

IBM Midrange System Storage Hardware Guide

DS4000 and DS5000 hardware
planning and configuration

Remote Support Manager (RSM)
configuration

Features of Storage Manager
V10.60



Sangam Racherla
Bruce Allworth
Alessio Bagnaresi
Chris Bogdanowicz
Corne Lottering
Pablo Pedrazas
Frank Schubert
John Sexton
Alexander Watson

Redbooks



International Technical Support Organization

IBM Midrange System Storage Hardware Guide

March 2010

Archived

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

Archived

Second Edition (March 2010)

This edition applies to:

- ▶ IBM Midrange System Storage running V7.60 firmware.
- ▶ IBM DS Storage Manager V10.60.

© Copyright International Business Machines Corporation 2010. All rights reserved.

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

Contents

Notices	ix
Trademarks	x
Preface	xi
The team who wrote this book	xi
Now you can become a published author, too!	xiv
Comments welcome	xiv
Stay connected to IBM Redbooks	xiv
Summary of changes	xv
March 2010, Second Edition	xv
Chapter 1. Introduction to IBM Midrange System Storage storage subsystems	1
1.1 Positioning the DS4000/DS5000 series	2
1.2 IBM Midrange System Storage storage server models	3
1.3 IBM Midrange System Storage expansion enclosure	6
1.4 IBM System Storage DS Storage Manager software	7
1.5 IBM Midrange System Storage hard disk drives	9
1.6 iSCSI basics	11
1.7 Fibre Channel direct/switch basics	12
Chapter 2. New features	15
2.1 Full Disk Encryption capable disk drive modules (DDM)	16
2.2 Solid State Drive (SSD) module	16
2.3 600 GB FC disk drive module	17
2.4 1 TB SATA enhanced disk drive module	17
2.5 IBM System Storage DS5020 storage subsystem and EXP520	17
2.5.1 Highlights	18
2.6 EXP5060 expansion enclosure	18
2.7 iSCSI host interface	19
2.8 8 Gbps FC host interface	19
2.9 16 port host port upgrade for the DS5100	19
2.10 64 GB cache upgrade for the DS5100 and DS5300	20
2.11 Remote Support Manager Hardware Model RS2	20
2.12 Increased drive support for the DS5100 and DS5300	20
2.13 New Storage Manager Subsystem Management window (SMW) design	21
Chapter 3. IBM System Storage DS4000 and DS5000 hardware	23
3.1 IBM System Storage DS4700 Express	24
3.1.1 DS4700 features	24
3.2 DS4700 model comparison	26
3.3 DS4000 series expansion enclosures	26
3.3.1 IBM Total Storage DS4000 EXP710 expansion enclosure	27
3.3.2 IBM System Storage EXP810 expansion enclosure	27
3.3.3 Intermixing EXP810 and EXP710	28
3.4 DS5100 and DS5300 storage subsystems	29
3.4.1 DS5100 and DS5300 controller architecture	31
3.4.2 DS5000 storage subsystem chassis design	40
3.4.3 DS5000 storage subsystem front view	40

3.4.4	Interconnect module and battery packs	42
3.4.5	DS5000 storage subsystem rear view	45
3.4.6	DS5000 storage subsystem LED indicator lights	46
3.4.7	DS5000 storage subsystem host-side connections	55
3.4.8	DS5000 storage subsystem drive-side connections	59
3.4.9	DS5100 and DS5300 storage subsystem drive-side cabling	61
3.4.10	DS5100 and DS5300 storage subsystem additional connections	65
3.5	DS5020 storage subsystem	66
3.5.1	DS5020 controller architecture	68
3.5.2	DS5020 components	70
3.5.3	DS5020 Storage Subsystem front view	74
3.5.4	DS5020 storage subsystem rear view	75
3.5.5	DS5020 storage subsystem LED indicator lights	79
3.5.6	DS5020 storage subsystem host-side connections	84
3.5.7	DS5020 storage subsystem drive-side connections	85
3.5.8	DS5020 storage subsystem drive-side cabling	87
3.5.9	DS5020 storage subsystem additional connections	91
3.6	DS5000 series product comparison	93
3.6.1	DS5020 product comparison	93
3.6.2	DS5300 product comparison	94
3.7	DS5000 series physical specifications	95
3.8	DS5000 supported operating systems	98
3.9	DS5000 storage subsystem disk enclosures	98
3.9.1	EXP5000 and EXP520 Storage Expansion Unit	99
	Chapter 4. IBM System Storage DS planning and configuration	103
4.1	Planning your DS storage structure	104
4.1.1	DS5000 arrays and RAID levels	104
4.1.2	Logical drives and controller ownership	114
4.1.3	Hot spare drive	114
4.1.4	Storage partitioning	116
4.1.5	Segment size	120
4.1.6	Cache parameters	122
4.1.7	Security planning: Full Disk Encryption	123
4.2	Planning for premium features	124
4.2.1	Disk Encryption or FDE	125
4.2.2	Storage partitioning	125
4.2.3	Drive slot limit	125
4.2.4	FlashCopy	125
4.2.5	VolumeCopy	125
4.2.6	Enhanced Remote Mirroring (ERM)	126
4.2.7	FC/SATA intermix	127
4.3	Planning your host attachment method	128
4.3.1	Fibre Channel: SAN or Direct Attach	128
4.3.2	Planning for iSCSI attachment	130
4.4	Host support and multipathing	132
4.4.1	Supported server platforms	132
4.4.2	Supported operating systems	133
4.4.3	Clustering support	133
4.4.4	Multipathing	133
4.4.5	Microsoft Windows MPIO	135
4.4.6	AIX MPIO	136
4.4.7	AIX Subsystem Device Driver Path Control Module (SDDPCM)	137

4.4.8 Linux: RHEL/SLES RDAC	137
4.4.9 Virtualization	138
4.4.10 Auto Logical Drive Transfer (ADT) feature	138
4.5 Operating system restrictions	141
4.5.1 Maximum file system size	142
4.5.2 Maximum number of LUNs per host	142
4.6 Storage Manager software	143
4.6.1 Storage subsystem management methods	145
4.6.2 Storage Manager client	148
4.6.3 Event Monitor service	151
4.6.4 Storage Manager utilities	152
4.7 Installing IBM System Storage DS Storage Manager	152
4.7.1 Installing Storage Manager in a Windows Server 2008 host	153
4.8 Preparing the DS5000 storage subsystem	159
4.8.1 Physical installation	159
4.8.2 Powering on the storage subsystem	160
4.8.3 Configuring IP addresses of the controllers	161
4.8.4 Using and configuring the DS Storage Manager client	166
4.8.5 Updating the controller microcode	174
4.9 Step-by-step configuration	175
4.9.1 Configuration planning	175
4.9.2 Enabling the premium features	176
4.9.3 Automatic configuration	179
4.9.4 Manual configuration	185
4.9.5 Configuring storage partitioning	201
4.9.6 Configuring mapped drives from Windows	219
4.9.7 Monitoring and alerting	223
4.9.8 Saving the configuration	227
4.10 Advanced functions	233
4.10.1 Expanding arrays	233
4.10.2 Defragment an array	235
4.10.3 Changing the RAID array level	236
4.10.4 Unconfiguring a storage subsystem and arrays	237
4.10.5 Performing advanced functions on logical drives (LUNs)	238
4.10.6 Cache parameters	249
4.10.7 Media scan	253
4.10.8 Failover alert delay	255
4.10.9 Persistent reservations	257
4.10.10 Automatic firmware synchronization	258
Chapter 5. Disk Security with Full Disk Encryption drives	259
5.1 The need for encryption	260
5.1.1 Encryption method used	260
5.2 Disk Security components	262
5.2.1 DS5000 Disk Encryption Manager	262
5.2.2 Full Data Encryption (FDE) disks	263
5.2.3 Premium feature license	263
5.2.4 Keys	263
5.2.5 Security key identifier	264
5.2.6 Passwords	264
5.3 Setting up and enabling a secure disk	265
5.3.1 FDE and premium feature check	265
5.3.2 Secure key creation	266

