC relationship of the consistency group:

```
svcinfo lsrcrelationshipprogress pprcmapaixdata
svcinfo lsrcrelationshipprogress pprcmapaixlog
```

This is shown in Example 11-2.

Example 11-12 Starting PPRC consistency group pprccongrpaix

```
SVC1:admin>svctask startrcconsistgrp pprccongrpaix

>>>>>>> the state will be “inconsistent_copying” until the complete synchronization

SVC1:admin>svcinfo lsrcconsistgrp pprccongrpaix
id 255
name pprccongrpaix
master_cluster_id 0000010030400364
master_cluster_name REDSTONE1
aux_cluster_id 0000010030C0012C
aux_cluster_name lodestoneA
primary master
state inconsistent_copying
relationship_count 2
freeze_time
status online
sync
RC_rel_id 12
RC_rel_name pprcmapaixlog
RC_rel_id 13
RC_rel_name pprcmapaixdata

>>>>>>>>>>> during the synchronization process

SVC1:admin>svcinfo lsrcrelationshipprogress pprcmapaixdata
id              progress
13              63
SVC1:admin>svcinfo lsrcrelationshipprogress pprcmapaixlog
id              progress
12              100

>>>>>>>>>> When the source (primary) and target (secondary) are synchronized

SVC1:admin>svcinfo lsrcrelationshipprogress pprcmapaixdata
id              progress
13
SVC1:admin>svcinfo lsrcrelationshipprogress pprcmapaixlog
id              progress
```

```
12

>>>>>>>>>> At the end the consistency group is consistent_synchronised

SVC1:admin>svcinfo lsrcconsistgrp  -delim /
id/name/master_cluster_id/master_cluster_name/aux_cluster_id/aux_cluster_name/primary/state
/relationship_count
255/pprccongrpaix/0000010030400364/REDSTONE1/0000010030C0012C/lodestoneA/master/consistent_
synchronised/2

SVC1:admin>svcinfo lsrcconsistgrp pprccongrpaix
id 255
name pprccongrpaix
master_cluster_id 0000010030400364
master_cluster_name REDSTONE1
aux_cluster_id 0000010030C0012C
aux_cluster_name lodestoneA
primary master
state consistent_synchronized
relationship_count 2
freeze_time
status online
sync
RC_rel_id 12
RC_rel_name pprcmapaixlog
RC_rel_id 13
RC_rel_name pprcmapaixdata

SVC1:admin>svcinfo lsrcrelationship pprcmapaixdata
id 13
name pprcmapaixdata
master_cluster_id 0000010030400364
master_cluster_name REDSTONE1
master_vDisk_id 13
master_vDisk_name c1_aix_data_src
aux_cluster_id 0000010030C0012C
aux_cluster_name lodestoneA
aux_vDisk_id 2
aux_vDisk_name
primary master
consistency_group_id 255
consistency_group_name pprccongrpaix
state consistent_synchronised
bg_copy_priority 50
progress
freeze_time
status online
sync

SVC1:admin>svcinfo lsrcrelationship pprcmapaixlog
id 12
name pprcmapaixlog
master_cluster_id 0000010030400364
master_cluster_name REDSTONE1
master_vDisk_id 12
master_vDisk_name c1_aix_log_src
aux_cluster_id 0000010030C0012C
aux_cluster_name lodestoneA
aux_vDisk_id 1
aux_vDisk_name
```

```
primary master
consistency_group_id 255
consistency_group_name pprccongrpaix
state consistent_synchronized
bg_copy_priority 50
progress
freeze_time
status online
sync
```

Using the GUI

Follow these steps:

1. As shown in Figure 11-28, under My Work, select **Manage Copy Services** and select **Remote Copy Consistency Groups**.
2. On the View Remote Copy Consistency Groups, under the Select column, select the remote copy consistency group you want to start.
3. Select **Start a Remote Copy** from the list and click **Go**.

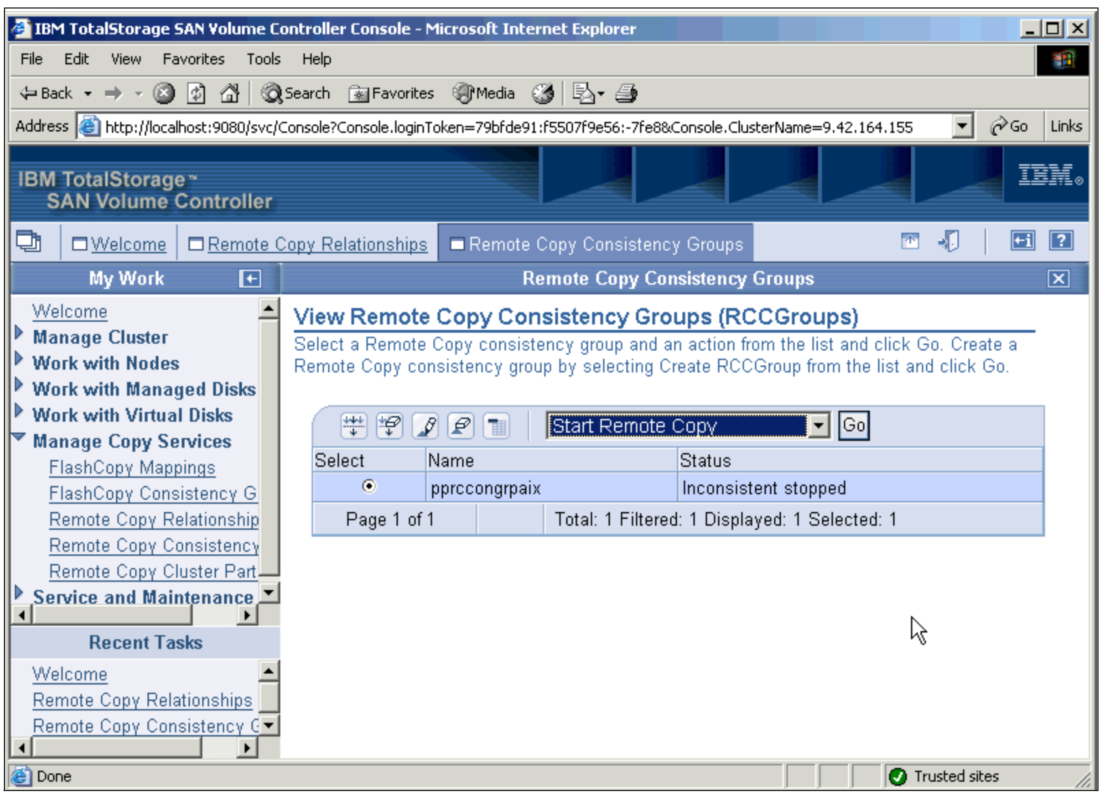


Figure 11-28 Starting PPRC consistency group pprccongrpaix

4. You see the Starting Remote Copy Consistency Group pprccongrpaix panel (Figure 11-29). In this example, under Copy Direction, we select **Do not set or change the copy direction** and click **OK**.

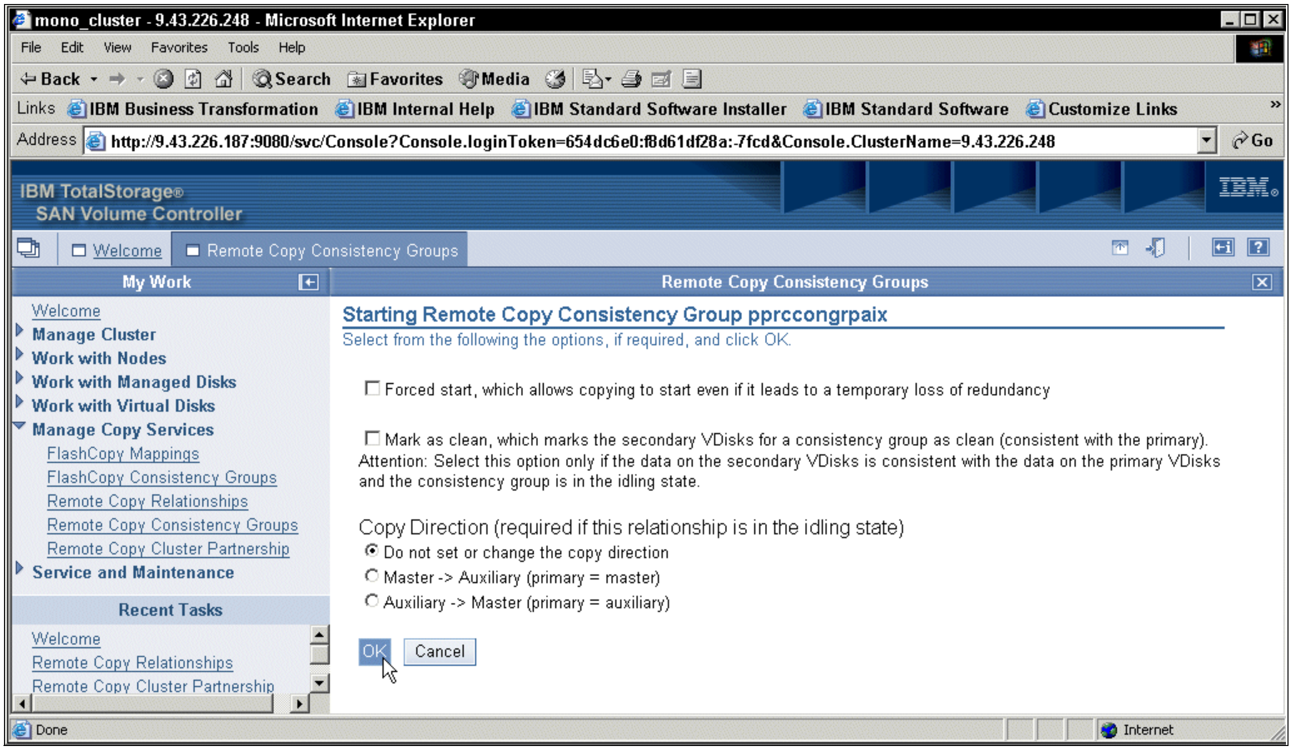


Figure 11-29 Starting Remote Copy Consistency Group pprccongrpaix

The Forced Copy flag is not required if the consistency is not lost by starting a copy operation. This is the case since it is the first time that remote copy synchronization is performed. See Figure 11-30.

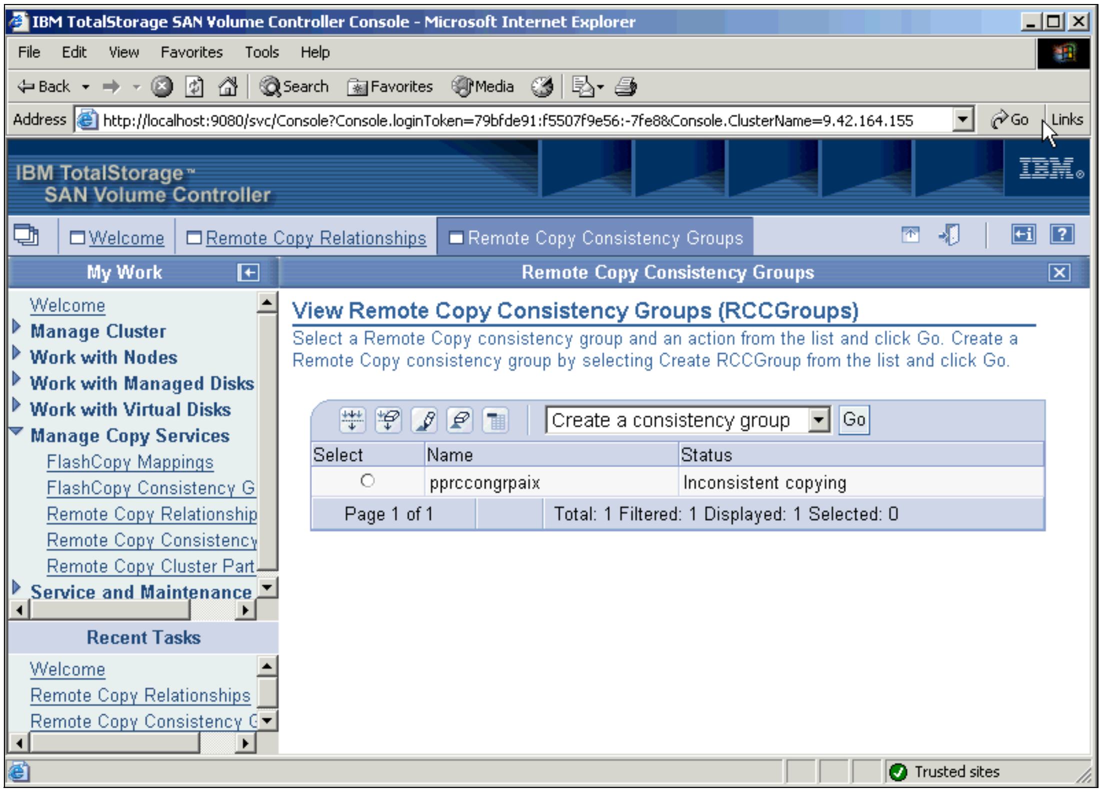


Figure 11-30 View Remote Copy Consistency Groups: Inconsistent copying status

When the copy is consistent, you see the results as shown in Figure 11-31.

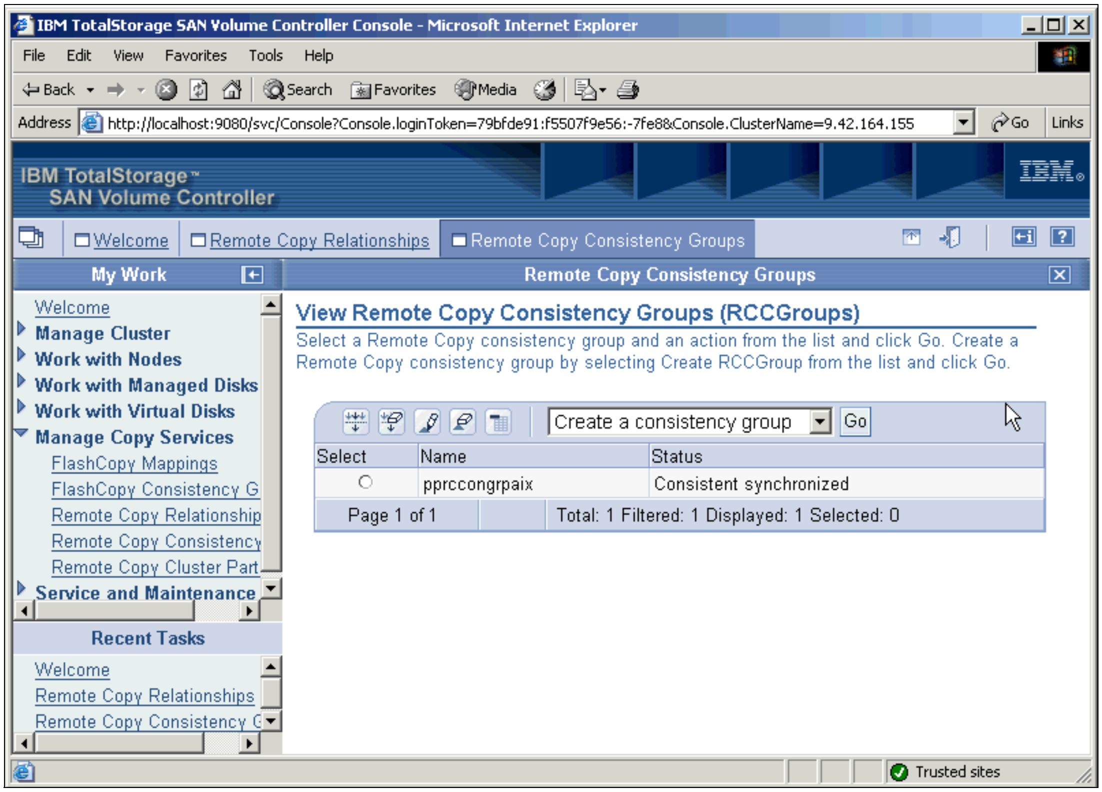


Figure 11-31 View Remote Copy Consistency Groups: Consistent synchronized status

The panel in Figure 11-32 shows the copy relationships.

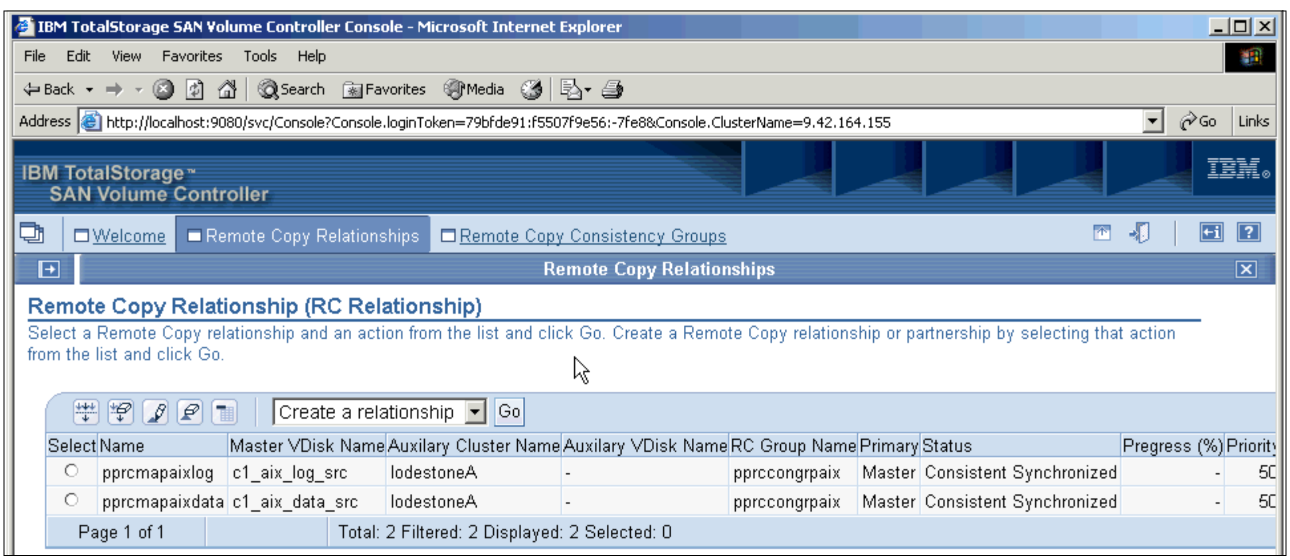


Figure 11-32 Remote Copy Relationships

You can check that the PPRC consistency group and the PPRC relationships are also synchronized on SVC2. See Figure 11-33.

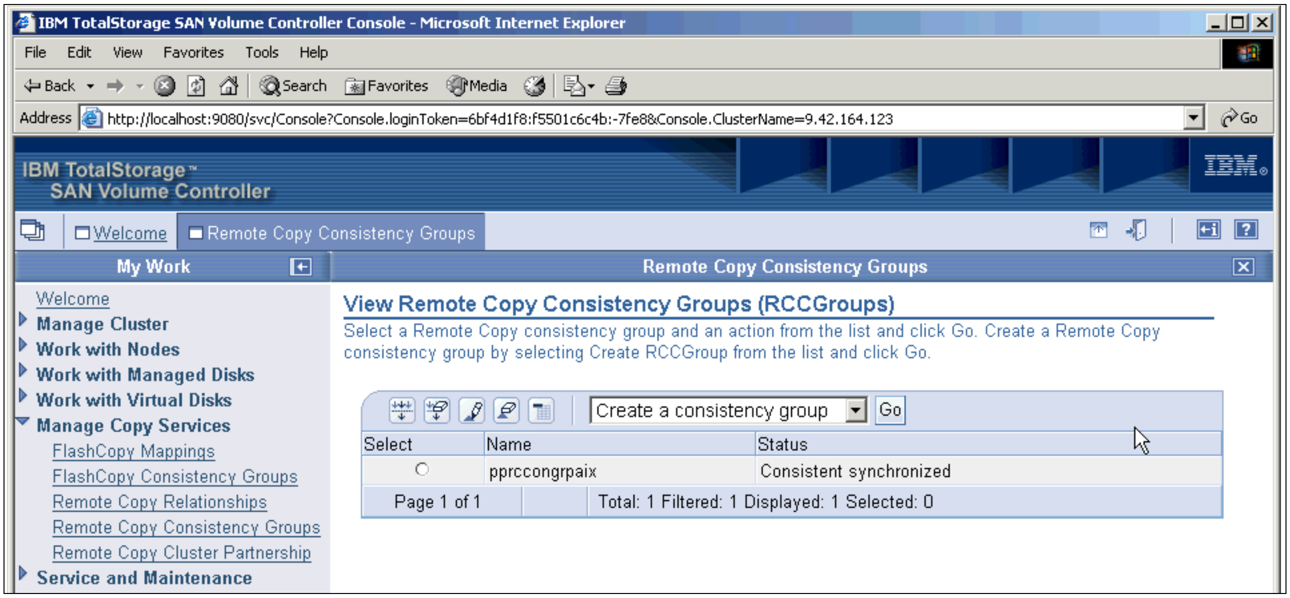


Figure 11-33 view Remote Copy Consistency Groups

The PPRC state is now synchronized as shown in Figure 11-34.

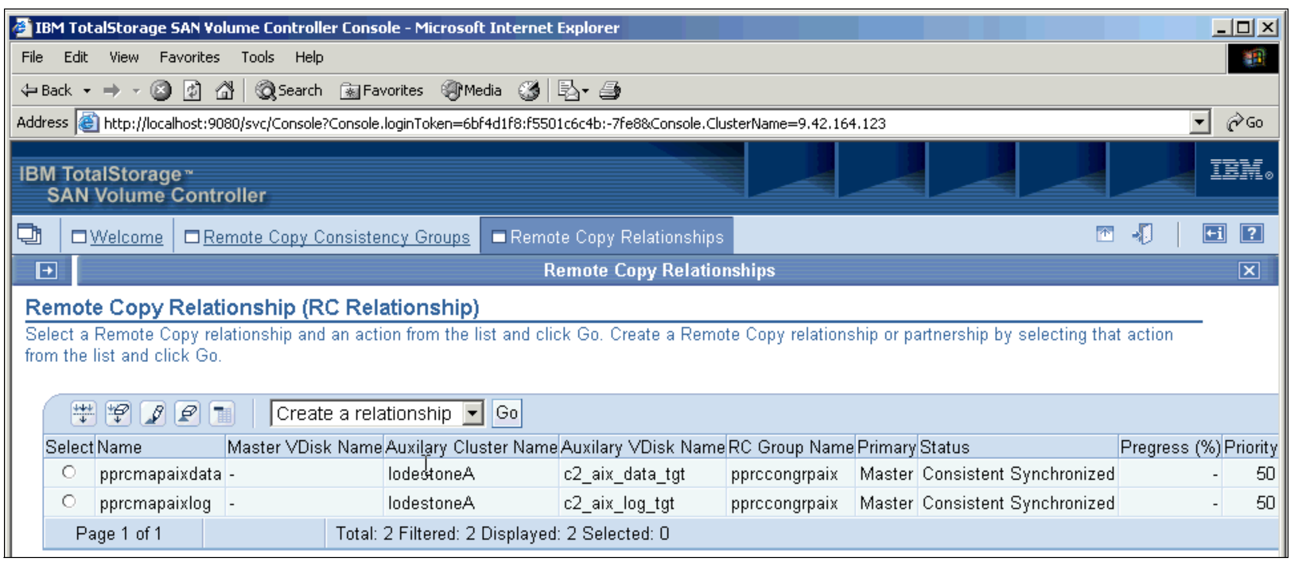


Figure 11-34 PPRC relationships synchronized on SVC2

11.3.7 Stopping the PPRC consistency group

You need to stop the PPRC consistency group using the access option to use the target or auxiliary virtual disks on a remote host system as aix2.

Using the CLI

To stop the PPRC consistency group pprccongrpaix, you can issue the following command:

```
svctask stoprcconsistgrp -access pprccongrpaix
```

If you do not specify the access option, the PPRC relationships will be in the consistent-stopped state with no write access to the secondary vDisks. If you specify the option access, the PPRC relationships will be in the idling state with write access to the secondary vDisks. This is shown in Example 11-13.

Example 11-13 Stopping a PPRC consistency group pprccongrpaix

```
SVC1:admin>svctask stoprcconsistgrp -access pprccongrpaix

SVC1:admin>svcinfo lsrcconsistgrp  -delim /
id/name/master_cluster_id/master_cluster_name/aux_cluster_id/aux_cluster_name/primary/state
/relationship_count
255/pprccongrpaix/0000010030400364/REDSTONE1/0000010030C0012C/lodestoneA//idling/2

SVC1:admin>svcinfo lsrcconsistgrp  pprccongrpaix
id 255
name pprccongrpaix
master_cluster_id 0000010030400364
master_cluster_name REDSTONE1
aux_cluster_id 0000010030C0012C
aux_cluster_name lodestoneA
primary
state idling
relationship_count 2
freeze_time
status
sync in_sync
RC_rel_id 12
RC_rel_name pprcmapaixlog
RC_rel_id 13
RC_rel_name pprcmapaixdata

SVC1:admin>svcinfo lsrcrelationship -delim /
id/name/master_cluster_id/master_cluster_name/master_vDisk_id/master_vDisk_name/aux_cluster
_id/aux_cluster_name/aux_vDisk_id/aux_vDisk_name/primary/consistency_group_id/consistency_g
roup_name/state/bg_copy_priority/progress
12/pprcmapaixlog/0000010030400364/REDSTONE1/12/c1_aix_log_src/0000010030C0012C/lodestoneA/1
///255/pprccongrpaix/idling/50/
13/pprcmapaixdata/0000010030400364/REDSTONE1/13/c1_aix_data_src/0000010030C0012C/lodestoneA
/2///255/pprccongrpaix/idling/50/

SVC1:admin>svcinfo lsrcrelationship pprcmapaixdata
id 13
name pprcmapaixdata
master_cluster_id 0000010030400364
master_cluster_name REDSTONE1
master_vDisk_id 13
master_vDisk_name c1_aix_data_src
aux_cluster_id 0000010030C0012C
aux_cluster_name lodestoneA
aux_vDisk_id 2
aux_vDisk_name
primary
consistency_group_id 255
consistency_group_name pprccongrpaix
state idling
bg_copy_priority 50
progress
```

```
freeze_time
status
sync in_sync

SVC1:admin>svcinfo lsrcrelationship pprcmapaixlog
id 12
name pprcmapaixlog
master_cluster_id 0000010030400364
master_cluster_name REDSTONE1
master_vDisk_id 12
master_vDisk_name c1_aix_log_src
aux_cluster_id 0000010030C0012C
aux_cluster_name lodestoneA
aux_vDisk_id 1
aux_vDisk_name
primary
consistency_group_id 255
consistency_group_name pprccongrpaix
state idling
bg_copy_priority 50
progress
freeze_time
status
sync in_sync

>>>>>>>>>>>> If you stop without the -access option

SVC1:admin>svctask stoprcconsistgrp pprccongrpaix

SVC1:admin>svcinfo lsrcconsistgrp  pprccongrpaix
id 255
name pprccongrpaix
master_cluster_id 0000010030400364
master_cluster_name REDSTONE1
aux_cluster_id 0000010030C0012C
aux_cluster_name lodestoneA
primary master
state consistent_stopped
relationship_count 2
freeze_time 2003/05/25/21/29/12
status online
sync in_sync
RC_rel_id 12
RC_rel_name pprcmapaixlog
RC_rel_id 13
RC_rel_name pprcmapaixdata
```

Using the GUI

Complete the following steps:

1. As shown in Figure 11-35, under My Work, select **Manage Copy Services** and select **Remote Copy Consistency Groups**.
2. In the View Remote Copy Consistency Groups panel, under the Select column, select the remote copy consistency group you want to start.
3. Select **Stop a Remote Copy** from the list and click **Go**.

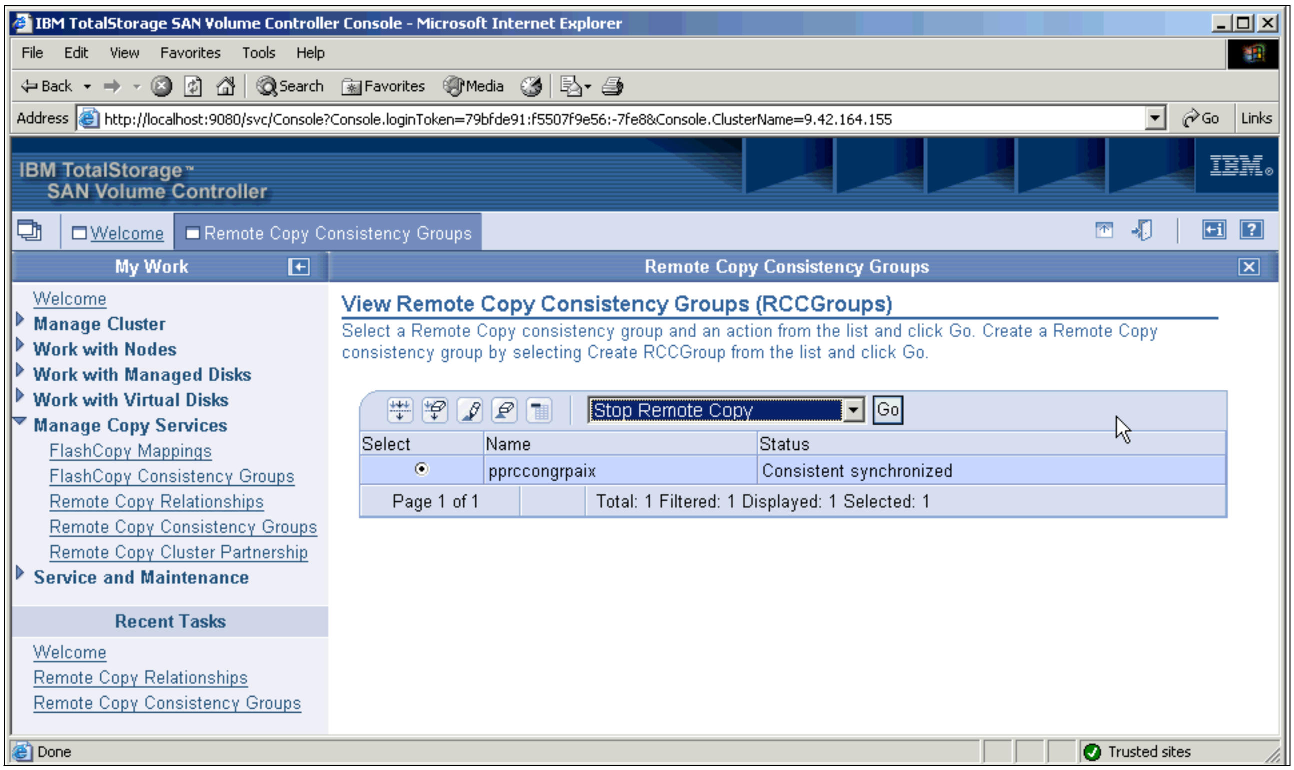


Figure 11-35 Stopping the consistency group pprccongrpaix

4. On the Stopping Remote Copy Consistency Group pprccongrpaix panel (Figure 11-36), we select the **Enable write to the secondary vDisks...** option and click **OK**.

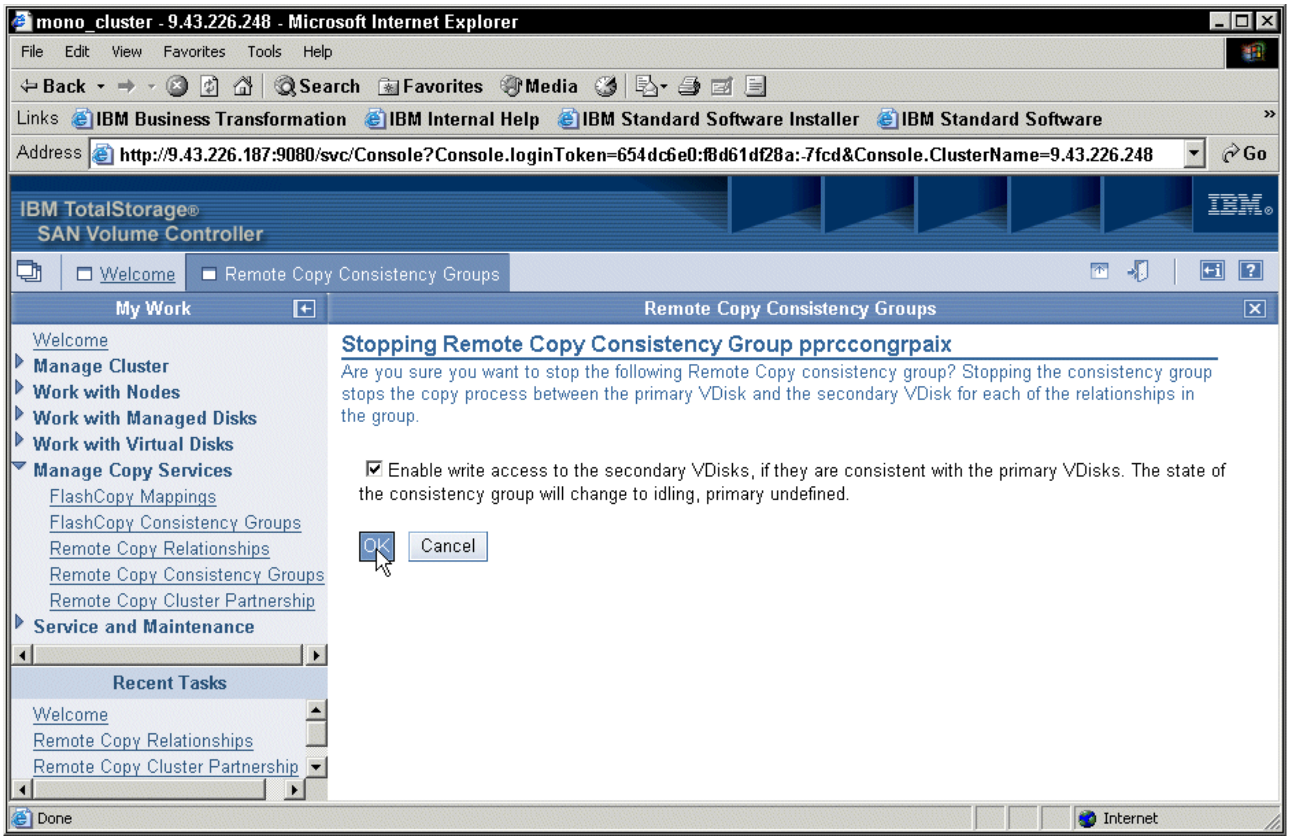


Figure 11-36 Stopping Remote Copy Consistency Group pprccongrpaix

The status of the PPRC consistency group is *idling* as shown in Figure 11-37.

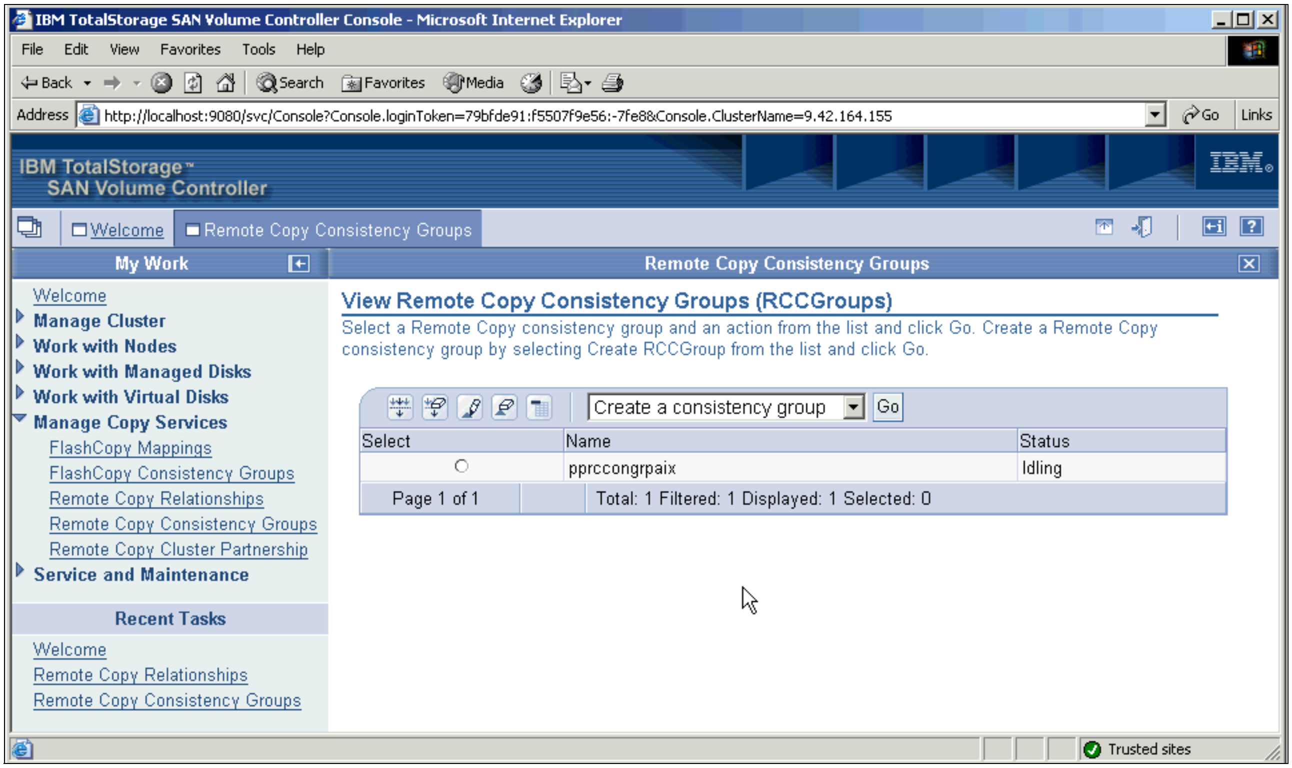


Figure 11-37 PPRC consistency group pprccongrpaix in idling state

The PPRC relationships within the consistency group are, as expected, also *idling* as shown in Figure 11-38.

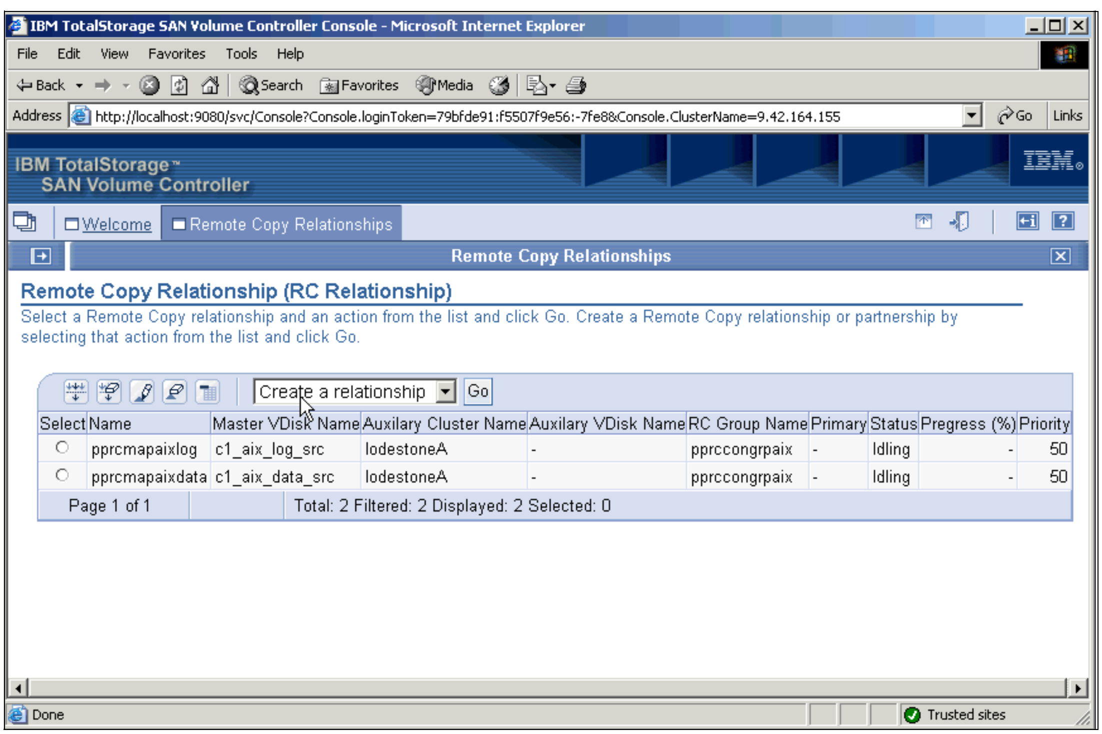


Figure 11-38 PPRC relationships in idling state

11.3.8 Using PPRC target vDisk on the host system

The secondary vDisks disks c2_aix_log_tgt and c2_aix_data_tgt are now available in the remote cluster. They are mapped to the aix2 host system.

Example 11-14 Mapping of PPRC secondary vDisks c2_aix_log_tgt and c2_aix_data_tgt to aix2

```
SVC2:admin>svcinfo lsvdisk -delim /
id/name/IO_group_id/IO_group_name/status/mDisk_grp_id/mDisk_grp_name/capacity/type/FC_id/FC
_name/RC_id/RC_name
0/svc2aix2test/0/io_grp0/online/0/SVC2_F1_MDG1/100.0MB/striped////
1/c2_aix_log_tgt/0/io_grp0/online/0/SVC2_F1_MDG1/500.0MB/striped///1/pprcmapaixlog
2/c2_aix_data_tgt/0/io_grp0/online/0/SVC2_F1_MDG1/1.0GB/striped///2/pprcmapaixdata

>>>>>>>>>>>>>>>> now we map the vDisks to aix2 in order to access them

SVC2:admin>svctask mkvdiskhostmap -host aix2 c2_aix_log_tgt
Host LUN to Host map, id [1], successfully created
SVC2:admin>svctask mkvdiskhostmap -host aix2 c2_aix_data_tgt
Host LUN to Host map, id [2], successfully created

SVC2:admin>svcinfo lshostvdiskmap aix2 -delim /
id/name/SCSI_id/vDisk_id/vDisk_name/wwpn/vDisk_UID
0/aix2/0/0/svc2aix2test/10000000C925841D/600507680186000960000000000000000
0/aix2/1/1/c2_aix_log_tgt/10000000C925841D/600507680186000960000000000000001
0/aix2/2/2/c2_aix_data_tgt/10000000C925841D/600507680186000960000000000000002
```

Now you can use the vDisks c2_aix_log_tgt and c2_aix_data_tgt on aix2, as shown in Example 11-15.

Example 11-15 Volume definition on aix2 after PPRC

```
AIX2#cfgmgr

AIX2#lspv
hdisk0        000b21dd2cf9b83d    rootvg
hdisk1        none                None
hdisk2        none                None
vpath0        000b21dd56751ac9    testvg
hdisk3        000790585b9d226d    None
hdisk4        000790585ba01755    None
hdisk5        000790585b9d226d    None
hdisk6        000790585ba01755    None
vpath1        none                None
vpath2        none                None

AIX2#lsdev -Cc disk
hdisk0 Available 10-80-00-2,0 16 Bit SCSI Disk Drive
hdisk1 Available 10-90-01     IBM TotalStorage SAN Volume Controller device
hdisk2 Available 10-90-01     IBM TotalStorage SAN Volume Controller device
vpath0 Available              Data Path Optimizer Pseudo Device Driver
hdisk3 Available 10-90-01     IBM TotalStorage SAN Volume Controller device
hdisk4 Available 10-90-01     IBM TotalStorage SAN Volume Controller device
hdisk5 Available 10-90-01     IBM TotalStorage SAN Volume Controller device
hdisk6 Available 10-90-01     IBM TotalStorage SAN Volume Controller device
vpath1 Available              Data Path Optimizer Pseudo Device Driver
vpath2 Available              Data Path Optimizer Pseudo Device Driver
```

```
AIX2#importvg -y pprcaix1logvg hdisk3
pprcaix1logvg
AIX2#importvg -y pprcaix1datavg hdisk4
pprcaix1datavg

AIX2#lspv
hdisk0         000b21dd2cf9b83d    rootvg
hdisk1         none                None
hdisk2         none                None
vpath0         000b21dd56751ac9    testvg
hdisk3         000790585b9d226d    pprcaix1logvg
hdisk4         000790585ba01755    pprcaix1datavg
hdisk5         000790585b9d226d    pprcaix1logvg
hdisk6         000790585ba01755    pprcaix1datavg
vpath1         none                None
vpath2         none                None

AIX2#hd2vp pprcaix1logvg
pprcaix1logvg
pprcaix1logvg contains these hdisks and pvids
   000790585b9d226d hdisk3
hdisk3 is part of vpath1.
pprcaix1logvg
Deleting pvid of hdisk3 from ODM database
Create vpath1 pvid of 000790585b9d226d in ODM database.
lname=vpath1, pvid=000790585b9d226d
Deleting pvid of hdisk5 from ODM database
varyon pprcaix1logvg was successful.
pprcaix1logvg is converted to dpo device vpathes successfully!

AIX2#hd2vp pprcaix1datavg
pprcaix1datavg
pprcaix1datavg contains these hdisks and pvids
   000790585ba01755 hdisk4
hdisk4 is part of vpath2.
pprcaix1datavg
Deleting pvid of hdisk4 from ODM database
Create vpath2 pvid of 000790585ba01755 in ODM database.
lname=vpath2, pvid=000790585ba01755
Deleting pvid of hdisk6 from ODM database
varyon pprcaix1datavg was successful.
pprcaix1datavg is converted to dpo device vpathes successfully!

AIX2#lspv
hdisk0         000b21dd2cf9b83d    rootvg
hdisk1         none                None
hdisk2         none                None
vpath0         000b21dd56751ac9    testvg
hdisk3         none                None
hdisk4         none                None
hdisk5         none                None
hdisk6         none                None
vpath1         000790585b9d226d    pprcaix1logvg
vpath2         000790585ba01755    pprcaix1datavg

AIX2#varyonvg pprcaix1logvg
AIX2#varyonvg pprcaix1datavg

AIX2#lsvg -o
```

```
pprcaix1datavg
pprcaix1logvg
testvg
rootvg

AIX2#fsck -y /pprcaix1logfs
** Checking /dev/rpprcaix1loglv (/pprca)
** Phase 0 - Check Log
log redo processing for /dev/rpprcaix1loglv
syncpt record at 2028
end of log 2a48
syncpt record at 2028
syncpt address 2028
number of log records  = 26
number of do blocks = 2
number of nodo blocks = 0
** Phase 1 - Check Blocks and Sizes
** Phase 2 - Check Pathnames
** Phase 3 - Check Connectivity
** Phase 4 - Check Reference Counts
** Phase 5 - Check Inode Map
** Phase 6 - Check Block Map
10 files 2664 blocks 79256 free

AIX2#fsck -y /pprcaix1datafs
** Checking /dev/rpprcaix1datalv (/pprca)
** Phase 0 - Check Log
log redo processing for /dev/rpprcaix1datalv
syncpt record at 2028
end of log 3420
syncpt record at 2028
syncpt address 2028
number of log records  = 52
number of do blocks = 2
number of nodo blocks = 0
** Phase 1 - Check Blocks and Sizes
** Phase 2 - Check Pathnames
** Phase 3 - Check Connectivity
** Phase 4 - Check Reference Counts
** Phase 5 - Check Inode Map
** Phase 6 - Check Block Map
10 files 2664 blocks 79256 free

AIX2#mount /pprcaix1logfs
AIX2#mount /pprcaix1datafs

AIX2#df -k
Filesystem    1024-blocks      Free %Used    Iused %Iused Mounted on
/dev/hd4            32768     26352   20%     1208     8% /
/dev/hd2           393216     79132   80%    14361    15% /usr
/dev/hd9var         32768     30268    8%      155     2% /var
/dev/hd3            32768     31048    6%       61     1% /tmp
/dev/hd1            32768     31684    4%       18     1% /home
/dev/testlv         40960     39632    4%       18     1% /testfs
/dev/pprcaix1loglv       40960     39628    4%       19     1% /pprcaix1logfs
/dev/pprcaix1datalv       40960     39628    4%       19     1% /pprcaix1datafs

AIX2#datapath query adapter
Active Adapters :1
Adpt#     Adapter Name   State     Mode     Select     Errors  Paths  Active
```

```
    0           fscsi0  NORMAL   ACTIVE       3032         0      6       6

AIX2#datapath query device
Total Devices : 3
DEV#:   0  DEVICE NAME: vpath0  TYPE: 2145          POLICY:    Optimized
SERIAL: 600507680186000960000000000000000
==========================================================================
Path#      Adapter/Hard Disk          State     Mode     Select     Errors
    0          fscsi0/hdisk1           OPEN   NORMAL       2076          0
    1          fscsi0/hdisk2           OPEN   NORMAL          0          0

DEV#:   1  DEVICE NAME: vpath1  TYPE: 2145          POLICY:    Optimized
SERIAL: 600507680186000960000000000000001
==========================================================================
Path#      Adapter/Hard Disk          State     Mode     Select     Errors
    0          fscsi0/hdisk3           OPEN   NORMAL          0          0
    1          fscsi0/hdisk5           OPEN   NORMAL        478          0

DEV#:   2  DEVICE NAME: vpath2  TYPE: 2145          POLICY:    Optimized
SERIAL: 600507680186000960000000000000002
==========================================================================
Path#      Adapter/Hard Disk          State     Mode     Select     Errors
    0          fscsi0/hdisk4           OPEN   NORMAL        478          0
    1          fscsi0/hdisk6           OPEN   NORMAL          0          0

AIX2#lsvpcfg
vpath0 (Avail pv testvg) 600507680186000960000000000000000 = hdisk1 (Avail ) hdisk2 (Avail )
vpath1 (Avail pv pprcaix1logvg) 600507680186000960000000000000001 = hdisk3 (Avail ) hdisk5
(Avail )
vpath2 (Avail pv pprcaix1datavg) 600507680186000960000000000000002 = hdisk4 (Avail ) hdisk6
(Avail )

AIX2#ls /pprcaix1logfs
createdonaix1.txt      createinsyncstate.txt  lost+found
AIX2#ls /pprcaix1datafs
createdonaix1.txt      createinsyncstate.txt  lost+found
AIX2#more /pprcaix1logfs/createinsyncstate.txt
this file was created after the start pprc consistency group on aix1 log
AIX2#more /pprcaix1datafs/createinsyncstate.txt
this file was created after the start pprc consistency group on aix1 data
```

11.3.9 Switching the PPRC consistency group

The master vDisks are in SVC1 (REDSTONE1), and the auxiliary vDisks are in SVC2 (lodestoneA). The first time we created the PPRC relationships, we used the default option in which:

- The primary (source) PPRC vDisk is the master vDisk.
- The secondary (target) PPRC is the auxiliary vDisk.

You can switch the primary and the secondary so that:

- The primary (source) PPRC vDisk is the auxiliary vDisk.
- The secondary (target) PPRC is the master vDisk.

Using the CLI

To reverse the role of primary and secondary virtual disk in the PPRC consistency group, you can use the following commands:

```
svctask switchrcconsistgrp -primary aux pprccongrpaix
svctask switchrcconsistgrp -primary master pprccongrpaix
```

This is shown in Example 11-16.

