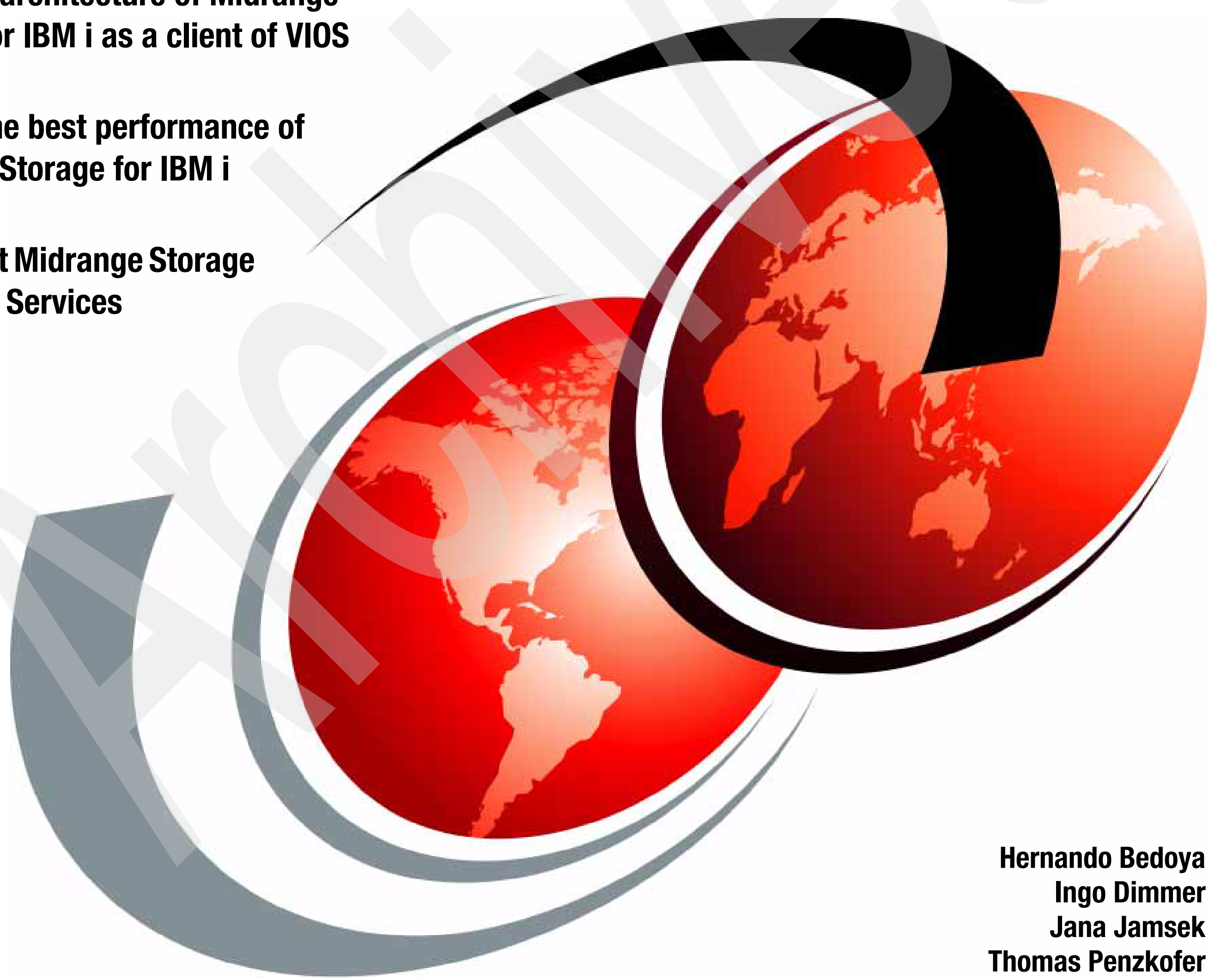
IBM i and Midrange External Storage

Learn the architecture of Midrange Storage for IBM i as a client of VIOS

Plan for the best performance of Midrange Storage for IBM i

Implement Midrange Storage with Copy Services



Hernando Bedoya
Ingo Dimmer
Jana Jamsek
Thomas Penzkofer

Redbooks

ibm.com/redbooks

International Technical Support Organization

IBM i and Midrange External Storage

January 2009

SG24-7668-00

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

First Edition (January 2009)

This edition applies to IBM i 6.1.

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

Trademarks

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

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

Active Memory™
AIX®
Alerts®
AS/400®
BladeCenter®
DB2®
DS4000™
DS6000™
DS8000™
eServer™
FlashCopy®
i5/OS®
IBM®
iSeries®
Lotus®
NetServer™
POWER™
Power Systems™
POWER5™
POWER6™
PowerVM™
Redbooks®
Redbooks (logo) ®
System i®
System p®
System Storage™
System Storage DS®
System z®
Tivoli®
Virtualization Engine™
xSeries®

The following terms are trademarks of other companies:

InfiniBand, and the InfiniBand design marks are trademarks and/or service marks of the InfiniBand Trade Association.

Disk Magic, IntelliMagic, and the IntelliMagic logo are trademarks of IntelliMagic BV in the United States, other countries, or both.

Snapshot, and the NetApp logo are trademarks or registered trademarks of NetApp, Inc. in the U.S. and other countries.

Novell, SUSE, the Novell logo, and the N logo are registered trademarks of Novell, Inc. in the United States and other countries.

VMware, the VMware "boxes" logo and design are registered trademarks or trademarks of VMware, Inc. in the United States and/or other jurisdictions.

RSM, Solaris, Sun, and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

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

Intel Xeon, Intel, Intel logo, Intel Inside logo, and Intel Centrino logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States, other countries, or both.

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

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

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

Preface

This IBM® Redbooks® publication describes the new IBM i Midrange External Storage solutions available for IBM POWER™ Systems POWER6™ servers with IBM i being a client of IBM Virtual I/O Server (VIOS). It introduces the VIOS virtualization concepts and IBM DS Midrange External Storage Systems architecture of the supported models DS3400, DS4700, DS4800, and DS5000, discusses planning and sizing for IBM i Midrange External Storage, and provides detailed implementation procedures including IBM DS Midrange Storage Copy Services. Finally, it provides monitoring, maintenance, and troubleshooting hints for the triumvirate of IBM i, VIOS, and IBM DS Midrange External Storage.

The information provided by this book will help customers, business partners, and IBM service professionals with planning and implementing IBM i Midrange External Storage solutions.

The team that wrote this book

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

Hernando Bedoya is an IT Specialist at the IBM ITSO in Rochester, Minnesota. He writes extensively and teaches IBM classes worldwide in all areas of DB2® for i5/OS®. Before joining the ITSO more than seven years ago, he worked for IBM Colombia as an IBM AS/400® IT Specialist doing pre-sales support for the Andean countries. He has 24 years of experience in the computing field and has taught database classes in Colombian universities. He holds a master's degree in computer science from EAFIT, Colombia. His areas of expertise are database technology, application development, and high availability.

Ingo Dimmer is an IBM Consulting IT Specialist for System i® and a PMI Project Management Professional working in the IBM STG Europe Storage Support Organization in Mainz, Germany. He has nine years of experience in enterprise storage support from working in IBM post-sales and pre-sales support. He holds a degree in electrical engineering from the Gerhard-Mercator University Duisburg. His areas of expertise include System i external disk and tape storage solutions, I/O performance, and tape encryption, for which he has been an author of several white paper and IBM Redbooks publications.

Jana Jamsek is an IT Specialist in IBM Slovenia. She works in Storage Advanced Technical Support for Europe as a specialist for IBM Storage Systems and i5/OS systems. Jana has eight years of experience in the System i and AS/400 areas, and six years of experience in storage. She has a master's degree in computer science and a degree in mathematics from the University of Ljubljana, Slovenia. She was among the authors of IBM Redpaper publication, *The LTO Ultrium Primer for IBM eServer iSeries Customers*, REDP-3580, and the IBM Redbooks publications *iSeries in Storage Area Networks*, *iSeries and IBM TotalStorageIBM i and IBM System Storage: A Guide to Implementing External Disks on IBM i*, SG24-7120; *IBM Virtualization EngineThe IBM Virtualization Engine TS7510: Getting Started with i5/OS and Backup Recovery and Media Services*, SG24-7510; *Implementing IBM Tape in i5/OS*, SG24-7440; *IBM System Storage DS8000: Copy Services in Open Environments*, SG24-6788; and *IBM System Storage DS6000 Series: Architecture and Implementation*, SG24-6781.

Thomas Penzkofer is a Product Field Engineer for IBM DS3000, DS4000™, and DS5000 Storage Servers. He works at the European Storage Competence Center in Germany, where he is doing pre-sales and post-sales support for IBM storage products. He has three years of PFE experience with IBM DS4000 power supply fan units and 10 years of experience with IBM storage products. He has a master's degree in electrical engineering from the University of Kaiserslautern, Germany.

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

Linda Robinson
Thomas Gray
Joanna Pohl-Miszczyk
Jenifer Servais
International Technical Support Organization, Rochester Center

Kris Whitney
Vess Natchev
Don Pischke
John Nordman
Ron Zacher
Vess Natchev
Ginny McCright
Joe Writz
Russ VanDuine
Jeff Palm
Donald Pischke
IBM Rochester

Bob Kovacs
Jim Pafumi
Steven Tee
George White
Jim Partridge

John Neemidge
Raj Patel
Omar Cardona
Sergio Reyes
IBM Austin

Siebo Friesenborg
David Navarro
Lee La Frese
Alex Osuna
Michael D Roll
Joshua Martin
IBM Tucson

Andy McNeill
IBM Raleigh

Bernd Baeuml
Mousa Hammad
IBM Mainz, Germany

A special thanks to the European Storage Competence Center in Mainz, Germany, for providing the required hardware for this project. For more information about ESCC see:

http://web.mainz.de.ibm.com/ESCC

Become a published author

Join us for a two- to six-week residency program! Help write a book dealing with specific products or solutions, while getting hands-on experience with leading-edge technologies. You will have the opportunity to team with IBM technical professionals, Business Partners, and Clients.

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

To learn more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

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

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

  ibm.com/redbooks

- Send your comments in an e-mail to:

  redbooks@us.ibm.com

- Mail your comments to:

  IBM Corporation, International Technical Support Organization
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400

1

Introduction

This chapter provides an overview of the new concepts introduced with IBM i POWER6 and the new operating system version IBM i 6.1 for using IBM Midrange Storage with IBM i as a client of the IBM Virtual I/O Server (VIOS):

- Section 1.1, “What is new” on page 2, summarizes the technical enhancements for using Midrange Storage with IBM i.
- Section 1.2, “Virtual I/O Server concepts” on page 3, introduces the VIOS virtualization concepts relevant for IBM i clients.

1.1 What is new

IBM i, formerly known as i5/OS, has a long heritage of its storage virtualization capability being a host partition for other clients like AIX®, Linux®, and Windows® by using network server storage spaces.

The new operating system version IBM i 6.1 now extends the virtualization support on IBM POWER Systems POWER6 models by IBM i being a client partition itself of either IBM Virtual I/O Server or another IBM i 6.1 hosting partition.

From a storage perspective, IBM i being a client partition of VIOS offers completely new possibilities for IBM i external storage attachment. Instead of proprietary IBM i 520-byte/sector formatted storage including 8-byte header information and 512-byte data per sector, VIOS uses industry-standard 512-byte/sector formatted storage. This means that ordinary 512-byte/sector Storage Systems like the supported IBM DS3400, DS4700, DS4800, and DS5000 Storage Systems can now be attached to IBM i via VIOS. (See Chapter 4, “Planning for Midrange Storage” on page 89, for supported configurations and requirements.)

To make IBM i compatible with 512-byte/sector storage IBM POWER6 hypervisor (PHYP) has been enhanced to support conversion from 8x520-byte/sector pages into 9x512-byte/sector pages. The additional ninth sector, called iSeries® Specific Information (ISSI) sector, is used to store the 8-byte header information from each of the 8x520-byte sectors of a page so that they fit into 9x512 bytes. To ensure data atomicity (that is, ensure that all nine sectors now representing a 4 KB page are processed as an atomic block) some 8 bytes of control information is added so that in addition to the headers also 64 bytes of user data are shifted into the ninth sector. Figure 1-1 illustrates the 520-byte to 512-byte sector page conversion.

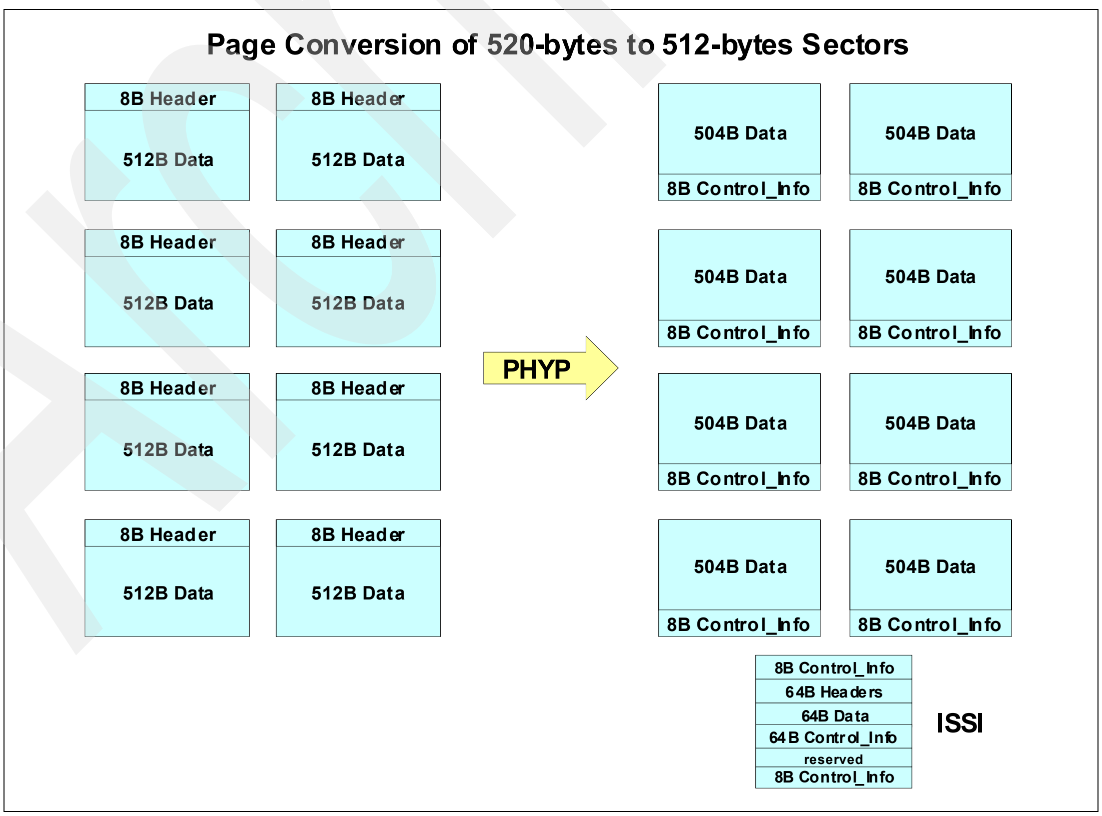


Figure 1-1 Page conversion of 520-byte to 512-byte sectors

1.2 Virtual I/O Server concepts

IBM Virtual I/O Server, first released in September 2004, is an AIX-based software appliance that has been specifically developed for virtualizing I/O resources on IBM POWER5™ and POWER6 systems for AIX, Linux, and now also IBM i client partitions. VIOS is included in the PowerVM™ (formerly known as Advanced POWER Virtualization) virtualization feature for IBM POWER Systems (IBM i, IBM p, and Blades) servers. Figure 1-2 shows the Virtual I/O Server with different client partitions communicating via virtual DMA across the POWER hypervisor.

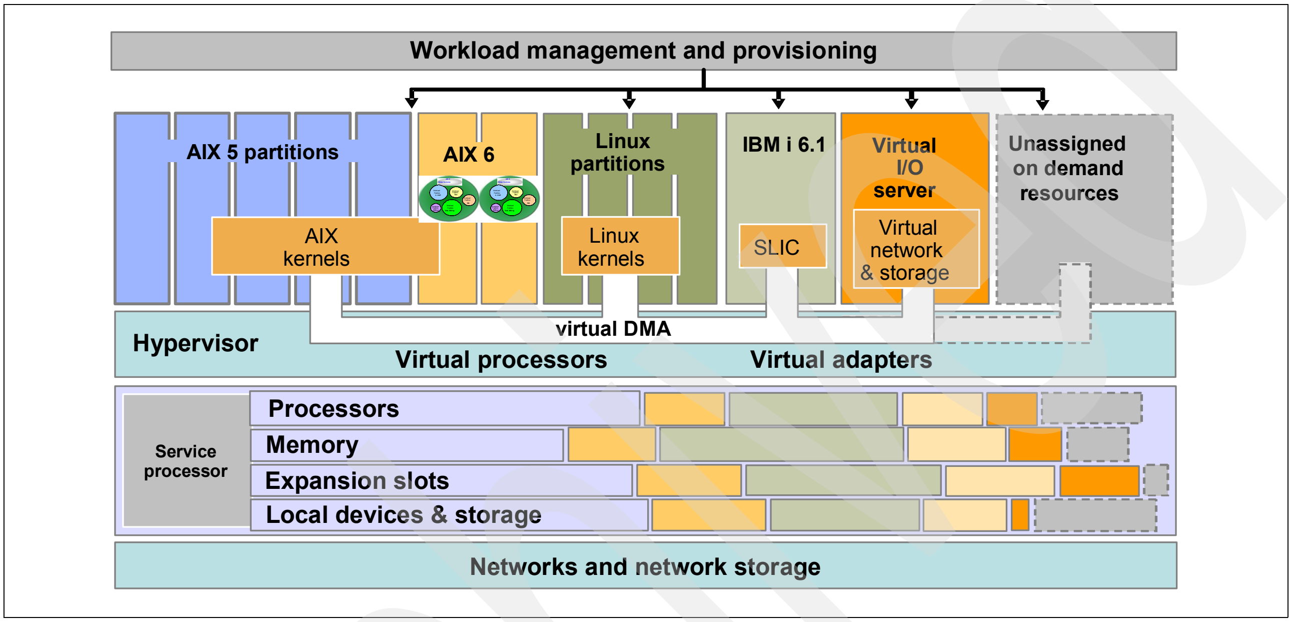


Figure 1-2 PowerVM VIrtual I/O Server

VIOS uses the para-virtualization method, meaning that changes were made to the client operating system kernels to make them aware of the virtualization done at the POWER hypervisor layer. With the operating systems using direct calls to the hypervisor, the sharing of common resources like shared processors is optimized, for example, for task switches.

The Virtual I/O Server is installed in its own logical partition (LPAR) and owns the physical I/O resources like Ethernet and SCSI/FC adapters. It virtualizes for its client LPARs to share them remotely using the built-in hypervisor services. These client LPARs can be quickly created, typically owning only real memory and shares of CPUs without any physical disks or physical Ethernet adapters.

VIOS Version 1.5 based on AIX 5.3 comprises the following virtualization features:

- Virtual SCSI disk storage and adapters
- Virtual Ethernet
- Virtual optical devices like DVD-RAM

The new VIOS Version 2.1 based on AIX 6.1 provides additional virtualization support with:

- Virtual tape

> **Note:** Virtual tape is currently not supported for IBM i clients of VIOS.

- PowerVM Active Memory™ (that is, virtual real memory (VRM)) using a shared, over-committed pool of physical memory, with overflow going to a VIOS paging disk
- N-port ID virtualization (NPIV) for virtualization of FiberChannel card ports

1.2.1 VIOS virtual SCSI support

Virtual SCSI support allows VIOS client partitions to share disk storage, optical devices, and, with VIOS 2.1, tape devices that are physically assigned to the Virtual I/O Server logical partition. This virtual SCSI support of VIOS is used to make IBM Midrange Storage devices that do not support the IBM i proprietary 520-byte/sectors format available to IBM i clients of VIOS. VIOS owns the physical adapters like the FiberChannel storage adapters connected to IBM Midrange Storage Systems. The LUNs of the physical storage devices seen by VIOS are mapped to VIOS virtual SCSI (VSCSI) server adapters created as part of its partition profile. The IBM i client partition with its corresponding virtual SCSI *client* adapters defined in its partition profile connects to the VIOS virtual SCSI *server* adapters via the hypervisor with VIOS performing SCSI emulation and acting as the SCSI target for IBM i. Figure 1-3 shows an example of the Virtual I/O Server owning the physical disk devices and its virtual SCSI connections to two IBM i clients.

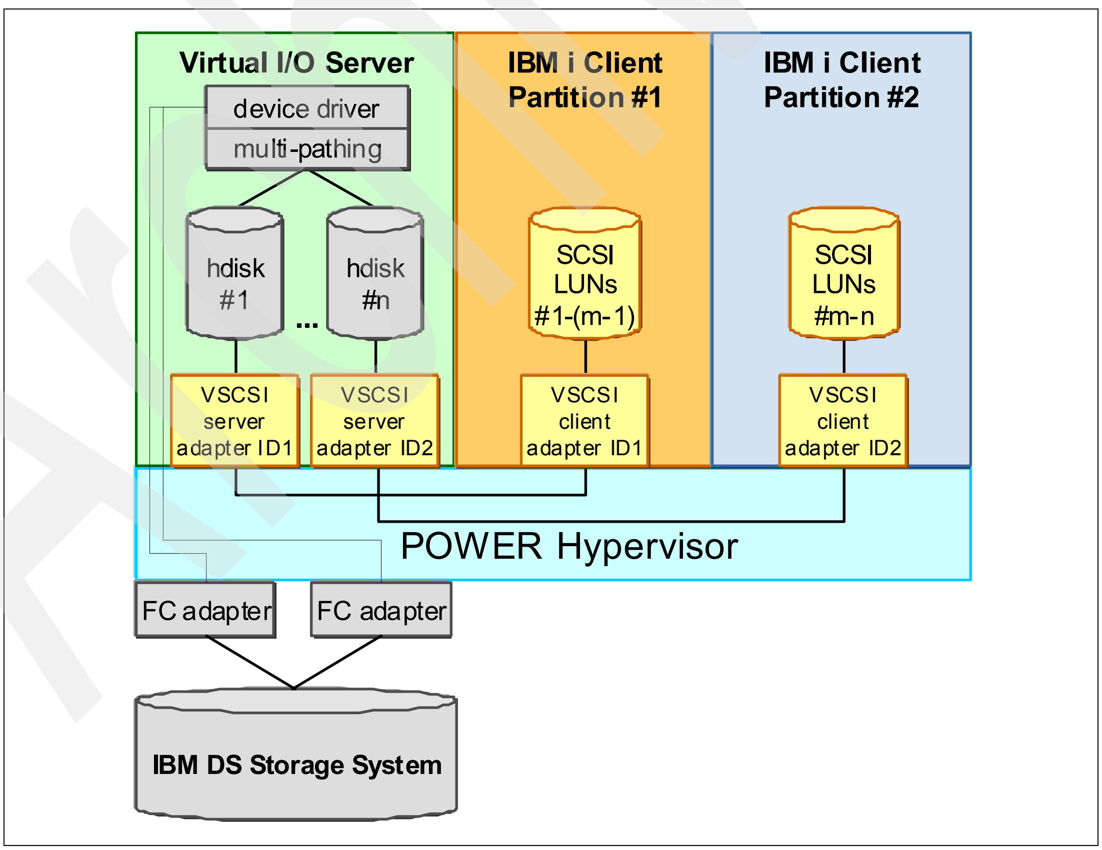


Figure 1-3 VIOS virtual SCSI support

A fast *remote direct memory access* (RDMA) I/O data transfer is used between the VIOS physical disk devices and the client virtual SCSI devices with VIOS itself processing only the SCSI command traffic. VIOS is capable of exporting a pool of heterogeneous physical storage as a homogeneous pool of block storage in the form of virtual SCSI devices. The device mapping on VIOS to provide generic SCSI devices and LUNs for its clients is either done on a physical device level like for hdisks and DVD drives or on a logical device level using logical volumes for virtual SCSI disks or files for virtual optical media. Figure 1-4 shows the block level virtualization concept of VIOS.

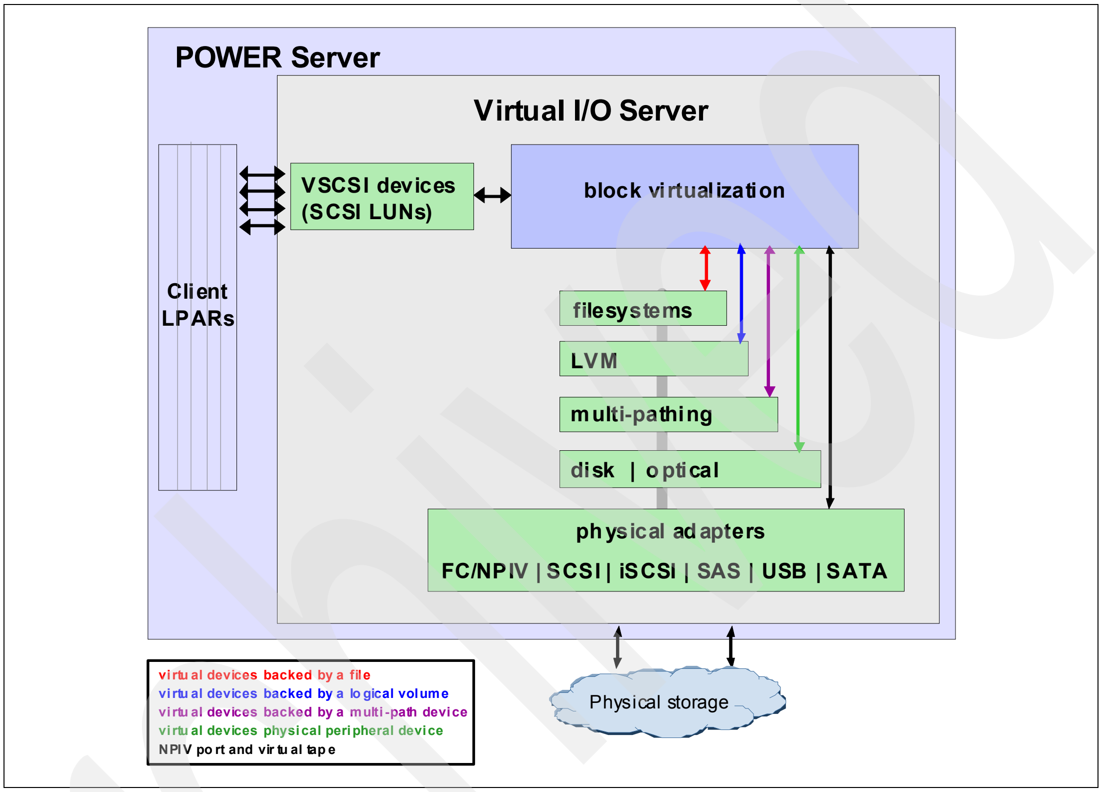


Figure 1-4 VIOS virtual SCSI block-level virtualization concept

1.2.2 Virtual Ethernet support

The VIOS virtual Ethernet support extends the virtual LAN support provided by the POWER hypervisor. While VLAN is used for inter-partition communication within the same IBM POWER Systems server through a virtual Ethernet switch in the hypervisor based on the IEEE 802.1Q VLAN standard, the Virtual I/O Server virtual Ethernet support for sharing physical Ethernet adapters allows the VIOS client LPARs to communicate with other systems in the local area network outside the server hardware unit without configuring any routing or assigning physical Ethernet ports to these client partitions.

Virtual Ethernet support by VIOS is enabled by creating a *Shared Ethernet Adapter* (SEA) serving as an Open System Interconnection (OSI) layer 2 bridge between a physical Ethernet adapter and one or more (up to 16) virtual Ethernet adapters as defined in the VIOS partition profile. The SEA forwards outbound packets received from a virtual Ethernet adapter to the external network and forwards inbound packets to the appropriate client partition's virtual Ethernet adapter. Virtual Ethernet adapters are created for a partition when creating the partition profile with the HMC assigning a unique hardware (MAC) address to them. With virtual Ethernet the logical network segmentation by VLANs is followed so that packages tagged with the VLAN ID in their Ethernet headers can only be delivered to virtual adapters that belong to the same VLAN or untagged ones to the required default virtual adapter with the assigned default port VLAN ID (PVID). Figure 1-5 shows a simple example of the VIOS virtual Ethernet support for three IBM i client partitions using a Shared Ethernet Adapter.

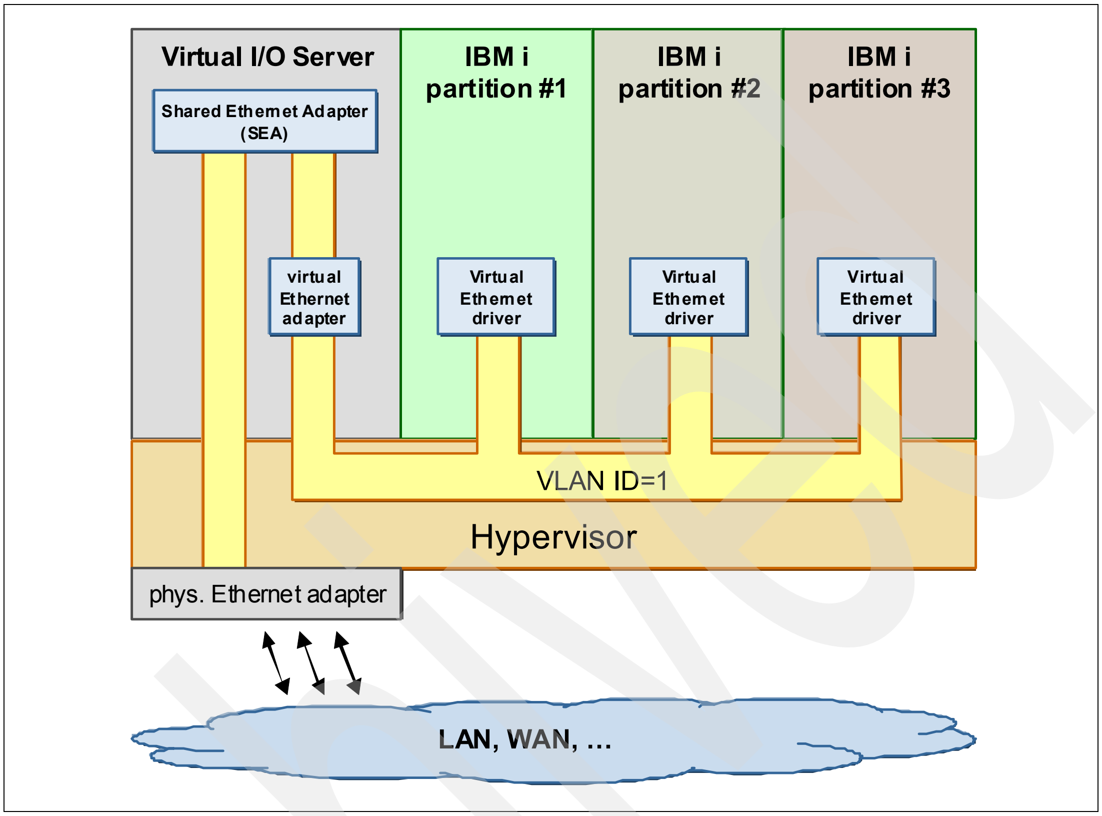


Figure 1-5 VIOS virtual Ethernet support via Shared Ethernet Adapter

The new *Integrated Virtual Ethernet* (IVE) adapter, which is a standard feature of the IBM POWER Systems POWER6 hardware, provides hardware accelerated virtualization of the physical ports on the *Host Ethernet Adapter* (HEA) GX+ bus expansion card. Unlike virtual Ethernet with packets being moved from physical to virtual adapters by VIOS, the IVE can communicate directly with the partitions without needing to interact with the hypervisor, resulting in high throughput and low latency. Each port of the HEA has its own logical OSI layer 2 switch and, with the default multi-core scaling up to four logical host Ethernet adapters (LHEA), can be connected to a physical port. The LHEAs are the logical adapters with their own unique hardware (MAC) address allocated to the partitions as defined in the partition profiles. All LHEAs assigned to the same physical HEA port can communicate with each other via inter-partition communication over the logical layer 2 switch. Otherwise, traditional VLANs with virtual Ethernet adapters can also be used for inter-partition communication.

Figure 1-6 shows the Host Ethernet Adapter hardware virtualization.

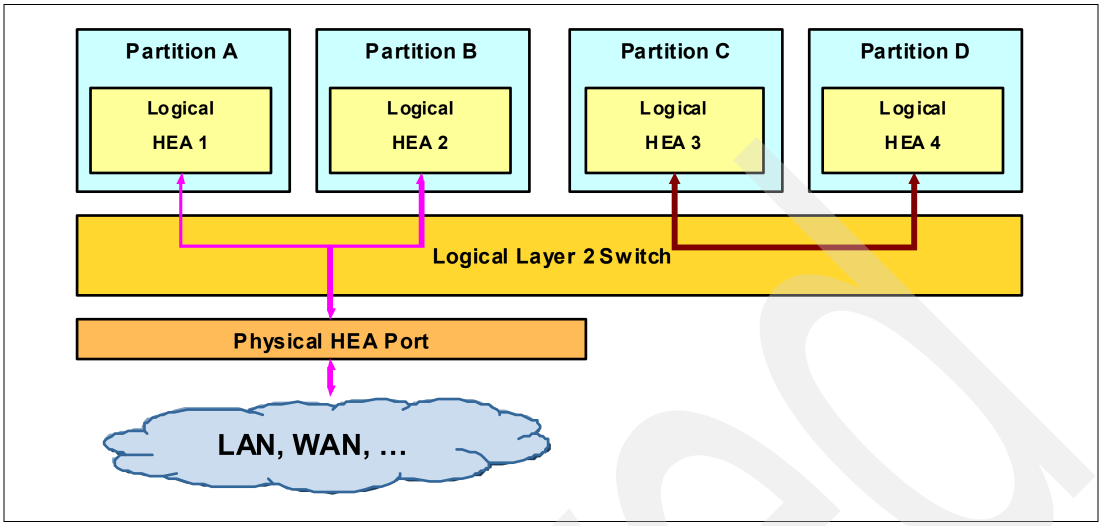


Figure 1-6 Host Ethernet Adapter hardware virtualization

Using the HEA could also be combined with implementing virtual Ethernet (SEA) by the Virtual I/O Server. However, in this case the HEA port must be set to promiscuous mode so that it is owned exclusively by the VIOS partition, that is, only a single LHEA of the VIOS LPAR can be connected to the HEA port.

With the exception noted below, we strongly recommend using the HEA over SEA, not only from a performance perspective because of the hardware-accelerated virtualization of the Host Ethernet Adapter, but also from an availability perspective because using HEA instead of SEA makes IBM i client independent from VIOS network-wise. This means that if VIOS for any reason temporarily became unavailable, the IBM i client would enter a SRC A6xx0266 hang state, but all network connections over HEA would remain established and all IBM i client operations including its remote sessions would resume as soon as VIOS became operational again.

Note: Unless more than the HEA maximum of 16 logical ports is needed for a physical Ethernet port for performance and availability reasons we strongly recommend using the POWER6 hardware accelerated virtualization of the Host Ethernet Adapter together with one or more Logical Host Ethernet Adapters for each partition instead of using virtual Ethernet (SEA) via VIOS with either a native PCI bus Ethernet adapter or the HEA.

For further information about IBM Virtual I/O Server and Integrated Virtual Ethernet refer to:

- *PowerVM Virtualization on IBM System p: Introduction and Configuration Fourth Edition*, SG24-7940

  http://www.redbooks.ibm.com/abstracts/sg247940.html?Open
- *IBM Systems Information Center*

  http://publib.boulder.ibm.com/infocenter/systems/scope/hw/index.jsp?topic=/iphb1/iphb1_vios_virtualioserveroverview.htm
- *Integrated Virtual Ethernet Adapter Technical Overview and Introduction*, REDP-4340

  http://www.redbooks.ibm.com/abstracts/redp4340.html?Open

2

Midrange Storage architecture

The IBM System Storage™ DS3000, DS4000, and DS5000 Series of storage servers use Redundant Array of Independent Disks (RAID) technology. RAID technology is used to protect user data from disk drive failures. DS3000/4000/5000 power supply fan units contain Fibre Channel (FC) or serial-attached SCSI (SAS) interfaces to connect both the host systems and the disk drive enclosures.

The storage servers in the DS3000, DS4000, and DS5000 Series provide high system availability through the use of hot-swappable and redundant components. This helps to prevent unplanned outages, especially for server consolidations on storage area networks (SANs).

All models offer autonomic functions such as dynamic volume expansion and dynamic capacity addition, allowing unused storage to be brought online without stopping operations.

Note: Dynamic volume expansion is not supported by IBM i.

This chapter contains an overview of the different IBM System Storage DS3000, DS4000, and DS5000 products that support the attachment of IBM i VIOS servers. It shows and compares the different models and features, and the available RAID levels and spare drive options to protect the system against data loss in case of hard disk drive (HDD) drive failures.

This chapter also introduces the different possibilities of managing DS3000, DS4000, and DS5000 power supply fan units with the Storage Manager GUI, the script editor or the command-line interface SMcli, and the possibility of remote alerts with the Remote Support Manager (RSM™).

Finally, it shows how to create arrays and logical drives and discusses the optional premium storage partitioning, FlashCopy®, Volume Copy, and Enhanced Remote Mirroring features.

2.1 IBM DS3400 models and features

The DS3000 product line is the entry level of the DS storage family. IBM i attachment is supported with the DS3400 only, which offers FC host connectivity and SAS drive connectivity. Figure 2-1shows the DS3400.

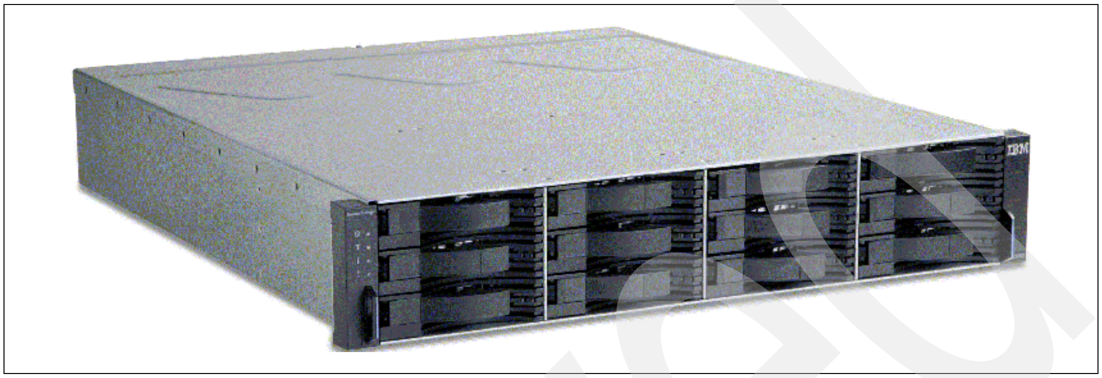


Figure 2-1 DS3400

The DS3400 is targeted at entry-level customers. It can hold a maximum of 12 disk drives inside the storage server enclosure and can attach up to three EXP3000 Expansion Units for a total of up to 48 disk drives. The DS3400 supports configurations of SAS or SATA disks, or a mix of both types of disk drives in the same enclosure. It is designed to deliver data throughput of up to 1600 MBps. The DS3400 has a total of four 4 Gbps FC host ports and 1 GB of cache memory. Like other DS3000 family members, the DS3400 is designed to efficiently handle the additional performance demands of FlashCopy and Volume Copy.

2.1.1 DS3400 model types and specifications

For the DS3400, there are two different models available. The DS3400 1726-42 is a dual controller model. The 1726-41 is a single controller model. Due to the missing redundancy, we do not recommend the single controller model 1726-41 in a high-availability environment or if good performance is essential. There is also a Telco model available that supports 48V DC power input. The Express models combine the DS3400 Storage System with the FC cables and SFP transceivers for a quick start. The characteristics of the DS3400 are:

- Models
  - 1726-41X: Single controller
  - 1726-42X: Dual controller
  - 1726-42T: Dual controller Telco
  - 1726-41E: Single controller Express
  - 1726-42E: Dual controller Express
- RAID controller: Dual active
- Cache per controller: 512 MB battery-backed cache with 1 GB upgrade option
- Host interface: Two host ports per controller, Fibre Channel (FC) 4 Gbps auto-sensing 1 Gbps/2 Gbps
- Drive interface: Serial-attached SCSI

- Supported drives
  - 3 Gbps SAS: 73 GB, 146 GB, and 300 GB SAS drives at 10,000 rpm and 36 GB, 73 GB, 146 GB, and 300 GB SAS drives at 15,000 rpm speeds
  - 3 Gbps SATA: 500 GB, 750 GB, and 1.0 TB SATA drives at 7,200 rpm speed
- RAID levels: RAID-0, 1, 3, 5, 10
- Storage partitions: 4, 16
- Maximum drives supported: 48 SAS or SATA drives (using three EXP3000 Expansion Units)
- Fans and power supplies: Dual-redundant, hot-swappable
- Rack support: 19-inch industry-standard rack
- Management software: IBM System Storage DS3000 Storage Manager
- SAN support: Supported IBM FC switches
- Warranty: Three-year parts and labor warranty
- Size
  - Height: 8.7 cm (3.4 in.)
  - Depth: 55.0 cm (21.6 in.)
  - Width: 44.7 cm (17.6 in.)
  - Weight: Approximately 17.2 kg (38.0 lb) for a standard unit; when fully configured: 29.2 kg (64.5 lb)
- Environment
  - Air temperature
    - DS3400 on: 10° to 35°C (50.0° to 95°F); altitude: 30.5 (100 ft.) below to 3,000 m (9840 ft) above sea level; temperature change: 10°C (18°F) per hour
    - DS3400 off: 10° to 50°C (14.0° to 120.0°F); maximum altitude: 3,000 m (9840 ft); temperature change: 15°C (27.0°F)
  - Per hour humidity
    - DS3400 on: 20 to 80%
    - DS3400 off: 10 to 90%
    - Maximum dew point: 26°C (79°F)
    - Maximum humidity gradient: 10% per hour
- Heat output: Approximate heat output in British thermal units (Btu) per hour
  - Minimum configuration: 205 Btu (60 watts)
  - Maximum configuration 1235 Btu (361 watts)
- Electrical input
  - Sine-wave input (50–60 Hz) required
  - Input voltage low range
    - Minimum: 90 V ac
    - Maximum: 136 V ac
  - Input voltage high range
    - Minimum: 198 V ac
    - Maximum: 264 V ac

– Approximate input kilovolt-amperes (kVA)
  - Minimum: 0.06 kVA
  - Maximum: 0.38 kVA

- Supported systems: For a list of currently supported servers, operating systems, host bus adapters, clustering applications, and SAN switches and directors, refer to the DS3000 series Interoperability Matrix available at:

  http://www-03.ibm.com/systems/storage/disk/ds3000/ds3400/

2.1.2 DS3400 optional features

There are multiple features available for the DS3400 that allow a customization of the system to the requirements. Figure 2-2 gives an overview on all available premium features.

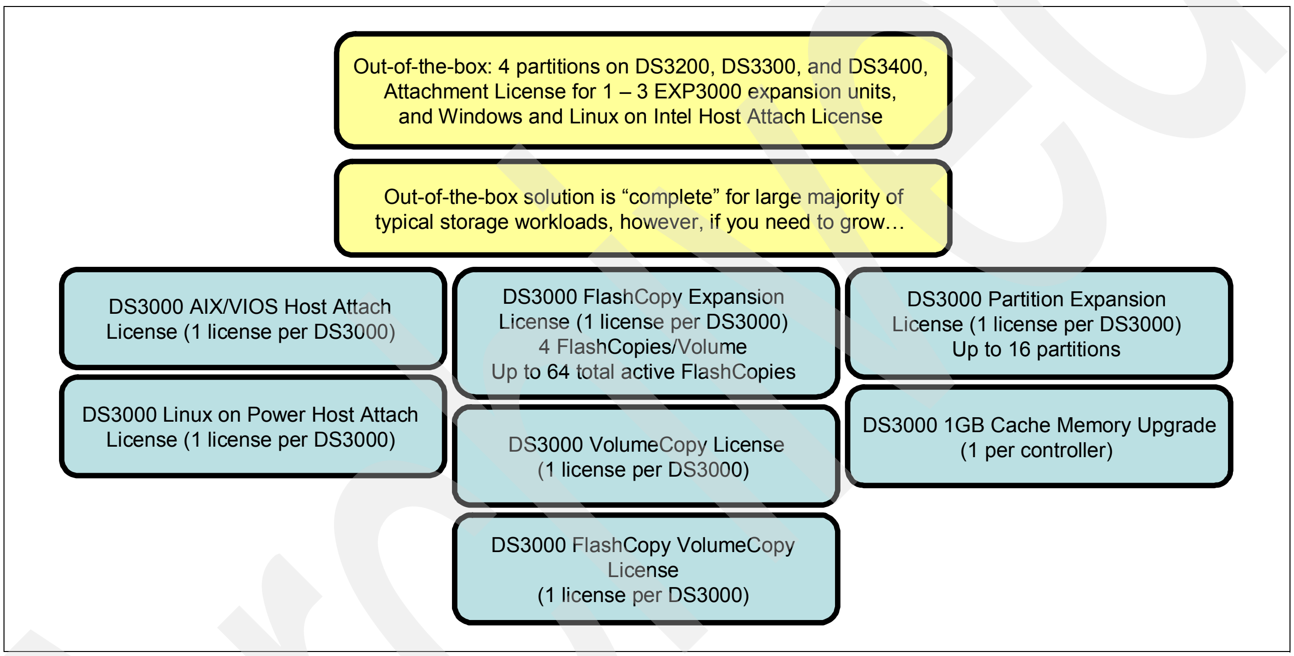


Figure 2-2 DS3400 optional features

The AIX/VIOS Host Attach feature is required to support IBM i attachment via VIOS.

The FlashCopy, Volume Copy, and partitioning features require an online activation on the storage server. After this activation, the new feature is immediately available.

The cache memory upgrade feature requires a hardware replacement, which is in fact the replacement of both controllers.

A detailed description of the FlashCopy, Volume Copy, and storage partitioning features can be found in 2.8, “Logical storage architecture” on page 64.

2.1.3 DS3400 architecture

This section explains the front and rear views of the DS3400 power supply fan unit.

Front view

The front of the DS3400 gives access to the internal drive slots, which can hold up to 12 hot-swapable SAS or SATA drives. The Bezel LEDs are located at the left side and provide information about AC power, whether the system temperature is above a critical level, and whether a fault condition exists, In addition, there is a system identifier light that can be toggled on and off through the management software to easily identify the enclosure in a multi-system environment. Figure 2-3 shows the front side of the DS3400.

Figure 2-3 DS3400 front view

Rear view

Figure 2-4 shows the rear side of the DS3400, which gives access to the hot swapable controllers and power-fan units.

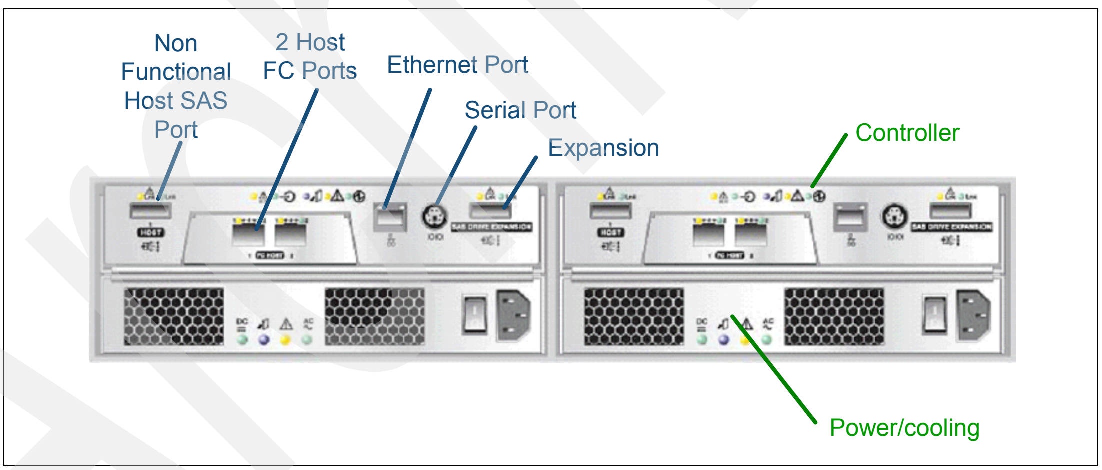


Figure 2-4 DS3400 rear view

Each controller contains two FC host ports for host attachment, an Ethernet port for system management, a serial port for maintenance access, and an expansion port for the attachment of up to three EXP3000 drive expansion units. The host SAS port at the left side is disabled on the DS3400. This is because the DS3400 shares the same chassis as the DS3200, which has SAS host connectivity.

The LEDs on the controller and power-fan units provide information about the module status.

For details on host and drive-side cabling, refer to the *Installation, User's, and Maintenance Guide - IBM System Storage DS3400,* MIGR-5069916, available at:

https://www-304.ibm.com/systems/support/supportsite.wss/docdisplay?lndocid=MIGR-5069916&brandind=5000028

The two controller support modules are located under the controllers and contain the power supplies and the cooling fans. The power-fan units are redundant, which means that only one power-fan module can provide power and cooling to both controllers and all drives within the enclosure.

2.2 IBM DS4000 models and features

The DS4000 product line is the midrange level of the DS storage family. IBM i attachment via VIOS is supported with the DS4700 and the DS4800. This section discusses both models.

2.2.1 IBM DS4700 models and features

The DS4700 Storage Server is designed to replace the DS4300 Storage Server. It is targeted at entry-level to mid-level storage customers. It can hold a maximum of 16 disk drives inside the power supply fan unit enclosure and can attach up to six EXP810 Expansion Units for a total of up to 112 Fibre Channel or SATA disk drives.

The DS4700 comes in two models, Model 72 and Model 70. The Model 72 has a total of eight 4 Gbps FC host ports and 4 GB of cache memory, while Model 70 has a total of four 4 Gbps FC host ports and 2 GB of cache memory. Like other DS4000 family members, the DS4700 supports existing customer infrastructures (helping protect investments), and it is a high-performance storage server for open systems. In addition, the DS4700 is a good choice for environments with intense replication requirements because it is designed to efficiently handle the additional performance demands of FlashCopy, Volume Copy, and Enhanced Remote Mirroring.

DS4700 model types and specifications

The DS4700 model types and specifications are:

- Models
  - 1814-70A/H/S/T, 1814-72A/H/S/T
  - 1814-70S/T, 1814-72S/T (DC power supplies)
- RAID controller: Dual active
- Cache
  - Model 70A/70S: 2 GB
  - Model 72A/72S: 4 GB
  - Battery-backed
- Host interface:
  - Eight host ports on model 72
  - Four host ports on model 70
  - Fibre Channel (FC) Switched and FC Arbitrated Loop (FC-AL) standard, Auto-sensing 1 Gbps/2 Gbps/4 Gbps

- Drive interface
  - Four drive ports: Fibre Channel (FC) Switched and FC Arbitrated Loop (FC-AL) standard
  - Auto-sensing 1 Gbps/2 Gbps/4 Gbps
- Supported drives:
  - 2 Gbps FC: 10K – 73.4 GB, 146.8 GB, 300 GB E-DDM,

    15K – 36.4 GB, 73.4 GB, 146.8 GB E-DDM
  - 4 Gbps FC: 15k – 300 GB, 146.8 GB, 73.4 GB, 36.4 GB E-DDM
  - 4 Gbps SATA: 7.2K 500 GB/750 GB and 1 TB E-DDM disk drives
- RAID levels: 0, 1, 3, 5, 6, 10
- Storage partitions: 4, 8, 16, 64, or 128 storage partitions
- Maximum drives supported:
  - Model 72A/H/S/T: 112 FC or 112 SATA drives (using six DS4000 EXP810 Expansion Units)
  - Model 70A//H/S/T: 112 FC or 112 SATA drives (using six DS4000 EXP810 Expansion Units)
- Fans and power supplies: Dual redundant, hot-swappable
- Rack support: 19-inch, industry-standard rack
- Management software: IBM System Storage DS4000 Storage Manager Version 10.10
- SAN support: Supported IBM FC switches and directors (product numbers 2005, 2006, 2109, 2026, 2027, 2031, 2032, 2034, 2042, 2054, 2061, and 2062, and IBM BladeCenter®)
- Warranty: Three-year parts and labor warranty, 9x5 next business day, upgradeable to 24x7 with four-hour response
- Physical characteristics
  - Dimensions
    - Height: 130.3 mm (5.13 in.)
    - Width: 447.0 mm (17.6 in.)
    - Depth: 563.8 mm (22.2 in.)
  - Weight: 36.38 kg (80.2 lb)
- Environment
  - Air temperature
    - DS4700 on: 10° to 35°C (50.0° to 95° F); altitude: 30.5 (100 ft.) below to 3000 m (9842 ft) above sea level; temperature change: 10°C (18°F) per hour
    - DS4700 off: - 10° to 65°C (14° to 149°F); maximum altitude: 3,000 m (9842 ft); temperature change: 15°C (27.0°F)
  - Per-hour humidity
    - DS4800 on: 8% to 80%
    - DS4800 off: 8% to 90%
    - Maximum dew point: 26°C (79°F)
    - Maximum humidity gradient: 10% per hour

- Heat output

  Approximate heat output in British thermal units (Btu) per hour: 1516 Btu

- Electrical input
  - Sine-wave input (50–60 Hz) required
  - Input voltage low range
    - Minimum: 90 V ac
    - Maximum: 136 V ac
  - Input voltage high range
    - Minimum: 198 V ac
    - Maximum: 264 V ac
  - Approximate input kilovolt-amperes (kVA): 0.454

- Supported systems

  For a list of currently supported servers, operating systems, host bus adapters, clustering applications, and SAN switches and directors, refer to the DS4700 Express Interoperability Matrix available at:

  http://www-03.ibm.com/systems/storage/disk/ds4000/ds4700/

DS4700 optional features

There are multiple features available for the DS4700 that allow a customization of the system to the requirements. The following premium features are available for the DS4700:

- DS4700 Advanced Features
  - 1814-7301 DS4700 FlashCopy Activation
  - 1814-7303 DS4700 Volume Copy Activation
  - 1814-7304 DS4700 FlashCopy/Volume Copy Activation
  - 1814-7305 DS4700 Enhanced Remote Mirror Activation
- DS4700 Expansion Unit attachments
  - 1814-7382 DS4700 Model 70 (1-3 EXPs) Attachment
  - 1814-7383 DS4700 Model 70 (4-6 EXPs) Attachment
  - 1814-7384 DS4700 Model 72 (4-6 EXPs) Attachment
- DS4700 Operating System Options
  - 1814-7700 DS4700 Windows Host Kit
  - 1814-7701 DS4700 Linux/Intel® Host Kit
  - 1814-7702 DS4700 Novell® NetWare Host Kit
  - 1814-7711 DS4700 AIX Host Kit (required for IBM i via VIOS attachment)
  - 1814-7712 DS4700 Sun™ Host Kit
  - 1814-7713 DS4700 HP-UX Host Kit
  - 1814-7714 DS4700 Linux on Power Host Kit
- DS4700 Storage Partition Upgrades
  - 1814-8860 DS4700 Model 70 2 to 4 Storage Partitions Upgrade
  - 1814-8861 DS4700 Model 70 2 to 8 Storage Partitions Upgrade
  - 1814-8862 DS4700 Model 70 4 to 8 Storage Partitions Upgrade
  - 1814-8863 DS4700 Model 70 4 to 16 Storage Partitions Upgrade
  - 1814-8864 DS4700 8 to 16 Storage Partitions Upgrade
  - 1814-8865 DS4700 8 to 64 Storage Partitions Upgrade
  - 1814-8866 DS4700 16 to 64 Storage Partitions Upgrade

DS4700: Front view

Figure 2-5 shows the front section of the DS4700 with the 16 pre-installed drive bays of hot-swap disk drives. The DS4700 supports an intermix of Fibre Channel (FC) or SATA drives in the same or different enclosures.

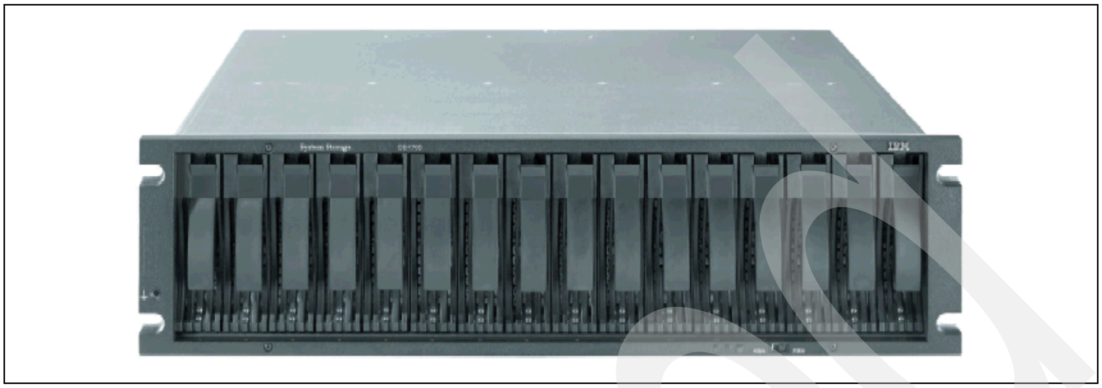


Figure 2-5 DS4700 front view

The hot-swap features of the DS4700 enable you to remove and replace the Fibre Channel or SATA hard disk drives without turning off the storage expansion enclosure. You can maintain the availability of your system while a hot-swap device is removed, installed, or replaced.

Each drive and carrier assembly is called a drive customer replaceable unit (CRU). Install drive CRUs in the 16 drive bays on the front of the storage expansion enclosure from the right-most slot (slot 16) to the left-most slot (slot 1).

Several LED indicators and the FC Link speed selector are also visible from the front of the storage unit.

The FC link speed selector is a physical switch that must be used to set the enclosure speed. It should be set to 2 Gbps if 2 Gbps drives are used. If it is set to 4 Gbps while 2 Gbps drives are used, the 2 Gbps drives will appear as bypassed drives in Storage Manager.)

Note: The drive link speed switch is only read out during the power-up sequence, so it should only be switched during power off.

DS4700: Rear view

Figure 2-6 shows the rear of the DS4700 with the two hot-swappable and redundant RAID controllers. The left controller is labeled controller A and the right controller is labeled controller B. Note that controller A is positioned upside-down relative to controller B. It is important to keep this in mind when connecting the back-end ports to hosts and drive-side expansion enclosures. The redundant power supply fan units are also visible from the back of the DS4700.

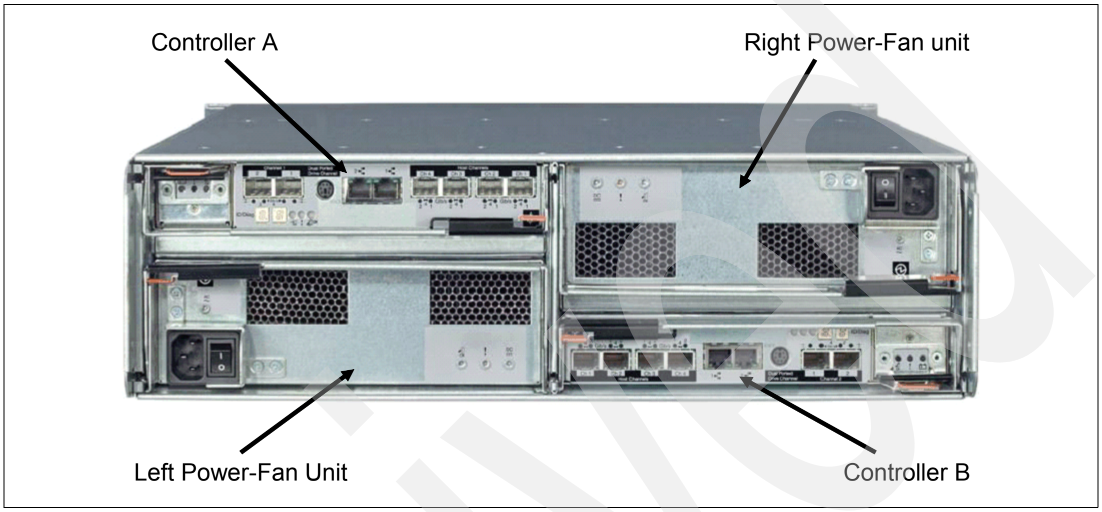


Figure 2-6 DS4700 rear view

RAID controllers

Each controller (Figure 2-7) includes the following features:

- Two drive ports
- One RS232 serial port
- Four host ports (Model 72) or two host ports (Model 70)
- Dual Ethernet ports
- One battery CRU

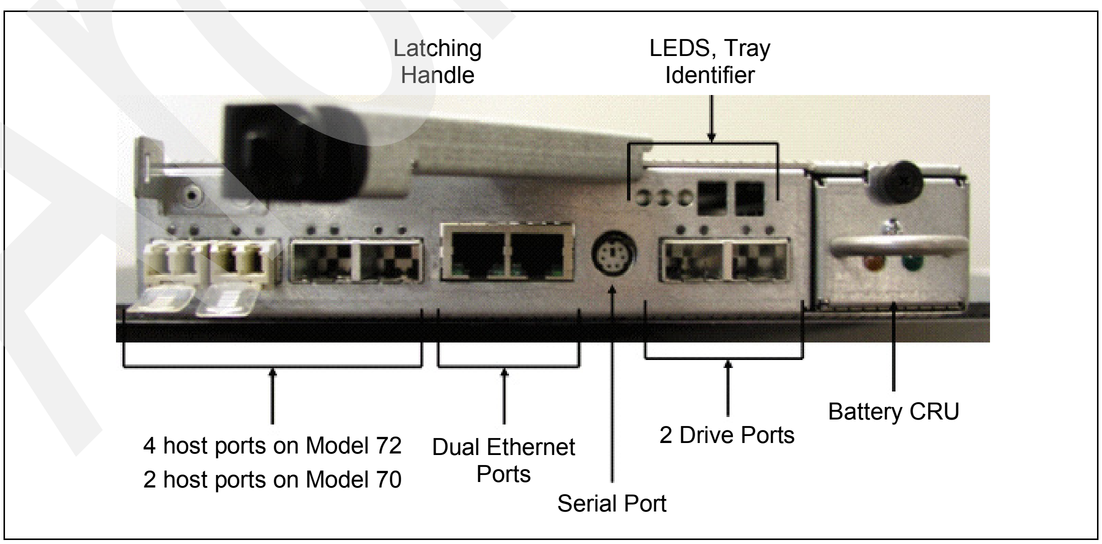


Figure 2-7 DS4700 RAID controller (may look a little different on the GA product)

The RS232 Serial Port is a *PS2 type* port. This serial port is used for management and diagnostic purposes. You can use a PC with a terminal emulation utility such as Hyper Terminal to access the command set.

The dual Ethernet ports accommodate an RJ-45 10BASE-Tx or 100BASE-Tx Ethernet connection. There are two ports per controller. One port is designed for out-of-band management and the other port is meant for serviceability. This feature is similar to DS4800.

Because of the extra port, it is now preferable to have two IP addresses per controller in order to manage and service the DS4700 appropriately. You can still operate the DS4700 with only one IP port active per controller. You can assign either port of the two ports for management or service on a given controller.

The default IP addresses for the controller A Ethernet ports 1 and 2 are 192.168.128.101 and 192.168.129.101, respectively. The default IP addresses for the controller B Ethernet ports 1 and 2 are 192.168.128.102 and 192.168.129.102, respectively. The default subnet mask for all four Ethernet ports is 255.255.255.0.

Important: When changing the default IP addresses, remember that port 1 and 2 of each controller must be in separate subnets.

DS4700 has a redundant battery system. Each controller has one battery CRU—a new Lithium Ion battery system for cache. Its hold-up time is up to seven days for Model 70 and three days for Model 72. The minimum life of the battery can be up to three years. The battery can be easily replaced on-site (Figure 2-8).

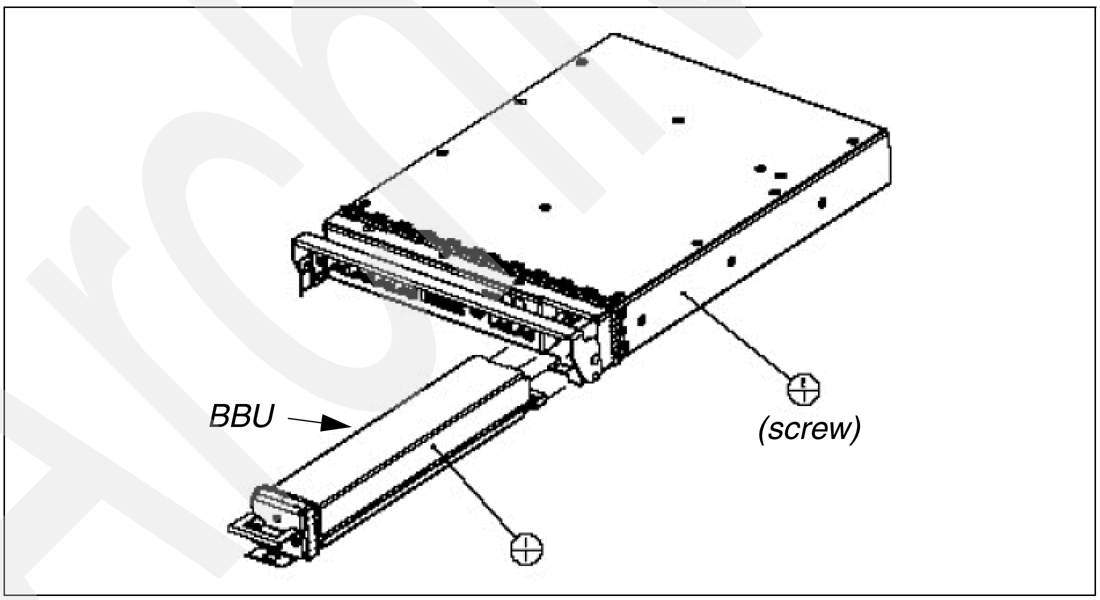


Figure 2-8 Replaceable backup battery unit

Power supply fan units

The DS4700 controller enclosure has two removable power supply fan units. Each power supply fan unit contains one power supply and two fans.

The four fans pull air through the drive from front to back across the drives. The fans provide redundant cooling, which means that if one of the fans in either fan housing fails, the remaining fans continue to provide sufficient cooling to operate the controller enclosure.

The power supplies provide power to the internal components by converting incoming AC voltage to DC voltage. If one power supply is turned off or malfunctions, the other power supply maintains electrical power to the controller enclosure.

To preserve the optimal airflow, do not remove a failed power supply fan CRU from the DS4700 controller enclosure chassis until you are ready to replace it with a new CRU.

Note: Although both power supply fan units (left and right) are identical, they are seated in the DS4700 controller enclosure chassis in opposite orientations. If the power supply fan cannot fully be inserted in the power supply fan bay, flip it 180 degrees and reinsert it.

2.2.2 IBM DS4800 models and features

IBM DS4800 power supply fan unit delivers breakthrough midrange disk performance and outstanding reliability for demanding applications in compute-intensive environments. The DS4800 is a key component of IBM business continuity solutions portfolio, delivering business resilience and continuity of operations. Figure 2-9 shows the front view of a DS4800 power supply fan unit.

Figure 2-9 DS4800 power supply fan unit

The DS4800 takes advantage of 4 Gbps Fibre Channel interface technology and can support up to 224 disk drives by attaching IBM System Storage EXP810, EXP710, or EXP100 disk units. It is a great choice for performance-oriented or capacity-oriented Midrange Storage requirements. Four models are available:

- The Model 80 with 4 GB of cache
- The 82A with 4 GB of cache
- The 84A with 8 GB of cache
- The Model 88A with 16 GB of cache

Additionally, support for high-performance Fibre Channel and high-capacity Serial ATA (SATA) disk drives help enable a single DS4800 Storage System to satisfy primary and secondary storage to accommodate the changing value of data over time while maintaining data availability.

The DS4800 disk Storage System supports disaster recovery strategies with the FlashCopy, Volume Copy, and Enhanced Remote Mirroring features.

DS4800 model types and specifications

There are four different DS4800 models available that only differ in the controller cache size:

- Models
  - 1815-80A
  - 1815-82A
  - 1815-84A
  - 1815-88A
- RAID controller: Dual active
- Cache
  - Model 80: 4 GB
  - Model 82A: 4 GB
  - Model 84A: 8 GB Model
  - 88: 16 GB
  - Battery-backed
- Host interface
  - Eighth host ports: Fibre Channel switched and FC Arbitrated Loop (FC-AL) standard
  - Auto-sensing 1 Gbps/2 Gbps/4 Gbps
- Drive interface
  - Eight drive ports: Fibre Channel switched and FC Arbitrated Loop (FC-AL) standard
  - Auto-sensing 2 Gbps/4 Gbps
- Supported drives with expansion units
  - 4 Gbps SATA: 7.2K rpm, 1 TB, 750 GB, and 500 GB
  - 4 Gbps FC: 15K rpm, 300 GB/146 GB/73 GB (E-DDM)
- RAID levels: 0, 1, 3, 5, 10
- Storage partitions: 8, 16, 32, 128, 256, or 512 storage partitions
- Maximum drives supported

  Note that DS4800 supports intermixing EXP810, EXP710, and EXP100 Expansion Units.
  - Model 80A: 224 drives (using 14 DS4000 EXP810, 16 EXP710, or EXP100 Expansion Units)
  - Model 82A: 224 drives (using 14 DS4000 EXP810, 16 EXP710, or EXP100 Expansion Units)
  - Model 84A: 224 drives (using 14 DS4000 EXP810, 16 EXP710, or EXP100 Expansion Units)
  - Model 88A: 224 drives (using 14 DS4000 EXP810, 16 EXP710, or EXP100 Expansion Units)

- Fans and power supplies: Dual redundant, hot-swappable
- Rack support: 19-inch, industry-standard rack
- Management software: IBM System Storage DS4000 Storage Manager Version 10.15
- SAN support

  Supported IBM FC switches and directors (product numbers 2005, 2006, 2109, 2026, 2027, 2031, 2032, 2034, 2042, 2054, 2061, and 2062, and IBM BladeCenter)
- Warranty: Three-year parts and labor warranty, 9x5 next business day, upgradeable to 24x7 with four-hour response
- Physical characteristics
  - Dimensions:
    - Height: 174.50 mm (6.87 in)
    - Width: 481.75 mm (18.97 in)
    - Depth: 634.92 mm (25.0 in)
  - Weight: 36.38 kg (80.2 lb)
- Environment
  - Air temperature
    - DS4800 on: 10° to 35°C (50.0° to 95°F); altitude: 30.5 (100 ft) below to 3048 m (10000 ft) above sea level; temperature change: 10°C (18°F) per hour
    - DS4800 off: 1° to 60°C (33° to 140°F); maximum altitude: 3048 m (10000 ft); temperature change: 15°C (27.0°F)
  - Per-hour humidity
    - DS4800 on: 20 to 80%
    - DS4800 off: 10 to 90%
    - Maximum dew point: 26°C (79°F)
    - Maximum humidity gradient: 10% per hour
- Heat output: Approximate heat output in British thermal units (Btu) per hour: 803.7 Btu
- Electrical input
  - Sine-wave input (50–60 Hz) required
  - Input voltage low range
    - Minimum: 90 V ac
    - Maximum: 132 V ac
  - Input voltage high range
    - Minimum: 198 V ac
    - Maximum: 264 V ac
  - Approximate input kilovolt-amperes (kVA): 0.240
- Supported systems

  For a list of currently supported servers, operating systems, host bus adapters, clustering applications, and SAN switches and directors refer to:

  http://www-03.ibm.com/systems/storage/disk/ds4000/ds4800/

DS4800 optional features

There are multiple features available for the DS4800 that allow a customization of the system to the requirements. The following premium features are available for the DS4800:

- DS4800 Model 80 Enhanced Performance Activation option 44E5459
- DS4800 Volume Copy
- DS4800 FlashCopy 4 to 16 FC upgrade 44X2431
- DS4800 Enhanced Remote Mirror 22R6621
- DS4800 Enhanced Remote Mirror: 64 to 128 Mirrors upgrade 44E5380
- DS4800 Storage Partition upgrades
  - DS4800 8 to 32 Storage Partitions upgrade-MES 44E5460
  - DS4800 16 to 32 Storage Partitions upgrade-MES 44E5461
  - DS4800 32 to 64 Storage Partitions upgrade-MES 44E5462
  - DS4800 32 to 128 Storage Partitions upgrade-MES 44E5463
  - DS4800 64 to 128 Storage Partitions upgrade-MES 44E5464
  - DS4800 64 to 256 Storage Partitions upgrade-MES 44E5465
  - DS4800 128 to 256 Storage Partitions upgrade-MES 44E5466
  - DS4800 128 to 512 Storage Partitions upgrade-MES 44E5467
  - DS4800 256 to 512 Storage Partitions upgrade-MES 44E5468
- Cache memory upgrades
  - Cache memory upgrade to 16 GB option for DS4800 (1815-84H) 23R0548
  - Cache memory upgrade to 8 GB option DS4800 (1815-82H) 23R0546
  - Cache memory upgrade to 16 GB option DS4800 (1815-82H) 23R0547
- DS4800 operating systems options
  - Linux/Intel Host Kit (CD-ROM) 22R4252 41Y5133
  - Novell NetWare Host Kit (CD-ROM) 22R4253 41Y5135
  - VMware® ESX Host Kit (CD-ROM) 22R4254 41Y5137
  - AIX Host Kit (CD-ROM) 22R4255 41Y5131 (required for IBM i via VIOS attachment)
  - SUN Solaris™ Host Kit (CD-ROM) 22R4256 41Y5136
  - HP/HP-UX Host Kit (CD-ROM) 22R4257 41Y5132
  - Linux on Power Host Kit (CD-ROM) 22R4258 41Y5134

DS4800: Front view

The front section of the DS4800 contains the two controller support modules and the interconnect module (Figure 2-10).

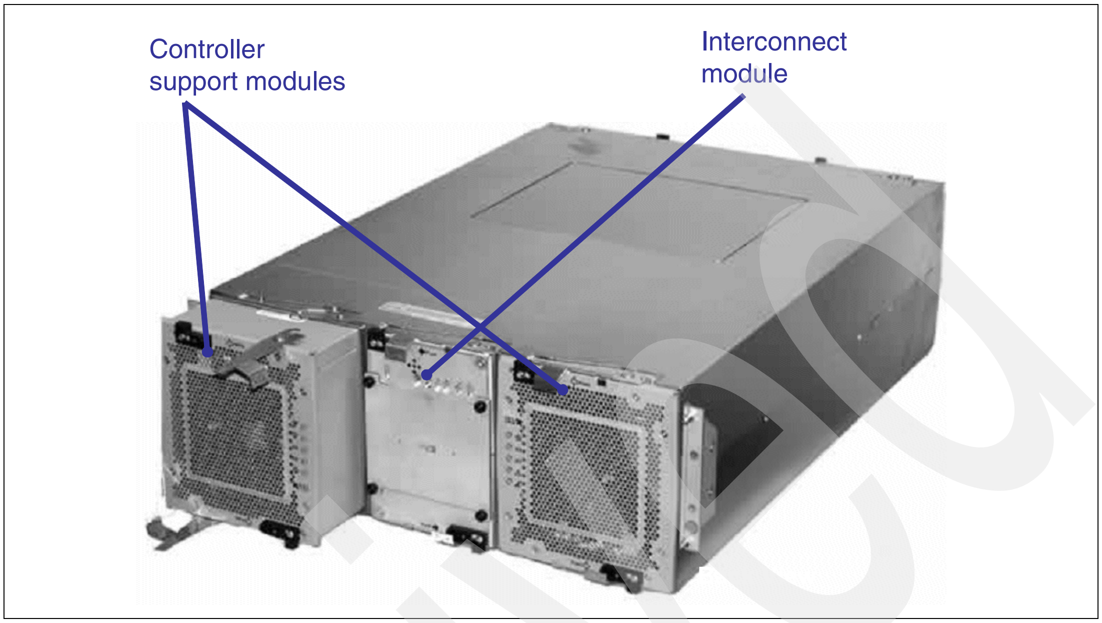


Figure 2-10 DS4800 front view

The controller support modules are the units (field replaceable units, FRUs) on the left and right in Figure 2-10. They each house a fan and a 375 W power supply. In order to replace the fan or power supply it is necessary to remove the controller support module and replace the broken or defective part.

The interconnect module is the unit (FRU) between the two controller support modules. The interconnect holds the cache batteries and the hot-swapable midplane. When the interconnect module is removed, the DS4800 automatically suspends controller B, fails over all the LUNs to controller A, and continues to operate.

When the interconnect module is put back in place, the DS4800 can revert back to normal operations. However, you have to manually redistribute the LUNs to their respective owning controller (A or B), since the DS4800 will not automatically redistribute for you.

All the components depicted in Figure 2-10 are hot-swappable and can be replaced on-the-fly to ensure that users do not incur any downtime.

Interconnect module

The interconnect module provides the electrical communication path between the power supply fan units and allows their power supplies to load-share and to charge the cache-backup battery packs. It houses two cache-backup battery packs. Each battery pack contains batteries for both controllers' caches.

For data-safety reasons the write-caching is disabled when either one of the backup battery packs fails. You should replace the failed battery pack as soon as possible to minimize any performance impact due to the disabling of the write-caching function.

Power supply fan units

The DS4800 controller enclosure has two removable power supply units (PSUs) that also contain the cooling fans for the entire enclosure.

The fans pull air through the ventilation holes of the PSUs and guide it through the controllers to the back side of the enclosure. The fans provide redundant cooling, which means that if one of the fans in either fan housing fails, the remaining fan continues to provide sufficient cooling to operate the controller enclosure.

The power supplies provide power to the internal components by converting incoming AC voltage to DC voltage. If one power supply is turned off or malfunctions, the other power supply maintains electrical power to the controller enclosure.

To preserve the optimal airflow, do not remove a failed power supply fan CRU from the DS4800 controller enclosure chassis until you are ready to replace it with a new CRU.

Note: Although both power supply fan units (left and right) are identical, they are seated in the DS4800 controller enclosure chassis in opposite orientations. If the power supply fan cannot fully be inserted in the power supply fan bay, flip it 180 degrees and reinsert it.

DS4800: Rear view

Figure 2-11 shows the rear of the DS4800 that has the two controllers stacked horizontally. Controller A is located on the top and controller B is located on the bottom. Controllers are hot-swappable.

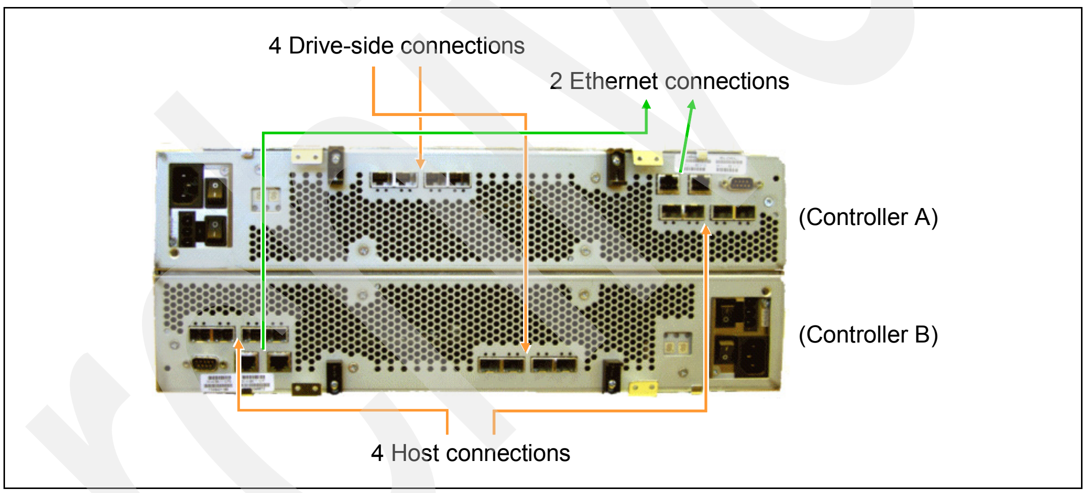


Figure 2-11 Rear view of the DS4800

Notice that controller A is positioned upside-down relative to controller B. It is important to keep this in mind when connecting the back-end ports to hosts and drive-side expansion enclosures, as we will discuss later.

Each controller is also equipped with two Ethernet RJ45 connectors and one serial port.

RAID controllers

Each controller includes the following features:

- Four drive ports
- One RS232 serial port
- Four host ports
- Dual Ethernet ports
- One power connector and a power switch for this power line

The RS232 Serial Port is a *PS2 type* port. This serial port is used for management and diagnostic purposes. You can use a PC with a terminal emulation utility such as Hyper Terminal to access the command set.

The dual Ethernet ports accommodate an RJ-45 10BASE-Tx or 100BASE-Tx Ethernet connection. There are two ports per controller. One port is designed for out-of-band management and the other port is meant for serviceability. This feature is similar to DS4700.

Because of the extra port, it is now preferable to have two IP addresses per controller in order to manage and service the DS4800 appropriately. You can still operate the DS4800 with only one IP port active per controller. You can assign either port of the two ports for management or service on a given controller.

The default IP addresses for the controller A Ethernet ports 1 and 2 are 192.168.128.101 and 192.168.129.101, respectively. The default IP addresses for the controller B Ethernet ports 1 and 2 are 192.168.128.102 and 192.168.129.102, respectively. The default subnet mask for all four Ethernet ports is 255.255.255.0.

Important: When changing the default IP addresses, remember that port 1 and 2 of each controller must be in separate subnets.

Refer to the IBM Support Web site for additional model-related information:

http://www.ibm.com/servers/storage/support/disk/ds4800/index.html

2.3 IBM DS5000 models and features

IBM DS5000 power supply fan unit is the DS4800 follow-on product and delivers breakthrough disk performance for demanding applications in compute-intensive environments and is architected to provide high reliability and availability.

The DS5000 offers linearly scalable performance to match growth requirements. The outstanding performance is supported by its seventh-generation controller technology, which features custom ASICs with built-in hardware-assist for RAID-5 XOR and RAID-6 p+q, PCI-Express buses to deliver 4 GBps of internal bandwidth, and two dedicated internal PCI-Express cache mirroring buses.

It features a new host attachment design that has sixteen 4 Gbps FC host ports at the beginning and is ready for 8 Gbps FC and 10 GBps iSCSI in the future.

The DS5000 can support up to 256 FC and SATA drives by attaching EXP510 disk expansion units. Future upgrades will be able to support up to 448 drives.

FC and SATA drive intermix is supported in the same or over different drive expansion enclosures.

An upgrade feature is available to attach existing EXP810 disk expansion units to simplify the migration from existing DS4000 systems.

Two models will be available:

- The DS5100 (Model 51A) with 8 GB of cache and eight 4 Gbps FC interfaces
- The DS5300 (model 53A) with 8 or 16 GB of cache, eight or sixteen 4 Gbps FC interfaces, and increased performance abilities

The DS5000 disk Storage System supports disaster recovery strategies with the FlashCopy, Volume Copy, and Enhanced Remote Mirroring features.

DS5000 model types and specifications

The DS5000 is initially available as two models:

- The base model DS5100 (1818-51A)
- The high performance model DS5300 (1818-53A).

If a feature or option is valid for both models, we will only refer to DS5000 in general.

- Models
  - DS5100 model 1818-51A
  - DS5300 model 1818-53A
- RAID controller: Dual active
- Cache
  - Model 51A: 8 GB
  - Model 53A: 8 or 16 GB
  - Battery-backed with permanent cache destaging to USB flash drive after power loss
- Host interface
  - 8 (51A and 53A) or 16 (53A) host ports: Fibre Channel switched and FC Arbitrated Loop standard
  - Auto-sensing 1 Gbps/2 Gbps/4 Gbps
- Drive interface
  - 16 drive ports: Fibre Channel switched and FC Arbitrated Loop standard
  - Auto-sensing 2 Gbps/4 Gbps
- Supported drives with expansion units
  - 4 Gbps SATA: 7.2K rpm, 1 TB, 750 GB and 500 GB
  - 4 Gbps FC: 15K rpm, 300 GB/146 GB/73 GB (E-DDM)
- RAID levels: 0, 1, 3, 5, 6, 10
- Storage partitions: 8, 16, 32, 128, 256 or 512 storage partitions
- Maximum drives supported

  Note that DS5000 supports intermixing EXP510 and EXP810 Expansion Units
  - 256 drives (using 16 EXP510 or EXP810 Expansion Units)
  - Future upgrade planned to support 448 drives
- Fans and power supplies: Dual redundant, hot-swappable
- Rack support: 19-inch, industry-standard rack
- Management software: IBM System Storage DS4000 Storage Manager Version 10.15
- SAN support: Supported IBM FC switches and directors
- Warranty: One year parts and labor warranty, 24x7 with four-hour response, maintenance extension available

- Physical characteristics
  - Dimensions:
    - Height: 174.50 mm (6.87 in)
    - Width: 481.75 mm (18.97 in)
    - Depth: 634.92 mm (25.0 in)
  - Weight: 40.9 kg (90.0 lb)
- Environment
  - Air temperature
    - DS5000 on: 10° to 35°C (50.0° to 95°F); altitude: 30.5 (100 ft) below to 3048 m (10000 ft) above sea level; temperature change: 10°C (18°F) per hour
    - DS5000 off: 1° to 60°C (33° to 140°F); maximum altitude: 3048 m (10000 ft); temperature change: 15°C (27.0°F)
  - Per-hour humidity
    - DS5000 on: 20 to 80%
    - DS5000 off: 10 to 93%
    - Maximum dew point: 26°C (79°F)
    - Maximum humidity gradient: 10% per hour
- Heat output: Approximate heat output in British thermal units (Btu) per hour: 1842 Btu
- Electrical input
  - Sine-wave input (50–60 Hz) required
  - Input voltage low range
    - Minimum: 90 V ac
    - Maximum: 132 V ac
  - Input voltage high range
    - Minimum: 198 V ac
    - Maximum: 264 V ac
  - Approximate input kilovolt-amperes (kVA): 0.562
- Supported systems

  For a list of currently supported servers, operating systems, host bus adapters, clustering applications, and SAN switches and directors refer to:

  `http://www-03.ibm.com/systems/storage/disk/`

DS5000 optional features

There will be multiple features available for the DS5000 that allow a customization of the system to the requirements. The following premium features will be available for the DS5000:

- DS5100 Model 51 Enhanced Performance Activation option
- Host Port Cards: Up to 16 host ports; 4 Gb FC initially, 8 Gb FC and 10 Gb EiSCSI in future release
- EXP810 upgrade for attachment: To encourage migration from DS4800s
- DS5000 Volume Copy
- DS5000 FlashCopy 4 to 16 FC upgrade

- DS5000 Enhanced Remote Mirror
- DS5000 Enhanced Remote Mirror: 64 to 128 Mirrors upgrade
- DS5000 Storage Partition upgrades
  - DS5000 8 to 32 Storage Partitions upgrade
  - DS5000 16 to 32 Storage Partitions upgrade
  - DS5000 32 to 64 Storage Partitions upgrade
  - DS5000 32 to 128 Storage Partitions upgrade
  - DS5000 64 to 128 Storage Partitions upgrade
  - DS5000 64 to 256 Storage Partitions upgrade
  - DS5000 128 to 256 Storage Partitions upgrade
  - DS5000 128 to 512 Storage Partitions upgrade
  - DS5000 256 to 512 Storage Partitions upgrade
- Cache memory upgrade to 16 GB option for DS5000
- DS5000 operating systems options
  - Linux/Intel Host Kit (CD-ROM)
  - Novell NetWare Host Kit (CD-ROM)
  - VMware ESX Host Kit (CD-ROM)
  - AIX Host Kit (CD-ROM) (required for IBM i via VIOS attachment)
  - SUN Solaris Host Kit (CD-ROM)
  - HP/HP-UX Host Kit (CD-ROM)
  - Linux on Power Host Kit (CD-ROM)

A list of all available features can be found on:

http://www-03.ibm.com/systems/storage/disk/

DS5000 front view

The front section of the DS5000 contains the two controller support modules and the Interconnect module (Figure 2-12).

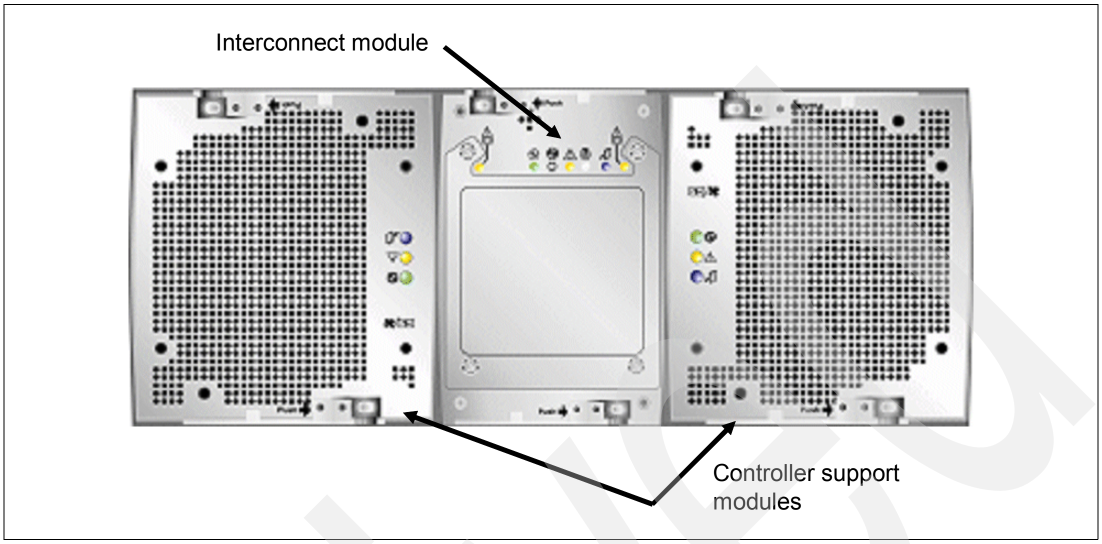


Figure 2-12 DS5000 front view

Interconnect module

The interconnect module is the unit (FRU) between the two controller support modules. The interconnect holds the cache batteries and the hot-swappable midplane. When the interconnect module is removed, the DS5000 automatically suspends controller B, fails over all the LUNs to controller A, and continues to operate.

When the interconnect module is put back in place, the DS5000 can revert back to normal operations. However, you have to manually redistribute the LUNs to their respective owning controller (A or B), since the DS5000 will not automatically redistribute for you.

The interconnect module provides the electrical communication path between the power supply fan units and allows their power supplies to load-share and to charge the cache-backup battery packs. It houses two cache-backup battery packs. Each battery pack contains batteries for both controllers' caches.

For data safety reasons the write-caching is disabled when either one of the backup battery packs fails. You should replace the failed battery pack as soon as possible to minimize any performance impact due to the disabling of the write-caching function.

Power supply fan units

The DS5000 controller enclosure has two removable power supply units that also contain the cooling fans for the entire enclosure.

The fans pull air through the ventilation holes of the PSUs and guide it through the controllers to the back side of the enclosure. The fans provide redundant cooling, which means that if one of the fans in either fan housing fails, the remaining fan continues to provide sufficient cooling to operate the controller enclosure.

The power supplies provide power to the internal components by converting incoming AC voltage to DC voltage. If one power supply is turned off or malfunctions, the other power supply maintains electrical power to the controller enclosure.

To preserve the optimal airflow, do not remove a failed power supply fan CRU from the DS5000 controller enclosure chassis until you are ready to replace it with a new CRU.

Note: Although both power supply fan units (left and right) are identical, they are seated in the DS5000 controller enclosure chassis in opposite orientations. If the power supply fan cannot fully be inserted in the power supply fan bay, flip it 180 degrees and reinsert it.

DS5000 rear view

Figure 2-13 shows the rear of the DS5000 with the two controllers stacked horizontally. Controller A is located on the top and controller B is located on the bottom. Controllers are hot-swappable.

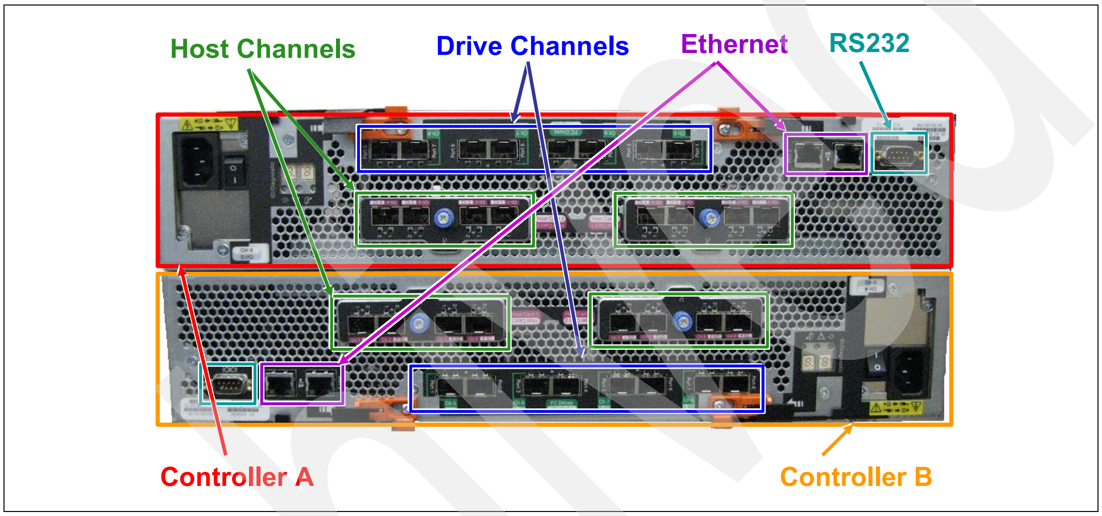


Figure 2-13 DS5000 rear view

Notice that controller A is positioned upside-down relative to controller B. It is important to keep this in mind when connecting the back-end ports to hosts and drive-side expansion enclosures.

Each controller is also equipped with two Ethernet RJ45 connectors and one serial port.

RAID controllers

Each controller includes the following features:

- Eight drive ports
- One RS232 serial port
- Eight host ports
- Dual Ethernet ports
- One power connector and a power switch for this power line

The RS232 Serial Port is a PS2 type port. This serial port is used for management and diagnostic purposes. You can use a PC with a terminal emulation utility such as Hyper Terminal to access the command set.

The dual Ethernet ports accommodate an RJ-45 10BASE-Tx or 100BASE-Tx Ethernet connection. There are two ports per controller. One port is designed for out-of-band management and the other port is meant for serviceability. This feature is similar to DS4800.

Because of the extra port, it is now preferable to have two IP addresses per controller in order to manage and service the DS5000 appropriately. You can still operate the DS5000 with only one IP port active per controller. You can assign either port of the two ports for management or service on a given controller.

The default IP addresses for the controller A Ethernet ports 1 and 2 are 192.168.128.101 and 192.168.129.101, respectively. The default IP addresses for the controller B Ethernet ports 1 and 2 are 192.168.128.102 and 192.168.129.102, respectively. The default subnet mask for all four Ethernet ports is 255.255.255.0.

Important: When changing the default IP addresses, remember that port 1 and 2 of each controller must be in separate subnets.

2.4 Positioning Midrange Storage

IBM has brought together into one family, known as the DS family, a broad range of disk systems to help small to large size enterprises select the correct solutions for their needs. The DS family combines the high-performance IBM System Storage DS8000™ and DS6000™ Series of enterprise servers that inherit from the ESS, the DS4000 and DS5000 series of mid-range systems, and the DS3000 entry-line systems.

The IBM System Storage DS4000 series are the IBM solutions for mid-range storage requirements. The DS4000 Midrange power supply fan units are flanked by the DS5000 Series Storage Servers for upper midrange requirements and the DS3000 Series Storage Servers for entry-level or lower midrange solutions. The positioning of the midrange products that support IBM i attachment via VIOS is shown in Figure 2-14.

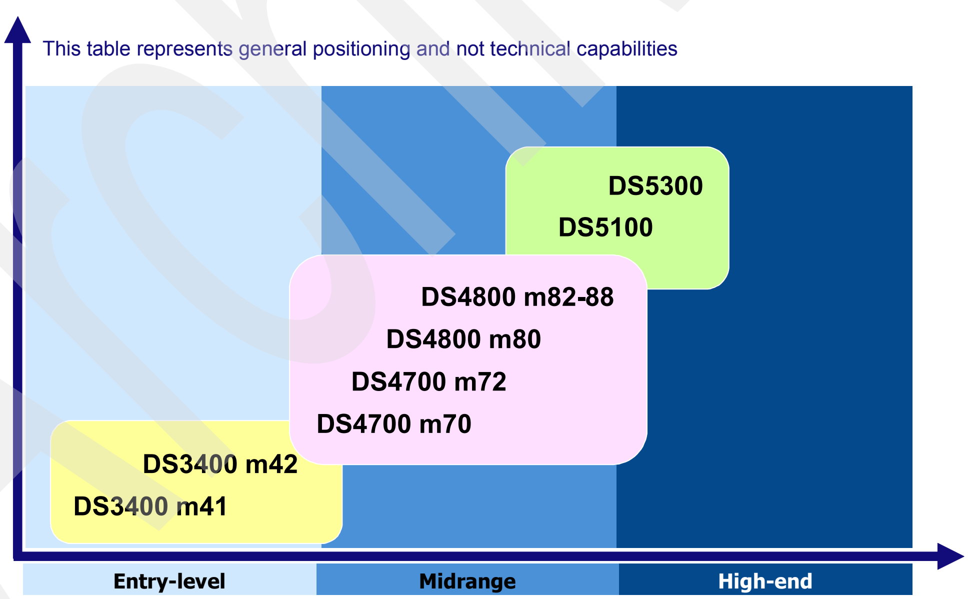


Figure 2-14 Midrange positioning

Within the DS family, the DS4000 and DS5000 series of servers supports both Fibre Channel and Serial disk drives. The DS3000 power supply fan units support serial-attached SCSI and SATA drives. The maximum raw SATA storage capacity of this family is over 256 TB (using 1 TB SATA drives). The maximum raw FC storage capacity is 76.8 TB. As the maximum usable capacity equals the maximum allowed number of drives multiplied with the capacity of the biggest drive available, these values will automatically increase as soon as bigger drives become available.

Table 2-1 highlights the major hardware differences between the DS3400, DS4700, DS4800, and DS5000.

Table 2-1 Hard feature comparison

	DS3400	DS4700	DS4800	DS5000
Machine type-model	1726-42X	1814-70A/72A	1815-8xA	1818-51A/53A
Host ports - max	4 FCP	4/8 FCP	8 FCP	16 FCP
Native host link speed	4 Gbps	4 Gbps	4 Gbps	4 Gbps
Supported host link speeds	1, 2, 4 Gbps	1, 2, 4 Gbps	1, 2, 4 Gbps	1, 2, 4 Gbps
Drive ports/loop pairs	2 SAS/1 loop	4 FC/2 loops	8 FC/4 loops	16 FC/8 loops
Native drive link speed	3 Gbps	4 Gbps	4 Gbps	4 Gbps
Processor type	Intel xScale 667 MHz	Intel xScale 667 MHz	Intel Xeon® 2.4 GHz	Intel Xeon 2.8 GHz
Cache per subsystem (dual controller)	1 GB/2 GB	2 GB/4 GB	4 GB/8 GB/16 GB	4 GB/8 GB/ 16 GB/32 GB/64 GB
Cache hold-up	Battery backed	Battery backed	Battery backed	Permanent
Cache mirroring	Back-end loops	Back-end loops	Back-end loops	Two dedicated busses
Maximum drives	48 (EXP3000)	112 (EXP810) 100 (EXP710)	224	256
Drives types supported	SAS, SATA	FC, SATA	FC, SATA	FC, SATA
FC/SAS drives	73, 146, 300 GB SAS	73, 146, 300 GB FC	73, 146, 300 GB FC	73, 146, 300 GB FC
SATA drives	500, 750 GB, 1TB SATA	500, 750 GB, 1TB SATA	500, 750 GB, 1TB SATA	500, 750 GB, 1TB SATA
Maximum capacity SAS/FC	14.4 TB SAS	33 TB FC	67 TB FC	77 TB FC
Maximum capacity SATA	48 TB SATA	112 TB SATA	224 TB SATA	256 TB SATA
RAID levels supported	0, 1, 3, 5, 10	0, 1, 3, 5, 6, 10	0, 1, 3, 5, 10	0, 1, 3, 5, 6, 10
Hot swappable disks/ components	Yes	Yes	Yes	Yes

The DS3400 has a completely different drive attachment technology compared to the DS4000 and DS5000 series. It uses SAS drive ports, which have a native 3 Gbps speed. This is slightly lower than the 4 Gbps drive speed supported by the FC ports of the DS4000 and DS5000.

The cache per subsystem is the sum of the cache from both controllers. For all models a cache upgrade feature is available. This feature requires the physical replacement of the controllers.

The cache mirroring of the DS3000 and DS4000 series is performed over the drive loops. This can cause an impact on performance, as the available drive loop bandwidth is shared with the host I/O and the host will see increased I/O service times if the drive loop utilization is near or at the physical limits.

The DS5000 features two dedicated internal cache mirroring busses for this purpose. So the complete drive loop bandwidth is available for host I/O on the DS5000. As a result, the DS5000 can handle more I/O on a drive loop as the DS4000, even if the drive loops operate at the same speed.

All models except the DS3400 and the DS4800 feature RAID-6 logical volume redundancy with firmware level 7.x and later. RAID-6 will also be available for the DS3400 with a future firmware release. As the RAID calculation of all models is hardware supported, this requires a special ASIC. The ASIC of the DS4800 does not support RAID-6, so this will probably never be available for this model.

Table 2-2 compares the software features of all models.

Table 2-2 Soft feature comparison

	DS3400	DS4700	DS4800	DS5000
FlashCopy	Yes	Yes	Yes	Yes
Volume Copy	Yes	Yes	Yes	Yes
Remote mirror, synchronous copy, up to 10 am (metro mirror)	No	Yes	Yes	Yes
Remote mirror, asynchronous copy, extended distance (global copy)	No	Yes	Yes	Yes
Remote mirror, asynch copy with write consistency, extended dist. (global mirror)	No	Yes	Yes	Yes
Call Home support	RSM for Storage	RSM for Storage	RSM for Storage	RSM for Storage
Non-call Home support	E-mail alerts, SNMP traps	E-mail alerts, SNMP traps	E-mail alerts, SNMP traps	E-mail alerts, SNMP traps

	DS3400	DS4700	DS4800	DS5000
Concurrent firmware upgrade	Yes (dual controller model only)	Yes	Yes	Yes
Warranty and support	Three-year parts and labor warranty	Three-year parts and labor warranty, 9x5 next business day, upgradeable to 24x7 with four-hour response	Three-year parts and labor warranty, 9x5 next business day, upgradeable to 24x7 with four-hour response	One-year parts and labor warranty, 24x7 with four-hour response Maintenance extension available

All copy features are available for all models. Only Enhanced Remote Mirroring is not available for the DS3000 series.

All systems can easily be configured for e-mail alerts or SNMP traps in case of a critical event or can use IBM Remote Support Manager (RSM) for automatic system call home. RSM and e-mail alerts are discussed in detail in 2.7, “Remote Support Manager” on page 54.

The warranty and support conditions differ between the DS3000, DS4000, and DS5000 to contribute to the different requirements of small, medium, or enterprise environments.

Table 2-3 compares the performance data of all models. As there is a major performance difference between the DS5100 and the DS5300, Table 2-3 shows both submodels.

Table 2-3 Performance Data Comparison

	DS3400	DS4700	DS4800	DS5100	DS5300
Random cache reads (IOPS)	120,486	120,000	575,000	~ 600,000	~ 700,000
Random disk reads (IOPS)	21,637	44,000	86,000	~ 115,000	~ 172,000
Random disk writes (IOPS)	4,651	9,000	22,000	~ 30,000	~ 45,000
Sequential cache reads (MBps)	1,633	1,500	1,700	~ 3,200	~ 6,400
Sequential disk reads (MBps)	945	990	1,600	~ 3,000	~ 6,400
Sequential disk writes (MBps)	726	850	1,300	~ 2,800	~ 5,200

The performance data shown in Table 2-3 were measured in a lab environment under standardized test environments. The numbers above will probably not be reached in a real customer environment, but Table 2-3 gives a good comparison between the performance abilities of each class.

Refer to 4.1, “Planning for IBM Midrange Storage series and models” on page 90, for guidelines on DS Midrange Storage model selection for IBM i specific workloads.

2.5 RAID technology and spare drives

This section discusses the RAID levels that can be chosen when creating a new array on the DS3000, DS4000, and DS5000. It shows the differences between the available RAID levels and the explicit advantages and disadvantages of each RAID level. See also 4.5, “Planning considerations for performance” on page 97.

Note: The DS Midrange Storage System’s RAID technology architectural information provided here is for general reference only. For performance-critical IBM i workload we recommend using RAID-1 or RAID-10 arrays.

RAID levels

The following RAID levels can be selected:

- RAID-0: For performance, but generally not recommended (Figure 2-15)

  RAID-0 is also known as *data striping*. It is well-suited for program libraries requiring rapid loading of large tables or, more generally, applications requiring fast access to read-only data or fast writing. RAID-0 is only designed to increase performance. There is no redundancy, so any disk failures require reloading from backups. Select RAID level 0 for applications that would benefit from the increased performance capabilities of this RAID level. Never use this level for critical applications that require high availability.

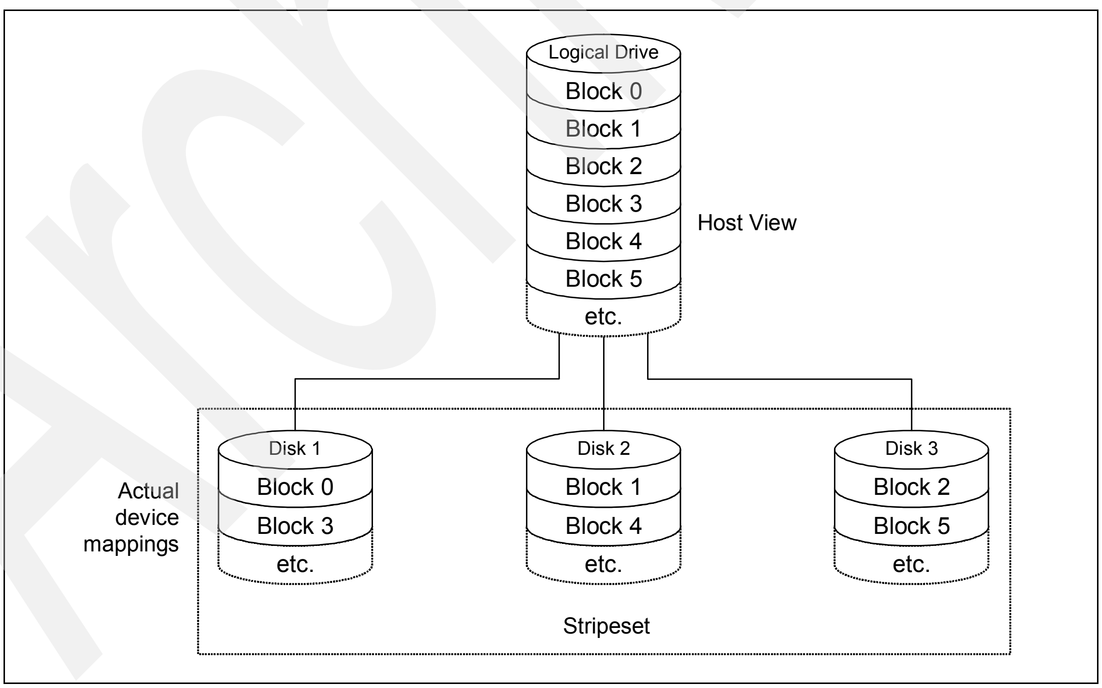


Figure 2-15 RAID-0

- RAID-1: For availability/good read response time (Figure 2-16)

  RAID-1 is also known as disk mirroring. It is most suited to applications that require high data availability, good read response times, and where cost is a secondary issue. The response time for writes can be somewhat slower than for a single disk, depending on the write policy. The writes can either be executed in parallel for speed or serially for safety. Select RAID level 1 for applications with a high percentage of read operations and where the cost is not the major concern. Because the data is mirrored, the capacity of the logical drive when assigned RAID level 1 is 50% of the array capacity.

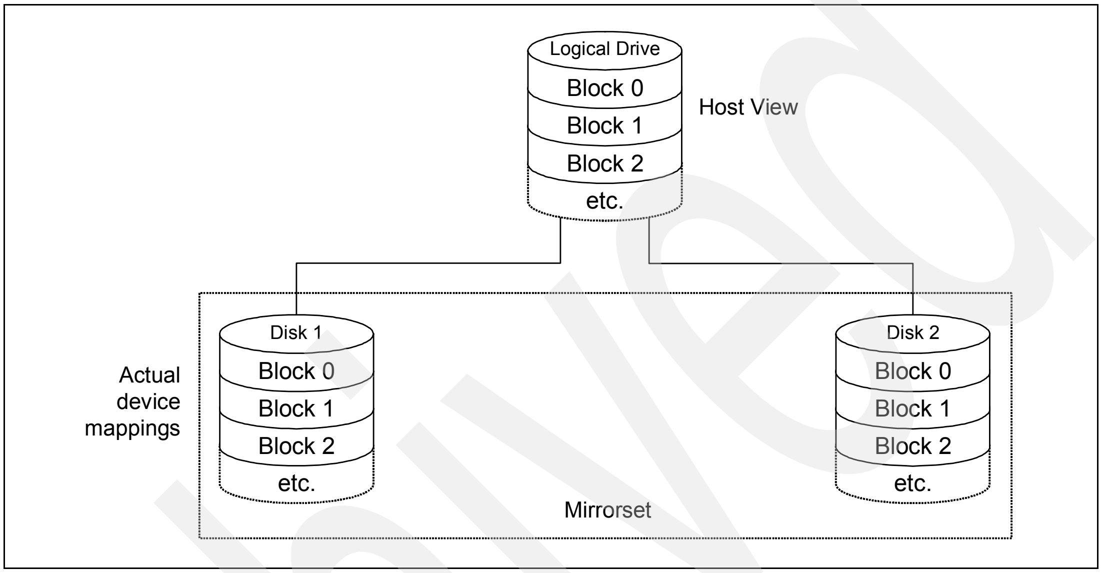


Figure 2-16 RAID-1

  Some recommendations when using RAID-1 include:

  – Use RAID-1 for the disks that contain your operating system. It is a good choice because the operating system can usually fit on one disk.

  – Use RAID-1 for transaction logs. Typically, the database server transaction log can fit on one disk drive. In addition, the transaction log performs mostly sequential writes. Only rollback operations cause reads from the transaction logs. Therefore, we can achieve a high rate of performance by isolating the transaction log on its own RAID-1 array.

  – Use write caching on RAID-1 arrays. Because a RAID-1 write will not complete until both writes have been done (two disks), performance of writes can be improved through the use of a write cache. When using a write cache, be sure that it is battery-backed up.

Note: RAID-1 is actually implemented only as RAID-10 (described below) on the DS4000 products.

- RAID-3: Sequential access to large files (Figure 2-17)

  RAID-3 is a parallel process array mechanism, where all drives in the array operate in unison. Similar to data striping, information to be written to disk is split into chunks (a fixed amount of data), and each chunk is written out to the same physical position on separate disks (in parallel). This architecture requires parity information to be written for each stripe of data.

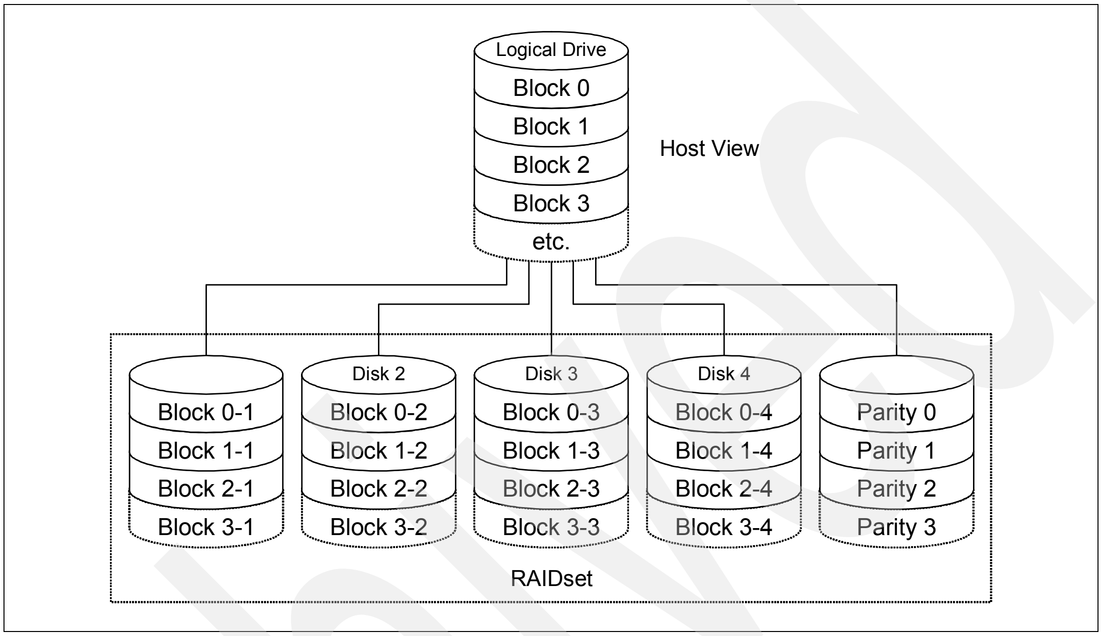


Figure 2-17 RAID-3

Performance is very good for large amounts of data, but poor for small requests because every drive is always involved, and there can be no overlapped or independent operation. RAID-3 is well-suited for large data objects such as CAD/CAM or image files, or applications requiring sequential access to large data files. Select RAID-3 for applications that process large blocks of data. It provides redundancy without the high overhead incurred by mirroring in RAID-1.

- RAID-5: High availability and fewer writes than reads (Figure 2-18)

  RAID level 5 stripes data and parity across all drives in the array. RAID level 5 offers both data protection and increased throughput. When you assign RAID-5 to an array, the capacity of the array is reduced by the capacity of one drive (for data-parity storage). RAID-5 gives you higher capacity than RAID-1, but RAID level 1 offers better performance.

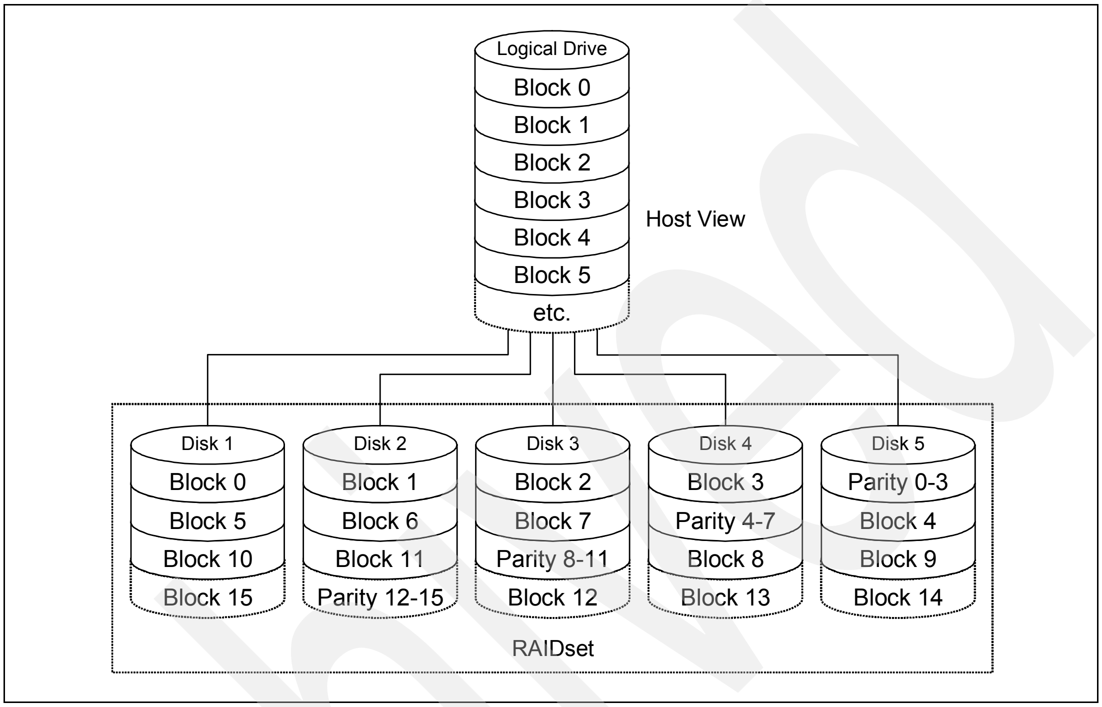


Figure 2-18 RAID-5

  RAID-5 is best used in environments requiring high availability with fewer writes than reads.

  RAID-5 is good for multi-user environments, such as database or file system storage, where typical I/O size is small, and there is a high proportion of read activity. Applications with a low read percentage (write-intensive) do not perform as well on RAID-5 logical drives because of the way that a controller writes data and redundancy data to the drives in a RAID-5 array. If there is a low percentage of read activity relative to write activity, consider changing the RAID level of an array for faster performance.

  Use write caching on RAID-5 arrays, because RAID-5 writes will not be completed until at least two reads and two writes (each for parity and data) have occurred. The response time of writes will be improved through the use of write cache (be sure that it is battery-backed up). RAID-5 arrays with caching can give as good a performance as any other RAID level and with some workloads the striping effect gives better performance than RAID-1.

- RAID-6: High availability with additional fault tolerance (Figure 2-19 on page 40)

  RAID-6 is a RAID level employing n+2 drives, which can survive the failure of any two drives.

  RAID-6 is available with firmware V7.10 and is only supported on the DS4700 and DS5000. RAID-6 stripes blocks of data and parity across an array of drives and it calculates two sets of information for each block of data (p+q). For the purposes of RAID-6 p+q, they can be used to generate up to two missing values from a set of data.The key to this method is the q, which is a codeword based upon Reed-Solomon error correction. As such, q is more like a CRC than parity. Based upon principles of set theory and linear algebra, Reed-Solomon codes are well-known codes that are also maximum distance separable (MDS).

The calculation of q is complex. In the case of the DS4000, this calculation is made by the hardware and thus performs better than the software-based implementation found in other Storage Systems.

By storing two sets of distributed parities, RAID-6 is designed to tolerate two simultaneous disk failures. This is a good implementation for environments using SATA disks.

Due to the added overhead of more parity calculations, in terms of writing data, RAID-6 is slower than RAID-5 but may be faster in random reads thanks to the spreading of data over one more disks.

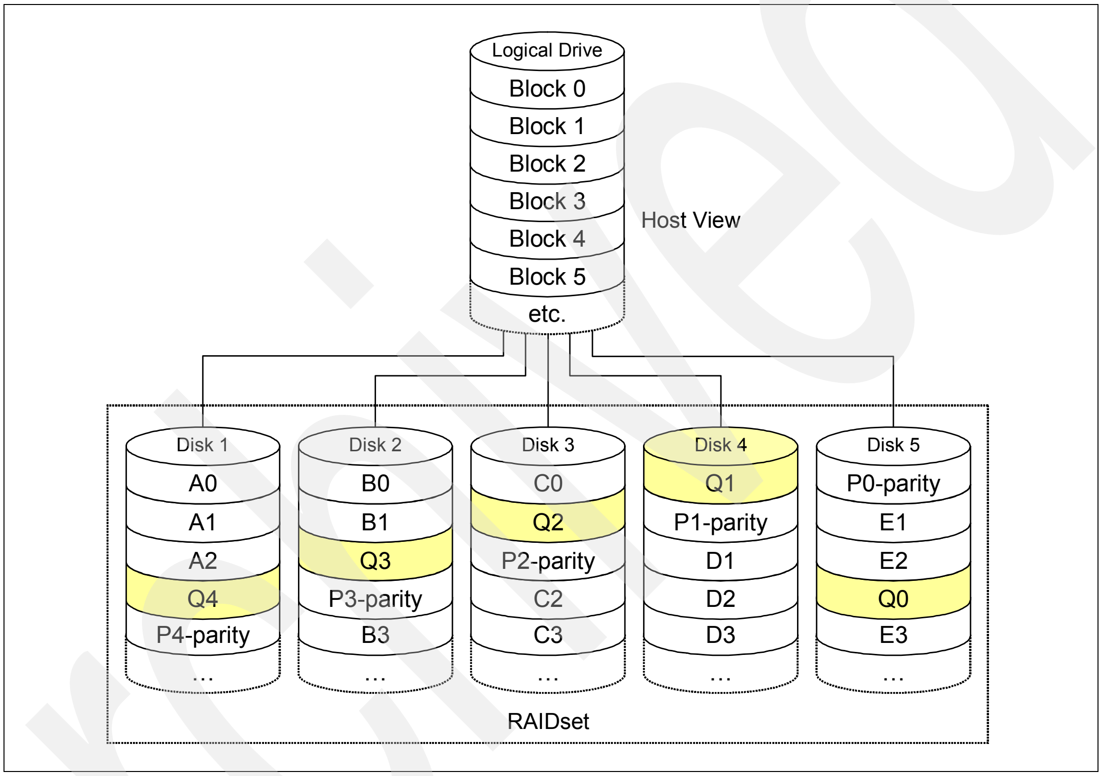


Figure 2-19 RAID-6

The q value allows the system to calculate up to two missing values (if you lost two disks) and can be used to recover the missing data.

> **Note:** RAID-6 is currently supported only on DS5000 and DS4700 with firmware Version 7.10 or later.

RAID-6 is good for multi-user environments, such as database or file system storage, where typical I/O size is small, and there is a high proportion of read activity, and in situations where additional fault tolerance is required.

- RAID-10: Higher performance than RAID-1 (Figure 2-20)

  RAID-10, also known as RAID-1+0, implements block-interleave data striping and mirroring. In RAID-10, data is striped across multiple disk drives, and then those drives are mirrored to another set of drives.

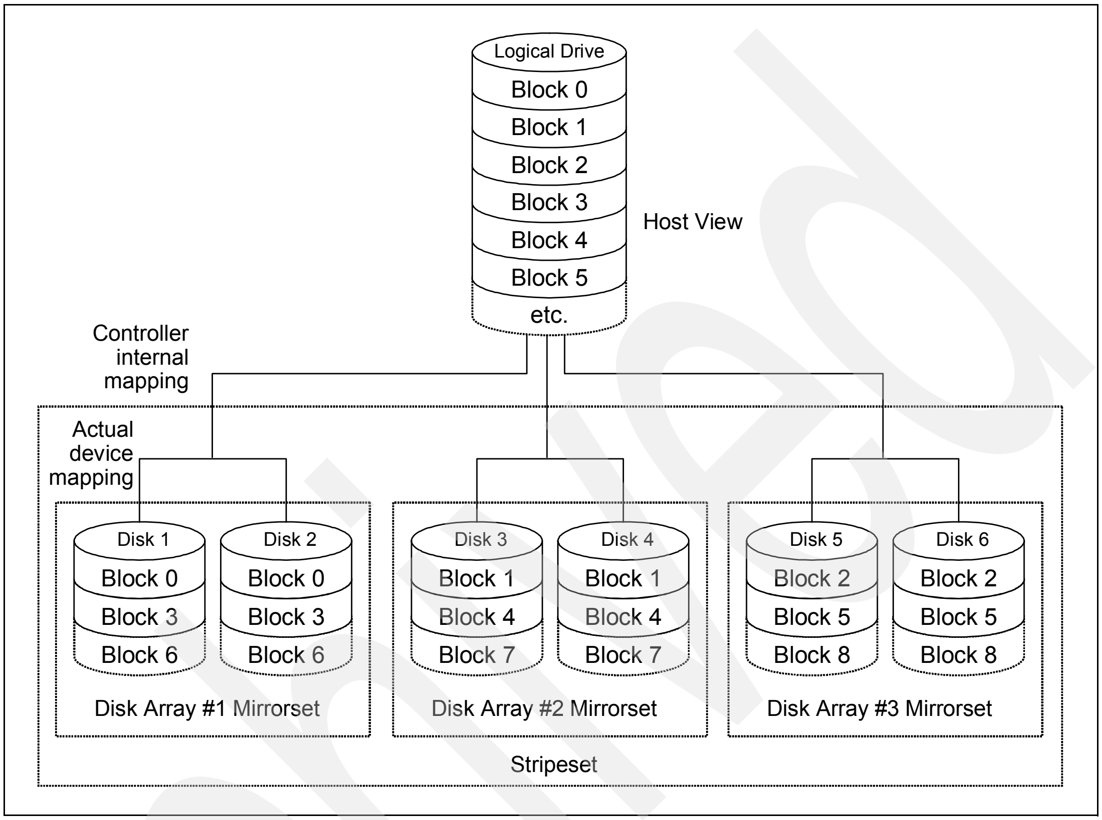


Figure 2-20 RAID-10

The performance of RAID-10 is approximately the same as RAID-0 for sequential I/Os. RAID-10 provides an enhanced feature for disk mirroring that stripes data and copies the data across all the drives of the array. The first stripe is the data stripe. The second stripe is the mirror (copy) of the first data stripe, but it is shifted over one drive. Because the data is mirrored, the capacity of the logical drive is 50% of the physical capacity of the hard disk drives in the array.

Our recommendations for using RAID-10 are:

- Use RAID-10 whenever the array experiences more than 10% writes. RAID-5 does not perform as well as RAID-10 with a large number of writes.
- Use RAID-10 when performance is critical. Use write caching on RAID-10. Because RAID-10 write will not be completed until both writes have been done, the performance of the writes can be improved through the use of a write cache (be sure that it is battery-backed up).

When comparing RAID-10 to RAID-5:

- RAID-10 writes a single block through two writes. RAID-5 requires two reads (read original data and parity) and two writes. Random writes are significantly faster on RAID-10.
- RAID-10 rebuilds take less time than RAID-5 rebuilds. If a real disk fails, RAID-10 rebuilds it by copying all the data on the mirrored disk to a spare. RAID-5 rebuilds a failed disk by merging the contents of the surviving disks in an array and writing the result to a spare.

RAID-10 is the best fault-tolerant solution in terms of protection and performance, but it comes at a cost. You must purchase twice the number of disks that are necessary with RAID-0.

Based on the respective level, RAID offers the following performance results:

- RAID-0 offers high performance, but does not provide any data redundancy.
- RAID-1 offers high performance for write-intensive applications.
- RAID-3 is good for large data transfers in applications, such as multimedia or medical imaging, that write and read large, sequential chunks of data.
- RAID-5 is good for multi-user environments, such as database or file system storage, where the typical I/O size is small, and there is a high proportion of read activity.
- RAID-6 is good for multi-user environments, such as database or file system storage, where the typical I/O size is small and there is a high proportion of read activity, and in situations where additional fault tolerance is required.
- RAID-10 offers higher performance than RAID-1 and more reliability than RAID-5.

See Table 2-4.

Table 2-4 RAID levels

RAID	Description	APP	Advantage	Disadvantage
0	Stripes data across multiple drives.	IOPS Mbps	Performance due to parallel operation of the access.	No redundancy. One drive fails, data is lost.
1	Disk's data is mirrored to another drive.	IOPS	Performance as multiple requests can be fulfilled simultaneously.	Storage costs are doubled.
10	Data is striped across multiple drives and mirrored to same number of disks.	IOPS	Performance as multiple requests can be fulfilled simultaneously. Most reliable RAID level on the DS4000 and recommended for IBM i performance critical workload.	Storage costs are doubled.
3	Drives operated independently with data blocks distributed among all drives. Parity is written to a dedicated drive.	Mbps	High performance for large, sequentially accessed files (image, video, graphical).	Degraded performance with 8–9 I/O threads, random IOPS, smaller more numerous IOPS.

RAID	Description	APP	Advantage	Disadvantage
5	Drives operated independently with data and parity blocks distributed across all drives in the group.	IOPS Mbps	Good for reads, small IOPS, many concurrent IOPS, and random I/Os.	Writes are particularly demanding.
6	Stripes blocks of data and parity across an array of drives and calculates two sets of parity information for each block of data.	IOPS Mbps	Good for multi-user environments, such as database or file system storage, where typical I/O size is small, and in situations where additional fault tolerance is required.	Slower in writing data, complex raid controller architecture.

Note: The IBM DS Midrange Storage DS3000, DS4000, and DS5000 series allows you to dynamically migrate an existing array from any RAID level to another. This operation can be performed concurrently with host I/O, but will degrade system performance during the RAID reconfiguration.

RAID reliability considerations

At first glance both RAID-3 and RAID-5 appear to provide excellent protection against drive failure. With today's high-reliability drives, it appears unlikely that a second drive in an array would fail (causing data loss) before an initial failed drive could be replaced. However, field experience has shown that when a RAID-3 or RAID-5 array fails, it is not usually due to two drives in the array experiencing complete failure. Instead, most failures are caused by one drive going bad and a single block somewhere else in the array that cannot be read reliably. This problem is exacerbated by using large arrays with RAID-5. This stripe kill can lead to data loss when the information to re-build the stripe is not available. The end effect of this issue will of course depend on the type of data and how sensitive it is to corruption. While most Storage Systems (including the DS4000 series) have mechanisms in place to try to prevent this from happening, they cannot work 100% of the time.

Any selection of RAID type should take into account the cost of downtime. Simple math tells us that RAID-3 and RAID-5 are going to suffer from failures more often than RAID-10. (Exactly how often is subject to many variables and is beyond the scope of this book.) The money saved by economizing on drives can be easily overwhelmed by the business cost of a crucial application going down until it can be restored from backup.

Naturally, no data protection method is 100% reliable, and even if RAID were 100% solid, it would not protect your data from accidental corruption or deletion by program error or operator error. Therefore, all crucial data should be backed up by appropriate software, according to business needs.

Hotspare drives

Hot spare disk drives provide additional protection that can be essential in case of a disk drive failure in a fault-tolerant array (RAID-1, 3, 5, or 6). A hot-spare drive is like a replacement drive installed in advance.

A hot spare drive defined on the DS3000, DS4000, and DS5000 power supply fan unit is always used as a so-called *global hot spare*. That is, a hot spare drive can always be used for a failed drive. It does not matter in which array or storage enclosure the hot spare is located. The spare and failed drives must, however, be of the same type (FC or SATA).

Previously, a maximum of up to 15 hot spare drives could be specified. With Storage Manager V10.10 (and firmware V7.10) you can now have an unlimited amount of hot spare drives.

Note: There is no definitive recommendation as to how many hot spares you should install, but it is common practice to use a ratio of one hot spare for two fully populated expansion enclosures.

Three methods exist to allocate hot spare drives in the storage subsystem:

- Automatic assignment: The storage subsystem automatically calculates the number of hot spare drives needed and allocates accordingly. This can be used on unconfigured storage subsystems.
- Automatic assignment by specification of the number of drives: The number of hot spare drives desired is specified and the storage subsystem automatically allocates them.
- Explicit assignment: The hot spare drives are manually selected and assigned.

With Storage Manager V10.10 (firmware V7.10), a new feature is available that allows you to select the preferred recovery option for data that has been written to the hot spare drive.

Three scenarios are presented:

- Once the failed drive is replaced with a new drive, the new replacement drive becomes an unassigned drive. This drive can then be selected as a replacement drive for the data stored on the hot spare drive, and data is copied back to the replaced drive. The original hot spare drive that is now in use becomes a free hot spare drive again. The location of a hot spare drive in this case is fixed and does not wander if it is used.
- Once the failed drive is replaced with a new drive, the new replacement drive becomes an unassigned drive. Any unassigned drive in the storage subsystem can now be selected as a replacement drive for the data stored on the hot spare drive, and data is copied back to this selected replacement drive. The original hot spare drive that is now in use becomes a free hot spare drive again. The location of a hot spare drive in this case is fixed and does not wander if it is used.
- Once the failed drive is replaced with a new drive, the new replacement drive becomes an unassigned drive. The original hot spare drive can be selected as the replacement drive, and no data copy occurs. The original hot spare drive becomes a permanent member of the array. The location of the hot spare drive in this case is not fixed. A new hot spare drive can now be assigned from the unassigned disk drive pool.

Important: Remember to allocate a new hot spare drive after a hot spare drive has been made a permanent member of an array.

A hot spare drive must be of the same type (FC or SATA) and at least of the capacity of the configured space on the failed drive (if the failed disk drive is larger than the hot spare, reconstruction is not possible). The DS4000 power supply fan unit can use a larger drive to recover a smaller failed drive to it. Then the remaining capacity is blocked.

Note: We recommend having your hot spare drives distributed across different disk drive channels to avoid the risk of a single disk drive channel failure causing loss of access to all hot spare drives in the storage subsystem.

The DS4000 attempts to find a hot spare drive in the enclosure with the failed drive first. It finds a drive that is at least the same size as the failed drive, but not necessarily giving preference to one the exact same size as the failed drive. If a match does not exist in the same enclosure, it looks for spares in the other enclosures that contain sufficient capacity to spare over.

The controller uses a free hot spare drive as soon as it finds one, even if there is another one that might be closer to the failed drive.

2.6 User interfaces

There are three different user interfaces available to manage the DS3000, DS4000, and DS5000 power supply fan units.

The first one is the graphical user interface (GUI), which is called Storage Manager. This provides a convenient and self-explanatory way to manage the storage server. DS4000 and DS5000 use the same Storage Manager. The DS3000 Storage Manager is a simplified version of the DS4000 Storage Manager and supports the basic functions of the DS4000 Storage Manager.

The second user interface is the Storage Manager Script Editor Tool. All functions of the DS4000 Storage Manager can be performed with the script command. Some specialized additional tasks are only available as the script command and cannot be done with the Storage Manager. The script commands for DS3000, DS4000, and DS5000 have identical syntax and functionality.

The third user interface is the SMcli command-line interface. It uses the same commands as the Storage Manager Script Tool, but it does not require the Storage Manager application to be started. In addition, it is able to run large command lists that are saved to a text file. This can be helpful for repeating tasks, that is, the creation of multiple volumes or the automated start of a copy service.

The following sections discuss the three user interfaces in detail.

2.6.1 Storage Manager

The DS4000 Storage Manager software is used primarily to configure RAID arrays and logical drives, assign logical drives to hosts, replace and rebuild failed disk drives, expand the size of the arrays and logical drives, and convert from one RAID level to another. It allows troubleshooting and management tasks, like checking the status of the Storage Server components, updating the firmware of the RAID controllers, and managing the Storage Server. Finally, it offers advanced functions such as FlashCopy, Volume Copy, and Enhanced Remote Mirroring.

The Storage Manager can be downloaded free of charge from the IBM Support Web page:

https://www-304.ibm.com/systems/support/supportsite.wss/selectproduct?taskind=2&brandind=5000028&familyind=5329626&typeind=0&modelind=0&osind=0&psid=dm&continue.x=1&matrix=Y#Storage%20Manager

The Storage Manager for DS4000 can manage DS3000, DS4000, and DS5000 Storage Servers. The Storage Manager for DS3000 can manage DS3000 systems only.

Note: If you want to upgrade your Storage Manager to a newer version, do not uninstall the current Storage Manager. Install the new Storage Manager at the same location of the current Storage Manager. This way all links to the existing storage servers are immediately available when you start the new Storage Manager for the first time. New Storage Manager versions are always downward compatible, so you also can manage Storage Systems that do not have the latest firmware installed to the controllers.

When opening the Storage Manager, the first window is the Enterprise Management view, which shows all DS3000, DS4000, and DS5000 systems that were added before and the system status of all storage servers (Figure 2-21).

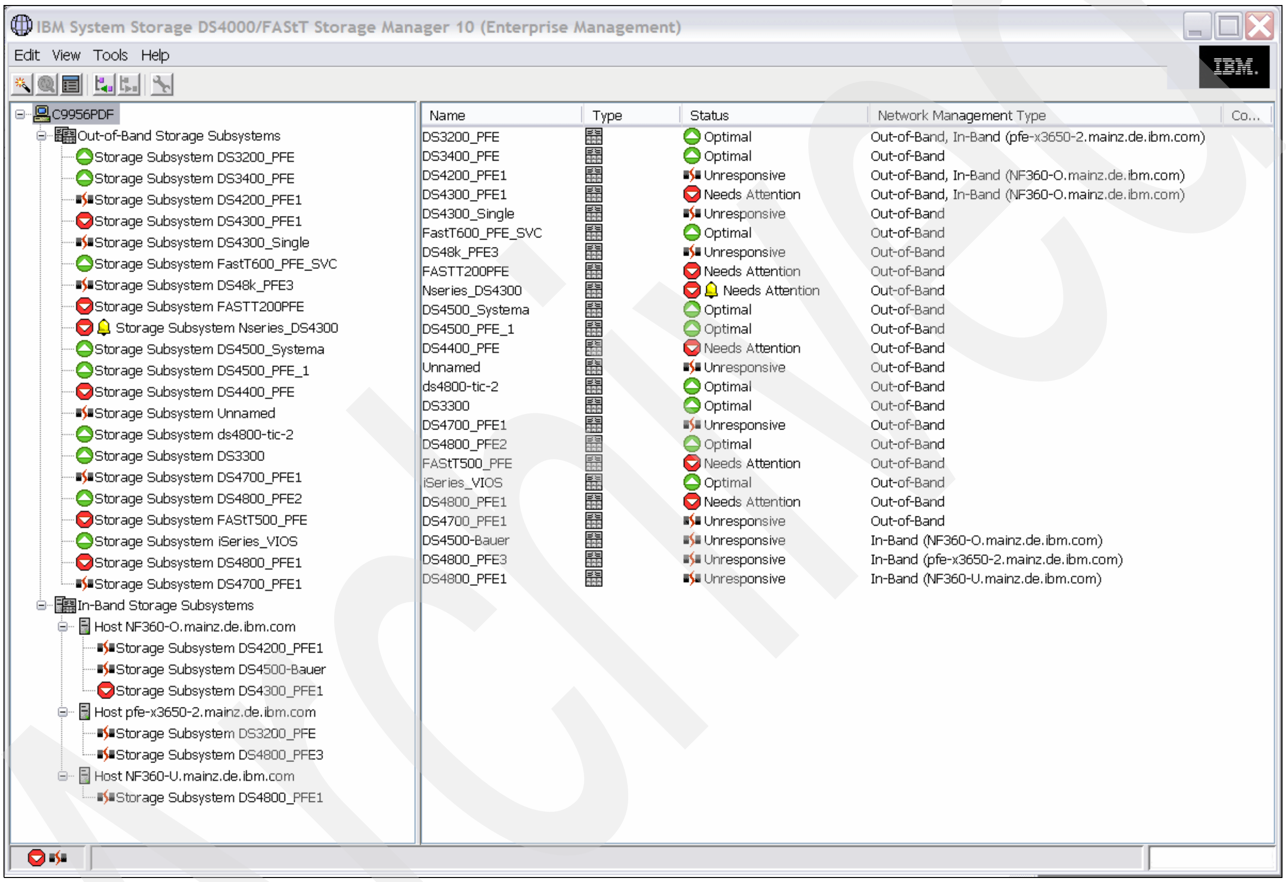


Figure 2-21 Storage Manager 10.15 Enterprise Management view

Systems can be added as out-of-band management, where the Storage Managercommunicates with the system over an established Ethernet connection, or as inband management, where the system communication takes place over the host Fibre Channel connections.

Double-clicking one of the listed Storage Systems will open the subsystem management window for this system. The subsystem management view is different for the DS3000, DS4000, and DS5000 Storage Systems.

Figure 2-22 shows the subsystem management view for the DS4000 and DS5000.

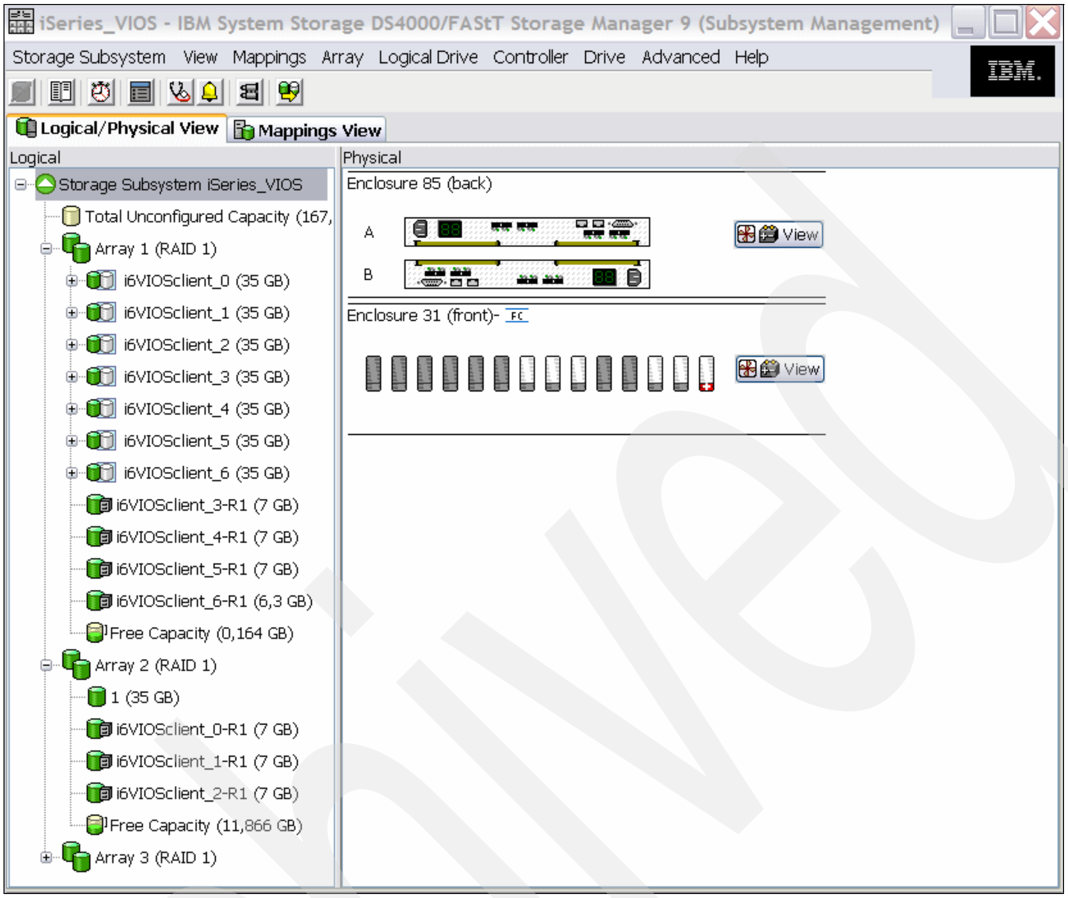


Figure 2-22 Storage Manager 10.15: Subsystem Management view of DS4800

All management tasks can be performed in this view and a detailed system status on all sub components is available.

Figure 2-23 shows the subsystem management view for the DS3000.

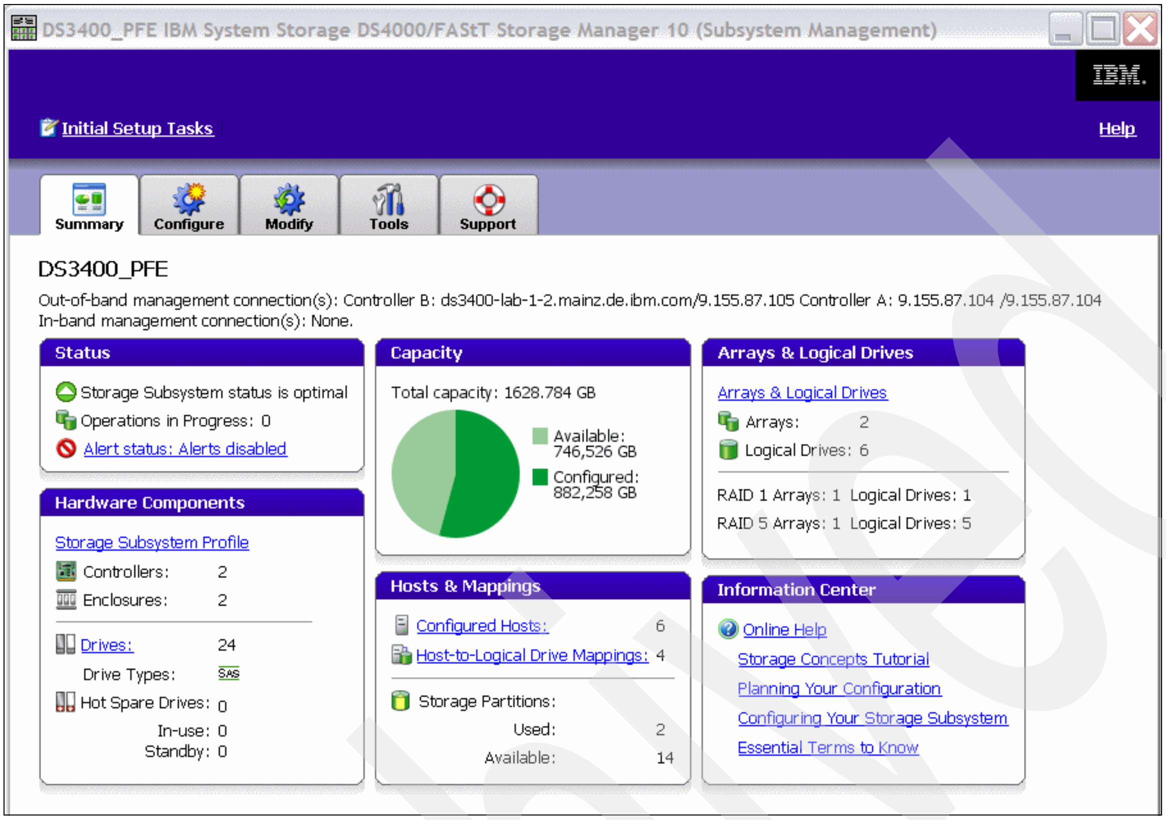


Figure 2-23 Storage Manager 10.15: DS3400 subsystem management view

The most important management tasks, like array, volume, and host management, can be performed in this view, and a detailed system status on all sub components is available.

All management tasks that are not available in the subsystem management window can be performed as script command. This is described in 2.6.2, “Script Editor” on page 49.

For examples using the Storage Manager GUI for configuring storage for IBM i refer to 7.3, “DS3400 Storage Configuration with the GUI” on page 228, and 7.4, “DS4000/DS5000 Storage configuration with the GUI” on page 252.

For more information about installing and using the Storage Manager refer to the *IBM System Storage DS4000 Storage Manager v10.10 Installation and Host Guide for AIX, HP-UX, Solaris, and Linux on POWER*, GC27-2170-00:

https://www-304.ibm.com/systems/support/supportsite.wss/docdisplay?lndocid=MIGR-61177&brandind=5000028

2.6.2 Script Editor

The Script Editor is a powerful tool to easily perform repeating, automated, or specific management tasks. To open the script editor, go to the Storage Manager enterprise management window, right-click the Storage System and select **Tools** → **Execute Scripts** (Figure 2-24).