5.3.3	Enable Disk Security on array	269
5.4	Additional secure disk functions	270
5.4.1	Changing the security key	270
5.4.2	Save security key file	272
5.4.3	Secure erase	273
5.4.4	FDE drive status	274
5.4.5	Hot spare drive	275
5.5	Migrating secure disk arrays	275
5.5.1	Planning checklist	275
5.5.2	Export the array	276
5.6	Import secure drive array	278
5.6.1	Unlock drives	280
5.6.2	Import array	281
Chapter 6. IBM Remote Support Manager for Storage		285
6.1	IBM Remote Support Manager for Storage	286
6.1.1	Hardware and software requirements	287
6.1.2	DS-RSM Model RS2	289
6.1.3	Installation choices for RSM for Storage	290
6.1.4	How RSM for Storage works	291
6.1.5	Notification e-mail and events filtering	292
6.1.6	Remote access methods	297
6.1.7	RSM management interface	298
6.1.8	RSM security considerations	299
6.2	Installing and setting up RSM	301
6.2.1	Installing the host OS	302
6.2.2	Installing RSM	302
6.2.3	Setting up RSM	302
6.2.4	Configuring SNMP traps in Storage Manager	314
6.2.5	Activating RSM	316
6.2.6	Remote access security	317
6.2.7	Managing alerts	322
Chapter 7. Advanced maintenance, troubleshooting, and diagnostics		327
7.1	Upgrades and maintenance	328
7.1.1	Displaying installed firmware versions	328
7.1.2	Obtaining updates	330
7.1.3	Planning for upgrades	331
7.1.4	Updating the DS5000 storage subsystem host software	332
7.1.5	Updating controller firmware	333
7.1.6	Controller Firmware Upgrade Tool	339
7.1.7	DbFix tool	343
7.1.8	Updating the ESM board firmware	347
7.1.9	Updating the hard disk drives' firmware	349
7.1.10	Updating host bus adapter (HBA) firmware	357
7.2	Handling premium features	362
7.2.1	Listing premium features/feature enabler	363
7.2.2	Enabling a premium feature	365
7.2.3	Disabling a premium feature	368
7.3	Saving and loading the configuration	368
7.3.1	Storage subsystem profile	371
7.4	Migrating arrays between DS storage subsystems	375
7.4.1	Intermixing EXP810 and EXP5000 storage expansion enclosures	376

7.4.2	Intermixing EXP520 and EXP810 storage expansion enclosures	376
7.4.3	Migration prerequisites	376
7.4.4	Migrating an array	378
7.4.5	Importing an array	382
7.5	Performing an upgrade from a DS4700 or DS4800 storage subsystem to a DS5000 storage subsystem	387
7.5.1	Planning the upgrade	388
7.5.2	Preparing the new storage subsystem	389
7.5.3	Preparing the original storage subsystem	389
7.5.4	Upgrading the controller firmware	390
7.5.5	Switching from the original to the new storage subsystem	392
7.5.6	Preparing the new storage subsystem for use	393
7.6	Securing the DS5000 storage subsystem client using remote management	394
7.7	Preventative maintenance and data collection	396
7.7.1	Storage Manager Enterprise Management window (EMW)	396
7.7.2	Storage Manager Subsystem Management window (SMW)	398
7.7.3	Storage subsystem profile	398
7.7.4	Recovery Guru	399
7.7.5	Major Event Log	400
7.7.6	Collect All Support Data option	401
7.7.7	Media Scan	403
7.7.8	Pre-read redundancy check	405
7.8	Problem determination	405
7.8.1	Diagnosing drive-side problems	406
7.8.2	Diagnosing host-side problems	418
7.8.3	Storage Manager communication problems	428
7.9	Replacement and maintenance procedures	429
7.9.1	Managing disk failures	429
7.9.2	Managing disks with an impending drive failure error	432
7.9.3	Monitoring Solid State Drives (SSD)	434
7.9.4	Managing battery issues	435
7.10	Replacing adapters (HBA) and storage controllers	435
7.11	HBAs and operating system tools	436
7.11.1	Brocade HBA and Brocade Host Configuration Manager (HCM)	436
7.11.2	Emulex HBA tools	441
7.11.3	Qlogic HBAs and SANsurfer (Windows/Linux)	443
7.11.4	Windows Server 2008	451
7.11.5	Linux	459
7.11.6	AIX	469
	Chapter 8. Command-line interface and Script Editor	473
8.1	Command-line interface (CLI)	474
8.1.1	Using CLI commands	474
8.1.2	CLI parameters	476
8.1.3	Syntax requirements	481
8.1.4	Error reporting	482
8.1.5	Commands overview	483
8.1.6	CLI examples	493
8.2	Script Editor	499
8.2.1	Using the Script Editor	499
8.2.2	Embedding commands in batch files	503
	Appendix A. Overview of IBM System Storage DS5000 RAID types	505

DS5000 arrays and RAID levels	506
RAID levels	506
RAID 0: For performance, but generally not recommended	506
RAID 1: For availability/good read response time	507
RAID 3: Sequential access to large files	508
RAID 5: High availability and fewer writes than reads	509
RAID 6: High availability with additional fault tolerance	510
RAID 10: Higher performance than RAID 1	511
RAID summary	512
RAID reliability considerations	513
Appendix B. Deploying iSCSI host interface controllers on the IBM System Storage DS5000 series	515
iSCSI technology	516
iSCSI Qualified Name (IQN)	516
iSCSI physical components	517
Network considerations	518
iSCSI configurations on the DS5000 series	519
Jumbo frames	519
Virtual Local Area Networks	520
Ethernet priority	521
Security	522
Internet Storage Name Service	522
Challenge Handshake Authentication Protocol	523
iSCSI performance considerations	524
Multipathing iSCSI	525
Other iSCSI performance considerations	525
Appendix C. Solid State Drives on the IBM System Storage DS5000 series	527
SSD technology	528
Solid State Drives in tiered storage	528
Implementing tiered storage	530
Solid State Drives on a DS5000 storage subsystem	530
Identifying SSD in Storage Manager	531
Wear life	531
SSD performance on DS5000 storage subsystems	532
A need for high performance disks	532
Initial lab tests of SSDs on a DS5000 storage subsystem	533
SSD summary	534
Related publications	535
IBM Redbooks	535
Other publications	535
Online resources	536
How to get Redbooks	536
Help from IBM	536
Index	537

Notices

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

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

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

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

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

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

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

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

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

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


COPYRIGHT LICENSE:

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

Trademarks

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

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

AIX 5L™	IBM®	System i®
AIX®	NUMA-Q®	System p®
BladeCenter®	Power Systems™	System Storage™
DB2®	PowerHA™	System Storage DS®
DS4000®	PowerVM™	System x®
DS8000®	Redbooks®	Tivoli®
FlashCopy®	Redbooks (logo)  ®	TotalStorage®
HACMP™	Solid®	XIV®

The following terms are trademarks of other companies:

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

AMD, the AMD Arrow logo, and combinations thereof, are trademarks of Advanced Micro Devices, Inc.