Example 11-16 Command line interface

```
IBM_2145:admin>svctask switchrcconsistgrp -primary aux pprccongrpaix
CMMVC5973E The operation was not performed because the consistency group is not
synchronized.
IBM_2145:admin>svctask switchrcconsistgrp -primary master pprccongrpaix
CMMVC5973E The operation was not performed because the consistency group is not
synchronized.

IBM_2145:admin>svctask startrcconsistgrp -primary aux pprccongrpaix
CMMVC5973E The operation was not performed because the consistency group is not
synchronized.
IBM_2145:admin>svctask startrcconsistgrp -force pprccongrpaix


IBM_2145:admin>svcinfo lsrcconsistgrp pprccongrpaix
id 255
name pprccongrpaix
master_cluster_id 0000010030400364
master_cluster_name REDSTONE1
aux_cluster_id 0000010030C0012C
aux_cluster_name lodestoneA
primary aux
state consistent_synchronised
relationship_count 2
freeze_time
status online
sync
RC_rel_id 12
RC_rel_name pprcmapaixlog
RC_rel_id 13
RC_rel_name pprcmapaixdata

IBM_2145:admin>svcinfo lsrcrelationship pprcmapaixlog
id 12
name pprcmapaixlog
master_cluster_id 0000010030400364
master_cluster_name REDSTONE1
master_vDisk_id 12
master_vDisk_name c1_aix_log_src
aux_cluster_id 0000010030C0012C
aux_cluster_name lodestoneA
aux_vDisk_id 1
aux_vDisk_name
primary aux
consistency_group_id 255
consistency_group_name pprccongrpaix
state consistent_synchronised
bg_copy_priority 50
progress
```

```
freeze_time
status online
sync

IBM_2145:admin>svcinfo lsrcrelationship pprcmapaixdata
id 13
name pprcmapaixdata
master_cluster_id 0000010030400364
master_cluster_name REDSTONE1
master_vDisk_id 13
master_vDisk_name c1_aix_data_src
aux_cluster_id 0000010030C0012C
aux_cluster_name lodestoneA
aux_vDisk_id 2
aux_vDisk_name
primary aux
consistency_group_id 255
consistency_group_name pprccongrpaix
state consistent_synchronised
bg_copy_priority 50
progress
freeze_time
status online
sync

IBM_2145:admin>svctask switchrcconsistgrp -primary aux pprccongrpaix

IBM_2145:admin>svctask switchrcconsistgrp -primary master pprccongrpaix
```

Using the GUI

Follow these steps:

1. On the Starting Remote Copy Consistency Group panel (Figure 11-39), you have the option to reverse the PPRC direction. In this example, we select the **Forced start** option. Under Copy Direction, we select the **Auxiliary --> Master** option. Then click **OK**.

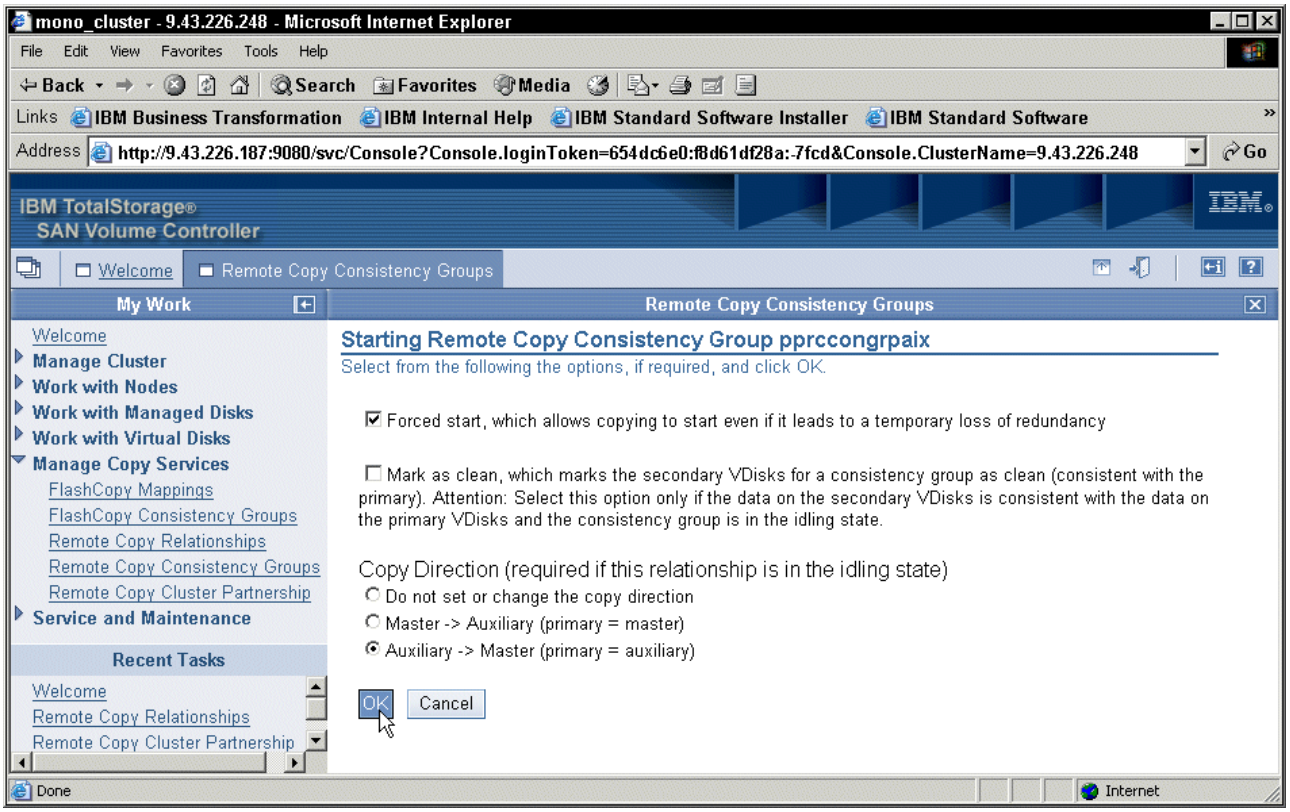


Figure 11-39 Switching the PPRC relationship direction

2. On the View Remote Copy Consistency Groups panel (Figure 11-40), select **Reverse the Copy Direction** from the menu and click **Go**.

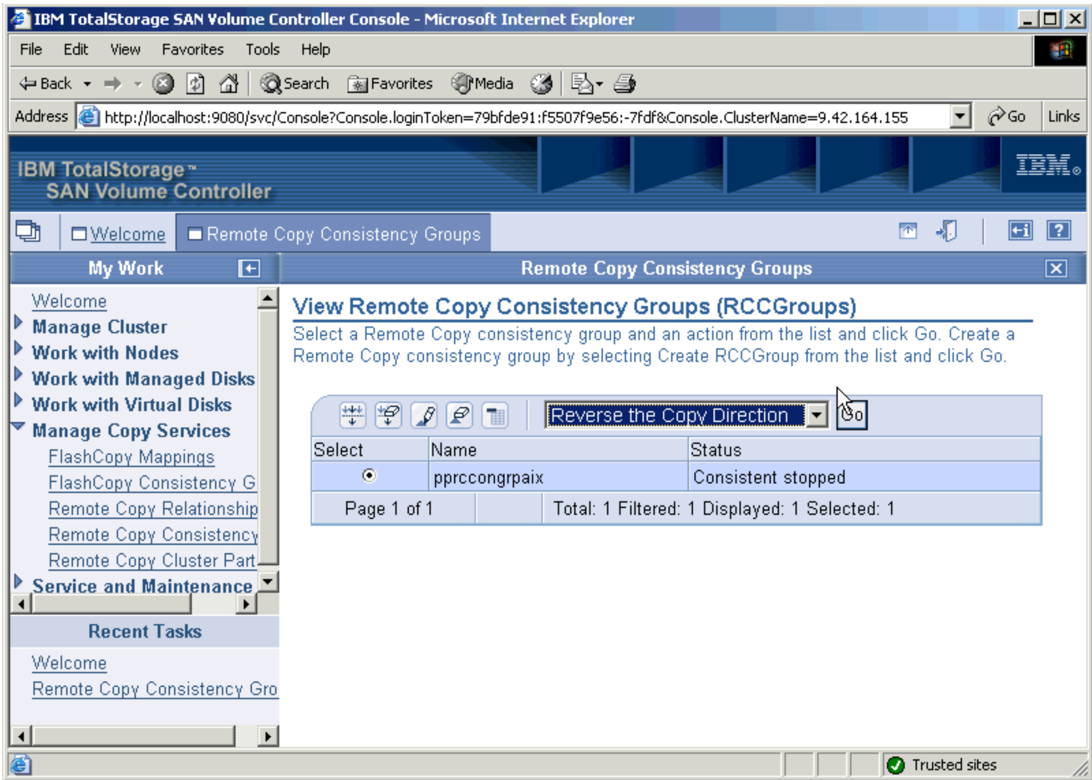


Figure 11-40 Switching the PPRC relationship direction

3. On the Set Copy Direction for Remote Copy relationship or Consistency Group panel (Figure 11-41), you can switch the copy direction of the PPRC consistency group. In this example, we select **Auxiliary --> Master**. Then click **OK**.

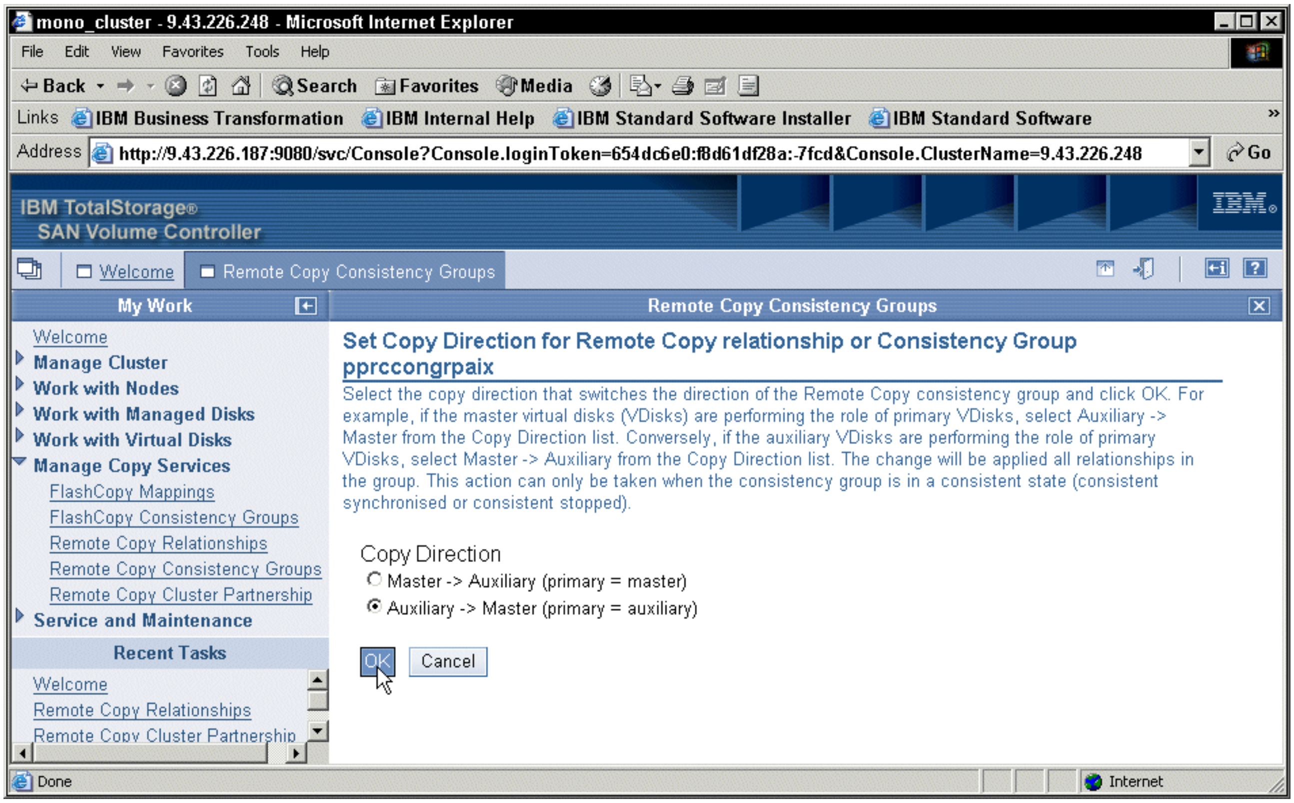


Figure 11-41   Switching the PPRC relationship direction

12

Migration to the SAN Volume Controller

This chapter explains how to migrate from conventional storage area network (SAN) storage to a virtualized storage system.

12.1 Migration overview

This facility allows the mapping of vDisk extents to managed disk extents to be changed, without interrupting a host's access to that virtual disk. This can be performed for any vDisk managed by IBM TotalStorage SAN Volume Controller (SVC). This may be used to:

- Redistribute workload within a cluster across back-end storage:
  - Moving workload onto newly installed storage
  - Moving workload off old/failing storage, ahead of decommissioning it
  - Moving workload to rebalance a changed workload
- Migrate data from legacy back-end storage to SVC managed storage.

12.2 Migration operations

Several different migration activities are supported:

- Migrate multiple extents.
- Migrate all the in-use extents from a managed disk.
- Migrate a virtual disk from one managed disk group (MDG) to another.

The user can inspect the progress of each migration as it is performed. The CLI command is:

```
svcinfo lsmigrate
```

The user can also know the extent allocation between managed disks and virtual disks. This is provided with two styles of information:

- A list of the vDisk ID and a number of extents: These are the vDisks that are using extents on the specified mDisk. The number of extents being used on each mDisk is also shown. The CLI command is:

  ```
  svcinfo lsmdiskextent <vdiskname>
  ```

- A list of mDisk ID and a number of extents: These are the mDisks that are providing extents for the specified vDisk. The number of extents being provided by each mDisk is also shown. The CLI command is:

  ```
  svcinfo lsvdiskextent <vdiskname>
  ```

The user can also know the free extents available on each mDisk. The CLI command is:

```
svcinfo lsfreeextents <mdiskname>
```

After they are started, there is no way for the user to stop a migration. The migration runs to completion unless it is stopped or suspended by an error condition or unless the vDisk being migrated is deleted.

12.2.1 Migrating multiple extents

A number of vDisk extents can be migrated at once using the **migrate extents** command. Extents are allocated on the destination mDisk using the algorithm described in "Allocation of free extents" on page 52.

When executed, this command migrates a given number of extents from the source, specified in terms of the vDisk, and the mDisk that contains some extents used to make up the vDisk. The target is specified in terms of an mDisk (within the same MDG).

If a large number of extents are being migrated, you can specify the number of threads which will be used in parallel (up to four).

If the virtual disk is operating in image mode, the first extent to be migrated is the last extent. After that extent has been migrated the virtual disk is operating in managed mode and behaves as does any other managed mode virtual disk.

The CLI command is:

```
svctask migrateexts -source <mdiskname> -target <mdiskname> -exts <#> -vdisk <vdiskname>
-thread <#>
```

12.2.2 Migrating extents off an mDisk which is being deleted

When an mDisk is deleted from a MDG using **rmmdisk -force** command as explained in 8.3.3, “MDGs” on page 180, it may be necessary to migrate extents off the managed disk before deleting it.

In this case, the extents that need to be migrated are moved onto the set of disks which is not being deleted using the algorithm described in “Allocation of free extents” on page 52. This statement holds true if multiple mDisks are being removed from the MDG at the same time. Disks which are being removed are not candidates for supplying free extents to the allocation of free extents algorithm.

If a vDisk uses one or more extents which need to be moved as a result of a Delete Managed Disk command, then the virtualization policy for that vDisk is set to striped.

If the mDisk is operating in image mode, the image mode vDisk is first converted into a managed mode vDisk.

12.2.3 Migrating a vDisk between MDGs

An entire virtual disk can be migrated to a new MDG using the Migrate Virtual Disk command as explained in 8.4.2, “Virtual disks” on page 188. The CLI command is **svctask migratevdisk**.

Pictorially, this can be represented as shown in Figure 12-1, which shows virtual disk V3 being migrated from MDG Y to MDG X.

> **Important:** In order for migration to be legal, the source and destination managed disk must have the same extent size.

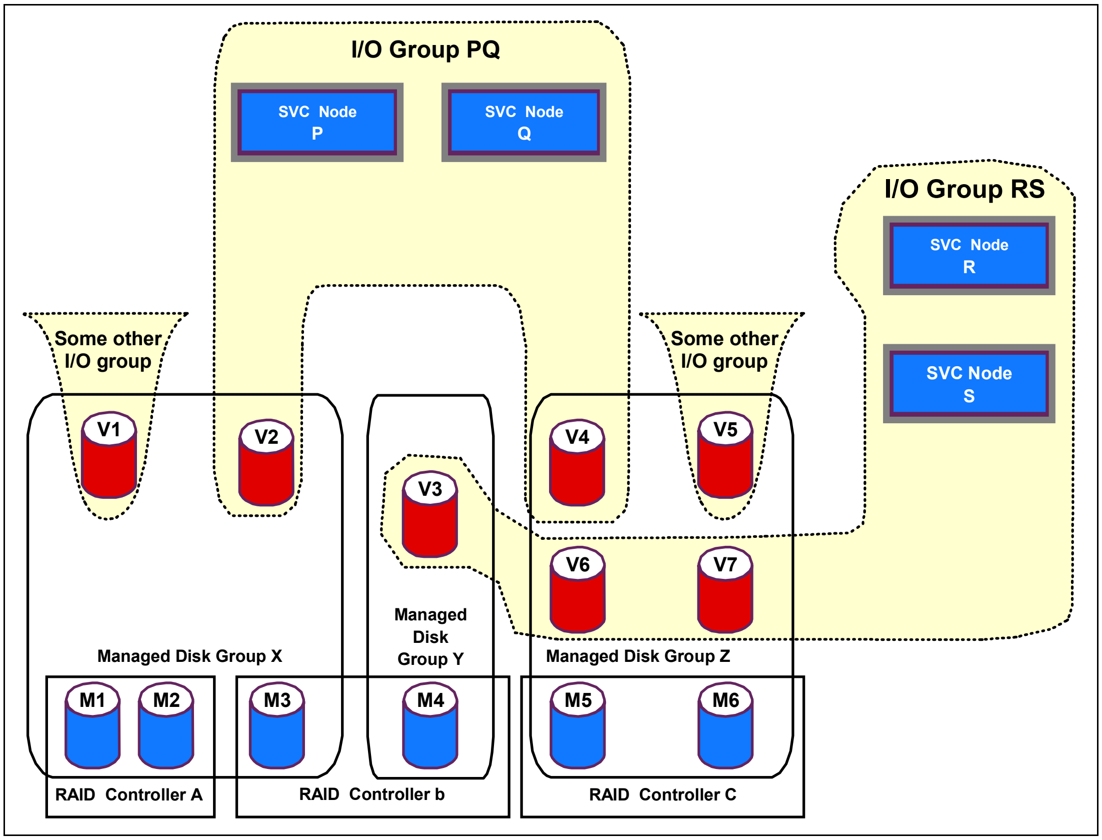


Figure 12-1 Virtual disk migration

Extents are allocated from the set of managed disks in the new MDG using the algorithm described in "Allocation of free extents" on page 52.

The process can be prioritized by specifying the number of threads to use while migrating, using only one thread will put the least background load on the system. If a large number of extents are being migrated, you can specify the number of threads which will be used in parallel (up to four).

After the move, the virtualization policy of the disk is set to striped.

For the duration of the move, the offline rules described in 3.5.7, "I/O handling and offline conditions" on page 53, apply to both MDGs. Thus, referring to Figure 12-1, if any of the managed disks M1, M2, M3, or M4 go offline, then virtual disk V3 goes offline. If managed disk M4 goes offline, then V3 goes offline but V1 and V2 remain online.

If the vDisk is operating in image mode, it is first converted into a managed mode vDisk as described in 12.4, "Migrating data from an image mode vDisk" on page 412.

For the duration of the move, the vDisk is listed as being a member of the original MDG. For the purposes of configuration, the vDisk moves to the new MDG instantaneously at the end of the migration.

12.3 Functional overview of migration

This section describes the functional view of data migration.

12.3.1 Parallelism

Some of the activities can be carried out in parallel.

Per cluster

A SVC cluster supports up to 32 active concurrent instances of members of the set of migration activities:

- Migrate multiple extents
- Migrate between MDGs
- Migrate off deleted managed disk

These high-level migration tasks operate by scheduling single extent migrations as described below.

Up to 256 single extent migrations can run concurrently. This number includes single extent migrates which result from the operations listed above and single extent operations requested by the user.

The Migrate Multiple Extents and Migrate Between MDGs command support a flag which allows the user to specify the number of “threads” to use between 1 and 4. This parameter affects the number of extents which will be concurrently migrated for that migration operation. Thus if the thread value is set to 4, up to four extents can be migrated concurrently for that operation, subject to other resource constraints.

Per managed disk

SVC supports up to four concurrent single extent migrates per managed disk. This limit does not take into account whether the managed disk is the source or the destination. If more than four single extent migrates are scheduled for a particular managed disk, further migrations are queued pending the completion of one of the currently running migrations.

12.3.2 Migration algorithm

This section describes the affect of the migration algorithm.

Chunks

Regardless of the extent size for the MDG, data is migrated in units of 16 MB. In this description, this unit is referred to as a *chunk*.

The algorithm followed to migrate an extent is:

1. Pause (means queue all new I/O requests in the virtualization layer in SVC and wait for all outstanding requests to complete) all I/O on the source managed disk on all nodes in the SVC cluster.
2. Release all I/O on all of the source managed disk apart from writes to the specific chunk which is being migrated.
3. On the node performing the migrate, for each 32 K section of the chunk:
    a. Synchronously read 32K from the source.
    b. Synchronously write 32K to the target.
    c. Wait for the write to complete before proceeding.

4. After the entire chunk is copied to the destination, the process repeats for the next chunk within the extent.

During the migration, the extent can be divided into three regions as shown in Figure 12-2. Region B is the chunk which is being copied. Writes 1 to region B are queued (paused) in the virtualization layer waiting for the chunk to be copied. Reads and writes to Region A are directed to the destination since this data has already been copied. Reads and writes to Region C are directed to the source because this region has yet to be migrated.

The migration of a chunk requires 64 synchronous reads and 64 synchronous writes. During this time, all writes to the chunk from higher layers in the software stack (such as cache destages) are held back. If the back-end storage is operating with significant latency, then it is possible that this operation may take some time (minutes) to complete. This can have an adverse affect on the overall performance of the SVC. To avoid this, if the migration of a particular chunk is still active after one minute, then the migration is paused for 30 seconds. During this time, writes to the chunk are allowed to proceed. After 30 seconds, the migration of the chunk is resumed. This algorithm is repeated as many times as necessary to complete the migration of the chunk.

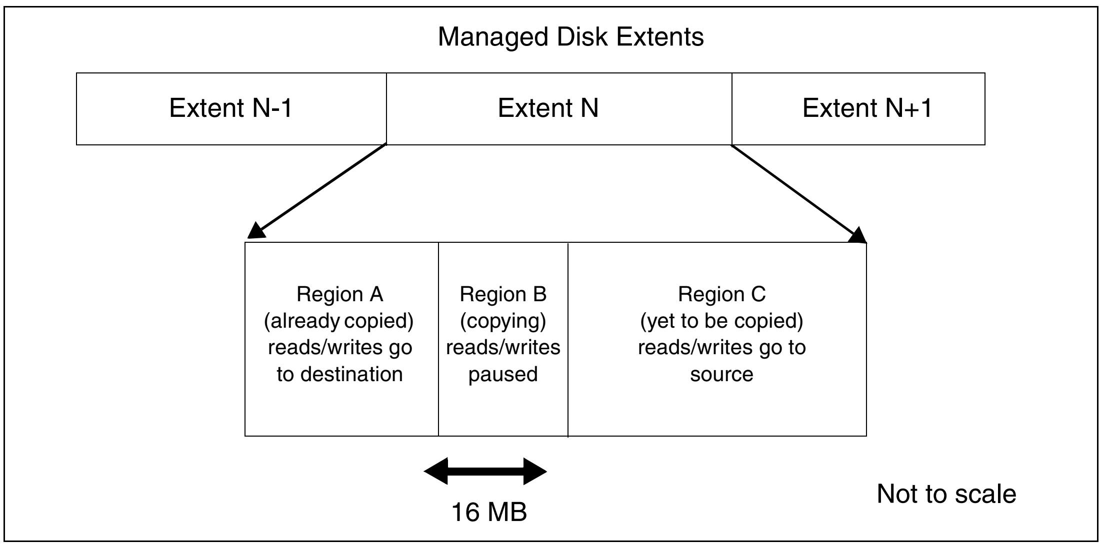


Figure 12-2 Migrating an extent

12.4 Migrating data from an image mode vDisk

This section describes how to migrate data from an image mode virtual disk to a virtual disk.

12.4.1 Image mode vDisk migration concept

Image mode vDisks have the special property that the last extent in the vDisk can be a partial extent. Managed mode disks do not have this property.

To perform any type of migration activity on an image mode vDisk, the image mode disk must first be converted into a managed mode disk. If the image mode disk has a partial last extent, then this last extent in the image mode virtual disk must be the first to be migrated. This migration is handled as a special case.

After this special migration operation has occurred, the vDisk becomes a managed mode virtual disk and is treated in the same way as any other managed mode vDisk. If the image mode disk does not have a partial last extent then no special processing is performed, the image mode virtual disk is simply changed into a managed mode vDisk and is treated in the same way as any other managed mode vDisk.

After data is migrated off a partial extent, there is no way to migrate data back onto the partial extent.

12.4.2 Migration tips

You have several methods to migrate an image mode virtual disk into a managed mode virtual disk:

- If your image mode virtual disk is in the same MDG than the managed disks on which you want to migrate the extents, you can:
  - Migrate a single extent. You have to migrate the last extent of the image mode virtual disk (number N-1).
  - Migrate multiple extents.
  - Migrate all the in-use extents from a managed disk. Migrate extents off a managed disk which is being deleted.
- If you have two MDGs, one for the image mode virtual disk and one for the managed mode virtual disks, you can migrate a virtual disk from one MDG to another.

The recommended method is to have one MDG for all the image mode virtual disks and other MDGs for the managed mode virtual disks and to use the migrate virtual disk facility.

Do not forget to check that enough extents are available in the target MDG.

After the image mode virtual disk is migrated, there is no way to migrate it back.

12.5 Data migration from FAStT to SVC using the CLI

This scenario consists of the migration of one hdisk of an AIX host from an IBM TotalStorage Fibre Array Storage Technology (FAStT) server to a SVC. Before the migration, the FAStT is connected directly to the AIX host. After the migration, the FAStT is only connected to the SVC nodes. Figure 12-3 shows the setup.

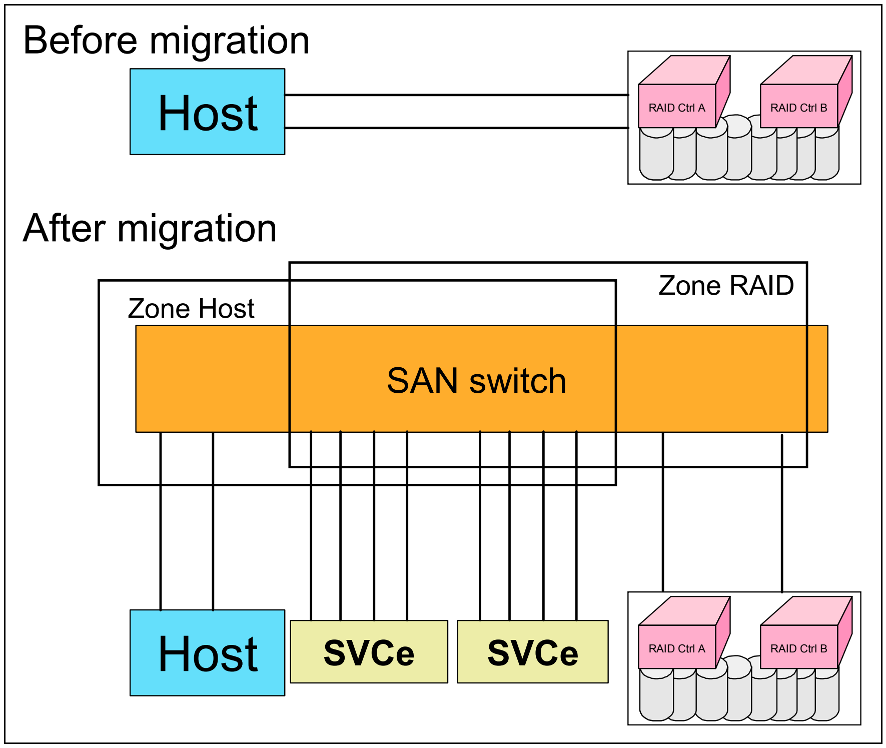


Figure 12-3 Data migration scenario with AIX host and FAStT

12.5.1 AIX host connected directly to the FAStT

The FAStT has two partitions:

- One partition for AIX host with one Logical disk of 2 GB (seen as 1742)
- One partition for SVC with seven logical disks

On the AIX host, the commands shown in Figure 12-4 are issued.

> **Note:** After **unmount**, **varyoffvg**, and **exportvg**, you can start the migration of the host to SVC.

```
lspv
hdisk0 0051a2dac3069dbc rootvg
hdisk1 none None
hdisk2 0051a2da822b17b1 dmvg

lsdev -Cc disk
hdisk0 Available 10-60-00-0,0 16 Bit LVD SCSI Disk Drive
hdisk1 Available 10-60-00-1,0 16 Bit LVD SCSI Disk Drive
hdisk2 Available 20-58-01 1742 Disk Array Device

lsdev -Cc adapter
fcs1 Available 20-58 FC Adapter
fcs0 Available 10-68 FC Adapter

lspv hdisk2
PHYSICAL VOLUME: hdisk2 VOLUME GROUP: dmvg
PV IDENTIFIER: 0051a2da822b17b1 VG IDENTIFIER
0051a2da00004c00000000f18242facf
PV STATE: active
STALE PARTITIONS: 0 ALLOCATABLE: yes
PP SIZE: 16 megabyte(s) LOGICAL VOLUMES: 2
TOTAL PPs: 127 (2032 megabytes) VG DESCRIPTORS: 2
FREE PPs: 62 (992 megabytes) HOT SPARE: no
USED PPs: 65 (1040 megabytes)
FREE DISTRIBUTION: 25..00..00..11..26
USED DISTRIBUTION: 01..25..25..14..00

lsvg -o
dmvg
rootvg

lsvg -l dmvg
dmvg:
LV NAME TYPE LPs PPs PVs LV STATE MOUNT POINT
dmlv jfs 64 64 1 open/syncd /dmfs
loglv00 jfslog 1 1 1 open/syncd N/A
```

Figure 12-4 AIX disk configuration before migration

On the SVC, the commands shown in Figure 12-5 are issued.

```
svcinfo lscontroller 0
id:0
type:4
WWN:200200A0B80F06F3
VLUN link count:8
degraded:no
peripheral dev type:0
peripheral qualifier:0
normal ACA:32
vendor id:IBM
product id low:1742
product id high:
product revision:0500
product serial number:
WWPN table:200200A0B80F06F4
path count:4
WWPN table:200300A0B80F06F4
path count:6

svcinfo lsmdisk -delim :
0:mDisk0:online:unmanaged:::3.0GB:
1:mDisk1:online:unmanaged:::4.0GB:
2:mDisk2:online:unmanaged:::10.0GB:
3:mDisk3:online:unmanaged:::50.0GB:
4:mDisk4:online:unmanaged:::60.0GB:
5:mDisk5:online:unmanaged:::89.0GB:
6:mDisk6:online:unmanaged:::256.3GB:

svcinfo lsnode -delim :
1:node1:10L3BIP:5005676801000050:online:0:io_grp0:yes:231304C33424950
2:node2:10L3BIT:5005676801000019:online:0:io_grp0:no:231304C33424954
```

Figure 12-5 SVC configuration before migration

12.5.2 The SVC is added between the AIX host and the FAStT

Perform these tasks:

1. Change partition's definition on FAStT.
2. Delete hdisk2 in AIX host by entering the following command:

   ```
   rmdev –dl hdisk2 -R
   ```

The FAStT has one partition for SVC with seven logical disks as before plus a new logical disk from the AIX partition (2 GB).

On the SVC

Complete these steps on the SVC:

1. Perform discovery of new managed disk (which was previously direct to AIX). Enter:

   ```
   svctask detectmdisk
   ```

The **svcinfo** command is issued as shown in Figure 12-6.

```
svcinfo lsmdisk
0:mDisk0:online:unmanaged:::3.0GB:
1:mDisk1:online:unmanaged:::4.0GB:
2:mDisk2:online:unmanaged:::10.0GB:
3:mDisk3:online:unmanaged:::50.0GB:
4:mDisk4:online:unmanaged:::60.0GB:
5:mDisk5:online:unmanaged:::89.0GB:
6:mDisk6:online:unmanaged:::256.3GB:
7:mDisk7:online:unmanaged:::2.0GB:
```

Figure 12-6 svcinfo command result

2. Create the mDiskgrp0 with the extent size set to 16 MB:

   `svctask mkmdiskgrp -name mDiskgrp0 -ext 16`

3. Add two mDisks (0 and 1) in mDiskgrp0:

   `svctask addmdisk -mDisk mDisk0:mDisk1 mDiskgrp0`

   The **svcinfo** command output is displayed as shown in Figure 12-7.

```
svcinfo lsmdisk
0:mDisk0:online:managed:0:mDiskgrp0:3.0GB:
1:mDisk1:online:managed:0:mDiskgrp0:4.0GB:
2:mDisk2:online:unmanaged:::10.0GB:
3:mDisk3:online:unmanaged:::50.0GB:
4:mDisk4:online:unmanaged:::60.0GB:
5:mDisk5:online:unmanaged:::89.0GB:
6:mDisk6:online:unmanaged:::256.3GB:
7:mDisk7:online:unmanaged:::2.0GB:
```

Figure 12-7 svinfo command output

4. Create the vDisk in image mode in diskgrp 0 with mDisk 7 and IOgroup 0:

   `svctask mkvdisk -mdiskgrp mDiskgrp0 -iogrp 0 -vtype image -mDisk mDisk7 vDisk0`

   The output of **svinfo** now is displayed as shown in Figure 12-8.

```
svcinfo lsvdisk
0:vDisk0:0:io_grp0:online:0:mDiskgrp0:2.0GB:image::::

svcinfo lsmdisk
0:mDisk0:online:managed:0:mDiskgrp0:3.0GB:
1:mDisk1:online:managed:0:mDiskgrp0:4.0GB:
2:mDisk2:online:unmanaged:::10.0GB:
3:mDisk3:online:unmanaged:::50.0GB:
4:mDisk4:online:unmanaged:::60.0GB:
5:mDisk5:online:unmanaged:::89.0GB:
6:mDisk6:online:unmanaged:::256.3GB:
7:mDisk7:online:image:0:mDiskgrp0:2.0GB:
```

Figure 12-8 Virtual disk in image mode on SVC

On the AIX host

The **shutdown -Fr** command was performed. You can obtain the same results by entering:

```
cfgmgr
```

Figure 12-9 shows the output of the **lspv** command.

```
lspv
hdisk0 0051a2dac3069dbc rootvg
hdisk1 none None
hdisk2 none None
hdisk3 none None
hdisk4 none None
hdisk5 none None
hdisk6 none None
hdisk7 none None
hdisk8 none None
hdisk9 none None
hdisk10 none None
hdisk11 none None
hdisk12 none None
hdisk13 none None
hdisk14 none None
hdisk15 none None
hdisk16 none None
hdisk17 none None
vpath0 0051a2da822b17b1 dmvg
```

Figure 12-9 lspv command output

Figure 12-10 shows the output of the **lsdev -Cc disk** command.

```
lsdev -Cc disk
hdisk0 Available 10-60-00-0,0 16 Bit LVD SCSI Disk Drive
hdisk1 Available 10-60-00-1,0 16 Bit LVD SCSI Disk Drive
hdisk2 Available 20-58-01 IBM TotalStorage SAN Volume Controller device
hdisk3 Available 20-58-01 IBM TotalStorage SAN Volume Controller device
hdisk4 Available 20-58-01 IBM TotalStorage SAN Volume Controller device
hdisk5 Available 20-58-01 IBM TotalStorage SAN Volume Controller device
hdisk6 Available 20-58-01 IBM TotalStorage SAN Volume Controller device
hdisk7 Available 20-58-01 IBM TotalStorage SAN Volume Controller device
hdisk8 Available 20-58-01 IBM TotalStorage SAN Volume Controller device
hdisk9 Available 20-58-01 IBM TotalStorage SAN Volume Controller device
hdisk10 Available 10-68-01 IBM TotalStorage SAN Volume Controller device
hdisk11 Available 10-68-01 IBM TotalStorage SAN Volume Controller device
hdisk12 Available 10-68-01 IBM TotalStorage SAN Volume Controller device
hdisk13 Available 10-68-01 IBM TotalStorage SAN Volume Controller device
hdisk14 Available 10-68-01 IBM TotalStorage SAN Volume Controller device
hdisk15 Available 10-68-01 IBM TotalStorage SAN Volume Controller device
hdisk16 Available 10-68-01 IBM TotalStorage SAN Volume Controller device
hdisk17 Available 10-68-01 IBM TotalStorage SAN Volume Controller device
vpath0 Available Data Path Optimizer Pseudo Device Driver
```

Figure 12-10 lsdev-Cc disk

Figure 12-11 shows the output of a number of AIX commands which are run on the attached host at this stage.

```
lsdev -Cc adapter | grep fcs
fcs1 Available 20-58 FC Adapter
fcs0 Available 10-68 FC Adapter

datapath query adapter
Active Adapters :2
Adpt# Adapter Name State Mode Select Errors Paths Active
0 fscsi1 NORMAL ACTIVE 35 0 8 8
1 fscsi0 NORMAL ACTIVE 46 0 8 8

datapath query device
Total Devices : 1
DEV#: 0 DEVICE NAME: vpath0 TYPE: 2145 POLICY: Optimized
SERIAL: 600567680180000280000000000000000
==========================================================================
Path# Adapter/Hard Disk State Mode Select Errors
0 fscsi1/hdisk2 OPEN NORMAL 0 0
1 fscsi1/hdisk3 OPEN NORMAL 5 0
2 fscsi1/hdisk4 OPEN NORMAL 4 0
3 fscsi1/hdisk5 OPEN NORMAL 3 0
4 fscsi1/hdisk6 OPEN NORMAL 3 0
5 fscsi1/hdisk7 OPEN NORMAL 4 0
6 fscsi1/hdisk8 OPEN NORMAL 9 0
7 fscsi1/hdisk9 OPEN NORMAL 7 0
8 fscsi0/hdisk10 OPEN NORMAL 4 0
9 fscsi0/hdisk11 OPEN NORMAL 6 0
10 fscsi0/hdisk12 OPEN NORMAL 5 0
11 fscsi0/hdisk13 OPEN NORMAL 6 0
12 fscsi0/hdisk14 OPEN NORMAL 6 0
13 fscsi0/hdisk15 OPEN NORMAL 3 0
14 fscsi0/hdisk16 OPEN NORMAL 8 0
15 fscsi0/hdisk17 OPEN NORMAL 8 0
lsvpcfg
vpath0 (Avail pv dmvg) 600567680180000280000000000000000 = hdisk2 (Avail )
hdisk3 (Avail ) hdisk4 (Avail ) hdisk5 (Avail ) hdisk6 (Avail ) hdisk7 (Avail
) hdisk8 (Avail ) hdisk9 (Avail ) hdisk10 (Avail ) hdisk11 (Avail ) hdisk12
(Avail ) hdisk13 (Avail ) hdisk14 (Avail ) hdisk15 (Avail ) hdisk16 (Avail )
hdisk17 (Avail )
```

Figure 12-11 AIX command output

Figure 12-12 shows the output of some of the LVM and file system commands.

```
lsvg -o
dmvg
rootvg

lsvg -l dmvg
dmvg:
LV NAME TYPE LPs PPs PVs LV STATE MOUNT POINT
dmlv jfs 64 64 1 open/syncd /dmfs
loglv00 jfslog 1 1 1 open/syncd N/A

df -k
Filesystem 1024-blocks Free %Used Iused %Iused Mounted on
/dev/hd4 65536 53764 18% 1391 5% /
/dev/hd2 720896 75328 90% 22303 13% /usr
/dev/hd9var 65536 56900 14% 463 3% /var
/dev/hd3 327680 316900 4% 91 1% /tmp
/dev/hd1 65536 63404 4% 18 1% /home
/proc - - - - - /proc
/dev/hd10opt 65536 48984 26% 596 4% /opt
/dev/lv00 262144 125912 52% 20019 31% /usr/vice/cache
AFS 72000000 72000000 0% 0 0% /afs
/dev/dmlv 1048576 1015596 4% 21 1% /dmfs

ls
lost+found
temp
titi.ttt
toto.ttt
```

Figure 12-12 LVM and file system commands

12.5.3 Migrating the vDisk from image mode to managed mode

The vDisk in image mode is migrated to managed mode by migrating several extents from one managed disk to another managed disk method. The output of the **svinfo** command is shown in Figure 12-13.

```
svcinfo lsvdiskextent vDisk0
id number extents
7 128
```

Figure 12-13 svinfo output

The value *128* is the number of extents that are used and exist for vDisk 0. They are all on the mDisk 7 (which is vDisk 0 since all the vDisk in image mode is only on one mDisk). No extents are used in mDisk 0 and mDisk 1

Note: If you want to migrate using one extent, then you have to migrate only extent 127 which is the last one. After that, the full vDisk is automatically in managed mode.

You can check the number of free extents on the mDisks that you want to migrate to using the following commands:

```
svcinfo lsfreeextents mDisk0
svcinfo lsfreeextents mDisk1
```

Attention: For migration using migrate extents, be sure that the number of free extents on the target mDisks is greater that the number of extents you want to migrate.

- Migration of vDisk in image mode (vDisk0) (source mDisk7) (number of extents 40) (target mDisk 0) to managed mode:

  ```
  svctask migrateexts -vDisk vDisk0 -source mDisk7 -target mDisk0 -exts 40
  ```

  The output from the **svinfo** command after this operation is shown in Figure 12-14.

```
svcinfo lsvdiskextent vDisk0
id: number extents
0 40 ---> 40 extents on mDisk0
1 0 ----> 0 extent on mDisk1
7 88 ---> 88 extents on mDisk7
```

Figure 12-14 svcinfo command output

- After a migration from mDisk 7 to mDisk 1 of 43 extents:

  ```
  svctask migrateexts -vDisk vDisk0 -source mDisk7 -target mDisk1 -exts 43
  ```

  Figure 12-15 shows the resulting output from the **svinfo** command with various flags.

```
svcinfo lsvdiskextent vDisk0
id:number extents
0:40
1:43
7:45
svcinfo lsmdiskextent mDisk0
id:number of extents
0:40

svcinfo lsmdiskextent mDisk1
id:number of extents
0:43

svcinfo lsmdiskextent mDisk7
id:number of extents
0:45
```

Figure 12-15 svcinfo command output

Figure 12-16 shows output from the **svinfo** command with different flags.

```
svcinfo lsmdisk -delim :
id:name:status:access mode:mDisk_grp id:mDisk_grp name:capacity:ctrl s/n
0:mDisk0:online:managed:0:mDiskgrp0:3.0GB:
1:mDisk1:online:managed:0:mDiskgrp0:4.0GB:
2:mDisk2:online:unmanaged:::10.0GB:
3:mDisk3:online:unmanaged:::50.0GB:
4:mDisk4:online:unmanaged:::60.0GB:
5:mDisk5:online:unmanaged:::89.0GB:
6:mDisk6:online:unmanaged:::256.3GB:
7:mDisk7:online:managed:0:mDiskgrp0:2.0GB:

svcinfo lsvdisk -delim :
id:name:IO group id:IO group name:status:mDisk_grp id:mDisk_grp
name:capacity:type:FC id:FC name:id:RC name
0:vDisk0:0:io_grp0:online:0:mDiskgrp0:2.0GB:seq::::

svcinfo lsmdiskgrp -delim :
id:name:status:mDisk count:vDisk count:capacity:extent size:free capacity
0:mDiskgrp0:online:3:1:9.0GB:16:7.0GB
```

Figure 12-16 svcinfo command output

12.6 Data migration from ESS to SVC using the CLI

This scenario (shown in Figure 12-17) consists in the migration of one hdisk (IBM TotalStorage Enterprise Storage Server (ESS) disk1) of an AIX host AIX1 from a ESS to a SVC. The ESS is before and after the migration connected directly the AIX host (via the switches 2109).

Before the migration, the logical unit number (LUN) masking is defined in the ESS in order to give access to the AIX host system for the ESS disk1 and ESS disk2. *After* the migration, the LUN masking is defined in the ESS to give access to the AIX host system for the ESS disk1 and access to the SVC nodes for the ESS disk2.

The following actions occur for the migration:

1. The hdisk is first discovered as “mDisk unmanaged” by the SVC. It is mDisk9.
2. A vDisk is created in a image mode using this mDisk mDisk9.
3. This new vDisk is mapped to the host system AIX1.
4. You can run the **cfgmgr** command on the AIX host system AIX1.
5. The vDisk is again available for the host system AIX1.
6. The vDisk can be migrated from image mode to managed mode, concurrently while it is being accessed by the AIX host system. The mode of migration is to migrate some extents (100) from an mDisk to another mDisk in an mDisk group.

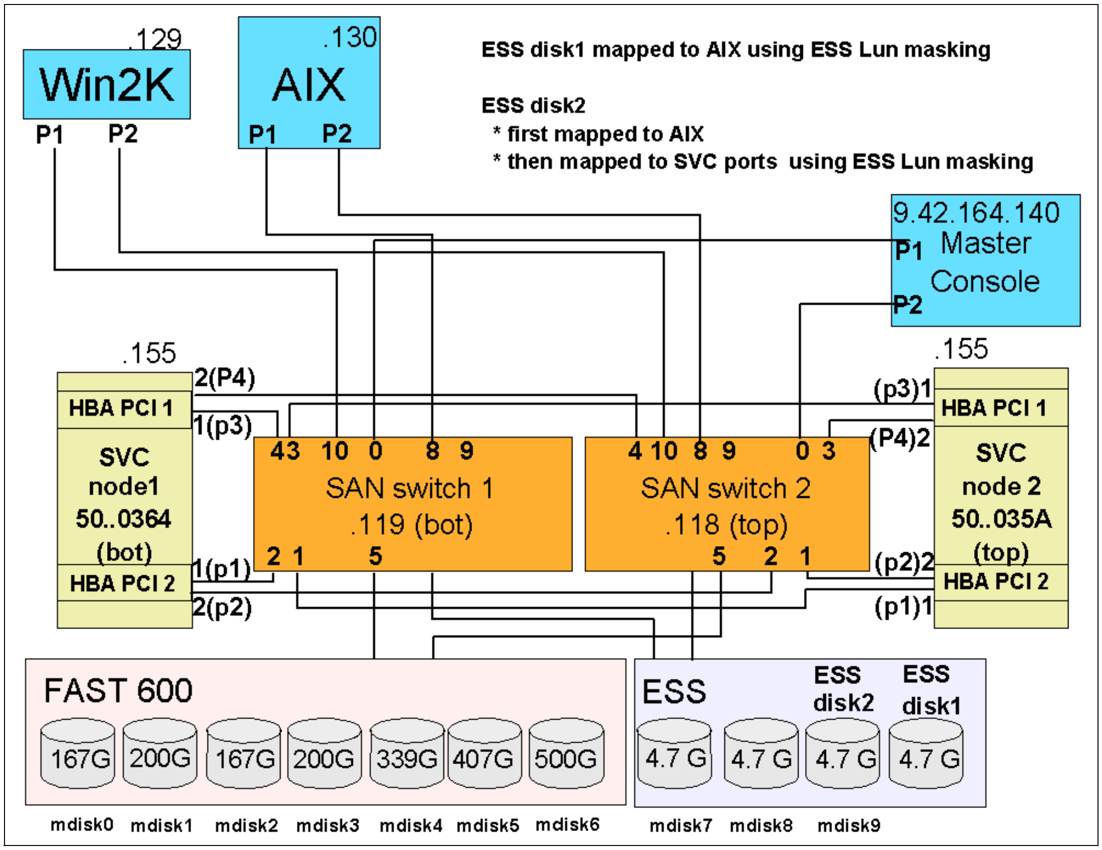


Figure 12-17 Migration scenario from ESS to SVC

12.6.1 AIX host connected directly to the ESS

The AIX host initial configuration is shown in Example 12-1.