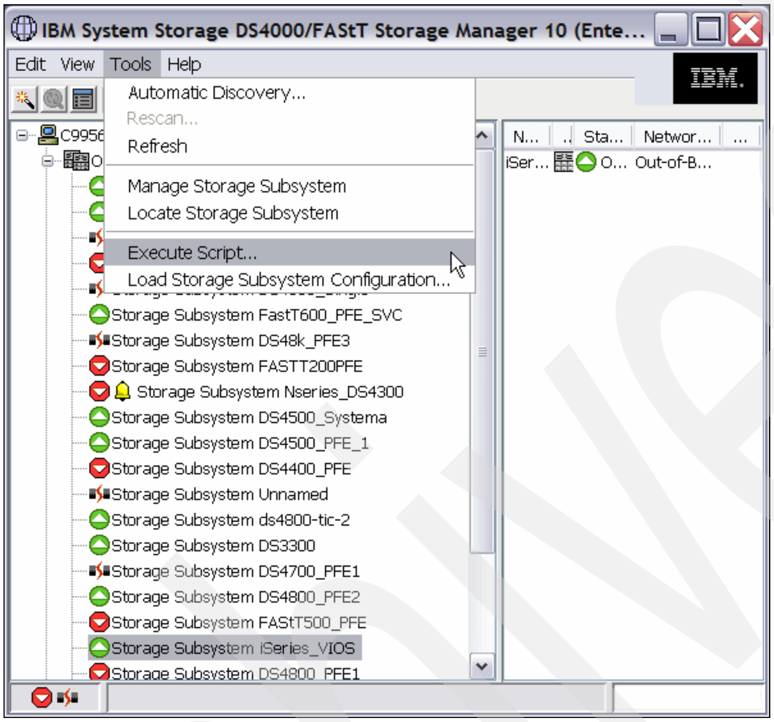


Figure 2-24 Start Script Editor from Enterprise Management Window

A new window opens that is divided into the script input window at the top and the script output window at the bottom. Insert the script command in the upper window or load a text file (**File** → **Load Script**) that contains the script and select **Tools** → **Verify and Execute**. After the script has finished, the result is shown in the lower part of the window (Figure 2-25).

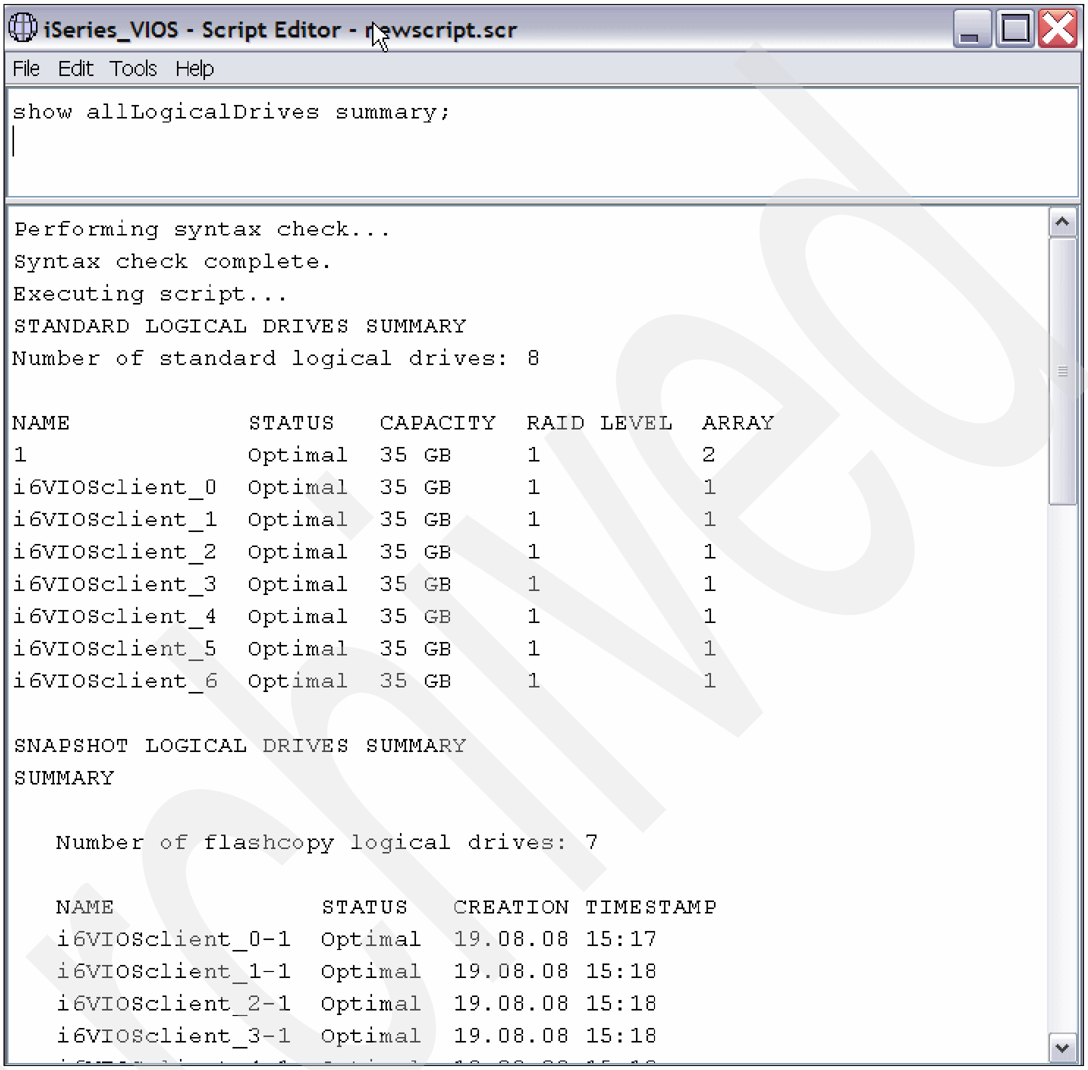


Figure 2-25 Storage Manager Script Editor

All script commands that can be executed with the Script Editor can also be run with the SMcli command-line interface, which we describe in the following section.

> **Note:** A complete list of all available script commands including syntax can be found under **Help** → **Command Reference** of the script editor window.

2.6.3 SMcli

The command-line interface (CLI) is a software tool that lets storage subsystem installers, developers, and engineers configure and monitor storage subsystems. Using the CLI, you can run commands from an operating system prompt, such as the Windows command prompt, a Linux operating system console, or a Solaris operating system console.

Each command performs a specific action for managing a storage subsystem or returning information about the status of a storage subsystem. You can enter individual commands or you can run script files when you need to perform operations more than once. For example, you can run script files when you want to install the same configuration on several storage subsystems. The CLI lets you load a script file from a disk and run the script file. The CLI provides a way to run storage management commands on more than one network storage subsystem.

The commands that you run on the CLI provide access to the script engine, specify the storage subsystem to receive the script commands, and set operation environment parameters.

A CLI command consists of these elements:

- The term SMcli
- The storage subsystem identifier
- Parameters
- Script commands

A CLI command takes this form:

```
SMcli storageSubsystem parameters script-commands;
```

- SMcli invokes the command line interface.

> **Note:** If you issue the command from the directory or folder that does not contain the SMcli.exe program, you need to include the appropriate path: *directoryName/SMcli*.

- *storageSubsystem* is the name or both IP addresses of the storage subsystem.
- *parameters* are CLI parameters that define the environment and the purpose for the command.
- *script-commands* are one or more script commands or the name of a script file that contains script commands. (The script commands are the storage subsystem configuration commands.)

> **Note:** Script commands must always end with a semicolon (;). For multiple commands (that is, in a batch file) each single command must end with a semicolon.

If you enter SMcli and a storage subsystem name but do not specify CLI parameters, script commands, or a script file, the command-line interface runs in interactive mode. Interactive mode lets you run individual commands without prefixing the commands with SMcli. In interactive mode, you can enter a single command, view the results, and enter the next command without typing the complete SMcli string. Interactive mode is useful for determining configuration errors and quickly testing configuration changes. Example 2-1shows the Interactive SMcli mode with the command:

```
show allLogicalDrives summary;
```

Example 2-1 SMcli Interactive mode

```
C:\Program Files\IBM_DS4000\client>SMcli 9.155.86.22 9.155.86.23
Entering interactive mode. Please type desired command.

show allLogicalDrives summary;
STANDARD LOGICAL DRIVES SUMMARY
Number of standard logical drives: 8
```

```
NAME            STATUS   CAPACITY  RAID LEVEL  ARRAY
1               Optimal  35 GB     1           2
i6VIOSclient_0  Optimal  35 GB     1           1
i6VIOSclient_1  Optimal  35 GB     1           1
i6VIOSclient_2  Optimal  35 GB     1           1
i6VIOSclient_3  Optimal  35 GB     1           1
i6VIOSclient_4  Optimal  35 GB     1           1
i6VIOSclient_5  Optimal  35 GB     1           1
i6VIOSclient_6  Optimal  35 GB     1           1

SNAPSHOT LOGICAL DRIVES SUMMARY
SUMMARY

   Number of flashcopy logical drives: 7

   NAME              STATUS   CREATION TIMESTAMP
   i6VIOSclient_0-1  Optimal  8/19/08 6:53 PM
   i6VIOSclient_1-1  Optimal  8/19/08 6:53 PM
   i6VIOSclient_2-1  Optimal  8/19/08 6:53 PM
   i6VIOSclient_3-1  Optimal  8/19/08 6:53 PM
   i6VIOSclient_4-1  Optimal  8/19/08 6:54 PM
   i6VIOSclient_5-1  Optimal  8/19/08 6:54 PM
   i6VIOSclient_6-1  Optimal  8/19/08 6:54 PM


SNAPSHOT REPOSITORY LOGICAL DRIVES SUMMARY
Number of flashcopy repositories: 7

NAME               CAPACITY USAGE(%)  THRESHOLD WARNING  FULL POLICY

i6VIOSclient_0-R1  Not Available      50% full           Fail flashcopy logical
drive
i6VIOSclient_1-R1  Not Available      50% full           Fail flashcopy logical
drive
i6VIOSclient_2-R1  Not Available      50% full           Fail flashcopy logical
drive
i6VIOSclient_3-R1  Not Available      50% full           Fail flashcopy logical
drive
i6VIOSclient_4-R1  Not Available      50% full           Fail flashcopy logical
drive
i6VIOSclient_5-R1  Not Available      50% full           Fail flashcopy logical
drive
i6VIOSclient_6-R1  Not Available      50% full           Fail flashcopy logical
drive

MIRROR LOGICAL DRIVES SUMMARY
Number of mirrored pairs: 7 of 64 used
Write consistency group: 7

P = Primary logical drive
S = Secondary logical drive

LOCAL LOGICAL DRIVE  REMOTE                      STATUS
i6VIOSclient_0 (P)   i6VIOSclient_0_mirror (S)   Suspended
```

```
i6VIOSclient_1 (P)   i6VIOSclient_1_mirror (S)  Suspended
i6VIOSclient_2 (P)   i6VIOSclient_2_mirror (S)  Suspended
i6VIOSclient_3 (P)   i6VIOSclient_3_mirror (S)  Suspended
i6VIOSclient_4 (P)   i6VIOSclient_4_mirror (S)  Suspended
i6VIOSclient_5 (P)   i6VIOSclient_5_mirror (S)  Suspended
i6VIOSclient_6 (P)   i6VIOSclient_6_mirror (S)  Suspended

MIRROR REPOSITORY LOGICAL DRIVES SUMMARY
   Number of mirror repositories: 2

   NAME                STATUS   CAPACITY  RAID LEVEL  ARRAY
   Mirror Repository 2 Optimal  0.126 GB  1           3
   Mirror Repository 1 Optimal  0.126 GB  1           3
```

Single commands can also be sent to the SMcli using the -c parameter. The command must be set between double quotation marks (“). Example 2-2 shows how to send single commands to the SMcli.

Example 2-2 Single commands with SMcli

```
C:\Program Files\IBM_DS4000\client>SMcli 9.155.86.22 9.155.86.23 -c “show
allLogicalDrives summary;”
```

For multiple commands we recommend using batch files. You can add an unlimited number of commands and save them in a text file using the Windows WordPad or any other editor. Then run the SMcli with the -f parameter and add the file name instead of the single commands. In Example 2-3 the command is stored to the file System_Summary.txt.

Example 2-3 SMcli for batch files

```
C:\Program Files\IBM_DS4000\client>SMcli 9.155.86.22 9.155.86.23 -f
“c:\temp\System_Summary.txt”
```

All available script commands can be found in the Storage Manager GUI help function.

> **Attention:** IBM recommends using the Storage Manager client GUI to manage your storage subsystems. The command-line interface does not have any mechanisms to prevent you from inadvertently making unwanted changes to the storage subsystem. Because the script commands are capable of damaging a configuration and causing loss of data access if not used correctly, IBM recommends using the Storage Manager client GUI to manage your storage subsystem configurations.

For examples using SMcli for configuring IBM Midrange Storage for IBM i refer to 7.5, “DS storage configuration with SMcli” on page 278.

For more details regarding the usage of the SMcli refer to *Command Line Interface and Script Commands Programming Guide for IBM System Storage DS3000 and DS4000*, GC52-1275-00:

https://www-304.ibm.com/systems/support/supportsite.wss/docdisplay?lndocid=MIGR-5076792&brandind=5000028

2.7 Remote Support Manager

The IBM Remote Support Manager for Storage (RSM for Storage) application is a no-charge application that provides problem reporting and remote access for IBM Service for the DS3000, DS4000, and DS5000 families.

The problem reporting provided by the RSM application automatically creates an entry in IBM call management system for each storage subsystem that reports a problem. Monitoring of storage subsystems is performed by your existing IBM Storage Manager application, which is configured to send SNMP traps to the Remote Support Manager when critical events are detected.

RSM must be installed on a dedicated workstation. IBM service personnel can remotely access the RSM server through a modem connection and a command-line interface, but only after access has been authorized.

Remote users cannot change any of the security features. An internal firewall managed by RSM isolates the server and other devices from local and remote users.

Files and logs needed for problem determination are sent to IBM using e-mail or an ftp connection using the RSM server Ethernet interface.

One RSM for Storage server can support up to 50 DS5000, DS4000, DS3000, FAStT 200, and 500 storage subsystems. Only IP connectivity to the Ethernet management ports of the subsystems is required. Serial cables are not needed for access to the disk subsystems. We recommend separating subsystems to be supported by an RSM for Storage System by country, and they must be covered under warranty or a current maintenance contract.

Note: There are no annual fees for RSM.

2.7.1 RSM and Alert Manager

RSM replaces DS4000 Alert Manager, which is no longer available and will replace DS4000 Service Alert, which is still available for a limited time.

Compared to DS4000 Alert Manager, RSM has several advantages:

- It provides IBM Service with more detailed information for each alert and improves response time.
- You are in control of the remote access interface and receive notification when remote users (IBM support) connect to the RSM system and when any changes are made that affect security of the system.
- One RSM for Storage System can support up to 50 DS4000 and FAStT 200 and 500 storage subsystems with IP connectivity to the Ethernet management ports of the subsystems. However, we recommend that subsystems be supported by an RSM for Storage System in the same country, and they must be covered under warranty or a current maintenance contract.
- Requires only Ethernet connection to the storage subsystems. No serial cables are required.
- There are no annual fees.

2.7.2 Hardware and software requirements

RSM for Storage has the following hardware and software requirements.

Hardware requirements

The RSM for Storage application is designed to run on an IBM xSeries® server. It has been tested with and is currently supported on the following xSeries servers:

- x3250 4364
- x306m 8849

It has these options:

- 512 MB memory.
- 80 GB hard disk drive.
- Serial port: This must be on the system board. The serial port on the Remote Supervisor Adapter (RSA) cannot be accessed by RSM.
- Ethernet port: Note that if your SAN devices are on a private management LAN, a second Ethernet port for accessing your company's SMTP server and the Internet will be required if your selected server has only a single Ethernet port.
- IBM 10/100/1000 Base-TX Ethernet PCI-X Adapter (52P8642) or equivalent.

The RSM for Storage application is designed to work with an external modem attached to the first serial port. The functional requirements for the modem are minimal, and most *Hayes-compatible* external modems can be used.

The RSM for Storage application has been tested with the following modems:

- Multitech Multimodem II MT5600BA
- Multitech MultiModem ZBA MT5634ZBA
- US Robotics Courier 56 K Business Modem USR3453B

Software requirements

The RSM for Storage requires the following prerequisite software:

- IBM Storage Manager V9.16 or later (latest version is recommended) with Event Monitor installed in a management station in a different server.
- Storage subsystems with controller firmware supported by the Storage Manager 9.16 or later. We recommend the latest supported firmware version.
- One of the following operating systems to install the RSM for Storage software:
  - Novell SLES 9 (Service Pack 3)
  - Novell SUSE® SLES 10
  - Red Hat RHEL 4 AS (Update 4)

> **Note:** Refer to IBM Remote Support Manager for Storage Compatibility Guide for the latest update of supported servers, modem, and operating systems. The document can be downloaded from the following Web page:
>
> http://www.ibm.com/support/docview.wss?uid=psg1MIGR-66062&rs=594

2.7.3 How RSM for Storage works

RSM for Storage uses an Ethernet connection for problem reporting and a modem for remote access by IBM Service, as shown in Figure 2-26.

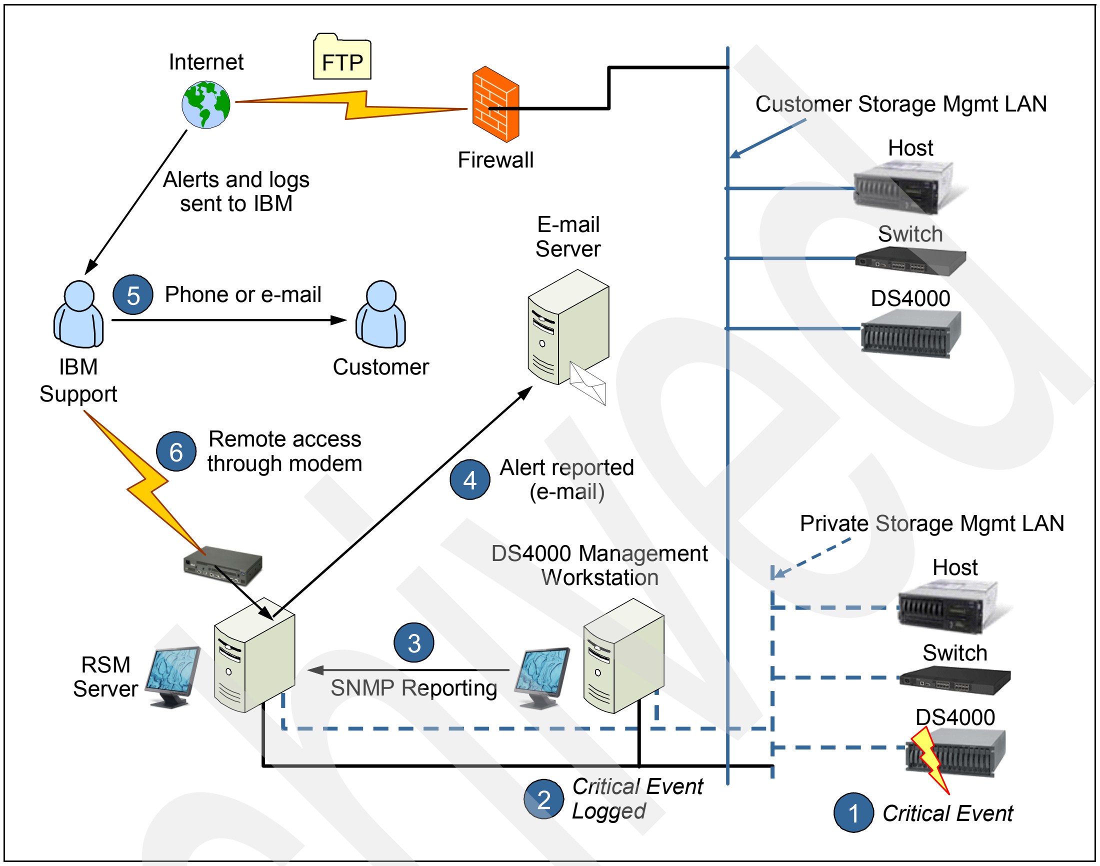


Figure 2-26 RSM for Storage connection diagram

The RSM for Storage server must have IP connectivity to the Ethernet management ports of the storage subsystems to be monitored and the management station running IBM Storage Manager's Event Monitor. It is also required that all storage subsystems, the management station, the e-mail server, and Internet gateway are accessible from the RSM server without requiring authorization through a firewall.

If your managed storage subsystems or other SAN devices are on a private management LAN, a second Ethernet port for accessing your company's SMTP server and the Internet will be required if your selected RSM server has only a single Ethernet port (see Figure 2-26).

Figure 2-26 shows that some storage subsystems and other SAN devices are managed through a private management LAN and the others are managed through the customer intranet. Therefore, the RSM server needs at least two network adapters.

Assuming that an RSM application is already installed, configured, and activated, below is an example scenario to show you how RSM works. Refer to Figure 2-26 to understand the flow.

1. An error occurs in one of the storage subsystems.
2. A critical event is logged in the management station (running Storage Manager).
3. The management station reports the critical event to the RSM server using an SNMP trap.

4. The RSM system receives notification of the critical event and sends an alert to IBM Service.

   When an alert is received from the management station, RSM downloads logs and other problem determination data such as major event log (MEL), read link status (RLS), and Storage System profile of the storage subsystem's controllers that reports the problem using the out-of-band Ethernet interfaces, and sends them along with the alert to IBM Service via e-mail.

   SNMP traps are sent by the IBM Storage Manager client or the IBM Storage Manager's Event Monitor service. As the Storage Manager Client may not always be running, we recommend that the Event Monitor be installed. Refer to Storage Manager documentation to check the installation of Event Monitor. The Storage Manager documentation is available under:

   – *Host Installation Support Guide - IBM System Storage DS4000 Storage Manager v10.15*, GA76-0422-00:

     https://www-304.ibm.com/systems/support/supportsite.wss/docdisplay?lndocid=MIGR-5075652&brandind=5000028

   – “Storage Manager v2 Installation and Support Guide for IBM AIX and Linux on POWER - IBM System Storage DS3000”:

     https://www-304.ibm.com/systems/support/supportsite.wss/docdisplay?lndocid=MIGR-5073850&brandind=5000028

   Refer to 2.7.4, “RSM management interface” on page 58, to configure the SNMP trap in Storage Manager.

5. IBM support does problem determination based on information sent by the alert along with the problem determination data such as MEL, RLS, and subsystem profile. If the problem can be fixed with existing information, IBM support contacts the customer either by phone or e-mail to tell how to resolve the problem, and either IBM support or the customer must indicate *Service Complete* for all alerts in the RSM after the problem is resolved. IBM support can dial to the RSM modem to *acknowledge* and indicate service complete for all alerts for the subsystem.

6. If the problem cannot be fixed with existing information, IBM support dials the RSM modem, acknowledges the alert, and does the troubleshooting by connecting to the storage subsystem using the command-line interface (SMcli) or RLOGIN. It may be required that IBM Support contacts the customer to obtain the password for the storage subsystem to use SMcli. IBM may also have to contact the customer to enable RLOGIN access. Indeed, we recommend normally disabling RLOGIN. Refer to “DS4000 subsystem security” on page 64 for setting up those security features.

   If IBM support needs to send or upload logs or other information from the RSM server to IBM, they can use FTP or e-mail commands from the Linux shell at the RSM server while connected through the modem line.

   The e-mail and the FTP data are sent to an IBM server through a customer network, not through the modem line (the modem connection is a TTY serial console, not an IP connection).

7. After the problem is resolved, all alerts must be closed either by IBM support or the customer before reporting will resume for that subsystem.

> **Note:** Once the RSM for Storage reports a problem to IBM for a given subsystem, no additional problems will be reported to IBM for that particular subsystem until all existing alerts are closed.

For details on configuring the DS Storage Manager for RSM and email alerts, refer to 7.2, “Configuring DS remote support” on page 222.

2.7.4 RSM management interface

The RSM management interface can be accessed through a Web browser pointing to the IP address or host name of the RSM server using HTTPS. You can use the Web interface to check the status and configure RSM settings. For IBM support, the interface is a command-line interface when connected to the RSM server via a modem.

Figure 2-27 is an example of a System Configuration menu for an already configured and activated RSM system.

Figure 2-27 RSM System Configuration menu

Under System Configuration, there are links at the top of the page that provide a summary status of the RSM system. Depending on the status, various icons may be displayed to the left of each link. The status values for each of these are:

- System
  - OK: RSM is correctly configured.
  - Incomplete: Format or content of configuration fields is incorrect or missing.
  - Problem: The configuration test has detected a problem.

- Reporting
  - Reporting: All subsystems are being monitored and no problems exist.
  - Suspended: There is a problem with the configuration that makes the system status not OK. No events will be processed by the RSM if the reporting status is suspended.
  - Partial: Reporting has been disabled for one or more subsystems.
  - Standby: Reporting has been disabled for all subsystems.
  - Storage Problem: One or more subsystems has a problem.
- Firewall
  - Disabled: No restrictions on access to the networks connected to the RSM.
  - Enabled:Close: This is the normal state when there are no active alerts present and the system is waiting for notification of a problem from Storage Manager.
  - Enabled:Open: Firewall is allowing access to one or more subsystems or other configured SAN devices.
- Remote Access
  - Disabled: Remote access via the modem is not allowed.
  - Enabled: The modem will answer and the remote service (rservice) user ID is valid for login via the remote connection.
  - Active: The rservice remote user is logged into the RSM.

2.7.5 RSM security considerations

RSM for Storage controls security for remote access by managing hardware and software components of the server it is installed on. Once installed, the server should be considered as a single-purpose appliance for problem reporting and remote access support for your storage subsystems. It should not be used for other applications.

User ID

During installation, RSM software creates three user IDs:

- admin: This is the administrative user that can perform management and configuration tasks.
- lservice: This is the local service user intended for use by IBM Service when on site. This user ID has some restrictions on directories that it can access. This is to prevent any configuration change that might affect the security of the system.
- rservice: This is the remote service (IBM Service) user that is used exclusively for remote access to the system and only valid when Remote Access is enabled. This user ID also does not have the ability to change any of the RSM security features.

Passwords for user IDs admin and lservice can be changed by the Linux *root* user and by invoking the command **`rsm-passwd admin`** or **`rsm-passwd lservice`**. We recommend setting a different password for each user ID.

For remote user (rservice) the password is automatically generated by RSM and it is changed daily at midnight Coordinated Universal Time (UTC). IBM Service has an internal tool that provides the current password. Therefore, you do not need to provide the current RSM password to IBM Service.

After validation of the initial login password, remote users are presented with a challenge string, which also requires access to an internal IBM tool in order to obtain the correct response. The response also includes an IBM employee user name that is recorded in the RSM for Storage security log.

Remote access security

Adding a modem to one of your systems creates a potential entry point for unauthorized access to your network. The RSM for Storage application modifies many characteristics and behaviors of the system that it is installed on to protect this entry point and to maximize the amount of control that you have in managing remote access.

In RSM, the modem send for remote access by IBM Service will not answer unless one of the storage subsystems has an active alert or Remote Access has manually been enabled.

Normally, Remote Access is enabled automatically when an alert is sent to IBM, but you can choose to wait for IBM Service to contact you when an alert is received and manually enable Remote Access at that time.

You can configure the Remote Access policy by selecting **Remote Access** from the Main Menu, as shown in Figure 2-28.

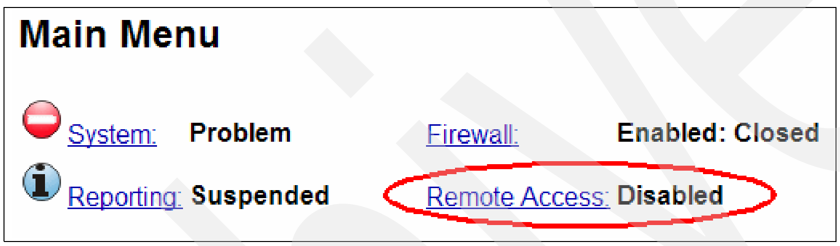


Figure 2-28 Accessing Remote access policy from main menu

In the Remote Access setting page, you can enable/disable the Remote Access service and enable/disable the option to automatically enable the Remote Access when an alert is sent to IBM. This is shown in Figure 2-29.

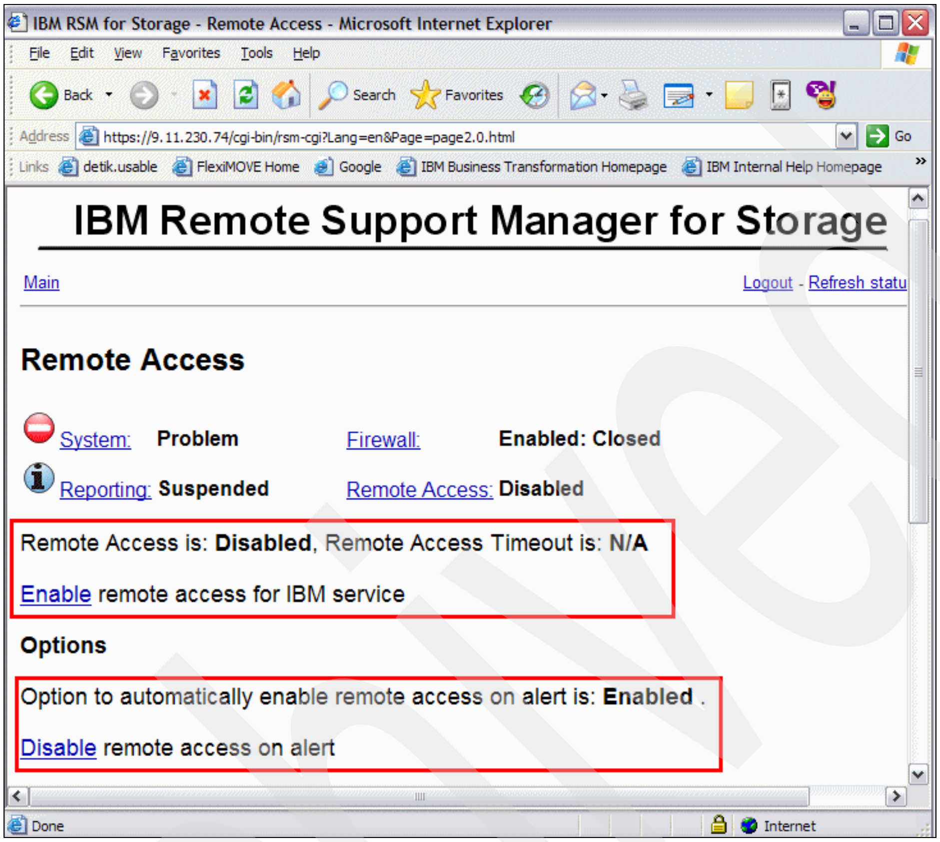


Figure 2-29 Remote Access settings

Remote Access also has a configurable time out between 12 to 96 hours. You can manually disable remote access when the service is complete or allow it to time out, thereby guaranteeing that the system will return to a secure state without intervention.

To configure the time-out value, scroll down the Remote Access settings page, select the desired time-out value, and click **Update Timeout Value**, as shown in Figure 2-30.

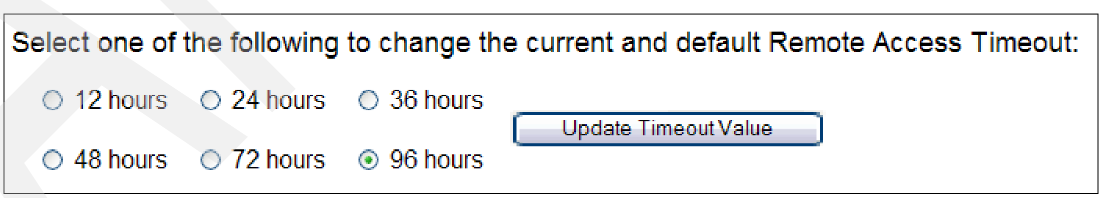


Figure 2-30 Configure remote access timeout

Note: You do not need to provide the rservice user ID password to IBM Service because IBM Service has an internal tool that provides the current rservice password. You only need to provide passwords of your storage subsystems or your other SAN devices if required.

Internal firewall

RSM for Storage includes an internal firewall to limit the scope of access a remote user has to your network. It also limits the IP destinations that can be accessed by local and remote users of the system. The rules for inbound and outbound IP traffic that control the internal firewall are managed dynamically by the RSM for Storage application.

There are three possible states of the RSM firewall:

- Disabled: No restrictions on access to the networks connected to the RSM. Remote access is not permitted if the firewall is in this state.
- Enabled:Close: This is the normal state when there are no active alerts present and the system is waiting for notification of a problem from Storage Manager.

  The following rules are applied when RSM is in this state:

  – Inbound SNMP traps, ping, traceroute, and HTTPS requests are accepted.

  – Outbound traffic for DNS, ping, traceroute, IBM WWW and FTP sites, and port 25 of your SMTP server are accepted.

  – Access to any of your other SAN devices is blocked.

  – IP traffic forwarding is disabled.

  – All other connections are blocked.
- Enabled:Open: Firewall is allowing access to one or more subsystems or other configured SAN devices. Access is allowed only to those devices that have active alerts or those that you have placed in Service Access mode.

> **Note:** Subsystems with active alerts are automatically allowed access from the Remote Support Manager while the alert is active and do not need to be enabled for Service Access.

To manage RSM internal firewall and service access of your managed storage subsystems (and other SAN devices) from the Web interface, click **Firewall** on the Main Menu, as shown in Figure 2-31.

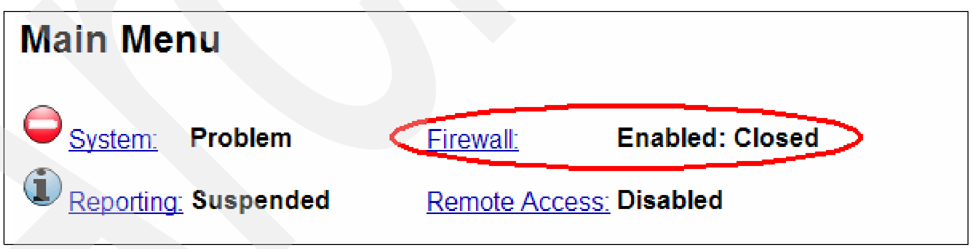


Figure 2-31 Accessing firewall configuration from main menu

On the Internal Firewall and Service Access page you can change the internal firewall status and service access mode of your managed storage subsystems, as shown in Figure 2-32.

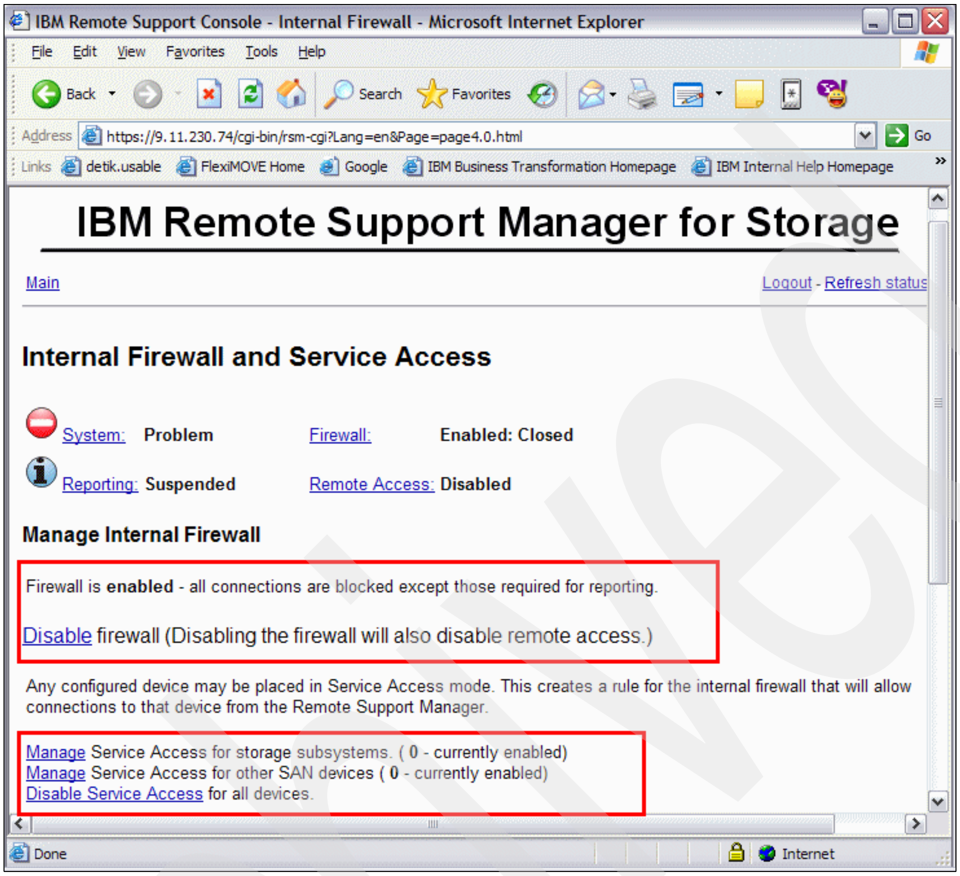


Figure 2-32 Internal Firewall and Service Access page

Placing a device into Service Access mode creates a rule for the internal firewall that allows connections to that device from the RSM server. For subsystems with active alerts, they are automatically allowed access from the Remote Support Manager while the alert is active and do not need to be enabled for Service Access.

Similar to Remote Access, you can also modify the service access timeout. To set the service access timeout, go to the Manage Service Access section in the Internal Firewall and Service Access page, select the desired Service Access Timeout value, and click **Update Access Timeout**, as shown in Figure 2-33.

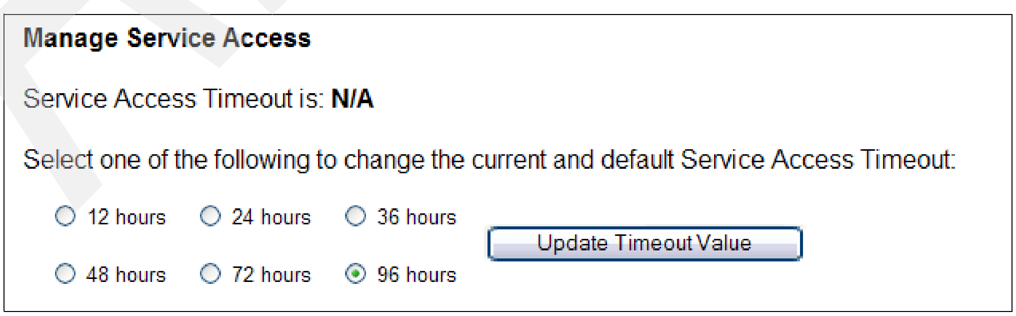


Figure 2-33 Manage Service Access

DS4000 subsystem security

Storage Manager has the ability to require an administrative password in order to make changes to the subsystem configuration. We recommend configuring this password.

DS4000 also has a controller shell environment that is accessible using a remote login (RLOGIN) client. IBM Storage Manager for DS4000 has an option to disable RLOGIN, and we normally recommend disabling RLOGIN.

For further information about RSM for Storage, a comprehensive Planning, Installation, and User's Guide is available to download from the following Web page:

http://www.ibm.com/support/docview.wss?uid=psg1MIGR-66062&rs=594

2.8 Logical storage architecture

This section explains how to create arrays and logical drives, how to use storage partitioning, and the premium features FlashCopy, Volume Copy and Enhanced Remote Mirroring.

2.8.1 Arrays and logical drives

An array is a set of physical drives that the controller logically groups together to provide one or more logical drives to an application host or cluster.

Controller firmware V7.10 introduces support for RAID-0 and 1 arrays with more than 30 drives, up to a maximum of 112 for DS4200 and DS4700 storage subsystems and up to a maximum of 224 drives per DS4800 storage subsystem. In certain DS4000 configurations, this could improve performance, provided that your system is optimally tuned. It also improves the data capacities of these arrays. Note that RAID-1 or 10 requires an even number of disk drives. RAID-3, 5, and 6 arrays can contain a maximum of 30 physical drives only.

Controller firmware V7.10 also introduces support for logical drives greater than 2 TB. With the ever-increasing size of physical disk drives, creating logical drives greater than 2 TB will become more common. This could typically be used with applications using large files requiring sequential access, like digital video, multimedia, and medical imaging.

Previously, if an application had a specific requirement for a logical drive greater than 2 TB, it was dependant on the operating system to provide the capability of combining several logical drives into one virtual drive presented to the application. This new feature shifts the management of greater than 2 TB logical drives to the DS4000 Storage Subsystem.

Table 2-5 reflects the maximum logical volume size supported per operating system.

Table 2-5 Maximum logical volume sizes

Operating system	Maximum logical volume size supported by operating system
Windows 2000	Approximately 274 TB
Windows Server® 2003	Approximately 274 TB
Solaris 8	1 TB
Solaris 9, 10	16 TB
AIX v5.2 and v5.3	1 TB on 32 bit, 2 TB on 64 bit

Operating system	Maximum logical volume size supported by operating system
HP-UX 11.11	2 TB
HP-UX v11.23	32 TB
Linux (2.6 kernel)	8 EB
NetWare	2 TB using NSS

Creating arrays and logical drives

Note: Very large logical drives could have a negative performance impact. See 5.2.2, “Size and allocation of LUNs” on page 111, for recommendations for IBM i attachment via VIOS.

A logical drive is a logical structure that you create on a Storage System for data storage. Creating arrays and logical drives is one of the most basic steps and is required before you can start using the physical disk space. That is, you divide your disk drives into arrays and create one or more logical drives inside each array, which are mapped to hosts as LUNs.

Note: Always leave a small amount of free space in the array after the LUNs have been created. This free space may be required for dynamic segment size changes or if future firmware upgrades require a larger DACstore area on the drives.

Through Storage Manager you have the option of letting the system automatically create arrays for you. This usually ensures the most optimal balance between capacity, performance, and redundancy. The automatic configuration provides you with a suggested configuration per RAID level (provided that there are enough unassigned drives to satisfy the minimum requirement per RAID level). Each configuration provides volume groups, volumes, and hot spare drives.

You also have the option to select your own characteristics for the automatic configuration specifying the RAID level required, number of logical drive groups, number of disk drives per array, number of hot spare drives required, and also to specify the I/O characteristics associated with the configuration.

On an already configured storage subsystem this feature lets you automatically configure the arrays using the remaining unconfigured storage capacity. Any free capacity on existing arrays within the storage subsystem will not be configured.

The Create Logical Drive wizard allows you to manually create one or more logical drives on the Storage System. Using the wizard, you select the capacity that you want to allocate for the logical drive (free capacity or unconfigured capacity) and then define basic and optional advanced logical drive parameters for the logical drive.

Tip: Storage Manager allows you to use the default setup through the automatic configuration, should you not have a specific setup requirement. This usually ensures the most optimal balance between capacity, performance, and redundancy. Using the manual method of creating logical drives allows greater flexibility for configuration settings, such as enclosure loss protection and utilizing all available drive loops.

For configuration examples refer to 7.3.3, “Creating RAID arrays and logical drives” on page 235.

In the simplest configuration, all of your drives would be assigned to one array, and multiple logical drives could be created within it. This configuration presents the drawback that if you experience a drive failure, the rebuild process will affect all logical drives and the overall system performance will go down.

Number of drives

The more physical drives that you have per array, the shorter the access time for read and write I/O operations.

You can determine how many physical drives should be associated with a RAID controller by looking at disk transfer rates (rather than at the megabytes per second). For example, if a hard disk drive is capable of 75 nonsequential (random) I/Os per second, then about 26 hard disk drives working together could (theoretically) produce 2,000 nonsequential I/Os per second, or enough to hit the maximum I/O handling capacity of a single RAID controller. If the hard disk drive can sustain 150 sequential I/Os per second, it then takes only about 13 hard disk drives working together to produce the same 2,000 sequential I/Os per second and keep the RAID controller running at maximum throughput.

Tip: Having more physical disks for the same overall capacity gives you:

- Performance: By doubling the number of the physical drives, you can expect up to a 50% increase in throughput performance.
- Flexibility: Using more physical drives gives you more flexibility to build arrays and logical drives according to your needs.
- Data capacity: When using RAID-5 logical drives, more data space is available with smaller physical drives because less space (capacity of a drive) is used for parity.

Enclosure loss protection

You achieve enclosure loss protection by spreading the array across multiple enclosures. Doing this increases redundancy and also performance. You reduce the chance of an array going offline by spreading it over many enclosures. If you have an enclosure failure, only one drive from the array is affected (assuming that the array uses no more than one drive per enclosure), leaving your data intact. For a RAID-1 array, you need two enclosures. For a RAID-5 array, you need a minimum of three enclosures. Figure 2-34 shows an array that is created with enclosure loss protection.

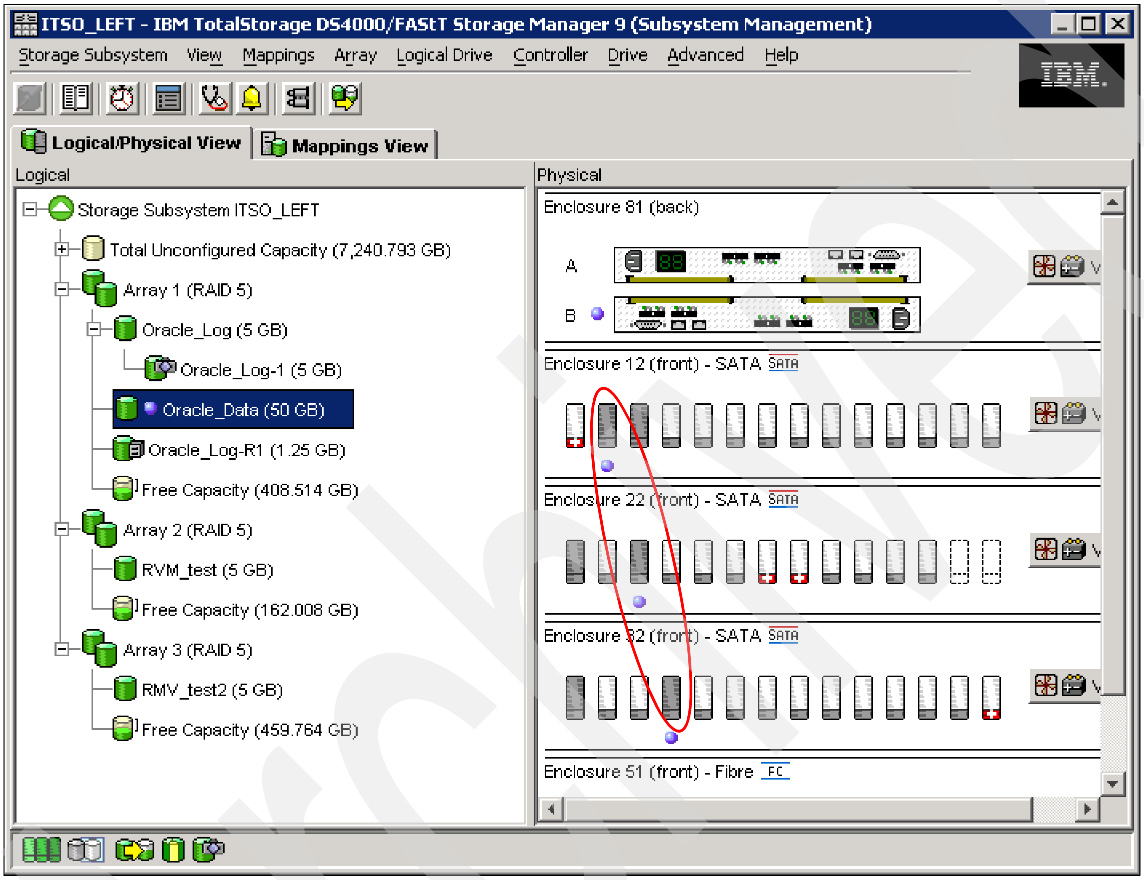


Figure 2-34 Array created with enclosure loss protection

You also increase performance because the I/O request is processed by both drive loops.

2.8.2 Storage partitioning

Storage partitioning adds a high level of flexibility to the DS4000 power supply fan unit. It enables you to connect to the same storage server multiple and heterogeneous host systems, either in stand-alone or clustered mode. The term storage partitioning is somewhat misleading, as it actually represents a host or a group of hosts and the logical disks that they access.

Without storage partitioning, the logical drives configured on a DS4000 power supply fan unit can only be accessed by a single host system or by a single cluster. This can surely lead to inefficient use of storage server hardware.

Storage partitioning, on the other hand, allows you to create sets, containing the hosts with their host bus adapters and the logical drives. We call these sets *storage partitions*. Now the host systems can only access their assigned logical drives, just as though these logical drives were locally attached to them. Storage partitioning adapts the SAN idea of globally accessible storage to the local-storage-minded operating systems.

Storage partitioning lets you map and mask LUNs (that is why it is also referred to as LUN masking). This means that once you have assigned the LUN to a host it is hidden from all other hosts connected to the same storage server. Therefore, the access to that LUN is exclusively reserved for that host.

It is a good practice to do your storage partitioning prior to connecting multiple hosts. Operating systems such as AIX or Windows 2000 like to write their signatures to any device that they can access.

> **Note:** The are limitations as to how many logical drives you can map per host. Most DS4000 Series power supply fan units can support up to 256 LUNs (including the access LUN) per partition.

Storage partition topology is a collection of topological elements (default group, host groups, hosts, and host ports) shown as nodes in the topology view of the mappings view. You must define the various topological elements if you want to define specific logical drive-to-LUN mappings for host groups or hosts.

A storage partition contains several components:

- Hosts groups
- Hosts
- Host ports
- Logical drive mappings

A host group is a collection of hosts that are allowed to access certain logical drives, for example, a cluster of two systems.

> **Note:** Host groups are not used for IBM i, as IBM i never shares access to its LUNs with other hosts.

A host is a single system that can be mapped to a logical drive.

A host port is the FC port of the host bus adapter on the host system. The host port is identified by its world-wide port name (WWPN). A single host can contain more than one host port. If you attach the servers in a redundant way (highly recommended), each server needs two host bus adapters. That is, it needs two host ports within the same host system.

The DS4000 power supply fan unit only communicates through the use of the WWPN. The Storage System is not aware of which host bus adapters are in the same server or in servers that have a certain relationship, such as a cluster. The host groups, the hosts, and their host ports actually reflect a logical view of the physical connections of your SAN, as well as the logical connection between servers, such as clusters.

With the logical setup defined previously, mappings are specific assignments of logical drives to particular host groups or hosts.

The storage partition is the combination of all these components. It ensures proper access to the different logical drives even if there are several hosts or clusters connected.

The default host group is a placeholder for hosts that are defined but have not been mapped. The default host group is also normally used only when storage partitioning is not enabled. If this is the case, then only one type of operating system should be sharing the logical drives.

Every unassigned logical drive is mapped to the undefined mappings group. This means that no host (or host port, to be precise) can access these logical drives until they are mapped.

With Storage Manager V10.10 (firmware V7.10), you can now have up to 128 partitions on a DS4700 (from 64 previously) and up to 512 storage partitions on a DS5000 and DS4800 (previously the limit was of 64 partitions for those systems). This allows the storage subsystem to serve storage capacity to a greater amount of heterogeneous hosts, allowing for greater scalability.

For the number of maximum storage partitions for a specific model, see Table 2-6. Note that on some DS4000 models, the number of partitions also depends on the licences that have been purchased.

Table 2-6 New supported maximum storage partitions with V7.10 firmware

DS4000 model	Number of supported partitions
DS3400	4, can upgrade to 16
DS4700	2, 4, 8, 16, 32, 64, or 128 (must order one), can upgrade to 4, 8, 16, 32, 64, or 128 (Mod 70A) 8, 16, 32, 64, or 128 (must order one), can upgrade to 16, 32, 64, or 128 (Mod 72A)
DS4800	8, 16, 32, 64, 128, 256, or 512 (must order one), can upgrade to 16, 32, 64, 128, 256, or 512
DS5000	8, 16, 32, 64, 128, 256, or 512 (must order one), can upgrade to 16, 32, 64, 128, 256, or 512

Note: With the increase in storage partitions under V7.10 firmware, the total number of host ports supported is now 512 ports for the DS4700 storage subsystem and 2,048 ports for the DS5000 and DS4800 storage subsystems.

Every mapping of a logical drive to a new host or host group creates a new storage partition. If you map additional logical drives to the same host or host group, this does not count as a new storage partition.

For a configuration example see 7.3.5, “Defining logical drive to LUN mapping” on page 249.

2.8.3 Controller ownership

Each logical drive has a preferred controller of ownership. This controller normally handles all I/O requests for this particular logical drive. In other words, each logical drive is owned by one and only one of both storage subsystem controllers. The alternate controller only takes over and handles the I/O requests in the case of a failure along the I/O path, for example, a defect host bus adapter or switch. When defining logical drives, the system normally alternates ownership between the two controllers.

Important: Be sure that the operating system using the logical drive uses a multipath I/O driver like RDAC or MPIO for VIOS. Otherwise, it loses access to the logical drive. We recommend enabling the RDAC device driver auto-recovery option to automatically re-assign the logical drives to their preferred path after recovery from an I/O controller path failure (see 6.5, "Configuring VIOS virtual devices" on page 191).

The Performance Monitor provides data useful for monitoring the I/O activity of a specific controller and a specific logical drive, which can help identify possible high-traffic I/O areas. Identify actual I/O patterns to the individual logical drives and compare those with the expectations based on the application. If a particular controller has considerably more I/O activity, consider rebalancing your IBM i ASP data instead of moving logical drives to the other controller in the Storage System.

For non-IBM i hosts consider changing the controller ownership of one or more logical drives to the controller with the lighter workload.

The preferred controller ownership of a logical drive or array is the controller of an active-active pair that is designated to own these logical drives. The preferred controller owner is the controller that currently owns the logical drive or array.

If the preferred controller is undergoing a firmware download, ownership of the logical drives is automatically shifted to the other controller, and that controller becomes the current owner of the logical drives. If the preferred controller needs to be replaced, you should disable the controller first. This will intentionally cause a failover of LUNs to the other controller and allow the removal and replacement of the preferred controller. This is considered a routine ownership change and is reported with an informational entry in the event log.

There can also be a forced failover from the preferred controller to the other controller because of I/O path errors. This is reported with a critical entry in the event log, and will be reported by the Enterprise Management software to e-mail and SNMP alert destinations.

Restriction: A secondary logical drive in a remote mirror does not have a preferred owner. Instead, the ownership of the secondary logical drive is determined by the controller owner of the associated primary logical drive. For example, if controller A owns the primary logical drive in the primary Storage System, then controller A owns the associated secondary logical drive in the secondary Storage System. Controller ownership changes of the primary logical drive cause a corresponding controller ownership change of the secondary logical drive.

2.8.4 FlashCopy and Volume Copy

This section gives an overview of the FlashCopy and the Volume Copy features.

FlashCopy

You use FlashCopy to create and manage FlashCopy logical drives. A FlashCopy logical drive is a *point-in-time* image of a *standard logical drive* in your storage subsystem and it is created much faster than a physical copy using the Volume Copy feature. It also requires less disk space than a physical copy. On the other hand, it is not a real physical copy, because it does not copy all the data. Consequently, if the FlashCopy relationship is disabled or the source logical drive is damaged, the FlashCopy logical drive becomes unusable as well.

The logical drive that is copied is called a *base logical drive*. When you make a FlashCopy, the controller suspends write operations to the base logical drive for a few seconds while it creates a *FlashCopy repository logical drive*. This is a physical logical drive where FlashCopy metadata and copy-on-write data are stored. Figure 2-35 on page 72 shows how FlashCopy works on the DS3000, DS4000, and DS5000.

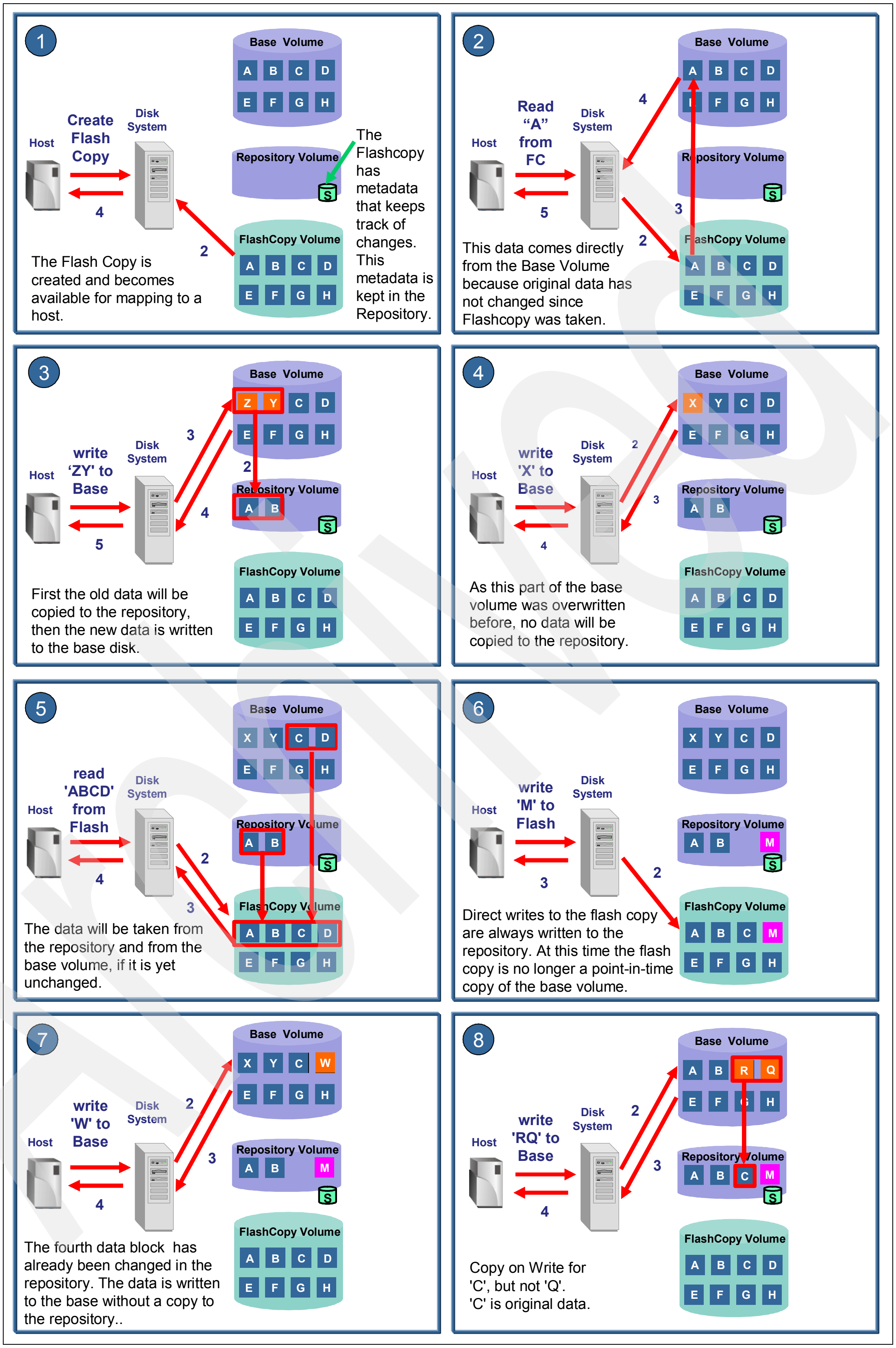


Figure 2-35 FlashCopy example

In picture 1 in Figure 2-35 on page 72 the FlashCopy is created. During the creation, the user defines the location and the size of the FlashCopy repository volume. The repository volume is a physical logical drive that contains metadata that keeps track of changes on the base logical drive after the FlashCopy is created, and it provides physical disk space for future copy-on-write operations. After creating a FlashCopy logical drive, this can immediately be mapped to a host.

In picture 2 data is read from the FlashCopy volume. Because the FlashCopy volume does not contain any data, the data is read from the base volume directly.

If data on the base volume is overwritten, a copy-on-write will first copy the original data to the repository. Then the new data is written to the base logical drive (picture 3).

If this data is overwritten again, no new copy-on-write will occur, because the original point-in-time data is already stored in the repository (picture 4).

If data is read from the FlashCopy logical drive, it will look up the repository volume first. If data exists in this area, it will be taken from the repository volume. If no data exists in the repository volume, it will be read from the base logical drive (picture 5).

If data is written to the FlashCopy volume, it will always be stored on the repository volume (picture 6).

If that data area is overwritten later on the base logical drive, no copy on write will occur, as the data area on the repository volume is already occupied (picture 7).

If the data area on the repository is only partially occupied, only the original data from the base logical drive will be copied to the repository (picture 8).

You can create up to four FlashCopies of a base logical drive. Once created FlashCopy logical drives are immediately available for full read and write access. For example, before upgrading a database management system, you can use FlashCopy logical drives to test different configurations. You can create FlashCopy logical drives for repeated reuse (backups or data analysis/reporting) or for one-time use (speculative change or upgrade testing).

The FlashCopy function is very space efficient because it only stores the data that is overwritten on the base logical volume since the creation of the FlashCopy logical drive and the data that is directly written to the FlashCopy volume. When creating the FlashCopy logical drive, the default value for the size of the repository is 20% of the size of the base logical drive. This value can be changed during the creation of the FlashCopy volume or dynamically at a later point in time to any percentage value. Depending on the write ratio to the base logical drive or the time the FlashCopy volume should remain active, it may be advisable to select a larger size for the repository volume.

If the repository reaches 50% of its capacity a warning message will be logged in the major event log (MEL).

If the repository volume exceeds its capacity, either the FlashCopy logical drive will be failed or the base logical drive will be set to a read-only mode. This depends on the FlashCopy failure policy that the user defines when creating the FlashCopy volume.

> **Note:** As IBM i always requires write access to its disk units, the only possible option for IBM i is to let the FlashCopy volumes fail if the repository capacity is exceeded.

Refer to 8.1, “Implementing FlashCopy” on page 284, for further information about the usage of FlashCopy with IBM i.

For more information regarding to the FlashCopy feature refer to the *Copy Services User's Guide - IBM System Storage DS4000 Storage Manager:*

https://www-304.ibm.com/systems/support/supportsite.wss/docdisplay?lndocid=MIGR-61173&brandind=5000028

Volume Copy

The Volume Copy premium feature is used to copy data from one logical drive (source) to another logical drive (target) in a single storage subsystem (Figure 2-36). The target logical drive is an exact copy or *clone* of the source logical drive.

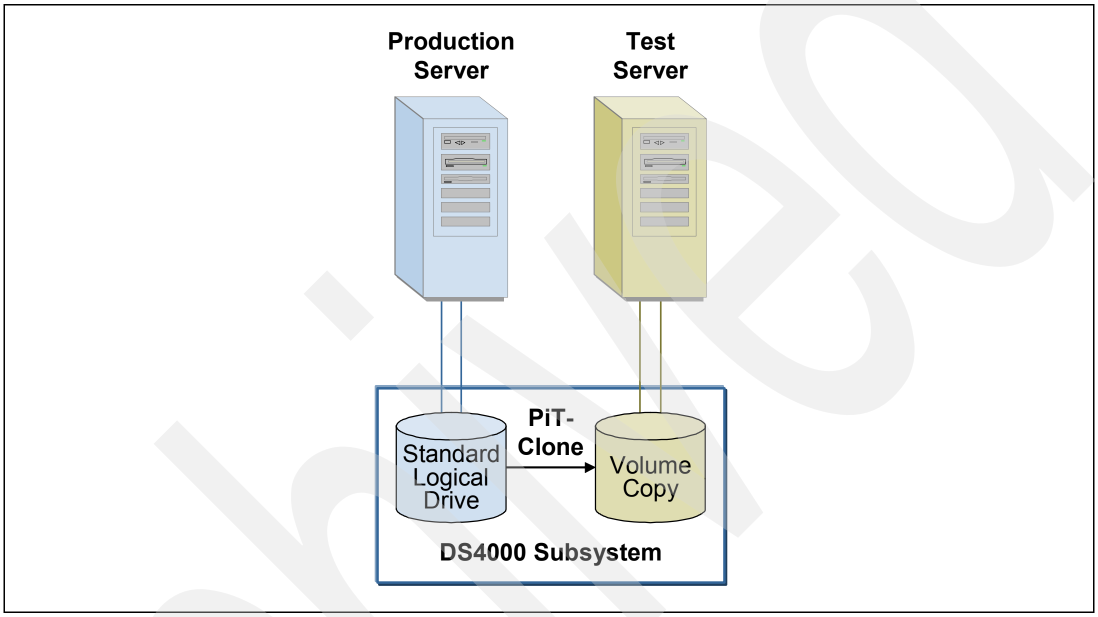


Figure 2-36 Volume Copy

The Volume Copy feature can be used for the following tasks:

- Data migration.

  As storage requirements for a logical drive change, the Volume Copy function can be used to copy data to a logical drive in an array that utilizes larger capacity disk drives within the same Storage System. This provides an opportunity to move data to larger drives (for example, 73 GB to 300 GB or 10 kRPM to 15 kRPM), change to drives with a higher data transfer rate (for example, 1 Gbps to 2 Gbps or to 4 Gbps), or to change to drives using different technologies (from SATA to FC) for higher performance and reliability.

- Back up data.

  The Volume Copy function allows you to create a backup of a logical drive by copying data from one logical drive to another logical drive in the same Storage System. The target logical drive can be used as a backup of the source logical drive, for system testing, or to do a backup without impacting the source logical drive performance to another device, such as a tape drive.

- Restore FlashCopy logical drive data to the base logical drive.

  If you need to restore data to the base logical drive from its associated FlashCopy logical drive, the Volume Copy function can be used to copy the data from the FlashCopy logical drive to the base logical drive. You can create a Volume Copy of the data on the FlashCopy logical drive, then copy the data to the base logical drive.

The Volume Copy premium feature includes a Create Copy Wizard to assist in creating a logical drive copy, and a Copy Manager to monitor logical drive copies after they have been created.

The Create Copy Wizard guides the user through the process of selecting a source logical drive from a list of available logical drives, selecting a target logical drive from a list of available logical drives, and setting the copy priority for the Volume Copy. All logical drives that have the same or a larger capacity as the source logical drive are potential targets for a Volume Copy operation.

> **Notes:** A Volume Copy overwrites data on the target logical drive and automatically makes the target logical drive *read-only* to hosts. The read-only status can be manually removed as soon a the copy operation has finished.
>
> For IBM i only Volume Copy target logical drives of the *same* capacity as the source logical drive are supported.

Already created Volume Copy operations can be monitored and managed through the Copy Manager. From the Copy Manager, a Volume Copy can be re-copied, stopped, or removed, and its attributes, including the copy priority and the target logical drive read-only attribute, can be modified. The status and progress of a Volume Copy can also be monitored with the Copy Manager.

The Volume Copy premium feature must be enabled by purchasing a feature key.

Volume Copy is a full point-in-time replication. It allows for analysis, data mining, and testing without any degradation of the production logical drive performance. It also brings improvements to backup and restore operations, making them faster and eliminating I/O contention on the primary (source) logical drive.

Users submit a Volume Copy request by specifying two compatible volumes (logical drives). One logical drive is designated as the source, and the other is the target.

Copying data is a background operation managed by the controller firmware, which reads the source logical drive and writes the data to the target logical drive. If the Storage System controller experiences a reset, like during concurrent firmware upgrades, the copy request is restored and the copy process resumes from the last known progress boundary.

After submitting a copy request, the source logical drive is only available for read I/O activity while a logical drive copy has a status of In progress, pending, or failed. Write requests are allowed after the logical drive copy is completed. Read and write requests to the target logical drive will not take place while the logical drive copy has a status of In progress, pending, or failed.

These restrictions on IBM DS Midrange Storage Systems are necessary to ensure the integrity of the point-in-time copy. If the logical drive being copied is large, this can result in an extended period of time without the ability for a production application to make updates or changes to the data.

> **Important:** During the Volume Copy data transfer operation, any write requests to the source logical drive will be rejected. If the source logical drive is used in a production environment, the FlashCopy feature must be enabled and the FlashCopy of the logical drive must be specified as the Volume Copy source logical drive instead of using the actual logical drive itself. This requirement is to prevent the logical drive from being inaccessible to the users.

As illustrated in Figure 2-37, FlashCopy, which allows a point-in-time copy to be made while maintaining read/write access, enables a complete copy to be created without interrupting the I/O activity of the production logical drive.

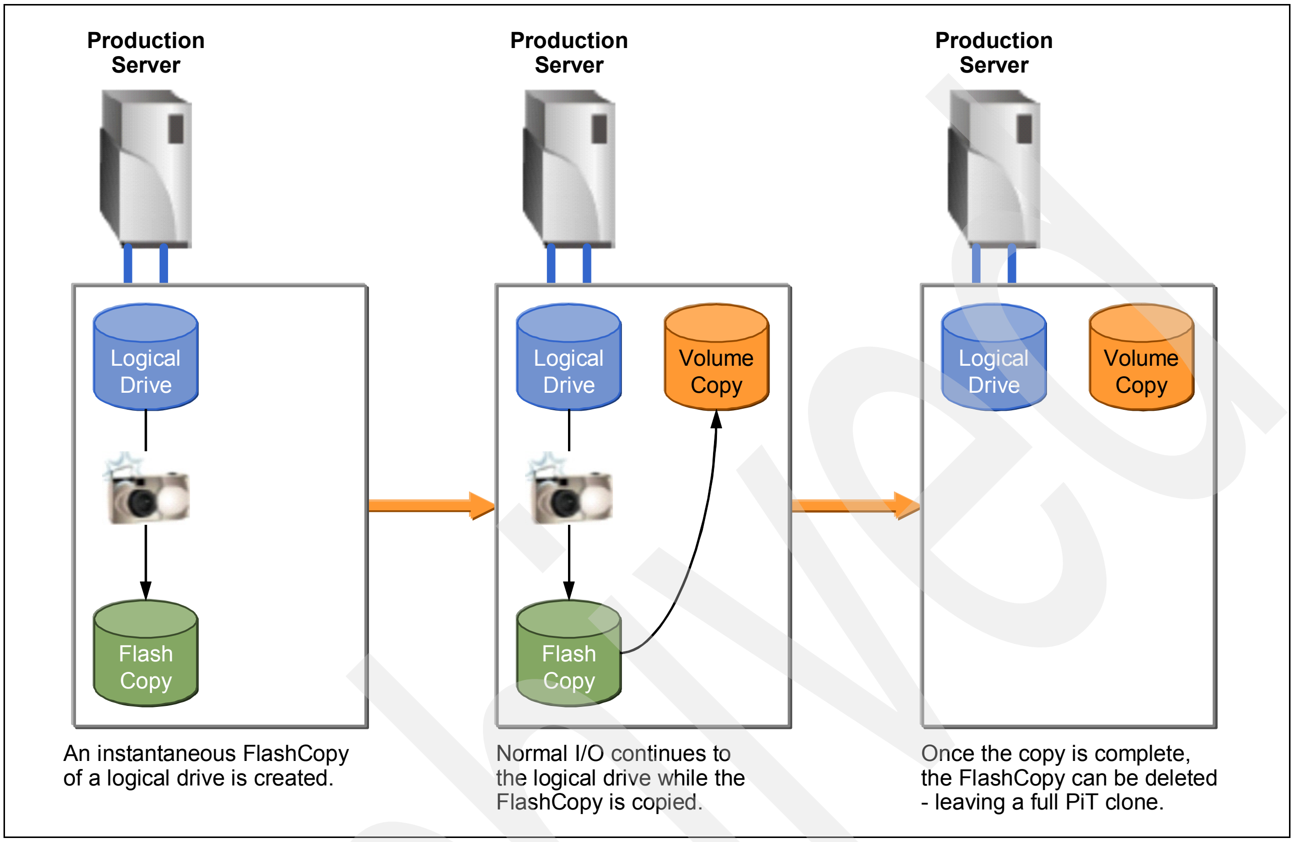


Figure 2-37 Volume Copy integration with FlashCopy

Restrictions on Volume Copy

The following restrictions apply to the source logical drive, target logical drive, and the Storage System:

- The source logical drive is available for read I/O activity only while a Volume Copy has a status of in progress or pending. Write requests are allowed after the Volume Copy is completed.

> **Tip:** In practice, Volume Copy should be used in combination with FlashCopy in order to avoid the long FlashCopy source logical drive read-only period.

- A logical drive can be used as a target logical drive in only one Volume Copy at a time.
- The maximum allowable number of logical drive copies per Storage System is dependent upon the number of target logical drives available on your Storage System.
- A Storage System can have up to eight VolumeCopies running in parallel. Any additional Volume Copy operations will have the status *pending* until the total number of ongoing Volume Copy operations is below 8.
- The target logical drive capacity must be equal to or greater than the source logical drive capacity. For IBM i only equal source and target capacities are supported.
- A source logical drive can be a standard logical drive, FlashCopy logical drive, FlashCopy base logical drive, or a Remote Mirror primary logical drive.

- A target logical drive can be a standard logical drive, a base logical drive of a disabled or failed FlashCopy logical drive, or a remote mirror primary logical drive.

Important: If you choose a base logical drive of a FlashCopy logical drive as your Volume Copy target logical drive, you must disable all FlashCopy logical drives associated with the base logical drive before you can select it as a target logical drive. Otherwise, the base logical drive cannot be used as a target logical drive.

Logical drives that have the following statuses cannot be used as a source logical drive or target logical drive:

- A logical drive that is reserved by the host cannot be selected as a source or target logical drive.
- A logical drive that is in a modification operation.
- A logical drive that is the source logical drive or target logical drive in another Volume Copy with a status of failed, in progress, or pending.
- A logical drive with a status of failed.
- A logical drive with a status of degraded.

For more information regarding the Volume Copy feature refer to the *Copy Services User's Guide - IBM System Storage DS4000 Storage Manager*:

https://www-304.ibm.com/systems/support/supportsite.wss/docdisplay?lndocid=MIGR-61173&brandind=5000028

2.8.5 Enhanced Remote Mirroring

Enhanced Remote Mirroring provides the ability to create remote logical drive mirror pairs by using a synchronous or asynchronous write mode with or without using the write consistency group option. Asynchronous mirroring with the consistency group option is referred to as global mirroring and asynchronous mirroring *without* the consistency group option is referred to as global copy. Mirroring that uses the synchronous write mode is referred to as metro mirroring.

Metro mirroring (synchronous mirroring)

Metro mirroring is a synchronous mirroring mode. This means that the controller does not send the I/O completion to the host until the data has been copied to both the primary and the secondary logical drives.

When the owning controller of the primary logical drive receives a write request from a host, the controller first logs information about the write request on the *mirror repository logical drive*. This information is actually placed in a queue on the mirror repository volumes. In parallel, it writes the data to the primary logical drive. The controller then initiates a remote write operation to copy the affected data blocks to the secondary logical drive at the remote site. When the remote write operation is complete, the primary controller removes the log record from the mirror repository logical drive, which means that the log entries are deleted from the queue. Finally, the controller sends an I/O completion indication back to the host system.

Note: The owning primary controller only writes status and control information to the repository logical drive. The repository is not used to store actual host data.

When write caching is enabled on either the primary or the secondary logical drive, the I/O completion is sent when data is in the cache on the site (primary or secondary) where write caching is enabled. When write caching is disabled on either the primary or secondary logical drive, then the I/O completion is not sent until the data has been stored to physical media on that site.

Figure 2-38 depicts how a write request from the host flows to both controllers to provide an instant copy.

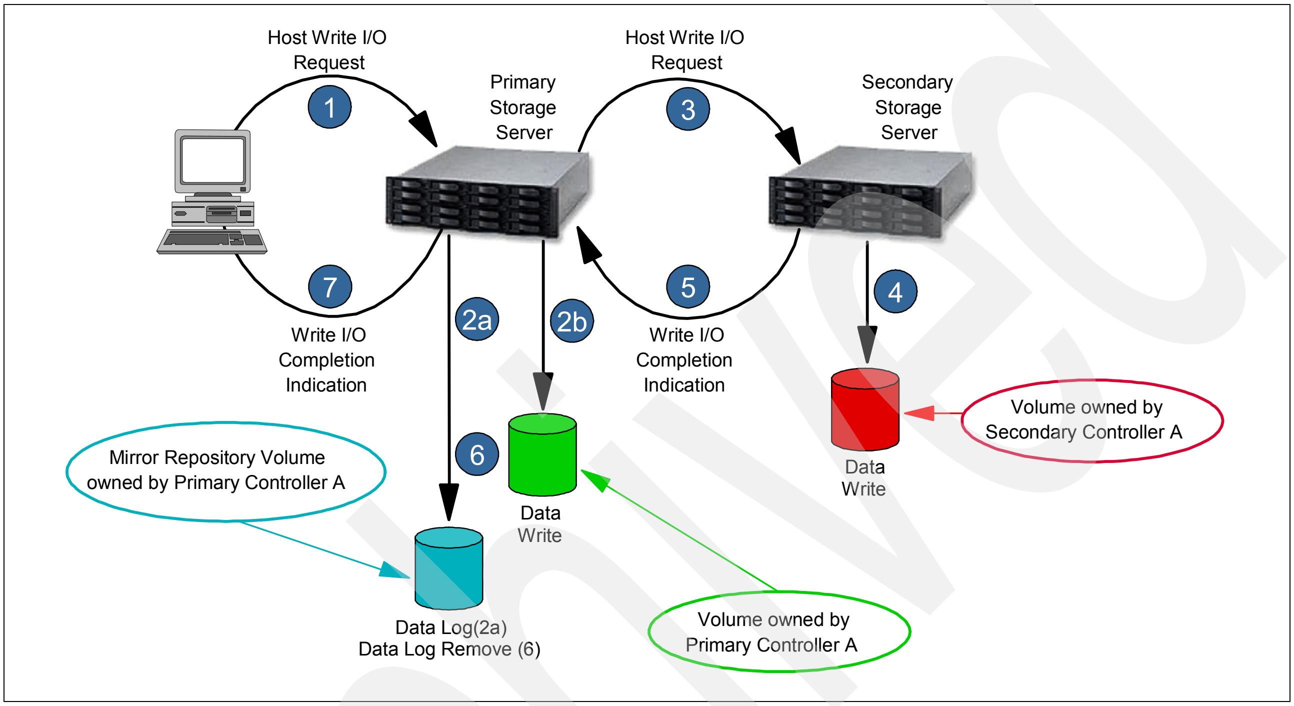


Figure 2-38 Metro mirroring mode (synchronous mirroring) data flow

When a controller receives a read request from a host system, the read request is handled on the primary Storage System and no communication takes place between the primary and the secondary Storage Systems.

Global copy (asynchronous mirroring without write consistency group)

Global copy is an asynchronous write mode. All write requests from the host are written to the primary (local) logical drive and immediately reported as completed to the host system. Regardless of when data was copied to the remote Storage System, the application does not wait for the I/O write request result from the remote site. However, global copy does not ensure that write requests at the primary site are processed in the same order at the remote site. As such, it is also referred as *asynchronous mirroring without write consistency group*.

When the owning controller of the primary logical drive receives a write request from a host, the controller first logs information about the write request on the *mirror repository logical drive*. This information is placed in a queue on the mirror repository volumes. In parallel, it writes the data to the primary logical drive (or cache). After the data has been written (or cached), the host receives an I/O completion from the primary controller. The controller then initiates a background remote write operation to copy the corresponding data blocks to the secondary logical drive at the remote site. After the data has been copied to the secondary logical drive at the remote site (or cached), the primary controller removes the log record on the mirror repository logical drive, that is, deletes it from the queue, as shown in Figure 2-39 on page 79.

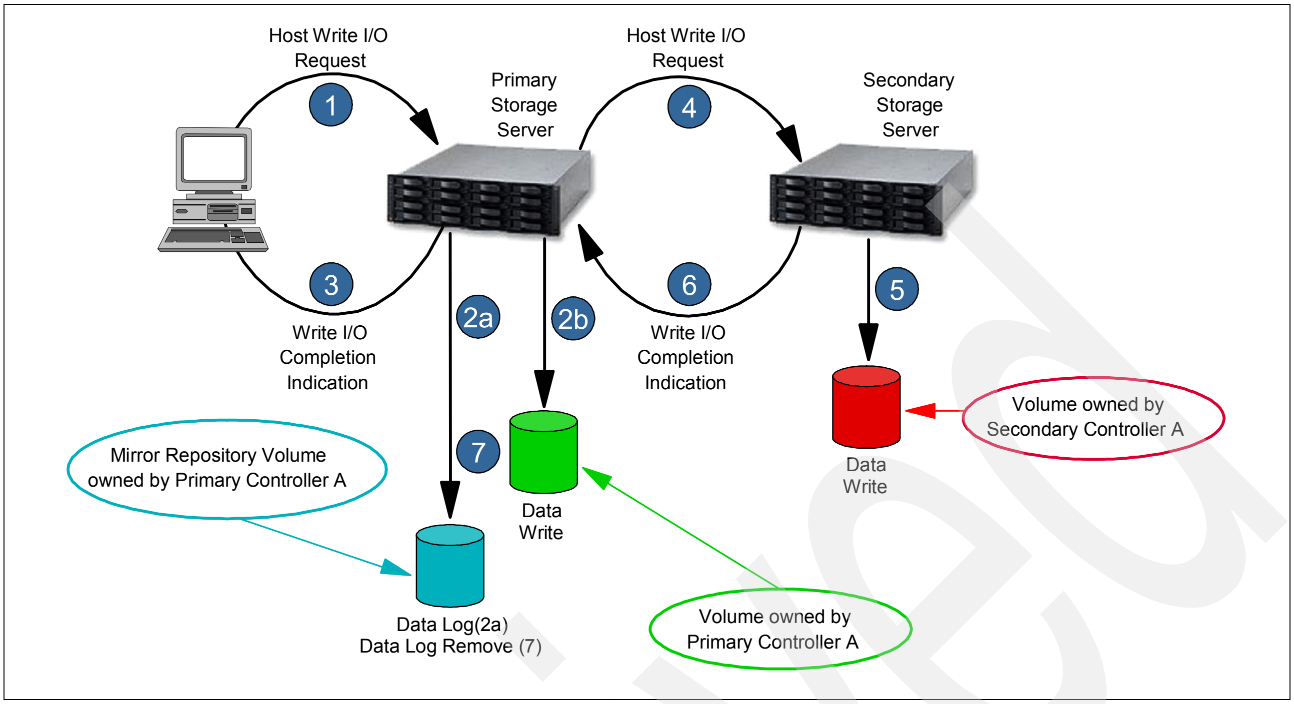


Figure 2-39 Global copy mode (asynchronous mirroring) data flow

When multiple mirror relationships are defined on the Storage System, the background synchronization of affected data blocks between the primary and secondary controller for the different relationships are conducted in parallel in a multi-threaded process.

Important: The write order for multiple volumes (for example, write requests to a database data volume and a database log volume on a data base server) is not guaranteed with the global copy mode. Therefore, the practical use case for GlobalCopy is limited to data migration.

When write caching is enabled on either the primary or the secondary logical drive, the I/O completion is sent when data is in the cache on the site (primary or secondary) where write caching is enabled. When write caching is disabled on either the primary or the secondary logical drive, then the I/O completion is not sent until the data has been stored to physical media on that site.

Note: The mirror repository logical drive can queue up to 128 I/O requests. Until the maximum number has been reached, the mirrored pair state is in a *synchronized* state. If the maximum number of unsynchronized I/O requests is exceeded, the state of the mirrored pair changes to *unsynchronized*.

The host can continue to issue write requests to the primary logical drive, but remote writes to the secondary logical drive will not take place. The requests are stored in the remote mirror repository on the primary site (*delta logging*).

When a controller receives a read request from a host system, the read request is handled on the primary Storage System and no communication takes place between the primary and the secondary Storage Systems.

Global mirroring (asynchronous mirroring with write consistency group)

Global mirroring is an asynchronous write mode where the order of host write requests at the primary site is preserved at the secondary site. This mode is also referred as *asynchronous mirroring with write consistency group*.

To preserve the write order for multiple mirrored volumes, global mirroring uses the *write consistency group* functionality. It tracks the order of the host write requests, queues the order of the I/O requests in the mirror repositories, and sends them to the remote controller in the same order.

> **Important:** Selecting write consistency for a single mirror relationship does not change the process in which data is replicated. More than one mirror relationship must reside on the primary Storage System for the replication process to change.

The volumes for which the write request order must be preserved have to be defined as members of a write consistency group. The write consistency group can be defined from the Storage Manager GUI.

When the owning controller of the primary logical drive receives a write request from a host, the controller first logs information about the write on the *mirror repository logical drive*. It then writes the data to the primary logical drive. The controller then initiates a remote write operation to copy the affected data blocks to the secondary logical drive at the remote site. The remote write request order corresponds to the host write request order.

After the host write to the primary logical drive is completed and the data has been copied to the secondary logical drive at the remote site, the controller removes the log record from the mirror repository logical drive. Refer to Figure 2-40 for a logical view of the data flow.

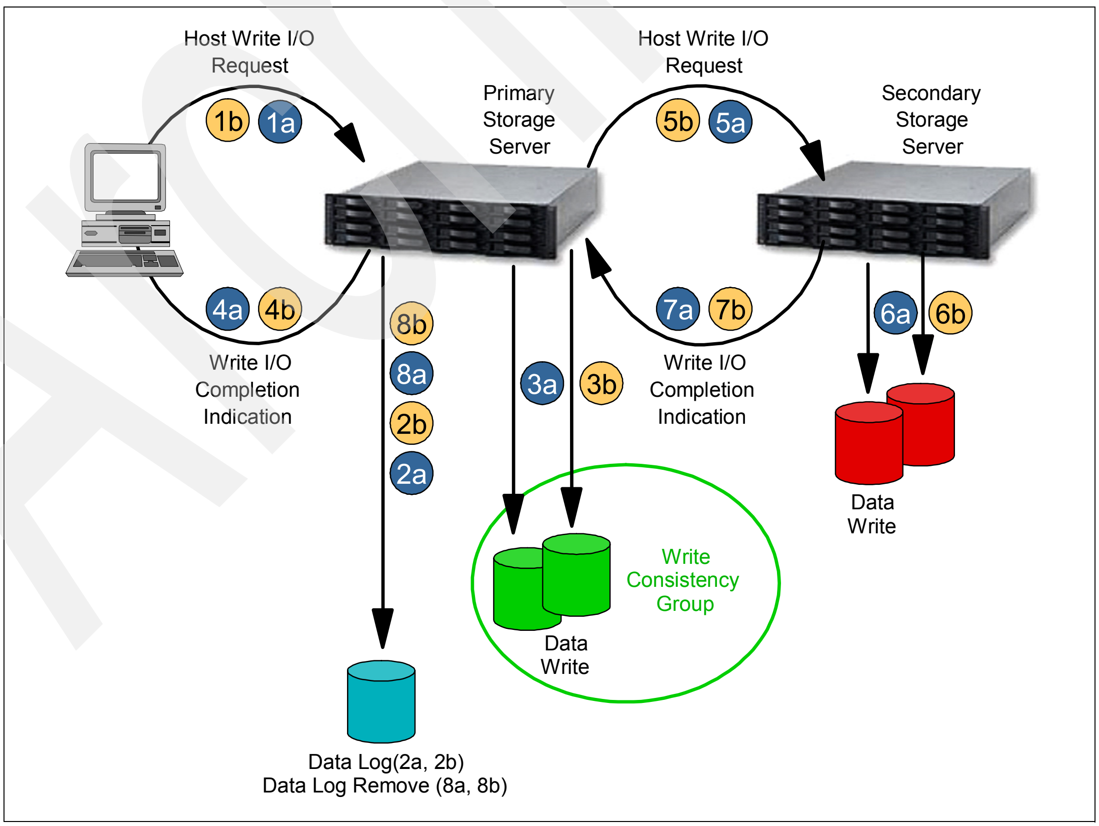


Figure 2-40 Global mirroring logical data flow

When write caching is enabled on either the primary or the secondary logical drive, the I/O completion is sent as soon as data is in the cache on the site (primary or secondary) where write caching is enabled. When write caching is disabled the I/O completion is not sent until the data has been stored to physical media.

> **Note:** The mirror repository logical drive can queue up to 128 I/O requests. Until the maximum number has been reached, the mirrored pair state is in a *synchronized* state. If the maximum number of unsynchronized I/O requests is exceeded, the state of the mirrored pair changes to *unsynchronized*.
>
> The host can continue to issue write requests to the primary logical drive, but remote writes to the secondary logical drive will not take place. The requests are stored in the remote mirror repository on the primary site (*delta logging*).

Whenever the data on the primary drive and the secondary drive becomes unsynchronized, the controller owner of the primary drive initiates a changed data synchronization.

For more information about the enhanced remote mirror feature refer to the *Copy Services User's Guide - IBM System Storage DS4000 Storage Manager*:

https://www-304.ibm.com/systems/support/supportsite.wss/docdisplay?lndocid=MIGR-61173&brandind=5000028

3

IBM i Midrange Storage solution examples

In this chapter we provide representative solution examples of using IBM System Storage DS® Midrange Storage Systems via IBM Virtual I/O Server with IBM i.

The scenarios that we cover in the following sections are:

- 3.1, "VIOS on IBM internal Drives" on page 84
- 3.2, "VIOS on SAN Storage" on page 85
- 3.3, "Dual VIOS with IBM i Mirroring" on page 86
- 3.4, "Disaster Recovery with Enhanced Remote Mirroring" on page 87

3.1 VIOS on IBM internal Drives

This solution scenario (Figure 3-1) describes the most common configuration with one IBM Virtual I/O Server (VIOS) being installed on IBM POWER Systems POWER6 server internal disk drives and IBM i client using virtualized storage from VIOS attached to IBM System Storage DS3400, DS4700, DS4800, or DS5000.

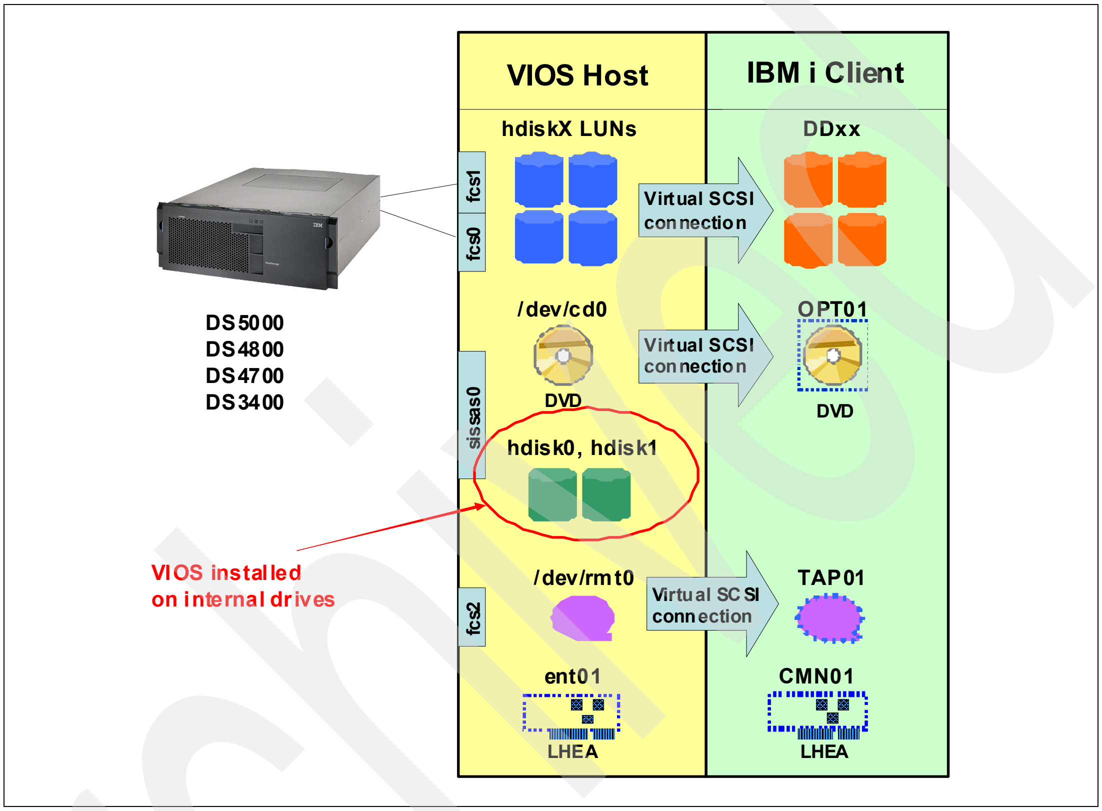


Figure 3-1 VIOS on internal disk drives

Since the IBM i client needs a DVD drive configured as an alternate restart device it makes sense to assign the CEC's embedded I/O controller to VIOS with installing VIOS on the CEC's internal SAS drives and using the virtual SCSI support of VIOS to virtualize the DVD drive for the IBM i client. Also, since VIOS does no virtual I/O caching and uses a fast remote DMA data transfer through the POWER hypervisor between its physical devices and its client partitions there are no high-performance demands for the disk units used for installation of VIOS, which is another argument for VIOS installation on SAS internal disk drives.

3.2 VIOS on SAN Storage

Here we show a configuration similar to that described in 3.1, “VIOS on IBM internal Drives” on page 84, except that VIOS is not installed on IBM POWER Systems internal disk drives, but instead on a SAN-attached IBM System Storage subsystem, as shown in Figure 3-2. The SAN Storage System used for VIOS installation itself can be the same IBM Midrange Storage System as used for IBM i virtualized storage or any other IBM Storage System supported by VIOS.

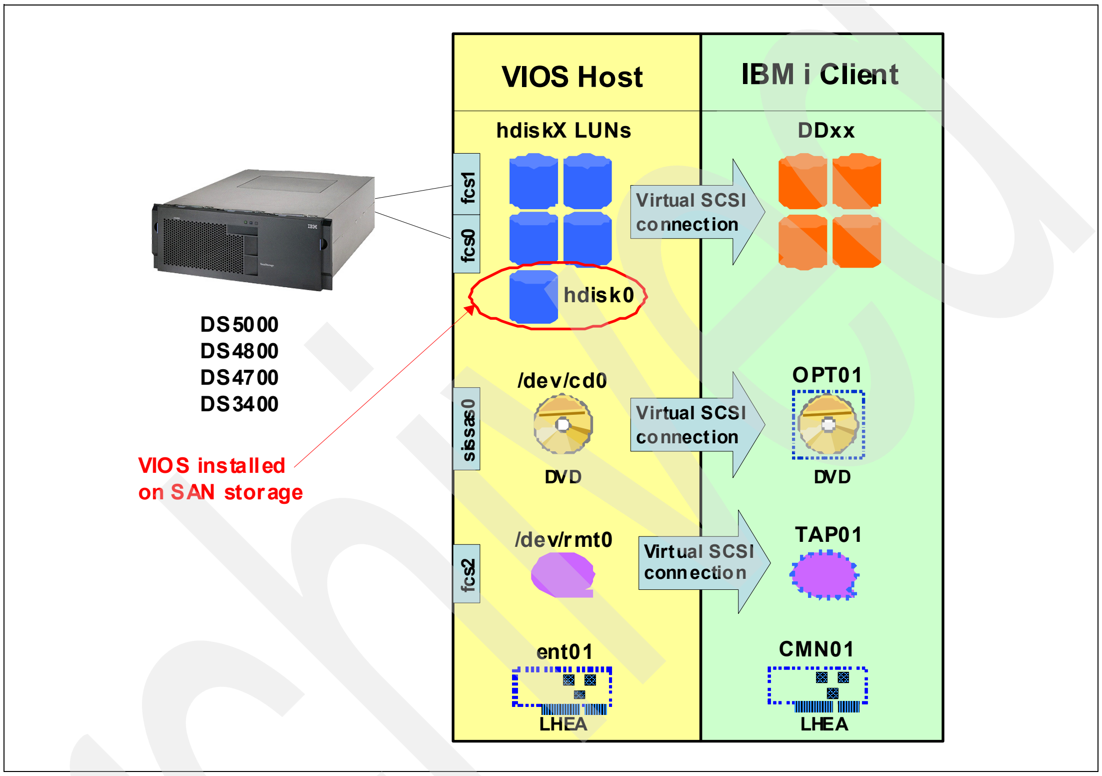


Figure 3-2 VIOS on SAN storage

From a functional perspective there is no difference between VIOS installed on internal disk drives and SAN-attached storage. However, since the IBM i client partition needs to be configured with a DVD drive as an alternate restart device and typically this DVD drive is controlled by the same I/O controller that attaches to the internal disk drives so that the internal disks and the DVD cannot be assigned to different partitions, the reasonability of choosing to install VIOS on SAN is questionable unless there are no internal disk drives attached to the DVD drive I/O controller.

3.3 Dual VIOS with IBM i Mirroring

The solution scenario shown in Figure 3-3 represents a configuration with two redundant IBM Virtual I/O Servers, each providing virtualization of different IBM System Storage logical volumes and IBM i performing software mirroring across them. Such a redundant VIOS solution provides the benefit of IBM i client protection against VIOS planned or unplanned outages, however, at the cost of additional resources required for installation and maintenance of a second VIOS server, which can be installed either on IBM POWER Systems internal drives or SAN storage, and also for twice the disk capacity required on the IBM Midrange Storage Systems. The two VIOS servers are ideally attached to two separate IBM Midrange Storage Systems to also protect against unplanned Storage System outages.

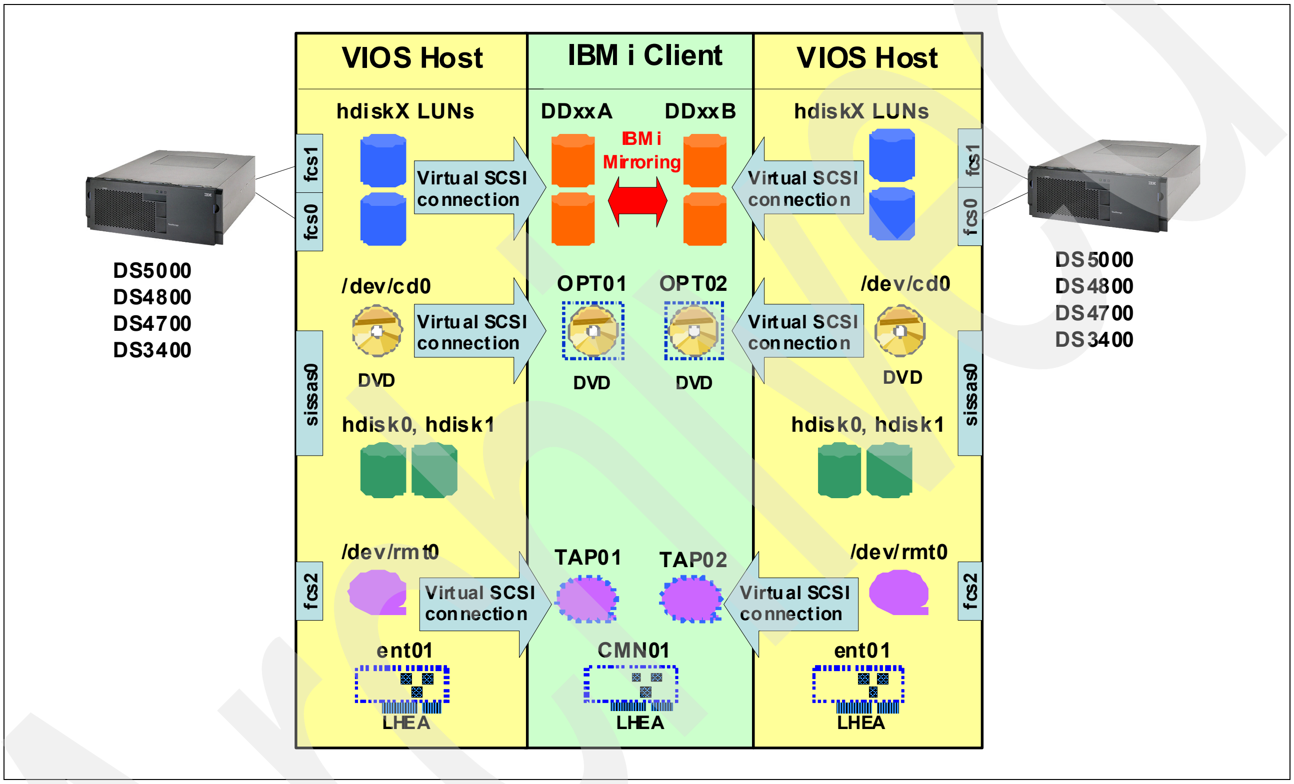


Figure 3-3 Dual VIOS with IBM i Mirroring

> **Note:** Using IBM i multi-pathing across two IBM Virtual I/O Servers is currently not supported.

3.4 Disaster Recovery with Enhanced Remote Mirroring

With the IBM System Storage DS4700, DS4800, or DS5000 Enhanced Remote Mirroring feature (see 2.8.5, "Enhanced Remote Mirroring" on page 77) allowing storage-based replication between two DS Storage Systems the complete disk space used by the IBM i client partition can be replicated to a remote location for implementing disaster recovery with IBM i Midrange Storage, as shown in Figure 3-4.

Note: Enhanced Remote Mirroring (that is, remote replication) is not available on the IBM System Storage DS3000 series models.

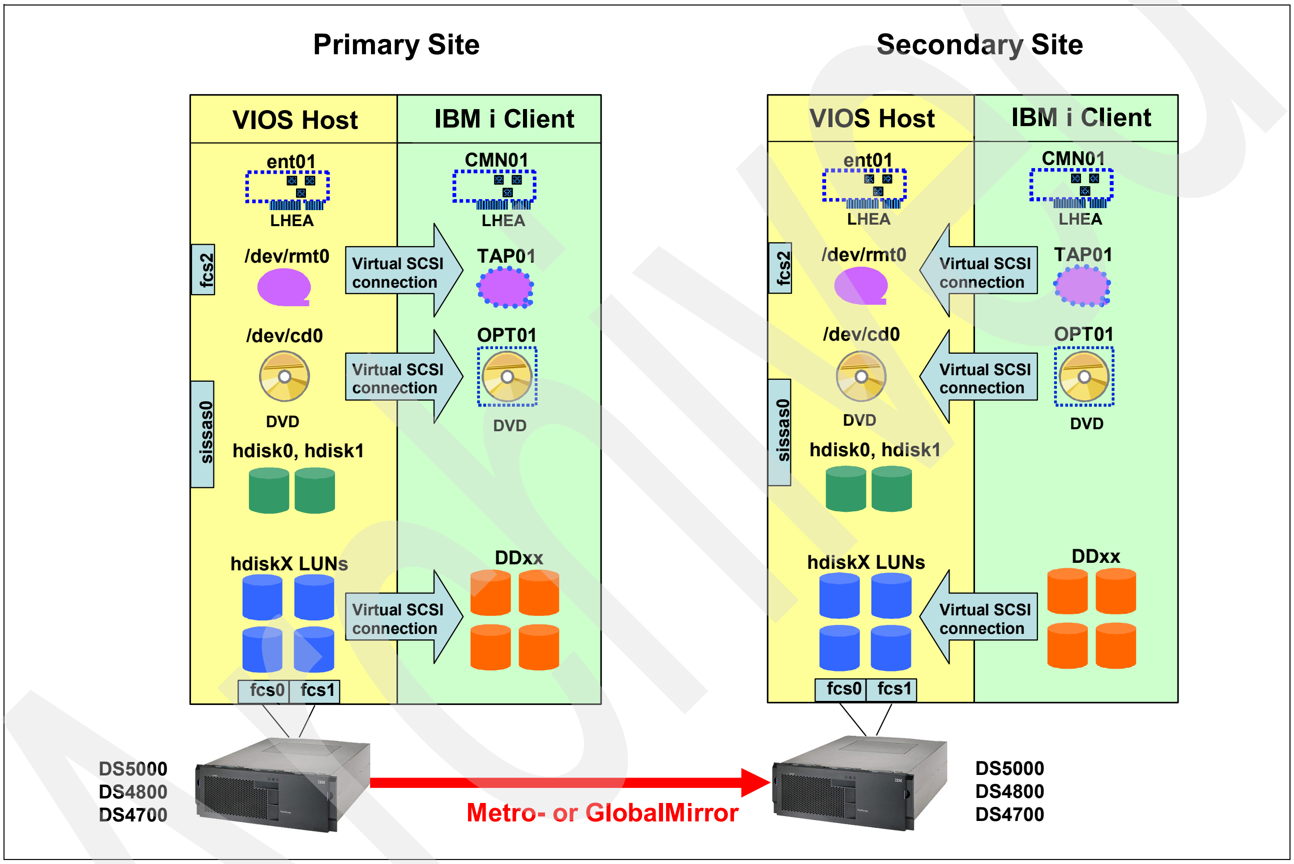


Figure 3-4 Disaster Recovery with DS4000 Enhanced Remote Mirroring

In case of an IBM i client, VIOS, or Storage System outage at the primary site, the standby IBM i client at the secondary (recovery) site can be IPLed from the metro mirror or global mirror target volumes.

For this disaster recovery solution scenario we recommend having a VIOS server at the secondary site up and running, that is, not including VIOS in the storage-based replication, so that when switching to the secondary site the recovery time is not extended by the need to IPL a VIOS server before being able to restart the production with IPLing the IBM i client.

For further information about using IBM Midrange Storage Copy Services with IBM i refer to 4.7, "Planning for Copy Services" on page 102, and Chapter 8, "Using Midrange Storage Copy Services" on page 283.

Note: Currently, only IBM i client full-system replication is supported with IBM i Midrange Storage, that is, no individual replication of independent auxiliary storage pool (IASP) is supported.

4

Planning for Midrange Storage

This chapter discusses important considerations for properly planning IBM System Storage DS3400, DS4700, DS4800, and DS5000 storage attachment to IBM i clients of IBM Virtual I/O Server.

Thoroughly planning the intended IBM i Midrange Storage solution is important to help ensure that the later implemented solution is an IBM supported configuration that meets the customer's requirements with regards to I/O connectivity, capacity, performance, usability, and reliability/availability.

Each of these topics is discussed in detail in the following sections:

4.1 Planning for IBM Midrange Storage series and models

To plan for IBM Midrange Storage you may first consider the series and model of Midrange Storage to use for a particular IBM i data center.

All models of DS3400, DS4700, DS4800, and DS5000 are supported by IBM i as a client of VIOS on IBM POWER system Power6 servers.

The minimal firmware levels to support IBM i partition are shown in Table 4-1.

Table 4-1 Firmware levels

	DS3400	DS4000	DS5000
Minimal firmware level	06.70.69.00	06.60.08.00	All firmware levels

To position different models refer to 2.2, "IBM DS4000 models and features" on page 14. Also, Table 4-2 may be useful in determining the most suitable DS4000 model for a System i workload. However, due to sensitivity of System i workload, we do not recommend loading the DS4000 to the maximum values specified in the table. These values should serve as a guideline for deciding for appropriate Midrange Storage series.

Table 4-2 Suitable DS model for an IBM i workload

Storage System	Recommended maximal capacity to use for System i server
DS3400	4 TB
DS4700	9 TB
DS4800	20 TB
DS5000	30 TB

Cache size is also a crucial factor to determine the Midrange Storage to plan for IBM i. For more information about deciding cache size refer to 5.2.4, "Cache size" on page 112.

For more information about planing the DS4000 refer to the *DS4000 Best Practices DS4000 Best Practices and Performance Tuning Guide*, SG24-6363.

Later in the chapter we discuss planning for the different areas of Midrange Storage connection to System i partition.

4.2 Planning for VIOS

IBM Virtual I/O Server (VIOS) is a special software appliance tied to IBM POWER Systems, that is, the converged IBM System i and System p® server platforms. It is included in the comprehensive virtualization offering of PowerVM Standard or Enterprise Edition, which are licensed on a POWER system processor basis. PowerVM comes shipped with a VIOS installation DVD and an authorization code that needs to be entered on the HMC before a VIOS partition can be created.

Note: While PowerVM and VIOS themselves are supported on both POWER5 and POWER6 systems, IBM i, being a client of VIOS, is supported only on POWER6 systems.

Minimum system hardware and software requirements

The minimum system hardware and software requirements for IBM i being a client of VIOS are:

- IBM POWER Systems POWER6 server model
- System firmware 320_040_031 or later
- PowerVM Standard (PID 5765-PVS) or Enterprise Edition (PID 5765-PVE)
- VIOS 1.5.0 with fixpack 10.1 or later
- HMC V7 R3.2.0 or later
- IBM i 6.1 or later

VIOS configuration best practices

For planning IBM Midrange Storage DS3400, DS4700, DS4800, or DS5000 with IBM i as a client of the IBM Virtual I/O Server, the following configuration best practices should be considered as rules of thumb with regards to VIOS:

- 0.25 CPUs per each 10,000 I/O of the virtual SCSI client. For lowest I/O latency preferably use a dedicated processor for VIOS.
- 1 GB main memory for VIOS.
- Two or more FC adapters assigned to VIOS for multi-pathing. (See also 4.3, “Planning for SAN connectivity” on page 91.)

For further information about VIOS sizing refer to the IBM Hardware Information Center at:

```
http://publib.boulder.ibm.com/infocenter/systems/scope/hw/index.jsp?topic=/iphb1/i
phb1_vios_planning_vscsi_sizing.htm
```

4.3 Planning for SAN connectivity

Since Midrange Storage connects to an IBM i partition via VIOS, actual SAN connectivity is done between Midrange Storage and VIOS, while client IBM i partition uses virtual I/O connectivity through the POWER hypevisor to VIOS. In this section we point to the SAN parts and features specific to Midrange Storage and VIOS, which are important for planning the connections.

4.3.1 Adapters, Fibre Channel protocols, distance

Host ports in DS3400, DS4700, DS4800, and DS5000 operate at a speed of 4 Gbps. However, a host port auto-negotiates also to support 2 Gbps and 1 Gbps speeds of connections. If multiple devices and switches with different speeds are connected to the host port, it automatically detects the lowest speed of connection and adjusts to it.

The following Fibre Channel adapters in VIOS in System i Power6 are supported to connect to DS3400, DS4700 and DS4800, and DS5000:

- 4 Gb Fibre Channel PCI-X 2.0 DDR Adapter (Single-Port), feature number 5758
- 4 Gb Fibre Channel PCI-X 2.0 DDR Adapter (Dual-Port), feature number 5759
- PCI-E 4x4 Gb Fibre Channel (Single Port), feature number 5773
- PCI-E 4x4 Gb Fibre Channel (Dual Port), feature number 5774

Host ports in DS3400, DS4700, DS4800, and DS5000 support switched fabric, arbitrated loop, and point-to-point protocols. All three FC protocols are also supported by host adapters in VIOS. Connection between Midrange Storage and adapters in VIOS can be done in switched fabrics and arbitrated loop protocols, direct connection being treated as a single connection arbitrated Loop. Host ports in DS3400, DS4700, DS4800, and DS5000 automatically connect in the Fibre Channel (FC) protocol that they detect in the connection.

4.3.2 Cables and connectors

Both shortwave and longwave FC cables are supported to connect DS3400, DS4700, DS4800, and DS5000 to VIOS host adapters.

Host ports in Midrange Storage and host adapters in VIOS are equipped with transceivers (Small Form Pluggable, SFP) used to convert the internal communication transport of gigabit transport. Lucent Connectors (LC connectors) are used to plug in to SFPs on both Midrange Storage host adapters and VIOS adapters to connect Fibre Channel cables. Both longwave an shortwave SFPs are supported on Storage System host adapters as well as adapters in VIOS.

4.3.3 Switches and zoning

Consider Midrange Storage interoperability matrixes to look for supported switches to connect Storage Systems to host adapters in VIOS. Interoperability matrixes for each Midrange Storage model can be found on the following Web page:

http://www-03.ibm.com/systems/storage/disk/index.html

By best practise we recommend connecting and zoning the switches as follows:

1. Spread VIOS host adapters or ports equally between the two Midrange Storagecontrollers.
2. Spread VIOS host adapters or ports equally between the switches.
3. For each host adapter or port create a separate zone containing the host adapter (or port) and the host port on the Midrange Storage controller.

An example of zoning is shown on Figure 4-1.

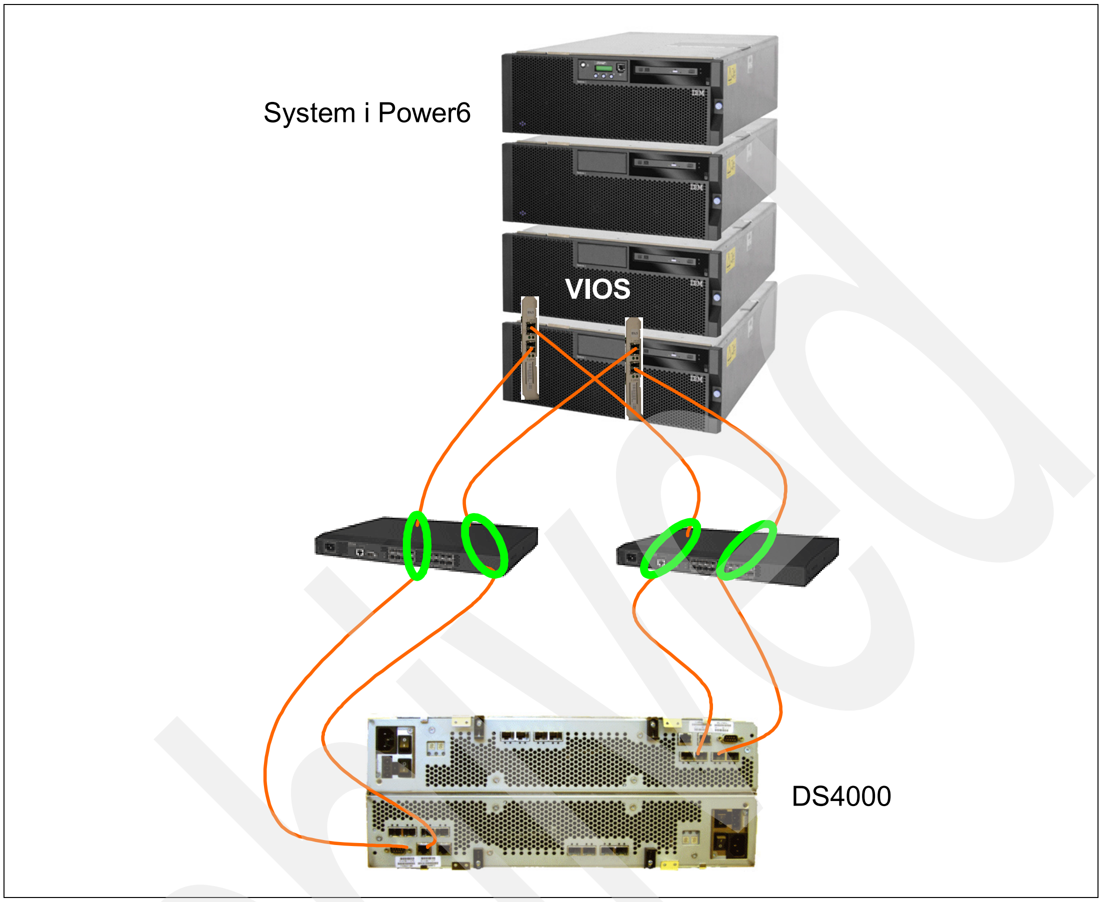


Figure 4-1 Zoning switches for DS4000 connection

4. We recommend connecting disk and tape via different host ports.
5. Create separate switch zones for disk and tape connection.

4.3.4 Multipath

Multipath, which is the ability to present logical volumes on a storage device to the host server via multiple connections, provides redundancy in case an I/O path fails. With Midrange Storage connected to IBM i client via VIOS, it is possible to implement multipath so that a logical volume connects to VIOS via multiple *physical* host adapters (or ports) in VIOS. However, *virtual* SCSI adapters are used in single path. Refer to 3.3, “Dual VIOS with IBM i Mirroring” on page 86, for more information about how to ensure redundancy of VIOS servers and consequently virtual SCSI adapters.

The following multipath drivers are supported in VIOS for Midrange Storage connection to a System i partition:

- Redundant Disk Array Controller (RDAC): VIOS supports RDAC for DS4000 series.
- Multipath Input/Output (MPIO): VIOS supports MPIO for the DS3000, DS4000, and DS5000 series.

Both drivers RDAC and MPIO provide the ability to detect a physical device in Midrange Storage by a LUN, manage all of the paths to a physical device, to present a single instance of a LUN to host system, and to detect a failure of a path to LUN and fail over to another path. For more information about RDAC and MPIO refer to *DS4000 Best Practices and Performance Tuning Guide*, SG24-6363.

When you create a logical drive in Midrange Storage, you assign one of the two active controllers to own the logical drive and to control the I/O between the logical drive and the host server along the path. This *preferred controller* normally receives the I/O requests from the logical drive. If a problem along the I/O path causes an I/O to fail, the multipath driver issues the I/O to the alternate controller.

An RDAC driver can manage no more than two paths from and to a LUN. MPIO, however, can use more than two paths from/to a LUN, and so provides additional redundancy in case an adapter in VIOS fails or a host port in the Storage System fails. Two paths supported by RDAC and multiple paths supported by MPIO are illustrated on Figure 4-2.

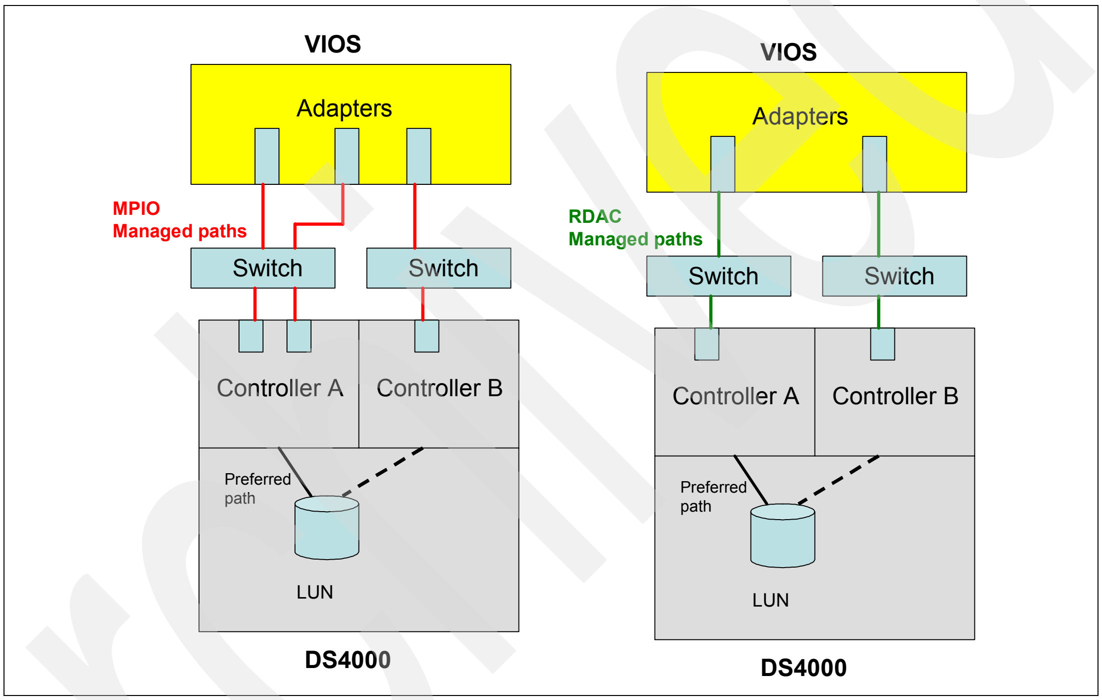


Figure 4-2 Paths managed by MPIO and RDAC

Regarding connection of an IBM i client, VIOS supports both drivers MPIO and RDAC for DS4700 and DS4800, and only MPIO for DS5000. VIOS 2.1 scratch install uses MPIO by default while VIOS V1.5.x uses RDAC by default for DS4000 attachment.

We expect that the customers will migrate from RDAC to MPIO while upgrading the version of VIOS from V1.5 to V2.1. For more information about the migration procedure refer to 9.2, “VIOS maintenance” on page 346.

4.4 Planning for capacity

When planning for capacity of Midrange Storage with IBM i, the most important question is: How much disk capacity is usable for System i partition after Midrange Storage is set up and connected via VIOS?

In this section we provide needed information to correctly plan the disk capacity in Midrange Storage so that the usable disk space for the System i partition meets customer requirements. We explain the net capacity in different RAID levels and usable capacity after connecting Midrange Storage via VIOS.

4.4.1 Disk drive capacities and spares

Disk drives in DS3000, DS4000, and DS5000 are located in the controller enclosure and expansion enclosures. The number of drives in an enclosure depends on the model of system storage and expansion. It can be 12, 14, or 16. For more information about the number of disk drives in an enclosure refer to Chapter 2, “Midrange Storage architecture” on page 9.

The following disk drive sizes are available in Midrange Storage:

- Fibre Channel drives, or SAS drives
  - 36.4 GB: Rotation speed 15 K RPM
  - 73.4 GB: Rotation speed 15 K RPM or 10 K RPM
  - 146.8 GB: Rotation speed 15 K RPM or 10 K RPM
  - 300 GB: Rotation speed 10 K RPM
- SATA drives
  - 250 GB: Rotation speed 7.2 K RPM
  - 400 GB: Rotation speed 7.2 K RPM
  - 500 GB: Rotation speed 7.2 K RPM
  - 1000 GB: Rotation speed 7.2 K RPM

The available disk capacities depend on the model of DS3000, DS4000, and DS5000.

> **Note:** We strongly recommend using Fibre Channel drives for System i daily workload, while SATA drives are appropriate for archiving a non I/O-intensive workload.
>
> SATA drives are supported by IBM i as a client of VIOS only on IBM POWER servers. They are not supported on BladeCenter JS22 and JS12.

The number and location of spare disk drives in Midrange Storage are determined by the user. You decide how many spare drives to define based on needed availability, RAID level used, number of enclosures used, and so on. For recommendations about spares refer to 2.5, “RAID technology and spare drives” on page 36.

4.4.2 RAID levels

The following RAID levels are available in Midrange Storage:

- RAID-0
- RAID-1
- RAID-3
- RAID-5
- RAID-6
- RAID-10

The available RAID levels depend on the model of DS300, DS4000, and DS5000. The number of disk drives in a RAID array is determined by the user. You decide the number of disk drives in an RAID array based on availability and capacity needs. However, take into account the minimal number of drives in each type of RAID level, for example, the minimal number for RAID-5 is three drives. For more information about RAID protection refer to 2.5, “RAID technology and spare drives” on page 36.

4.4.3 DACstore

DACstore is an area on each drive in a Midrange Storage subsystem or expansion enclosure where configuration information is stored. The disk capacity belonging to DACStore is reserved, invisible to a user, and contains information about the Midrange Storage configuration. The DACStore size depends on the level of Midrange Storage firmware, with current microcode level 512 MB reserved for it.

> **Note:** Since the DSCStore size may increase with future microcodes we recommend that you not fully utilize the available space in an array.

4.4.4 Usable capacity in Storage System

After you decide the number of spares, RAID level, and size of RAID arrays, calculate the usable capacity of Storage System with the following steps:

1. From the number of all disk drives subtract the number of spares and the possible capacity of parity or mirrored data to get the equivalent number of usable data disk drives.

   For example, calculate the equivalent data drives of DS4800 with four expansion enclosures EXP810 with 16 disk drives each, four spare drives, and three RAID-10 arrays of 20 disk drives.

   From all available drives subtract spares and mirrored capacity:

   ```
   64 - 4 - 3 * 10 = 30
   ```

2. For each data drive use the following formula to calculate usable capacity:

   a. Convert decimal capacity of the drive (in GB) to binary:

      ```
      Decimal capacity * 0.9313
      ```

   b. Convert DacStora capacity 0.512 GB to binary:

      ```
      0.512 * 0.9313 = 0.477 GB
      ```

   c. Subtract 0.477 for DACStore to get binary usable capacity.

   d. Convert binary usable capacity to decimal:

      ```
      Binary usable capacity * 1.0737
      ```

   For example, a 73.4 GB disk drive provides the following usable capacity:

   ```
   73.4 * 0.9313 - 0.477 = 67.8
   67.8 * * 1.0737 = 72.8
   ```

IBM Storage lab provides the space calculator for Midrange Storage to help you calculate usable capacity of a specific configuration. It can be downloaded from the following Web page:

http://w3-03.ibm.com/sales/support/ShowDoc.wss?docid=D272029U21918Z20&infotype=SK&infosubtype=S0&node=doctype,S0%7Cdoctype,STL%7Cbrands,B5000%7Cgeography,AMR&appname=CC_CFSS

> **Note:** Disk capacity available to the System i partition is about 11% smaller than calculated usable capacity, as is discussed in 4.4.5, “Available capacity to IBM i partition” on page 97.

4.4.5 Available capacity to IBM i partition

The existing IBM i storage portfolio includes integrated disks that support 520 bytes per sector, and FC-attached Storage Systems that support 520 bytes per sector, such as IBM DS8000. Midrange Storage only supports 512 bytes per sector. Therefore, it is needed to convert 520 bytes per sector data layout to 512 byte per sector when connecting Midrange Storage to IBM i client.

Basically, conversion is done the following way: For every page in IBM i (8*512-byte sectors) an extra 512-byte sector is allocated. The extra sectors contain the information previously stored in the 8-byte sector headers. For more information about sector conversion refer to 1.1, “What is new” on page 2.

Because of this process, you must allocate nine sectors of Midrange Storage for every eight sectors in IBM i. From a capacity point of view, the usable capacity in Midrange Storage is therefore multiplied by 8/9 or app.0.89, which means that it is reduced by about 11%, when the logical volumes report to an IBM i client.

For example, a System i data center implements DS4000 with 60 * 73.4 GB disk drives in RAID-10. Their usable capacity is 30 * 72.8 (usable capacity of one 73 GB data drive) = 2184 GB. After the logical volumes are created and connected to IBM i, the capacity available to System i will be 0.89 * 2184 GB = 1944 GB.

4.5 Planning considerations for performance

To achieve good performance of System i server with Midrange Storage, you must carefully plan the disk drives, RAID levels, sector sizes in Storage System, size of logical drives, and number of adapters in VIOS and System i partition. In this section we discuss each of these.

4.5.1 Disk drives in Midrange Storage

It is very important to provide IBM i production workload with enough disk arms. The reason for this is that System i architecture uses built-in storage management to manage I/O and data on disk devices. For more information about System i Storage Manager functions refer to 5.2, “Sizing rules for Midrange Storage” on page 110.

To ensure that enough disk arms are used by System i, and at the same time stay within reasonable capacity limits, we recommend planning smaller sizes of disk drives (for example, 36.4 GB, 73.4 GB, or 146.8 GB) for IBM i.

System i transaction processing workload usually requires high rotation speed of disk drives for good performance.

> **Note:** We generally recommend using 15 K RPM Fibre Channel drives for performance-critical production workload.

During the planning phase of DS3000, DS4000, or DS5000 for System i, we recommend obtaining characteristics of System i workload, like I/O rate, read/write ratio, cache hits, and transfer size. Based on them you may want to precisely calculate the needed number of disk drives for a specific RAID level, as is described in Chapter 5, “Sizing for Midrange Storage” on page 107.

Some System i data centers may decide to connect Midrange Storage for archiving their data. In this case it is more important to ensure enough capacity for performance, so it may be a good idea to plan for SATA drives.

4.5.2 RAID arrays

When considering the RAID level of Midrange Storage with System i partition, take into account general RAID performance guidelines:

- RAID-0 offers high performance, but does not provide any data redundancy.
- RAID-1 offers high performance for write-intensive applications.
- RAID-3 is good for large data transfers in applications, such as multimedia or medical imaging, that write and read large sequential chunks of data.
- RAID-5 is good for environments where the typical transfer size is small and there is a high proportion of read activity.
- RAID-6 is suitable for workloads with a very high proportion of read activity and a small transfer size. This offers better availability than RAID-5.
- RAID-10 provides better reliability than RAID-5.

The maximum array size is 30 data drives.

For further information about IBM RAID technology refer to 2.5, “RAID technology and spare drives” on page 36.

4.5.3 Sector sizes

From a performance perspective we suggest choosing the DS storage subsystem sector size (strip size on one RAID member disk) such that it is bigger than the IBM i I/O transfer size, to not engage more than one disk arm per I/O request. For example:

- If the customer’s workload experiences maximum blocksize 12 KB, you may want to set up a sector size of 16 KB.
- Many installations experience a blocksize lower than 16 KB during daily transaction workload, while their nightly batch job results in transfer sizes about 100 KB to 120 KB. For them it may be a good idea to create RAID-10 arrays of four disk drives and define a sector size 64 KB. This way the transfers of transaction workload are accommodated in one disk drive, and large batch job blocksizes fit into a stripe across two drives.

4.5.4 Logical drives

The maximum available size of a logical drive for System i is 2 TB - 512 bytes. But for performance reasons we recommend much smaller sizes, as is described further in this section.

We expect that many installations will use a RAID-10 or RAID-1 level for high availability. As was measured in the System i lab in Rochester, smaller RAID-10 arrays provide better performance than bigger RAID-10 arrays, so we recommend creating RAID-1 arrays with two disk drives, or RAID-10 arrays with four disk drives.

Figure 4-3 shows the location of logical drives on a physical drive and movements of read/write head. On one big logical drive the head makes small movements usually in one direction, so the seek time is minimal. While with multiple smaller logical devices the head has to move from one device to another, which causes longer seek time. From this aspect we recommend defining larger logical drives—only a few drives per physical disk or per array.

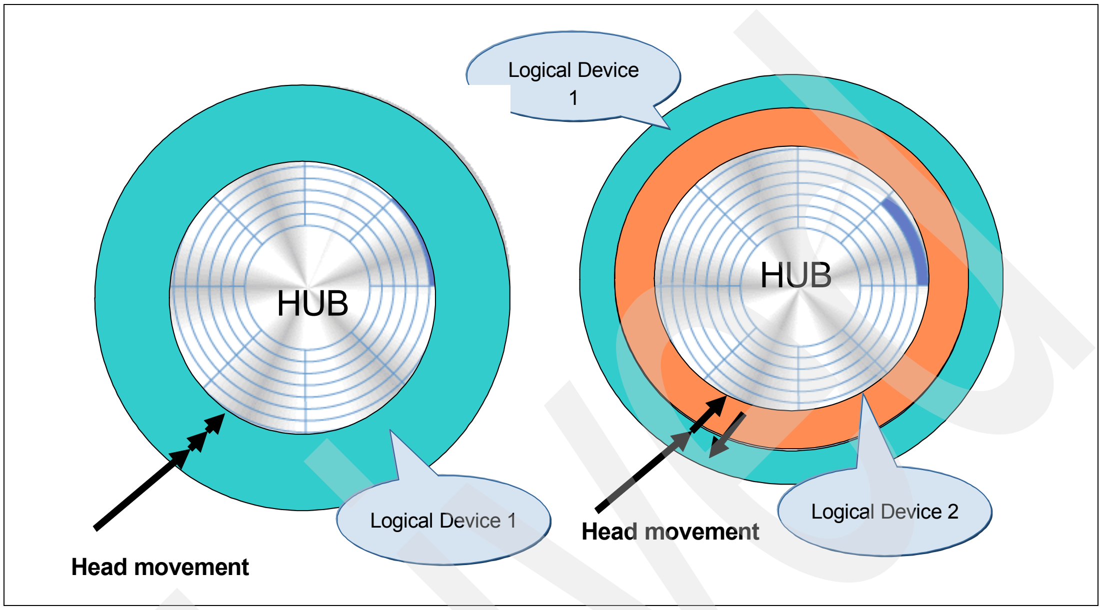


Figure 4-3 Logical drives on a physical disk

On the other hand, we consider the number of System i I/O operations that one logical drive can handle. IBM i provides *SCSI Command Tag Queuing* for LUNs on Midrange Storage, which enables up to 32 I/O operations to one LUN at the same time (in other words, *queue depth* of a LUN is 32). If more than 32 I/O operations are experienced on the same LUN at the same time, the I/Os over 32 are queued for this LUN in a *deferred* queue. The bigger that the deferred queue is the higher the wait time of an I/O operation is. Due to the queue depth 32, we can consider reasonable, large LUNs. Still, LUNs that are too large would cause I/Os to queue in the deferred queue and therefore a longer wait time. Also, large LUNs would cause high LUN utilization and decreased performance. So we recommend defining LUNs no bigger than 300 GB.

Combining both aspects, a) defining large enough LUNs to minimize seek time, and b) still keeping the size of LUNs within the limit to avoid deferred queues and high utilization, we provide the following guidelines:

- Use 73 GB or 146 GB physical disks.
- Make RAID-1 arrays of tow physical disks and create one logical drive per RAID-1 array.
- If the number of LUNs is limited to 16 (for example, connecting to IBM i on BladeCenter) you may want to make a RAID-10 array of four physical disks and create one LUN per array.

For example, define a RAID-1 array of 2 * 146 GB disks and create one LUN per array, or define a RAID-10 array of 4 * 146 GB disks and create one LUN per array.

When using RAID-.5 protection we recommend defining about eight physical disks per array. This is because a larger RAID-5 array will need a long time to rebuild the spare if one disk drives fails, and also because the probability that two drives will fail is lower with bigger arrays.

Make sure that logical drives for an IBM i client are evenly spread between the two controllers in Storage System.

4.5.5 Adapters in VIOS and IBM i client

It is possible to connect up to 4,095 LUNs per target and up to 510 targets per port in a physical adapter in VIOS. However, for performance reasons you should plan the number of LUNs for one port the following way: first determine the number of ports needed for a workload so that the ports do not get saturated. Based on this, decide how many LUNs to assign to one port.

To determine the needed number of physical adapters in VIOS, consider the guidelines in Table 4-3. They are modelled by Disk Magic™ for DS4700 with 4 GB cache, using 4 Gb connections to the server.

Table 4-3 Physical Adapters in VIOS

% of cache hits	Blocksize (KB)	Maximum IO/sec per on port in VIOS	Max MBps per one port in VIOS
50%	16	8929	140
30%	16	7783	122
50%	120	2253	264
30%	120	1976	231

Regarding the modelled values, we expect that two physical adapters in VIOS, one port in each of them connected to Storage System, are sufficient for most of the System i workloads. You may consider more than two adapters only for significant workloads of big blocksizes.

It is possible to assign 16 LUNs to each virtual SCSI adapter in VIOS and in IBM i client. We do not consider the throughput of virtual adapters a performance issue, so we recommend connecting up to 16 LUNs to each virtual SCSI adapter.

4.6 Planning for backup and recovery

We recommend that you plan regular System i backups at the same time that you plan connecting Midrange Storage to System i partition. This is because using VIOS provides more possibilities for backups. On the other hand, there are some limitations for IBM i client on Blade that you should be aware of. Later in this section we describe the possible ways to back up and restore IBM i client with Midrange Storage attached via VIOS.

4.6.1 Back up to physical tape drive in IBM i client partition

When System i partition that has Midrange Storage attached, it resides on the Power6 System i model. You can use the usual way of backup to the *physical* tape drive attached to this partition. You may want to use one tape drive for the backup of multiple System i partitions so that you can move the tape drive from one partition to another by Dynamic Logical Partitioning (DLPAR).

Note that this type of backup is not possible when running IBM i client on Power6 Blade.

4.6.2 Back up to virtual optical volume in VIOS

This option for backing up System i data is available to an IBM i client in System i POWER6, as well as in POWER6 Blade.

In order to save from IBM i client you must set up a *Virtual Media Library* with *virtual optical media* in VIOS, and assign a virtual optical volume to the virtual optical device in IBM i partition. Then you use BRMS or IBM i commands to perform backup to the IBM i virtual optical device to which the virtual optical volume in VIOS is assigned. After saving to virtual optical media is done, save from virtual optical media to physical tape drive attached to VIOS.

This two-phase process of saving IBM client on Power6 Blade is shown on Figure 4-4.

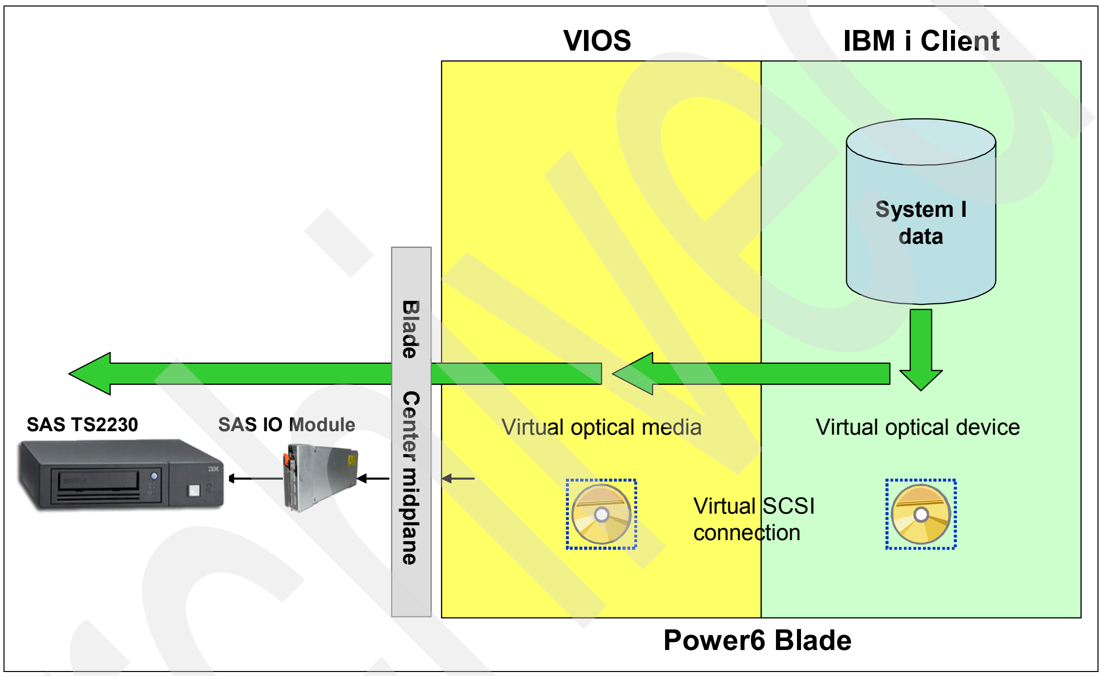


Figure 4-4 Saving to virtual optical media in VIOS

To restore data to the IBM i client, first restore from the physical tape drive to the virtual optical media in VIOS, then restore from the virtual optical device in the IBM client.

Alternatively, to save to physical tape in VIOS you can use Tivoli® Storage Manager (TSM) to perform a save from virtual optical media to real tape. For this, VIOS needs to be set up as a TSM client.

When planning for this type of backup consider the following:

- With VIOS and IBM i clients on System i Power6, any physical tape drive that attaches to VIOS through any supported FC adapter can be used to save data from virtual optical media.
- With VIOS and IBM i client on POWER6 Blade, the physical tape drive to which we save from VIOS is limited to SAS connected TS2230.
- This scenario can be used for both full system backup and backup of just IBM i libraries and objects.

- The virtual optical device in IBM i client to which a virtual optical media is assigned can be used as an alternate IPL device.
- Plan for enough disk capacity in VIOS to accommodate virtual optical volumes for save and restore.

4.6.3 Back up to virtual tape in System i partition

This option for backing up System i data is available to an IBM i client in System i POWER6, as well as in POWER6 Blade.

With this option you use IBM i virtual tape to save data in a System i partition. For more information about saving to virtual tape refer to *i5/OS V5R4 virtual Tape: A Guide to Planning and Implementation,* SG24-7164, available at:

http://w3.itso.ibm.com/abstracts/sg247164.html?Open

After data are saved to the virtual tape (image catalog) you may FTP the images to another IBM i partition.

When considering this option note that you need disk capacity for virtual tape in the IBM i client and in another IBM i partition, so you need double the capacity of saved data.

4.7 Planning for Copy Services

In this section we provide planning guidelines for Midrange Storage Copy Services with the System i partition. For more information about usage and implementation refer to Chapter 8, “Using Midrange Storage Copy Services” on page 283.

4.7.1 FlashCopy

FlashCopy is a function in Midrange Storage that provides an immediate point-in-time image of a logical drive. The image can be used by host servers for saving data to tape. FlashCopy is available in DS3400, DS4000, and DS5000. The source logical drive from which FlashCopy is done is called *base logical drive*, and the drive that contains the image is called *FlashCopy logical drive*. For more information about FlashCopy in Midrange Storage refer to 2.8.4, “FlashCopy and Volume Copy” on page 70. For instructions about how to implement FlashCopy for System i, refer to 8.1, “Implementing FlashCopy” on page 284.

When using FlashCopy in Midrange Storage with the System i server you are required to copy *all* logical drives that are used by the IBM i client, due to System i single-level storage architecture. After FlashCopy of all logical drives belonging to the production IBM i client is done, a stand-by IBM i client performs IPL from FlashCopy repository logical drives, so the stand-by system becomes a *clone* of the production system. Typically, the clone IBM i client is used to save production data to tape so that save downtime in the production system is minimized. For testing, developing, exploring data, and similar purposes it may be more appropriate to use Volume Copy, as described in 4.7.3, “Volume Copy” on page 105.

We recommend that you consider the points discussed in the following sections when planning for FlashCopy in Midrange Storage with an IBM i client.

Connecting via VIOS

If you plan to protect the production System i partition by connecting it via two VIOS servers and using IBM i mirroring, as is described in 3.3, “Dual VIOS with IBM i Mirroring” on page 86, you will probably want to FlashCopy only one half of the mirrored logical volumes. Plan to connect FlashCopy logical drives via another VIOS server (not any of the VIOS servers used for the production System i partition).

If you plan to connect the production IBM i client via only one VIOS, establish another VIOS for connecting FlashCopy logical drives and so provide protection of data being saved in case VIOS of the production IBM i client fails.

Capacity of FlashCopy repository logical drives

Since the FlashCopy repository logical drive holds only the blocks that were changed in the base logical drive, its capacity can be smaller than the capacity of base logical drive. When planning for FlashCopy, the usual question is: How much disk capacity do I need for FlashCopy repository logical drives? The needed capacity of repository logical drive depends on the number of writes in the production System i partition, the number of re-writes (writing multiple times to the same block), and the duration that the FlashCopy repository logical drives will be needed. A general guideline is to use 20% of base logical drive capacity for the FlashCopy repository logical drive. This is also the default value when establishing FlashCopy in Storage Manager. Twenty percent of the base drive is usually an adequate logical drive capacity for most situations. Two factors might influence making the repository logical drive a higher percentage of the base logical drive:

- Heavy write I/O activity that causes change to a larger percentage of data blocks on the base logical drive.
- The FlashCopy logical drive is needed for a longer period of time.

The size of the FlashCopy repository logical drive can be estimated by calculating and adding together the *copy-on-write space* and the *management overhead space*:

- Copy-on-write space is based on the percentage of data blocks on the base logical drive that are expected to change. For example:

  ```
  30% blocks expected to change * 5 GB total base logical drive size = 1.5 GB
  copy-on-write space needed for the FC repository logical drive
  ```

- Management overhead space is the space required to store FlashCopy data on the FlashCopy repository logical drive. It us calculated using the formula:

  ```
  192 KB + capacity of the base logical drive in bytes / 2000
  ```

The majority of the needed capacity is represented by copy-on-write space, since the space for management overhead can be usually measured in MB. The percentage of blocks expected to change depends on characteristics of the System i workload, that is, the number of writes/sec, re-writes, and the duration of FlashCopy.

Refer also to *DS4000 Best Practices DS4000 Best Practices and Performance Tuning Guide*, SG24-6363, for information about sizing FlashCopy repository logical drives.

Plan for FlashCopy failure policy

If the capacity of the FlashCopy repository logical drive is correctly estimated, we do not expect that the repository will ever become full. Still, if it does become full, one of the following actions can be performed:

- Fail FlashCopy logical drive.

  This is the default action. The FlashCopy repository logical drive becomes full, the software fails the FlashCopy repository logical drive, and the base logical drive remains online and accessible for users. The FlashCopy logical drive is no longer valid.

- Fail writes to base logical drive.

  When the FlashCopy repository logical drive becomes full, the software fails writes to the base logical drive. The FlashCopy logical drive is preserved and remains valid since no new copy-on-write data is generated for the FlashCopy repository logical drive.

The FlashCopy failure policy, which you specify during implementation, determines which of the listed actions should be performed if the FlashCopy repository drive becomes full.

4.7.2 Enhanced Remote Mirroring

The Enhanced Remote Mirroring (ERM) option of Midrange Storage is used for online, real-time replication of data between Storage Systems over a remote distance. In case of failure of the production site a server on a remote site can take over the production work using the copy of data on remote Storage System.

The following mirroring methods are provided by Enhanced Remote Mirroring:

- Metro mirroring: Synchronous mirroring mode.
- Global copy: Asynchronous write mode.
- Global mirroring: Asynchronous mirroring with write consistency group.
- Metro Mirroring: Synchronous mirroring mode is supported for an IBM I partition. It is required to replicate *all* logical drives of a System I partition. In case a production site fails a stand-by System i partition connected to Storage System on the remote site via VIOS is IPLed from a copy of data on the remote Storage System. This brings up a clone of the production IBM i partition on the remote site, and it continues with the production workload.

For more information about Enhanced Remote Mirroring refer to 2.8.5, “Enhanced Remote Mirroring” on page 77. For instructions about how to implement Enhanced Remote Mirroring with System i, refer to 8.3, “Implementing metro mirroring” on page 310.

Take the following points into consideration when planning for metro mirroring of Midrange Storage:

- Metro mirroring is supported on Midrange Storage models DS4700, DS4800, and DS5000.
- You must purchase a ERM Premium Feature License from IBM and enable the Premium Feature on each DS4000 power supply fan unit that is supposed to hold primary or secondary mirrored logical drive.
- On Ds4700 it is possible to have up to 64 mirrored pairs of logical drives. On DS4800 the maximum number of mirrored pairs is 128.
- For metro mirroring, plan one Fibre Channel link to connect each controller in the Storage System to a corresponding controller on the remote Storage System.

- To use metro mirroring it is mandatory to attach the storage servers to a SAN fabric, using Fibre Channel switches or Fibre Channel-IP (FC-IP) routers. This Fibre Channel connection must be a dedicated for metro mirroring between the subsystems.
- SAN switches should be zoned so that the Storage System host ports are separated from ports for metro mirroring. Also, mirroring ports in one Storage System controller should be separated from mirroring ports in the other controller.
- If the distance between local and remote Midrange Storage does not exceed 10 km plan for Fibre Channel links for metro mirroring. For distances longer than 10 km consider FC over IP connection and Fibre Channel-IP routers.
- For estimating needed bandwidth of links for metro mirroring refer to Chapter 5, “Sizing for Midrange Storage” on page 107.
- The Ethernet ports in each Storage System should be on the same subnet or part of a Virtual Private Network (VPN) connection. Such Ethernet configurations are needed in order to enable the server with DS4000 Storage Manager software to access both Storage Systems simultaneously, which is a requirement for metro mirroring.

4.7.3 Volume Copy

Volume Copy is a function in Midrange Storage that is used to copy data from one logical drive (source) to another logical drive (target) in a single Storage System. The target logical drive is an exact copy of the source logical drive. While FlashCopy provides an image of the source drive through metadata pointers, Volume Copy uses the actual copying of data from the source to the target logical drive.