Emulex, HBAnyware, SBOD, and the Emulex logo are trademarks or registered trademarks of Emulex Corporation.

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

QLogic, SANsurfer, and the QLogic logo are registered trademarks of QLogic Corporation. SANblade is a registered trademark in the United States.

ACS, Red Hat, and the Shadowman logo are trademarks or registered trademarks of Red Hat, Inc. in the U.S. 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.

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

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

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

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 consolidates, in one document, detailed descriptions of the hardware configurations and options offered as part of the IBM Midrange System Storage™ servers, which include the IBM System Storage DS4000® and DS5000 families of products.

This edition covers updates and additional functions available with the IBM System Storage DS® Storage Manager Version 10.60 (firmware level 7.60). This book presents the concepts and functions used in planning and managing the storage servers, such as multipathing and path failover. The book offers a step-by-step guide to using the Storage Manager to create arrays, logical drives, and other basic (as well as advanced) management tasks.

This publication also contains practical information about diagnostics and troubleshooting, and includes practical examples of how to use scripts and the command-line interface.

This publication is intended for customers, IBM Business Partners, and IBM technical professionals who want to learn more about the capabilities and advanced functions of the DS4000 series of storage servers with Storage Manager Software V10.60. It also targets those who have a DS4000 and DS5000 storage subsystem and need detailed advice about how to configure it.

This publication is designed specifically to address the hardware features and configuration of the IBM Midrange System Storage servers and can be used in conjunction with the following IBM Midrange System Storage Redbooks publications:

- ▶ *IBM Midrange System Storage Implementation and Best Practices Guide*, SG24-6363
- ▶ *IBM Midrange System Storage Copy Services Guide*, SG24-7822

The team who wrote this book

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

Sangam Racherla is an IT Specialist and Project Leader working at the International Technical Support Organization, San Jose Center. He holds a degree in electronics and communication engineering and has nine years of experience in the IT field. He has been with the International Technical Support Organization for the past six years and has extensive experience installing and supporting the ITSO lab equipment for various IBM Redbooks publications projects. His areas of expertise include Microsoft® Windows®, Linux®, AIX®, IBM System x® and IBM System p® servers, and various SAN and storage products.

Bruce Allworth is a Senior IT Specialist working in IBM Americas Storage Advanced Technical Support (ATS). He is a Subject Matter Expert and the ATS Team Leader for the DS5000, DS4000, and DS3000 product lines. He has many years of experience with these products, including management, solution design, advanced problem determination, and disaster recovery. He works closely with various IBM divisions and LSI in launching new products, creating critical documentation, including Technical and Delivery Assessment Checklists, and developing and delivering technical training for a wide range of audiences.

Alessio Bagnaresi is a Senior Solution Architect and Technical Sales Manager at Infracom, a major IBM Business Partner in Italy. Currently, he is working on customer assessments and proof of concepts about desktop, server, and storage virtualization, consolidation, infrastructure optimization, and platform management. He is certified on several platforms, such as AIX, Linux, VMware, Citrix, Xen, IBM Tivoli® software, and IBM Enterprise System Storage products. His job includes the planning, design, and delivery of Platform Management, Business Continuity, Disaster Recovery, Backup/Restore, and Storage/Server/Desktop Virtualization solutions involving IBM Director, IBM System p, IBM System x, and IBM System Storage platforms (mostly covering IBM San Volume Controller, IBM DS4000/DS5000 Midrange Storage Server, IBM DS8000® Enterprise Storage, and IBM N series). He previously worked at IBM as a Cross-Brand System Architect. He worked in customer projects about Server Consolidation (PowerVM™, VMware, Hyper-V, and Xen), Business Continuity (DS8000 Advanced Copy Services, Power HA XD, AIX Cross-site Mirroring, DB2® High Availability and Disaster Recovery, and DS4000/DS5000 Enhanced Remote Mirror), Disaster Recovery (TSM DRM, ProtecTier TS7650G), and Storage Virtualization (SVC and N series).

Chris Bogdanowicz has over 20 years of experience in the IT industry. He joined Sequent Computer Systems 15 years ago, initially specializing in symmetric multiprocessing UNIX® platforms and later NUMA-Q® technology. He remained in a support role when Sequent merged with IBM in 1999. He is currently a member of the IBM MTS SAN and midrange storage hardware support team in the UK. In addition, he is part of a Virtual EMEA Team (VET), providing Level 2 support for DS4000 and DS5000 products in Europe. He also maintains a keen interest in performance and configuration issues through participation in the Storage Solution Expert (SSE) program.

Corne Lottering is a Systems Storage Sales Specialist in the IBM Sub-Saharan Africa Growth Market Region for Systems and Technology Group. His primary focus is sales in the Central African countries, but also provides pre-sales support to the IBM Business Partner community across Africa. He has been with IBM for nine years and has experience in a wide variety of storage technologies, including the IBM System Storage DS4000, DS5000, DS8000, and XIV®, and IBM SAN switches, IBM Tape Systems, and storage software. Since joining IBM, he has been responsible for various implementation and support projects for customers across Africa.

Pablo Pedrazas is a Hardware Specialist working with Power Servers and Storage Products at IBM Argentina Support Center, doing post-sales second level support for Spanish Speaking Latin American countries in the Maintenance & Technical Support Organization. He has 21 years of experience in the IT industry, with expertise in UNIX servers and storage products. He holds a Bachelor's degree in Computer Science and a Master of Science degree in Information Technology and Telecommunications Management from the EOI of Madrid.

Frank Schubert is an IBM Certified Systems Expert and Education Specialist for DS4000 family of storage systems. He works for IBM Global Technology Services (GTS) in the Technical Education and Competence Center (TECC) in Mainz, Germany. His focus is on deploying education and training IBM service personnel in EMEA to maintain, service, and implement IBM storage products, such as the DS4000, DS5000, and IBM N series. He has been with IBM for the last 14 years and has gained storage experience since 2003 in different support roles.