Example 12-1 AIX host system configuration before the migration

```
AIX1#lspv
hdisk0          0007905873b5c2fb                    rootvg
hdisk1          0007905873b5c364                    rootvg
hdisk2          none                                None
hdisk3          none                                None
hdisk4          none                                None
hdisk5          none                                None
vpath0          00079058573c2177                    essdisk1vg
vpath1          00079058573c65a9                    essdisk2vg

AIX1#lsdev -Cc disk
hdisk0 Available 10-60-00-8,0 16 Bit SCSI Disk Drive
hdisk1 Available 10-60-00-9,0 16 Bit SCSI Disk Drive
hdisk2 Available 20-58-01     IBM FC 2105F20
hdisk3 Available 20-58-01     IBM FC 2105F20
hdisk4 Available 20-60-01     IBM FC 2105F20
hdisk5 Available 20-60-01     IBM FC 2105F20
vpath0 Available              Data Path Optimizer Pseudo Device Driver
vpath1 Available              Data Path Optimizer Pseudo Device Driver
```

```
AIX1#datapath query adapter
Active Adapters :2
Adpt#     Adapter Name   State     Mode     Select     Errors  Paths  Active
    0          fscsi0  NORMAL   ACTIVE       6607         12      2       2
    1          fscsi1  NORMAL   ACTIVE       6757         12      2       2

AIX1#datapath query device
Total Devices : 2
DEV#:   0  DEVICE NAME: vpath0  TYPE: 2105F20         POLICY:    Optimized
SERIAL: 40514830
==========================================================================
Path#      Adapter/Hard Disk          State     Mode     Select     Errors
    0          fscsi0/hdisk2           OPEN   NORMAL       1797          0
    1          fscsi1/hdisk4           OPEN   NORMAL       1825          0
DEV#:   1  DEVICE NAME: vpath1  TYPE: 2105F20         POLICY:    Optimized
SERIAL: 40914830
==========================================================================
Path#      Adapter/Hard Disk          State     Mode     Select     Errors
    0          fscsi0/hdisk3           OPEN   NORMAL       1792          0
    1          fscsi1/hdisk5           OPEN   NORMAL       1796          0

AIX1#lspv vpath1
PHYSICAL VOLUME:    vpath1                   VOLUME GROUP:     essdisk2vg
PV IDENTIFIER:      00079058573c65a9 VG IDENTIFIER     0007905800004c00000000f5573c6ecd
PV STATE:           active
STALE PARTITIONS:   0                        ALLOCATABLE:      yes
PP SIZE:            8 megabyte(s)            LOGICAL VOLUMES:  2
TOTAL PPs:          595 (4760 megabytes)     VG DESCRIPTORS:   2
FREE PPs:           584 (4672 megabytes)     HOT SPARE:        no
USED PPs:           11 (88 megabytes)
FREE DISTRIBUTION:  119..108..119..119..119
USED DISTRIBUTION:  00..11..00..00..00

AIX1#lsvpcfg
vpath0 (Avail pv essdisk1vg) 40514830 = hdisk2 (Avail ) hdisk4 (Avail )
vpath1 (Avail pv essdisk2vg) 40914830 = hdisk3 (Avail ) hdisk5 (Avail )

AIX1#lsvg -o
essdisk2vg
essdisk1vg
rootvg

AIX1#lsvg -l essdisk1vg
essdisk1vg:
LV NAME             TYPE       LPs   PPs   PVs  LV STATE      MOUNT POINT
essdisk1lv          jfs        10    10    1    open/syncd    /essdisk1fs
loglv00             jfslog     1     1     1    open/syncd    N/A

AIX1#lsvg -l essdisk2vg
essdisk2vg:
LV NAME             TYPE       LPs   PPs   PVs  LV STATE      MOUNT POINT
essdisk2lv          jfs        10    10    1    open/syncd    /essdisk2fs
loglv01             jfslog     1     1     1    open/syncd    N/A

AIX1#lsvg essdisk2vg
VOLUME GROUP:   essdisk2vg               VG IDENTIFIER:  0007905800004c00000000f5573c6ecd
VG STATE:       active                   PP SIZE:        8 megabyte(s)
VG PERMISSION:  read/write               TOTAL PPs:      595 (4760 megabytes)
MAX LVs:        256                      FREE PPs:       584 (4672 megabytes)
LVs:            2                        USED PPs:       11 (88 megabytes)
```

```
OPEN LVs:       2                        QUORUM:         2
TOTAL PVs:      1                        VG DESCRIPTORS: 2
STALE PVs:      0                        STALE PPs:      0
ACTIVE PVs:     1                        AUTO ON:        yes
MAX PPs per PV: 1016                     MAX PVs:        32
LTG size:       128 kilobyte(s)          AUTO SYNC:      no
HOT SPARE:      no

AIX1#df -k
Filesystem    1024-blocks      Free %Used    Iused %Iused Mounted on
/dev/hd4            16384      7556   54%     1296    16% /
/dev/hd2          1867776   1265400   33%    20526     5% /usr
/dev/hd9var         16384      9912   40%      436    11% /var
/dev/hd3            32768     29156   12%       65     1% /tmp
/dev/essdisk1lv       81920     79308    4%       17     1% /essdisk1fs
/dev/essdisk2lv       81920     79308    4%       17     1% /essdisk2fs

AIX1#pwd
/essdisk1fs
AIX1#ls
file_on_ess.txt  lost+found
AIX1#more file_on_ess.txt
file created on ess with direct connexion on disk1

AIX1#pwd
/essdisk2fs
AIX1#ls
file_on_ess.txt  lost+found
AIX1#more file_on_ess.txt
file created on ess with direct connexion on disk2

>>>>>>>>>>>> Before the migration of ess disk 2 >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>><<<

AIX1#varyoffvg essdisk2vg

AIX1#exportvg essdisk2vg

AIX1#lsvg -o
essdisk1vg
rootvg

AIX1#df -k
Filesystem    1024-blocks      Free %Used    Iused %Iused Mounted on
/dev/hd4            16384      7552   54%     1290    16% /
/dev/hd2          1867776   1265400   33%    20526     5% /usr
/dev/hd9var         16384      9912   40%      436    11% /var
/dev/hd3            32768     29148   12%       66     1% /tmp
/dev/essdisk1lv       81920     79304    4%       18     1% /essdisk1fs

AIX1#lsvpcfg
vpath0 (Avail pv essdisk1vg) 40514830 = hdisk2 (Avail ) hdisk4 (Avail )
vpath1 (Avail pv ) 40914830 = hdisk3 (Avail ) hdisk5 (Avail )

AIX1#lsdev -Cc disk
hdisk0 Available 10-60-00-8,0 16 Bit SCSI Disk Drive
hdisk1 Available 10-60-00-9,0 16 Bit SCSI Disk Drive
hdisk2 Available 20-58-01     IBM FC 2105F20
hdisk3 Available 20-58-01     IBM FC 2105F20
hdisk4 Available 20-60-01     IBM FC 2105F20
hdisk5 Available 20-60-01     IBM FC 2105F20
```

```
vpath0 Available              Data Path Optimizer Pseudo Device Driver
vpath1 Available              Data Path Optimizer Pseudo Device Driver

AIX1#lspv
hdisk0         0007905873b5c2fb                    rootvg
hdisk1         0007905873b5c364                    rootvg
hdisk2         none                                None
hdisk3         none                                None
hdisk4         none                                None
hdisk5         none                                None
vpath0         00079058573c2177                    essdisk1vg
vpath1         00079058573c65a9                    None

AIX1#rmdev -dl vpath1
vpath1 deleted
AIX1#rmdev -dl hdisk3
hdisk3 deleted
AIX1#rmdev -dl hdisk5
hdisk5 deleted

AIX1#lsdev -Cc disk
hdisk0 Available 10-60-00-8,0 16 Bit SCSI Disk Drive
hdisk1 Available 10-60-00-9,0 16 Bit SCSI Disk Drive
hdisk2 Available 20-58-01     IBM FC 2105F20
hdisk4 Available 20-60-01     IBM FC 2105F20
vpath0 Available              Data Path Optimizer Pseudo Device Driver

AIX1#lspv
hdisk0         0007905873b5c2fb                    rootvg
hdisk1         0007905873b5c364                    rootvg
hdisk2         none                                None
hdisk4         none                                None
vpath0         00079058573c2177                    essdisk1vg

AIX1#lsvpcfg
vpath0 (Avail pv essdisk1vg) 40514830 = hdisk2 (Avail ) hdisk4 (Avail )

AIX1#datapath query adapter
Active Adapters :2
Adpt#     Adapter Name   State     Mode     Select     Errors  Paths  Active
    0           fscsi0  NORMAL   ACTIVE       8029         12      1       1
    1           fscsi1  NORMAL   ACTIVE       8228         12      1       1

AIX1#datapath query device
Total Devices : 1
DEV#:   0  DEVICE NAME: vpath0  TYPE: 2105F20         POLICY:    Optimized
SERIAL: 40514830
==========================================================================
Path#      Adapter/Hard Disk          State     Mode     Select     Errors
    0         fscsi0/hdisk2            OPEN   NORMAL       2502          0
    1         fscsi1/hdisk4            OPEN   NORMAL       2566          0
```

Now if you look at the SVC, there are already two mDisks (mDisk7 and mDisk8) from an ESS in unmanaged mode as shown Example 12-2.

Example 12-2 Configuring SVC back-end storage

```
IBM_2145:admin>svcinfo lsmdisk -delim /
id/name/status/mode/mDisk_grp_id/mDisk_grp_name/capacity/ctrl_LUN_#/controller_name
0/mDisk0/online/unmanaged///167.0GB/0000000000000000/F600_23C1436
1/mDisk1/online/unmanaged///200.4GB/0000000000000001/F600_23C1436
2/mDisk2/online/unmanaged///167.0GB/0000000000000002/F600_23C1436
3/mDisk3/online/unmanaged///200.4GB/0000000000000003/F600_23C1436
4/mDisk4/online/unmanaged///339.3GB/0000000000000004/F600_23C1436
5/mDisk5/online/unmanaged///407.2GB/0000000000000005/F600_23C1436
6/mDisk6/online/unmanaged///500.0GB/0000000000000006/F600_23C1436
7/mDisk7/online/unmanaged///4.7GB/0000000000000000/ESS
8/mDisk8/online/unmanaged///4.7GB/0000000000000001/ESS
```

We use these two mDisks to create a MDG named “ess_mDiskgrp0” with an extent size of 16 MB and then we add the two mDisks from ESS in the mDisk group as shown in Example 12-3.

Example 12-3 Creating an mDisk group with two mDisks from ESS

```
>>>>>>>>>>>>>   Creation of an mDisk group for ESS back-end storage

IBM_2145:admin>svctask mkmDiskgrp -name ess_mDiskgrp0 -ext 16
mDisk Group, id [1], successfully created

IBM_2145:admin>svcinfo lsmdiskgrp -delim :
id:name:status:mDisk_count:vDisk_count:capacity:extent_size:free_capacity
1:ess_mDiskgrp0:online:0:0:0:16:0


>>>>>>>>> add 2 mDisks 7 and 8 from ESS in ess_mDiskgrp0

IBM_2145:admin>svctask addmDisk -mDisk mDisk7:mDisk8 ess_mDiskgrp0

IBM_2145:admin>svcinfo lsmdiskgrp -delim :
id:name:status:mDisk_count:vDisk_count:capacity:extent_size:free_capacity
1:ess_mDiskgrp0:online:2:0:9.3GB:16:9.3GB

IBM_2145:admin>svcinfo lsmdiskgrp ess_mDiskgrp0
id 1
name ess_mDiskgrp0
status online
mDisk_count 2
vDisk_count 0
capacity 9.3GB
extent_size 16
free_capacity 9.3GB

IBM_2145:admin>svcinfo lsmdisk -delim :
id:name:status:mode:mDisk_grp_id:mDisk_grp_name:capacity:ctrl_LUN_#:controller_name
0:mDisk0:online:unmanaged:::167.0GB:0000000000000000:F600_23C1436
1:mDisk1:online:unmanaged:::200.4GB:0000000000000001:F600_23C1436
2:mDisk2:online:unmanaged:::167.0GB:0000000000000002:F600_23C1436
```

```
3:mDisk3:online:unmanaged:::200.4GB:0000000000000003:F600_23C1436
4:mDisk4:online:unmanaged:::339.3GB:0000000000000004:F600_23C1436
5:mDisk5:online:unmanaged:::407.2GB:0000000000000005:F600_23C1436
6:mDisk6:online:unmanaged:::500.0GB:0000000000000006:F600_23C1436
7:mDisk7:online:managed:1:ess_mDiskgrp0:4.7GB:0000000000000000:ESS
8:mDisk8:online:managed:1:ess_mDiskgrp0:4.7GB:0000000000000001:ESS
```

12.6.2 The SVC added between the AIX host and the ESS

After changing the LUN masking definition on the ESS, add the SVC between the AIX host and the ESS.

On the SAN Volume Controller

Follow these steps:

1. Perform discovery of the new managed disk (previously mapped to AIX) by entering the following command:

   `svctask detectmdisk`

2. Enter the **`svcinfo`** command as shown Example 12-4.

Example 12-4 A new mDisk from ESS mDisk9 which corresponds to ESS disk2 from AIX

```
IBM_2145:admin>svcinfo lsmdisk -delim :
id:name:status:mode:mDisk_grp_id:mDisk_grp_name:capacity:ctrl_LUN_#:controller_name
0:mDisk0:online:unmanaged:::167.0GB:0000000000000000:F600_23C1436
1:mDisk1:online:unmanaged:::200.4GB:0000000000000001:F600_23C1436
2:mDisk2:online:unmanaged:::167.0GB:0000000000000002:F600_23C1436
3:mDisk3:online:unmanaged:::200.4GB:0000000000000003:F600_23C1436
4:mDisk4:online:unmanaged:::339.3GB:0000000000000004:F600_23C1436
5:mDisk5:online:unmanaged:::407.2GB:0000000000000005:F600_23C1436
6:mDisk6:online:unmanaged:::500.0GB:0000000000000006:F600_23C1436
7:mDisk7:online:managed:1:ess_mDiskgrp0:4.7GB:0000000000000000:ESS
8:mDisk8:online:managed:1:ess_mDiskgrp0:4.7GB:0000000000000001:ESS
9:mDisk9:online:unmanaged:::4.7GB:0000000000000002:ESS

IBM_2145:admin>svcinfo lsfreeextents mDisk7
id 7
number_of_extents 297

IBM_2145:admin>svcinfo lsfreeextents mDisk8
id 8
number_of_extents 297

IBM_2145:admin>svcinfo lsfreeextents mDisk9
id 9
number_of_extents 0
```

3. Create the vDisk in image mode using the mDisk mDisk9 as shown in Example 12-5.

Example 12-5 Creating the vDisk in image mode

```
>>>>>>>>> create the vDisk in image mode using mDisk9 in mdisgrp ess

IBM_2145:admin>svctask mkvdisk -mDiskgrp ess_mDiskgrp0 -iogrp 0 -vtype image -mDisk mDisk9
aix1_ima_vd1
Host LUN, id [0], successfully created

IBM_2145:admin>svcinfo lsmdisk -delim :
id:name:status:mode:mDisk_grp_id:mDisk_grp_name:capacity:ctrl_LUN_#:controller_name
0:mDisk0:online:unmanaged:::167.0GB:0000000000000000:F600_23C1436
1:mDisk1:online:unmanaged:::200.4GB:0000000000000001:F600_23C1436
2:mDisk2:online:unmanaged:::167.0GB:0000000000000002:F600_23C1436
3:mDisk3:online:unmanaged:::200.4GB:0000000000000003:F600_23C1436
4:mDisk4:online:unmanaged:::339.3GB:0000000000000004:F600_23C1436
5:mDisk5:online:unmanaged:::407.2GB:0000000000000005:F600_23C1436
6:mDisk6:online:unmanaged:::500.0GB:0000000000000006:F600_23C1436
7:mDisk7:online:managed:1:ess_mDiskgrp0:4.7GB:0000000000000000:ESS
8:mDisk8:online:managed:1:ess_mDiskgrp0:4.7GB:0000000000000001:ESS
9:mDisk9:online:image:1:ess_mDiskgrp0:4.7GB:0000000000000002:ESS

IBM_2145:admin>svcinfo lsvdisk -delim :
id:name:IO_group_id:IO_group_name:status:mDisk_grp_id:mDisk_grp_name:capacity:type:FC_id:FC
_name:RC_id:RC_name
0:vDisk0:0:io_grp0:online:1:ess_mDiskgrp0:4.7GB:image::::

IBM_2145:admin>svcinfo lsmdisk mDisk9
id 9
name mDisk9
status online
mode image
mDisk_grp_id 1
mDisk_grp_name ess_mDiskgrp0
capacity 4.7GB
quorum_index
block_size 512
controller_name ESS
ctrl_type 4
ctrl_WWNN 5005076300C00BE8
controller_id 1
path_count 2
max_path_count 2
ctrl_LUN_# 2

IBM_2145:admin>svcinfo lsvdisk vDisk0
id 0
name vDisk0
IO_group_id 0
IO_group_name io_grp0
status online
mDisk_grp_id 1
mDisk_grp_name ess_mDiskgrp0
capacity 4.7GB
type image
formatted no
mDisk_id 9
mDisk_name mDisk9
```

```
FC_id
FC_name
RC_id
RC_name
throttling 0
preferred_node_id 2

>>> number of extents of the vDisk. mDisk 9 is providing extents for the specified vDisk 0
IBM_2145:admin>svcinfo lsvdiskextent vDisk0
id              number_extents
9               299

>>> vDisk 0 that is using extents on the specified mDisk mDisk9.

IBM_2145:admin>svcinfo lsmdiskextent mDisk7

IBM_2145:admin>svcinfo lsmdiskextent mDisk8

IBM_2145:admin>svcinfo lsmdiskextent mDisk9
id              number_of_extents
0               299
```

4. Map the vDisk to the host system AIX (LUN masking) as shown in Example 12-6.

Example 12-6 Host mapping of the image vDisk to host system AIX

```
IBM_2145:admin>svcinfo lshbaportcandidate
id
210000E08B08AFD6
210100E08B291541
210100E08B28AFD6
210000E08B091541
210100E08B24D751
210000E08B04D751
10000000C9295A9A
10000000C9266F5B

IBM_2145:admin>svcinfo lshost
id              name            port_count
0               AIX1            2

IBM_2145:admin>svcinfo lsvdisk -delim :
id:name:IO_group_id:IO_group_name:status:mDisk_grp_id:mDisk_grp_name:capacity:type:FC_id:FC
_name:RC_id:RC_name
0:vDisk0:0:io_grp0:online:1:ess_mDiskgrp0:4.7GB:image::::

IBM_2145:admin>svcinfo lshostvdiskmap -delim :
id:name:SCSI_id:vDisk_id:vDisk_name:wwpn:vDisk_UID
0:AIX1:0:0:vDisk0:10000000C9266F5B:600507680183001AD00000000000000A

IBM_2145:admin>svcinfo lsvdiskhostmap -delim :
id:name:SCSI_id:host_id:host_name:wwpn:vDisk_UID
0:vDisk0:0:0:AIX1:10000000C9266F5B:600507680183001AD00000000000000A
0:vDisk0:0:0:AIX1:10000000C9295A9A:600507680183001AD00000000000000A
```

On the AIX host

On the AIX host system after you run the **cfgmgr** command, you see the results as shown in Example 12-7.

Example 12-7 Discovery of the ESS disk2 on the AIX host system via the SVC

```
>>>>>>>>>>>>>>>>>>>>>>>>>>>>  after cfgmgr >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>><

AIX1#lspv
hdisk0          0007905873b5c2fb                    rootvg
hdisk1          0007905873b5c364                    rootvg
hdisk2          none                                None
hdisk4          none                                None
vpath0          00079058573c2177                    essdisk1vg
hdisk3          00079058573c65a9                    None
hdisk5          00079058573c65a9                    None
hdisk6          00079058573c65a9                    None
hdisk7          00079058573c65a9                    None
vpath1          none                                None

AIX1#lsdev -Cc disk
hdisk0 Available 10-60-00-8,0 16 Bit SCSI Disk Drive
hdisk1 Available 10-60-00-9,0 16 Bit SCSI Disk Drive
hdisk2 Available 20-58-01     IBM FC 2105F20
hdisk4 Available 20-60-01     IBM FC 2105F20
vpath0 Available              Data Path Optimizer Pseudo Device Driver
hdisk3 Available 20-58-01     IBM TotalStorage SAN Volume Controller device
hdisk5 Available 20-58-01     IBM TotalStorage SAN Volume Controller device
hdisk6 Available 20-60-01     IBM TotalStorage SAN Volume Controller device
hdisk7 Available 20-60-01     IBM TotalStorage SAN Volume Controller device
vpath1 Available              Data Path Optimizer Pseudo Device Driver

AIX1#importvg -y essdisk2vg hdisk3
essdisk2vg
AIX1#lspv
hdisk0          0007905873b5c2fb                    rootvg
hdisk1          0007905873b5c364                    rootvg
hdisk2          none                                None
hdisk4          none                                None
vpath0          00079058573c2177                    essdisk1vg
hdisk3          00079058573c65a9                    essdisk2vg
hdisk5          00079058573c65a9                    essdisk2vg
hdisk6          00079058573c65a9                    essdisk2vg
hdisk7          00079058573c65a9                    essdisk2vg
vpath1          none                                None

AIX1#hd2vp essdisk2vg
essdisk2vg
essdisk2vg contains these hdisks and pvids
   00079058573c65a90000000000000000 hdisk3

hdisk3 is part of vpath1.
essdisk2vg
Deleting pvid of hdisk3 from ODM database
Create vpath1 pvid of 00079058573c65a9 in ODM database.
lname=vpath1, pvid=00079058573c65a9
Deleting pvid of hdisk5 from ODM database
Deleting pvid of hdisk6 from ODM database
```

```
Deleting pvid of hdisk7 from ODM database
varyon essdisk2vg was successful.
essdisk2vg is converted to dpo device vpathes successfully!

AIX1#lspv
hdisk0          0007905873b5c2fb                    rootvg
hdisk1          0007905873b5c364                    rootvg
hdisk2          none                                None
hdisk4          none                                None
vpath0          00079058573c2177                    essdisk1vg
hdisk3          none                                None
hdisk5          none                                None
hdisk6          none                                None
hdisk7          none                                None
vpath1          00079058573c65a9                    essdisk2vg

AIX1#datapath query adapter
Active Adapters :2
Adpt#     Adapter Name   State     Mode     Select     Errors  Paths  Active
    0          fscsi0  NORMAL   ACTIVE        386          0      3       3
    1          fscsi1  NORMAL   ACTIVE        449          0      3       3

AIX1#datapath query device
Total Devices : 2
DEV#:   0  DEVICE NAME: vpath0  TYPE: 2105F20         POLICY:    Optimized
SERIAL: 40514830
==========================================================================
Path#      Adapter/Hard Disk          State     Mode     Select     Errors
    0          fscsi0/hdisk2           OPEN   NORMAL         31          0
    1          fscsi1/hdisk4           OPEN   NORMAL         37          0

DEV#:   1  DEVICE NAME: vpath1  TYPE: 2145            POLICY:    Optimized
SERIAL: 600507680183001AD00000000000000A
==========================================================================
Path#      Adapter/Hard Disk          State     Mode     Select     Errors
    0          fscsi0/hdisk3           OPEN   NORMAL        355          0
    1          fscsi0/hdisk5           OPEN   NORMAL          0          0
    2          fscsi1/hdisk6           OPEN   NORMAL        412          0
    3          fscsi1/hdisk7           OPEN   NORMAL          0          0

AIX1#lsvpcfg
vpath0 (Avail pv essdisk1vg) 40514830 = hdisk2 (Avail ) hdisk4 (Avail )
vpath1 (Avail pv essdisk2vg) 600507680183001AD00000000000000A = hdisk3 (Avail ) hdisk5
(Avail ) hdisk6 (Avail ) hdisk7 (Avail )

AIX1#varyonvg essdisk2vg

AIX1#lsvg -o
essdisk2vg
essdisk1vg
rootvg

AIX1#lsvg -l essdisk2vg
essdisk2vg:
LV NAME             TYPE       LPs    PPs    PVs  LV STATE      MOUNT POINT
essdisk2lv          jfs        10     10     1    open/syncd    /essdisk2fs
loglv01             jfslog     1      1      1    open/syncd    N/A

AIX1#lsvg essdisk2vg
VOLUME GROUP:   essdisk2vg              VG IDENTIFIER:  0007905800004c00000000f5573c6ecd
```

```
VG STATE:       active                  PP SIZE:        8 megabyte(s)
VG PERMISSION:  read/write              TOTAL PPs:      595 (4760 megabytes)
MAX LVs:        256                     FREE PPs:       584 (4672 megabytes)
LVs:            2                       USED PPs:       11 (88 megabytes)
OPEN LVs:       2                       QUORUM:         2
TOTAL PVs:      1                       VG DESCRIPTORS: 2
STALE PVs:      0                       STALE PPs:      0
ACTIVE PVs:     1                       AUTO ON:        yes
MAX PPs per PV: 1016                    MAX PVs:        32
LTG size:       128 kilobyte(s)         AUTO SYNC:      no
HOT SPARE:      no

AIX1#fsck -y /essdisk2fs
** Checking /dev/ressdisk2lv (/essdi)
** Phase 0 - Check Log
log redo processing for /dev/ressdisk2lv
** Phase 1 - Check Blocks and Sizes
** Phase 2 - Check Pathnames
** Phase 3 - Check Connectivity
** Phase 4 - Check Reference Counts
** Phase 5 - Check Inode Map
** Phase 6 - Check Block Map
9 files 5232 blocks 158608 free

AIX1#mount /essdisk2fs

AIX1#df -k
Filesystem    1024-blocks      Free %Used    Iused %Iused Mounted on
/dev/hd4            16384      7560   54%     1301    16% /
/dev/hd2          1867776   1265392   33%    20526     5% /usr
/dev/hd9var         16384      9880   40%      437    11% /var
/dev/hd3            32768     29136   12%       67     1% /tmp
/dev/hd1            16384     15820    4%       18     1% /home
/proc                   -         -    -         -     -  /proc
/dev/hd10opt        32768     26592   19%      294     4% /opt
/dev/essdisk2lv       81920     79304    4%       18     1% /essdisk2fs

AIX1#cd /essdisk2fs
AIX1#ls
file_on_ess.txt  lost+found
AIX1#more file_on_ess.txt
file created on ess with direct connexion on disk2
```

12.6.3 Migrating the vDisk from image mode to managed mode

The vDisk in image mode is still completely made of extents (299) from mDisk9. mDisk7 and mDisk8, with 277 free extents each, are in the same mDisk group.

The purpose of the migration is to migrate some extents from mDisk9 to mDisk8 and mDisk7 using the command:

```
svctask migrateexts
```

This command also migrates the vDisk from image mode to managed mode. See Example 12-8.

Example 12-8 Migrating a vDisk from image mode to managed mode

```
IBM_2145:admin>svcinfo lsvdiskextent vDisk0
id               number_extents
9                299

>>>>  299 extents for the vDisk which are in mDisk 9

IBM_2145:admin>svcinfo lsfreeextents mDisk9
id 9
number_of_extents 0

IBM_2145:admin>svcinfo lsfreeextents mDisk8
id 8
number_of_extents 297

IBM_2145:admin>svcinfo lsfreeextents mDisk7
id 7
number_of_extents 297

>>>>>>>>>>>>>>>>>>>>  you know that 299 extents of mDisk 9 are in vDisk0 (number 0)

IBM_2145:admin>svcinfo lsmdiskextent mDisk9
id               number_of_extents
0                299
>>>>>>> no vDisk use extents from mDisk 7 and 8
IBM_2145:admin>svcinfo lsmdiskextent mDisk8
IBM_2145:admin>svcinfo lsmdiskextent mDisk7
>>>>>>>>>>>>  start migration

IBM_2145:admin>svcinfo lsmigrate
>>>>>> no migration have started yet

IBM_2145:admin>svctask migrateexts -source mDisk9 -exts 100 -target mDisk7 -vDisk vDisk0

>>>>>>>>>>>>> The migration has started

IBM_2145:admin>svcinfo lsmigrate
migrate_type mDisk_Extents_Migration
progress 1
migrate_vDisk_index 0
migrate_source_mDisk_index 9
migrate_target_mDisk_index 7
number_extents 100
max_thread_count 4

IBM_2145:admin>svcinfo lsmigrate
migrate_type mDisk_Extents_Migration
progress 5
migrate_vDisk_index 0
migrate_source_mDisk_index 9
migrate_target_mDisk_index 7
number_extents 100
max_thread_count 4

IBM_2145:admin>svcinfo lsmigrate
migrate_type mDisk_Extents_Migration
progress 7
migrate_vDisk_index 0
migrate_source_mDisk_index 9
migrate_target_mDisk_index 7
```

```
number_extents 100
max_thread_count 4

IBM_2145:admin>svcinfo lsmigrate
migrate_type mDisk_Extents_Migration
progress 35
migrate_vDisk_index 0
migrate_source_mDisk_index 9
migrate_target_mDisk_index 7
number_extents 100
max_thread_count 4

IBM_2145:admin>svcinfo lsmigrate
migrate_type mDisk_Extents_Migration
progress 85
migrate_vDisk_index 0
migrate_source_mDisk_index 9
migrate_target_mDisk_index 7
number_extents 100
max_thread_count 4

IBM_2145:admin>svcinfo lsmigrate
migrate_type mDisk_Extents_Migration
progress 92
migrate_vDisk_index 0
migrate_source_mDisk_index 9
migrate_target_mDisk_index 7
number_extents 100
max_thread_count 4

IBM_2145:admin>svcinfo lsmigrate

>>>>>>>>>>>>>>>  migration done the vDisk is now managed >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>

IBM_2145:admin>svcinfo lsmdisk -delim :
id:name:status:mode:mDisk_grp_id:mDisk_grp_name:capacity:ctrl_LUN_#:controller_name
0:mDisk0:online:unmanaged:::167.0GB:0000000000000000:F600_23C1436
1:mDisk1:online:unmanaged:::200.4GB:0000000000000001:F600_23C1436
2:mDisk2:online:unmanaged:::167.0GB:0000000000000002:F600_23C1436
3:mDisk3:online:unmanaged:::200.4GB:0000000000000003:F600_23C1436
4:mDisk4:online:unmanaged:::339.3GB:0000000000000004:F600_23C1436
5:mDisk5:online:unmanaged:::407.2GB:0000000000000005:F600_23C1436
6:mDisk6:online:unmanaged:::500.0GB:0000000000000006:F600_23C1436
7:mDisk7:online:managed:1:ess_mDiskgrp0:4.7GB:0000000000000000:ESS
8:mDisk8:online:managed:1:ess_mDiskgrp0:4.7GB:0000000000000001:ESS
9:mDisk9:online:managed:1:ess_mDiskgrp0:4.7GB:0000000000000002:ESS

>>>>>>>>>>>>>>>>> Status of the extents after the first migration

IBM_2145:admin>svcinfo lsmdiskextent mDisk7
id               number_of_extents
0                100

IBM_2145:admin>svcinfo lsmdiskextent mDisk8

IBM_2145:admin>svcinfo lsmdiskextent mDisk9
id               number_of_extents
0                199

IBM_2145:admin>svcinfo lsvdiskextent vDisk0
```

```
id               number_extents
7                100
9                199

>>>>>>>>>>>>  new migration from mDisk 9 to mDisk 8 for vDisk vDisk0 with 100 extents

IBM_2145:admin>svctask migrateexts -source mDisk9 -exts 100 -target mDisk8 -vDisk vDisk0

IBM_2145:admin>svcinfo lsmigrate
migrate_type mDisk_Extents_Migration
progress 4
migrate_vDisk_index 0
migrate_source_mDisk_index 9
migrate_target_mDisk_index 8
number_extents 100
max_thread_count 4

IBM_2145:admin>svcinfo lsmigrate
migrate_type mDisk_Extents_Migration
progress 51
migrate_vDisk_index 0
migrate_source_mDisk_index 9
migrate_target_mDisk_index 8
number_extents 100
max_thread_count 4

IBM_2145:admin>svcinfo lsmigrate
migrate_type mDisk_Extents_Migration
progress 65
migrate_vDisk_index 0
migrate_source_mDisk_index 9
migrate_target_mDisk_index 8
number_extents 100
max_thread_count 4

IBM_2145:admin>svcinfo lsmigrate

>>>> Migration is done >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>

IBM_2145:admin>svcinfo lsvdiskextent vDisk0
id               number_extents
7                100
8                100
9                99

IBM_2145:admin>svcinfo lsmdiskextent mDisk7
id               number_of_extents
0                100

IBM_2145:admin>svcinfo lsmdiskextent mDisk8
id               number_of_extents
0                100

IBM_2145:admin>svcinfo lsmdiskextent mDisk9
id               number_of_extents
0                99

IBM_2145:admin>svcinfo lsfreeextents mDisk7
id 7
number_of_extents 197
```

```
IBM_2145:admin>svcinfo lsfreeextents mDisk8
id 8
number_of_extents 197

IBM_2145:admin>svcinfo lsfreeextents mDisk9
id 9
number_of_extents 198

IBM_2145:admin>svcinfo lsmdiskgrp -delim :
id:name:status:mDisk_count:vDisk_count:capacity:extent_size:free_capacity
1:ess_mDiskgrp0:online:3:1:13.9GB:16:9.2GB

IBM_2145:admin>svcinfo lsmdisk -delim :
id:name:status:mode:mDisk_grp_id:mDisk_grp_name:capacity:ctrl_LUN_#:controller_name
0:mDisk0:online:unmanaged:::167.0GB:0000000000000000:F600_23C1436
1:mDisk1:online:unmanaged:::200.4GB:0000000000000001:F600_23C1436
2:mDisk2:online:unmanaged:::167.0GB:0000000000000002:F600_23C1436
3:mDisk3:online:unmanaged:::200.4GB:0000000000000003:F600_23C1436
4:mDisk4:online:unmanaged:::339.3GB:0000000000000004:F600_23C1436
5:mDisk5:online:unmanaged:::407.2GB:0000000000000005:F600_23C1436
6:mDisk6:online:unmanaged:::500.0GB:0000000000000006:F600_23C1436
7:mDisk7:online:managed:1:ess_mDiskgrp0:4.7GB:0000000000000000:ESS
8:mDisk8:online:managed:1:ess_mDiskgrp0:4.7GB:0000000000000001:ESS
9:mDisk9:online:managed:1:ess_mDiskgrp0:4.7GB:0000000000000002:ESS
10:mDisk10:online:unmanaged:::4.7GB:0000000000000003:ESS

IBM_2145:admin>svcinfo lsvdisk -delim :
id:name:IO_group_id:IO_group_name:status:mDisk_grp_id:mDisk_grp_name:capacity:type:FC_id:FC
_name:RC_id:RC_name
0:vDisk0:0:io_grp0:online:1:ess_mDiskgrp0:4.7GB:striped::::

IBM_2145:admin>svcinfo lsvdisk vDisk0
id 0
name vDisk0
IO_group_id 0
IO_group_name io_grp0
status online
mDisk_grp_id 1
mDisk_grp_name ess_mDiskgrp0
capacity 4.7GB
type striped
formatted no
mDisk_id
mDisk_name
FC_id
FC_name
RC_id
RC_name
throttling 0
preferred_node_id 2
```

During the migration, access to essdisk2fs was possible and there was no impact on the AIX host system. The migration process is transparent to the host system.

12.7 Data migration from ESS to SVC using the GUI

This configuration is the same as the one shown in Figure 12-17 on page 423. This scenario consists of the migration of one volume (from ESS label S:) of a Windows 2000 host from a ESS to a SVC. The ESS is before and after the migration connected directly the Windows 2000 host (via the 2109 switches).

Before the migration, the LUN masking is defined in the ESS to give access to the Windows 2000 host system for the volume from ESS label S:. *After* the migration, LUN masking is defined in the ESS to give access to the SVC nodes for the volume from ESS label S:.

The following actions occur for the migration:

1. Shut down the Windows 2000 host system before changing LUN masking in the ESS.
2. The volume is first discovered as a "mDisk unmanaged" by the SVC. It is mDisk10.
3. A vDisk is created in image mode using this mDisk.
4. This new vDisk is mapped to the host system Win2K.
5. Restart the Windows 2000 host system.
6. The vDisk is again available for the host system Win2K.

The vDisk can be migrated from image mode to managed mode, concurrently to the access to the Win2K host system. The mode of migration is to migrate a vDisk from an mDisk group to another mDisk group.

12.7.1 Windows 2000 host system connected directly to the ESS

Figure 12-18 shows the disk from the ESS.

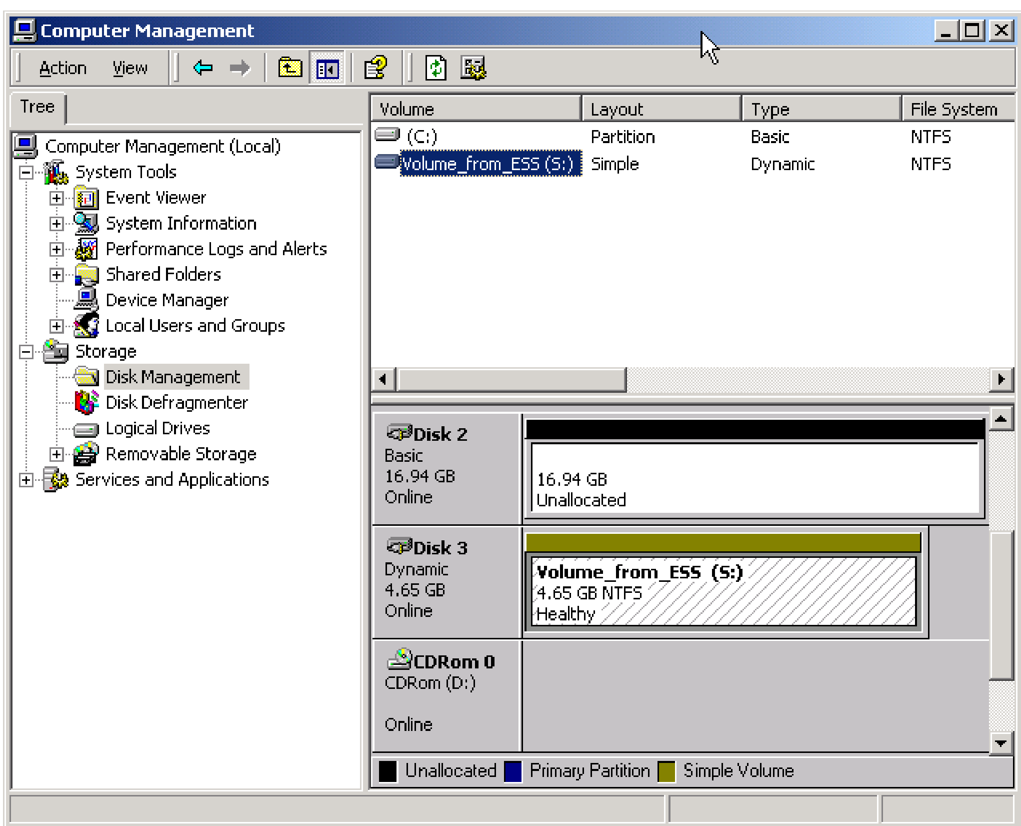


Figure 12-18 Disk management: One volume from ESS with label S:

Figure 12-19 shows the properties using the Subsystem Device Driver (SDD).

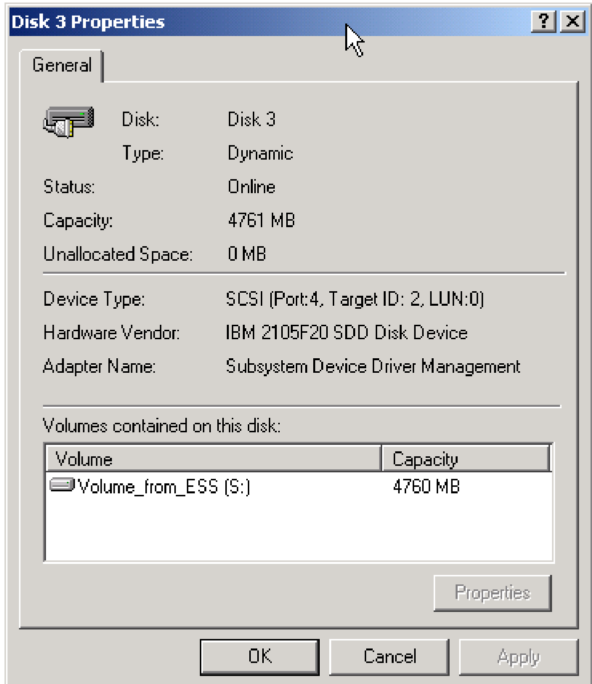


Figure 12-19 Drive S: from ESS with SDD

Figure 12-20 shows the volume properties.

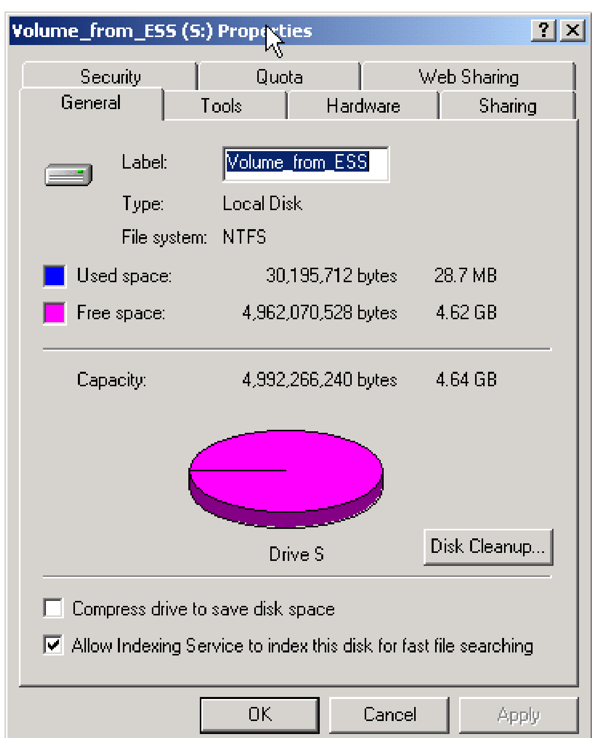


Figure 12-20 Volume properties of Drive S

Figure 12-21 shows the files on volume S.

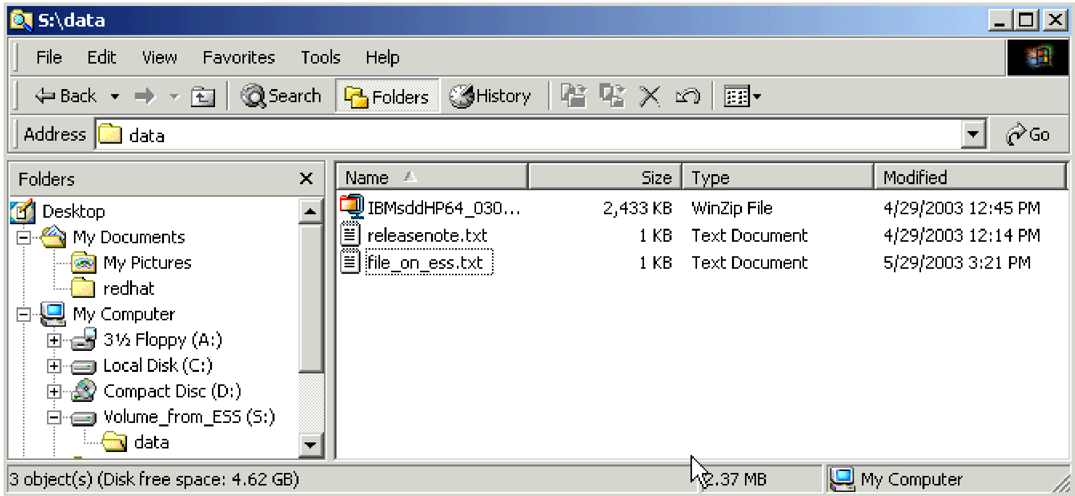


Figure 12-21 Files on volume S: (Volume from ESS)

12.7.2 The SVC added between the Windows 2000 host system and the ESS

After you change LUN masking in the ESS, a new mDisk is discovered in the SVC—mDisk10. A vDisk named vDisk0 already exists. It was created after migrating one hdisk from a host system.

Create a new vDisk named winimagevDisk1 in image mode using the mDisk10 in the mDisk group ess_mDiskgrp0:

1. As shown in Figure 12-22, select **Create a vDisk** from the list and click **Go**.

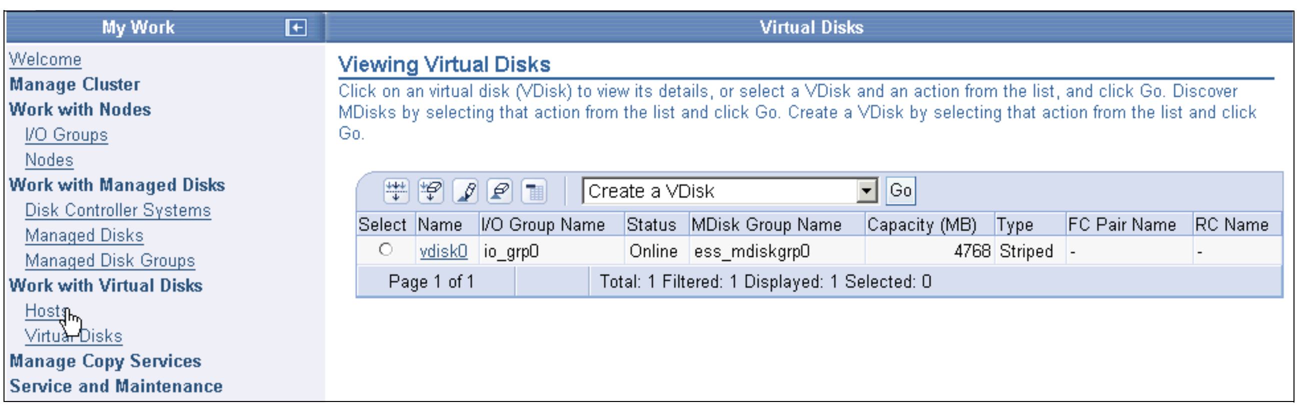


Figure 12-22 Viewing Virtual Disks

2. The Create Virtual Disks panel (Figure 12-23) of the wizard is displayed. Click **Next**.

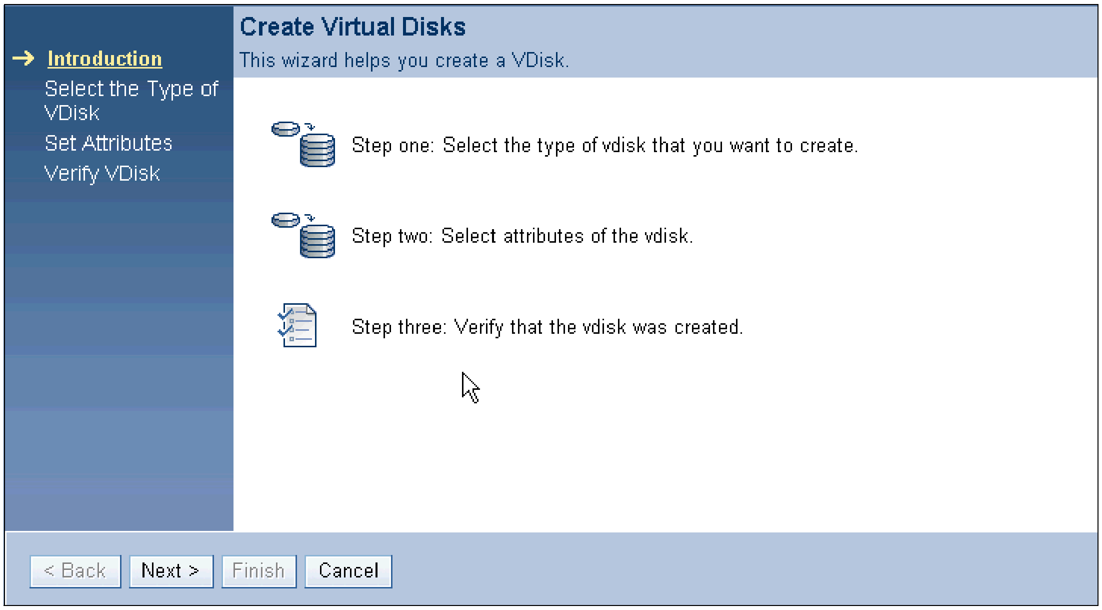


Figure 12-23 Create Virtual Disks

3. On the Select the Type of vDisk panel (Figure 12-24), type the name of the disk, select the I/O group, select the mDisk group, and select the type of vDisk. Then click **Next**.

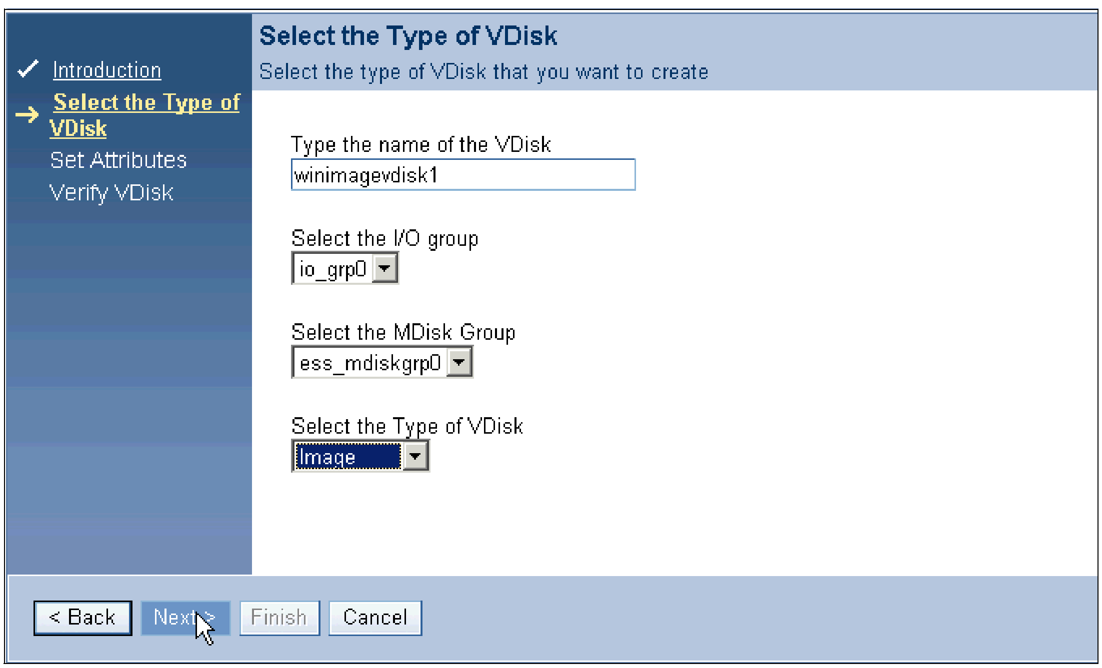


Figure 12-24 Select the Type of vDisk

4. On the Select Attributes for Image-mode vDisk panel (Figure 12-25), select the preferred node for I/O and the managed disk used to create the vDisk. Then click **Next**.