While copying of logical drives is in progress the source logical drive is only available for read I/Os, while write requests are allowed after the logical drive copy is completed. Employing both FlashCopy and Volume Copy enables a complete copy to be created without interrupting the production workload.

> **Note:** For a System i workload it is essential to always use Volume Copy in conjunction with FlashCopy in order to prevent source logical drives from being locked for write I/O operations during the Volume Copy relationship.

For more information about Volume Copy refer to 2.8.4, “FlashCopy and Volume Copy” on page 70. For implementation refer to 8.2, “Implementing Volume Copy” on page 302.

5

Sizing for Midrange Storage

Good and stable performance of IBM i applications is becoming more and more vital for business enterprises using System i servers. To ensure that performance of External Midrange Storage with IBM i will meet customer expectations, we must carefully size resources such as disk arms, adapters, cache, links for data replication, and so on.

In this chapter we provide sizing guidelines for important resources and explain how to use the Disk Magic tool to model performance of the planned configuration. We also describe which information about IBM i partition is needed for sizing and modelling with Disk Magic.

First of all, we describe System i specific features important to understand scenarios and performance of Midrange Storage with this server platform.

5.1 Single-level storage and input/output operations

The Power6 system with IBM i uses the same architectural component as used by i5, iSeries, and AS/400: single-level storage. It sees all disk space and the main memory as one storage area, and uses the same set of virtual addresses to cover both main memory and disk space. Paging in this virtual address space is performed in 4 KB pages. Single-level storage is shown in Figure 5-1.

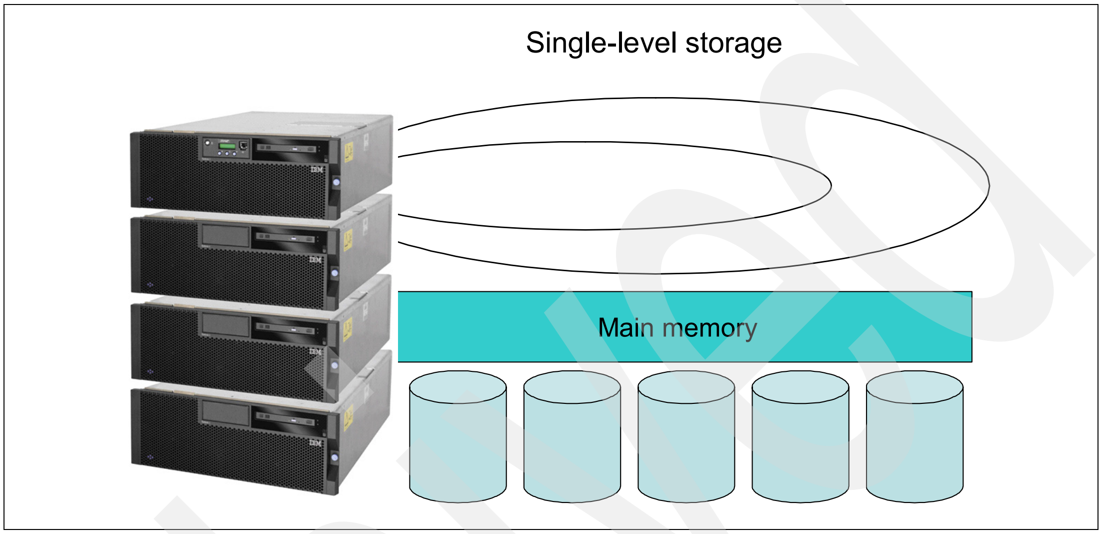


Figure 5-1 Single-level storage

When the application performs an I/O operation, the portion of the program that contains read or write instructions is first brought into main memory where the instructions are then executing.

With the read request, the virtual addresses of the needed record are resolved and for each needed page, storage management first looks to see whether it is in the main memory. If the page is there it is used for resolving the read request. But if the corresponding page is not in the main memory, it must be retrieved from disk (page fault). When a page is retrieved, it replaces a page that was recently not used. The replaced page is swapped to disk.

Similarly, writing a new record or updating an existing record is done in main memory, and the affected pages are marked as changed. A changed page remains in main memory until it is swapped to disk as a result of a page fault. Pages are also written to disk when a file is closed or when write to disk is forced by a user through commands and parameters. Also, database journals are written to the disk.

When a page must be retrieved from disk or a page is written to disk, IBM i storage management translates the virtual address to the real address of a disk location and builds an I/O request to disk. The amount of data that is transferred to disk at one I/O request is called *blocksize* or *transfer size*.

Therefore, there are two considerations that you should keep in mind when sizing for Midrange Storage with IBM i:

- The amount of I/O operations and transfer size does not depend only on IBM i workload, but also on other resources such as main memory and expert cache. Sizing of all System i hardware components and Midrange Storage should be coordinated to provide good performance.
- Due to single-level architecture and storage management, consider all System i disk space as a big entity. IBM i storage management determines which addresses on disk space will be used for IBM i objects such as libraries, database files, programs, and so on. A user does not have this information. It balances the disk I/O workload by striping the data across the disk units in an auxiliary storage pool. It is important to provide enough disk arms and adapters for all System i disk storage.

5.1.1 Flow of input/output operations

IBM i client sends a request to read data from disk. The request (SCSI command) proceedes through the virtual SCSI adapter in IBM i client, the virtual SCSI adapter in VIOS, and by AIX code to the device driver of a physical FC adapter in VIOS. The FC adapter then communicates a request to external storage. The requested block of data is transferred from the Storage System to th eFC adapter in VIOS. VIOS and Hypervisor then communicate to transfer the data to main memory in IBM i client through Power6 hardware, such as multi adapter bridge, RIO-G, and memory buffer. Acknowledgement about successful transfer goes the same way as the request, but in the opposite direction—through the adapter in VIOS, the device driver in VIOS, it is proceeded by AIX code to the virtual SCSI adapter in VIOS, the virtual SCSI adapter in IBM i, and further to IBM i storage management.

5.1.2 Influence of disk performance

It is important for sizing to understand which type of workload is critical for the customer and how disk performance influences it. Therefore, we briefly describe the critical response times:

- Application response time: This is the response time of an application transaction. This is usually important for the customer.
- Duration of batch job: Batch jobs usually run during the night. The duration of a batch job is critical for the customer because it must be finished before the regular daily transactions start.
- Disk response time: This is the time needed for an I/O operation to complete. Disk response time consists of service time and wait time:
  - Service time is the time between when the request for data is sent to the adapter and the time when data are transferred to the system.
  - Wait time is the time that the I/O operation spends in the deferred queue. For more information about the deferred queue refer to 4.5.4, “Logical drives” on page 98.

  Note that with IOP-based adapters service and wait times are defined differently than with IOP-less and virtual SCSI adapters.

The application response time and the duration of the batch job depend on many resources, not just disk response time, so it is very difficult to give a general rule as to how disk response time influences any of them.

However, performance measurements done in the IBM lab in Tucson show how disk response time relates to throughput (that is, the number of transactions per second in the internal development laboratory’s test workload). Dependence of throughput on disk response time is roughly linear.

5.2 Sizing rules for Midrange Storage

In this section we provide sizing guidelines for external Midrange Storage to achieve performance that a customer expects. These complement planning for performance, described in 4.5, “Planning considerations for performance” on page 97.

5.2.1 Number of disk drives

It is extremely important to provide IBM i workload with enough disk arms for the following reasons. When a page or a block of data is written to disk space, storage management spreads it over multiple disks to so that multiple disk arms work in parallel for any request to this piece of data. IBM i storage management *thinks* in terms of physical disk drives that are installed in Power6 system hardware, so it takes one disk drive as one disk arm. However, when Midrange Storage is used with IBM i, what storage management sees as a disk drive is actually a logical drive, or LUN. Depending on the RAID level in the Storage System, a logical drive uses multiple physical disk arms, but the same arms are typically used also by other LUNs. Proper sizing of physical disk drives is needed to ensure that enough disk arms are provided to a set of LUNs.

We recommend keeping disk drive utilization (% of time disk drive is busy) below a certain limit to provide stable response time even if the workload varies or grows a bit. With enough physical disk arms we can be confident that utilization will not exceed the recommended value.

Use the following rules of thumb for the disk drives in Midrange Storage: For critical production workload implement 15 K RPM Fibre Channel (FC) disk drives of size 73.4 GB or 146.8 GB. 300 GB FC drives may be used for less important workload with low access density. SATA drives should be considered for archiving only.

To determine the number of needed disk drives for a System i workload follow the guidelines in Table 5-1. The rules require knowing System i I/O per second and percentage of read/sec.

Table 5-1 Host I/O per disk drive

	Host I/O per second per disk drive	
	70% read	**50% read**
15 K RPM disk drive		
RAID-1 or RAID-10	110	98
RAID-5	77	60
RAID-6	59	43
10 K RPM disk drive		
RAID-1 or RAID-10	74	66
RAID-5	51	40
RAID-6	40	29
SATA disk drive		

	Host I/O per second per disk drive	
RAID-1 or RAID-10	49	44
RAID-5	34	27
RAID-6	26	19

Note that the number of disk drives calculated from the Table 5-1 on page 110 includes parity capacity (parity disk drives) in RAID-5 and RAID-6 or mirrored drives in RAID-1 and RAID-10. For example, a customer is looking for DS4000 to connect to System i partition. Currently, his workload experiences an average of 1764 IO/sec with about 70% reads/sec during the batch job, which is customer's main concern. The client is considering RAID-1 or RAID-10 protection.

To determine the number of needed 15 K RPM disk drives search Table 5-1 on page 110 under 70% read and 15 K RPM Disk drive, RAID-1 or RAID-10. The relevant number is 110 IO/sec per disk drive. Divide the customer's IO/sec by this number to get the needed disk drives:

```
1764 / 110 = 16
```

So, to accommodate the client's batch job you must configure DS4000 with at least 16 disk drives, excluding hot-spares. Note that the number 16 includes mirrored drives, so eight disk drives are available for the customer's data.

The numbers in Table 5-1 on page 110 are based on measurements take in the IBM lab in Tucson, with the assumption of 60% disk utilization, 20% read cache hit, and 30% write cache efficiency, and calculation of disk operations at each parity protection.

Measurements in the IBM lab show the following values:

- One 15 K RPM disk drive can handle a maximum of 180 disk operations/sec, or 108 disk operations/sec at 60% utilization.
- One 10 K RPM disk drive can handle a maximum of 120 disk operations/sec, or 72 disk operations/sec at 60% utilization.
- One SATA disk drive can handle a maximum of 80 disk operations/sec, or 48 disk operations/sec at 60% utilization.

To calculate the amount of disk operation that a host workload experiences at a specific RAID level, we use the following formula:

```
(Reads/sec - reads/sec that hit the cache) + raid penalty factor * (writes/sec -
writes/sec that are cache efficient)
```

The number of host IO/sec for one drive is calculated with the equation:

```
disk operations/sec from host = max disk operations for a drive at 60% utilization
```

5.2.2 Size and allocation of LUNs

For more information about the recommended size of LUNs refer to 4.5.4, "Logical drives" on page 98. We recommend distributing logical drives for a System i partition evenly between the two controllers in the Storage System.

5.2.3 Number of Fibre Channel adapters and Virtual SCSI adapters

For more information about the recommended number of physical adapters in VIOS and Virtual SCSi adapters refer to 4.5.5, “Adapters in VIOS and IBM i client” on page 100.

5.2.4 Cache size

Performance of Storage Systems highly depends on the cache hits. The higher the percentage of read cache hits and write cache efficiency the better that response times. On the other hand, the percentage of cache hits depends on the size of the cache, the cache algorithm, and the workload.

Generally, System i workload on the Storage System experiences about the same or a higher percentage of read cache hits than with internal disk, but sometimes significantly lower write cache efficiency than on the cache of internal adapters. It seems that System i workload profits from big cache on external storage. Therefore, we recommend configuring Midrange Storage with big cache sizes to use with System i workload. As a rough guideline you may follow the specifications below, which are based on the best practice with System i and System Storage:

- Plan 4 GB cache for System i workload that experiences no more than 1,000 IO/sec (light workload).
- Plan 8 GB cache for workload with no more than 2,000 IO/sec (intermediate workload).
- If the workload exceeds 2,000 IO/sec plan for 16 GB cache (heavy workload).

Also, note that the choice of Midrange Storage model described in Chapter 4, “Planning for Midrange Storage” on page 89, depends mostly on the cache size. For customers with light workload who do not plan to grow their applications in the near future, consider DS4700 with 4 GB cache. The data centers for which 4 GB cache are sufficient at the moment, but that plan to grow, size for DS4800 or DS5000 with 4 GB cache and the option to upgrade to bigger cache in future. Enterprises with intermediate workload should consider DS4800 or DS5000 with 8 GB cache, while heavy workload requires DS4800 or DS5000 with 16 GB cache. DS4700 with 2 GB cache might be considered for very non-demanding applications.

5.2.5 Links for Enhanced Remote Mirroring

If you plan Enhanced Remote Mirroring at a distance that can be covered with Fibre Channel links consider two links—one from each controller in primary (local) Storage System to the corresponding controller in secondary (remote) storage. Two links are recommended for redundancy and to provide enough bandwidth for ERM replication.

The usual connection for ERM on long distances is Fibre Channel over IP. If you plan for such connection you should carefully size the IP links for replicating System i workload. Use the following guidelines to determine the needed bandwidth:

1. Collect System i performance data. It is best to collect it over one week and, if needed, during heavy workload such as when running end-of-month jobs.

   We recommend using SQL query for performance collection database to obtain precise write MBps. For more information refer to 5.3.3, “Using SQL queries on collected performance data” on page 120.

2. Multiply the reported writes/sec by transfer size to get the write rate (MBps) for the entire period of collecting performance data.

3. Look for the highest reported write rate. Calculate the needed bandwidth as follows:
   a. Assume 10 bits per byte for network overhead.
   b. If the compression of devices for remote links is known you may apply it.
   c. Assume a maximum of 80% utilization of the network.
   d. Apply a 10% uplift factor to the result to account for peaks in the 5-minute intervals of collecting data, and a 20–25% uplift factor for 15-minute intervals

As an example we show how to calculate the required bandwidth for a given write workload:

1. The highest reported write rate at an System i data center is 40 MBps.
2. Assume 10 bits per byte for network overhead: 40 MB/sec * 1.25 = 50 MBps.
3. Assume a maximum of 80% utilization of the network: 50 MB/sec * 1.25 = 62.5 MBps.
4. Apply a 10% uplift for 5-minute intervals: 62.5 MB/sec * 1.1 = app 69 MBps.
5. The needed bamdwidth is 69 MBps.

> **Note:** Proper sizing of the secondary Storage System is equally important to the sizing of the primary Storage System.
>
> When estimating disk response time with ERM take into account latency of devices (SAN directors, converters) for remote links.

For information about sizing links for ERM refer also to the IBM Redbooks publication *DS4000 Best Practices and Performance Tuning Guide*, SG24-6363.

5.3 Gathering information for sizing

To properly use sizing guidelines and Disk Magic modeling for Midrange Storage with System i partition, obtain the IBM i workload characteristics in peak period—the period that is the most critical for the customer, and in which the heaviest I/O load is placed upon the external disk.

Later in this section we explain how to decide on the peak period, which information from System i workload is needed, and how to obtain it.

5.3.1 Sizing for the peak period

It is very important to determine the correct peak period of System i workload. Both criteria—how the period is critical to the customer, and how large the I/O rate is experienced in this interval—should be taken into account.

We recommend the following approach:

1. Together with the customer, define the time frame that is most important to them. In some enterprises successful business depends on fast response time in application transaction workload, typically during the day. Other data centers are mostly concerned about duration of their nightly batch job. In many cases, however, both transaction workload and batch job are critical to the client, so it is a good idea to consider both of them.
2. Once you have defined the critical time frames obtain workload characteristics for the entire time frames. Then look for the period with the highest I/O rate and the period with the highest data rate (MBps). The peak period should not be longer than 15 minutes.
3. Apply sizing rules for each peak period, and plan Midrange Storage configuration so that all peak periods are accommodated.

4. Run the Disk Magic model for all determined peak periods.

   Note that after you insert System i performance reports to Disk Magic, the tool presents you with the possibility to select the interval with the highest I/O rate, highest MB/sec, highest write ratio, and so on.

5.3.2 IBM i performance data

To obtain workload characteristics you must collect System i performance data with the licensed product IBM Performance Tools for i5/OS (LPP 5761-PT1) during critical time frames or during the entire day. You may keep a collection interval 15 minutes, but consider changing it to 5 minutes for critical workloads or for sizing Enhanced Remote Mirroring links. Configuring and starting of performance collection is shown in Figure 5-2 and Figure 5-3 on page 115.

```
                        Configure Perf Collection (CFGPFRCOL)

 Type choices, press Enter.

 Default interval . . . . . . . .   15.00         *SAME, .25, .50, 1.0, 5.0...
 Collection library . . . . . . .   QPFRDATA      Name, *SAME
 Default collection profile . . .   *MINIMUM      *SAME, *MINIMUM, *STANDARD...
 Cycle time . . . . . . . . . . .   000000        Time, *SAME
 Cycle interval . . . . . . . . .   24            *SAME, 1-24 hours
 Collection retention period:
   Number of units  . . . . . . .   00720         *SAME, 1-720, *PERM
   Unit of time . . . . . . . . .   *HOURS        *HOURS, *DAYS
 Create database files  . . . . .   *YES          *SAME, *YES, *NO








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

Figure 5-2 Configure performance collection on IBM i

```
                          Collect Performance Data                  I6PFE
                                                       08/08/08  00:08:31
Collection Services status:
  Status  . . . . . . . . . . . . . :   Stopped











Select one of the following:

     1. Start Performance Collection
     2. Configure Performance Collection
     3. End Performance Collection

Selection or command
===> 1

F3=Exit   F4=Prompt   F5=Refresh   F9=Retrieve   F12=Cancel
```

Figure 5-3 Start performance collection on IBM i

To better understand the workload pattern and discover the peaks we recommend gathering performance data in a few consecutive days and during heavy end-of-month jobs.

With IBM Performance Tools for i5/OS make the following reports from the collected data:

- System report, sections Storage Pool Utilization and Disk Utilization
- Resource report, section Disk Utilization
- Component report, section Disk Activity

When creating reports select the intervals in the critical time frame.

Selecting the system report and the needed sections is shown in Figure 5-4 and Figure 5-5 on page 117.

```
                      Print Performance Report - Sample data

Library . . . . . .   QPFRDATA

Type option, press Enter.
  1=System report   2=Component report   3=Job report   4=Pool report
  5=Resource report


Option   Member       Text                                  Date       Time
  1      Q031154212                                         01/31/08   15:42:12
         Q030000002                                         01/30/08   00:00:02
         Q029000002                                         01/29/08   00:00:02
         Q028122842                                         01/28/08   12:28:42
         Q028000002                                         01/28/08   00:00:02
         Q027000002                                         01/27/08   00:00:02
         Q026000002                                         01/26/08   00:00:02
         Q025000002                                         01/25/08   00:00:02
         Q024000002                                         01/24/08   00:00:02
         Q023000002                                         01/23/08   00:00:02
                                                                        More...
F3=Exit   F5=Refresh   F11=Work with your spooled output files   F12=Cancel
F15=Sort by member     F16=Sort by text
```

Figure 5-4 Selecting system report

```
                            Select Sections for Report

Member . . . . . . . . . . . :   Q031154212

Type options, press Enter.  Press F6 to print entire report.
  1=Select

Option     Section
           Workload
           Resource Utilization
           Resource Utilization Expansion
  1        Storage Pool Utilization
  1        Disk Utilization
           Communication Summary
           TCP/IP Summary
           HTTP Server Summary
           Logical Partitions Summary




                                                                        Bottom
F3=Exit   F6=Print entire report   F12=Cancel
```

Figure 5-5 Selecting sections

Selecting the intervals for the report is shown in Figure 5-6 and Figure 5-7 on page 119.

```
                         Select Categories for Report

Member . . . . . . . . . . . :   Q031154212

Type options, press Enter.  Press F6 to print entire report.
  1=Select

Option     Category
  1        Time interval
           Job
           User ID
           Subsystem
           Pool
           Communications line
           Control unit
           Functional area




                                                                      Bottom
F3=Exit   F6=Print entire report   F12=Cancel
```

Figure 5-6 Specifying selection by intervals

```
                            Select Time Intervals

 Library . . . . . :   QPFRDATA      Performance data . . . . . :   Q031154212

 Type options, press Enter.
   1=Select
 O                                          Int    High        Pool
 p              Transaction  -CPU Util--  Feat  --Util--  -Fault/Sec-
 t  Date   Time   Count  Resp  Tot Int Bch  Util  Dsk Unit  Mch User ID    Excp
    01/31  18:30      0   .00    0   0   0     0    1 0007    0    0 02     433
    01/31  18:45     10   .03    0   0   0     0    1 0007    0    0 03     450
 1  01/31  19:00      0   .00    0   0   0     0    1 0007    0    0 02     434
 1  01/31  19:15      0   .00    0   0   0     0    1 0007    0    0 02     450
 1  01/31  19:30      0   .00    0   0   0     0    1 0007    0    0 02     432
 1  01/31  19:45      0   .00    0   0   0     0    1 0007    0    0 02     433
 1  01/31  20:00      0   .00    0   0   0     0    1 0007    0    0 02     453
 1  01/31  20:15     19   .01    0   0   0     0    1 0007    0    0 03     432
 1  01/31  20:30      0   .00    0   0   0     0    1 0007    0    0 02     451
    01/31  20:45      0   .00    0   0   0     0    1 0007    0    0 02     432
                                                                        More...

  F3=Exit                    F5=Refresh                  F12=Cancel
  F13=Sort (date/time)       F14=Sort (count)            F24=More keys
```

Figure 5-7 Selecting the intervals in critical time frame

Reports are created as IBM i spoolfiles. Transfer them to the PC as text files in order to use them for Disk Magic modelling. To transfer them, open IBM Systems Director Navigator for i5/OS and click **Basic operations** → **Printer Output**. On the Printer Output tab select the spoolfile that you want to transfer to your PC, select **Export** from the Actions menu, and click **Go**, as shown in Figure 5-8.

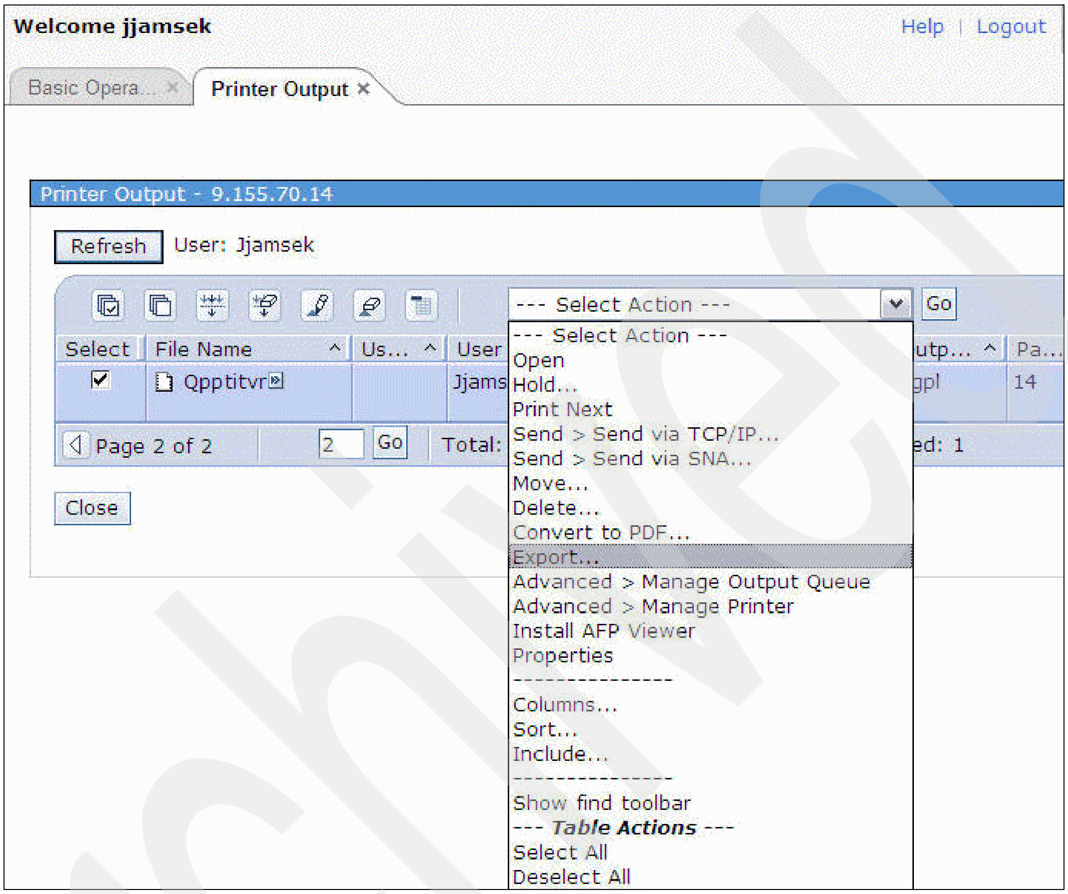


Figure 5-8 Export spoolfile in Director Navigator

On the File Download window that is shown next, select **Save**, select the directory to save to, and click **Save**.

Note that you can also use System i Navigator to transfer IBM i spoolfiles to text files in your PC.

5.3.3 Using SQL queries on collected performance data

Performance reports show only some of the values that are contained in database files collected by IBM Performance Tools for i5/OS. In some cases you may want to obtain also the performance data not shown in the reports. For this, use IBM i SQL queries to collected data. A typical case in which we recommend using SQL Query is sizing for IP links for Enhanced Remote Mirroring on long distance. With SQL query for performance collection data you get more precise information about the IBM i write rate than you can get from the reports.

The following is an example of SQL query to collect read MB/sec and write MB/sec:

```
- select intnum, DSASP,
       sum(DSBLKR) * 520 / 1024 / max(intsec) as "Read KB/sec",
       sum(DSBLKW) * 520 / 1024 / max(intsec) as "Write KB/sec"
from qapmdisk
group by intnum, dsasp
order by intnum, dsasp
```

Note that before using SQL query you should make sure that the query is performed on the correct member of the collection services database. This can be achieved by command override with the database file, as follows:

```
OVRDBF  QAPMDISK  <library>/QAPMDISK  <member>
```

5.4 Sizing example with Disk Magic

Disk Magic is a tool for sizing and modelling disk systems for various servers. It can be used for modelling IBM and other disk systems attached to System i, System z®, System p, and other systems. It is developed by the company IntelliMagic™.

IBM employees can access Disk Magic by visiting the following internal Web page (this page is not available to IBM Business Partners or Customers):

http://w3-03.ibm.com/sales/support/information/stg/

Business Partners can sign on to the IBM Partner World Web page and do a search for Disk Magic:

http://wow-1.ibm.com/partnerworld/pwhome.nsf/weblook/index_emea_en.html

Usually, we run Disk Magic modelling to predict the response time of IBM i workload with connecting the external Storage System. However, Disk Magic can be also used for more complex modelling, such as predicting how disk response time will change if the customer shares storage arrays among many IBM i partitions, modelling how disk performance are influenced by workload growth, how upgrading cache or adding more disk drives enhance performance, and so on.

In this section we show an example of modelling DS4800 with a System i workload.

To install Disk Magic follow the instructions on the Web page referred to in this section.

To provide to Disk Magic information about System i workload, you may import System i Performance Tools reports in Disk Magic. Alternatively, you may manually type in the necessary values. We recommend importing reports whenever possible. Consider inserting information manualy only in cases when reports cannot be read by Disk Magic, for example, when they are in a language other than English.

Before starting Disk Magic prepare the needed System i performance reports. For more information about how to obtain them refer to 5.3.2, “IBM i performance data” on page 114.

Insert System i performance reports

To launch Disk Magic click its icon on your desktop. In the first Disk Magic window, Welcome to Disk Magic, select **Open and iSeries Automated Input (*IOSTAT, *TXT, *CSV)**, as shown in Figure 5-9, and click **OK.**

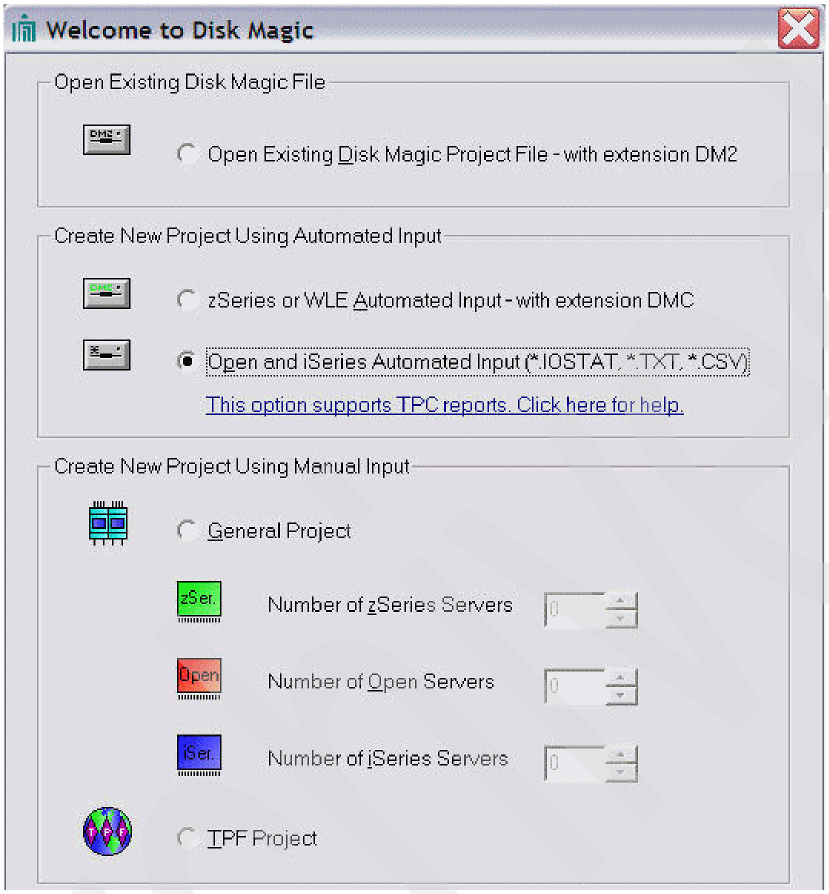


Figure 5-9 Automated input to Disk Magic

On the next window select the System i reports that you want to insert in Disk Magic and click **Open**, see Figure 5-10.

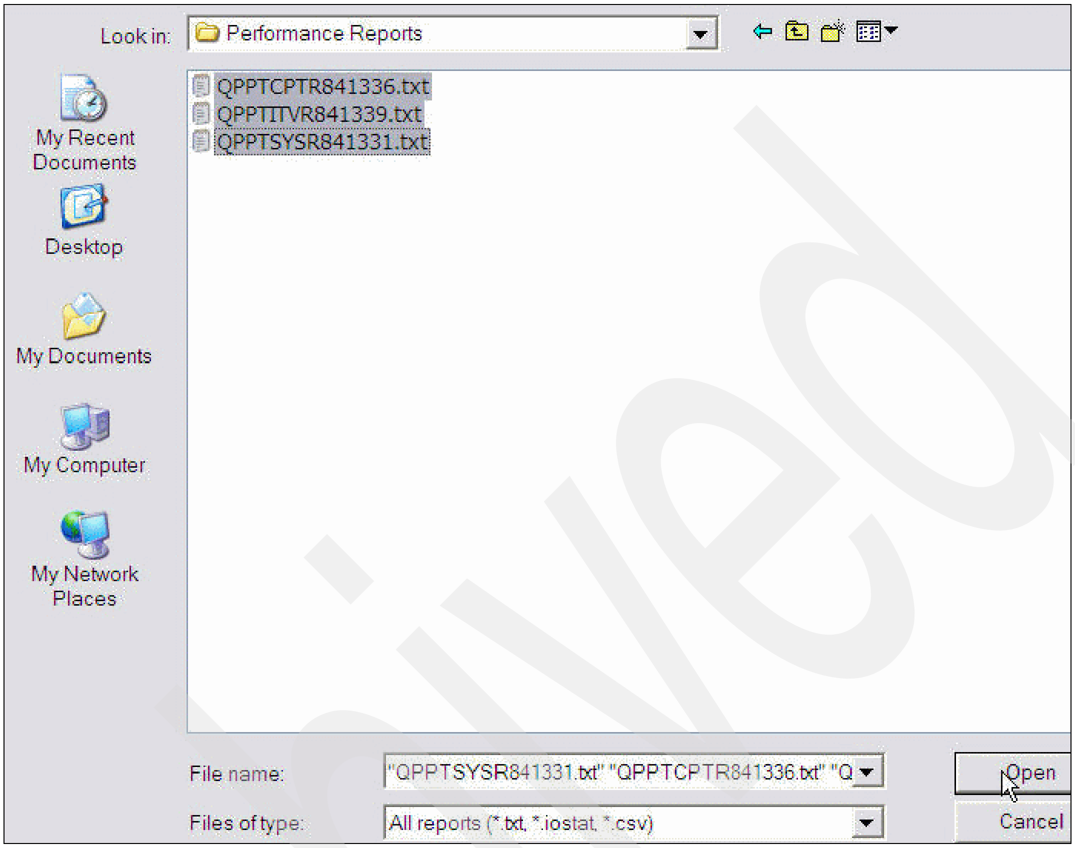


Figure 5-10 Select reports to insert

You may want to adjust the way taht data will be processed in Disk Magic, like to separate Auxiliary Storage Pools (ASPs) or use IBM i mirroring. You may also need to correct the date and time of the collected data. To do so click **Edit Properties** on the Multiple File Open - File Overview window. Then adjust the processing by selecting the appropriate System i properties, as can be seen in Figure 5-11, and click **OK**. In our example we un-check the box **Discern ASP level** since we want Disk Magic to model overall response times for System i workload, not separately for each ASP.

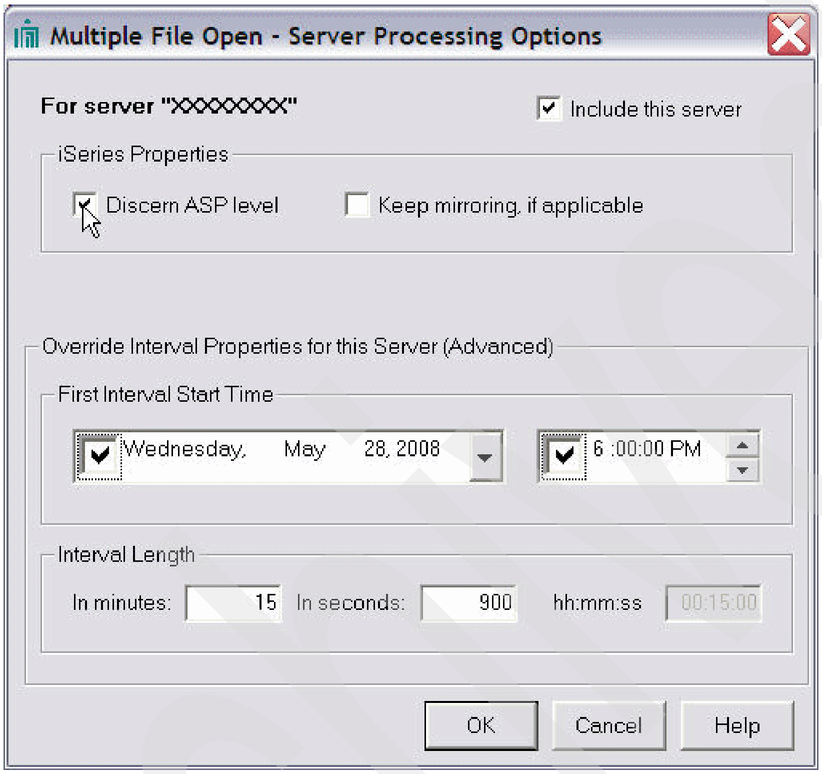
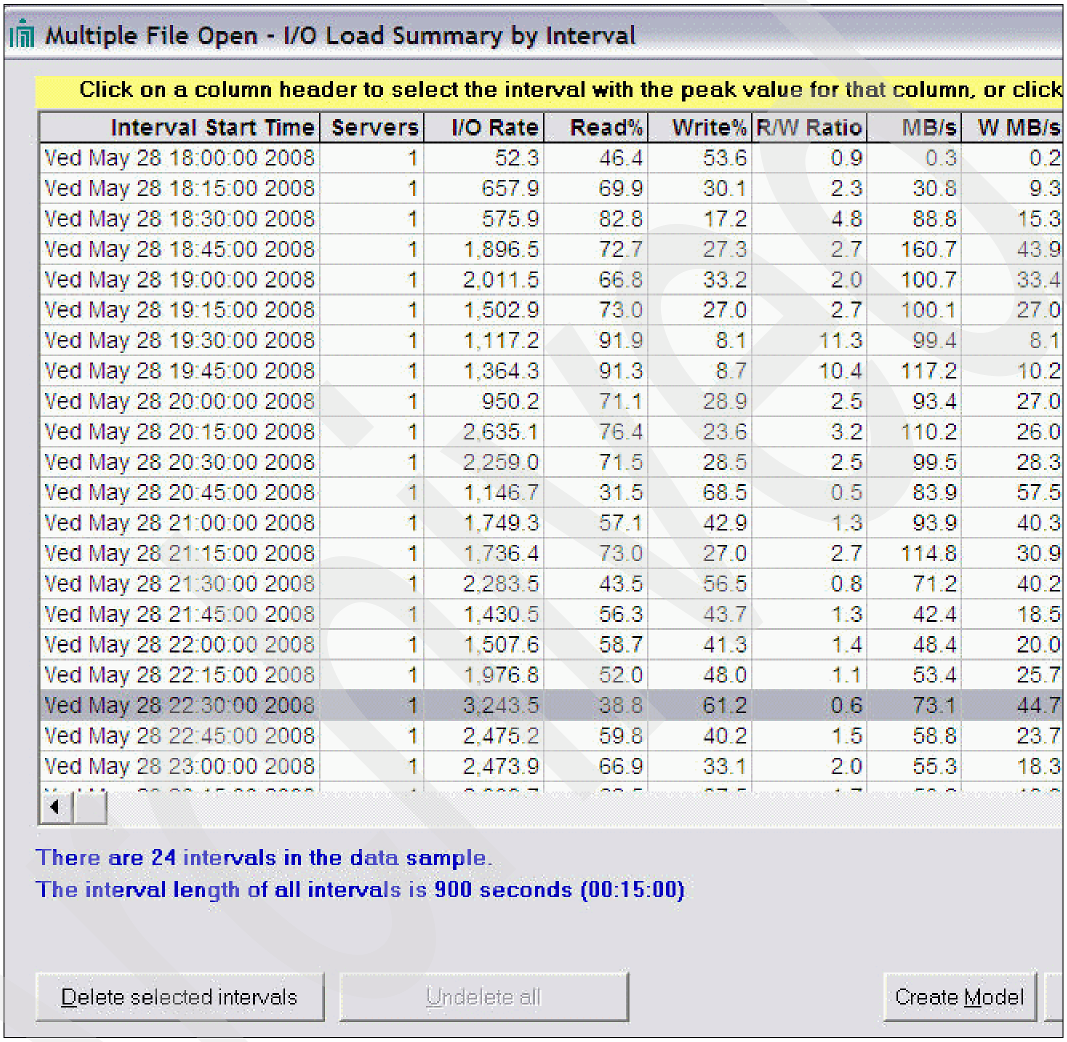


Figure 5-11 Adjust System i properties

Next click **Process all Files,** which brings up the Multiple File Open - I/O Load Summary by Interval window. Select the peak interval and click **Create Model**, as shown in Figure 5-12. For more information about how to decide for the peak interval refer to 5.3.1, “Sizing for the peak period” on page 113.

Note that when you click a column header Disk Magic automatically selects the peak interval for that column.



Interval Start Time	Servers	I/O Rate	Read%	Write%	R/W Ratio	MB/s	W MB/s
Ved May 28 18:00:00 2008	1	52.3	46.4	53.6	0.9	0.3	0.2
Ved May 28 18:15:00 2008	1	657.9	69.9	30.1	2.3	30.8	9.3
Ved May 28 18:30:00 2008	1	575.9	82.8	17.2	4.8	88.8	15.3
Ved May 28 18:45:00 2008	1	1,896.5	72.7	27.3	2.7	160.7	43.9
Ved May 28 19:00:00 2008	1	2,011.5	66.8	33.2	2.0	100.7	33.4
Ved May 28 19:15:00 2008	1	1,502.9	73.0	27.0	2.7	100.1	27.0
Ved May 28 19:30:00 2008	1	1,117.2	91.9	8.1	11.3	99.4	8.1
Ved May 28 19:45:00 2008	1	1,364.3	91.3	8.7	10.4	117.2	10.2
Ved May 28 20:00:00 2008	1	950.2	71.1	28.9	2.5	93.4	27.0
Ved May 28 20:15:00 2008	1	2,635.1	76.4	23.6	3.2	110.2	26.0
Ved May 28 20:30:00 2008	1	2,259.0	71.5	28.5	2.5	99.5	28.3
Ved May 28 20:45:00 2008	1	1,146.7	31.5	68.5	0.5	83.9	57.5
Ved May 28 21:00:00 2008	1	1,749.3	57.1	42.9	1.3	93.9	40.3
Ved May 28 21:15:00 2008	1	1,736.4	73.0	27.0	2.7	114.8	30.9
Ved May 28 21:30:00 2008	1	2,283.5	43.5	56.5	0.8	71.2	40.2
Ved May 28 21:45:00 2008	1	1,430.5	56.3	43.7	1.3	42.4	18.5
Ved May 28 22:00:00 2008	1	1,507.6	58.7	41.3	1.4	48.4	20.0
Ved May 28 22:15:00 2008	1	1,976.8	52.0	48.0	1.1	53.4	25.7
Ved May 28 22:30:00 2008	1	3,243.5	38.8	61.2	0.6	73.1	44.7
Ved May 28 22:45:00 2008	1	2,475.2	59.8	40.2	1.5	58.8	23.7
Ved May 28 23:00:00 2008	1	2,473.9	66.9	33.1	2.0	55.3	18.3

Figure 5-12 Select peak interval

Now Disk Magic creates the model from inserted reports. You may look to the log for any errors while processing them. The model consists of the System i server (blue icon) and the Disk Subsystem (green icon). The created model and part of the Disk Magic log are shown in Figure 5-13.

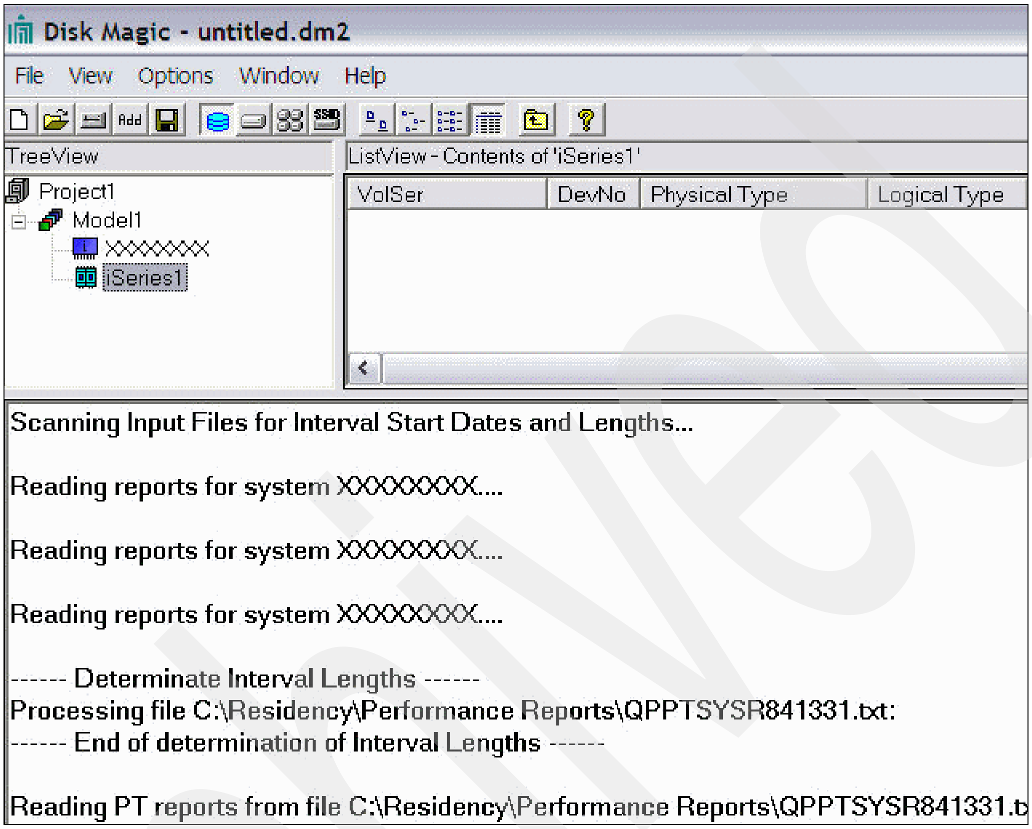


Figure 5-13 Created model

Save base

Double-click the disk subsystem in the tree view. This brings up the Disk Subsystem - iSeries1 window. A corresponding tab in the window contains values about the Storage System, disk drives being used, System i workload, and interfaces. As you can see in Figure 5-14, the current model contains internal System i disks as read from System i performance reports. Save this model as the base for further Disk Magic modelling by clicking **Base**.

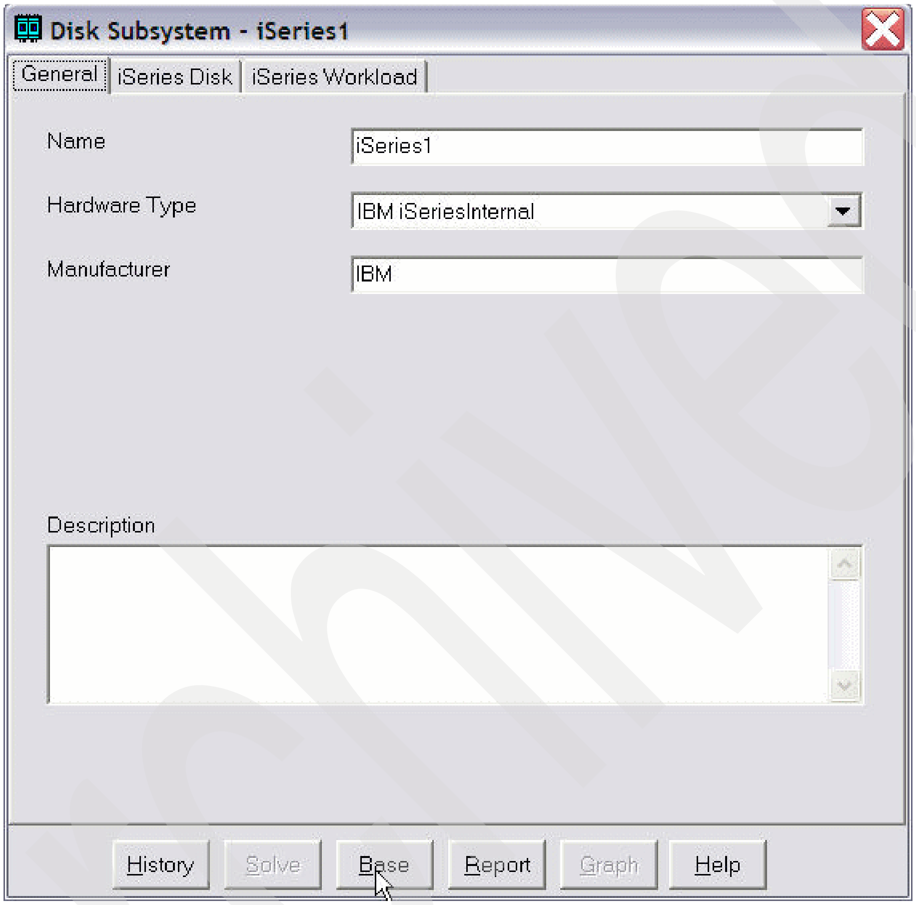


Figure 5-14 Save base

Still in the model for the internal disk, click the **Series Workload** tab, observe the disk response time (service time + wait time) and percentage of cache hits experienced on internal disk. In our example the disk response time is 1.4 ms, the percentage of read cache hit is 24.3%, and write cache efficiency is 55.2%, as shown in Figure 5-15 and Figure 5-16.

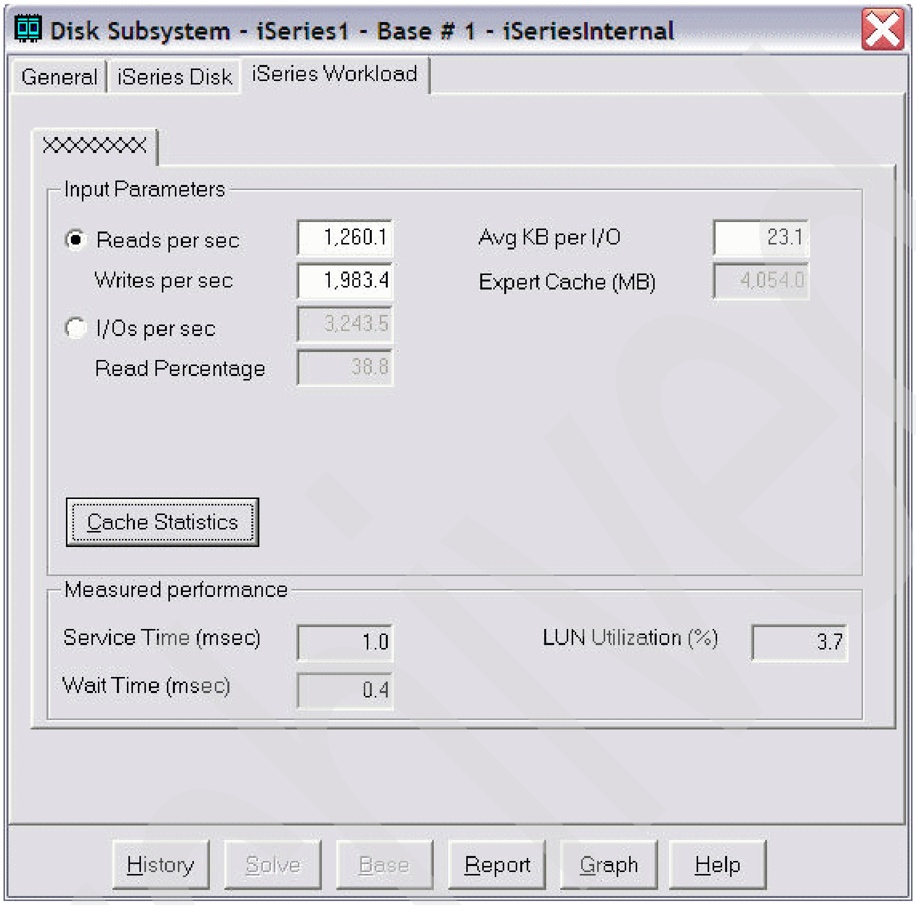


Figure 5-15 Disk response time on internal disk

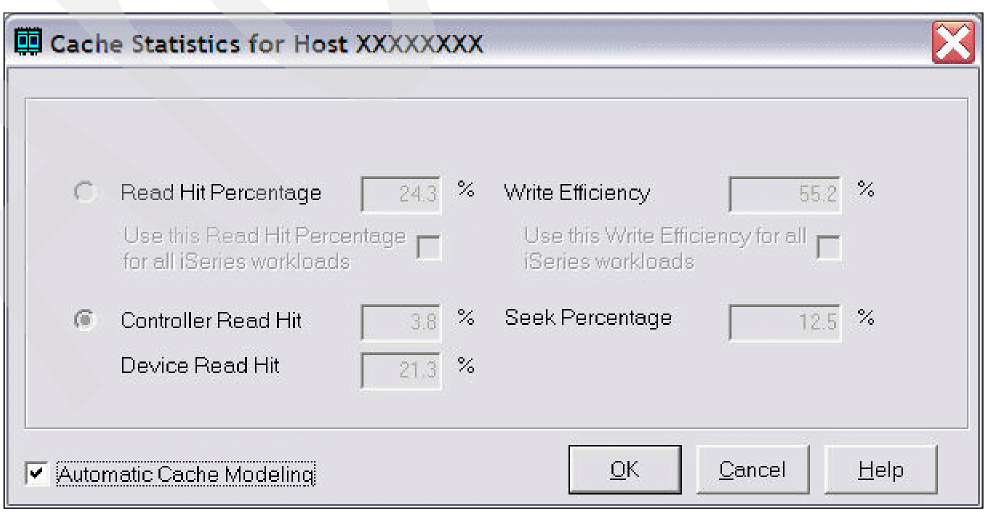


Figure 5-16 Cache statistics on internal disk

Model DS4000 for System i workload

To model DS4000 and DS5000 for the inserted workload, click the **General** tab and select the planned model of DS4000 from the Hardware Type pull-down, as shown in Figure 5-17. For our example we select **IBM DS4800**.

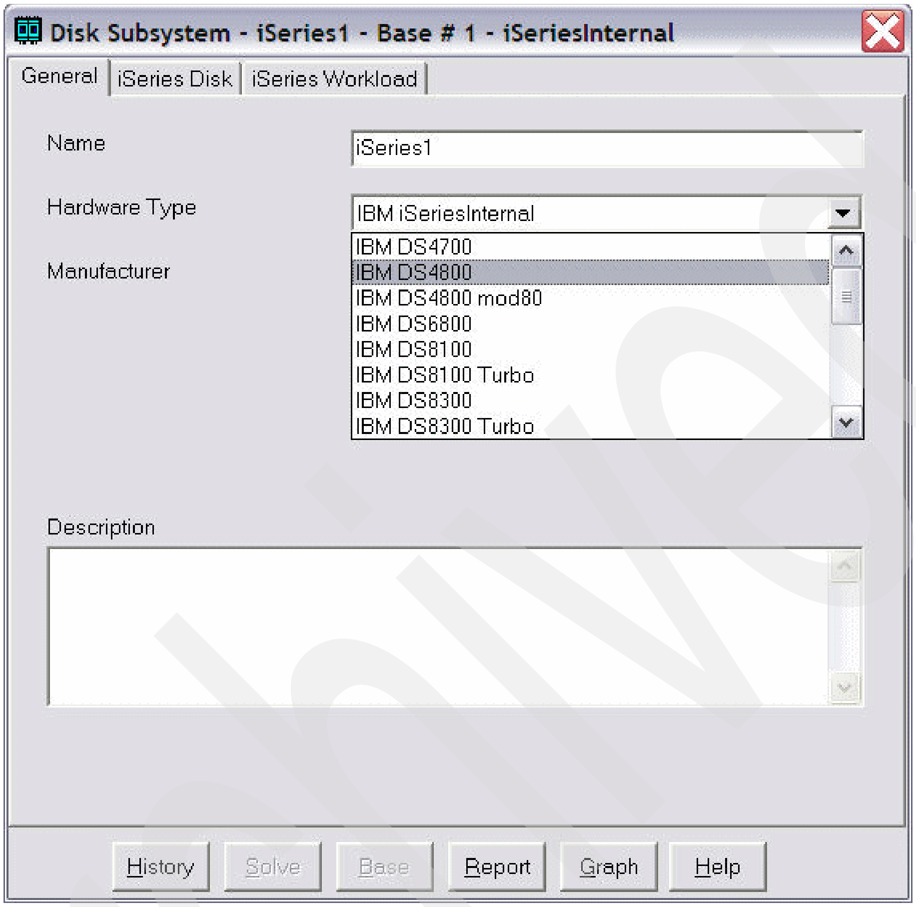


Figure 5-17 Select DS4000 model

We recommend using a large amount of cache for System i workload, so we adjust the DS4000 cache to 16 GB (Figure 5-18).

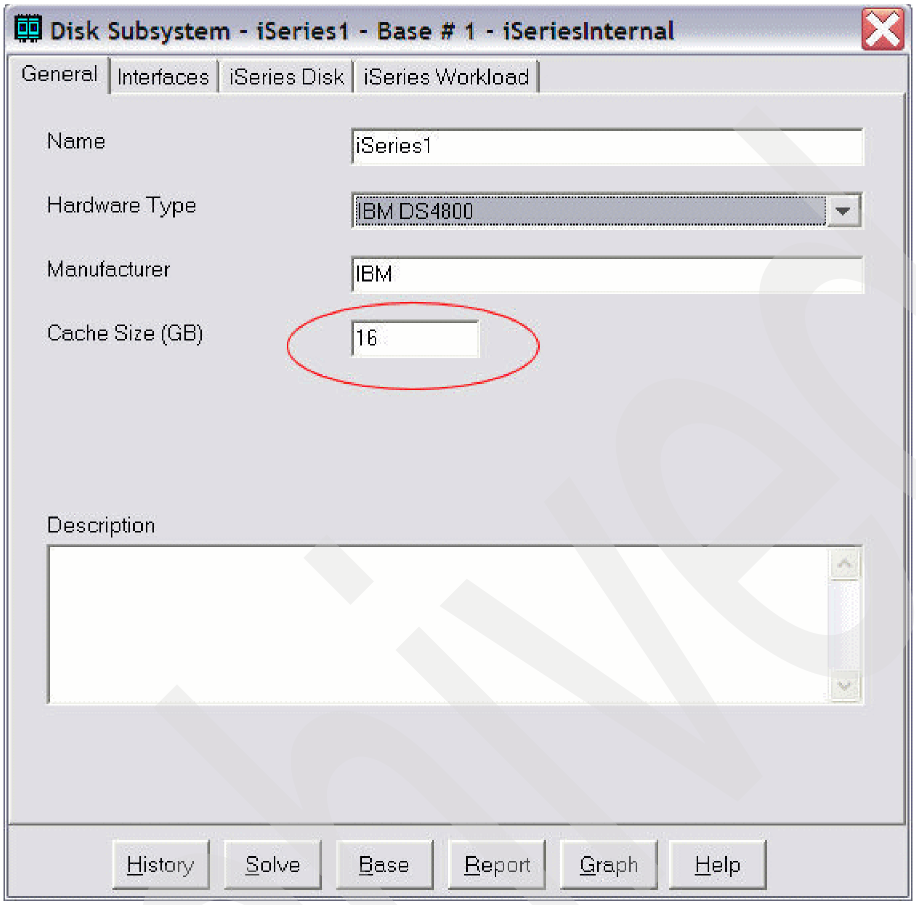


Figure 5-18 Adjusting cache size

On the Interfaces tab, adjust the number and speed of host ports in DS4000 (Interfaces from Disk Subsystem) and adapters in VIOS (Interfaces from Servers), as shown in Figure 5-19. For more information about sizing the number of adapters in VIOS refer to 5.2.3, “Number of Fibre Channel adapters and Virtual SCSI adapters” on page 112. For this example we plan two host ports in DS4800 and two adapters in VIOS.

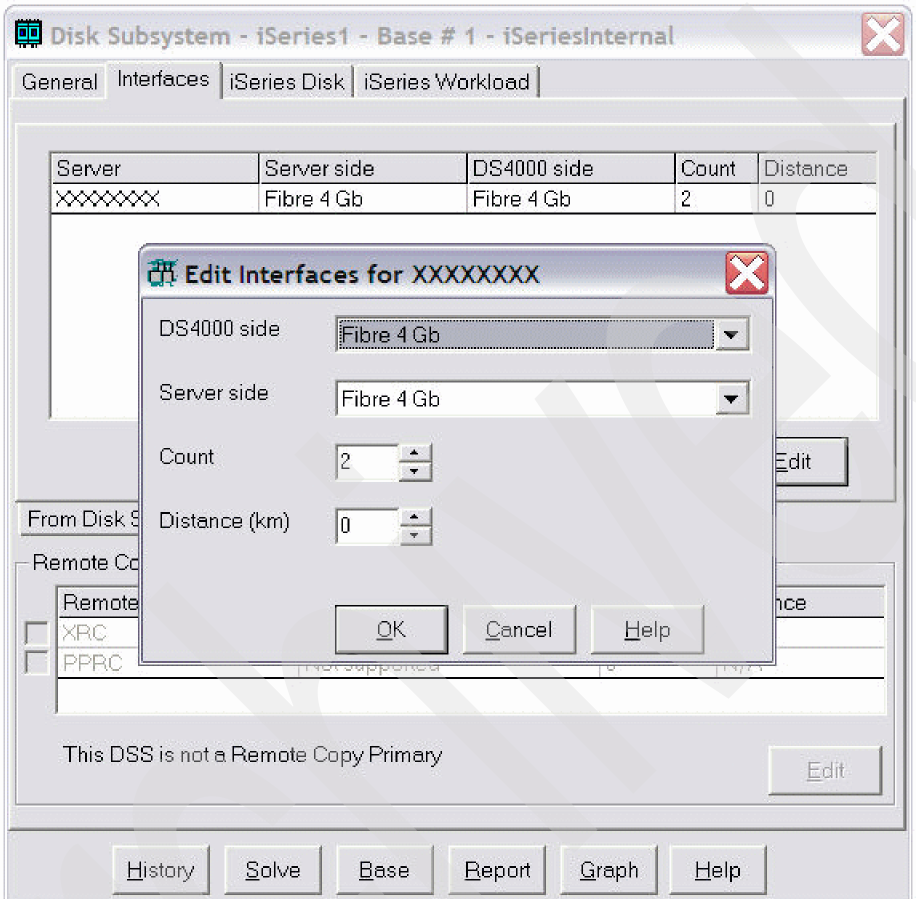


Figure 5-19 Adjusting interfaces

On the iSeries Disk tab, adjust the type of disk drives, RAID level, and capacity, based on the sizing guideline for the number of disk drives and the desired capacity.

In our example we use the calculation for needed disk drives described in 5.2.1, “Number of disk drives” on page 110, for selected peak interval of the customer’s workload. 1260 reads/sec and 1983 writes/sec require 36 disk drives in RAID-1 or RAID-10 to ensure that there is no more than 60% disk drive utilization. We decide for 146.8 GB 15 K RPM disk drives. By only using the rule for number of disk arms, 18 * 146.2 GB = 2632 GB usable capacity would be provided, but the customer’s current capacity is more than this. It is 2955 GB. To ensure that the IBM i partition has enough capacity we size 50 disk drives in RAID-10. This gives the following usable capacity for IBM i workload:

```
25 * 146.2 GB * 0.89 = 3253 GB
```

We specify IBM 146 GB 15 K RPM disk drives and RAID-10 level, and we insert capacity 3253, as shown in Figure 5-20.

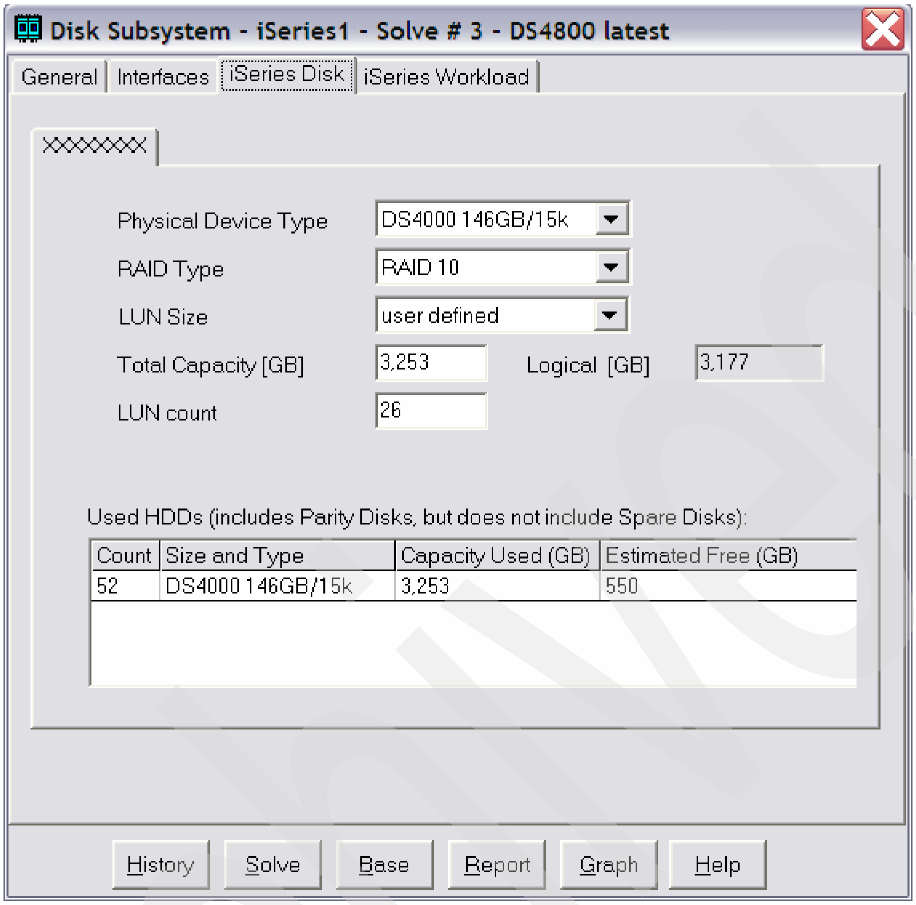


Figure 5-20 Specify disk

In our example Disk Magic calculates 52 needed disk drives because of rounding up the calculated decimal number to the next integer, for number of disk drives.

Solve the DS4000 model by clicking **Solve** on any tab. After the model is solved, observe the predicted disk response time on the iSeries Workload tab. Figure 5-21 shows predicted response time in our example, which is 1.8 ms.

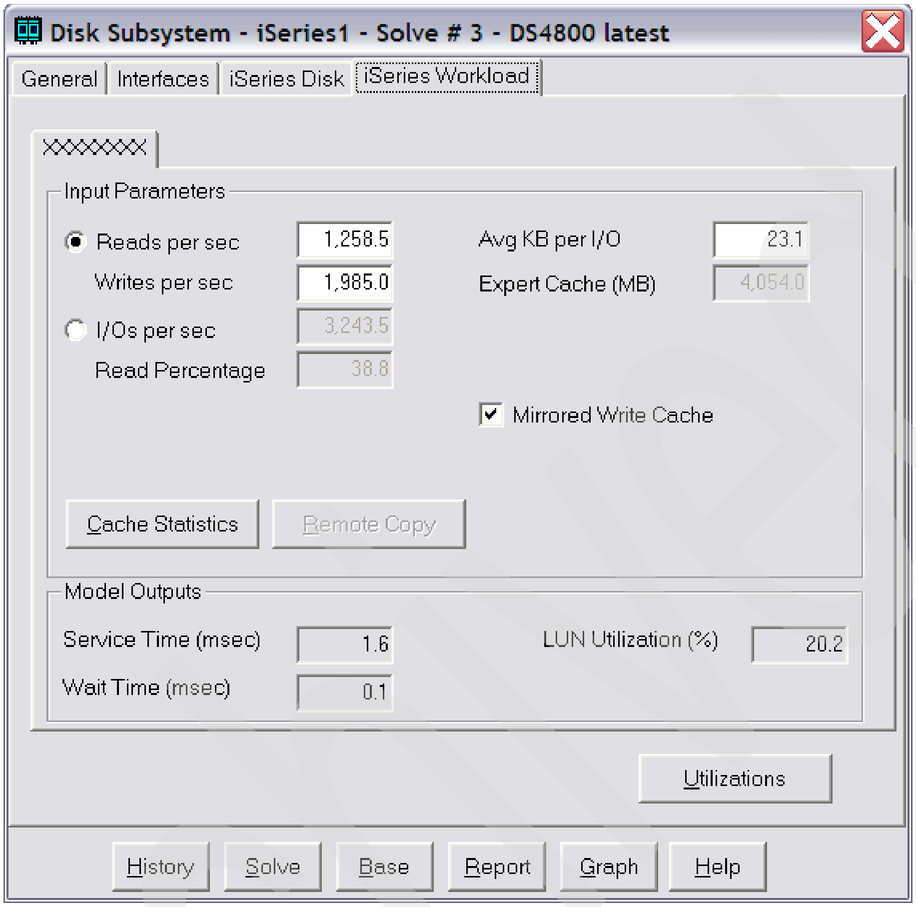


Figure 5-21 Predicted response time on DS4800

It is important to observe also other modelled values like cache hits and utilizations of resources in the Storage System. You can see them by clicking the relevant button.

As shown in Figure 5-22 and Figure 5-23, predicted read cache hits on DS4800 in our modelling are higher than on internal disk, while the percentage of write cache efficiency is lower.

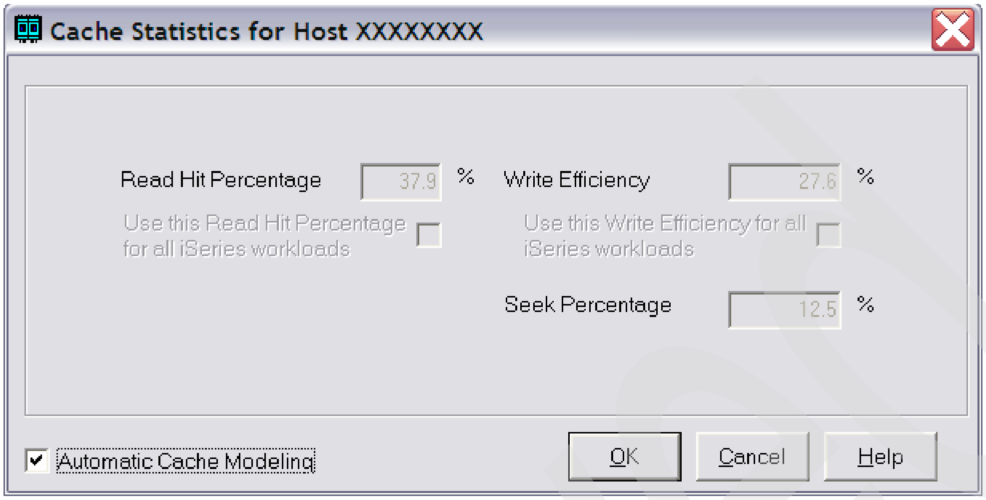


Figure 5-22 Predicted cache hits on DS4800

Utilization of disk drives, interfaces, and processors in DS4800 is within the recommended limits. This is indicated by green color in Figure 5-23.

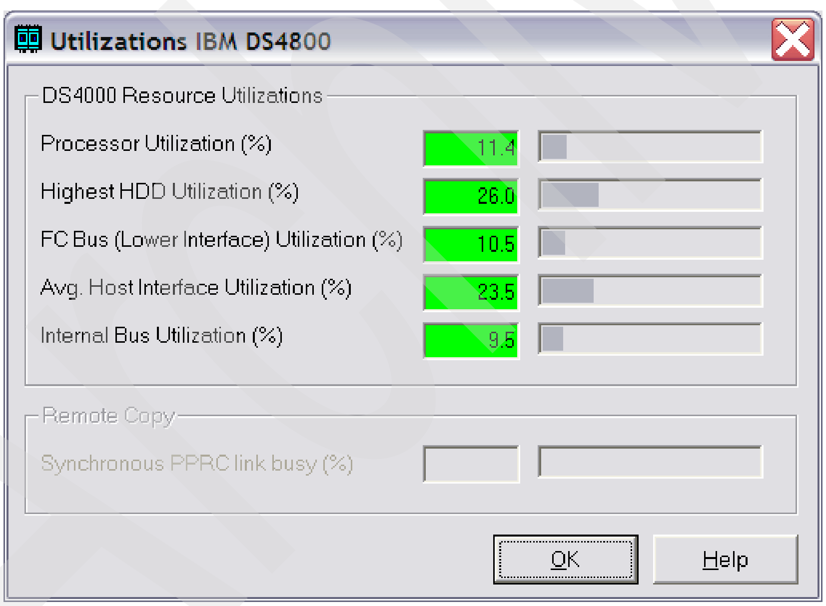


Figure 5-23 Predicted utilization on DS4800

Modelled disk response time 1.7 ms is a bit higher than the experienced response time 1.4 ms on internal disk. If the customer expects the same or a lower response time than is currently experienced, you may want to model more disk drives to reach the desired values. Note also that our example was done for absolute peak in IO/sec, while modeling average values of a six-hour period gave a better response time on DS4800 than on internal disk.

6

Implementing Midrange Storage

In this chapter we guide you through the implementation of IBM Midrange Storage for IBM i as a client of Virtual I/O Server (VIOS).

Based on an example configuration described in 6.1, “Setup example” on page 136, we show each configuration step, as follows:

- Section 6.2, “Setting up the host environment” on page 137
- Section 6.3, “Attaching Midrange Storage to VIOS” on page 190
- Section 6.4, “Configuring storage for IBM i” on page 191

  For the storage configuration we refer to 7.1, “Installation of DS Storage Manager Client” on page 214, where we show two alternatives for configuring IBM i Midrange Storage either using the DS Storage Manager GUI or the SMcli command-line interface.
- Section 6.5, “Configuring VIOS virtual devices” on page 191
- Section 6.6, “Adding DS Storage to IBM i configuration” on page 197

Finally, for those customers willing to migrate from an IBM i internal storage configuration to an IBM i external storage solution with IBM System Storage DS3400, DS4700, DS4800, or DS5000 we outline the migration procedure in 6.7, “Migration to Midrange Storage” on page 212.

6.1 Setup example

Figure 6-1 shows our setup example of the required steps for implementing IBM Midrange External Storage with IBM i as a client of VIOS.

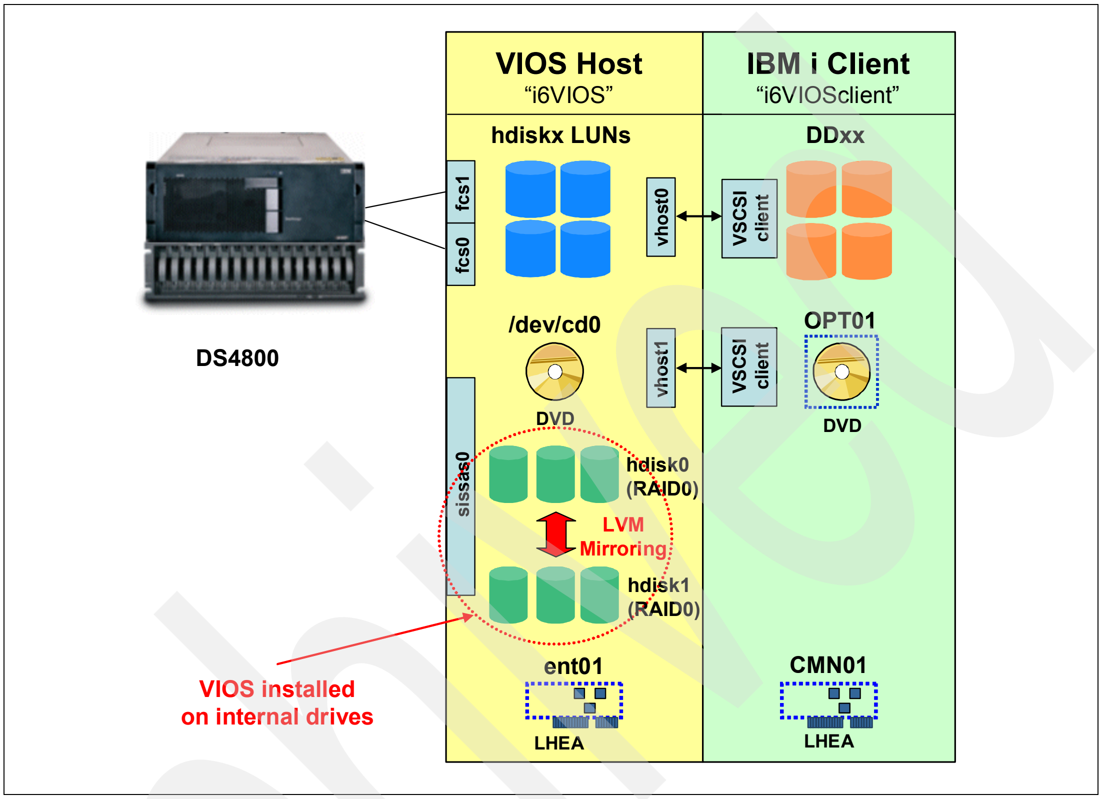


Figure 6-1 Setup example used for guidance through the implementation steps

We install VIOS 1.5.x on the IBM POWER Systems POWER6 model 9406-MMA CEC's internal SAS drives on which we create two RAID0 arrays, each consisting of three SAS disk drives using SAS controller hardware striping. Using LVM mirroring we mirror VIOS across both RAID0 arrays for protection against disk failures. The VIOS partition i6VIOS is configured with two Fibre Channel adapters, one of them connected to the IBM System Storage DS4800 controller A and the other connected to DS4800 controller B. We use the VIOS 1.5.x default RDAC multi-path device driver for our DS4800 attachment. For using MPIO instead of RDAC on VIOS for DS4000 attachment refer to "Migrating from RDAC to MPIO for DS4000 series attachment" on page 351. In total, two virtual SCSI server adapters are configured on the VIOS partition, with one of them used for mapping the virtualized SCSI storage devices (LUNs) to the IBM i client and the other one used for mapping the DVD drive owned by the VIOS partition as a virtual SCSI optical device to the IBM i client. For network connectivity both the VIOS and the IBM i client partition i6VIOS_client are configured with a logical host Ethernet adapter (LHEA) using the POWER6 system's Integrated Virtual Ethernet (IVE) feature (see 1.2.2, "Virtual Ethernet support" on page 5).

On the IBM System Storage DS4800 side we configured one RAID-10 array with 7x35 GB LUNs.

Note: On the IBM i client virtual LUNs are assigned a queue depth of 32 by IBM i SLIC storage management compared to a queue depth of 6 used for native attached SCSI LUNs, so making a large number of LUNs accessible for the IBM i client of VIOS is not as important from a performance perspective as it would be for IBM i native storage attachment (see also 4.5, “Planning considerations for performance” on page 97).

6.2 Setting up the host environment

In this section we describe the configuration steps on the IBM POWER Systems POWER6 server side for creating the VIOS and IBM i client partitions and installing the IBM Virtual I/O Server appliance.

6.2.1 Creating the VIOS LPAR

Before being able to create a VIOS partition on the IBM POWER Systems POWER6 server, have your paper with the PowerVM activation code available to be entered on the HMC by selecting the **Servers** view, marking the server to be used for creating the VIOS partition, selecting **Tasks** → **Capacity on Demand (CoD)** → **Advanced POWER Virtualization** → **Enter Activation Code**, as shown in Figure 6-2, entering your activation code, and clicking **OK** to confirm.

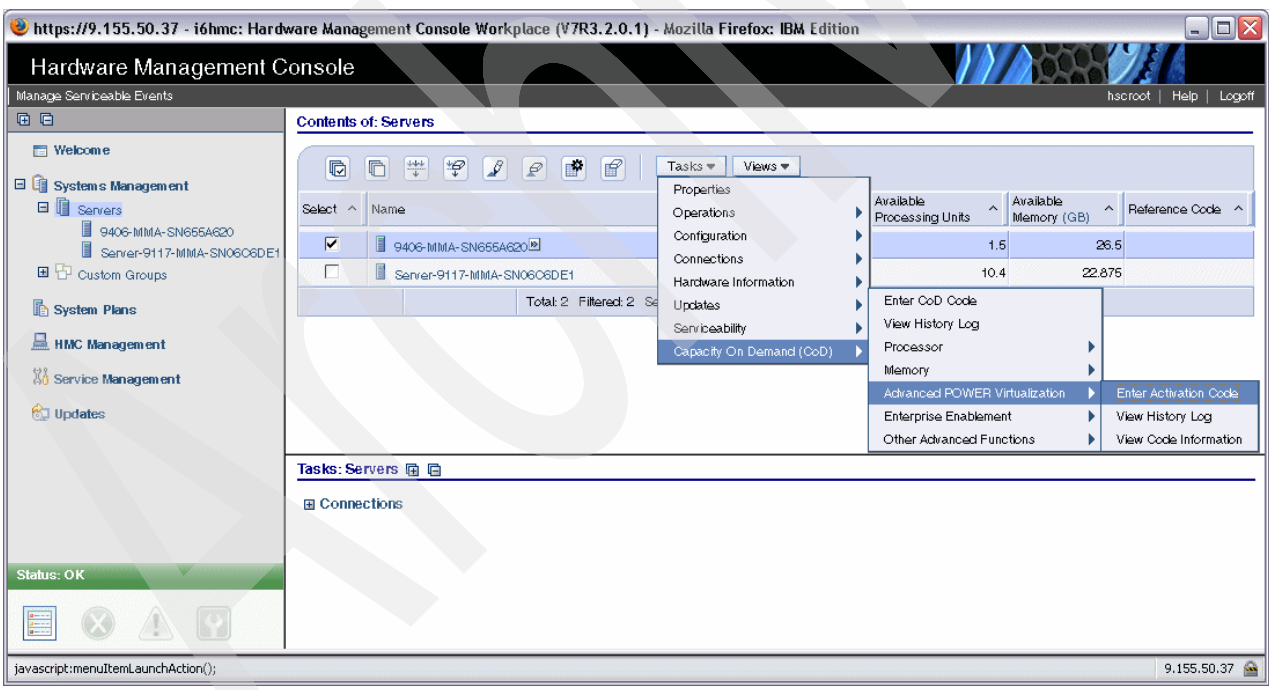


Figure 6-2 Entering PowerVM activation code

Now we are able create the VIOS partition i6VIOS for our IBM POWER Systems server:

1. From the HCM Systems Management → Servers view we mark the server for creating the new VIOS partition on and select **Tasks** → **Configuration** → **Create Logical Partition** → **VIO Server**, as shown in Figure 6-3.

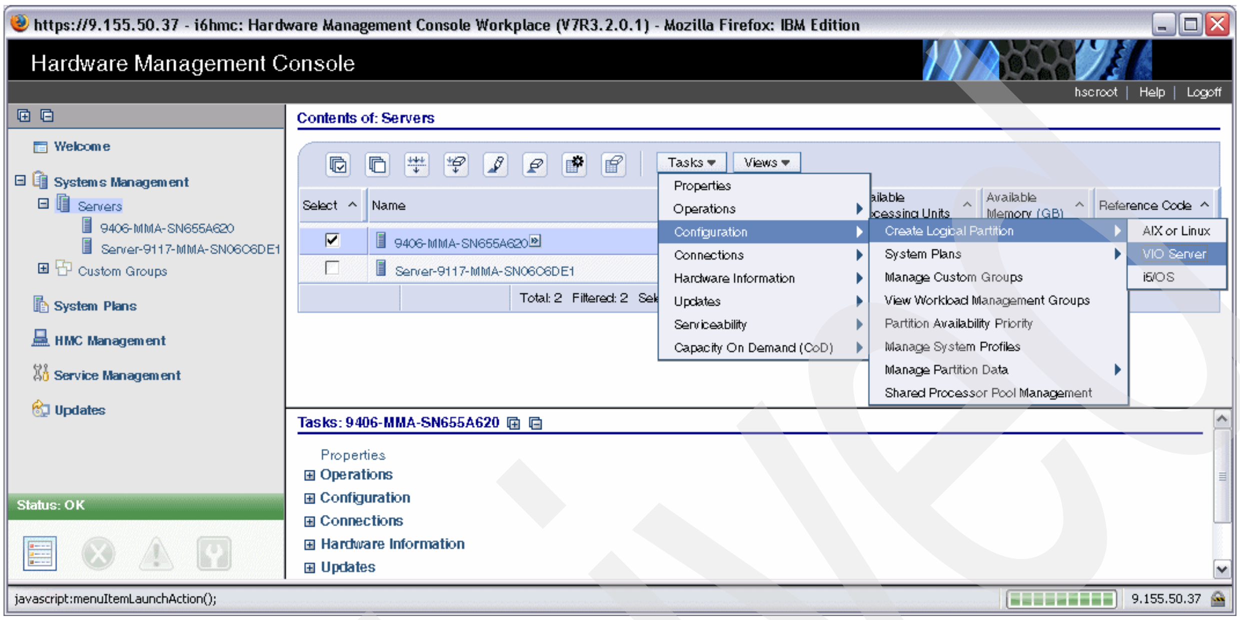


Figure 6-3 Create Logical Partition for VIOS

2. We enter an unused partition ID and partition name and click **Next** to proceed, as shown in Figure 6-4.

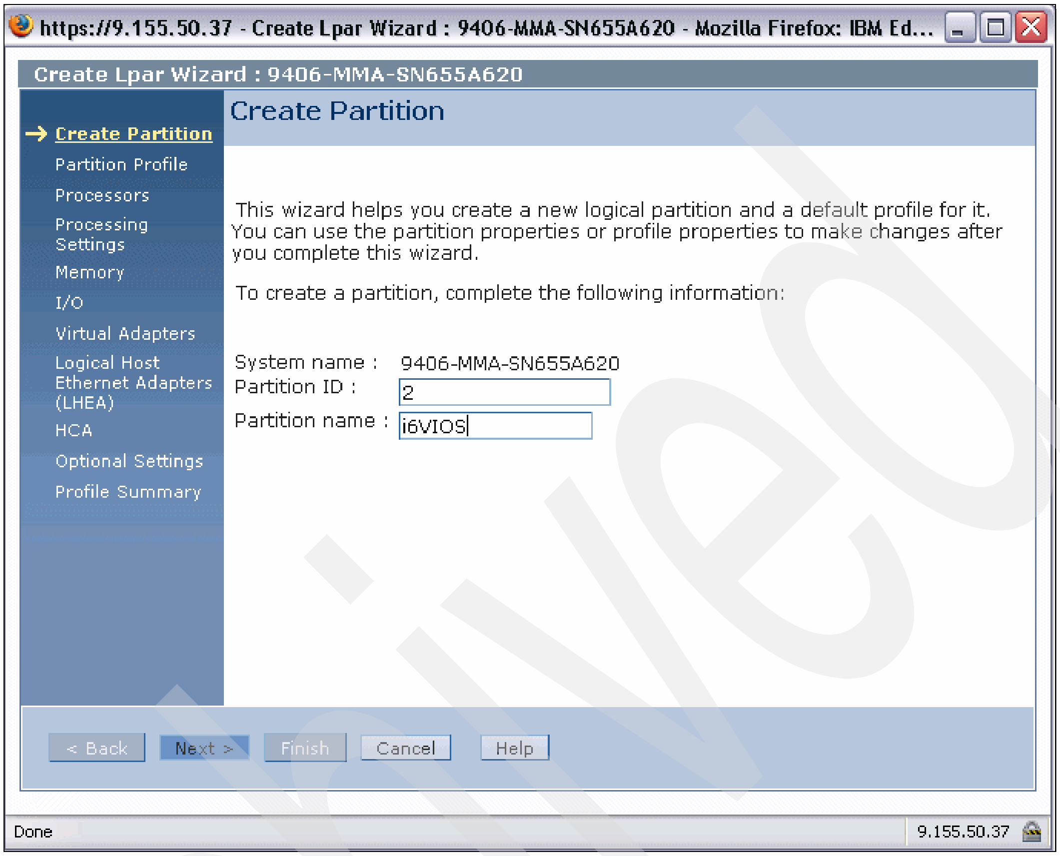


Figure 6-4 Create LPAR Create Partition dialog

3. We enter a profile name and click **Next** to proceed, as shown in Figure 6-5.

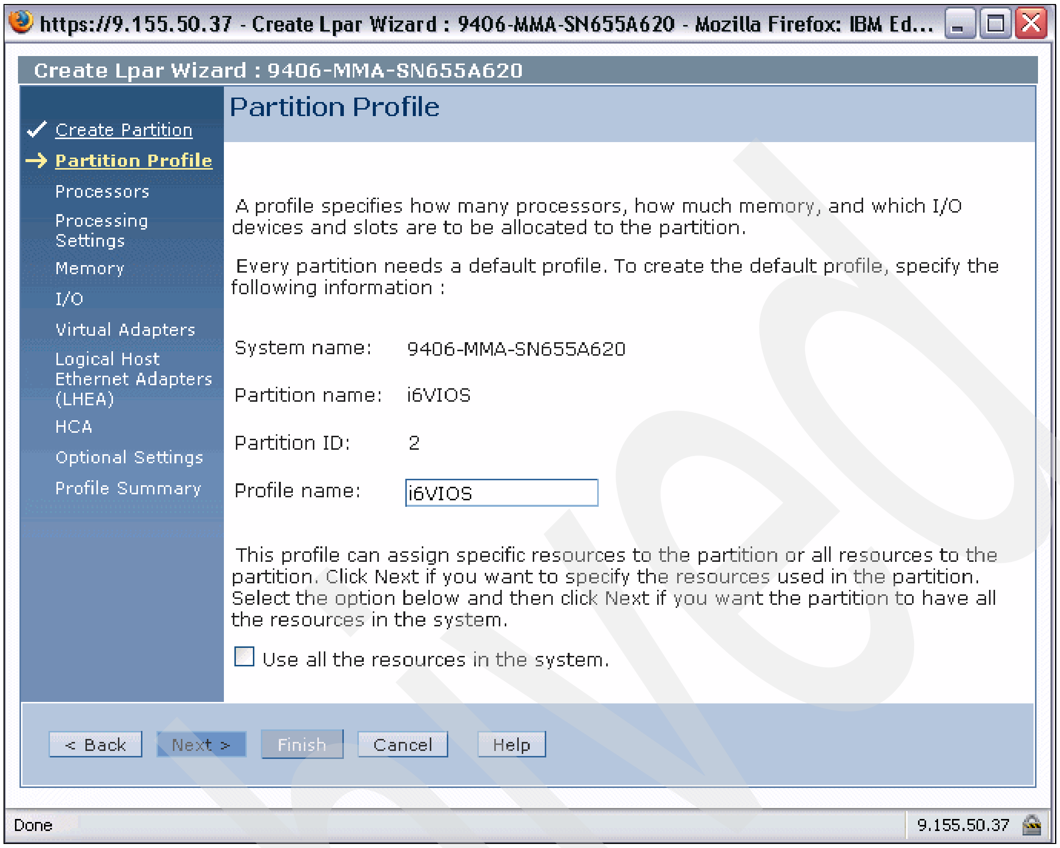


Figure 6-5 Create LPAR Partition Profile dialog

4. For best VIOS partition performance (see 4.2, “Planning for VIOS” on page 90) we select **Dedicated** processors and click **Next** to continue, as shown in Figure 6-6.

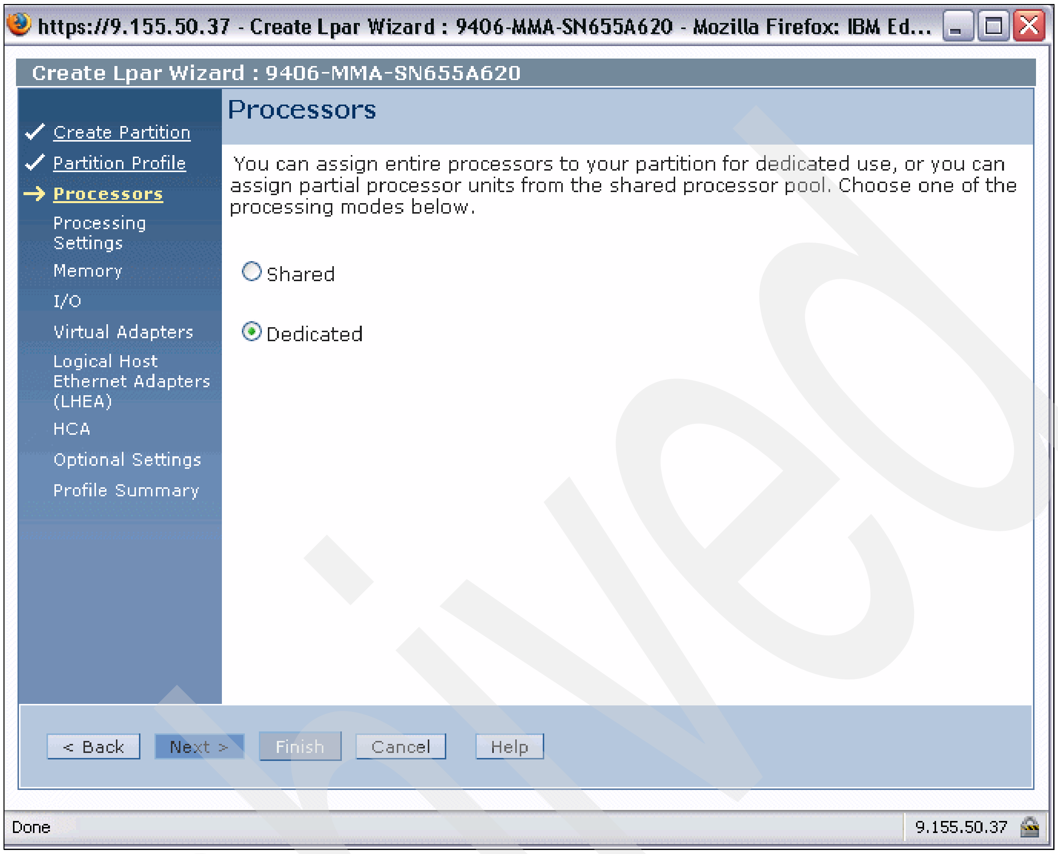


Figure 6-6 Create LPAR Processors dialog

5. We accept the default of one processor for the minimum, desired, and maximum processors settings and select **Next** to continue, as shown in Figure 6-7.

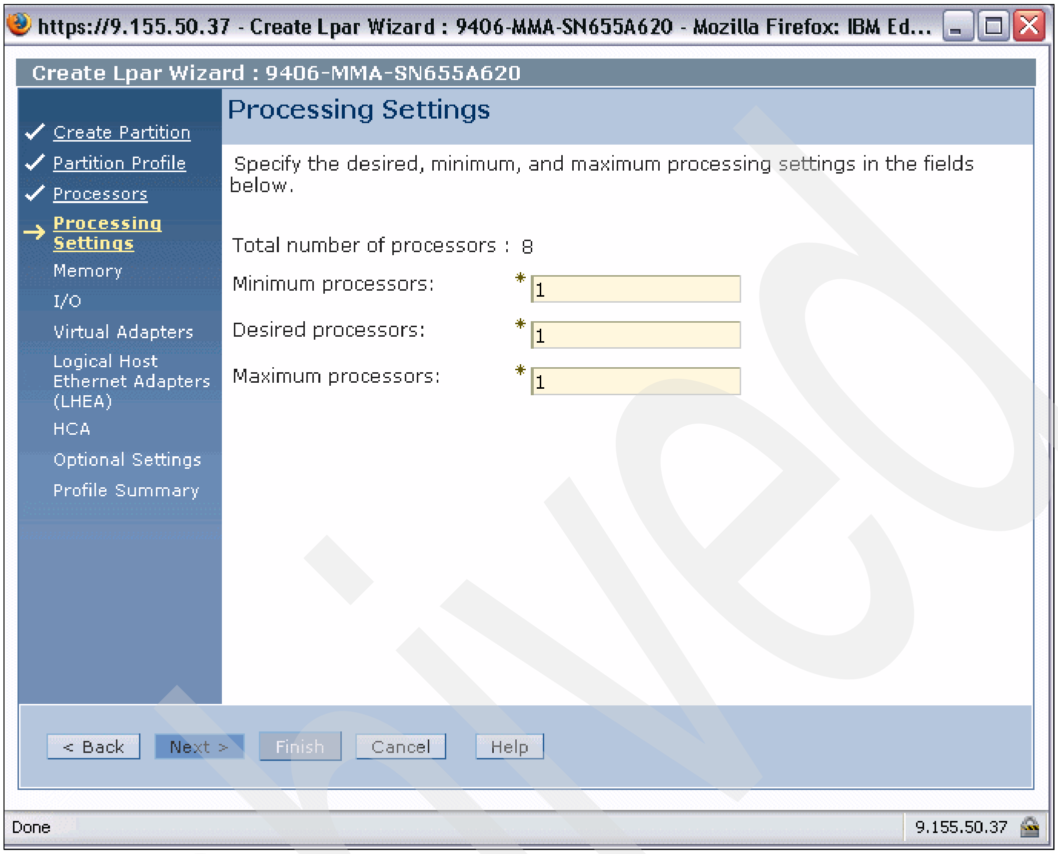


Figure 6-7 Create LPAR Processing Settings dialog

6. We select 1 GB memory for the minimum, desired, and maximum memory settings for the new VIOS partition and click **Next** to continue, as shown in Figure 6-8.

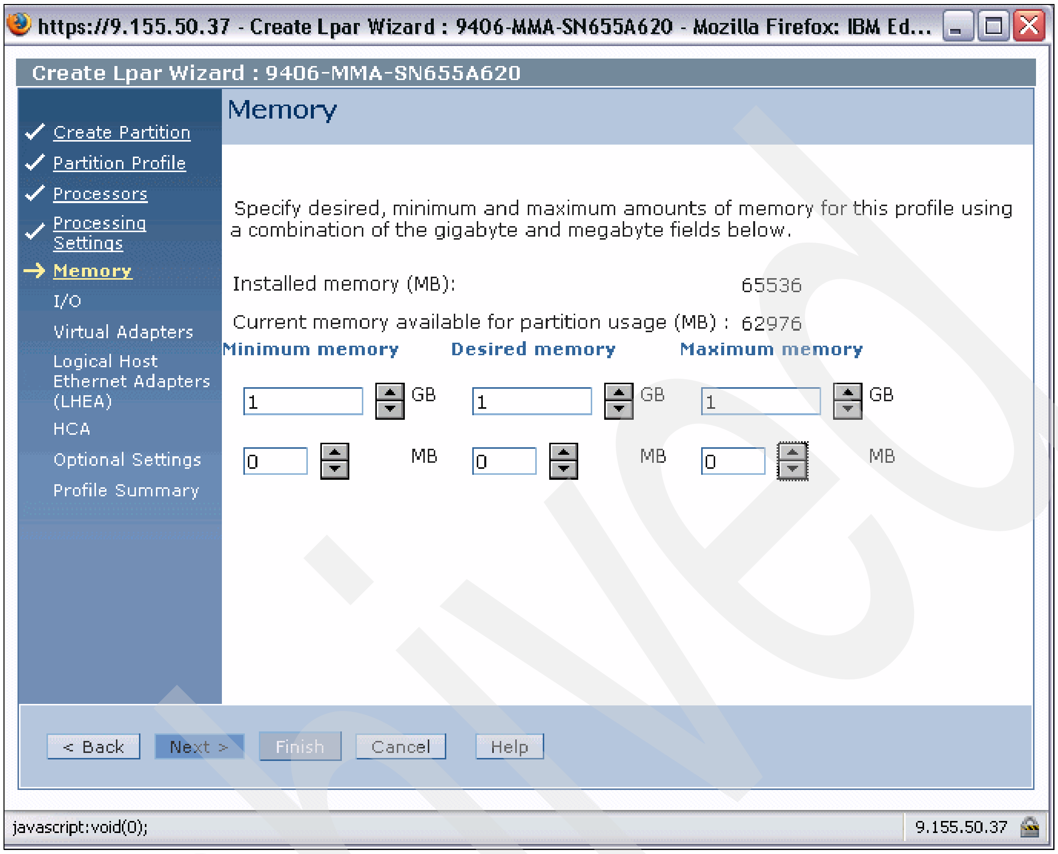


Figure 6-8 Create LPAR Memory dialog

7. In the I/O dialog we select the CEC's internal RAID controller in slot T3 and two FiberChannel adapters, one from each CEC's slot C3 before choosing **Add as required** and clicking **Next** to continue, as shown in Figure 6-9. The RAID controller provides us the with internal SAS drives that we use for installing VIOS and also the DVD-RAM drive.

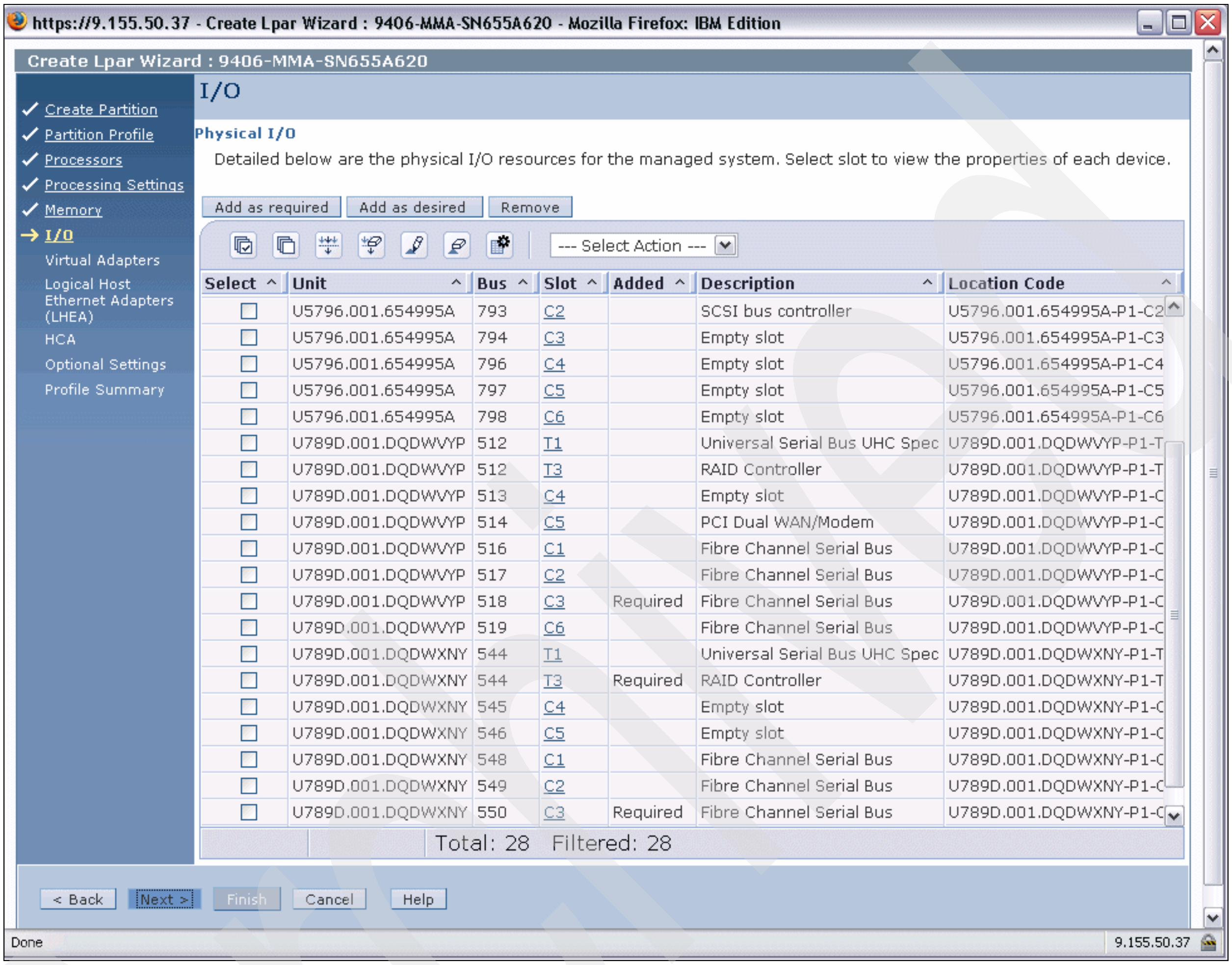


Figure 6-9 Create LPAR I/O dialog

8. In the Virtual Adapters dialog we choose **Actions** → **Create** → **Ethernet Adapter** (as shown in Figure 6-10) to create a virtual Ethernet adapter for inter-partition VLAN. This way we do not need to share the physical HEA port for inter-partition communication (see 1.2.2, “Virtual Ethernet support” on page 5).

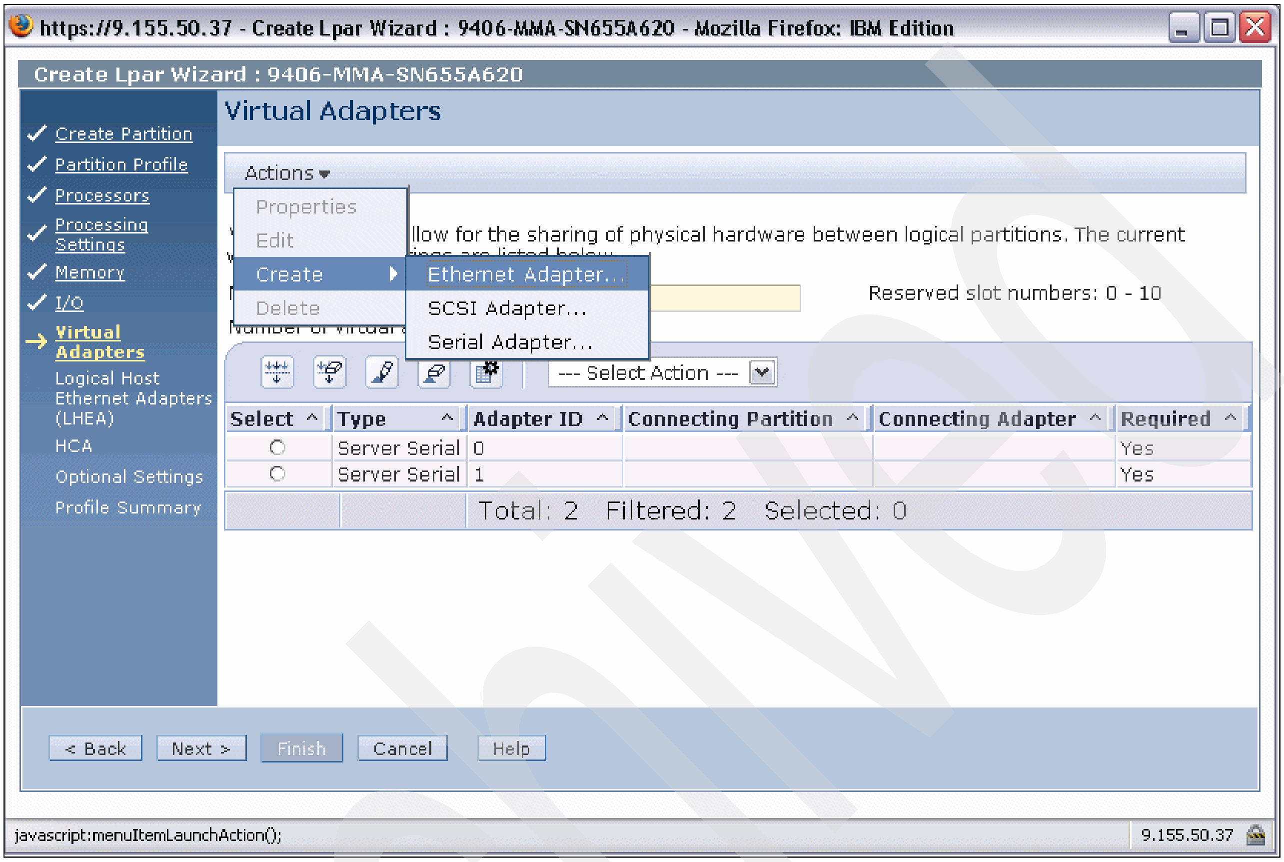


Figure 6-10 Create LPAR Virtual Adapters dialog

9. In the Create Virtual Ethernet Adapter dialog we accept the suggested values for the adapter ID of 11 and the VLAN ID of 1, check mark **This adapter is required for partition activation**, and select **OK** to proceed, as shown in Figure 6-11.

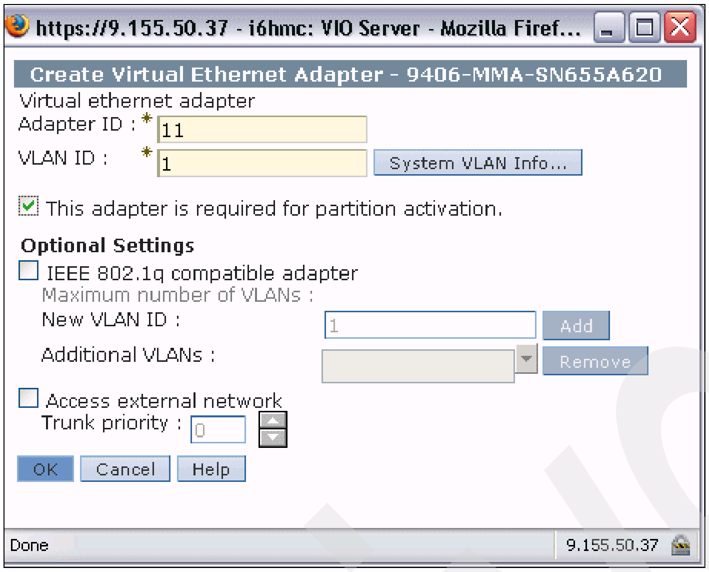


Figure 6-11 Create Virtual Ethernet Adapter dialog

10.The Virtual Adapters dialog now shows the newly created Ethernet adapter and we select **Actions** → **Create** → **SCSI Adapter**, as shown in Figure 6-12, to create virtual SCSI server adapters on the VIOS partition that the IBM i client partition can connect to.

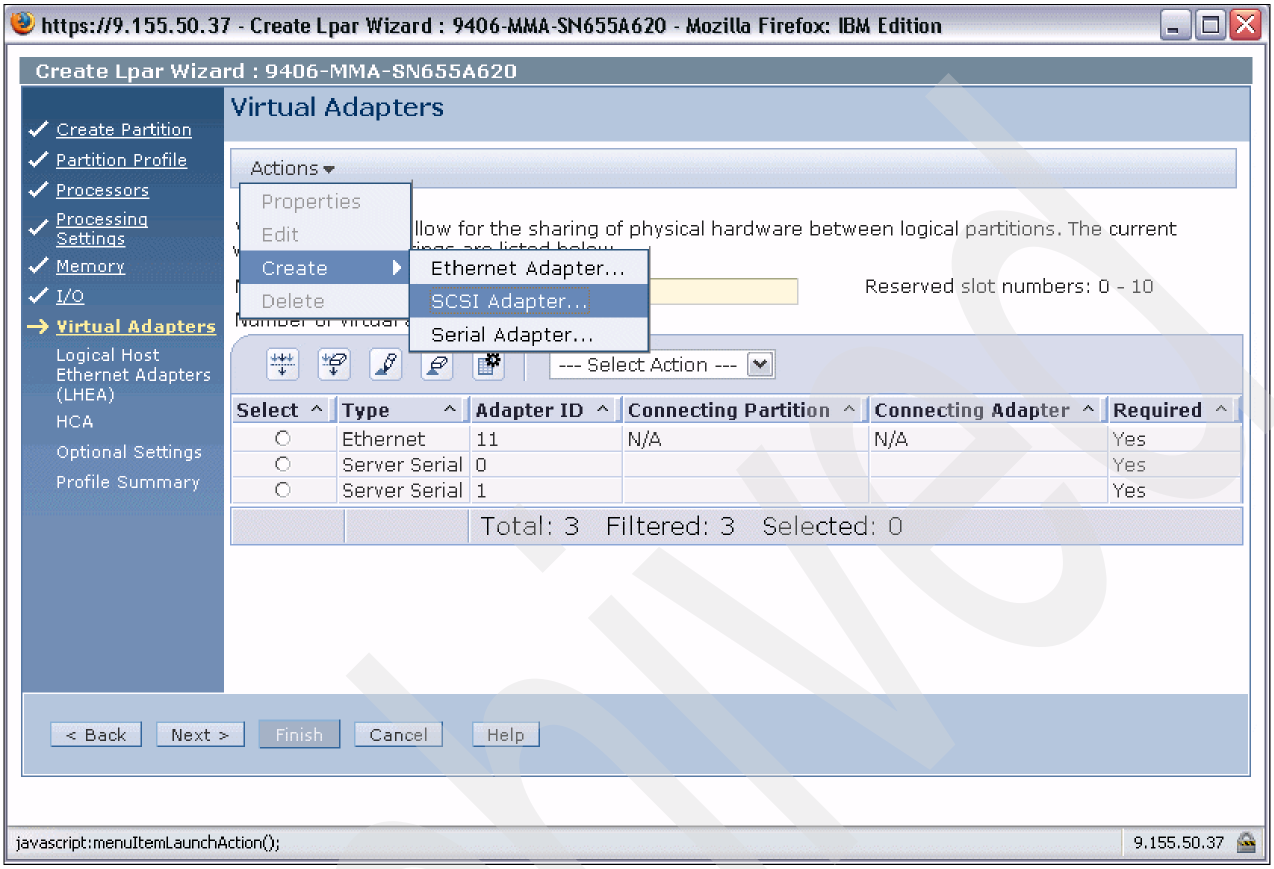


Figure 6-12 Create LPAR Virtual Adapters dialog

11.We accept the suggested virtual SCSI adapter ID of 12, leave the setting Any client partition can connect for the moment, check mark **This adapter is required for partition activation**, and select **OK** to continue (Figure 6-13).

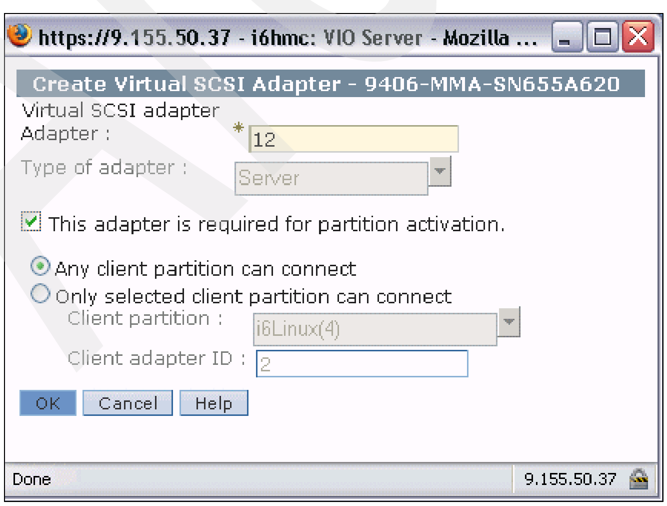


Figure 6-13 Create Virtual SCSI Adapter dialog

12. Now that we have created one virtual SCSI server adapter for VIOS we can use for the virtualized IBM System Storage DS4800 SCSI disk devices. Since we like to have a dedicated virtual SCSI server adapter for VIOS to virtualize the DVD-RAM we repeat step 11 above to create another virtual SCSI adapter. Figure 6-14 shows our virtual adapter configuration with the one Ethernet adapter and the two SCSI server adapters that we just created. Select **Next** to proceed.

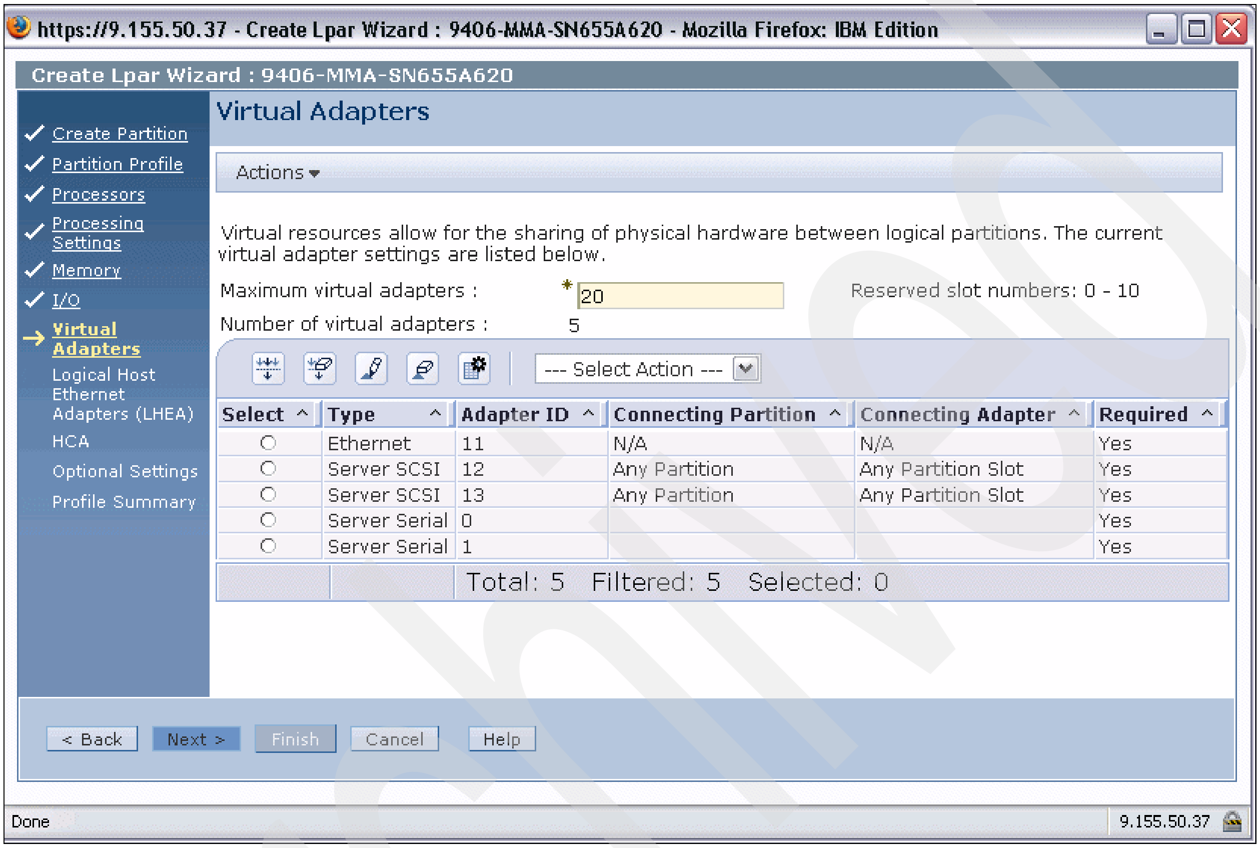


Figure 6-14 Create LPAR Virtual Adapters dialog

13. In the Logical Host Ethernet Adapters dialog shown in Figure 6-15 we select the CEC's HEA that we want to use from the drop-down list box, select the physical port, and choose **Configure** to proceed.

> **Note:** We leave the default LHEA capability setting of Base Minimum, as the LHEA capability queuing options are not implemented yet by PHYP or the operating system (OS).

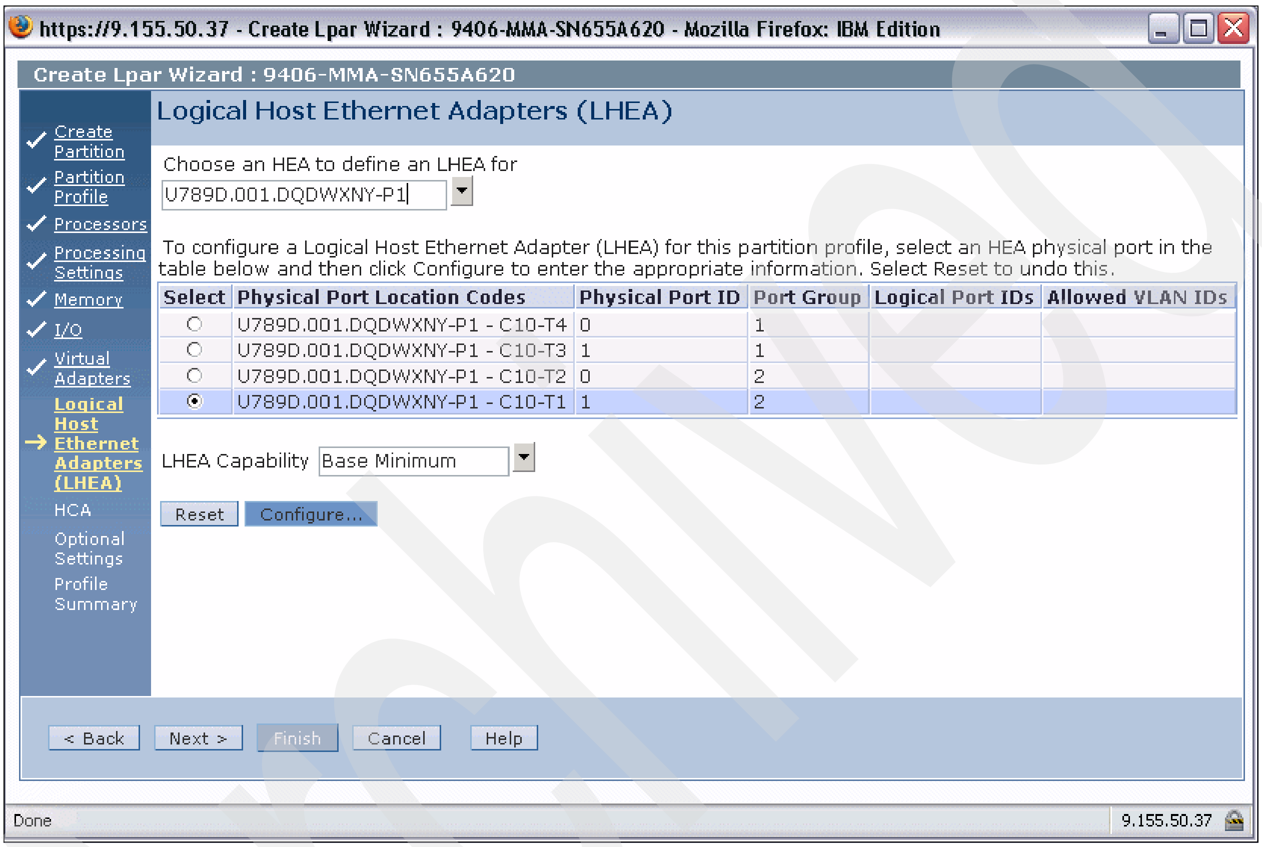


Figure 6-15 Create LPAR Logical Host Ethernet Adapters dialog

14. The Logical Host Ethernet Adapter Configuration dialog (Figure 6-16) shows that there are no other partitions using the HEA port T1 we selected, so we choose logical port ID 1, select **Allow all VLAN IDs**, and click **OK** to proceed.

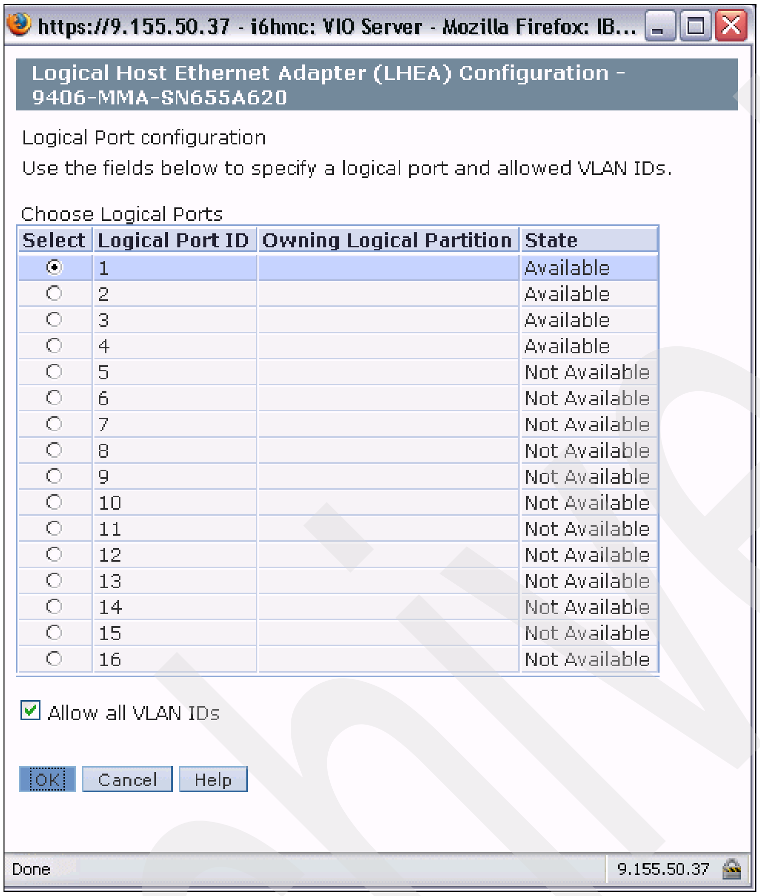


Figure 6-16 Logical Host Ethernet Adapter Configuration dialog

15. The Logical Host Ethernet Adapters dialog now reflects our HEA port logical configuration, as shown in Figure 6-17, and we select **Next** to proceed.

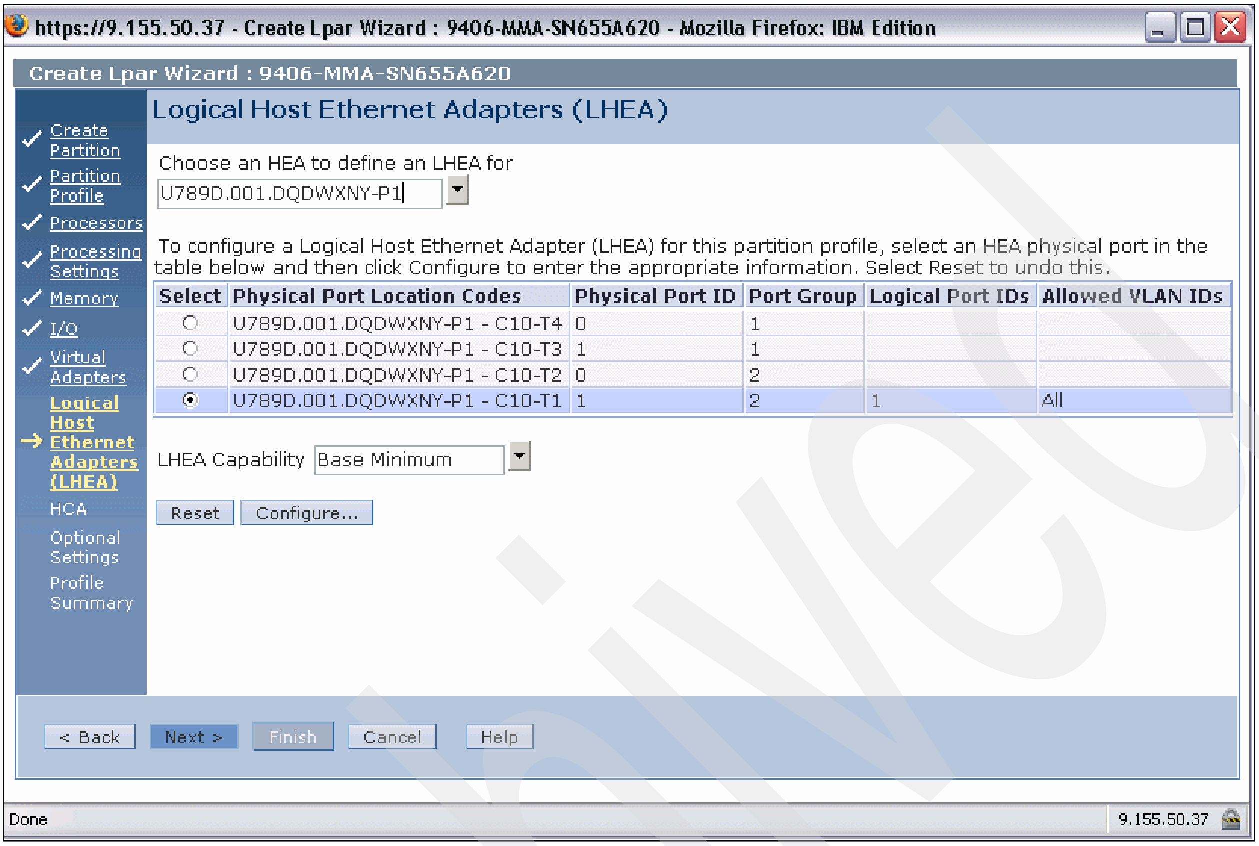


Figure 6-17 Logical Host Ethernet Adapters dialog

16. We select **Next** on the HCA dialog shown in Figure 6-18 to skip configuring a Host Channel Adapter since we do not use InfiniBand® clustering.

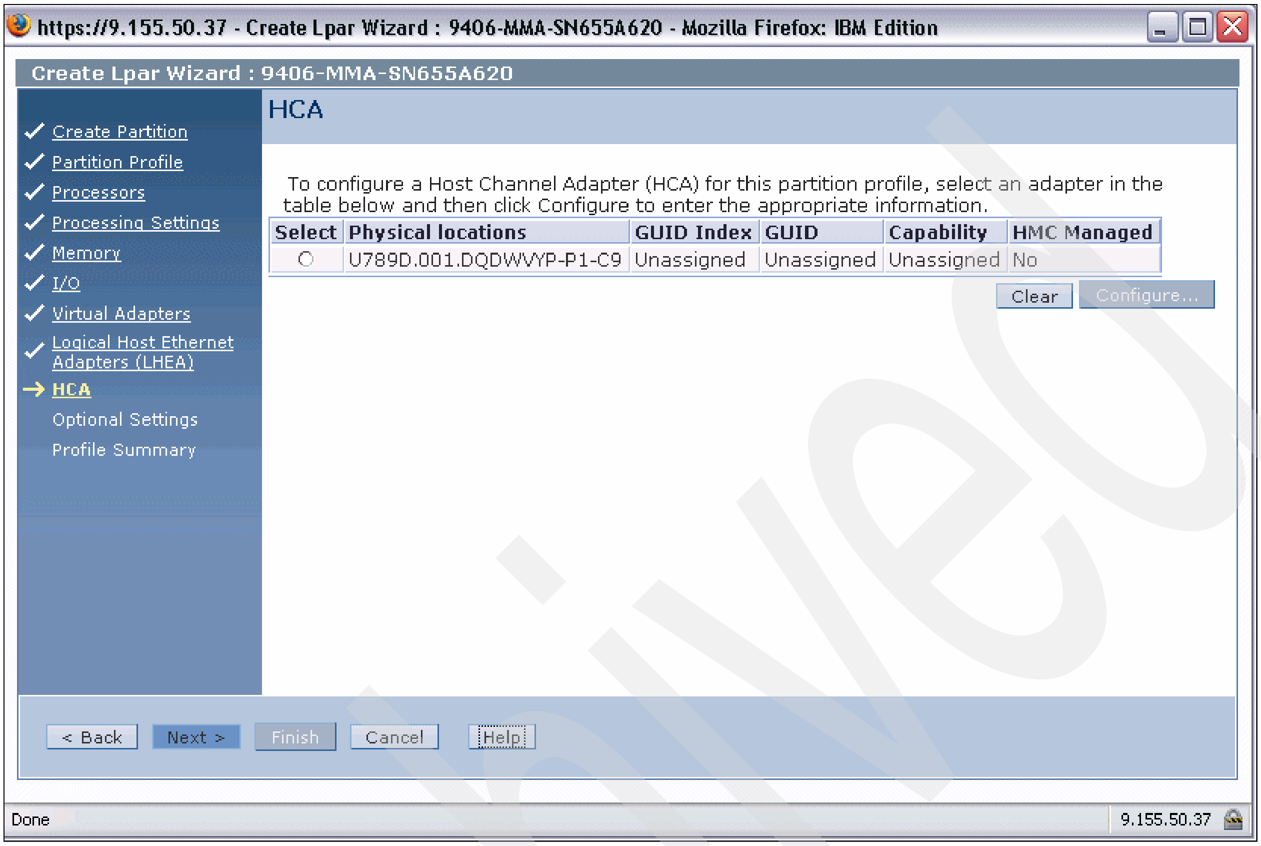


Figure 6-18 Create LPAR HCA dialog

17. In the Optional Settings dialog shown in Figure 6-19 we accept the default Normal setting for the boot mode and select **Next** to proceed.

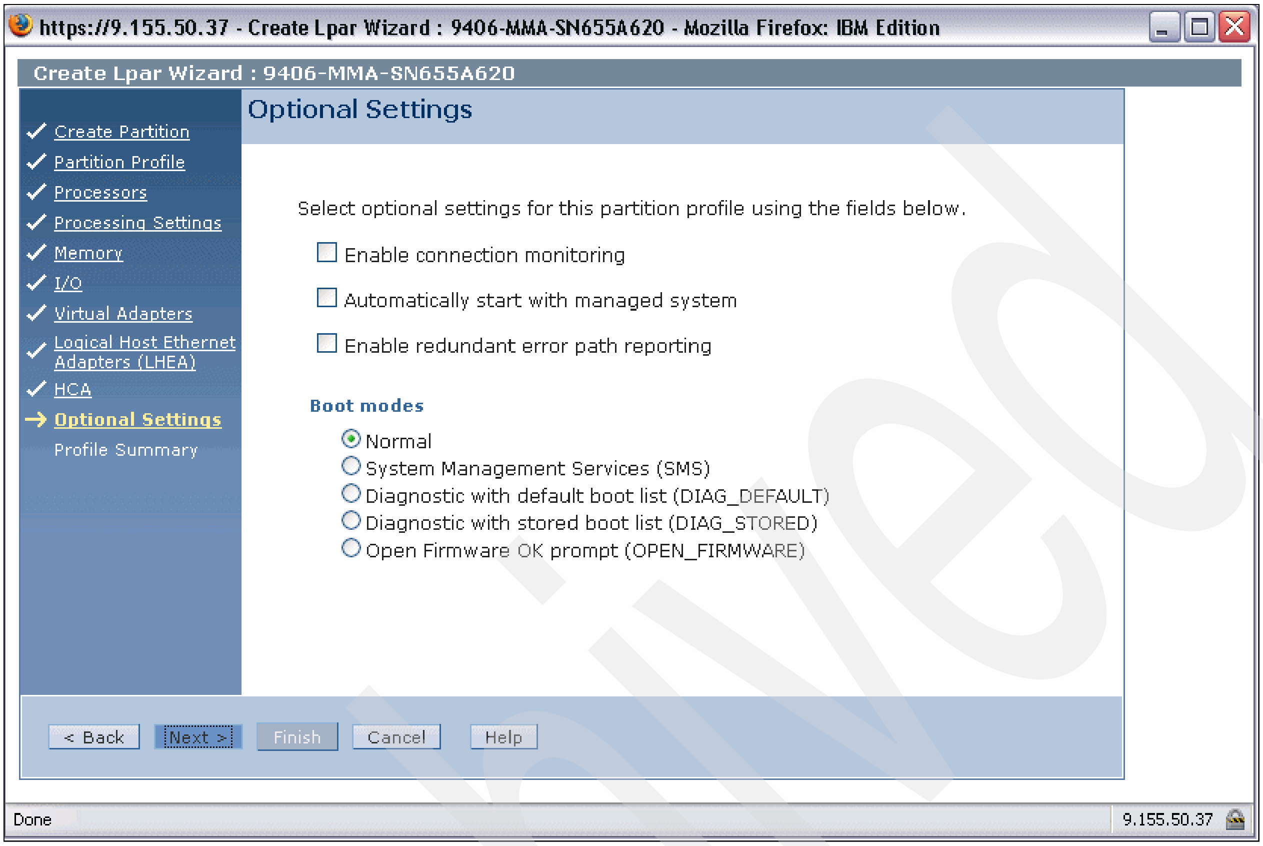


Figure 6-19 Create LPAR Optional Settings dialog

18. We review the information in the Profile Summary dialog shown in Figure 6-20 and select **Finish** to finally create our new VIOS partition.

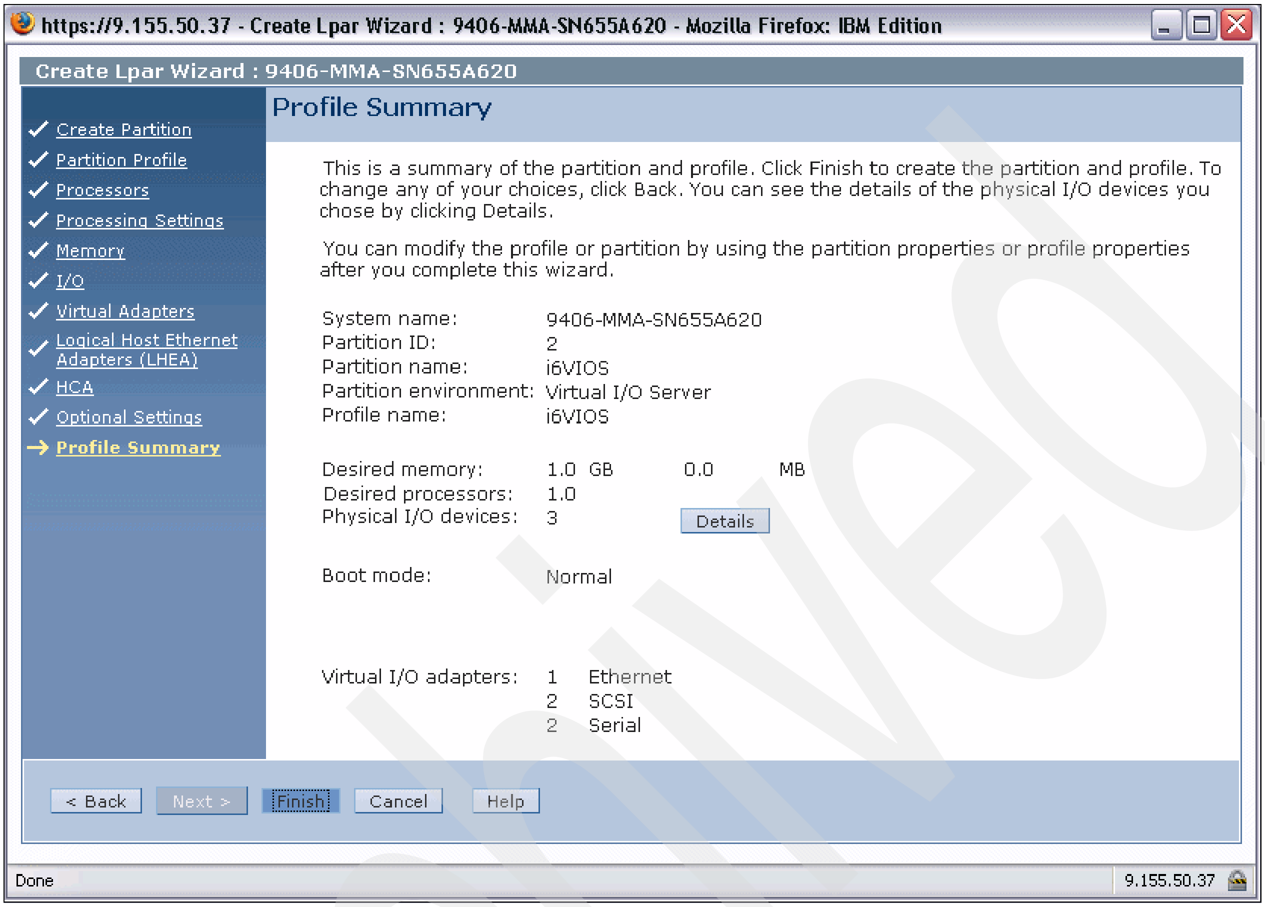


Figure 6-20 Create LPAR Profile Summary

6.2.2 Creating the IBM i Client LPAR

After creating the partition for the IBM Virtual I/O Server as described in 6.2.1, “Creating the VIOS LPAR” on page 137, we proceed with creating the IBM i client partition on our IBM POWER systems POWER6 server as described below.

1. From the HMC Systems Management → Servers view we mark our IBM i server that we want to create the IBM i partition on and select **Tasks** → **Configuration** → **Create Logical Partition** → **i5/OS**, as shown in Figure 6-21.

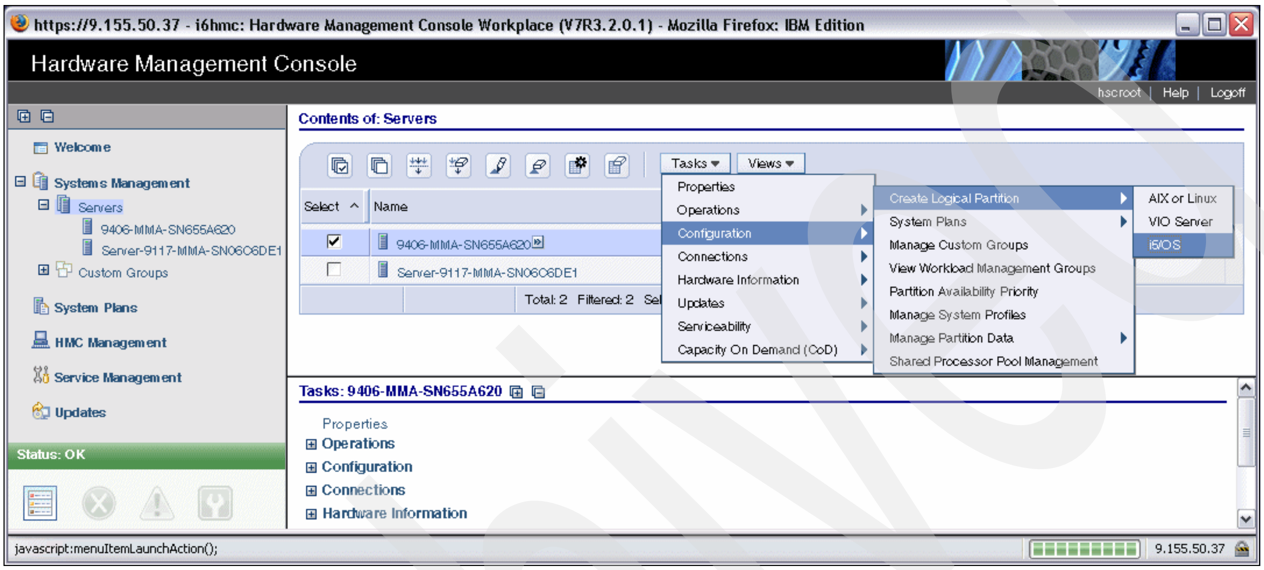


Figure 6-21 Create Logical Partition for IBM i

2. We enter an unused partition ID and partition name and click **Next** to proceed, as shown in Figure 6-22.

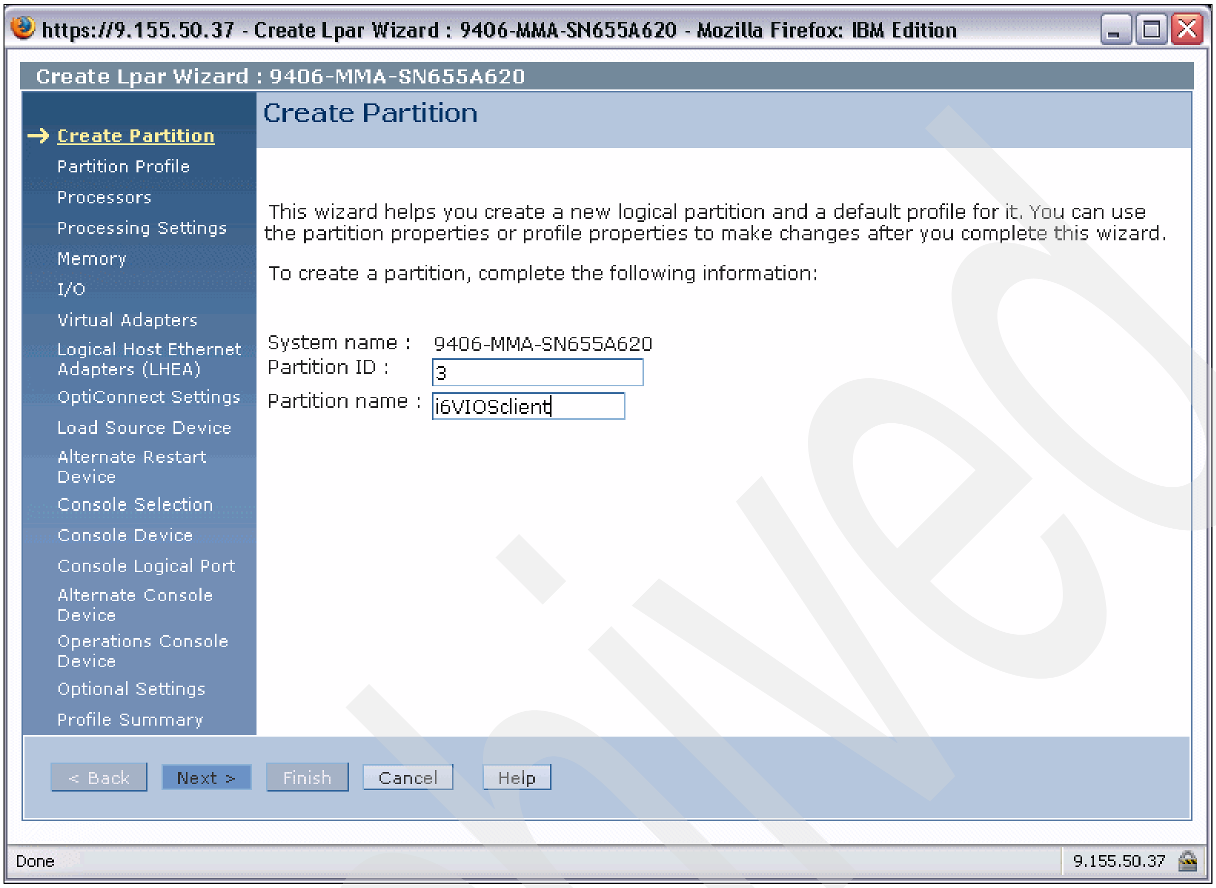


Figure 6-22 Create LPAR Create Partition dialog

3. We enter a profile name and click **Next** to proceed, as shown in Figure 6-23.

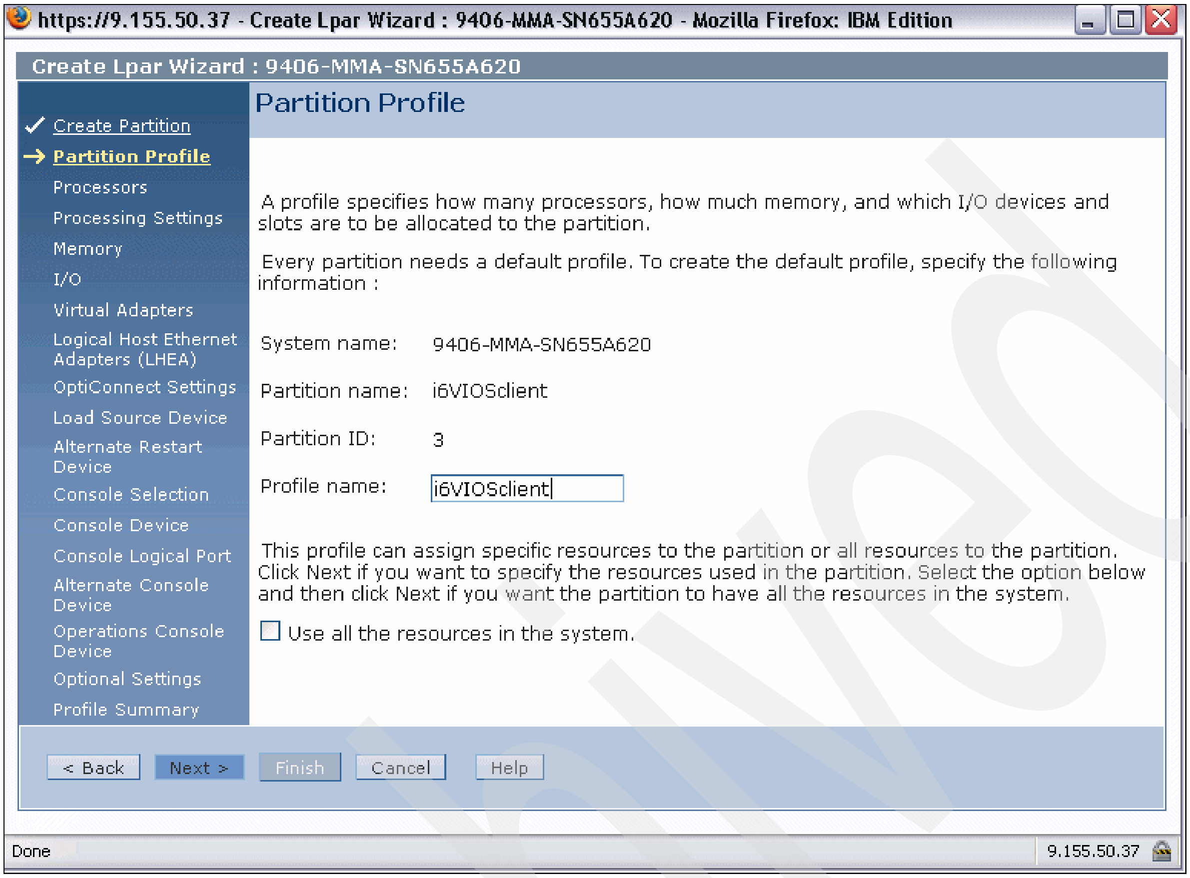


Figure 6-23 Create LPAR Partition Profile dialog

4. For our IBM i partition we prefer dedicated processors, so we select **Dedicated** processors and click **Next** to continue, as shown in Figure 6-24

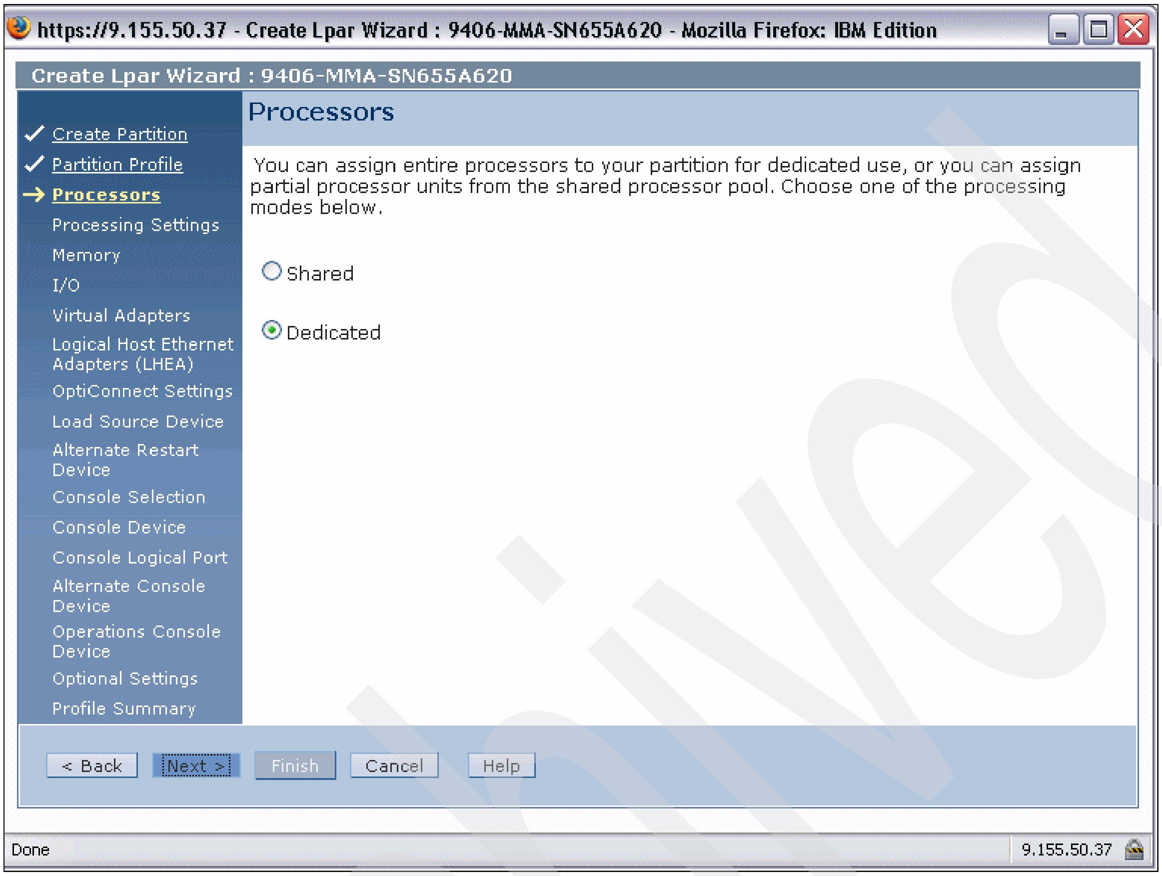


Figure 6-24 Create LPAR Processors dialog

5. We like to have two dedicated processors for our IBM i client partition so we enter the value of 2 for the minimum, desired, and maximum processors settings and click **Next** to proceed, as shown in Figure 6-25.

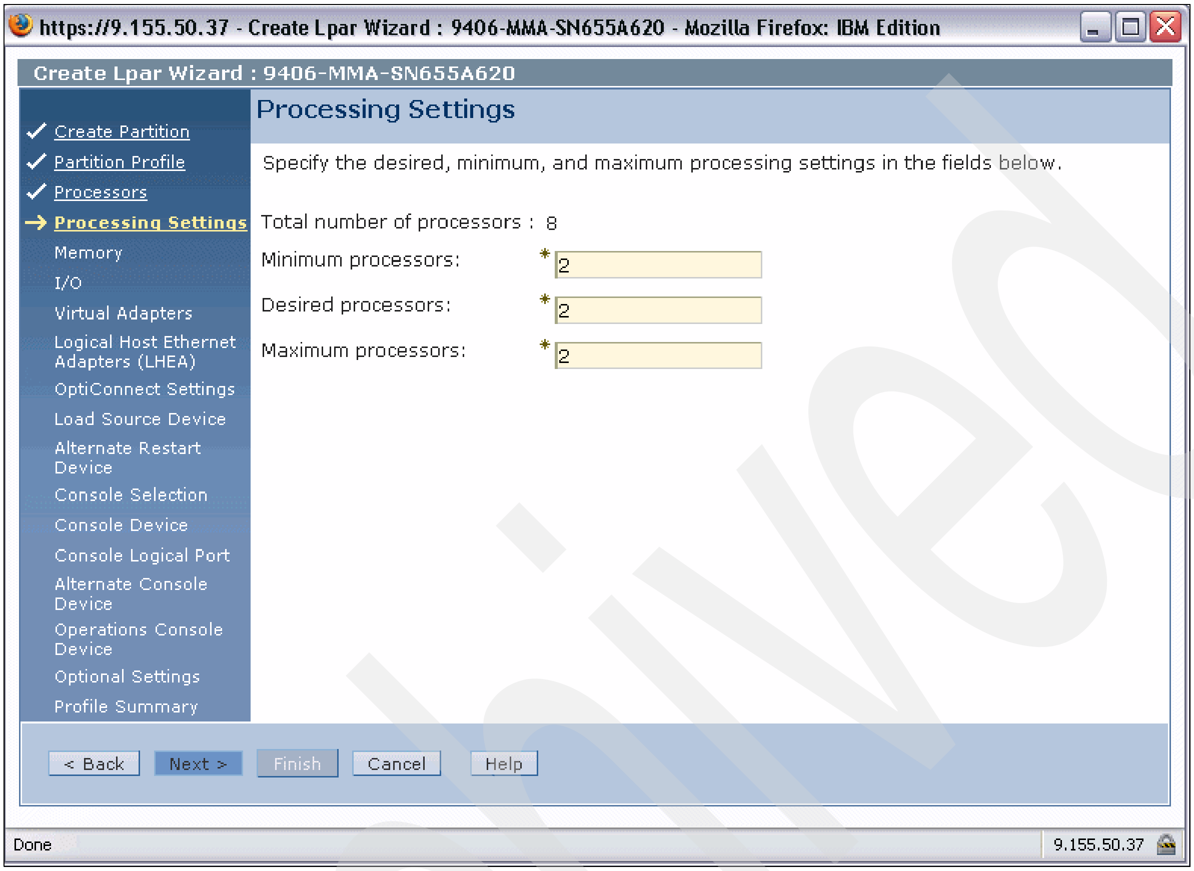


Figure 6-25 Create LPAR Processing Settings dialog

6. For our IBM i client partition we like to have 8 GB of main memory, so we enter the value of 8 for the minimum, desired, and maximum memory settings and click **Next** to proceed, as shown in Figure 6-26.

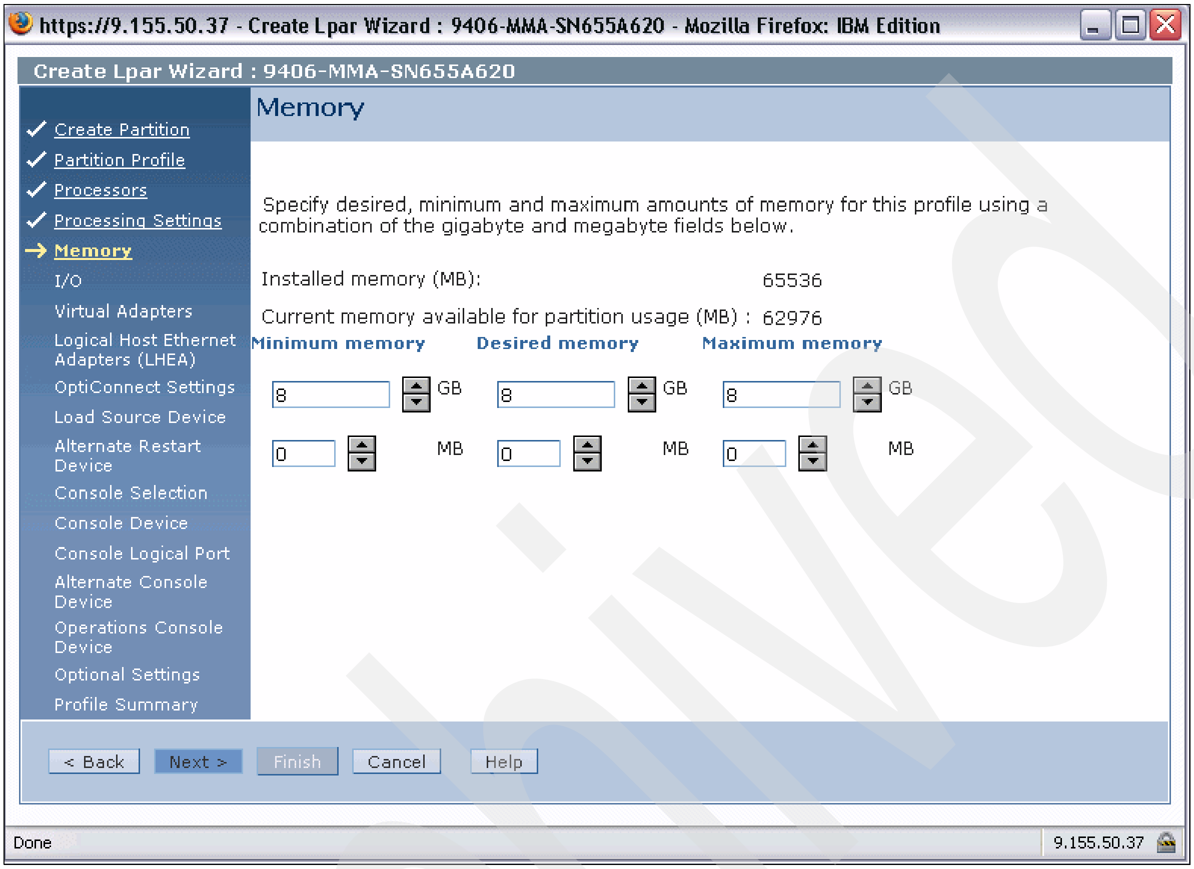


Figure 6-26 Create LPAR Memory dialog

7. Our IBM i client partition is not supposed to own any physical I/O hardware so we simply select **Next** in the I/O dialog, as shown in Figure 6-27.

> **Note:** If we wanted to attach a physical tape drive to our IBM i client partition we would have selected here a tape IOA to be added as desired to the partition profile.

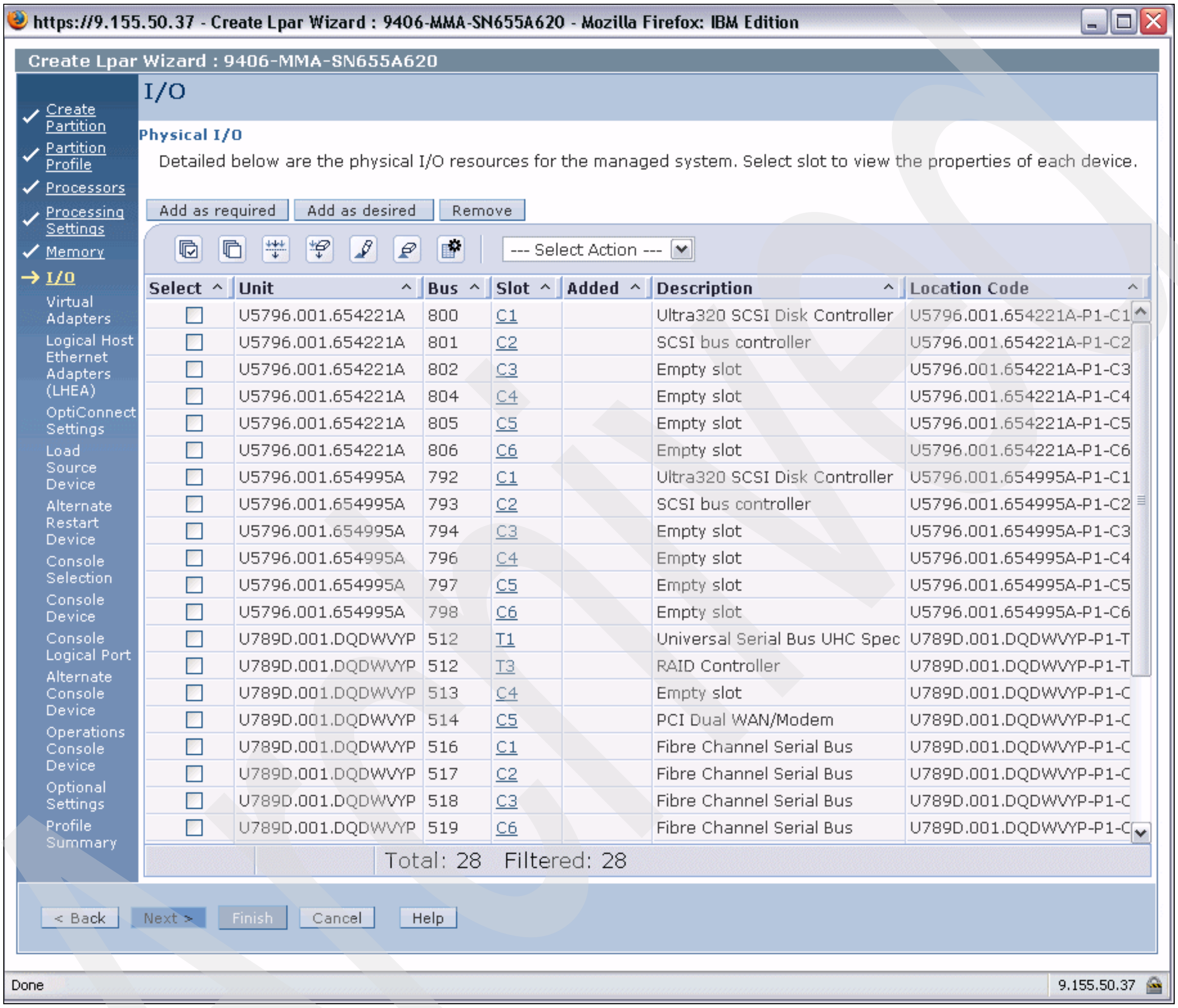


Figure 6-27 Create LPAR I/O dialog

8. Similar to the VIOS partition, we like to have a virtual Ethernet adapter on our IBM i client partition for inter-partition VLAN communication. To allow us to use the same virtual adapter IDs between VIOS and our IBM i client partition we first increase the maximum virtual adapters value to 20 and select **Actions** → **Create** → **Ethernet Adapter**, as shown in Figure 6-28.

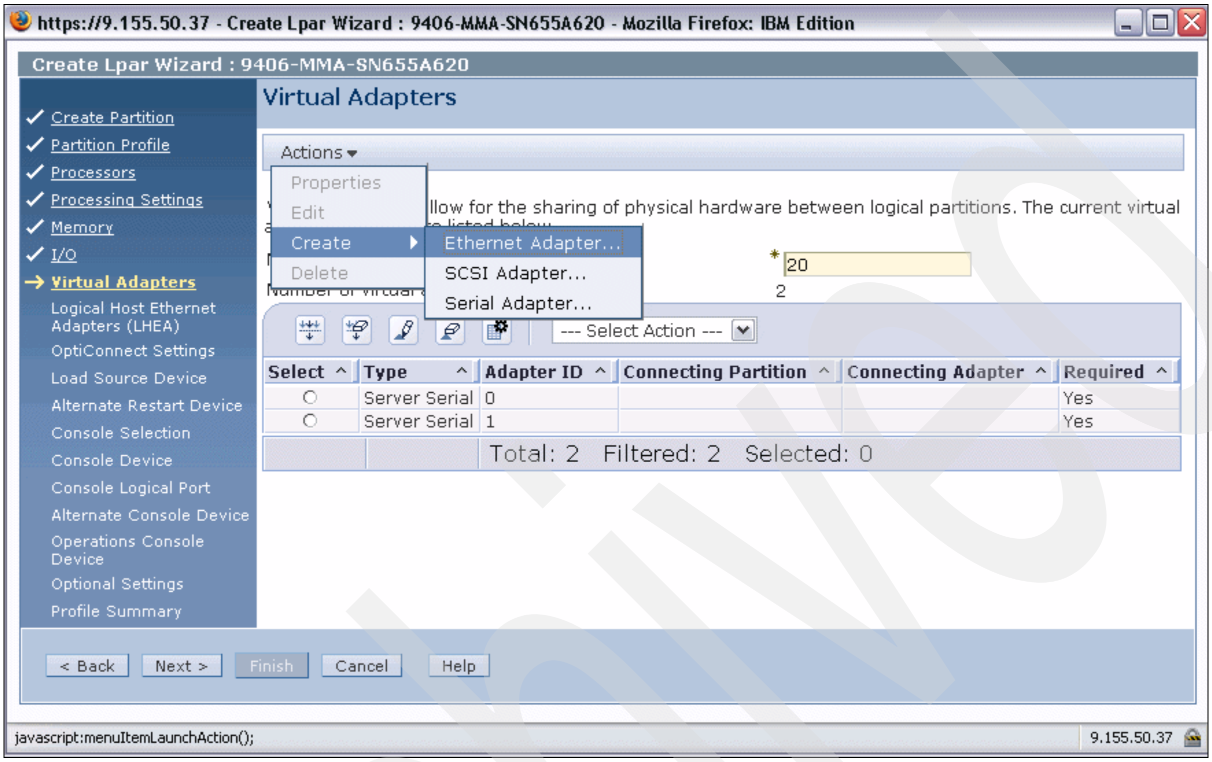


Figure 6-28 Create LPAR Virtual Adapters dialog

9. In the Create Virtual Ethernet Adapter dialog shown in Figure 6-29 we accept the suggested adapter ID of 2 and the VLAN ID of 1, mark **This adapter is required for partition activation**, and select **OK** to continue.

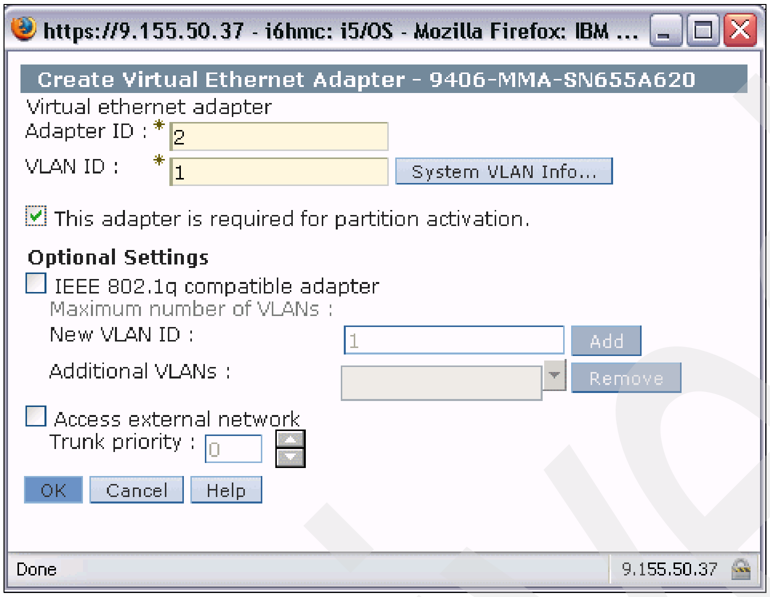


Figure 6-29 Create Virtual Ethernet Adapter dialog

10. The Virtual Adapters dialog now shows the newly created Ethernet adapter and we select **Actions** → **Create** → **SCSI Adapter** (as shown in Figure 6-30) to create virtual SCSI client adapters on our IBM i client partition used for connecting to the corresponding VIOS virtual SCSI server adapters.

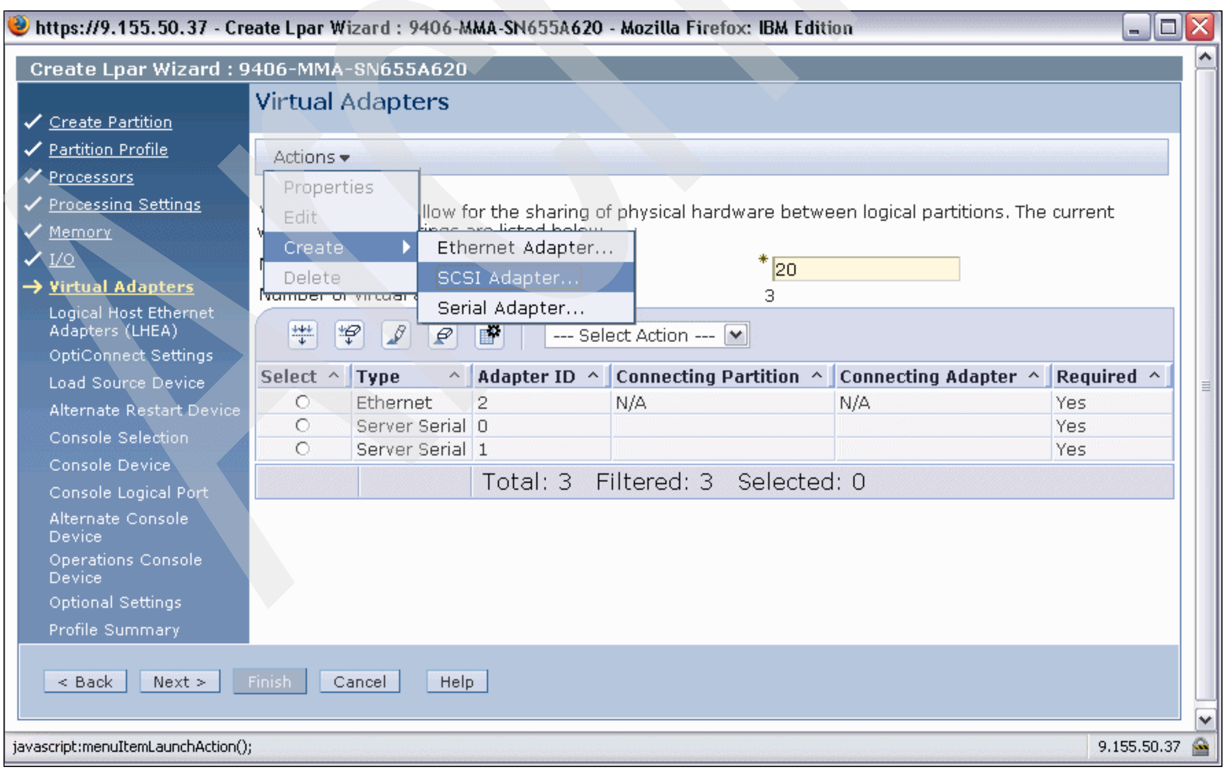


Figure 6-30 Create LPAR Virtual Adapters dialog

11. For the Virtual SCSI client adapter ID we specify the same ID that we used before when creating the virtual SCSI server adapter of the VIOS LPAR. For type of adapter we select **Client**, check mark **This adapter is required for partition activation**, select our previously created VIOS partition **i6VIOS** for the server partition, enter server adapter ID 12, and click **OK** to proceed, as shown in Figure 6-31.

> **Note:** Using the same adapter ID for the virtual SCSI server adapter and the corresponding virtual SCSI client adapter is not required but is a best practice to help keep an overview of the virtual SCSI server and client adapter connections.

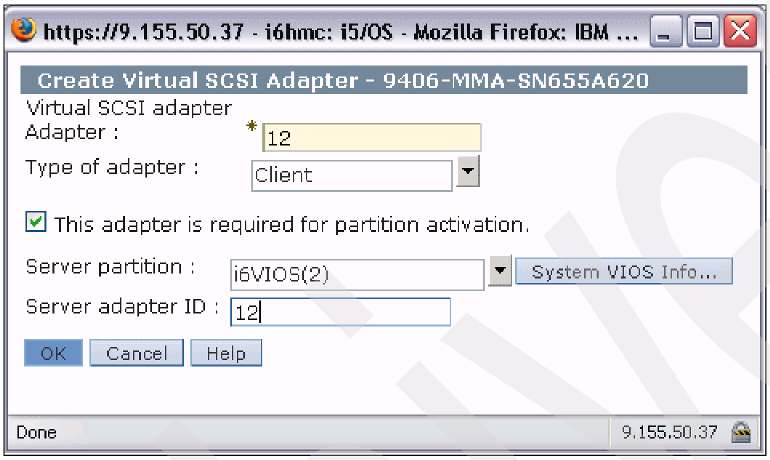


Figure 6-31 Create Virtual SCSI Adapter dialog

12. Now we have created one virtual SCSI client adapter for connecting to the virtual SCSI server adapter in our VIOS partition used for the virtualized IBM System Storage DS4800 SCSI disk devices. We repeat step 11 above to create another virtual SCSI client adapter to connect to the VIOS virtual SCSI server adapter used for virtualizing the DVD-RAM. Figure 6-32 shows our virtual adapter configuration with the one Ethernet adapter and the two SCSI client adapters that we just created. Select **Next** to proceed.

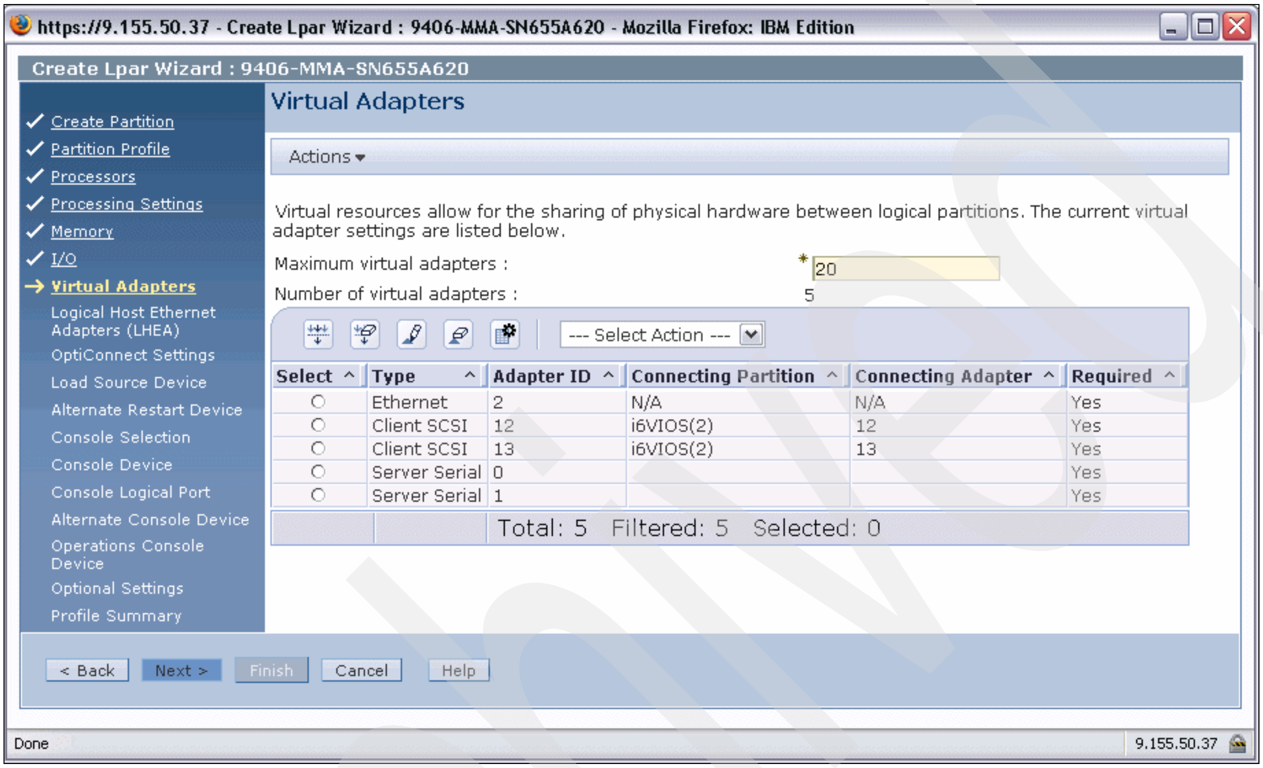


Figure 6-32 Create LPAR Virtual Adapters dialog

13. in the Logical Host Ethernet Adapters dialog shown in Figure 6-33 we select the CEC's HEA that we want to use from the drop-down list box, select the physical port, and choose **Configure** to proceed.

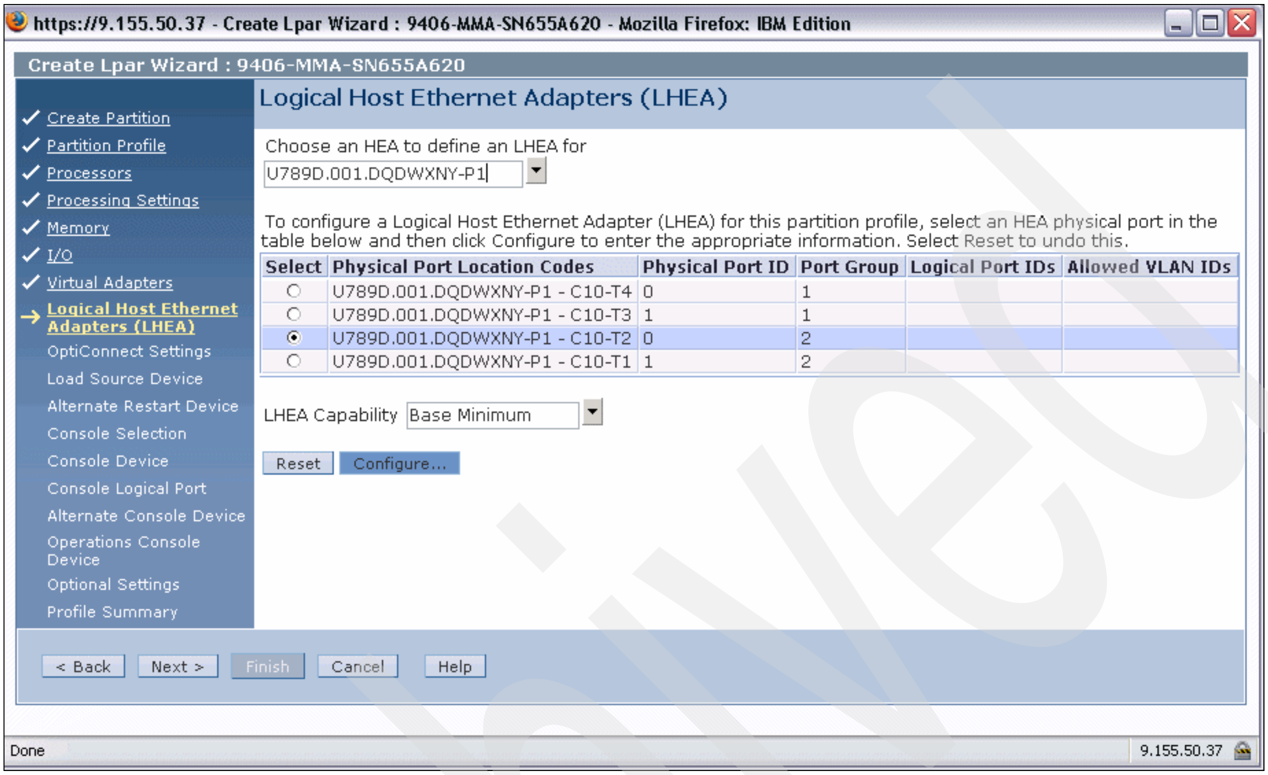


Figure 6-33 Logical Host Ethernet Adapters dialog

14. The Logical Host Ethernet Adapter Configuration dialog in Figure 6-34 shows that there are no other partitions using the HEA port T2 that we selected, so we choose the first available logical port with ID 1, select **Allow all VLAN IDs**, and click **OK** to proceed.

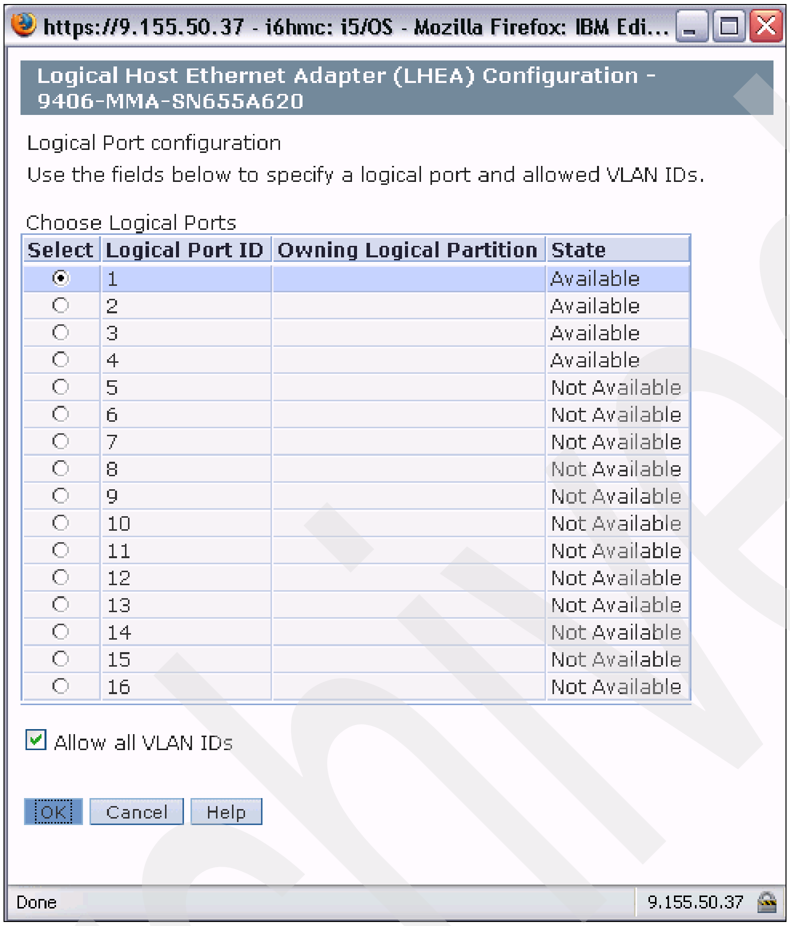


Figure 6-34 Logical Host Ethernet Adapter Configuration dialog

15. The Logical Host Ethernet Adapters dialog now reflects our HEA port logical configuration, as shown in Figure 6-35, and we select **Next** to proceed.

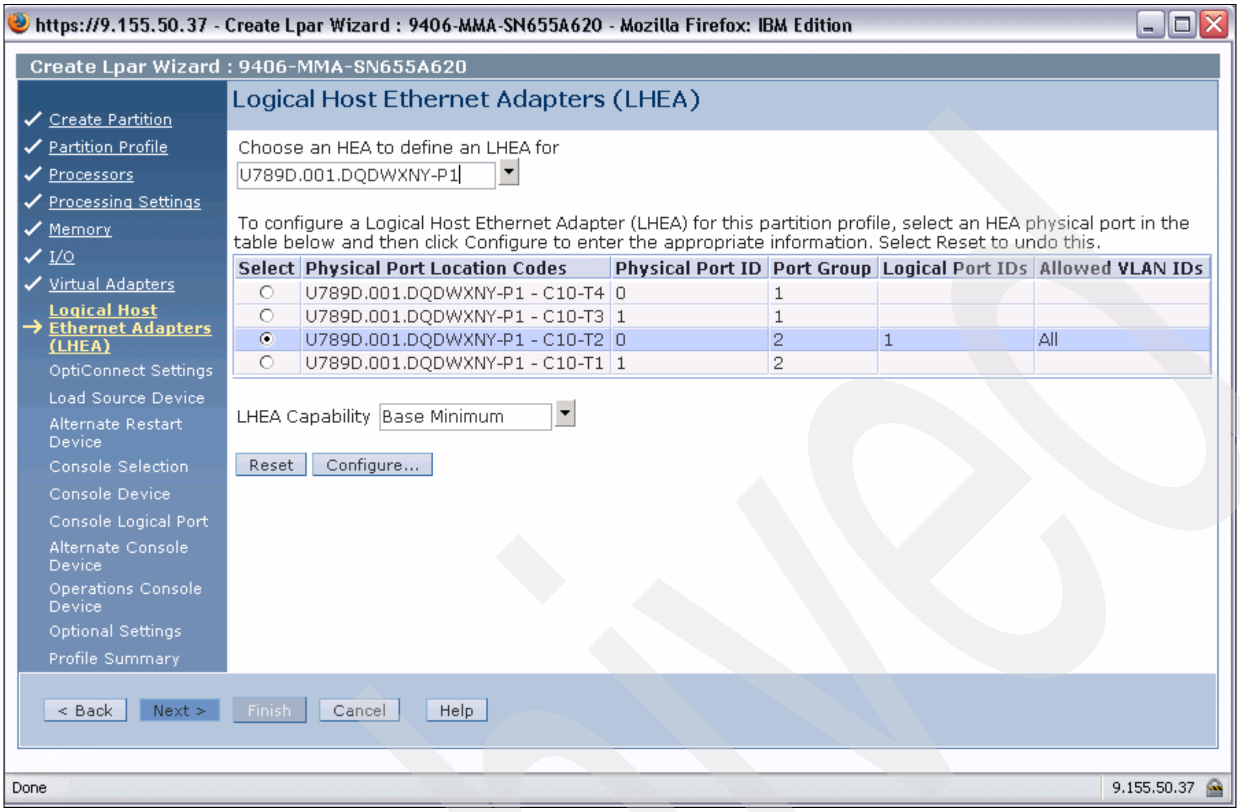


Figure 6-35 Logical Host Ethernet Adapters updated dialog

16.We do not want to use OptiConnect connectivity for our IBM i client partition so we click **Next**, as shown in Figure 6-36, to proceed.

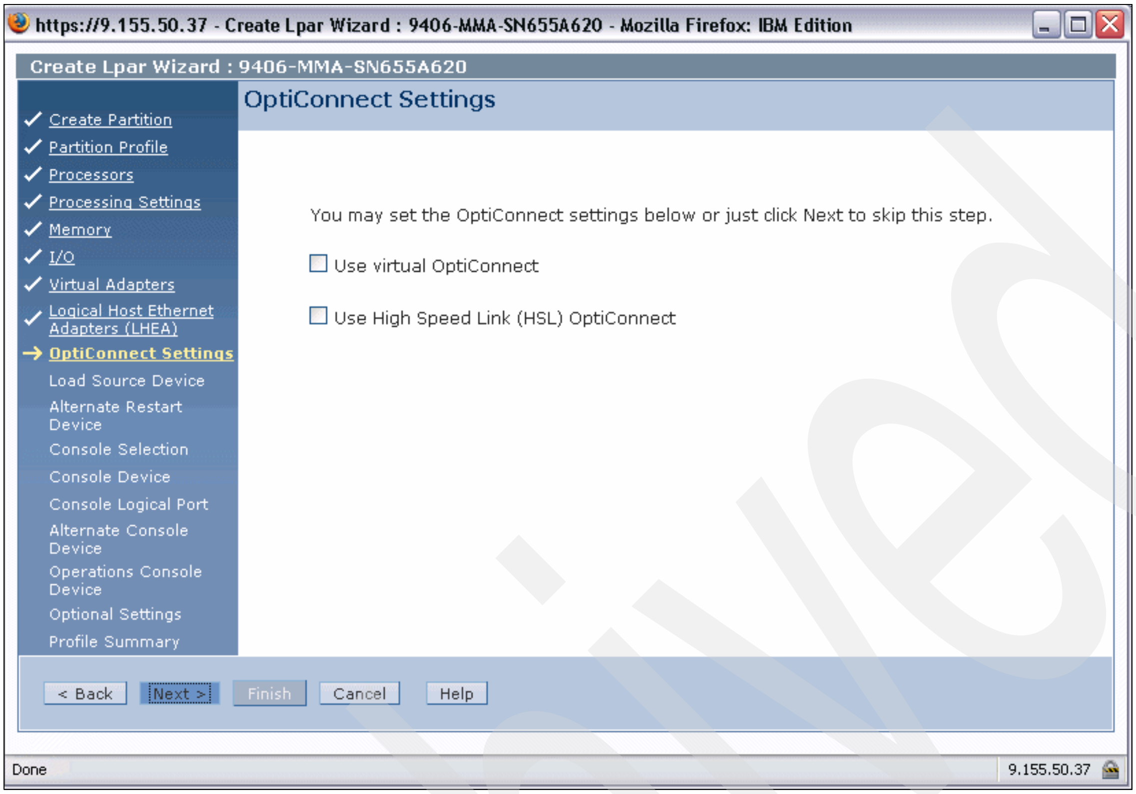


Figure 6-36 Create LPAR OptiConnect Settings dialog

17. For the load source device of our IBM i client partition we choose **Virtual Adapter** from the Adapter Type drop-down list box, select adapter ID **12** to use the first virtual adapter as the load source adapter to which we will later map the DS4800 SCSI devices to (see 6.5, “Configuring VIOS virtual devices” on page 191), and click **Next** to proceed, as shown in Figure 6-37.

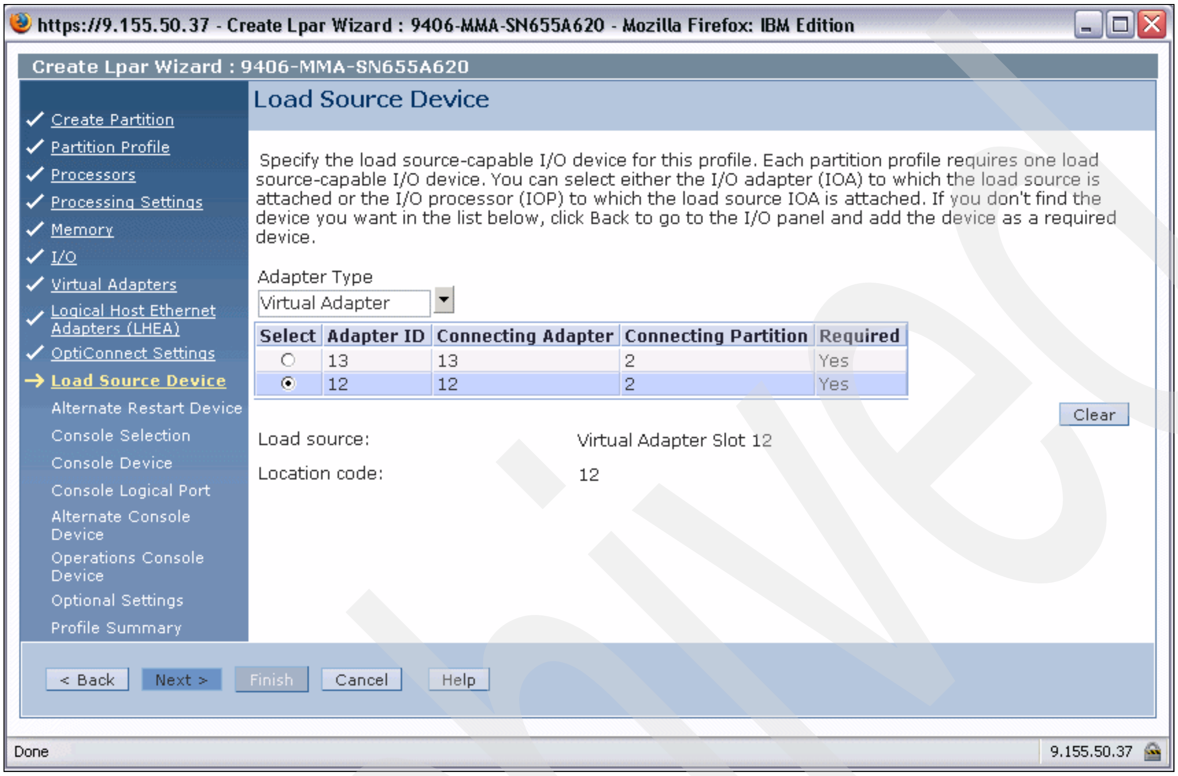


Figure 6-37 Create LPAR Load Source Device dialog

18. For specifying the alternate restart device we also choose **Virtual Adapter** from the Adapter Type drop-down list box, select adapter ID **13** to which we will later map the virtual DVD-RAM drive to (see also 6.5, “Configuring VIOS virtual devices” on page 191), and select **Next** to proceed, as shown in Figure 6-38.

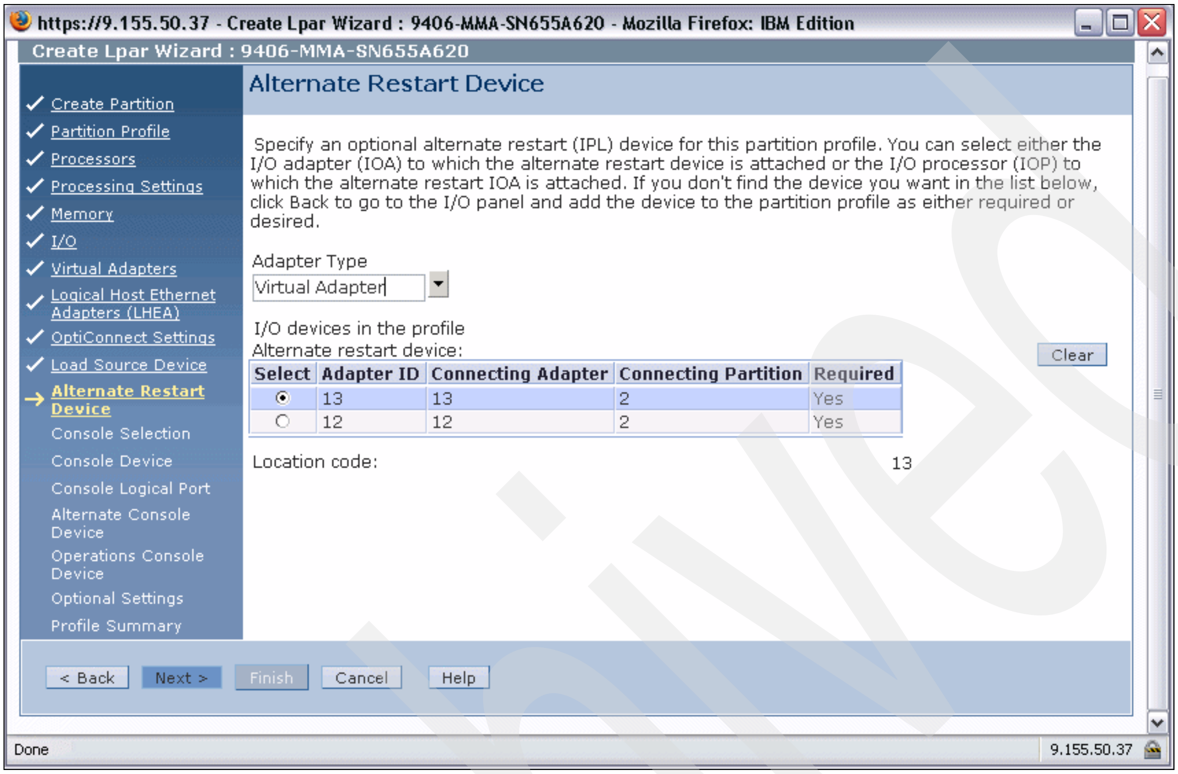


Figure 6-38 Create LPAR Alternate Restart Device dialog

19. We accept the default selection of **HMC** to be used as the console device and click **Next** to proceed, as shown in Figure 6-39.

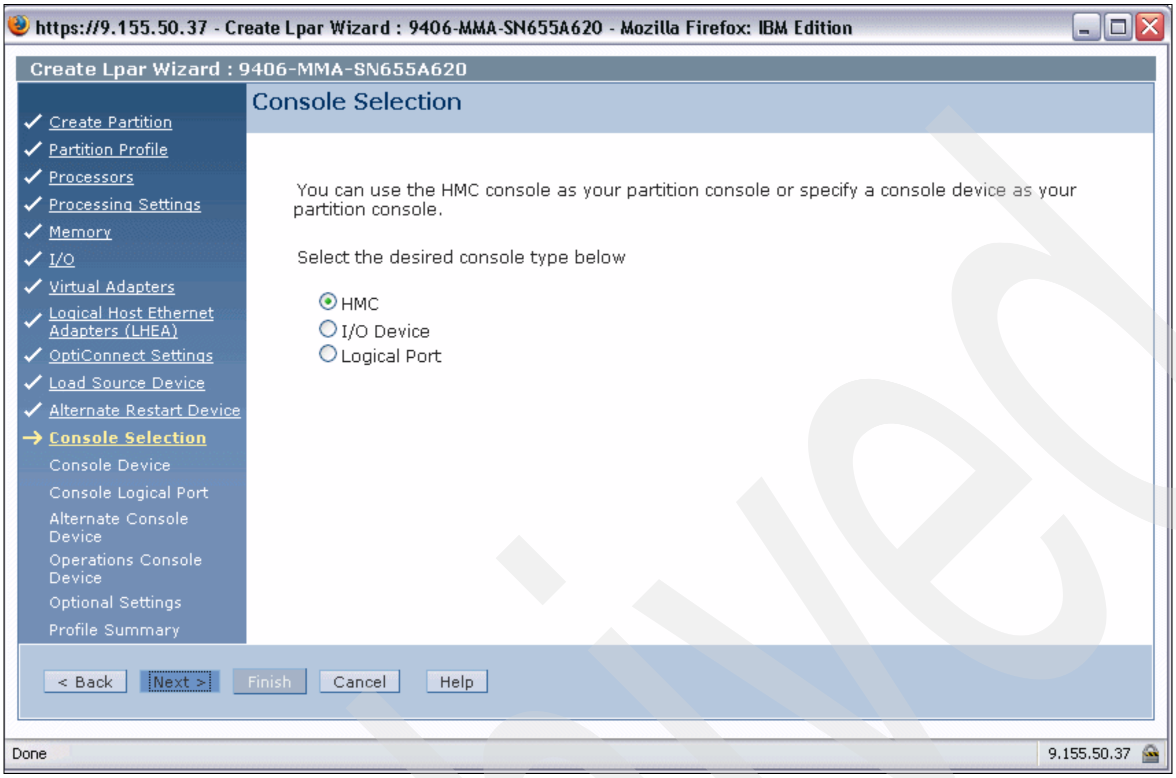


Figure 6-39 Create LPAR Console Selection dialog

20. We neither like to configure an Alternate Console Device nor an Operation Console Device, nor do we like to specify any optional settings, so we click **Next** again on the following three panels to skip these configurations until we reach the Profile Summary panel (as shown in Figure 6-40) where we double check our configuration before selecting **Finish** to finally create our IBM i client LPAR.

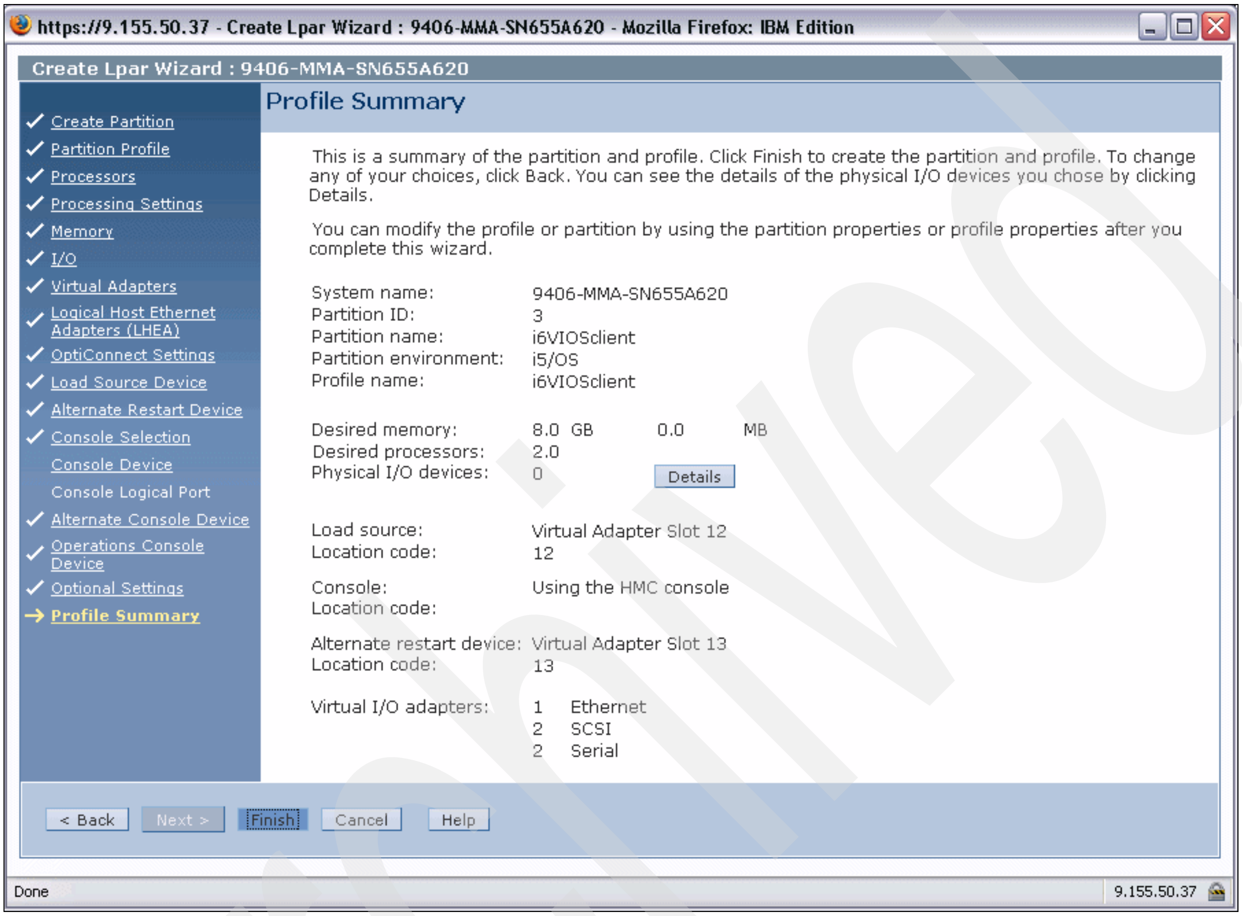


Figure 6-40 Create LPAR Profile Summary dialog

6.2.3 Installing VIOS

The procedures that we describe in this section show the following steps that we used to install and configure the IBM Virtual I/O Server (VIOS) on the POWER Systems POWER6 model 570 partition that we created earlier in 6.2.1, “Creating the VIOS LPAR” on page 137:

- “Preparing the IBM i internal Disk Drives for VIOS installation” on page 174
- “Performing the VIOS software installation” on page 175
- “Using LVM mirroring for VIOS” on page 188
- “Upgrading VIOS to the latest fixpack” on page 189

Preparing the IBM i internal Disk Drives for VIOS installation

Since our IBM POWER Systems POWER6 model 570 internal disk drives that we want to use for VIOS installation were used by a no-longer-existing IBM i partition before, that is, they got low-level formatted to 520 bytes per sector, we must reformat them to 512 byte/sector for usage with VIOS, as we describe below:

1. Insert the standalone AIX diagnostics CD-ROM into the DVD drive of the VIOS partition.
2. Activate the VIOS partition in SMS boot mode and select to open a terminal window.
3. After the POST indicator word *keyboard* displays press 5 before the word *speaker* displays and choose **VT320** for the terminal type.
4. From the FUNCTION SELECTION panel select option **3. Task Selection (Diagnostics, Advanced Diagnostics, Service Aids, and so on)**.
5. Select **RAID Array Manager** → **IBM SAS Disk Array Manager** → **Change/Show SAS pdisk Status** → **Delete an Array Candidate pdisk and Format to 512 Byte Sectors**.

   From an AIX perspective this converts pdisks to hdisks.
6. Since we want to use RAID adapter hardware RAID0 (striping) for creating two RAID0 arrays out of our CEC's six internal SAS drives for later LVM mirroring of VIOS we:
   a. Choose **IBM SAS Disk Array Manager** → **Create an Array Candidate pdisk and Format to 528 Byte Sectors**
   b. Create a SAS disk array.

      We repeat this step for creating a second RAID0 set to which we can later mirror.

Figure 6-41 shows the final RAID array configuration with two RAID0 arrays of our six internal SAS drives that we will use for VIOS installation.