John Sexton is a Certified Consulting IT Specialist, based in Auckland, New Zealand, and has over 20 years of experience working in IT. He has worked at IBM for the last 13 years. His areas of expertise include IBM System p, AIX, HACMP™, virtualization, storage, IBM Tivoli Storage Manager, SAN, SVC, and business continuity. He provides pre-sales support and technical services for clients throughout New Zealand, including consulting, solution implementation, troubleshooting, performance monitoring, system migration, and training. Prior to joining IBM in New Zealand, John worked in the United Kingdom supporting and maintaining systems in the financial and advertising industries.

Alexander Watson is a Senior IT Specialist for Storage ATS Americas in the United States. He is a Subject Matter Expert on SAN switches and DS4000 products. He has over ten years of experience in planning, managing, designing, implementing, analyzing, and tuning of SAN environments. He has worked at IBM for ten years. His areas of expertise include SAN fabric networking, Open System Storage I/O, and the IBM Midrange Storage Subsystems family of products.

The authors want to express their thanks to the following people, whose expertise and support were integral to the writing of this book:

Jed Bless
Brian Steffler
Brocade Communications Systems, Inc.

Shawn Andrews
Mark Brougher
Mark S. Fleming
Paul Goetz
Harsha Gunatilaka
Richard Hutzler
Doris Konieczny
Harold Pike
Scott Rainwater
Michael D Roll
Pete Urbisci
Bill Wilson
IBM

Brad Breault
Stacey Dershem
Ryan Leonard
Amanda Ryan
David Worley
LSI Corporation

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

Bertrand Dufrasne
Ann Lund
Alex Osuna
Jon Tate
International Technical Support Organization, San Jose Center

Thanks to the authors of the previous edition of this book.

Now you can become a published author, too!

Here's an opportunity to spotlight your skills, grow your career, and become a published author - all at the same time! Join an ITSO residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

Find out more about the residency program, browse the residency index, and apply online at: ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

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

- ▶ Use the online **Contact us** review Redbooks form found at:
ibm.com/redbooks
- ▶ Send your comments in an 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

Stay connected to IBM Redbooks

- ▶ Find us on Facebook:
<http://www.facebook.com/pages/IBM-Redbooks/178023492563?ref=ts>
- ▶ Follow us on twitter:
<http://twitter.com/ibmredbooks>
- ▶ Look for us on LinkedIn:
<http://www.linkedin.com/groups?home=&gid=2130806>
- ▶ Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks weekly newsletter:
<https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm>
- ▶ Stay current on recent Redbooks publications with RSS Feeds:
<http://www.redbooks.ibm.com/rss.html>

Summary of changes

This section describes the technical changes made in this edition of the book and in previous editions. This edition might also include minor corrections and editorial changes that are not identified.

Summary of Changes
for SG24-7676-01
for IBM Midrange System Storage Hardware Guide
as created or updated on March 16, 2010.

March 2010, Second Edition

This revision reflects the addition, deletion, or modification of new and changed information described below.

New information

- ▶ Controller firmware Version 7.60
- ▶ IBM System Storage DS5020 and EXP520
- ▶ Full Disk Encryption (FDE) drives
- ▶ Solid® State Drive (SSD)
- ▶ iSCSI host interface on the DS5000 models
- ▶ 8 Gbps FC host interface
- ▶ Remote Support Manager Hardware Model - RS2

Changed information

- ▶ Host Port Upgrade for DS5100
- ▶ Cache Upgrade option for DS5100 and DS5300
- ▶ 448 Drive support for DS5100 and DS5300

Archived



Introduction to IBM Midrange System Storage storage subsystems

This chapter introduces the IBM Midrange System Storage storage subsystems and positions them within the overall IBM System Storage Disk Systems (DS) family. The IBM Midrange System Storage storage subsystems include the IBM System Storage DS4000 and IBM System Storage DS5000 models. The current hardware models are briefly described, as well as the Storage Manager software. Detailed information is given in subsequent chapters of this book.

1.1 Positioning the DS4000/DS5000 series

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 right solutions for their needs. The DS family combines the high-performance IBM System Storage DS6000/DS8000 series of enterprise servers with the IBM System Storage DS4000 and DS5000 series of midrange systems, and other line-of-entry systems (IBM System Storage DS3000 series).

The IBM System Storage DS4000/DS5000 series is composed of products that fit different requirements in terms of performance and scalability, and are ready for multiple environments ranging from departmental to bandwidth-intensive and transaction-heavy. Moreover, these products are designed for business continuity and high availability, are ready for the challenges of IT Optimization (Consolidation, Virtualization, and Adaptability), and are designed for a longer life cycle with investment protection.

The IBM System Storage DS4000/DS5000 series of disk storage systems that this IBM Redbooks publication addresses is the IBM solution for midrange/departmental storage requirements and controlling change through adaptability. The positioning of the products within the Midrange DS4000/DS5000 series is shown in Figure 1-1.