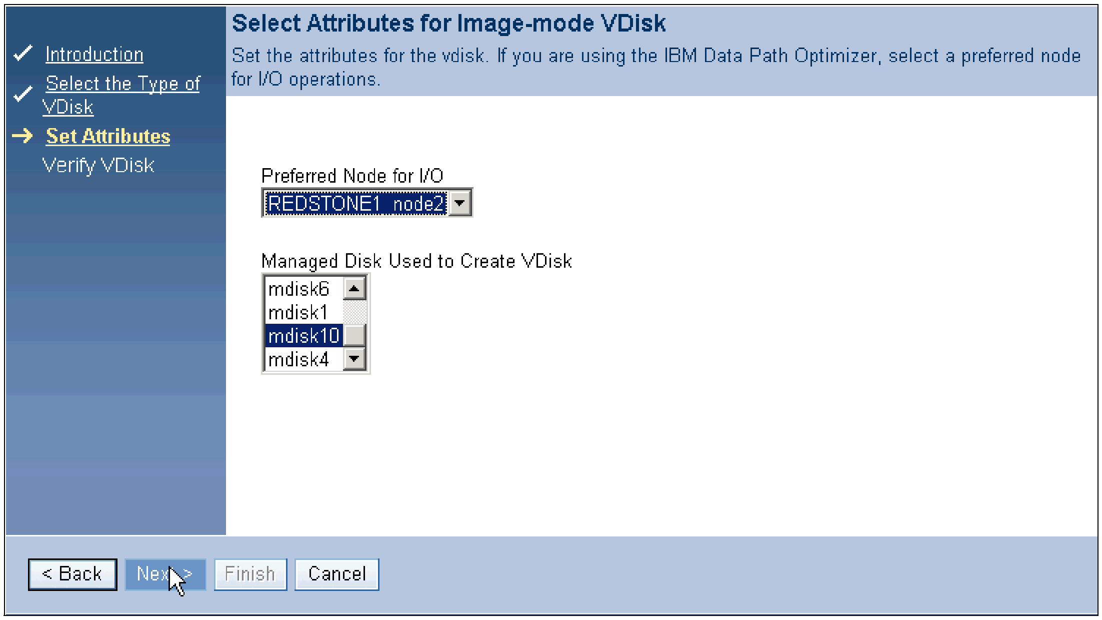


Figure 12-25 Select Attributes for Image-mode vDisk

5. Verify the options that you selected as shown in Figure 12-26. Click **Finish**.

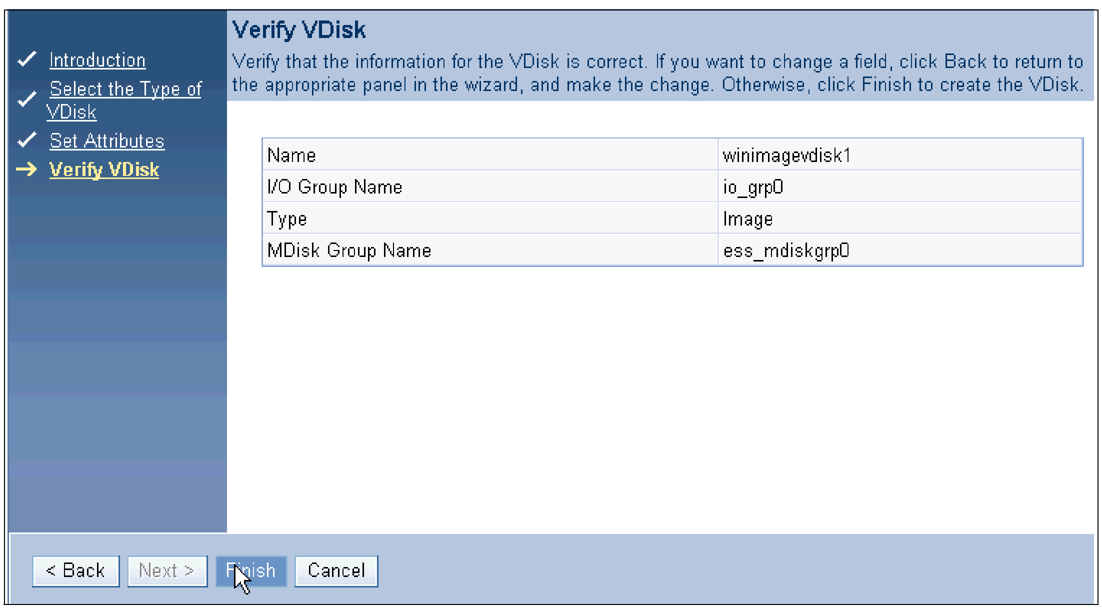


Figure 12-26 Verify vDisk

6. You can view the vDisk that you created as shown in Figure 12-27.

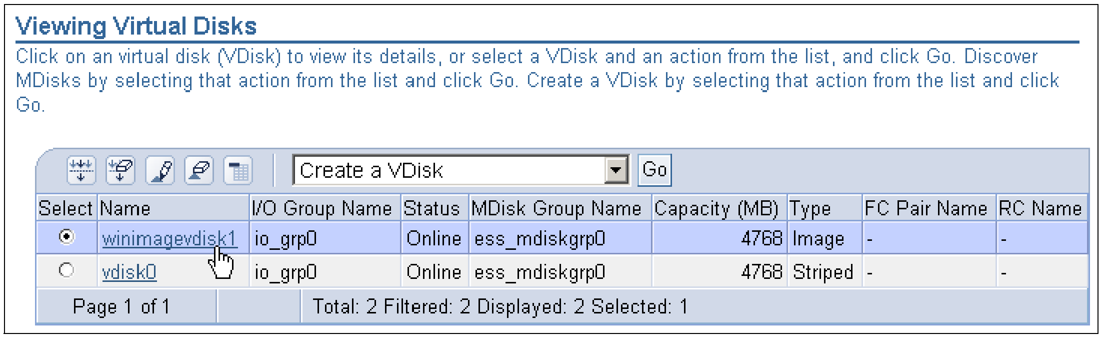


Figure 12-27 Viewing Virtual Disks

The managed disk view is shown in Figure 12-28.

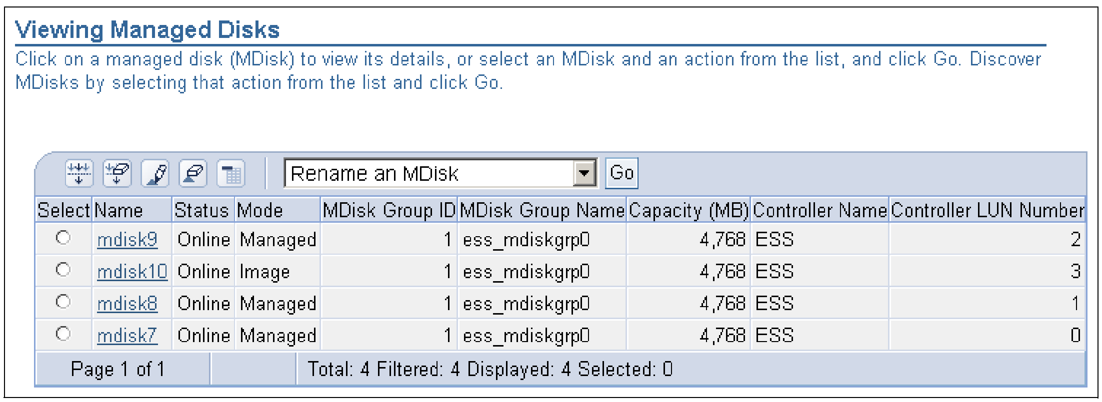


Figure 12-28 Viewing Managed Disks

7. Map the vDisk again to the Windows 2000 host system “WIN2K”. As shown in Figure 12-29, select the name of the vDisk. Then select **Map a vDisk to a host** and click **Go**.

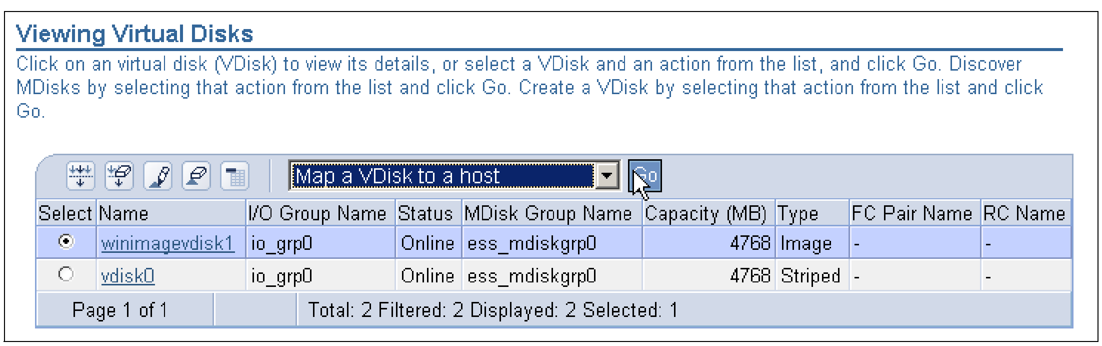


Figure 12-29 Viewing Virtual Disks

8. Select the target host as shown in Figure 12-30 and click **OK**.

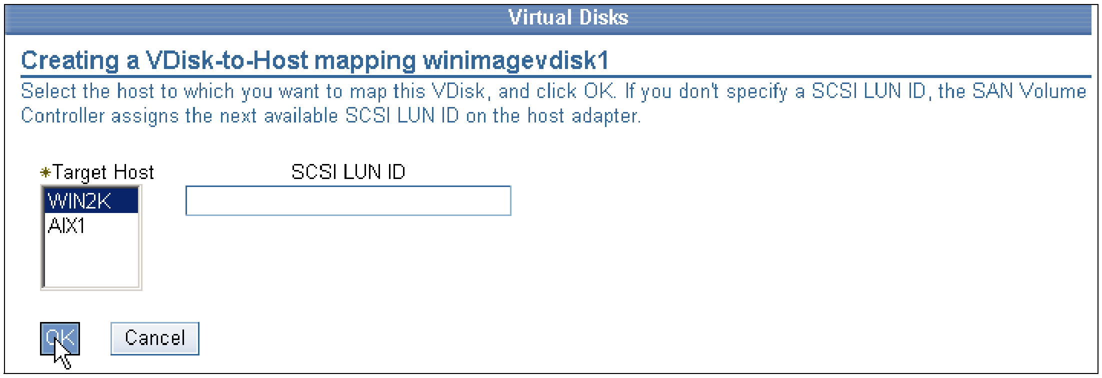


Figure 12-30 Creating a vDisk-to-Host mapping winimagevDisk1

9. Restart the Windows 2000 host system. Figure 12-31 shows the result.

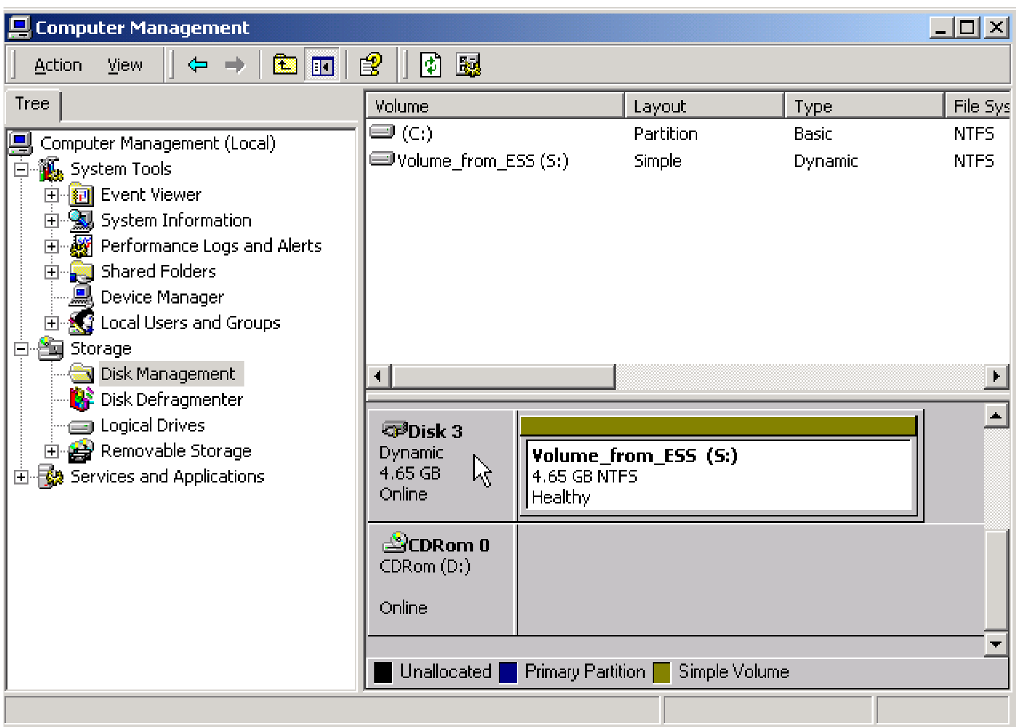


Figure 12-31 The volume S: Volume_from_ESS is online

Figure 12-32 shows the disk properties.

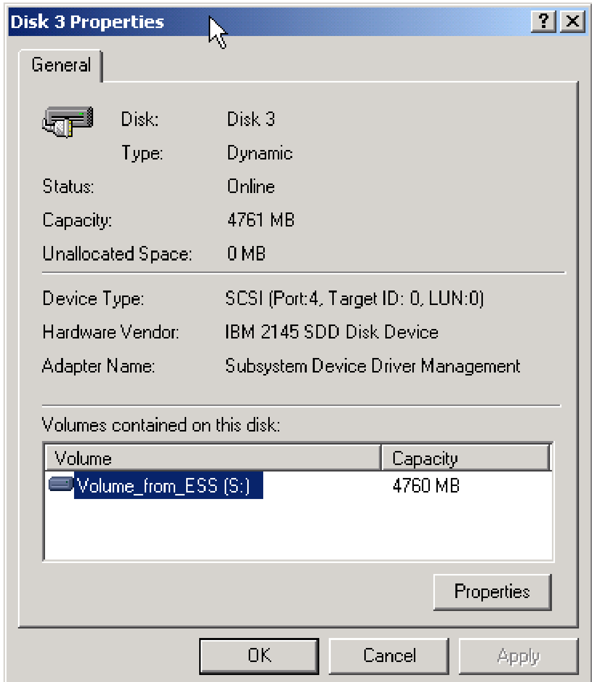


Figure 12-32 The volume S: Volume_from_ESS is online and is 2145 SDD Disk device

The volume is now online with the same data as before the migration. See Figure 12-33.

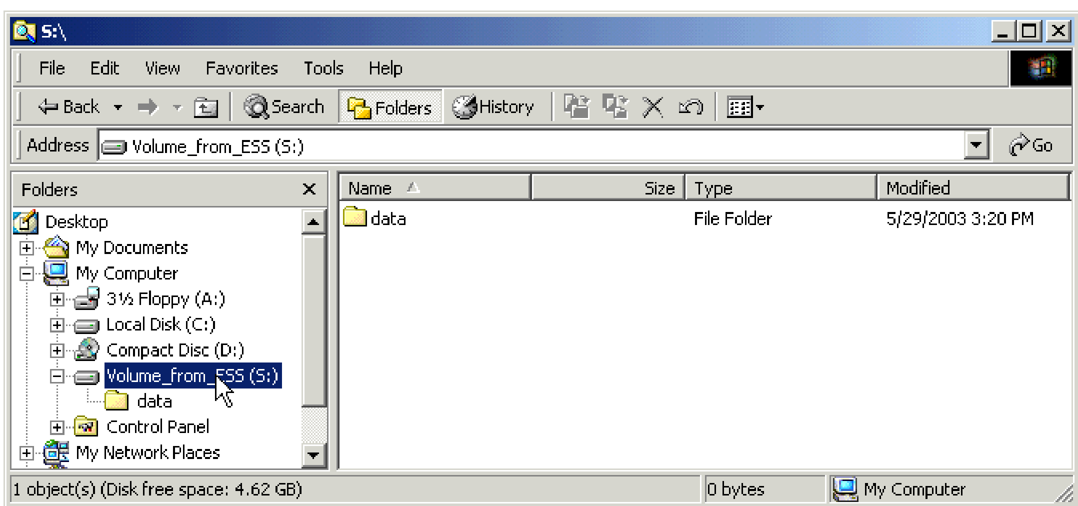


Figure 12-33 The volume S: Volume_from_ESS is online with the same data

Figure 12-34 shows the properties.

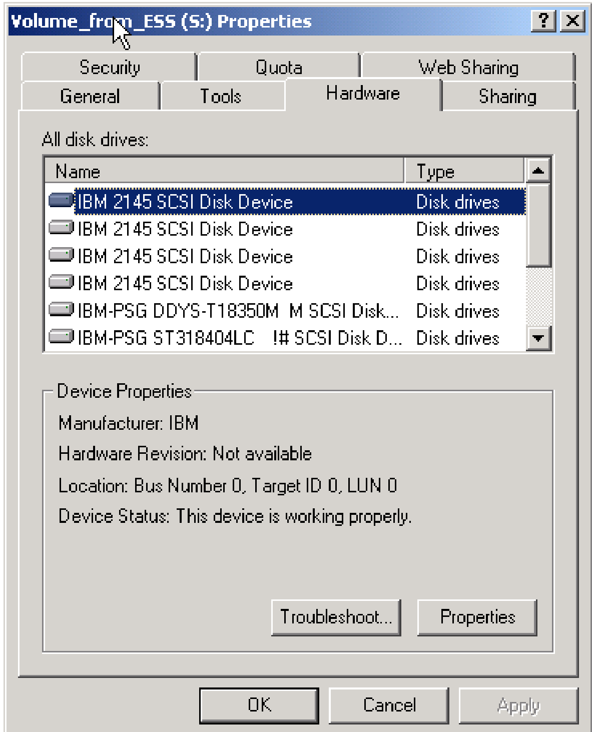


Figure 12-34 The volume S: Volume_from_ESS is online and have four paths to SVC 2145

12.7.3 Migrating the vDisk from image mode to managed mode

Now the vDisk is migrated to managed mode by migrating the completed vDisk from the MDG ess_mDiskgrp0 to the MDG Migrated _vDisk. This MDG is based on FAStT 600 back-end storage instead of ESS. It consists of mDisk4 and mDisk5.

1. As shown in Figure 12-35, select the vDisk. Then select **Migrate a vDisk** from the list and click **Go**.

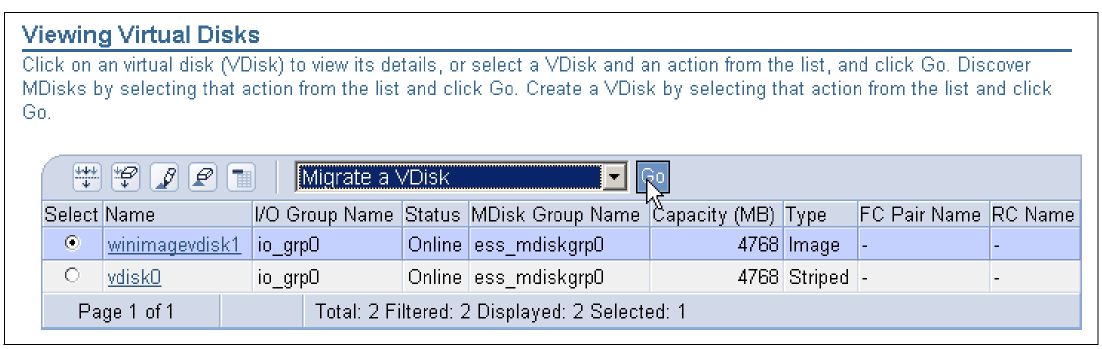


Figure 12-35 Viewing Virtual Disks

2. Select the MDG to which to migrate the disk as shown in Figure 12-36. Click **OK**.

Figure 12-36 Migrating Virtual Disks-winimagevdisk1

3. You can now view the MDG as shown in Figure 12-37.

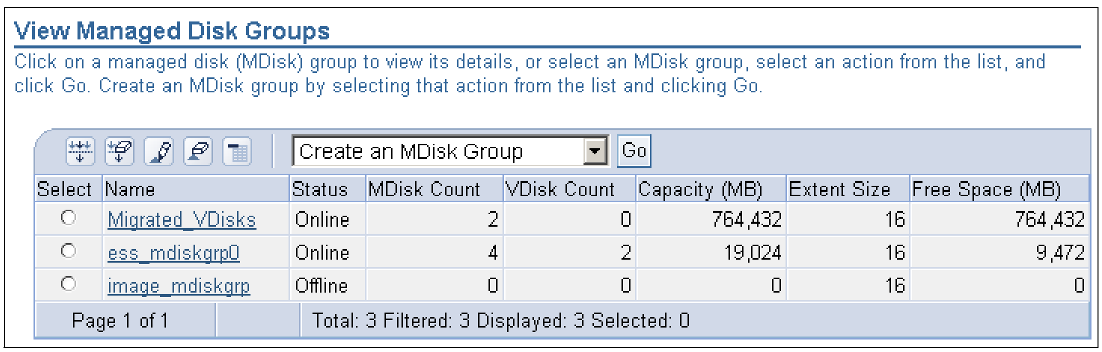


Figure 12-37 View Managed Disk Groups

Before the migration is complete, the vDisk still belongs to ess_mDiskgrp0 as shown in Figure 12-38.

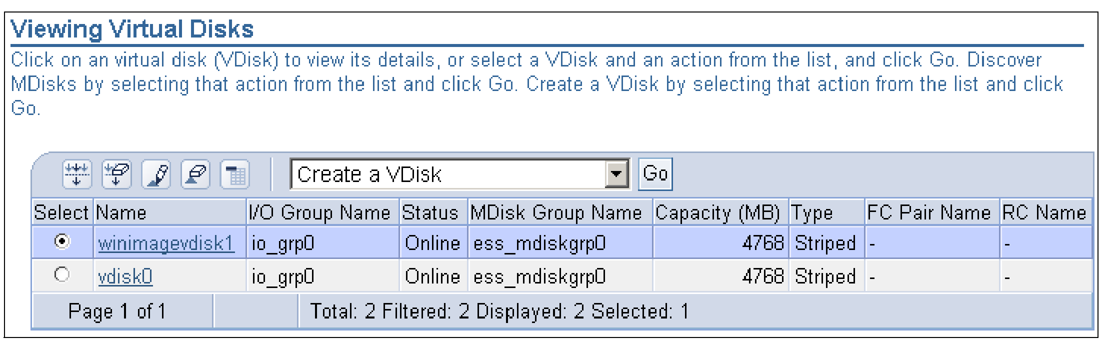
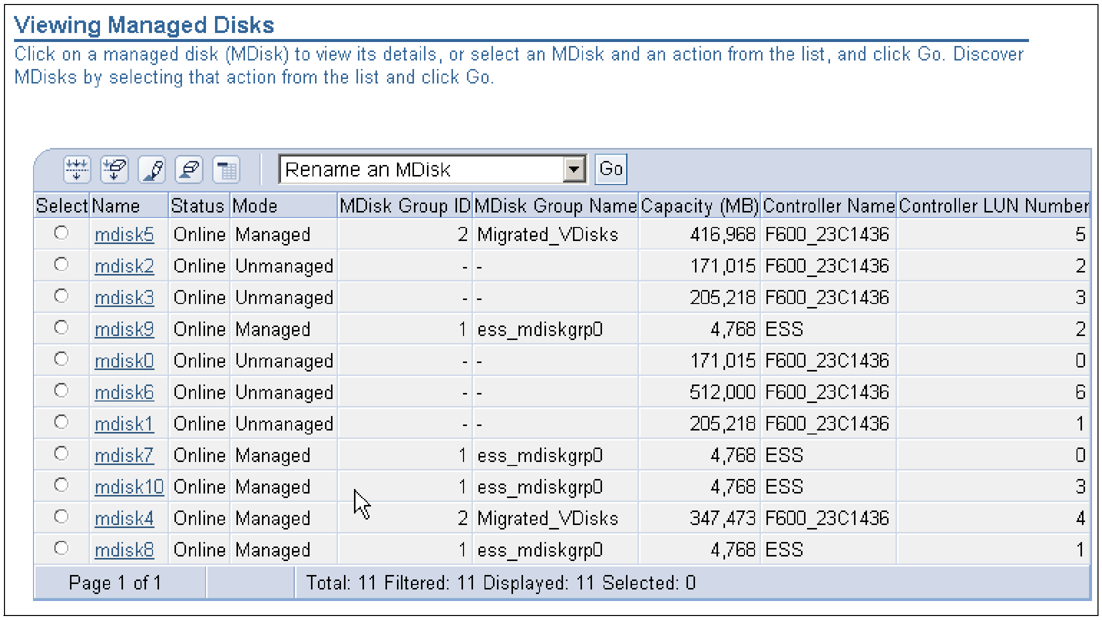


Figure 12-38 Viewing Virtual Disks

After the migration is complete, you see the results shown in Figure 12-39.



Viewing Managed Disks

Click on a managed disk (MDisk) to view its details, or select an MDisk and an action from the list, and click Go. Discover MDisks by selecting that action from the list and click Go.

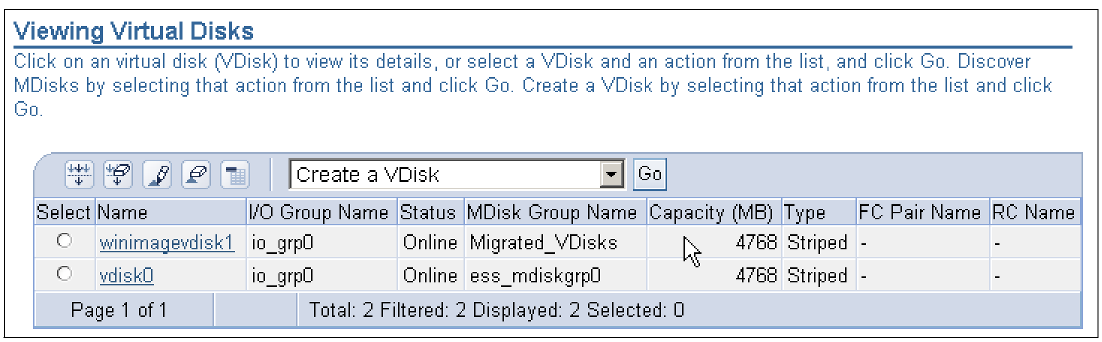
Rename an MDisk | Go

Select	Name	Status	Mode	MDisk Group ID	MDisk Group Name	Capacity (MB)	Controller Name	Controller LUN Number
○	mdisk5	Online	Managed	2	Migrated_VDisks	416,968	F600_23C1436	5
○	mdisk2	Online	Unmanaged	-	-	171,015	F600_23C1436	2
○	mdisk3	Online	Unmanaged	-	-	205,218	F600_23C1436	3
○	mdisk9	Online	Managed	1	ess_mdiskgrp0	4,768	ESS	2
○	mdisk0	Online	Unmanaged	-	-	171,015	F600_23C1436	0
○	mdisk6	Online	Unmanaged	-	-	512,000	F600_23C1436	6
○	mdisk1	Online	Unmanaged	-	-	205,218	F600_23C1436	1
○	mdisk7	Online	Managed	1	ess_mdiskgrp0	4,768	ESS	0
○	mdisk10	Online	Managed	1	ess_mdiskgrp0	4,768	ESS	3
○	mdisk4	Online	Managed	2	Migrated_VDisks	347,473	F600_23C1436	4
○	mdisk8	Online	Managed	1	ess_mdiskgrp0	4,768	ESS	1

Page 1 of 1 | Total: 11 Filtered: 11 Displayed: 11 Selected: 0

Figure 12-39 Viewing Managed Disks

Viewing the vDisks gives the results shown in Figure 12-40.



Viewing Virtual Disks

Click on an virtual disk (VDisk) to view its details, or select a VDisk and an action from the list, and click Go. Discover MDisks by selecting that action from the list and click Go. Create a VDisk by selecting that action from the list and click Go.

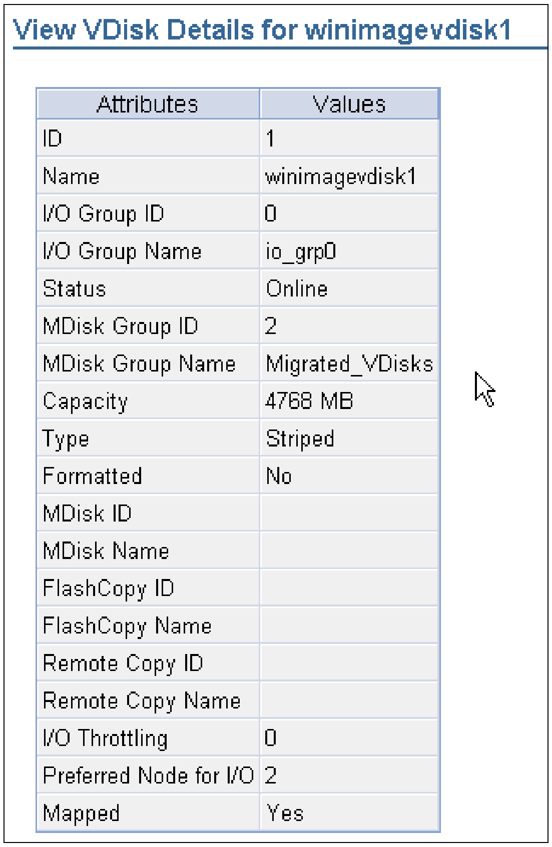
Create a VDisk | Go

Select	Name	I/O Group Name	Status	MDisk Group Name	Capacity (MB)	Type	FC Pair Name	RC Name
○	winimagevdisk1	io_grp0	Online	Migrated_VDisks	4768	Striped	-	-
○	vdisk0	io_grp0	Online	ess_mdiskgrp0	4768	Striped	-	-

Page 1 of 1 | Total: 2 Filtered: 2 Displayed: 2 Selected: 0

Figure 12-40 vDisk is now in Migrated_vDisks instead of ess_mDiskgrp0

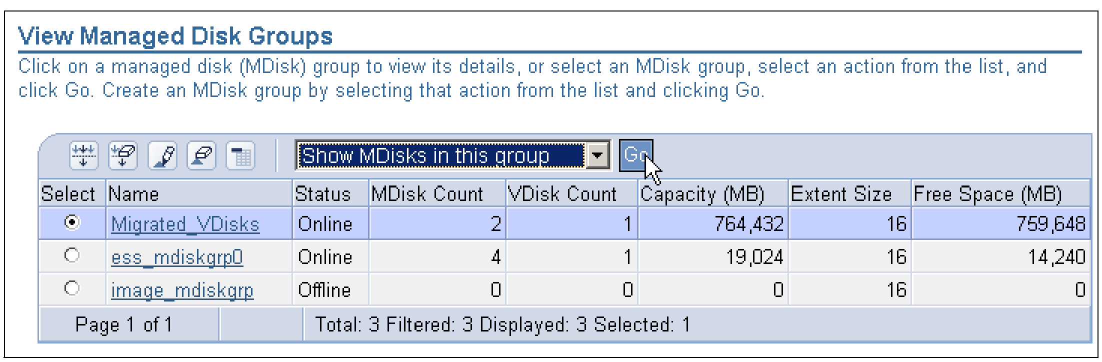
Figure 12-41 shows the details of the vDisk.


View VDisk Details for winimagevdisk1

Attributes	Values
ID	1
Name	winimagevdisk1
I/O Group ID	0
I/O Group Name	io_grp0
Status	Online
MDisk Group ID	2
MDisk Group Name	Migrated_VDisks
Capacity	4768 MB
Type	Striped
Formatted	No
MDisk ID	
MDisk Name	
FlashCopy ID	
FlashCopy Name	
Remote Copy ID	
Remote Copy Name	
I/O Throttling	0
Preferred Node for I/O	2
Mapped	Yes

Figure 12-41 Details for vDisk winimagevDisk1

Viewing the MDGs after the migration, you see the results shown in Figure 12-42.


View Managed Disk Groups

Click on a managed disk (MDisk) group to view its details, or select an MDisk group, select an action from the list, and click Go. Create an MDisk group by selecting that action from the list and clicking Go.

Show MDisks in this group | Go

Select	Name	Status	MDisk Count	VDisk Count	Capacity (MB)	Extent Size	Free Space (MB)
◉	Migrated_VDisks	Online	2	1	764,432	16	759,648
○	ess_mdiskgrp0	Online	4	1	19,024	16	14,240
○	image_mdiskgrp	Offline	0	0	0	16	0

Page 1 of 1 | Total: 3 Filtered: 3 Displayed: 3 Selected: 1

Figure 12-42 The MDGs after complete migration

Finally viewing the managed disks, you see the information as shown in Figure 12-43.

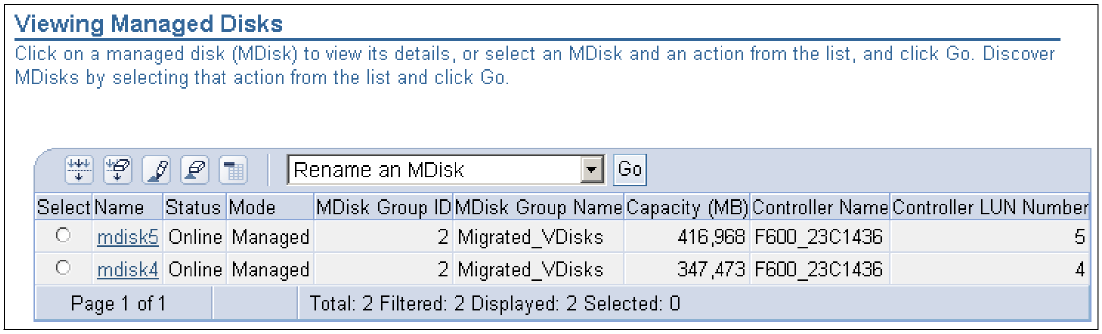


Figure 12-43 The managed disks after migration is complete

13

IBM TotalStorage SAN Integration Server

This chapter discusses the IBM TotalStorage SAN Integration Server (SIS) solution. It looks at the configurations and implications on configuring back-end storage.

13.1 Introduction

The SAN Integration Server (Figure 13-1) is a preconfigured, complete storage area network (SAN) solution with storage virtualization. A SIS offers the function of a storage subsystem housed in one or more 19-inch racks. It contains the SAN Volume Controller technology along with Fibre Channel switches, IBM TotalStorage Fibre Array Storage Technology (FAStT) storage servers, master console, and Ethernet switches. These features are all delivered and installed as a single unit, although SIS includes IBM components that are available individually.

Figure 13-1 IBM TotalStorage SAN Integration Server

Orderable SIS configurations address different storage capacity and SAN connectivity requirements. They contain either one or two pairs of SAN Volume Controllers (SVCs). Standard upgrade paths are available to provide storage capacity, connectivity, and performance growth. The solution can leverage the built-in SVC to virtualize the storage from other supported external storage subsystems. SIS can also be integrated with other SIS or SVC configurations for Remote Copy functions.

SAN Integration Server offers the following features to attached hosts:

- A single view of attached storage
- Logical Unit virtualization
- Large, scalable cache
- Scalable performance
- Point-in-time copy (optional)
- Synchronous Remote Copy (optional)
- Applications are provided to individually manage the devices that comprise a SAN Integration Server configuration

13.2 Configuration

The SAN Integration Server ships preconfigured to match client’s requirements for:

- Storage capacity
- Number of available Fibre Channel connections to attach hosts, external storage subsystems or other SIS/SVC
- Performance (one or two SVC node pairs)

The SIS components are preinstalled in cabinets. SVC nodes, uninterruptible power supply units, SAN and Ethernet switches and master console will be installed into the primary cabinet. The primary cabinet (SIS model V1A) only accepts one FAStT 600 controller and up to three EXP700 expansion units. Each expansion cabinet (SIS model V1B) accepts up to three FAStT 600 controllers and up to nine EXP700. Up to four expansion cabinets can be configured.

Figure 13-2 shows the components that can be configured in the SIS cabinets, including the optional ones.

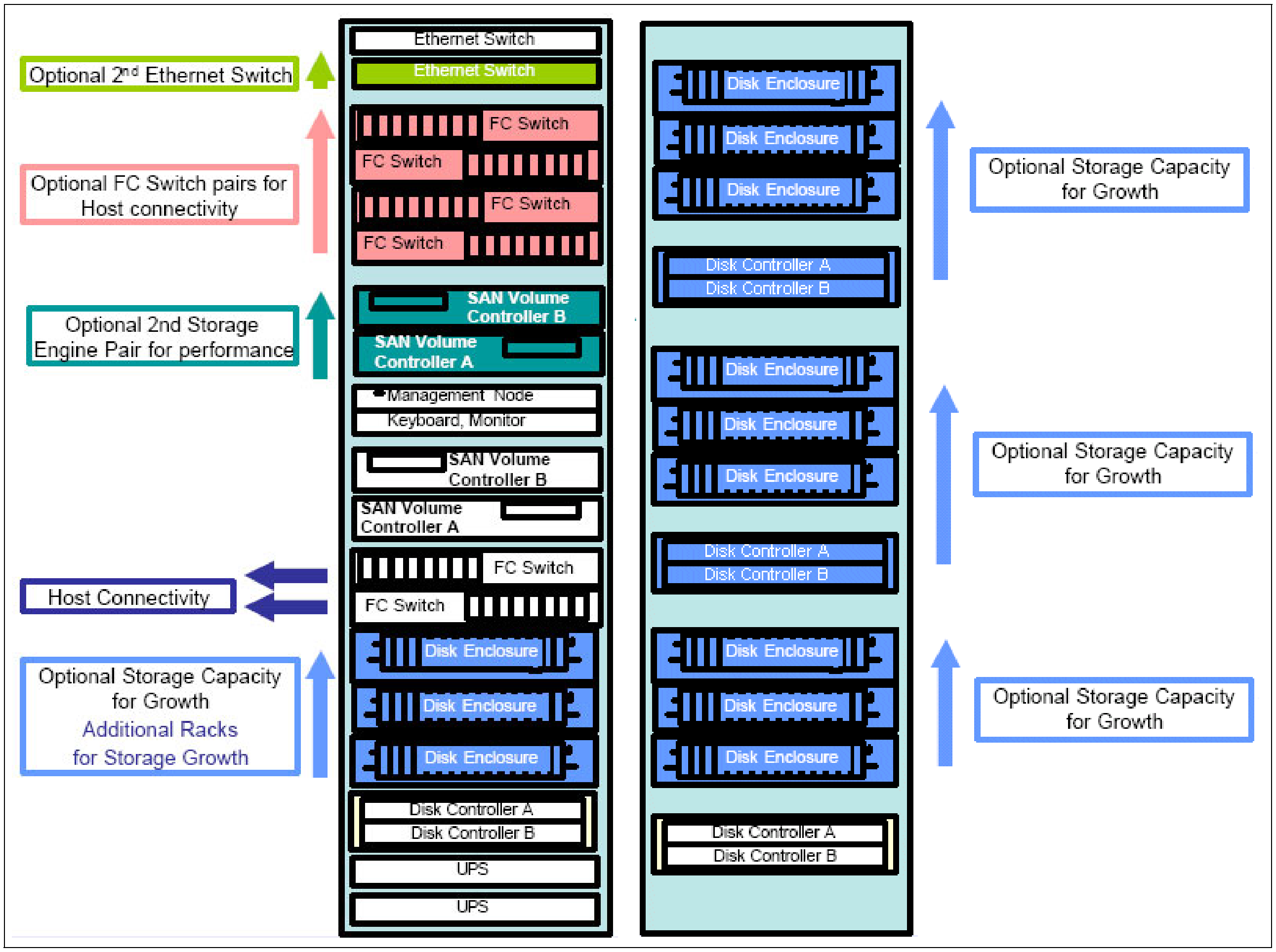


Figure 13-2 Primary and expansion SIS cabinets

All of the components have network Internet Protocol (IP) addresses assigned to them. The Fibre Channel switches are preconfigured and zoned. FAStT storage that is purchased with the SIS comes with standard arrays and volumes already created. If you need to change the preconfigured network addresses, see the *IBM TotalStorage 2146 SAN Integration Server Model V1A User's Guide*, SC26-7556.

13.2.1 FAStT storage

The storage controller that SIS uses is the FAStT 600. The SIS model V1A contains one FAStT 600 controller and up to three EXP700 drawers (56 disks maximum). The model V1B is an expansion rack that can hold up to three FAStT 600 controllers and up to nine EXP700 expansion units (168 disks maximum).

The maximum configurations for SIS are:

- Maximum number of expansion racks: four
- Maximum number of FAStT 600 controllers: 13
- Maximum number of EXP700 expansion units: 39
- Maximum number of physical disks: 728

Each FAStT 600 and EXP700 is set up for RAID 5 arrays as follows:

- One array with five disks + parity (total of six disks)
- One array with six disks + parity (total of seven disks)
- One hot spare disk

Each array has only one logical unit number (LUN) defined and is presented to the SVC as one mDisk. The predefined naming convention that is used describes the size and array type on which the LUN mDisk was created:

- 36 GB 5+p: ARR36P5N01 to ARR36P5N52
- 36 GB 6+p: ARR36P6N01 to ARR36P6N52
- 73 GB 5+p: ARR73P5N01 to ARR73P5N52
- 73 GB 6+p: ARR73P6N01 to ARR73P6N52
- 146 GB 5+p: ARR146P5N01 to ARR146P5N52
- 146 GB 6+p: ARR146P6N01 to ARR146P6N52

See the *IBM TotalStorage 2146 SAN Integration Server Model V1A User's Guide*, SC26-7556, for details. This guide also has recommendations regarding drive performance selection.

13.2.2 SAN switches

SIS can be configured with a minimum of two and a maximum of six SAN switches to provide redundant internal Fibre Channel connectivity between SVC nodes, FAStT controllers, and a master console. Additionally, available ports on those switches for connecting hosts, external storage subsystems, and other SIS or SVC for Remote Copy must be configured.

Current SIS options for SAN switches in the initial order are:

- Two IBM TotalStorage SAN Switches 3534 Model F08 (8 ports)
- Two IBM TotalStorage SAN Switches 2109 Model F16 (16 ports)
- Two IBM TotalStorage SAN Switches 2109 Model F32 (32 ports)

You can optionally upgrade the SIS configuration to include more IBM TotalStorage SAN Switches up to a total of six, in predefined combinations. Other supported external SAN switches can be connected to the internal ones, subject to changes to the zoning.

13.2.3 Ethernet connectivity

One or two Ethernet switches are configured with SIS, depending on the number of FAStT 600 controllers being installed in the SIS cabinets. The purpose of these switches is to provide internal communication between the SIS components for reporting and alert management, independently of the user’s local area network (LAN). The internal SIS LAN can be configured and connected to other LANs by the client to have external access and remotely manage the storage resources of this solution.

13.2.4 SVC software

SIS comes with SVC engines and SVC software. SVC software is licensed to the users in tiers, based on the total managed storage capacity (virtualization) and on the optional copy services functions: FlashCopy, PPRC, or both. Copy services are also charged by the capacity they manage within the SVC Cluster. Two software maintenance options of one or three years are also available.

The SVC software license can be upgraded to reflect capacity growth or to include other functions that were previously not available. Software maintenance options can be renewed when they expire. We recommend that you renew the software maintenance before they expire for continuous support.

> **Note:** If you are connecting external storage subsystems to SIS, their storage capacity being managed by SVC software must be included in your licenses, both for storage virtualization and Copy Services capacities.

13.3 Compatibility

The SIS is capable of supporting Windows NT/2000, AIX, Red Hat Linux, Sun Solaris, and HP-UX hosts. Although SIS is preconfigured with integrated FAStT storage servers, it can also be attached to other supported IBM or non-IBM external storage subsystems to virtualize their capacities or for data migration purposes.

For the latest supported operating systems, hosts, host bus adapters (HBAs), and storage subsystems, see the Supported Hardware List for SIS at:

http://www.storage.ibm.com/support/2146

13.4 Client responsibilities

There are some pre and post-installation tasks required of the client:

- Location/environmental preparation
- Site-specific IP addressing
- Host HBA and Subsystem Device Driver (SDD) installation
- Host Fibre Channel links marked and in place
- Fibre Channel links for Remote Copy installed if required
- Virtual disks (vDisks) creation and mapping to hosts
- Call Home and VPN site customization
- Powering on and off the components in proper order

> **Restriction:** Clients are allowed to make any changes to the SIS within the guidelines for the internal components. However, be aware that any modifications from the factory setting complicate upgrade paths and void the use of IBM tools to configure future upgrade orders. Warranty terms are not affected, but users may be required to remove non-supported attachments during problem resolution.

For more information, see *IBM TotalStorage SAN Integration Server 2146 Installation & Service Guide*, SC26-7557, and *IBM TotalStorage 2146 SAN Integration Server Model V1A User's Guide*, SC26-7556.

14

Master console

This chapter provides an overview of the master console. The master console is included with both the IBM TotalStorage SAN Volume Controller (SVC) and SAN Integration Server (SIS) products. It is provided to help manage your virtualization environment.

This console can be used to manage the SVC clusters using the CIM Agent for SVC by means of ICAT. A trial copy of Tivoli's Basic Management Suite is included with the solution for evaluation. IBM Tivoli SAN Manager Bonus Pack, a part of the suite, discovers and renders your entire storage area network (SAN) into an easily understood topology. All the applications of the rendered storage devices can be launched from the topology, further simplifying the data center. The master console provides the ability to diagnose problems remotely. You can configure it to call a specified phone number or send an e-mail or electronic page if a configured event occurs.

This chapter discusses the hardware and software components, and customization of the master console. See the *IBM TotalStorage Virtualization Family SAN Volume Controller: Configuration Guide*, SC26-7543, for more detailed information. You can locate this book via the SVC Support page by clicking the **Documentation** link at:

`http://www.ibm.com/storage/support/2145`

14.1 Hardware

Figure 14-1 shows the master console installed in the rack.

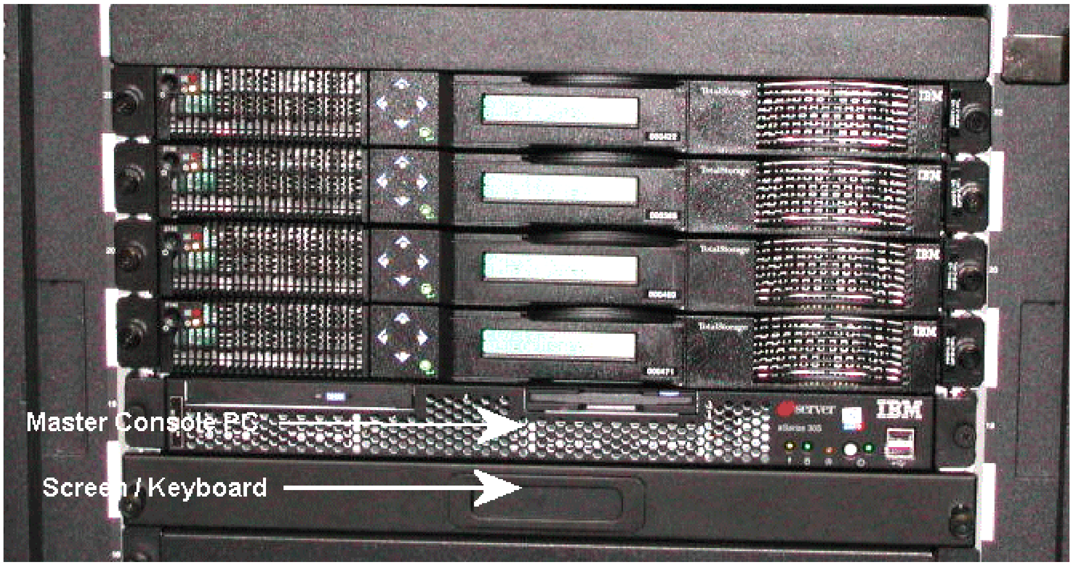


Figure 14-1 Master console in rack

The hardware platform used for the master console is the IBM @server xSeries Model 305. Figure 14-2 shows the screen and keyboard.

Figure 14-2 Master console deployed

The hardware that is supplied as part of the master console includes:

- Pearl (Rack, IU), Uni Northwood/512KB L2
- 2 GHz/QP133 MHz (533 MHz FSB) 256 MB-4 GB 200/266 MHz DDR (four DIMMs) Dual Gigabit Ethernet
- 2 (64-bit/100, 1 half, 1 half LP) PCI-X slots CD, FDD, two Non H/S HDD (IDE)
- IBM 51 MB PC 133 ECC SDRAM module (2 off) = 1 GB
- IBM 40 GB IDE HDD X2
- Netbay 1U Flat Panel Monitor Console Kit
- Space Saver NLS keyboard (English)
- QLogic 2342 FC-2 Host Bus Adapter
- FC: two ports required, one for each of the dual fabrics

14.2 Management console software

The software that is installed in the master console is:

- Microsoft Windows 2000 Server Edition with the latest Service Pack
- Windows 2000 Security Hotfixes
- Windows Internet Explorer
- IBM Tivoli SAN Manager Bonus Pack version
- IBM TotalStorage Fibre Array Storage Technology (FAStT) Storage Manager Client and Utilities
- Host Bus Adapter Driver (QLogic 2343)
- PuTTY utility package
- IBM Director Server
- SAN Volume Controller Console package
- Adobe Acrobat Reader
- Connection Manager (VPN)
- Java™ 2 Runtime Environment Standard Edition

For a complete and current list of the supported software levels in the master console, refer to the SVC Support page at:

http://www.ibm.com/storage/support/2145

Click the **Supported Software Levels -> Supported Master Console Software Levels** links.

14.3 IBM Tivoli SAN Manager

The master console ships with IBM Tivoli SAN Manager Bonus Pack version agent and manager installed. It is installed to help with SAN problem determination. Tivoli SAN Manager monitors and manages switches and hubs, storage, and servers in a SAN. If a SAN problem occurs, clients and service representatives can use the SAN manager to understand the SAN configuration.

The IBM Tivoli SAN Manager Bonus Pack has a 64-switch port limit for the SANs it manages. If the SANs you manage are less than 64 ports, then the Bonus Pack works. To manage your SAN with more than 64 ports, you can purchase a license to extend the number of ports. Consider running Tivoli SAN Manager in a more powerful server. For Tivoli SAN Manager supported levels, device matrixes and updates, see the following Web site:

http://www-3.ibm.com/software/sysmgmt/products/support/IBMTivoliStorageAreaNetworkManager.html

For other information about Tivoli SAN Manager, see *IBM Tivoli Storage Area Network Manager: A Practical Introduction*, SG24-6848.

Two methods of communication are used for Tivoli SAN Manager: In-band communication over the Fibre Channel fabric and out-band communication over the Ethernet. Tivoli SAN Manager discovers SAN information by performing the following operations:

- Communicates with Tivoli SAN Manager agents. The agents run on monitored hosts (in-band discovery).
- Sends Management Information Base (MIB) queries directly to switches and other devices (out-of-band discovery).

Figure 14-3 shows the discovery paths.

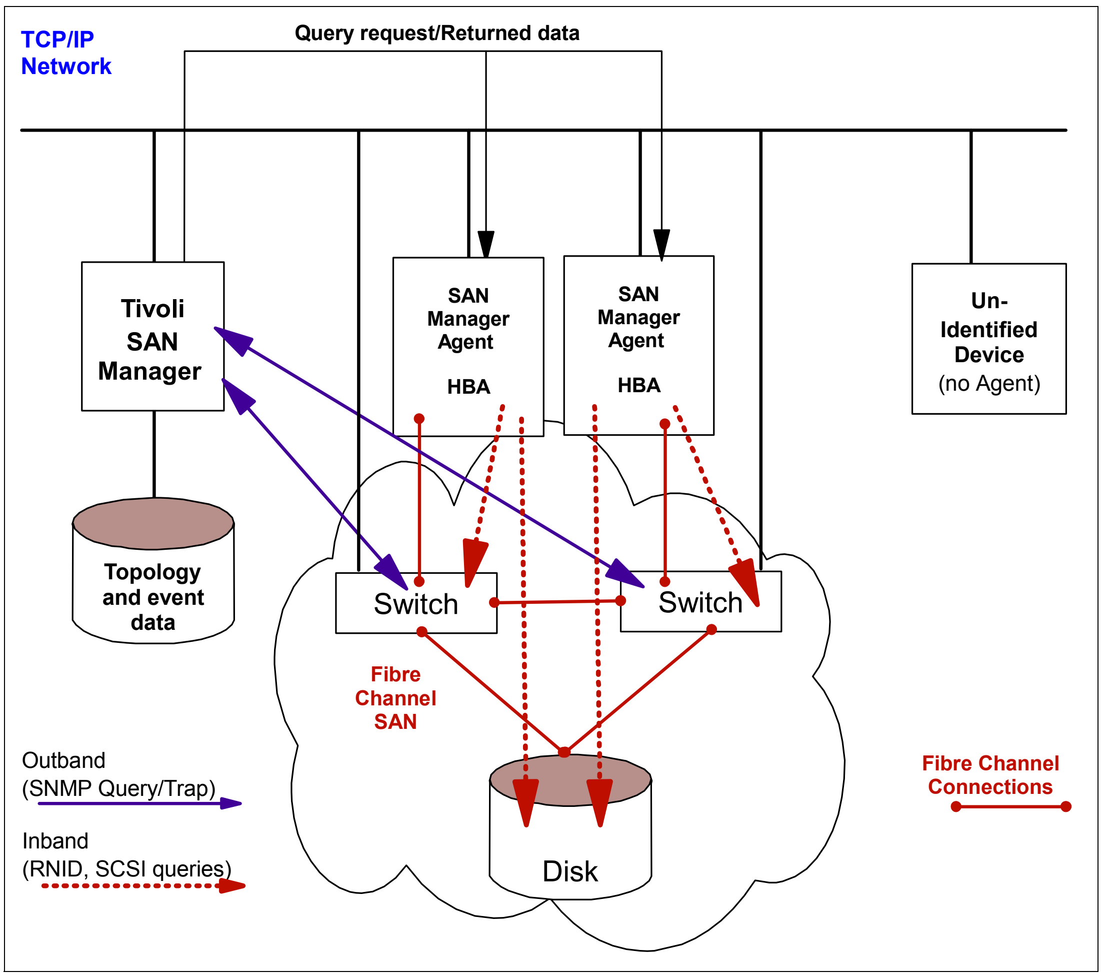


Figure 14-3 Discovery paths

In-band events and Simple Network Management Protocol (SNMP) traps provide the same information, such as a change that has occurred in the SAN by indicating that a discovery operation should be performed. The in-band events and SNMP traps let Tivoli SAN Manager know when something has happened in the SAN. Then a discovery is done to identify the changes. The Tivoli SAN Manager agents that are running on the monitored hosts detect in-band events.