```
                                COMMAND STATUS

Command: OK            stdout: yes          stderr: no

Before command completion, additional instructions may appear below.

[TOP]
------------------------------------------------------------------------
Name      Resource  State      Description             Size
------------------------------------------------------------------------
sissas0   FFFFFFFF  Available  PCI-X266 Planar 3Gb SAS Adapter

hdisk0    00FF0000  Optimal    RAID 0 Array          209.3GB
 pdisk0   00000200  Active     Array Member           69.7GB
 pdisk1   00000300  Active     Array Member           69.7GB
 pdisk2   00000400  Active     Array Member           69.7GB

hdisk1    00FF0100  Optimal    RAID 0 Array          209.3GB
 pdisk3   00000500  Active     Array Member           69.7GB
 pdisk4   00000600  Active     Array Member           69.7GB
 pdisk5   00000700  Active     Array Member           69.7GB


[BOTTOM]

F1=Help             F2=Refresh          F3=Cancel           Esc+6=Command
Esc+8=Image         Esc+9=Shell         Esc+0=Exit          /=Find
n=Find Next
```

Figure 6-41 RAID array configuration for VIOS installation

Performing the VIOS software installation

We used the following generic steps for installing the IBM Virtual I/O Server software from the DVD installation media:

1. Activate the VIOS partition in SMS boot mode.
2. Use System Management Services (SMS) to set the install/boot device.
3. Perform the VIOS software installation.

The procedural steps we used are:

1. Insert the VIOS installation DVD into the VIOS partition's DVD drive.
2. Activate the VIOS partition by selecting its server from the **Systems Management** → **Servers** view, selecting the VIOS partition, and choosing **Operations** → **Activate** from the pop-up menu, as shown in Figure 6-42.

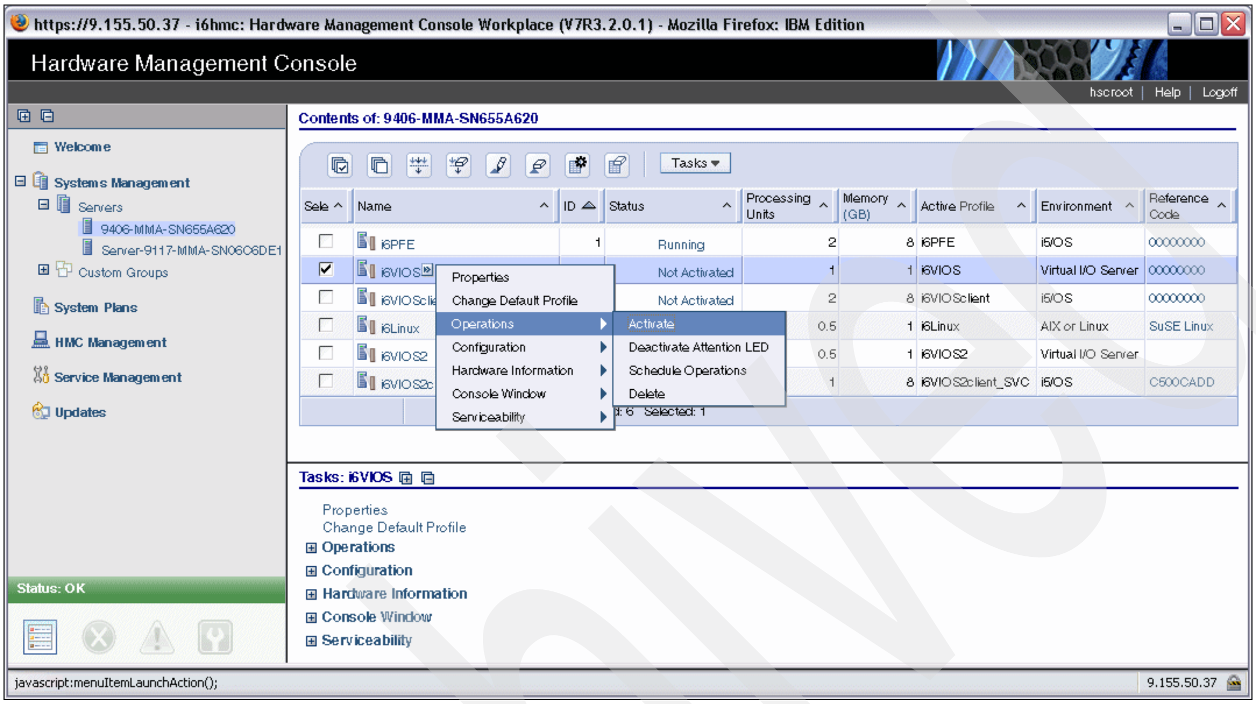


Figure 6-42 VIOS partition activation

3. Select the partition profile, mark **Open a terminal window or console session**, and click **Advanced**, as shown in Figure 6-43.

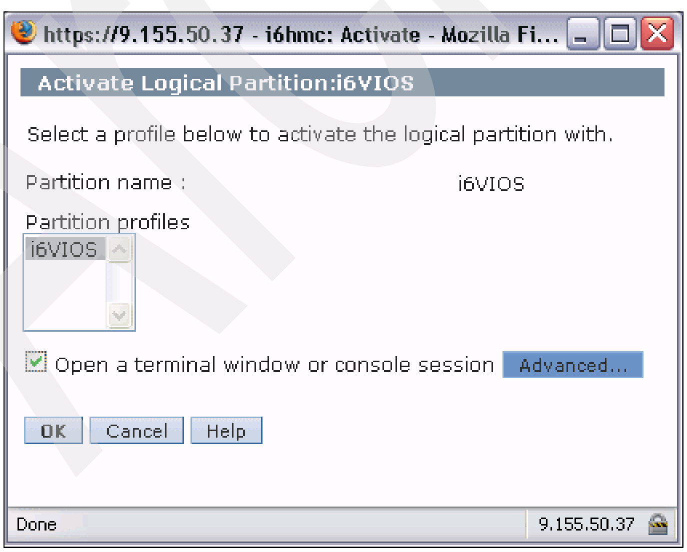


Figure 6-43 Activate Logical Partition dialog

4. Choose boot mode **SMS** and click **OK** to, as shown in Figure 6-44.

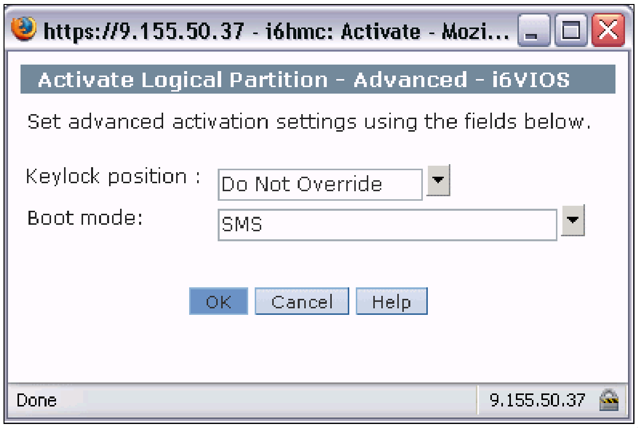


Figure 6-44 Activate Logical Partition: Advanced dialog

5. Select **OK** (as shown in Figure 6-45) to proceed to activate the VIOS partition in SMS boot mode while having a terminal session window opened.

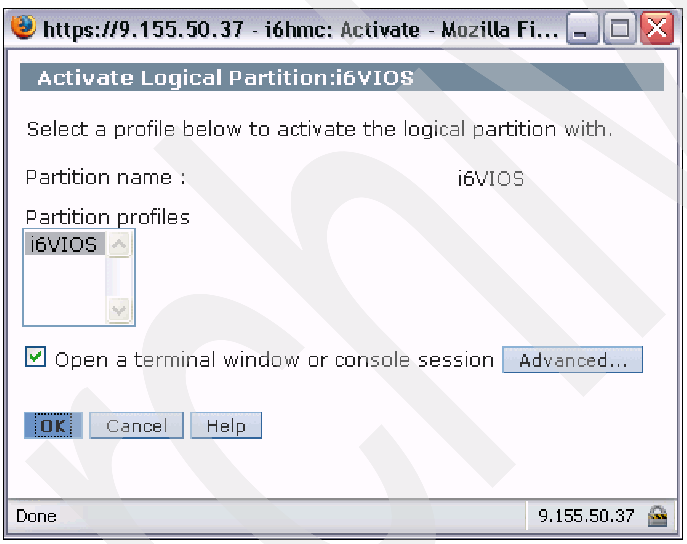


Figure 6-45 Activate Logical Partition dialog

6. Select option **5. Select Boot Options** from the console window's SMS main menu, as shown in Figure 6-46.

```
 PowerPC Firmware
 Version EM320_031
 SMS 1.7 (c) Copyright IBM Corp. 2000,2007 All rights reserved.
-------------------------------------------------------------------------------
 Main Menu
 1.   Select Language
 2.   Setup Remote IPL (Initial Program Load)
 3.   Change SCSI Settings
 4.   Select Console
 5.   Select Boot Options









 -------------------------------------------------------------------------------
 Navigation Keys:

                                            X = eXit System Management Services
 -------------------------------------------------------------------------------
 Type menu item number and press Enter or select Navigation key:5
```

Figure 6-46 SMS main menu

7. Select option **1. Select Install/Boot Device** from the Multiboot menu, as shown in Figure 6-47.

```
PowerPC Firmware
Version EM320_031
SMS 1.7 (c) Copyright IBM Corp. 2000,2007 All rights reserved.
-------------------------------------------------------------------------------
Multiboot
1.   Select Install/Boot Device
2.   Configure Boot Device Order
3.   Multiboot Startup <OFF>









-------------------------------------------------------------------------------
Navigation keys:
M = return to Main Menu
ESC key = return to previous screen         X = eXit System Management Services
-------------------------------------------------------------------------------
Type menu item number and press Enter or select Navigation key:1
```

Figure 6-47 SMS Multiboot menu

8. Selection option **3. CD/DVD** from the Select Device Type menu, as shown in Figure 6-48.

```
PowerPC Firmware
Version EM320_031
SMS 1.7 (c) Copyright IBM Corp. 2000,2007 All rights reserved.
-------------------------------------------------------------------------------
Select Device Type
1.   Diskette
2.   Tape
3.   CD/DVD
4.   IDE
5.   Hard Drive
6.   Network
7.   List all Devices








-------------------------------------------------------------------------------
Navigation keys:
M = return to Main Menu
ESC key = return to previous screen         X = eXit System Management Services
-------------------------------------------------------------------------------
Type menu item number and press Enter or select Navigation key:3
```

Figure 6-48 SMS Select Device Type menu

9. Select option **5. SATA** from the Select Media Type menu, as shown in Figure 6-49.

```
PowerPC Firmware
Version EM320_031
SMS 1.7 (c) Copyright IBM Corp. 2000,2007 All rights reserved.
-------------------------------------------------------------------------------
Select Media Type
1.   SCSI
2.   SSA
3.   SAN
4.   SAS
5.   SATA
6.   USB
7.   IDE
8.   ISA
9.   List All Devices






-------------------------------------------------------------------------------
Navigation keys:
M = return to Main Menu
ESC key = return to previous screen         X = eXit System Management Services
-------------------------------------------------------------------------------
Type menu item number and press Enter or select Navigation key:5
```

Figure 6-49 SMS Select Media Type menu

10. Select the media adapter from the Select Media Adapter menu, as shown in Figure 6-50.

```
PowerPC Firmware
Version EM320_031
SMS 1.7 (c) Copyright IBM Corp. 2000,2007 All rights reserved.
-------------------------------------------------------------------------------
Select Media Adapter
1.          /pci@800000020000200/pci@2/pci1014,02BD@1/sata
2.   List all devices











-------------------------------------------------------------------------------
Navigation keys:
M = return to Main Menu
ESC key = return to previous screen        X = eXit System Management Services
-------------------------------------------------------------------------------
Type menu item number and press Enter or select Navigation key:1
```

Figure 6-50 SMS Select Media Adapter menu

11. Select the device from the Select Device menu, as shown in Figure 6-51.

```
PowerPC Firmware
Version EM320_031
SMS 1.7 (c) Copyright IBM Corp. 2000,2007 All rights reserved.
-------------------------------------------------------------------------------
Select Device
Device  Current  Device
Number  Position  Name
1.       1      SATA CD-ROM
                ( loc=U789D.001.DQDZPVH-P4-D1 )










-------------------------------------------------------------------------------
Navigation keys:
M = return to Main Menu
ESC key = return to previous screen        X = eXit System Management Services
-------------------------------------------------------------------------------
Type menu item number and press Enter or select Navigation key:1
```

Figure 6-51 SMS Select Device menu

12. Select option **2. Normal boot mode** from the Select Task menu, as shown in Figure 6-52.

```
 PowerPC Firmware
 Version EM320_031
 SMS 1.7 (c) Copyright IBM Corp. 2000,2007 All rights reserved.
-------------------------------------------------------------------------------
 Select Task

SATA CD-ROM
    ( loc=U789D.001.DQDZPVH-P4-D1 )

 1.   Information
 2.   Normal Mode Boot
 3.   Service Mode Boot









 -------------------------------------------------------------------------------
 Navigation keys:
 M = return to Main Menu
 ESC key = return to previous screen         X = eXit System Management Services
 -------------------------------------------------------------------------------
 Type menu item number and press Enter or select Navigation key:2
```

Figure 6-52 SMS Select Task menu

13.Select option **1. Yes** for the “Are you sure you want to exit System Management Services?” question, as shown in Figure 6-53.

```
PowerPC Firmware
Version EM320_031
SMS 1.7 (c) Copyright IBM Corp. 2000,2007 All rights reserved.
-------------------------------------------------------------------------------
Are you sure you want to exit System Management Services?
1.   Yes
2.   No













-------------------------------------------------------------------------------
Navigation Keys:

                                            X = eXit System Management Services
-------------------------------------------------------------------------------
Type menu item number and press Enter or select Navigation key:1
```

Figure 6-53 SMS exit question

14. After automatic reboot of the VIOS partition in normal boot mode the VIOS installation panel comes up. At the prompt for the system console enter 1 to use this terminal as the system console, as shown in Figure 6-54.

```
-------------------------------------------------------------------------------
                       Welcome to the Virtual I/O Server.
                       boot image timestamp: 22:15 10/08
                 The current time and date: 18:18:51 07/24/2008
               number of processors: 1    size of memory: 1024MB
boot device: /pci@800000020000200/pci@2/pci1014,02BD@1/sata/disk@40000:\ppc\chrp
\bootfile.exe
                     kernel size: 12155218; 32 bit kernel
-------------------------------------------------------------------------------



 ******* Please define the System Console. *******

Type a 1 and press Enter to use this terminal as the
  system console.
Pour definir ce terminal comme console systeme, appuyez
  sur 1 puis sur Entree.
Taste 1 und anschliessend die Eingabetaste druecken, um
  diese Datenstation als Systemkonsole zu verwenden.
Premere il tasto 1 ed Invio per usare questo terminal
  come console.
Escriba 1 y pulse Intro para utilizar esta terminal como
  consola del sistema.
Escriviu 1 1 i premeu Intro per utilitzar aquest
  terminal com a consola del sistema.
Digite um 1 e pressione Enter para utilizar este terminal
  como console do sistema.



1
```

Figure 6-54 VIOS Define System Console prompt

15. At the install language prompt enter 1 to use English as the install language, as shown in Figure 6-55.

```
>>>  1 Type 1 and press Enter to have English during install.
     2 Entreu 2 i premeu Intro per veure la instal? laci?  en catal
     3 Entrez 3 pour effectuer l'installation en fran? ais.
     4 F? r Installation in deutscher Sprache 4 eingeben
        und die Eingabetaste dr? cken.
     5 Immettere 5 e premere Invio per l'installazione in Italiano.
     6 Digite 6 e pressione Enter para usar Portugu? s na instala? ? o.
     7 Escriba 7 y pulse Intro para la instalaci? n en espa? ol.




    88  Help ?


>>> Choice [1]: 1
```

Figure 6-55   VIOS install language prompt

16. Select option **1 Start Install Now with Default Settings** from the VIOS welcome panel, as shown in Figure 6-56.

```
 Welcome to Base Operating System
                      Installation and Maintenance

Type the number of your choice and press Enter.  Choice is indicated by >>>.

>>> 1 Start Install Now with Default Settings

    2 Change/Show Installation Settings and Install

    3 Start Maintenance Mode for System Recovery

    4 Configure Network Disks (iSCSI)




    88  Help ?
    99  Previous Menu

>>> Choice [1]: 1
```

Figure 6-56   VIOS welcome panel

17.Select option **0 Install with the settings listed above** from the VIOS System Backup Installation and Settings menu, as shown in Figure 6-57, to proceed with installing VIOS on hdisk0.

```
 System Backup Installation and Settings

Either type 0 and press Enter to install with the current settings, or type the
number of the setting you want to change and press Enter.

      Setting:                                 Current Choice(s):

    1 Disk(s) where you want to install ...... hdisk0
         Use Maps............................. No
    2 Shrink File Systems..................... No
    3 Import User Volume Groups............... No
    4 Recover Devices......................... No

>>> 0 Install with the settings listed above.




                       +-----------------------------------------------------
    88  Help ?         |    WARNING: Base Operating System Installation will
    99  Previous Menu  |    destroy or impair recovery of ALL data on the
                       |    destination disk hdisk0.
>>> Choice [0]: 0
```

Figure 6-57 VIOS System Backup Installation and Settings menu

18. Installation of VIOS now starts with its progress being shown on the Installing Base Operation System panel, as shown in Figure 6-58.

```
Installing Base Operating System



       Please wait...










       Approximate     Elapsed time
    % tasks complete   (in minutes)


         50               20      58% of mksysb data restored.
```

Figure 6-58 VIOS installation progress panel

19. After the successful installation of VIOS login with the prime administrator user ID padmin, enter a new password and enter `a` to accept the software terms and conditions, as shown in Figure 6-59.

```
IBM Virtual I/O Server

                          login: padmin
[compat]: 3004-610 You are required to change your password.
        Please choose a new one.

padmin's New password:
Enter the new password again:

Indicate by selecting the appropriate response below whether you
accept or decline the software maintenance terms and conditions.
[Accept (a)] |  Decline (d) |  View Terms (v)
a
$
```

Figure 6-59 VIOS login prompt

20. Before we can run any VIOS command other than **chlang** to change the language setting we must accept the SW license terms by entering the command:

```
license -accept
```

To view the license before accepting it, enter the command:

```
license -view
```

Using LVM mirroring for VIOS

In this section we describe how we used LVM mirroring to mirror the VIOS root volume group (rootvg) across our two RAID0 arrays of hdisk0 and hdisk1 to help protect VIOS from potential CEC internal SAS disk drive failures:

1. Enter the command **`smit extendvg`** to extent the VIOS **`rootvg`** from current hdisk0 to hdisk1.
2. Enter the command **`smit lvm`**, select **Volume Groups** → **Mirror a Volume Group** to mirror the VIOS **`rootvg`** volume group, selecting **Mirror sync mode = background** to start mirroring in the background choosing **PHYSICAL VOLUME names = hdisk0 hdisk1**.
3. Enter the command **`lsvg -l rootvg`** to verify that there are two physical partitions (PPs) allocated per logical partition for each logical volume of the VIOS rootvg, as shown in Example 6-1.

Example 6-1 Mirrored logical volumes of VIOS

```
# lsvg -l rootvg
rootvg:
LV NAME             TYPE       LPs     PPs     PVs  LV STATE      MOUNTPOINT
hd5                 boot       1       2       2    closed/syncd  N/A
hd6                 paging     2       4       2    open/syncd    N/A
paging00            paging     4       8       2    open/syncd    N/A
hd8                 jfs2log    1       2       2    open/syncd    N/A
hd4                 jfs2       1       2       2    open/syncd    /
hd2                 jfs2       10      20      2    open/syncd    /usr
hd9var              jfs2       3       6       2    open/syncd    /var
hd3                 jfs2       9       18      2    open/syncd    /tmp
hd1                 jfs2       40      80      2    open/syncd    /home
hd10opt             jfs2       4       8       2    open/syncd    /opt
lg_dumplv           sysdump    4       4       1    open/syncd    N/A
```

4. Enter the command **`bosboot -ad /dev/hdisk0`** to write the new boot image to hisk0.
5. Enter the command **`bootlist -m normal -o`** to list the current AIX bootlist.
6. Enter the command **`bootlist -m normal hdisk0 hdisk1`** to add hdisk1 to the AIX bootlist.
7. Enter the command **`varyonvg hdisk0`** to make the changes effective.

Configuring VIOS network connectivity

We used the following procedure to configure TCP/IP network connectivity on our VIOS partition:

1. Being connected to VIOS with our HMC terminal window and logged in as padmin, we enter the command **`lsdev -type adapter | grep ent`** (as shown in Example 6-2). We look for the *logical* host Ethernet adapter resources. In our example it is ent1, as ent0 is the VLAN adapter.

Example 6-2 Displaying the Ethernet adapter resources

```
$ lsdev -type adapter | grep ent
ent0             Available   Virtual I/O Ethernet Adapter (l-lan)
ent1             Available   Logical Host Ethernet Port (lp-hea)
```

2. We configure TCP/IP for the logical Ethernet adapter ent*X* using the **mktcpip** command syntax and specifying the corresponding *interface* resource en*X*, as shown in Example 6-3.

Example 6-3 mktcpip command syntax

```
$ mktcpip -hostname hostname -inetaddr host_IP_address -interface en1 -netmask
network_mask -gateway gateway_IP_address -nsrvaddr DNS_IP_address -nsrvdomain
DNS_domain_name -start
```

3. To verify a proper connection we ping our TCP/IP network gateway IP address 9.155.50.1 using the **ping** command and specifying 5 echo requests, as shown in Example 6-4.

Example 6-4 Pinging the network gateway

```
$ ping 9.155.50.1 5
PING 9.155.50.1: (9.155.50.1): 56 data bytes
64 bytes from 9.155.50.1: icmp_seq=0 ttl=255 time=0 ms
64 bytes from 9.155.50.1: icmp_seq=1 ttl=255 time=0 ms
64 bytes from 9.155.50.1: icmp_seq=2 ttl=255 time=0 ms
64 bytes from 9.155.50.1: icmp_seq=3 ttl=255 time=0 ms
64 bytes from 9.155.50.1: icmp_seq=4 ttl=255 time=0 ms

----9.155.50.1 PING Statistics----
5 packets transmitted, 5 packets received, 0% packet loss
round-trip min/avg/max = 0/0/0 ms
```

Upgrading VIOS to the latest fixpack

To do this:

1. We check the VIOS level by entering the **ioslevel** command, as shown in Example 6-5.

Example 6-5 Displaying the VIOS level

```
$ ioslevel
1.5.0.0
```

2. With this VIOS 1.5.0.0 installation from our distribution media we do not meet the minimum VIOS level requirement for IBM i client attachment so we upgrade VIOS to the latest available fixpack, as described in 9.2, “VIOS maintenance” on page 346, which is recommended.
3. After installation of the latest fixpack we run the **ioslevel** command again to verify our VIOS level, as shown in Example 6-6.

Example 6-6 Displaying the VIOS level after its upgrade

```
$ ioslevel
1.5.2.1-FP-11.1
```

After successful VIOS installation and upgrade to the latest fixpack we are now ready to proceed with attaching our IBM external Midrange Storage to VIOS as described in 6.3, “Attaching Midrange Storage to VIOS” on page 190.

Note: The VIOS online help can be accessed by entering the command **help** to get a list of the available VIOS commands. The command **man *command*** displays the online command reference incl. syntax information for the specified VIOS command.

For further information about IBM Virtual I/O Server commands refer to *Power Systems Virtual I/O Server and Integrated Virtualization Manager Command Reference*, SA76-0101, available at:

http://publib.boulder.ibm.com/infocenter/systems/topic/iphdx/sa76-0101.pdf

6.3 Attaching Midrange Storage to VIOS

For attaching the IBM System Storage DS4800 external Storage System to our VIOS partition with two PCIe 4 Gb dual-port Fibre Channel adapters #5774 we used a direct-connection attaching one port from the first VIOS FC adapter to DS4800 controller A and one port from the second VIOS FC adapter to DS4800 controller B, as shown in Figure 6-60.

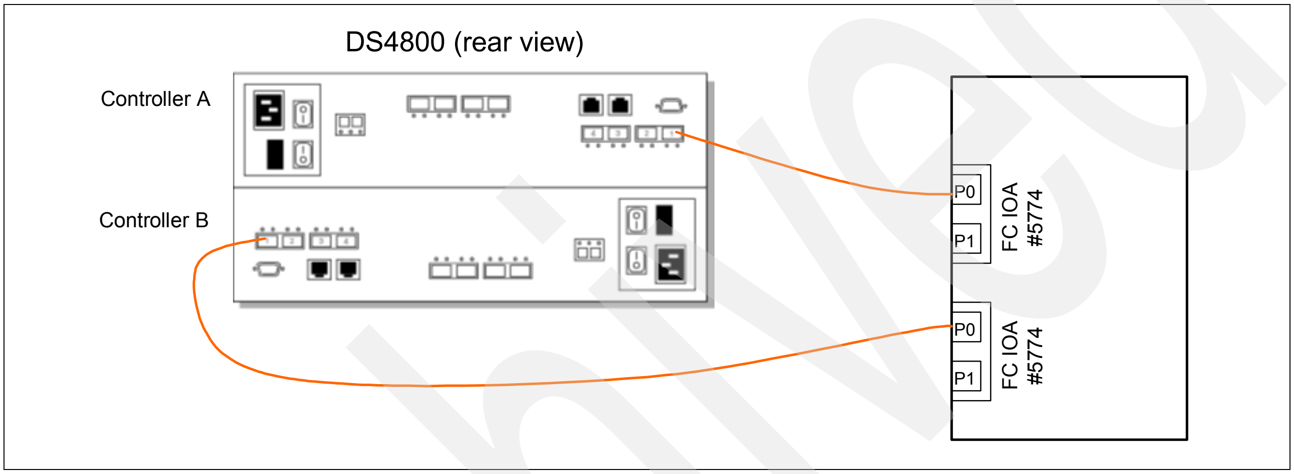


Figure 6-60 VIOS external storage attachment

There is no possibility of changing the protocol being used by the DS4800 host port since the DS4800 automatically adapts to the FC protocol. It will be provided by the SAN or host port and it will either be FC arbitrated-loop (FC-AL) or switched-fabric. VIOS uses the switched-fabric protocol when connected to a SAN switch and FC-AL when directly connected to the DS4800 Storage System.

Note: In contrast to IBM i, if VIOS is being attached to a SAN fabric or storage, a manual action of running the configuration manager (**`cfgdev`** command) is required to trigger the VIOS-owned Fibre Channel adapter initiators to log into the fabric or storage device and discover its SCSI devices.

VIOS has both the DS4000 series RDAC multi-path device driver and the MPIO multi-path device driver for DS3000 and DS5000 series included. For our IBM i DS4800 implementation example with VIOS 1.5 we simply use the default RDAC driver. A VIOS 2.1 *scratch* installation based on AIX 6.1 kernel uses MPIO as the default device driver also for DS4000 series attachment. For migrating an existing VIOS DS4000 series attachment from RDAC to MPIO refer to “Migrating from RDAC to MPIO for DS4000 series attachment” on page 351.

6.4 Configuring storage for IBM i

We used the following steps for performing the logical storage configuration on the IBM System Storage DS4800 storage subsystem for attaching the IBM Virtual I/O Server and defining storage capacity for the IBM i client:

1. Created a hot spare drive on the DS4800: See 7.4.1, “Creating hot spares” on page 252.
2. Created a RAID-10 array with logical drives on the DS4800: See 7.4.2, “Creating RAID arrays and logical drives” on page 257.
3. Created the VIOS host definition on the DS4800: See 7.4.3, “Creating hosts” on page 265.
4. Mapped the logical drives to the VIOS host on the DS4800: See 7.4.4, “Defining logical drive to LUN mapping” on page 274.

6.5 Configuring VIOS virtual devices

After completing the DS4800 logical storage configuration we perform the following steps, making the DS4800 logical volumes accessible as virtualized SCSI LUNs by our IBM i client:

1. We connect to our VIOS partition via a secure shell (SSH) connection over TCP/IP port 22 from our PuTTY client, as shown in Figure 6-61.

> **Note:** PuTTY is a free Telnet and SSH client for Windows and UNIX® operating systems available for download at:
>
> http://www.chiark.greenend.org.uk/%7Esgtatham/putty/

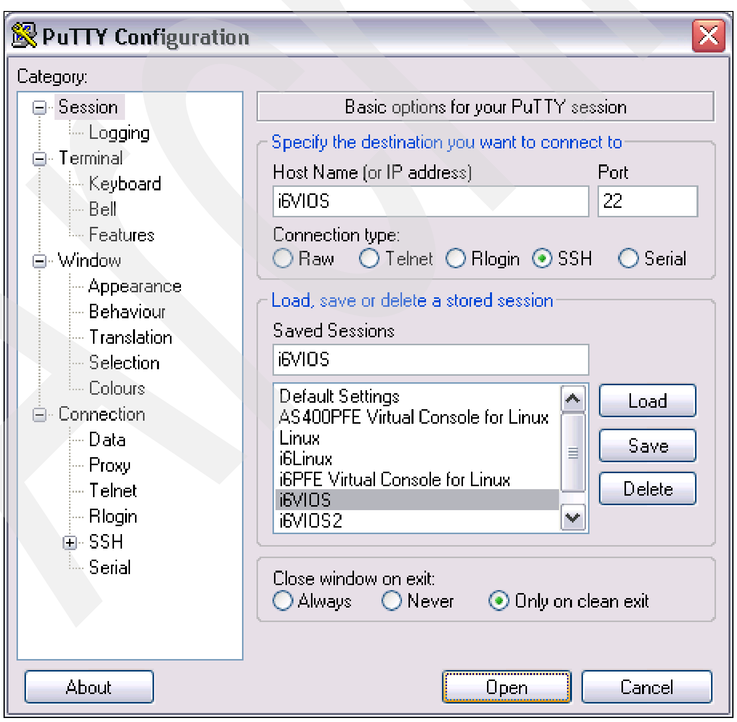


Figure 6-61 Using PuTTY to connect to VIOS via SSH

2. We trust our VIOS host's RSA key and add it to the Windows registry to avoid being asked for it again at the next connection by selecting **Yes** in the PuTTY Security alert dialog, as shown in Figure 6-62.

Figure 6-62 PuTTY Security Alert dialog

3. After logging in as padmin we look at our VIOS disk configuration by entering the **`lsdev -type disk`** command, as shown in Example 6-7.

Example 6-7 Displaying the VIOS disk devices using the lsdev command

```
$ lsdev -type disk
name            status      description
hdisk0          Available   SAS RAID 0 Disk Array
hdisk1          Available   SAS RAID 0 Disk Array
$
```

4. So far VIOS only sees its two internal disk devices hdisk0 and hdisk1, which represent the two RAID0 arrays that we installed VIOS on. To have VIOS scan for newly attached disk devices under our Fibre Channel controllers fcs0 and fcs2 we need to use the VIOS **`cfgdev`** command before our DS4800 devices are detected by VIOS, as shown in Example 6-8.

Example 6-8 Displaying the VIOS disk devices after running cfgmgr

```
$ cfgdev -dev fcs0
$ cfgdev -dev fcs2
$ lsdev -type disk
name            status      description
hdisk0          Available   SAS RAID 0 Disk Array
hdisk1          Available   SAS RAID 0 Disk Array
hdisk2          Available   1815      DS4800 Disk Array Device
hdisk3          Available   1815      DS4800 Disk Array Device
hdisk4          Available   1815      DS4800 Disk Array Device
hdisk5          Available   1815      DS4800 Disk Array Device
hdisk6          Available   1815      DS4800 Disk Array Device
hdisk7          Available   1815      DS4800 Disk Array Device
hdisk8          Available   1815      DS4800 Disk Array Device
$
```

5. We now look at the virtual device mapping on our VIOS partition by entering the **lsmap -all** command, as shown in Example 6-9.

Example 6-9 Displaying the VIOS virtual device mapping using the lsmap command

```
$ lsmap -all
SVSA            Physloc                                      Client Partition ID
--------------- -------------------------------------------- ------------------
vhost0          U9406.MMA.655A620-V2-C12                     0x00000003

VTD                   NO VIRTUAL TARGET DEVICE FOUND

SVSA            Physloc                                      Client Partition ID
--------------- -------------------------------------------- ------------------
vhost1          U9406.MMA.655A620-V2-C13                     0x00000003

VTD                   NO VIRTUAL TARGET DEVICE FOUND

$
```

The **lsmap** output shows that on our newly installed VIOS server we currently do not have any virtual target device (VTD) mapped to our two virtual SCSI server adapters vhost0 and vhost1.

We now remember that we tagged the IBM i client's virtual SCSI client adapter in slot 12, which we mapped to our VIOS virtual SCSI server adapter in slot 12 (see 6.2.2, "Creating the IBM i Client LPAR" on page 155) as the load source I/O adapter, so we know that we have to map the DS4800 LUN that we are going to use as a load source to the virtual SCSI server adapter in location C12, which is vhost0.

6. To figure out which VIOS hdisk corresponds to which DS4800 LUN we use the **fget_config** command from the VIOS root command line, as shown in Example 6-10.

Example 6-10 Displaying the VIOS DS4800 device mapping using the fget_config command

```
$ oem_setup_env
# fget_config -Av

---dar0---

User array name = 'iSeries_VIOS'
dac0 ACTIVE dac1 ACTIVE

Disk     DAC   LUN Logical Drive
hdisk2   dac1    0 i6VIOSclient_0
hdisk3   dac0    1 i6VIOSclient_1
hdisk4   dac1    2 i6VIOSclient_2
hdisk5   dac0    3 i6VIOSclient_3
hdisk6   dac1    4 i6VIOSclient_4
hdisk7   dac0    5 i6VIOSclient_5
hdisk8   dac1    6 i6VIOSclient_6
# exit
$
```

Note: If using MPIO as the VIOS multi-path I/O driver use the **mpio_get_config -Av** command instead.

7. We like to use LUN 0 as the load source, and since in our case we only created seven LUNs on the DS4800 Storage System for our IBM i client we can map all of them (up to

16) to the same virtual SCSI server adapter that we tagged as the load source IOA, which is vhost0, using the **mkvdev** command, as shown in Example 6-11.

Example 6-11 Mapping the VIOS partition's hdisks to the IBM i client

```
$ i=2
$ while [ $i -lt 9 ]
> do
> mkvdev -vdev hdisk$i -vadapter vhost0
> i=$((i+1))
> done
vtscsi0 Available
vtscsi1 Available
vtscsi2 Available
vtscsi3 Available
vtscsi4 Available
vtscsi5 Available
vtscsi6 Available
$
```

Now we have created for our seven DS4800 LUNs seven corresponding virtual SCSI target devices, vtscsi0 to vtscsi6, which are available to our IBM i client.

8. The remaining step before we are able to activate our IBM i client for installing system licensed internal code (SLIC) and the IBM i 6.1 operating system is to use the **mkvdev** command again for mapping the VIOS partition's DVD-RAM drive as a virtual SCSI optical device to our IBM i client. For this we remember that we tagged our IBM i client's alternate restart device to the virtual SCSI client adapter in slot 13, which connects to our VIOS virtual SCSI server adapter in slot 13, so we know that we have to map our VIOS partition's DVD-RAM drive cd0 to vhost1 (that is, the adapter in slot C13, as shown in Example 6-12).

Example 6-12 Mapping the VIOS partition's optical drive to the IBM i client

```
$ lsdev -type optical
name             status      description
cd0              Available   SATA DVD-RAM Drive
$ mkvdev -vdev cd0 -vadapter vhost1 -dev vcd
vcd Available
$
```

9. Our final DS4800 and DVD-RAM virtual device mapping on VIOS for our IBM i client partition is shown in Example 6-13.

> **Note:** The physical location information in the **lsmap** output shows the SCSI LUN ID *X* in its L*X*000000000000 suffix.

Example 6-13 VIOS device mapping for the IBM i client

```
$ lsmap -all
SVSA            Physloc                                      Client Partition ID
--------------- -------------------------------------------- ------------------
vhost0          U9406.MMA.655A620-V2-C12                     0x00000003

VTD                   vtscsi0
Status                Available
LUN                   0x8100000000000000
Backing device        hdisk2
Physloc               U789D.001.DQDWVYP-P1-C3-T1-W201500A0B811F4C0-L0
```

```
VTD                 vtscsi1
Status              Available
LUN                 0x8200000000000000
Backing device      hdisk3
Physloc             U789D.001.DQDWVYP-P1-C3-T1-W201500A0B811F4C0-L1000000000000

VTD                 vtscsi2
Status              Available
LUN                 0x8300000000000000
Backing device      hdisk4
Physloc             U789D.001.DQDWVYP-P1-C3-T1-W201500A0B811F4C0-L2000000000000

VTD                 vtscsi3
Status              Available
LUN                 0x8400000000000000
Backing device      hdisk5
Physloc             U789D.001.DQDWVYP-P1-C3-T1-W201500A0B811F4C0-L3000000000000

VTD                 vtscsi4
Status              Available
LUN                 0x8500000000000000
Backing device      hdisk6
Physloc             U789D.001.DQDWVYP-P1-C3-T1-W201500A0B811F4C0-L4000000000000

VTD                 vtscsi5
Status              Available
LUN                 0x8600000000000000
Backing device      hdisk7
Physloc             U789D.001.DQDWVYP-P1-C3-T1-W201500A0B811F4C0-L5000000000000

VTD                 vtscsi6
Status              Available
LUN                 0x8700000000000000
Backing device      hdisk8
Physloc             U789D.001.DQDWVYP-P1-C3-T1-W201500A0B811F4C0-L6000000000000

SVSA            Physloc                                      Client Partition ID
--------------- -------------------------------------------- ------------------
vhost1          U9406.MMA.655A620-V2-C13                     0x00000003

VTD                 vcd
Status              Available
LUN                 0x8100000000000000
Backing device      cd0
Physloc             U789D.001.DQDWXNY-P4-D1

$
```

10. We enable the RDAC autorecovery to automatically redistribute the LUNs across both DS4000 series controllers after a controller path failure using the **chdev** command from the VIOS root-command line, as shown in Example 6-14.

Example 6-14 Enabling RDAC autorecovery

```
$ lsdev -dev dar0 -attr
act_controller dac0,dac1 Active Controllers                        False
aen_freq       600       Polled AEN frequency in seconds           True
all_controller dac0,dac1 Available Controllers                     False
autorecovery   no        Autorecover after failure is corrected    True
balance_freq   600       Dynamic Load Balancing frequency in seconds True
```

```
cache_size      2048      Cache size for both controllers           False
fast_write_ok   yes       Fast Write available                      False
held_in_reset   none      Held-in-reset controller                  True
hlthchk_freq    600       Health check frequency in seconds         True
load_balancing  no        Dynamic Load Balancing                    True
switch_retries  5         Number of times to retry failed switches  True

$ chdev -dev dar0 -attr autorecovery=yes
dar0 changed

$ lsdev -dev dar0 -attr autorecovery
value

yes
$
```

Note: We generally recommend enabling the RDAC or MPIO autorecovery for automatic redistribution of the LUNs across both DS storage controllers after recovery of a controller path failure. However, keep in mind that a VIOS Fibre Channel HBA hot-plug replacement is only supported with autorecovery disabled (see 9.2, “VIOS maintenance” on page 346).

Now we are ready to proceed with activating our IBM i client for a SLIC install and adding the DS4800 LUNs to its configuration, as described in 6.6, “Adding DS Storage to IBM i configuration” on page 197.

6.6 Adding DS Storage to IBM i configuration

After we completed the mapping of the DS4800 LUNs and the VIOS-owned DVD-RAM as virtual devices for our IBM i client we start the IBM i software installation on our IBM i client partition of VIOS and add the virtual LUNs to the configuration, as described in the following steps:

1. We insert the IBM i 6.1 I_BASE CD containing the System Licensed Internal Code (SLIC) into the VIOS partition's DVD-RAM drive.
2. We use the HMC to activate our newly created IBM i client partition i6VIOSclient by marking this partition from the Systems Management → Servers view and selecting **Operations** → **Activate** from the pop-up menu, as shown in Figure 6-63.

> **Note:** We do not need to change the partition's properties for setting a D-mode manual IPL for SLIC install because this is already the default for a newly created IBM i partition.

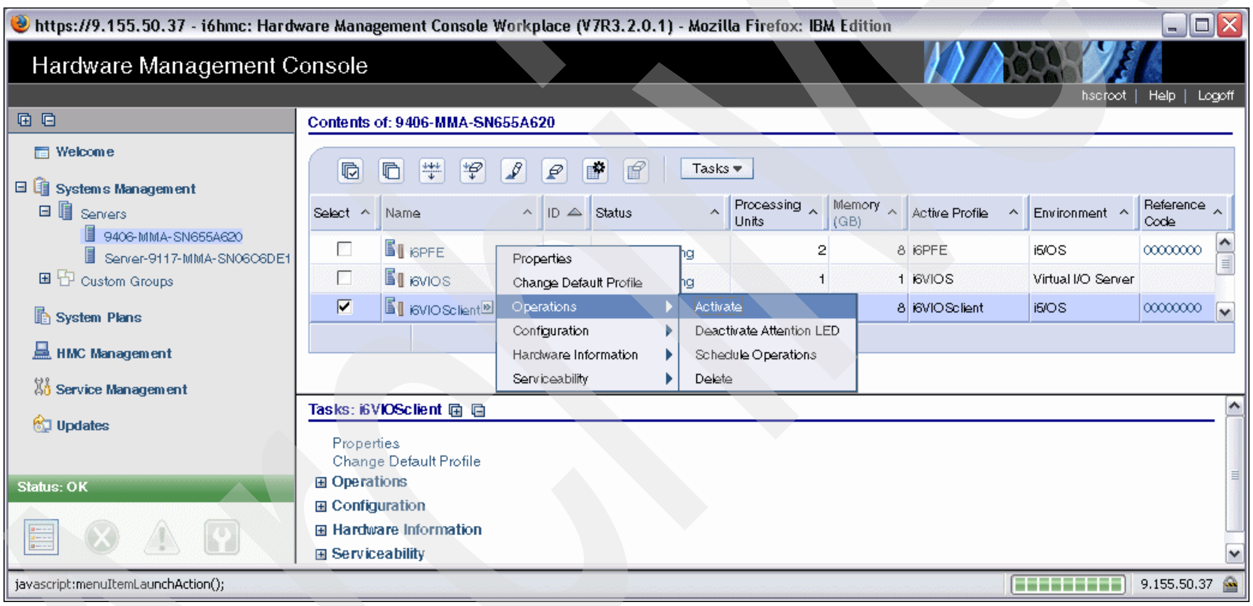


Figure 6-63 HMC IBM i client partition activation

3. We select **OK** on the partition activation dialog, as shown in Figure 6-64.

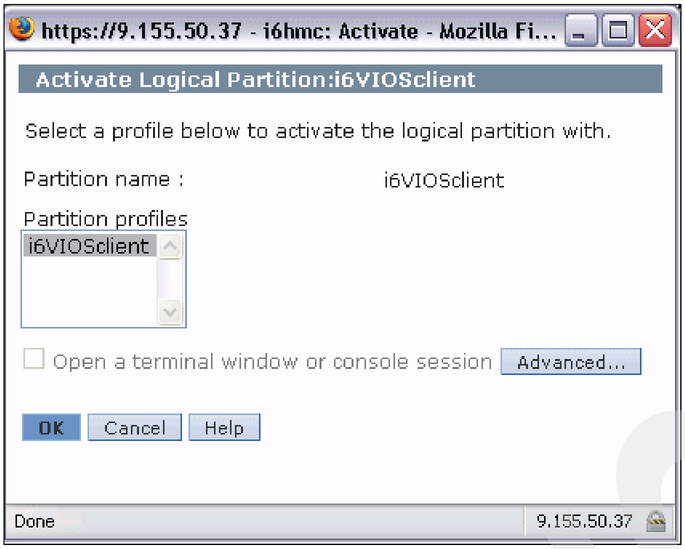


Figure 6-64 HMC Activate Logical Partition dialog

4. After the IBM i partition has booted from the I_BASE CD-ROM we enter the language feature code `2924` (English), as shown in Figure 6-65.

```
     II           //      000000         SSSSS
                 //     00      00      SS     SS
                //     00        00    SS
               //     00          00     SS
     II       //     00            00       SS
     II      //      00            00         SS
     II     //        00          00           SS
     II    //          00        00            SS
     II   //            00      00      SS     SS
     II  //               000000          SSSSS



 LANGUAGE FEATURE ===>   2924
```

Figure 6-65 IBM i language feature panel

5. We press Enter to confirm our language feature choice, as shown in Figure 6-66.

```
Confirm Language Group


  Language feature  . . . . . . . . . . . . . . :   2924

  Press Enter to confirm your choice for language feature.
  Press F12 to change your choice for language feature.
```

Figure 6-66 IBM i language feature confirmation panel

6. We select option **1. Install Licensed Internal Code**, as shown in Figure 6-67.

```
Install Licensed Internal Code
                                                            System:   C605A620
Select one of the following:

     1. Install Licensed Internal Code
     2. Work with Dedicated Service Tools (DST)
     3. Define alternate installation device








Selection
     1

Licensed Internal Code - Property of IBM 5761-999 Licensed
Internal Code (c) Copyright IBM Corp.  1980, 2007.  All
rights reserved.  US Government Users Restricted Rights -
Use duplication or disclosure restricted by GSA ADP schedule
Contract with IBM Corp.
```

Figure 6-67   IBM i Install Licensed Internal Code panel

7. We use option 1 to select the device with Ctl=1 as our load source device corresponding to LUN 0, matching LUN 0x81 on VIOS, which is vtscsi0 backed by hdisk2, and press Enter to proceed, as shown in Figure 6-68.

> **Notes:** Virtual LUNs always show up as device type 6B22 model 050 on the IBM i client, no matter from which storage subsystem they are.
>
> The controller information “Ctl” XOR 0x80 corresponds to the vtscsi*X* “LUN” information (see Example 6-13 on page 194).
>
> The system card information Sys Card shows the IBM i virtual SCSI client adapter ID as defined in its partition configuration on the HMC.
>
> Make sure that the selected load source volume has large enough capacity (>= 17 GB) to be used as a load source for IBM i 6.1. You can check the capacity of the corresponding hdisk under VIOS root command line using the command **`lsattr -El hdiskX -a size`** but should consider that due to the 8-to-9 sector conversion (see 1.1, “What is new” on page 2) the shown and usable capacity by the IBM i client is only 8/9 of the capacity available under VIOS.

```
Select Load Source Device

Type 1 to select, press Enter.

                                       Sys  Sys   I/O        I/O
Opt  Serial Number   Type  Model    Bus  Card  Adapter   Bus   Ctl   Dev
     YB97LL8VVYKP    6B22   050     255    12        0     0     7     0
     YQMYF57E28VZ    6B22   050     255    12        0     0     3     0
     YHJCEWV5ZJUT    6B22   050     255    12        0     0     5     0
 1   YKKRPYR7RRGP    6B22   050     255    12        0     0     1     0
     YWDRYMVEPM7F    6B22   050     255    12        0     0     6     0
     YVSGM9U6H55B    6B22   050     255    12        0     0     2     0
     YQ6J5E7MV8WH    6B22   050     255    12        0     0     4     0




F3=Exit          F5=Refresh             F12=Cancel
```

Figure 6-68 IBM i Select Load Source Device panel

8. We choose F10=Continue to confirm our load source device selection and continue, as shown in Figure 6-69.

```
Confirm Load Source Device

You have chosen the following device to be the load source.
Press F10 to confirm your choice.

                                   Sys  Sys   I/O       I/O
     Serial Number    Type  Model  Bus  Card  Adapter   Bus   Ctl   Dev
     YKKRPYR7RRGP     6B22   050   255   12       0       0     1     0









F3=Exit           F10=Continue           F12=Cancel
```

Figure 6-69 IBM i Confirm Load Source Device panel

9. We select option **2. Install Licensed Internal Code and Initialize system** to proceed with installing the SLIC on our new IBM i client partition of VIOS, as shown in Figure 6-70.

```
Install Licensed Internal Code (LIC)

Disk selected to write the Licensed Internal Code to:
         Serial Number   Type   Model   I/O Bus   Controller   Device
       YKKRPYR7RRGP      6B22    050        0          1          0

Select one of the following:

     1. Restore Licensed Internal Code
     2. Install Licensed Internal Code and Initialize system
     3. Install Licensed Internal Code and Recover Configuration
     4. Install Licensed Internal Code and Restore Disk Unit Data
     5. Install Licensed Internal Code and Upgrade Load Source






Selection
     2

F3=Exit           F12=Cancel
```

Figure 6-70 IBM i Install Licensed Internal Code (LIC) panel

10. We press F10=Continue to confirm our intention to proceed with the SLIC installation, as shown in Figure 6-71.

```
Install LIC and Initialize System - Confirmation

Warning:
    All data on this system will be destroyed and the Licensed
    Internal Code will be written to the selected disk if you
    choose to continue the initialize and install.

    Return to the install selection screen and choose one of the
    other options if you want to perform some type of recovery
    after the install of the Licensed Internal Code is complete.



    Press F10 to continue the install.
    Press F12 (Cancel) to return to the previous screen.
    Press F3 (Exit) to return to the install selection screen.






F3=Exit            F10=Continue            F12=Cancel
```

Figure 6-71 IBM i Install LIC and Initialize System - Confirmation panel

11. We wait until the load source initialization (formatting) is complete, as shown in Figure 6-72. The actual time that it takes for the initialization is typically much less than the estimated time.

```
Initialize the Disk - Status

The load source disk is being initialized.


Estimated time to initialize in minutes :       180

Elapsed time in minutes . . . . . . . . :       9.5









Please wait.


Wait for next display or press F16 for DST main menu
```

Figure 6-72 IBM i Initialize the Disk - Status panel

12. After load source initialization we wait for the System Licensed Internal Code installation to complete, as shown in Figure 6-73.

```
Install Licensed Internal Code - Status

Install of the Licensed Internal Code in progress.

            +--------------------------------------------------+
Percent     !                       55%                        !
complete    +--------------------------------------------------+


Elapsed time in minutes  . . . . . . . . :       1.5









Please wait.


Wait for next display or press F16 for DST main menu
```

Figure 6-73 Install Licensed Internal Code - Status panel

13. The Disk Configuration Attention Report informs about a new disk configuration because unconfigured disk units were found in the configuration. We accept by pressing F10=Accept the problems and continue, as shown in Figure 6-74.

```
Disk Configuration Attention Report

Type option, press Enter.
  5=Display Detailed Report

Press F10 to accept all the following problems and continue.
The system will attempt to correct them.

Opt  Problem
     New disk configuration











F3=Exit   F10=Accept the problems and continue   F12=Cancel
```

Figure 6-74 IBM i Disk Configuration Attention Report panel

14. After the SLIC install has completed, the IPL or Install the System panel appears. We insert the IBM i 6.1 B2924_01 DVD into the VIOS partition's DVD-RAM drive and select option **2. Install the operating system** to proceed with our IBM i 6.1 client installation, as shown in Figure 6-75.

```
IPL or Install the System
                                                          System:   C605A620
Select one of the following:

     1. Perform an IPL
     2. Install the operating system
     3. Use Dedicated Service Tools (DST)
     4. Perform automatic installation of the operating system
     5. Save Licensed Internal Code





Selection
     2

Licensed Internal Code - Property of IBM 5761-999 Licensed
Internal Code (c) Copyright IBM Corp.  1980, 2007.  All
rights reserved.  US Government Users Restricted Rights -
Use duplication or disclosure restricted by GSA ADP schedule
Contract with IBM Corp.
```

Figure 6-75 IBM i IPL or Install the System panel

15. We select **2. Optical** on the Install Device Type Selection panel to proceed with installing the IBM i 6.1 operating system from the VIOS partition's virtualized DVD-RAM, drive as shown in Figure 6-76.

```
Install Device Type Selection
                                                          System:   C605A620
Select the installation device type:

     1. Tape
     2. Optical
     3. Virtual device - preselected image catalog
     4. Current alternate selected device           None
     5. Network device









Selection
     2

F3=Exit   F12=Cancel
```

Figure 6-76 Install Device Type Selection panel

16.We press Enter to confirm the installation of the operating system, as shown in Figure 6-77.

```
Confirm Install of the Operating System
                                                           System:   C605A620
Press Enter to confirm your choice to install the operating
  system.

Press F12 to return and cancel your choice to install the
  operating system.















F12=Cancel
```

Figure 6-77 IBM i Confirm Install of the Operating System panel

17.We confirm the language feature selection by pressing Enter again, as shown in Figure 6-78.

```
Confirm Language Feature Selection
                                                           System:   C605A620
Language feature  . . . . . . . . . . . . . . :   2924

Press Enter to confirm your choice for language feature.
  Installing the system will continue.

Press F12 to return to change your choice for
  language feature.














F3=Exit   F12=Cancel
```

Figure 6-78 IBM i Confirm Language Feature Selection panel

18. We select option **2. Perform disk configuration using DST**, as shown in Figure 6-79, so that we can look at our disk configuration and select the virtual LUNs that we want to add to the IBM i auxiliary storage pool configuration.

> **Note:** Usually, it makes sense to add additional disk units to the system ASP (ASP1) before proceeding to help speed up the installation of the IBM i operating system by using more disk arms. We could have used option 3 to add all units to the system ASP, but for demonstration purposes showing how the virtual LUNs show up in the disk configuration of our IBM i client, we prefer using a manual detour via DST.

```
Add All Disk Units to the System
                                                          System:   C605A620
Select one of the following:

     1. Keep the current disk configuration
     2. Perform disk configuration using DST
     3. Add all disk units to the system auxiliary storage pool
     4. Add all disk units to the system ASP and balance data

Selection
     2

F3=Exit   F12=Cancel
```

Figure 6-79 IBM i Add All Disk Units to the System panel

19. In the Dedicated Service Tools (DST) we select the options **4. Work with disk units** → **1. Work with disk configuration** → **1. Display disk configuration** → **1. Display disk configuration status** and see our current IBM i client disk configuration with only the load source (unit 1) being configured in ASP1, as shown in Figure 6-80.

```
Display Disk Configuration Status

          Serial                     Resource
ASP Unit  Number          Type Model Name       Status
  1                                             Unprotected
       1  YKKRPYR7RRGP    6B22  050  DD001      Configured









Press Enter to continue.

F3=Exit        F5=Refresh         F9=Display disk unit details
F11=Disk configuration capacity   F12=Cancel
```

Figure 6-80 IBM i DST Display Disk Configuration Status panel

20. We press F12=Cancel until we are back on the DST Work with Disk Configuration panel where we select option **3. Work with ASP configuration** → **3. Add units to ASPs** → **3. Add units to existing ASPs**. We are presented with the remaining six virtual LUNs that we created for our IBM i client that are still unconfigured and that we select via option 1 to be added to ASP1 *before* installing the IBM i 6.1 operating system.

Note: Virtual LUNs show up with a DPH*xxx* resource name on the IBM i client, as they are not readable until they get initialized, by which they become DD*xxx* devices.

```
Specify ASPs to Add Units to

Specify the existing ASP to add each unit to.

Specify  Serial                                  Resource
  ASP    Number          Type  Model  Capacity  Name
  1      YB97LL8VVYKP    6B22   050      33405  DPH005
  1      YQMYF57E28VZ    6B22   050      33405  DPH002
  1      YHJCEWV5ZJUT    6B22   050      33405  DPH006
  1      YWDRYMVEPM7F    6B22   050      33405  DPH004
  1      YVSGM9U6H55B    6B22   050      33405  DPH001
  1      YQ6J5E7MV8WH    6B22   050      33405  DPH003

F3=Exit     F5=Refresh     F11=Display disk configuration capacity
F12=Cancel
```

Figure 6-81   IBM i DST Specify ASPs to Add Units to panel

21. We press F10=Ignore problems and continue on the warning panel about the units possibly configured for POWER PC AS, as shown in Figure 6-82, which wants to alert us to not erroneously destroy data from disk units possibly used by another OS like AIX.

```
Problem Report

Note: Some action for the problems listed below may need to
be taken.  Please select a problem to display more detailed
information about the problem and to see what possible
action may be taken to correct the problem.

Type option, press Enter.
  5=Display Detailed Report

OPT  Problem
     Unit possibly configured for Power PC AS










F3=Exit      F10=Ignore problems and continue     F12=Cancel
```

Figure 6-82 IBM i DST Problem Report panel

22. We press Enter at the Confirm Add Units panel, as shown in Figure 6-83.

```
Confirm Add Units

Add will take several minutes for each unit.  The system will
have the displayed protection after the unit(s) are added.

Press Enter to confirm your choice for Add units.
Press F9=Capacity Information to display the resulting capacity.
Press F10=Confirm Add and Balance data on units.
Press F12=Cancel to return and change your choice.

          Serial                      Resource
ASP Unit  Number          Type Model  Name       Protection
  1                                              Unprotected
       1  YKKRPYR7RRGP    6B22  050   DD001      Unprotected
       2  YB97LL8VVYKP    6B22  050   DPH005     Unprotected
       3  YQMYF57E28VZ    6B22  050   DPH002     Unprotected
       4  YHJCEWV5ZJUT    6B22  050   DPH008     Unprotected
       5  YWDRYMVEPM7F    6B22  050   DPH004     Unprotected
       6  YVSGM9U6H55B    6B22  050   DPH001     Unprotected
       7  YQ6J5E7MV8WH    6B22  050   DPH003     Unprotected

F9=Resulting Capacity             F10=Add and Balance
F11=Display Encryption Status     F12=Cancel
```

Figure 6-83 IBM i DST Confirm Add Units panel

23. After all our disk units got added successfully to the system ASP, as shown in Figure 6-84, we return to the IPL or Install the System panel, selecting option 2. Install the operating system again to start overflowing the standard procedure for installing the IBM i operating system as described in the IBM Systems Information Center at:

http://publib.boulder.ibm.com/infocenter/systems/scope/i5os/index.jsp?topic=/rzahc/newnos.htm

```
Display Disk Configuration Status

          Serial                         Resource
ASP Unit  Number           Type Model  Name        Status
  1                                                Unprotected
       1  YKKRPYR7RRGP     6B22  050   DD001       Configured
       2  YB97LL8VVYKP     6B22  050   DD016       Configured
       3  YQMYF57E28VZ     6B22  050   DD017       Configured
       4  YHJCEWV5ZJUT     6B22  050   DD008       Configured
       5  YWDRYMVEPM7F     6B22  050   DD014       Configured
       6  YVSGM9U6H55B     6B22  050   DD013       Configured
       7  YQ6J5E7MV8WH     6B22  050   DD015       Configured








Press Enter to continue.