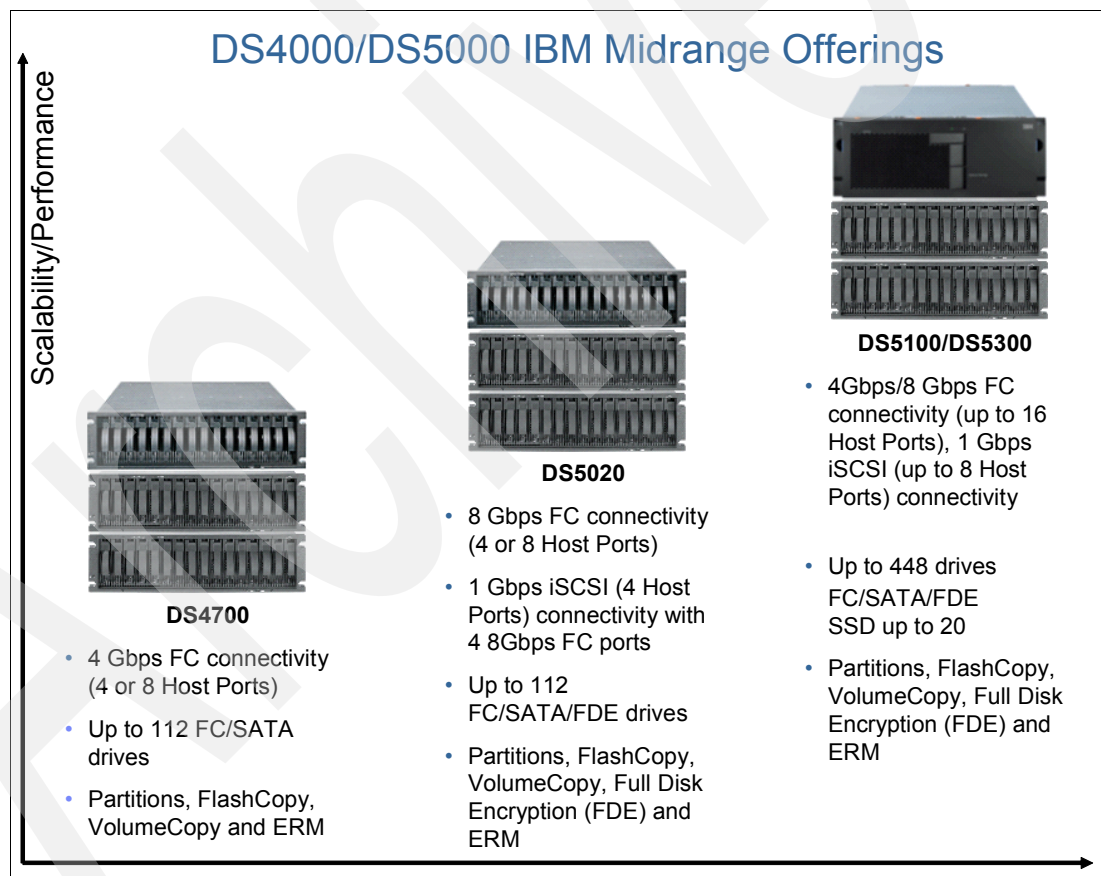


Figure 1-1 Product positioning within the IBM Midrange System Storage DS4000/DS5000 series

Note: In this publication, we only discuss the DS4700 model. More information about the older models of the DS4000 series can be found in *IBM System Storage DS4000 and Storage Manager V10.30*, SG24-7010.

The overall positioning of the DS4000/DS5000 series within the IBM System Storage DS family is shown in Figure 1-2. It expands the IBM Midrange System Storage offering in terms of performance and scalability.

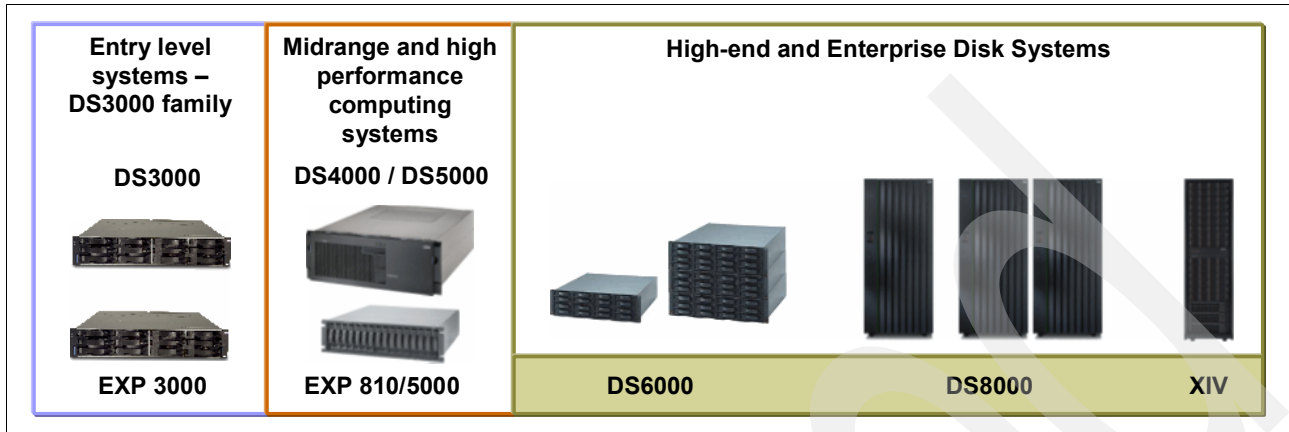


Figure 1-2 DS4000/DS5000 series positioning within the IBM System Storage family

Within the IBM Midrange storage series, the DS5000 models of storage servers support Fibre Channel (FC), Serial ATA (SATA), Full Disk Encryption (FDE) disk drives, and Solid State Drives (SSD), and the DS4000 models supports both FC and SATA disk drives.

In terms of capacity, the maximum raw SATA storage capacity is 448 TB, using 1TB SATA DDM drives. The maximum raw FC storage capacity is 268.8 TB (using 600 GB 15K 4 Gbps FC DDM).

1.2 IBM Midrange System Storage storage server models

The IBM Midrange System Storage series of storage servers (DS4000 and DS5000) uses Redundant Array of Independent Disks (RAID) technology. RAID technology is used to protect the user data from disk drive failures. DS storage subsystems contain Fibre Channel (FC) interfaces to connect the disk drive enclosures, which contain iSCSI or Fibre Channel (FC) interfaces to connect the host systems.

The storage servers in the DS5000 series provide high system availability through the use of hot-swappable and redundant components. This is crucial when the storage server is placed in high-end customer environments, such as server consolidation on Storage Area Networks (SANs).

At the time of writing, the IBM Midrange System Storage storage series is composed of five products that are available in specific models. These products are named DS4700, DS5020, DS5100, and DS5300, which offers balanced performance, and linear IOPS scalability supports workloads ranging from departmental to bandwidth-intensive.

In terms of intermixing among different hard disk drive technologies, we have the following supported configurations:

- ▶ DS5100 and DS5300 storage servers support the intermixing of high performance FC, high capacity SATA drives, Full Disk Encryption (FDE), and Solid State Drives (SSD) within single expansion units.
- ▶ DS5020 storage servers support the intermixing of high performance FC, high capacity SATA drives, Full Disk Encryption (FDE), and Solid State Drives (SSD) within single expansion units and even within the controller enclosure as well.
- ▶ DS4700 storage servers support the intermixing of high performance FC and high capacity SATA drives within single expansion units and even within the controller enclosure as well.

Currently, the DS4000 and the DS5000 series supports host connectivity through the following interfaces:

- ▶ 4 Gbps FC host ports (DS4700, DS5100, and DS5300)
- ▶ 8 Gbps FC host ports (DS5020, DS5100, and DS5300)
- ▶ 1 Gbps iSCSI host ports (DS5020, DS5100, and DS5300)

We briefly describe the characteristics of the four products previously mentioned:

- ▶ IBM System Storage DS4700 server

The DS4700 storage server is targeted at entry-level to mid-level customers. It can hold a maximum of 16 disk drives inside the storage server 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, and the 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, helps protect investments, is a higher performance storage server for open systems.

- ▶ IBM System Storage DS5020 server

The DS5020 is the newest member of the DS5000 series and is designed to help address midrange or departmental storage requirements. The DS5020 has a 3U rack-mountable enclosure, has four 4 Gbps FC drive interfaces, and can consist of a maximum of six EXP520 expansion units for a total of up to 112 disk drives. Through a specific activation feature, six EXP810 expansions can be used in place of the EXP520s.

The DS5020 can be configured with 2 or 4 GB of cache memory and different host connectivity options as listed below:

- Two 8 Gbps FC host ports on each of its two controllers
- Four 8 Gbps FC host ports on each of its two controllers
- Two 8 Gbps FC host ports and, optionally, two 1 Gbps iSCSI on each of its two controllers

At the time of writing one model is available:

- Model 20A with 8 GB or 16 GB of cache memory and one of the combinations of host interfaces specified in the list above.