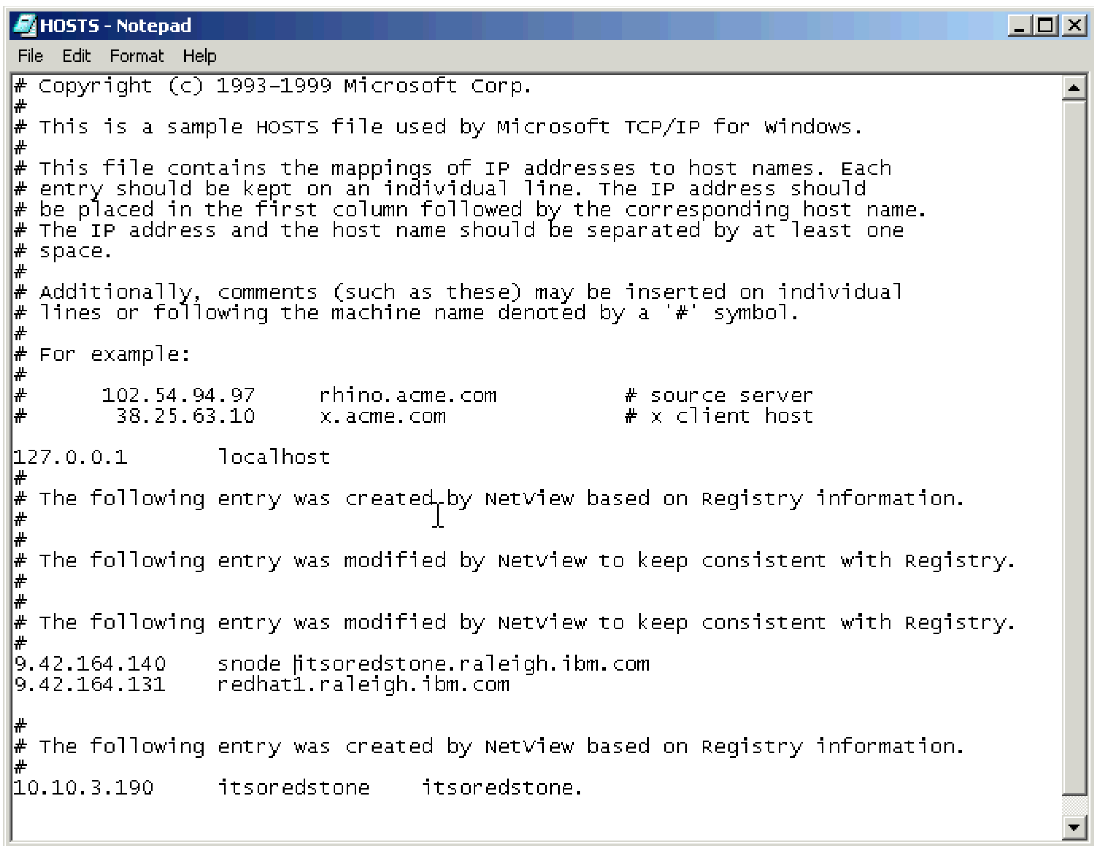
The agent sends commands through its host bus adapters (HBA) and the Fibre Channel network to gather information about the switches. The switch returns information through the

Fibre Channel network and HBA to the agent. The agent queries the devices using RNID and SCSI protocols. The agent then returns the information over the IP network to the Tivoli SAN Manager manager.

If you change the name of the master console, the following Tivoli SAN Manager configuration files may need to be changed (see the *IBM TotalStorage Virtualization Family SAN Volume Controller: Configuration Guide*, SC26-7543):

- c:\tivoli\itsanm\manager\bin\w32-ix86\setenv.bat
- c:\tivoli\itsanm\manager\bin\w32-ix86\setDeployEnv.bat
- c:\tivoli\itsanm\manager\conf\tsnmdbparms.properties
- c:\tivoli\itsanm\manager\conf\user.properties
- c:\tivoli\itsanm\manager\apps\was\java\jre\lib\orb.properties
- c:\tivoli\itsanm\manager\apps\was\config\cells\DefaultNode\nodes\DefaultNode\serverindex.xml
- c:\WINNT\system32\drivers\etc\HOSTS

These are shown in Figure 14-4.


HOSTS - Notepad
File Edit Format Help

```
# Copyright (c) 1993-1999 Microsoft Corp.
#
# This is a sample HOSTS file used by Microsoft TCP/IP for Windows.
#
# This file contains the mappings of IP addresses to host names. Each
# entry should be kept on an individual line. The IP address should
# be placed in the first column followed by the corresponding host name.
# The IP address and the host name should be separated by at least one
# space.
#
# Additionally, comments (such as these) may be inserted on individual
# lines or following the machine name denoted by a '#' symbol.
#
# For example:
#
#      102.54.94.97     rhino.acme.com          # source server
#       38.25.63.10     x.acme.com              # x client host

127.0.0.1       localhost
#
# The following entry was created by NetView based on Registry information.
#
#
# The following entry was modified by NetView to keep consistent with Registry.
#
#
# The following entry was modified by NetView to keep consistent with Registry.
#
9.42.164.140    snode itsoredstone.raleigh.ibm.com
9.42.164.131    redhat1.raleigh.ibm.com

#
# The following entry was created by NetView based on Registry information.
#
10.10.3.190     itsoredstone    itsoredstone.
```

Figure 14-4 Sample HOSTS file

If you are not using a dynamic name server, you must define your host for name resolution as in the previous example.

Follow these steps:

1. To start Tivoli SAN Manager, click the **Tivoli NetView** icon on the master console desktop area. This opens the default (Root) panel, as shown in Figure 14-5. Now we configure the agents.

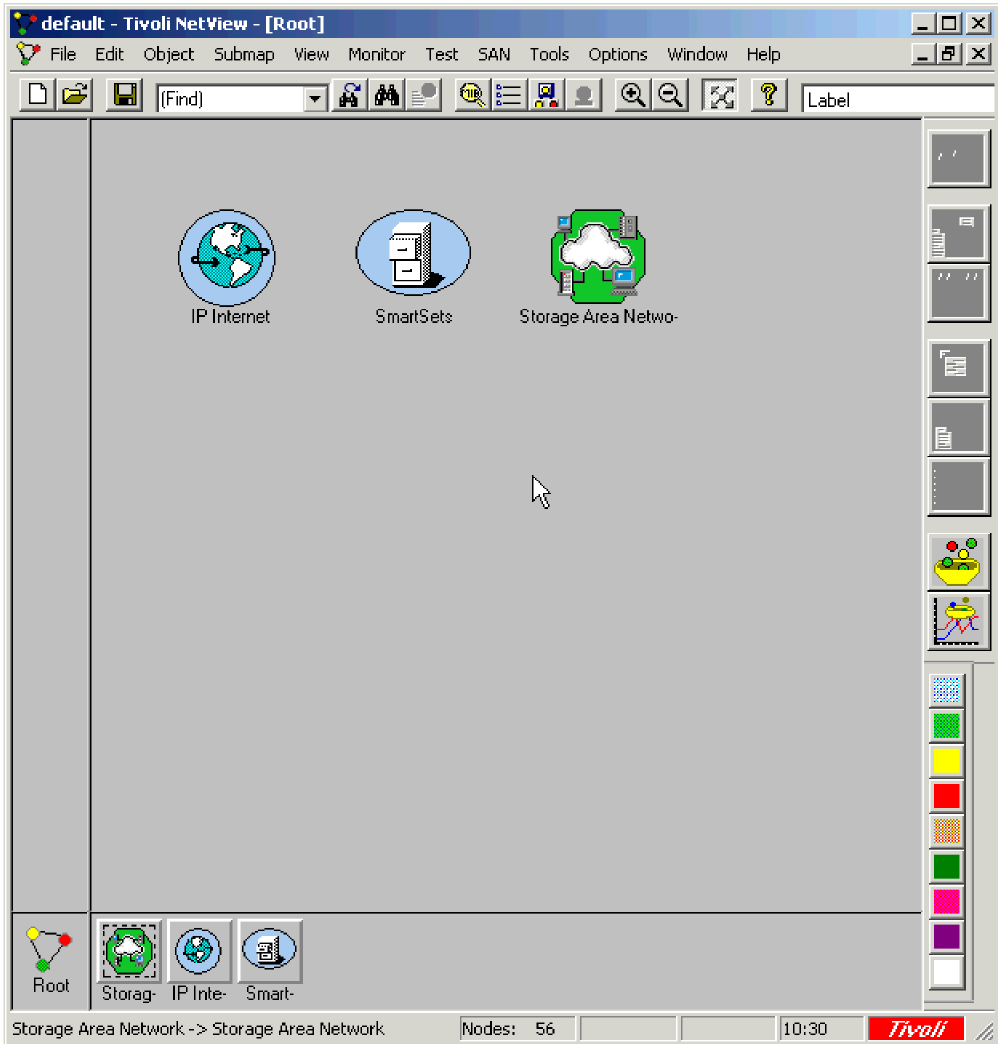


Figure 14-5 Tivoli SAN Manager start panel

2. Select **SAN -> Configure Agents** to start adding your switch IP addresses and the servers on which you installed the Tivoli SAN Manager agent (see Figure 14-6).

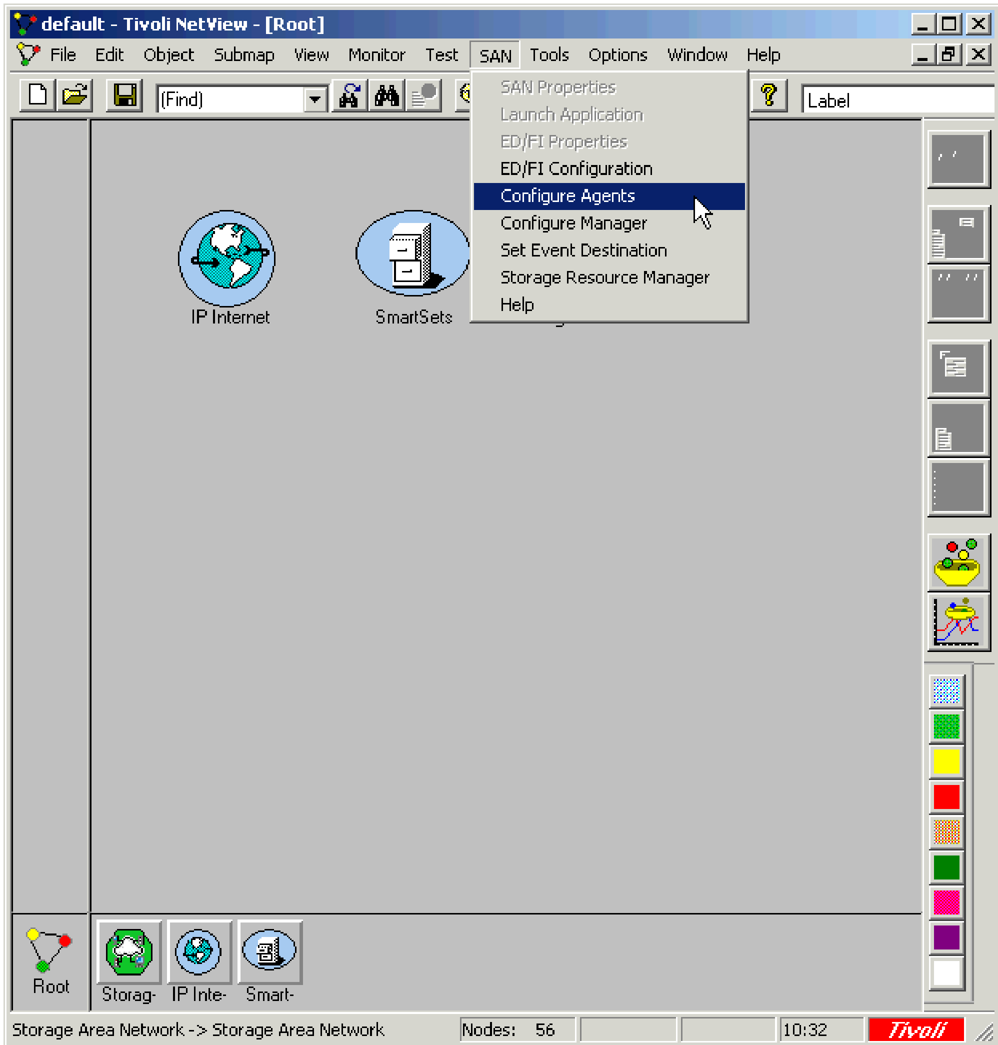


Figure 14-6 Configure agents

3. The top half of the Agent Configuration window in Figure 14-7 accommodates the agents installed on your host systems. In this example, we installed the agent on the master console by default and, in our example, on the windows host SRVE. When you install the Tivoli SAN Manager agent on your host, it asks for the Tivoli SAN Manager manager's host name and address. We have four Fibre Channel switches that we want Tivoli SAN Manager to report on, so we added them as well.

   On the Agent Configuration window, select the **Advanced** button. Enter the IDs and passwords for our switches. This allows you to launch the switch management interface.

   Click **OK** to close the Agent Configuration panel.

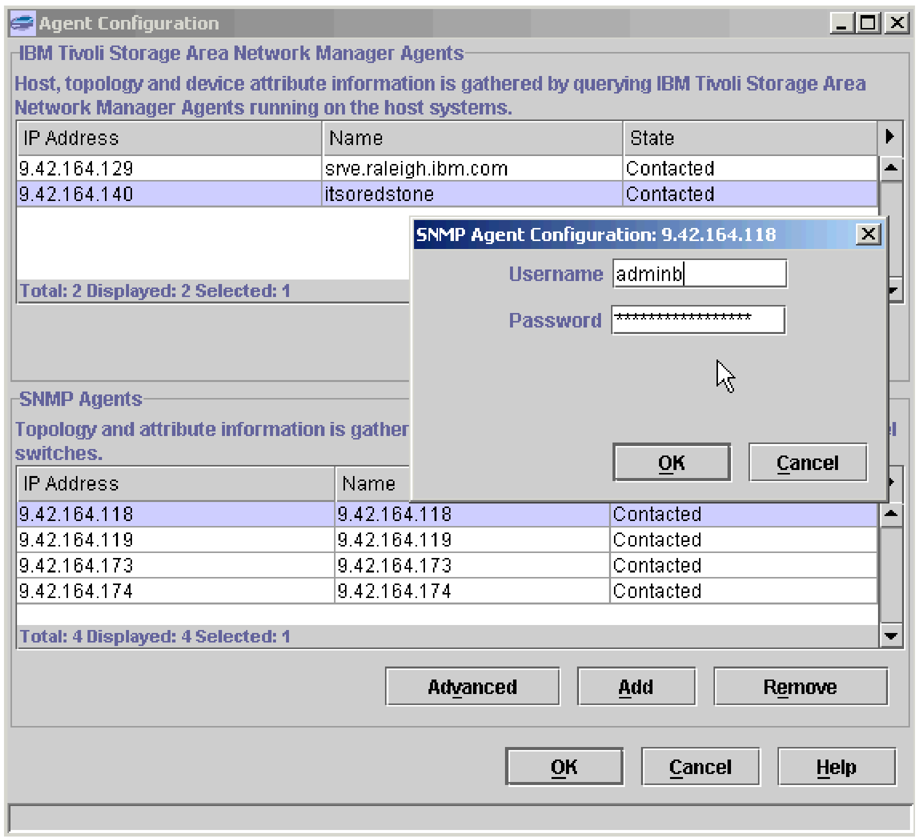


Figure 14-7 SNMP agent configuration: Switch user ID and password

4. From the Tivoli NetView [Root] window (Figure 14-8), select **SAN -> Configure Manager**.

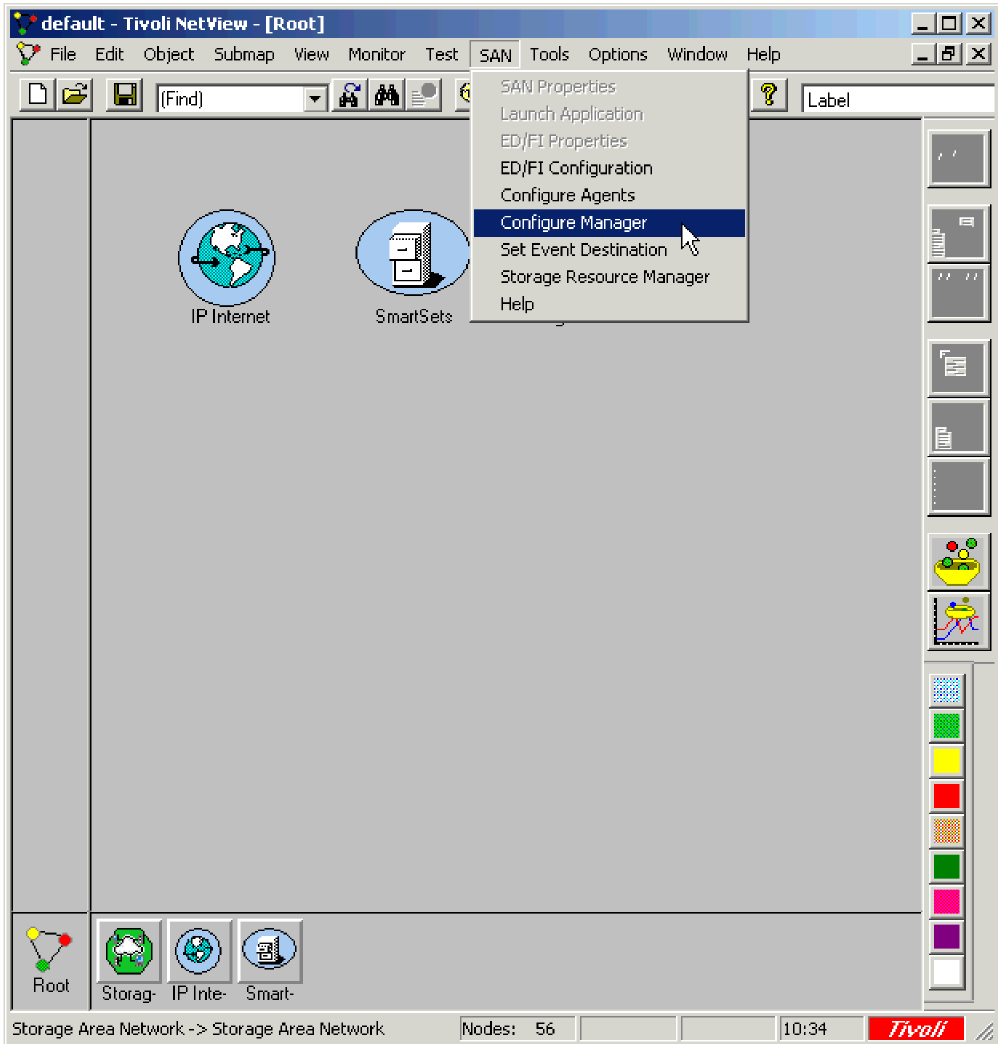


Figure 14-8 Selecting Configure Manager

5. On the SAN Configuration window (Figure 14-9), configure the polling interval for Tivoli SAN Manager automatically gather information about the SAN topology and other information.

   To clear the history file of the previously discovered topology, click **Clear History** and then rediscover the topology by clicking the **Poll Now** button.

   Click **OK** to close this window.

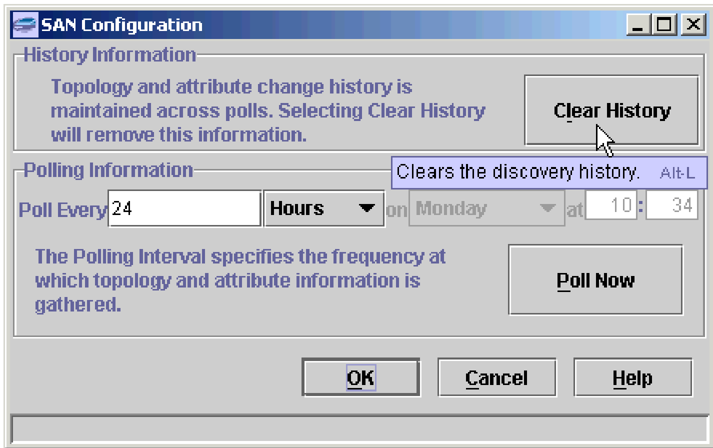


Figure 14-9 Polling interval

6. On the *Root* panel (Figure 14-10), right-click the **Storage Area Network** icon, and select **Explore**.

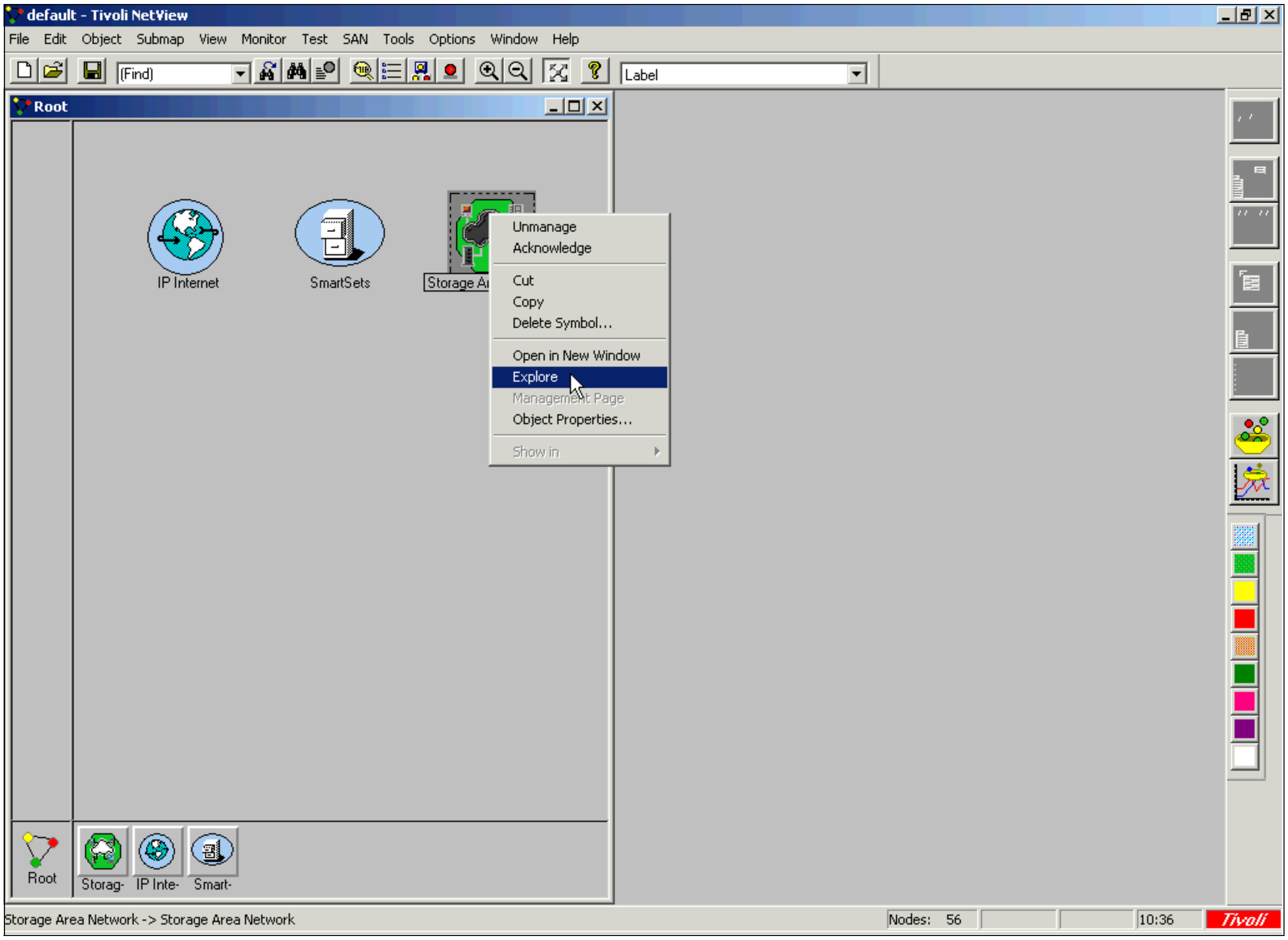


Figure 14-10 Exploring storage area network

Figure 14-11 shows the results.

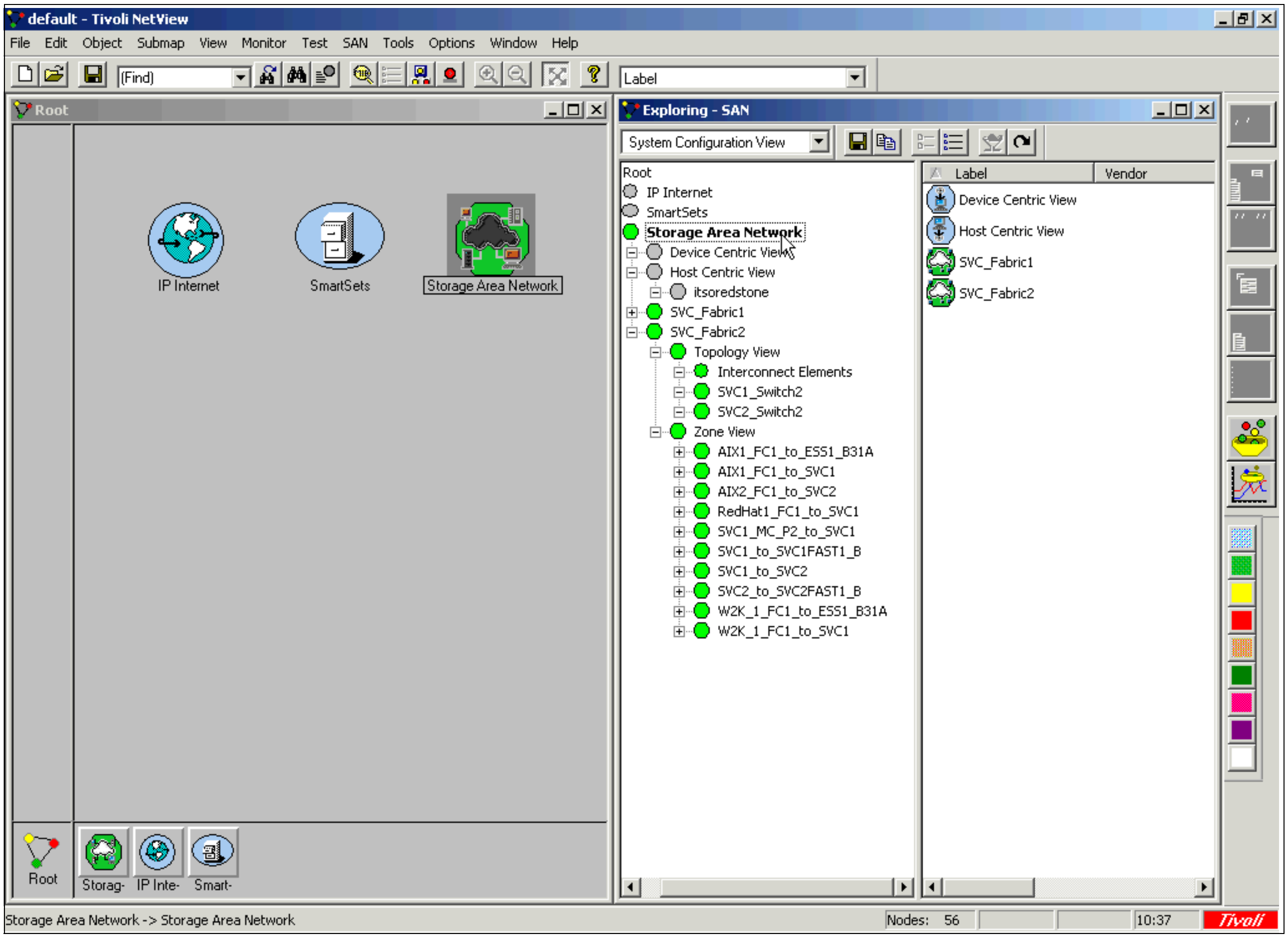


Figure 14-11 Exploring SAN

7. The Explorer panel (Figure 14-12) opens to display all the devices, topology, and zones in a tree format. The different views of the fabric components can then be shown, such as the Fabric View.

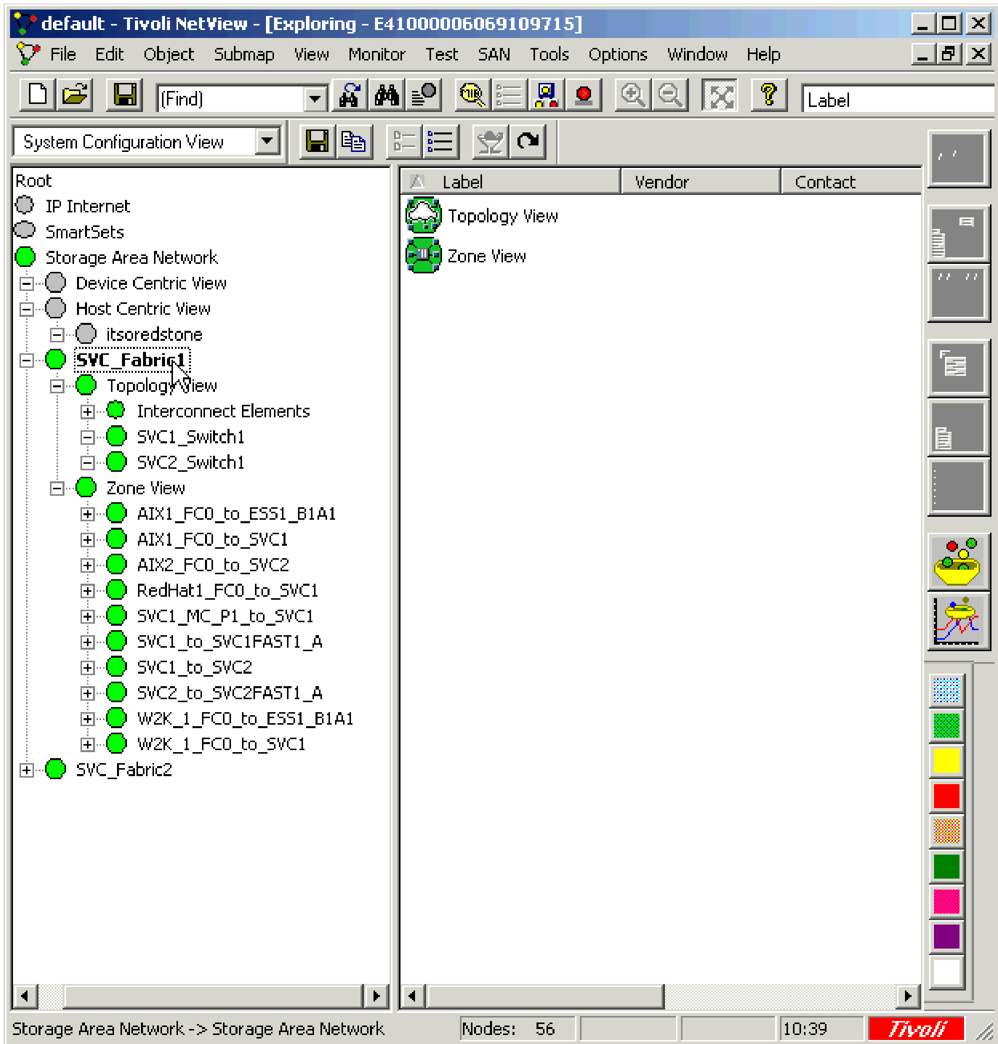


Figure 14-12 Fabric view

Figure 14-13 shows the Switch view.

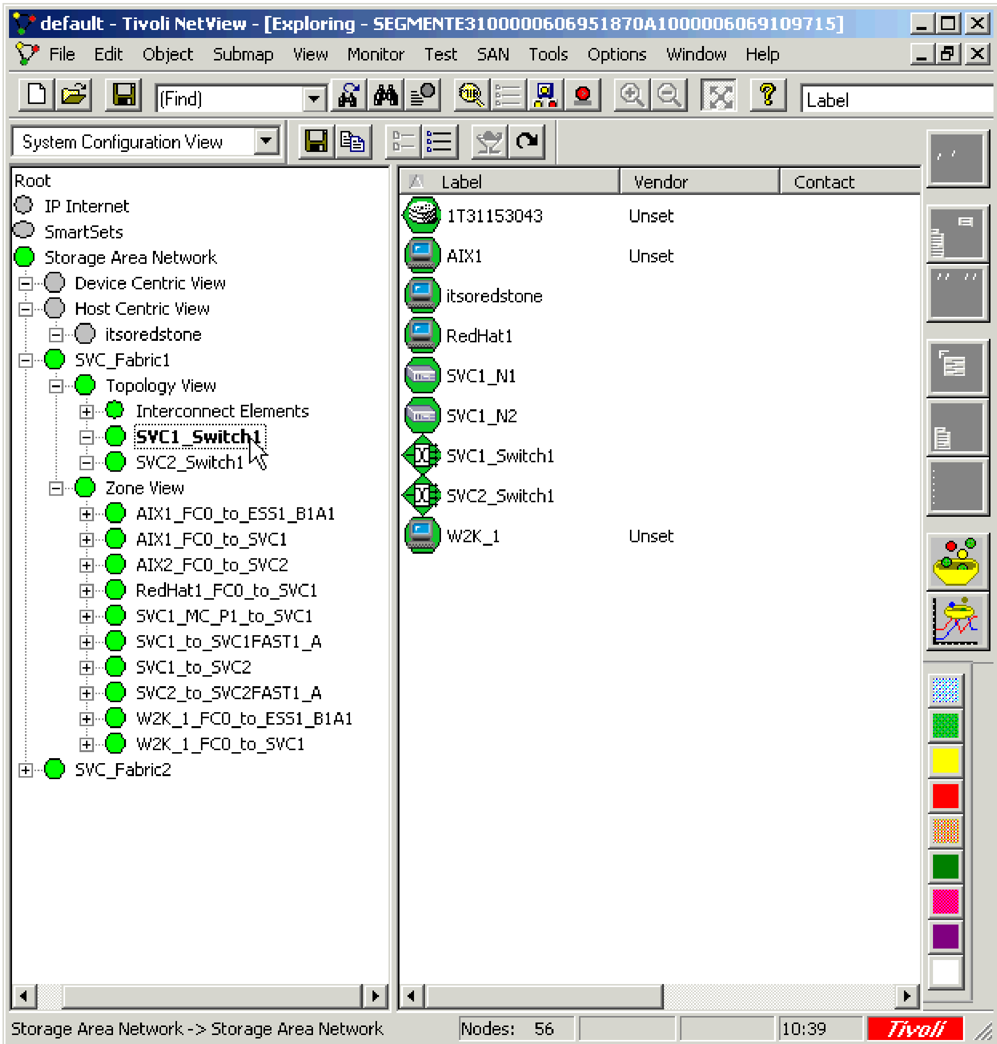


Figure 14-13 Switch view

To focus on one of the hosts, click it once and selecting **SAN -> SAN Properties** from the menu, as shown in Figure 14-14.

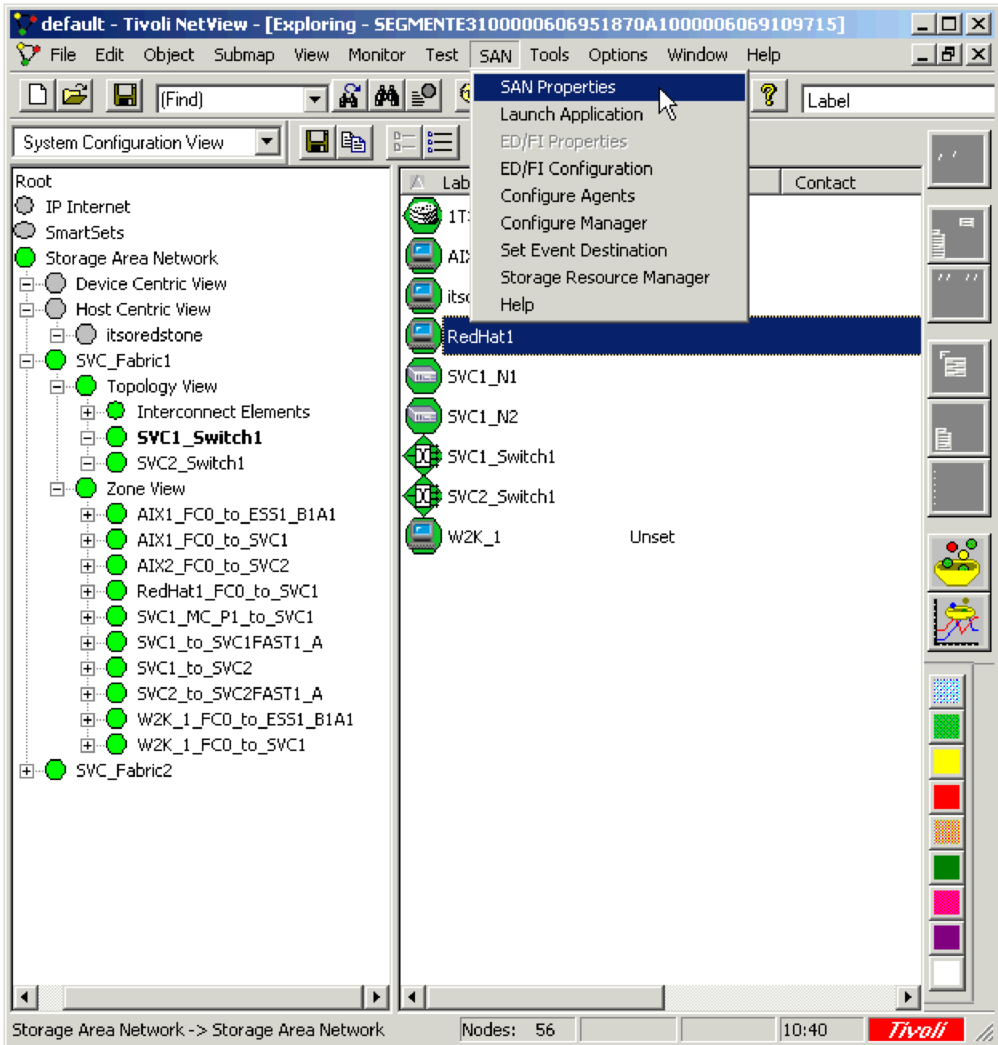


Figure 14-14 SAN properties: Host

Figure 14-15 shows the details for the host.

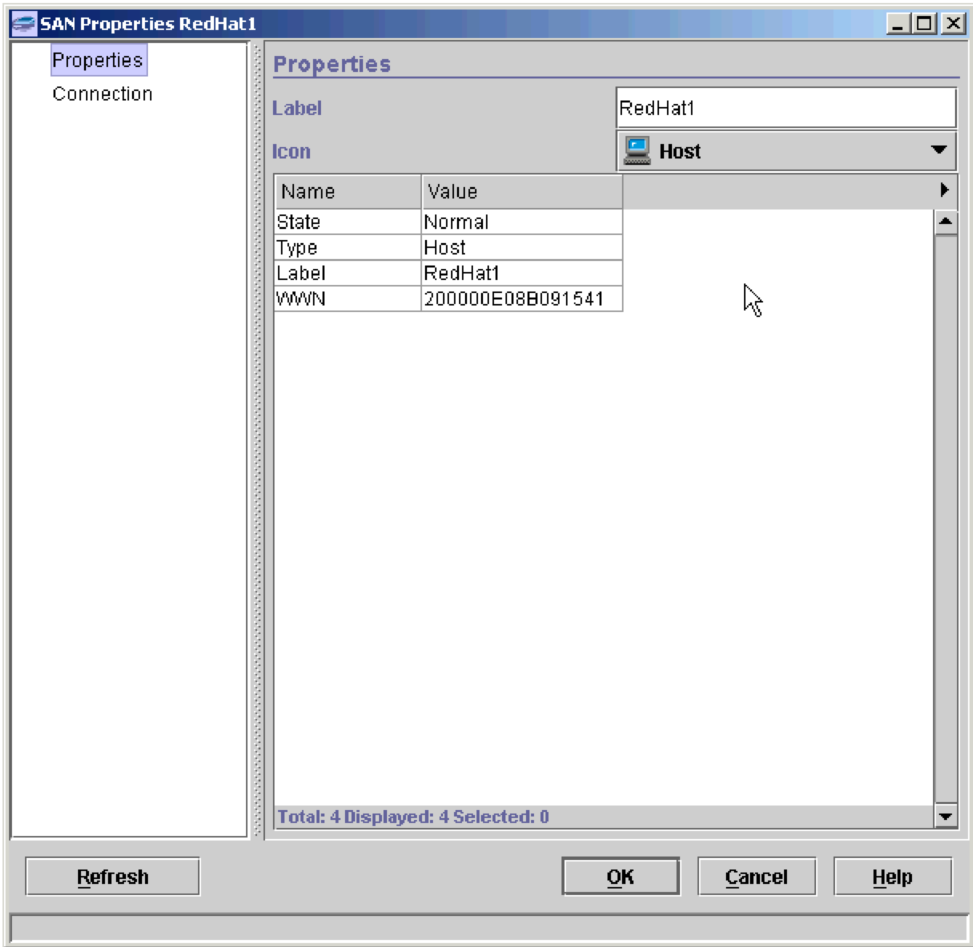


Figure 14-15 SAN properties for a host

If you select one of the switches in the Explore panel and look at its SAN properties, you see its firmware levels, as shown in Figure 14-16.

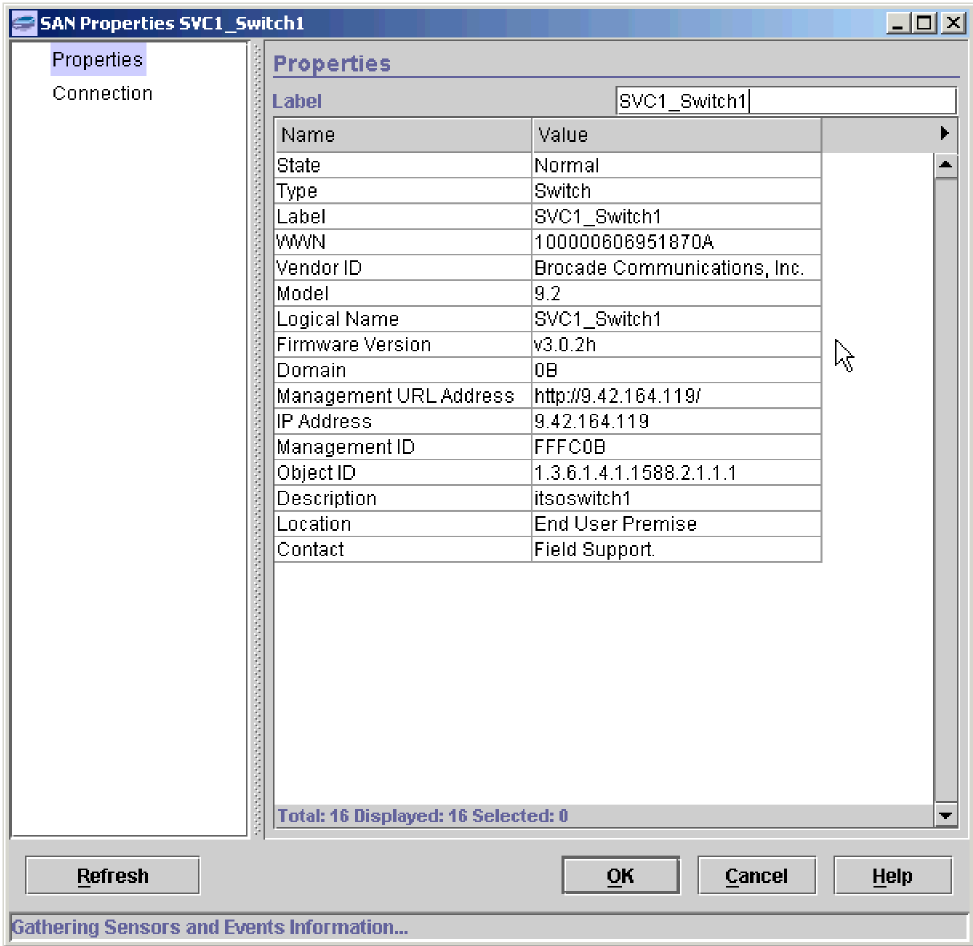


Figure 14-16 Switch1 SAN properties

Select **Connections** in the left panel of the Properties display, and you see the switch ports and devices, as shown in Figure 14-17.

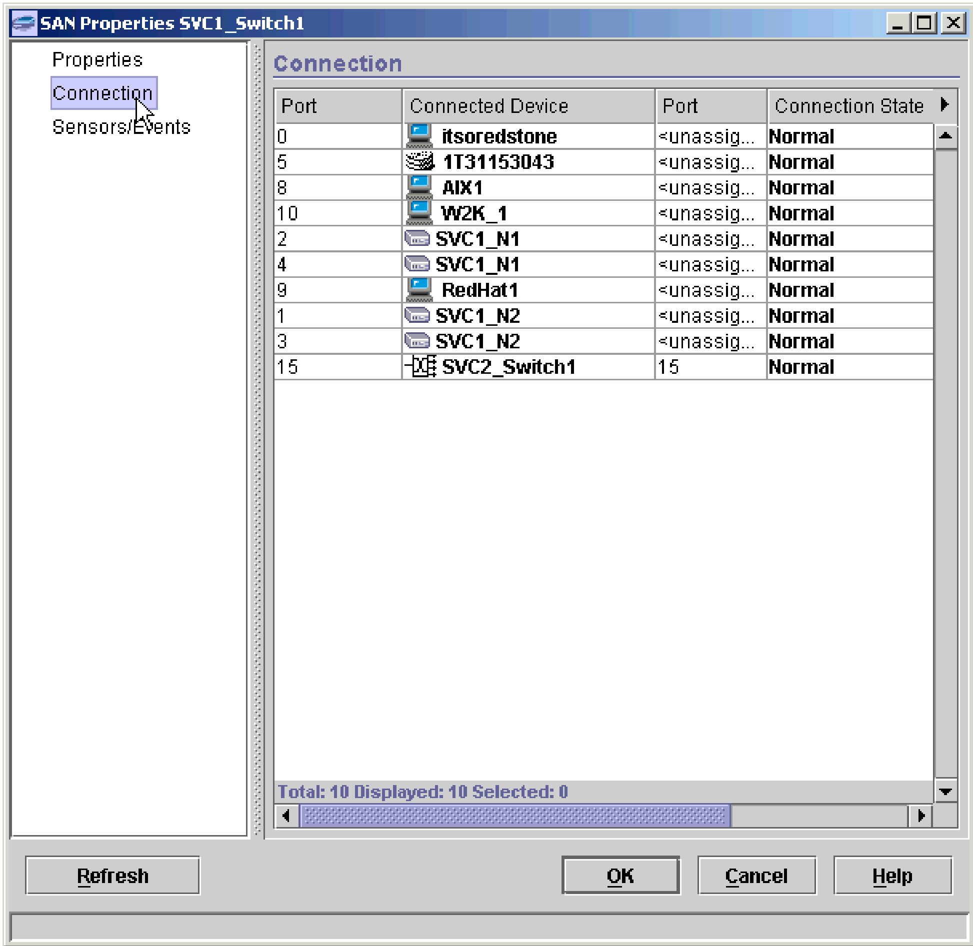


Figure 14-17 Switch1 connection properties

You can then explore the fabric zones (see Figure 14-18). In this example, we selected the zone view for SVC_fabric1. These are the effective zones in SVC_fabric1.

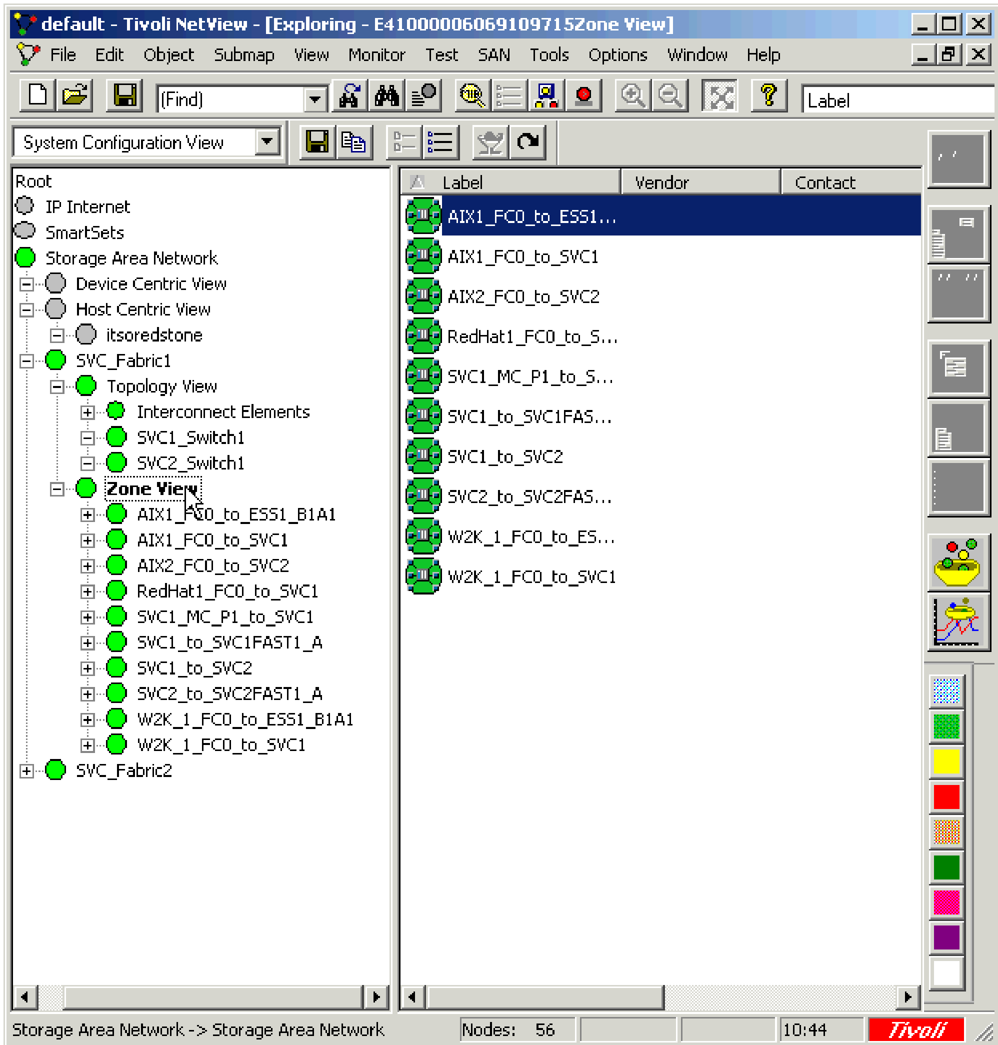


Figure 14-18 Exploring the zone view

Now we display the connection to SCV2_switch2 as shown in Figure 14-19. Notice that some of the items are displayed as a question mark “?”. These items require you to complete their properties so that they display a meaningful description. After you enter all the appropriate information, all of the links and nodes are kept in a history file. The history file is used to show links or nodes with problems or new Fibre Channel links and devices.