F3=Exit        F5=Refresh          F9=Display disk unit details
F11=Disk configuration capacity    F12=Cancel
```

Figure 6-84 IBM i DST Display Disk Configuration Status after units added to ASP

6.7 Migration to Midrange Storage

For migration of existing IBM i 6.1 systems to IBM i Midrange External Storage attached via VIOS there are basically two options:

- *System save/restore* where a system save of the existing IBM i 6.1 system is done to tape and a system restore from this tape to the new IBM i client of VIOS attached to IBM i Midrange External Storage.
- For existing IBM i 6.1 systems on POWER6 using either internal or other external storage a virtual SCSI client adapter can be added to connect to a VIOS partition providing virtualized IBM i Midrange External Storage LUNs, which can be added to the configuration with data being migrated via the STRASPBAL *ENDALC and *MOVDTA options and load source migration being done via the DST *Copy disk unit data* option.

7

Using the DS Storage Manager Client

In this chapter we show using IBM DS Storage Manager GUI for DS3000, DS4000, and DS5000 series as well as using the Storage Manager command-line interface (SMcli) for configuring storage for IBM i clients of Virtual I/O Server (VIOS).

7.1 Installation of DS Storage Manager Client

Note: At this point we assume that the IBM Midrange External Storage System has already been physically installed, connected to a TCP/IP network, and is powered on.

For further information about physical installation and startup refer to the following IBM System Storage quick-start guides:

- *IBM System Storage DS4800 Quick Start Guide*, GC27-2148-00, available at:

  ftp://ftp.software.ibm.com/systems/support/system_x_pdf/44e5421.pdf

- *IBM System Storage DS4200 and DS4700 Storage Subsystem Quick Start Guide*, GC27-2147-00, available at:

  ftp://ftp.software.ibm.com/systems/support/system_x_pdf/gc27214701.pdf

- *IBM System Storage DS3400 Storage Subsystem Installation, User's, and Maintenance Guide*, P/N 43W7847, available at:

  ftp://ftp.software.ibm.com/systems/support/system_x_pdf/44w1263.pdf

The following procedure describes the installation of the IBM System Storage DS Storage Manager client on a Microsoft® Windows operating system for out-band managing of one or more IBM Midrange External Storage Systems over a TCP/IP network.

For installation of the DS Storage Manager client on other supported management platforms like AIX, Linux, or Solaris refer to the *IBM System Storage DS4000 Storage Manager Installation and Host Support Guide*, GA76-0422, available at:

ftp://ftp.software.ibm.com/systems/support/storage_disk_pdf/ga76042200.pdf

Note: The IBM System Storage DS4000 Storage Manager client can be used for both managing the IBM System Storage DS3000 series and DS4000/DS5000 series Storage Systems and includes both the GUI and the Storage Manager command-line interface.

1. Download the IBM DS4000 Storage Manager client from the following IBM support Web site:

   http://www-304.ibm.com/systems/support/supportsite.wss/selectproduct?taskind=2&brandind=5000028&familyind=5329626&typeind=0&modelind=0&osind=0&psid=sr&continue.x=1&matrix=Y#Storage%20Manager

2. Start the DS Storage Manager client install wizard by running the SMIA-WS-xx.xx.xx.xx.exe executable to invoke the install wizard, as shown in Figure 7-1.

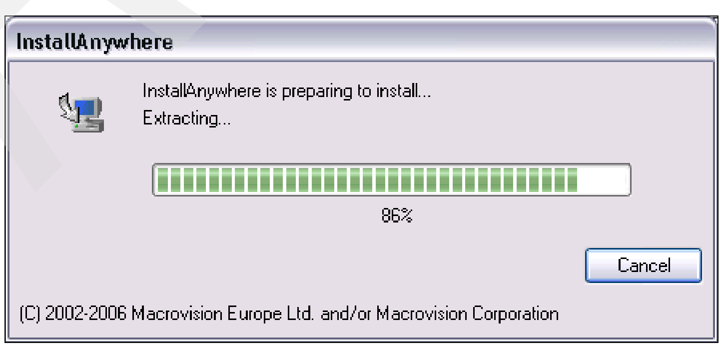


Figure 7-1 SM InstallAnywhere

3. Select the installation language of your choice from the drop-down list box and click **OK** to proceed, as shown in Figure 7-2.

Figure 7-2 SM Install Wizard welcome dialog

4. Select **Next** to continue with the installation on the DS Storage Manager Introduction dialog, as shown in Figure 7-3.

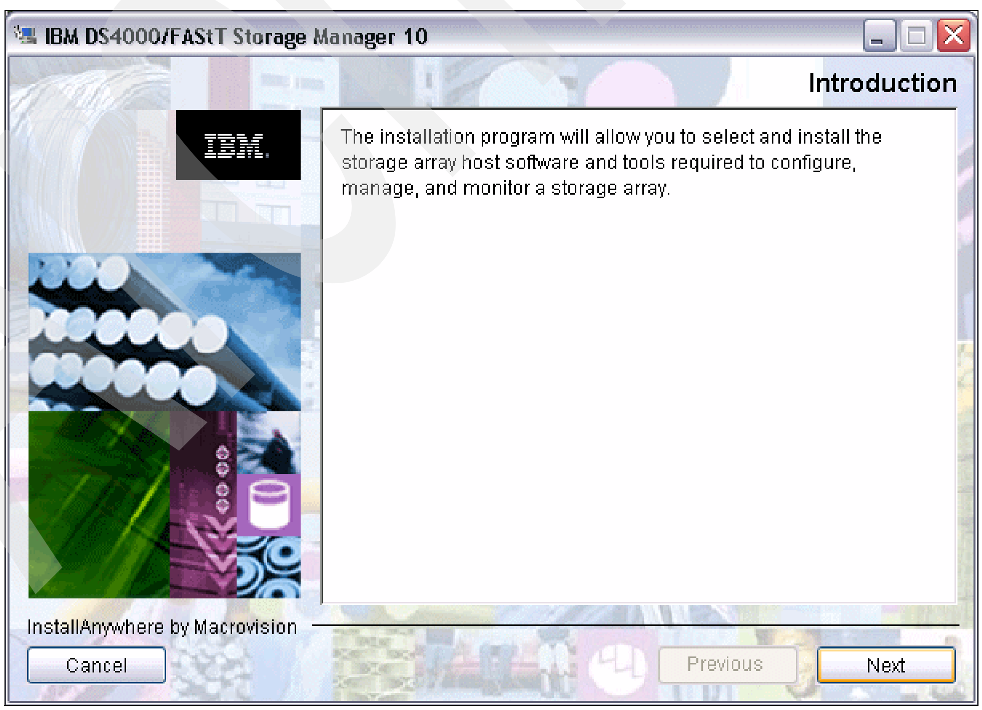


Figure 7-3 SM Introduction dialog

5. Select **Next** again on the Copyright Statement dialog, as shown in Figure 7-4.

Figure 7-4 SM Copyright Statement dialog

6. Select **I accept the terms of the License Agreement** on the License Agreement dialog and click **Next** to continue, as shown in Figure 7-5.

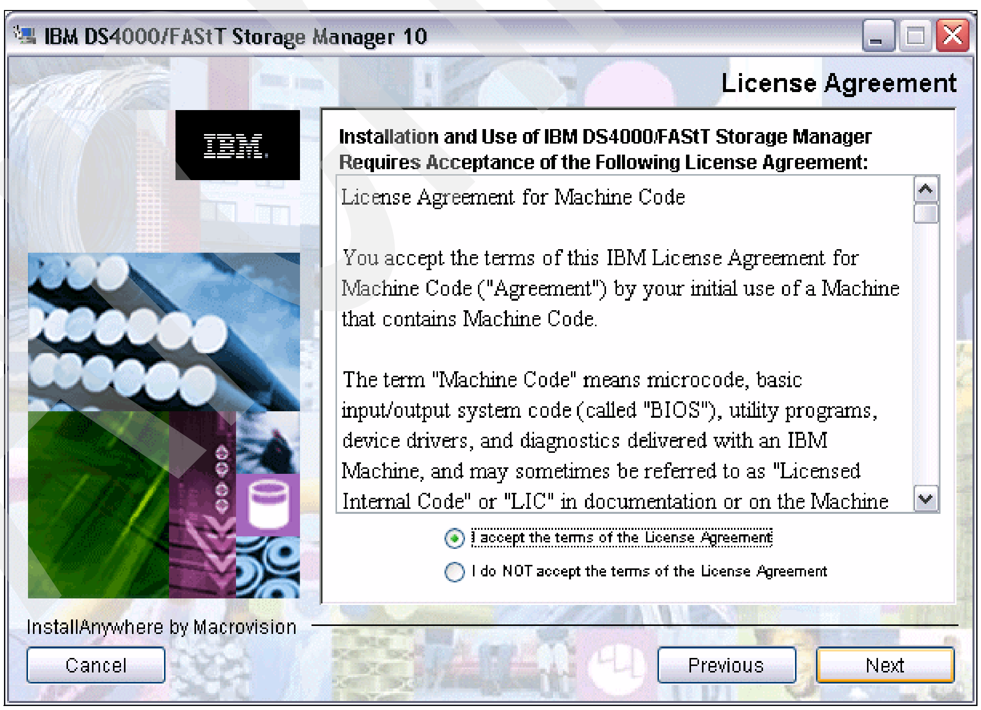


Figure 7-5 SM License Agreement dialog

7. Select **Management Station** on the Select Installation Type dialog and select **Next** to proceed, as shown in Figure 7-6.

> **Note:** Choosing Full installation would mean that also the software packages for host I/O connectivity to the DS storage subsystem like multi-path drivers would be installed, which we do not need on the Windows server. We only want to use this as a management station.

Figure 7-6 SM Select Installation Type dialog

8. We select **Do not Automatically Start the Monitor** and click **Next** on the Automatically Start Monitor dialog, as shown in Figure 7-7.

> **Note:** If this would be the only management system used for managing the IBM DS Storage System selecting to automatically start the event monitor service would be recommended to have SNMP/e-mail notifications sent out by the management station for potential Storage System problems (see also 2.7, “Remote Support Manager” on page 54, and 7.2, “Configuring DS remote support” on page 222).

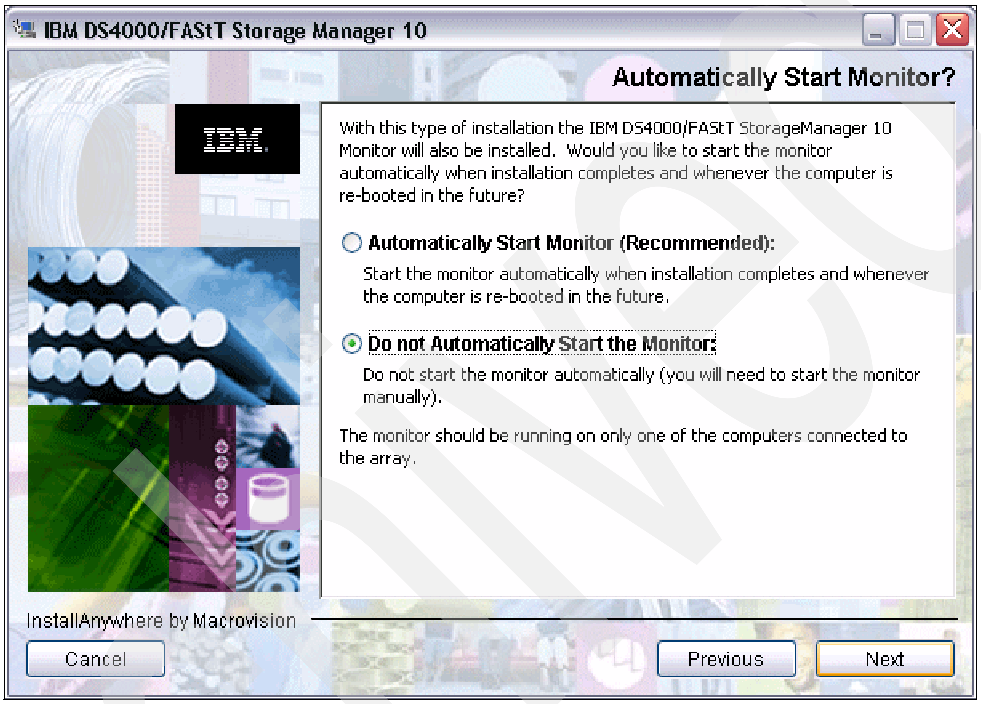


Figure 7-7 SM Automatically Start Monitor dialog

9. Review the information in the Pre-Installation Summary dialog and click **Install** to start the DS Storage Manager client installation, as shown in Figure 7-8.

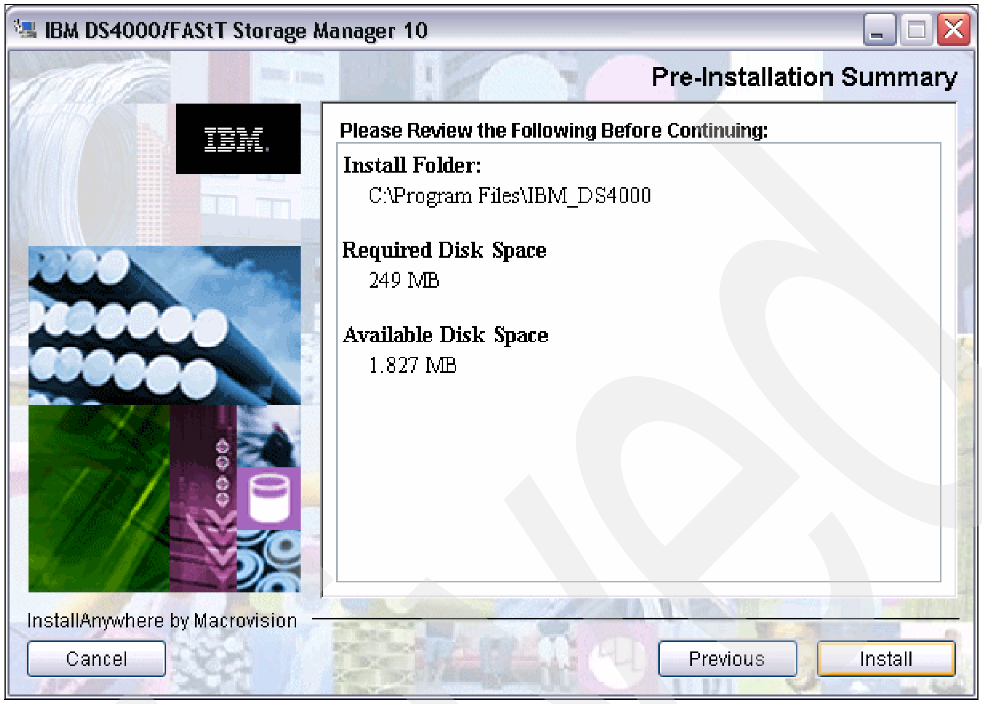


Figure 7-8 SM Pre-Installation Summary dialog

10. The install wizard shows the installation progress. After successful completion of the installation click **Done**, as shown in the Install Complete dialog in Figure 7-9.

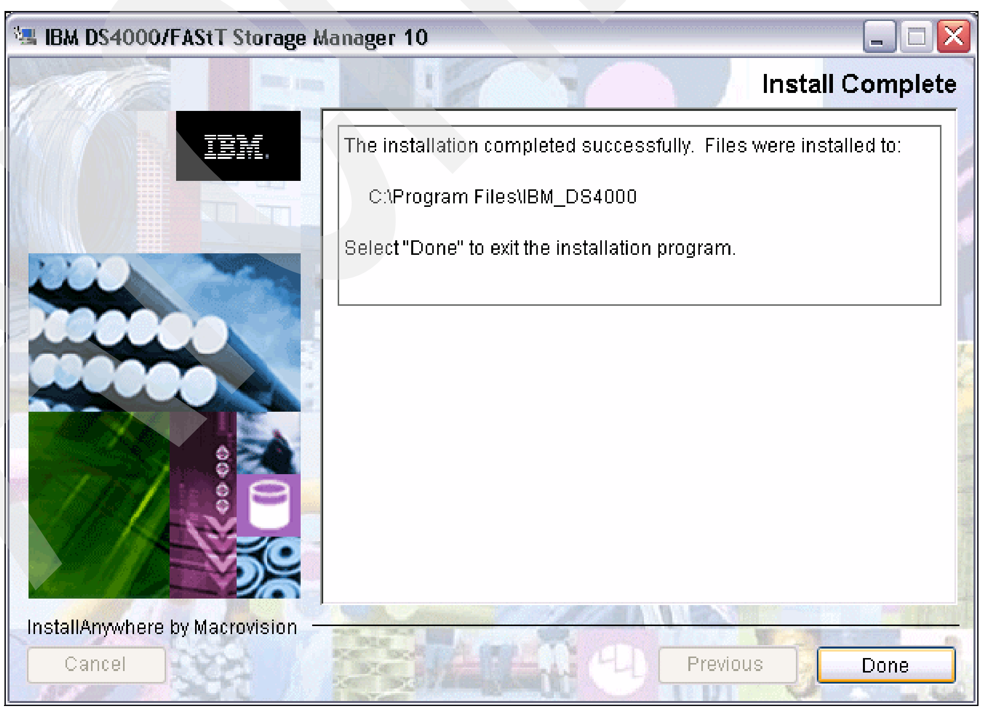


Figure 7-9 SM Install Complete dialog

11. Start the DS Storage Manager client from the Windows' start menu by selecting **Start** → **Programs** → **Storage Manager 10 Client** → **Storage Manager 10 Client**.

12. At its first start the IBM System Storage DS4000 Storage Manager presents the Initial Automatic Discovery dialog, as shown in Figure 7-10, where we select **No** since we prefer to add our DS4800 Storage System manually, as we know the IP addresses of both controllers.

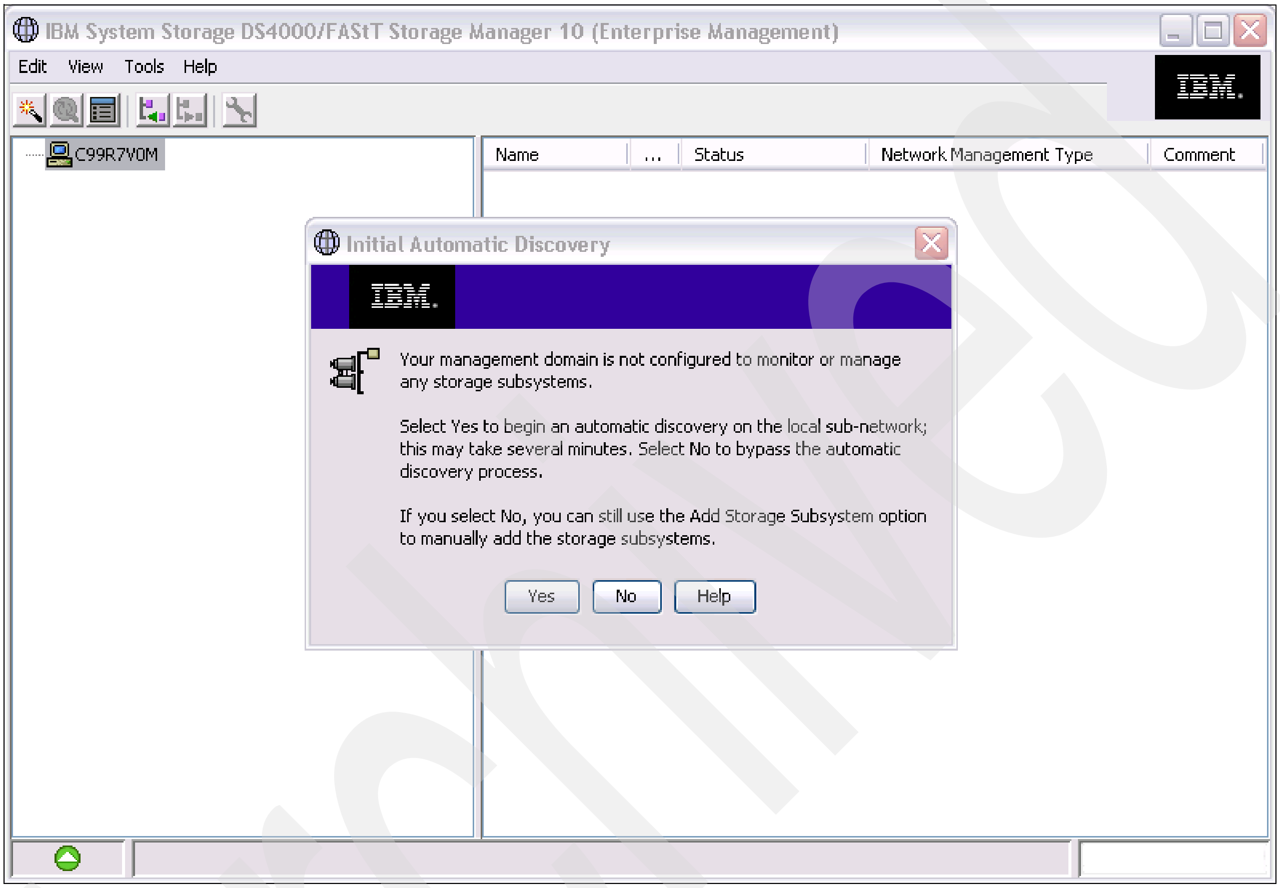


Figure 7-10 DS Storage Manager Initial Automatic Discovery dialog

13. To manually add our DS4800 Storage System to the management domain of our DS Storage Manager client we select from the menu **Edit** → **Add Storage Subsystem**, as shown in Figure 7-11.

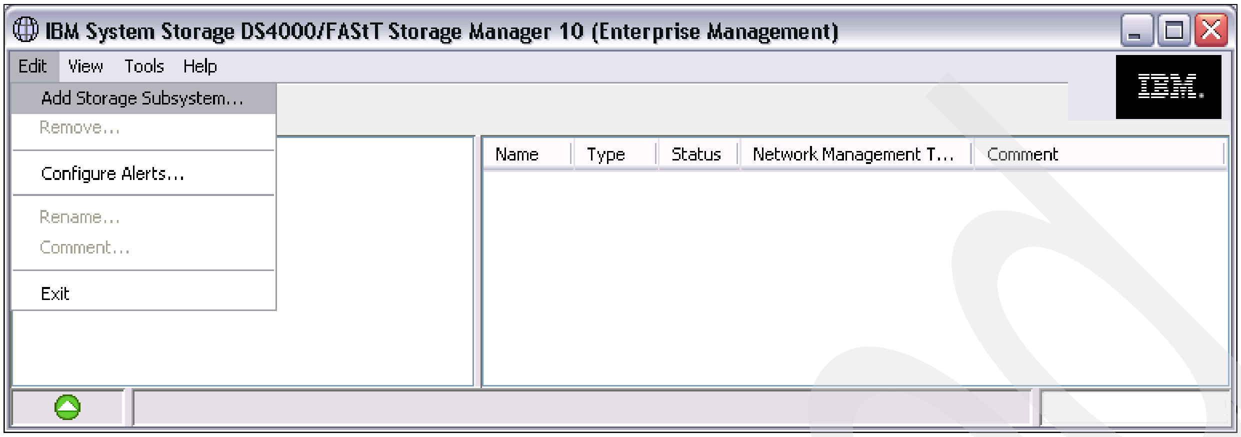


Figure 7-11 DS Storage Add Storage Subsystem menu

14. In the Add New Storage Subsystem dialog we select **Out-of-band management**, enter the IP addresses of our DS4800 storage subsystem controllers A and B, and click **Add**, as shown in Figure 7-12.

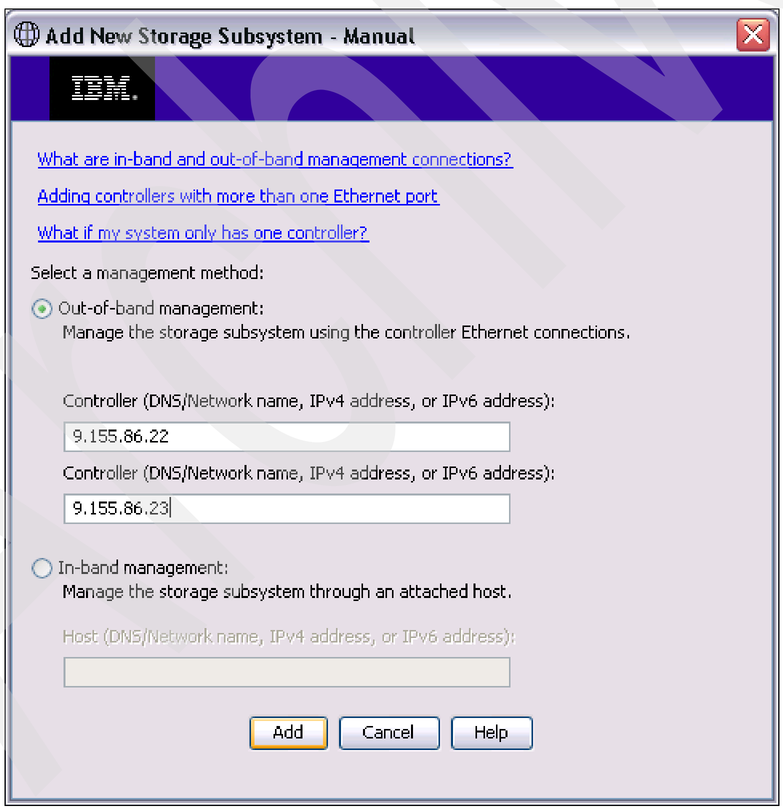


Figure 7-12 DS Storage Manager Add New Storage Subsystem dialog

15. As we do not want to add another Storage System to the management domain we choose **No** in the Storage Subsystem Added confirmation dialog, as shown in Figure 7-13.

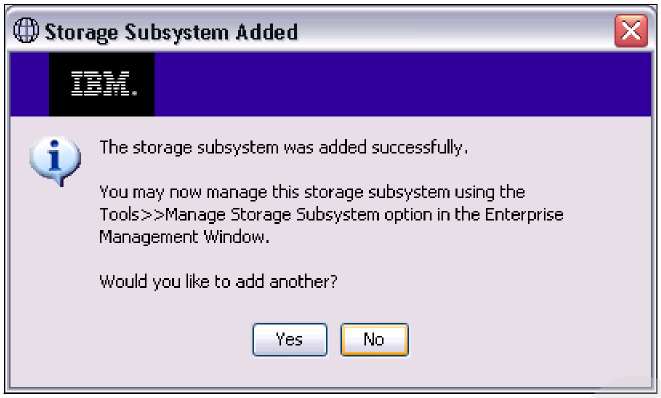


Figure 7-13 DS Storage Manager Storage Subsystem Added dialog

16. The DS Storage Manager window is now showing our added DS4800 Storage System iSeries_VIOS, as shown in Figure 7-14.

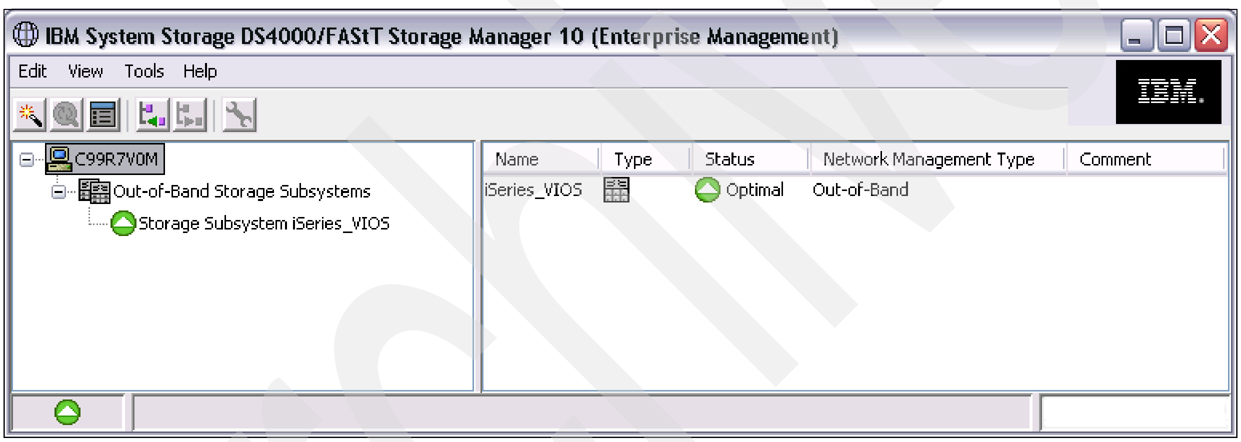


Figure 7-14 DS Storage Manager with added storage subsystem

Now we are all set up for our DS4800 logical storage configuration for our VIOS attached IBM i client, as described in 7.4, "DS4000/DS5000 Storage configuration with the GUI" on page 252.

7.2 Configuring DS remote support

In this section we outline how to configure remote support for IBM System Storage DS Midrange Storage subsystems.

7.2.1 Configuring SNMP trap in Storage Manager

To allow the DS4000 management workstation to send SNMP traps to the RSM server, set the RSM server as your SNMP traps destination in the Storage Manager client. Open Storage Manager's Enterprise Management window and select **Edit** → **Configure Alerts®**, as shown in Figure 7-15. This configuration will also be applied to the event monitor (if installed) on that same system.

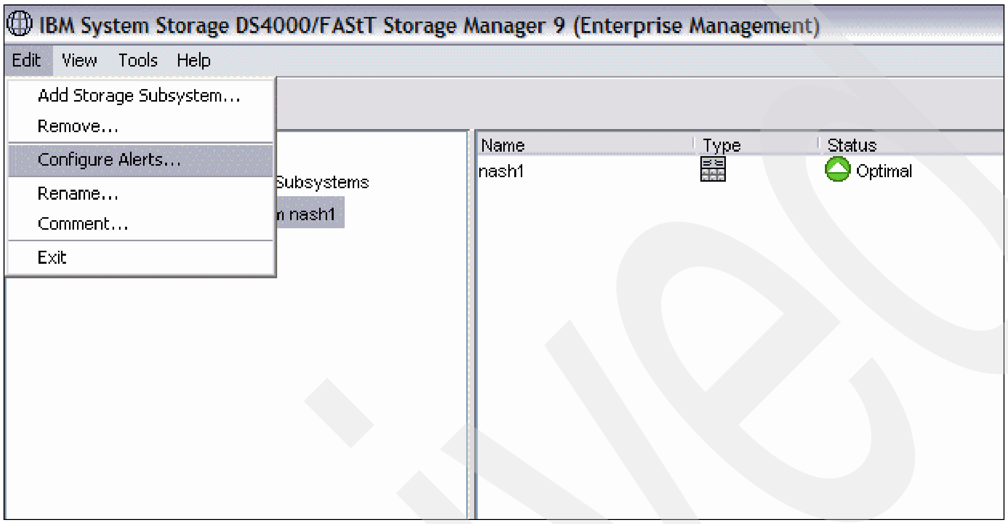


Figure 7-15 Configure Alert in Storage Manager

If you have more than one storage subsystem managed by the Storage Manager, select **All storage subsystems** and then click **OK** to bring up a Configure Alert window. Click the **SNMP** tab and type the IP address of the RSM server in the Trap Destination field, as shown in Figure 7-16. Keep the default community name *public*.

If you have an existing SNMP infrastructure and there is already an SNMP trap destination set, you can add the IP address of the RSM server as an additional SNMP trap destination without having to delete the existing SNMP trap destination setting.

To send an alert validation to test the SNMP trap, click **Validate**, as shown in Figure 7-16. This test sends a trap message to the RSM server as the SNMP trap destination. Click **OK** to close the dialog, then click **Add** and **OK** to finish.

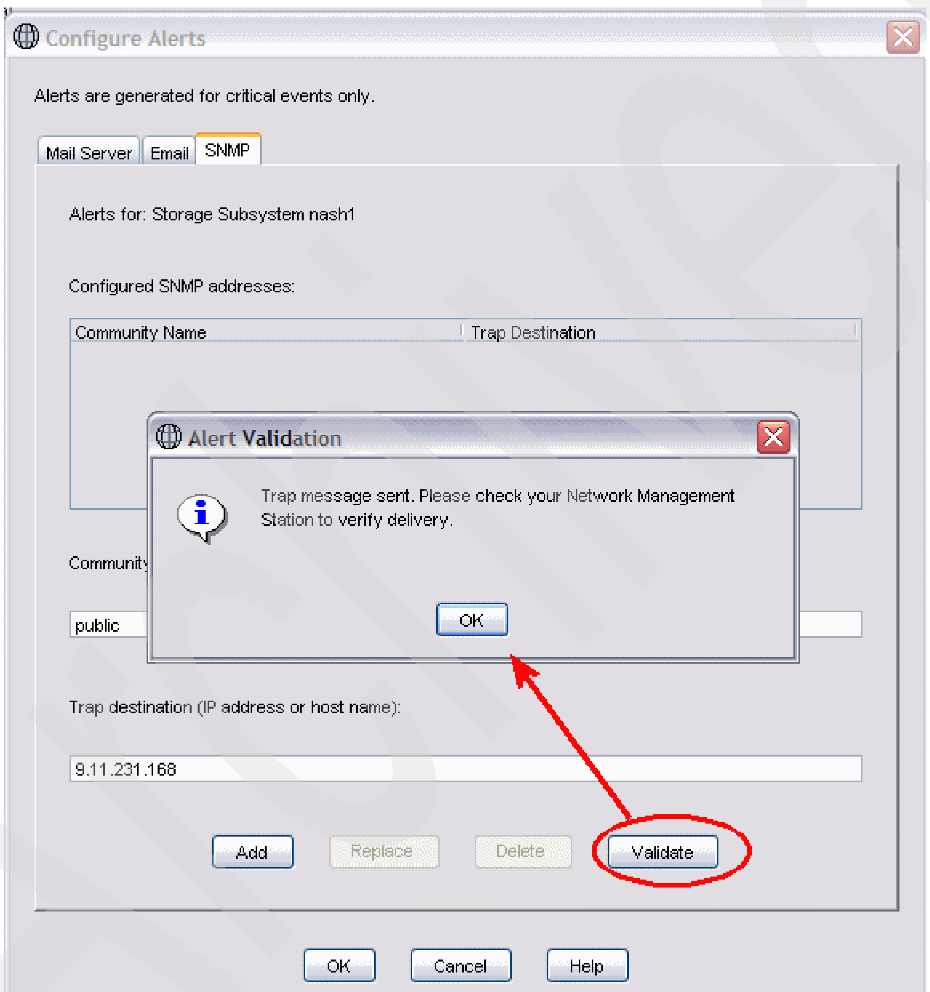


Figure 7-16 Configure SNMP trap and send a trap message

To check whether the RSM server receives the trap message, log in to the RSM server and click the activity log icon on the RSM server desktop. You should see a message indicating reception of the test alert, as depicted in Figure 7-17.

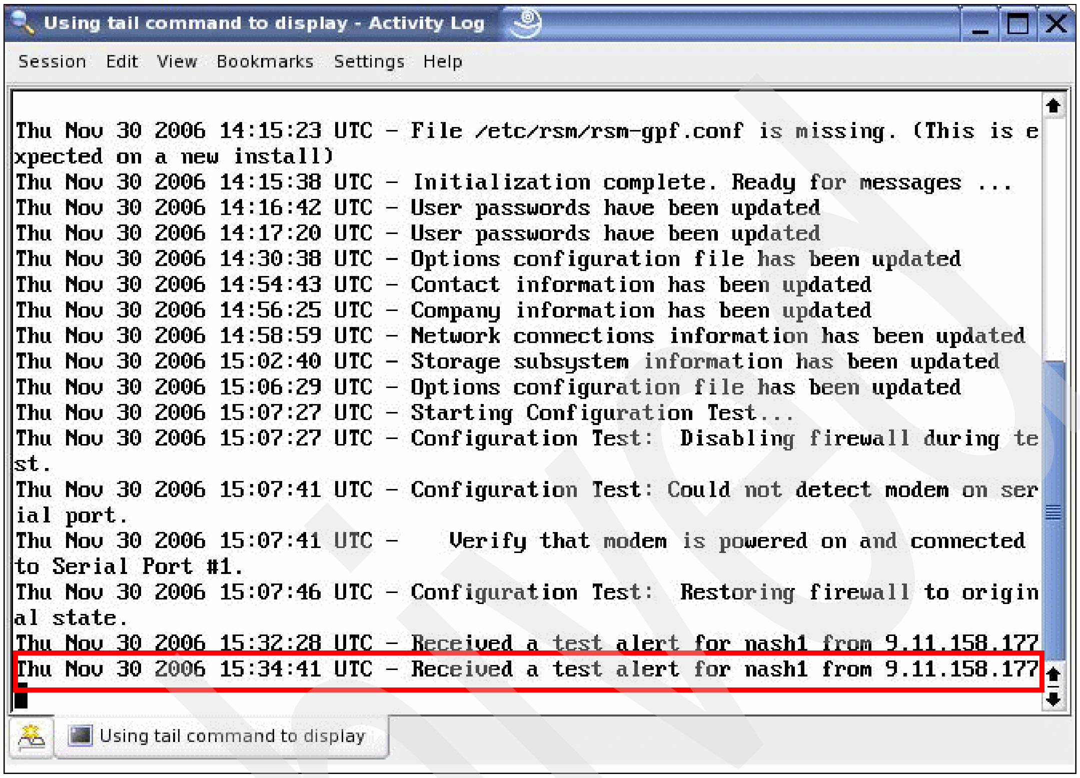


Figure 7-17 Activity Log to check the test alert sent from the management station

7.2.2 Notification e-mail and events filtering

The RSM for Storage application sends e-mails to notify you about important changes in status. There are several types of notification sent by RSM to the primary contact or subsystem contact, as configured in the RSM:

- Remote access notifications, sent when:
  - Remote access is enabled or disabled.
  - A remote user connects or disconnects from the system.
  - The remote access automatic time out is about to expire and the system has one or more active alerts.
- Alerts status notifications, sent when an alert has been sent to IBM Service.
- Daily status e-mails, which serve as a heartbeat notification that the RSM for Storage System is operational. These include the summary status for the system and the status of alerts that may be active for storage subsystems.
- Firewall status notifications, sent when any change is made to the internal firewall settings for the RSM for Storage System.

- Ignored Event notifications, sent when an event is received that is configured to be ignored by the RSM for Storage System, and therefore not reported to IBM Service. These are events for which a response by IBM Service is not usually required, as listed in Table 7-1.

Table 7-1 Filtered events

Event code	Event text
6200	FlashCopy repository logical drive capacity - threshold.
6202	FlashCopy logical drive failed.
none	The persistent monitor running on Host xxxxxxx cannot reach the indicated Storage System.
none	The persistent monitor running on Host xxxxxxx can now reach the indicated Storage System.
4011	Logical drive not on preferred path due to ADT/RDAC.

RSM and Storage Manager e-mail alerts

Storage Manager in the management station can be configured to send e-mail notifications when a problem is detected. However, this feature *should be disabled* when RSM for Storage is installed and the e-mail contact configured in the Storage Manager is the same as the e-mail contact configured in the RSM contact list. Otherwise, you will receive multiple notification e-mails about the same problem—one notification from RSM and another one from Storage Manager's e-mail alerts.

To disable e-mail alerts in Storage Manager:

1. Select **Configure Alerts** in Storage Manager's Task Assistant and select **All Storage Subsystems**.
2. On the E-mail tab, delete any configured e-mail destinations.

If there are e-mail addresses already configured to receive e-mails from Storage Manager but are not listed in the RSM contact list, it is not necessary to delete them in the Storage Manager.

DS4000 subsystem security

Storage Manager has the ability to require an administrative password in order to make changes to the subsystem configuration. We recommend configuring this password. The password can be set in the Storage Manager subsystem view under **Storage Subsystem → Set Password** (Figure 7-18).

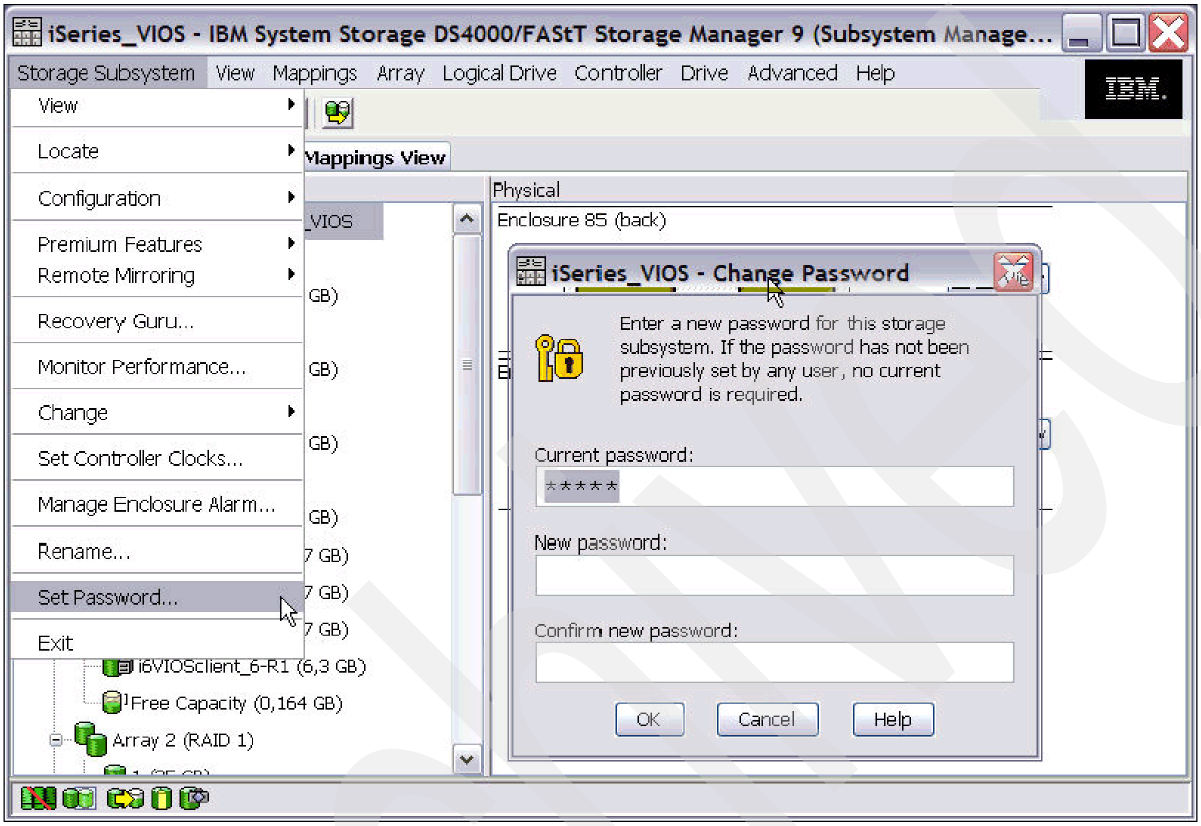


Figure 7-18 Set Storage Manager password

DS4000 also has a controller shell environment that is accessible using a remote login (RLOGIN) client. IBM Storage Manager for DS4000 has an option to disable RLOGIN, and we normally recommend disabling RLOGIN.

To disable RLogin, first click one controller to highlight it. Then select **Controller** → **Change** → **Network Configuration**. Select a controller and click the **Advanced** button. Unmark the decision box for enabling RLogin (Figure 7-19). Repeat this procedure for the second controller.

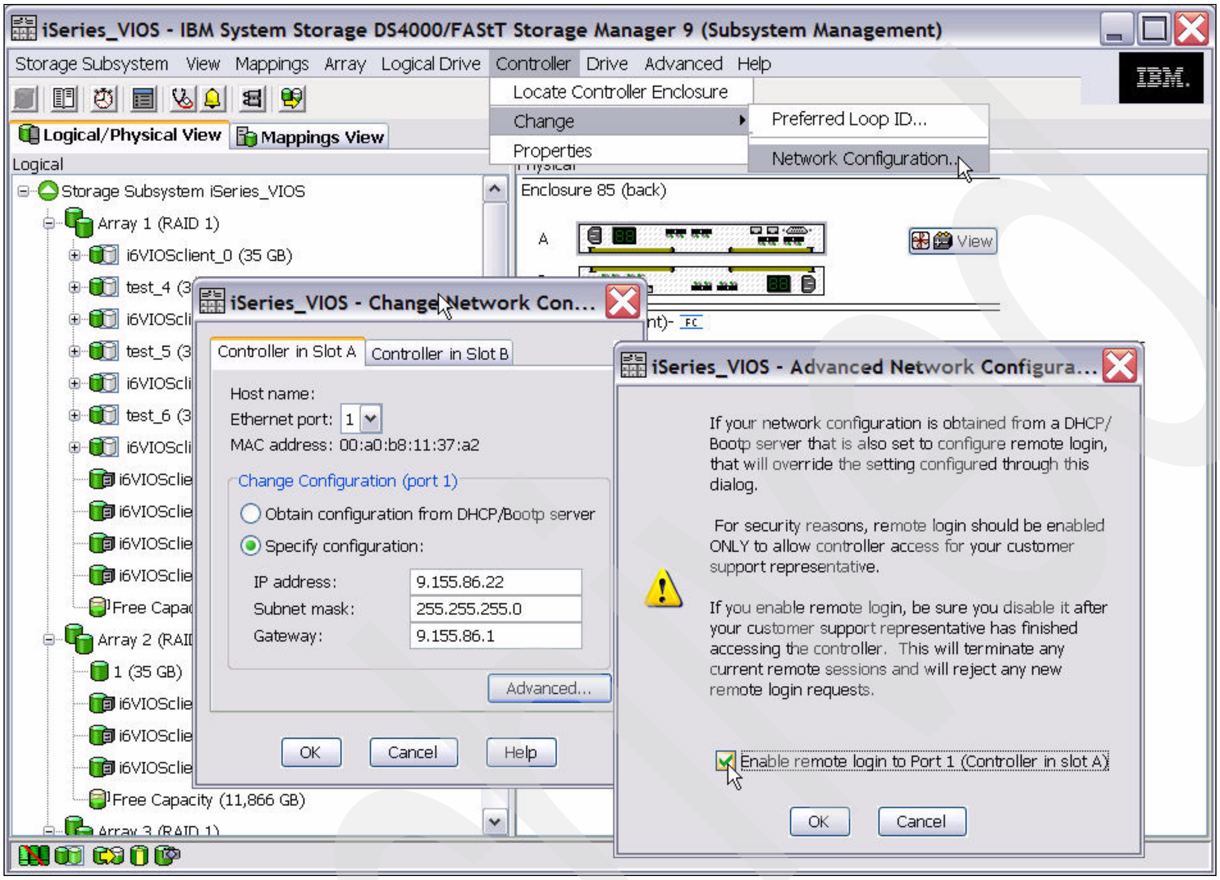


Figure 7-19 Disabling RLogin for the controllers

For further information about RSM for Storage, a comprehensive Planning, Installation, and User's Guide is available to download from the following Web page:

http://www.ibm.com/support/docview.wss?uid=psg1MIGR-66062&rs=594

7.3 DS3400 Storage Configuration with the GUI

In this section we show how to configure an IBM System Storage DS3400 with the DS Storage Manager Client for attachment of an IBM i client of VIOS.

7.3.1 Adding DS3400 to DS Storage Manager Client

To add DS3400 to the DS Storage Manager Client launch the client on the PC. Expand **Edit** → **Add Storage Subsystem**. In the Add New Storage Subsystem window we keep the button selected for out-of-band management and insert the IP addresses of controllers A and B in DS3400, as shown in Figure 7-20. Then we click **Add**.

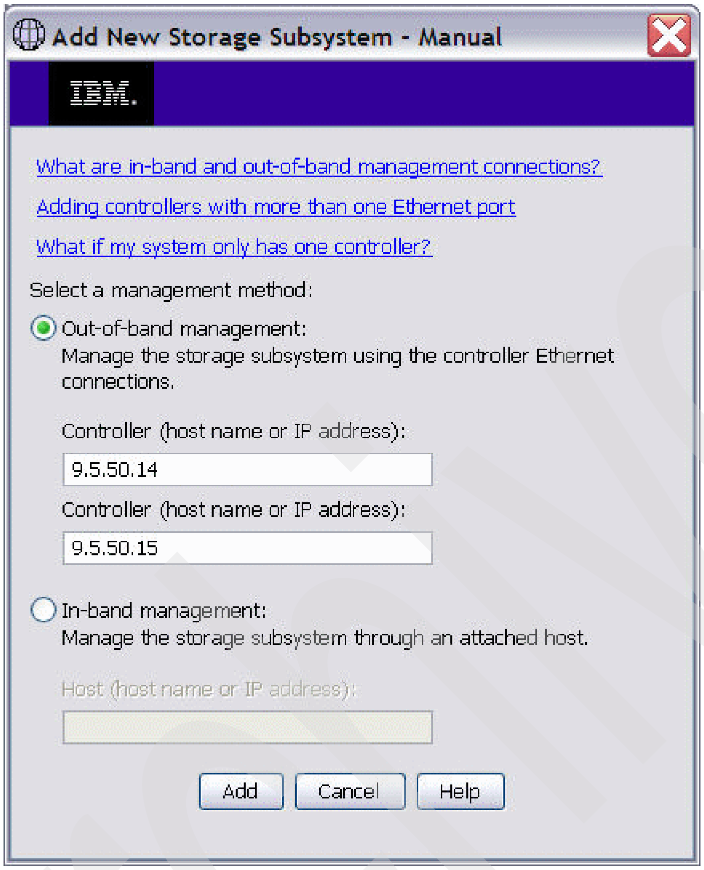


Figure 7-20 Add DS3400 to DS Storage Manager Client

The information window pops up, confirming that the Storage System was added successfully. The added DS3400 is now shown in the tree view of the left pane in DS Storage Manager Client.

To launch the DS Storage Manager Client window for DS3400D, we double-click **DS3400 storage subsystem** in the tree view. The DS3400 window is shown in Figure 7-21.

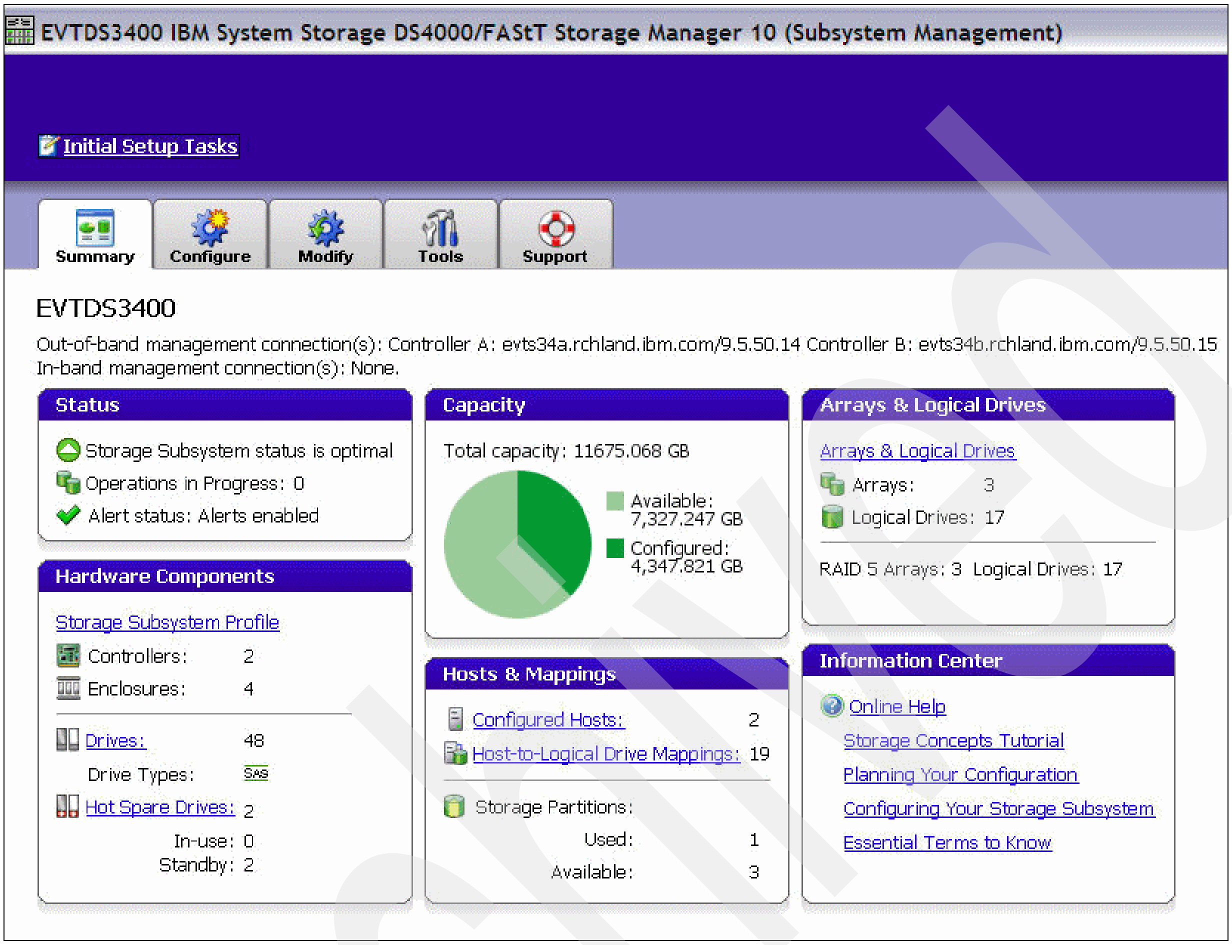


Figure 7-21 DS Storage Manager Client: DS3400 window

We also see the window EVTD3400 - Initial Setup tasks, as shown in Figure 7-22. You may want to use the tasks in this window if you are new to DS3400 configuration. In our setup we prefer to do configuration tasks from the DS3400 menu, so we close the Initial Setup Tasks window.

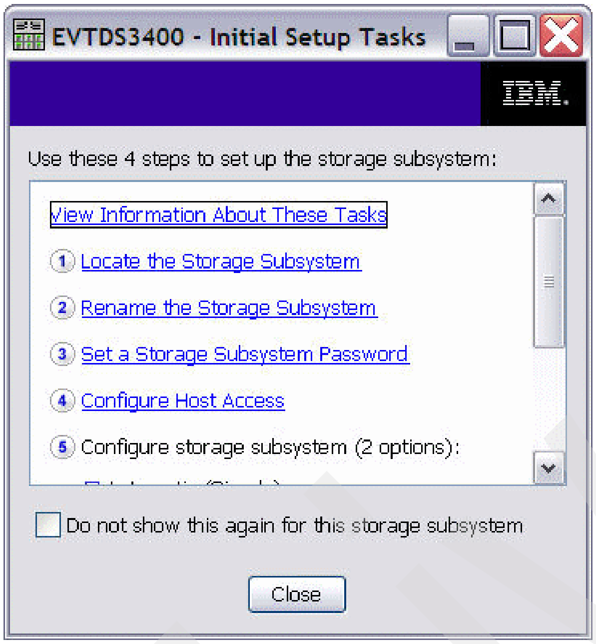


Figure 7-22 Initial Setup Tasks window

7.3.2 Creating hot spares

Hot spare drives are used to give an additional level of protection in case a disk drive fails, since a hot spare automatically replaces a failed disk drive that is part of an array. The following rules apply when creating hot spares:

- It must be a disk drive type that matches the other disk drives in the array. For instance, SAS disk drives in an array require SAS hot spare. SATA drives need SATA hot spare.
- It must be as large as or larger than the largest disk drive in the array.
- At least one hot spare should be created in Storage System, but we recommend creating more of them. A good number is one spare per two enclosures.

We create two hot spares by using the following steps:

1. In DS3400 window we use **Configure** → **Configure Hot Spares** → **Configure Hot Spares (Manual)**. In the Configure Hot Spares window we select the type of disk drive for the spare and click **Assign**, as can shown in Figure 7-23.

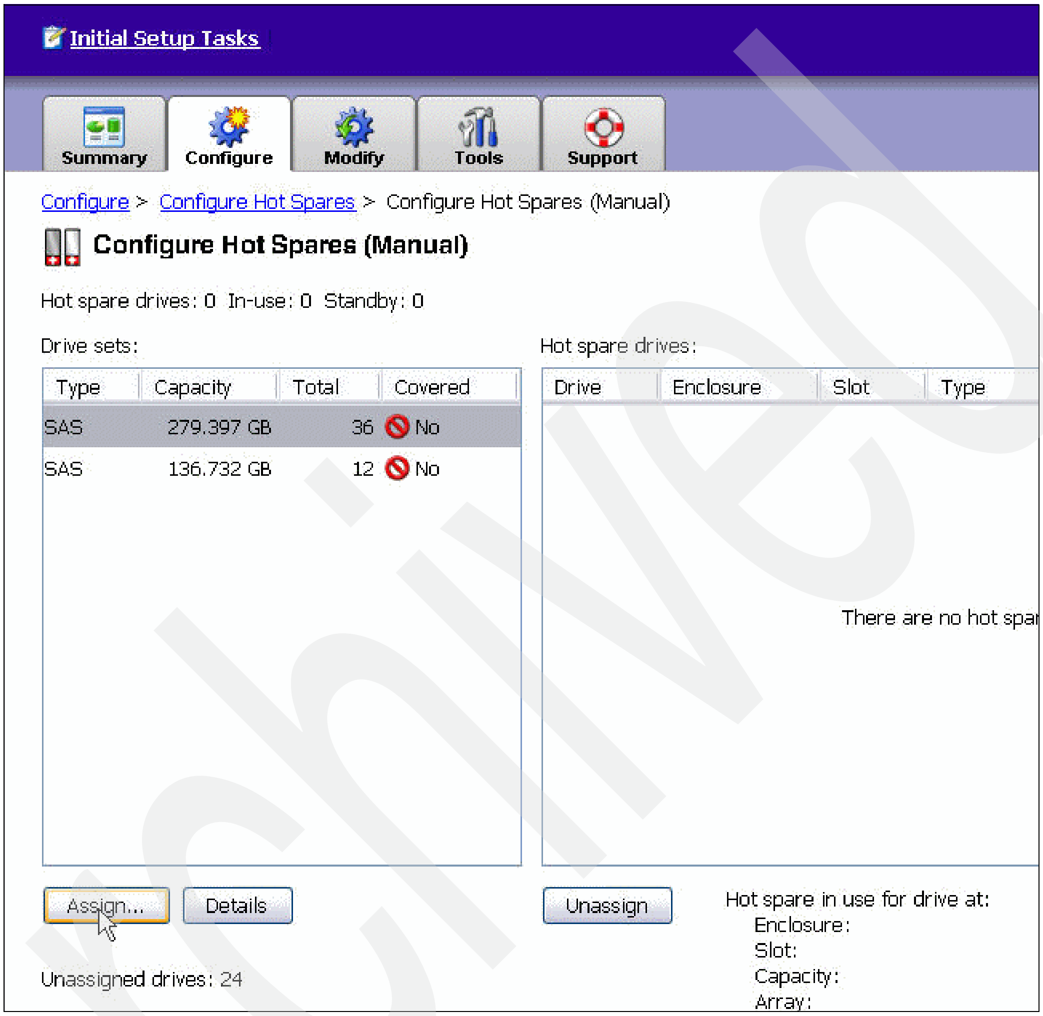


Figure 7-23 Assigning disk drives for hot spares

2. In the next window we get a list of disk drives that can be used as hot spare. As shown in Figure 7-24, we select a drive in each enclosure 1 and 85, and confirm our selection by clicking **OK**.

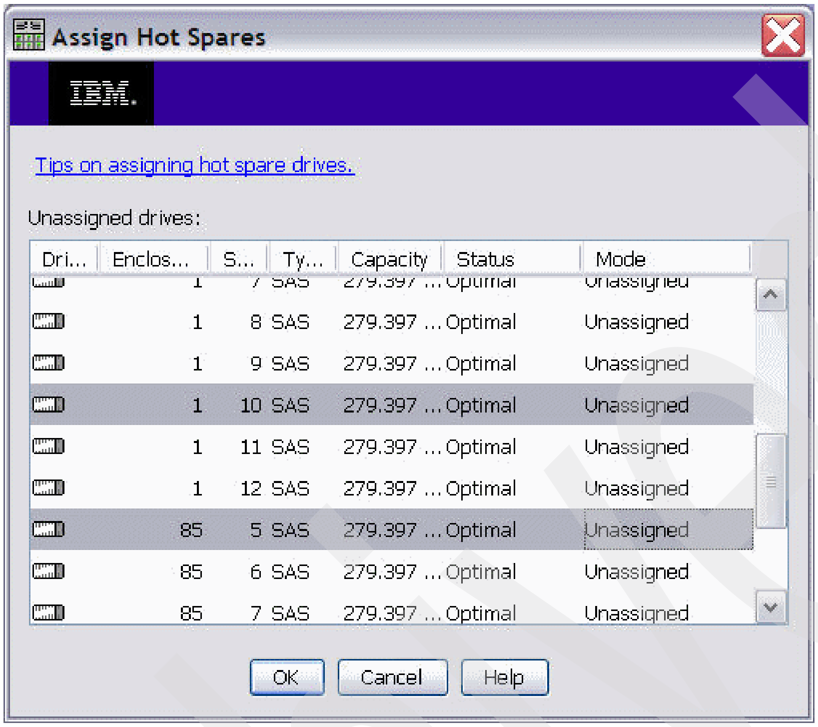


Figure 7-24 Select drives for hot spares

3. After a few seconds new hot spares show as assigned in the Configure Hot Spares window. They are in standby mode, currently not replacing any disk drive. See Figure 7-25.

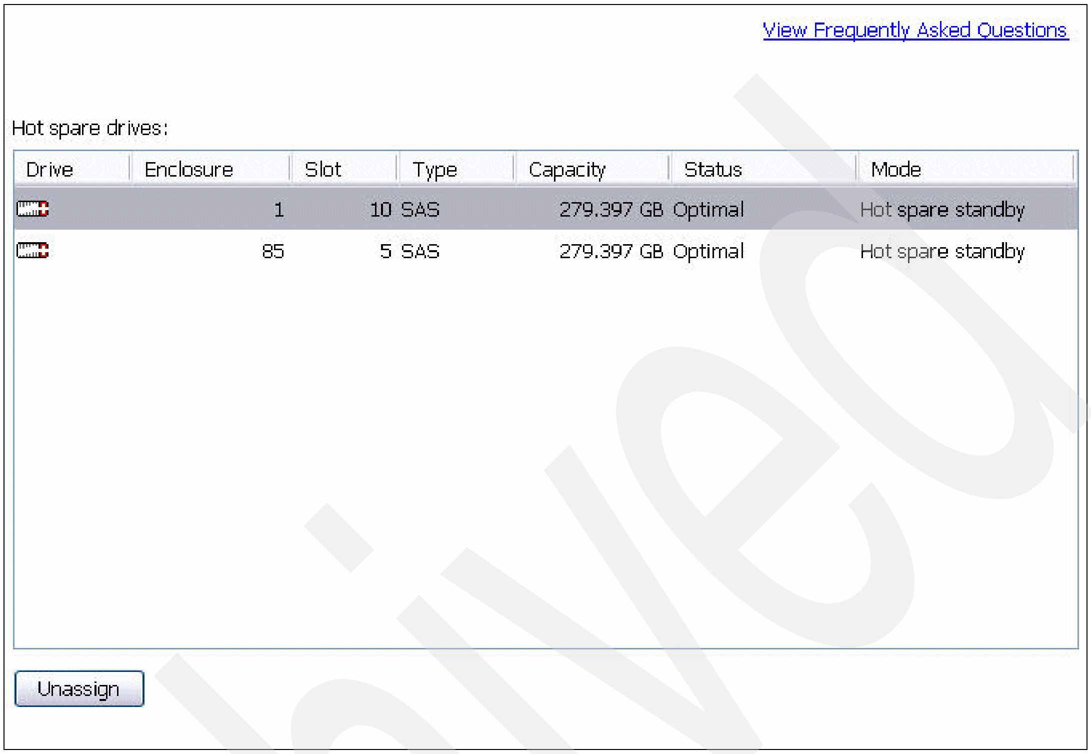


Figure 7-25 Assigned hot spares

7.3.3 Creating RAID arrays and logical drives

By using the following steps we create a RAID-10 array of four disk drives and seven logical drives:

1. We use Configure → Create Logical Drives. We keep Unconfigured Capacity: Create a new an array and logical drive checked and click **Next** (Figure 7-26).

Figure 7-26 Select capacity type for the new array

2. In the Create Logical Drives - Drive Selection Choices window (Figure 7-27) we select **Manual (Advanced)** and click **Next**.

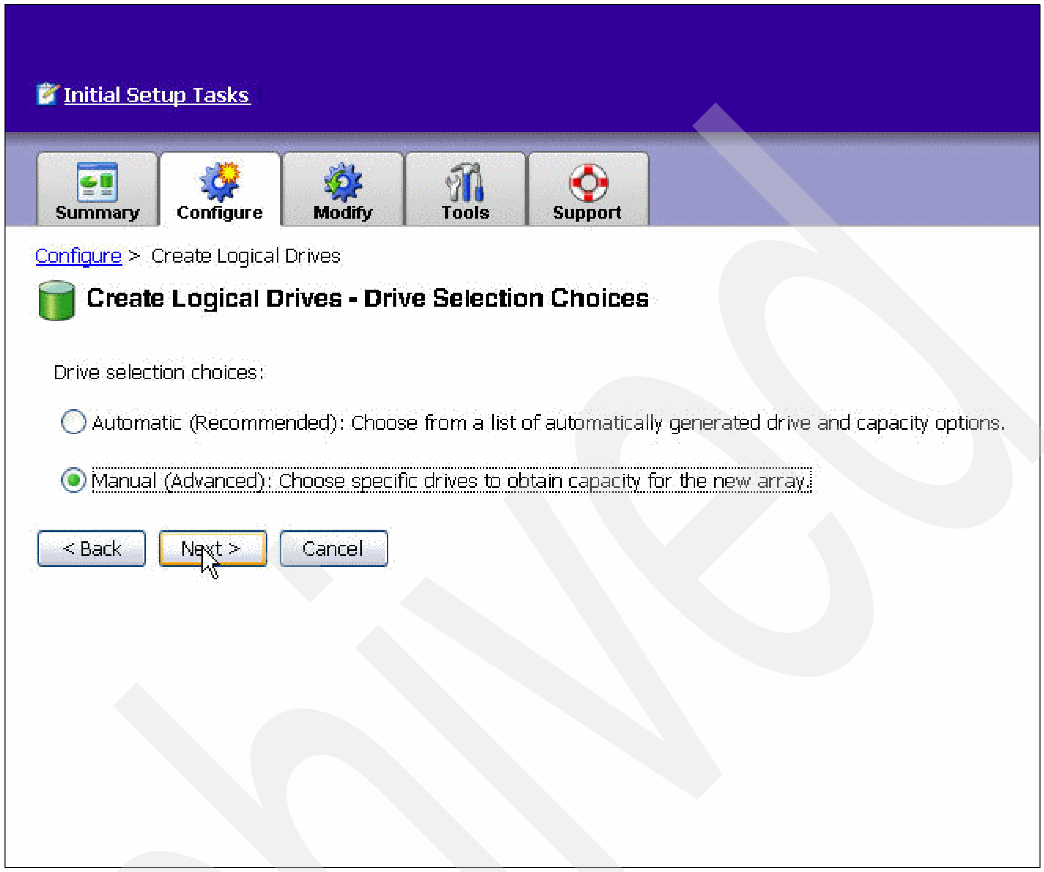


Figure 7-27 Manual drive choice selected

3. In the Create Logical Drives - Manual Drive Selection window we select RAID level RAID-1. We select the a disk drive in enclosure 1, slot 8, and a drive in enclosure 85, slot 6, and click **Add**, as shown on Figure 7-28.

   Note that if you choose RAID-1 and select four or more disk drives, RAID-10 is automatically configured across the array. You choose drives in pairs of two at a time.

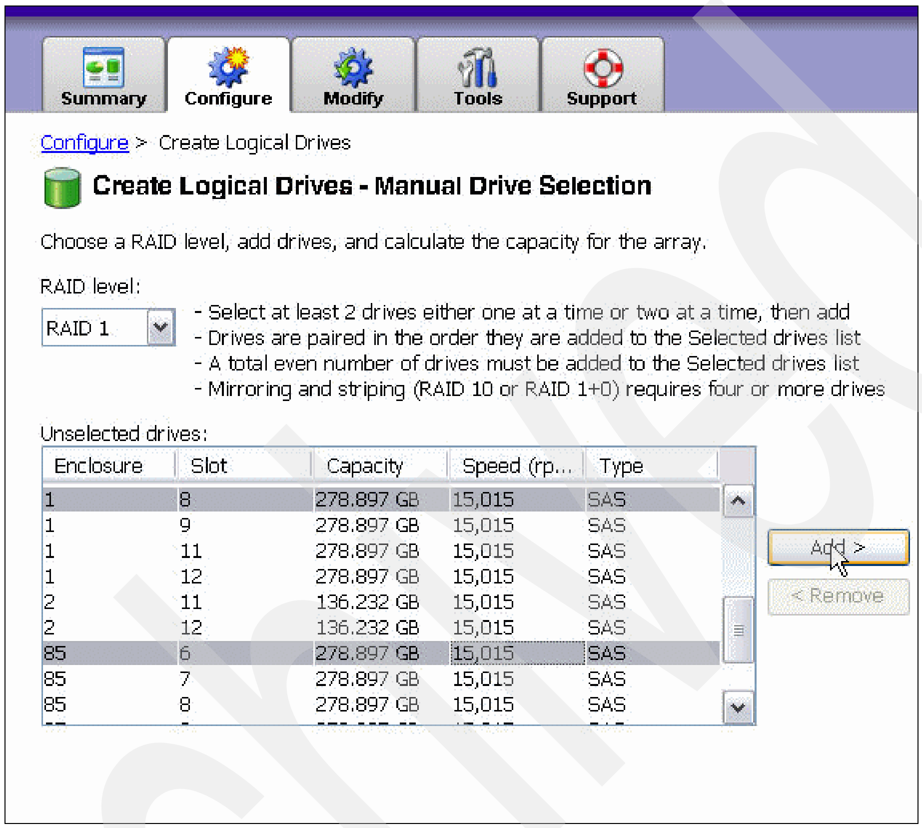


Figure 7-28 Select drives for RAID-10

4. We select and add the next pair of disk drives each from one enclosure, the same way as we did in step 3. Creating a RAID-10 array form disk drives in two different enclosures ensures availability in case one enclosure fails.

5. After clicking **Calculate Capacity** we see the available capacity in the new RAID-10 array, as shown on Figure 7-29. We proceed by clicking **Next**.

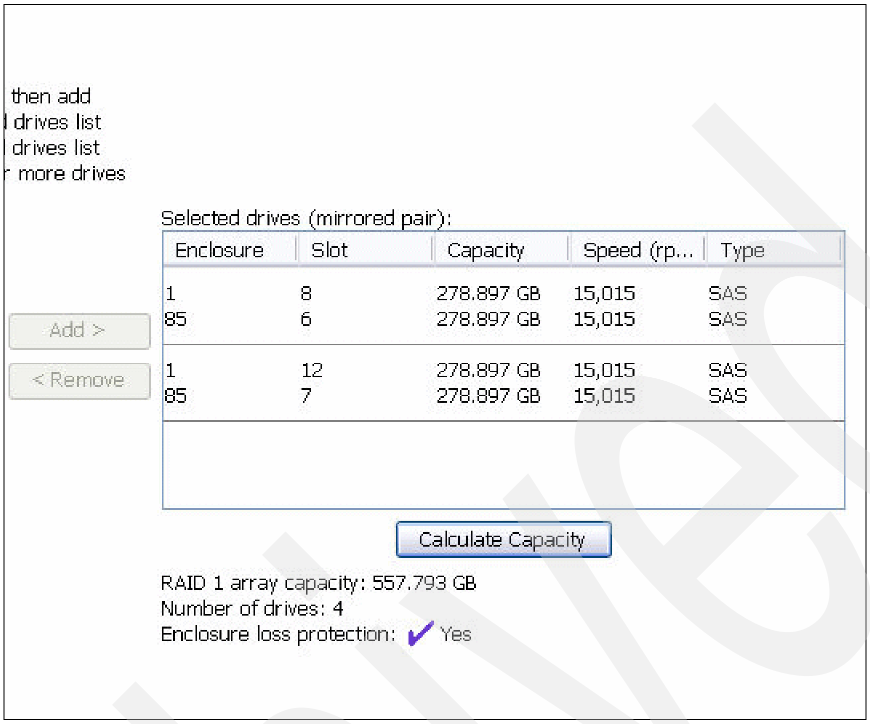


Figure 7-29 Capacity of created array

6. In the Create Logical Drives - Specify Logical Drive window we define a logical drive of capacity 35 GB. We specify Database as Logical Drive I/O characteristics and click **Next**.

> **Note** With DS3400 we have the option of choosing between three predefined types of logical drives based on I/O characteristics:
>
> - File-system
> - Database
> - Multimedia
>
> All three types use cache pre-fetch. File-system and database have segment sizes of 128 KB, while multimedia uses segment size 256 KB.
>
> For System i connection we define database logical drive because its segment size is the best suited to IBM i blocksizes.

For more information about specifying segment size refer to 7.3.3, “Creating RAID arrays and logical drives” on page 235 and 4.5, “Planning considerations for performance” on page 97.

Defining the logical drive can be seen on Figure 7-30.

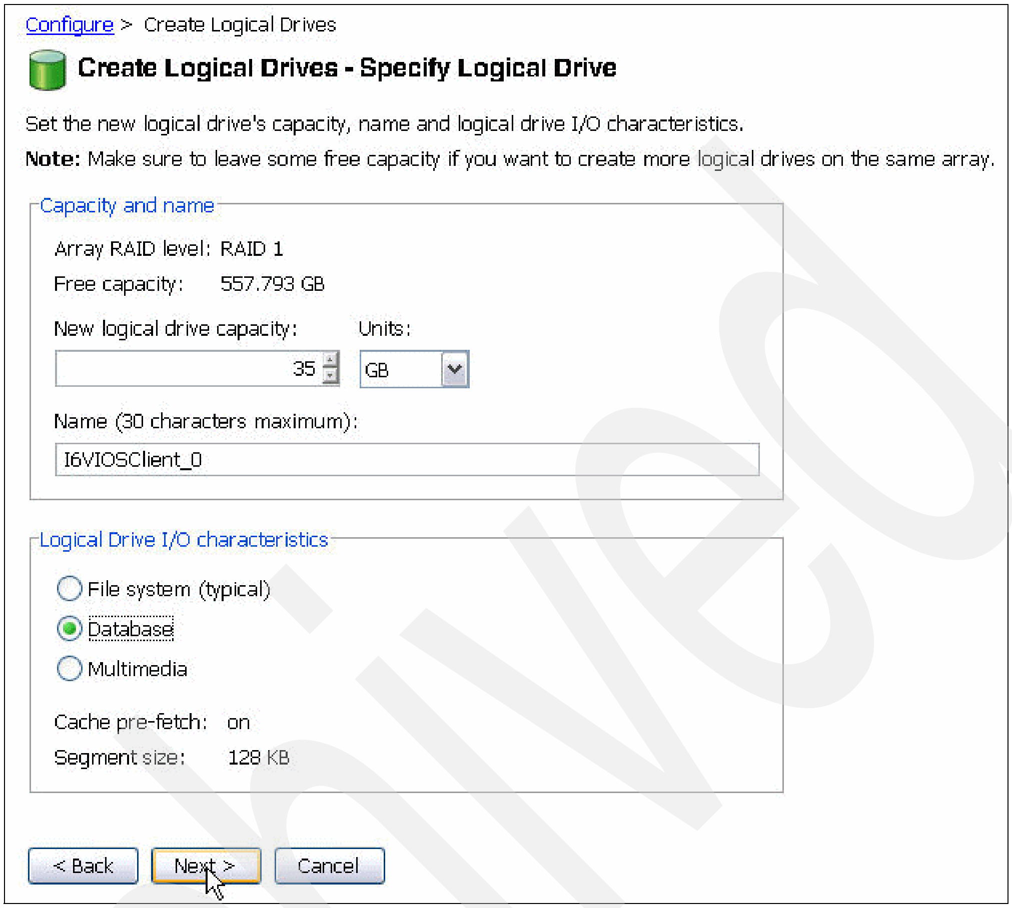


Figure 7-30 Define Logical Drive

7. In the Create Logical Drives - Map Logical Drive To Host window we select **Map later** since we will later map all created logical drives to the host, and we click **Finish**. See Figure 7-31.

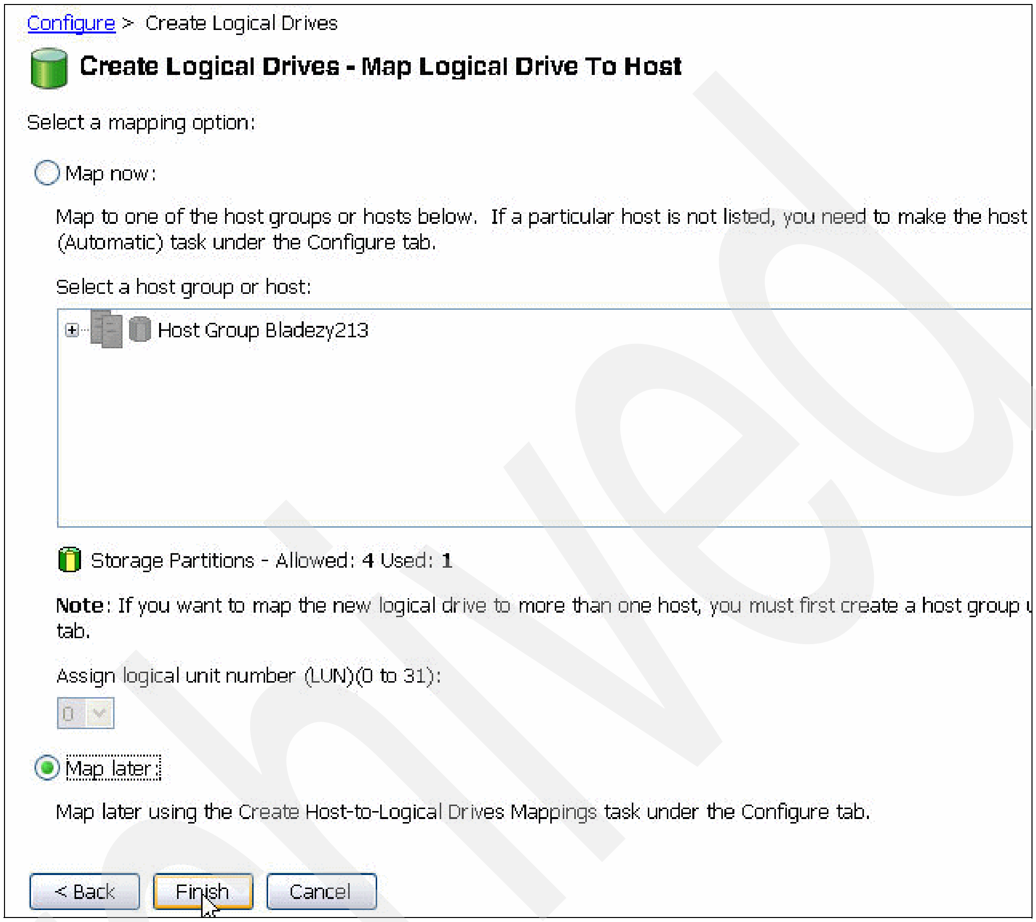


Figure 7-31 Finish defining logical drive

8. On the Create Logical Drives - Complete window we receive confirmation that the drive is successfully created and have the option to define another logical drive. We click **Yes** to create the next logical drive, as shown in Figure 7-32.

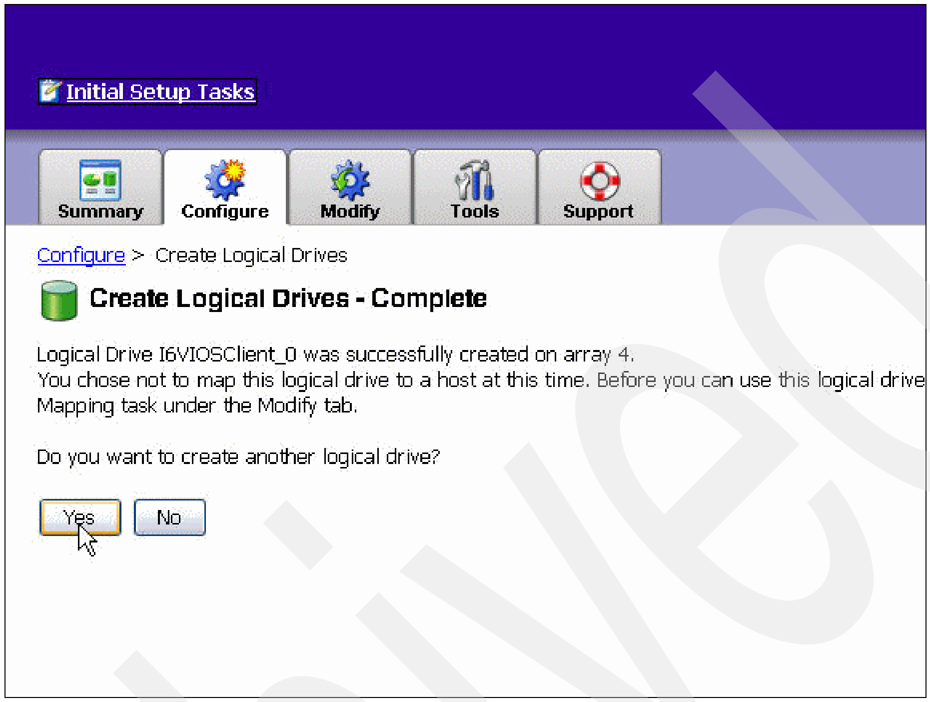


Figure 7-32 Start creating next logical drive

9. To create the next logical drive we select **Free Capacity** in the Create Logical Drives - Select Capacity Type window, then we expand the created array in RAID-1 and select **Free Capacity**, as shown on Figure 7-33. We continue the same way as described in step 6.

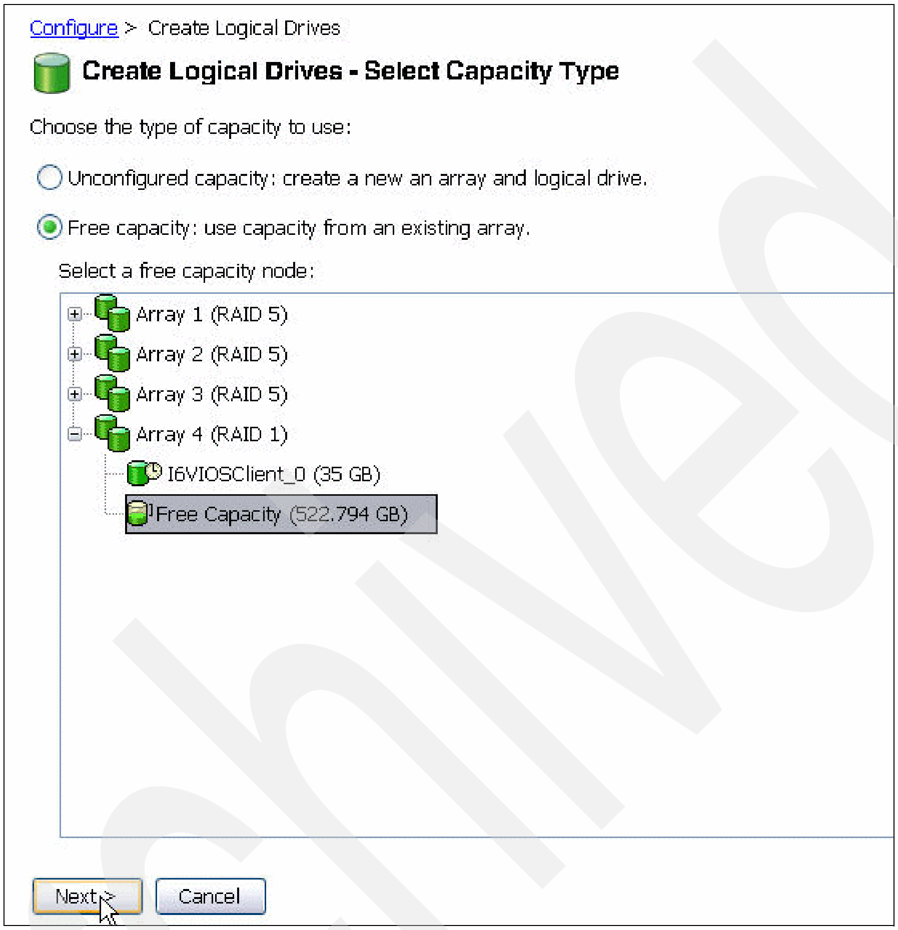


Figure 7-33 Create next logical drive

To check the created logical drives we use the Summary tab window ECTDS3400 → Arrays & Logical Drives. Expand **Array 4**. Created logical drives and free capacity can be seen in the Arrays and logical drives window, which is shown in Figure 7-34.

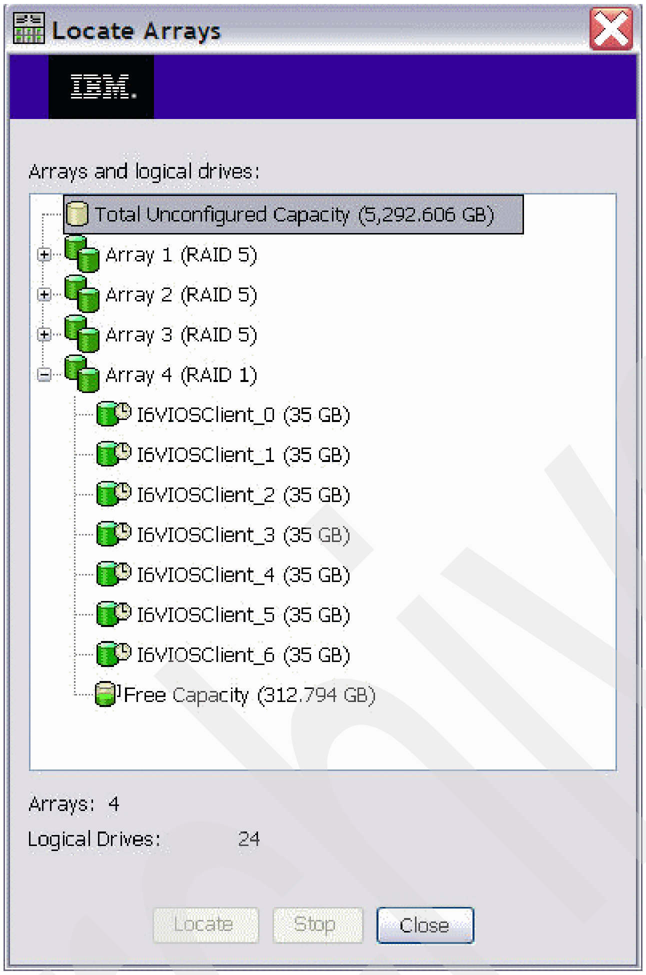


Figure 7-34 Created logical drives

When the logical drives are created they get automatically assigned to a controller in DS3400. The assigned controller has the ownership and preferred path for the logical drives. We recommend that the logical drives for a System i partition are evenly assigned to both controllers A and B. We check the ownership and preferred path and, if needed, we modify it so that half of the logical drives are assigned to each controller. For this we use the following steps:

1. We use the tab Modify → Change Local Drive Ownership/Preferred Path. In the Change Local Drive Ownership/Preferred Path window we expand the array of logical drives that we created and click each of them to see controller ownership and preferred path, as shown in Figure 7-35.

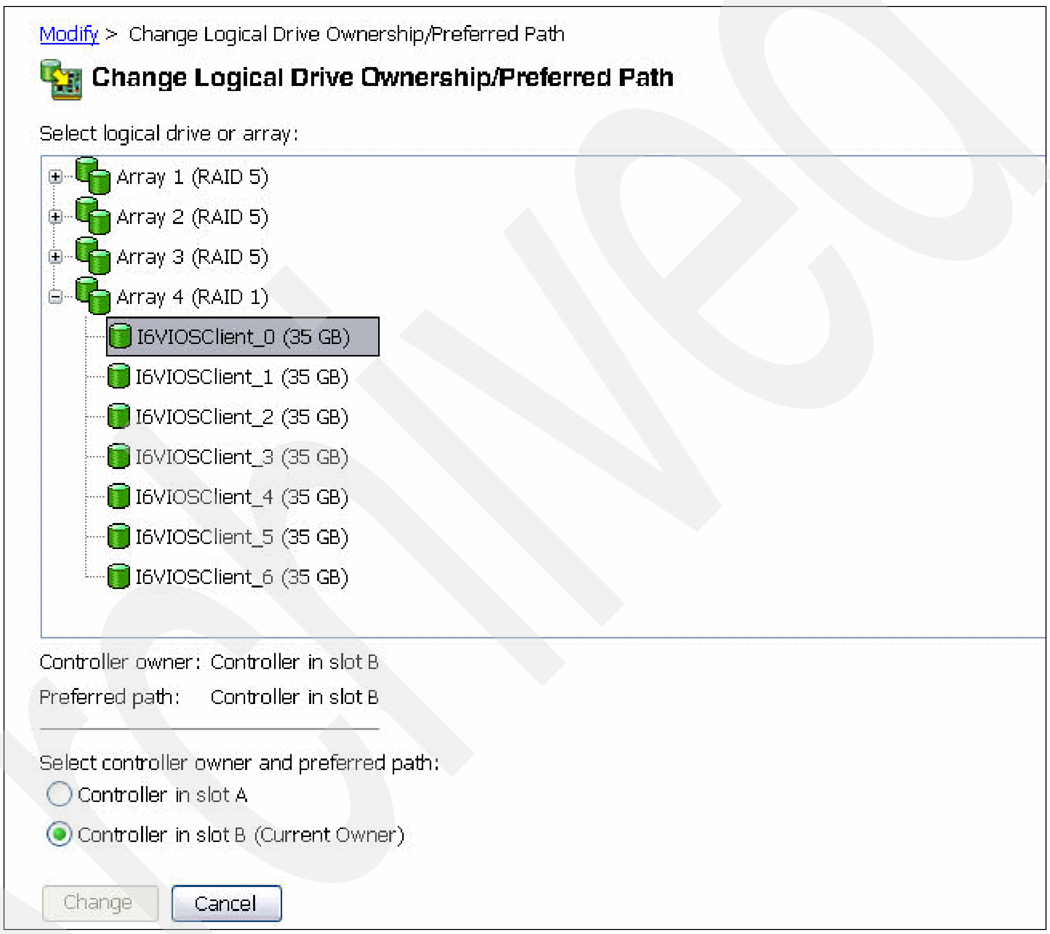


Figure 7-35 Controller ownership and preferred path

2. All seven defined logical drives are assigned to controller B. Therefore, we change assignment of three drives to controller A. For this, we check the box at controller A and click **Change** at each drive that we want to be assigned to A. Changing drive assignment is shown in Figure 7-36.

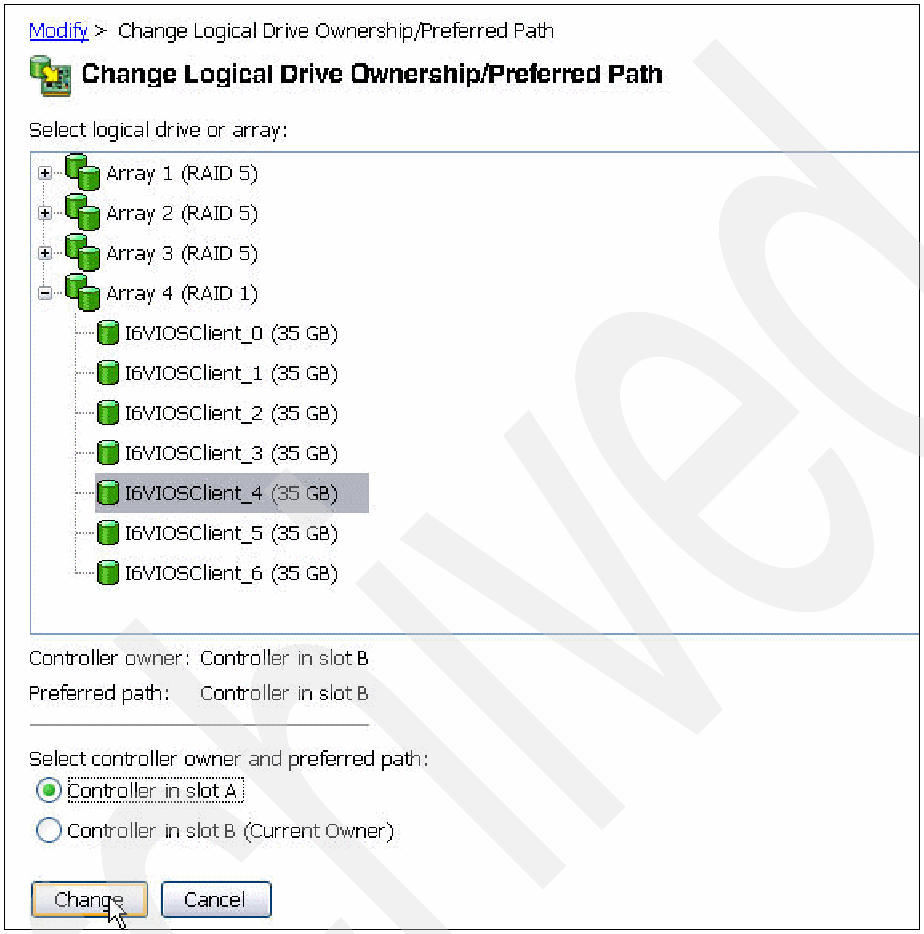


Figure 7-36 Change controller ownership and preferred path

7.3.4 Creating hosts

The logical drives for System i client will be assigned to VIOS, so we create host definition in DS34000 for VIOS:

1. We look for the world-wide port names (WWPNs) of ports in Fibre Channel adapters planned to connect DS3400 form VIOS. For more information about this refer to “Creating hosts” on page 265.
2. We open DS Storage Manager Client → DS3400 Storage Subsystem → Configure, which opens the Configure Storage Subsystem window → Configure Host Access (Manual).

In the next window we specify host name, select **AIX** as the host type, and click **Next** to proceed, as shown in Figure 7-37.

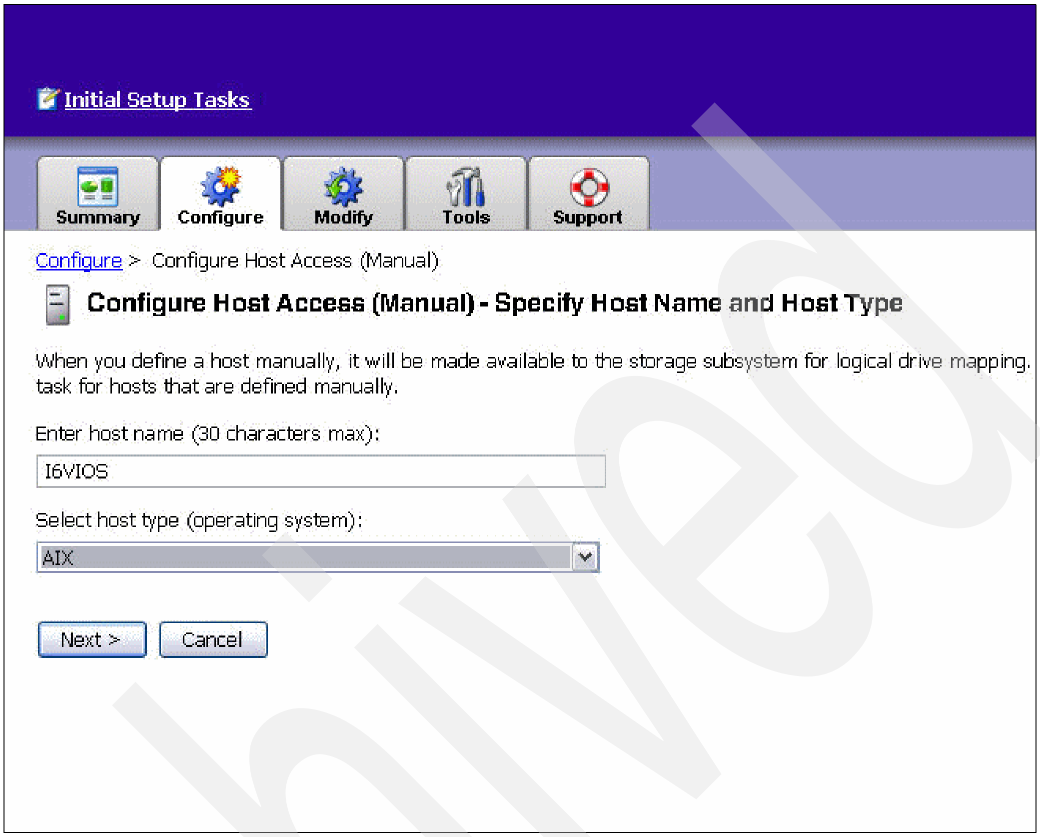


Figure 7-37 Specify host name

3. In the next window we select the WWPns of two adapter ports in VIOS that connect the DS3400 and click **Add**, as can be seen in Figure 7-38.

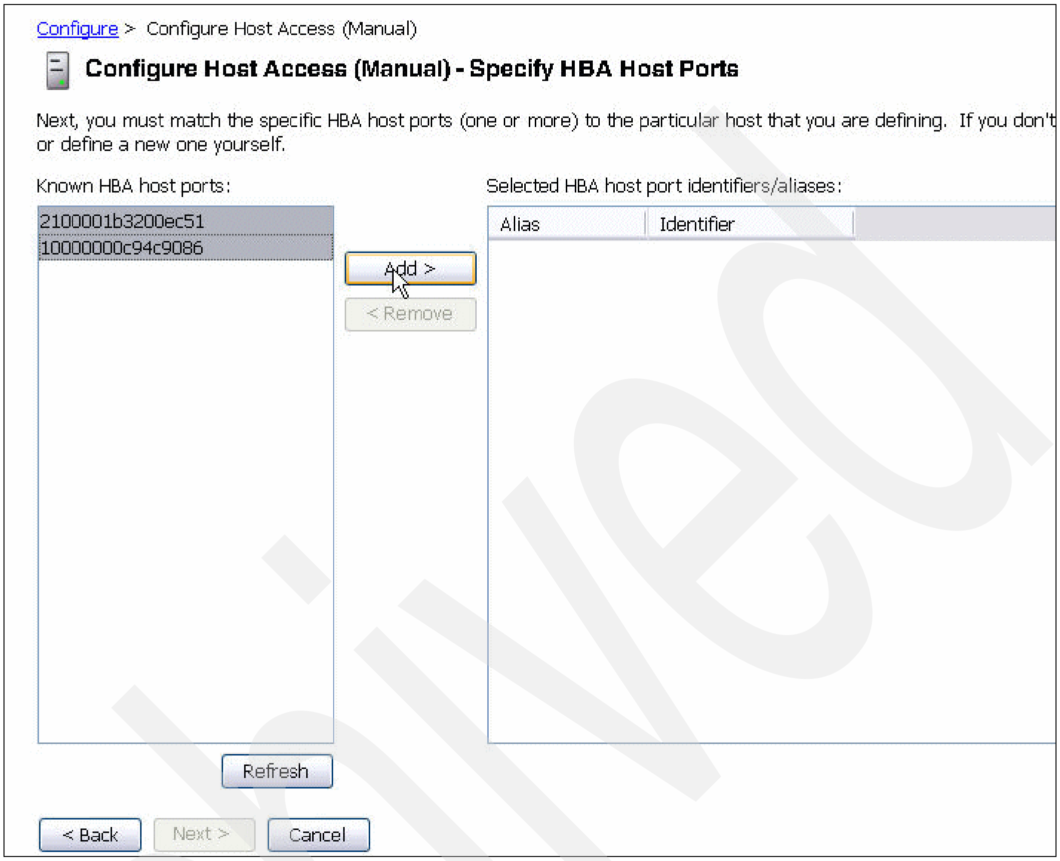


Figure 7-38 Select and add WWPNs

4. As shown in Figure 7-39, DS Storage Manager Client automatically assign alias names for the added WWPNs. We change the names so that they contain a suffix describing the physical adapter location. For information about how to obtain the location of the adapter refer to “Creating hosts” on page 265.

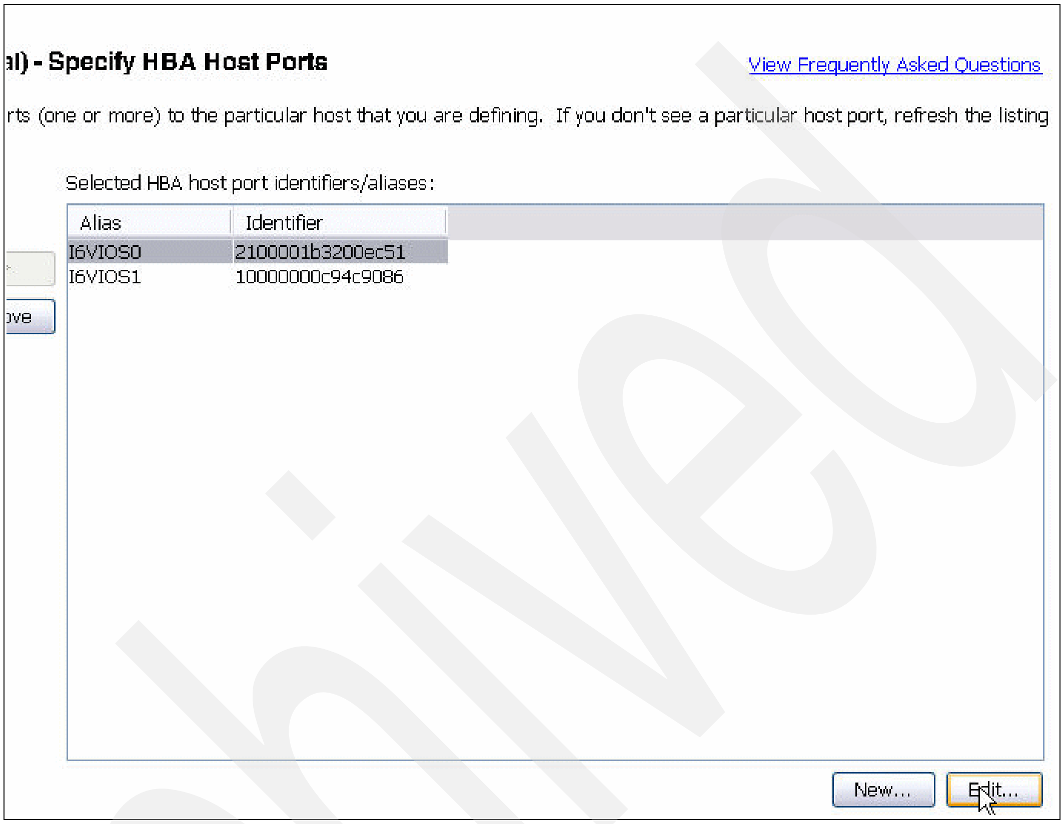


Figure 7-39 WWPN aliases

To change an alias we select a WWPN and click **Edit**. In the Edit HBA Host Port window we change the alias name, as shown in Figure 7-40. After changing the aliases we click **Next** in the Configure Host Access(Manual) - Specify HBA Host Ports window.

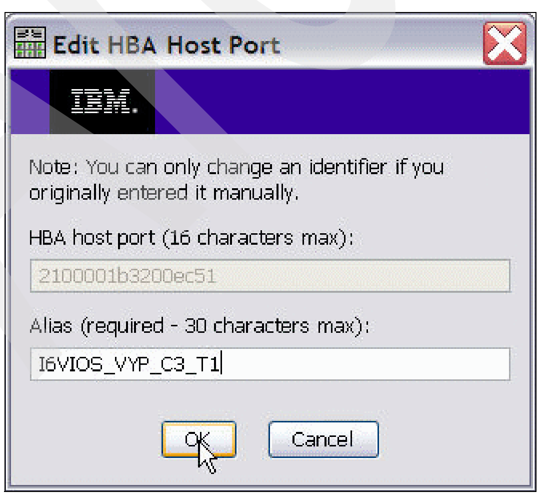


Figure 7-40 Change the alias of WWPN

5. In the next window we can select whether to share the logical drives with another host. Since we do not want this we keep the **No** box selected and click **Next**, as shown in Figure 7-41.

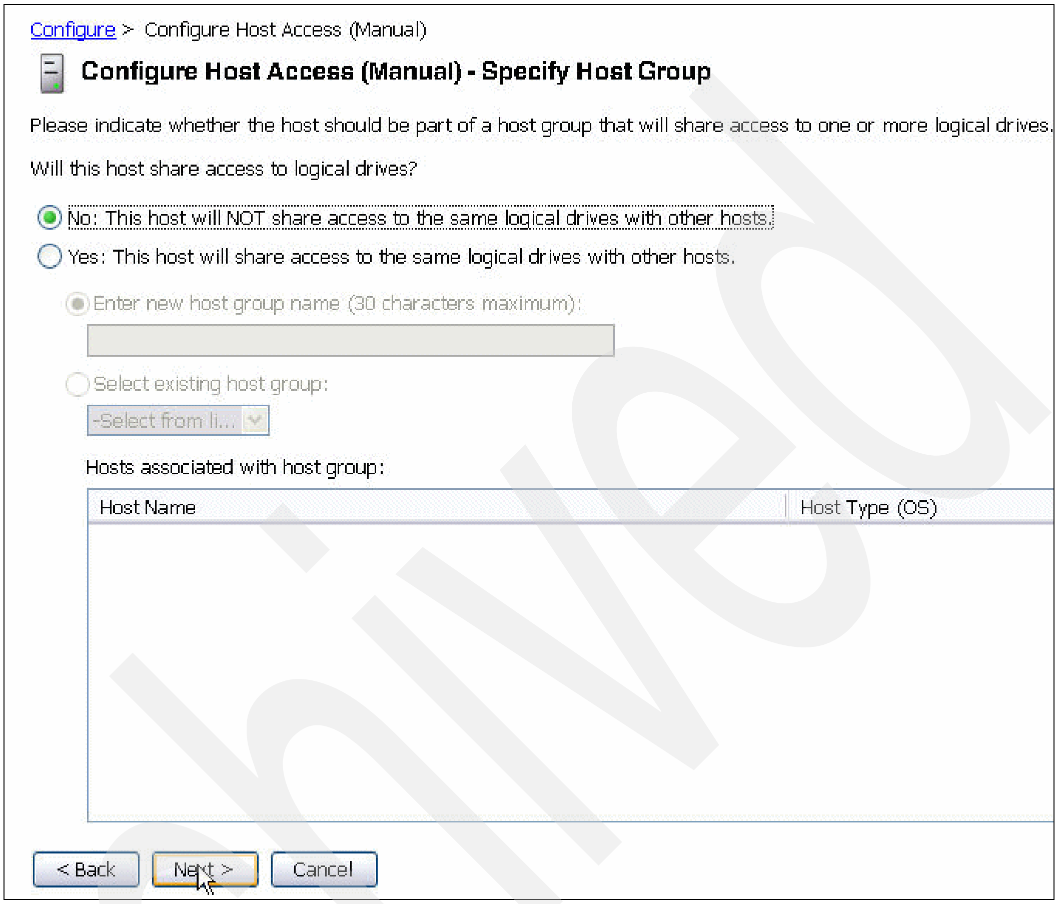


Figure 7-41 Un-select host sharing

6. In the conformation window we check that the WWPNs and aliases are correct, then we click **Finish** to define the host name.

7.3.5 Defining logical drive to LUN mapping

After VIOS is defined as host to DS3400 we map created System i logical drives to the VIOS host name. To do this:

1. We open **DS Storage Manager Client** → **DS3400 Storage Subsystem** → **Configure**, which opens the Configure Storage Subsystem → Create Host-to-Logical Drive Mappings window. This starts the Create Host-to-Logical Drive Mappings wizard.

2. In the wizard window Select Host we select the host name for the VIOS that we created and click **Next**, as shown in Figure 7-42.

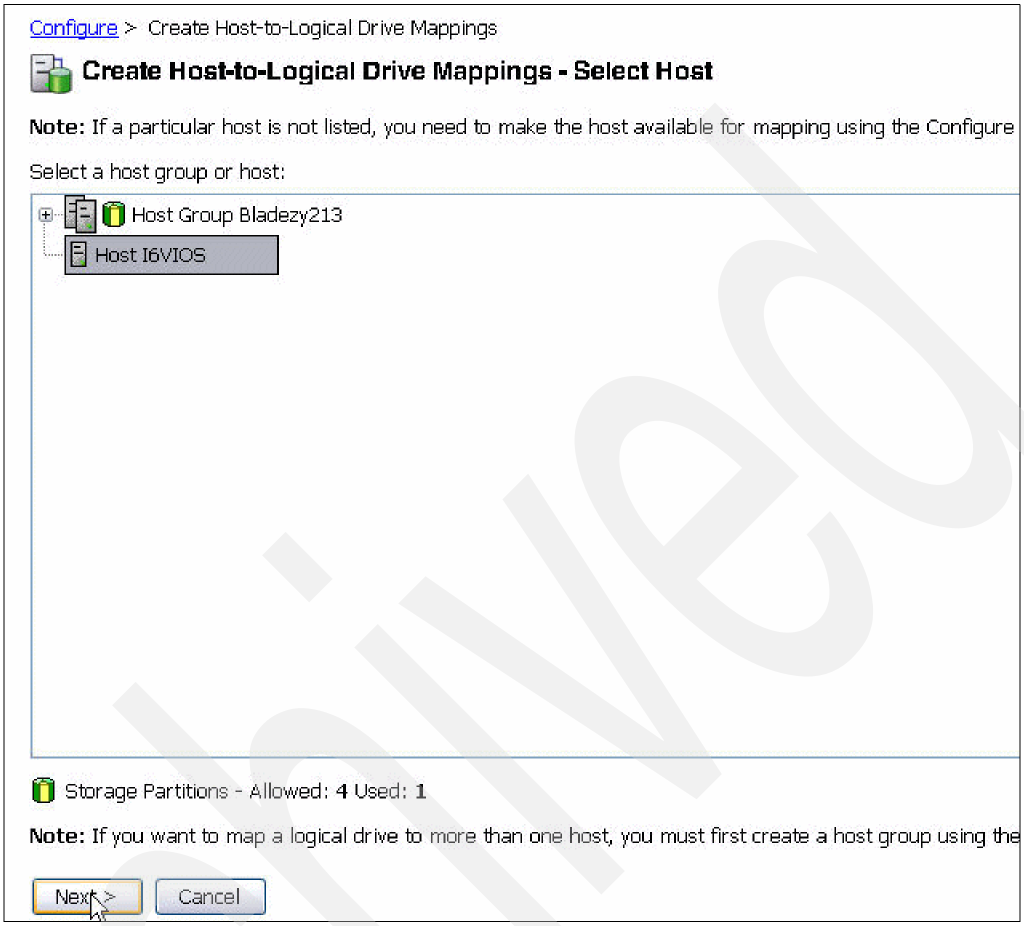


Figure 7-42 Select host name for logical drive mapping

3. In the wizard window Select Logical Drives we check the box **Select All Logical Drives** since we want to map all of them to the VIOS host name, and click **Finish**. See Figure 7-43.

   Note that if you map all logical drives by checking the Select All Logical Drives box they will be automatically assigned a LUN number starting with 0. But when mapping individual logical drives you can choose which LUN number they will receive.

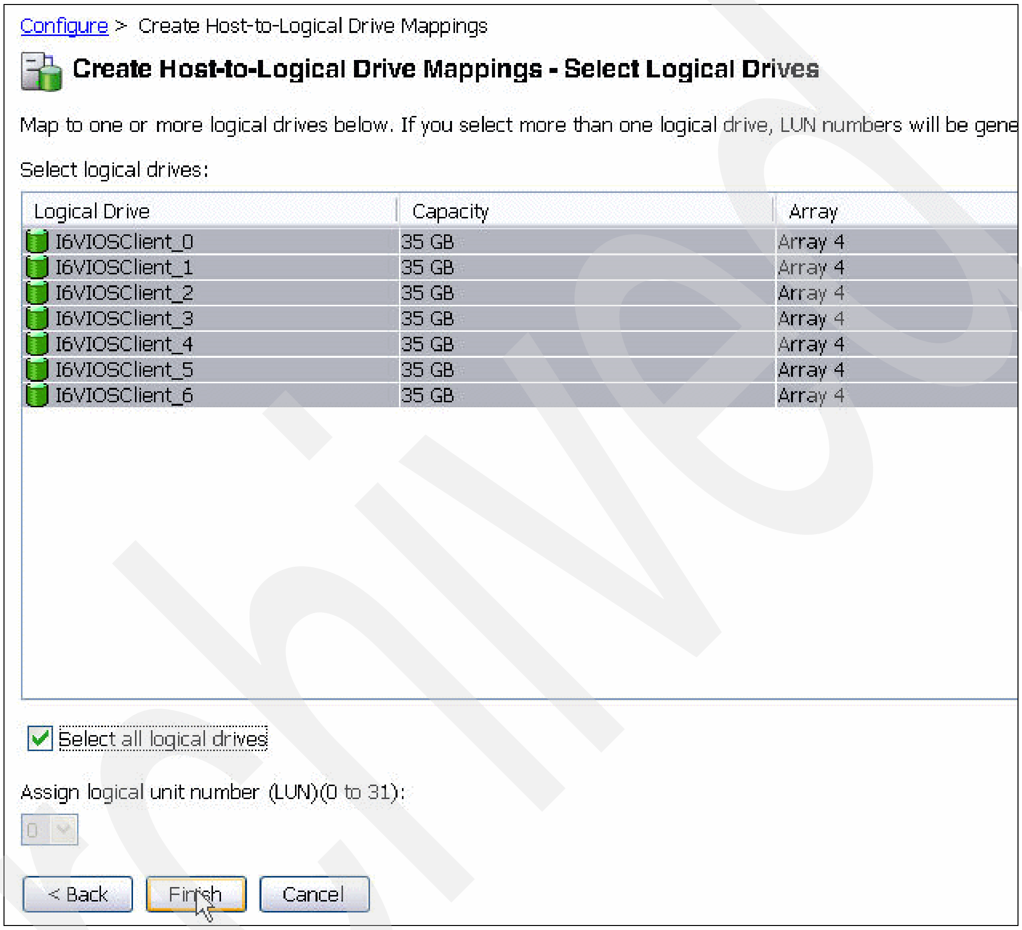


Figure 7-43 Select logical drives to map to host

4. We observe the progress of logical drives being mapped to the host in the wizard progress window. It also informs us of successful completion of mapping.

As explained in 7.3.4, “Creating hosts” on page 245, we added to the host name I6VIOS two ports each from a different FC adapter in VIOS. Each FC adapter is connected to one controller in DS3400. When creating logical drives we assigned three of them to controller A and 4 to controller B, as documented in 7.3.3, “Creating RAID arrays and logical drives” on page 235. We mapped all seven logical drives to host name I6VIOS.

With such connection and specifications we achieve balance of the I/O traffic to LUNs:

- I/Os to LUNs assigned to controller A will go via the adapter connected to this controller, and will fail over to controller B only when A is not operational.
- I/Os to LUNs assigned to controller B will primarily go via adapter connected to B. It will fail over to controller A only when B fails.

Such connection and assignment of a LUN are shown in Figure 4-2 on page 94.

7.4 DS4000/DS5000 Storage configuration with the GUI

In this section we introduce the logical storage configuration of an IBM System Storage DS4800 storage subsystem using the DS Storage Manager client graphical user interface for our IBM i client of VIOS setup as described in the configuration example in 6.1, “Setup example” on page 136.

7.4.1 Creating hot spares

To ensure that full redundancy of the RAID array is re-established after a drive failure as soon as possible before another drive failure may lead to data loss we first create a hot spare drive on the DS4000 storage subsystem:

1. We start the DS Storage Manager client by selecting from the Windows’ taskbar **Start** → **Programs** → **Storage Manager 10 Client** → **Storage Manager 10 Client** and double-clicking the DS4800 Storage System **iSeries_VIOS**, as shown in Figure 7-44.

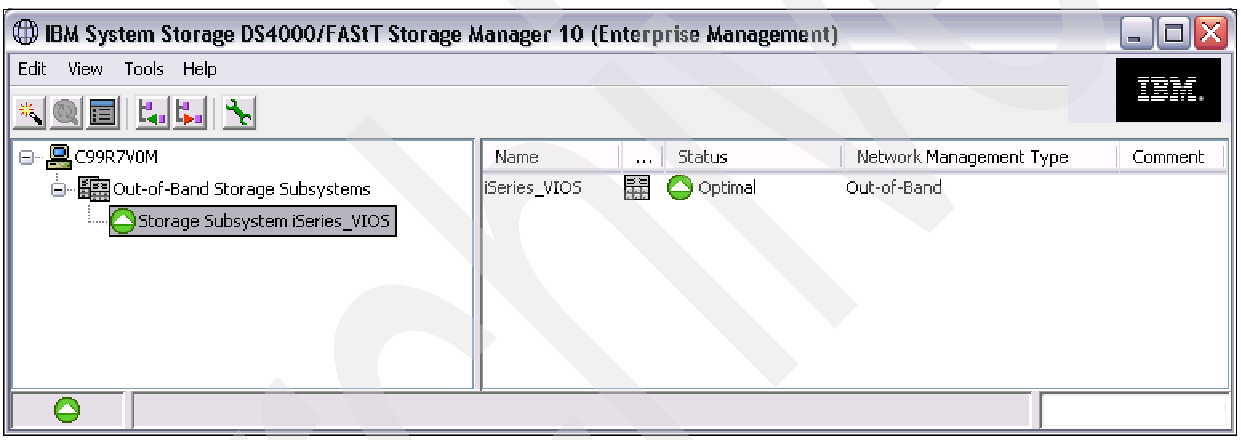


Figure 7-44   DS Storage Manager window

2. We close the Subsystem Management Window Task Assistant by selecting **Close**, as shown in Figure 7-45.

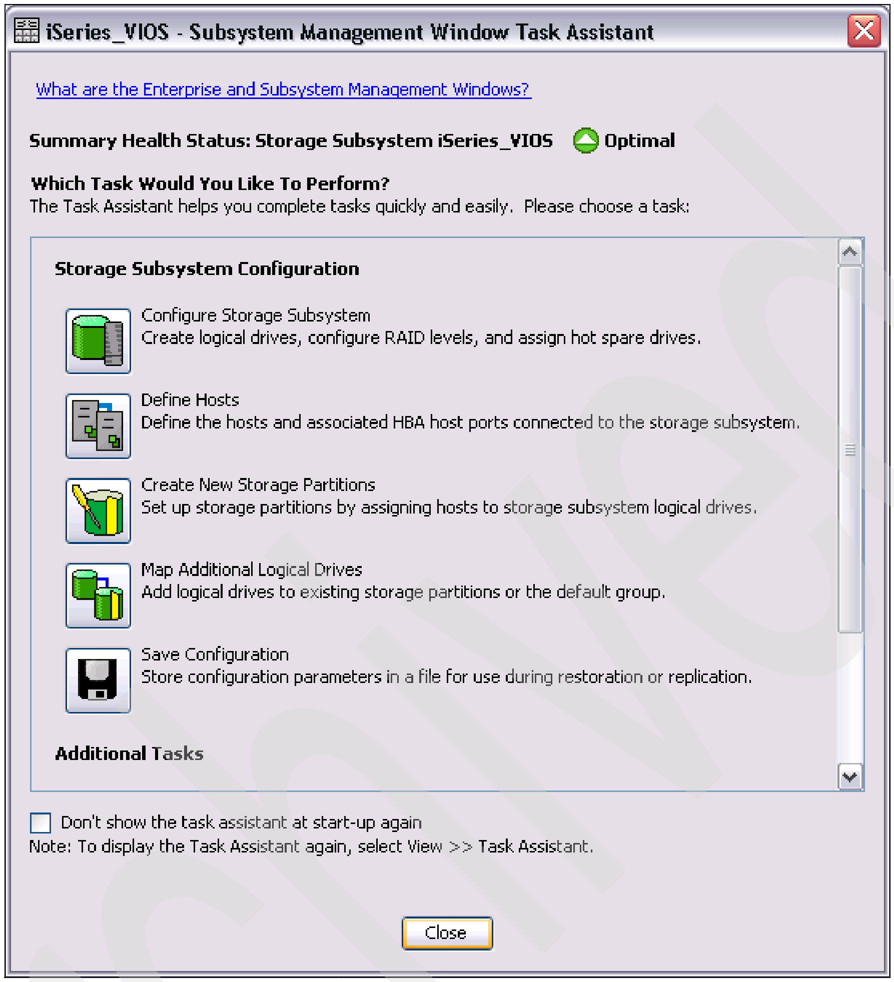


Figure 7-45 DS Storage Manager Task Assistant

3. In the DS Storage Manager Logical/Physical View of our storage subsystem we right-click the drive that we want to make a hot spare and select **Hot Spare Coverage** from the context menu, as shown in Figure 7-46.

> **Note:** If you have mixed drive capacities in the DS Storage System make sure to select a hot spare with equal (or larger) capacity of the RAID array drives that will be protected.

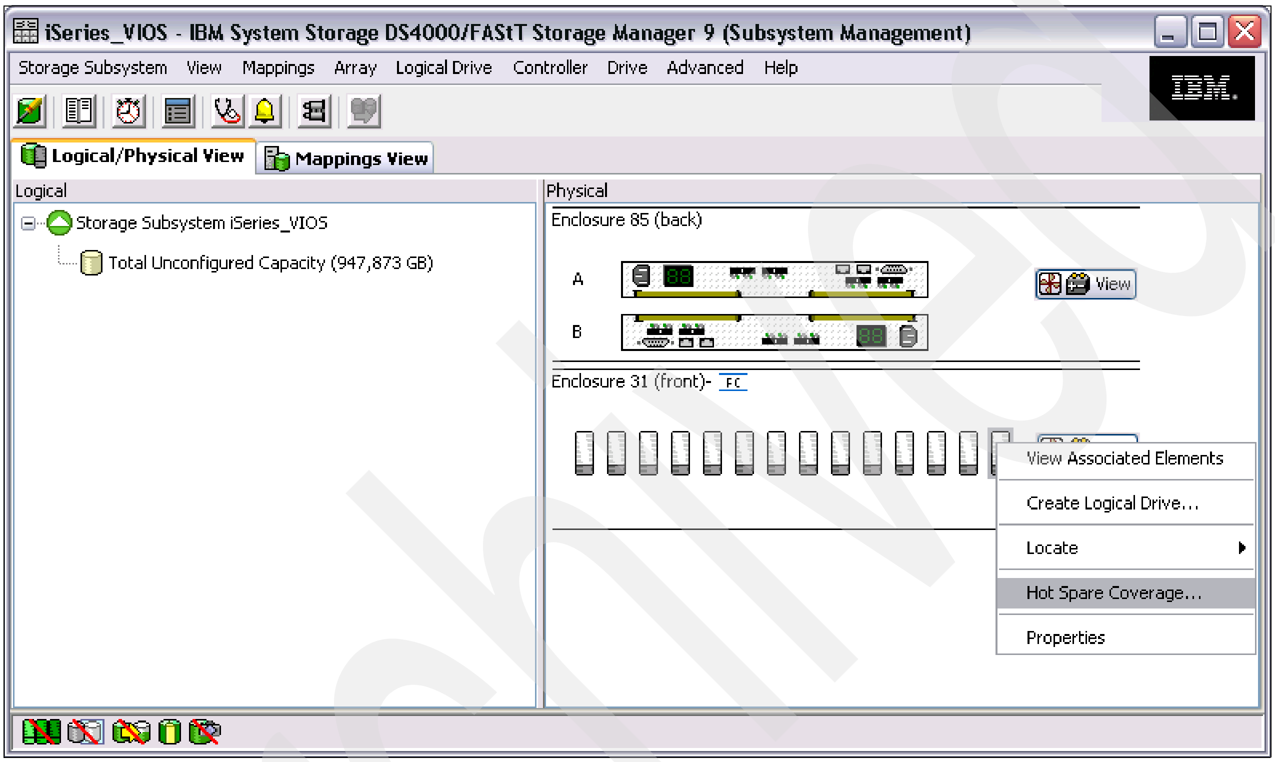


Figure 7-46 DS Storage Manager Subsystem Management Logical/Physical View

4. We accept the default option **Manually assign individual drives** and select **OK**, as shown in Figure 7-47.

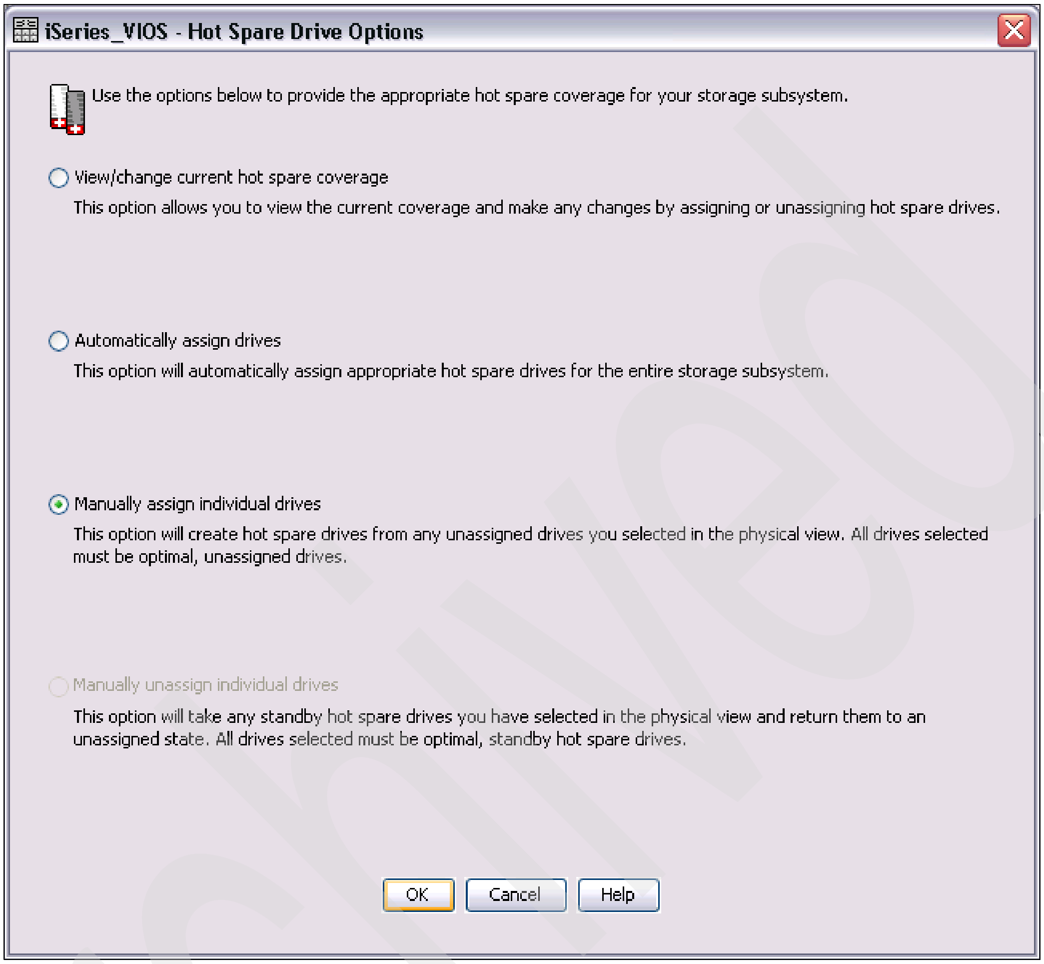


Figure 7-47 DS Storage Manager Hot Spare Drive Options

5. Our previously selected drive has now become a hot spare drive as indicated by the white on red cross, as shown in Figure 7-48.

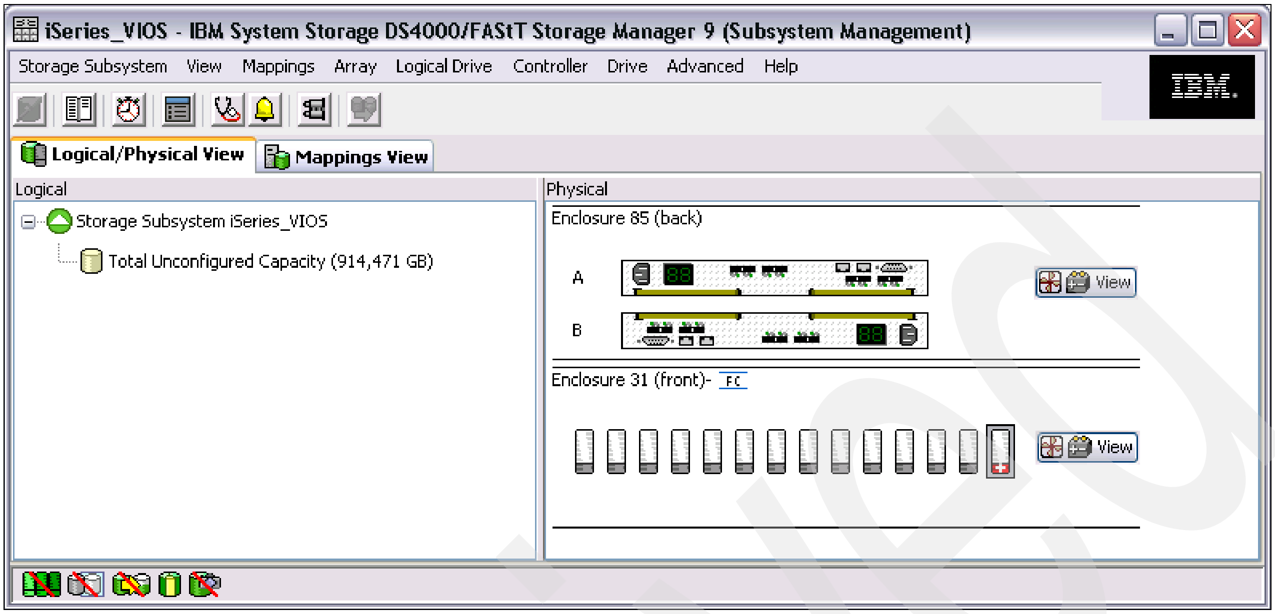


Figure 7-48 DS Storage Manager view after hot spare creation

After creating the hot spare drive we are now ready to create RAID arrays and logical volumes.

7.4.2 Creating RAID arrays and logical drives

To create a RAID-10 array and logical volumes for our IBM i client of VIOS we performed the following steps using the DS Storage Manager client GUI:

1. In the DS Storage Manager client we right-clicked **Total Unconfigured Capacity** of our storage subsystem and selected **Create Logical Drive**, as shown in Figure 7-49.

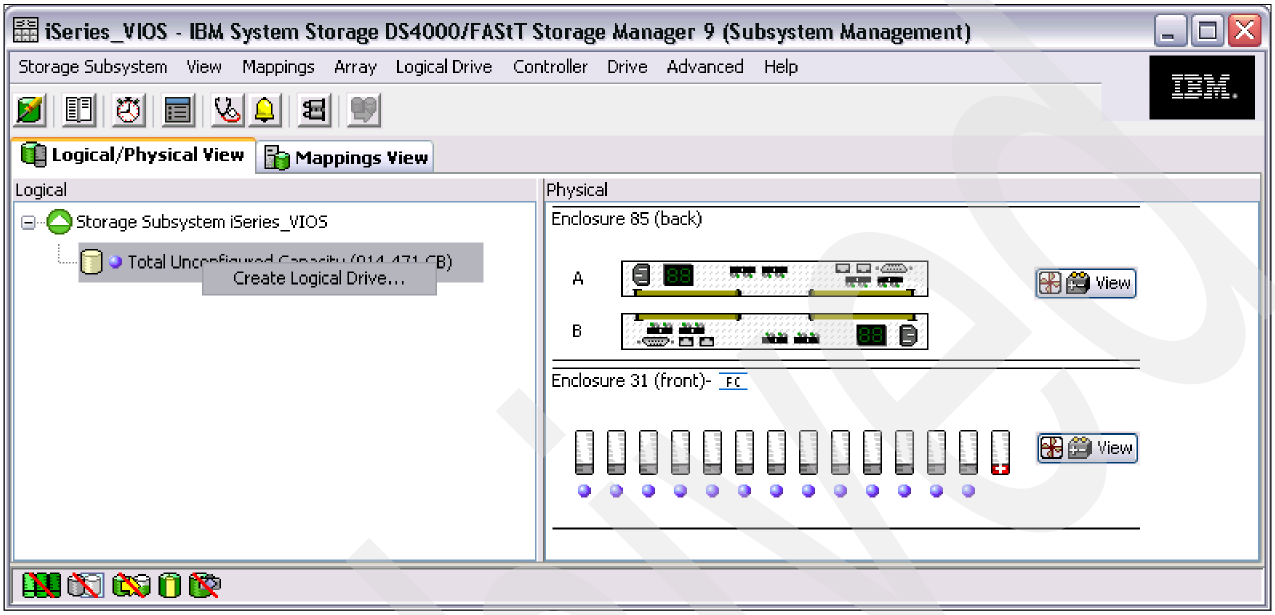


Figure 7-49 DS Storage Manager Create Logical Drive option

2. We selected **Next** in the Create Logical Drive dialog, as shown in Figure 7-50.

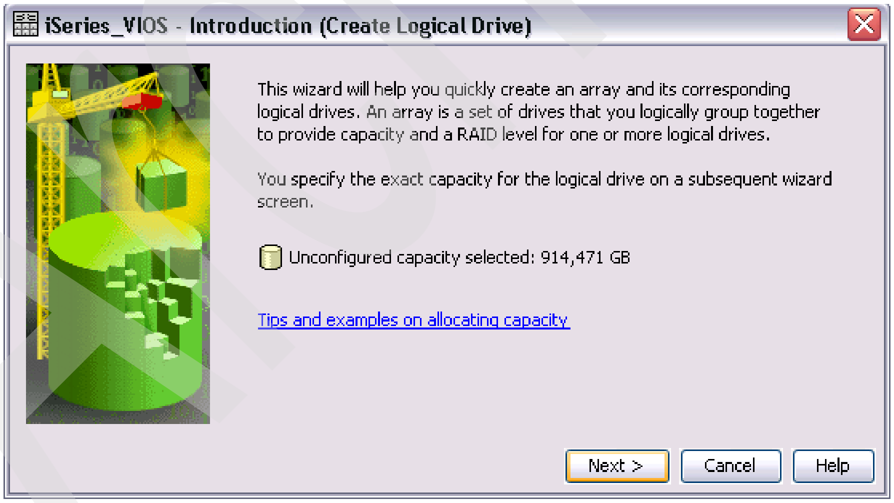


Figure 7-50 DS Storage Manager Create Logical Drive dialog

3. For the RAID level we select **RAID 1** from the drop-down list box and for Drive selection choices we choose **Manual - select drives to obtain array capacity**, as shown in Figure 7-51.

> **Note:** A RAID-1 array will automatically become a RAID-10 array as soon as more than one drive pair is added.

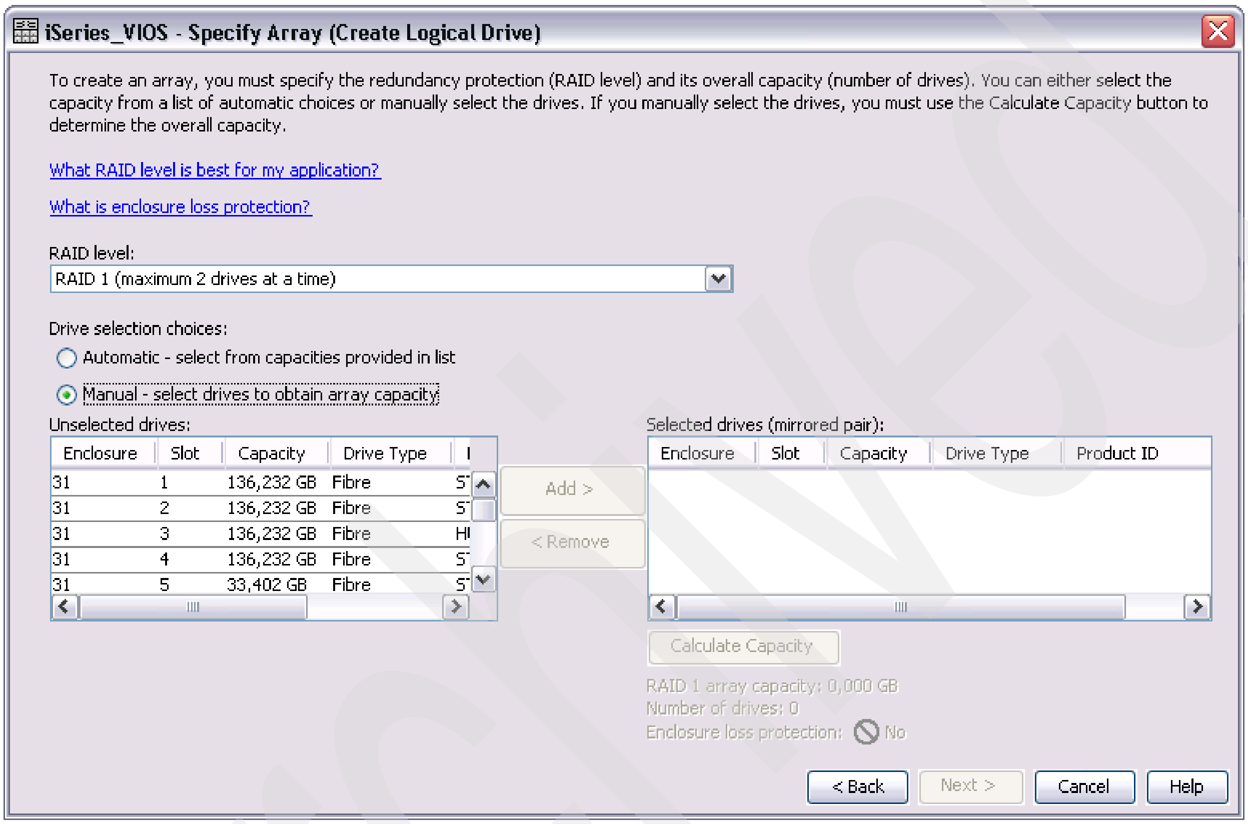


Figure 7-51 DS Storage Manager Specify Array dialog

4. We select the first two 136 GB drives for our RAID-10 array from the Unselected drives list and click **Add**, select the second two 136 GB drives, and click **Add** again. Having two mirrored pairs now in the Selected drives list we select **Calculate Capacity**, as shown in Figure 7-52.

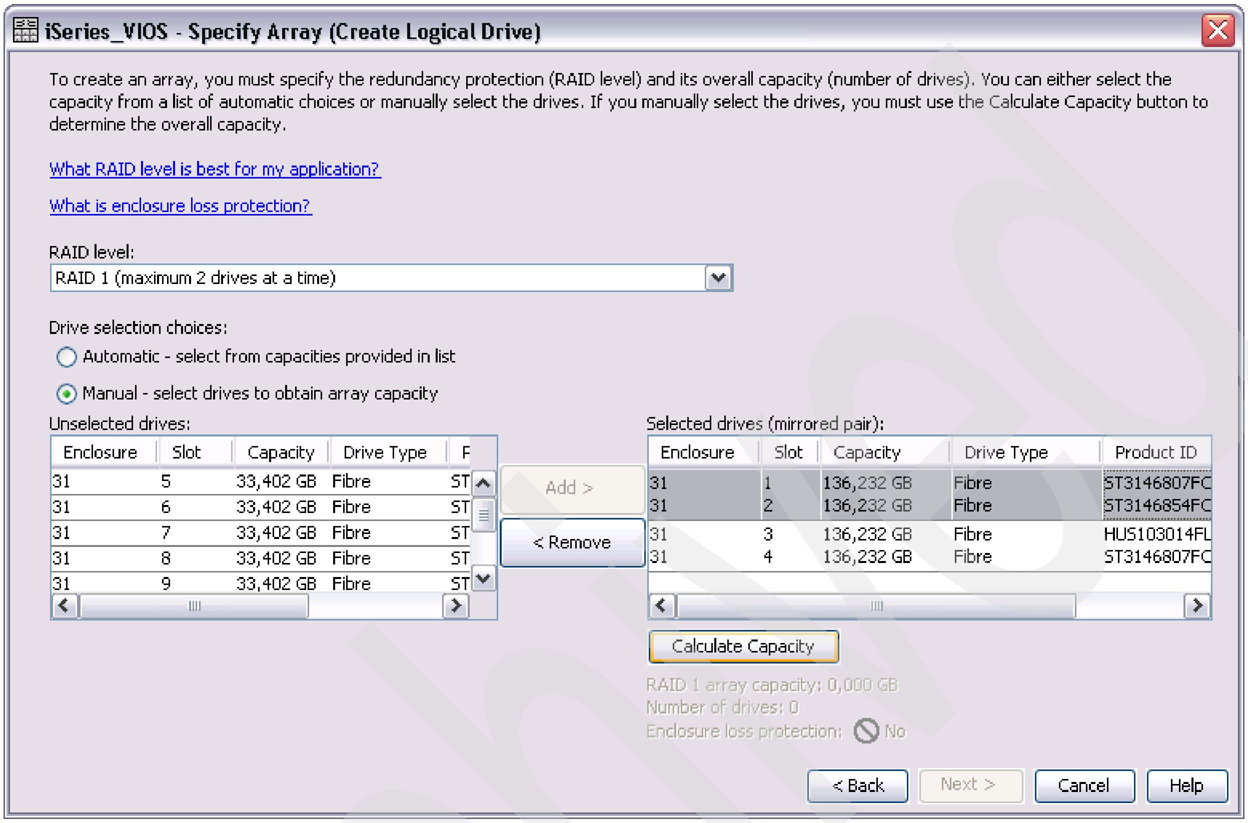


Figure 7-52 DS Storage Manager Create Logical Drive dialog after adding drives

5. We get the resulting capacity of our RAID-10 array of ~272 GB and click **Next** to proceed, as shown in Figure 7-53.

> **Note:** We generally recommend spreading the mirrored pairs across different drive enclosures to be protected against a potential enclosure failure. We had only one disk enclosure in our DS4800 so we could not do this.

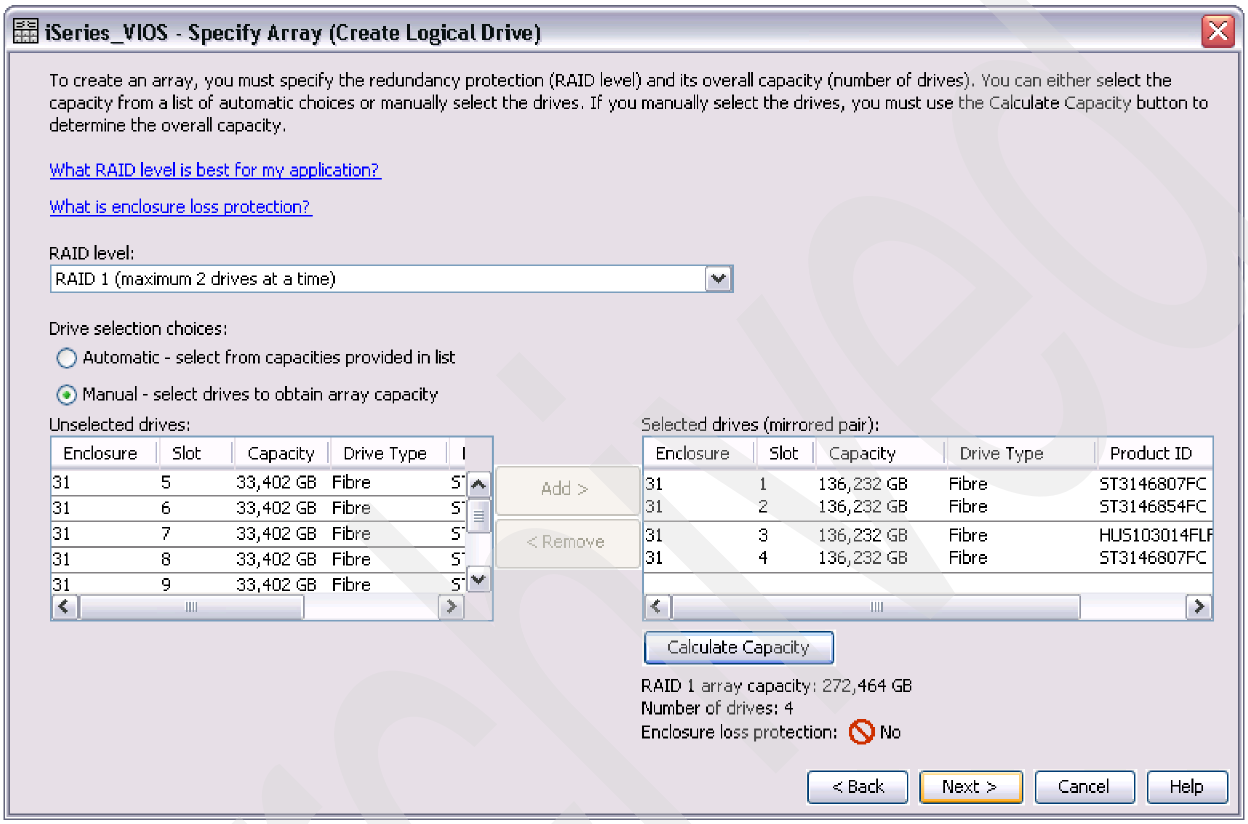


Figure 7-53 DS Storage Manager Create Logical Drive dialog after Calculate Capacity

6. We select **OK** in the Array Success dialog, as shown in Figure 7-54, to continue with creating logical drives.

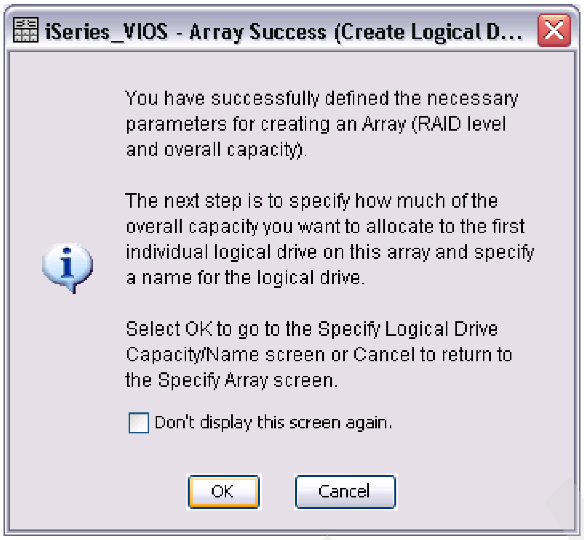


Figure 7-54 DS Storage Manager Array Success dialog

7. For the new logical drive to be created we specify a LUN size of 35 GB, provide it the name `i6VIOSclient` of our IBM i client partition, choose **Customize settings (I/O characteristics and controller ownership)**, and click **Next**, as shown in Figure 7-55.

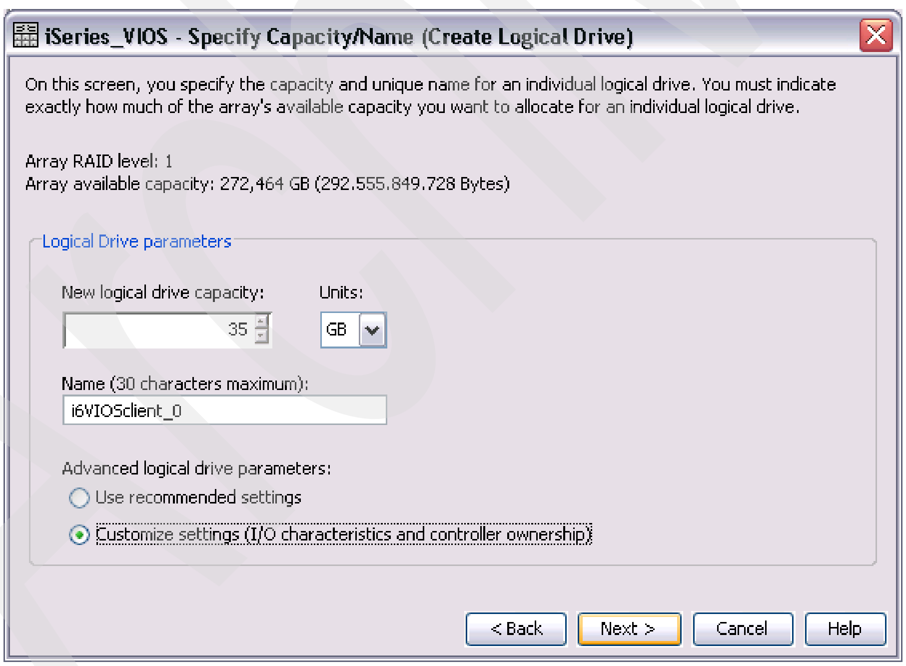


Figure 7-55 DS Storage Manager Create Logical Drive dialog

8. We choose **Custom** for the logical drive I/O characteristics type to select a segment (strip) size of **128 KB** and click **Next** to continue, as shown in Figure 7-56.

> **Notes:** For performance reasons we recommend using a segment size of 128 KB per drive as it helps to ensure that both an IBM i transaction type small block or a save/restore large block I/O transfer can be handled by a single drive only instead of multiple drives, which would need to be engaged for a single I/O transfer if the segment size would be smaller than the host transfer size (see also 4.5, “Planning considerations for performance” on page 97).
>
> The DS Storage Manager automatically alternates the proposed *preferred controller ownership* setting for the next logical drive being created to help make sure that the logical drives are evenly distributed across both DS4800 controllers A and B for balancing its I/O processing workload.

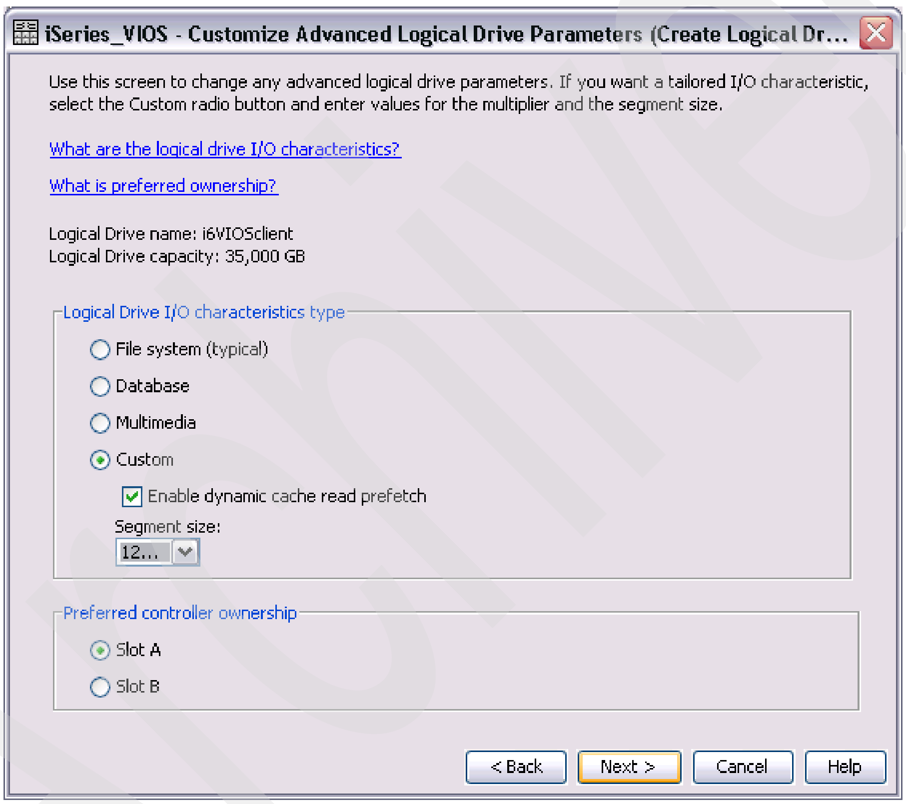


Figure 7-56 DS Storage Manager Customize Advanced Logical Drive Parameters dialog

9. In the Specify Logical Drive-to-LUN Mapping dialog we accept the default **Map later using the Mappings View** and click **Finish**, as shown in Figure 7-57.

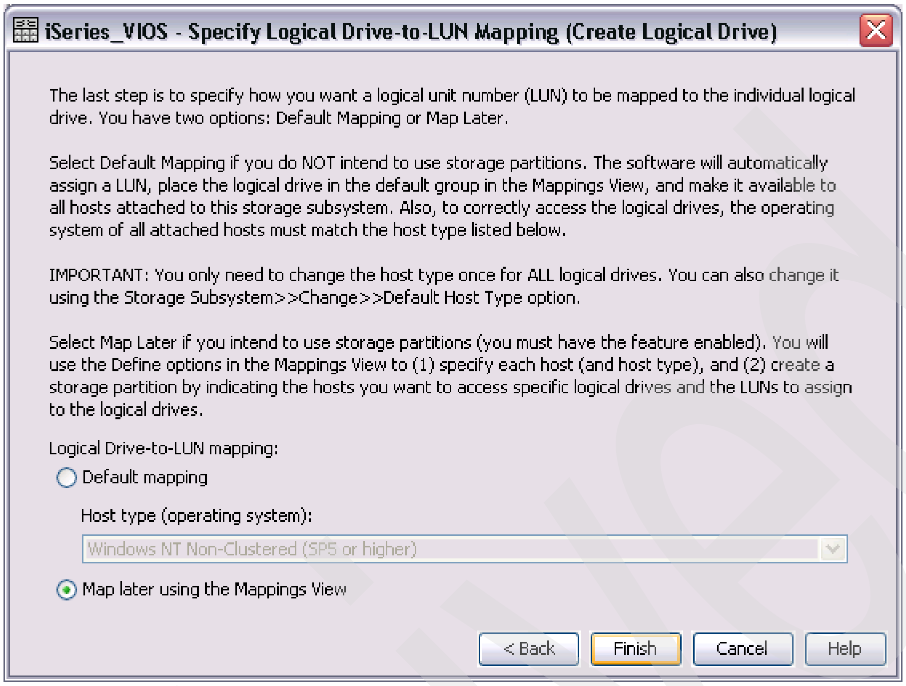


Figure 7-57 DS Storage Manager Specify Logical Drive-to-LUN Mapping dialog

10. Since we want to create more than one logical drive for our IBM i client we answer the question “Do you want to create another logical drive?” by selecting **Yes**, as shown in Figure 7-58.

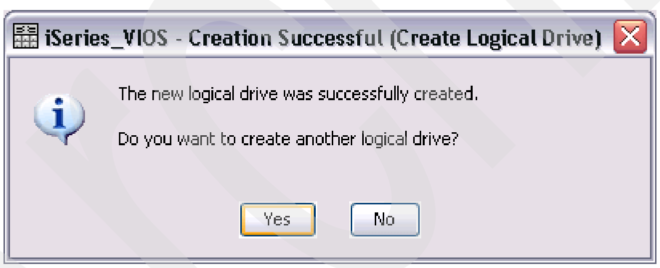


Figure 7-58 DS Storage Manager Creation Successful dialog

11. We choose **Free capacity on the same array (create individual logical drive)** and click **Next**, as shown in Figure 7-59, to continue creating logical volumes on our RAID-10 array following the same steps 7 to 10 that we did before for the previous logical drive that we created.

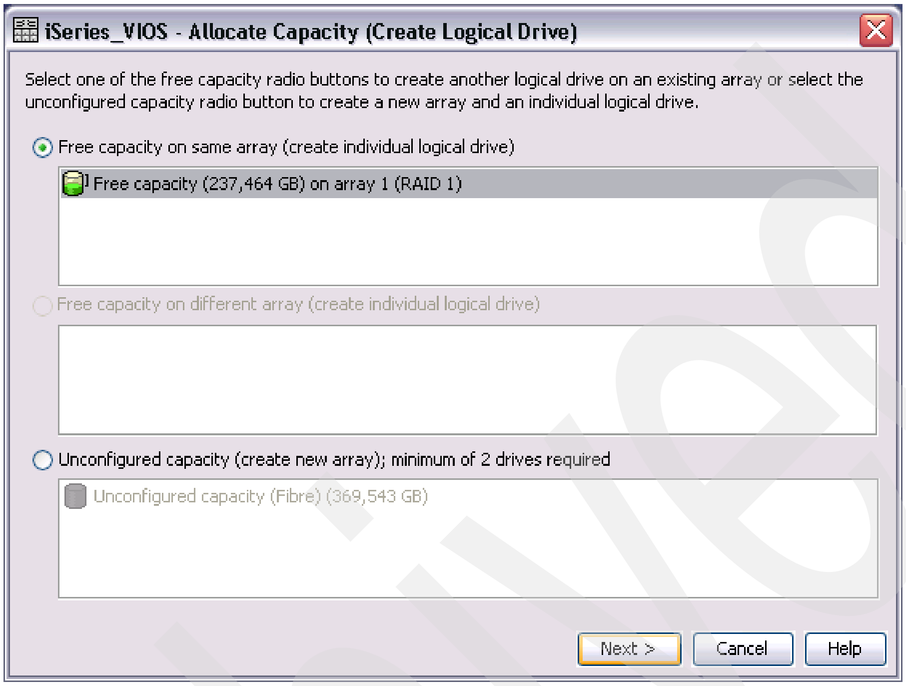


Figure 7-59 DS Storage Manager Allocate Capacity dialog

12. The logical drive configuration that we created for our IBM i client partition of VIOS (consisting of seven 35 GB logical drives on a RAID-10 array that are not mapped to a host yet) is shown in Figure 7-60.

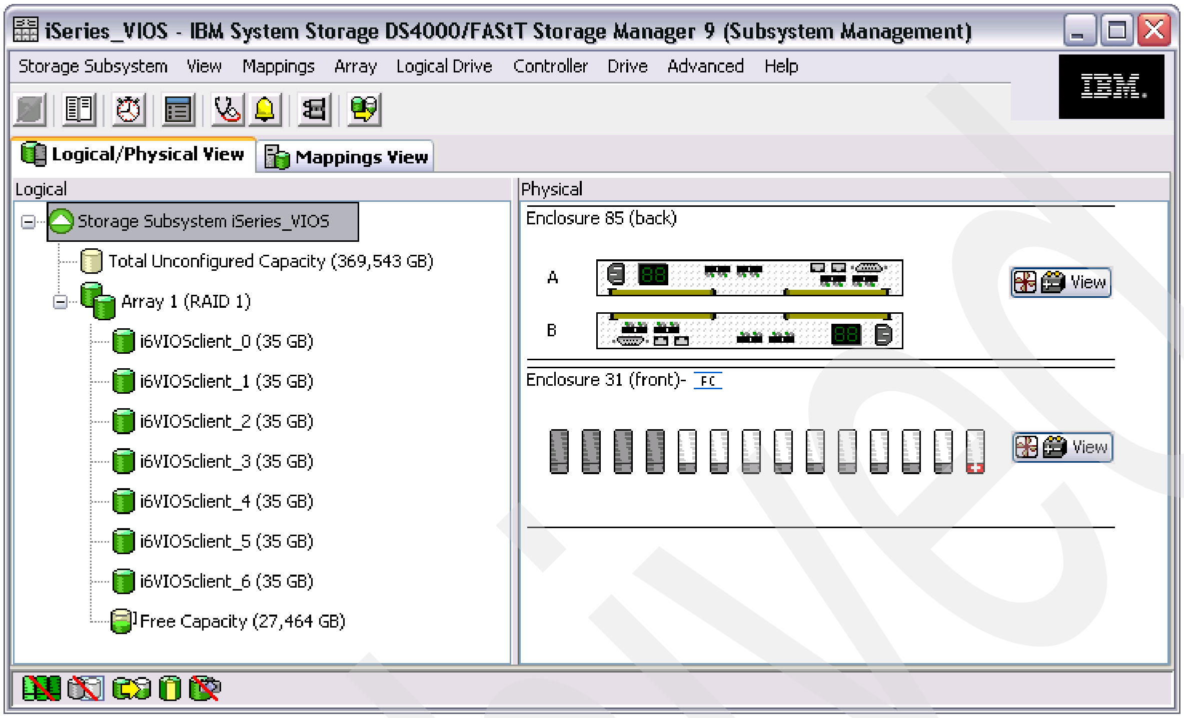


Figure 7-60 DS Storage Manager logical drive configuration

7.4.3 Creating hosts

In this section we describe how we created the host definitions on the DS4800 for our VIOS partition to allow us in a further step to map the logical drives to our VIOS host:

1. Before we create the host definition for the VIOS partition on the DS4800 Storage System we look for the world-wide port name (WWPN) information of the Fibre Channel adapter cards from our VIOS partition using the **lscfg -vl fcs*X*** command, as shown in Example 7-1.

Example 7-1 Displaying the FC adapter resources and configuration

```
$ lsdev -type adapter | grep fcs
fcs0            Available   4Gb FC PCI Express Adapter (df1000fe)
fcs1            Available   4Gb FC PCI Express Adapter (df1000fe)
fcs2            Available   4Gb FC PCI Express Adapter (df1000fe)
fcs3            Available   4Gb FC PCI Express Adapter (df1000fe)
$ lsdev -dev fcs0 -vpd
fcs0            U789D.001.DQDWVYP-P1-C3-T1  4Gb FC PCI Express Adapter (df1000fe)

        Part Number.................10N7255
        Serial Number...............1F7440CD4D
        Manufacturer................001F
        EC Level....................A
        Customer Card ID Number.....5774
        FRU Number.................. 10N7255
        Device Specific.(ZM)........3
        Network Address.............10000000C96DA75E
        ROS Level and ID............02E82774
```

```
        Device Specific.(Z0)........2057706D
        Device Specific.(Z1)........00000000
        Device Specific.(Z2)........00000000
        Device Specific.(Z3)........03000909
        Device Specific.(Z4)........FFE01231
        Device Specific.(Z5)........02E82774
        Device Specific.(Z6)........06E12715
        Device Specific.(Z7)........07E12774
        Device Specific.(Z8)........20000000C96DA75E
        Device Specific.(Z9)........ZS2.71X4
        Device Specific.(ZA)........Z1F2.70A5
        Device Specific.(ZB)........Z2F2.71X4
        Device Specific.(ZC)........00000000
        Hardware Location Code......U789D.001.DQDWVYP-P1-C3-T1