► IBM System Storage DS5100 server

The DS5100 is targeted at high-end customers. This storage subsystem is a 4U rack-mountable enclosure, has sixteen 4 Gbps FC drive interfaces, and can hold a maximum of twenty-eight EXP5000 expansion units, a maximum of twenty-eight EXP810 expansion units or, for migration purposes, up to twenty-eight expansion units composed of a mix of EXP5000 and EXP810 for a total of up to 448 disk drives.

The DS5100 can have up to 32 GB of cache memory and different host connectivity options as listed below:

- Four 4 Gbps FC host ports on each of its two controllers
- Four 8 Gbps FC host ports on each of its two controllers
- Two 1 Gbps iSCSI host ports on each of its two controllers
- Eight 4 Gbps FC host ports on each of its two controllers
- Eight 8 Gbps FC host ports on each of its two controllers
- Four 8 Gbps FC host ports and Two 1 Gbps iSCSI host ports on each of its two controllers (waiting for DST)
- Four 1 Gbps iSCSI host ports on each of its two controllers

At the time of writing one model is available:

- Model 51A with 8 GB, 16 GB, or 32 GB of cache memory and one of the combinations of host interfaces specified in the list above.

► IBM System Storage DS5300 server

The DS5300 server has greater scalability than the DS5100. This storage subsystem is a 4U rack-mountable enclosure, has sixteen 4 Gbps FC drive interfaces, and can hold a maximum of twenty-eight EXP5000 expansion units, a maximum of twenty-eight EXP810 expansion units, or, for migration purposes, up to twenty-eight expansion units composed of a mix of EXP5000 and EXP810 for a total of up to 448 disk drives. It is designed to deliver data throughput of up to 400 MBps per drive port.

The DS5300 can mount up to 64 GB of cache memory and different host connectivity options as listed below:

- Four 4 Gbps FC host ports on each of its two controllers
- Four 8 Gbps FC host ports on each of its two controllers
- Four 1 Gbps iSCSI host ports on each of its two controllers
- Eight 4 Gbps FC host ports on each of its two controllers
- Eight 8 Gbps FC host ports on each of its two controllers
- Four 8 Gbps FC host ports and four 1 Gbps iSCSI host ports on each of its two controllers
- Eight 1 Gbps iSCSI host ports on each of its two controllers

One model is available:

- Model 53A with 8 GB, 16 GB, 32 GB, or 64 GB of cache memory and one of the combinations of host interfaces specified in the list above.

1.3 IBM Midrange System Storage expansion enclosure

At the time of writing, the IBM Midrange System Storage expansion enclosures offer a 4 Gbps FC interface, and four models are available:

- ▶ EXP810 Expansion Enclosure

This expansion unit is packaged in a 3U rack-mountable enclosure, supports up to 16 FC disk drives or E-DMM SATA drives. It contains 16 drive bays, dual-switched 4 Gbps ESMS, and dual power supplies and cooling components. Fully populated with 450 GB FC disk drive modules, this enclosure offers up to 7.2 TB of raw storage capacity or up to 16 TB when populated with the 1000 GB E-DDM SATA drives. The EXP810 expansion unit is the only one that may be connected to every storage subsystem of the DS4000/DS5000 family. Through the proper firmware level, this expansion unit is able to host both FC and SATA Drives. Intermix of FC and SATA drives is supported within this expansion enclosure.

- ▶ EXP5000 Expansion Enclosure

This expansion unit is packaged in a 3U rack-mountable enclosure, supports up to 16 FC disk drives, E-DMM SATA drives, Full Disk Encryption (FDE) drives, and up to 20 SSD. It contains 16 drive bays, dual-switched 4 Gbps ESMS, and dual power supplies and cooling components. Fully populated with 600 GB FC disk drive modules, this enclosure offers up to 9.6 TB of raw storage capacity or up to 16 TB when populated with the 1000 GB E-DDM SATA drives. The EXP5000 expansion unit may be connected to the DS5100 or DS5300 storage server. Through the proper firmware level, this expansion unit is able to host both FDE, FC, SATA drives, and SSD as well. An intermix of FC, SATA, FDE, and SSD drives is supported within this expansion enclosure.

- ▶ EXP520 Expansion Enclosure

This expansion unit is packaged in a 3U rack-mountable enclosure, supports up to 16 FC disk drives, E-DMM SATA drives, or Full Disk Encryption (FDE) drives. It contains 16 drive bays, dual-switched 4 Gbps ESMS, and dual power supplies and cooling components. Fully populated with 600 GB FC disk drive modules, this enclosure offers up to 9.6 TB of raw storage capacity or up to 16 TB when populated with the 1000 GB E-DDM SATA drives. The EXP520 expansion unit may be connected to the DS5020 storage server. Through the proper firmware level, this expansion unit is able to host both FDE, FC, and SATA drives. An intermix of FC, SATA, and FDE drives is supported within this expansion enclosure.

- ▶ EXP5060 Expansion Enclosure

The IBM System Storage EXP5060 storage expansion enclosure provides high-capacity SATA disk storage for the DS5100 and DS5300 storage subsystems. The storage expansion enclosure provides continuous, reliable service, using hot-swap technology for easy replacement without shutting down the system and supports redundant, dual-loop configurations. External cables and Small Form-Factor Pluggable (SFP) modules connect the DS5100 or DS5300 storage subsystem to the EXP5060 storage expansion enclosure. The EXP5060 uses redundant 4 Gbps Fibre Channel connections to make connections to the DS5100 or DS5300 storage subsystem, and another EXP5060 storage expansion enclosure in a cascading cabling configuration, offering reliability and performance.

Note: A maximum of eight EXP5060 storage expansion enclosures (with 480 hard drives) can be attached only to the DS5100 and DS5300 storage subsystems and only SATA disks are supported in the EXP5060 expansion.

The EXP5060 is a 4U rack-mountable enclosure that supports up to 60 SATA Disk Drive Modules (DDMs), offering up to 60 TB of SATA disk space per enclosure using 1 TB SATA DDMs. The expansion enclosure contains 60 drive bays (arranged on five stacked drawers with twelve drives for each drawer), dual-switched 4 Gbps ESMs, and dual power supplies and cooling components. Coupled with a storage subsystem (DS5100 or DS5300), you can configure RAID-protected storage solutions of up to 480 TB when using 1 TB SATA DDMs and eight EXP5060 storage expansion enclosures, providing economical and scalable storage for your rapidly growing application needs. The Attach up to 8 EXP5060s feature pack must be purchased for the DS5100/DS5300 storage subsystem to enable it to be connected to up to eight EXP5060 storage expansion enclosures.

1.4 IBM System Storage DS Storage Manager software

The IBM System Storage DS Storage Manager software (see Figure 1-3) is used to configure, manage, and troubleshoot the DS4000 and DS5000 storage subsystems.