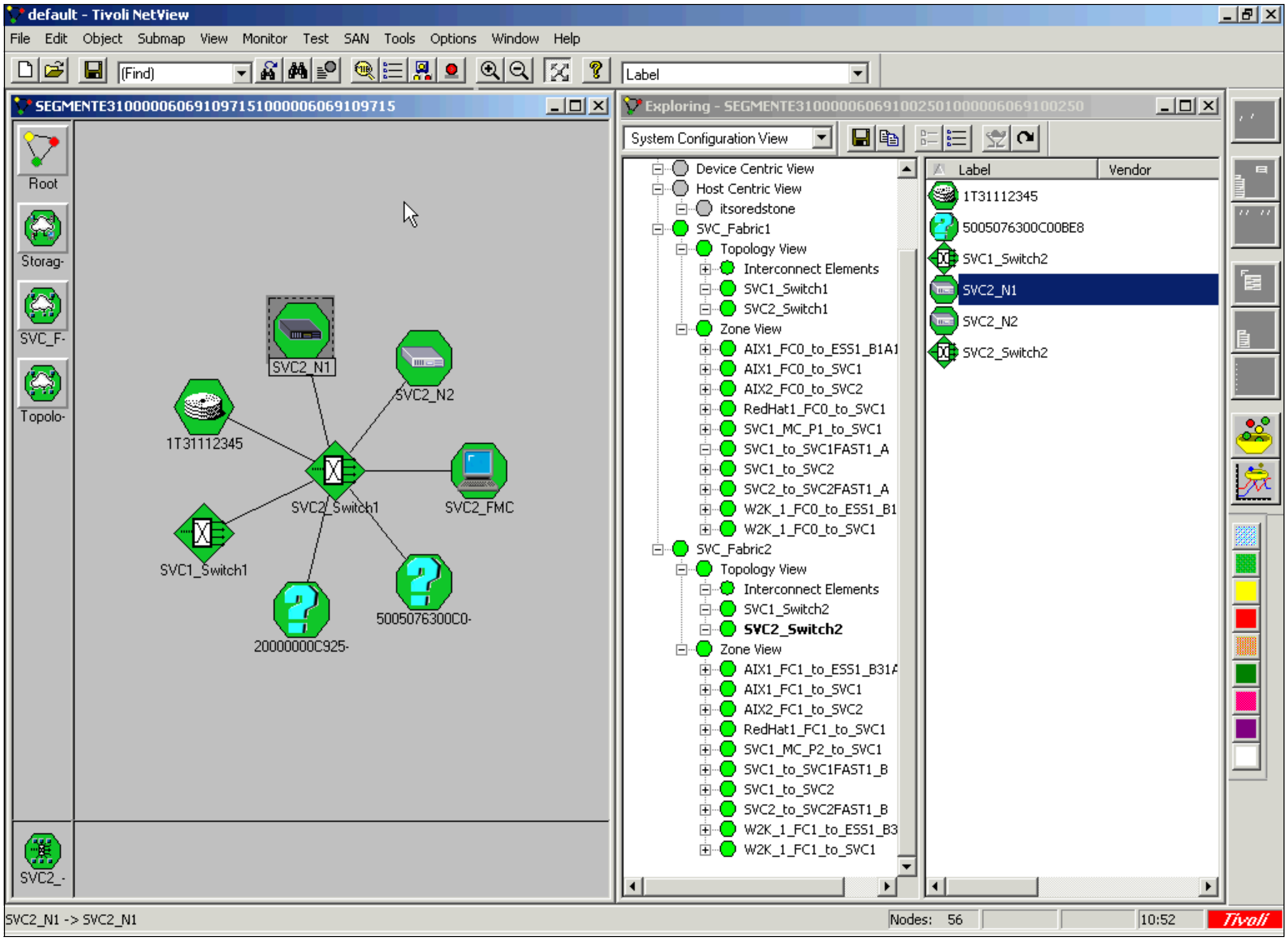


Figure 14-19 Switch2 unknown items

Right-click and select **Object Properties** for the objects with a question mark, as shown in Figure 14-21.

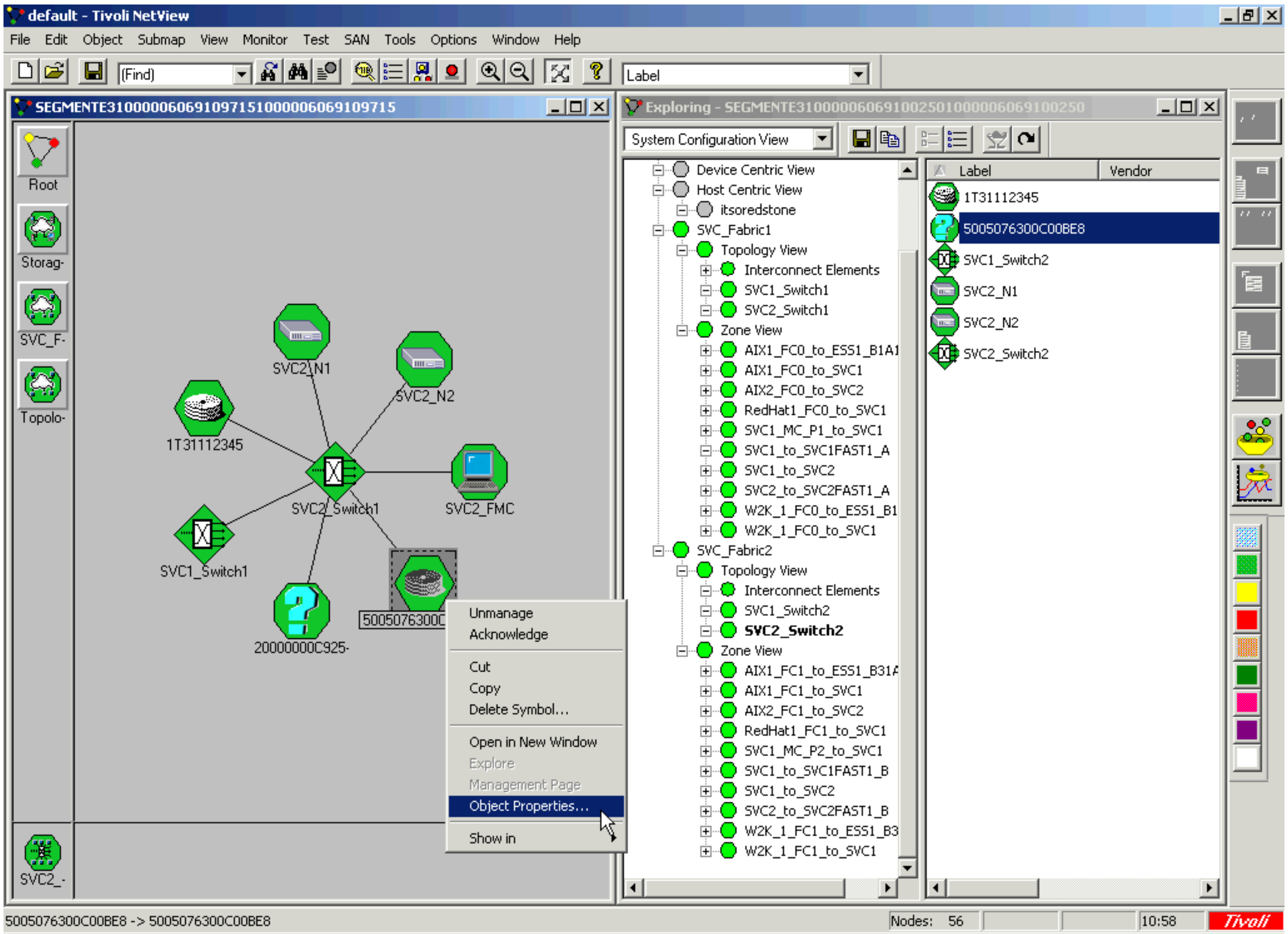


Figure 14-20 Selecting Object Properties

Now you can customize the object icon and properties as shown in Figure 14-21. In this example, we know it is a storage device from its worldwide node name (WWNN). Click the **Type** tab and select the **device** and **harddisk** options.

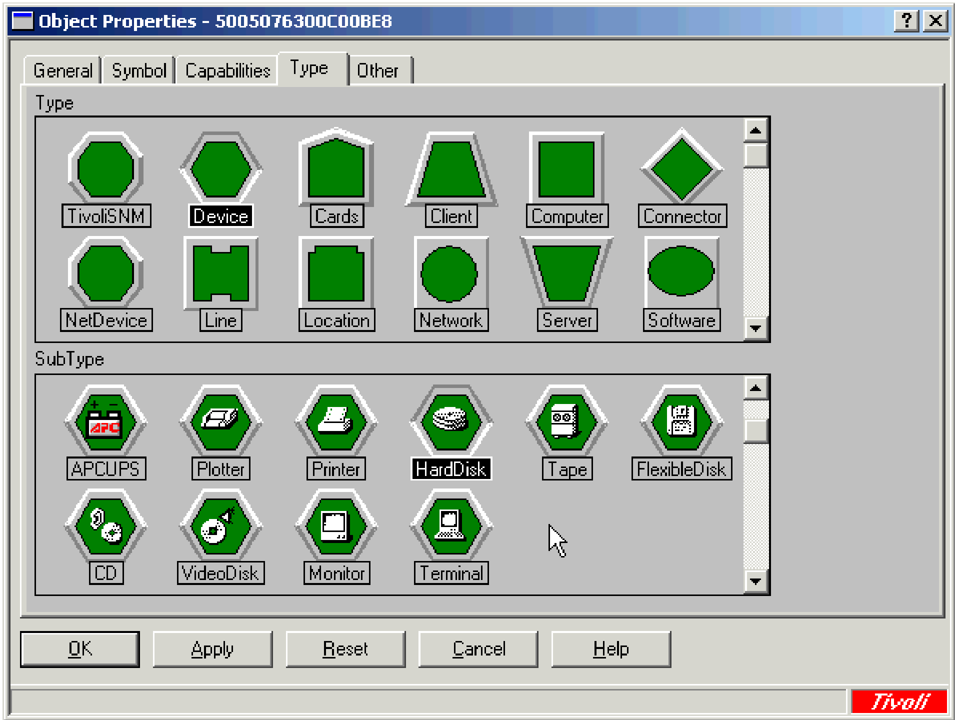


Figure 14-21 Properties type

Click the **General** tab to enter other information that identifies the object, as shown in Figure 14-22. In this example, we typed `ESS1_B1A` for the name of our disk.

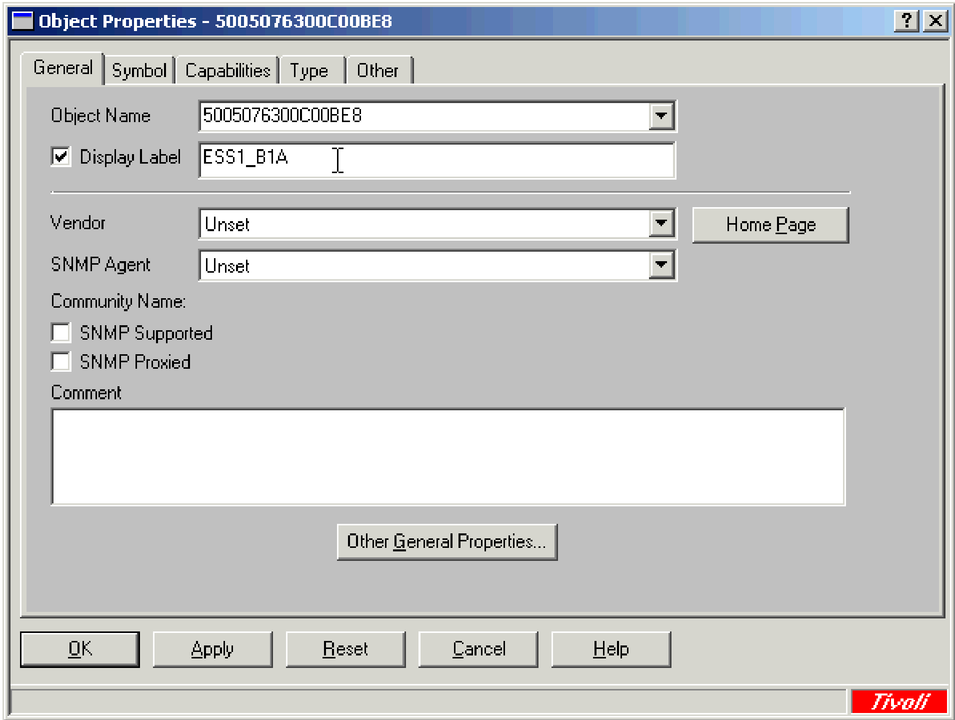


Figure 14-22 General properties

Now the name of our disk is displayed on the explorer window as shown in Figure 14-23.

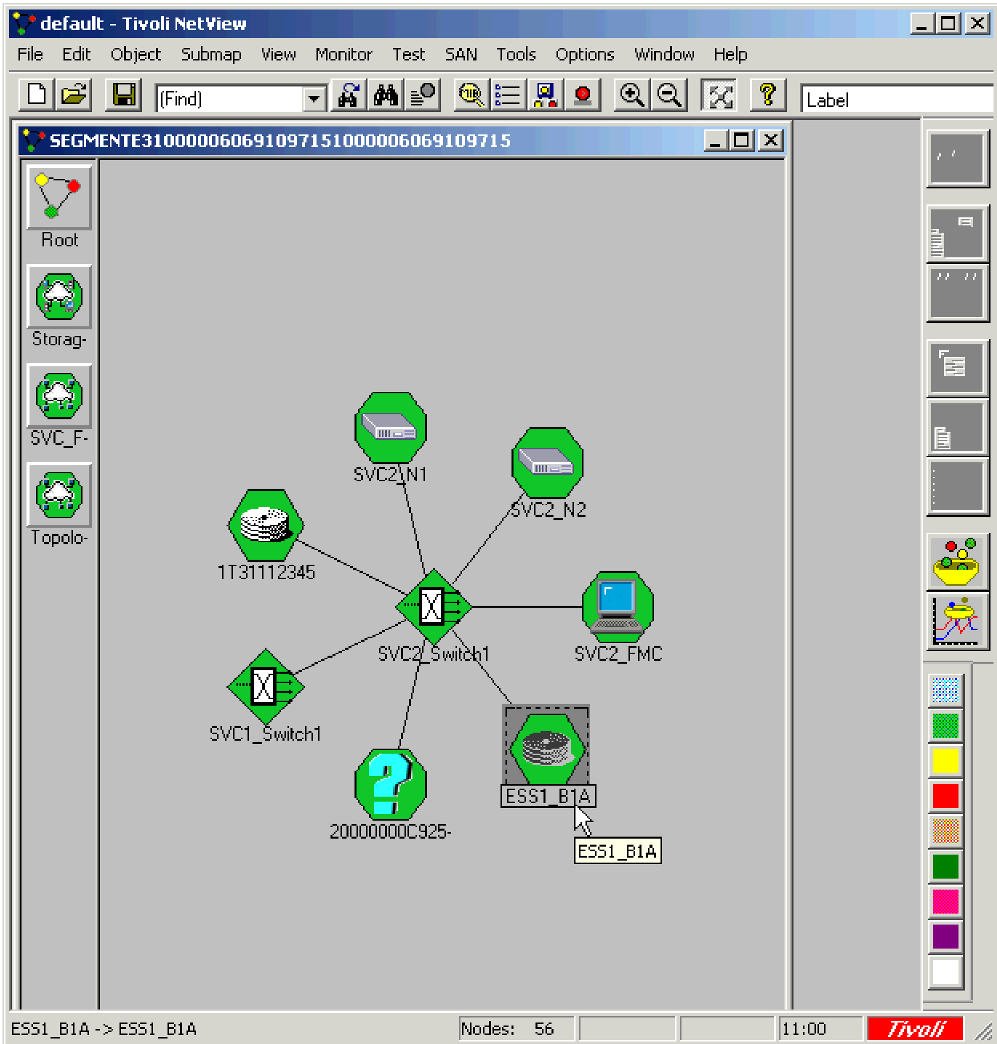


Figure 14-23 Object renamed

To simulate a problem, we removed one of the fibre cables to the Redhat1 server. This creates an event and Tivoli SAN Manager acquires the change as shown in Figure 14-24. The icon has a red background in the right panel meaning it is *offline*. The switch to which it is connected shows a yellow caution in the left panel, meaning an item is *suspect*. The three indication colors are:

- Green: Object is online.
- Yellow: Something is suspect.
- Red: Object is offline.

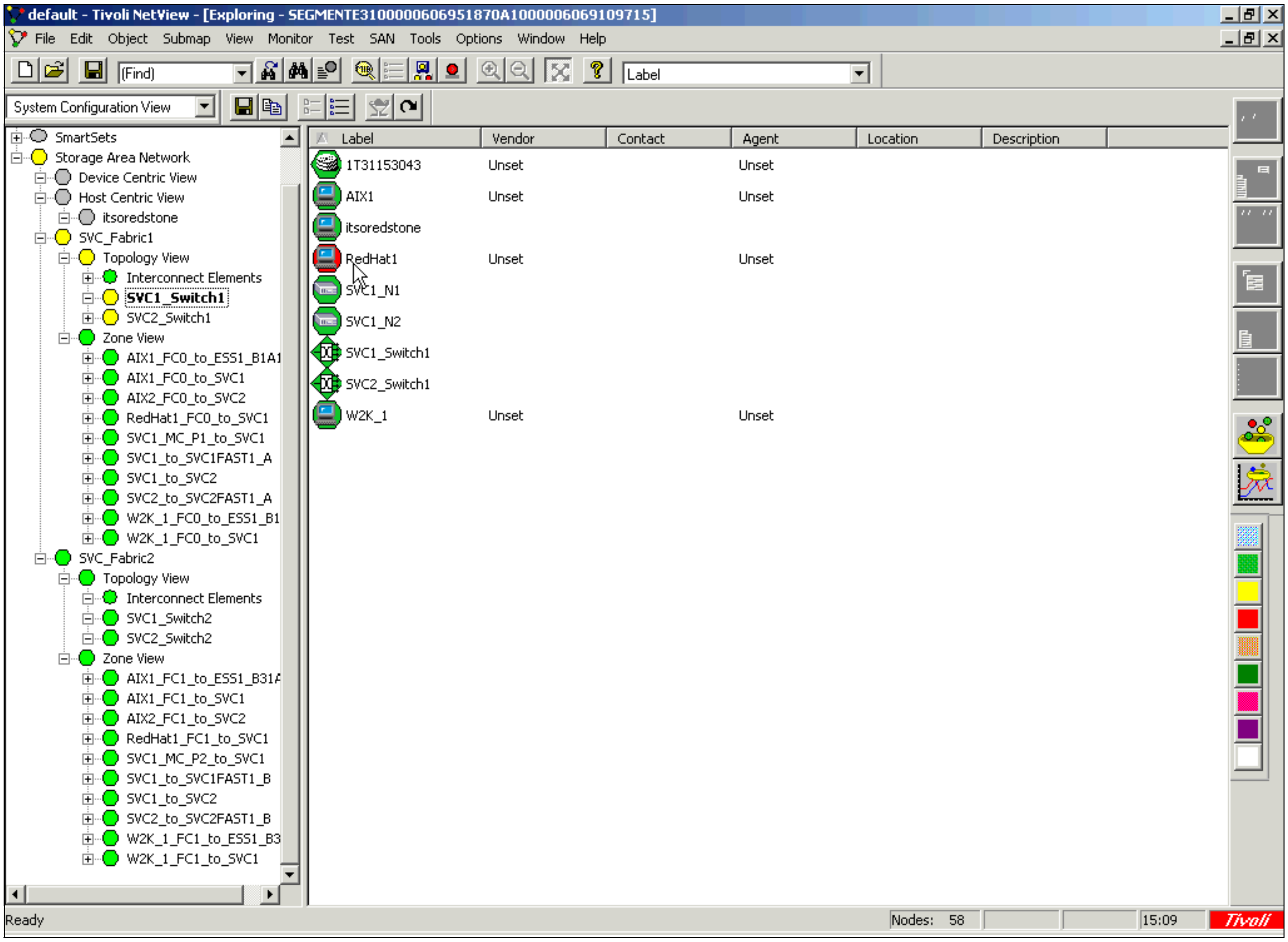


Figure 14-24 Fibre Channel error on host

The IBM Tivoli Storage Resource Manager general information panel (Figure 14-25) displays both IBM Tivoli Storage Resource Manager and Tivoli SAN Manager running processes and levels of the Agent Service Manager. To obtain this information, point your browser to:

`http://`*`IP address of master console`*`:9570/ITSRM/ServiceManager`

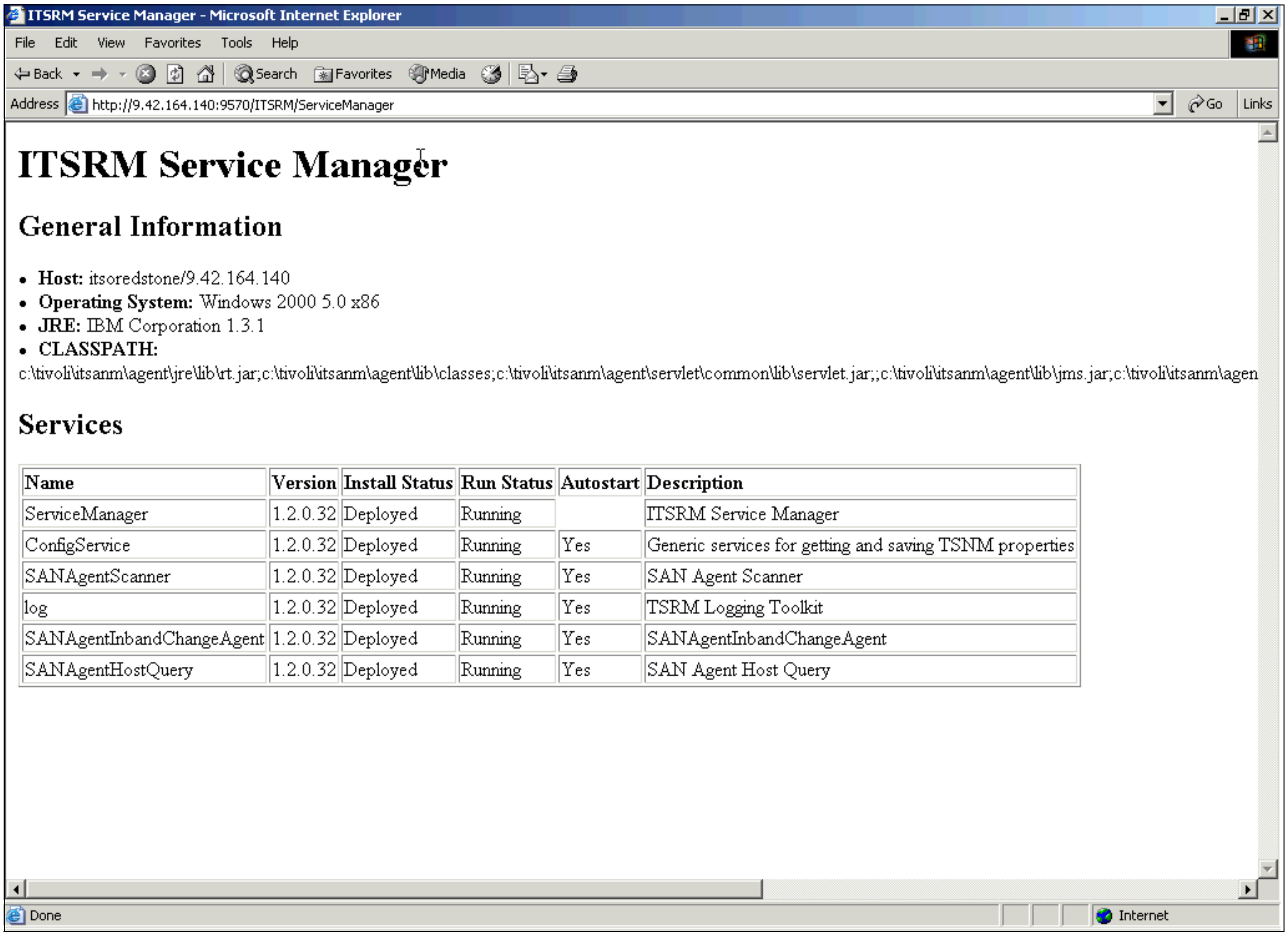
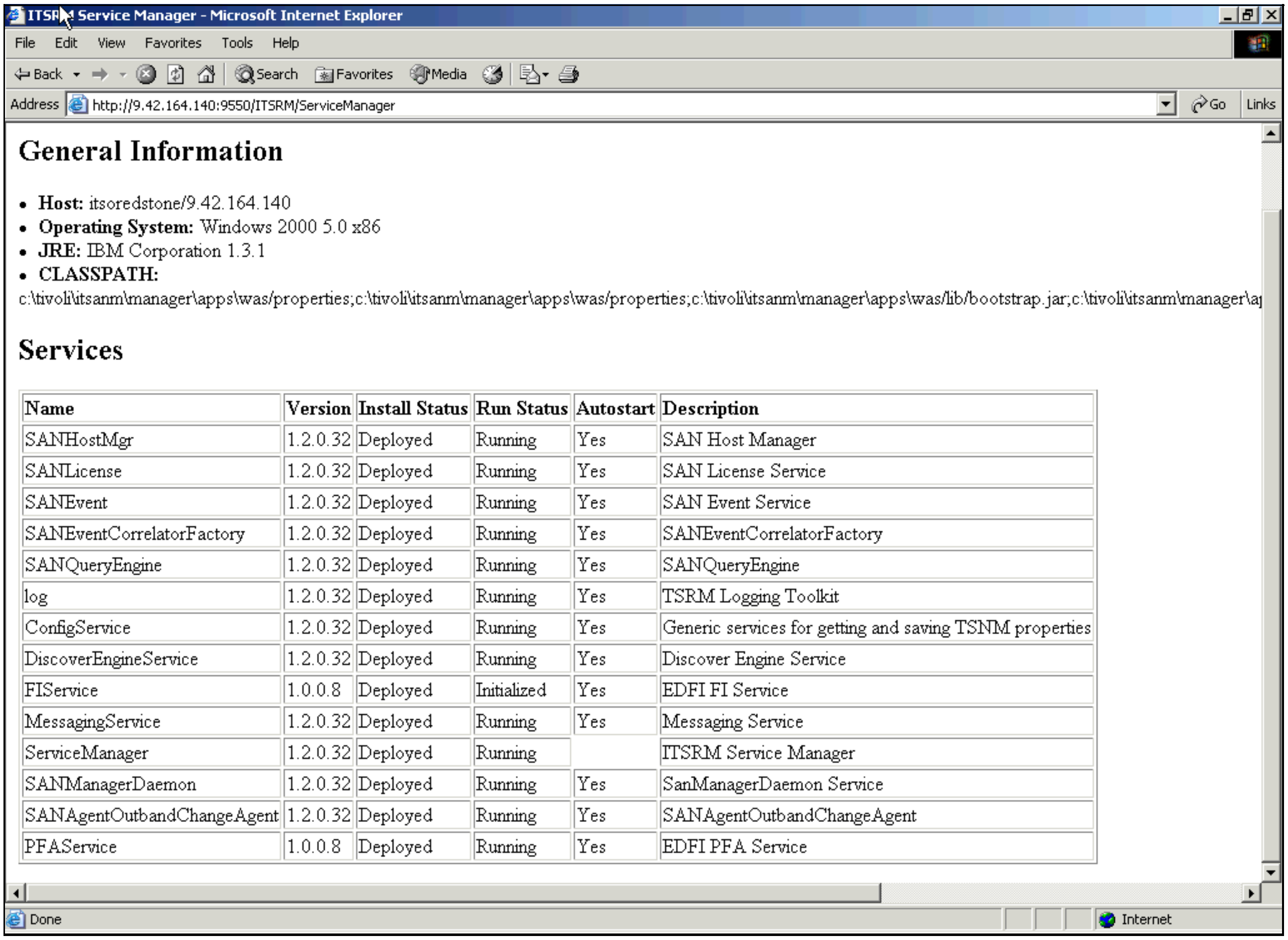


Figure 14-25 IBM Tivoli Storage Resource Manager agent service manager

To display the levels of the Tivoli SAN Manager server software installed, as shown in Figure 14-26, point your browser to the following address:

```
http://IP address of master console:9550/ITSRM/ServiceManager
```



General Information

- **Host:** itsoredstone/9.42.164.140
- **Operating System:** Windows 2000 5.0 x86
- **JRE:** IBM Corporation 1.3.1
- **CLASSPATH:**

c:\tivoli\itsanm\manager\apps\was/properties;c:\tivoli\itsanm\manager\apps\was/properties;c:\tivoli\itsanm\manager\apps\was/lib/bootstrap.jar;c:\tivoli\itsanm\manager\a

Services

Name	Version	Install Status	Run Status	Autostart	Description
SANHostMgr	1.2.0.32	Deployed	Running	Yes	SAN Host Manager
SANLicense	1.2.0.32	Deployed	Running	Yes	SAN License Service
SANEvent	1.2.0.32	Deployed	Running	Yes	SAN Event Service
SANEventCorrelatorFactory	1.2.0.32	Deployed	Running	Yes	SANEventCorrelatorFactory
SANQueryEngine	1.2.0.32	Deployed	Running	Yes	SANQueryEngine
log	1.2.0.32	Deployed	Running	Yes	TSRM Logging Toolkit
ConfigService	1.2.0.32	Deployed	Running	Yes	Generic services for getting and saving TSNM properties
DiscoverEngineService	1.2.0.32	Deployed	Running	Yes	Discover Engine Service
FIService	1.0.0.8	Deployed	Initialized	Yes	EDFI FI Service
MessagingService	1.2.0.32	Deployed	Running	Yes	Messaging Service
ServiceManager	1.2.0.32	Deployed	Running		ITSRM Service Manager
SANManagerDaemon	1.2.0.32	Deployed	Running	Yes	SanManagerDaemon Service
SANAgentOutbandChangeAgent	1.2.0.32	Deployed	Running	Yes	SANAgentOutbandChangeAgent
PFAService	1.0.0.8	Deployed	Running	Yes	EDFI PFA Service

Figure 14-26 Tivoli SAN Manager server general information

14.4 Installation planning information for the master console

The following items need to be available to help in planning for your master console installation:

- Network IP address: Need a static IP address to use Tivoli SAN Manager
- Network host name
- Space or location in the rack where the machine will be installed
- Two fiber ports connect to the SAN redundant fabric

> **Attention:** The master console ports must be included in a zone to see all the SVC Fibre Channel adapter ports and no other ports.

For detailed installation guidance, see the *IBM TotalStorage Virtualization Family SAN Volume Controller: Configuration Guide*, SC26-7543.

14.5 Administration tasks

Each time you want to add a SVC cluster to the collection of clusters managed by the SVC master console, you must store the SSH public key for the master console and any other systems that require access to the SVC on the SAN Volume Controller cluster.

This is a long-term administrative task and not just a post-installation task. We outline how to perform this task using your Web browser. The same information is included in the "Post Installation Tasks" section in the *IBM TotalStorage Virtualization Family SAN Volume Controller: Configuration Guide*, SC26-7543.

> **Attention:** If you do not store the SSH public key on the SVC cluster, SVC access software cannot connect to the cluster.

Here is a summary of the steps:

1. Start the browser to access the SVC console.
2. Log onto the SVC console using the default `superuser` name and password.
3. Access user assistance (optional).
4. Identify the SVC clusters to the SVC master console.
5. Store the SSH public key on the SVC master console.
6. Launch the secondary browser window to manage your specific cluster.

Detailed procedure follows. This subject is discussed in detail in the "SAN Volume Controller Console" chapter in the *IBM TotalStorage Virtualization Family SAN Volume Controller: Configuration Guide*, SC26-7543.

1. Start the browser to access the SVC Console. Access the SVC console from a Web browser. You must use a browser to log onto the system on which the SVC console is installed to complete uploading the public SSH key for each cluster you will manage. You can access the SVC console by pointing a Web browser to the uniform resource locator (URL):

   `http://localhost:9080/ica`

2. Log onto the SAN Volume Controller Console using the `superuser` name and password. The default superuser name is `superuser` and the default superuser password is `passw0rd` (with a zero replacing the letter "o").

3. Access user assistance (optional). You can access help for the specific task on which you are working by clicking the small information icon just beneath the banner in the upper right section of the Web page. You can also launch a separate user assistance panel by clicking the small question mark icon just beneath the banner in the upper right section of the Web page. A secondary browser window opens that provides access to extensive user assistance.

4. Identify the SVC clusters to the SVC master console. The steps you may need to perform to add SVC clusters to the SVC console collection of managed clusters depend on the current status of the cluster in which you are interested:

   – **A SVC cluster which has not yet been initialized**: If you have not yet created a SVC cluster using the front panel of the SVC cluster, you need to perform that phase of the cluster creation first. See 4.3.1, "Creating the cluster (first time) using the service panel" on page 74. You are given a special password that is displayed in the SVC front panel for 60 seconds to be used in later steps to initialize the SVC master console.

     After the first phase to create the SVC cluster using the front panel of a SVC node, you need to complete the creation of the cluster by using the SVC console Web interface, as described in 4.4, “Completing the initial cluster setup using the Web interface” on page 76.

– **Previously initialized SVC cluster**: If the SVC cluster has completed the initialization (creation) process but is not yet registered with the SVC console, you log on with the `superuser` ID and password, select **Add Cluster** from the list, and then you simply add the cluster in the SVC Welcome page. You need to inform the IP address of the cluster you are adding but do not select the Create (Initialize) Cluster check box which is above the OK button. When you click the OK button, you are taken to the page to provide the SSH public key for upload to the cluster. Step 5 continues with the SSH key input description.

  As part of this process the program prompts you to enter the network password. You should type the `admin` user name and the password which is configured for the cluster.

5. Store the SSH public key on the SAN Volume Controller master console. Each key is associated with an ID string that you define that can consist of up to 30 characters. Up to 100 keys can be stored on a cluster. You can add keys to provide either administrator access or service access as described in 9.6.3, “Maintaining SSH keys” on page 280.
6. Launch the secondary browser window to manage your specific cluster. After you identify the SVC clusters to the SVC console, you can see a summary of all clusters previously unidentified. From this point, you can select the specific cluster in which you are interested and then launch the browser window specifically for that cluster.
   a. Under the My Work section of your browser window, click **Clusters**. A new view is displayed in the work area (main frame).
   b. In the Select column, select the check box to the left of the cluster in which you are interested. Select **Launch the SAN Volume Controller Application** from the list of the work area and click **Go**. A secondary browser window opens to the SVC application. Now you can work with the specific SVC cluster which you selected. Notice the ClusterName parameter in the browser location URL, which identifies the IP address of the cluster with which you are working as shown here:

```
http://9.43.147.38:9080/ica/Console?Console.loginToken=79334064:f46d035f31:-7ff1
&Console.ClusterName=9.43.225.208
```

While using the SVC console, there is an issue with Windows registration of an SSH key. When a cluster is first created and then added to the ICAT for management, SVC will send a public key to the master console. This key is saved in the Windows registry. If you deleted and recreated the cluster, again SVC will send a public key. Since a key for this device is already saved, the Windows registry is not updated with the new key and the cluster cannot be managed from the ICAT. This is for security reasons as it thinks another device is attempting to access it. The work-around is to delete the key from the registry after deleting the cluster and before it is recreated. There is a function (Reset SSH Fingerprint) provided in the dropdown list to correct this situation. This is not an issue with the command line SSH as you are prompted to overwrite the keys.

To establish SSH connection to the cluster, the master console stores the public key sent by SVC in the following path \HKEY_USERS\.DEFAULT\Software\SimonTatham\PuTTY\SshHostKeys. The name of the registry key is rsa2@22:*cluster_IP_address*.

The reset function fixes the registry to use the correct public SSH key sent from SVC.

14.6 Secure Shell

Secure Shell (SSH) is used to secure data flow between the SVC cluster (SSH server) and a client (either a command line client (CLI), or the CIMOM). The connection is secured by the means of a private/public key pair:

- Public key is uploaded to the SSH server.
- Private key identifies the client and is checked against the public key during the connection. The private key must be protected.
- The two keys are generated together.
- The SSH server must also identify itself with a specific host key.
- If the client does not have that key yet, it is added to a list of known hosts.
- If the client already has a key related to that server's IP address:
  - The client can overwrite the existing key.
  - The client can refuse the connection to allow checking of that key.

These mechanisms (public/private key pair and host key) are used so that each party is sure about the identity of the other one, as shown in Figure 14-27.

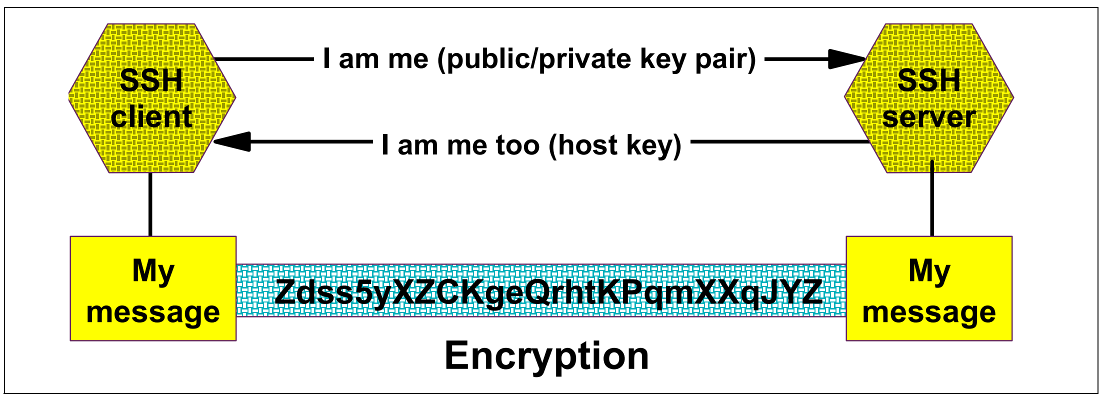


Figure 14-27 SSH client/server

The communication interfaces are shown in Figure 14-28.

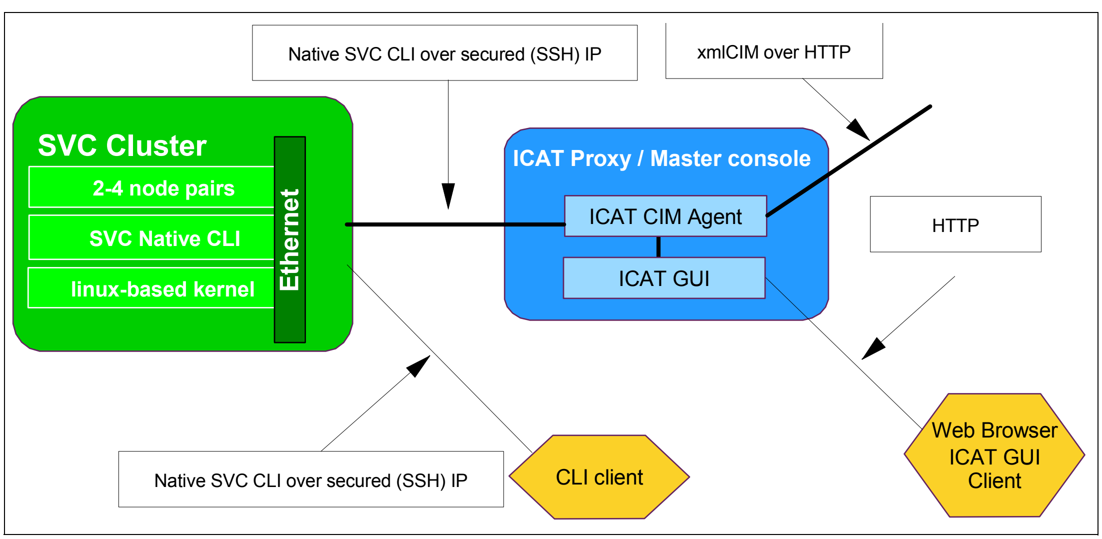


Figure 14-28 Communication interfaces

14.7 Call Home

SVC does not directly "call home". However, SVC makes all error reports available to network managers using industry standard error reporting techniques. Call Home is supported by providing the error data needed by an external call home agent. The necessary machine information to generate a call home record is sent to network management software via an SNMP trap.

SVC uses *e-mail to RETAIN®* to Call Home. The agent that receives the SNMP trap and generates the e-mail in the correct format is the IBM Director, which comes pre-installed in the SVC master console. The e-mail to RETAIN Call Home system works by sending an e-mail to an IBM catcher machine which reformats the data based on the machine type and model information and then opens the RETAIN record. The e-mail address for Call Home is `mailto:callhome1@de.ibm.com®` for United States and Canada, and `mailto:callhome0@de.ibm.com` for all other countries (regions).

To configure IBM Director settings for the SVC Call Home feature, complete these steps:

1. To launch the IBM Director, click the **IBM Director** icon on the SVC console desktop area.
2. You need to login using the master console Windows user ID and password.
3. After IBM Director becomes active, select **Options -> Discovery Preferences** as shown in Figure 14-29.

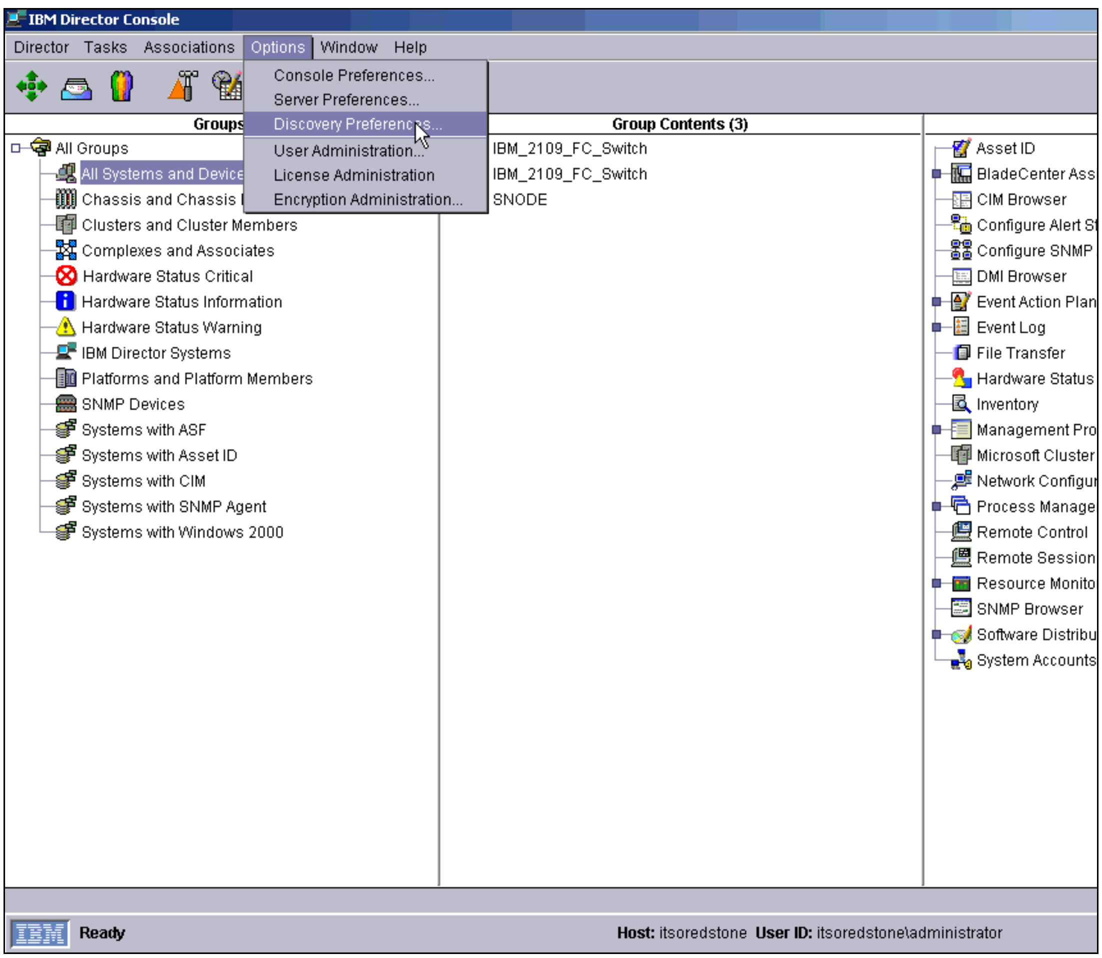


Figure 14-29 IBM Director Discovery Preferences

4. The Discovery Preferences window (Figure 14-30) opens. Select the **SNMP Discovery** tab. Change the entry to the IP address of the master console and the switches using a subnet mask of all zeros. Then click **Add** or **Replace**.

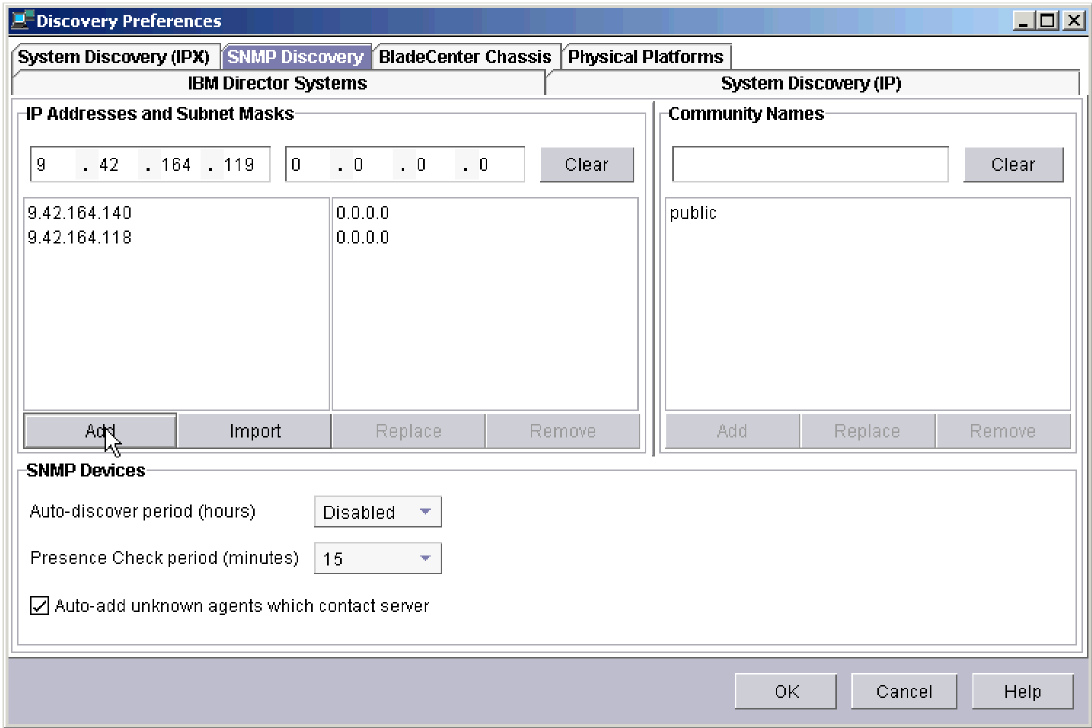


Figure 14-30   Network addresses for SNMP discovery

5. With the Director console application launched, select **Task -> Event Action Plan Builder**, as shown in Figure 14-31.

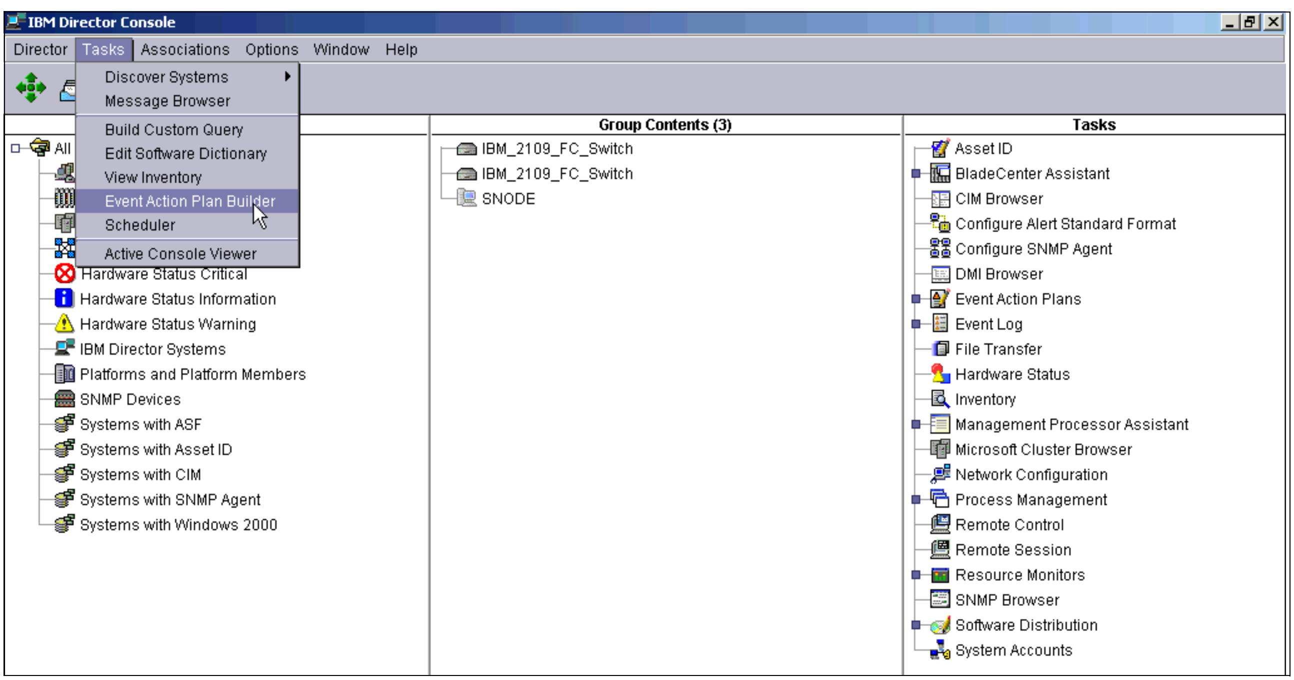


Figure 14-31   Director Action Plan Builder

6. On the Event Action Plan Builder window (Figure 14-32), in the Actions panel, select **Send an Internet (SMTP) E-mail**. Then right-click **2145** and select **Update**.