  PLATFORM SPECIFIC

  Name:  fibre-channel
    Model:  10N7255
    Node:  fibre-channel@0
    Device Type:  fcp
    Physical Location: U789D.001.DQDWVYP-P1-C3-T1
```

Our example shows four FC adapter resources (fcs0, fcs1, fcs2, and fcs3), one for each FC port of our two dual-port PCIe FC adapters that we assigned to our VIOS partition. The upper FC port of a card, port0, is designated with a suffix of T" in the Hardware Location Code field and its WWPN is shown under the Network Address field.

2. With knowing the WWPN information from our two dual-port PCIe FC adapters in VIOS (in our configuration example we only use port0 of each card) we are ready to create the host definition on the DS4800 by selecting the **Mappings View** for our storage subsystem in the DS Storage Manager GUI, as shown in Figure 7-61.

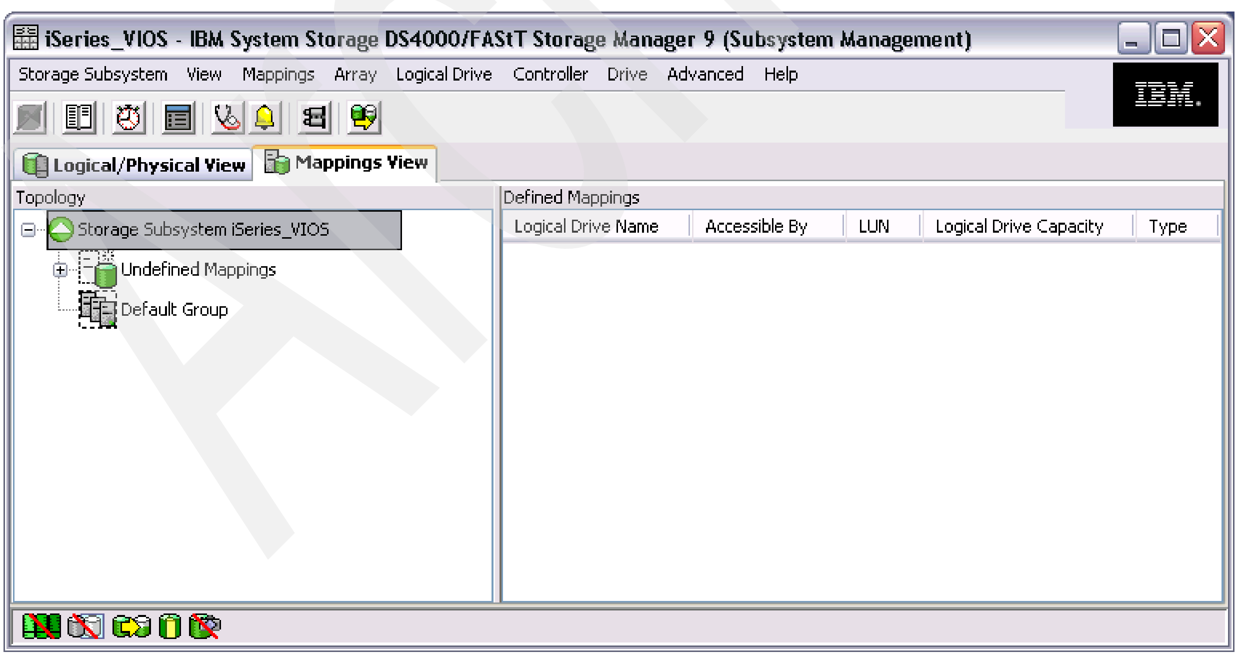


Figure 7-61 DS Storage Manager Mappings View

3. The Mapping Start-Up Help dialog appears, which we close by selecting **Close**, as shown in Figure 7-62.

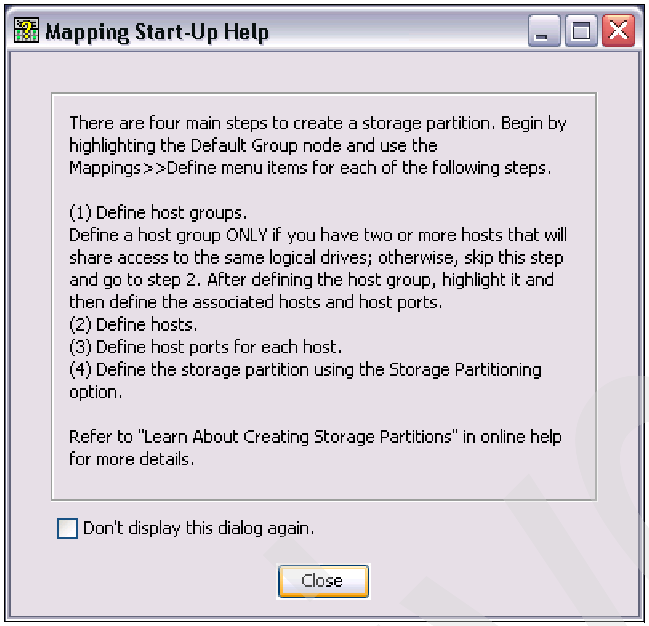


Figure 7-62 DS Storage Manager Mapping Start-Up Help dialog

4. We right-click our Storage System and select **Define Host** from the context menu, as shown in Figure 7-63.

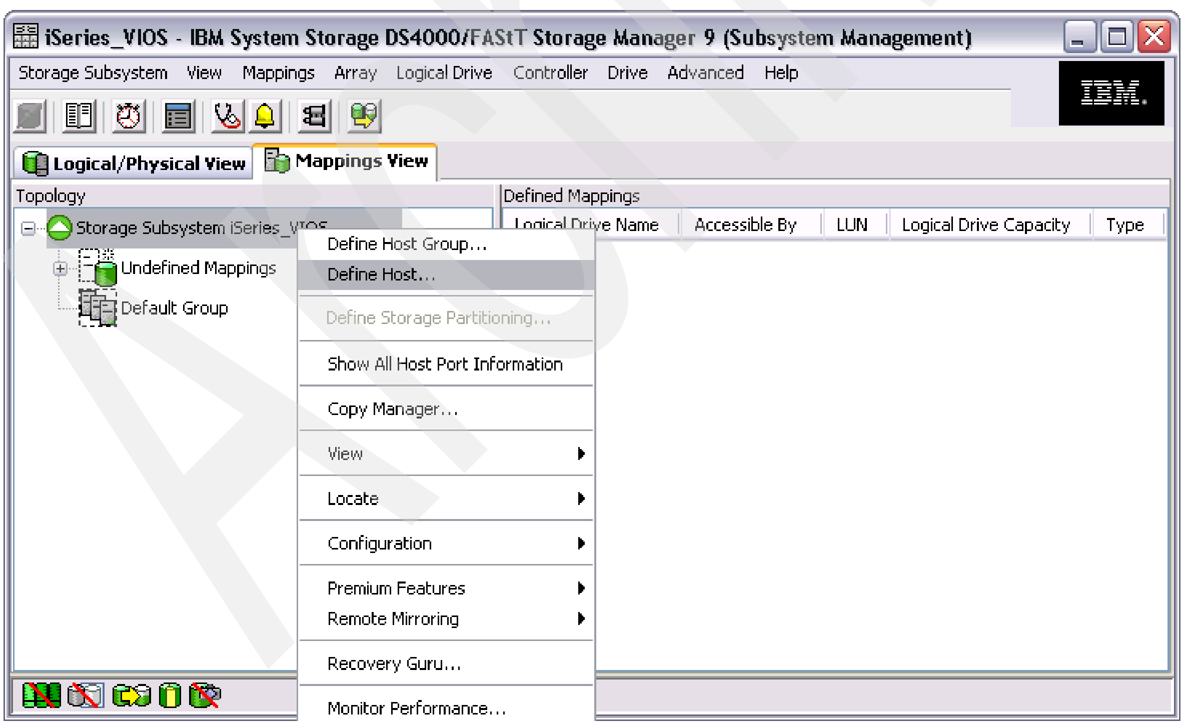


Figure 7-63 DS Storage Manager Define Host context menu

5. We accept the default setting of **Yes** for planning to use storage partitions on our storage subsystem and select **Next** to proceed, as shown in Figure 7-64.

> **Note:** A *storage partition* is a logical entity created by mapping logical drives to LUNs and describes which host or host group can access them. Without storage partitioning all logical drives configured on the DS Storage System would be accessible by any attached host.

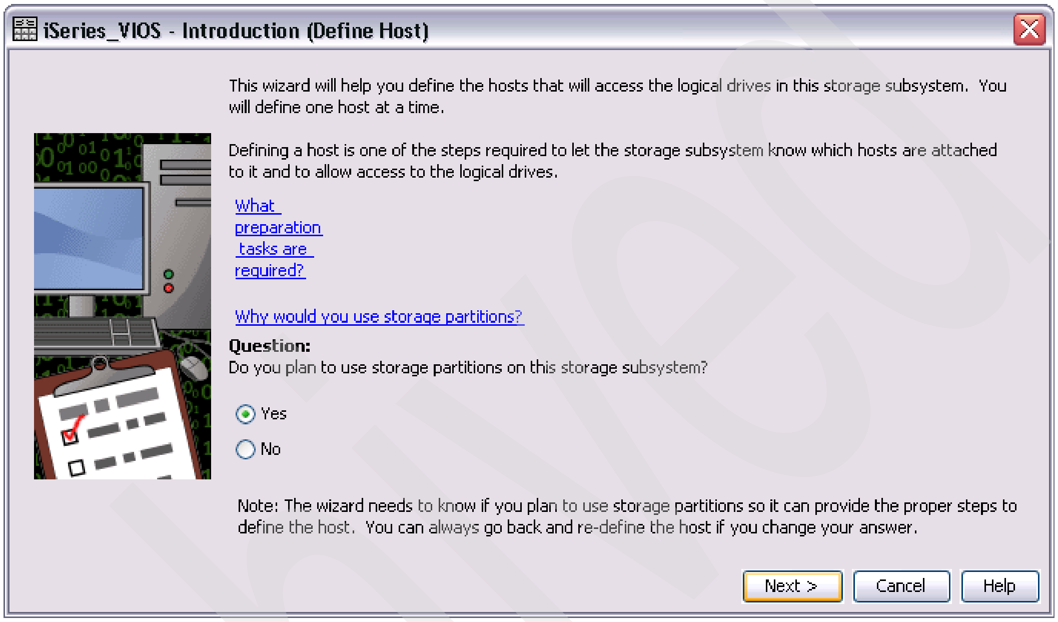


Figure 7-64 DS Storage Manager Introduction (Define Host) dialog

6. We enter the host name `i6VIOS` for our VIOS partition, select the two Fibre Channel adapter WWPNs from our VIOS partition in the Known HBA host port identifiers list box, and click **Add** to proceed, as shown in Figure 7-65.

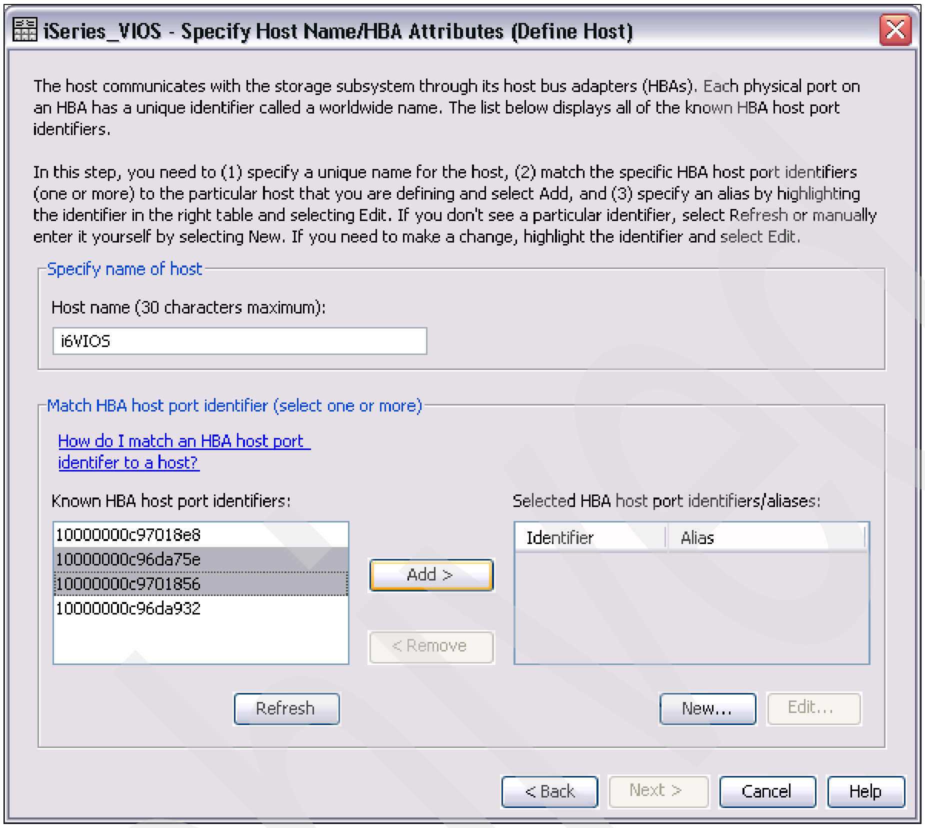


Figure 7-65 DS Storage Manager Specify Host Name/HBA Attributes dialog

7. Our selected WWPNs from our VIOS partition got added to the Selected HBA host port identifiers/aliases list box and we select **Edit** to proceed, as shown in Figure 7-66.

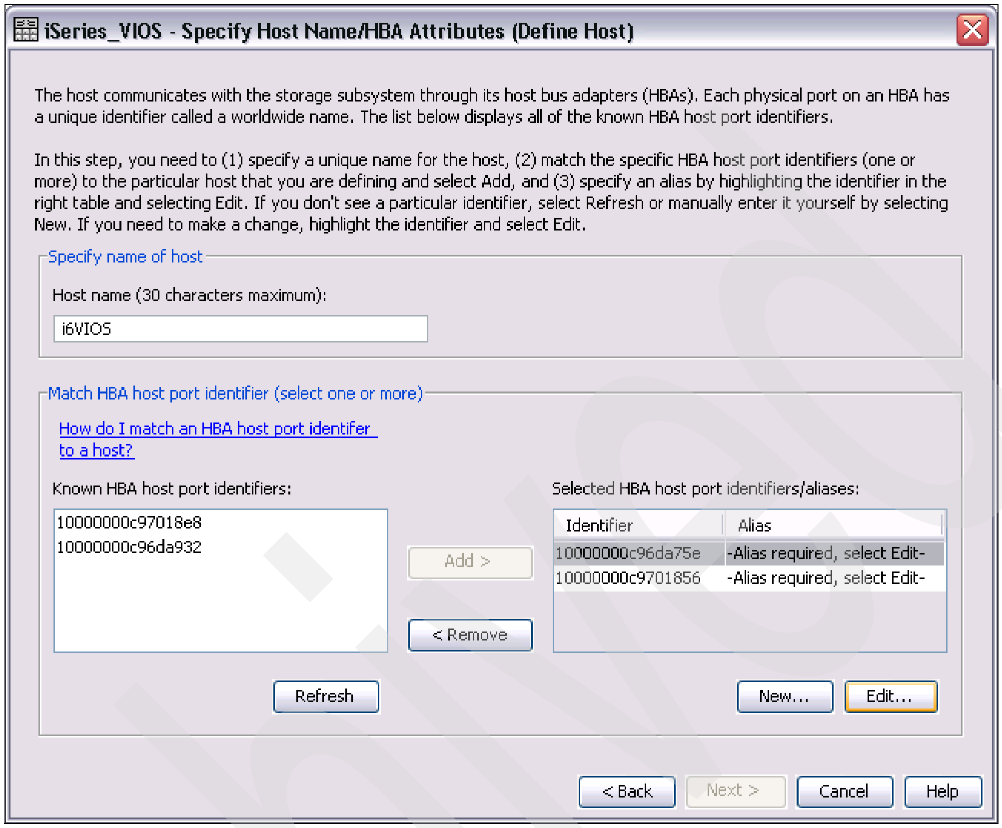


Figure 7-66 DS Storage Manager Specify Host Name/HBA Attributes dialog after adding WWPNs

8. We specify an alias name for each of our VIOS partition's host ports using the hostname and adding a suffix describing the physical adapter location information that we retrieved earlier from the lsdev output in step 1, as shown in Figure 7-67.

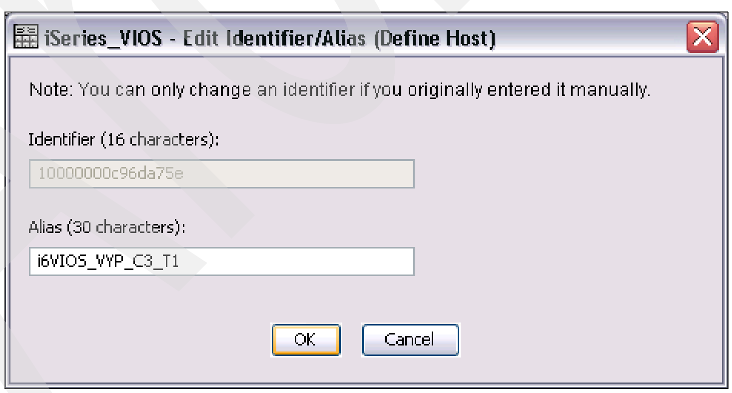


Figure 7-67 DS Storage Manager Edit Identifiers/Aliases (Define Host) dialog

9. After having defined the aliases for both of our VIOS partition’s host ports we select **Next** to proceed, as shown in Figure 7-68.

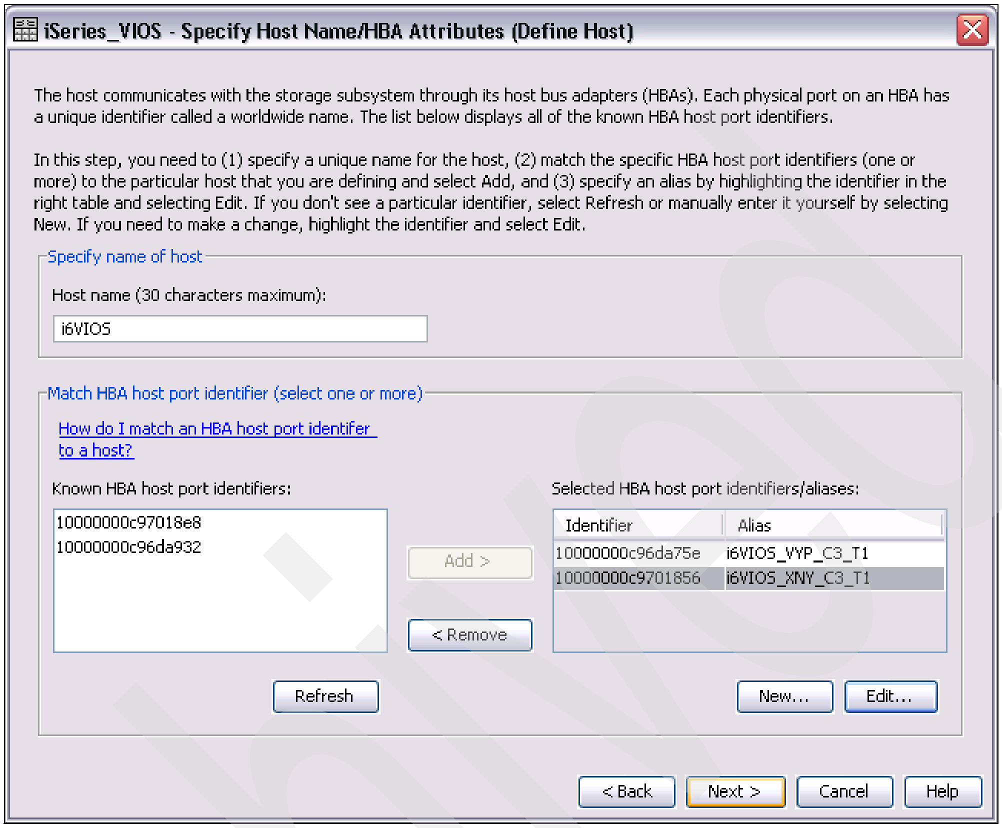


Figure 7-68 DS Storage Manager Specify Host Name/HBA Attributes dialog after aliases definition

10. For our VIOS partition we choose **AIX** as the host type and click **Next**, as shown in Figure 7-69.

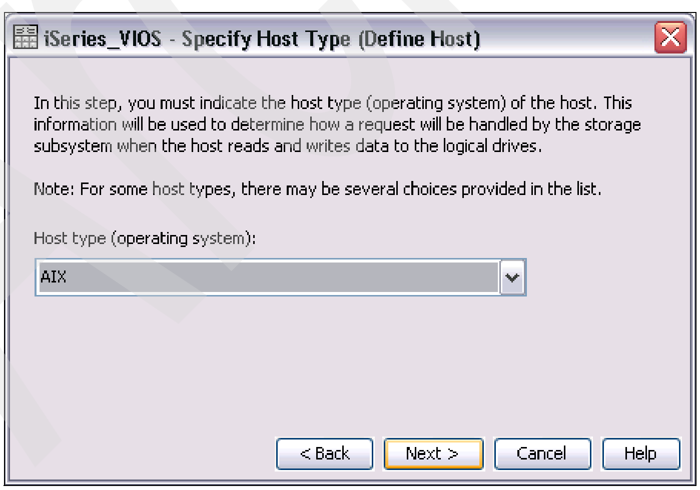


Figure 7-69 DS Storage Manager Specify Host Type (Define Host) dialog

11. In the Host Group Question dialog we select **No - This host will NOT share access to the same logical drives with other hosts** and click **Next** to proceed, as shown in Figure 7-70.

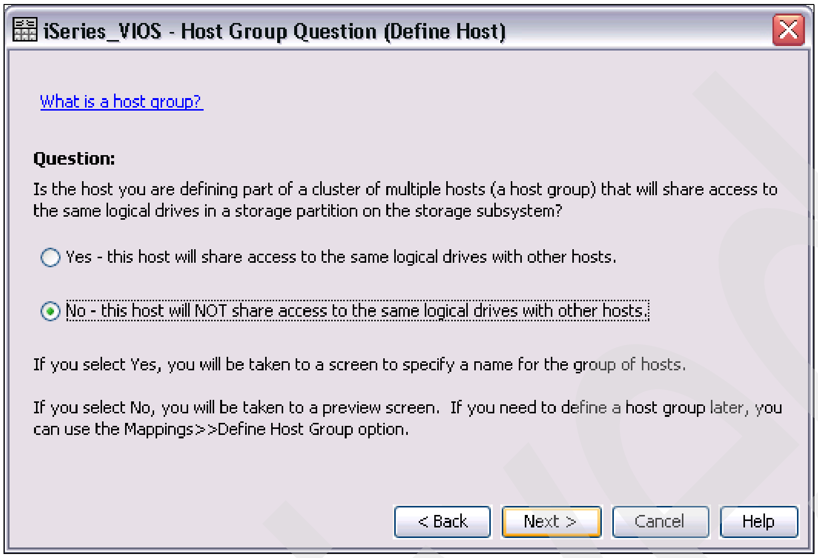


Figure 7-70 DS Storage Manager Host Group Question dialog

12. After reviewing the host configuration in the Preview dialog we click **Finish** to finally create the host definition for our host i6VIOS, as shown in Figure 7-71.

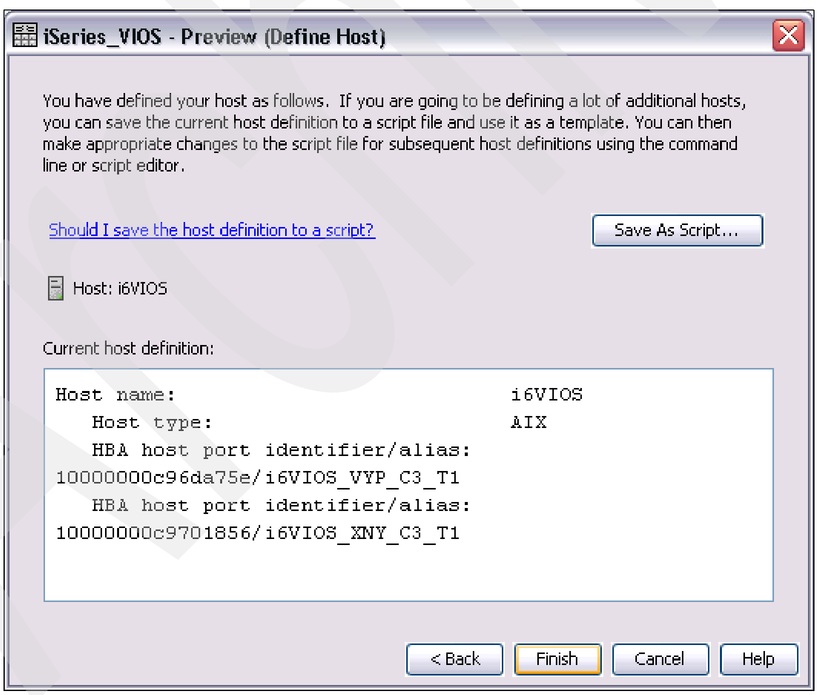


Figure 7-71 DS Storage Manager Preview (Define Host) dialog

13. Since we do not want to create another host definition at this time we choose **No** in the confirmation dialog about the newly created host definition, as shown in Figure 7-72.

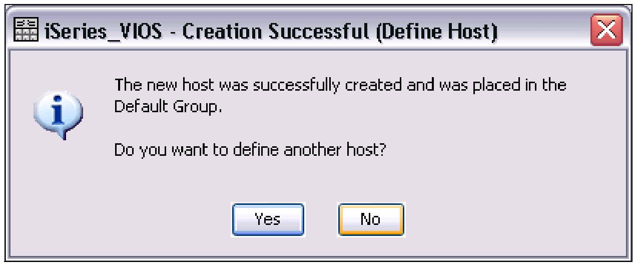


Figure 7-72 DS Storage Manager Creation Successful (Define Host) dialog

14. The newly created host definition for our VIOS host i6VIOS with its two Fibre Channel adapter ports is now listed under the Default Group in the DS Storage Manager Mappings View, as shown in Figure 7-73.

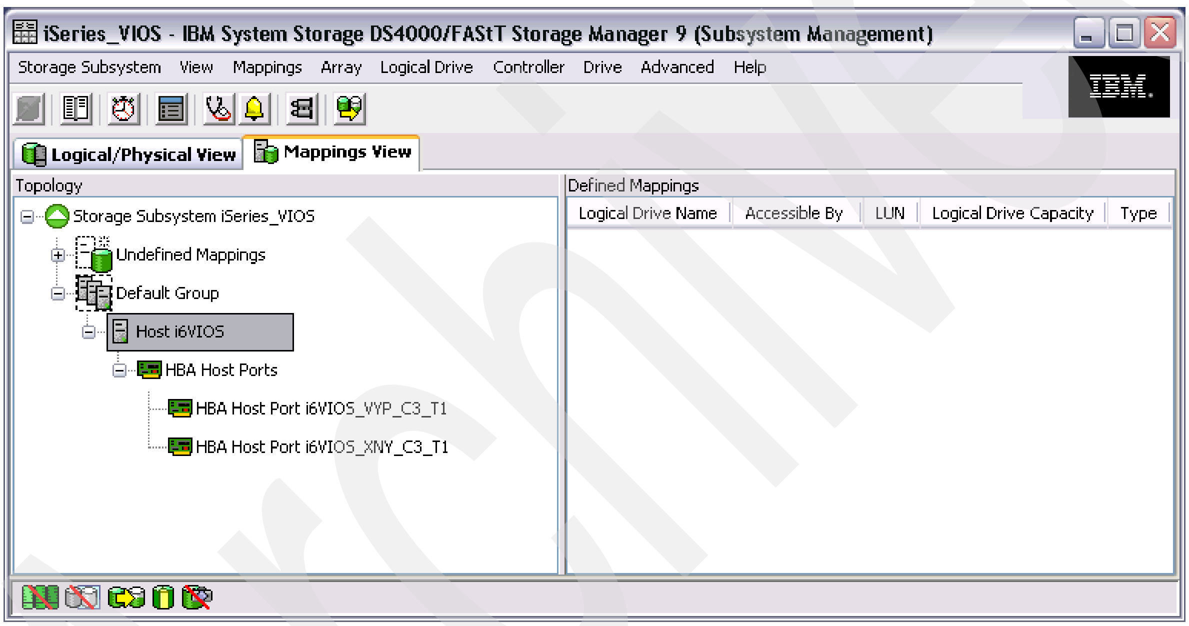


Figure 7-73 DS Storage Manager Mappings View after host definition

After having created the host definition for our VIOS host on the DS4800 we are now prepared to continue mapping the logical drives to LUNs, as described in 7.4.4, “Defining logical drive to LUN mapping” on page 274.

7.4.4 Defining logical drive to LUN mapping

In this section we describe the steps that we performed using the DS Storage Manager client to map the logical drives to LUNs for making them accessible to our previously defined VIOS host.

1. We perform the initial logical drives to LUN mapping by using the Storage Partitioning wizard, which we invoke by right-clicking our VIOS host in the default group of the DS Storage Manager Mappings View and selecting **Define Storage Partitioning** from the context menu, as shown in Figure 7-74.

> **Note:** If we wanted to add additional logical drives to our VIOS host after it got included already in a storage partition we would choose the option **Add additional mapping** instead.

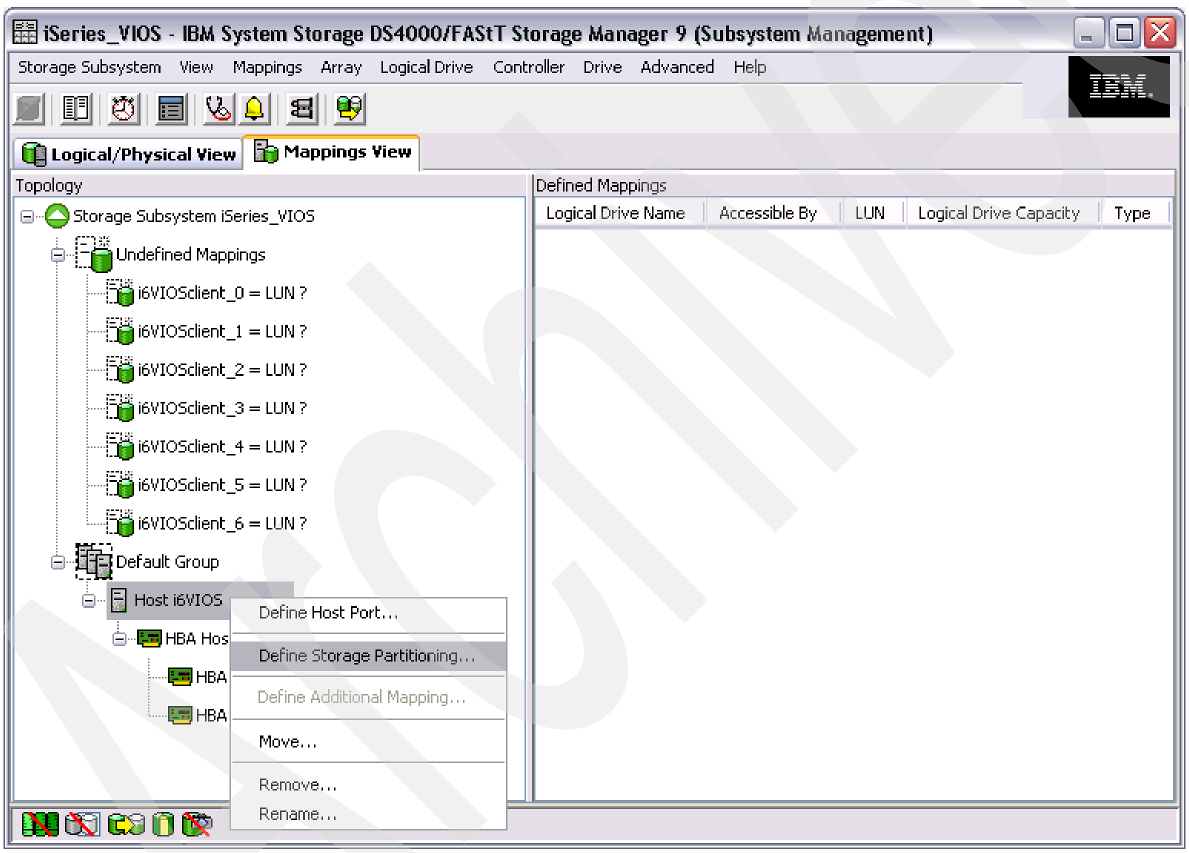


Figure 7-74 DS Storage Manager Mappings View Define Storage Partitioning context menu

2. We select **Next** to proceed in the Storage Partitioning Wizard - Introduction dialog, as shown in Figure 7-75.

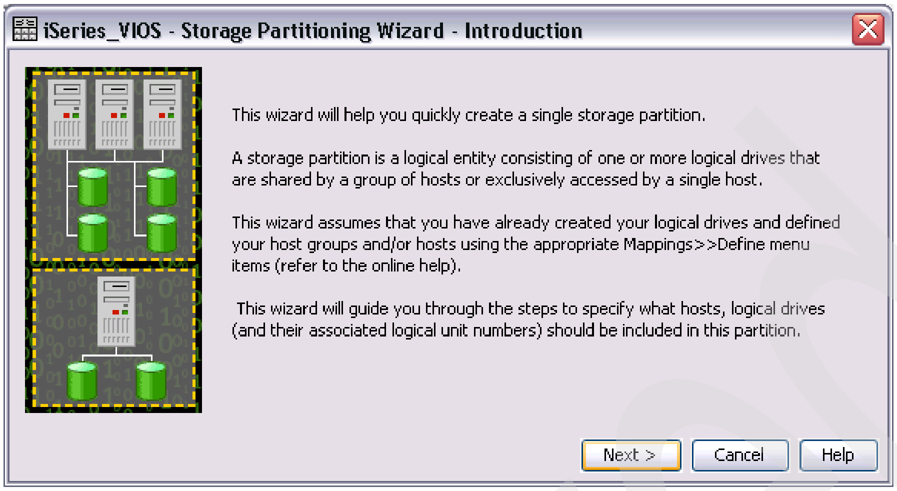


Figure 7-75 DS Storage Manager Storage Partitioning Wizard - Introduction dialog

3. Since our i6VIOS host is the only host configured so far it got pre-selected by the wizard so we do not need to select it and can simply click **Next** again to proceed, as shown in Figure 7-76.

Figure 7-76 DS Storage Manager Storage Partitioning Wizard - Select Host Group or Host dialog

4. We map the first logical drive i6VIOSclient_0 as LUN 0 to our i6VIOS host by selecting this logical drive, accepting the default **LUN 0** selection, and clicking **Add**, as shown in Figure 7-77.

> **Note:** We do not map the access LUN for in-bound management of the DS Storage System to our VIOS host as VIOS does not support any software installation. We decided to use a Windows host for out-band management, as described in 7.1, “Installation of DS Storage Manager Client” on page 214.

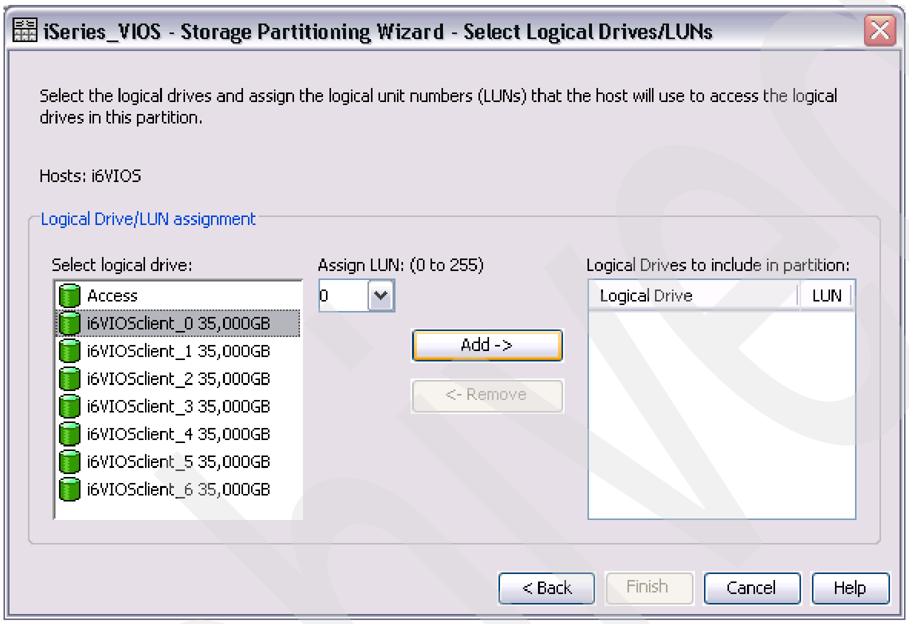


Figure 7-77 DS Storage Manager Storage Partitioning Wizard - Select Logical Drives/LUNs

5. We repeat step 4 above and map the six remaining logical drives by using the default next-available LUN number until we have all seven logical drives mapped to a LUN and select **Finish**, as shown in Figure 7-78.

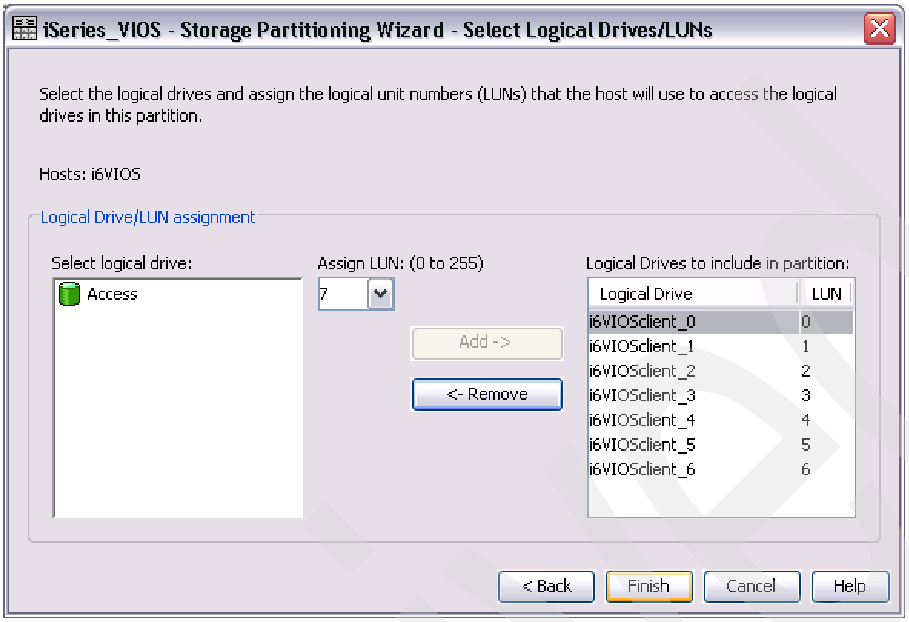


Figure 7-78 DS Storage Manager Storage Partitioning Wizard after mapping all host LUNs

6. After successful completion of the logical drives to LUN mapping we click **OK** to close the Define Storage Partitioning Wizard and see our i6VIOS host now removed from the Default Group, similar to the logical drives, which are removed now from the Undefined Mappings view.

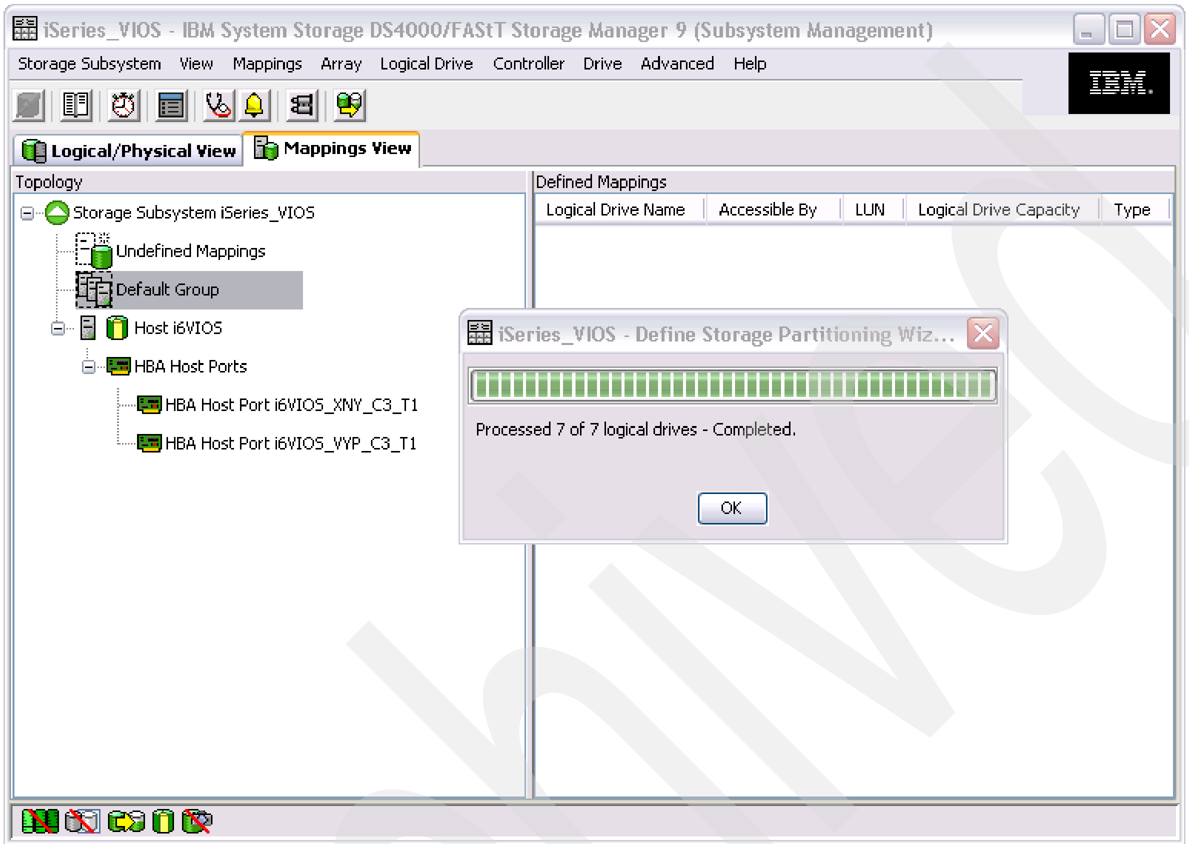


Figure 7-79 DS Storage Manager Mappings View after storage partitioning

At this point we are done with the logical storage configuration on the DS4800 and can proceed with mapping our newly defined DS4800 LUNs to virtual SCSI LUNs for the IBM i client on our VIOS host, as described in 6.5, “Configuring VIOS virtual devices” on page 191.

7.5 DS storage configuration with SMcli

IBM officially recommends using the DS Storage Manager client GUI for configuring the IBM System Storage DS Storage System (see 7.4, “DS4000/DS5000 Storage configuration with the GUI” on page 252, and 7.4, “DS4000/DS5000 Storage configuration with the GUI” on page 252) because in contrast to the Storage Manager command-line interface (SMcli) it provides precaution mechanisms to warn you about making unwanted changes to your configuration that may lead to erroneous data loss.

We still like to introduce you to using SMcli for configuring the IBM Midrange External Storage Systems, especially since you may decide that you like to create SMcli scripts for re-occurring configuration changes or CopyServices functions (see Chapter 8, “Using Midrange Storage Copy Services” on page 283).

Invoking the Storage Manager command-line interface

SMcli is invoked from its installation directory, which for a Windows Storage Manager client would be C:\Program Files\IBM_DS4000\client using the following syntax:

```
SMcli storageSubsystem parameters script-commands;
```

- *storageSubsystem* denotes the names or IP addresses of both DS storage subsystem controllers.
- *script-commands* denotes one or more storage configuration commands or a script text file including storage configuration commands.

SMcli can be used in three different modes:

- *Interactive mode*, which lets you run single commands in an interactive session when invoking SMcli without specifying any parameters or script-commands
- *Command mode*, allowing you to run one or more storage configuration commands when invoking SMcli with parameters and script-commands from the shell
- *Script mode*, specifying a script for the script-commands parameter from a shell

Here we show how we performed our DS4800 logical storage configuration for our environment, as described in 6.1, “Setup example” on page 136 using SMcli in interactive mode. With SMcli we purposely perform the same DS4800 logical configuration as we describe in 7.4, “DS4000/DS5000 Storage configuration with the GUI” on page 252, so the two configuration alternatives SMcli and GUI can easily be compared.

> **Note:** The Storage Manager client allows you to easily save the existing DS storage configuratoin in a text file on a DS3000 series machine with gathering support information. This can be used later for scripting purposes or for duplicating DS Storage System configurations.

For further information about using the Storage Manager command-line interface refer to *IBM System Storage DS3000 and DS4000 Command Line Interface and Script Commands Programming Guide*, GC52-1275, available at:

ftp://ftp.software.ibm.com/systems/support/storage_disk_pdf/gc52127500.pdf

7.5.1 Creating hot spares

Corresponding to the hot spare assignment we performed with the DS Storage Manager client GUI (as described in 7.4.1, “Creating hot spares” on page 252) we now use the SMcli **`set [`*`enclosureID,slotID`*`] hotspare=true`** command for making the drive in enclosure 31, slot14 a hot spare, as shown in Example 7-2.

Example 7-2 SMcli hot spare creation

```
C:\Program Files\IBM_DS4000\client>SMcli 9.155.86.22 9.155.86.23
Entering interactive mode. Please type desired command.

set drive[31,14] hotSpare=true;
```

7.5.2 Creating RAID arrays and logical drives

In the same way in which we created a RAID-10 array with seven 35 GB logical drives from the disk drives in enclosure 31, slots 1 to 4, with the GUI in 7.4.2, “Creating RAID arrays and logical drives” on page 257, we now show this array and logical drive configuration using the SMcli **create logicaldrive** command, as shown in Example 7-3

Example 7-3 SMcli RAID and logical drive creation

```
create logicaldrive drives[31,1 31,3 31,2 31,4] raidLevel=1 userLabel="i6VIOSclient_0"
owner=A segmentSize=128 capacity=35 GB;

show allLogicalDrives summary;
STANDARD LOGICAL DRIVES SUMMARY
Number of standard logical drives: 1

NAME            STATUS   CAPACITY  RAID LEVEL  ARRAY
i6VIOSclient_0  Optimal  35 GB     1           1

create logicaldrive array[1] raidLevel=1 userLabel="i6VIOSclient_1" owner=B segmentSize=128
capacity=35 GB;
create logicaldrive array[1] raidLevel=1 userLabel="i6VIOSclient_2" owner=A segmentSize=128
capacity=35 GB;
create logicaldrive array[1] raidLevel=1 userLabel="i6VIOSclient_3" owner=B segmentSize=128
capacity=35 GB;
create logicaldrive array[1] raidLevel=1 userLabel="i6VIOSclient_4" owner=A segmentSize=128
capacity=35 GB;
create logicaldrive array[1] raidLevel=1 userLabel="i6VIOSclient_5" owner=B segmentSize=128
capacity=35 GB;
create logicaldrive array[1] raidLevel=1 userLabel="i6VIOSclient_6" owner=A segmentSize=128
capacity=35 GB;

show allLogicalDrives summary;
STANDARD LOGICAL DRIVES SUMMARY
Number of standard logical drives: 7

NAME            STATUS   CAPACITY  RAID LEVEL  ARRAY
i6VIOSclient_0  Optimal  35 GB     1           1
i6VIOSclient_1  Optimal  35 GB     1           1
i6VIOSclient_2  Optimal  35 GB     1           1
i6VIOSclient_3  Optimal  35 GB     1           1
i6VIOSclient_4  Optimal  35 GB     1           1
i6VIOSclient_5  Optimal  35 GB     1           1
i6VIOSclient_6  Optimal  35 GB     1           1
```

7.5.3 Creating hosts

After creating the logical drives we now create the host and host port definitions for our i6VIOS host using the SMcli **create host** and **create hostport** commands, as shown in Example 7-4.

Example 7-4 SMcli host and host port definition

```
create host userLabel="i6VIOS";
create hostPort host="i6VIOS" userLabel="i6VIOS_XNY_C3_T1" identifier="10000000c9701856"
hostType=6;
create hostPort host="i6VIOS" userLabel="i6VIOS_VYP_C3_T1" identifier="10000000c96da75e"
hostType=6;

show allHostPorts;
HOST PORT IDENTIFIER     HOST PORT NAME    HOST TYPE
10:00:00:00:c9:70:18:56  i6VIOS_XNY_C3_T1  AIX
10:00:00:00:c9:6d:a7:5e  i6VIOS_VYP_C3_T1  AIX
```

```
10:00:00:00:c9:70:18:e8  Undefined        Undefined
10:00:00:00:c9:6d:a9:32  Undefined        Undefined
```

7.5.4 Defining logical drive to LUN mapping

Finally, we create the storage partitioning for our “i6VIOS” host by mapping the logical drives to LUNs like we did before with the Storage Manager client GUI (see 7.4.4, “Defining logical drive to LUN mapping” on page 274), now using the SMcli **set logicaldrive** command, as shown in Example 7-5.

Example 7-5 SMcli logical drive to LUN mapping

```
set logicaldrive ["i6VIOSclient_0"] logicalUnitNumber=0 host="i6VIOS";
set logicaldrive ["i6VIOSclient_1"] logicalUnitNumber=1 host="i6VIOS";
set logicaldrive ["i6VIOSclient_2"] logicalUnitNumber=2 host="i6VIOS";
set logicaldrive ["i6VIOSclient_3"] logicalUnitNumber=3 host="i6VIOS";
set logicaldrive ["i6VIOSclient_4"] logicalUnitNumber=4 host="i6VIOS";
set logicaldrive ["i6VIOSclient_5"] logicalUnitNumber=5 host="i6VIOS";
set logicaldrive ["i6VIOSclient_6"] logicalUnitNumber=6 host="i6VIOS";

show storageSubsystem lunMappings;
MAPPINGS (Storage Partitioning - Enabled (1 of 8 used))-------------------

   VOLUME NAME     LUN  CONTROLLER  ACCESSIBLE BY  VOLUME STATUS
   i6VIOSclient_0  0    A           Host i6VIOS    Optimal
   i6VIOSclient_1  1    B           Host i6VIOS    Optimal
   i6VIOSclient_2  2    A           Host i6VIOS    Optimal (Initialization)
   i6VIOSclient_3  3    B           Host i6VIOS    Optimal (Initialization)
   i6VIOSclient_4  4    A           Host i6VIOS    Optimal (Initialization)
   i6VIOSclient_5  5    B           Host i6VIOS    Optimal (Initialization)
   i6VIOSclient_6  6    A           Host i6VIOS    Optimal (Initialization)
```

At this point we are done with the DS4800 logical configuration and can proceed with mapping the DS4800 LUNs under VIOS to the IBM i client, as described in 6.5, “Configuring VIOS virtual devices” on page 191.

8

Using Midrange Storage Copy Services

Many System i data centers are looking for Business Continuity (BC) solutions with Copy Services of System Storage. They implement FlashCopy for minimizing downtime to save their system to tape, they establish remote replication to provide a copy of their system at remote site, to recover from in case of disaster on production site.

Solutions with Copy Services are appreciated among System i customers because of some important advantages over the other Business Continuity solutions available the them. For example: Copy Services provide relatively simple management, unified BC scenario for System i and other servers, low total cost of ownership, and so on.

For more information about Business Continuity solutions for a System i server with external storage, refer to:

- *IBM i and IBM System Storage: A Guide to Implementing External Disks on IBM i*, SG24-7120
- *Implementing Tivoli Data Warehouse V 1.2*, SG24-71003
- *IBM System Storage DS8000: Copy Services in Open Environments*, SG24-6788

In this section we describe usage and implementation of Midrange Storage FlashCopy, Volume Copy, Metro Mirroring and Global Mirroring for a System i installation.

8.1 Implementing FlashCopy

To implement a System i Business Continuity solution with Midrange Storage, FlashCopy of *all disk units in an IBM i partition* should be performed.

> **Note:** In our setup we refer to the System i partition that uses FlashCopy base drives as a *production partition*, and we call the partition that brings up a clone from FlashCopy logical drives a *backup partition*.
>
> For our setup we connect both production and backup partitions to the same VIOS. However, we recommend that you connect FlashCopy base logical drives and FlashCopy target logical drives through different VIOSes.
>
> For more information about planning for a backup partition refer to 4.7, “Planning for Copy Services” on page 102.

When implementing FlashCopy for minimizing the backup window, we recommend creating FlashCopy logical drives only the first time that you use it, and for every future backup just *disable* and *recreate* FlashCopy. In this section we describe the steps for the first implementation and the steps for future usage.

8.1.1 Creating FlashCopy the first time

To implement FlashCopy of an entire IBM i disk space:

1. Obtain the keys for the premium feature FlashCopy for the Storage System. The Web page from which to download keys and instructions for downloading are provided to the customer that purchased FlashCopy premium feature. After the keys are stored in a file on PC, enable the FlashCopy by taking the following steps:

   a. Start the DS Storage Manager client and select **Logical / Physical view** of the relevant Storage System. Select **Storage Subsystems** → **Premium Features** → **Enable**. See Figure 8-1.

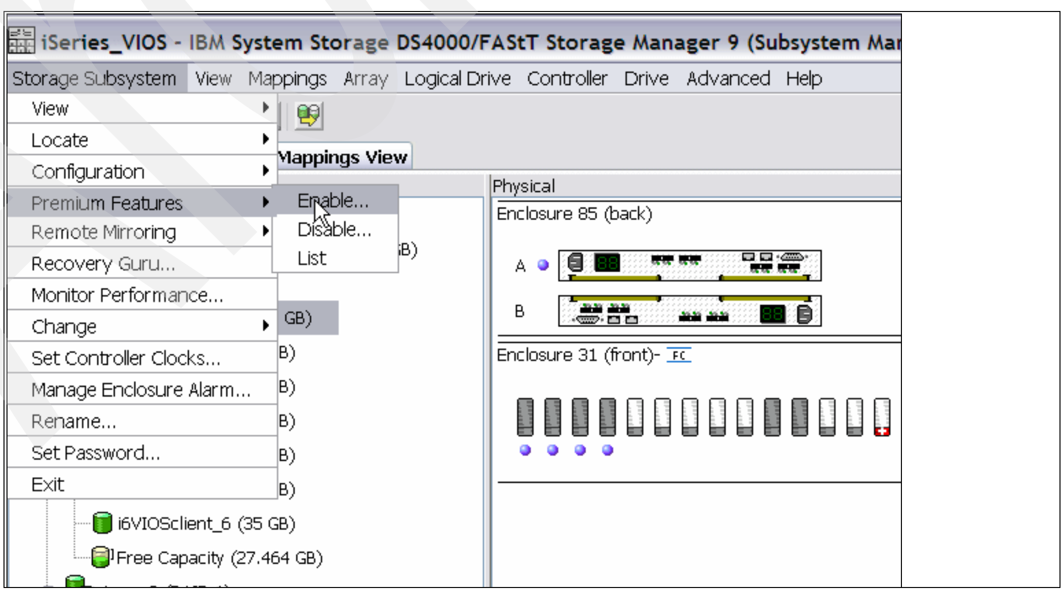


Figure 8-1 Enable FlashCopy premium feature

b. This brings up the Select Feature Key File window, where you select the file that contains the downloaded keys and click **OK**, as shown in Figure 8-2. On the next window confirm that you want to enable the premium feature for FlashCopy.

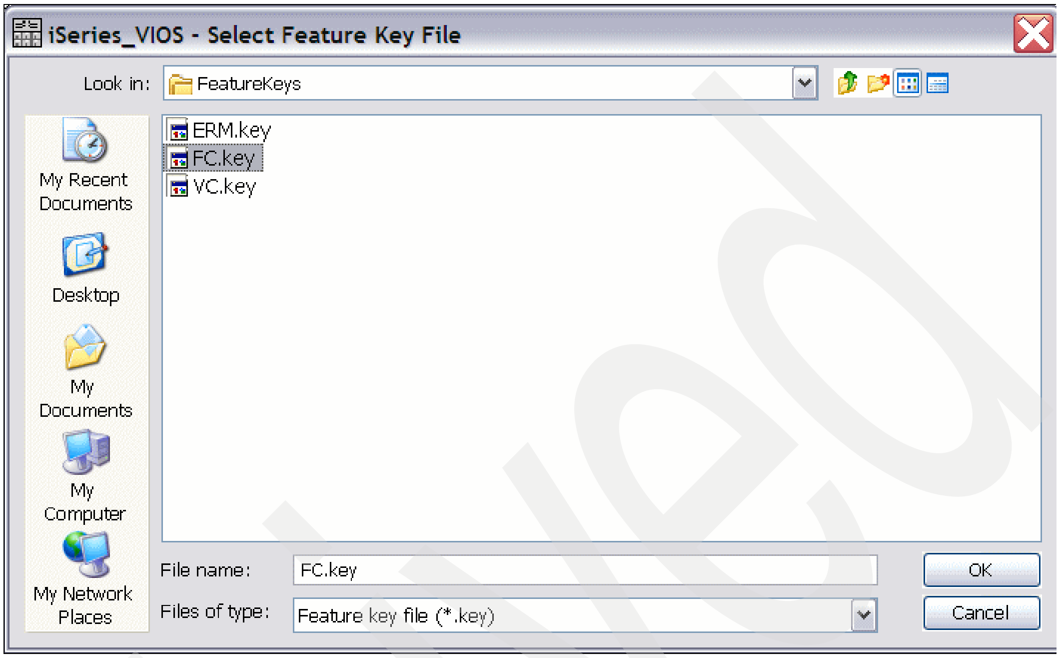


Figure 8-2 Select the file with keys

2. In the System i partition (client partition of VIOS) observe the disk units in SST, as shown in Figure 8-3, and identify the corresponding logical drives in Midrange Storage. For more information about how to get to the Display Disk Configuration Status panel and how to identify relevant logical drives refer to "Configuring storage for IBM i" on page 191 and "Adding DS Storage to IBM i configuration" on page 197.

```
                      Display Disk Configuration Status

          Serial                      Resource
ASP Unit  Number          Type Model  Name       Status
  1                                              Unprotected
       1  Y47UBW7RK862    6B22  050   DD001      Configured
       2  YWVTCLJL99U5    6B22  050   DD006      Configured
       3  YK8VKNB3PU8F    6B22  050   DD005      Configured
       4  YFS2CFWAHZ7U    6B22  050   DD002      Configured
       5  YFTGZ27PJMH2    6B22  050   DD003      Configured
       6  Y699U2SWAVYA    6B22  050   DD007      Configured
       7  YX6T78XTA2ZM    6B22  050   DD004      Configured









Press Enter to continue.

F3=Exit        F5=Refresh         F9=Display disk unit details
F11=Disk configuration capacity   F12=Cancel
```

Figure 8-3 Disk units in System i partition

3. Power down the production partition by using IBM i command PWRDWNSYS.
4. Select **DS Storage Manager client** → **Logical / Physical view** of the relevant Storage System.

5. Right-click the logical drive for IBM i and select **Create FlashCopy logical drive** from the pull-down, as shown in Figure 8-4. This starts the wizard to create the FlashCopy logical drive.

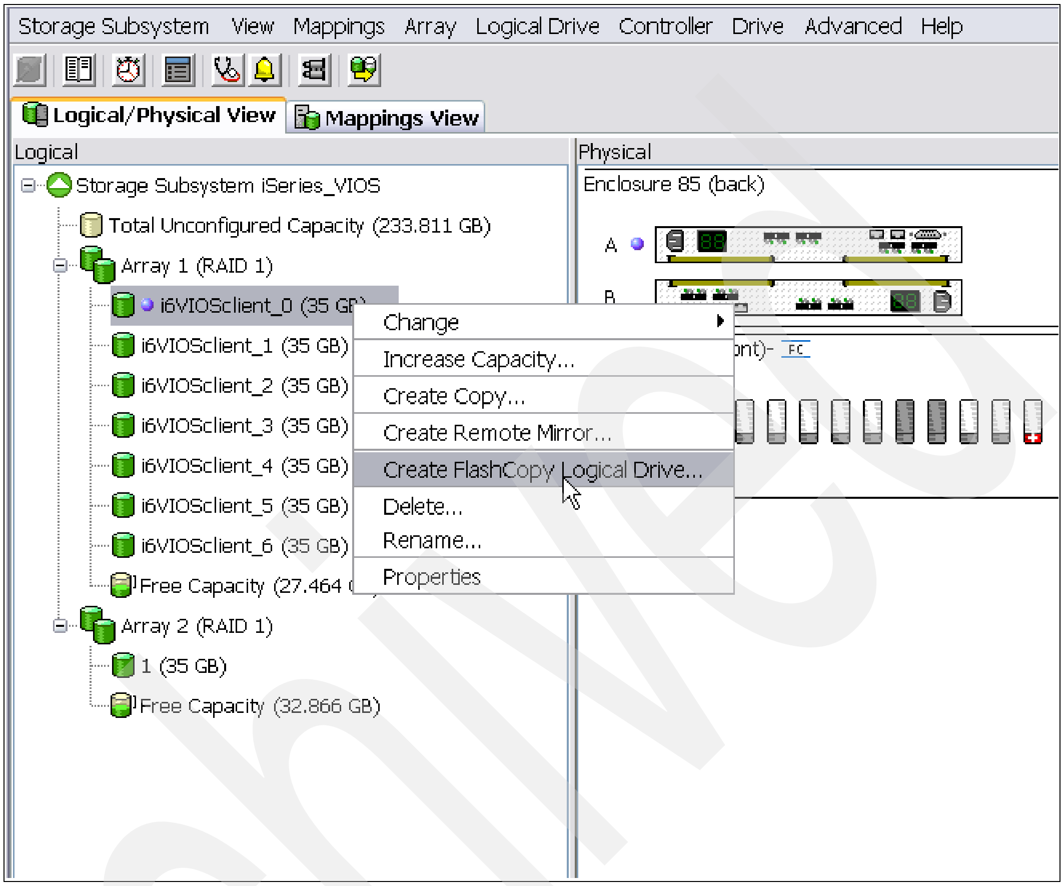


Figure 8-4 Create FlashCopy Logical Drive

6. In the first wizard window you choose to perform either the simple or the advanced way of creating FlashCopy. When you select the simple way the default values, such as percentage of capacity for FlashCopy repository logical volume, RAID array to define FlashCopy logical volumes, and so on, are automatically chosen. With the advanced way, however, you can change the default values according to your needs. After making a selection click **Next**. In our setup we choose the advanced way, as shown in Figure 8-5.

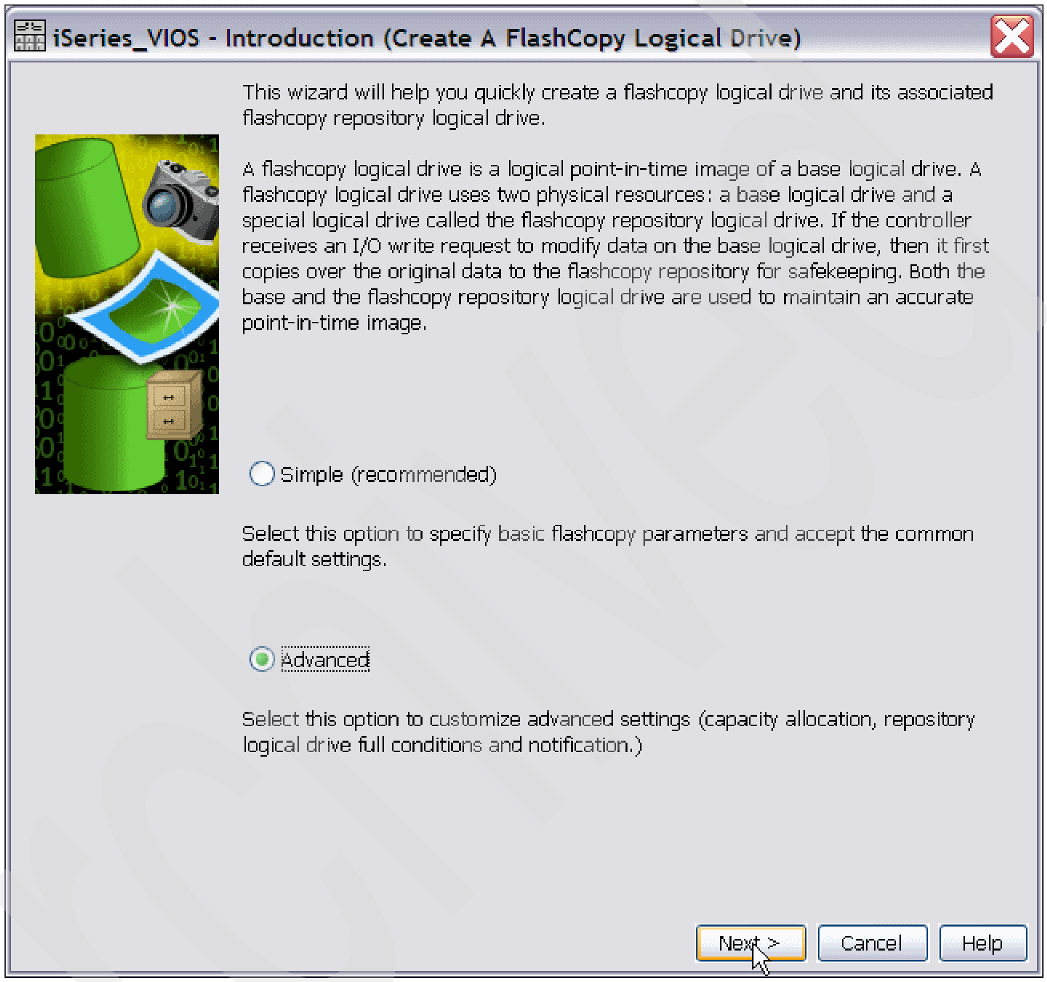


Figure 8-5 Choose the way to create FlashCopy

7. On the next wizard panel specify the names of the FlashCopy logical drive and the FlashCopy repository logical drive. You can also keep the default names, as shown on Figure 8-6. Click **Next**.

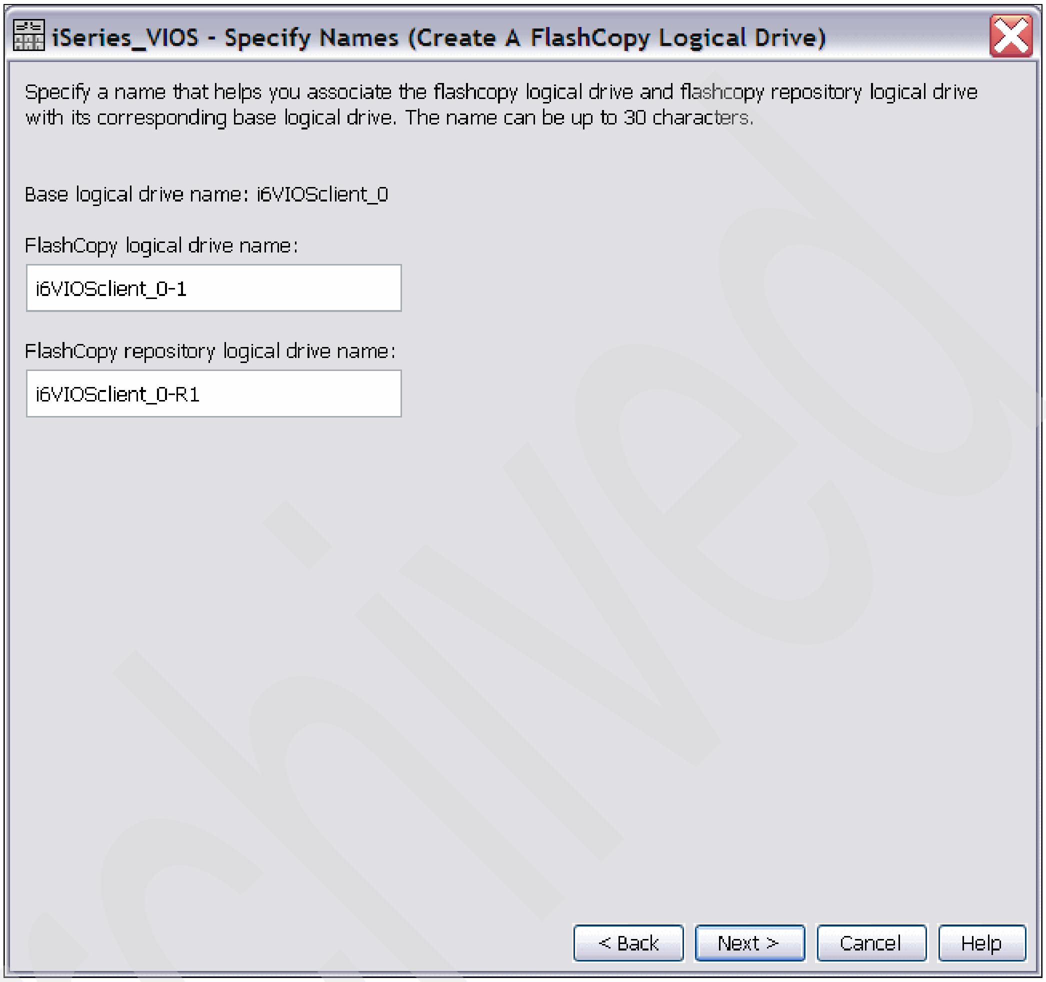


Figure 8-6 Names for FlashCopy drives

8. In the next wizard window select whether to define the FlashCopy repository logical drive from the same array as the base drive or to use another array. In our setup we do not have enough free capacity for all repository drives on the same array on which base drives are defined, so we choose another array. See Figure 8-7.

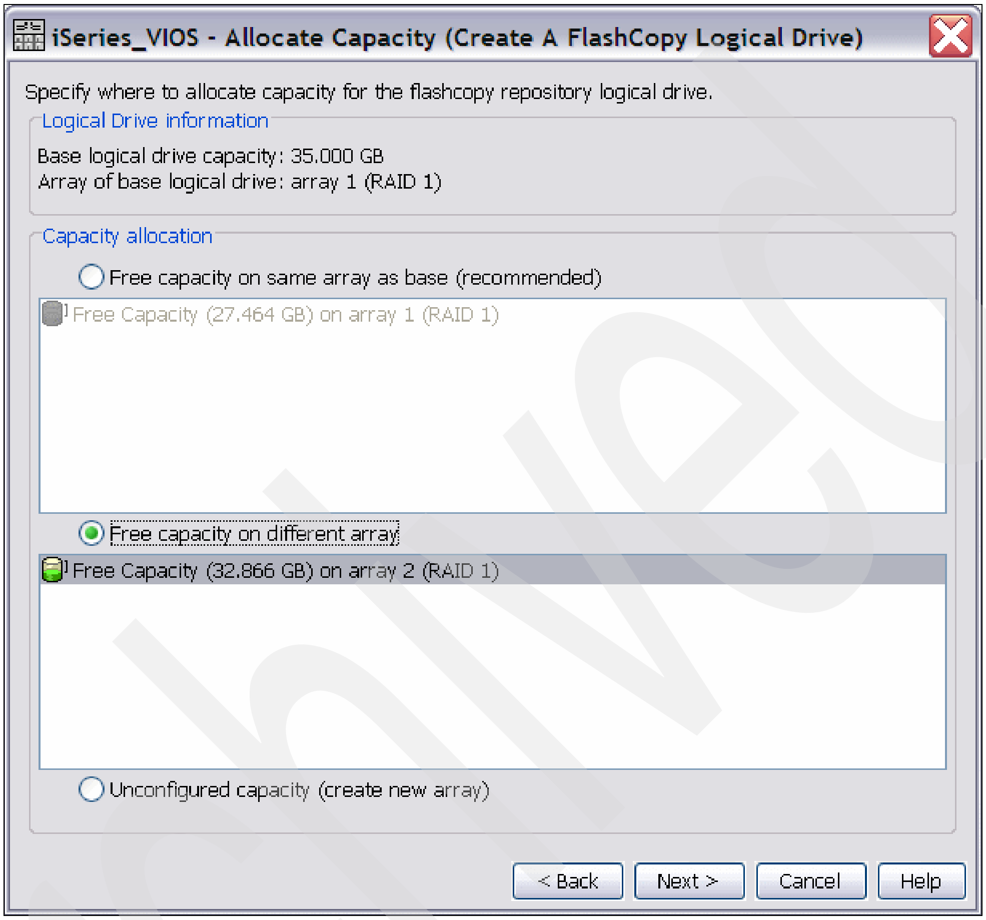


Figure 8-7 Select array for FlashCopy repository logical drive

9. On the next wizard panel specify the percentage of base drive capacity to use for the FlashCopy repository drive. We expect that the default value of 20% will be enough for most System i workloads. However, we strongly recommend that you estimate the needed capacity, as described in "Planning for Copy Services" on page 102. In our example we keep default value 20% (Figure 8-8).

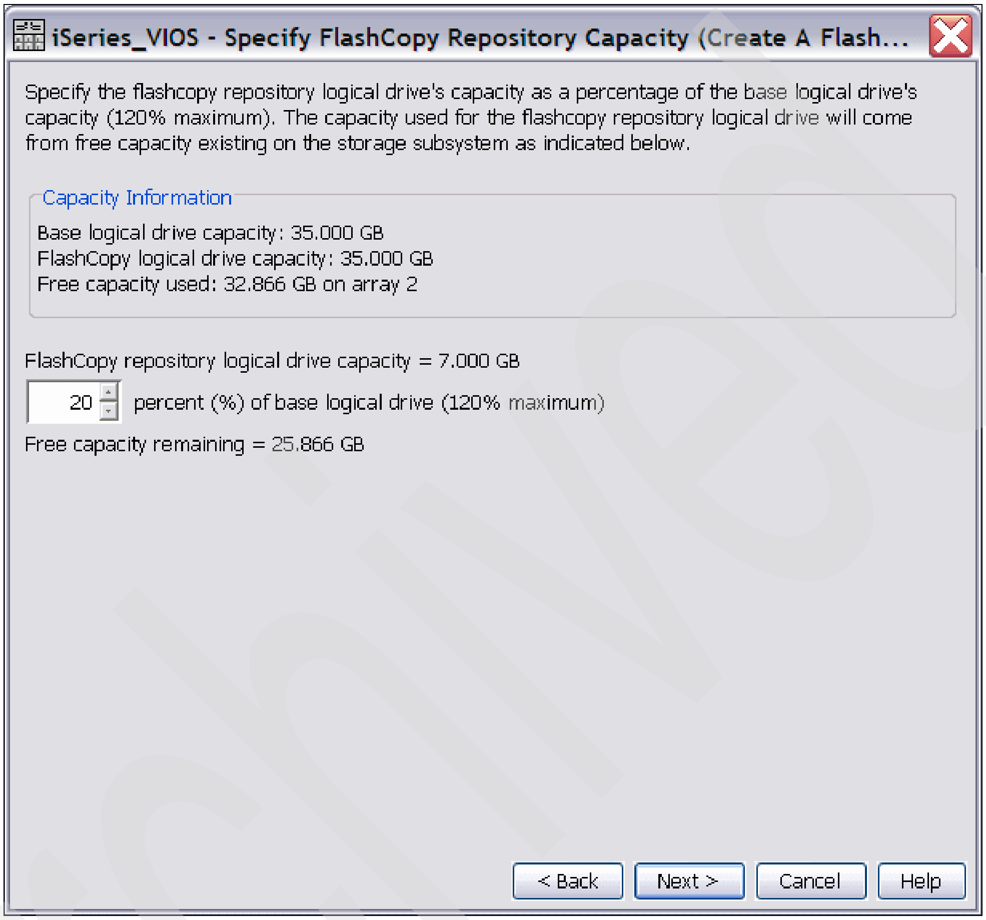


Figure 8-8 Capacity for FlashCopy repository logical drive

10. Specify whether to immediately map FlashCopy logical drives to LUNs or to map later. Specify also the treshold occupation of the repository logical drive at which the alert is sent. See Figure 8-9.

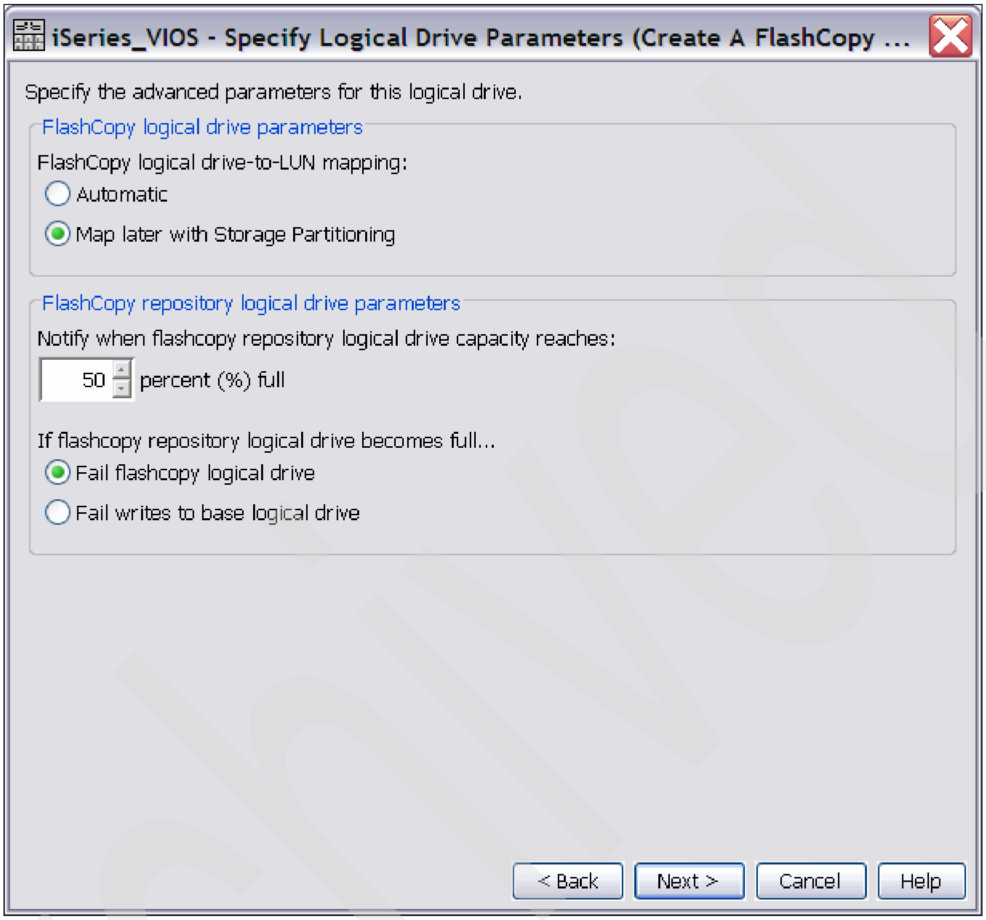


Figure 8-9 LUN mapping and treshold capacity

11. Check the characteristics of the FlashCopy logical drive being created. If they are okay, click **Finish** to create the drive, as shown in Figure 8-10. Click **OK** on the Completed (Create a FlashCopy Logical Drive) panel.

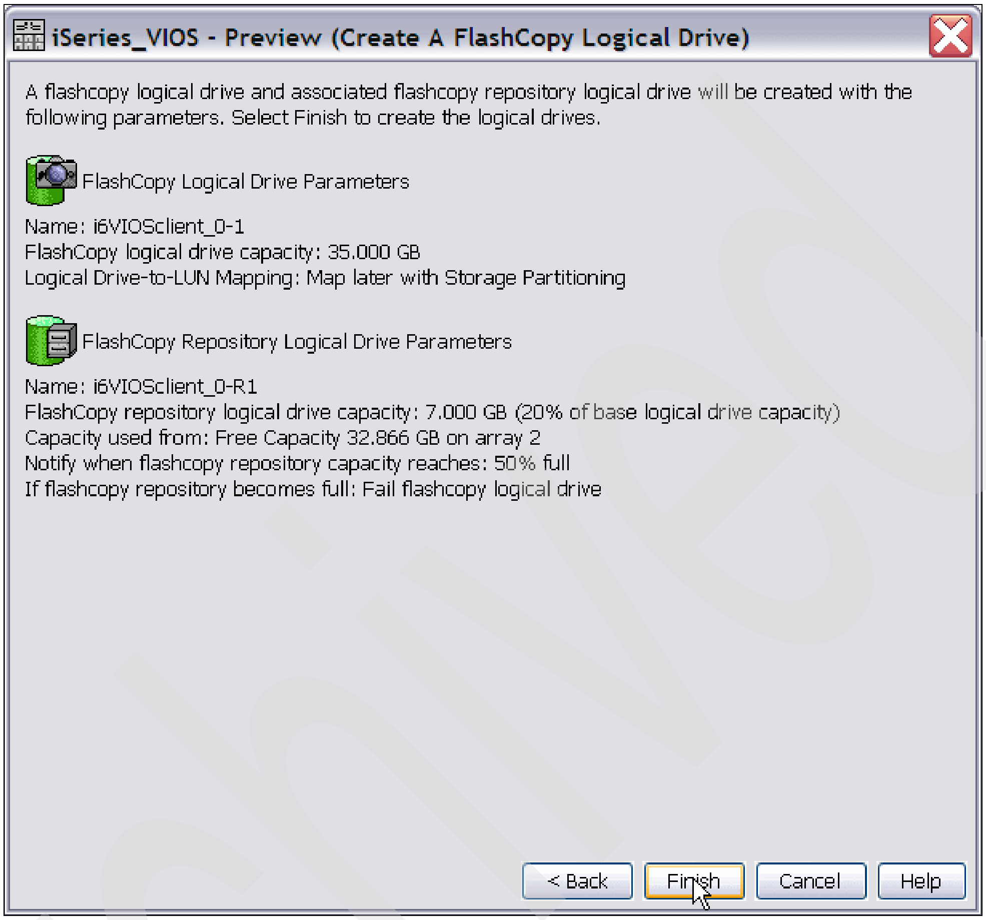


Figure 8-10 Finish creating FlashCopy logical drive

12. Repeat steps 4–11 for each logical drive in the IBM i client partition. After all logical drives are created, you may observe FlashCopy logical drives and FlashCopy repository logical drives in the Logical/Physical View of the Storage System. See Figure 8-11.

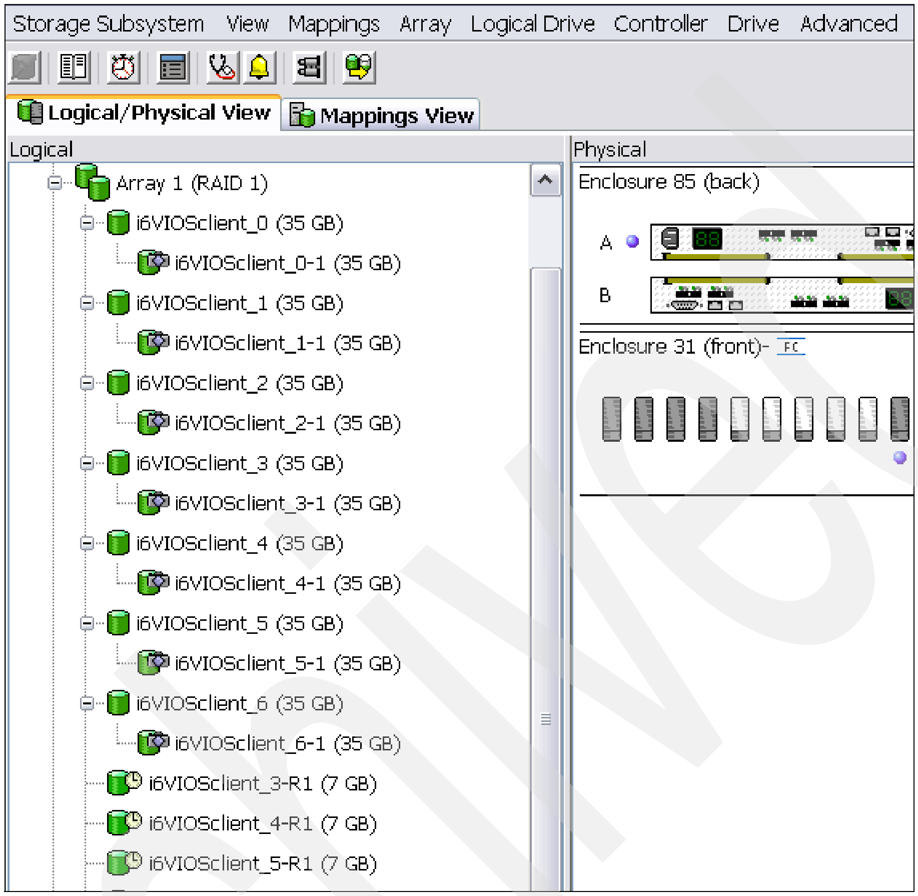


Figure 8-11 FlashCopy logical drives

13. Activate the System i Production partition to resume production workload.

14. To connect FlashCopy logical drives to the backup partition:

   a. Map the FlashCopy logical drives to LUNs. For more information about mapping logical drives refer to “Defining logical drive to LUN mapping” on page 281.

   b. In VIOS run the command CFGDEV to recognize the added LUNs, then map the FlashCopy LUNs to the virtual adapter assigned to the backup IBM i client. For more detailed instructions refer to “Configuring VIOS virtual devices” on page 191.

15. In the backup IBM i client, make sure that the virtual adapter assigned to the VIOS virtual adapter is tagged for load source. The profile of our backup partition with a tagged virtual adapter is shown in Figure 8-12. Also, make sure that the IPL source is set to B. For more information about settings in the IBM i client, refer to “Creating the IBM i Client LPAR” on page 155.

    Activate the backup partition.

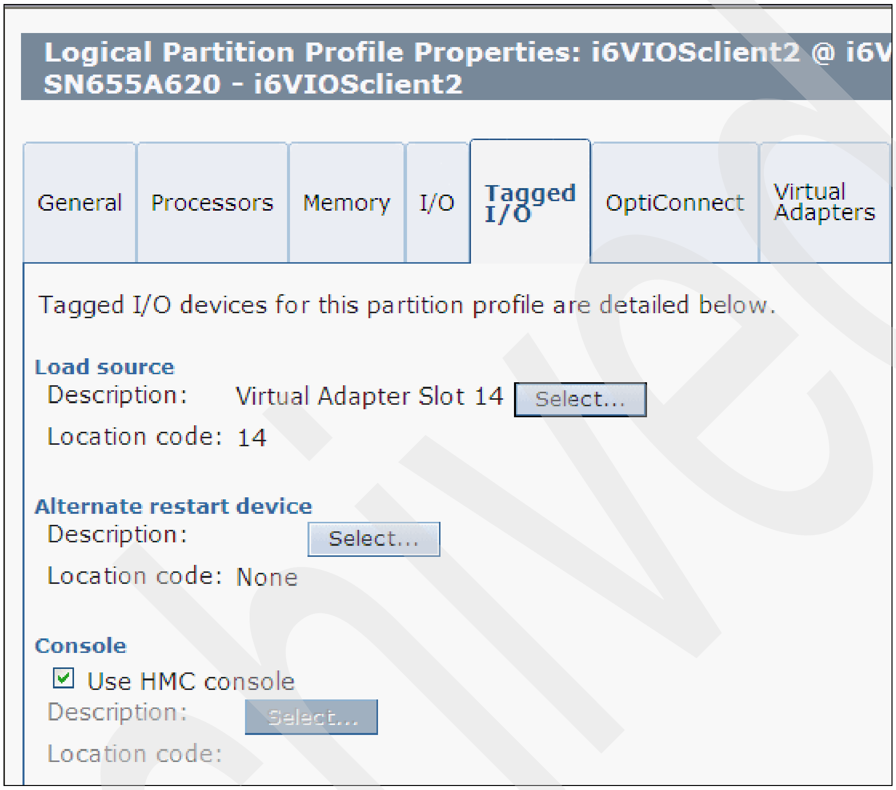


Figure 8-12 Tagged virtual adapter for load source

16. Once the backup partition is activated, you may look in SST for the virtual adapter and connected FlashCopy logical drives, the load source unit being one of them, as shown in Figure 8-13.

Now the backup IBM i client contains the clone of the production partition. For more information about adjustments and using BRMS in the backup partition, refer to 8.1.3, “Considerations when cloning IBM i system” on page 300.

```
                Logical Hardware Resources Associated with IOP

Type options, press Enter.
  2=Change detail    4=Remove    5=Display detail    6=I/O debug
  7=Verify          8=Associated packaging resource(s)

                                                              Resource
Opt Description                   Type-Model  Status          Name
    Virtual IOP                    290A-001   Operational     CMB09
     Virtual Storage IOA           290A-001   Operational     DC05
      Disk Unit               *    6B22-050   Operational     DD001
      Disk Unit                    6B22-050   Operational     DD007
      Disk Unit                    6B22-050   Operational     DD005
      Disk Unit                    6B22-050   Operational     DD004
      Disk Unit                    6B22-050   Operational     DD002
      Disk Unit                    6B22-050   Operational     DD003
      Disk Unit                    6B22-050   Operational     DD006



F3=Exit    F5=Refresh    F6=Print    F8=Include non-reporting resources
F9=Failed resources      F10=Non-reporting resources
F11=Display serial/part numbers      F12=Cancel
```

Figure 8-13 FlashCopy logical drives in the backup partition

8.1.2 Regular usage of FlashCopy

When using FlashCopy for daily backups, we recommend implementing the functions *disable* FlashCopy and *re-create* FlashCopy. Disable just removes the FlashCopy relation but it keeps the created FlashCopy logical drives and repository drives. It also clears the repository occupation. Re-create establishes again the FlashCopy relation on existing FlashCopy logical drives. With these functions you do not need to change any LUN mapping and VIOS assignments for everyday backup.

To disable and re-create FlashCopy:

1. After the save in backup partition is finished, power-down the partition.
2. Use the DS Storage Manager. Double-click the relevant Storage System to open the Subsystem Management window Logical /Physical view. Right-click each FlashCopy logical drive of the production partition and select **Disable** from the pull-down, as shown in Figure 8-14. Confirm disabling of the logical drive by typing `Yes` in the confirmation window, as shown in Figure 8-15 on page 298.

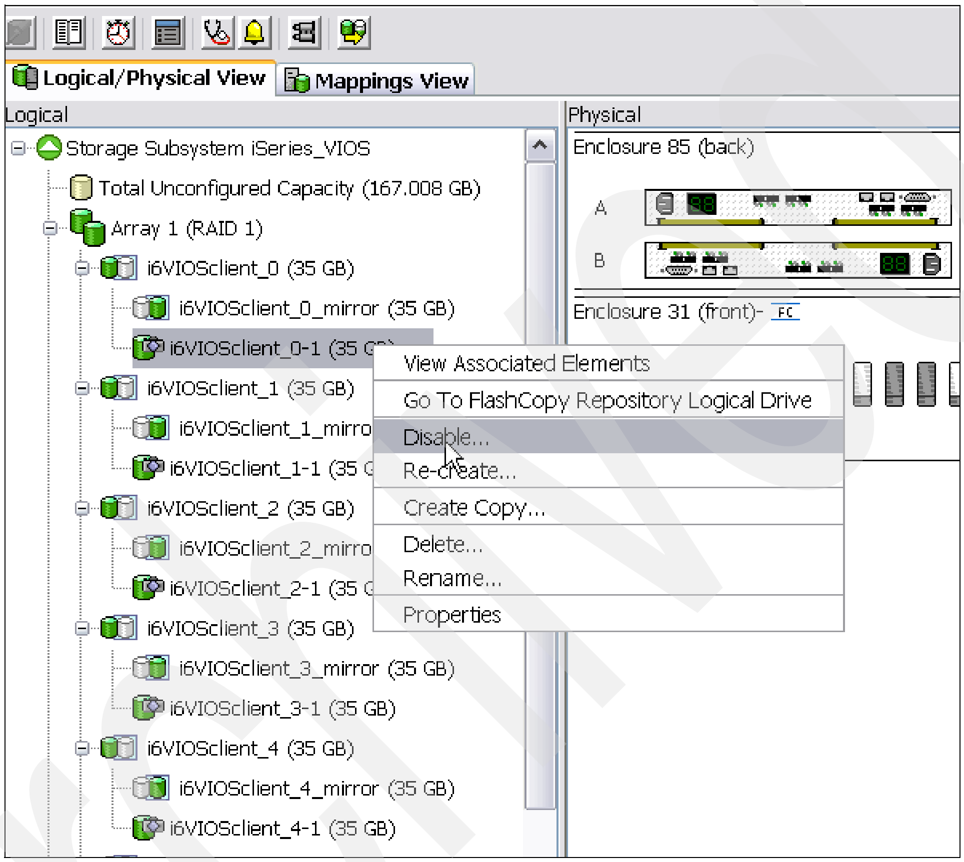


Figure 8-14 Disable FlashCopy logical drive

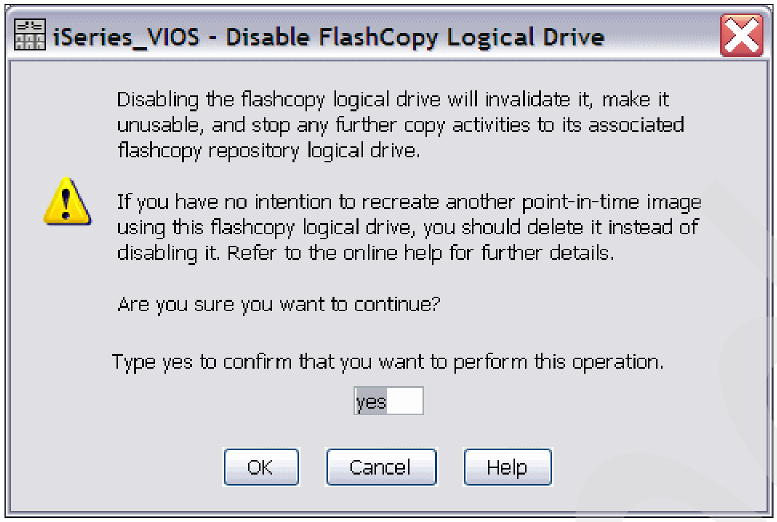


Figure 8-15 Conform disable of FlashCopy Logical drive

After a FlashCopy logical drive is disabled its icon changes form dark green to light green. Disabled FlashCopy logical drives in our setup are shown in Figure 8-16.

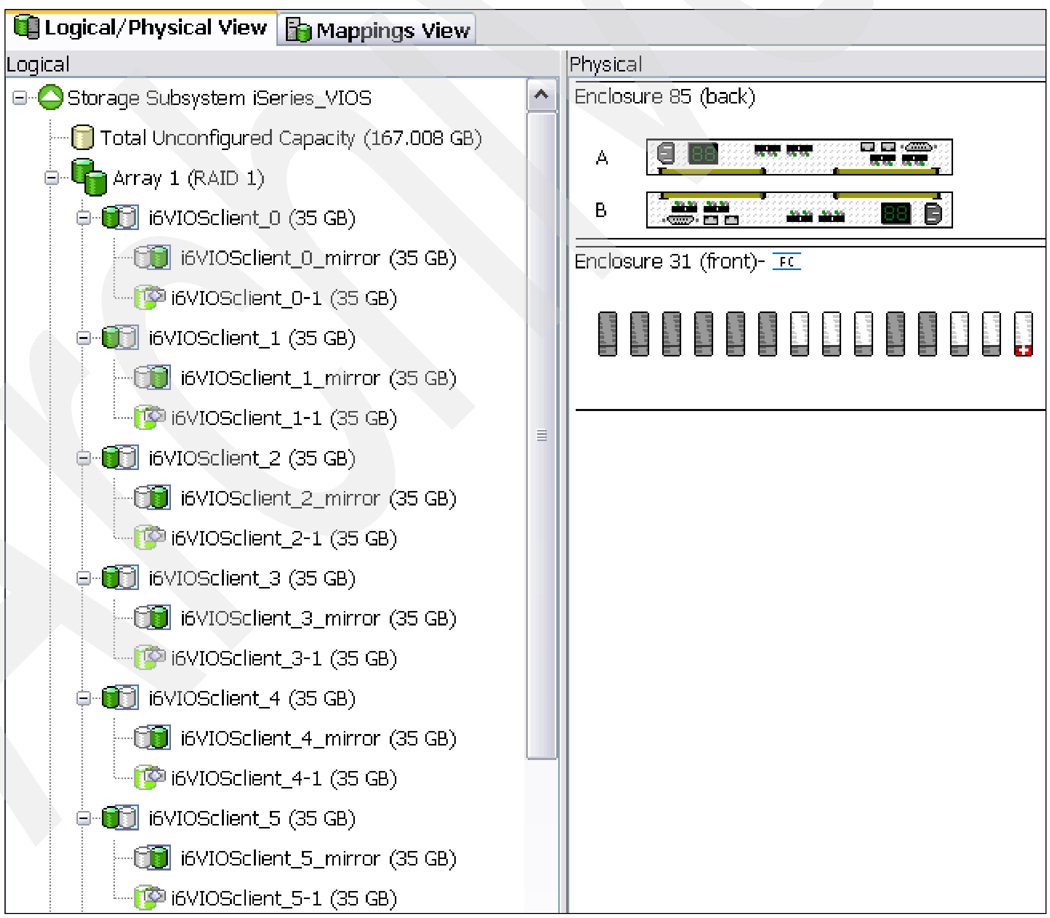


Figure 8-16 Disabled FlashCopy logical drives

3. Before the next save, power down the system production partition.

4. Use DS Storage Manager and double-click the relevant Storage System to open the Subsystem Management window Logical/Physical view. Right-click each FlashCopy logical drive of the production partition and select **Re-create** from the pull-down, as shown in Figure 8-17. Type `Yes` in the confirmation window to confirm the re-creation of FlashCopy logical drive.

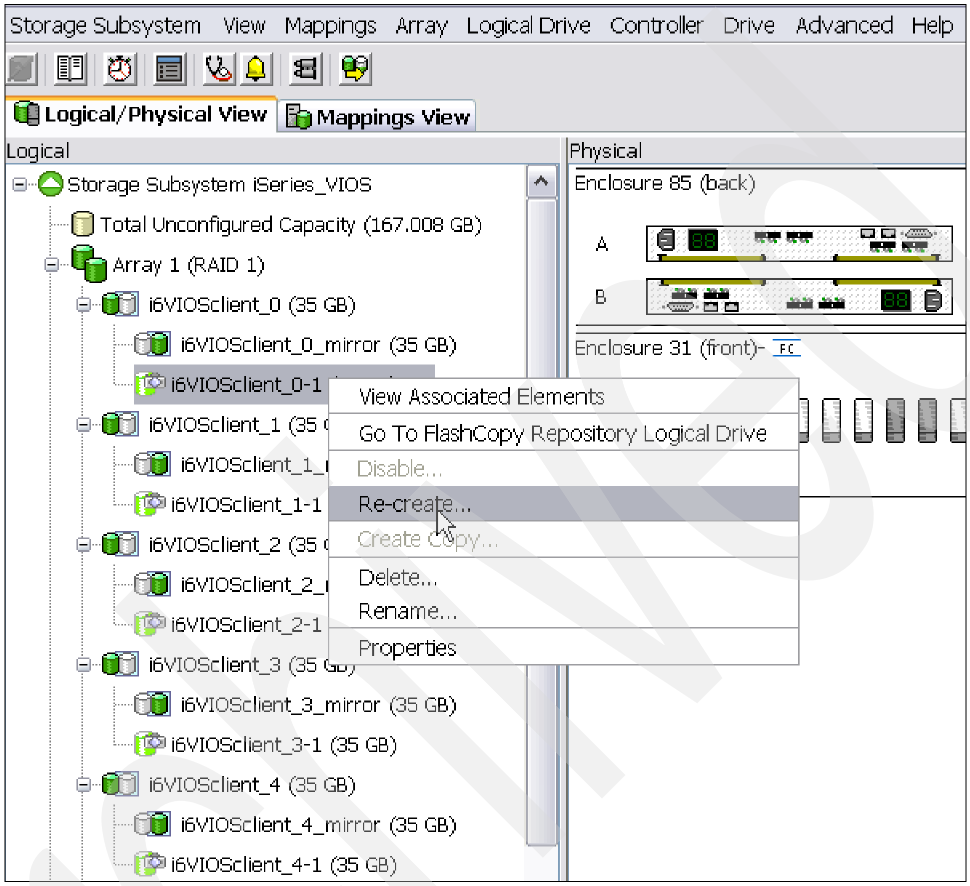


Figure 8-17 Re-create FlashCopy logical drive

5. Activate the production partition to resume production workload.

6. In VIOS, perform the command **cfgdev**. This is needed to enable the backup partition to recognize logical drives after they are changed by FlashCopy. Inserting this command via Putty is shown in Figure 8-18.

```
Using username "padmin".
padmin@9.155.50.45's password:
Last unsuccessful login: Fri Aug  8 11:35:47 CDT 2008 on ssh from
c99r7v0m.rchland.ibm.com
Last login: Thu Aug 21 09:22:38 CDT 2008 on /dev/pts/0 from
c99r7v0m.rchland.ibm.com
$ cfgdev


Some error messages may contain invalid information
for the Virtual I/O Server environment.

Method error (/usr/sbin/fcppcmmap > /etc/essmap.out):
        0514-001 System error:
$
```

Figure 8-18 Command cfgdev in VIOS

7. The FlashCopy logical volumes are now ready for use. IPL the backup partition and perform the next save to tape.

8.1.3 Considerations when cloning IBM i system

Once the backup partition is IPLed from FlashCopy logical drives it contains a clone of the production partition, which you use to save production data to tape. Performing backups from the clone is a very convenient way to minimize the save window in the production system. However, you should keep in mind that:

- A clone is an exact copy of the original source system in every respect.
- The system name and network attributes are identical.
- The TCP/IP settings are identical.
- The BRMS network information is identical.
- The NetServer™ settings are identical.
- User profiles and passwords are identical.
- The job schedule entries are identical.
- Relational database entries are identical.

Extreme care should be taken when you activate a partition that has been built from a FlashCopy of total disk space. In particular, you must ensure that it does not automatically connect to the network, as this would cause substantial problems within both the copy and its parent system. For more information about cloning IBM i system refer to *IBM i and IBM System Storage: A Guide to Implementing External Disks on IBM i*, SG24-7120. See also 8.1.6, “Automation of FlashCopy” on page 302.

8.1.4 FlashCopy and BRMS

Backup Recovery and Media Services (BRMS) is the IBM strategic solution for performing backups and recovering IBM i systems. BRMS has a wealth of features, including the ability to work in a network with other systems to maintain a common inventory of tape volumes.

The BRMS implementation of FlashCopy provides a way to perform a backup on a FlashCopied clone partition, and BRMS history would appear as the backup is performed in the production partition. The BRMS FlashCopy function requires the BRMS-Network Feature.

Before taking a FlashCopy of disks in the production partition you should enable the BRMS FlashCopy function and perform pre-FlashCopy actions on the production system. After the backup partition is up and running, you should perform needed pre-backup steps on the clone system. Once backup is finished there are some post-FlashCopy steps that need to be done in the backup partition.

For more information about the steps for FlashCopy with BRMS refer to *Implementing Tivoli Data Warehouse V 1.2*, SG24-71003.

8.1.5 Occupation of FlashCopy repository logical drive

During the time that FlashCopy is enabled, the occupation of FlashCopy repository logical drive grows along with write operation to FlashCopy base drives. If the capacity of FlashCopy repository logical dives is properly planned they do not get 100% occupied during the time needed to save from the backup partition.

However, when the occupation reaches the percentage specified during FlashCopy setup (default is 50%) the alert is sent to DS Storage Manager. It can be seen with the Recovery Guru. Such an alert is shown in Figure 8-19. You can also set up DS Storage Manager to send the SNMP alerts as e-mail to the specified mail server. It may be a good idea to enable IBM i mail to receive the alerts.

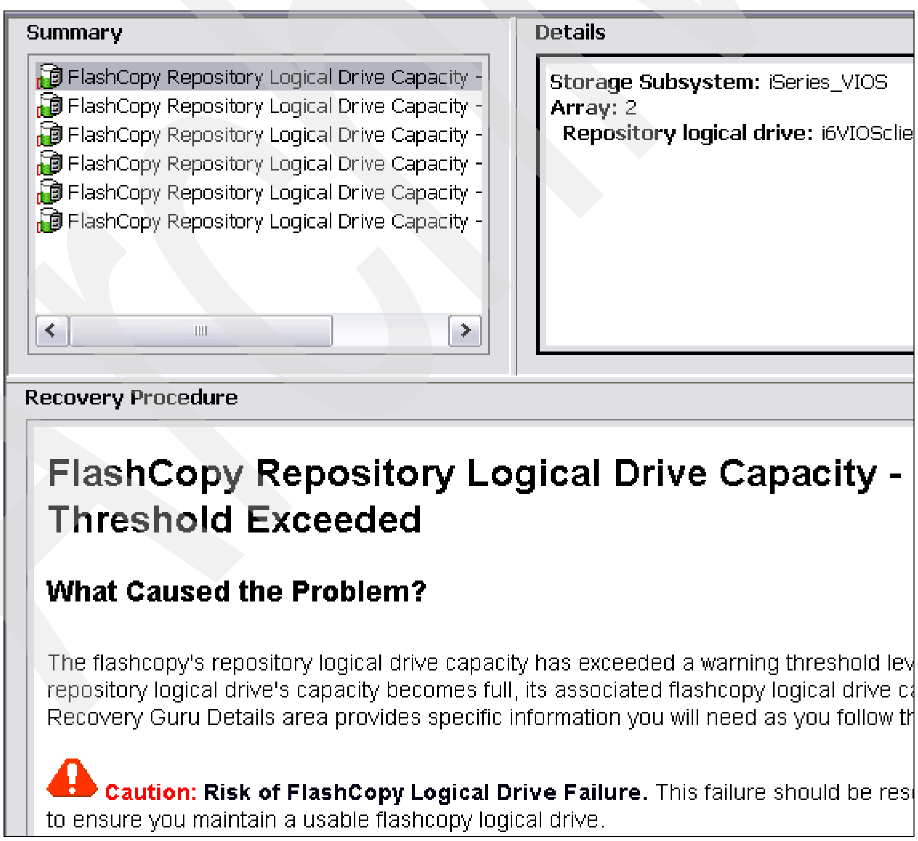


Figure 8-19 Treshold of repository drive reached

If, by any chance, the occupation of the repository drive reaches 100%, the FlashCopy logical drive is automatically disabled or the writes to the FlashCopy base drive fail, depending which of these two actions you specify when creating FlashCopy. For FlashCopy with the System i partition, we strongly recommend that you chose failure of the FlashCopy logical drive. This is because the failure to write on a FlashCopy base drive would cause the production partition to claim the disk as missing, and the partition would stop with SRC code A6020266. It would start again only when the writes are enabled on FlashCopy base logical drive.

8.1.6 Automation of FlashCopy

It is convenient to automate the procedure for regular FlashCopy for full disk space and backups. Since this scenario requires commands in the DS Storage Manager client, in VIOS, and in both the production and the backup partition, it may be a good idea to use remote commands from a third IBM i partition (not production or backup) and automate them in a CL program. Below we give you an idea of which commands to use for such automation:

- Use a remote command to power down the production partition.
- Issue a remote command to the workstation where DS Storage Manager is installed to perform FlashCopy with SMCli or scripts.
- Use an secure shell command to HMC to IPL the production partition.
- Issue a secure shell command to VIOS to perform CFGDEV.

  The licensed product Portable Utilities for i5/OS, 5733-SC1, is the prerequisite for this. It is needed to first create public keys in IBM i for access to VIOS.
- Use an secure shell command to HMC to IPL the backup partition.

8.2 Implementing Volume Copy

Volume Copy is a full point-in-time replication, where data are actually copied from source to target logical drives. It is suitable for testing, analysis, and data mining, since it does not bring any degradation of the production logical drive performance. On the other hand, FlashCopy is usually considered for daily backups.

While the copying of logical drives is in progress the source logical drive is only available for read I/O activity, while write requests are allowed after the logical drive copy is complete. Employing both FlashCopy and Volume Copy, as shown in Figure 8-20, enables a complete copy to be created without interrupting the production workload.

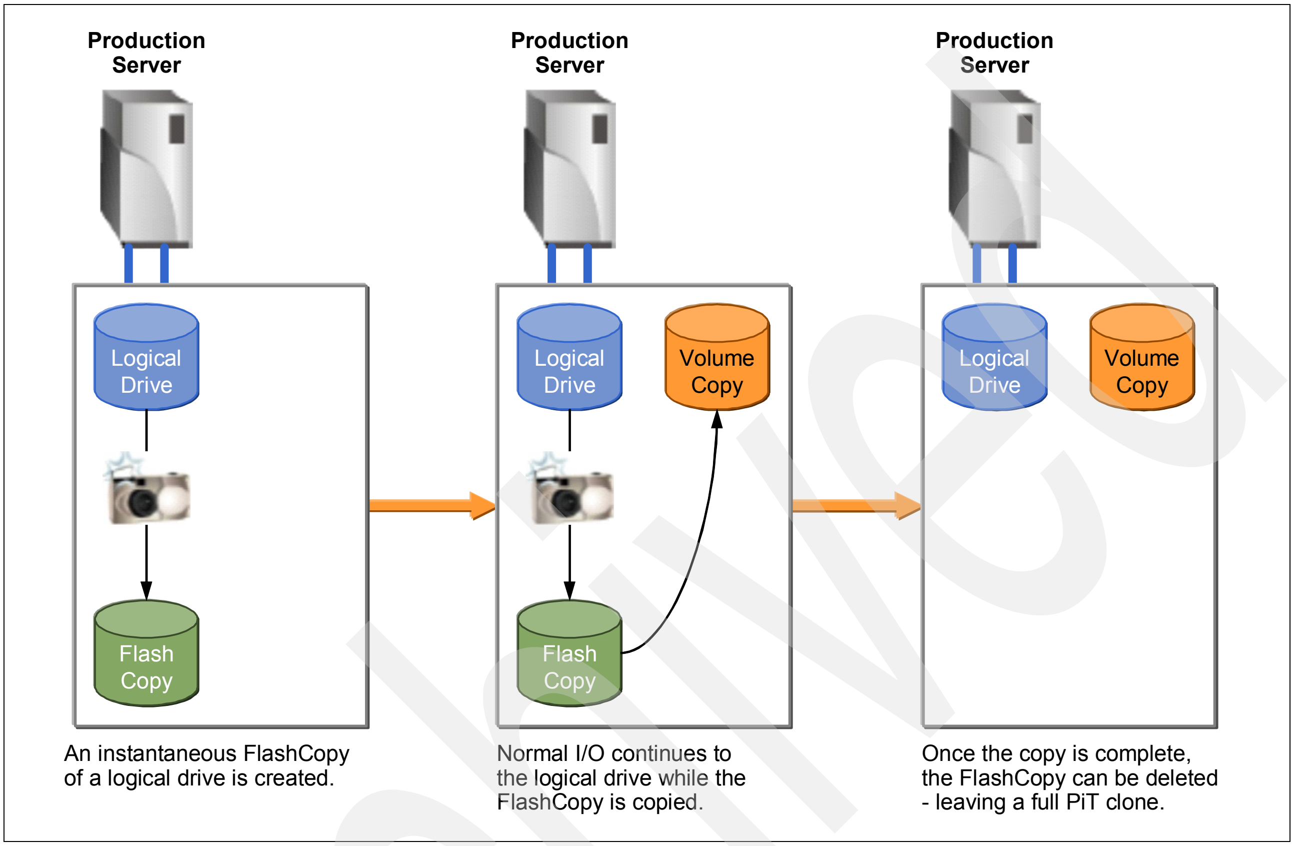


Figure 8-20 Volume Copy integration with FlashCopy

> **Note:** For System i implementations it is essential to use Volume Copy always in conjunction with FlashCopy. Using only Volume Copy would cause the production system to stop.

It makes sense to copy all logical drives (IBM i disk units) of the System i production partition. Once the logical drives are copied, another IBM i partition can be IPLed from Volume Copy logical drives. It will contain a clone of the production partition, which can be used for testing, data mining, and similar tasks. Some installations may also consider Volume Copy for migration purposes.

To implement Volume Copy with FlashCopy for a System i partition, perform the following steps for each logical drive in the production partition:

1. Power down the production partition.
2. Make a FlashCopy of the logical drive, as is described in “Implementing FlashCopy” on page 284.
3. Activate the production partition to resume the production workload.

4. In DS Storage Manager → Subsystem Management Logical / physical view right-click **FlashCopy logical volume** and select **Create Copy**. See Figure 8-21.

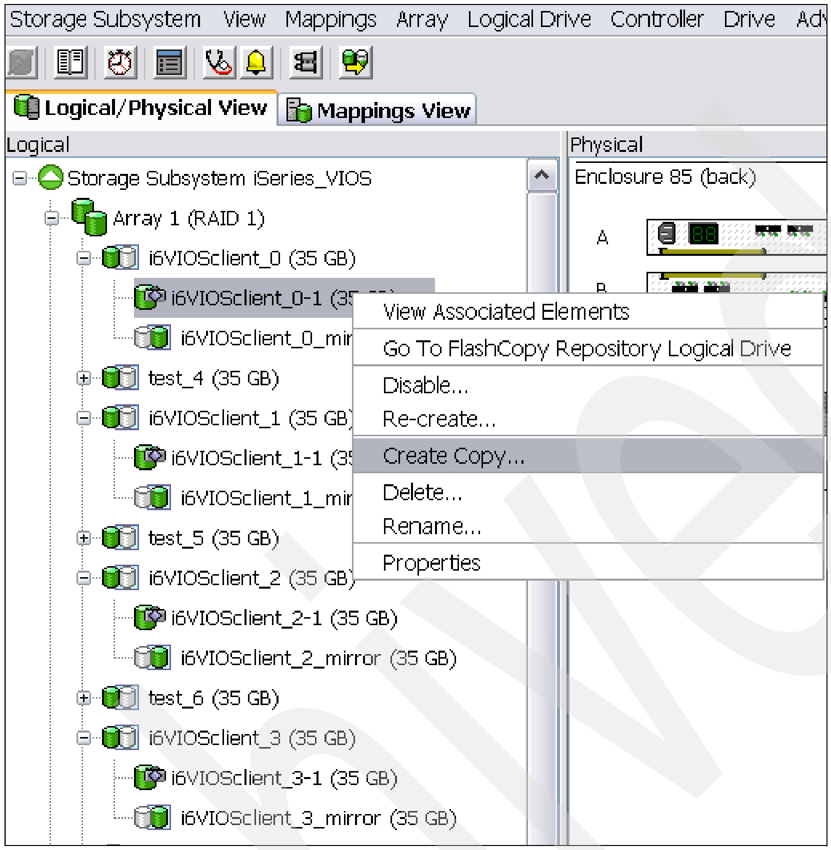


Figure 8-21 Create Copy

5. This starts the Create Copy wizard. On the first wizard window select the Volume Copy source logical drive (FlashCopy logical drive), as shown in Figure 8-22.

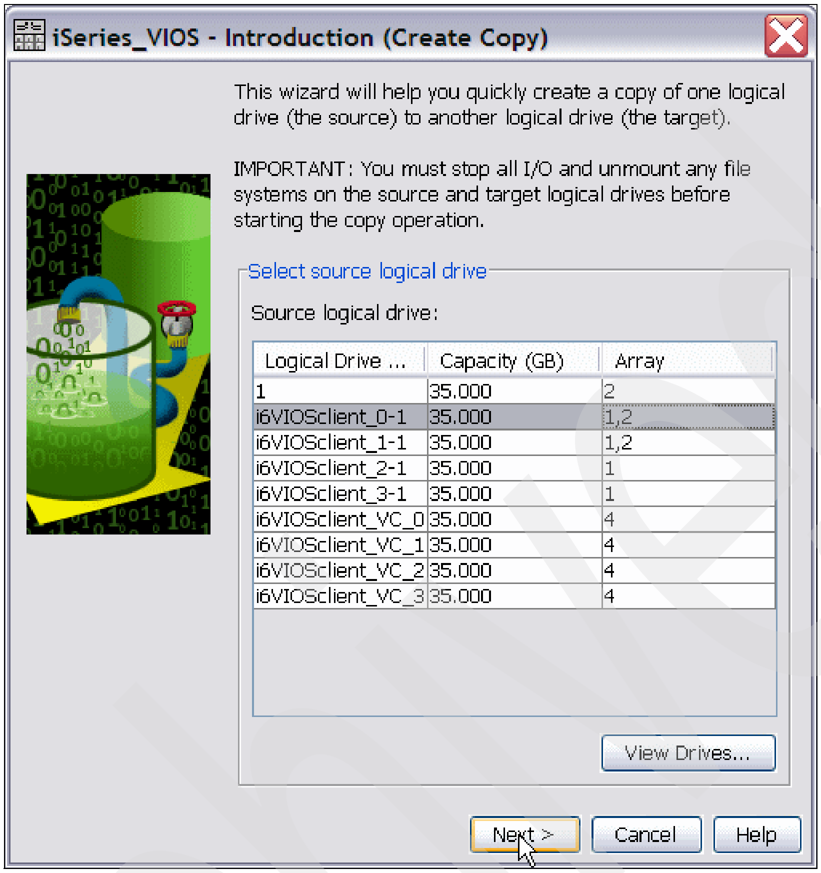


Figure 8-22 Select the Volume Copy source drive

6. On the next window select the Volume Copy target logical drive. See Figure 8-23.

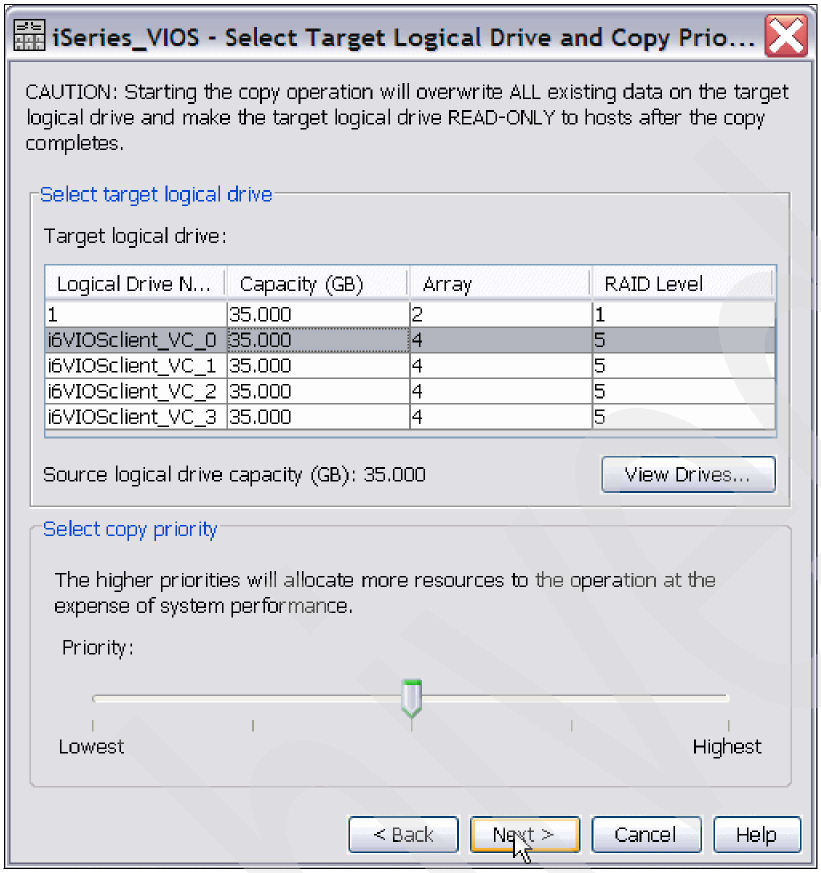


Figure 8-23 Select Volume Copy target drive

7. Type `Yes` in the confirmation window (Figure 8-24). In next window we specify Yes to copy another logical drive.

.

Figure 8-24 Confirm to create Volume Copy

Repeat steps 2–5 for each FlashCopy logical drive. During copying you may observe the progress by right-clicking **Volume Copy target drive** and selecting **Properties**. In the Logical drive properties window click the **Copying** tab (Figure 8-25).

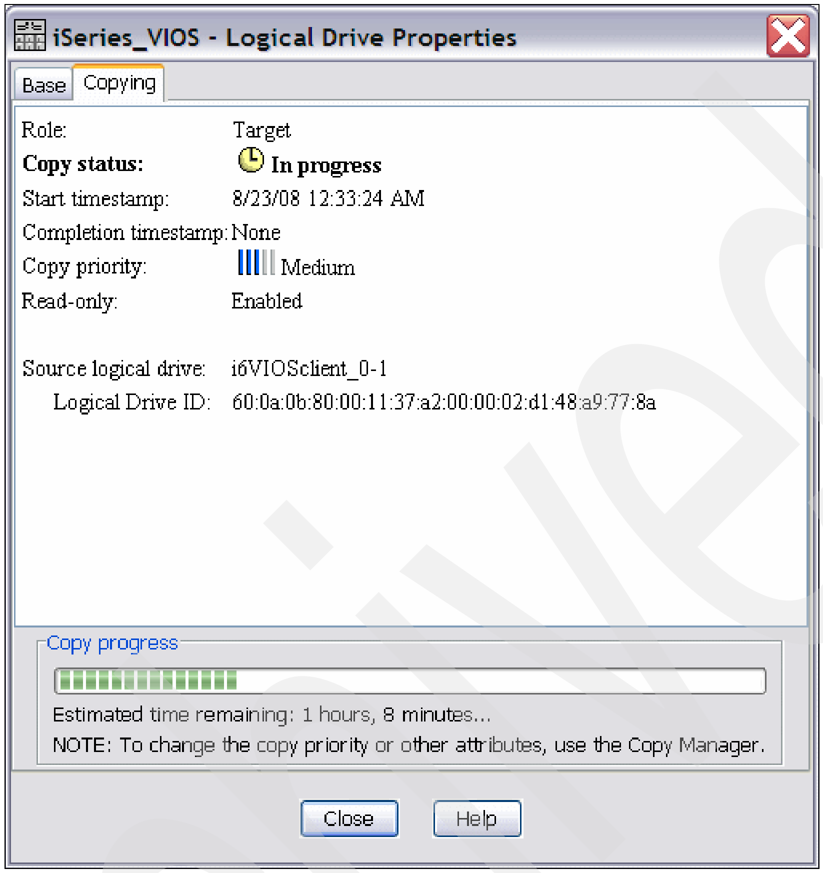


Figure 8-25 Copying in progress

8. After data are copied to Volume Copy logical drives. The Volume Copy drives are locked for writes. To make these available for both reads and writes, right-click a Volume Copy drive in the Subsystem Management Physical/Logical view, select **Copy Manager** from the pull-down, in the Copy Manager window select each logical drive, and click **Change** → **Target Logical Drive Permissions** → **Disable Read-Only**, as can shown in Figure 8-26.

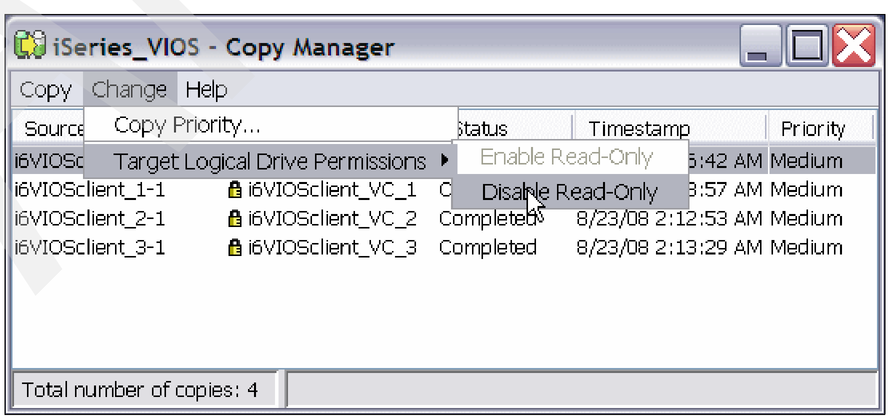


Figure 8-26 Disable Read-Only on Volume Copy logical drives

9. Connect Volume Copy logical drives to a System i partition. For this you must first map them to a VIOS host. In the Subsystem Management Mappings view right-click **VIOS host** and select **Define Additional Mapping**. Select the LUN ID and the drive to map, and click **Add**, as is shown in Figure 8-27. Repeat this step for each Volume Copy logical drive.

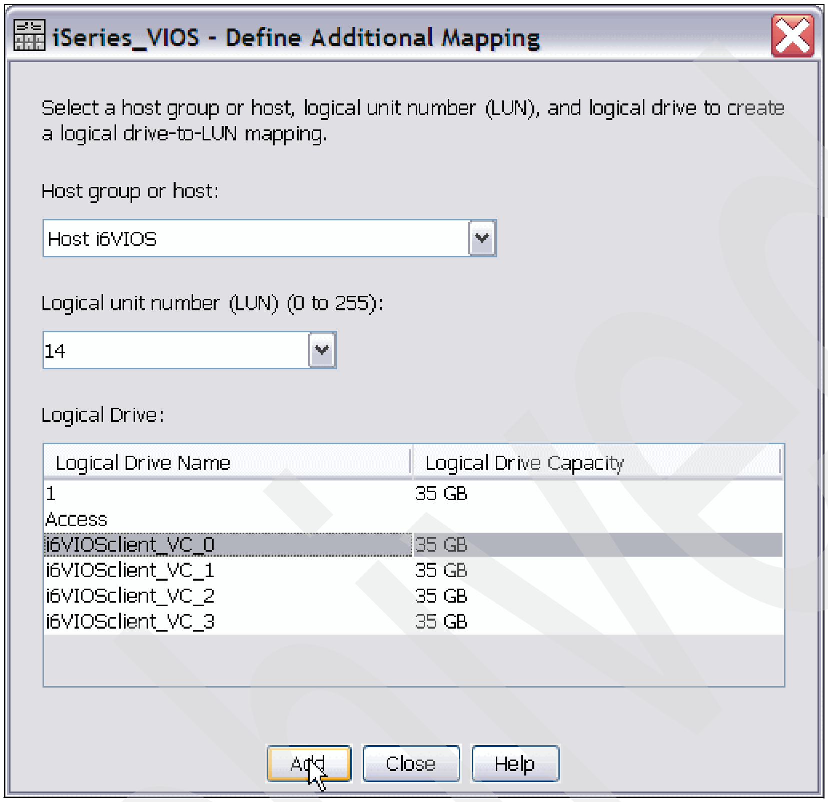


Figure 8-27 Map Volume Copy logical drives to VIOS

10. In VIOS, issue the command **`cfgdev`**, then map the newly added logical volumes to the virtual SCSI adapter assigned to the virtual adapter of System i partition by using the VIOS command:

```
mkdev -vdev hdiskxx -vadapter vhostx
```

For more information about VIOS commands refer to 6.5, “Configuring VIOS virtual devices” on page 191. In our setup we use Putty to connect to VIOS and map the virtual devices. Mapped logical disks in our setup are shown in Figure 8-28.

```
$ lsmap -vadapter vhost4
SVSA            Physloc                                        Client Partition
ID
--------------- -------------------------------------------- ------------------
vhost4          U9406.MMA.655A620-V2-C16                       0x00000008

VTD                   vtscsi7
Status                Available
LUN                   0x8100000000000000
Backing device        hdisk16
Physloc
U789D.001.DQDWVYP-P1-C3-T1-W201900A0B8470488-L14000000000000

VTD                   vtscsi8
Status                Available
LUN                   0x8200000000000000
Backing device        hdisk17
Physloc
U789D.001.DQDWVYP-P1-C3-T1-W201900A0B8470488-L15000000000000

VTD                   vtscsi9
Status                Available
LUN                   0x8300000000000000
Backing device        hdisk18
Physloc
```

Figure 8-28 Mapping of virtual devices in VIOS

11. Make sure that the relevant virtual SCSI adapter is tagged as the load source unit of the System i partition and activate the partition to IPL from Volume Copy logical drives.

    Note that after IPL the partition contains a clone of the production system. Therefore, mind the considerations described in 8.1.3, “Considerations when cloning IBM i system” on page 300.

8.3 Implementing metro mirroring

A Business Continuity solution with metro mirroring requires two site scenario. We expect that metro mirroring is implemented in an installation when the distance between two sites does not exceed 50 km. You may consider global mirroring for distances longer than 50 km.

The production system i partition reside on a local site. It uses disk space on the local Storage System connected via VIOS. On the remote site we implement a stand-by System i partition to which a Storage System is connected via VIOS. Metro mirroring is established between the two Storage Systems. In case of disaster or failure on the local site, the stand-by System i partition IPLs from the metro mirroring copy of the disk space and so brings up a clone of the production system.

Note: In our setup, we refer to the System i partition that uses FlashCopy base drives as the production partition. We call the partition that IPLs from the metro mirroring copy of logical drives the recovery partition.

Note: The Storage System that is being replicated is referred to as the *primary* Storage System, and its logical drives are refired to as *primary* logical drives. The Storage System to which the replication is done is called the *secondary* Storage System, and the logical drive *secondary* logical drives.

The solution with metro mirroring for all disk space of a System i partition consists of the following activities:

- Establish metro mirroring.
- Switch to the remote site at planned outages.
- Handle unplanned outages.
- Switch back to the local site.

In this section we describe each of these.

8.3.1 Establish metro mirroring

To establish metro mirror of System i full disk space:

1. Enable the premium feature for Enhanced Remote Mirroring. Do this the same way in which you enable the premium feature for FlashCopy, as described in “Creating FlashCopy the first time” on page 284. The only difference is that you insert the feature key for Enhance Remote Mirroring this time (Figure 8-2 on page 285).
2. Use Storage Manager and double-click the relevant Storage System to open the Subsystem Management window Logical/Physical view. Expand **Storage Subsystem** → **Remote Mirroring** → **Activate**, as shown in Figure 8-29. This starts the wizard to activate the Remote Mirroring feature.

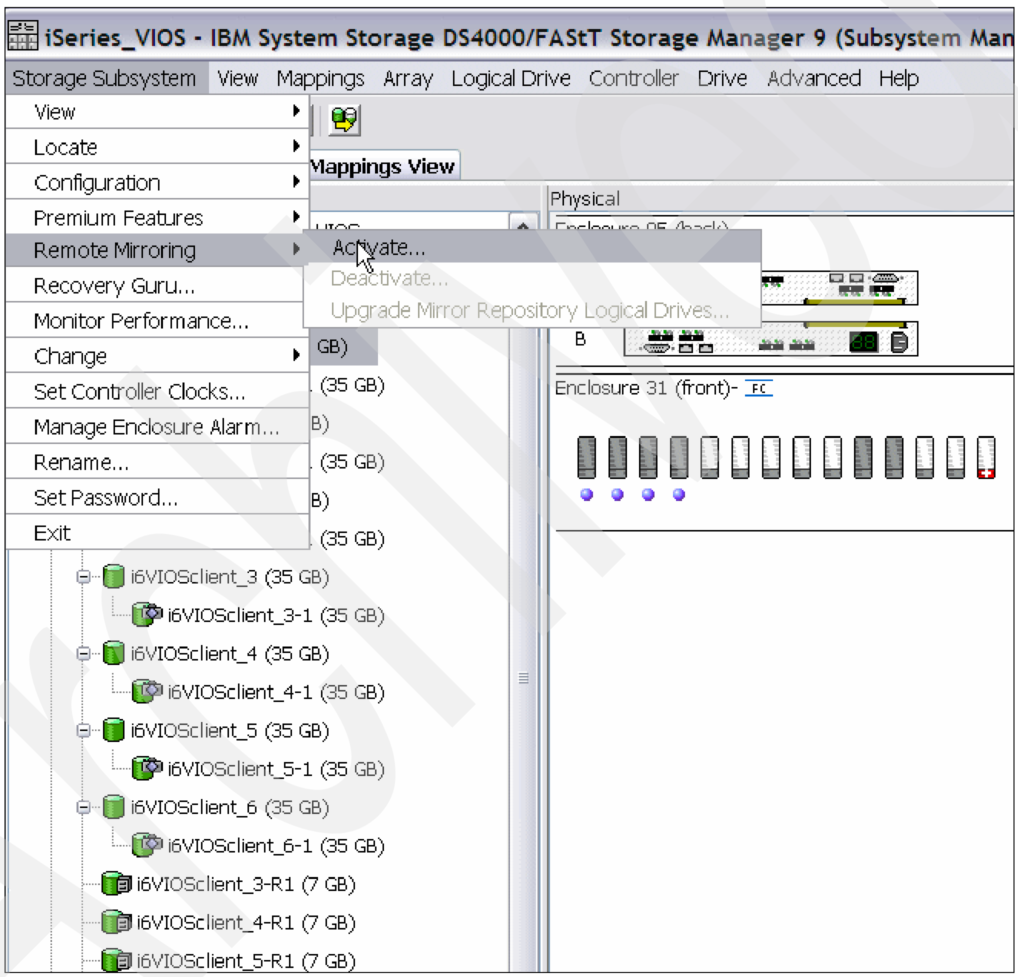


Figure 8-29 Activate remote mirroring

> **Note:** The term remote mirroring denotes all metro mirroring, global mirroring, and global copy.

3. While the remote mirroring feature is activated, the system assigns space for Remote Mirroring repository logical drives. Remote mirroring repository logical drives store the mirroring information, including information about the remote write request that has not yet been written to the secondary logical drive. Figure 8-30 shows the window where you specify from which array to take capacity for remote mirroring repository logical drives.

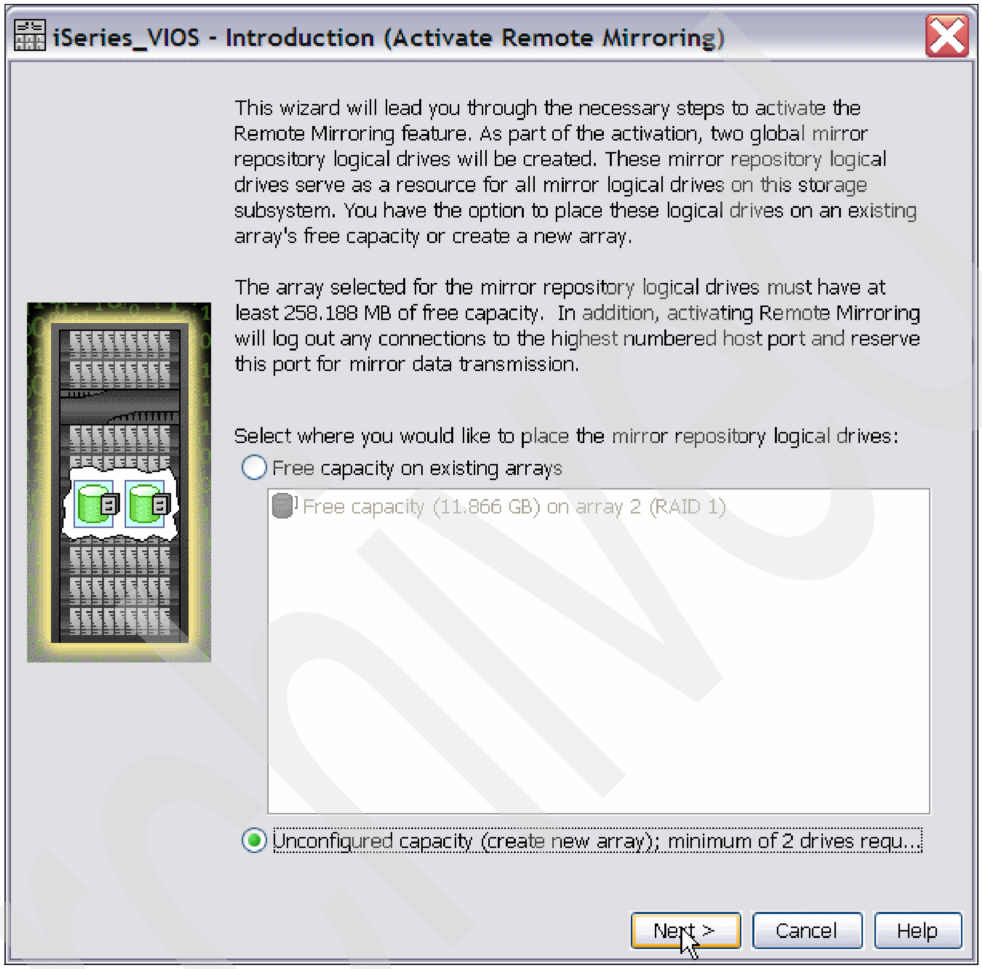


Figure 8-30 Specifying capacity for remote mirroring repository logical drives

4. On the next window you confirm the activating of the remote mirroring feature, as shown in Figure 8-31. When activating this feature two host ports with the highest port number, one on each controller, are assigned for remote mirroring replication.

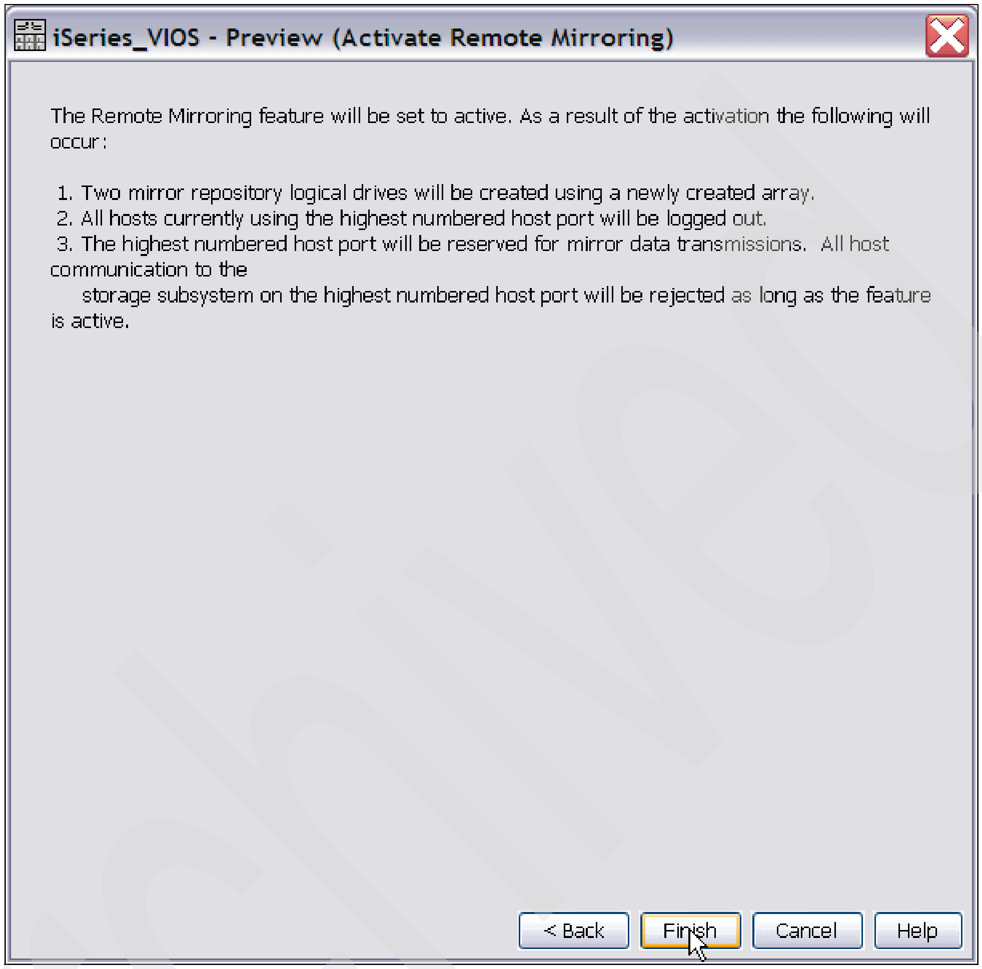


Figure 8-31 Confirm activating of the remote mirroring feature

5. To start metro mirroring of a logical drive right-click the logical drive in the Subsystem Management window Logical / Physical view and select **Create Mirror** from the pull-down, as shown in Figure 8-32. This invokes the wizard to create the remote mirror replication.

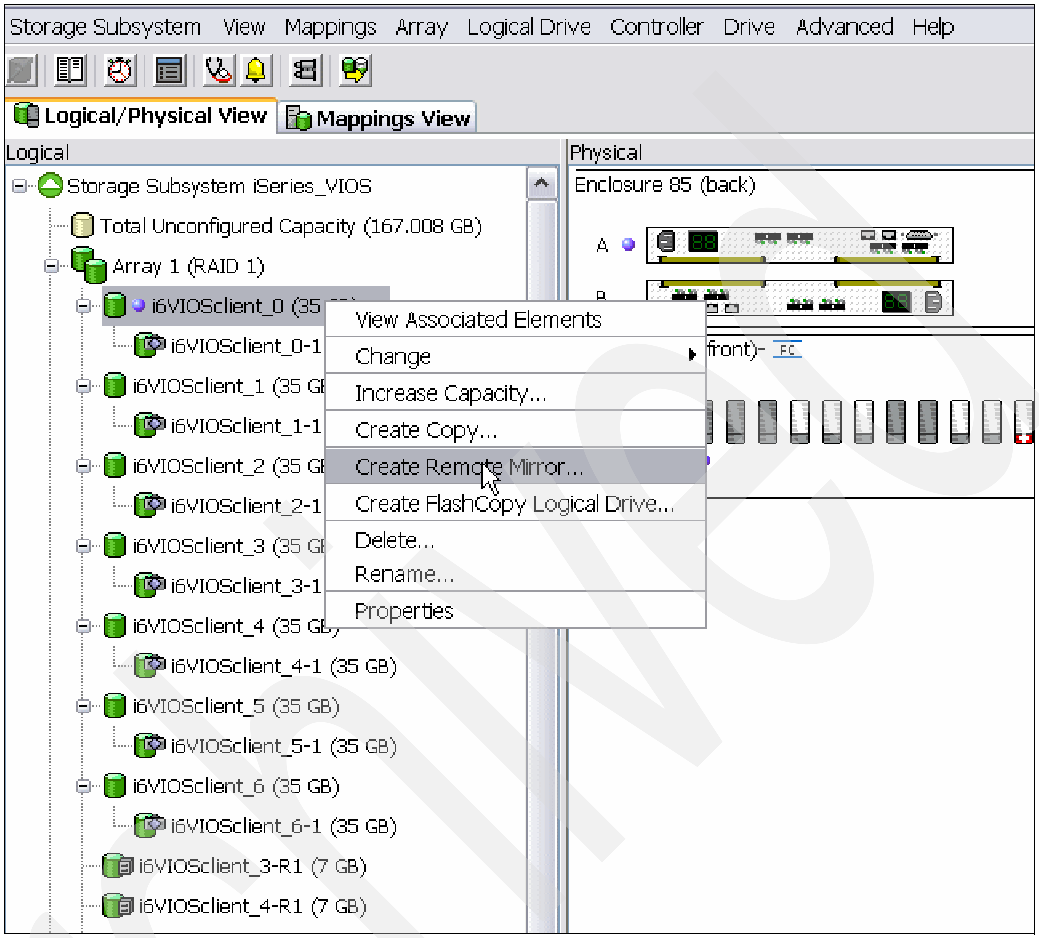


Figure 8-32 Create Remote Mirror

6. On next wizard window select the Storage System to contain secondary logical drives, as shown in Figure 8-33.

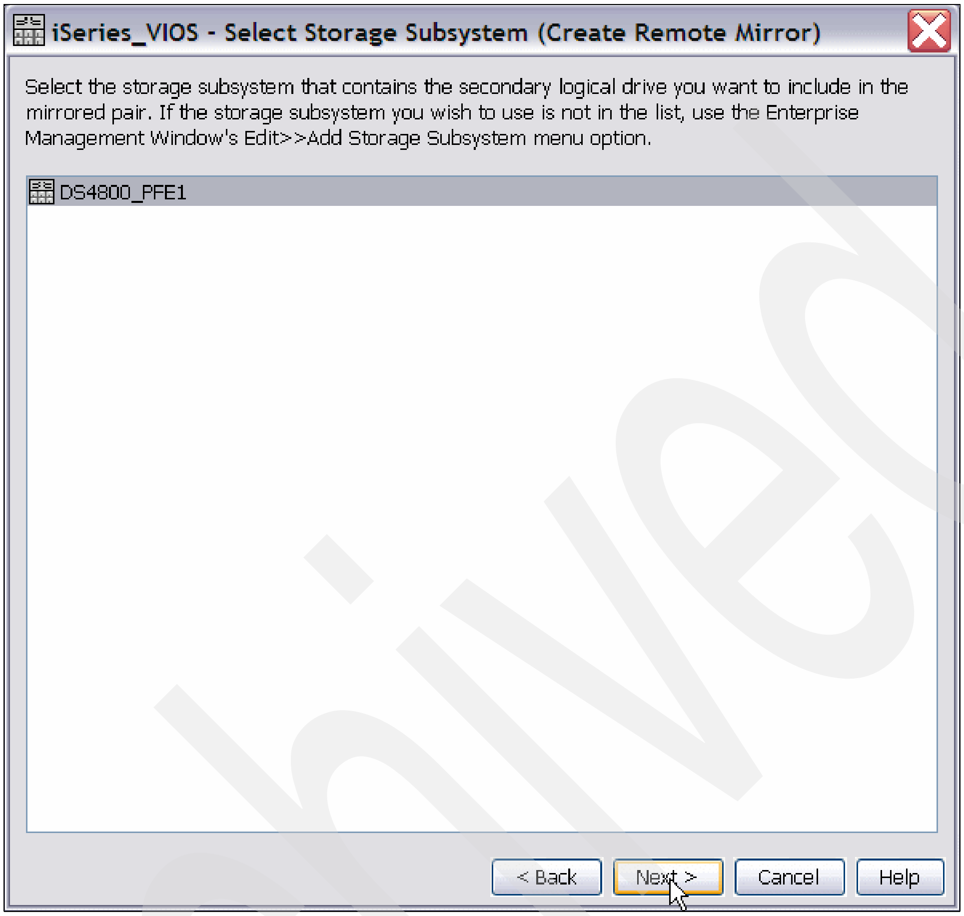


Figure 8-33 Specify secondary Storage System

7. In the next window select the secondary logical volume, as shown in Figure 8-34.

> **Note:** The secondary logical drives must be created before starting remote mirroring.

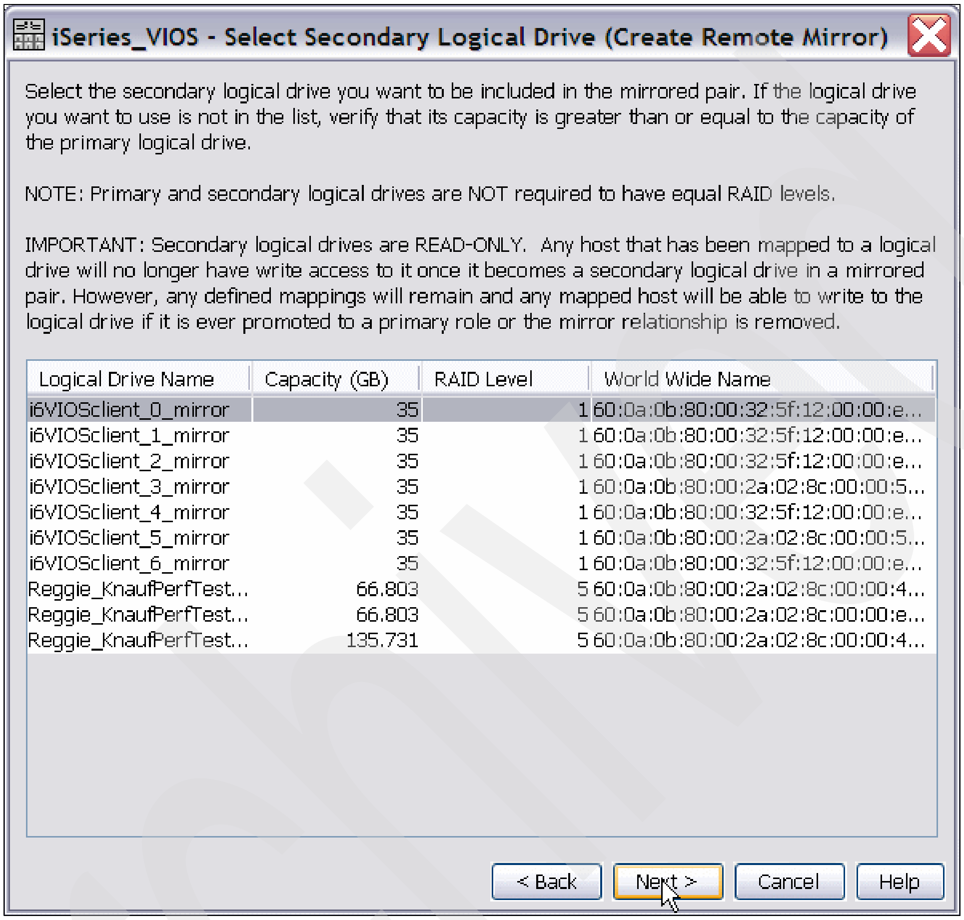


Figure 8-34 Select secondary logical drive

8. Choose the type of remote mirroring that you want to start. To establish metro mirroring select **Synchronous**, as shown in Figure 8-35.

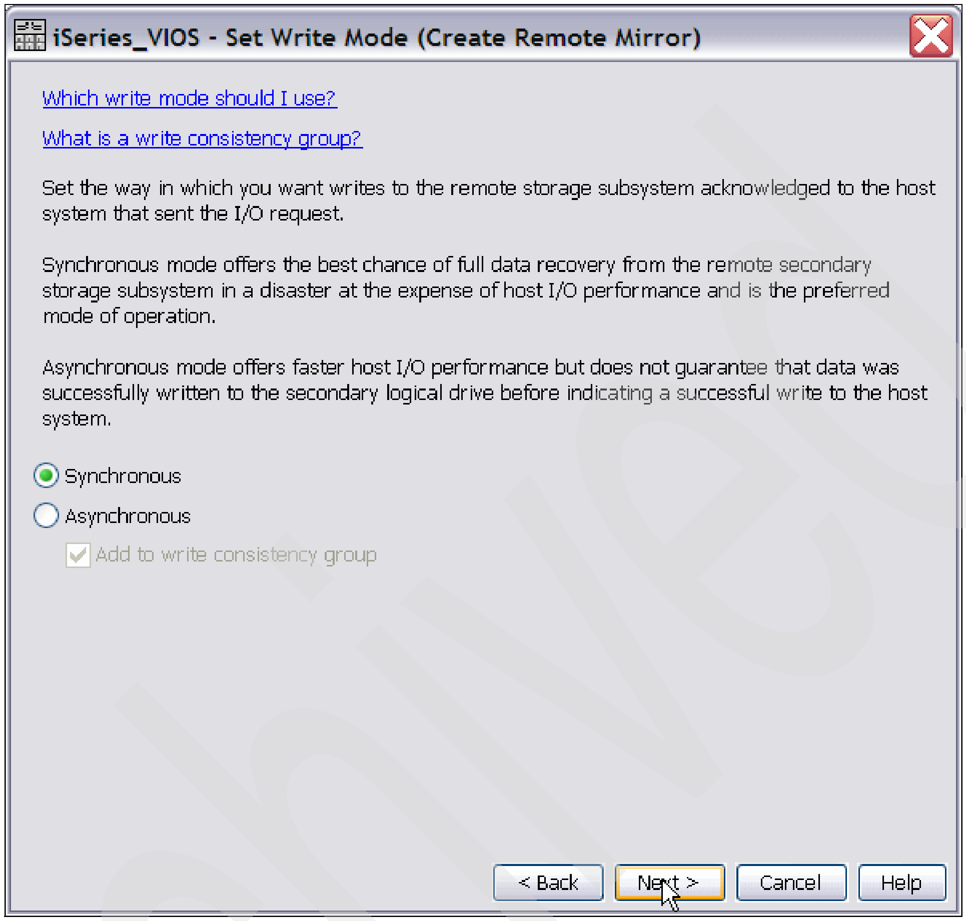


Figure 8-35 Select type of remote mirroring

9. On next window, specify the priority at which the system will allocate resources for initial Metro Mirroring synchronization, and for re-synchronization after suspend. Also, specify whether the re-synchronization after a suspend at link failure should be automatic, or you prefer to start it manually. See Figure 8-36.

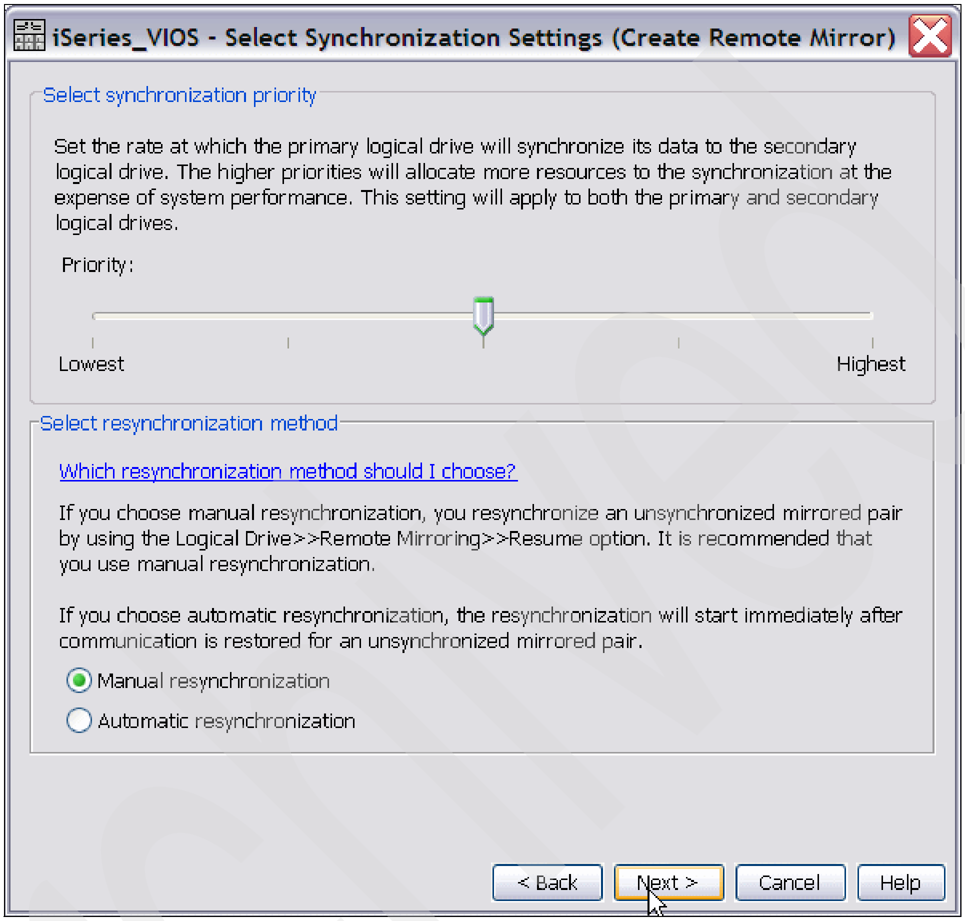


Figure 8-36 Specify priority of synchronization

10. On the next window you are asked to confirm the creating of metro mirroring. To confirm, type in `Yes` and click **Finish**, as shown in Figure 8-37.

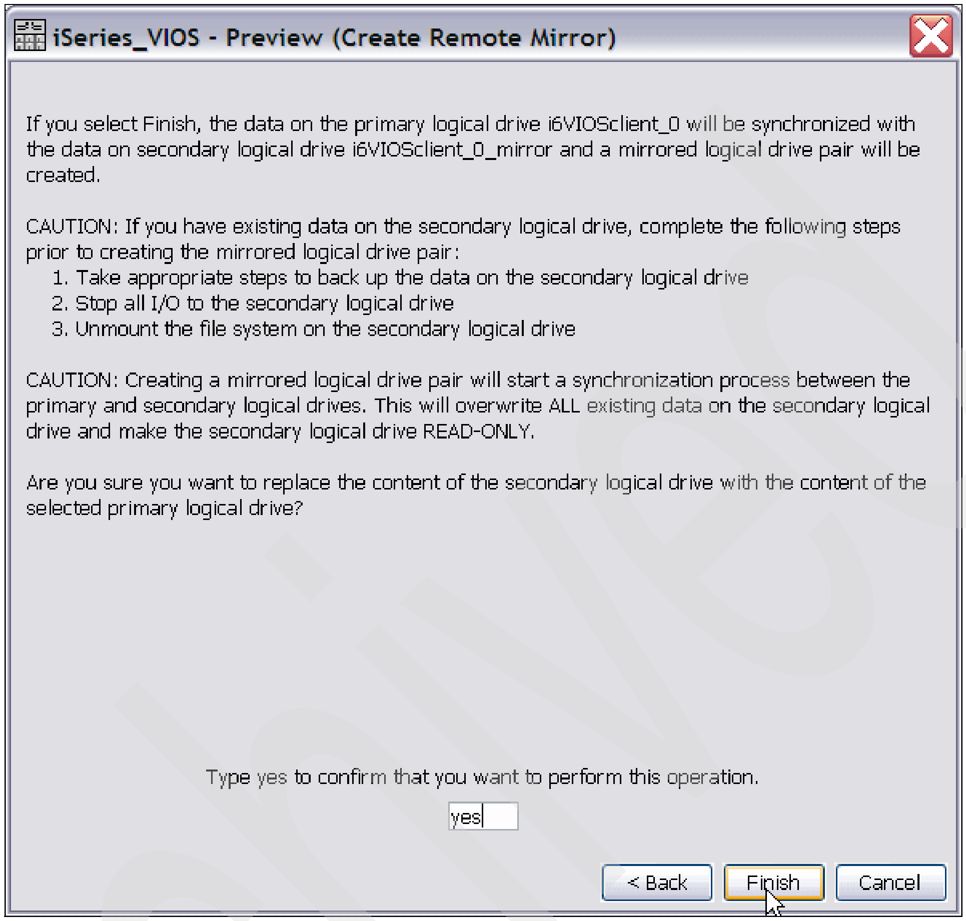


Figure 8-37 Confirm start of metro mirroring

The wizard informs you that metro mirroring is created as is shown on Figure 8-38.

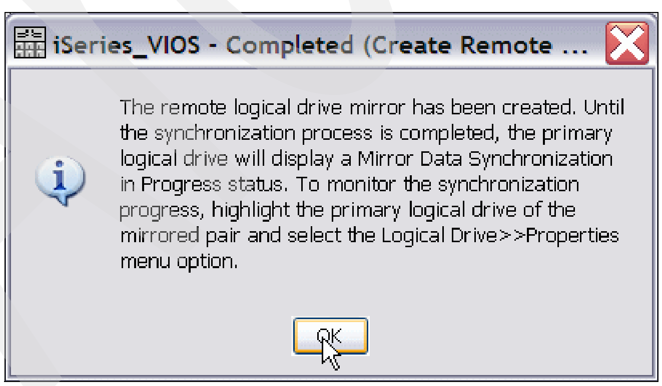


Figure 8-38 Created metro mirroring

11. After Metro Mirroring of the logical drive is established, you can choose to start it for another logical drive by specifying **Yes** in the next window. See Figure 8-39.

Figure 8-39 Create metro mirroring for another logical drive

12. Choose the logical drive for which you want to establish metro mirroring, as shown in Figure 8-40.

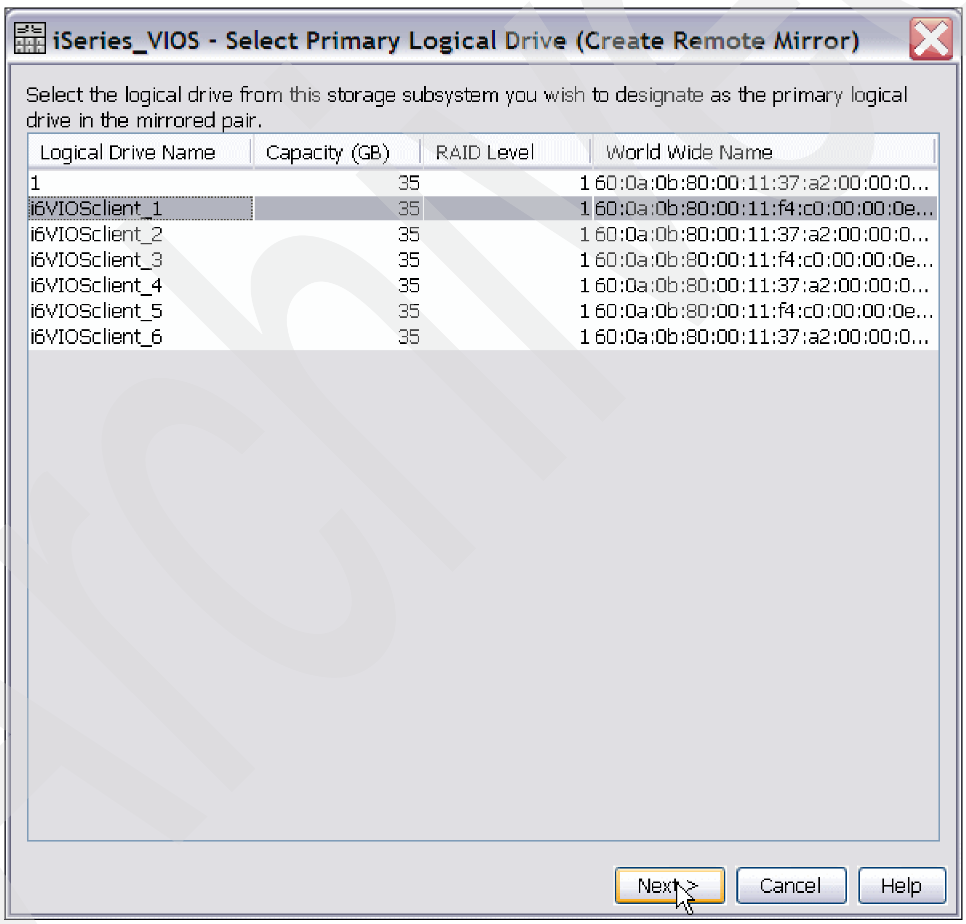


Figure 8-40 Select the logical drive for metro mirroring

13. Repeat steps 1–10 for each logical drive of the production System i partition.

After metro mirroring is established an entry for the secondary logical drive appears at each primary drive in the Subsystem Management Logical/Physical view.

During initial synchronization of metro mirroring, the secondary logical drives are in synchronizing status. You can observe this by right-clicking the entry of the secondary logical drive, selecting **Properties** from the pull-down, and selecting the **Mirroring** tab. A logical drive in synchronizing status is shown in Figure 8-41. After synchronization, if finished, the secondary logical drives show the status synchronized.

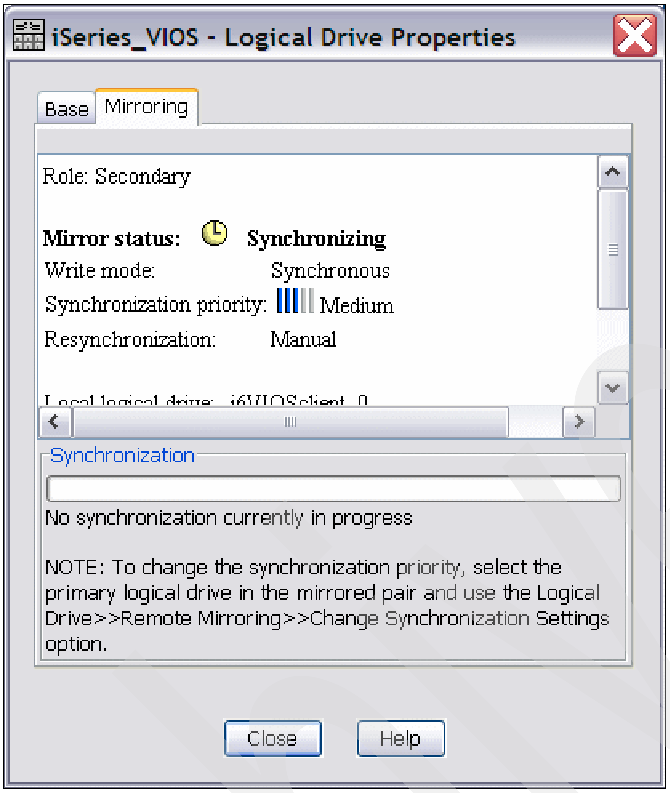


Figure 8-41 Metro mirroring secondary logical drive in synchronizing status

14. Map the secondary logical volumes to VIOS on remote Power6. For more information about logical drive mapping, refer to “Defining logical drive to LUN mapping” on page 274.

15. In VIOS, on the remote site map the LUNs to virtual SCSI adapters, as is described in 6.5, "Configuring VIOS virtual devices" on page 191. The virtual devices in VIOS in our setup are shown in Figure 8-42.

    Note that in our setup we established production and recovery partitions on the same POWER6 hardware. We use one VIOS to connect the production, backup, and recovery partitions, so the virtual devices of metro mirror secondary drives reports with higher numbers and are assigned higher LUN IDs.