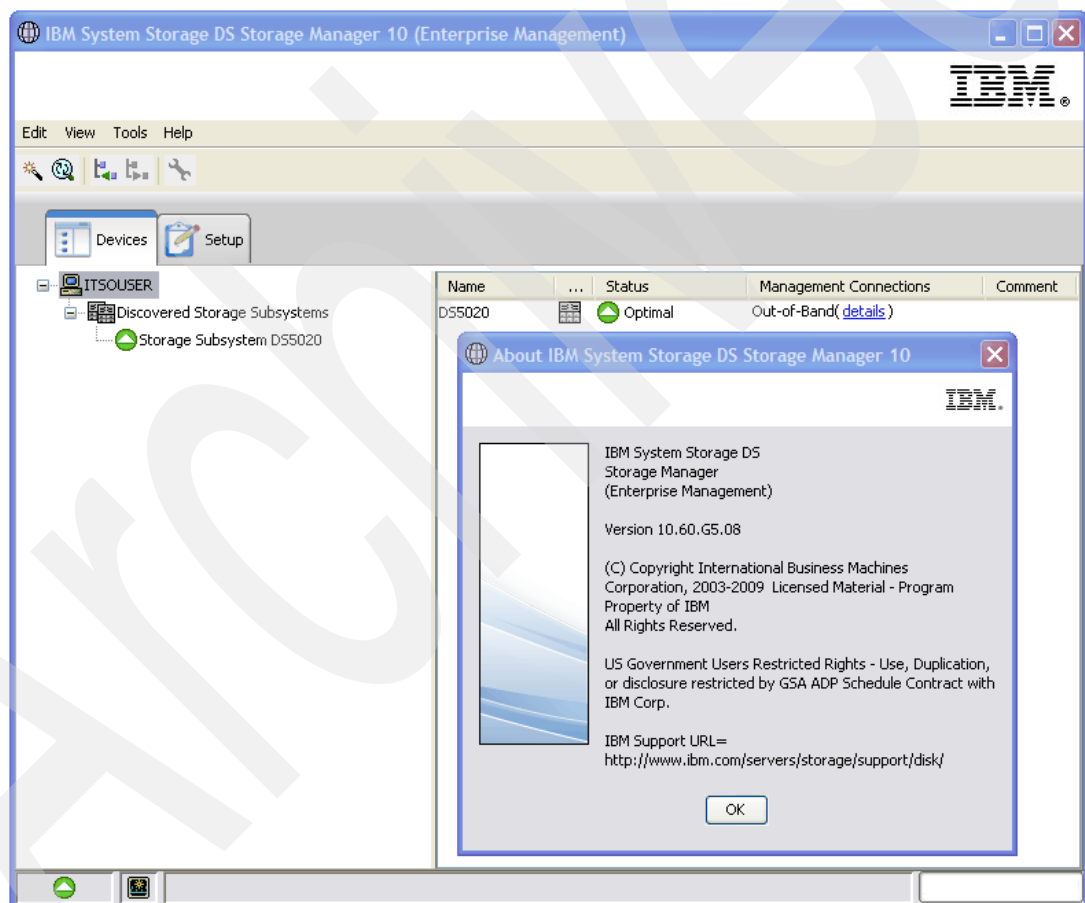


Figure 1-3 Storage Manager

It 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 for troubleshooting and management tasks, such as 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, Enhanced Remote Mirroring, and Disk Encryption.

The Storage Manager software is now packaged as the following combinations:

► *Host-based software*

– Storage Manager Client (SMclient)

The SMclient component provides the graphical user interface (GUI) for managing storage systems through the Ethernet network or from the host computer.

– Storage Manager Runtime (SMruntime)

The SMruntime is a Java™ runtime environment that is required for the SMclient to function. It is not available on every platform as a separate package, but in those cases, it has been bundled into the SMclient package.

– Storage Manager Agent (SMagent)

The SMagent package is an optional component that allows in-band management of the DS4000 and DS5000 storage subsystems.

– Storage Manager Utilities (SMutil)

The Storage Manager Utilities package contains command-line tools for making logical drives available to the operating system.

– Failover driver support

Storage Manager host based software includes an optional failover driver. It is a multipath driver built on MPIO technology.

– Storage Management Initiative - Specification Provider (SMI-S Provider)

An SMI-S provider is a vendor-specific module that is used so that independent management software, such as IBM TotalStorage® Productivity Center (TPC), can manage a vendor device using a standard interface based on the Common Information Model (CIM) protocol

► *Controller-based software*

– DS4000 and DS5000 storage subsystem controller firmware and NVSRAM

The controller firmware and NVSRAM are always installed as a pair and provide the “intelligence” of the Midrange System Storage server.

– DS4000 and DS5000 storage subsystem Environmental Service Modules (ESM) firmware

The ESM firmware controls the interface between the controller and the drives.

– DS4000 and DS5000 storage subsystem drive firmware:

The drive firmware is the software that instructs the Fibre Channel (FC) drives about how to operate on the FC loop.

The Storage Manager functions are reviewed in detail in 4.9, “Step-by-step configuration” on page 175.

1.5 IBM Midrange System Storage hard disk drives

Every storage subsystem within the IBM Midrange System Storage series is a multi-tiered storage server that is capable of supporting multiple disk drive technologies even within the single expansion unit. This capability enables you to have a Total Cost of Ownership (TCO) reduction and a storage consolidation by putting together different workloads with different performance requirements. Figure 1-4 shows a possible hard disk drive (HDD) layout within four expansion units.