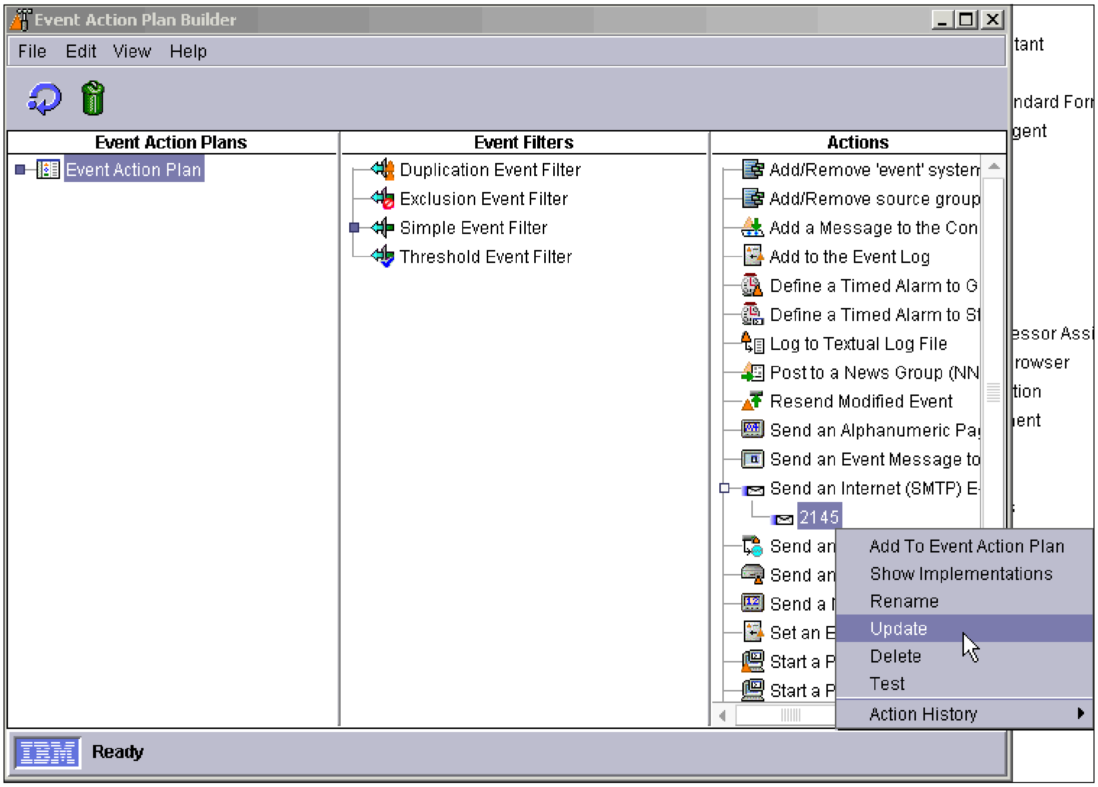


Figure 14-32 Updating the 2145 event

7. The Customize action:2145 window (Figure 14-33) opens. Complete the following items:

   – Internet e-mail address to which the SVC alerts are sent: Type the IBM Retain e-mail address:

     - CALLHOME1@de.ibm.com for U.S.A. and Canada clients
     - CALLHOME0@de.ibm.com for all other countries (regions)

   – Reply to: Type the e-mail address to which you want any replies to be directed.

   – SMTP E-mail server: Type the address of your e-mail (SMTP) server.

   – SMTP Port: Change this, if required, to your SMTP server port number.

   – Subject of E-mail Message: Type `2145 Error Notification`.

   – Body of the E-mail Message: Type the following information:

     - Contact name
     - Contact Phone number
     - Off-shift phone number
     - Machine location
     - `Record Type = 1`

> **Important:** Do not change the line with text `Record Type = 1`.

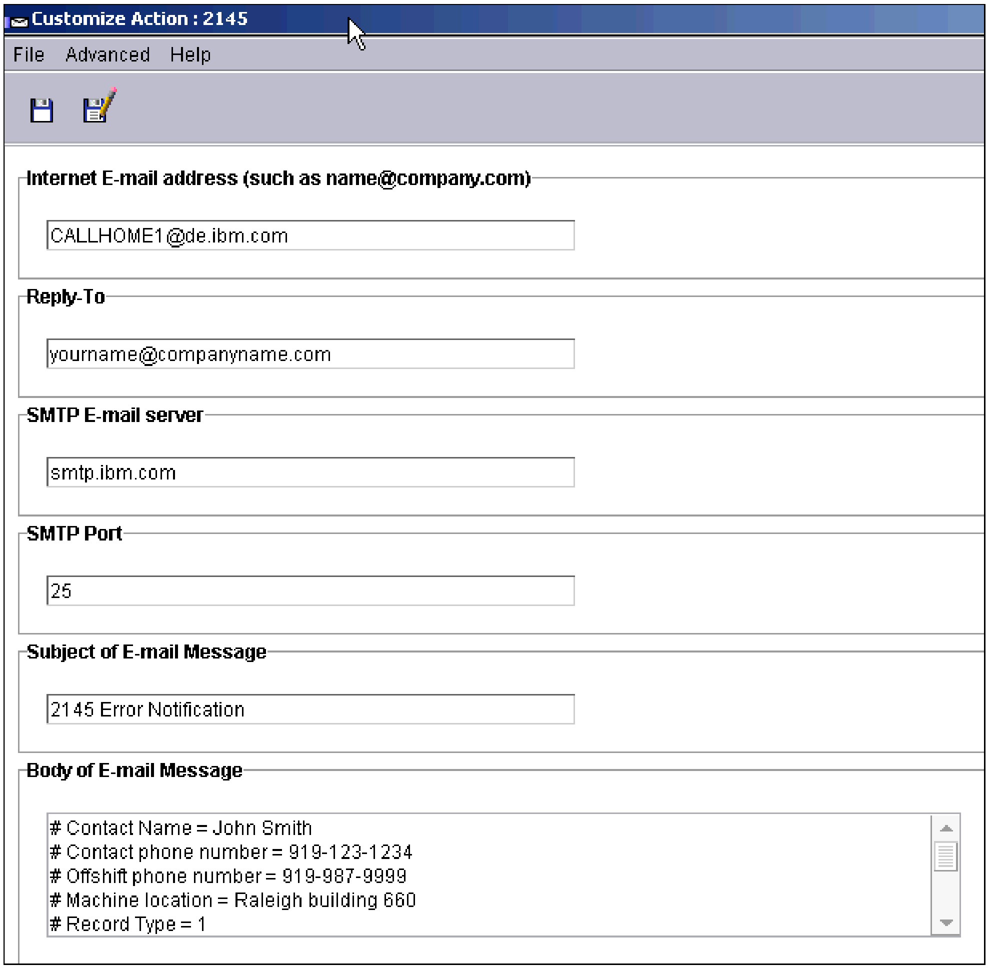


Figure 14-33 Customize Action 2145

14.8 Remote support

Remote support is provided by means of a Virtual Private Network (VPN). This allows the IBM Support Team to access your master console for diagnosis and problem determination purposes. The VPN is client controlled where the client sets up a connection to IBM. IBM Support needs the client-supplied password to tunnel back to their master console.

See the “Enhanced Remote Support” document on the SVC support Web site to set up the configuration for this facility:

`http://www.ibm.com/storage/support/2145`

A

Performance and capacity planning

Virtualization greatly simplifies the management and configuration of your storage infrastructure. It does not remove the fundamental requirement to have storage to meet your business requirements. We recommend that you have performance and capacity planning in place to achieve the desired end architecture.

This appendix discusses IBM TotalStorage SAN Volume Controller (SVC) performance and capacity planning.

Performance considerations

The virtualization features of the SVC allow the user to choose how to divide their storage and present it to hosts. While virtualization provides the user a great deal of flexibility, it does not remove the necessity to have an adequate number of hard disks.

As with all storage subsystem there, are three basics: capacity, input/output (I/O) rate, and throughput (MB/Sec). One is the limiting factor and you need enough back-end storage to meet this demand.

Maximum capability of one node pair (I/O group)

The SVC or IBM TotalStorage SAN Integration Server (SIS) are built from pairs of nodes (also called *storage engines*). Together, a pair of nodes is called an *I/O group*. One or two I/O groups can operate together as a SVC *cluster*.

A single I/O group can handle up to 140,000 read requests of 4 KB, where all I/Os are cache hits, or up to 39,900 write hit requests of 4 KB. The write number is lower because all the writes are mirrored to both nodes in a pair and the SVC must destage data to disk after it is written to both caches. Normal client workloads achieve less throughput than these numbers because real workloads include cache miss I/Os and may also transfer more than 4 KB.

For sequential read operations, such as database scans or backup operations, a single I/O group can achieve up to 890 MB/s provided that the back-end disk configuration is properly configured to provide this level of throughput. Close to this level of throughput is possible with client workloads.

A SVC or SIS configuration with two I/O groups (four nodes) achieves approximately double the throughput of a single I/O group cluster.

Latency

Latency is delay added to the response time for an I/O operation. All in-band storage virtualization solutions add some latency to cache miss I/Os. This is not a unique characteristic of the SVC or SIS.

However, the SVC latency is very low. For a 4 KB read operation, the SVC introduces approximately 60 µs (microseconds, millionths of a second) of additional delay. When you consider that a typical cache miss response time might be around 10 ms (milliseconds, thousandths of a second), you can see that the delay typically caused by the SVC is negligible (less than 1% of total response time).

In the real world, the effect of latency is normally even less. All writes to the SVC or SIS are cache hits, so they add no latency. Because of the cache, many reads are also cache hits. Only cache miss I/Os add latency.

In some cases, the SVC may be able to improve the performance of older disk systems with slow controllers or uncached disk systems. This can happen because of caching in the SVC or because the SVC is just faster than older disk system controllers.

SVC nodes planning guidelines

When considering the number of nodes to use per SVC cluster or in the SIS, note that the 890 MB/s sequential throughput capability of an I/O group is usually the limiting factor when configuring attached storage. If you have information about the workloads that you plan to use with the SVC or SIS, you can use that information to size the amount of capacity you can configure per I/O group. To be conservative, assume a throughput ability of about 800 MB/s per I/O group. Achieving the maximum of 890 MB/s requires careful configuration.

SVC paths to disk planning guidelines

When configuring IBM TotalStorage Fibre Array Storage Technology (FAStT) storage servers for use with the SVC, you should normally configure two controller ports in the FAStT 600 and have two SVC ports available to communicate with the FAStT system. This configuration, coupled with at least two host ports, is intended to meet the throughput requirements of most workloads.

When attaching existing disk systems to the SVC, normally keep the same number of paths into the disk system and the same number of ports on the host servers. Some modification to this guidance may be needed while migrating a disk system from native attachment to attachment through the SVC.

While keeping the host and disk system ports the same, you effectively deploy the SVC as a middle layer between the two, with access controlled through zoning or logical unit number (LUN) masking.

SVC managed and virtual disk layout planning

When configuring managed disks with the SVC, create managed disk groups (MDGs) to use the largest practical extent size. Doing so maximizes the learning ability of the SVC adaptive cache. Remember that managed disk space is allocated in units of whole extents, so virtual disks (vDisks) whose size is not a multiple of the extent size can waste space.

For example, an 18 GB vDisk uses 36 extents of 512 MB, but an 18.1 GB vDisk uses 37 such extents. Of course, a 17.9 GB vDisk also uses 36 extents of 512 MB and wastes little space.

If most of your vDisks are a multiple of 512 MB in size, use 512 MB extents. If you expect to have many LUNs that are not a multiple of 512MB, use the largest extent size that results in acceptable space utilization. Recent performance testing has shown that using an extent size of 128 MB has no appreciable effect on performance.

When configuring FAStT storage servers, consider using a segment size of 256 KB. This size helps to optimize sequential performance without hurting random I/O performance.

When creating vDisks, the default choice of striped allocation is normally the best choice. This option helps to balance I/Os across all the managed disks in a MDG, which tends to optimize overall performance and helps to reduce hot spots.

The SVC or SIS allocates a preferred path for each vDisk. This is the node within an I/O group that is normally used to handle I/Os. The default, which is to alternate vDisks between the nodes of an I/O group in the order the vDisks are created, normally produces good results. it also normally balances load well across both nodes of an I/O group. In highly unusual cases, such as where vDisks vary greatly in size or where the I/O load to different vDisks varies widely, it is possible that a significant imbalance can arise, which may deliver less than optimal performance. In these cases, when creating vDisks, you can use the -node parameter of the **mkvdisk** command (or the equivalent graphical user interface (GUI) function) to specify which node of an I/O group should be the preferred path.

There is no need to change the LUN definitions in existing disk systems when placing them under the control of the SVC. Simply include the LUNs in MDGs and start to create vDisks. When installing new storage to be managed by the SVC, you need define only a simple LUN arrangement. We recommend that you create one LUN per disk array in the disk system.

vDisks are created using managed disks within a MDG. Accordingly, all the managed disks in a single MDG should have the same (or similar) performance characteristics. If you mix managed disks with different performance characteristics, vDisks may exhibit “lumpy” performance where I/Os to different portions of the vDisk perform differently.

Application database scalability, performance, and capacity planning

The SVC is capable of delivering very high I/O demand. And as SVC capability to scale from two to four nodes initially, more in the future, it can handle approximately 60,000 I/Os per second for and SPC-1 type workload. The planners are encouraged to compare data posted on the SPC Web site at:

http://www.storageperformance.org/

Application database scalability, performance, and capacity planning require lots of other planning requirements such as a business plan, operational and maintenance plan, queuing models and tools, and benchmarking. Use Table A-1 as a performance comparison guideline.

Table A-1 SVC, SIS and different FAStT performance comparison

Product	Maximum cache	Maximum IOP	Sequential throughput (read MB/s)
SVC SIS with two node	8 GB	140,000	890
SVC SIS with four node	16 GB	280,000	1,780
FAStT 200	256 MB	11,800	170
FAStT 500	512 MB	45,000	380
FAStT 600	1 GB	60,000	392
FAStT 700	2 GB	110,000	390
FAStT 900	2 GB	148,000	790

B

Open system specifics

This appendix describes the basic tasks that you need to perform on the individual host systems when using IBM TotalStorage SAN Volume Controller (SVC) Copy Services. It explains how to bring FlashCopy target volumes online to the same host as well as to a second host. The appendix covers various AIX and Windows Intel platforms.

Attention: Most of this appendix is based on work carried out using IBM TotalStorage Fibre Array Storage Technology (FAStT) and Copy Services.

AIX specifics

The following section describes what is necessary to use Copy Services volumes on AIX systems.

AIX and FlashCopy

The FlashCopy functionality in SVC Copy Services copies the entire contents of a source volume to a target volume. If the source volume is defined to the AIX Logical Volume Manager (LVM), all of its data structures and identifiers are copied to the target volume as well. This includes the Volume Group Descriptor Area (VGDA), which contains the physical volume identifier (PVID) and volume group identifier (VGID).

For AIX LVM, it is currently not possible to activate a volume group with a disk that contains a VGID in the VGDA and a PVID that is already used in a active volume group on the same server, even if the hdisk PVID is cleared and reassigned with the following two commands:

```
chdev -l <hdisk#> -a pv=clear
chdev -l <hdisk#> -a pv=yes
```

Therefore, it is necessary to redefine the volume group information on the FlashCopy target using low level LVM commands or the **recreatevg** command. Refer to “AIX recreatevg command” on page 498. This alters the VGID on the VGDA of the FlashCopy target so that there are no conflicts with existing VGIDs on active volume groups. If you do not redefine the volume group information prior to importing the volume group, then the **importvg** command fails.

Accessing FlashCopy target on another AIX host

Accessing a FlashCopy target on another AIX host poses some problems. As a result of copying the entire disk contents, all the data structures and identifiers used by the LVM. If you FlashCopy only one half of a mirror, 50% of the hdisks in the VGDA would be *missing*. Therefore, it is necessary to either turn off quorum for the volume group that you are point-in-time copying, or you may need to force the varyon of the volume group on the target volumes.

The following procedure makes the data of the FlashCopy target volume available to another AIX that has no prior definitions in its configuration database:

1. The target volume (hdisk) is new to AIX. Therefore the Configuration Manager should be run on the specific SCSI or Fibre Channel adapter:

   ```
   cfgmgr -l <host_adapter>
   ```

2. Check which hdisk is your FlashCopy target:

   ```
   lsdev -Cc disk | grep 2105
   ```

3. Import the volume group:

   ```
   importvg -y <volume_group_name> <hdisk#>
   ```

4. Vary on the volume group. The **importvg** command should vary on the volume group:

   ```
   varyonvg <volume_group_name>
   ```

5. Verify consistency of all file systems on the FlashCopy target:

   ```
   fsck -y <filesystem_name>
   ```

6. Mount all the target file systems:

   ```
   mount <filesystem_name>
   ```

The data is now available. For example, you can create a backup of the data on the FlashCopy Volume to a tape device. This procedure may be run after the relationship between FlashCopy source and target is established, even if data is still being copied from the source to the target in the background.

It may be the case that the disk containing the target volume were previously defined to an AIX system, for example, if you periodically do backups from the same volume. In this case, it may depend on whether any LVM modifications have been made to the volume group. If no volume group, file system or logical volume changes were made, then perform the following actions on the target system:

1. Unmount all file systems in the source volume group:
   ```
   umount <src_filesystem>
   ```
2. Unmount all file systems in the target volume group:
   ```
   umount <tgt_filesystem>
   ```
3. Vary off the target volume group:
   ```
   varyoffvg <tgt_volume_group_name>
   ```
4. Perform the FlashCopy to the target volume.
5. Mount all file systems in the source volume group:
   ```
   mount <src_filesystem>
   ```
6. Vary on the target volume group:
   ```
   varyonvg <tgt_volume_group_name>
   ```
7. Perform a file system consistency check on the target:
   ```
   fsck -y <tgt_filesystem>
   ```
8. Mount all target file systems in the volume group:
   ```
   mount <tgt_filesystem>
   ```

If volume group modifications were made, such as changing file system size or the modification of logical volumes (LV), then we recommend that you do not perform the previous set of tasks. Instead, we recommend that you perform the following tasks on the target system:

1. Unmount all file systems in the target volume group:
   ```
   umount <tgt_filesystem>
   ```
2. Vary off the volume group:
   ```
   varyoffvg <tgt_volume_group_name>
   ```
3. Export the volume group:
   ```
   exportvg <tgt_volume_group_name>
   ```
4. Delete the target volumes:
   ```
   rmdev -dl <hdisk#>
   ```
5. Perform the FlashCopy to the target volumes.
6. Perform tasks as though the volumes were new to the system as previously described.

Accessing FlashCopy source and target on the same AIX host

This section describes a method to access the FlashCopy target volume on a single AIX host while the source is active on the same server. The procedure is intended to be used as a guide and may not cover all scenarios.

AIX recreatevg command

Point-in-time copying a source volume's content in all cases causes all of the data structures and identifiers used by the AIX LVM to be duplicated to the target volume. The duplicate definitions, in turn, cause conflicts within LVM. Until recently, none of the existing set of LVM commands have had the capability to access the logical volumes or file systems on the target disk. This problem is solved now using the AIX command **`recreatevg`**.

The **`recreatevg`** command is packaged as a PTF for AIX 4.3.3 in APAR IY10456 and higher. It is officially available in:

- AIX 4.3.3 Recommended Maintenance Level 05 (RML05) or higher
- AIX 5L Version 5.1

The **`recreatevg`** command overcomes the problem of duplicated LVM data structures and identifiers caused by a disk copying process such as FlashCopy. It is used to recreate an AIX volume group (VG) on a set of disks that are copied from a set of disks belonging to a specific VG. The command allocates new PVIDs for the member disks and a new VGID to the volume group. The command also provides options to rename the logical volumes with a prefix you specify, and options to rename "labels" to specify different mount points for file systems.

Here is the AIX *man* page synopsis.

```
recreatevg [-y VGname] [-p] [-f] [-Y lv_prefix | -l LvNameFile] [-L label_prefix] [-n] \
PVname...
```

You can use this command to recreate a VG on a set of disks that are mirrored from a set of disks belonging to a specific VG. This command allocates new PVID for the member disks since the PVIDs are also duplicated by the disk mirroring. Similarly, other LVM logical members that are duplicated are also changed to new names with the specified prefixes.

Note the following flags:

- **-y VGname** specifies the volume group name rather than having the name generated automatically. Volume group names must be unique system wide and can range from one to 15 characters. The name cannot begin with a prefix already defined in the PdDv class in the device configuration database for other devices. The volume group name that is created is sent to standard output.
- **-p** disables the automatic generation of the new PVIDs. If -p flag is used, you must ensure that there are no duplicated PVIDs on the system. All the disks that were hardware mirrored must have had their PVIDs changed to an unique value.
- **-Y lv_prefix** causes the logical volumes on the VG being recreated renamed with this prefix. For the number of characters in the prefix, the total length of the prefix and the logical volume name must be less than or equal to 15 characters. If the length exceeds 15 characters, the logical volume is renamed with the default name. The name cannot begin with a prefix already defined in the PdDv class in the device configuration database for other devices, nor can it be a name already used by another device.
- **-L label_prefix** causes the labels of logical volumes on the VG being recreated changed with this prefix. User must modify the /etc/file systems stanza manually if a simple modification of the mount point is not enough to define the stanza uniquely.
- **-l LvNameFile** entries in the LvNameFile must be in the format `LV1:NEWLV1`. After **`recreatevg`**, LV1 is renamed with NEWLV1. All the logical volumes that are not included in the LvNameFile will be recreated with the default system generated name.
- **-f** allows a volume group to be recreated that does not have all disks available.
- **-n**: After **`recreatevg`**, the volume group is imported but varied off. The default is imported and vary on.

Note the following points:

- To use this command, you must have root user authority.
- All the member physical volumes of the volume group must be specified on the command line. The command fails if the input list does not match with the list compiled from the VGDA.
- If you perform a Copy Services function on one half of a RAID-1 pair to reduce the capacity required for FlashCopy targets or Remote Copy secondary volumes, then use the -f option to force the creation of the volume group. Otherwise the VGDA has PVIDs of volumes that made up the other half of the mirror at the source or primary site.

Consider these examples:

- To recreate a volume group that contains three physical volumes, enter this command:

  ```
  recreatevg hdisk1 hdisk2 hdisk3
  ```

  The volume group on hdisk1, hdisk2, and hdisk3 is recreated with an automatically generated name, which is displayed.
- Enter the following command:

  ```
  recreatevg -y testvg hdisk1
  ```

  The volume group on hdisk1 is recreated with the new name testvg.
- Enter the following command:

  ```
  recreatevg -Y newlv hdisk14
  ```

  The volume group on hdisk14 is recreated. All logical volumes in that volume group are recreated and renamed with the prefix newlv.

Accessing FlashCopy target with recreatevg

For example, we have a volume group that contains two volumes (hdisks). We want to FlashCopy the volumes for the purpose of creating a backup. To achieve this, we must have two target LUNs available of size equal to or greater than the sources in the same LSS.

The source volume group is fc_source_vg containing hdisk4 and hdisk5. The target volume group is fc_target_vg containing hdisk8 and hdisk9.

Perform the following tasks to create the FlashCopy and make the target volumes available to AIX:

1. Stop all applications that access the FlashCopy source volumes.
2. Unmount all related file systems for the short period of FlashCopy establishment.
3. Establish the FlashCopy pairs with the copy parameter set to zero. This is the equivalent of the IBM TotalStorage Enterprise Storage Server (ESS) message "`Do not perform background copy`".
4. Mount all related file systems.
5. Restart applications that access the FlashCopy source volumes.
6. The target volumes, hdisk8 and hdisk9, now have the same volume group data structures as the source volumes hdisk4 and hdisk5. Clear the PVIDs from the target hdisks to allow a new volume group to be made:

   ```
   chdev -l hdisk8 -a pv=clear
   chdev -l hdisk9 -a pv=clear
   ```

   The output from **`lspv`** shows the result. See Figure B-1.

```
# lspv
hdisk4        000567992d4c9024    fc_source_vg
hdisk5        000567995abe005e    fc_source_vg
hdisk8        none                None
hdisk9        none                None
```

Figure B-1 lspv after pv=clear

7. Create the target volume group and prefix all file system path names with `/backup` and prefix all AIX logical volumes with `bkup`*:*

   `recreatevg -y fc_target_vg -L /backup -Y bkup hdisk8 hdisk9`

   Specify the hdisk names of all disk volumes participating in the volume group. The output from **`lspv`** illustrates the new volume group definition. See Figure B-2.

```
# lspv
hdisk4        000567992d4c9024    fc_source_vg
hdisk5        000567995abe005e    fc_source_vg
hdisk8        000567995abdf345    fc_target_vg
hdisk9        00056799b2d831b9    fc_target_vg
```

Figure B-2 Recreated FlashCopy target volumes

An extract from /etc/filesystems shows how **`recreatevg`** generates a new file system stanza. The file system named /u01 in the source volume group is renamed to /backup/u01 in the target volume group. Also, the directory /backup/u01 is created. Notice, also, that the logical volume and JFS log logical volume are renamed. The remainder of the stanza is the same as the stanza for /u01. See Figure B-3.

```
/backup/u01:
        dev             = /dev/bkupelv001
        vfs             = jfs
        log             = /dev/bkupelvlog001
        mount           = true
        check           = false
        options         = rw
        account         = false
```

Figure B-3 Target file system stanza

8. Mount the new file systems that belong to the target to make them accessible.

AIX and Remote Copy

When you have Remote Copy primary and secondary volumes in a copy pair relationship, it is not possible to read the secondary. Therefore, if you are configuring the secondary volumes onto the target server, it is necessary to terminate the copy pair relationship.

After the volumes are in simplex state, the secondary volumes can be configured (**`cfgmgr`**) into the target systems customized device class (CuDv). This brings in the secondary volumes as hdisks and contains the same PVID as the primary volumes. Because these volumes are new to the system, there is no conflict with existing PVIDs. The volume group on the secondary volumes containing the LV and file system information can now be imported into the Object Data Manager (ODM) and the /etc/filesystems file using the **`importvg`** command.

If the Remote Copy secondary volumes were previously defined on the target AIX system as hdisks or vpaths, but the original volume group was destroyed on the primary volumes, you must remove the old volume group and disk definitions (using **exportvg** and **rmdev**) and redefine (**cfgmgr**) them before running the **importvg** command to attain the new volume group definitions. If this is not done first, the **importvg** command imports the volume group improperly. The volume group data structures (PVIDs, and so on) in ODM differ from the data structures in the VGDAs and disk volume super blocks. The file systems are not accessible.

When you execute the **lsdev -Cc disk** command, you observe that the state of the original Remote Copy secondary volumes becomes *Defined* during reboot (shown in Figure B-4).

```
# lsdev -Cc disk
hdisk6  Defined   20-58-01     IBM FC 2105F20
hdisk7  Defined   20-58-01     IBM FC 2105F20
hdisk8  Defined   20-58-01     IBM FC 2105F20
hdisk9  Defined   20-58-01     IBM FC 2105F20
hdisk13 Available 20-58-01     IBM FC 2105F20
hdisk14 Available 20-58-01     IBM FC 2105F20
hdisk15 Available 20-58-01     IBM FC 2105F20
hdisk16 Available 20-58-01     IBM FC 2105F20
```

Figure B-4 Phantom hdisks

It is important to execute both the **lspv** and **lsdev** commands back-to-back, so that you can be certain which disks are the phantoms. From the **lspv** output, the phantom disks have no PVIDs and are not assigned to a volume group. From the **lsdev** output, the phantom is in an available state. The original disks have PVIDs, are assigned to a volume group, and are marked in a defined sate.

To remove the phantom hdisks from the configuration database, run the **rmdev -dl** command on each phantom disk device as shown here:

```
# for i in 13 14 15 16
do
rmdev -dl hdisk$i
done
```

Set the original hdisks to an available state with the **mkdev** command as shown here:

```
# for i in 6 7 8 9
do
mkdev -l hdisk$i
done
```

You can reactivate the volume group, evg001 (**varyonvg**) and mount its file systems.

Making updates to the LVM information

When performing Remote Copy between primary and secondary volumes, the primary AIX host may create, modify, or delete existing LVM information from a volume group. However, because the secondary volume is not accessible when in a Remote Copy relationship and when the read-from- secondary has not been enabled, the LVM information in the secondary AIX host is out-of-date. Therefore, you need to allot scheduled periods where write I/Os to the primary Remote Copy volume can be quiesced and file systems unmounted. At this point, the copy pair relationship can be terminated, and the secondary AIX host can perform a learn on the volume group (**importvg -L**).

After the updates are read into the secondary AIX hosts ODM, you can establish the Remote Copy copy pair again. Select **Do not copy volume** from Select copy options when establishing the Remote Copy copy pair.

As soon as the Remote Copy pair is established, immediately suspend the Remote Copy relationship. Because there was no write I/O to the primary volumes, both primary and secondary are consistent. Now that the primary volume is suspended, you can remount the file systems and resume write I/O. After the write I/O has been going for a while, you can re-establish the relationship with the primary and secondary by choosing the **Copy out-of-sync cylinders only** option.

If the read from secondary option was selected during the Remote Copy copy pair establish, then we recommend that you suspend the primary volume and then perform the import learn function. When complete, re-establish the copy pair only copying the out-of-sync cylinders.

Figure B-5 shows two systems, sanf50 and sanh70, where sanf50 has the primary volume vpath5 and sanh70 has the secondary volume vpath16. Both systems have had their ODMs populated with the volume group itsovg from their respective Remote Copy volumes. Prior to any modification of both systems, ODMs have the same time stamp.

```
root@sanf50:/ > getlvodm -T itsovg
3d99d8911542ab68

root@sanh70:/ > getlvodm -T itsovg
3d99d8911542ab68
```

Figure B-5 Original time stamp

Volumes vpath5 and vpath16 are in a Remote Copy duplex state. The volume group itsovg on sanf50 is updated with a new logical volume. The time stamp on the VGDA of the volumes gets updated and so does the ODM on sanf50, but not sanh70 as shown in Figure B-6.

```
root@sanf50:/ > lqueryvg -p vpath5 -Tt
Time Stamp:     3d99e5de077392e4
root@sanf50:/ > getlvodm -T itsovg
3d99e5de077392e4

root@sanh70:/ > lqueryvg -p vpath16 -Tt
Time Stamp:     3d99e5de077392e4
root@sanh70:/ > getlvodm -T itsovg
3d99d8911542ab68
```

Figure B-6 Updated source time stamp

To update the ODM on the secondary server, we recommend that you suspend the Remote Copy copy pair prior to performing the **`importvg -L`** command. This stops any conflicts from LVM actions occuring on the primary server. Figure B-7 shows the updated ODM entry on sanh70.

```
root@sanh70:/ > importvg -L itsovg vpath16
root@sanh70:/ > lqueryvg -p vpath16 -Tt
3d99e5de077392e4
root@sanh70:/ > getlvodm -T itsovg
3d99e5de077392e4
```

Figure B-7 Update secondary server's ODM

After the `importvg -L` has completed, you can reestablish the Remote Copy copy pairs and copy only the out-of-sync cylinders.

Windows NT and 2000 specifics

This section discusses the tasks that are necessary when performing Copy Services operations on volumes owned by Microsoft Windows NT and 2000 operating systems.

Windows NT and Copy Services

This section discusses the actions that you need to perform on Remote Copy and FlashCopy volumes owned by Microsoft Windows NT operating systems.

Windows NT handles disks in a way that is not similar to any other operating system covered in this book. The need to reboot a server to scan for new disks and the need to run a GUI-based Disk Administrator to manipulate the disks are the main factors that restrict the routine use of Remote Copy and FlashCopy and make automation virtually impossible. It is possible to automate the actions of the GUI-based Disk Administrator using third-party software to remotely reboot the server. It is also possible to remotely assign the drive letter from the server that starts the Copy Services task. This was not tested during our project.

If you are going to create an automated script with Windows NT, you need to be careful about data consistency. It may be that some part of the automation process may run a script on a source server, and subsequent actions maybe taken by a script on a target server. Therefore, interprocess communication across servers may be required for timing. Otherwise, you may get inconsistent data. Not all applications allow this.

You have two options on how to make the Remote Copy or FlashCopy target available to the server: *with reboot* or *without reboot*. We recommend that you reboot the server. It is safer because then it is guaranteed that all the registry entries are created. However, using Remote Copy or FlashCopy without rebooting is faster.

Registering the Remote Copy and FlashCopy volumes to Windows NT

If you are going to reboot the server, you do not have to make the target disks known to Windows NT before you do the Remote Copy or FlashCopy. However, we recommend that you preassign and register them in the server. The “assign disk and run Remote Copy or FlashCopy” approach is useful for a *non-routine* Remote Copy or FlashCopy, for example, for testing or migration.

For routine purposes, we recommend that you have target disks already present in Disk Administrator with partitions created and partition information saved. Click **Start -> Programs -> Administrative Tools -> Disk Administrator**. Then follow these steps:

1. If the target disk was not previously seen by the system, Disk Administrator issues a pop-up message saying “`No signature on Disk X. Should I write a signature?`”, where *X* is the number assigned to the newly present disk.

   Click **OK** to save the signature on the target disk.

2. The Disk Administrator opens. Click the disk that is to be used as the Remote Copy or FlashCopy target (it should be gray and marked as free space) and select **Create**.
3. Confirm the partition parameters and click **OK**. The partition appears as *Unknown*.
4. Click the newly created partition and select **Commit Changes Now**.
5. Right-click the partition and select **Assign Drive letter**.

6. Assign a drive letter and click **OK**.
7. Exit Disk Administrator.

After this procedure, the Remote Copy or FlashCopy target is properly registered in the Windows NT.

Bringing down the target server

Bring down the server that will use the target if you want to use the safer method. Also keep in mind that if you assign the volume to the host just before you perform the Remote Copy or FlashCopy, you must use the volume serial number for the target.

Performing a Remote Copy or FlashCopy

Stop all applications using the source volume. Now flush the data to the source volume. Click **Start ->Programs -> Administrative Tools -> Disk Administrator**. Then follow these steps:

1. Right-click the disk that is to be used as the Remote Copy or FlashCopy source. It should have a drive letter assigned and be formatted. Then select **Assign Drive letter**.
2. From the pop-up window, select **Do not assign a drive letter** and click **OK**.
3. Now the data is flushed to the source. You can start the Remote Copy or FlashCopy task from the SVC Copy Services Web Interface or from any server CLI.
4. Observe the GUI. Or enter the following command to see if the Remote Copy or FlashCopy task successfully started:

   ```
   lsfcmapprogress/lsrcrelationshipprogress
   ```

5. Reassign the drive letter to the source volume. Right-click the disk that is a Remote Copy or FlashCopy source and select **Assign Drive Letter**.
6. Assign a drive letter and click **OK**.
7. Exit Disk Administrator.

You can resume using the source volume.

Bringing up the target server

Next you may boot up the target server. In this case, you just assigned the target volumes to the host that will create the disk entry in the Windows NT registry. To verify that the registry entry is created, complete these tasks:

1. Click **Start -> Settings -> Control Panel -> Hardware -> Device Manager**.
2. In Control Panel, double-click **Disk Drives**.
3. Click the adapter that has the target volume attached.
4. A list of targets opens. Verify the list includes the target ID and LUN of the volume you just made available to the server. If you are using SDD, you see each disk entry several times `[(# of vDisks) x (# of Nodes) x (4 Ports/Node) x (# of HBAs/host)]`, which is the number of paths to the volume that you have.

You may also run the **`datapath query device`** command from the SDD command line to check whether the Remote Copy or FlashCopy targets are listed between the volumes. This command also enables you to check volume serial numbers and gives you a more understandable overview of the volumes and their paths.

Making the Remote Copy or FlashCopy target available

Log in, start the Windows NT Disk Administrator, write a signature if necessary (do not write a signature if data was already copied into this volume), and assign a drive letter. To begin, click **Start -> Programs -> Administrative Tools -> Disk Administrator**. Then follow these steps:

1. If the disk was not previously seen by this system, Disk Administrator issues the "`No signature on Disk X. Should I write a signature?`" message, where `X` is the number assigned to the newly present disk. Click **OK** to save the signature on the target disk.
2. The Disk Administrator opens. Click the disk that is a Remote Copy or FlashCopy target. You should see a formatted partition on it. Select **Assign Drive Letter**.
3. If you cannot assign a drive letter, the target may be corrupt. Try repeating the whole process and consider the scenario that includes reboot.
4. Assign a drive letter and click **OK**. Exit Disk Administrator.
5. From a Windows NT command prompt, run the following command, where x is the letter assigned to the Remote Copy or FlashCopy target:

   ```
   chkdsk x: /f /r
   ```

   An option is to run the disk check from **Properties** of a disk in Windows NT Explorer.

After you complete this procedure, the Remote Copy or FlashCopy target is available to the Windows NT and can be handled like normal disk.

Windows 2000 and Copy Services

Windows 2000 handles its disks differently than Windows NT does. Windows 2000 incorporates a stripped-down version of the Veritas Volume Manager, called the *logical disk manager* (LDM).

With the LDM, you can create logical partitions, perform disk mounts, and create dynamic volumes. There are five types of dynamic volumes: simple, spanned, mirrored, striped, and RAID-5.

On Windows NT, the information relating to the disks was stored in the Windows NT registry. With Windows 2000, this information is stored on the disk drive itself in a partition called the *LDM database*, which is kept on the last few tracks of the disk. Each volume has its own 128-bit Globally Unique Identifier (GUID). This is similar to the disk PVID in AIX. Since the LDM is stored on the physical drive itself, with Windows 2000, it is possible to move disk drives between different computers.

Copy Services limitations with Windows 2000

Having the drive information stored on the disk itself imposes some limitations when using Copy Services functionality on a Windows 2000 system:

- The source and target volumes *must* be of the same physical size for two reasons:
  - The LDM database holds information relating to the size of the volume. Since this is copied from the source to the target, if the target volume is a different size from the source, then the database information is incorrect, and the host system returns an exception.
  - The LDM database is stored at the end of the volume. The copy process is a track-by-track copy, unless the target is an identical size to the source, the database is not at the end of the target volume.
- It is *not* possible to have the source and target FlashCopy Volume on the same Windows 2000 System, when they were created as Windows 2000 dynamic volumes. The reason is

that each dynamic volume must have its own 128-bit GUID. As its name implies, the GUID must be unique on one system. When you perform FlashCopy, the GUID is copied as well. This means that if you tried to mount the source and target volume on the same host system, you would have two volumes with exactly the same GUID. This is not allowed, and you are not able to mount the target volume.

Mounting a Copy Services target volume

To see target volumes on a second Windows 2000 host, you have to complete these tasks:

1. Perform the Remote Copy or FlashCopy function onto the target volume. Ensure that when using Remote Copy, that the primary and secondary volumes were in consistent mode, and write I/O was ceased prior to terminating the copy pair relationship.
2. Reboot the host machine on which you want to mount the Copy Services target volume.
3. Click **Computer Management-> Disk Management**.
4. Find the Disk that is associated with your volume. There are two "panes" for each disk. The left pane should read *Dynamic and Foreign*. It is likely that no drive letter is associated with that volume.
5. Right-click that pane, and select **Import Foreign Disks**. Select **OK**, and then **OK** again. The volume now has a drive letter assigned to it. It is of Simple Layout and Dynamic Type. You can read and write to that volume.

> **Tip:** Disable the *fast-indexing* option on the source disk. Otherwise, operations to that volume are cached to speed up disk access. However, this means that data is not flushed from memory and the target disk may have copies of files or folders that were deleted from the source system.

When performing subsequent Remote Copy or FlashCopy copies to the target volume, it is not necessary to perform a reboot because the target volume is still known to the target system. However, to detect any changes to the contents of the target volume, it is necessary to run the following command on the target volume:

```
chkdsk.exe /F
```

Copy Services with Windows Volume Sets

The following section explains how to perform Copy Services functions with Windows Volume Sets and volume groups (Windows 2000).

Copy Services with Windows NT Volume Sets

Both Remote Copy and FlashCopy are supported when using normal disks and Volume Sets. When using either Remote Copy or FlashCopy with Volume Sets, because these outboard copy features do not copy the Volume Set information in the Windows Registry, certain limitations exist and a special procedure is required as outlined in below. After SP6, it is possible to have the FlashCopy source and target volumes accessible by the same server. Prior to SP6, the FlashCopy source and target volumes must be attached to different servers. Remote Copy primary and secondary volumes must be attached to different servers.

Procedure for using Remote Copy and FlashCopy with Volume Sets

This special procedure is required to FlashCopy or Remote Copy a Windows NT volume set. The procedure can also be applied to other Windows NT fault tolerant disk configurations, such as mirrored sets, striped sets, and striped sets with parity.

Consider the case where the target disks are in the same order as the source disks, *and* the target disks are contiguous (that is all the disks are next to each other as viewed by the target machine's Disk Administrator). Then simply create an identical volume set on the target machine and reboot prior to performing the FlashCopy. You do this before you perform FlashCopy or Remote Copy for the first time. Subsequent copies should work as expected, provided that the file system is unmounted (the drive letter is unassigned) on the target prior to performing a copy.

If the target disks do not appear contiguous to Windows NT or appear in a different order than on the source machine, then a different procedure must be used. Microsoft's FTEDIT, available on the NT Resource Kit, is a Microsoft supported tool designed to write volume set information into the registry. Using FTEDIT is much safer than editing the registry directly.

> **Attention:** Incorrect use of FTEDIT could result in loss of access to software RAID arrays. We recommend that you use Disk Administrator to save your disk configuration before using FTEDIT. In general, most errors made using FTEDIT are recoverable. For more information about how to recover from FTEDIT errors, and on FTEDIT in general, see the *Microsoft Knowledge Base* article for Q131658:
>
> `http://support.microsoft.com/default.aspx?scid=kb;en-us;131658`

The following procedure explains how to use FlashCopy and Remote Copy with FTEDIT.

Preparation

On the target machine, complete the following tasks:

1. Back up the disk data using Disk Administrator, and registry information using REGEDIT.
2. If the target disks were previously used, delete all of the target disks in Disk Administrator. Do not simply unmount them, but delete all of the partitions on the target disks. Commit the changes.
3. In the control panel, double-click **Devices**. Make sure that Ftdisk is started and set to start on boot. Ftdisk is the driver used by Windows NT to identify and access fault tolerant drives, such as volume sets. If there are any fault tolerant drives in use on the system, Ftdisk is started and set to start on boot. If it is not started, one way to start it is to create a fault tolerant drive on a couple of spare disks. This requires a reboot.

On the source machine, obtain the order in which the disks were added to the volume set. One way to do this is to use a freeware utility called *diskkey.exe*, available from:

`http://www.sysinternals.com`

This utility is not supported by IBM and is known to report disk numbering and other information that is different than what Disk Administrator reports. However, the order in which the disks are included in the volume set is correct. Also the correct ordering of the disks is the information required to create a duplicate volume set on the target server.

Map the disks on the source machine to the disks on the target machine. For example, determine that Disk6 on the source is FlashCopy copied to Disk9 on the target.

Performing the Remote Copy and FlashCopy

On the target machine, follow these steps:

1. Run the FlashCopy establish or Remote Copy terminate tasks.
2. Start Disk Administrator. If it asks you to write a signature on any of the disks, click **No** (except in the special cases, see the following Attention box). After Disk Administrator is up, commit the changes (this is *very* important), and close Disk Administrator.

> **Attention:** Disk Administrator asks to write a signature when the FlashCopy is performed to the same machine because it detects a duplicate disk signature (the source and target volumes have the same disk signature) and needs to write a new one. It is safe to do this, but be sure that you are writing the signature to the FlashCopy target disk. If a signature is written to the wrong disk, it may cause data corruption.
>
> When FlashCopying to a different machine, usually the disk signature on the target machine's disks are different than the FlashCopy source disks' signature, so Disk Administrator does not need to write a new signature to the target disks to use it. It is unlikely, but possible, that by coincidence the disk signature one of the source disks is the same as one of the disks on the target machine. In this case, you must write a signature on the target disk before you use it. Again, it is safe to do this, but be sure that you are writing the signature to the right disk.

3. Start FTEDIT. Select **Start -> Resource Kit 4.0 -> Disk Tools -> Fault Tolerance Editor**.
4. Read the warning and click **OK**.
5. There are two panes in the FTEDIT window. On the left pane is a list of the disks in the system. On the right pane is the list of partitions on that disk. You must add the disks to the volume set in the right order. Use the results of diskkey.exe to determine the order in which the disks were added on the source volume set.

> **Note:** If active Remote Copy target volumes are on the target, then the disk numbering used in FTEDIT may differ from the disk numbering used in the Disk Administrator. The Remote Copy target volumes are not seen by FTEDIT and are not included in the disk numbering scheme. Adjust your disk choices accordingly.

6. Click **Make FT** set in the lower left corner.
7. When it asks you what kind of set you want, choose **Volume set** and click **OK**.
8. Click the first **target disk** in the left pane.
9. The list of partitions on that disk should appear in the right pane. Choose the partition that contains the volume set on that disk (usually Partition 1). Double-click **Partition 1** in the right pane. This adds this disk or partition to the volume set, in order.
10. Repeat Steps 8 and 9 for the rest of the disks. If you make a mistake, you can cancel and start from scratch. The disks *must* be added in the correct order.
11. After you add all of the disks, click **Save FT set** at the bottom.
12. Click **Edit->Save Changes to System**.
13. Close Ftedit.
14. Reboot the system.
15. When Windows NT restarts, start Disk Administrator. The target disks should be yellow now, indicating that they are in a volume set. Assign a drive letter and commit the

changes. If the drives are not usable at this point, then the disks were probably added in the wrong order.

As long as the disk configuration does not change on the source or target, FlashCopy should work as expected. If the disk configuration is changed in anyway, such as adding an additional disk to the volume set or rearranging the disks, then you have to perform this procedure again.

Copy Services with Windows 2000 Volumes

Basic disks are the same as the Windows NT disks with the same restrictions. Dynamic disks are supported for both Remote Copy and FlashCopy and the primary source and secondary target volumes must be attached to different servers. We also support the use of spanned volumes with Remote Copy and FlashCopy on dynamic disks with the same attachment restriction.

Extending simple volume

The Copy Services source may initially be a single *simple* volume. However, as requirements change on the application server, the logical volume may be extended over two or more volumes. You should not independently extend the target volumes, but let Windows 2000 detect the correct sequence of the extended volumes during the import process.

For this reason the target volumes should be reverted back to *basic* disks prior to the initial FlashCopy after the source has been extended. The target server should also be rebooted for disk manager to pick up the new volumes. After the reboot, the volumes are recognized as *foreign* disks, and you can proceed to import them. Reboot of the target system on subsequent FlashCopy copies is not necessary until the source volume is further extended. However, following each FlashCopy, run the **`chkdsk /F`** command against the target.

Enlarging an extended and spanned volumes

When you have an *extended* or *spanned* disks, the logical drive may in time grow to include more of the initial volume (extended disk) or include additional volumes. When this occurs, it is necessary, as before, to remove the target volume group information and revert the target volumes back to basic disks.

On the initial FlashCopy, reboot the target server to configure the additional disks. Then import all the foreign disks that are part of the volume group. On subsequent FlashCopy copies to the target volume group, run only a **`chkdsk /F`** command on the target volume.

Remote Copy and Windows spanned volumes

We followed this procedure when carrying out the Remote Copy of a Windows 2000 spanned volume set from Server (A) to Server (B):

1. On the source server (A), we created a Windows spanned volume set of multiple dynamic disks.
2. We rebooted the target server (B), imported multiple target disks, and wrote a disk signature on each as basic disks.
3. We established Remote Copy between the source (A) and target volumes (B).
4. After the source and target volumes were synchronized, we terminated Remote Copy.
5. We rebooted the target host (B.)
6. We started Disk Manager, the Remote Copy target volumes were seen as *Foreign Dynamic Disks.*
7. The disks were imported into the target host and were seen as a spanned volume.

To demonstrate failback to the original setup, we carried out the following steps:

1. We removed the original paths and re-established them in the reverse direction from (B) to (A).
2. We removed the spanned volume drive letter from the original source, the spanned volume on server (A).
3. We established Remote Copy from (B) to (A) and wrote some data onto the spanned volume.
4. Remote Copy was terminated.
5. We restored the drive letter to the spanned volume on server (A).
6. The contents of the spanned volume could now be read from server (A).

C

FAStT migration scenarios

This appendix gives a high-level overview of migrating from normal storage area network (SAN)-attached storage to virtualized storage. In all of the examples, we use the IBM TotalStorage Fibre Array Storage Technology (FAStT) as the storage system.

Initial considerations

Here are some basic factors that you must take into account in all situations before starting:

- Each partition can contain up to 32 logical unit numbers (LUNs).
- Host device drivers must be changed, so *all* LUNs in a host partition must be moved to the SVC partition in one step.
- Each partition can only access a unique set of host bus adapter (HBA) ports (as defined by worldwide port names (WWPNs)).
- Only one storage partition must be created that includes any IBM TotalStorage SAN Volume Controller (SVC) ports of nodes that are in the same SVC cluster.
- The contents of an existing partition *must* be moved to the SVC partition at the same time. Some configurations may require backup, reconfigure, and restore.
- Some versions of FAStT allow RAID arrays to be expanded allowing their capacity to increase, which is not recommended, although it may be helpful for some configurations.

Important: If you have more than 32 logical units, then some spare storage is required to allow for temporary migration of the data while the FAStT is re-configured to have fewer logical units.

Scenario 1: Less than 32 logical units in all partitions

Existing host or hosts attached to FAStT have less than 32 LUNs on FAStT. Only one partition per host or both partitions on a single host have less than 32 LUNs when combined. Figure C-1 shows the initial configuration.

Note: Access LUN (A) is used by Host A for inband configuration of the FAStT. This is deleted and FAStT is configured to use SVC master console over Ethernet. Access LUN is not required by SVC.

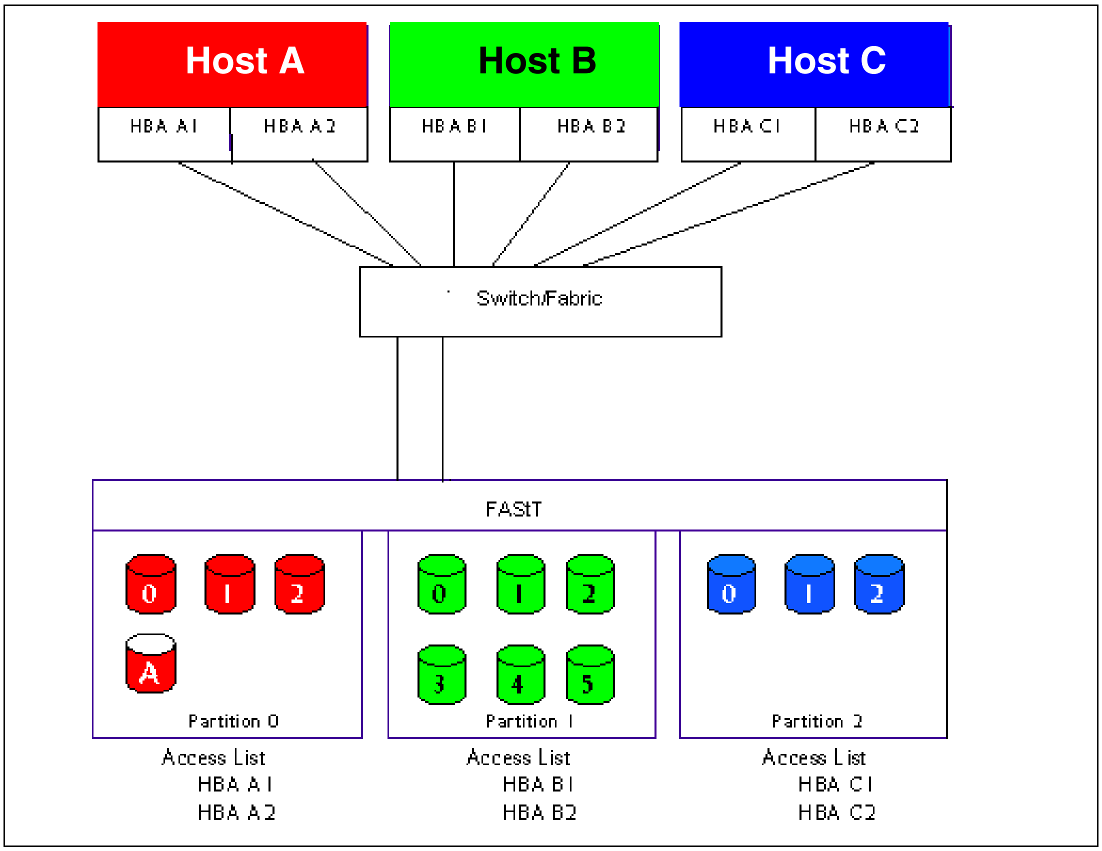


Figure C-1 Initial configuration

Then we add the SVC and create an SVC partition on the FAStT. See Figure C-2. The following steps are required for this task:

1. Modify the zoning so that SVC can “see” the FAStT *only*. This allows partition 3 to be created, and access to partitions 0, 1, and 2 can continue.
2. The port or ports of the SVC master console must not be in any partition.