```
vhost4         U9406.MMA.655A620-V2-C16                     0x00000008

VTD                   vtscsi7
Status                Available
LUN                   0x8100000000000000
Backing device        hdisk16
Physloc
U789D.001.DQDWVYP-P1-C3-T1-W201900A0B8470488-L14000000000000

VTD                   vtscsi8
Status                Available
LUN                   0x8200000000000000
Backing device        hdisk17
Physloc
U789D.001.DQDWVYP-P1-C3-T1-W201900A0B8470488-L15000000000000

VTD                   vtscsi9
Status                Available
LUN                   0x8300000000000000
Backing device        hdisk18
Physloc
U789D.001.DQDWVYP-P1-C3-T1-W201900A0B8470488-L16000000000000

VTD                   vtscsi10
Status                Available
```

Figure 8-42 Metro mirroring secondary logical drives as virtual devices in VIOS

8.3.2 Switch to remote site at planned outages

To switch to the remote site at planned outages of the production partition:

1. Power-down production partaking.
2. Reverse the direction of metro mirroring. In DS Storage Manager double-click the secondary Storage System to open the Subsystem Management window Logical/Physical view.

For each secondary logical drive right-click the drive, select **Change** from the pull-down, and select **Role to Primary** from the next pull-down, as shown in Figure 8-43. Confirm changing the role to primary by clicking **Yes** in the confirmation window (Figure 8-44 on page 325).

Note: After the metro mirroring direction is reversed the primary logical drives become unusable for the production partition, and the secondary logical drives can be used by the recovery partition.

Note: Reversing the direction of metro mirroring is instant, since it does not require synchronization of metro mirror.

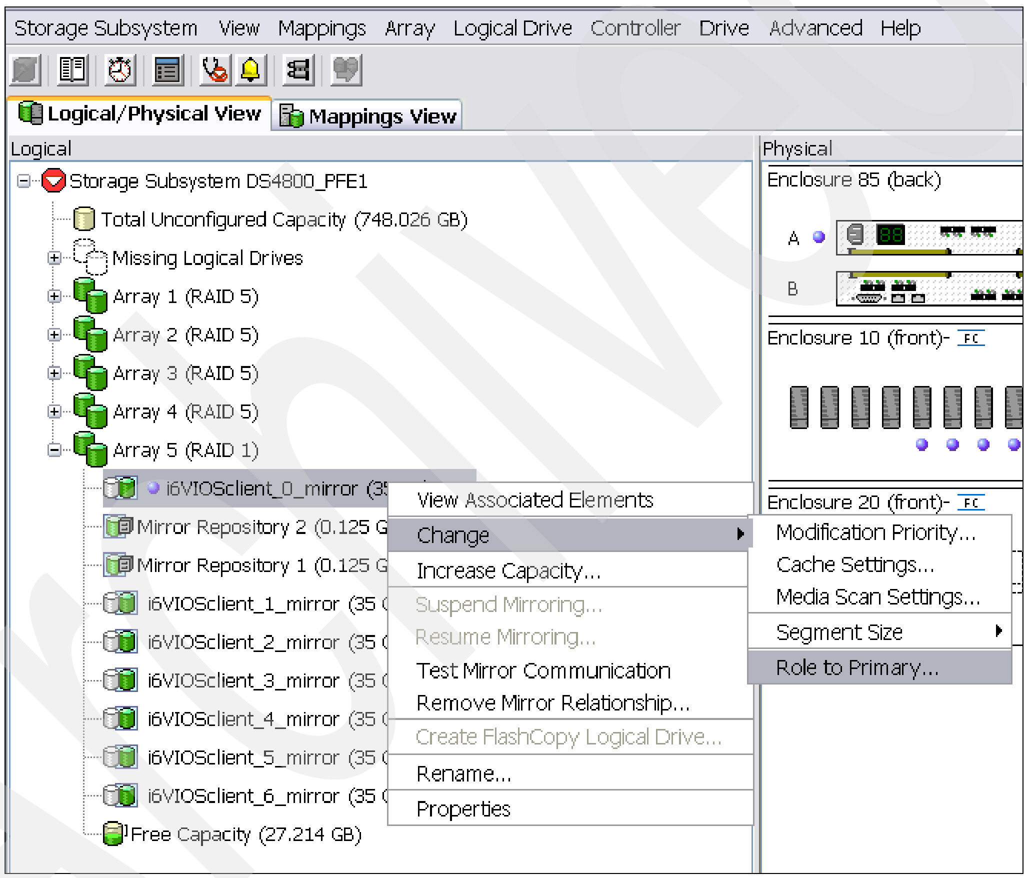


Figure 8-43 On remote Storage System change role to primary

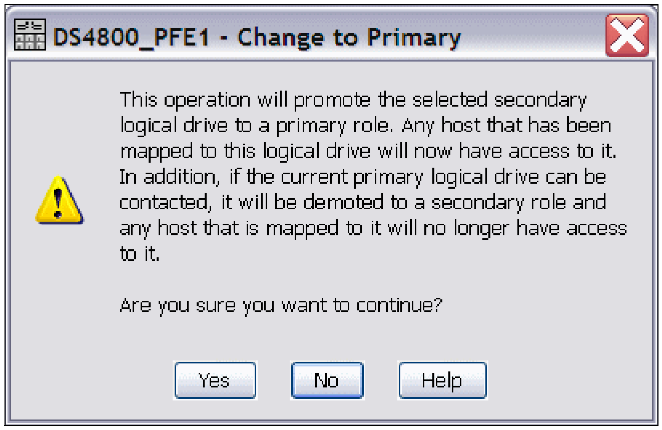


Figure 8-44 Confirm Change to Primary

After the role to primary is changed on all logical volumes, their icons show that the roles are reversed, as shown in Figure 8-45.

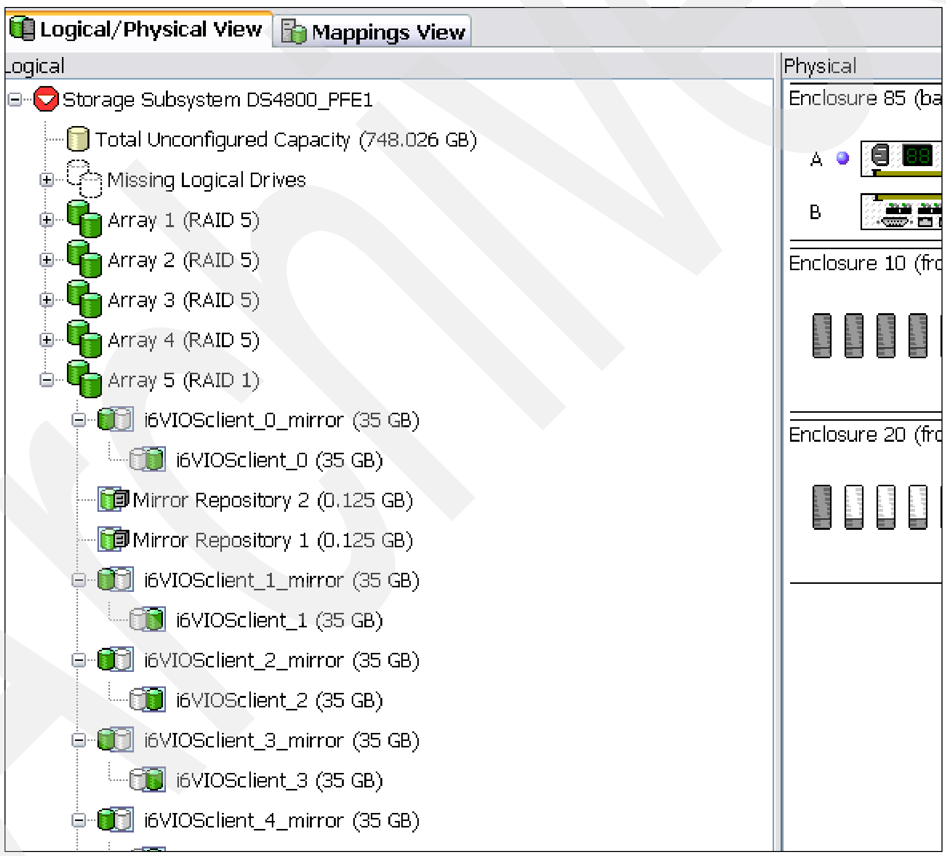


Figure 8-45 Changed role to primary

3. Since the secondary logical drives are already mapped to VIOS and virtual devices are created, you only need to enter the command **`cfgdev`** in VIOS to enable the changed logical drives to be recognized by the recovery partition.

4. IPL the recovery partition that is connected to metro mirroring secondary logical drives via VIOS. In HMC, check that the virtual SCSI adapter in the recovery partition is tagged for load source, as shown in Figure 8-46, and IPL the recovery partition.

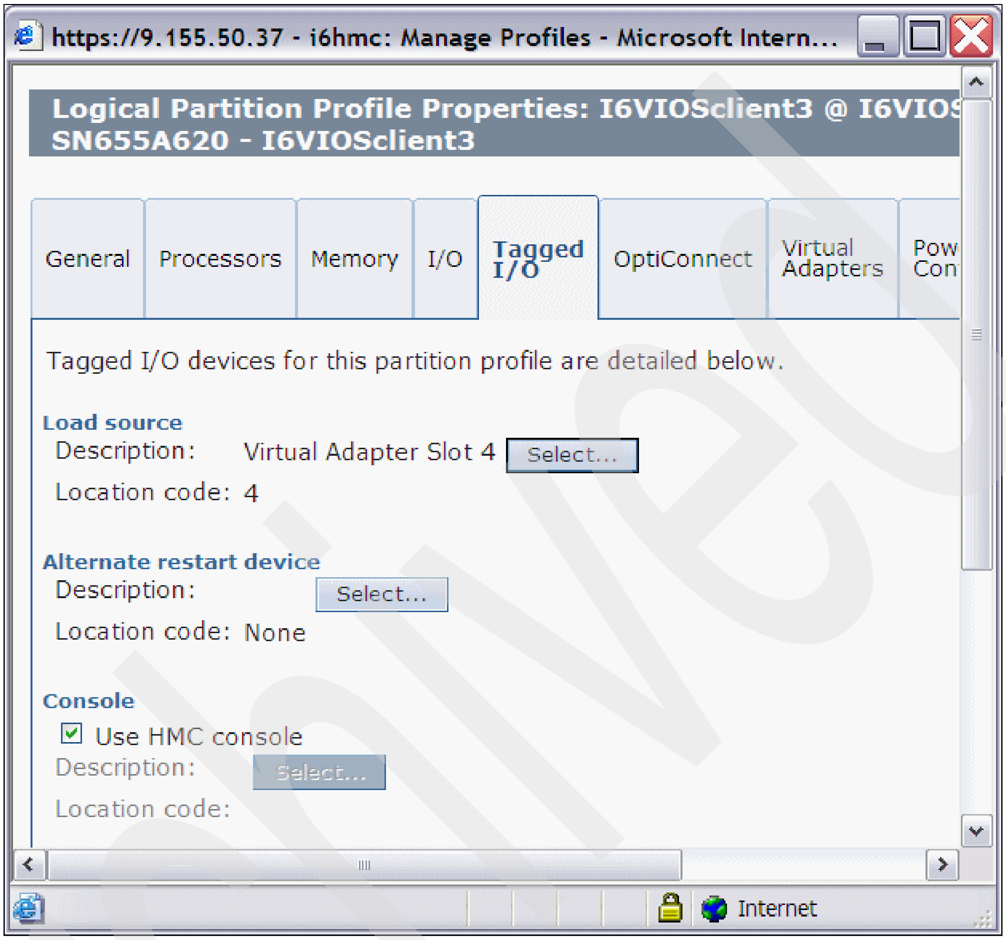


Figure 8-46 Virtual SCSI adapter tagged for load source

5. Once the recovery system is up and running it contains a clone of the production system. You should take into account the considerations for cloning the System i partition, described in 8.1.3, “Considerations when cloning IBM i system” on page 300.

8.3.3 Handling unplanned outages

In this section we describe the actions that you must perform at unplanned outages. We present some of unplanned outages that we consider the most typical.

Failure of metro mirroring link

If the Fibre Channel or IP link between the primary and the secondary Storage Systems fail, both primary and secondary logical drives become unsynchronized. Metro mirror primary logical drives continue to be available to the production partition, so production workload keeps running without any interruption. While the logical drives are unsynchronized, the writes to the primary Storage System are not replicated to the secondary Storage System. However, the changed data blocks are logged in the mirroring repository logical drive.

After the link is available again, re-synchronization of metro mirroring volumes starts automatically, or it must be started manually, depending which option was chosen when establishing metro mirroring. Since the changed blocks are kept in the repository volumes on the primary Storage System, synchronization is incremental. Only the data that was changed after the metro mirroring link failed are replicated to the secondary Storage System.

Failure of production System i partition

In case the production partition fails but the primary Storage System is running, perform the following actions:

1. Change the role of secondary logical drives to primary.
2. Issue the command **`cfgdev`** in VIOS on the remote site.
3. IPL the recovery partition.

This is described more in 8.3.2, “Switch to remote site at planned outages” on page 323. Most likely IPL will be abnormal since the secondary logical drives (System i disk units) on the remote site reflect the failed System i partition.

Failure of primary Storage System or disaster of entire local site

In case the primary Storage System fails or the entire local site is unavailable, metro mirror logical drives get unsynchronized. To recover on the remote site:

1. In DS Storage Manager double-click the secondary Storage System to open the Subsystem Management window Logical/Physical view.

   For each secondary logical drive right-click the drive, select **Change** from the pull-down, select **Role to Primary** from the next pull-down, click **Yes** on the confirmation window, and in the next window confirm to force the change to primary, as shown in Figure 8-47.

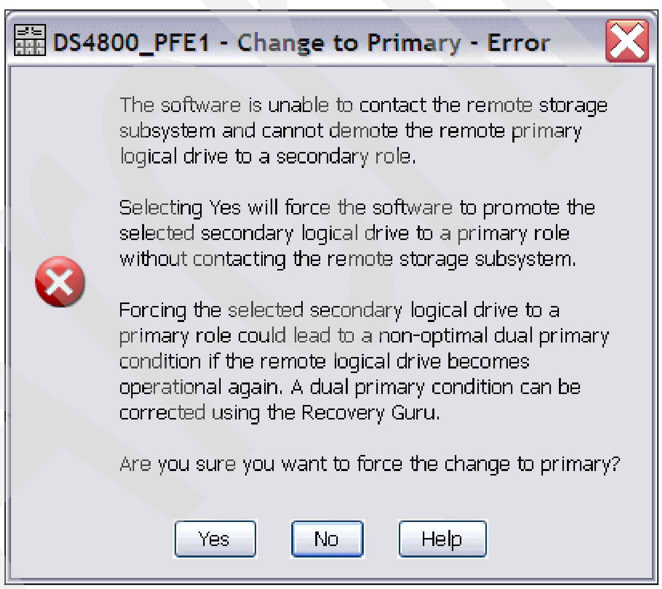


Figure 8-47 Force the change to primary

   Both metro mirroring logical drives on the local and the remote sites now have the primary role.

2. In VIOS on the remote site perform the command **`cfgdev`** to enable recovery partition to recognize metro mirror secondary LUNs.
3. IPL the recovery partition to bring up the clone of the production partition.

8.3.4 Switch back to local site

To switch back to the local site after planned outages or after failed local System i is working again:

1. Power down the recovery partition on the remote site.
2. Reverse the direction of metro mirroring. In DS Storage Manager double-click **local (current secondary)** Storage System to open the Subsystem Management window Logical/Physical view.

   For each secondary logical drive right-click the drive, select **Change** from the pull-down, and select **Role to Primary** from the next pull-down, as shown in Figure 8-48. Confirm to change role to primary by clicking **Yes** in the confirmation window.

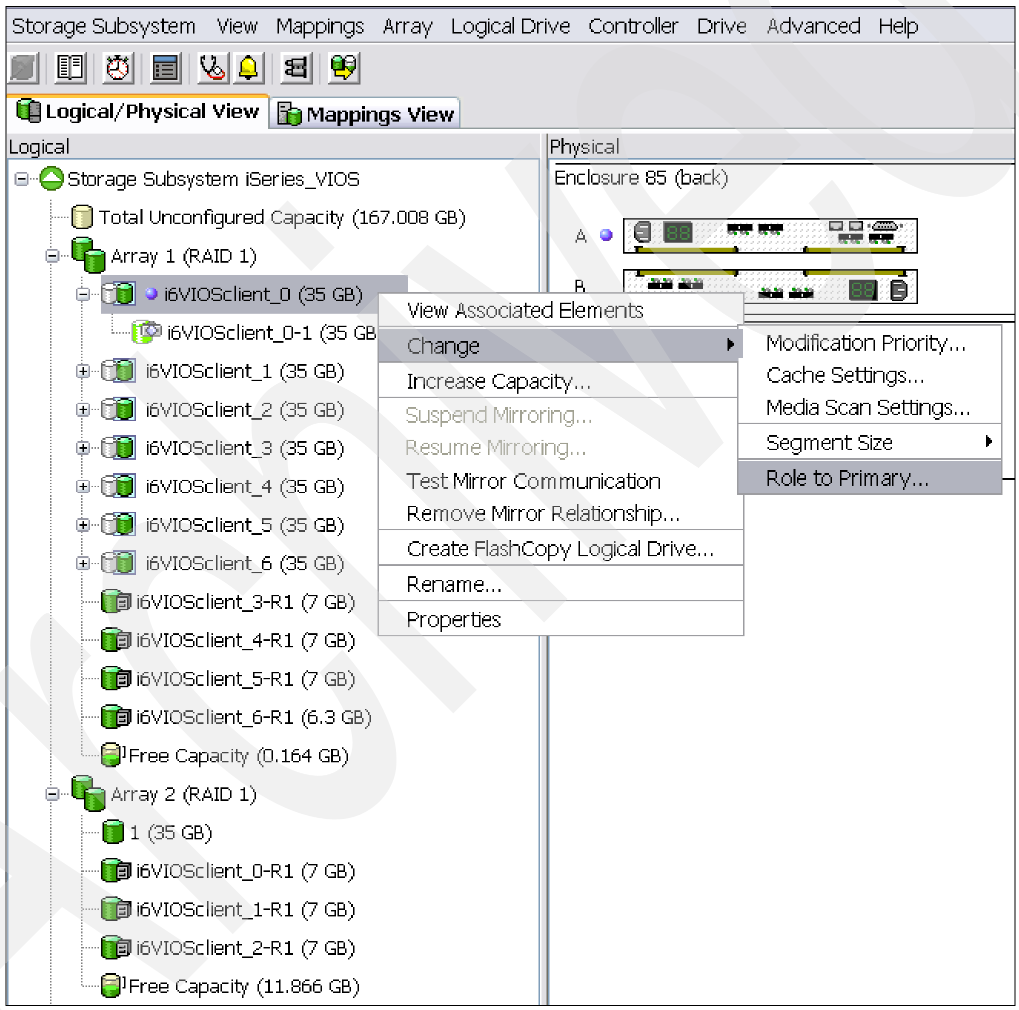


Figure 8-48 On local Storage System change role to primary

Alternatively, you can reverse the direction of metro mirroring by changing the role to secondary on each relevant logical drive on the remote (current primary) Storage System. To do this, right-click the logical drive in the Subsystem Management window Logical / Physical view, select **Change** from the pull-down, and select **Role to Secondary**, as shown in Figure 8-49. Confirm the change of the role to secondary by clicking **Yes** in the confirmation window.

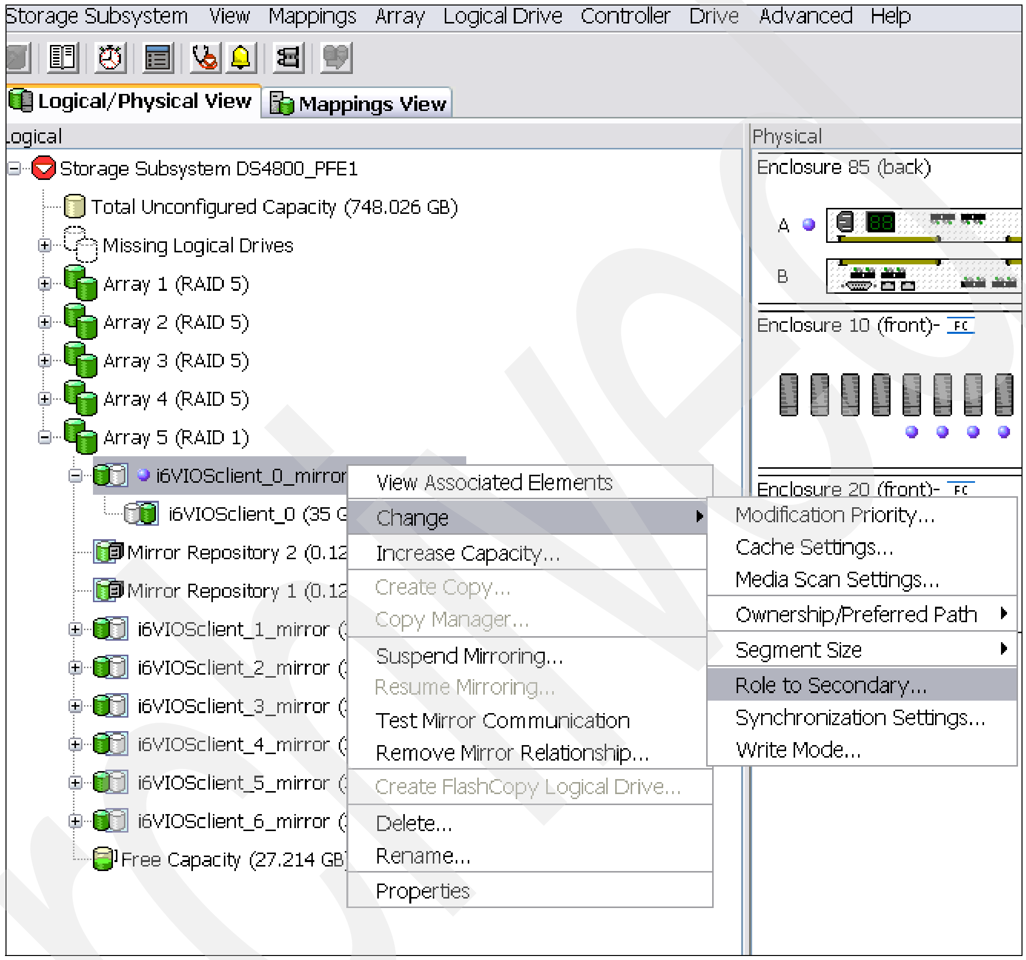


Figure 8-49 On remote Storage System change role to secondary

3. In VIOS, on the local site issue the command **cfgdev** to enable the logical drives to be recognized by the production partition.
4. IPL the production partition.

After failed local Storage System i is back again, or when the damaged site is back and configured, and the logical drives are still available on local site, perform the following steps to switch back to the local site:

1. Change the role to secondary on each relevant logical drive on the local site. For this, right-click the logical drive in the Subsystem Management window Logical / Physical view, select **Change** from the pull-down, and select **Role to Secondary**. Confirm that you want to change the role to secondary by clicking **Yes** in the confirmation window.

2. Resume metro mirroring on the remote site. To do this, right-click the logical drive in the Subsystem Management window Logical / Physical view and select **Resume Mirroring** from the pull-down, as shown in Figure 8-50. Confirm that you want to change the role to secondary by clicking **Yes** in the confirmation window.

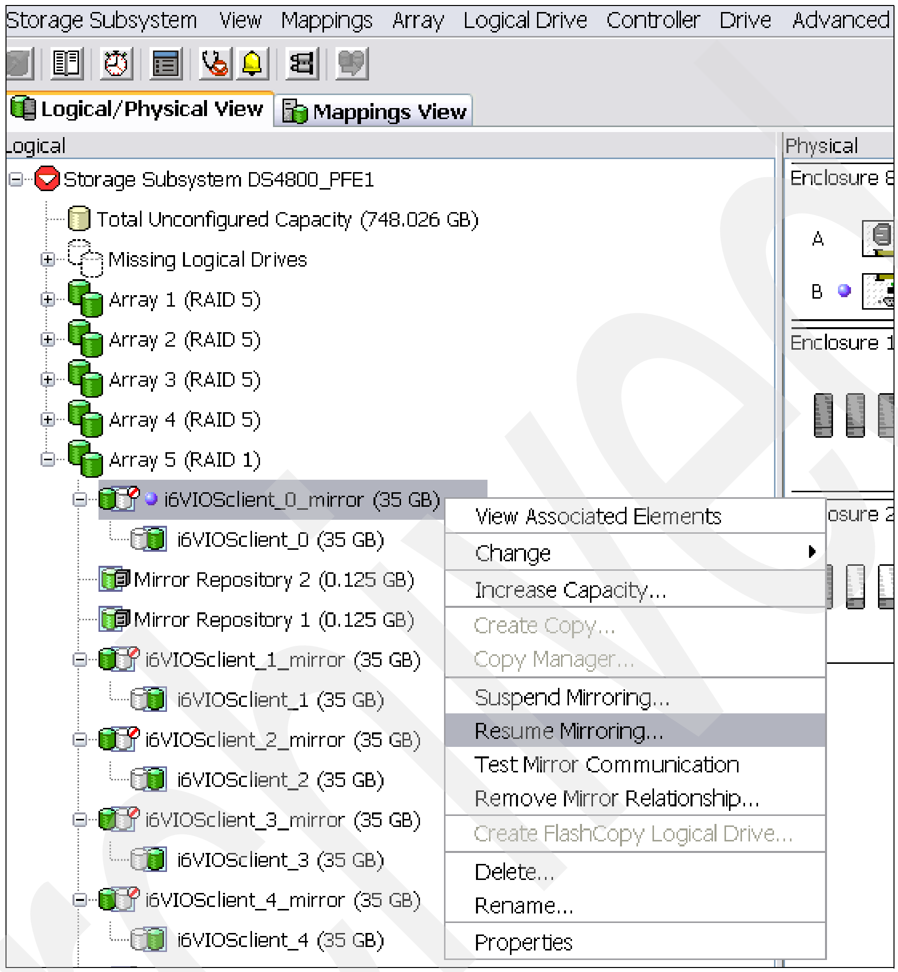


Figure 8-50 Resume metro mirroring

During metro mirroring pairs are unsynchronized. The changed data are recorded in repository logical drives on each Storage System. Resynchronization at resuming metro mirroring will therefore be incremental.

3. After logical drives are synchronized again, change the role of each relevant LUN in the local Storage System to primary, issue **`cfgdev`** in VIOS on the local site, and IPL the production partition, as described earlier in this section.

After the failed Storage System is back again, or the local site is available and configured again, and local Storage System is new or reconfigured, perform the following steps to fail back to the local site:

1. On the remote site terminate metro mirroring on each secondary logical drive by right-clicking the drive in the Storage Subsystem window and clicking **Remove Mirror Relationship**, as shown in Figure 8-51.

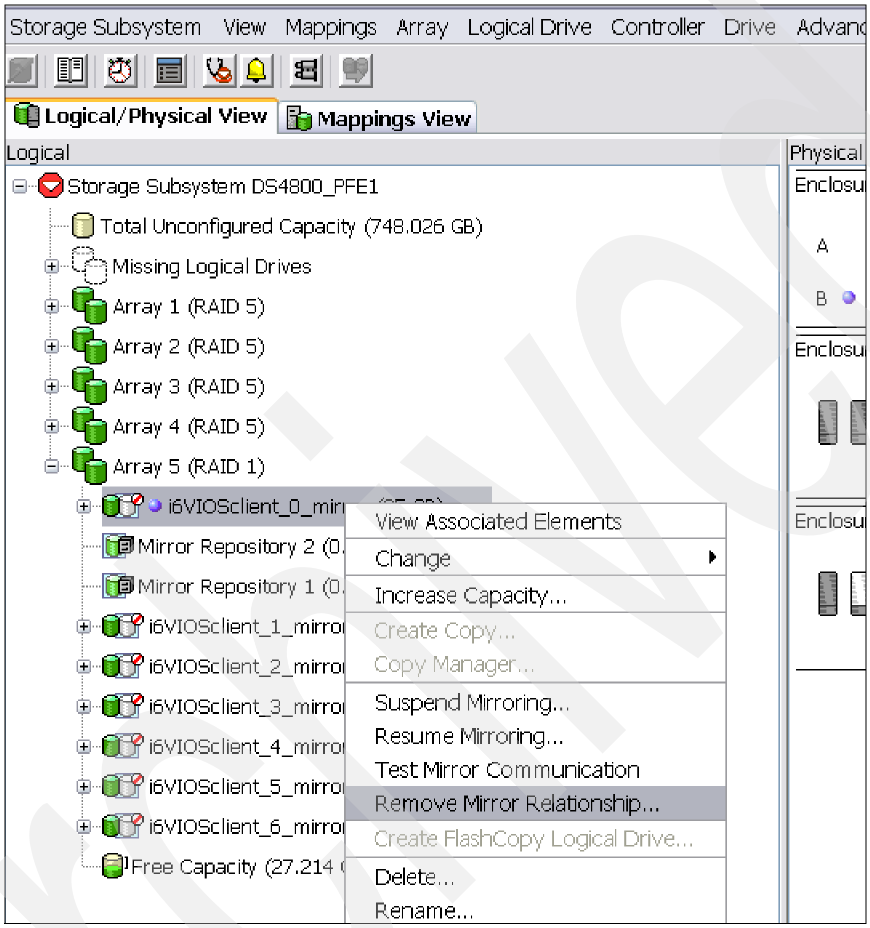


Figure 8-51 Remove Mirror Relationship

2. Establish metro mirroring from the remote to the local Storage System. For information about how to establish metro mirroring refer to “Establish metro mirroring” on page 312.
3. Switch from the remote to the local site. For information how to switch refer to “Switch to remote site at planned outages” on page 323.

9

Monitoring and managing Midrange Storage Solutions

In this chapter we describe common monitoring and managing tasks like performance monitoring, maintenance, and troubleshooting for IBM i, Virtual I/O Server (VIOS), and IBM Midrange External Storage.

9.1 Performance monitoring

In this section we introduce the tools that we recommend for monitoring I/O performance on IBM i, the IBM Virtual I/O Serve,r and the IBM DS Midrange External Storage subsystem.

9.1.1 Performance monitoring for IBM i

Consider using the tools discussed in this section for monitoring performance of an IBM i partition.

IBM Performance Tools for i5/OS

For IBM i performance monitoring we recommend using IBM Performance Tools for i5/OS (LPP 5761PT1) for collecting performance data via collection services at predefined intervals. We recommend choosing 5-minute intervals. All basic Performance Tools management tasks are available from the PERFORM menu, as shown in Figure 9-1.

```
PERFORM               IBM Performance Tools for i5/OS
                                                          System:   C605A620
Select one of the following:

     1. Select type of status
     2. Collect performance data
     3. Print performance report

     5. Performance utilities
     6. Configure and manage tools
     7. Display performance data
     8. System activity
     9. Performance graphics
    10. Advisor

    70. Related commands


Selection or command
===>

F3=Exit   F4=Prompt   F9=Retrieve   F12=Cancel   F13=Information Assistant
F16=System main menu
(C) COPYRIGHT IBM CORP. 1981, 2007.                                        +
```

Figure 9-1 IBM Performance Tools for i5/OS PERFORM menu

For further information about using IBM Performance Tools for i5/OS refer to the IBM Systems Information Center at:

http://publib.boulder.ibm.com/infocenter/systems/scope/i5os/index.jsp?topic=/rzahx/rzahxperftoolsdesc.htm

From the collection services data diverse disk performance reports can be generated out from the QAPMDISK member files of which a *system report* for disk utilization showing the service times and a *resource report* for disk utilization showing the I/O throughput per interval are of particular interest for IBM i virtual LUN I/O performance monitoring, as shown in Example 9-1 and Example 9-2.

Example 9-1 i5/OS Performance Tools System Report for disk utilization

```
System Report                                          8/21/08  1:01:1
                                                          Disk Utilization                                          Page 000
                                                               DS4800
 Member  . . . : Q233234620 Model/Serial  . : MMA/65-5A620       Main storage . . : 8192.0 MB  Started  . . . . : 08/20/08 23:46:2
   Library . . : QPFRDATA   System name . . : C605A620           Version/Release  :   6/ 1.0   Stopped  . . . . : 08/21/08 00:20:0
 Partition ID  : 003        Feature Code  . : 5461-7380          Int Threshold  . : 100.00 %
 Virtual Processors:   1    Processor Units :   1.00
          Unit              Size  IOP  IOP        Dsk CPU --Percent--  Op Per    K Per    - Average Time Per I/O --
   Unit   Name       Type   (M)   Util Name       Util    Full   Util  Second     I/O     Service   Wait   Response
   ----   ---------- ---- ------- ---- ---------- ------- ----   ---- -------- ---------  -------  ------  --------
  ASP ID/ASP Rsc Name:  1/
   0001   DD001      6B22  33,405   .0 CMB01           .0 37.4    4.3    89.29      12.2    .0005   .0000     .0005
   0002   DD006      6B22  33,405   .0 CMB01           .0 23.8    9.0   188.65      11.5    .0004   .0000     .0004
   0003   DD005      6B22  33,405   .0 CMB01           .0 23.8    9.6   223.25      10.8    .0004   .0000     .0004
   0004   DD002      6B22  33,405   .0 CMB01           .0 23.8    9.2   205.09      11.0    .0004   .0000     .0004
   0005   DD003      6B22  33,405   .0 CMB01           .0 23.8    9.5   222.62      10.7    .0004   .0000     .0004
   0006   DD007      6B22  33,405   .0 CMB01           .0 23.8    9.0   195.87      10.7    .0004   .0000     .0004
   0007   DD004      6B22  33,405   .0 CMB01           .0 23.8    8.5   199.05      11.2    .0004   .0000     .0004
 Total for ASP ID:  1     233,835

 Average                                                  25.8    8.5   185.19      11.1    .0004   .0000     .0004
 Total                    233,835
 Average                                                  25.8    8.5   185.19      11.1    .0004   .0000     .0004
 Unit                    --  Disk arm identifier
 Unit Name               --  Disk arm resource name
 Type                    --  Type of disk
 Size (M)                --  Disk space capacity in millions of bytes
 IOP Util                --  Percentage of utilization for each Input/Output Processor
 IOP Name                --  Input/Output Processor resource name
 Dsk CPU Util            --  Percentage of Disk Processor Utilization
 ASP Rsc Name            --  ASP resource name to which the disk unit was allocated at collection time
 ASP ID                  --  Auxiliary Storage Pool ID
 Percent Full            --  Percentage of disk space capacity in use
 Percent Util            --  Average disk operation utilization (busy)
 Op per Second           --  Average number of disk operations per second
 K Per I/O               --  Average number of kilobytes (1024) transferred per disk operation
 Average Service Time    --  Average disk service time per I/O operation
 Average Wait Time       --  Average disk wait time per I/O operation
 Average Response Time   --  Average disk response time per I/O operation
```

Example 9-2 i5/OS Performance Tools Resource Report for disk utilization summary

```
Resource Interval Report                                    08/21/08 00:38:3
                                                       Disk Utilization Summary                                          Page
                                                               DS4800
  Member  . . . : Q233234620 Model/Serial  . : MMA/65-5A620        Main storage  . . : 8192.0 MB Started . . . : 08/20/08 23:46:20
    Library . . : QPFRDATA   System name . . : C605A620            Version/Release . :     6/1.0 Stopped . . . : 08/21/08 00:20:00
  Partition ID  : 003        Feature Code  . : 5461-7380
                             Average      Average       Average                             High      High      High       Disk
   Itv        Average         Reads        Writes        K Per       Avg       High       Util      Srv       Srv       Space
   End        I/O /Sec        /Sec          /Sec          I/O        Util      Util       Unit      Time      Unit     Used (GB)
  -----      ---------      --------      --------      -------     -----     -----      ----      -----     ----     -----------
  23:50       1,310.6       1,146.4        164.1         11.2        8.1        9.6      0002      .0005     0002         56.407
  23:55       1,339.2       1,148.3        190.8         11.0        8.2        9.6      0003      .0004     0007         56.379
  08/21/08
  00:00       1,315.9       1,137.3        178.5         11.1        8.2        9.9      0005      .0004     0007         56.348
  00:05       1,335.1       1,127.7        207.3         11.1       10.6       12.3      0003      .0010     0001         56.321
  00:10       1,316.0       1,138.4        177.6         11.1        8.1        9.5      0005      .0004     0007         56.291
  00:15       1,318.3       1,141.2        177.0         11.1        8.0        9.2      0004      .0004     0007         56.245
  00:20       1,328.1       1,150.1        177.9         11.0        8.0        9.7      0004      .0004     0007         56.214
             ---------      --------      --------      -------     -----
  Average:    1,323.8       1,141.1        182.6         11.1        8.5
  Itv End                -- Interval end time (hour and minute)
  Average Phys I/O /Sec  -- Average number of physical I/O operations per second
  Average Reads / Sec    -- Average number of physical reads per second
  Average Writes /Sec    -- Average number of physical writes per second
  Average K Per I/O      -- Average number of kilobytes (1024) per I/O operation
  Avg Util               -- Average percent utilization of all disk arms
```

```
High Util              -- Highest percent utilization for a disk arm
High Util Unit         -- Disk arm with the highest utilization percent
High Srv Time          -- Highest average service time in seconds
High Srv Unit          -- Disk arm with the highest service time
Disk Space Used        -- Total disk space used in Gigabytes (1024)
```

For performance monitoring we recommend observing the following values in collection services reports:

- Service and wait time

  Service and wait time constitute the disk response time and are an immediate indication of possible disk performance problems. A high amount of wait time compared to the service time typically indicates an I/O concurrency problem causing I/O queuing on the IBM i host.

- (Percent) disk utilization

  The disk utilization percentage value reflects the utilization (better called *busy rate*) of the complete I/O path for a LUN. A high disk utilization may indicate a storage subsystem performance problem if service time is also high, but could also be related to an improperly sized configuration of too few IBM i LUNs configured but not providing enough I/O concurrency.

- I/O per second

  This value is an important indicator about the I/O workload on external storage. You may want to follow IO/sec during five consecutive days during heavy end-of-month jobs to capture the pattern of I/O rate and peaks.

- Read/write ratio and KB per I/O

  These values provide useful write data throughput information for sizing remote replication links. The read/write ratio is also an important factor to consider when deciding for a certain RAID level configuration. For example, for a high write percentage using RAID-5 may not be a good choice for performance.

IBM i5/OS Disk Watcher and IBM i commands

It may be a good idea to use IBM i5/OS Disk Watcher and IBM i commands WRKDSKSTS, WRKSYSSTS, and WRKSYSACT, along with IBM Performance Tools for i5/OS, or in case Performance Tools is not installed on the relevant System i partition.

IBM i5/OS Disk Watcher is provided as a part of IBM i operating system. It allows you to obtain data concerning I/O operations to disk units, along with frequently needed run-time data to determine which objects, files, processes, threads, and tasks are being accessed. This tool surfaces data beyond what is provided by such tools as WRKDSKSTS, WRKSYSSTS, and WRKSYSACT. Disk Watcher provides a mechanism to use short and longer duration traces to collect disk I/O data along with the associated task and object name.

Some potential uses of this tool are:

- Evaluating the performance of I/O operations on multi-path disk units
- Evaluating the performance of I/O queuing
- Determining how performance may be improved by re-spreading data across units
- Determining the optimal placement of devices, IOAs, or buses

IBM i5/OS Disk Watcher can be configured and managed through the IBM Systems Director Navigator for i5/OS Performance interface or CL commands.

For more information about IBM i5/OS Disk Watcher refer to the System i Information Center on the following Web page:

http://publib.boulder.ibm.com/iseries/

IBM iDoctor for IBM i

You may consider using iDoctor along with Performance Tools for solving complex System i performance issues.

The iDoctor for IBM i consists of a variety of software tools for managing performance, two of them being as follows:

- Job Watcher: Job Watcher displays real-time tables and graphical data that represent, in a very detailed way, what a job is doing and why it is not running.
- PEX Analyzer: PEX Analyzer evaluates the overall performance of the system. It condenses volumes of trace data into reports that can be graphed or viewed.

For more information about IBM iDoctor for IBM i refer to the following Web page:

https://www-912.ibm.com/i_dir/idoctor.nsf/jw.html

9.1.2 Performance monitoring for VIOS

For CPU usage monitoring and disk I/O performance monitoring under VIOS we recommend using the **viostat** command after enabling disk I/O statistics data collection, as shown in Example 9-3.

The parameters that we used for the **viostat** command are:

- The -tty parameter for monitoring CPU usage by the VIOS kernel reported in the %sys value
- The -extdsk parameter for extended disk I/O monitoring including service times
- The -time parameter to add the time stamp of the ending interval to each output
- 60 for specifying a 60-second I/O statistics interval

Example 9-3 VIOS performance monitoring using the viostat command

```
$ lsdev -dev sys0 -attr iostat
value

false
$ chdev -dev sys0 -attr iostat=true
sys0 changed

$ lsdev -dev sys0 -attr iostat
value

true

$ viostat -tty -time 60
System configuration: lcpu=2

tty:      tin         tout    avg-cpu: % user % sys % idle % iowait  time
          0.0          0.0               0.1    4.7   95.3      0.0  14:00:45
          0.0          3.1               0.0    3.6   96.4      0.0  14:01:45
          0.0          1.3               0.1    3.8   96.1      0.0  14:02:45

$ viostat -extdisk 60
System configuration: lcpu=2 drives=21 paths=2 vdisks=19

hdisk0         xfer:  %tm_act      bps      tps      bread      bwrtn      time
                         0.1      1.1K     0.2        0.0        1.1K  14:19:42
```

```
                read:      rps  avgserv  minserv  maxserv   timeouts      fails
                           0.0      0.0      0.0      0.0          0          0
               write:      wps  avgserv  minserv  maxserv   timeouts      fails
                           0.2      4.5      1.3     20.0          0          0
               queue:  avgtime  mintime  maxtime  avgwqsz    avgsqsz     sqfull
                           0.0      0.0      2.5      0.0        0.0        0.0
hdisk1          xfer:  %tm_act      bps      tps      bread      bwrtn       time
                           0.1     1.1K      0.2        0.0       1.1K   14:19:42
                read:      rps  avgserv  minserv  maxserv   timeouts      fails
                           0.0      0.0      0.0      0.0          0          0
               write:      wps  avgserv  minserv  maxserv   timeouts      fails
                           0.2      5.5      1.7     21.2          0          0
               queue:  avgtime  mintime  maxtime  avgwqsz    avgsqsz     sqfull
                           0.0      0.0      4.2      0.0        0.0        0.0
dac1            xfer:  %tm_act      bps      tps      bread      bwrtn       time
                           0.0     8.0M    681.9       2.6M       5.4M   14:19:42
                read:      rps  avgserv  minserv  maxserv   timeouts      fails
                         567.9      0.4      0.1    310.5          0          0
               write:      wps  avgserv  minserv  maxserv   timeouts      fails
                         114.0      0.6      0.2     23.6          0          0
               queue:  avgtime  mintime  maxtime  avgwqsz    avgsqsz     sqfull
                           0.0      0.0      0.1      0.0        0.0        0.0
dac1utm         xfer:  %tm_act      bps      tps      bread      bwrtn       time
                           0.0      0.0      0.0        0.0        0.0   14:19:42
                read:      rps  avgserv  minserv  maxserv   timeouts      fails
                           0.0      0.0      0.0      0.0          0          0
               write:      wps  avgserv  minserv  maxserv   timeouts      fails
                           0.0      0.0      0.0      0.0          0          0
               queue:  avgtime  mintime  maxtime  avgwqsz    avgsqsz     sqfull
                           0.0      0.0      0.0      0.0        0.0        0.0
hdisk2          xfer:  %tm_act      bps      tps      bread      bwrtn       time
                           2.4   946.1K     54.4     205.5K     740.7K   14:19:42
                read:      rps  avgserv  minserv  maxserv   timeouts      fails
                          44.5      0.3      0.1    112.5          0          0
               write:      wps  avgserv  minserv  maxserv   timeouts      fails
                           9.9      0.7      0.2     23.6          0          0
               queue:  avgtime  mintime  maxtime  avgwqsz    avgsqsz     sqfull
                           0.0      0.0      0.0      0.0        0.0        0.0
hdisk3          xfer:  %tm_act      bps      tps      bread      bwrtn       time
                           9.1     2.2M    195.2     829.6K       1.4M   14:19:42
                read:      rps  avgserv  minserv  maxserv   timeouts      fails
                         180.0      0.4      0.1    122.3          0          0
               write:      wps  avgserv  minserv  maxserv   timeouts      fails
                          15.1      0.8      0.2      3.4          0          0
               queue:  avgtime  mintime  maxtime  avgwqsz    avgsqsz     sqfull
                           0.0      0.0      1.5      0.0        0.0        0.0
hdisk4          xfer:  %tm_act      bps      tps      bread      bwrtn       time
                           7.7     2.4M    217.9     829.7K       1.6M   14:19:42
                read:      rps  avgserv  minserv  maxserv   timeouts      fails
                         180.1      0.4      0.1    100.7          0          0
               write:      wps  avgserv  minserv  maxserv   timeouts      fails
                          37.8      0.5      0.2      7.7          0          0
               queue:  avgtime  mintime  maxtime  avgwqsz    avgsqsz     sqfull
                           0.0      0.0      0.0      0.0        0.0        0.0
hdisk5          xfer:  %tm_act      bps      tps      bread      bwrtn       time
                           7.9     2.2M    178.3     746.8K       1.5M   14:19:42
                read:      rps  avgserv  minserv  maxserv   timeouts      fails
                         162.1      0.4      0.1    251.8          0          0
               write:      wps  avgserv  minserv  maxserv   timeouts      fails
                          16.2      1.1      0.2    494.1          0          0
```

```
                 queue:  avgtime  mintime  maxtime  avgwqsz    avgsqsz     sqfull
                            0.0      0.0      0.0      0.0        0.0        0.0
hdisk6            xfer:  %tm_act      bps      tps      bread      bwrtn       time
                           10.4     2.4M    211.5     824.5K       1.5M   14:19:42
                  read:      rps  avgserv  minserv  maxserv   timeouts      fails
                          178.9      0.4      0.1    158.5          0          0
                 write:      wps  avgserv  minserv  maxserv   timeouts      fails
                           32.5      0.6      0.2      4.1          0          0
                 queue:  avgtime  mintime  maxtime  avgwqsz    avgsqsz     sqfull
                            0.0      0.0      0.0      0.0        0.1        0.0
hdisk7            xfer:  %tm_act      bps      tps      bread      bwrtn       time
                           10.4     2.5M    231.4     921.1K       1.5M   14:19:42
                  read:      rps  avgserv  minserv  maxserv   timeouts      fails
                          199.9      0.4      0.1    258.0          0          0
                 write:      wps  avgserv  minserv  maxserv   timeouts      fails
                           31.5      0.6      0.2     41.5          0          0
                 queue:  avgtime  mintime  maxtime  avgwqsz    avgsqsz     sqfull
                            0.0      0.0      0.0      0.0        0.1        0.0
hdisk8            xfer:  %tm_act      bps      tps      bread      bwrtn       time
                            9.2     2.3M    198.1     757.5K       1.5M   14:19:42
                  read:      rps  avgserv  minserv  maxserv   timeouts      fails
                          164.4      0.4      0.1    310.5          0          0
                 write:      wps  avgserv  minserv  maxserv   timeouts      fails
                           33.8      0.6      0.2      4.9          0          0
                 queue:  avgtime  mintime  maxtime  avgwqsz    avgsqsz     sqfull
                            0.0      0.0      0.1      0.0        0.0        0.0
cd0               xfer:  %tm_act      bps      tps      bread      bwrtn       time
                            0.0      0.0      0.0        0.0        0.0   14:19:42
dac0              xfer:  %tm_act      bps      tps      bread      bwrtn       time
                            0.0     6.9M    604.8       2.5M       4.4M   14:19:42
                  read:      rps  avgserv  minserv  maxserv   timeouts      fails
                          542.0      0.4      0.1    258.0          0          0
                 write:      wps  avgserv  minserv  maxserv   timeouts      fails
                           62.8      0.8      0.2    494.1          0          0
                 queue:  avgtime  mintime  maxtime  avgwqsz    avgsqsz     sqfull
                            0.0      0.0      1.5      0.0        0.0        0.0
dac0utm           xfer:  %tm_act      bps      tps      bread      bwrtn       time
                            0.0      0.0      0.0        0.0        0.0   14:19:42
                  read:      rps  avgserv  minserv  maxserv   timeouts      fails
                            0.0      0.0      0.0      0.0          0          0
                 write:      wps  avgserv  minserv  maxserv   timeouts      fails
                            0.0      0.0      0.0      0.0          0          0
                 queue:  avgtime  mintime  maxtime  avgwqsz    avgsqsz     sqfull
                            0.0      0.0      0.0      0.0        0.0        0.0
```

The following list is an explanation from the **man viostat** command output of the most interesting statistical performance values being reported:

- % sys

  Shows the percentage of CPU utilization that occurred while executing at the system level (kernel).

- % idle

  Shows the percentage of time that the CPUs were idle and the system did not have an outstanding disk I/O request.

- % iowait

  Shows the percentage of time that the CPUs were idle during which the system had an outstanding disk I/O request.

- % tm_act

  Indicates the percentage of time that the physical disk was active (bandwidth utilization for the drive).

  The % utilization field for the physical volumes indicates how evenly the file activity is spread across the drives. A high % utilization on a physical volume is a good indication that there may be contention for this resource. Since the CPU utilization statistics are also available with the viostat report, the percentage of time that the CPU is in I/O wait can be determined at the same time. Consider distributing data across drives if the I/O wait time is significant and the disk utilization is not evenly distributed across volumes.

- Kbps

  Indicates the amount of data transferred (read or written) to the drive in KB per second.

- tps

  Indicates the number of transfers per second that were issued to the physical disk. A transfer is an I/O request to the physical disk. Multiple logical requests can be combined into a single I/O request to the disk. A transfer is of indeterminate size.

- Kb_read

  The total number of KB read.

- Kb_wrtn

  The total number of KB written.

9.1.3 Performance monitoring for DS Storage

There are two ways to monitor the performance of the DS Midrange Storage servers. The first one uses the DS4000 Storage Manager GUI and is not available for the DS3400. The other uses script commands and is used to monitor the DS power supply fan unit performance over time.

Storage Manager Performance Monitor

The Storage Performance Monitor is a tool built into the DS4000 Storage Manager client. It monitors performance on each logical drive and collects information such as:

- Total I/Os
- Read percentage
- Cache hit percentage
- Current KBps and maximum KBps
- Current I/O per sec and maximum I/O per sec

This section describes how to use the Performance Monitor and how to analyze the collected data.

Using the Storage Manager GUI

Launch the Performance Monitor from the Storage Manager Subsystem Management window by either of the following methods:

- Method 1:
  a. Select the Monitor Performance icon.
  b. Select the **Storage Subsystem** → **Monitor Performance** pull-down menu option (Figure 9-2).
- Select the storage subsystem node in the Logical View or Mappings View, then choose **Monitor Performance** from the right-mouse pop-up menu.

The Performance Monitor window opens up with all logical drives displayed, as shown in Figure 9-2.

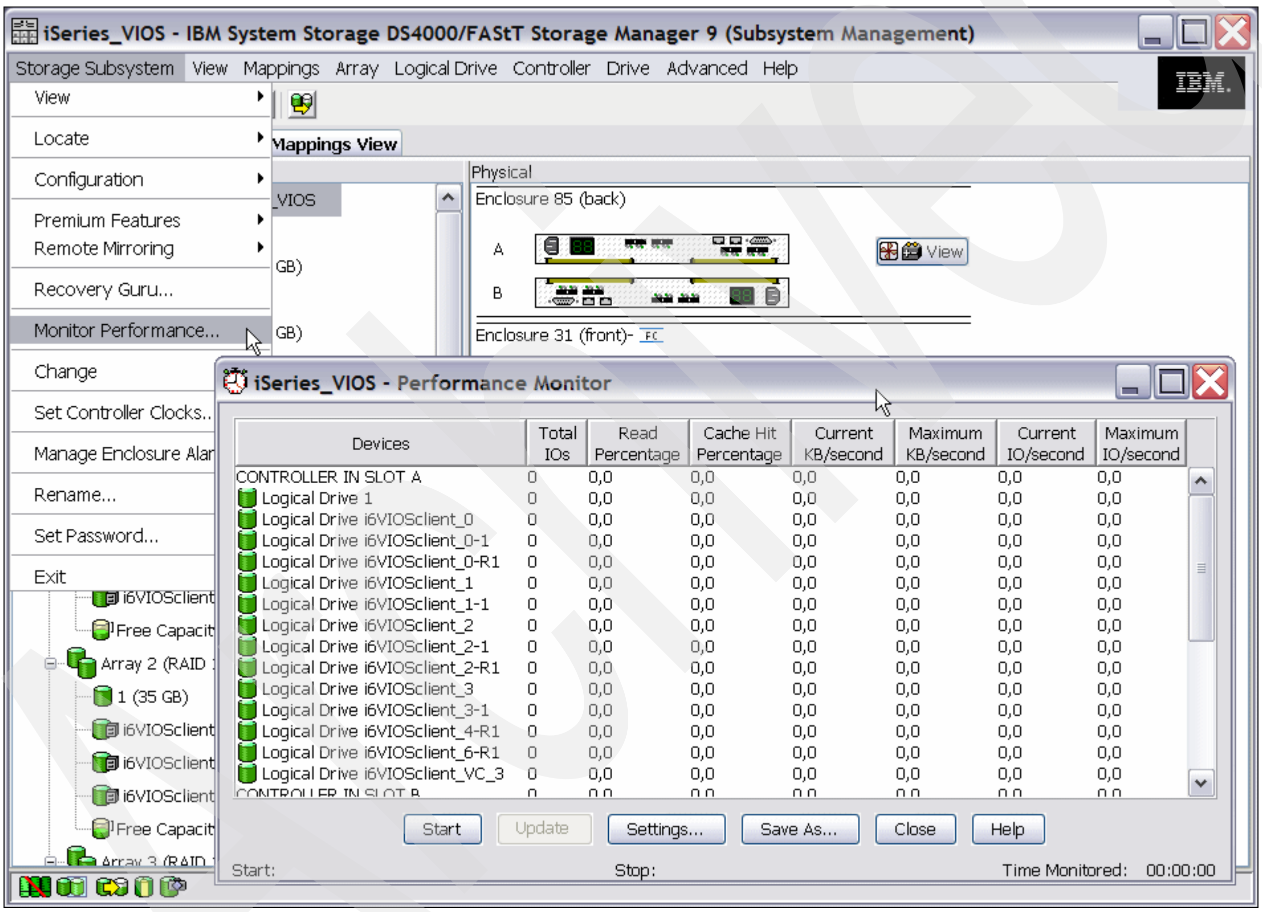
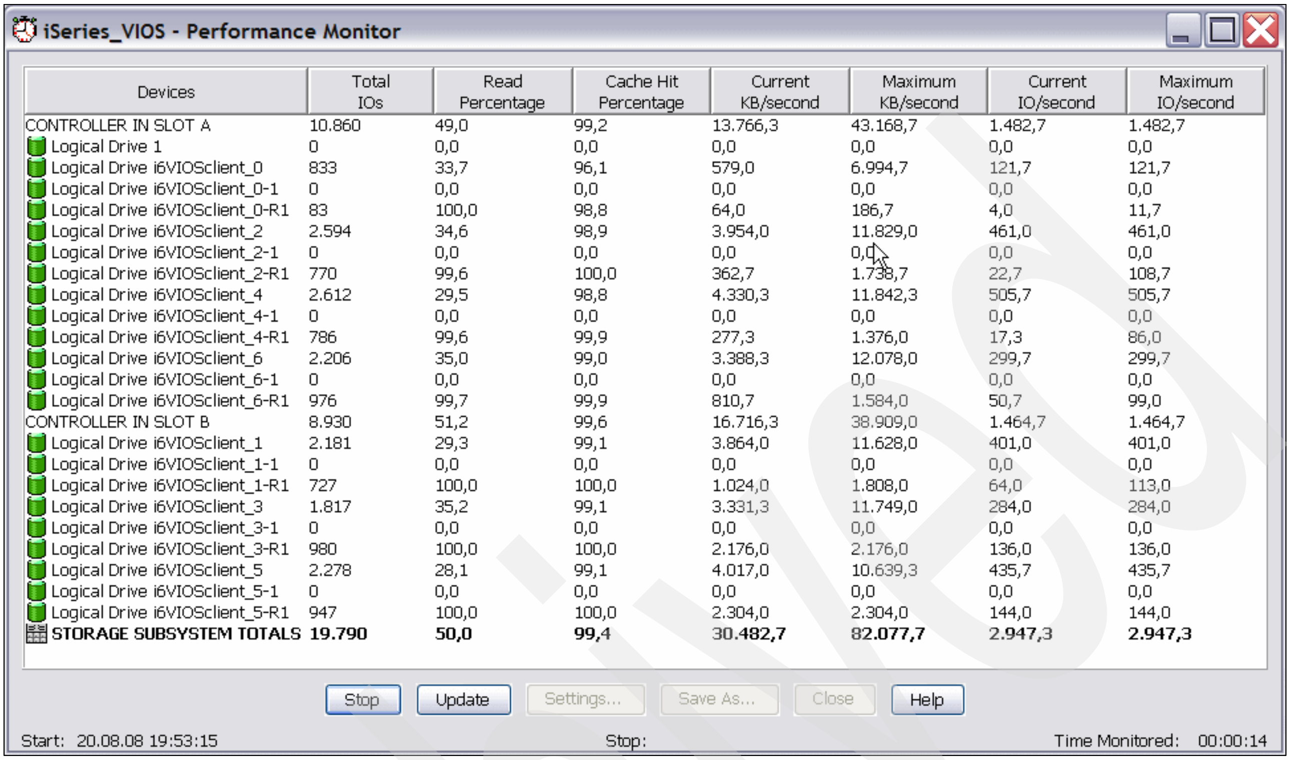


Figure 9-2 DS4000 Performance Monitor

Click the **Settings** button to select the logical drives that you want to monitor and to define the polling interval. The polling interval will define the time period of how often the Performance Monitor will be updated with new data.

Click the **Start** button to start the Performance Monitor.

After the first polling interval time has passed, the Performance Monitor window will show the first results (Figure 9-3).



Devices	Total IOs	Read Percentage	Cache Hit Percentage	Current KB/second	Maximum KB/second	Current IO/second	Maximum IO/second
CONTROLLER IN SLOT A	10.860	49,0	99,2	13.766,3	43.168,7	1.482,7	1.482,7
Logical Drive 1	0	0,0	0,0	0,0	0,0	0,0	0,0
Logical Drive i6VIOSclient_0	833	33,7	96,1	579,0	6.994,7	121,7	121,7
Logical Drive i6VIOSclient_0-1	0	0,0	0,0	0,0	0,0	0,0	0,0
Logical Drive i6VIOSclient_0-R1	83	100,0	98,8	64,0	186,7	4,0	11,7
Logical Drive i6VIOSclient_2	2.594	34,6	98,9	3.954,0	11.829,0	461,0	461,0
Logical Drive i6VIOSclient_2-1	0	0,0	0,0	0,0	0,0	0,0	0,0
Logical Drive i6VIOSclient_2-R1	770	99,6	100,0	362,7	1.738,7	22,7	108,7
Logical Drive i6VIOSclient_4	2.612	29,5	98,8	4.330,3	11.842,3	505,7	505,7
Logical Drive i6VIOSclient_4-1	0	0,0	0,0	0,0	0,0	0,0	0,0
Logical Drive i6VIOSclient_4-R1	786	99,6	99,9	277,3	1.376,0	17,3	86,0
Logical Drive i6VIOSclient_6	2.206	35,0	99,0	3.388,3	12.078,0	299,7	299,7
Logical Drive i6VIOSclient_6-1	0	0,0	0,0	0,0	0,0	0,0	0,0
Logical Drive i6VIOSclient_6-R1	976	99,7	99,9	810,7	1.584,0	50,7	99,0
CONTROLLER IN SLOT B	8.930	51,2	99,6	16.716,3	38.909,0	1.464,7	1.464,7
Logical Drive i6VIOSclient_1	2.181	29,3	99,1	3.864,0	11.628,0	401,0	401,0
Logical Drive i6VIOSclient_1-1	0	0,0	0,0	0,0	0,0	0,0	0,0
Logical Drive i6VIOSclient_1-R1	727	100,0	100,0	1.024,0	1.808,0	64,0	113,0
Logical Drive i6VIOSclient_3	1.817	35,2	99,1	3.331,3	11.749,0	284,0	284,0
Logical Drive i6VIOSclient_3-1	0	0,0	0,0	0,0	0,0	0,0	0,0
Logical Drive i6VIOSclient_3-R1	980	100,0	100,0	2.176,0	2.176,0	136,0	136,0
Logical Drive i6VIOSclient_5	2.278	28,1	99,1	4.017,0	10.639,3	435,7	435,7
Logical Drive i6VIOSclient_5-1	0	0,0	0,0	0,0	0,0	0,0	0,0
Logical Drive i6VIOSclient_5-R1	947	100,0	100,0	2.304,0	2.304,0	144,0	144,0
STORAGE SUBSYSTEM TOTALS	**19.790**	**50,0**	**99,4**	**30.482,7**	**82.077,7**	**2.947,3**	**2.947,3**

Figure 9-3 Monitoring performance

Click **Stop** to stop the monitoring. You can now save the actual window data to a file. The saved data only contains the current I/O and KBps values from the last polling interval.

Table 9-1 describes the information collected by the Performance Monitor.

Table 9-1 Information collected by Performance Monitor

Data field	**Description**
Total I/O	Total I/Os performed by this device since the beginning of the polling session.
Read Percentage	The percentage of total I/Os that are read operations for this device. Write percentage can be calculated as 100 minus this value. The read percentage value is the average value of the complete monitoring interval.
Cache Hit Percentage	The percentage of reads that are processed with data from the cache rather than requiring a read from disk. The cache hit percentage value is the average value of the complete monitoring interval.
Current KB/second	The average transfer rate during the last polling interval. The transfer rate is the amount of data in Kilobytes that can be moved through the I/O data connection in a second (also called throughput).
Maximum KBps	The maximum transfer rate that was achieved during the Performance Monitor polling session.
Current IO/second	The average number of I/O requests serviced per second during the last polling interval (also called I/O request rate).

Data field	Description
Maximum IO/second	The maximum number of I/O requests serviced during a one-second interval over the entire polling session.

The collected data can be used to analyze and improve the performance of Storage System. The following list can help to interpret the results:

- Total I/Os

  This data is useful for monitoring the I/O activity of a specific controller and a specific logical drive, which can help identify possible high-traffic I/O areas. If I/O rate is slow on a logical drive, try increasing the array size by adding more drives to this array. You might notice a disparity in the total I/Os (workload) of controllers, for example, the workload of one controller is heavy or is increasing over time, while that of the other controller is lighter or more stable. In this case rebalance the workload on the IBM i side. If you notice that the workload across the storage subsystem (storage subsystem totals total I/O statistic) continues to increase over time while application performance decreases, this might indicate the need to add additional storage subsystems to your installation so that you can continue to meet application needs at an acceptable performance level.

- Read percentage

  Use the read percentage for a logical drive to determine actual application behavior. If there is a low percentage of read activity relative to write activity, consider changing the RAID level of an array from RAID-5 to RAID-1 or RAID-10 for faster performance. We recommend the usage of RAID-1 or RAID-10 for IBM i attachment via VIOS.

- Cache hit percentage

  A higher percentage is desirable for optimal application performance. There is a positive correlation between the cache hit percentage and I/O rates. The cache hit percentage of all of the logical drives might be low or trending downward. This might indicate inherent randomness in access patterns, or at the storage subsystem or controller level, this can indicate the need to install more controller cache memory if you do not have the maximum amount of memory installed. If an individual logical drive is experiencing a low cache hit percentage, consider enabling cache read-ahead for that logical drive. Cache read-ahead can increase the cache hit percentage for a sequential I/O workload.

- Current KB/sec and maximum KB/sec

  The current KB/sec value is the average size of the amount of data that was transferred during one second during a particular interval period that was monitored. The maximum KB/sec value is the highest amount that was transferred over any one-second period during all of the interval periods in the number of iterations that were run for a specific command. This value can show you when peak transfer rate period was detected during the command runtime.

  The transfer rates of the controller are determined by the application I/O size and the I/O rate. Generally, small application I/O requests result in a lower transfer rate, but provide a faster I/O rate and shorter response time. With larger application I/O requests, higher throughput rates are possible. Understanding your typical application I/O patterns can help you determine the maximum I/O transfer rates for a given storage subsystem.

- Current I/O per second and maximum I/O per second

  The Current IO/sec value is the average number of I/Os serviced in one second during a particular interval period that was monitored. The maximum IO/sec value is the highest number of I/Os serviced in any one-second period during all of the interval periods in the number of iterations that were run for a specific command. This value can show you when the peak I/O period was detected during the command runtime.

Factors that affect I/Os per second include access pattern (random or sequential), I/O size, RAID level, segment size, and number of drives in the arrays or storage subsystem. The higher the cache hit rate, the higher the I/O rates. Performance improvements caused by changing the segment size can be seen in the I/Os per second statistics for a logical drive. Experiment to determine the optimal segment size, or use the file system or database block size.

Higher write I/O rates are experienced with write caching enabled compared to disabled. When deciding whether to enable write caching for a logical drive, consider the current and maximum I/Os per second. You should expect to see higher rates for sequential I/O patterns than for random I/O patterns. Regardless of your I/O pattern, we recommend that write caching be enabled to maximize I/O rate and shorten application response time.

Using the script editor to collect performance data over time

The Performance Monitor is a real-time tool. It is not possible to collect performance data over time with the Storage Manager GUI. You must use a script to collect performance data over some period of time and analyze it later.

Example 9-4 shows the script to collect performance data.

Example 9-4 Script to collect performance data

```
on error stop;
set performanceMonitor interval=20 iterations=180;
upload storageSubsystem file="c:\perf01.txt" content=performanceStats;
```

The *interval* specifies the polling interval time in seconds. This value can be set between 3 and 3600. The *iterations* value can be set between 1 and 3600. The monitoring time equals the number of iterations multiplied by the interval time. In the example above the monitoring time is 20s x 180 iterations = 3600s or 1 hour.

The performance data will be stored in a comma-separated text file at the location specified in the script. The path and file name can be changed freely.

The script can be run from the command line with SMcli or from the Storage Manager Script Editor GUI. From the Storage Manager Enterprise management window, select **Tools** → **Execute Script** (Figure 9-4).

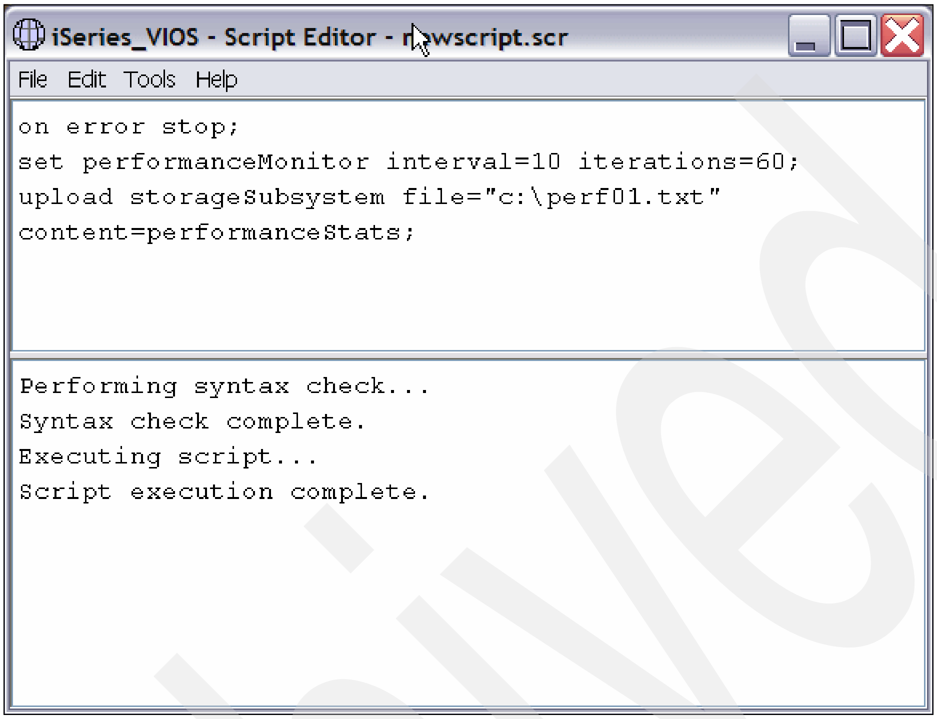


Figure 9-4 Collect performance data with script

In this example the performance data will be stored to a file perf01.txt. Example 9-5 shows the content of this file.

Example 9-5 Performance data collected by script

```
Performance Monitor Statistics for Storage Subsystem: iSeries_VIOS
Date/Time: 20.08.08 19:59:10
Polling interval in seconds: 10

Storage Subsystems,Total,Read,Cache Hit,Current,Maximum,Current,Maximum
,IOs,Percentage,Percentage,KB/second,KB/second,IO/second,IO/second
Capture Iteration: 1
Date/Time: 20.08.08 19:59:12
CONTROLLER IN SLOT A,10507.0,40.5,99.3,35406.5,35406.5,875.6,875.6,
Logical Drive 1,0.0,0.0,0.0,0.0,0.0,0.0,0.0,
Logical Drive i6VIOSclient_0,879.0,29.8,96.6,3774.3,3774.3,73.2,73.2,
Logical Drive i6VIOSclient_2,2284.0,16.8,98.4,9208.1,9208.1,190.3,190.3,
Logical Drive i6VIOSclient_4,2321.0,11.2,97.3,9219.0,9219.0,193.4,193.4,
Logical Drive i6VIOSclient_6,1927.0,13.4,98.1,9079.7,9079.7,160.6,160.6,
CONTROLLER IN SLOT B,9233.0,44.4,99.7,29794.5,29794.5,839.4,839.4,
Logical Drive i6VIOSclient_1,2088.0,14.0,100.0,8389.2,8389.2,189.8,189.8,
Logical Drive i6VIOSclient_3,1804.0,19.8,97.5,8333.2,8333.2,164.0,164.0,
Logical Drive i6VIOSclient_5,2153.0,11.9,98.8,8435.0,8435.0,195.7,195.7,
STORAGE SUBSYSTEM TOTALS,19740.0,42.3,99.5,65201.0,65201.0,1714.9,1714.9,

Capture Iteration: 2
Date/Time: 20.08.08 19:59:23
```

```
CONTROLLER IN SLOT A,19737.0,41.4,99.3,30368.6,35406.5,839.1,875.6,

Logical Drive 1,0.0,0.0,0.0,0.0,0.0,0.0,0.0,
Logical Drive i6VIOSclient_0,1484.0,17.9,95.5,4005.5,4005.5,55.0,73.2,
Logical Drive i6VIOSclient_2,4501.0,22.8,98.5,7569.0,9208.1,201.5,201.5,
...
```

This data can be easily imported to any spreadsheet for further analysis.

See also the IBM Redbooks publication *DS4000 Best Practices and Performance Tuning Guide*, SG24-6363, for more information about performance monitoring and system tuning.

9.2 VIOS maintenance

In this section we provide an introduction to the following VIOS maintenance tasks:

- “Upgrading VIOS” on page 346
- “VIOS backup/recovery” on page 347
- “Cleaning up orphan resources after a storage configuration change” on page 349
- “Migrating from RDAC to MPIO for DS4000 series attachment” on page 351
- “Migrating from RDAC to MPIO for DS4000 series attachment” on page 351

Upgrading VIOS

To upgrade VIOS to a higher level by installing a VIOS fixpack:

1. Download the latest fixpack *.iso CD images from the VIOS support Web site at:

   http://www14.software.ibm.com/webapp/set2/sas/f/vios/download/home.html

2. Follow the online instructions on the VIOS support Web site for installing the fixpack.

   For example:

   – To upgrade VIOS from a local directory on VIOS to which the CD image files were FTP transferred run the command:

   ```
   updateios -accept -install -dev <directory_of_CD_image_files>
   ```

   – To upgrade VIOS from burned CD images insert the first VIOS upgrade CD and run the command:

   ```
   updateios -accept -install -dev /dev/cd0
   ```

> **Notes:** VIOS upgrade CDs cannot be used for a VIOS scratch-install.
>
> We recommend installing the fixpack from files on either a local directory or a mounted network drive, as burning and repeatedly changing CDs may become cumbersome especially considering that due to a large amount of AIX filesets being updated the VIOS upgrade process may take around two hours.
>
> VIOS requires a reboot for the installed updates to be applied.

VIOS backup/recovery

There are different choices for backing up and restoring the IBM Virtual I/O Server:

- Back up to tape or DVD using the **backupios** command, as shown in Example 9-6, and restore from the tape/DVD install drive selected in SMS boot mode.

Example 9-6 VIOS backup to either tape or DVD

```
$ backupios -tape rmt0
$ backupios -cd cd0
```

- Back up to a network drive by creating a nim_resources.tar file and restoring from the HMC command line.
- Back up/restore to an AIX Network Installation Manager (NIM) server, as we describe below.

> **Note:** For further information about VIOS backup and restore please to the document *Methods to Backup and Restore the Virtual I/O Server* available at:
>
> http://www14.software.ibm.com/webapp/set2/sas/f/vios/documentation/backupios_mod.doc

VIOS backup to an AIX NIM server

To do this:

1. Ensure that there is enough space (~4 GB) in the NIM server file system:

   a. Run the command **df -gI** on VIOS to get an idea about the required space (used GB blocks), as shown in Example 9-7.

   Example 9-7 Displaying the root file system disk usage of VIOS

```
$ df -gI
Filesystem    GB blocks      Used      Free %Used Mounted on
/dev/hd4           0.25      0.05      0.20   19% /
/dev/hd2           4.50      3.50      1.00   78% /usr
/dev/hd9var        0.75      0.05      0.70    7% /var
/dev/hd3           2.25      0.10      2.15    5% /tmp
/dev/hd1          10.00      0.01      9.99    1% /home
/proc                 -         -         -     -  /proc
/dev/hd10opt       2.25      0.68      1.57   31% /opt
```

   b. Run the command **df -k** on NIM to check the available space for /export/mksysb, as shown in Example 9-8.

   Example 9-8 File system disk usage on the NIM server

```
# df -g
Filesystem    GB blocks      Free %Used    Iused %Iused Mounted on
/dev/hd4           0.50      0.43   15%     1893     2% /
/dev/hd2           2.00      0.96   52%    27086    11% /usr
/dev/hd9var        0.06      0.05   24%      405     4% /var
/dev/hd3           0.06      0.04   30%      163     2% /tmp
/dev/hd1           4.00      1.17   71%     1012     1% /home
/proc                 -         -    -         -     -  /proc
/dev/hd10opt       0.06      0.03   45%      678     8% /opt
/dev/fslv00       19.00      8.58   55%     4484     1% /export/lpp_source
/dev/fslv01        2.00      1.23   39%    22223     8% /export/spot
/dev/fslv02        8.00      6.49   22%        8     1% /export/mksysb
```

2. Export the NIM server directory /export/mksysb. To do this run the command **`smitty nfs`** and select **Network File System** → **Add a Directory to Exports List**.
3. Mount your NIM server NFS directory to the VIOS server using the **`mount`** command, as shown in Example 9-9.

Example 9-9 Mounting the NIM server directory to VIOS

```
$ mount i5nim:/export/mksysb /mnt
```

4. Back up your VIOS server to your NIM server from the VIOS padmin command line using the **`backupios`** command, as shown in Example 9-10.

Example 9-10 VIOS backup to NIM server

```
$ backupios -file /mnt/i6vios_1.5.2.1-FP-11.1_080709 -mksysb
```

VIOS install/restore from an AIX NIM server

To do this:

1. Create the NIM client for your VIOS server on your AIX NIM server. To do this run the command **`smitty nim`** and select **Configure the NIM Environment** → **Advanced Configuration** → **Define NIM Client Machines** → **Add a NIM Client** to add your VIOS server that you want to restore to as a NIM client machine.
2. Create a mksysb and SPOT resource on the NIM server. To do this run the command **`smitty nim`**, select **Perform NIM Administration Tasks** → **Manage Resources** → **Define a Resource**, and choose **spot** as well as **mksysb**.
3. Set up VIOS restore via TFTPBOOT from the NIM server by running the command **`smitty nim_bosinst`** and selecting your VIOS server NIM client as the target.

> **Note:** The installation settings are configured in the bosinst.data and image. files contained in the mksysb backup file, which could be manually adapted.

4. Activate your VIOS partition that you want to restore to in SMS boot mode.
5. Change the boot device to the network card.
6. Set up the IP configuration for the network card and the NIM master.
7. Set up network *normal boot* for your VIOS partition and the restore from the mksysb VIOS system backup on your NIM server will start via TFTBOOT.

Note: If your VIOS mksysb was obtained from an LVM mirrored VIOS partition and you did not adapt the installation settings for the target disks in “bosinst.data” make sure to select option **2. Change/Show Installation Settings and Install** from the Welcome to Base Operating System Installation and Maintenance window to select both hdisk0 and hdisk1 for the installation device, as shown below. Any prior existing VGs like on hdisk0 below will be deleted automatically:

```
                          Change Disk(s) Where You Want to Install

Type one or more numbers for the disk(s) to be used for installation and press
Enter. To cancel a choice, type the corresponding number and Press Enter.
At least one bootable disk must be selected. The current choice is indicated
by >>>.

        Name      Location Code   Size(MB)  VG Status   Bootable   Maps

>>>  1  hdisk1   03-08-00        199680   none           Yes    No
>>>  2  hdisk0   03-08-00        199680   other vg       Yes    No




>>>  0   Continue with choices indicated above
    55  More Disk Options
    66  Disks not known to Base Operating System Installation
    77  Display Alternative Disk Attributes
    88  Help ?
    99  Previous Menu

>>> Choice [0]: 0
```

Cleaning up orphan resources after a storage configuration change

In case you did remove virtual LUNs from the IBM i client ASP configuration and want to remove the corresponding virtual SCSI device mapping to the IBM i virtual SCSI server adapter on VIOS, which makes sense if you are not planning to re-use them by the same IBM i client, use the **rmdev -dev vtscsi*X*** command on VIOS, as shown in Example 9-11.

Important: Make sure that you are going to remove the correct vtscsi*X* device that corresponds to an unconfigured unit on your IBM i client by XORing the “Ctl” controller information with 0x80 for the unconfigured disk unit shown under IBM i with the vtscsi*X* LUN ID under VIOS (see also 6.6, “Adding DS Storage to IBM i configuration” on page 197).

Example 9-11 Removing virtual SCSI devices on VIOS

```
$ lsdev -type disk
name             status      description
hdisk0           Available   SAS RAID 0 Disk Array
hdisk1           Available   SAS RAID 0 Disk Array
hdisk2           Available   1815     DS4800 Disk Array Device
hdisk3           Available   1815     DS4800 Disk Array Device
hdisk4           Available   1815     DS4800 Disk Array Device
hdisk5           Available   1815     DS4800 Disk Array Device
hdisk6           Available   1815     DS4800 Disk Array Device
```

```
hdisk7          Available   1815     DS4800 Disk Array Device
hdisk8          Available   1815     DS4800 Disk Array Device
vtscsi0         Available   Virtual Target Device - Disk
vtscsi1         Available   Virtual Target Device - Disk
vtscsi2         Available   Virtual Target Device - Disk
vtscsi3         Available   Virtual Target Device - Disk
vtscsi4         Available   Virtual Target Device - Disk
vtscsi5         Available   Virtual Target Device - Disk
vtscsi6         Available   Virtual Target Device - Disk

$ lsdev -dev vtscsi6 -attr
attribute       value              description          user_settable

LogicalUnitAddr 0x8700000000000000 Logical Unit Address False
aix_tdev        hdisk8             Target Device Name   False
client_reserve  no                 Client Reserve       True

$ rmdev -dev vtscsi6
vtscsi6 deleted
$
$ lsdev -type disk
name            status      description
hdisk0          Available   SAS RAID 0 Disk Array
hdisk1          Available   SAS RAID 0 Disk Array
hdisk2          Available   1815     DS4800 Disk Array Device
hdisk3          Available   1815     DS4800 Disk Array Device
hdisk4          Available   1815     DS4800 Disk Array Device
hdisk5          Available   1815     DS4800 Disk Array Device
hdisk6          Available   1815     DS4800 Disk Array Device
hdisk7          Available   1815     DS4800 Disk Array Device
hdisk8          Available   1815     DS4800 Disk Array Device
vtscsi0         Available   Virtual Target Device - Disk
vtscsi1         Available   Virtual Target Device - Disk
vtscsi2         Available   Virtual Target Device - Disk
vtscsi3         Available   Virtual Target Device - Disk
vtscsi4         Available   Virtual Target Device - Disk
vtscsi5         Available   Virtual Target Device - Disk
```

When you are sure that you also do not want to map the corresponding DS Storage System LUN to another client of the same VIOS and thus have removed the mapping of the LUN to the VIOS partition by the DS Storage Manager client you should also remove the hdisk device from the Object Data Manager (ODM) database on the VIOS partition, as shown in Example 9-12. Otherwise you would end up with orphan resources in defined status on VIOS.

Example 9-12 Removing hdisk devices on VIOS

```
$ rmdev -dev hdisk8
hdisk8 deleted
$ lsdev -type disk
name            status      description
hdisk0          Available   SAS RAID 0 Disk Array
hdisk1          Available   SAS RAID 0 Disk Array
hdisk2          Available   1815     DS4800 Disk Array Device
hdisk3          Available   1815     DS4800 Disk Array Device
hdisk4          Available   1815     DS4800 Disk Array Device
hdisk5          Available   1815     DS4800 Disk Array Device
hdisk6          Available   1815     DS4800 Disk Array Device
hdisk7          Available   1815     DS4800 Disk Array Device
vtscsi0         Available   Virtual Target Device - Disk
vtscsi1         Available   Virtual Target Device - Disk
```

```
vtscsi2         Available   Virtual Target Device - Disk
vtscsi3         Available   Virtual Target Device - Disk
vtscsi4         Available   Virtual Target Device - Disk
vtscsi5         Available   Virtual Target Device - Disk
```

If you have multiple devices to remove you may want to use a *while* loop using the syntax that we used in Example 6-11 on page 194.

Note: Remember that any newly attached DS Storage System LUNs to the IBM Virtual I/O Server first need to be discovered by manually running the **cfgdev** command from VIOS.

Migrating from RDAC to MPIO for DS4000 series attachment

We used the following disruptive procedure for migrating an existing VIOS 1.5.x DS4000 series attachment using the default RDAC device driver to MPIO. A VIOS 2.1 *scratch* install based on AIX 6.1 kernel uses MPIO as the default device driver not only for DS3000 and DS5000 series but also for DS4000 series attachment:

1. We power down all our VIOS *client* partitions.
2. We look at our VIOS current disk configuration and mapping using the **lsdev** and **lsmap** commands, as shown in Example 9-13.

Example 9-13 Looking at the VIOS current disk configuration and mapping

```
$ lsdev -type disk
hdisk0  Available  SAS RAID 0 Disk Array
hdisk1  Available  SAS RAID 0 Disk Array
hdisk2  Available  DS4800 Disk Array Device
hdisk3  Available  DS4800 Disk Array Device
hdisk4  Available  DS4800 Disk Array Device
hdisk5  Available  DS4800 Disk Array Device
hdisk6  Available  DS4800 Disk Array Device
hdisk7  Available  DS4800 Disk Array Device
hdisk8  Available  DS4800 Disk Array Device

$ lsmap -all
SVSA            Physloc                                      Client Partition ID
--------------- -------------------------------------------- ------------------
vhost0          U9406.MMA.655A620-V2-C12                     0x00000003

VTD                   vtscsi0
Status                Available
LUN                   0x8100000000000000
Backing device        hdisk2
Physloc               U789D.001.DQDWVYP-P1-C3-T1-W201500A0B811F4C0-L0

VTD                   vtscsi1
Status                Available
LUN                   0x8200000000000000
Backing device        hdisk3
Physloc               U789D.001.DQDWVYP-P1-C3-T1-W201500A0B811F4C0-L1000000000000

VTD                   vtscsi2
Status                Available
LUN                   0x8300000000000000
Backing device        hdisk4
Physloc               U789D.001.DQDWVYP-P1-C3-T1-W201500A0B811F4C0-L2000000000000

VTD                   vtscsi3
```

```
Status              Available
LUN                 0x8400000000000000
Backing device      hdisk5
Physloc             U789D.001.DQDWVYP-P1-C3-T1-W201500A0B811F4C0-L3000000000000

VTD                 vtscsi4
Status              Available
LUN                 0x8500000000000000
Backing device      hdisk6
Physloc             U789D.001.DQDWVYP-P1-C3-T1-W201500A0B811F4C0-L4000000000000

VTD                 vtscsi5
Status              Available
LUN                 0x8600000000000000
Backing device      hdisk7
Physloc             U789D.001.DQDWVYP-P1-C3-T1-W201500A0B811F4C0-L5000000000000

VTD                 vtscsi6
Status              Available
LUN                 0x8700000000000000
Backing device      hdisk8
Physloc             U789D.001.DQDWVYP-P1-C3-T1-W201500A0B811F4C0-L6000000000000

SVSA            Physloc                                      Client Partition ID
--------------- -------------------------------------------- ------------------
vhost1          U9406.MMA.655A620-V2-C13                     0x00000003

VTD                 vcd
Status              Available
LUN                 0x8100000000000000
Backing device      cd0
Physloc             U789D.001.DQDWXNY-P4-D1
```

3. We remove all our client partitions' vtscsi devices using the **rmdev** command, as shown in Example 9-14, because when changing the multi-path driver the characteristics of the hdisk backing devices will be changed.

Example 9-14 Removing all VIOS client vtscsi devices

```
$ i=0
$ while [$i -lt 7]
> do
> rmdev -dev vtscsi$i
> i=$((i+1)
> done
vtscsi0 deleted
vtscsi1 deleted
vtscsi2 deleted
vtscsi3 deleted
vtscsi4 deleted
vtscsi5 deleted
vtscsi6 deleted
```

4. We use the **manage_disk_drivers** command from the VIOS root-command line to look at our current VIOS multi-path driver configuration for DS4800, which is currently set to using RDAC, as shown in Example 9-15.

Example 9-15 Looking at the VIOS multi-path driver configuration

```
$ oem_setup_env
# manage_disk_drivers
1: DS4100: currently RDAC; supported: RDAC/fcparray, MPIO
2: DS4300: currently RDAC; supported: RDAC/fcparray, MPIO
3: DS4500: currently RDAC; supported: RDAC/fcparray, MPIO
4: DS4700/DS4200: currently RDAC; supported: RDAC/fcparray, MPIO
5: DS4800: currently RDAC; supported: RDAC/fcparray, MPIO
```

5. To change from our current RDAC device driver setting for DS4800 storage attachment to MPIO we use the **manage_disk_drivers -c 5** command to change the multi-path driver for item 5 (which is DS4800), write a new boot image, and reboot VIOS, as shown in Example 9-16.

Example 9-16 Changing the VIOS multi-path driver

```
# manage_disk_drivers -c 5
DS4800 currently RDAC/fcparray
Change to alternate driver? [Y/N] Y
DS4800 now managed by MPIO

It is necessary to perform a bosboot before rebooting the system in
order to incorporate this change into the boot image.

In order to change to the new driver, either a reboot or a full
unconfigure and reconfigure of all devices of the type changed
must be performed.

# bosboot -a
bosboot: Boot image is 38563 512 byte blocks.

# shutdown -Fr
SHUTDOWN PROGRAM
Thu Aug 21 14:48:29 CDT 2008
0513-044 The sshd Subsystem was requested to stop.

Wait for 'Rebooting...' before stopping.
Error reporting has stopped.
```

6. After the VIOS reboot our devices now show up as MPIO Other DS4K Array Disk, as shown in Example 9-17.

Example 9-17 Looking at the VIOS MPIO devices

```
$ lsdev -type disk
name            status      description
hdisk0          Available   SAS RAID 0 Disk Array
hdisk1          Available   SAS RAID 0 Disk Array
hdisk2          Available   MPIO Other DS4K Array Disk
hdisk3          Available   MPIO Other DS4K Array Disk
hdisk4          Available   MPIO Other DS4K Array Disk
hdisk5          Available   MPIO Other DS4K Array Disk
hdisk6          Available   MPIO Other DS4K Array Disk
hdisk7          Available   MPIO Other DS4K Array Disk
hdisk8          Available   MPIO Other DS4K Array Disk
```

7. Similar to the **fget_config** command for RDAC, we can now use the **mpio_get_config** command for MPIO showing our hdisk to LUN assignments, as in Example 9-18.

Example 9-18 Looking at the MPIO hdisks LUN assignments

```
$ oem_setup_env
# mpio_get_config -Av
Frame id 0:
    Storage Subsystem worldwide name: 60ab800139f690000478b4c7
    Controller count: 2
    Partition count: 1
    Partition 0:
    Storage Subsystem Name = 'iSeries_VIOS'
        hdisk      LUN #   Ownership          User Label
        hdisk2         0   A (preferred)      i6VIOSclient_0
        hdisk3         1   B (preferred)      i6VIOSclient_1
        hdisk4         2   A (preferred)      i6VIOSclient_2
        hdisk5         3   B (preferred)      i6VIOSclient_3
        hdisk6         4   A (preferred)      i6VIOSclient_4
        hdisk7         5   B (preferred)      i6VIOSclient_5
        hdisk8         6   A (preferred)      i6VIOSclient_6
```

8. Analogous to the recommendation we gave for RDAC, we also recommend enabling the autorecovery option for MPIO to allow for automatic LUN re-assignment to the preferred controller after recovery of a DS4000 controller path failure. In contrast to RDAC, which does a grouping of LUNs under a "dar*X*" device, MPIO treats each LUN individually, so changing MPIO settings like the autorecovery feature needs to be done on an individual LUN basis. Example 9-19 shows how we enabled the MPIO autorecovery option for each LUN of our IBM i client after we unconfigured the hdisks first.

Example 9-19 Enabling the MPIO autorecovery feature

```
$ i=2
$ while [ $i -lt 9 ]
> do
> rmdev -dev hdisk$i -ucfg
> i=$((i+1))
> done
hdisk2 Defined
hdisk3 Defined
hdisk4 Defined
hdisk5 Defined
hdisk6 Defined
hdisk7 Defined
hdisk8 Defined

$ i=2
$ while [ $i -lt 9]
> do
> chdev -dev hdisk$i -attr autorecovery=yes
> i=$((i+1))
> done
hdisk2 changed
hdisk3 changed
hdisk4 changed
hdisk5 changed
hdisk6 changed
hdisk7 changed
hdisk8 changed
```

```
$ cfgdev

$ lsdev -dev hdisk2 -attr
attribute       value                   description                       user_settable

PCM             PCM/friend/otherapdisk Path Control Module               False
PR_key_value    none                    Persistant Reserve Key Value      True
algorithm       fail_over               Algorithm                         True
autorecovery    yes                     Path/Ownership Autorecovery       True
clr_q           no                      Device CLEARS its Queue on error True
cntl_delay_time 0                       Controller Delay Time             True
cntl_hcheck_int 0                       Controller Health Check Interval True
dist_err_pcnt   0                       Distributed Error Percentage      True
dist_tw_width   50                      Distributed Error Sample Time     True
hcheck_cmd      inquiry                 Health Check Command              True
hcheck_interval 60                      Health Check Interval             True
hcheck_mode     nonactive               Health Check Mode                 True
location                                Location Label                    True
lun_id          0x0                     Logical Unit Number ID            False
max_transfer    0x40000                 Maximum TRANSFER Size             True
node_name       0x200400a0b811f4c0      FC Node Name                      False
pvid            none                    Physical volume identifier        False
q_err           yes                     Use QERR bit                      True
q_type          simple                  Queuing TYPE                      True
queue_depth     10                      Queue DEPTH                       True
reassign_to     120                     REASSIGN time out value           True
reserve_policy  single_path             Reserve Policy                    True
rw_timeout      30                      READ/WRITE time out value         True
scsi_id         0x962e00                SCSI ID                           False
start_timeout   60                      START unit time out value         True
ww_name         0x201500a0b811f4c0      FC World Wide Name                False
```

9. Now we can recreate our virtual SCSI target devices using the **mkvdev** command, as shown in Example 9-20.

Example 9-20 Mapping of hdisks to virtual SCSI target devices

```
$ i=2
$  while [ $i -lt 9 ]
> do
> mkvdev -vdev hdisk$i -vadapter vhost0
> i=$((i+1))
> done
vtscsi0 Available
vtscsi1 Available
vtscsi2 Available
vtscsi3 Available
vtscsi4 Available
vtscsi5 Available
vtscsi6 Available
```

10. At this point we are done with the RDAC to MPIO migration and can reactivate (that is, IPL) our IBM i client.

VIOS hot-plug Fibre Channel adapter replacement

The following steps outline the supported procedure for hot-plug replacing a Fibre Channel HBA in the VIOS partition.

> **Notes:** *Hot-plug* replacement of a defective Fibre Channel HBA is only supported when replacing it with the *same* adapter model in the *same* PCI slot and with the RDAC or MPIO autorecovery option disabled.

1. If the RDAC or MPIO multi-path device driver is used for IBM DS Midrange Storage attachment disable autorecovery by using the **chdev** command, as shown in Example 9-21 and Example 9-22.

   Example 9-21 Disabling RDAC autorecovery

```
$ chdev -dev dar0 -attr autorecovery=no
dar0 changed
```

   Example 9-22 Disabling MPIO autorecovery

```
$ i=0
$ while [ $i -lt 7 ]
> do
> rmdev -dev vtscsi$i -ucfg
> i=$((i+1))
> done
vtscsi0 Defined
vtscsi1 Defined
vtscsi2 Defined
vtscsi3 Defined
vtscsi4 Defined
vtscsi5 Defined
vtscsi6 Defined
$ i=2
$ while [ $i -lt 9 ]
> do
> chdev -dev hdisk$i -attr autorecovery=no
> i=$((i+1))
> done
hdisk2 changed
hdisk3 changed
hdisk4 changed
hdisk5 changed
hdisk6 changed
hdisk7 changed
hdisk8 changed
```

2. Unconfigure each port of the Fibre Channel adapter that you want to replace using the **rmdev** command, as shown in Example 9-23.

   Example 9-23 Unconfiguring Fibre Channel adapters

```
$ rmdev -dev fcs0 -recursive -ucfg
fcnet0 Defined
dac0 Defined
fscsi0 Defined
fcs0 Defined
```

```
$ rmdev -dev fcs1 -recursive -ucfg
fcnet1 Defined
fscsi1 Defined
fcs1 Defined
```

3. Initiate the hot-plug adapter replacement by entering the **`smitty devdrpci`** command, selecting the menu option **Replace/Remove a PCI Hot Plug Adapter**, and choosing your FC adapter to replace by following the instructions on the smit panels.
4. When the new FC adapter is powered on and its fcs*X* devices are in the defined state reattach the FC cable.
5. Configure the new FC adapter by entering the **`cfgdev -dev fcsX`** command.
6. Re-enable the RDAC or MPIO autorecovery using the **`chdev`** command, as shown in Example 9-25 and Example 9-24.

Example 9-24 Re-enabling RDAC autorecovery

```
$ chdev -dev dar0 -attr autorecovery=yes
dar0 changed
```

Example 9-25 Re-enabling MPIO autorecovery

```
$ i=2
$ while [ $i -lt 9 ]
> do
> chdev -dev hdisk$i -attr autorecovery=yes
> i=$((i+1))
> done
hdisk2 changed
hdisk3 changed
hdisk4 changed
hdisk5 changed
hdisk6 changed
hdisk7 changed
hdisk8 changed
$ i=0
$ while [ $i -lt 7 ]
> do
> cfgdev -dev vtscsi$i
> i=$((i+1))
> done
$
```

7. Use the **`lsdev -dev fcsX -vpd`** command on VIOS to get the WWPN information from the new card and change the WWPN information in the DS Storage Manager client to the one of the new Fibre Channel HBAs by right-clicking the HBA ports of the replaced FC adapter in the mappings view and selecting the **Replace** option from the context menu.

For further information refer to *IBM System Storage DS4000 Storage Manager Installation and Host Support Guide*, GA76-0422, available at:

ftp://ftp.software.ibm.com/systems/support/storage_disk_pdf/ga76042200.pdf

9.3 DS maintenance

This section addresses maintenance topics for the DS3000, DS4000, and DS5000.

We explain how to check for and upgrade to the latest firmware level for the different storage server components.

This section also includes information about other management and maintenance tasks such as handling the premium features and how to back up your configuration

9.3.1 Displaying installed firmware versions

You need to know the firmware level currently installed on your system. This information is needed when you report a problem to your support representative, if you plan an update after receiving notification of the availability of a new level, or when you want to use a new feature only available with a particular level.

Use the profile data to find the different components' firmware versions in your system:

1. On the DS4000 and DS5000, select **Storage Subsystem** → **View** → **Profile** from the Subsystem Management window, or **Storage Subsystem** → **View Profile** if using older levels of Storage Manager client.

   On the DS3000, select the **Storage Subsystem Profile** link on the summary page.

2. Click the **Controller** tab to display a window, as shown in Figure 9-5. This gives the controller firmware and NVSRAM version installed.

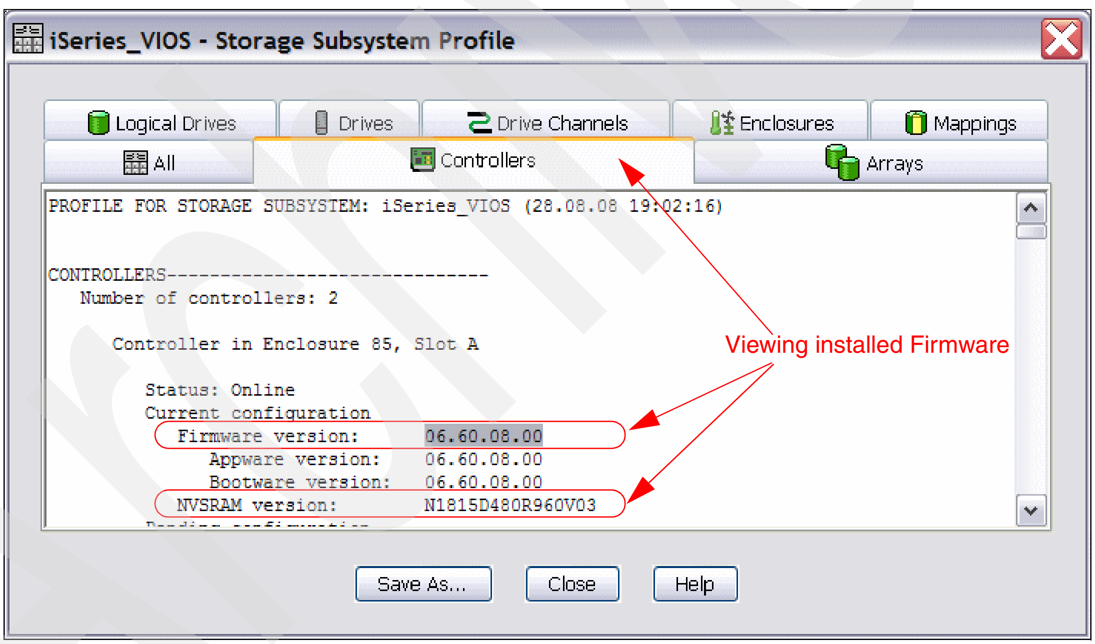


Figure 9-5 Viewing firmware and NVSRAM versions

3. Click the **Enclosures** tab to see the ESM current firmware level. As shown in Figure 9-6, you must use the scroll bar to show all of the ESM firmware (two per enclosure), as circled in red in Figure 9-6.

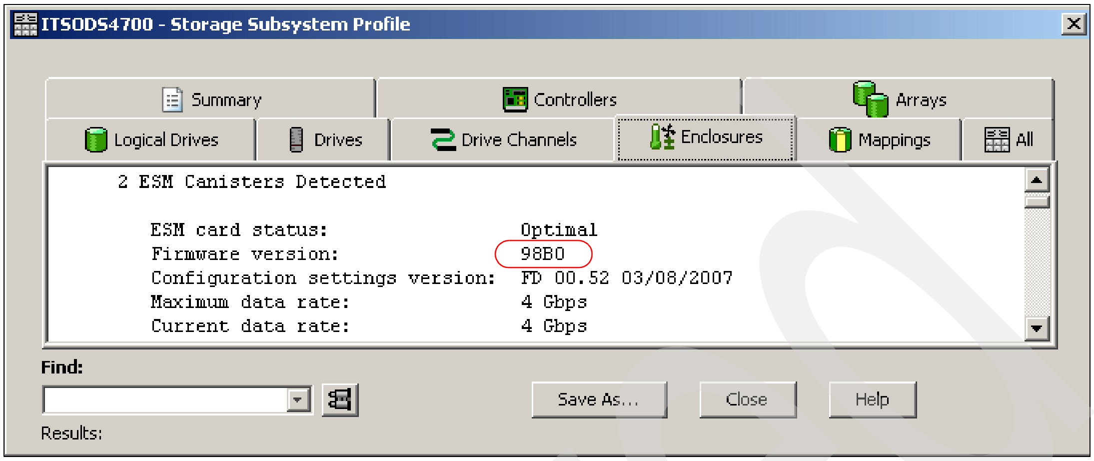


Figure 9-6 Viewing ESM firmware

4. Click the **Drives** tab to see the drives information. Scroll to the right side of the window until the Firmware Version column is visible. Be aware that you might have different drive types in your enclosures, so you can find multiple versions. See Figure 9-7.

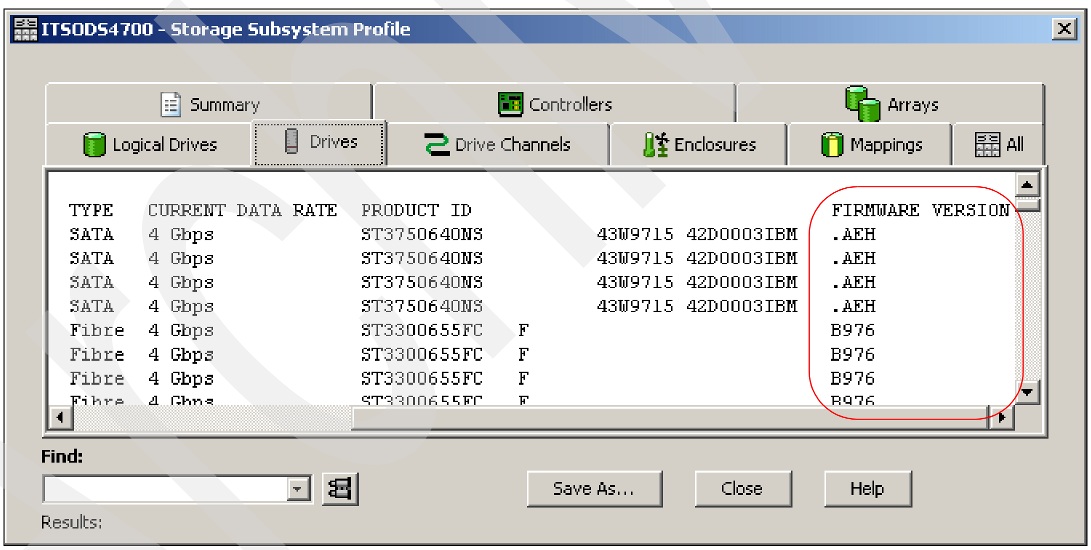


Figure 9-7 Viewing disks firmware

Obtaining updates

To download the latest firmware and drivers, and for updates, hints, and tips on known problems, fixes, technical notes, and product flashes (alerts), consult the IBM Support Web site at:

http://www.ibm.com/servers/storage/support/disk

Select your specific DS3000, DS400,0 or DS5000 model and click the **Download** tab to get to the latest versions of Storage Manager, firmware, HBA tools, tips, and publications available. There might be more than a single version of firmware available for download. Always review the readme file to make sure that it is the correct version for your product and configuration. If the readme contains specific enhancements or fixes specifically for your system, then consider updating your Storage System.

Support Notifications subscription service

There is a Support Notifications subscription service launched on the IBM System Storage product support Web site. This service is designed to keep you informed of new or updated System Storage support site information, without the need to periodically check the IBM Support Web site for updates. To subscribe to the Support Notifications service, go to:

http://www.ibm.com/jct01004c/systems/support/storage/news/05072007SupportNotif.html

System Storage Interoperation Center (SSIC)

If you are planning to add a new host to your storage, add expansions, change HBAs in your hosts, or perform other hardware configuration changes, check also the latest supported combinations at the following Web site:

http://www.ibm.com/systems/support/storage/config/ssic/index.jsp

9.3.2 Planning for upgrades

Upgrading the firmware and management software for the DS3000, DS4000, and DS5000 power supply fan unit is a relatively simple procedure, but some preparation and caution are needed to ensure a smooth upgrade. Depending on your specific model and version to install, you should follow specific instruction described in that version's readme file.

The components that typically need to be updated are:

- Storage Manager software
- Controller firmware
- NVSRAM firmware
- ESM firmware
- Disk drives firmware

Depending on the component to update, the current version, and the destination level, the procedure may allow a concurrent upgrade or be a disruptive update.

For a disruptive upgrade, before you start, make sure that you have an adequate maintenance window to do the procedure. Upgrading large storage configurations can be time consuming. Time estimates for upgrading all the associated firmware and software are listed in Table 9-2.

Table 9-2 Upgrade times

Element being upgraded	Approximate time of upgrade
Storage Manager and associated drivers and utilities	35 minutes
ESM firmware	5 minutes per ESM
DS4000 power supply fan unit firmware and NVSRAM	5 to 35 minutes
Hard drives	2 minutes per drive (but it is possible to do a parallel firmware upgrade)

These times were observed in the lab. They are approximate and may vary from system to system. Activating controller firmware or NVSRAM maximum times are 15 minutes per controller.

It is critical that if you update one component, you also update other components to the matching levels (refer to the readme file for the release to which you are upgrading). Running a mismatched set is not supported. A temporary firmware mismatch during the upgrade period may be necessary and should be limited to the shortest possible duration.

9.3.3 Updating the DS4000 host software

In this section we discuss updating the DS4000 host software on a Windows management server.

> **Important:** Storage Manager host software has to be updated to V10.10 to manage DS4000 with firmware V7.10. You must update the SM host software before performing the firmware upgrade.

Code update for the Windows environment

DS4000 Storage Manager (using the InstallAnywhere installation wizard) supports install upgrade. Launch the install program and it automatically upgrades all the components that you previously installed. For details on the installation procedures, refer to the *DS4000 Storage Manager Guide for Windows 2000/Server 2003, NetWare, ESX Server and Linux*, GC26-7847, at:

https://www-304.ibm.com/systems/support/supportsite.wss/docdisplay?lndocid=MIGR-57808&brandind=5000028

For other operating systems refer to the *IBM System Storage DS4000 Storage Manager v10.10 Installation and Host Guide for AIX, HP-UX, Solaris, and Linux on POWER*, GC27-2170, at:

https://www-304.ibm.com/systems/support/supportsite.wss/docdisplay?lndocid=MIGR-61177&brandind=5000028

If your storage is running firmware level 6.12 or earlier, refer to this book for specific instructions, since you might need to uninstall a previous Storage Manager version.

The Storage Manager upgrade for the DS3400 is identical. Refer to the following documents:

- *Installation and Support Guide for Microsoft Windows Server 2003, Linux, Novell NetWare, and VMware ESX - IBM System Storage DS3200, DS3300, DS3400, BladeCenter Boot Disk (Type 1726):*

  https://www-304.ibm.com/systems/support/supportsite.wss/docdisplay?lndocid=MIGR-5069917&brandind=5000028

- *Storage Manager v2 Installation and Support Guide for IBM AIX and Linux on POWER - IBM System Storage DS3000*:

  https://www-304.ibm.com/systems/support/supportsite.wss/docdisplay?lndocid=MIGR-5073850&brandind=5000028

9.3.4 Updating the Controller Firmware

The upgrade from the 6.x to the 7.x firmware level is not described here. For this upgrade, a special upgrade utility is contained in the firmware bundle that performs a system health check first and the performs the firmware upgrade automatically. If you plan to upgrade from the 6.x to the 7.x firmware level refer to the special documentation in the 7.10.23 firmware bundle.

We describe the regular procedure for upgrading the firmware using the Storage Manager client interface. This procedure is used for the following cases:

- Upgrading DS3400, DS4700, and DS4800 to the latest release of Version 6.x
- Upgrading DS5000, DS4700, and DS4800 from Version 7.x to a later version

> **Important:** You cannot use the Storage Manager interface to upgrade to firmware Version 7, but you should continue using Storage Manager to update up to the latest Version 6.X and to update from any 7.X to any later level.

In general, the controller, NVSRAM, and ESM firmware upgrades can be done online during non-peak periods of traffic if your storage server model has redundant controllers, and if a redundant driver is installed in all the hosts being serviced by your DS.

> **Important:** Before upgrading the controller firmware, make sure that the system is in an optimal state. If not, run the Recovery Guru to diagnose and fix the problem before you proceed with the upgrade.

In the above conditions, once the upgrade is initiated, the files are transferred to one controller, which then copies them to the other. A consistency check is done on both controllers, and if the checksum is okay, the uploaded firmware is marked as ready and available for activation.

The activation procedure can be done immediately after the transfer, or later during a period with less I/O access. During the activation, the first controller moves all logical drives to the second one. Then it reboots and activates new firmware. After that, it takes ownership of all logical drives, and the second controller is rebooted in order to have its new firmware activated. When both controllers are up again, the logical drives are redistributed to the preferred paths.

If you choose to not activate the transferred image at the same moment of the firmware transfer, remember that a normal reboot of a controller or a power cycle of the DS4000 does not activate the new firmware. It is only activated after the user specifically chooses to activate the firmware.

To perform the staged firmware and NVSRAM update:

1. Open the Subsystem Management window for the DS4000 power supply fan unit that you want to upgrade. To download the firmware select **Advanced** → **Maintenance** → **Download** → **Controller Firmware**, as shown in Figure 9-8.

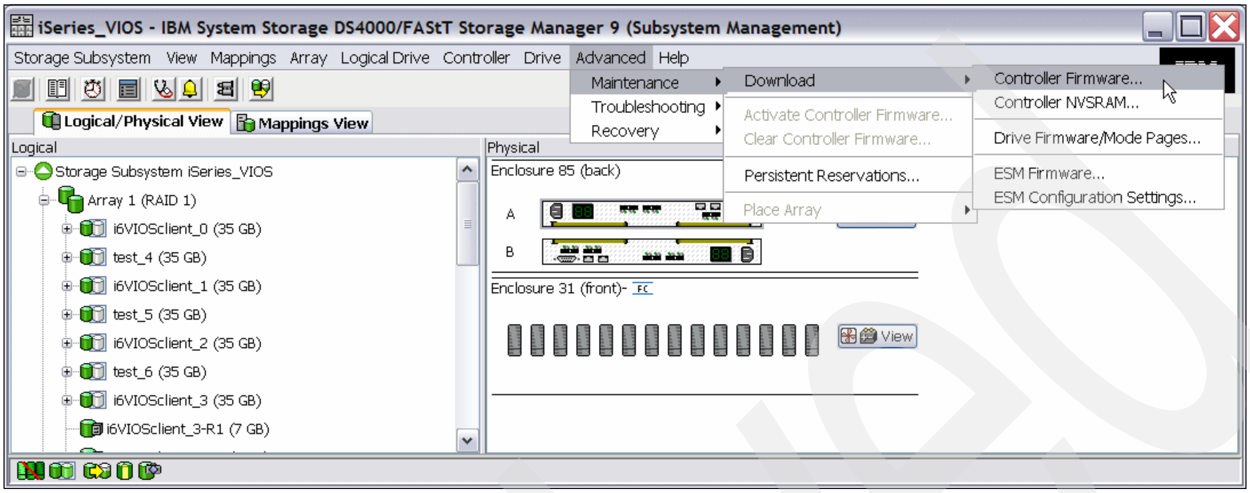


Figure 9-8 Subsystem Management window: Controller firmware update on DS4000 and DS5000

On the DS3000, go to the support page on the Subsystem Management view and select **Download Firmware**, as shown in Figure 9-9.

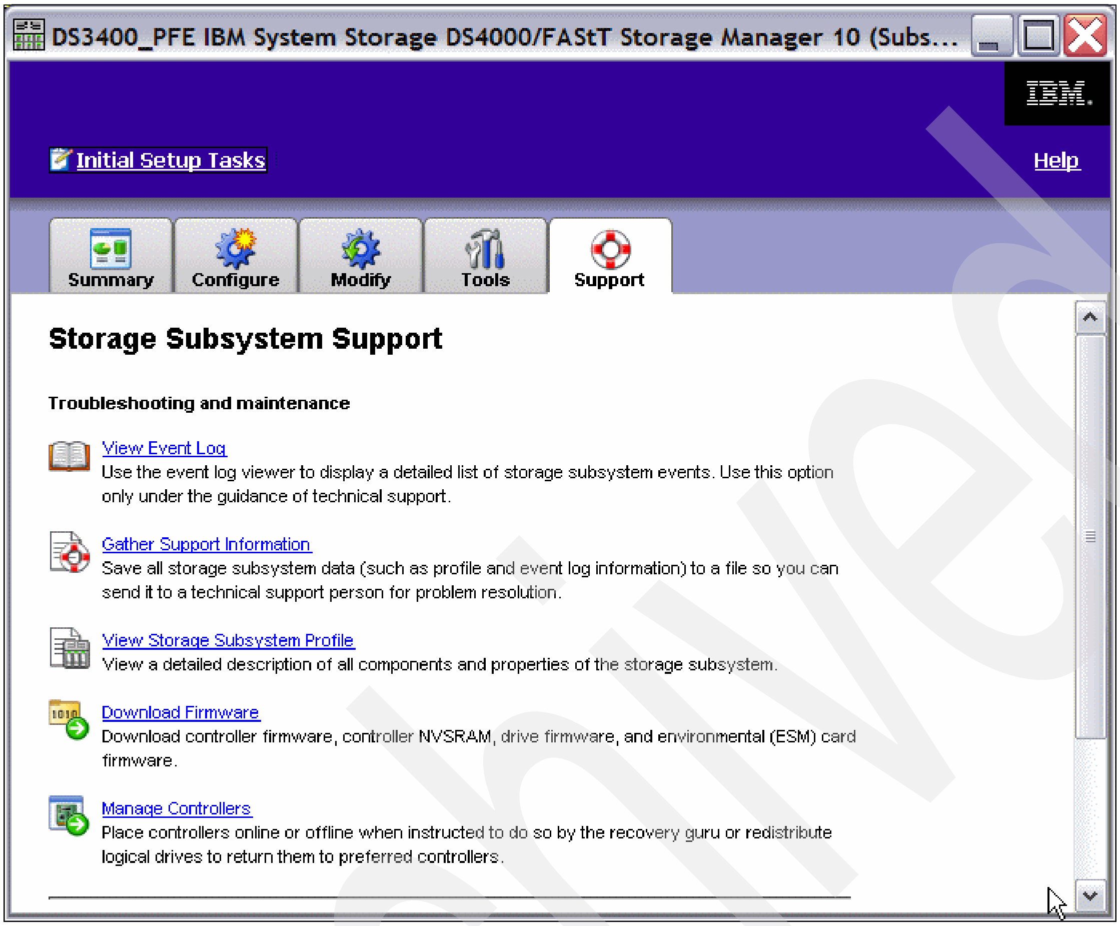


Figure 9-9 Subsystem Management window: Controller firmware update on DS3000

2. The Download Firmware window opens showing the current firmware and NVSRAM versions. Select the correct firmware and NVSRAM files, as shown in Figure 9-10. Mark the check box to download the NVSRAM file as well.

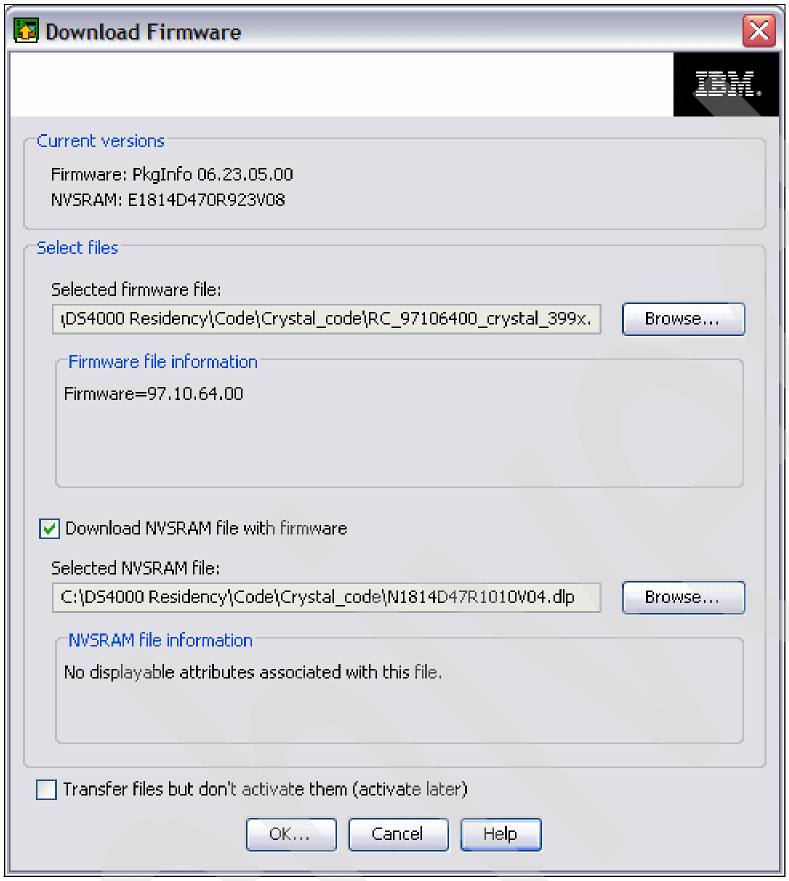


Figure 9-10 Download Firmware window

There is a “Transfer files but do not activate them” check box at the bottom of the window. Select this check box if you want to activate the new firmware at a later time. Then click **OK** to continue.

3. The next window instructs you to confirm the firmware and NVSRAM download because the process cannot be cancelled once it begins.

Confirm by clicking **Yes**. The firmware/NVSRAM transfer begins and you can watch the progress. When the process finishes, the `Transfer Successful` message is displayed, as shown in Figure 9-11.

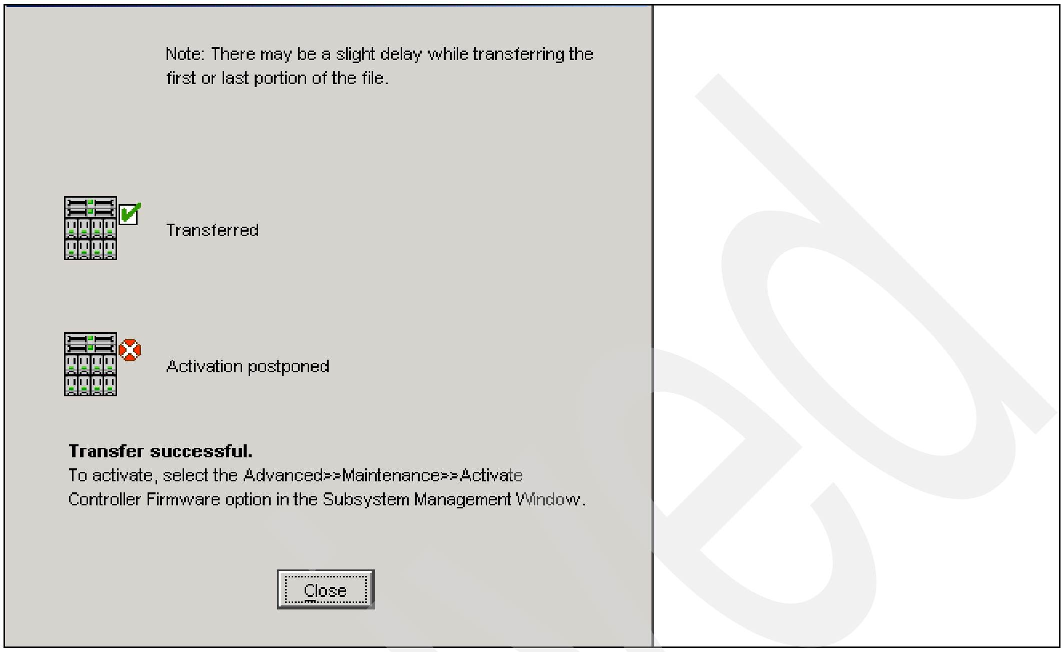


Figure 9-11 Firmware/NVSRAM download finished

4. After clicking **Close**, you are back in the Subsystem Management window. Because this is a staged firmware upgrade, the new firmware is now ready for activation. This is indicated by an icon (blue 101) next to the Storage System name, as shown in Figure 9-12.

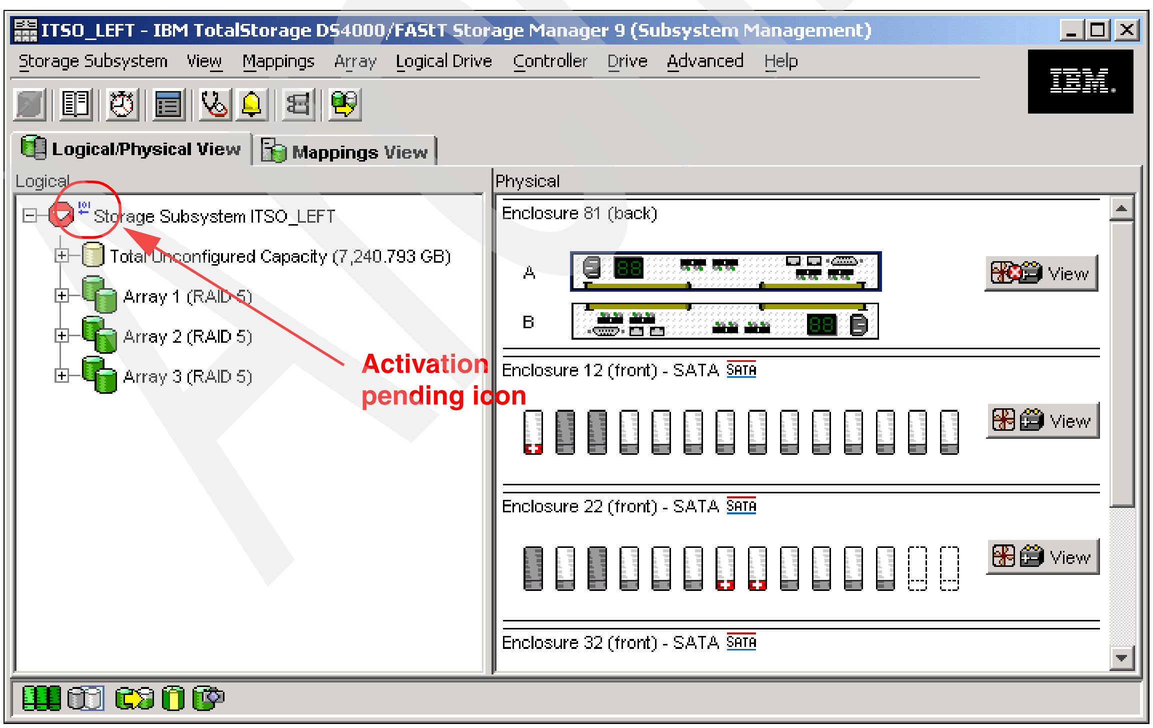


Figure 9-12 Subsystem Management window: Firmware ready for activation

5. To activate the new firmware, select **Advanced** → **Maintenance** → **Activate Controller Firmware**, as shown in Figure 9-13.

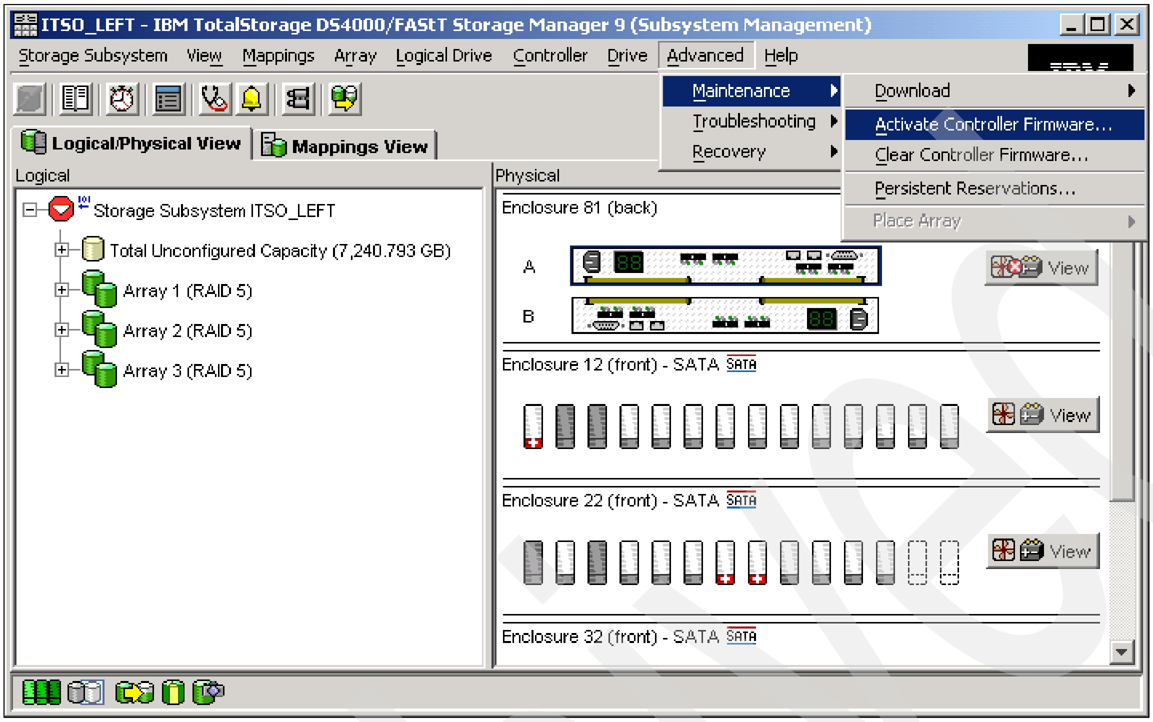


Figure 9-13 Subsystem Management window: Firmware activation

The Activate Firmware window opens and asks you for confirmation to continue. After you click **Yes** the activation process starts. The activation applies the transferred code to the controllers, rebooting first one, then the other. If you have all your hosts with path redundancy, you should not see any problem more than the disks going through one controller to the other while they are rebooted. We recommend scheduling this activation during a period of low I/O access.

You can monitor the progress in the Activation window, as shown in Figure 9-14.

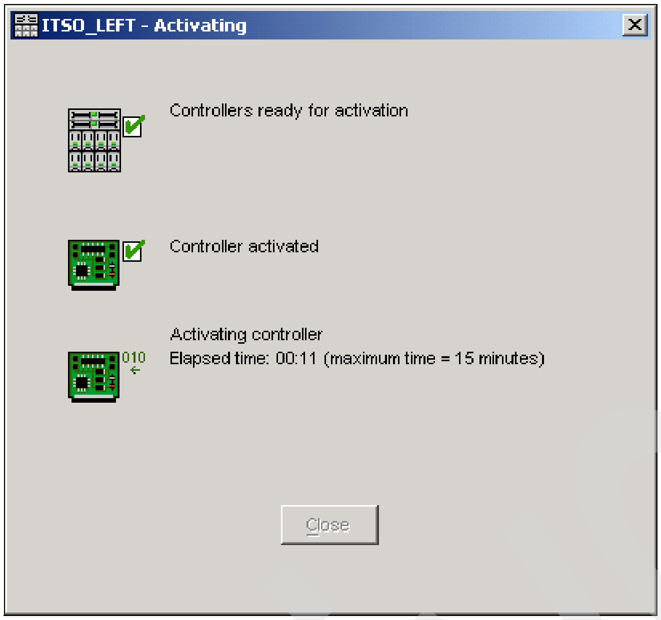


Figure 9-14 Activating the firmware

6. When the new firmware is activated on both controllers, you will see the `Activation successful` message. Click **Close** to return to the Subsystem Management window.

9.3.5 Updating the ESM board firmware

An environmental services module (ESM) card monitors the environmental condition of the components in that enclosure. It also provides the interface between the Fibre Channel drives in a given storage expansion enclosure with other ESM cards and controller blades in a drive loop.

Before performing any upgrade, read the specific firmware version readme file for details of installation. Pay attention to any dependencies between controller and ESM, or drives, since there may be a specific sequence in which each of the components should be updated.

In case an ESM is replaced in an EXP810, there is a firmware synchronization feature that ensures that new ESM is automatically synchronized with the firmware in the existing ESM. This resolves any ESM firmware mismatch conditions automatically. You still have the option to update the ESM firmware manually, to a new level, or in the case of a failure that could not let the code synchronization identify which is the original ESM and which is the replacement to synchronize.

For EXP710 and EXP810 there is an option to set the enclosure settings using the Storage Manager by selecting **Maintenance** → **Download** → **ESM Configuration Settings**. This option should be used when ESM reports different versions of configuration settings and the software could not resolve the mismatch problem by automatically synchronizing the configuration settings between the two ESMs automatically.

> **Important:** Before upgrading the ESM firmware make sure that the system is in an optimal state. If not, run the Recovery Guru to diagnose and fix the problem before you proceed with the upgrade.

If you have expansion enclosures, use the following instructions to transfer a firmware file to the ESM boards:

1. From the Subsystem Management window select the **Advanced** → **Maintenance** → **Download** → **ESM Firmware**.

   On the DS3000, go to the Support page on the Subsystem Management view and select **Download Firmware**. Then select **Download Environmental (ESM) Card Firmware**.

2. The Download Environmental (ESM) Card Firmware main window is displayed (see Figure 9-15).

   – The select enclosures table lists all the enclosures found attached to the storage array that contains ESM cards.

   – The select file allows you to specify the ESM firmware file to use as the source of the upgrade.

> **Note:** If an ESM card does not appear in the list (because of a loss of redundancy or some other problem), run the Recovery Guru to diagnose and fix the problem before continuing with the download to avoid loosing both paths to the disks.

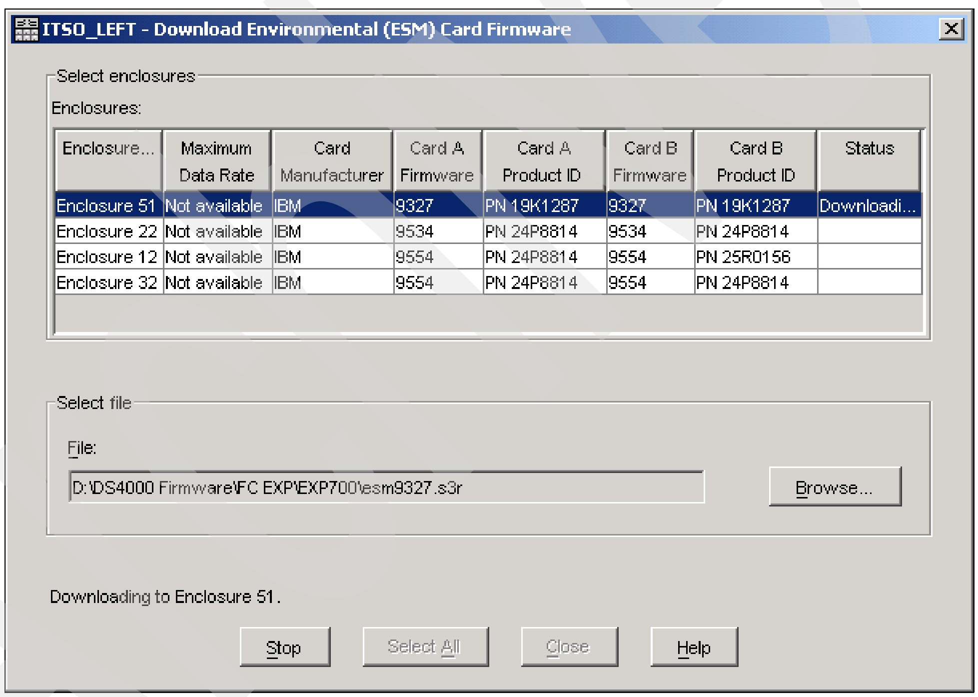


Figure 9-15 Download ESM firmware dialog

3. In the Select enclosures area, highlight each enclosure to which you wish to download firmware or select the **Select All** button to highlight all drive enclosures in the Storage System (each drive enclosure selected should have the same product ID).

4. Enter the firmware file to download in the Select file area by either entering the location and name of the file in the Select file text box, or selecting **Browse** and getting the firmware file from a local or network drive. (The Browse button is unavailable until an enclosure has been selected.)

5. Select **Start**. Confirm your selections and then select **Yes** to continue with the firmware download or **No** to quit.

6. The Status field in the Select enclosures table changes from pending to downloading for the ESM card firmware operation in progress.

   Monitor the progress and completion status of the download to the enclosures. The progress and status of each drive enclosure participating in the download is displayed in the status field of the Select enclosures table.

9.3.6 Updating hard disk drives firmware

Updating the hard disk drives firmware is sometimes required after upgrading the ESM firmware or the controller firmware/NVSRAM. Always refer to the readme file associated with the hard drive and ESM firmware upgrade package for precise instructions.

Parallel drive firmware download

Storage Manager allows the download of hard disk drive firmware to several drives in parallel. This way, large configurations with multiple expansion enclosures are not affected by the download time that updating a large amount of drives might generate.

However, make sure that you are not using an old version of Storage Manager, since versions prior to 9.10 manage the download to the drives in a serial way, thus impacting the total amount of time required.

You can update up to four different drive types (for example, ST373307FC, ST336732FC, ST336704FC, and ST318451FC) simultaneously. It does not matter whether there are different types or different firmware versions.

> **Note:** All I/O has to be stopped while downloading the firmware to the drives.
>
> In the DS4000 storage subsystems with the FC/SATA intermix premium feature enabled do not download the drive firmware to both SATA-technology hard drives and Fibre Channel technology hard drives at the same time. Complete the drive firmware download to all of the drives of a drive technology (either SATA or FC) before downloading the drive firmware to the drives of the other drive technology (either FC or SATA).

You can find the firmware files and full instructions at:

http://www-1.ibm.com/servers/storage/support/disk

To update the firmware of the disk drives:

1. To start the hard disk drives firmware update process, select **Advanced** → **Maintenance** → **Download** → **Drive Firmware/Mode Pages**.

   On the DS3000, go to the Support page on the Subsystem Management view and select **Download Firmware**. Then select **Download Drive Firmware**.

2. After taking the option from the Storage Manager, the window shown in Figure 9-16 appears. This window lists current drive firmware package information and allows you to specify up to four firmware update packages. Click the **Add** button to select these drive firmware packages.

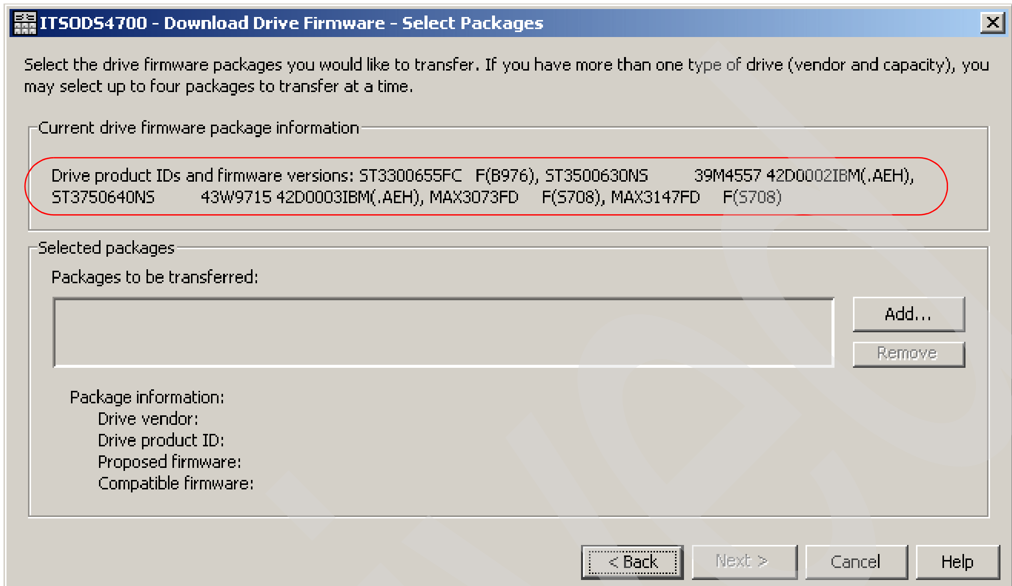


Figure 9-16 Drive firmware: Adding packages

Notice that at the top there is a section describing all the disk types detected in your system with their current firmware levels. This helps you choose which of the files to select, since you will have a large list of files downloaded with the hard drive microcode package.

3. In the following window, browse to the directory where you downloaded the drive firmware files. Select a package based in your disk types. You can use the top section while browsing through the different files. It shows whether each of them is compatible with the disks that you have in your enclosures. Select a compatible file and click **OK**.

   You can repeat steps 2 and 3 to select up to four drive packages.

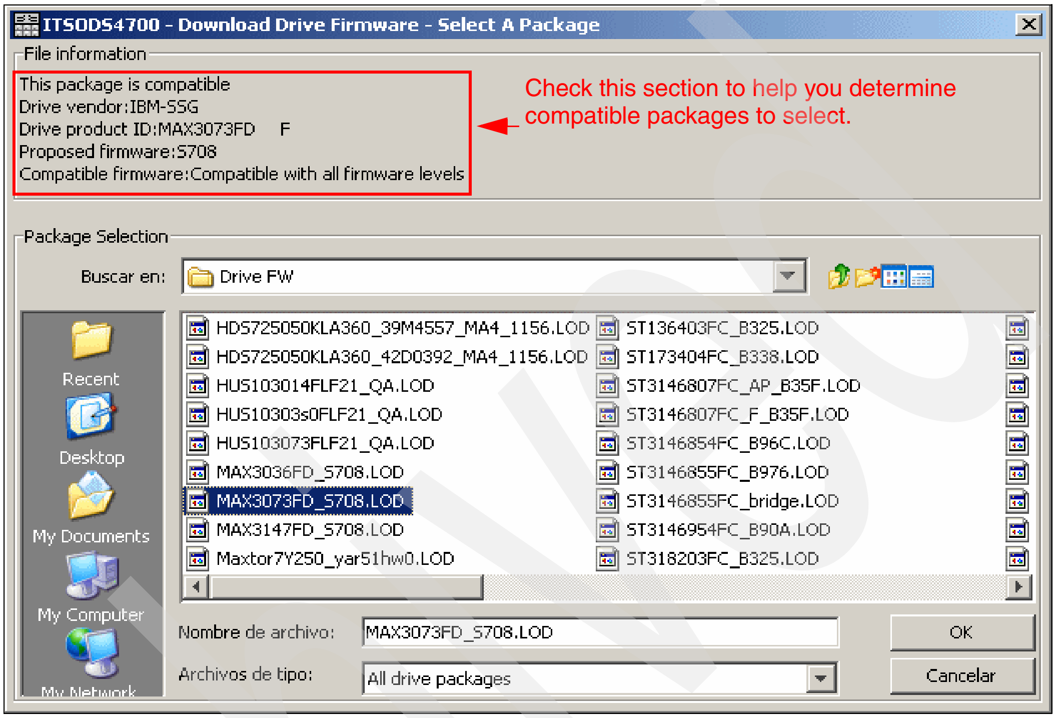


Figure 9-17 Drive Firmware - Selecting package

4. The next step is to select the down-level drives and mark them for update (see Figure 9-18).

   Select a drive from the Compatible Drives section and click **Finish** to continue. You must type `Yes` in order to confirm that you are sure that you want to update the firmware.

   As a suggestion, update only one drive firmware to check the image file, and then if successful, install all the other drives of the same type in parallel.

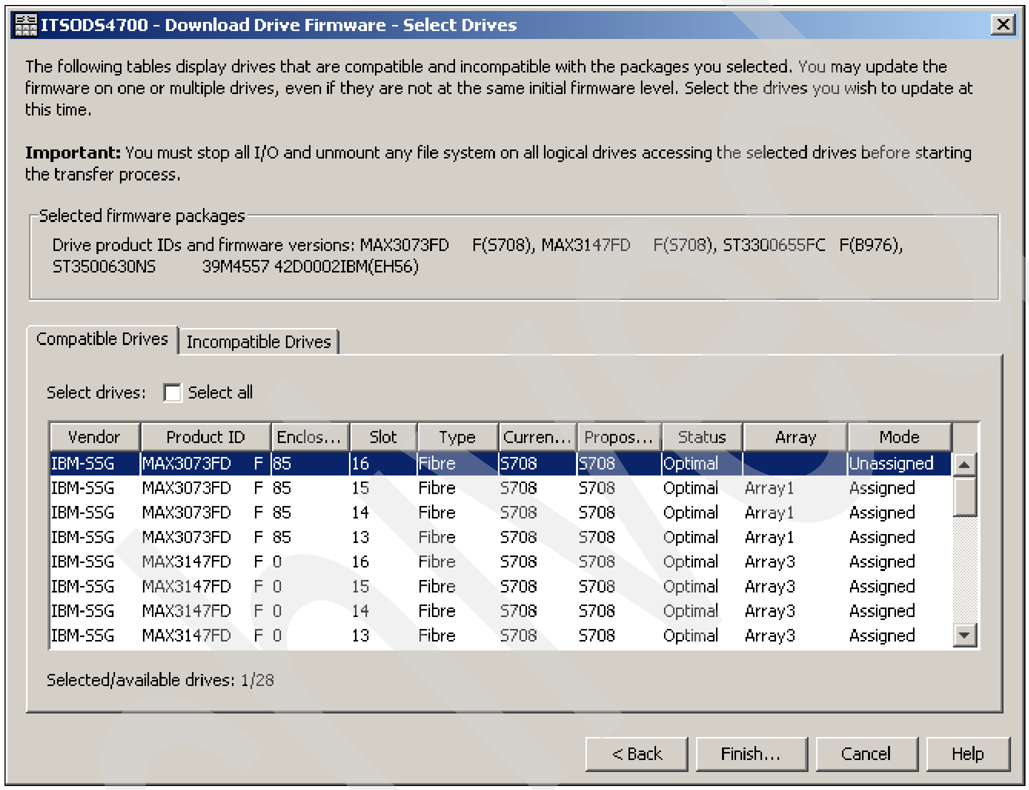


Figure 9-18 Drive firmware update: Select drives

5. Once the drive firmware download procedure starts you can monitor the progress, as shown in Figure 9-19. This gives you information about the current status of the download. After the process is finished you can also see whether all drives were successfully updated.

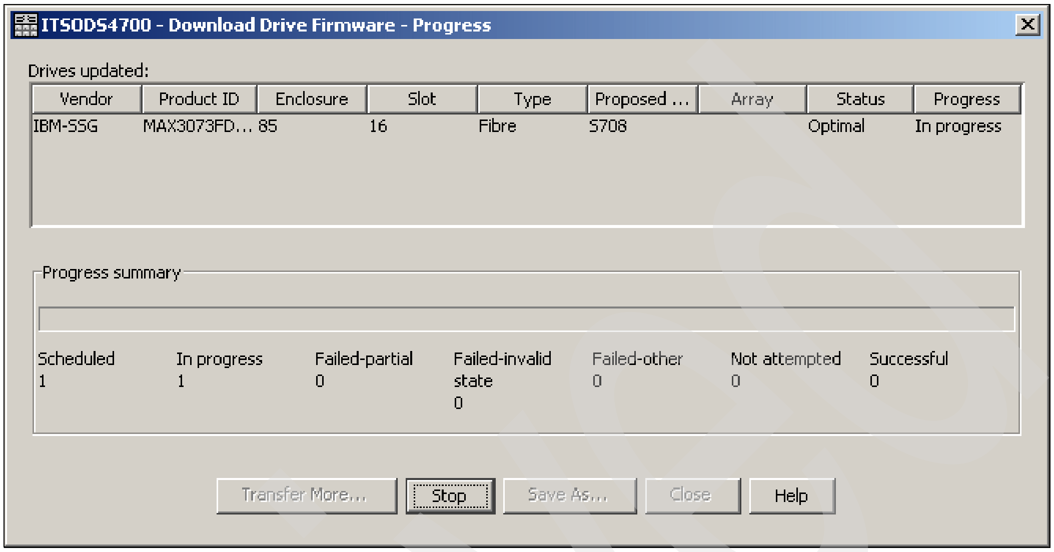


Figure 9-19 Drive firmware update: Download progress window

9.3.7 Handling premium features

Depending on your specific DS3000, DS4000, or DS5000 model, you might want to use a specific function or capability of your subsystem. Some of these capabilities, although included in the firmware, are not enabled by default. These are called premium features, and you must activate them by entering an activation key. You must buy a license for the premium feature to get a key and instructions about how to activate it.

You can activate an individual premium feature by providing its corresponding key, or activate a pack or bundle of multiple features.

Below we discuss how to order, list, and install any of the premium features available for your DS4000. The available premium features are:

- Storage partitioning
- FlashCopy
- Enhanced Remote Mirroring (ERM)
- Volume Copy
- FC/SATA Intermix
- Enhanced performance (DS4800 mod80 and DS5100)

Not all the premium features are available for all the models, and there are several new options with SM 10.10 (controller firmware V7.10). For example, the storage partitioning premium feature could be ordered to allow up to 32, 64, 96, 128, 256, or 512 partitions, depending on your DS4000 model and needs.

Listing premium features/feature enabler

From the Storage Manager client GUI, you can check what premium features are already activated in your DS4000 subsystem.

New with Storage Manager 10 and its related firmware, this information, with additional details, is generated as a text file. The file is available within the zip package generated by selecting the option Collect all Support Data from Storage Manager V10.10 and higher.

1. Select **Storage Subsystem** → **Premium Features** → **List** from the Subsystem Management window, or select **Premium Features** → **List** from the pop-up menu. As a result, the Premium Features dialog is displayed with a list of enabled premium features (Figure 9-20).

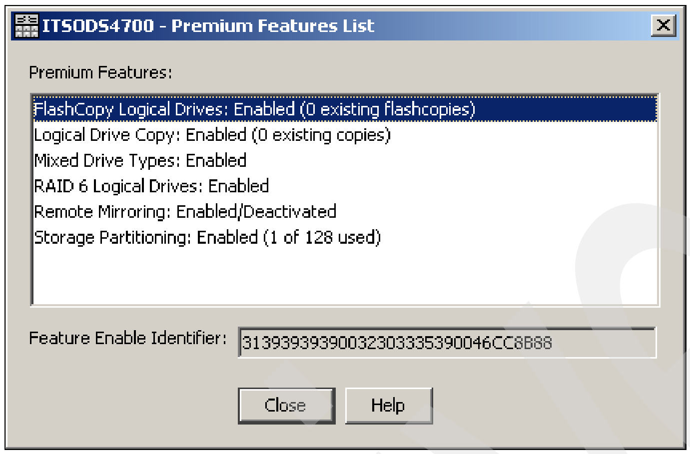


Figure 9-20 Listing of premium features

   If you receive a `Premium Features - Out of Compliance` error message during a management session, use the Recovery Guru to resolve the problem.

2. Collect all support data by selecting **Advanced** → **Troubleshooting** → **Support Data** → **Collect**. Provide a file name and directory where you want to store the zip file generated.

3. Once the file is generated, open the zip file and then look at the file named featureBundle.txt to see the status, limits, and in-usage capacities for the premium features, as shown in Example 9-26.

Example 9-26 Collect support data featureBundle.txt file

```
FEATURE BUNDLE FOR STORAGE SUBSYSTEM:     ITSODS4700 (28/08/07 20:34:11)
Total logical drives allowed:           1024
Total logical drives allowed per partition:  256
Management application client:           1
Drives supported:                    B; J; S; BM
Drive enclosures supported:                 IBM,EXP100,dSATA; IBM,EXP700,ESM_CARD;
IBM,EXP710,SBOD_ESM; IBM,EXP810,HUSKER
Drive enclosure expansion:             7
High Performance Tier:                      No data is available about supported values of this
feature
Storage Partitioning
   Enabled by default:      Yes
   Default limit:           8
   Enable/Upgrade via key:  Yes
   Limit with feature key:  128
FlashCopy Logical Drives
   Enabled by default:         No
   Enable/Upgrade via key:     Yes
   Total flashcopies allowed:  512
Logical Drive Copy
   Enabled by default:      No
   Enable/Upgrade via key:  Yes
   Total copies allowed:    1024
Remote Mirroring
   Enabled by default:      No
   Enable/Upgrade via key:  Yes
   Total mirrors allowed:   32
Mixed Drive Types
   Enabled by default:      No
   Enable/Upgrade via key:  Yes
```

4. Gather the following data along with the Feature Enable Identifier:
   – Machine type
   – Model
   – Serial number
   – Controller firmware version (optional but recommended since some premium features are only available for specific versions)

> **Note:** Machine type, model, and serial information are printed on a label on the back of your DS4000 controller unit. To view the installed level of firmware see 9.3.1, "Displaying installed firmware versions" on page 358.

Enabling a premium feature

How you obtain a feature key varies depending upon DS4000 packaging procedures and time of order:

- If you bought any premium feature together with the DS4000, you will receive the feature keys with the DS4000 hardware, but you still have to activate them separately.
- If you are purchasing a new premium feature you can get the key from your local storage support (IBM or Business Partner).

In any case, a premium feature is a chargeable option for every DS4000, and you have to request the feature from your sales contact person for a specific machine type, model, and serial number.

Once you have purchased a premium feature, you receive an envelope containing a license activation card for that particular feature, with detailed instructions on how to proceed to activate it.

Important: All the premium features are activated immediately and do not require a reboot of the controllers. The exception to that is the new enhanced performance feature. Make sure to follow the instructions provided and schedule the activation of the enhanced performance feature as a non-concurrent update.

The procedure below will help you activate the new feature obtained or reactivate it if for any reason it is out of compliance:

1. On the card that you received locate the feature activation code and make sure that the instructions received are for your machine type and model. The feature activation code (XX-XXXX-XXXXX) is located at the top of the license activation card.
2. From the Subsystem Management window, select either the **Storage Subsystem** → **Premium Features** → **List** pull-down menu option, or **Premium Features** → **List** from the right-click pop-up menu (see Figure 9-20 on page 375).
3. Write down the 32-digit number next to the feature enable identifier, or preferably copy and paste the complete number to a text file to avoid typing errors.

   You can also find the feature enable identifier in the profile data, either by selecting **View-Profile** or by opening the storageArrayProfile.txt file contained in the zip file generated collecting all support data.
4. Go to the following Web site to personalize the key received to your specific system, and generate the activation key file:

   `http://www.ibm.com/Premium Features/`

   Select the option to activate a premium feature.
5. Complete the form in the Web page with the following:
   - Feature activation code
   - Feature enable identifier
   - Machine type
   - Model number
   - Unit serial number

   The Web page generates a an activation key file that you have to download to your management station. It also sends the file to the specified e-mail address.

6. Once you have received the key file, go to Storage Manager and click **Storage Subsystem** → **Premium Features** → **Enable** in the Subsystem Management window, as shown in Figure 9-21.

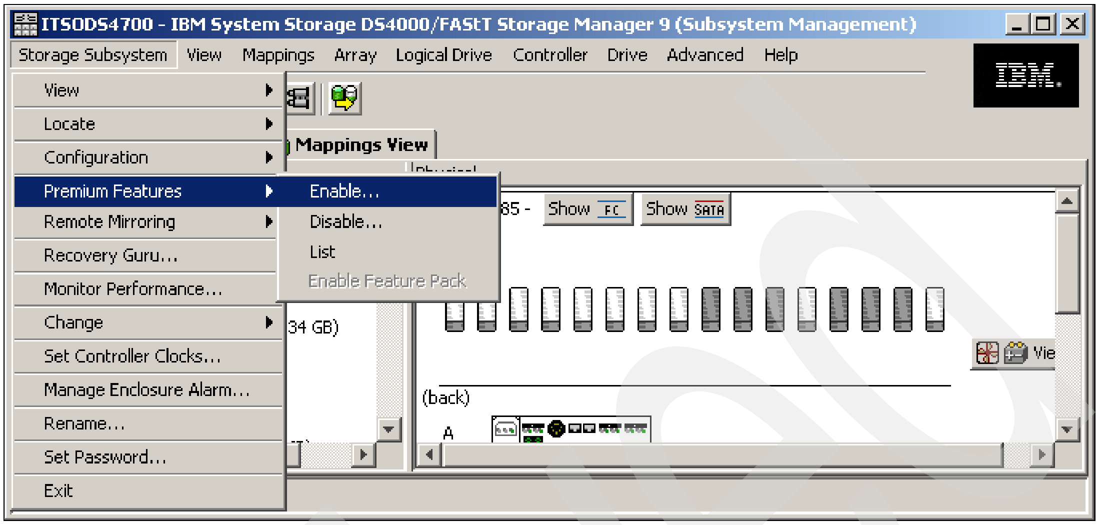


Figure 9-21 Option to enable a premium feature

7. In the dialog window, point to the location where the key file is stored. You must confirm whether to enable the premium feature selected (as shown in Figure 9-22).

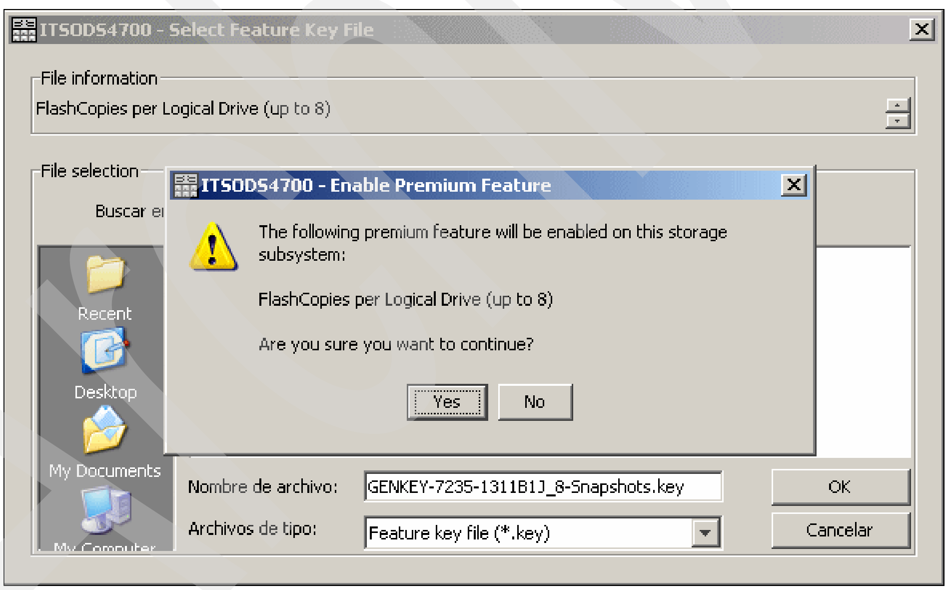


Figure 9-22 Enabling a premium feature

The DS4000 subsystem validates the code supplied to make sure that it is suitable for the specific serial number and is compatible with the machine type and model. It also checks that it is not already installed.

If everything is okay, the feature is applied and immediately available for use (with the exception for the enhanced performance feature that first requires a reboot of the controllers).

If the feature enable identifier does not match the DS4000 power supply fan unit, or it is already installed, you receive a notification and the key will not be installed.

Disabling a premium feature

To disable a premium feature, click **Storage Subsystem** → **Premium Features** → **Disable** in the Subsystem Management window. Choose the feature that you want to disable from the list, as shown in Figure 9-23, and confirm.

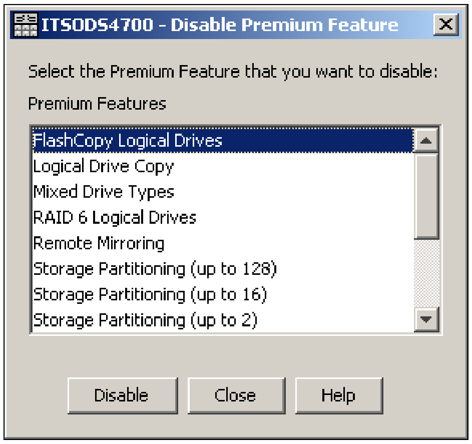


Figure 9-23 Disable Premium Feature

Keep in mind that the change happens immediately. If you use storage partitioning and disable the premium feature, then you cannot create new partitions. However, any existing partitions remain operational.

9.3.8 Saving and loading the configuration

Once your DS4000 storage subsystem is configured and running, you should save this configuration to be able to restore it in case of problems.

The saved configuration includes the array and logical drive configuration, the name of the subsystem, its cache settings, and other parametrization, including the storage partitioning configuration.

The saved file can be used to restore the configuration data to the same DS4000, or also to other DS4000 Storage subsystems in case you want to set up multiple Storage Systems with the same configuration. To allow that, the destination subsystem must have the same hardware layout, number of enclosures and drives, and drive capacities.

All information is stored in a file that contains a script for the script editor. To save the configuration of the subsystem, open the Subsystem Management window, highlight the subsystem, and click **Storage Subsystem** → **Configuration** → **Save** (Figure 9-24).

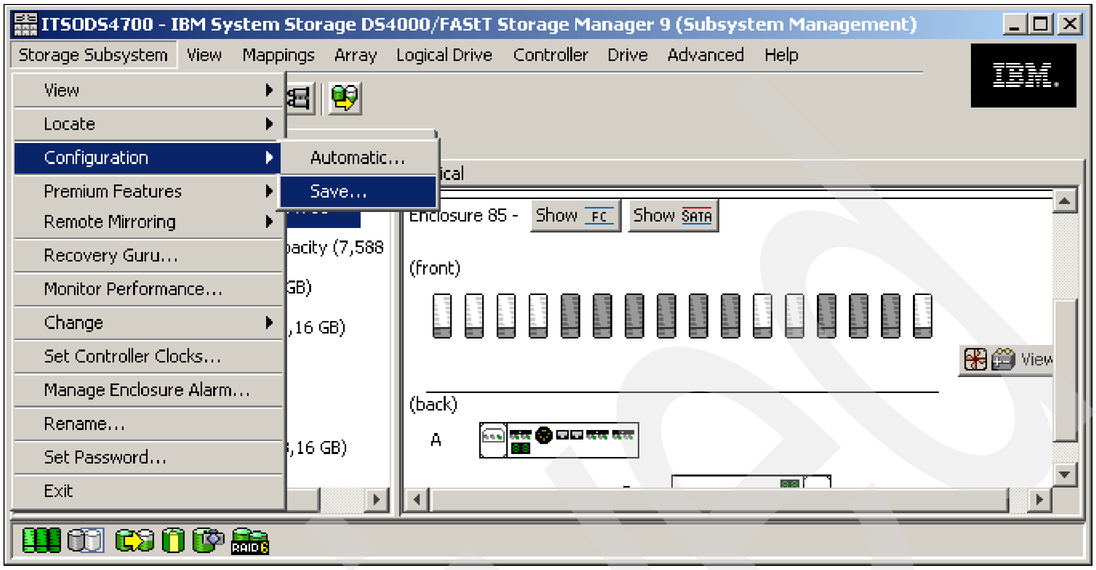


Figure 9-24 Saving DS4000 configuration

We can choose to save specific elements of the configuration (Figure 9-25).

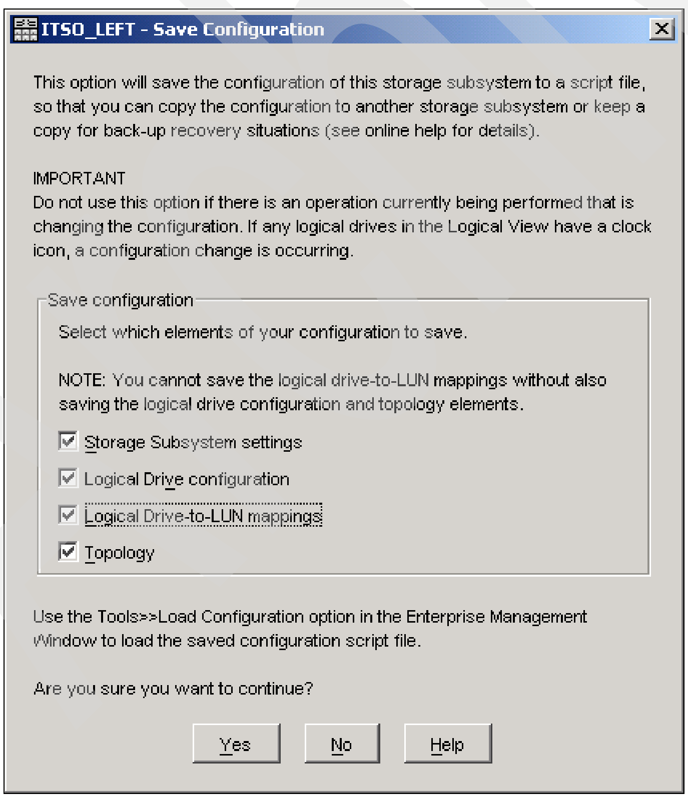


Figure 9-25 Saving configuration elements

Select the desired configuration elements, click **Yes**, and select a file name and destination folder in which to the save file. Make sure not to use a directory located on a DS400 disk, otherwise, you might not be able to access it when needed.

The script created can be used to replicate the configuration of the DS4000. You can apply the configuration to the destination subsystem for all the saved elements or any particular element. The script consists of the following information:

- Storage subsystem settings
  - User label
  - Media scan rate
  - Cache block size
  - Cache flush start
  - Cache flush stop
  - Default host type
  - Failover alert delay
- Logical drive configuration
  - RAID level
  - User label
  - Owning controller
  - Segment size
  - Capacity
  - Cache flush modifier
  - Read-ahead multiplier
  - Modification priority
  - Caching without batteries enabled/disabled
  - Cache mirroring enabled/disabled
  - Read caching enabled/disabled
  - Write caching enabled/disabled
  - Media scan enabled/disabled
  - Redundancy check enabled/disabled
- Logical drive-to-LUN mappings
- Topology
  - Host groups
  - Hosts and parent host groups
  - Host ports, associated host type, and parent hosts

You have two possibilities to load the configuration to a DS4000 with the Storage Manager.

> **Attention:** This procedure replaces any configuration on the Storage System. All data stored on the DS4000 power supply fan unit is lost because all logical drives are initialized. It is a good idea to save the configuration every time that a change on the Storage System is made.
>
> Do not attempt to load a saved configuration on our DS4000 unless you fully understand the consequences.

- Open the Enterprise Management window and select the subsystem. Click **Tools** → **Load Storage Subsystem Configuration** from the tool bar menu. Point to the file containing the configuration and load it. The script editor and a warning message appear. To load the configuration onto the DS4000 power supply fan unit, choose **Execute**. You can also edit the script before executing.

- Open the Enterprise Management window and select the subsystem. Click **Tools** → **Load Configuration** from the menu. Point to the file containing the configuration and load it. The script editor and a warning message appear. To load the configuration onto the DS4000 power supply fan unit, choose **Execute**. You can also edit the script before executing it.

The procedure can take a long time, depending on the number of arrays and logical drives defined. When the procedure finishes, the subsystem contains the same configuration as the source.

While the configuration file allows an immediate recreation of all the parameters configured, it does not provide a friendly reading file to list all the configuration. To read the current configuration, we use the view profile option described in the next section.

9.3.9 Storage subsystem profile

The profile is one of the most important items needed for support to help you solve whatever problem you have with your DS4000. It is a simple text file that includes data about the various firmware levels, array and volume configuration, storage partitioning, and the status of all the hardware components of the DS4000.

It is also presented through the Storage Manager interface in a readable format, as opposed to the save configuration option (script), which is mostly used for backup and restore purposes.

> **Note:** You should always save your profile after you have changed the configuration of the DS4000. For example, if you create or delete logical drives, change the mapping, or add new disks or enclosures to the DS4000, you should save a new profile, as the IBM Support might need it to help you in case of any problems.

To save the profile click **Storage Subsystem** → **View** → **Profile** in the Subsystem Management window. There are seven different sections (controllers, arrays, logical drives, drives, drive channels, enclosures, and mappings). The section *all* simply shows all seven sections on one page. Click **Save as** to continue saving the profile.

Reading the profile

Reading and interpreting the profile is a task usually done by IBM Support, but there are some common information and failure situations that can be used, analyzed, and fixed easily by an administrator.

Controller and NVSRAM information

In this example you can see the firmware and NVSRAM versions of controller A:

```
Controller in Enclosure 82, Slot A

         Status: Online
         Current configuration
            Firmware version:       06.16.63.00
               Appware version:     06.16.63.00
               Bootware version:    06.16.63.00
            NVSRAM version:         N1815D480R916V06
         Pending configuration
            Firmware version:       None
               Appware version:     None
               Bootware version:    None
            NVSRAM version:         None
```

```
         Transferred on:          None
```

HDD information

The Drives tab in the profile data provides the following information:

```
DRIVES------------------------------

SUMMARY
   Number of drives: 54
      Supported drive types: Fibre (14), Serial ATA (SATA) (40)

   BASIC:
      TRAY, SLOT  STATUS   CAPACITY    CURRENT DATA RATE  PRODUCT ID FIRMWARE VERSION
      12, 1       Optimal  232.882 GB  2 Gbps             Maxtor 7Y250M0 YAR51HW0
      12, 2       Optimal  232.882 GB  2 Gbps             Maxtor 7Y250M0 YAR51HW0
      12, 3       Optimal  232.882 GB  2 Gbps             Maxtor 7Y250M0 YAR51HW0
      12, 4       Optimal  232.882 GB  2 Gbps             Maxtor 7Y250M0 YAR51HW0
      12, 5       Optimal  232.882 GB  2 Gbps             Maxtor 7Y250M0 YAR51HW0
.......
```

Mappings information

The following is what you find in the profile Mappings tab, similar to the mapping section of the Storage Manager client:

```
MAPPINGS (Storage Partitioning - Enabled (2 of 128 used))-------------------

   Logical Drive Name    LUN  Controller  Accessible by       Logical Drive status
   Access Logical Drive  31   A,B         Host Group AIX      Optimal
   Data1                 0    A           Host Group AIX      Optimal
   Data2                 1    A           Host Group AIX      Optimal
   Access Logical Drive  31   A,B         Host Group Windows  Optimal
   1                     1    A           Host Group Windows  Optimal
   2                     2    B           Host Group Windows  Optimal
   3                     3    A           Host Group Windows  Optimal
   4                     4    B           Host Group Windows  Optimal
```

Reading profile in a failed disk scenario

Using the profile data, selecting the **Arrays** tab we see the information shown in Example 9-27.

Example 9-27 Failed drive protected by hot spare

```
ARRAYS------------------------------

   Number of logical drive groups: 3
   Name:                         Data
         Status:                     Degraded - partially complete array
         Capacity                    930,523 GB
         RAID level:                 3
         Drive type:                 Serial ATA (SATA)
         Enclosure loss protection:  No
         Current owner:              Controller in slot B
         Associated logical drives and free capacity
         Logical Drive     Capacity
           AIX1              9,000 GB
         Free Capacity     921,523 GB


         Associated drives - present (in piece order)
         Enclosure     Slot
```

```
         85           1
         85           2
         85           3
         85           4 [hot spare drive is sparing for drive at 85, 1]
         85           1
```

You can see that the array named data is a RAID-3 built of the disks in slots 1, 2, and 3 of enclosure 85. However, the drive in 85,1 seems to be failed, as a hot spare at location 85,4 is sparing for the drive in 85,1.

The status of the array is degraded, so we guess that the reconstruction is still in progress to the hot spare drive, but we must check the logical drives (AIX1and AIX2 in this case) that are part of this array.

Example 9-28 Profile data: Logical drives

```
STANDARD LOGICAL DRIVES------------------------------
SUMMARY
Number of standard logical drives: 9
   See other Logical Drives sub-tabs for premium feature information.

   NAME     STATUS   CAPACITY  RAID LEVEL  ARRAY   DRIVE TYPE
   1        Optimal  2,0 GB    5           Array1  Fibre
   2        Optimal  4,0 GB    5           Array1  Fibre
   3        Optimal  4,0 GB    5           Array1  Fibre
   4        Optimal  2,0 GB    5           Array1  Fibre
   5        Optimal  2,0 GB    5           Array1  Fibre
   AIX1     Optimal  9,0 GB    3           Data    SATA
   AIX2     Degraded 9,0 GB    3           Data    SATA
```

AIX1 has already finished the reconstruction to the hot spare as the status is optimal, but the AIX2 logical disk is still reconstructing since it still shows degraded status. We can cross check the information about the drives by using the Drives tab of the profile data (Example 9-29).

Example 9-29 Profile data: Drives

```
DRIVES------------------------------

SUMMARY
   Number of drives: 32
      Current drive types: Fibre (24), Serial ATA (SATA) (8)

   BASIC:
     TRAY, SLOT  STATUS   CAPACITY    TYPE   CURRENT DATA RATE  PRODUCT ID            FIRMWARE VERSION
    0, 1       Optimal  698,638 GB  SATA   4 Gbps          ST3750640NS        43W9715 42D0003IBM   .AEH
    0, 2       Optimal  698,638 GB  SATA   4 Gbps          ST3750640NS        43W9715 42D0003IBM   .AEH
...
    0, 16      Optimal  136,732 GB  Fibre  4 Gbps           MAX3147FD    F                        S708
    85, 1      Failed   465,762 GB  SATA   4 Gbps          ST3500630NS        39M4557 42D0002IBM   .AEH
    85, 2      Optimal  465,762 GB  SATA   4 Gbps          ST3500630NS        39M4557 42D0002IBM   .AEH
    85, 3      Optimal  465,762 GB  SATA   4 Gbps          ST3500630NS        39M4557 42D0002IBM   .AEH
    85, 4      Optimal  465,762 GB  SATA   4 Gbps          ST3500630NS        39M4557 42D0002IBM   .AEH
...
    85, 16     Optimal  68,366 GB   Fibre  4 Gbps           MAX3073FD    F                        S708
```

Here we see that the drive in 85,1 has the status failed and needs to be replaced.

9.4 Troubleshooting

In this section we discuss the different ways to do troubleshooting on:

- IBM i
- VIOS
- DS Storage

9.4.1 Troubleshooting IBM i

In this section we describe the recommended actions to take in the System i partition when you experience a problem with Midrange Storage connected to IBM i.

We suggest that you check the following settings in a System i partition:

- Tagging for the load source unit
- Status of disk units (logical drives in Midrange Storage)
- Assignments of virtual SCSI adapters
- Possible error logs in

Tagging the load source unit

When experiencing a problem with IPL from a logical drive in Storage System, check whether the correct adapter is tagged as load source unit. For this, select the relevant System i partition in HMC, select **Configuration** → **Manage Profiles** from the pull-down, select **Edit** from Actions pull-down, and click the **Tagged I/O** tab. See Figure 9-26. Check whether the tagged adapter is assigned to VIOS and that the virtual disks are mapped to the relevant virtual SCSI adapter.

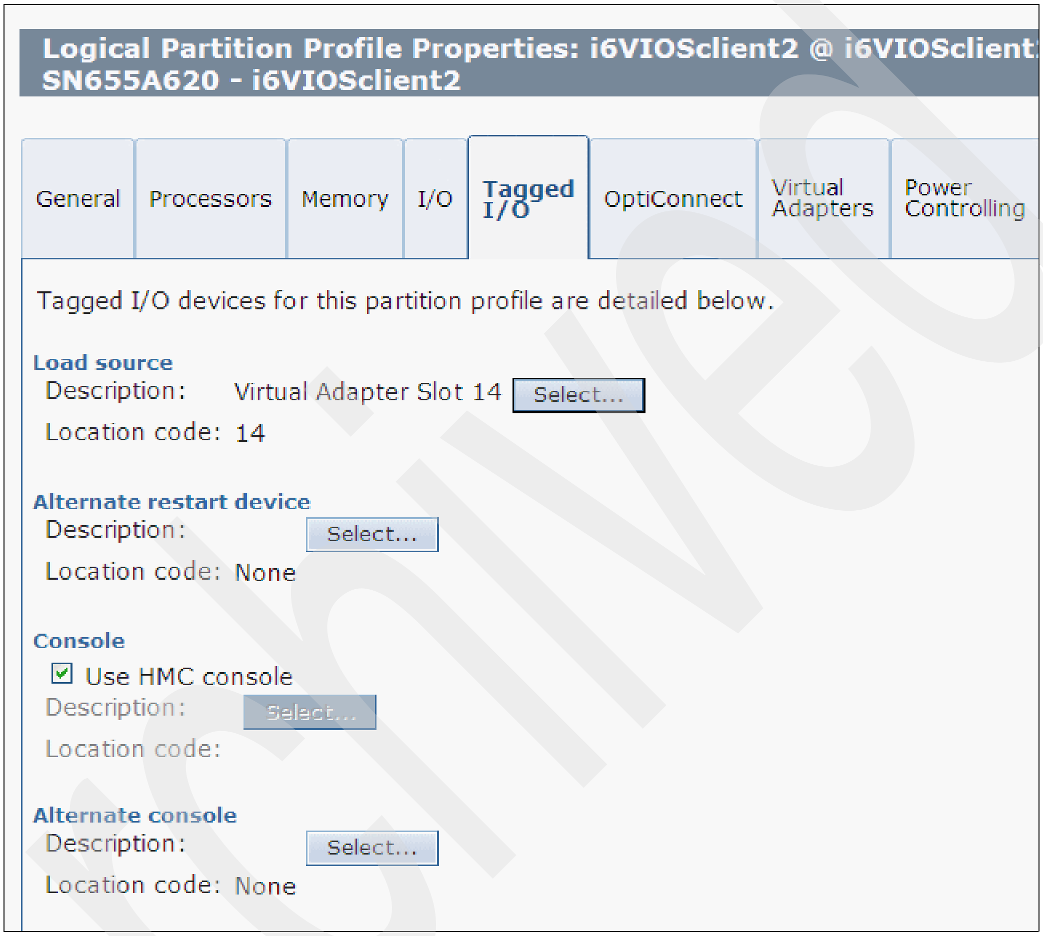


Figure 9-26 Tagged load source unit

Status of disk units

To check the status of disk units start Service Tools (SST) in the System i partition, select option **1. Start a Service Tool**, select option **7. Hardware Service Manager**, select option **2. Logical Hardwrae Resources**, and select option **1. System bus resources**. Look for the virtual IOP to whcih the external disks are assigned and use option **9. Resources associated with IOP)**. Check whether the disk status is operational, as shown in Figure 9-27.

```
                Logical Hardware Resources Associated with IOP

  Type options, press Enter.
    2=Change detail    4=Remove    5=Display detail    6=I/O debug
    7=Verify           8=Associated packaging resource(s)

                                                                Resource
  Opt Description                  Type-Model  Status           Name
      Virtual IOP               *    290A-001   Operational      CMB10
       Virtual Storage IOA           290A-001   Operational      DC06
        Disk Unit               *    6B22-050   Operational      DD001
        Disk Unit                    6B22-050   Operational      DD005
        Disk Unit                    6B22-050   Operational      DD002
        Disk Unit                    6B22-050   Operational      DD006






  F3=Exit    F5=Refresh    F6=Print    F8=Include non-reporting resources
  F9=Failed resources      F10=Non-reporting resources
  F11=Display serial/part numbers      F12=Cancel
```

Figure 9-27 Status of external disk units

Assignment of Virtual SCSI adapters

When experiencing problems with connected Midrange Storage it may be a good idea to check the assignments of Virtual SCSI adapters in the System i partition (IBM i client) and in VIOS. For this, use the following steps:

1. Select the relevant System i partition in HMC, select **Configuration** → **Manage Profiles** from the pull-down, select **Edit** from the **Actions** pull-down, select the **Virtual Adapters** tab, look for the relevant Client SCSI adapter, check that the Connecting Partition column contains the correct name of VIOS, and observe the number of the Connecting Adapter column. See Figure 9-28.

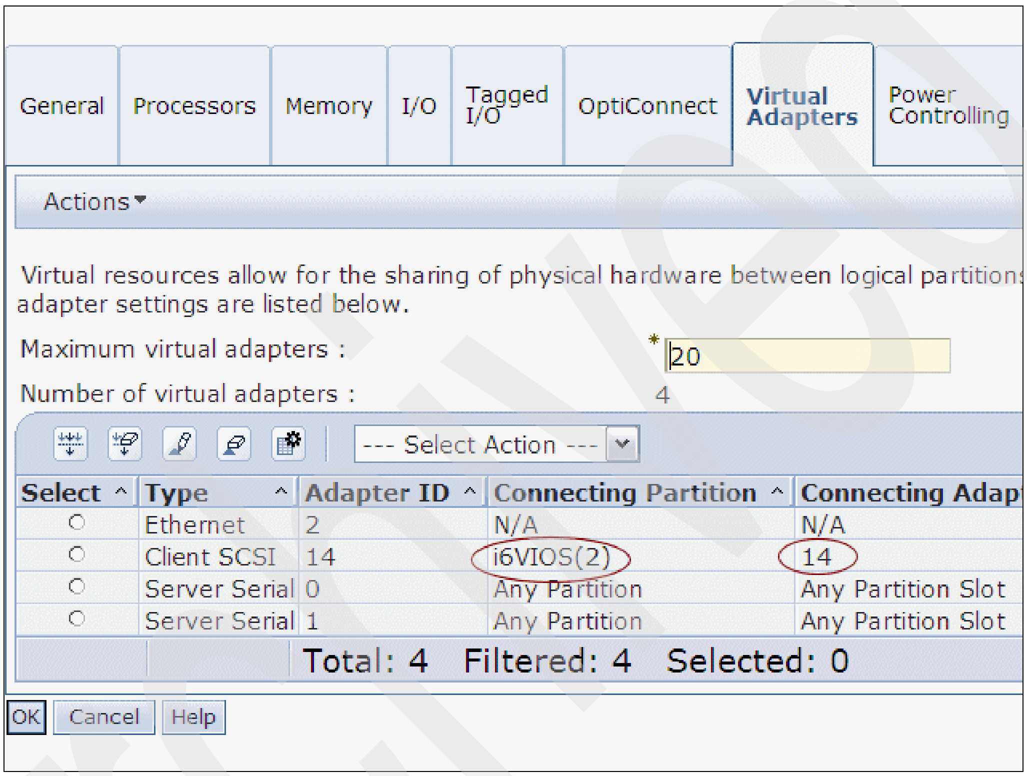


Figure 9-28 Assigned virtual adapters in IBM i client

2. Select the relevant VIOS partition in HMC, select **Configuration** → **Manage Profiles** from the pull-down, select **Edit** from the **Actions** pull-down, select the **Virtual Adapters** tab, look for the relevant Server SCSI adapters, check that one of them has your System i partition as the connecting partition, and check that the number of the connecting adapter is the same as you observed in step 1. See Figure 9-29.

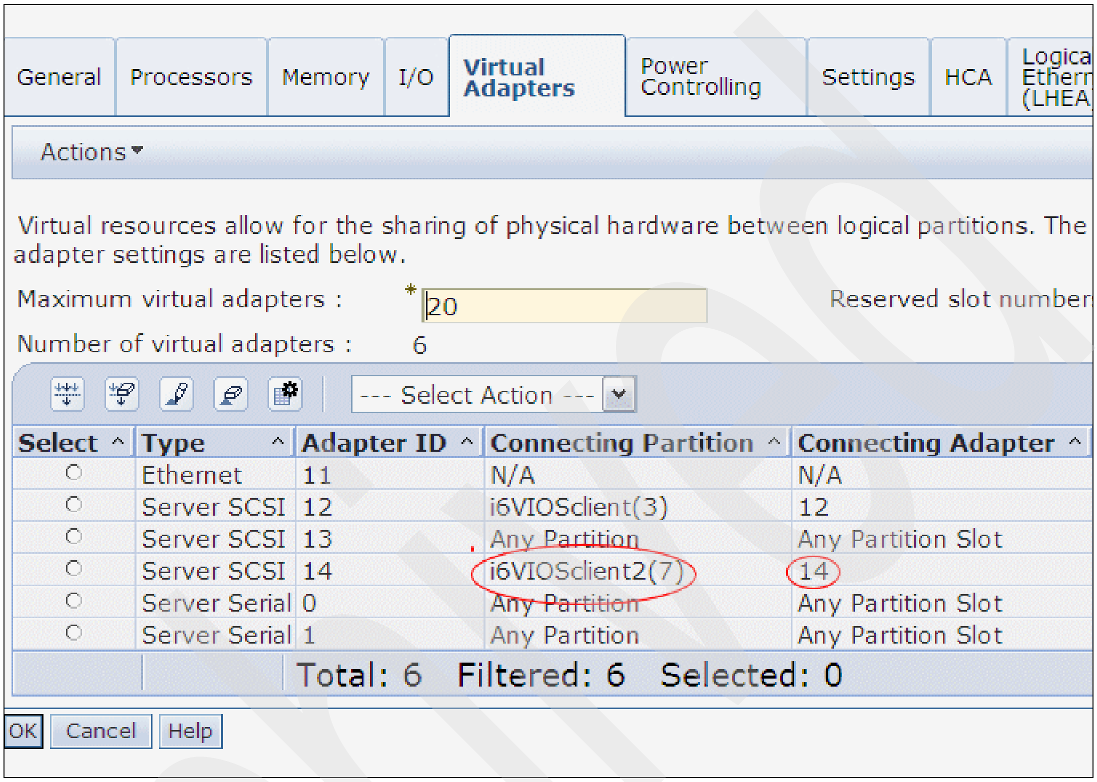


Figure 9-29 Assigned adapters in VIOS

Using product activity log (PAL)

In case of problems with connected disk units (logical volumes on Storage System) you may want to look at relevant PAL entries and also provide their contents to System i defect support. To do this:

1. Start SST. Select option **1. Start a Service Tool**, select option **1. Product activity log**, and select **option 1. Analyze log**. In the Select Subsystem Data panel choose the type of log to look to and select the date and time frame of log entries, as shown in Figure 9-30.

```
                              Select Subsystem Data

 Type choices, press Enter.

   Log . . . . . . . . . .  1          1=All logs
                                       2=Processor
                                       3=Magnetic media
                                       4=Local work station
                                       5=Communications
                                       6=Power
                                       7=Cryptography
                                       8=Licensed program
                                       9=Licensed Internal Code

   From:
     Date . . . . . . . .   08/25/08   MM/DD/YY
     Time . . . . . . . .   17:05:21   HH:MM:SS

   To:
     Date . . . . . . . .   08/26/08   MM/DD/YY
     Time . . . . . . . .   17:05:21   HH:MM:SS

 F3=Exit              F5=Refresh            F12=Cancel
```

Figure 9-30 Product activity log: Select subsystem data

2. On the next panel select the appropriate report type, option to include the reference code, and the appropriate option for devices. The selection in our example is shown in Figure 9-31.

```
Select Analysis Report Options

Type choices, press Enter.

  Report type . . . . . . . . .   1   1=Display analysis, 2=Display summary,
                                      3=Print options
  Optional entries to include:
    Informational . . . . . . .   Y   Y=Yes, N=No
    Statistic . . . . . . . . .   N   Y=Yes, N=No

  Reference code selection:
    Option  . . . . . . . . . .   1   1=Include, 2=Omit
    Reference codes
    *ALL                                                  *ALL...

  Device selection:
    Option  . . . . . . . . . .   2   1=Types, 2=Resource names
    Device types or Resource names
    *ALL                                                  *ALL...



F3=Exit       F5=Refresh       F9=Sort by ...       F12=Cancel
```

Figure 9-31 Selecting options for display PAL

3. On next panel select option 5 at the appropriate device or SRC code to display a report (Figure 9-32).

```
                              Log Analysis Report

From  . . :   08/25/08  19:18:38      To . . :   08/26/08  19:18:38

Type options, press Enter.
  5=Display report   6=Print report

     System                                         Resource      Resource
Opt  Ref Code     Date      Time      Class  Name          Type
 5   B2004158     08/25/08  22:07:17  Info                 9406










 F3=Exit
 F11=View Description                          F12=Cancel
```

Figure 9-32 Display Log Analysis Report

The displayed report in our example is shown in Figure 9-33.

```
 Display Detail Report for Resource

                                        Serial                  Resource
Name             Type      Model       Number                  Name
SYSTEM           9406      MMA         65-5A620

Log ID  . . . . . . . . . :   C800F7BA   Sequence . . . . . . . :        3900
Date  . . . . . . . . . . :   08/25/08   Time . . . . . . . . . :   22:07:17
Reference code  . . . . . :   4158       Secondary code . . . . :   00000000
Table ID  . . . . . . . . :   B200FA00   IPL source/state . . . :   B/7

Class . . . . . . . . . . :   Informational
System Ref Code . . . . . :   B2004158
No description is available





Press Enter to continue.

F3=Exit         F4=Additional Information               F6=Hexadecimal report
                                   F10=Previous detail report       F12=Cancel
```

Figure 9-33 Displayed report details

9.4.2 Troubleshooting VIOS

In this section we outline basic troubleshooting steps on VIOS like checking the error log and basic VIOS problem data collection.

Checking the VIOS error log

The error log on the VIOS partition usually shows relevant information regarding IBM DS Midrange Storage System problems like a RAID array disk drive failure, as shown in Example 9-30.

Example 9-30 VIOS error log entries for RAID array drive failure

```
$ errlog | more
IDENTIFIER TIMESTAMP  T C RESOURCE_NAME  DESCRIPTION
09890235   0812175008 P H hdisk8         ARRAY DRIVE FAILURE
B9735AF4   0812175008 P H hdisk8         SUBSYSTEM COMPONENT FAILURE
B9735AF4   0812175008 P H hdisk7         SUBSYSTEM COMPONENT FAILURE
09890235   0812175008 P H hdisk6         ARRAY DRIVE FAILURE
B9735AF4   0812175008 P H hdisk6         SUBSYSTEM COMPONENT FAILURE
B9735AF4   0812175008 P H hdisk5         SUBSYSTEM COMPONENT FAILURE
09890235   0812175008 P H hdisk4         ARRAY DRIVE FAILURE
B9735AF4   0812175008 P H hdisk4         SUBSYSTEM COMPONENT FAILURE
B9735AF4   0812175008 P H hdisk3         SUBSYSTEM COMPONENT FAILURE
09890235   0812175008 P H hdisk2         ARRAY DRIVE FAILURE
B9735AF4   0812175008 P H hdisk2         SUBSYSTEM COMPONENT FAILURE
```

Also for diagnosing SAN storage connectivity problems the error log may show helpful information like loss of a Fibre Channel link to a DS Storage System controller with a failure of the RDAC device driver controller health check run every 10 minutes by default, as shown in Example 9-31.

Example 9-31 VIOS error log entries for DS Storage System controller access loss

```
BC669AA7   0814184208 P H dac1          CONTROLLER HEALTH CHECK FAILURE
3074FEB7   0814184208 T H fscsi2        ADAPTER ERROR
3074FEB7   0814184208 T H fscsi2        ADAPTER ERROR
3074FEB7   0814184208 T H fscsi2        ADAPTER ERROR
3074FEB7   0814184208 T H fscsi2        ADAPTER ERROR
BC669AA7   0814183208 P H dac1          CONTROLLER HEALTH CHECK FAILURE
3074FEB7   0814183208 T H fscsi2        ADAPTER ERROR
3074FEB7   0814183208 T H fscsi2        ADAPTER ERROR
3074FEB7   0814183208 T H fscsi2        ADAPTER ERROR
3074FEB7   0814183208 T H fscsi2        ADAPTER ERROR
B8113DD1   0814183008 T H fcs2          LINK ERROR
```

Detailed error log information from VIOS is available by using the **errlog -ls** command, as shown in Example 9-32.

Example 9-32 VIOS detailed error log entry

```
$ errlog -ls | more
---------------------------------------------------------------------------
LABEL:          FCP_ARRAY_ERR23
IDENTIFIER:     672766BB

Date/Time:       Wed Aug 20 17:47:43 CDT 2008
Sequence Number: 1263802
Machine Id:      00C5A6204C00
Node Id:         i6vios
Class:           H
Type:            PERM
Resource Name:   hdisk9
Resource Class:  disk
Resource Type:   array
Location:        U789D.001.DQDWVYP-P1-C3-T1-W201500A0B811F4C0-L7000000000000

Description
 "SNAPSHOT OPERATION NOT ALLOWED"

Probable Causes
 "SNAPSHOT VOLUME'S REPOSITORY FULL"

Failure Causes
"SNAPSHOT VOLUME HAS BEEN FAILED"

        Recommended Actions
        "DELETE OR RECREATE SNAPSHOT"

Detail Data
SENSE DATA
0600 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 1E40 0102 0000 7000 0400
0000 0098 0000 0000 8400 0000 0000 0000 0100 0000 0000 0000 0000 0000 0000 0000
0002 0900 0000 0000 0000 0000 0000 0000 0000 3154 3532 3536 3839 3032 2020 2020
2020 0660 0800 0007 0000 0D00 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000
00FF 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000
0000 0000 0599 10E9 3038 3230 3038 2F31 3534 3730 3200 0000 0000 0000 0000 0000
```

```
0000 0000 1A97 0000 F205 3704 0000 0200 0000 0000 0000 0000 0000 0000 0000 0000
0000 0000
```

The label information from the detailed error log like “FCP_ARRAY_ERR23” in Example 9-32 on page 394 can help with further problem isolation referring to the list of DS Storage System-reported “FCP_ARRAY_ERR*X*” / “SC_DISK_PCM_ERR*X*” error messages on VIOS when using either RDAC or MPIO (see Appendix A, “VIOS error log message for DS Midrange Storage Systems” on page 427).

VIOS storage resources in defined state

Devices showing up in defined state indicate that VIOS’s AIX kernel was not able to configure them the last time that VIOS was booted or **cfgdev** was run. If a Fibre Channel adapter (fcs*X*) or protocol device (fscsi*X*) shows up in defined state it may indicate a device driver or hardware problem, in which case we recommend following the problem isolation procedures in the *Fibre Channel Planning and Integration: User’s Guide and Service Information*, SC23-4329, available at:

http://www-1.ibm.com/support/docview.wss?uid=pub1sc23432903

A hdisk in defined state like hdisk8 with the corresponding parent fcs0 adapter being available, as shown in Example 9-33, indicates that the LUN is not accessible for any reason, for example, because it was erroneously removed from the host mapping on the storage side. After the LUN access problem has been corrected the device should be rediscovered by running **cfgdev** again.

Example 9-33 VIOS hdisk in defined state

```
$ lsdev -type disk -field name status physloc
name             status      physloc
hdisk0           Available   U789D.001.DQDWXNY-P1-T3-LFF0000-L0
hdisk1           Available   U789D.001.DQDWXNY-P1-T3-LFF0100-L0
hdisk2           Available   U789D.001.DQDWVYP-P1-C3-T1-W201500A0B811F4C0-L0
hdisk3           Available   U789D.001.DQDWVYP-P1-C3-T1-W201500A0B811F4C0-L1000000000000
hdisk4           Available   U789D.001.DQDWVYP-P1-C3-T1-W201500A0B811F4C0-L2000000000000
hdisk5           Available   U789D.001.DQDWVYP-P1-C3-T1-W201500A0B811F4C0-L3000000000000
hdisk6           Available   U789D.001.DQDWVYP-P1-C3-T1-W201500A0B811F4C0-L4000000000000
hdisk7           Available   U789D.001.DQDWVYP-P1-C3-T1-W201500A0B811F4C0-L5000000000000
hdisk8           Defined     U789D.001.DQDWVYP-P1-C3-T1-W201500A0B811F4C0-L6000000000000

$ lsdev -type adapter -field name status physloc | grep fcs
fcs0             Available   U789D.001.DQDWVYP-P1-C3-T1
fcs1             Available   U789D.001.DQDWVYP-P1-C3-T2
fcs2             Available   U789D.001.DQDWXNY-P1-C3-T1
fcs3             Available   U789D.001.DQDWXNY-P1-C3-T2

$ cfgdev -dev hdisk8

$ lsdev -dev hdisk8
name             status      description
hdisk8           Available   1815      DS4800 Disk Array Device
```

VIOS problem data collection

The basic problem data collection you should perform *before* contacting IBM support for further assistance with a VIOS-related problem is:

- Detailed problem description including date and time of occurrence
- VIOS fixpack level (See **`ioslevel`** command output.)
- Any hardware/software changes on the POWER6 system prior to the problem occurrence
- Storage I/O subsystem including its firmware level used
- **`snap`** command output stored by default in “/home/padmin/snap.pax.Z”
- **`lsmap -all`** and **`lsdev -virtual`** output

Once you have a problem management record (PMR) number from IBM support create a compressed TAR archive from the above-collected data and upload it to the IBM testcase FTP server as described in:

http://www-1.ibm.com/support/docview.wss?uid=isg3T1000214

9.4.3 Troubleshooting DS Storage

In this section we discuss Storage Manager integrated capabilities that can help in problem determination. We explain, based on examples, how to use the Recovery Guru, the Major Event Log (MEL), Read Link Status, and other diagnostic and recovery tools.

The presented figures in this section were taken from the Storage Manager 10.10 on a DS4000 power supply fan unit. Not all functions are available on the DS3000, but the Recovery Guru, Major Event Log, and collecting support data is very similar and the functions are easy to find in the Storage Manager GUI.

For more information about some topics addressed in this section, refer to the installation, users and maintenance guide for your specific DS model, and consult the *IBM System Storage DS4000 - Problem Determination Guide*, GC26-7703-01, available at:

http://www.ibm.com/support/docview.wss?rs=1157&context=HW27J&uid=psg1MIGR-57817

Storage Manager error reporting and diagnostics

The DS3000, DS4000, and DS5000 power supply fan unit and Storage Manager offer several functions to detect failures and log them for further problem determination, notify users of faults, and guide them through the necessary recovery steps. We cover some of the functions in the following section.

Preventive maintenance and support notifications

The DS3000, DS4000, and DS5000 have integrated functions to collect diagnostic information as they arise and notify you about the problems so that you can take corrective action without delay. Even if the product does not require specific maintenance, we recommend that the storage administrator consult the IBM Support Web site on a regular basis and keep the product up to date with the latest firmware and notices published.

The IBM Storage Support Web site is:

http://www.ibm.com/servers/storage/support/disk

A Support Notifications subscription service is available from the IBM System Storage product support Web site. This service is designed to keep you informed of new or updated System Storage support site information, such as publications, hints and tips, technical notes, product flashes (alerts), and downloads and drivers.

Support Notifications subscription can be done at:

http://www-304.ibm.com/jct01004c/systems/support/storage/news/05072007SupportNotif.html

> **Tip:** We recommend subscribing to the IBM System Storage product Web site to receive notifications directly to your e-mail address.

We recommend keeping your systems at the latest available (and compatible) firmware level. For instructions on how to update DS Storage to the latest firmware level, refer to 9.3.2, “Planning for upgrades” on page 360.

Recovery Guru

If there is an error condition on your storage server, the Recovery Guru will explain the cause of the problem and provide necessary actions to recover. It will guide you to perform specific actions, depending on the event, referring you also to the product documentation, or in particular situations will direct you to contact your IBM Service Representative.

Once the problem is solved, use the Recovery Guru to check that there is no longer a problem impacting the system. If this is the case, you should get the message `No failures were detected on the storage subsystem`. Note that a green icon displaying next to the Storage Subsystem in the SM GUI also indicates that the subsystem is clear of any errors.

Use the recover from failures icon in the Subsystem Management window to open the Recovery Guru. Figure 9-34 shows the recovery guru icon and the problem.

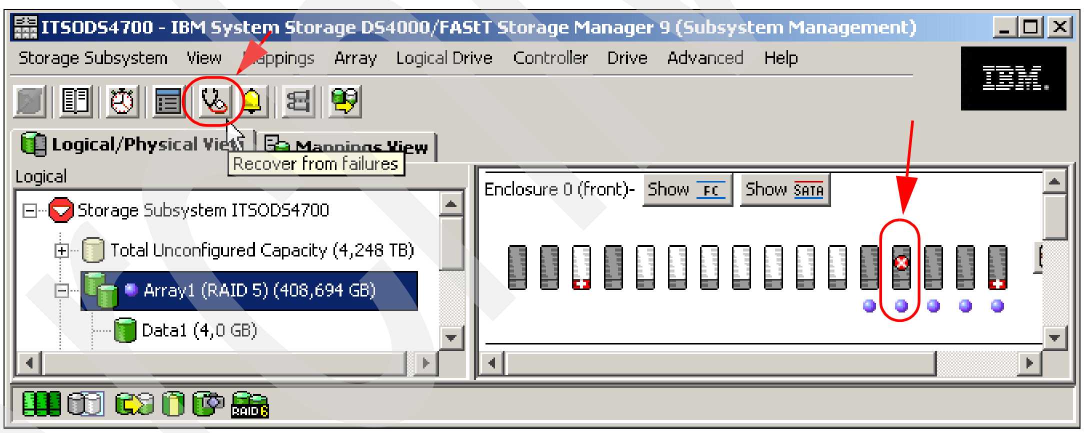


Figure 9-34 Recovery Guru Icon

Alternatively, select **Storage Subsystem** → **Recovery Guru** from the pull-down menu.

If no problems are detected, the guru returns a `No Problem Found` message. But in a case as shown in Figure 9-34 on page 397, where there is a failed drive, the Recovery Guru will display the information shown in Figure 9-35.

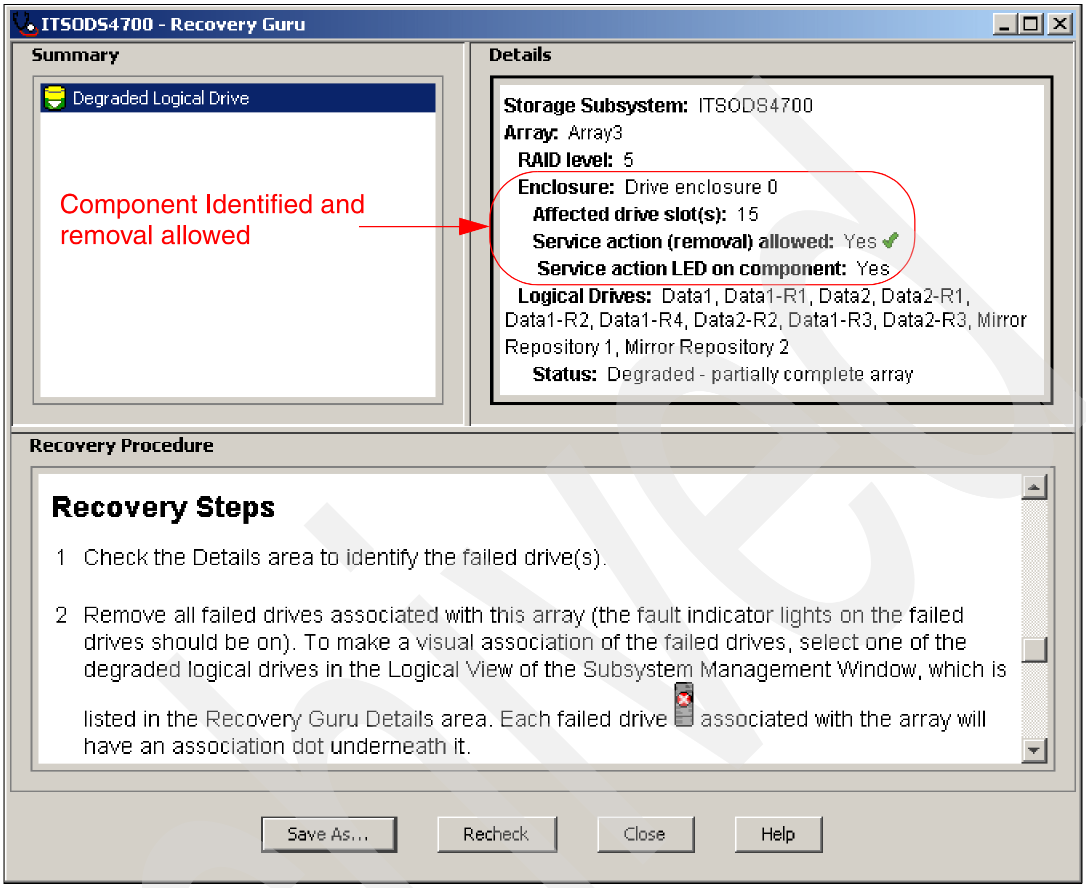


Figure 9-35 Recovery Guru actions

The details area tells you in this case that a component was identified in enclosure0, slot 5 and that it is safe to remove the component. In the unlikely event of multiple failed components, the *Service action (removal) allowed* indicator will take care about the correct order of the part replacements if the order is essential to keeping the system accessible (that is, a failed controller and a failed power supply on a DS4800 may require the power supply to be replaced first).

The recovery steps explain in detail the actions to take for each particular problem.

For another problem case scenario, we cover a common drive failure in "Failed drive example" on page 403.

Major event log

The major event log (MEL) is the primary source for troubleshooting a DS4000 power supply fan unit. To access the MEL select **Advanced** → **Troubleshooting** → **View Event Log**.

By default only the last 100 critical events will be shown, but you can choose how many events you want to have listed. The maximum number is 8191.

> **Important:** Remember to configure event monitoring and alerts to notify you of critical events as soon as they occur.

The MEL has two options:

- View only critical events (default option): For a quick overview of all events that might affect the operational status of your DS4000. This is an effective way of only viewing critical events rather than looking through the entire log file.
- View all events: Detailed information about all events logged by the controller. The list may be very long because most of the entries will be informational events about media scan tasks or battery charging cycles.

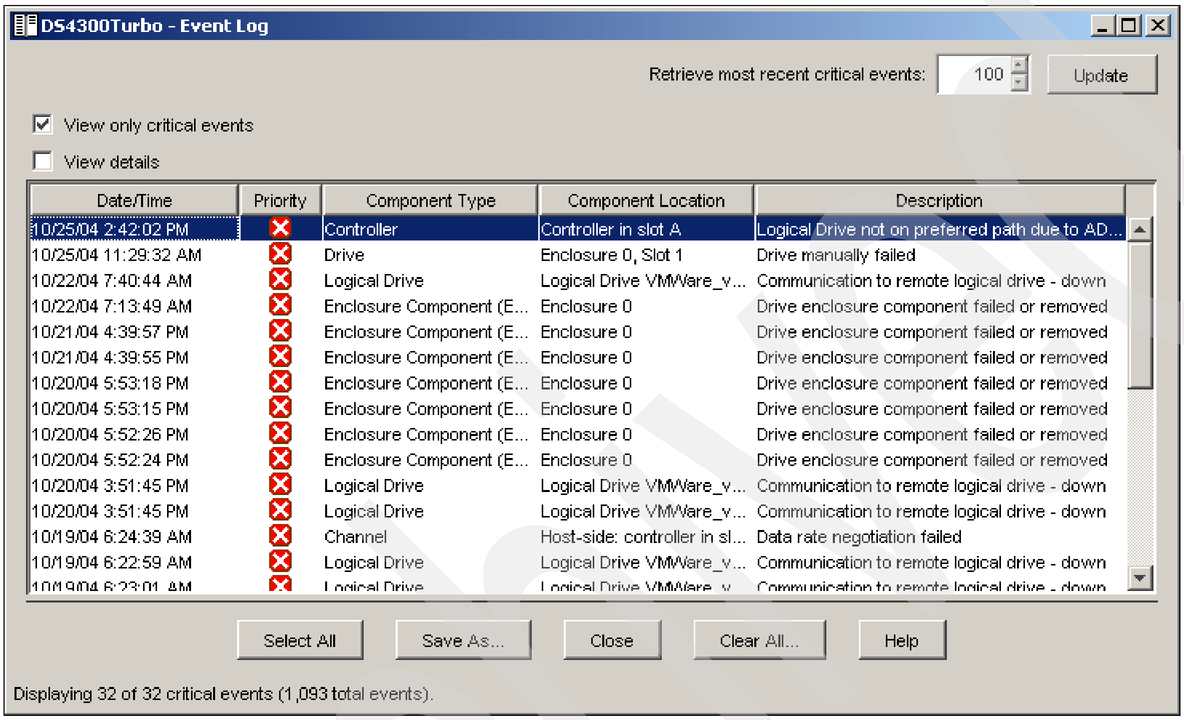


Figure 9-36 Major event log

If you want to troubleshoot your system, you should use the full event log, as it will include information about actions that took place before the actual critical event happened, thus giving you the complete history of the problem.

Another reason to check the MEL is preemptive in nature. If, for instance, you find many destination driver errors to a specific drive, you should check the read link status diagnostics and, if possible, the operating system error logs for problems such as time outs. An increased number of destination driver errors can indicate a faulty device. However, these logs are part of normal logging operation and should be examined in conjunction with other MEL entries.

Support data: Collecting information

In this section we describe the options available to collect information normally required by your support representative, for troubleshooting purposes.

Automatic support data collection

Storage Manager V10.10 offers the option to enable automatic support data collection. When enabled, a support data file is collected as soon as a critical event occurs and transferred to the directory specified. Doing this, the information relevant for troubleshooting by your support representative is stored in that directory. If a new event occurs, this will not overwrite the already collected file.

Tip: We recommend that you enable the automatic support data collection option, to have a support data file automatically generated and saved to the specified location after occurrence of a critical events. Make sure that:

- You specify a directory outside your DS4000 system to collect the information.
- The SM Monitor process has to run on the workstation or host where you want to collect the logs.

To enable automatic captures:

1. Select **Advanced** → **Troubleshooting** → **Support Data** → **Automatic Settings**.
2. Click the check mark to enable the automatic data collection and specify the destination folder, as shown in Figure 9-37.

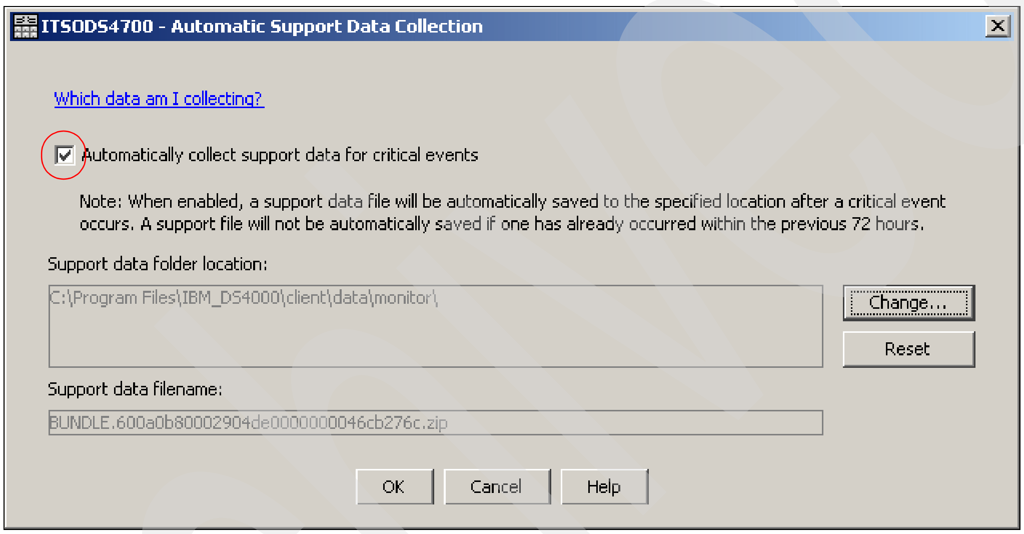


Figure 9-37 Enabling automatic data collection

Collect support data

In addition to automatically collect support data, you can manually generate a new collection at any time from the Storage Manager:

1. Select **Advanced** → **Troubleshooting** → **Support Data** → **Collect**.
2. The Collect All Support Data window (Figure 9-38) opens. Specify the name and location of the zip file that you want to create and then click **Start** to begin the collection.

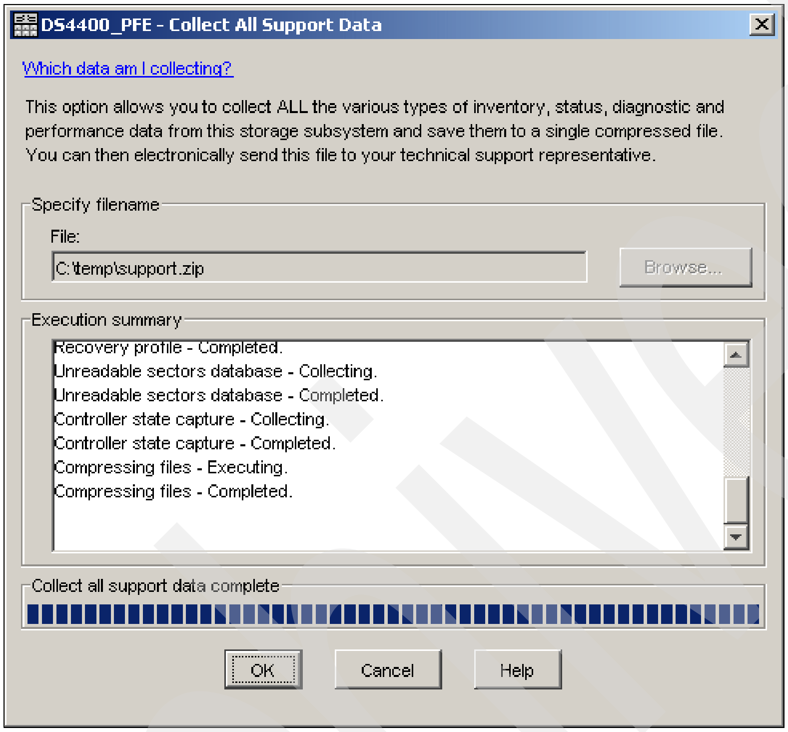


Figure 9-38   Collect All Support Data

The zip file that is produced consists of the following files:

- storageArrayProfile.txt: Equivalent to the View-Profile data option. Contains a description of all the components and properties of a Storage System.
- majorEventLog.txt: A detailed list of events that occur on the Storage System. The list is stored in reserved areas on the disks in the Storage System and records configuration events and Storage System component failures.
- readLinkStatus.csv: A detailed list of error counters accumulated over a period of time detected in the traffic flow between devices on the Fibre Channel loop. A file of historical read link status data might also be included with the file.
- NVSRAMdata.txt: A controller file that specifies the default settings for the controllers.
- recoveryGuruProcedures.html: A detailed list of all the warnings posted by executing the Recovery Guru in response to problems detected on the Storage System.
- performanceStatistics.csv: A detailed description of how your Storage System is performing. Data collected includes the I/O activity of specific controllers or logical drives, the transfer rate of the controller, the current I/O per second, and the maximum I/O per second.
- persistentReservations.txt: A detailed list of logical drives on the Storage System with persistent reservations and registrations.

- objectBundle: A detailed description of the status of your Storage System and its components, valid at the time the file was generated. Only since FW 7.10.
- driveDiagnosticData.txt: A detailed list of log sense data from all the drives in your Storage System.
- recoveryProfile.csv: A detailed description of the latest recovery profile record and historical data.
- unreadableSectors.txt: A detailed list of all the unreadable sectors that have been logged to the Storage Systems.
- stateCaptureData.dmp: A detailed description of the current state of your Storage System.
- storageArrayConfiguration.cfg: Backup file of the current DS4000 configuration.
- socStatistics.csv: A detailed list of error counters found in the traffic between controllers and switched FC devices. This list, used together with RLS diagnostics, helps to determine problems that may arise in the interconnection between the disks and controllers in your subsystem.
- featureBundle.txt: Details of all the premium features with installation status, current used capacities, and maximum supported.
- ExpansionTrayLog: Attached expansion event logs with details per ESM.
- Connections.txt: Detail of expansions connections.

Note: We recommend collecting support data before performing major changes on the Storage System (that is, firmware upgrades, drive migrations, adding new expansion units, and so on). In the unlikely event of a problem during the operation, all information about the last known good status is available for support.

Collect disk drive data

This option collects log sense data from all the disk drives on your DS4000 subsystem. Log sense data consists of statistical information maintained by each of the disk drives in your storage subsystem. Your support representative may request this information to analyze the performance of your disks and troubleshoot particular problems that might arise. Select **Advanced** → **Troubleshooting** → **Collect Drive Data**.

Important: Use this option only under the guidance of a customer and technical support representative.

Capture state information

The state capture option captures information about the current state of your DS4000 and saves the captured information to a text file. The captured information can then be sent to your support representative for analysis. Select **Advanced** → **Troubleshooting** → **Capture State Information**.

Diagnostic data capture (DDC)

For certain unusual events in the controller firmware, diagnostic data capture is available on the Storage System. The Recovery Guru will indicate whether diagnostic data was collected and will provide the necessary steps to download the DDC data to the management workstation. The Storage System will not return to an optimal state until the DDC data is downloaded.

Managing drive problems

Starting with Storage Manager V10.10 there is a new function, replace drives, that allows failed drives be replaced in the array configuration with an unassigned free drive or with the assigned hot spare.

Replace drive option

In a case of a failure of a physical disk, the array to which the disk belongs will become degraded (if it has RAID protection). If there is a hot spare available, it will take over the failed drive and start a reconstruction process. Once the reconstruction is finished, the array will become optimal.

With the new replace drive option provided in Storage Manager V10.10 you can redefine the array marking the assigned hot spare to be a permanent part of the array. This eliminates the copyback function previously needed to perform a second reconstruction on the new drive when it is inserted.

The new replacement drive functionality also allows you to not wait for your service provider to send a new drive. If you do not have a hot spare drive configured, but have free drives, you can select an unassigned drive as a replacement for the failed drive without swapping parts. The array will start reconstructing with the replacement drive now being part of the array. The failed drive becomes a failed unassiged drive.

The drive used to replace a failed drive should meet the following conditions:

- Hot spare assigned to cover the array exposure
- Unassigned optimal drive
  - Same technology (FC/SATA)
  - Adequate capacity

Failed drive example

Assume that we have a failed disk drive in a RAID-protected array and hot spare, as shown in Figure 9-39.

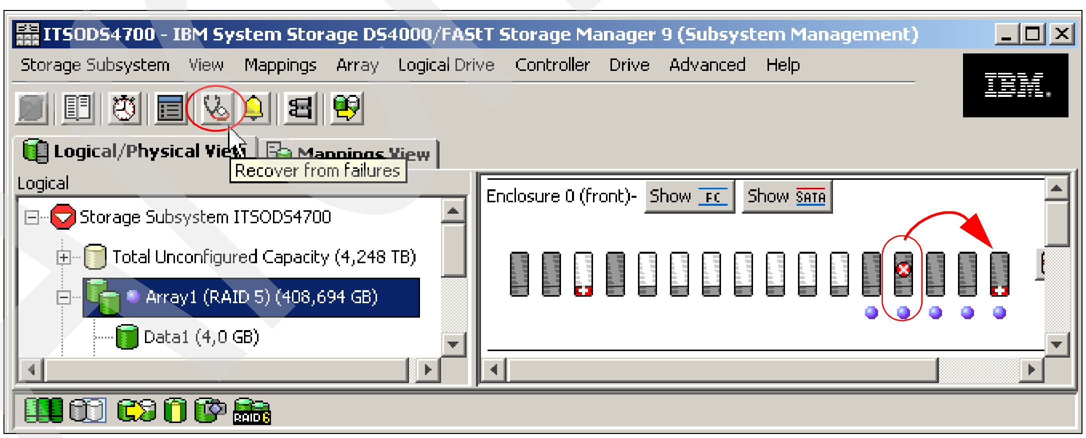


Figure 9-39 Drive problem

Click the recover from failures icon and the Recovery Guru displays a description of the problem, with additional information, as shown in Figure 9-40.

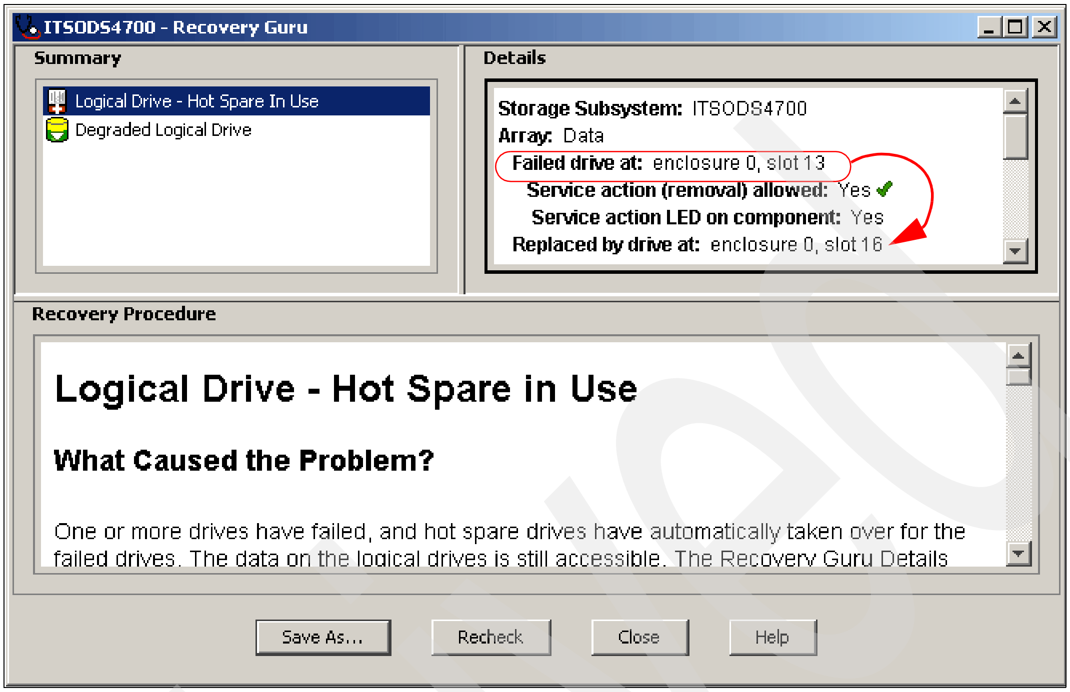


Figure 9-40 Recovery Guru

In this example, the failed drive in slot 13 of enclosure 0 was replaced by the hot spare configured in slot 16. The logical drives in the affected array will reconstruct to the hot spare. Once finished, all the logical drives will be shown as optimal instead of reconstructing.

Starting with SM 10.10 (firmware V7.10), you can now mark the assigned spare drive as part of the array to avoid the copyback function from the array to the replacement drive. Also, you can replace it by a free drive if you have one.

If the failed drive is replaced, instead of starting a copyback from the drive in slot 16 to the new drive in slot 13, it will be presented as unassigned-optimal. You can choose what to do, either preserving the original slots drives/assignments or avoiding the copyback to the original drive and its potential impact to the subsystem performance.

Perform the following steps:

1. Right-click the failed drive or select the drive and then select **Drive** → **Replace Drives**.

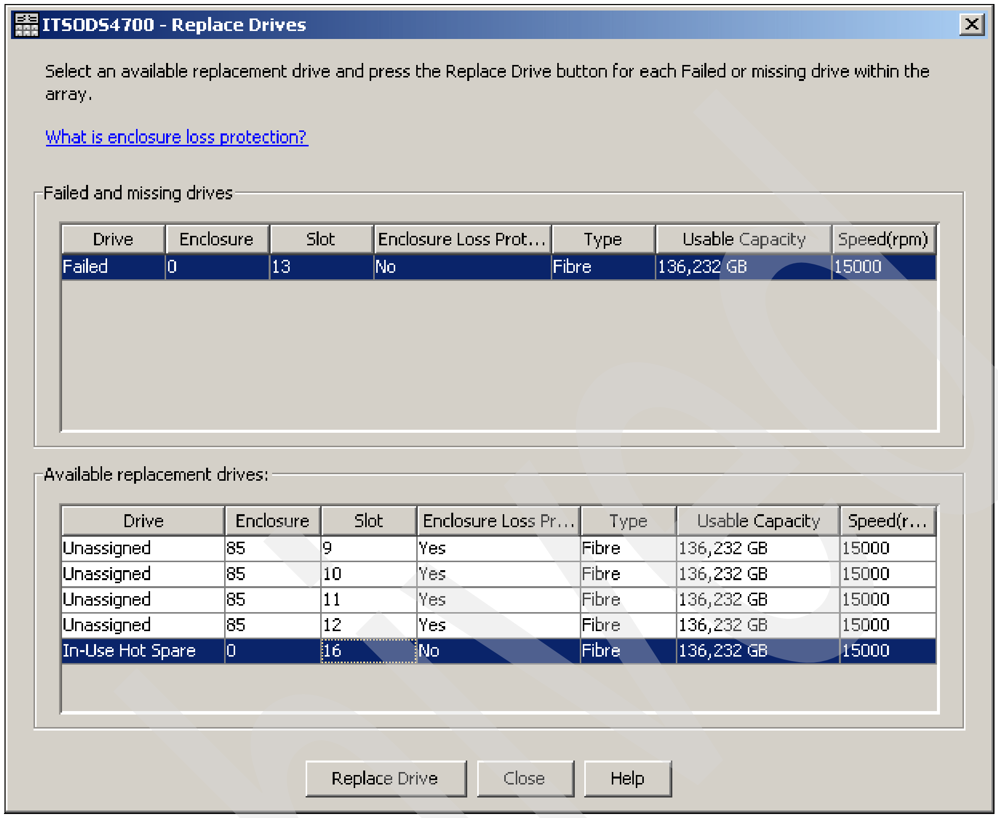


Figure 9-41 Replace Drives option

2. Select the available drive for replacement and then click **Replace Drive**. Remember that a replacement drive should be:
   – Any unassigned drive (same type and capacity)
   – Any hot spare in use

   Note: If the second option is selected, the hot spare covering the failed drive will be marked as a permanent member of the array, replacing the failed drive, thus avoiding the copyback function.

3. The failed drive will change to unassigned-failed and no longer be part of the array.
4. If the problem requests the drive to be replaced, remember after physically replacing the drive to configure the new one as hot spare (to ensure that you still have adequate hot spare coverage).

Important: If you have multiple failed drives you should open a call with IBM Support, as it is very likely that the error is not related to the drives but rather to a problem on the FC loop on the drive side.

Missing logical drives

A missing logical drive is a placeholder node displayed in the logical view (Figure 9-42). It indicates that the Storage System has detected inaccessible drives associated with a logical drive. Typically, this is the result of removing drives associated with an array, or when a power loss to one or more drive enclosures has occurred.

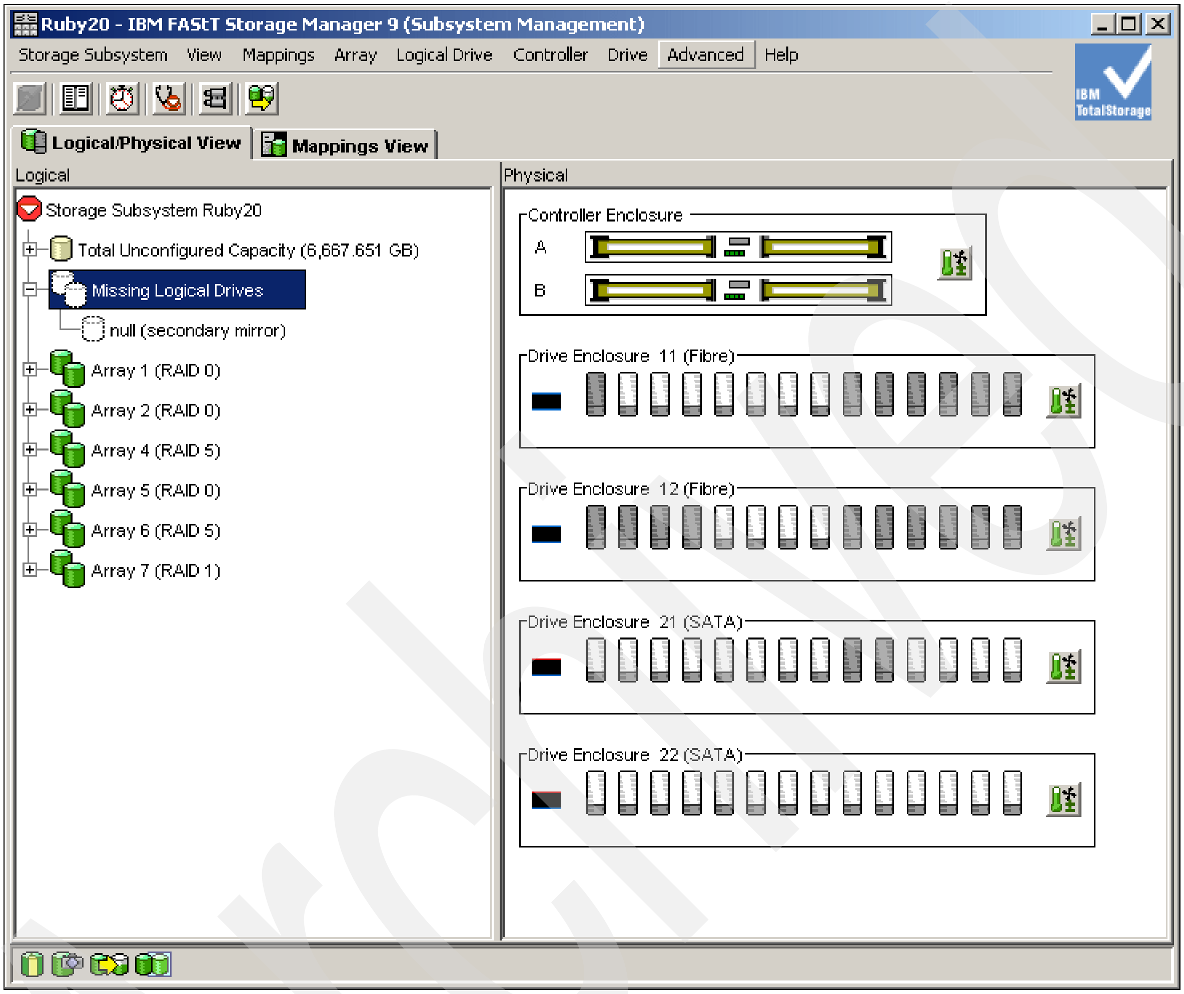


Figure 9-42 Missing logical drives

Missing logical drives are only displayed in the logical view if they are standard logical drives or repository logical drives. In addition, one of the following conditions must exist:

- The logical drive has an existing *logical drive-to-LUN mapping*, and drives associated with the logical drive are no longer accessible.
- The logical drive is participating in a *remote volume mirror* as either a primary logical drive or secondary logical drive, and drives associated with the logical drive are no longer accessible.
- The logical drive is a *mirror repository logical drive*, and drives associated with the logical drive are no longer accessible. The Recovery Guru has a special recovery procedure for this case. Two mirror repository logical drives are created together on the same array when the Remote Volume Mirroring premium feature is activated and one is used for each controller in the Storage System. If drives associated with the array are no longer accessible, then both mirror repository logical drives are missing, and all remote volume mirrors are in an unsynchronized state.

- The logical drive is a *base logical drive* with associated FlashCopy logical drives, and drives associated with the logical drive are no longer accessible.
- The logical drive is a *FlashCopy repository logical drive*, and drives associated with the logical drive are no longer accessible.

If missing logical drives are detected by the Storage System, a missing logical drives group is created in the logical view of the Subsystem Management window. Each missing logical drive is shown and identified by its world-wide name and logical drive type. Missing logical drives are identified as being either a standard logical drive, base logical drive, FlashCopy repository logical drive, primary logical drive, secondary logical drive, or mirror repository logical drive.

> **Note:** Missing logical drives, in most cases, are recoverable. Do not delete missing logical drives without confirming that the logical drives are no longer required, because they will be permanently removed from the configuration.

If missing logical drives are detected because drives are accidentally removed, or are detected as missing due to a loss of power to the drive enclosures, recovery of these logical drives is possible by:

- Re-inserting the drives back into the drive enclosure
- Ensuring that the drive enclosure's power supplies are properly connected to an operating power source and have an optimal status

Diagnostics options

We describe here the available diagnostic options integrated in the Storage Manager client interface, with particular emphasis on the read link status diagnostics.

Controller diagnostics

The controller diagnostics feature allows you to verify that a controller is functioning properly by performing various internal tests.

The diagnostics use a combination of three different tests: Read test, write test, and data loopback test. You should run all three tests when suspecting a controller problem or under guidance of your support representative.

> **Important:** During the diagnostics, the controller on which the tests are run will *not* be available for I/O. When diagnostics are completed, the controller should automatically allow data to be transferred to it. However, if there is a situation where data transfer is not re-enabled, select the controller and use the **Data Transfer** → **Enable** option.

To run the diagnostics, highlight a controller from the Subsystem Management window. Then select the **Advanced** → **Troubleshooting** → **Run Diagnostics** → **Controller**. The Run Diagnostics dialog (shown in Figure 9-43) is displayed.

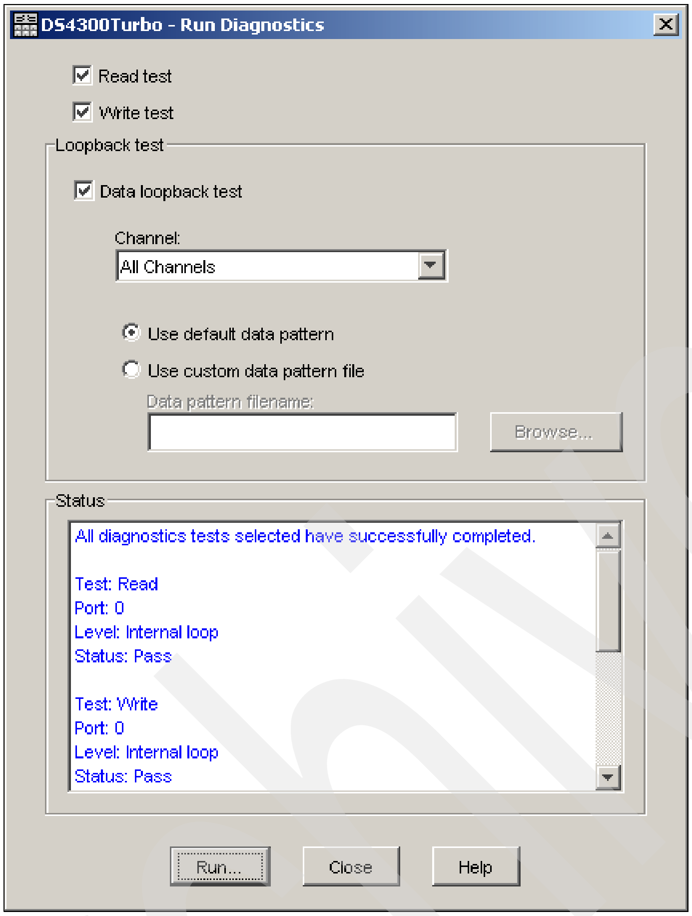


Figure 9-43 Controller diagnostics

The Run Diagnostics window includes the following check boxes:

- Read test: Initiates a read command as it would be sent over an I/O data path. It compares data with a known, specific data pattern, checking for data integrity and redundancy errors. If the read command is unsuccessful or the data compared is not correct, the controller is considered to be in error and is failed.
- Write test: Initiates a write command as it would be sent over an I/O data path (to the diagnostics region on a specified drive). This diagnostics region is then read and compared to a specific data pattern. If the write fails or the data compared is not correct, the controller is considered to be in error and is failed and placed offline. In this case, you would then use the Recovery Guru to replace the controller.
- Data loopback test: This test can be run only on controllers that have Fibre Channel connections between the controller and the drives. The test passes data through each controller's drive-side channel, mini-hub, out onto the loop, and then back again. Enough data is transferred to determine error conditions on the channel. If the test fails on any channel, then this status is saved so that it can be returned if all other tests pass.

All test results are displayed in the diagnostics dialog status area.

Events are written to the event log when diagnostics are started and when the tests have completed. These events will help you evaluate whether diagnostics testing was successful or failed, and the reason for the failure. To view the event log, click **Advanced** → **Troubleshooting** → **View Event Log** in the Subsystem Management window.

Read Link Status diagnostics

During communication between devices, read link status (RLS) error counts are detected within the traffic flow of the loop. Error count information is accumulated over a period of time for every component and device including:

- Drives
- ESMs
- Fibre Channel ports

Error counts are calculated from a baseline, which describes the error count values for each type of device in the Fibre Channel loop. Calculation occurs from the time when the baseline was established to the time at which the error count information is requested. The baseline is automatically set by the controller. However, a new baseline can be set manually through the Read Link Status Diagnostics dialog.

Read link status error counts refer to link errors that have been detected in the traffic flow of a Fibre Channel loop. The errors detected are represented as a count (32-bit field) of error occurrences accumulated over time and help to provide a coarse measure of the integrity of the components and devices on the loop.

By analyzing the error counts retrieved, it is possible to determine the components or devices within the Fibre Channel loop that might be experiencing problems communicating with the other devices on the loop. A high error count for a particular component or device indicates that it might be experiencing problems and should be given immediate attention.

Analyzing read link status results

Use the RLS diagnostics data together with the file socStatistics.csv provided in the support data zip file to debug problems on the drive side of your storage server.

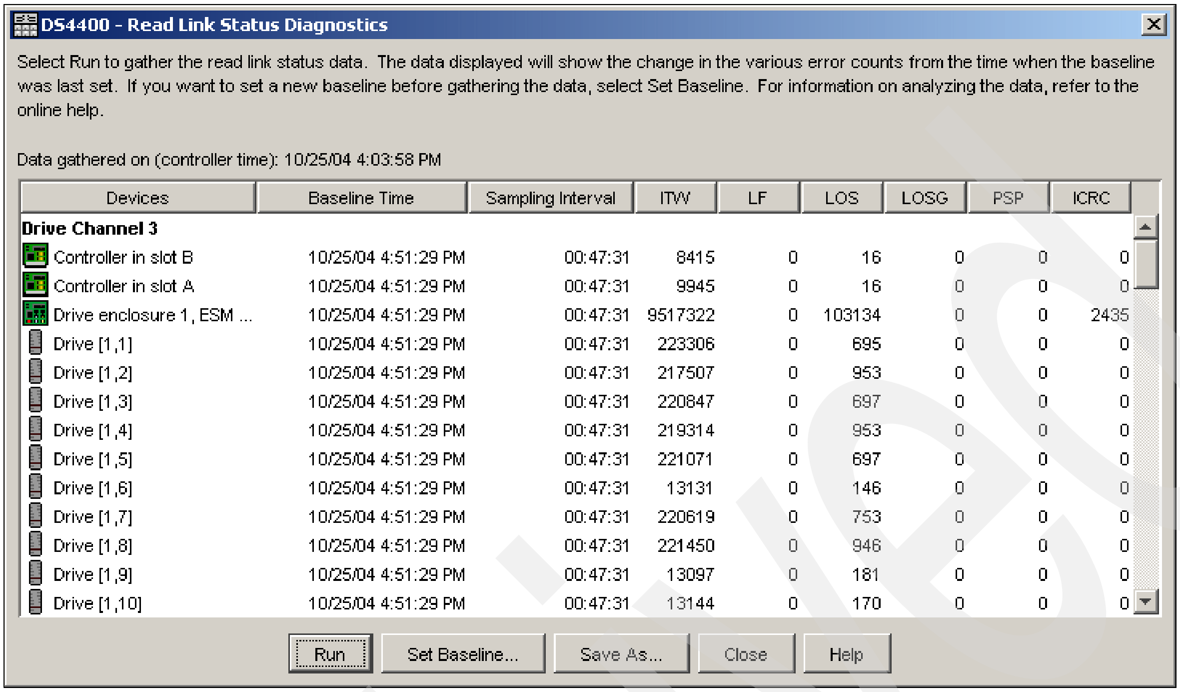
Analysis of the RLS error count data is based on the principle that the device immediately *downstream* of the problematic component should see the largest number of Invalid Transmission Word (ITW) error counts.

> **Important:** Because the current error counting standard is vague about when the ITW count is calculated, different vendor devices calculate errors at different rates. Take this into account in your analysis of the data. ITW errors can occur during normal operations but should not exceed 3-4 counts in 24 hours.

Usually, RLS diagnostics should be run over a period of 24 hours with workload applied to the controller. Otherwise, with no I/O it is very unlikely that RLS will record anything.

From the Storage System menu, right-click **Advanced** and select **Troubleshooting** → **Run Diagnostics** → **Read Link Status**.

Figure 9-44 shows the RLS Diagnostics window. It lists all devices for each redundant loop. The channel numbers correspond with the drive-side mini-hubs.



Devices	Baseline Time	Sampling Interval	ITW	LF	LOS	LOSG	PSP	ICRC
Drive Channel 3								
Controller in slot B	10/25/04 4:51:29 PM	00:47:31	8415	0	16	0	0	0
Controller in slot A	10/25/04 4:51:29 PM	00:47:31	9945	0	16	0	0	0
Drive enclosure 1, ESM ...	10/25/04 4:51:29 PM	00:47:31	9517322	0	103134	0	0	2435
Drive [1,1]	10/25/04 4:51:29 PM	00:47:31	223306	0	695	0	0	0
Drive [1,2]	10/25/04 4:51:29 PM	00:47:31	217507	0	953	0	0	0
Drive [1,3]	10/25/04 4:51:29 PM	00:47:31	220847	0	697	0	0	0
Drive [1,4]	10/25/04 4:51:29 PM	00:47:31	219314	0	953	0	0	0
Drive [1,5]	10/25/04 4:51:29 PM	00:47:31	221071	0	697	0	0	0
Drive [1,6]	10/25/04 4:51:29 PM	00:47:31	13131	0	146	0	0	0
Drive [1,7]	10/25/04 4:51:29 PM	00:47:31	220619	0	753	0	0	0
Drive [1,8]	10/25/04 4:51:29 PM	00:47:31	221450	0	946	0	0	0
Drive [1,9]	10/25/04 4:51:29 PM	00:47:31	13097	0	181	0	0	0
Drive [1,10]	10/25/04 4:51:29 PM	00:47:31	13144	0	170	0	0	0

Figure 9-44 Storage Manager 10.10 Read Link Status Diagnostics

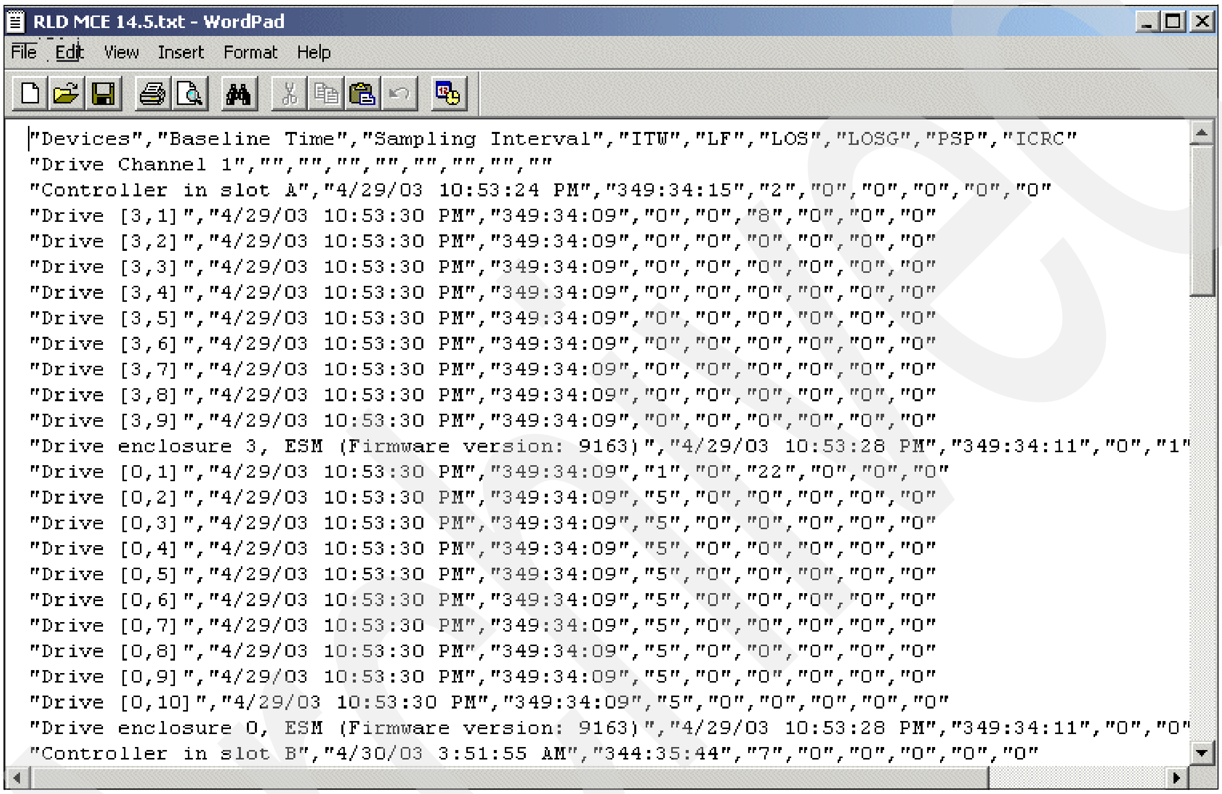
Each of the columns in the window is explained in the following list:

- Devices: A list of all the devices on the Fibre Channel loop. The devices are displayed in channel order. Within each channel, they are sorted according to the device's position within the loop.
- Baseline Time: The date and time of when the baseline was last set.
- Elapsed Time: The elapsed time between when the baseline time was set and when the read link status data was gathered using the run option.
- ITW: The total number of invalid transmission word errors detected on the Fibre Channel loop from the baseline time to the date and time. ITW can also be referred to as the *received bad character count*.
- Link Failure (LF): The total number of link failure errors detected on the Fibre Channel loop from the baseline time to the current date and time. When detected, link failures indicate that there has been a failure within the media module laser operation. Link failures might also be caused by a link fault signal, a loss of signal, or a loss of synchronization.
- Loss of Synchronization (LOS): The total number of loss of synchronization errors detected on the Fibre Channel loop from the baseline time to the current date and time. This indicates that the receiver cannot acquire symbol lock with the incoming data stream due to a degraded input signal. If this condition persists, the number of loss of signal errors increases.
- Loss of Signal (LOSG): The total number of loss of signal errors detected on the Fibre Channel loop from the baseline date to the current date and time. This indicates a loss of signal from the transmitting node or physical component within the Fibre Channel loop. Errors logged against a single drive usually point to a drive failure. If you see an increase in numbers against multiple drives across one channel the problem probably lies with cabling or connection (including mini-hubs or SFPs).

Errors logged against the ESM board can point to bad ESM boards or bad connectivity (cables, SFT, mini-hubs).

If you experience an increase on the controller blades there might be a problem at the midplane or the controller itself (Figure 9-44 on page 410).

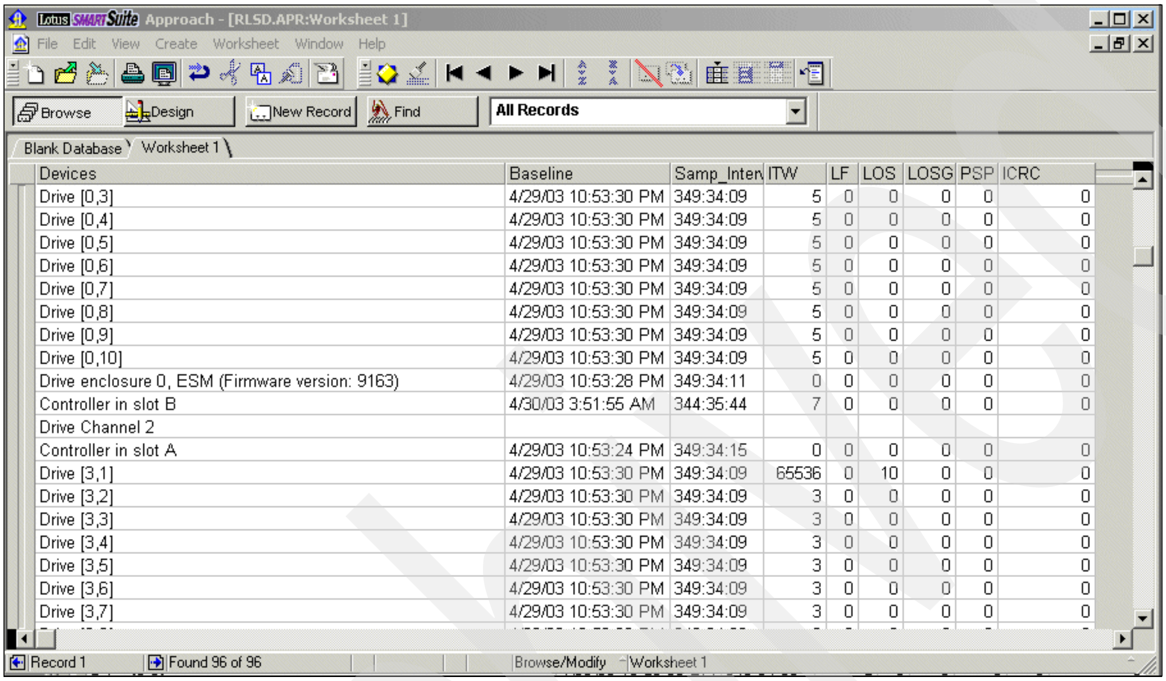
In our example in Figure 9-44 on page 410 the baseline was set to October 25. When you open the RLSD window, the most recent data is displayed. If you click **Run** the window will be updated (refreshed). In order to reset the counters, click **Set Baseline**. If you need to submit diagnostic data to the support staff, save the file as a .csv file. Figure 9-45 shows an example of an RLSD file in .csv format.




```
"Devices","Baseline Time","Sampling Interval","ITW","LF","LOS","LOSG","PSP","ICRC"
"Drive Channel 1","","","","","","","",""
"Controller in slot A","4/29/03 10:53:24 PM","349:34:15","2","0","0","0","0","0"
"Drive [3,1]","4/29/03 10:53:30 PM","349:34:09","0","0","8","0","0","0"
"Drive [3,2]","4/29/03 10:53:30 PM","349:34:09","0","0","0","0","0","0"
"Drive [3,3]","4/29/03 10:53:30 PM","349:34:09","0","0","0","0","0","0"
"Drive [3,4]","4/29/03 10:53:30 PM","349:34:09","0","0","0","0","0","0"
"Drive [3,5]","4/29/03 10:53:30 PM","349:34:09","0","0","0","0","0","0"
"Drive [3,6]","4/29/03 10:53:30 PM","349:34:09","0","0","0","0","0","0"
"Drive [3,7]","4/29/03 10:53:30 PM","349:34:09","0","0","0","0","0","0"
"Drive [3,8]","4/29/03 10:53:30 PM","349:34:09","0","0","0","0","0","0"
"Drive [3,9]","4/29/03 10:53:30 PM","349:34:09","0","0","0","0","0","0"
"Drive enclosure 3, ESM (Firmware version: 9163)","4/29/03 10:53:28 PM","349:34:11","0","1"
"Drive [0,1]","4/29/03 10:53:30 PM","349:34:09","1","0","22","0","0","0"
"Drive [0,2]","4/29/03 10:53:30 PM","349:34:09","5","0","0","0","0","0"
"Drive [0,3]","4/29/03 10:53:30 PM","349:34:09","5","0","0","0","0","0"
"Drive [0,4]","4/29/03 10:53:30 PM","349:34:09","5","0","0","0","0","0"
"Drive [0,5]","4/29/03 10:53:30 PM","349:34:09","5","0","0","0","0","0"
"Drive [0,6]","4/29/03 10:53:30 PM","349:34:09","5","0","0","0","0","0"
"Drive [0,7]","4/29/03 10:53:30 PM","349:34:09","5","0","0","0","0","0"
"Drive [0,8]","4/29/03 10:53:30 PM","349:34:09","5","0","0","0","0","0"
"Drive [0,9]","4/29/03 10:53:30 PM","349:34:09","5","0","0","0","0","0"
"Drive [0,10]","4/29/03 10:53:30 PM","349:34:09","5","0","0","0","0","0"
"Drive enclosure 0, ESM (Firmware version: 9163)","4/29/03 10:53:28 PM","349:34:11","0","0"
"Controller in slot B","4/30/03 3:51:55 AM","344:35:44","7","0","0","0","0","0"
```

Figure 9-45 Sample .csv file format

You can use the .csv format to import the data in a database for better readability. As an example, Figure 9-46 shows the display, under Lotus® Approach, of error information collected for a bad drive that generated errors across the drive loop. The event log shows that many destination drive errors were logged. Not shown in Figure 9-46 is that drive channel 1 was virtually error free. Since the drives have two ports, we suspect that the port connected to drive channel 2 has a problem. In a fiber environment it is likely that the device that is downstream (AL-PA ID) from the error logging device is the faulty component. In this case, however, the MEL pointed to drive 3,1 (which was ultimately replaced).



Devices	Baseline	Samp_Inten	ITW	LF	LOS	LOSG	PSP	ICRC
Drive [0,3]	4/29/03 10:53:30 PM	349:34:09	5	0	0	0	0	0
Drive [0,4]	4/29/03 10:53:30 PM	349:34:09	5	0	0	0	0	0
Drive [0,5]	4/29/03 10:53:30 PM	349:34:09	5	0	0	0	0	0
Drive [0,6]	4/29/03 10:53:30 PM	349:34:09	5	0	0	0	0	0
Drive [0,7]	4/29/03 10:53:30 PM	349:34:09	5	0	0	0	0	0
Drive [0,8]	4/29/03 10:53:30 PM	349:34:09	5	0	0	0	0	0
Drive [0,9]	4/29/03 10:53:30 PM	349:34:09	5	0	0	0	0	0
Drive [0,10]	4/29/03 10:53:30 PM	349:34:09	5	0	0	0	0	0
Drive enclosure 0, ESM (Firmware version: 9163)	4/29/03 10:53:28 PM	349:34:11	0	0	0	0	0	0
Controller in slot B	4/30/03 3:51:55 AM	344:35:44	7	0	0	0	0	0
Drive Channel 2								
Controller in slot A	4/29/03 10:53:24 PM	349:34:15	0	0	0	0	0	0
Drive [3,1]	4/29/03 10:53:30 PM	349:34:09	65536	0	10	0	0	0
Drive [3,2]	4/29/03 10:53:30 PM	349:34:09	3	0	0	0	0	0
Drive [3,3]	4/29/03 10:53:30 PM	349:34:09	3	0	0	0	0	0
Drive [3,4]	4/29/03 10:53:30 PM	349:34:09	3	0	0	0	0	0
Drive [3,5]	4/29/03 10:53:30 PM	349:34:09	3	0	0	0	0	0
Drive [3,6]	4/29/03 10:53:30 PM	349:34:09	3	0	0	0	0	0
Drive [3,7]	4/29/03 10:53:30 PM	349:34:09	3	0	0	0	0	0

Figure 9-46 RLSD Approach database

As a general guidance, you can use Table 9-3 to find the most likely defective part on the DS4000 drive loop.

Table 9-3 RLS diagnostics table

Location of device A (first device that detects a link error)	**Location of device B** (first device upstream of device A)	**Possible candidates**
Enclosure X	Enclosure X	Device B (drive), ESM of dnclosure X, device A (drive)
Enclosure X	Enclosure Y	Cable or SFPs between enclosure X and Y, ESM of enclosure Y, device B (drive), ESM of enclosure X, device A (drive)
Enclosure X	Controller module	Cable or SFPs between enclosure X and controller module, mini-hub, device B (controller), ESM of enclosure X, device A (drive), controller chassis
Controller module	Enclosure Y	Mini-hub, any cable or SFP in the channel, ESM of enclosure X, device B (drive), device A (controller), controller chassis
Controller module	Controller module	Mini-hub, device B (controller), device A (controller)

In Example 9-34 you can see that device A (th efirst device with an error detected) corresponds to drive 5 in enclosure 1. The upstream device is drive 4 in enclosure 1. So this would be the enclosure X/enclosure X type of problem. Candidates for replacement would be drive 1,4 or the ESM in enclosure 1. It is also possible that drive 1,5 is the reason for the problem, but it is fairly unlikely.

Example 9-34 RLSD output usage example

```
Devices        Baseline Time Sampling Interval ITW    LF     LOS    LOSG   PSP    ICRC
Drive Channel 3
....
Drive [1,2]    10/25/2004 16:51       0:16:02 0      0      0      0      0      0
Drive [1,3]    10/25/2004 16:51       0:16:02 0      0      0      0      0      0
Drive [1,4]    10/25/2004 16:51       0:16:02 0      0      0      0      0      0
Drive [1,5]    10/25/2004 16:51       0:16:02 425    0      0      0      0      0
Drive [1,6]    10/25/2004 16:51       0:16:02 425    0      0      0      0      0
Drive [1,7]    10/25/2004 16:51       0:16:02 422    0      0      0      0      0
Drive [1,8]    10/25/2004 16:51       0:16:02 418    0      0      0      0      0
Drive [1,9]    10/25/2004 16:51       0:16:02 418    0      0      0      0      0
Drive [1,10]   10/25/2004 16:51       0:16:02 418    0      0      0      0      0
....
```

Discrete lines diagnostics

The discrete lines are dedicated control and status lines between the controllers that reside in the interconnect-battery canister. The Recovery Guru will guide you when you have to select this option as part of the recovery steps, although it is a tool mainly to be used by your support representative.

This option is to execute diagnostics to the controllers to further identify a possible failure when there is not enough information from the error logged on which controller has failed, or if the problem was with the interconnect.

In the above case, the Recovery Guru guides you to place a specific controller offline and then execute this diagnostics by selecting **Advanced** → **Troubleshooting** → **Run Diagnostics** → **Discrete Lines**.

Degraded drive channels

We saw in “Read Link Status diagnostics” on page 409 that the RLS output is a good way to find problems on the FC loop of the DS4000.

If an error condition is detected by the controller in one of the drive channels, it will change its status from optimal to degraded, routing the traffic to the remaining optimal redundant channel path. If you are using Storage Manager with an older firmware, make sure to update your DS4000 to the latest supported level, since under some particular situation, older firmwares might keep retrying to recover the channel, creating an impact on the I/O traffic and a possible access problem. After firmware V6.10, a user intervention is required to set the channel drive back to optimal.

A drive channel is set to degraded when a certain threshold of errors in a certain period of time is exceeded. These errors include time-out errors, controller-detected errors, drive-detected errors, and link-down errors. The I/O is stopped on the degraded channel and only the optimal channel is used for I/O. A controller always selects an optimal channel over a degraded one. However, if both channels are degraded, the controller will arbitrarily choose one of the two.

To check the drive channels, select **Advanced** → **Troubleshooting** → **Drive Channels**, which displays the Drive Channels Summary window (Figure 9-47).

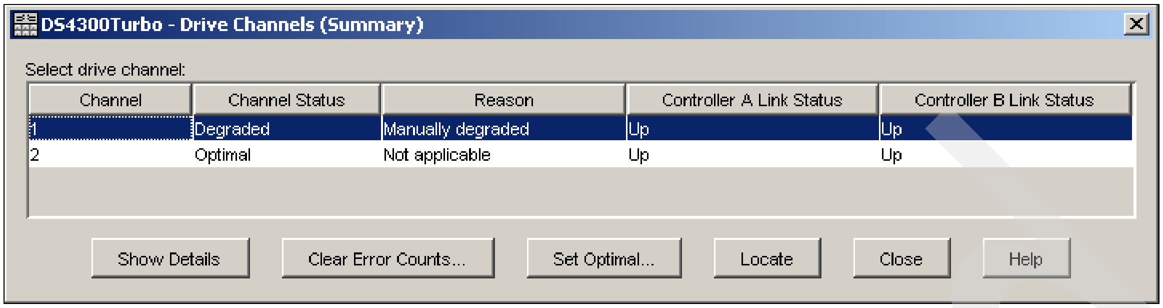


Figure 9-47 Drive Channel Summary

The link status is fully controlled by the controller. There are two states for the link status:

- Up: The link is currently able to handle I/Os.
- Down: The link is currently not able to handle I/Os.

The link status cannot be manually changed. The controller will set the link status to down if the controller cannot do a loop initialization within a certain amount of time. When the link is down, the controller will continuously try to bring the link up.

Consequently, this means that even if a channel is degraded and the link is up, there is the possibility that I/O will occur on this link. However, the controller will first choose the channel that is not in a degraded state.

To get a more detailed view of the drive channel status, click **Show Details**. This will show the error counters on this specific channel during the sample period (Figure 9-48).

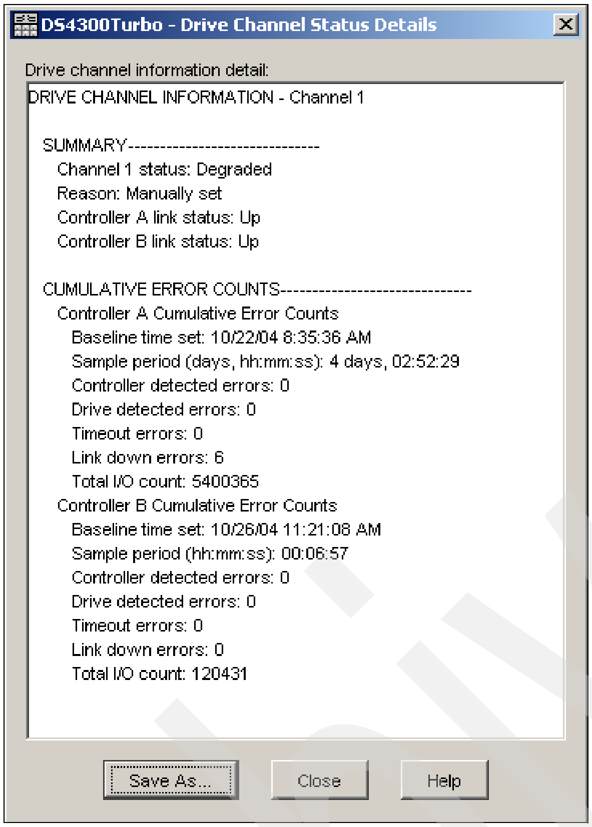


Figure 9-48 Drive Channel Status Details

See also “Expansion unit connections” on page 415 for an additional tool for problem determination on the drive channels side.

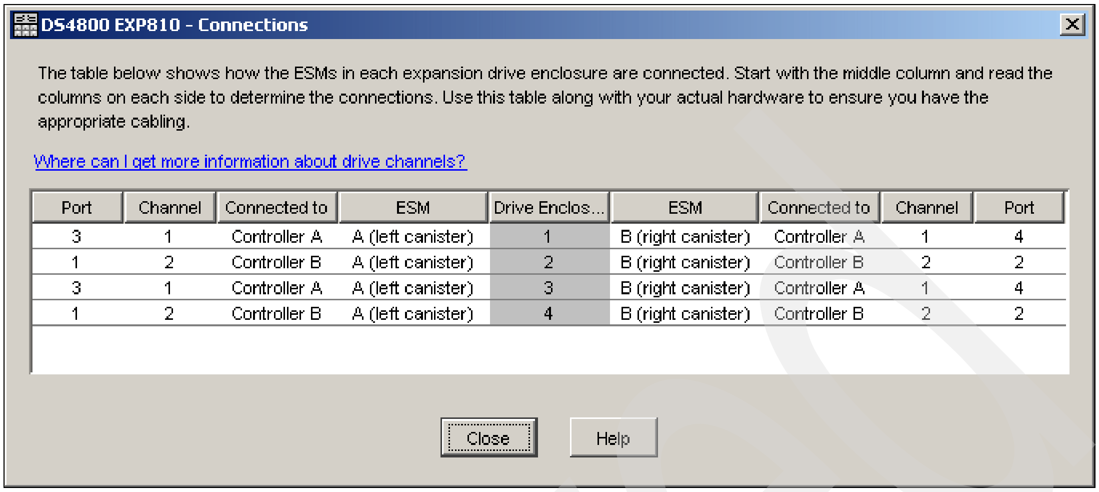
Managing ESMs

We describe here some additional Storage Manager options available for ESM problem determination and recovery.

Expansion unit connections

To debug problems in the drive channel side of the cabling, Storage Manager incorporates an option to detail out how the interconnection was made.

Select **Storage Subsystems** → **View** → **Connection** and you will get a table describing your expansion enclosures cabling, as shown in Figure 9-49.


DS4800 EXP810 - Connections

The table below shows how the ESMs in each expansion drive enclosure are connected. Start with the middle column and read the columns on each side to determine the connections. Use this table along with your actual hardware to ensure you have the appropriate cabling.

Where can I get more information about drive channels?

Port	Channel	Connected to	ESM	Drive Enclos...	ESM	Connected to	Channel	Port
3	1	Controller A	A (left canister)	1	B (right canister)	Controller A	1	4
1	2	Controller B	A (left canister)	2	B (right canister)	Controller B	2	2
3	1	Controller A	A (left canister)	3	B (right canister)	Controller A	1	4
1	2	Controller B	A (left canister)	4	B (right canister)	Controller B	2	2

Close Help

Figure 9-49 EXP connections

ESM replacement

Whenever an ESM is replaced in an EXP810 or EXP710, a firmware synchronization feature ensures that the firmware on the new ESM is automatically synchronized with the firmware in the existing ESM. This eliminates any ESM firmware mismatch conditions automatically. This is a feature implemented with firmware V6.16 for EXP710 and EXP810. You still have the option to update ESM firmware manually to a new level or in the case of a failure that would not let the code synchronization identify which is the original ESM and which is the replacement to synchronize.

The ESMs also hols configuration settings as NVSRAM, which are automatically set. If, for a a particular problematic situation, one of the ESMs on an enclosure has a different configuration than the other and could not resolve the synchronization automatically, a *needs attention* condition would be displayed. For EXP420, EXP710, and EXP810, there is an option to set the enclosure settings by selecting **Maintenance** → **Download** → **ESM Configuration Settings** in the Storage Manager Subsystem Management window.

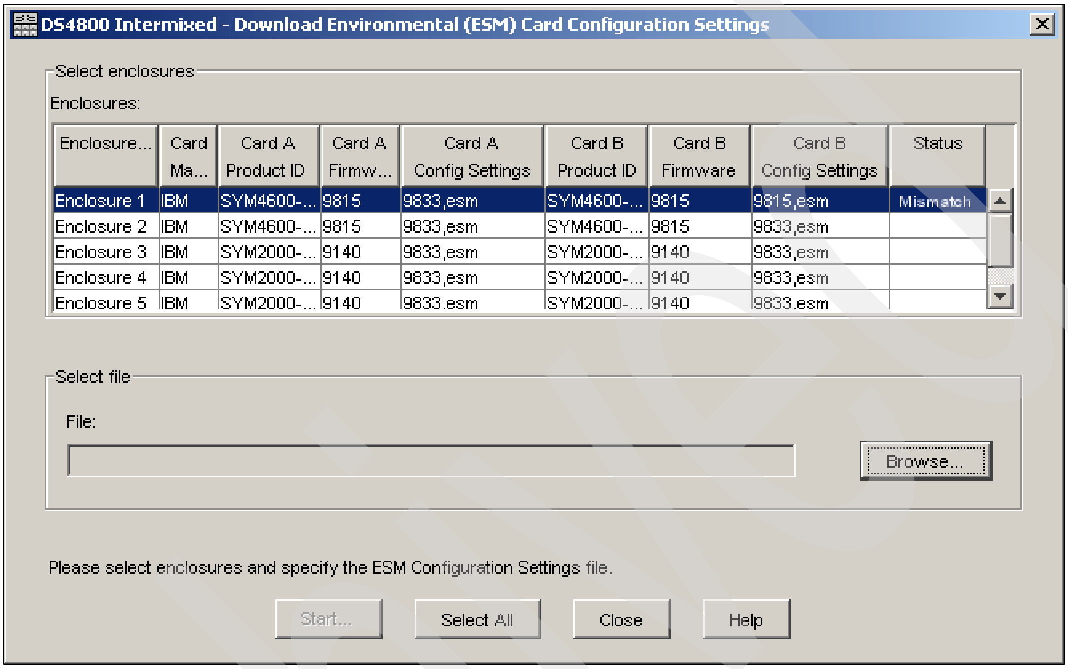


Figure 9-50 ESM configuration mismatch problem

This option should be used under supervision of your support representative when ESMs are reporting different versions of configuration settings. A configuration file is provided to reset the ESM NVSRAM settings.

Locating and removing parts

The DS4000 is a modular system containing customer replaceable units (CRUs). Most of the parts contain LED indications that allow a clear indication of the affected part. The Recovery Guru option will guide you in determine the problem and, if necessary, to replace the affected part.

“Recovery Guru” on page 397 discusses how to identify a part and whether it is safe to proceed with the removal. We now cover other options to identify a component in your DS4000 subsystem.

Locate option

This option allows you to physically identify a component by turning on a LED indication. To locate a drive, right-click the drive, then select **Locate** → **Drive**, as in Figure 9-51.

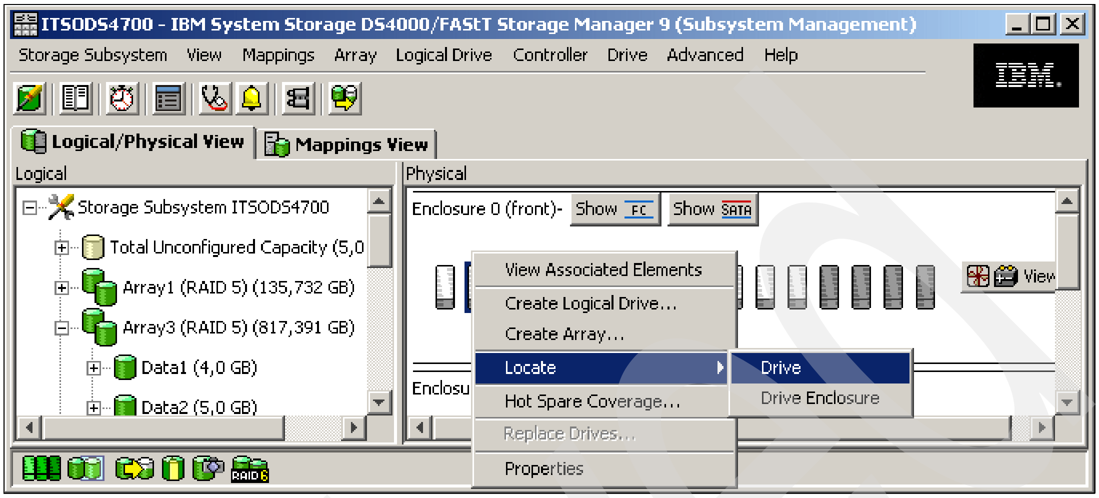


Figure 9-51 Locating a drive

This action will turn on the LED indication on the component selected, so if you have to perform any action you can physically identify it by the flashing LED.

The components that have identification LEDs to allow a physical identification are:

- Arrays: will turn on all drives LEDs
- Controller
- Drives
- Enclosures
- Storage Subsystem

Use the locate option especially in highly populated racks with multiple subsystems and enclosures. Also make sure that your Storage Manager enclosures are shown in the graphical interface in the same physical order as installed in the racks. This eliminates any possible confusion.

Prepare for removal option

Use this option also to locate a particular drive and perform removal preparation. During this action, the identify LED will turn on and the drive will be checked for safe removal of the component. The tool will not let you select a drive in use, for example. Use this option only when guided by the Recovery Guru or as requested by your support representative.

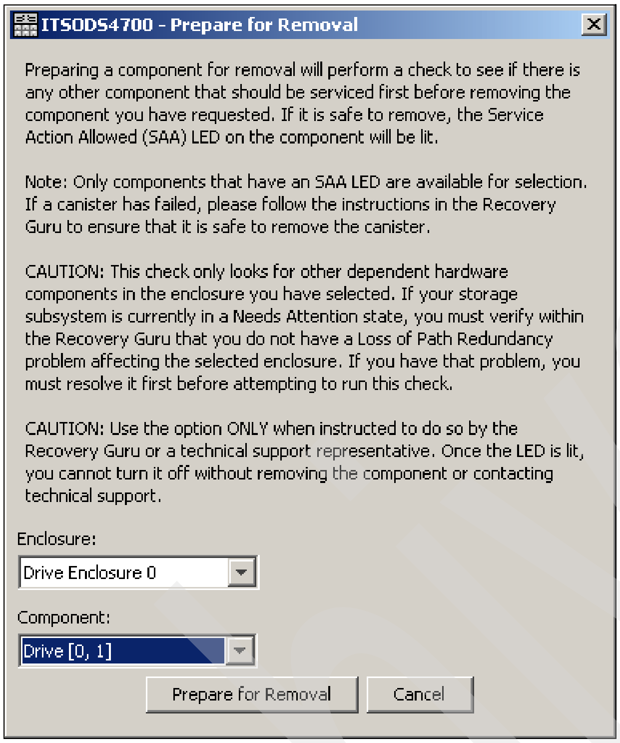


Figure 9-52 Prepare for Removal option

Storage Manager advanced recovery options

Here we discuss options that are accessible from the Advanced menu in the Storage Manager Subsystem Management window.

> **Important:** Most of these operations should not be attempted unless you have been instructed to by a technical support representative.

The Advanced Menu has the structure shown in Table 9-4.

Table 9-4 Storage Manager advanced menu structure

Maintenance	Troubleshooting	Recovery
► Export Array ► Import Array ► Download – Controller Firmware – Controller NVSRAM – Drive Firmware/Mode Pages – ESM Firmware – ESM Configuration Settings ► Activate Controller Firmware ► Clear Controller Firmware ► Persistant Reservations	► Support Data – Collect – Automatic settings ► Collect Drive Data ► View Event Log ► Drive Channels ► Capture State Information ► Run Diagnostics – Controller – Link status – Discrete Lines ► Prepare for Removal	► Initialize – Array – Drive – Logical Drive ► Revive Drive ► Clear Configuration – Storage Subsystem – Array ► Place Controller – Online – Offline – In Service Mode ► Reset controller ► Enable Controller Data Transfer ► Redistribute Logical drives ► Fail Drive ► Reconstruct Drive ► Defragment Array ► Check array Redundancy ► Unreadable Sectors

The different functions of the recovery menu is discussed in detail on the following pages.

Initializing a drive

Use the initialize drive option only when you have moved a subset of drives that were previously part of an array from one Storage System to another, to add space, not to migrate data. Because you did not move the entire set of drives, the array information about the drives is incomplete.

Therefore, to erase all logical drive/array information on these drives and make them available again for new logical drive creation, you must initialize these drives. This action erases the logical drive/array information and returns the selected drives to an unassigned state, adding new or additional unconfigured capacity to the storage subsystem. You can use this capacity to create additional logical drives. Perform the following steps. All data on the drives will be lost.

1. Select the drives to be initialized and select **Advanced** → **Recovery** → **Initialize** → **Drive**.
2. Type `yes` and click **OK** or press Enter.

Initializing a logical drive

A logical drive is automatically initialized when it is first created. All data is lost on the logical drive that is initialized with this function. Only use this option if you want to reuse an existing logical drive without the previously stored data.

To initialize a logical drive:

1. Select the logical drive to be initialized in the Logical view and select **Advanced** → **Recovery** → **Initialize** → **Logical Drive**.
2. Type `yes` and click **OK** or press Enter.

Important notes regarding this operation:

- You cannot cancel the operation after it begins.
- Do not use this option if any modification operations are in progress on the logical drive or the array.
- Do not change cache parameters of the logical drive while the initialize operation is in progress.
- All logical drive data is destroyed.

Initializing an array

All data is lost on the array that is initialized with this function. Only use this option if you want to reuse an existing array without the previously stored data.

1. Select the logical drive to be initialized in the Logical view and select **Advanced** → **Recovery** → **Initialize** → **Array**.
2. Type `yes` and click **OK** or press Enter.

Important notes regarding this operation:

- You cannot cancel the operation after it begins.
- Do not use this option if any modification operations are in progress on the logical drive or the array.
- Do not change cache parameters of the logical drive while the initialize operation is in progress.
- All logical drive data is destroyed.

Revive a drive

This option attempts to spin up a failed drive. However, the use of this function may affect the data on all the physical disks participating in the array. Use this process only under the direction of your support representative.

1. Select the drives to be initialized and select **Advanced** → **Recovery** → **Revive** → **Drive**.
2. Type `yes` and click **OK** or press Enter.

This should not be used as a substitute for replacing a failed drive unless multiple drives were marked as failed due to another problem.

Revive an array

This option is not available in Storage Manager 10. It was used in previous versions to revive (attempt to spin up) the drives in a selected array. The Recovery Guru advises using this option when a failed LUN or a failed canister causes a multi-drive failure. In either case, it is possible that the data on the drives is not corrupt, and that the array can be successfully revived. Perform the following steps:

> **Note:** Correct use of this option depends on the data configuration on all physical disks in the array. This operation fails if one or more of the drives is unusable and cannot be revived.

1. To revive a failed array select the array to be revived and select **Advanced** → **Recovery** → **Revive** → **Array**.
2. Type `yes` and click **OK** or press Enter.

Never attempt to revive an array unless supervised by a technical support representative.

Clear configuration

Use this option with caution. This option also lets you correct configuration errors that cannot be corrected by using any other method. You have the following options:

- Clear all storage subsystem configuration:

  a. Select **Advanced** → **Recovery** → **Clear Configuration** → **Storage Subsystem**.
  b. Type `yes` and click **OK** or press Enter.

  Important notes on resetting the configuration:

  – The current configuration is deleted, destroying any existing data on the storage subsystem.
  – Do not attempt this operation unless all data on the storage subsystem has been backed up to tape or to another storage subsystem.
  – All logical drives and arrays are deleted.
  – The user-supplied name is deleted.
  – If you have password protected your storage subsystem to prevent destructive operations, that password protection is lost.

- Clear all array configuration.

  This option clears all the arrays in your DS4000, destroying existing data on the storage subsystem. If you wan to delete one specific array, select it from the Subsystem Management, right-click above it, and select the **Delete** option.

  a. Select **Advanced** → **Recovery** → **Clear Configuration** → **Array**.
  b. Type `yes` and click **OK** or press Enter.

In both cases make sure that you have backed up your data or that it is no longer needed.

Place controller

Each controller has three states:

- Online

  When a controller is active it is called online. To place a controller online from either an offline or service mode, select the controller. Select **Advanced** → **Recovery** → **Place Controller** → **Online**. Select **Yes** to confirm that the controller is online. This removes the red cross through the controller if the controller was offline, or reinstates the look of the controller if it was in service mode.

  Placing the controller online from being in service mode redistributes the logical drives back to their preferred controller.

  Make sure that the problem that caused the controller to go offline was corrected before placing the controller online.

- Offline

  To place a controller offline, select the controller and select **Advanced** → **Recovery** → **Place Controller** → **Offline**. Select **Yes** to confirm that the controller is offline. This places a red cross through the controller to signify that it is offline.

- In service mode

  Placing a controller in service mode forces all logical drives to the active controller. This mode provides for continuous data access. The preferred logical drive owner does not change. When the controller is removed from service mode, the logical drives are redistributed back to their preferred owner.

  A controller can only be put into service mode if the storage server has two controllers and they are both online or active. Placing a controller in service mode is a persistent setting. Even if the controller is reset it will return to the service mode state. A controller that is in service mode cannot be made the owner of any logical drive.

Reset controller

Use the Reset option to reset a controller from within the storage management software. Resetting a controller makes the controller unavailable for I/O until the reset is complete. If a host is using logical drives owned by the controller being reset, the I/O directed to the controller is rejected, so before proceeding, either verify that the logical drives owned by the controller are not in use or ensure that there is a multipath driver properly installed on all hosts using these logical drives. This is done by selecting the controller to be reset:

1. Select **Advanced** → **Recovery** → **Reset** → **Controller**.
2. Click **OK** to reset the controller.

Controller data transfer

Disabling data transfer ensures that the controller is no longer available for I/O operations.

Select a controller in the Physical view. Then select **Advanced** → **Recovery** → **Disable Controller Data Transfer**. The Confirm Data Transfer Disable dialog is displayed. Type `Yes` and select **OK**.

The controller no longer accepts I/O and is marked with a dashed red border in the Physical view. If a multi-path driver is in use, I/O is sent to the other controller path, initiating a logical drive ownership transfer.

To enable data transfer and enable the controller for I/O operations, select a controller in the Physical view, then select **Advanced** → **Recovery** → **Enable Controller Data Transfer**.

Redistribute logical drives

This happens when multipath drivers move logical drives from their preferred controller owner or when a problem occurs along the data path between the host and the storage server. When the problem is corrected, you should move all the logical drives back to their preferred controllers to maintain controller balance. Be aware that manually moving back to their preferred path, since checking first the path is working, might generate an access problem.

Select a controller in the Physical view, then select **Advanced** → **Recovery** → **Redistribute Logical Drives**.

Fail drive

Use this option to fail a selected drive. Select one or more drives in the Physical view, then select **Advanced** → **Recovery** → **Fail Drive**.

In the confirm Fail Drive dialog type `Yes` and select **OK**.

> **Note:** It is a good practice to fail an unused drive with this option before the drive is physically removed, that is, to be migrated to a different drive slot or another Storage System. If you fail a drive before moving it, you ensure that it will not be in use when you remove it. After inserting the drive to a new slot, it will become optimal automatically.

Reconstruct drive

Use this option to manually start reconstruction of a drive only if instructed to do so by the Recovery Guru. This option is available only when a drive is assigned to a RAID-1, 3, 5, or 6 array, with either a failed status or a replaced status, and reconstruction did not automatically start. Normally, drive reconstruction should begin automatically after a drive replacement, so if it is not, make sure there is not another problem by using MEL or RLS before continuing.

1. Select one drive in the Physical view, then select **Advanced** → **Recovery** → **Reconstruct Drive**.
2. In the confirm Reconstruct dialog select **Yes**.
3. The confirm Reconstruct dialog is closed and the drive reconstruction starts. To view the reconstruction progress, select a logical drive in the Logical view that is associated with this drive, then select **Logical Drive** → **Properties**.

Defragment array

A logical drive can be deleted anytime to free the space in the array. The free space might be fragmented within the array in different free space nodes.

Because new logical drives cannot spread across several free space nodes, the logical drive size is limited to the greatest free space node available, even if there is more free space in the logical drive. The array needs to be defragmented first to consolidate all free space nodes into one free space node for the array. Then all new logical drives can use the entire available free space.

Open the Subsystem Management window. Highlight the array to defragment and click **Advanced** → **Recovery** → **Defragment** to start the procedure. The defragmentation can run concurrently with normal I/O, but it impacts performance because data of the logical drives must be moved within the array. Depending on the array configuration, this process continues to run for a long period of time. Once the procedure is started, it cannot be stopped again. During this time, no configuration changes can be performed on the array.

The defragmentation done on the DS4000 power supply fan unit only applies to the free space nodes on the array. It is not connected to a defragmentation of the file system used by the host operating systems in any way.

Check array redundancy

Use this option to check the redundancy on a selected array only when instructed to do so by the Recovery Guru.

- This option cannot be used on RAID-0 arrays that have no redundancy.
- If you use this option on a RAID-1 array, the redundancy check compares the data on the mirrored drives.
- If you perform this operation on a RAID-3, 5, or 6 array, the redundancy check inspects the parity information that is striped across the drives.

To successfully perform this operation:

- All the logical drives in the array must be in optimal status.
- The array must have no logical drive modification operations in progress.
- This option can be performed only on one array at a time.

The steps are:

1. Select a single array in the Logical view, then select **Advanced** → **Recovery** → **Check Array Redundancy**.
2. A confirmation dialog is displayed, warning that this option should only be used when instructed to do so by Recovery Guru. Select **Yes** to continue. Select **Start**.
3. The check redundancy operation begins. The logical drives in the array are sequentially scanned. Select **Done** after the last logical drive in the array has been checked.

Unreadable sectors

New since microcode level 7.10, you now have the option from the Storage Manager graphical interface to display a detailed list of all known unreadable sectors detected on the storage subsystem.

Unreadable sectors are specified as a logical drive block address that is no longer readable due to a disk media-related error, double fault if under RAID protection different from RAID-0, or single fault error for non-redundant RAID-0 logical volumes. Since the small portion of data is no longer readable, it is important to indicate where the volume is being affected. This log is now available from the graphical interface. To view the unreadable sectors log, select **Advanced** → **Recovery** → **Unreadable Sectors**.

Unreadable sectors are detected during normal I/O and during modification operations, such as reconstructions. When unreadable sectors are detected, a notification is posted to the event log and a storage subsystem attention indication icon is displayed for the storage subsystem.

Run the Recovery Guru to determine which unreadable sector condition needs attention. Furthermore, the Recovery Guru suggests the appropriate procedures for resolving the condition. The same list is also collected with the collect all support data option.

A

VIOS error log message for DS Midrange Storage Systems

This appendix provides reference information for analyzing VIOS error log messages logged for DS Midrange Storage System configuration changes or problems either for RDAC (FCP_ARRAY) or MPIO (SC_PCM) attached storage.

The source of this reference information is I*IBM System Storage DS4000 and Storage Manager V10.10*, SG24-7010.

VIOS error log information for RDAC attached DS Midrange Storage

The error log information is:

- FCP_ARRAY_ERR1 *ARRAY OPERATION ERROR*

  A permanent hardware error involving the disk array media.
- FCP_ARRAY_ERR2 *ARRAY OPERATION ERROR*

  A permanent hardware error.
- FCP_ARRAY_ERR3 *ARRAY OPERATION ERROR*

  A permanent error detected by the array adapter.
- FCP_ARRAY_ERR4 *ARRAY OPERATION ERROR*

  A temporary error within the array, communications, adapter, and so on.
- FCP_ARRAY_ERR5 *UNDETERMINED ERROR*

  An undetermined error has occurred.
- FCP_ARRAY_ERR6 *SUBSYSTEM COMPONENT FAILURE*

  A degradation condition has occurred other than a disk drive.
- FCP_ARRAY_ERR7 *CONTROLLER HEALTH CHECK FAILURE*

  A health check on the passive controller has failed.
- FCP_ARRAY_ERR8 *ARRAY CONTROLLER SWITCH*

  One array controller has become unavailable, so I/O has moved to the other controller.
- FCP_ARRAY_ERR9 *ARRAY CONTROLLER SWITCH FAILURE*

  An array controller switch has failed.
- FCP_ARRAY_ERR10 *ARRAY CONFIGURATION CHANGED*

  A logical unit has been moved from one controller to the other (most likely by the action of an alternate host).
- FCP_ARRAY_ERR11 *IMPROPER DRIVE TYPE FOR DUAL ACTIVE MODE*

  This error should not be possible on the 2102 array and exists for history reasons only. FCP_ARRAY_ERR11 might be reused for a different error in the future.
- FCP_ARRAY_ERR12 *POLLED AEN FAILURE*

  An automatic error notification has failed.
- FCP_ARRAY_ERR13 *ARRAY INTER-CONTROLLER COMMUNICATION FAILURE*

  The controllers are unable to communicate with each other. This could result from one of the controllers being rebooted while the error log was being generated. However, it could be a much more serious error that indicates a problem with the Fibre Channel connections.
- FCP_ARRAY_ERR14 *ARRAY DRIVE FAILURE*

  A serious or unrecoverable error has been detected on a physical disk within the DS4000 subsystem. A system engineer might be able to obtain the exact cause from an analysis of the sense data.

- FCP_ARRAY_ERR15 *CACHE BATTERY LOW/DATA LOSS POSSIBLE*

  If a controller card is replaced, it is likely that the cache batteries will be flat. It can take two days for the cache batteries to be fully recharged. During this time errors are logged in the error log. Do not replace the controller.

- FCP_ARRAY_ERR16 *CACHE BATTERY CHARGE BELOW 87.5%*

  If a controller card is replaced, it is likely that the cache batteries will be flat. It can take two days for the cache batteries to be fully recharged. During this time errors are logged in the error log. Do not replace the controller.

- FCP_ARRAY_ERR17 *WORLDWIDE NAME CHANGED*

  A controller has changed worldwide names (most likely either it was replaced without placing it in the reset state first, or the cabling was changed so that a different controller with the same SCSI ID is on the loop.

- FCP_ARRAY_ERR18 *RESERVATION CONFLICT*

  An operation failed because the disk array logical drive (LUN) is reserved by another host.

- FCP_ARRAY_ERR19 *SNAPSHOT VOLUME'S REPOSITORY FULL*

  The repository capacity limit has been reached. To resolve this error you can increase the repository capacity.

- FCP_ARRAY_ERR20 *SNAPSHOT OPERATION STOPPED BY ADMIN*

  The FlashCopy (snapshot) operation has been disabled or stopped. To resolve this error you can recreate the FlashCopy.

- FCP_ARRAY_ERR21 *SNAPSHOT REPOSITORY METADATA ERROR*

  There was a problem with the metadata of the FlashCopy (snapshot) repository during the FlashCopy operation. To resolve this error you can recreate the FlashCopy.

- FCP_ARRAY_ERR22 *REMOTE VOL MIRRORING: ILLEGAL I/O ORIGIN*

  The primary logical drive received I/O from a remote array, or the secondary logical drive received I/O from other than the primary logical drive. To resolve this error you can try the operation again.

- FCP_ARRAY_ERR23 *SNAPSHOT OPERATION NOT ALLOWED*

  The repository capacity limit has been reached, so the FlashCopy (snapshot) operation has failed. To resolve this error you can delete or recreate the FlashCopy.

- FCP_ARRAY_ERR24 *SNAPSHOT VOLUME'S REPOSITORY FULL*

  The repository capacity limit has been reached. To resolve this error you can delete or recreate the FlashCopy (snapshot).

- FCP_ARRAY_ERR25 *CACHED DATA WILL BE LOST IF CONTROLLER FAILS*

  This message is a warning that a disk array logical drive (LUN) is running with write cache enabled and cache mirroring disabled. The warning displays when the LUN is opened, and it displays again every 24 hours until cache mirroring is enabled again. If a controller failure or a power down occurs while the LUN is running in this mode, data that is in the write cache (but not written to the physical disk media) might be lost. This can result in corrupted files, file systems, or databases.

- FCP_ARRAY_ERR26 *LOGICAL VOLUME IS WRITE PROTECTED*

  The status of the logical drive is read-only. The probable reason is that it is a secondary logical drive of a FlashCopy, Volume Copy, or remote mirror pair. Check which relationship applies to the logical drive.

  – For FlashCopy, a status of read-only on the secondary logical drive usually indicates that the repository is full.

  – For Volume Copy, both the primary and secondary logical drives are read-only during the copy. The secondary logical drive is read-only when the copy is stopped but the copy pair had not been deleted.

  – For remote mirroring, the secondary logical drive is always read-only as long as the mirror is active.

- FCP_ARRAY_ERR27 *SINGLE CONTROLLER RESTARTED*

  The subsystem is operating as a single controller, and an error has been repaired. The error might have been a communication or hardware problem, or it might have occurred because a LUN was moved to a controller that does not have a path to the current host. If this is a dual-controller subsystem, find the reason that the subsystem is operating in single-controller mode and resolve the problem. Possible reasons include:

  – An HBA, switch port, switch, DS4000 port, or DS4000 controller was unavailable during the last system reboot or the last time that the **cfgmgr** command was run.

  – A user removed a path (dac) as part of a Fibre Channel adapter hot swap operation.

- FCP_ARRAY_ERR28 *SINGLE CONTROLLER RESTART FAILURE*

  The subsystem is operating as a single controller and the error has not been repaired. There is a problem with the path between this host and the subsystem or with the subsystem itself. The host has attempted to communicate with the subsystem and that communication has failed. If the number of retries that is specified in the ODM attribute switch_retries is reached, the I/O is failed back to the user. Repair the error. Then if this is a dual-controller subsystem, find the reason that the subsystem is operating in single-controller mode and resolve that problem. Possible reasons include:

  – An HBA, switch port, switch, DS4000 port, or DS4000 controller was unavailable during the last system reboot or the last time that the **cfgmgr** command was run.

  – A user removed a path (dac) as part of a Fibre Channel adapter hot swap operation.

VIOS error log information for MPIO attached DS Midrange Storage

The error log information is:

- SC_DISK_PCM_ERR1 *Subsystem Component Failure*

  The storage subsystem has returned an error indicating that some component (hardware or software) of the storage subsystem has failed. The detailed sense data identifies the failing component and the recovery action that is required. Failing hardware components should also be shown in the Storage Manager software, so the placement of these errors in the error log is advisory and is an aid for your technical-support representative.

- SC_DISK_PCM_ERR2 *Array Active Controller Switch*

  The active controller for one or more hdisks associated with the storage subsystem has changed. This is in response to some direct action by the AIX host (failover or autorecovery). This message is associated with either a set of failure conditions causing a failover or, after a successful failover, with the recovery of paths to the preferred controller on hdisks with the autorecovery attribute set to yes.

- SC_DISK_PCM_ERR3 *Array Controller Switch Failure*

  An attempt to switch active controllers has failed. This leaves one or more paths with no working path to a controller. The AIX MPIO PCM will retry this error several times in an attempt to find a successful path to a controller.

- SC_DISK_PCM_ERR4 *Array Configuration Changed*

  The active controller for an hdisk has changed, usually due to an action not initiated by this host. This might be another host initiating failover or recovery, for shared LUNs, a redistribute operation from the Storage Manager software, a change to the preferred path in the Storage Manager software, a controller being taken offline, or any other action that causes the active controller ownership to change.

- SC_DISK_PCM_ERR5 *Array Cache Battery Drained*

  The storage subsystem cache battery has drained. Any data remaining in the cache is dumped and is vulnerable to data loss until it is dumped. Caching is not normally allowed with drained batteries unless the administrator takes action to enable it within the Storage Manager software.

- SC_DISK_PCM_ERR6 *Array Cache Battery Charge Is Low*

  The storage subsystem cache batteries are low and need to be charged or replaced.

- SC_DISK_PCM_ERR7 *Cache Mirroring Disabled*

  Cache mirroring is disabled on the affected hdisks. Normally, any cached write data is kept within the cache of both controllers so that if either controller fails there is still a good copy of the data. This is a warning message stating that loss of a single controller will result in data loss.

- SC_DISK_PCM_ERR8 *Path Has Failed*

  The I/O path to a controller has failed or gone offline.

- SC_DISK_PCM_ERR9 *Path Has Recovered*

  The I/O path to a controller has resumed and is back online.

- SC_DISK_PCM_ERR10 *Array Drive Failure*

  A physical drive in the storage array has failed and should be replaced.

- SC_DISK_PCM_ERR11 *Reservation Conflict*

  A PCM operation has failed due to a reservation conflict. This error is not currently issued.

- SC_DISK_PCM_ERR12 *Snapshot™ Volume's Repository Is Full*

  The snapshot volume repository is full. Write actions to the snapshot volume will fail until the repository problems are fixed.

- SC_DISK_PCM_ERR13 *Snapshot Op Stopped By Administrator*

  The administrator has halted a snapshot operation.

- SC_DISK_PCM_ERR14 *Snapshot repository metadata error*

  The storage subsystem has reported that there is a problem with snapshot metadata.

- SC_DISK_PCM_ERR15 *Illegal I/O - Remote Volume Mirroring*

  The I/O is directed to an illegal target that is part of a remote volume mirroring pair (the target volume rather than the source volume).

- SC_DISK_PCM_ERR16 *Snapshot Operation Not Allowed*

  A snapshot operation that is not allowed has been attempted.

- SC_DISK_PCM_ERR17 *Snapshot Volume's Repository Is Full*

  The snapshot volume repository is full. Write actions to the snapshot volume will fail until the repository problems are fixed.
- SC_DISK_PCM_ERR18 *Write Protected*

  The hdisk is write-protected. This can happen if a snapshot volume repository is full.
- SC_DISK_PCM_ERR19 *Single Controller Restarted*

  The I/O to a single-controller storage subsystem is resumed.
- SC_DISK_PCM_ERR20 *Single Controller Restart Failure*

  The I/O to a single-controller storage subsystem is not resumed. The AIX MPIO PCM will continue to attempt to restart the I/O to the storage subsystem.

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this IBM Redbooks publication.

IBM Redbooks publications

For information about ordering these publications see “How to get Redbooks publications” on page 433. Note that some of the documents referenced here may be available in softcopy only.

- *IBM BladeCenter JS12 and JS22 Implementation Guide*, SG24-7655
- *IBM System Storage Copy Services and IBM i: A Guide to Planning and Implementation*, SG24-7103
- *IBM i and IBM System Storage: A Guide to Implementing External Disks on IBM i*, SG24-7120
- *DS4000 Best Practices and Performance Tuning Guide*, SG24-6363
- *DS5000 Disk Storage Subsystem Architecture. Implementation and Usage*, SG24-7676

Online resources

The IBM Systems Information Center Web site is also relevant as a further information source:

http://publib.boulder.ibm.com/eserver/

How to get Redbooks publications

You can search for, view, or download Redbooks publications, Redpapers publications, Technotes, draft publications, and additional materials, as well as order hardcopy Redbooks publications, at this Web site:

ibm.com/redbooks

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

Index

IBM

Redbooks

IBM i and Midrange External Storage

IBM i and Midrange External Storage

Learn the architecture of Midrange Storage for IBM i as a client of VIOS

Plan for the best performance of Midrange Storage for IBM i

Implement Midrange Storage with Copy Services

This IBM® Redbooks® publication describes the new IBM i Midrange External Storage solutions available for IBM POWER™ Systems POWER6™ servers with IBM i being a client of IBM Virtual I/O Server (VIOS). It introduces the VIOS virtualization concepts and IBM DS Midrange External Storage Systems architecture of the supported models DS3400, DS4700, DS4800, and DS5000, discusses planning and sizing for IBM i Midrange External Storage, and provides detailed implementation procedures including IBM DS Midrange Storage Copy Services. Finally, it provides monitoring, maintenance, and troubleshooting hints for the triumvirate of IBM i, VIOS, and IBM DS Midrange External Storage.

The information provided by this book will help customers, business partners, and IBM service professionals with planning and implementing IBM i Midrange External Storage solutions.

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

SG24-7668-00 ISBN 0738432083