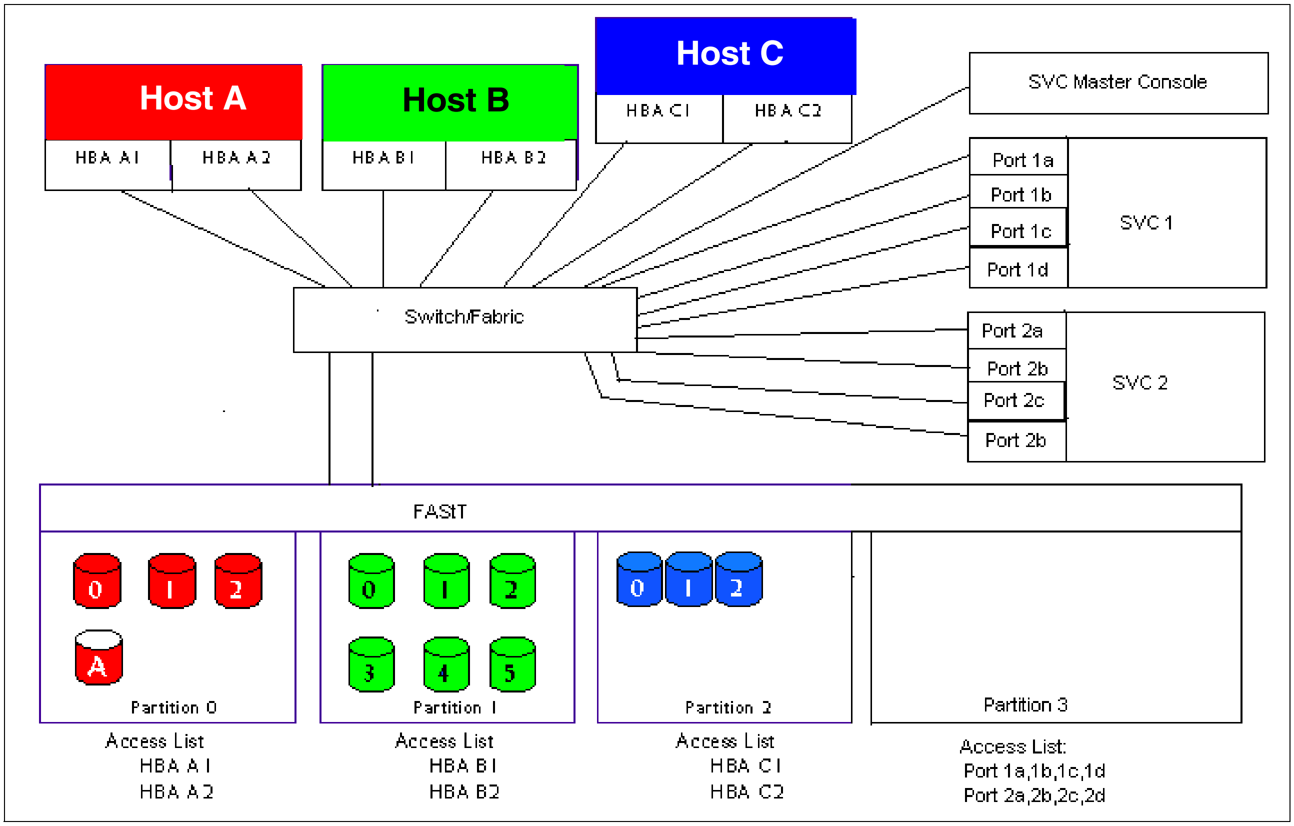


Figure C-2 Partition 3 created

We move the storage for host C from host partition 2 to partition 3 (SVC partition), to be managed by the SVC. See Figure C-3. Note the following points:

- Concurrent access from host C to its logical units is not possible.
- Host C requires reconfiguration from FAStT device drivers to SDD. Changes to the adapter configuration and microcode levels, settings, etc. may also be required.
- Switch zoning changes are also required to prevent host C from “seeing” the FAStT ports and instead “seeing” the SVC ports.
- Logical units from host partition 2, that are now in partition 3, must be configured as image mode mDisks in SVC and mapped to host C.

Note: Partition 2 should now be deleted after all logical units are moved to partition 3.

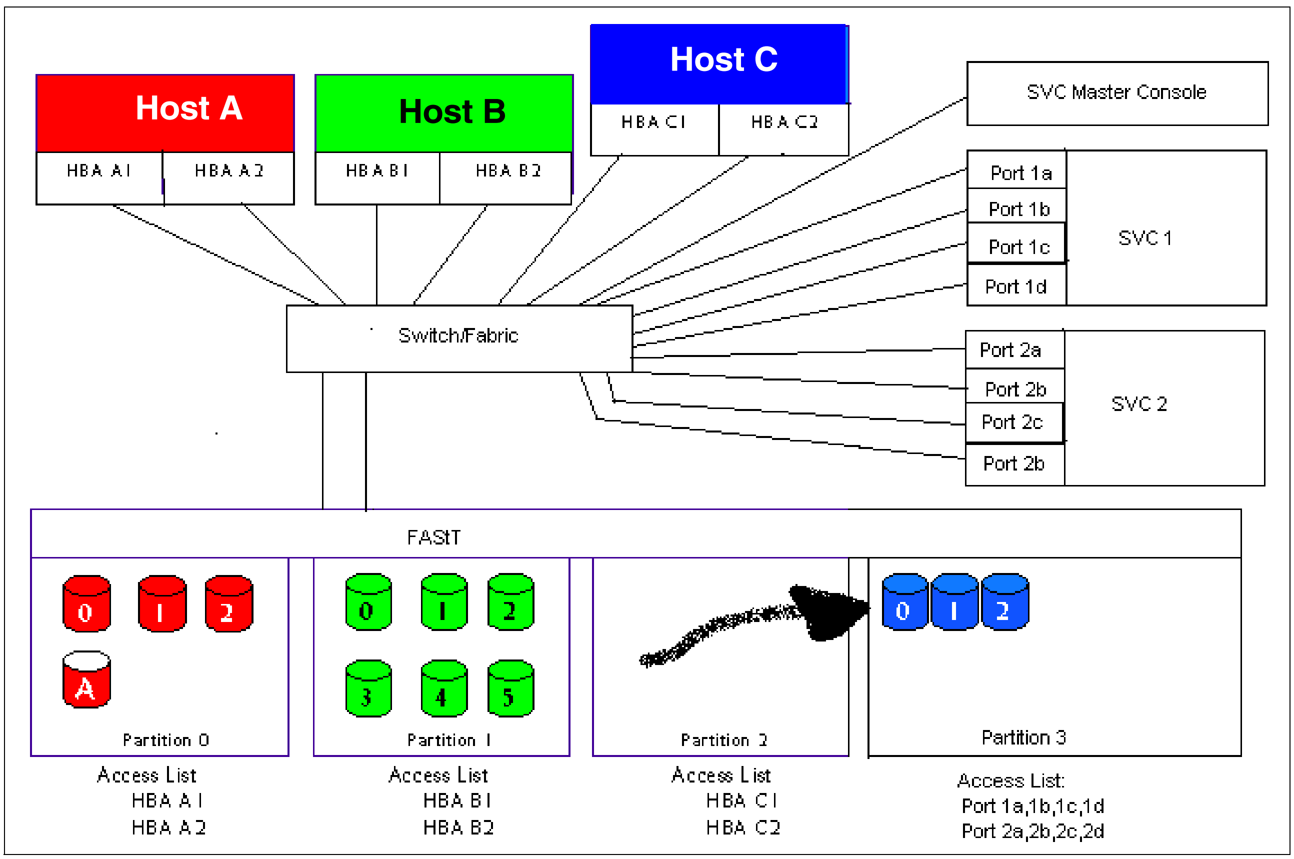


Figure C-3 Storage moved from partition 2 to partition 3

Then we move the storage for host B from host partition 1 to partition 3 (SVC partition), to be managed by the SVC. This is shown in Figure C-4.

The following steps are required to do this:

1. Stop access from host B to its logical units.
2. Host B requires reconfiguration from FAStT device drivers to SDD, changes to adapter configuration and microcode levels, settings, etc.
3. Switch zoning changes are also required to prevent host B from “seeing” the FAStT ports and instead of “seeing” the SVC ports.
4. Logical units from host partition 1, which are now in partition 3, must be configured as image mode mDisks in SVC and mapped to host C, using their original logical unit numbers.
5. Partition 3 can be deleted if required, after all logical units are moved to partition. Note that logical Units moved from partition 1 to partition 3 have different logical unit numbers.

> **Note:** Access LUN (A) can no longer be used by host A for inband configuration of FAStT. This can deleted and FAStT configured to use SVC master console over the Ethernet. The access LUN is not required by SVC.

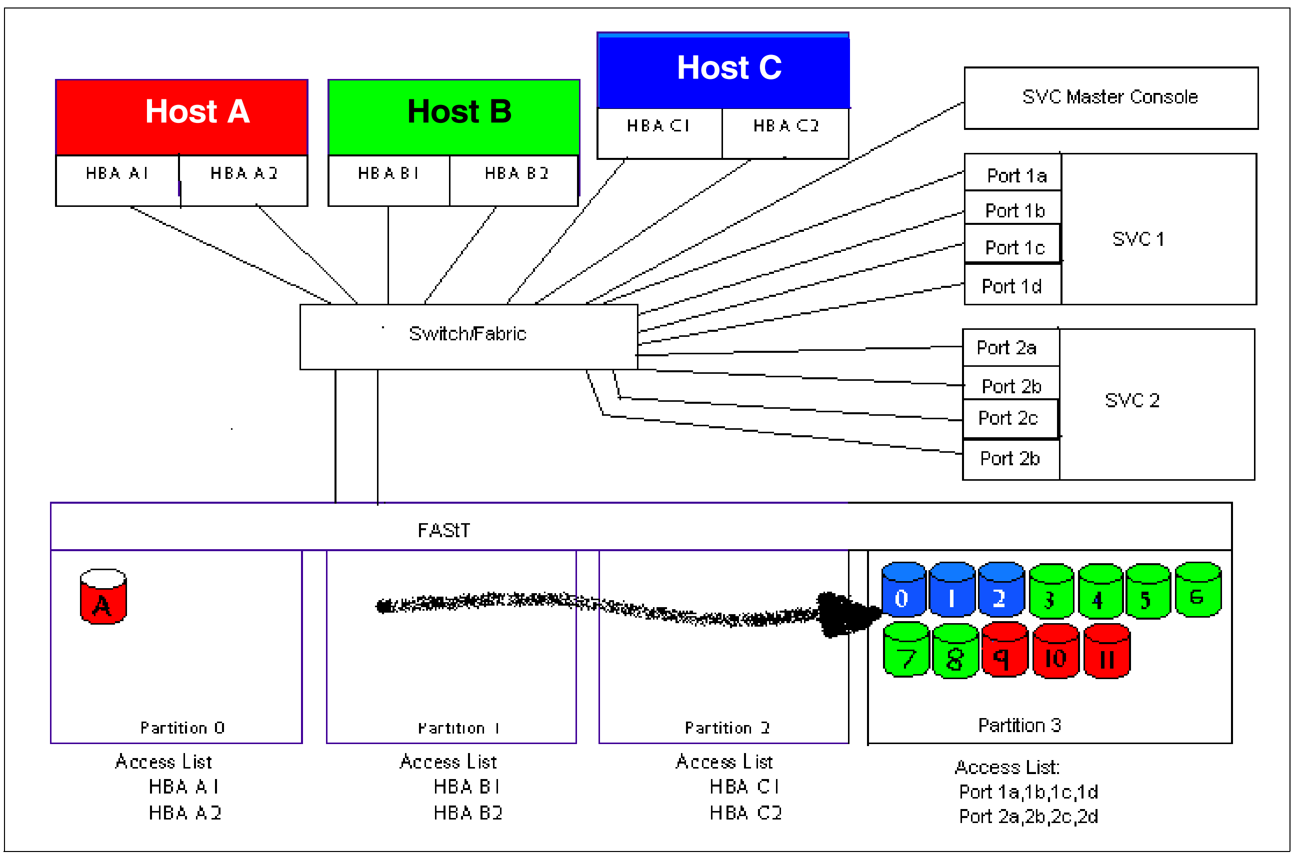


Figure C-4 Storage moved from partition 1 to partition 3

We must now move any remaining host storage moved from host partitions to partition 3 (SVC partition). We use the previous steps to accomplish this. This gives us the configuration shown in Figure C-5. Image mode mDisks can now be converted to managed mDisks using data migration commands as required.

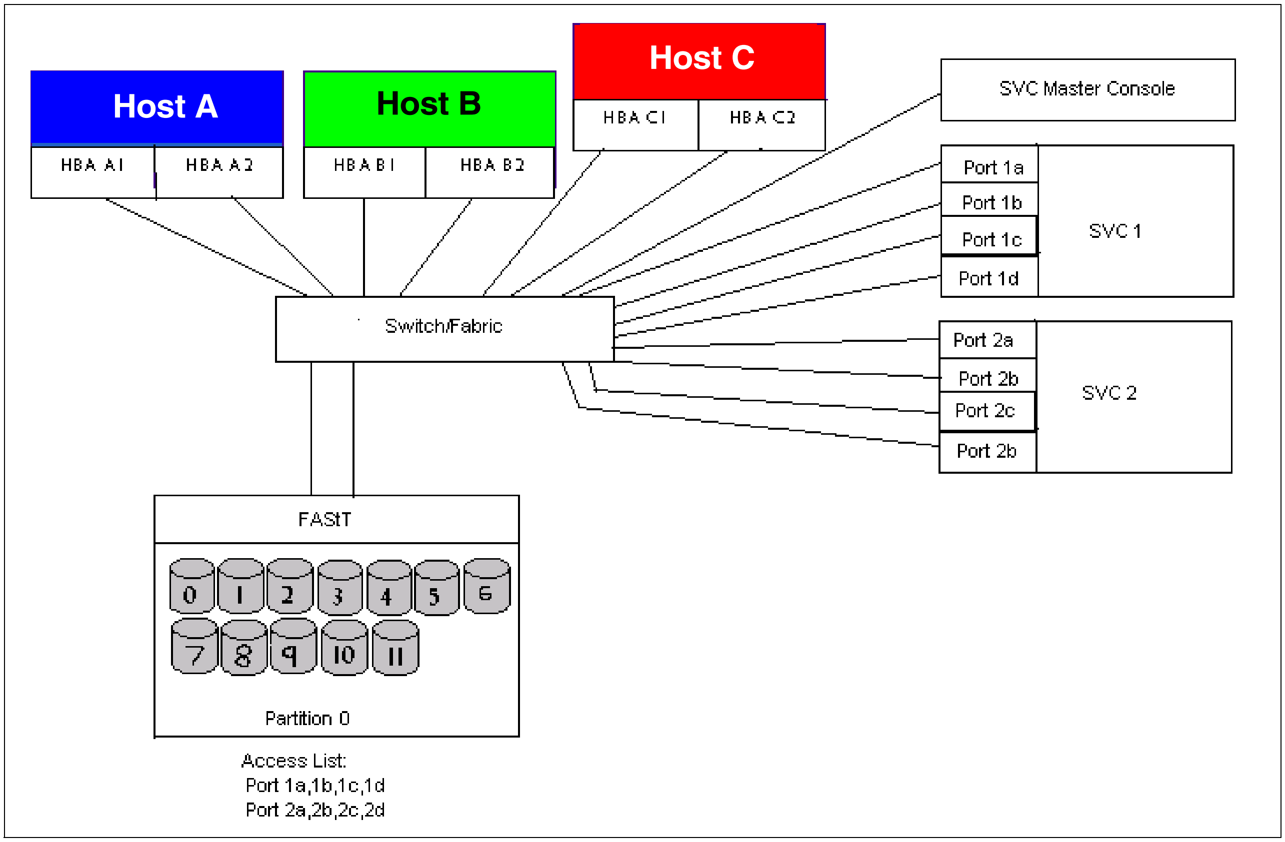


Figure C-5 All storage under the SVC

Scenario 2: Migrating storage (more than 32 LUNs in all partitions)

The initial configuration is shown in Figure C-6. Note the following points:

- More than 32 LUNs on FAStT 1
- New FAStT providing new storage, larger than or equal to the capacity of FAStT 1
- Only one partition per host

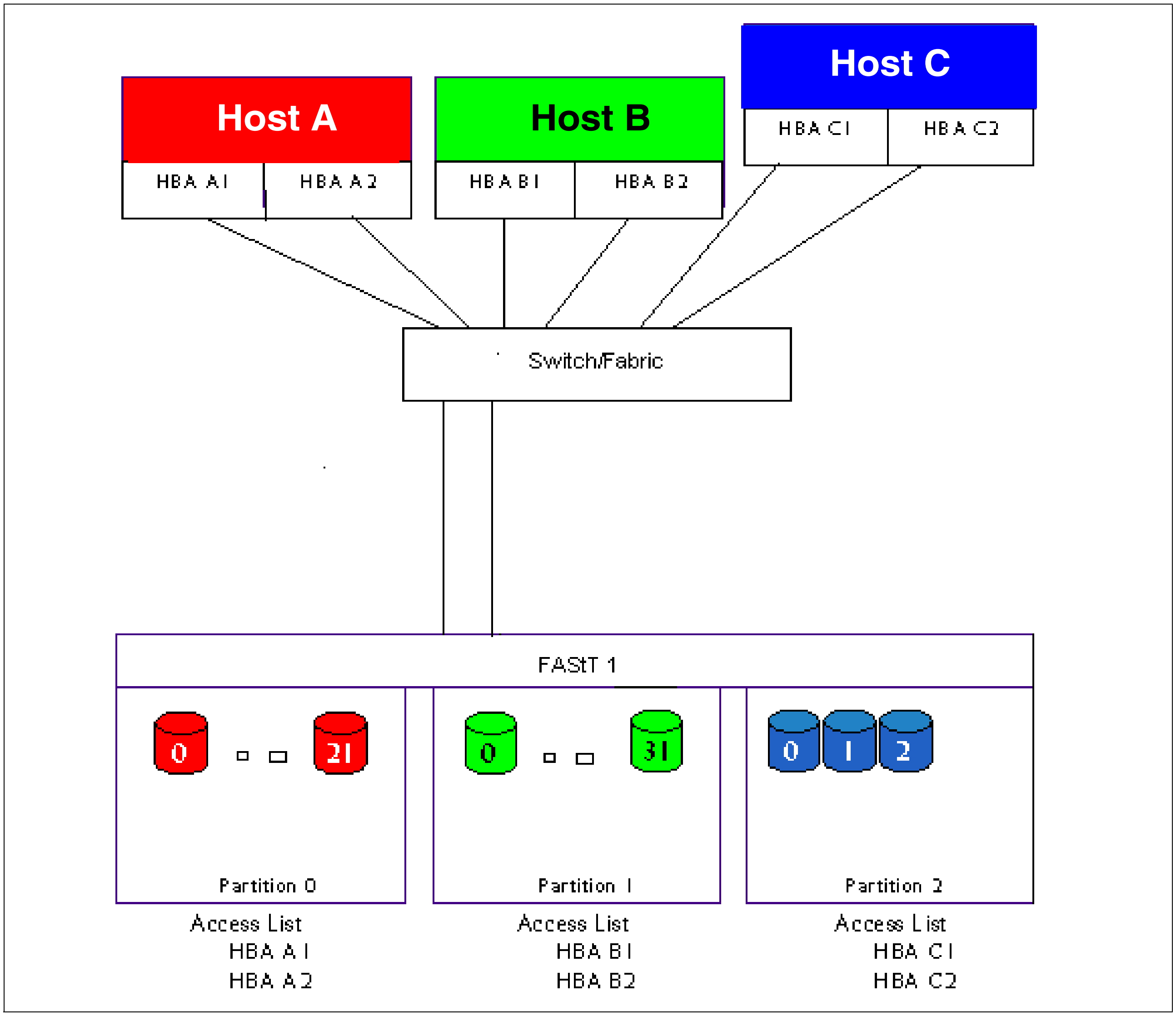


Figure C-6 Scenario 2 initial configuration

We then add another FAStT and carry out the following steps:

1. Create RAID arrays on FAStT 2, one LUN per array.
2. Rezone the switch to allow SVC ports to access FAStT 2 ports.
3. Create the partition, including all LUNs and SVC ports.

This is shown in Figure C-7.

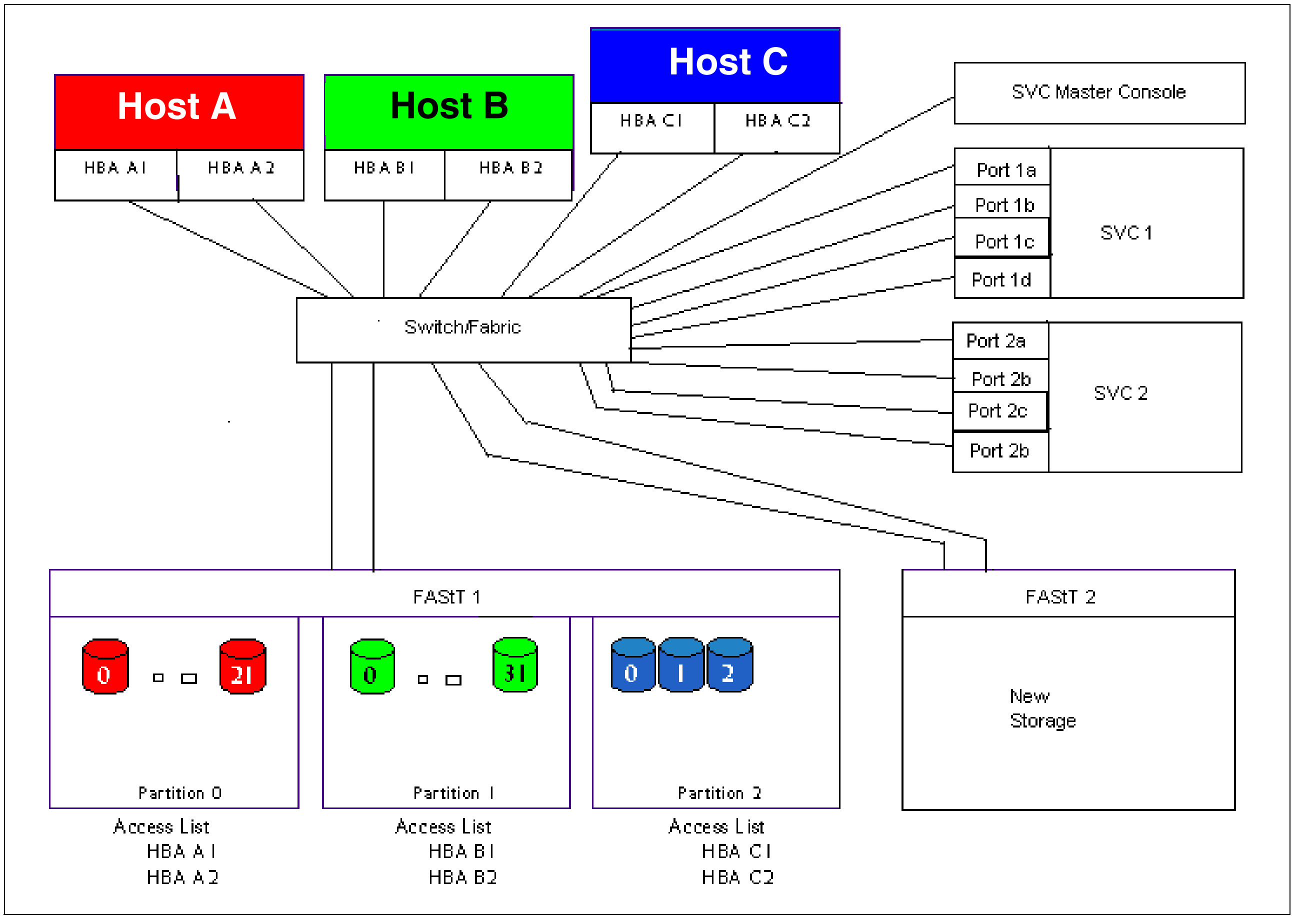


Figure C-7 Second FAStT added

We then move the storage for host C under the control of the SVC. This is shown in Figure C-8.

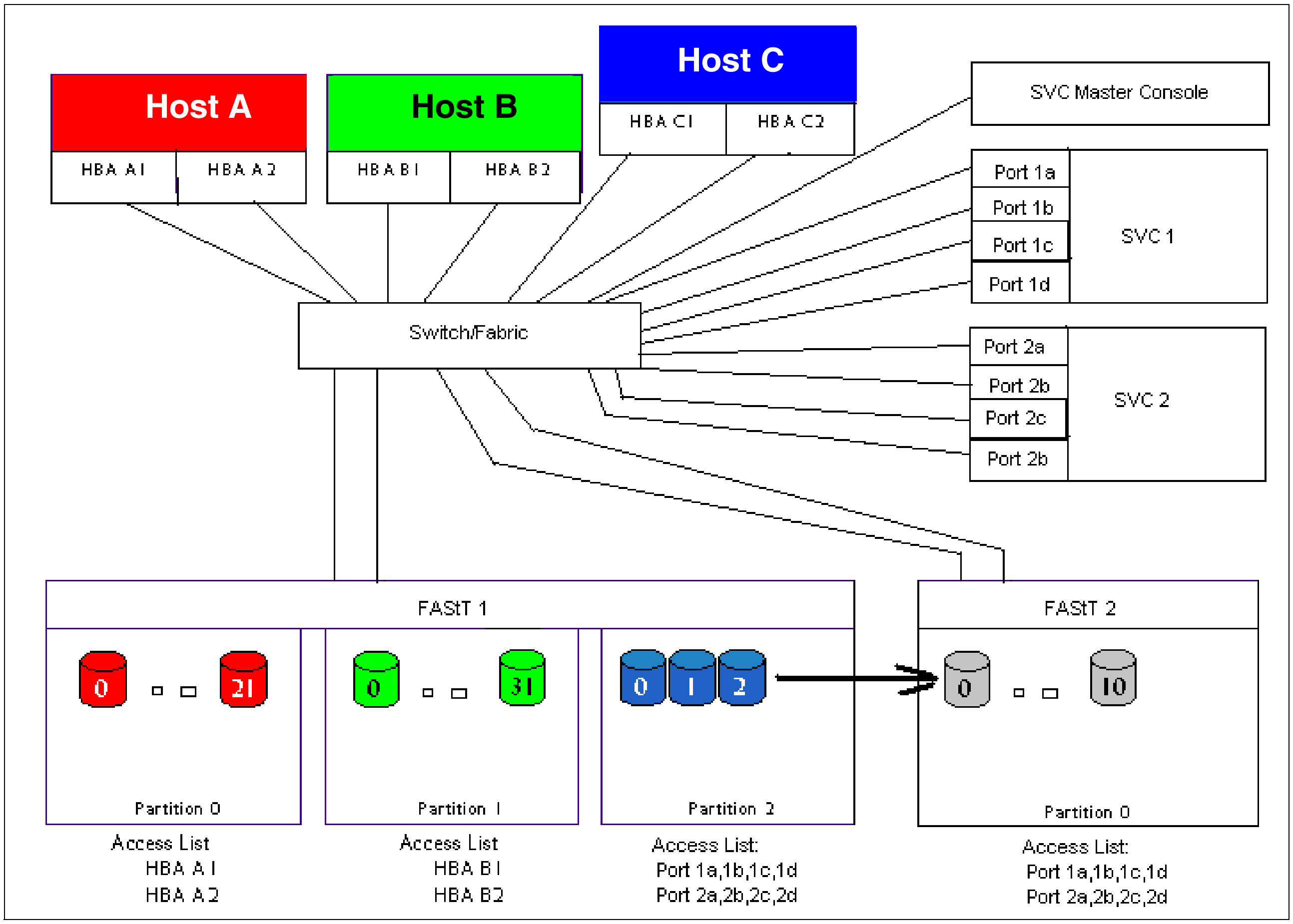


Figure C-8 Portions created on FAStT2

The following steps are carried out:

1. Stop host C.
2. Rezone the switch so that host C port accesses the SVC ports as required, *not* the FAStT 1 ports.
3. Rezone the switch to allow the SVC ports to access FAStT 1 ports.
4. Change host C device drivers, settings, software etc., to support the SVC.
5. Change partition 2 to SVC *host type* and change port names to SVC ports removing ports of host C.
6. Create SVC managed mode disks from storage in partition 0 on FAStT 2.
7. Create SVC image mode disks from storage in partition 2 on FAStT 1.
8. Migrate image mode vDisks for host C to managed disks on FAStT 2.
9. When migration completes, delete LUNs and partition 2 on FAStT 1.

Figure C-9 shows the result of this.

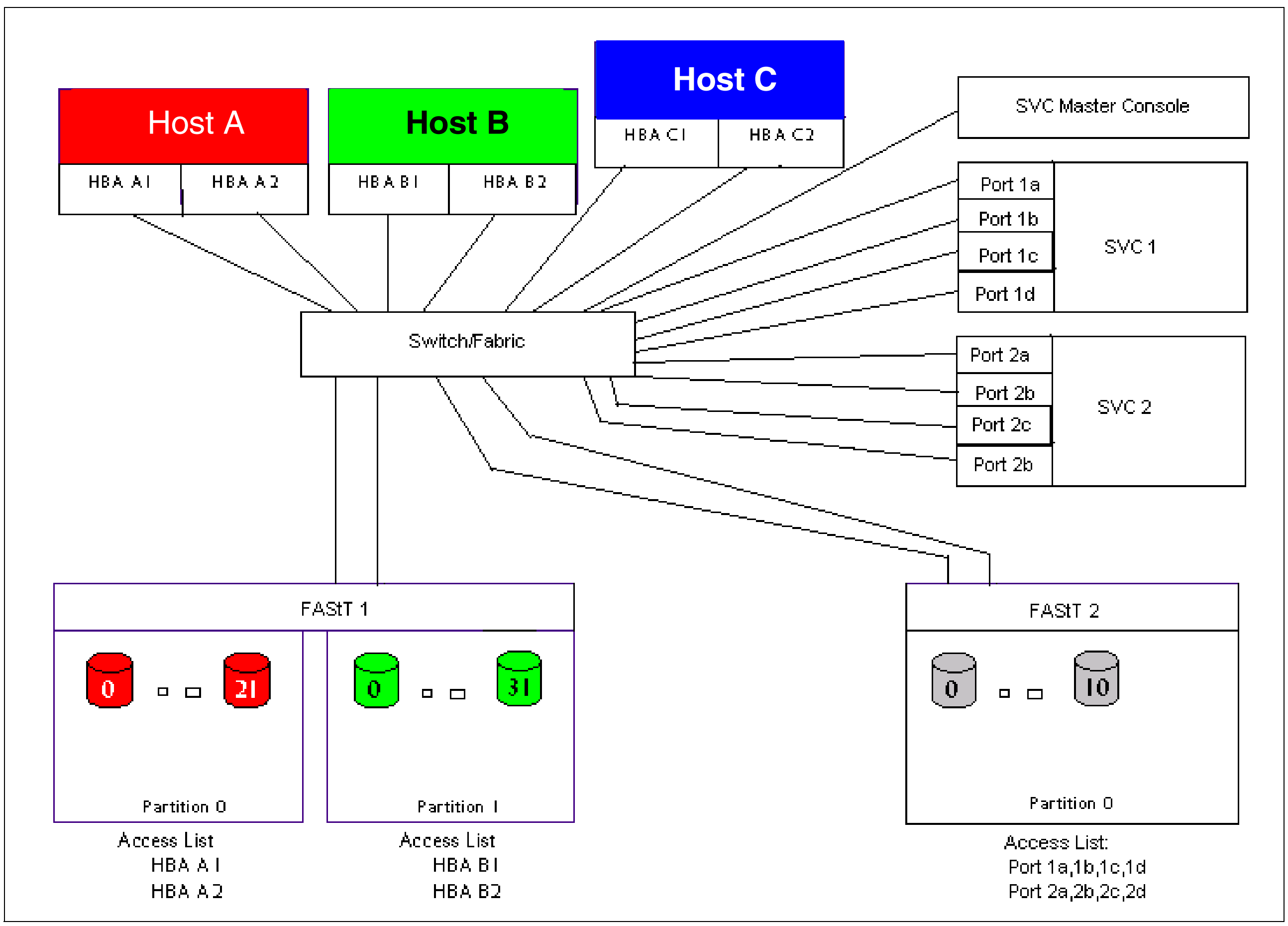


Figure C-9 Storage for host C migrated to FAStT 2

The procedure is repeated for the remaining host until all the storage is migrated to the control of the SVC. This is shown in Figure C-10. FAStT1 is now unused.

> **Note:** Although we used a second FAStT in this scenario, it is possible to carry out a similar procedure if there is enough spare capacity on FAStT 1.

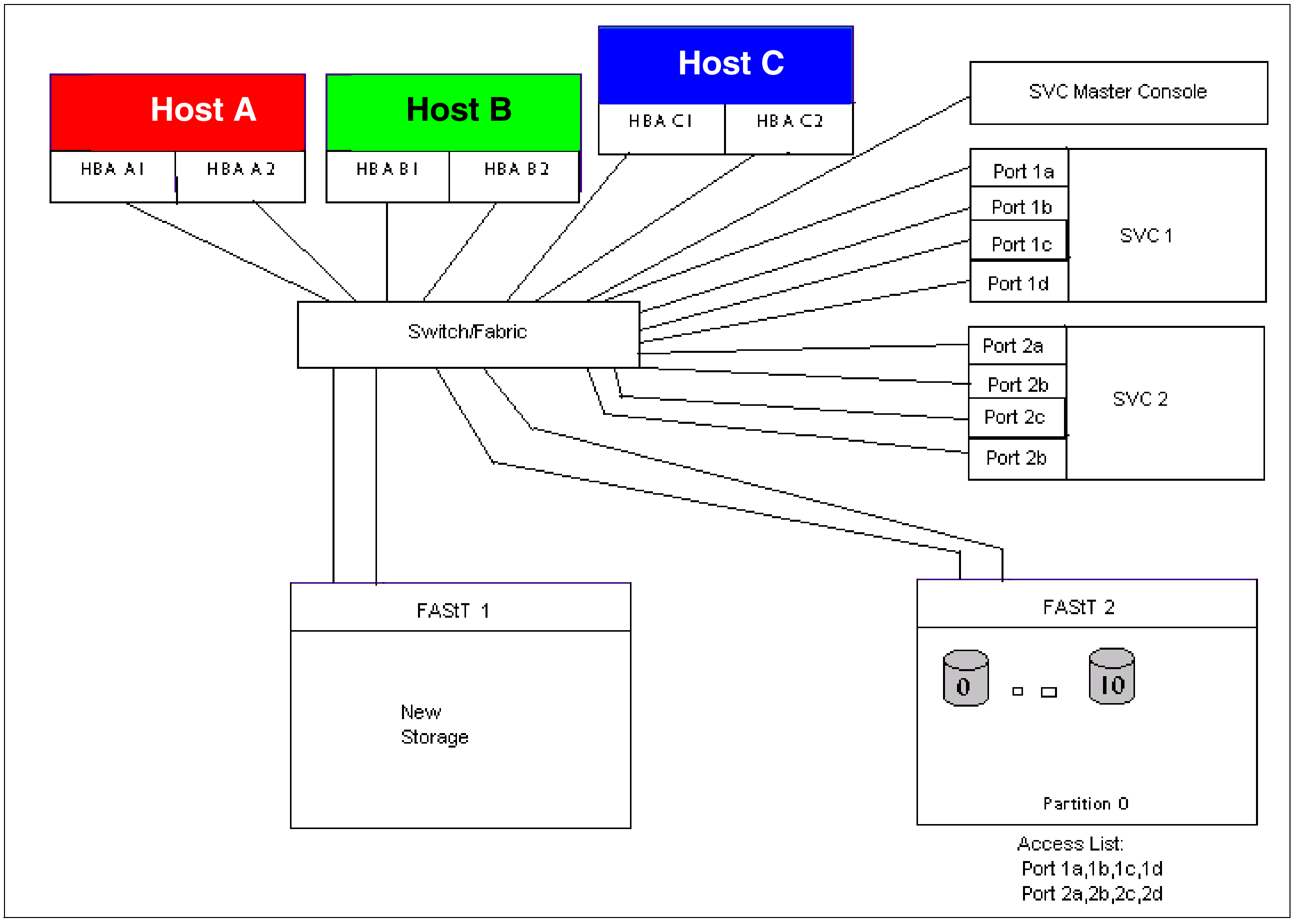


Figure C-10 All storage under control of the SVC

D

Standby node and cloning

If an IBM TotalStorage SAN Volume Controller (SVC) node fails, the SVC cluster continues to operate with degraded performance until the failed node is repaired. If the repair operation is likely to take an unacceptably long time, it may be useful to replace the failed node with a *spare*.

However, to replace a failed node without interrupting I/O and without any risk to data integrity when a repaired node is reconnected to the storage area network (SAN) fabric, you must follow various procedures and take precautions. The procedure involves changing the worldwide node name (WWNN) of a node. You must follow this procedure with care since duplicate WWNNs can cause data corruption.

This appendix describes the basic tasks involved in setting up a node as a standby node and how to clone a node, particularly a failed node in the cluster.

Attention: The advantage of cloning a node as opposed to simply adding a new node is that, by replicating the failed node's WWNN, we make the new node's worldwide port names (WWPNs) the same as the failed node. In this manner, all logical unit number (LUN) maskings, persistent bindings, and zonings that are setup in the storage, hosts, and switches attached to the cluster can remain unchanged.

Prerequisites

Before you attempt to replace a failed node, you must comply with these prerequisites:

1. Have SVC software Version 1.1.1 or higher installed on the SVC cluster and on the spare node.
2. Know the name of the cluster that contains the failed node.
3. Have a spare node in the same rack as the SVC cluster that contains the failed node.
4. Make a record of the last five characters of the original WWNN of the spare node since this may be needed again if you want to stop using a spare node and use this node as a normal node that can be assigned to any cluster. Complete the following steps to display the WWNN of the node:
   a. Display the node status on the front panel display of the node.
   b. With the node status displayed on the front panel, press and hold the **Down** button. Press and release the **Select** button. Release the **Down** button.
   c. The text "WWNN" is displayed on line one of the display. Line 2 of the display contains the last five characters of the WWNN.
   d. Record this number in a safe place. It is needed if you want to stop using a spare node.

Replacing a failed SVC node

When a node is swapped, the following actions may occur:

- *Front Panel ID* can be changed. This is the number that is printed on the front of the node and used to select the node that is to be added to a cluster.
- *Node Name* can be changed. If you permit the SVC application to assign default names when adding nodes to the cluster, it creates a new name each time a node is added. If you choose to assign your own names, then you need to type in the node name that you want to use. If you are using scripts to perform management tasks on the cluster and those scripts use the node name, then, by assigning the original name to a replacement node, you avoid the need to make changes to the scripts following service activity on the cluster.
- *Node ID* changes. A new node ID is assigned each time a node is added to a cluster. The node ID or the node name can be used when performing management tasks on the cluster. If scripts are being used to perform those tasks, we recommend that you use the node name instead of the node ID. This is because the node name remains unchanged following service activity on the cluster.
- *World Wide Node Name* does not change. The WWNN is used to uniquely identify the node and the Fibre Channel ports. The node replacement procedure changes the WWNN of the spare node to match that of the failed node. The node replacement procedures must be followed exactly to avoid any duplication of WWNNs.
- The *World Wide Port Name* of each Fibre Channel port does not change. The WWPNs are derived from the WWNN that is written to the replacement node as part of this procedure.

Complete these steps:

1. Use the SVC console or the command line interface (CLI) to gather and record the following information about the failed node:

   – **Node name**: To display the node name using the console:

     i. From the Welcome window, select **Work with Nodes -> Nodes**.
     ii. The failed node is offline. Note the name.

     To display the node name using the CLI:

     i. Use the command:

        ```
        svcinfo lsnode
        ```

     ii. The failed node is offline. Note the name.

   – **I/O group name**: To display the I/O group name using the console:

     i. From the Welcome window, select **Work with Nodes -> Nodes**.
     ii. The failed node is offline. Note the I/O group name.

     To display the node name using the CLI:

     i. Use the command:

        ```
        svcinfo lsnode
        ```

     ii. The failed node is offline. Note the I/O_group_name.

   – **The last five characters of the WWNN**: To display the WWNN using the console:

     i. From the Welcome window, select **Work with Nodes -> Nodes**.
     ii. The failed node is offline. Note the last five characters of the WWNN.

     To display the WWNN using the CLI:

     i. Use the command:

        ```
        svcinfo lsnodevpd <node_name>
        ```

        Here <*node_name*> is the name recorded in Step 1.

     ii. Find the WWNN field in the output. Note the last five characters of the WWNN.

   – **Front panel ID**: To display the front panel id using the console:

     i. From the Welcome window, select **Work with Nodes -> Nodes**.
     ii. The failed node is offline. Click the name of the offline node.
     iii. Select the **Vital Product Data** tab.
     iv. The front panel ID is under the Front panel assembly section of the VPD. Note the front panel ID.

     To display the front panel ID using the CLI:

     i. Use the command:

        ```
        svcinfo lsnodevpd <node_name>
        ```

        Here <*node_name*> is the name recorded in the step above.

     ii. Find the front_panel_id field in the output. Note the front panel ID.

   – **The uninterruptible power supply serial number**: To display the uninterruptible power supply serial number using the console:

     i. From the Welcome window, select **Work with Nodes -> Nodes**.
     ii. The failed node is offline. Click the name of the offline node.
     iii. Select the **Vital Product Data** tab.
     iv. The uninterruptible power supply serial number is in the **UPS** section of the VPD. Note the serial number.

To display the node VPD using the CLI:

i. Enter the command:

```
svcinfo lsnodevpd <node_name>
```

Here <*node_name*> is the name recorded in Step 1.

ii. Find the UPS_serial_number field in the output. Note the uninterruptible power supply serial number.

2. Use the front panel ID to locate the failed node. Disconnect all four Fibre Channel cables from the node.

Important: The cables must not be reconnected until the failed node is replaced and the WWNN of the spare node is changed to match the WWNN of the failed node.

3. Connect the power or signal cable from the spare node to the uninterruptible power supply with the serial number noted in Step 1.

Note: The signal cable can be plugged into any vacant position on the top row of the serial connectors on the uninterruptible power supply. If no spare serial connectors are available on the uninterruptible power supply, disconnect the cables from the failed node.

4. Power on the spare node.
5. Display the node status on the service panel.
6. Complete the following steps to change the WWNN of the spare node:

   a. With the node status displayed on the front panel, press and hold the **Down** button. Press and release the **Select** button. Release the **Down** button.

      The text "WWNN" is displayed on line one of the display. Line two of the display contains the last five characters of the WWNN.

   b. With the WWNN displayed on the service panel, press and hold the **Down** button. Press and release the **Select** button. Release the **Down** button. This switches the display into edit mode.

   c. Change the displayed number to match the WWNN recorded in Step 1. To edit the displayed number, use the Up and Down buttons to increase or decrease the numbers displayed. Use the left and right buttons to move between fields. When the five characters match the number recorded in Step 1, press the **Select** button twice to accept the number.

7. Connect the four Fibre Channel cables that were disconnected from the failed node to the spare node.
8. Depending on how the original node failed, the replacement node may or may not automatically join the cluster. If the node does not rejoin the cluster, delete the offline node on the master console. Then, add the spare node into the cluster on the master console.
9. Use the Subsystem Device Driver (SDD) management tool on the host systems to verify that all paths are now online.

When the failed node is repaired, do not connect the Fibre Channel cables to it. Connecting the cables may cause data corruption. When the node repair returns the node to an operational state, perform the following steps:

1. Display theW node status on the service panel.
2. With the SVC status displayed on the front panel, press and hold the **Down** button. Press and release the **Select** button. Release the **Down** button.

   The text "WWNN" is displayed on line one of the display. Line two of the display contains the last five characters of the WWNN.
3. With the WWNN displayed on the service panel, press and hold the **Down** button, press and release the **Select** button, and release the **Down** button. This switches the display into edit mode.
4. Change the displayed number to `00000`. To edit the displayed number, use the Up and Down buttons to increase or decrease the numbers displayed. Use the left and right buttons to move between fields. When the number is set to 00000, press the **Select** button twice to accept the number.

You can now use this SVC as a spare node. If this SVC is no longer required as a spare and is to be used for normal attachment to a cluster, you must first use the procedure described previously to change the WWNN to the number saved when a spare was being created. See “Prerequisites” on page 524. Using any other number may cause data corruption.

Attention: Never connect a node with a WWNN of "00000" to the cluster.

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this redbook.

IBM Redbooks

For information about ordering these publications, see “How to get IBM Redbooks” on page 530.

- *IBM SAN Survival Guide*, SG24-6143
- *Introducing Hosts to the SAN Fabric*, SG24-6411
- *IBM Tivoli Storage Area Network Manager: A Practical Introduction*, SG24-6848
- *IBM TotalStorage: Introducing the SAN File System*, SG24-7057
- *Implementing the IBM TotalStorage SAN Volume Controller Storage Software on the Cisco MDS 9000*, SG24-7059
- *Virtualization in a SAN*, REDP3633

Other resources

These publications are also relevant as further information sources:

- *IBM TotalStorage Virtualization Family SAN Volume Controller: Planning Guide*, GA22-1052
- *Subsystem Device Driver User's Guide for the IBM TotalStorage Enterprise Storage Server and the IBM TotalStorage SAN Volume Controller*, SC26-7540
- *IBM TotalStorage Virtualization Family SAN Volume Controller: Installation Guide*, SC26-7541
- *IBM TotalStorage Virtualization Family SAN Volume Controller: Service Guide*, SC26-7542
- *IBM TotalStorage Virtualization Family SAN Volume Controller: Configuration Guide*, SC26-7543
- *IBM TotalStorage Virtualization Family SAN Volume Controller: Command-Line Interface User's Guide*, SC26-7544
- *IBM TotalStorage 2146 SAN Integration Server Model V1A User's Guide*, SC26-7556
- *IBM TotalStorage SAN Integration Server 2146 Installation & Service Guide*, SC26-7557
- *IBM TotalStorage Virtualization Family SAN Volume Controller: Host Attachment Guide*, SC26-7563
- *IBM TotalStorage ESS, SAN Volume Controller, SAN Volume Controller for Cisco*, SC26-7608
- *IBM TotalStorage Virtualization Family SAN Volume Controller: Service Guide Supplement for SAN Volume Controller V1.1.1*, P/N 64P7990

Referenced Web sites

These Web sites are also relevant as further information sources:

- IBM TotalStorage home page

  http://www.storage.ibm.com
- San Volume Controller supported platform

  http://www.storage.ibm.com/support/2145
- Download site for Windows SSH freeware

  http://www.chiark.greenend.org.uk/~sgtatham/putty
- IBM site to download SSH for AIX

  http://oss.software.ibm.com/developerworks/projects/openssh
- Open source site for SSH for Windows and Mac

  http://www.openssh.com/windows.html
- Cygwin Linux-like environment for Windows

  http://www.cygwin.com
- IBM Tivoli Storage Area Network Manager site

  http://www-3.ibm.com/software/sysmgmt/products/support/IBMTivoliStorageAreaNetworkManager.html
- Microsoft Knowledge Base Article 131658

  http://support.microsoft.com/support/kb/articles/Q131/6/58.asp
- Microsoft Knowledge Base Article 149927

  http://support.microsoft.com/support/kb/articles/Q149/9/27.asp
- Sysinternals home page

  http://www.sysinternals.com
- Subsystem Device Driver download site

  http://www-1.ibm.com/servers/storage/support/software/sdd.html
- IBM TotalStorage SAN Integration Server support site:

  http://www.ibm.com/storage/support/2146

How to get IBM Redbooks

You can order hardcopy Redbooks, as well as view, download, or search for Redbooks at the following Web site:

ibm.com/redbooks

You can also download additional materials (code samples or diskette/CD-ROM images) from that site.

IBM Redbooks collections

Redbooks are also available on CD-ROMs. Click the CD-ROMs button on the Redbooks Web site for information about all the CD-ROMs offered, as well as updates and formats.

Index

D

E

F

J

L

M

N

T

U

V

W

Z

IBM TotalStorage SAN Volume Controller and SAN Integration Server

(1.0” spine)
0.875”<->1.498”
460 <-> 788 pages

IBM TotalStorage SAN Volume Controller and SAN Integration Server

IBM®

Redbooks

Install, use, and troubleshoot SAN Volume Controller and SAN Integration Server

Learn how to implement block virtualization

Perform backup and restore on a cluster

This IBM Redbook is a detailed technical guide to the IBM TotalStorage® SAN Volume Controller (SVC) and IBM TotalStorage SAN Integration Server (SIS). It covers the November 2003 enhancements. Although much of the discussion in this redbook is based upon the SVC, it equally applies to the SIS. The SIS is a specially packaged SVC with the addition of storage, storage area network (SAN) switches, and an Ethernet switch.

The book covers the following areas:

- Storage virtualization: This is only a high-level overview since this topic is described in detail in other Redbooks.
- Architecture of the SVC and SIS
- Implementing and configuring the SVC and SIS
- Using virtualization and advanced functions such as FlashCopy® and Remote Copy
- Migrating existing storage to the SVC and SIS
- Backup and recovery

You can now combine the virtualization functions of the SAN Volume Controller storage software with the Cisco MDS 9000 Series Caching Services Module from the Cisco MDS 9000 Family. The virtualization functions of this new product are similar to those described in this book. For a full discussion of this new product, read the IBM Redbook, *Implementing the IBM TotalStorage SAN Volume Controller Storage Software on the Cisco MDS 9000*, SG24-7059.

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

SG24-6423-01 ISBN 0738498661