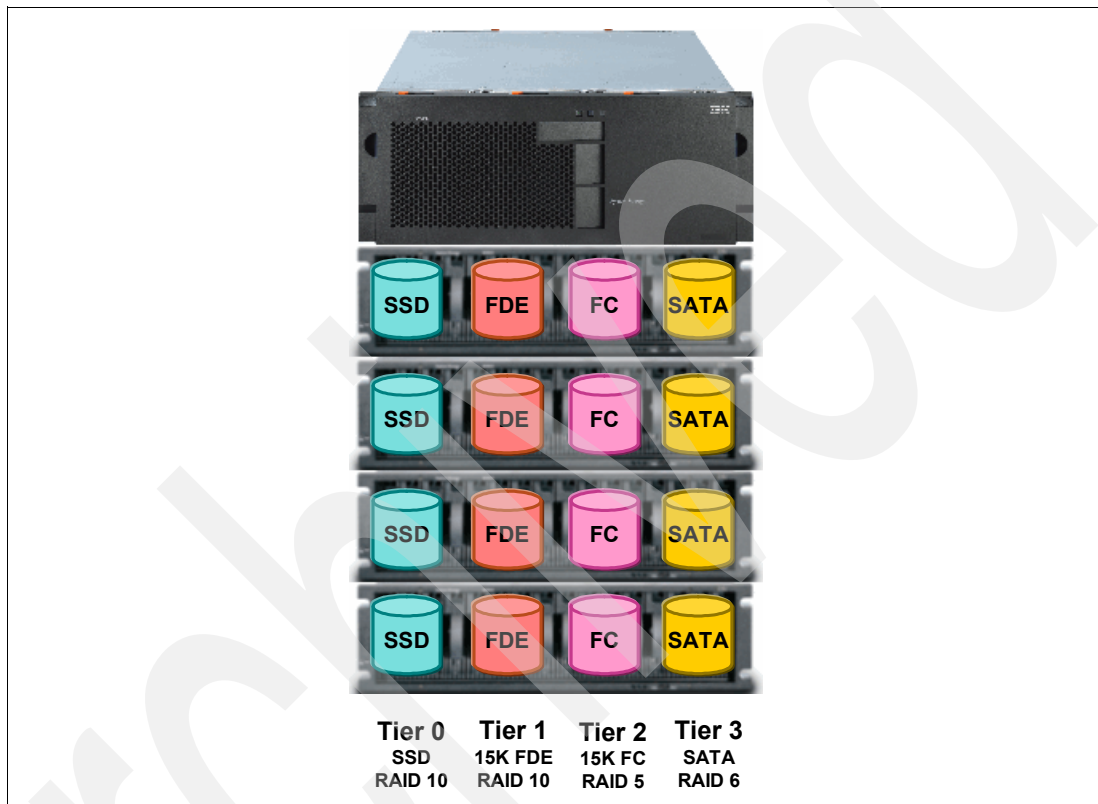


Figure 1-4 Multi-Tier HDD layout

The currently supported HDD technologies are briefly described below:

► **Solid State Drives (SSDs)**

Solid State Drives use semiconductor devices (solid state memory) to store data and have no moving parts. An SSD is a flash memory device mimicking a disk drive. SSDs are available with the same interfaces used by hard disk drives (for example, SAS, SATA, and Fibre Channel) and they are packaged in the same form factors as hard disk drives (3.5-in., 2.5-in., and 1.8-in.). SSDs are designed to plug into the same environments as those supported by hard disk drives. As already mentioned, the simple fact that there are no moving parts (disk platters, magnetic heads, or motor) in an SSD results in:

- **Faster data access and throughput:** The access to the data with an SSD is faster because again, there is no read/write head to move and no magnetic platters need to spin up (no latency). On an SSD, the data can be read almost immediately.
- **Better reliability:** Again, the lack of moving and rotating parts almost eliminates the risk of mechanical failure. SSDs have the ability to tolerate extreme shocks, higher altitudes, vibration, and extremes of temperature. However, they can still fail and must be RAID protected like traditional drives.

- Less power consumption: Because there is no power for the motor required to spin up the magnetic platters and to move the heads, the drive uses less energy than a traditional hard disk drive. Each SSD uses about half of the power of a 15K RPM HDD. The savings can be substantial if a few SSDs can replace many HDDs to deliver the same performance. This is particularly true for applications that were forced, for performance reasons, to use large quantities of HDDs to get as many spindles as they can, though they were only using a small portion of the disk capacity. Besides power consumption savings, the overall system will weigh much less because SSDs are already much lighter than HDDs.

SDD is supported in an EXP5000 expansion unit behind a DS5100 or a DS5300 storage server. At the time of writing, one SSD model with 73 GB of capacity is supported.

► Full Disk Encryption Hard Disk Drives (FDE HDD)

FDE is a technology that performs encryption on the hard disk drive at the hardware level. The Fibre Channel hard drive contains a custom chip or application specific integrated circuit (ASIC) that is used to encrypt every bit of data as it is written and also decrypts data as it is being read. ASIC requires a security key to allow encryption and decryption to begin.

FDE disk drives encrypt all the data on the disk. The secured drive requires that a security key be supplied before read or write operations can occur. The encryption and decryption of data is processed entirely by the drive and is not apparent to the storage subsystem.

FDE HDDs are supported on DS5000 models (DS5020, DS5100, and DS5300).

The FDE HDDs that are currently supported include:

- 4 Gbps FC, 146.8 GB/15k rpm
- 4 Gbps FC, 300 GB/15k rpm
- 4 Gbps FC, 450 GB/15k rpm
- 4 Gbps FC, 600 GB/15k rpm

► Fibre Channel Hard Disk Drives (FC HDD)

Fibre Channel is an industry-standard interface that supports very high data rates as well as many more devices than traditional SCSI or ATA/IDE technologies. It is also a serial interface, and uses a loop topology. FC HDDs feature a full-duplex dual-port active/active 4 Gbps Fibre Channel interface. The DS4000 and DS5000 series storage servers can mount different types of FC HDDs in terms of size and revolutions per minute (RPM):

- 4 Gbps FC, 73.4 GB/15k rpm (only DS4000 models)
- 4 Gbps FC, 146.8 GB/15k rpm
- 4 Gbps FC, 300 GB/15k rpm
- 4 Gbps FC, 450 GB/15k rpm
- 4 Gbps FC, 600 GB/15k rpm

► Serial ATA Hard Disk Drives (SATA HDD)

Serial ATA is the hard disk standard created to replace the parallel ATA interface (also called IDE). Serial ATA transmits data in serial mode and implements two separated datapaths, one for transmitting and another for receiving data. The standard transfer rate is 1.5 Gbps for Serial ATA standard, although Serial ATA II (second generation) provides new features, such as Native Command Queuing (NCQ), plus a higher speed rate of 3 Gbps. NCQ increases the hard disk drive performance by reordering the commands send by the host. Through the speed-matching technology provided by the expansion enclosures mentioned in 1.3, “IBM Midrange System Storage expansion enclosure” on page 6, SATA hard disk drives are able to work with a transfer rate of 4 Gbps, enabling the intermixing of high performance FC and high capacity SATA drives.

The DS4000 and DS5000 series can mount different types of FC HDDs in terms of size and revolutions per minute (RPM):

- 500 GB/7.2k rpm SATA
- 750 GB/7.2k rpm SATA II (all the models)
- 1000 GB/7.2k rpm SATA II (all the models)
- 2000 GB/7.2k rpm SATA II (only EXP5000)

1.6 iSCSI basics

All the DS5000 series models now support iSCSI host connectivity and in this section we briefly describe the basics of the iSCSI protocol. Consult Appendix B, “Deploying iSCSI host interface controllers on the IBM System Storage DS5000 series” on page 515 for more detailed information.

iSCSI is an industry standard developed to enable transmission of SCSI block commands over the existing IP network by using the TCP/IP protocol. The TCP/IP protocol provides iSCSI with inherent reliability along with byte by byte, in order delivery, built-in security, and no interoperability barriers.

When a user or application sends a request, the operating system generates the appropriate SCSI commands and data request, which then go through encapsulation. A packet header is added before the resulting IP packets are transmitted over an Ethernet connection. When a packet is received, it is disassembled, separating the SCSI commands and request. The SCSI commands are sent on to the SCSI controller, and from there to the SCSI storage device. Because iSCSI is bi-directional, the protocol can also be used to return data in response to the original request.

The logical link that carries the commands and data to and from TCP/IP endpoints is called an *iSCSI session*. A session is made up of at least one TCP/IP connection from an *initiator* (*host*) to a *target* (storage device), as shown in Figure 1-5. The iSCSI initiator can be either an iSCSI HBA inside a host server, or you can define a software iSCSI initiator by using an iSCSI stack and an Ethernet network adapter. An example of an iSCSI stack is the Microsoft iSCSI Software Initiator, which runs on Windows Server 2003 and Windows Server 2008.

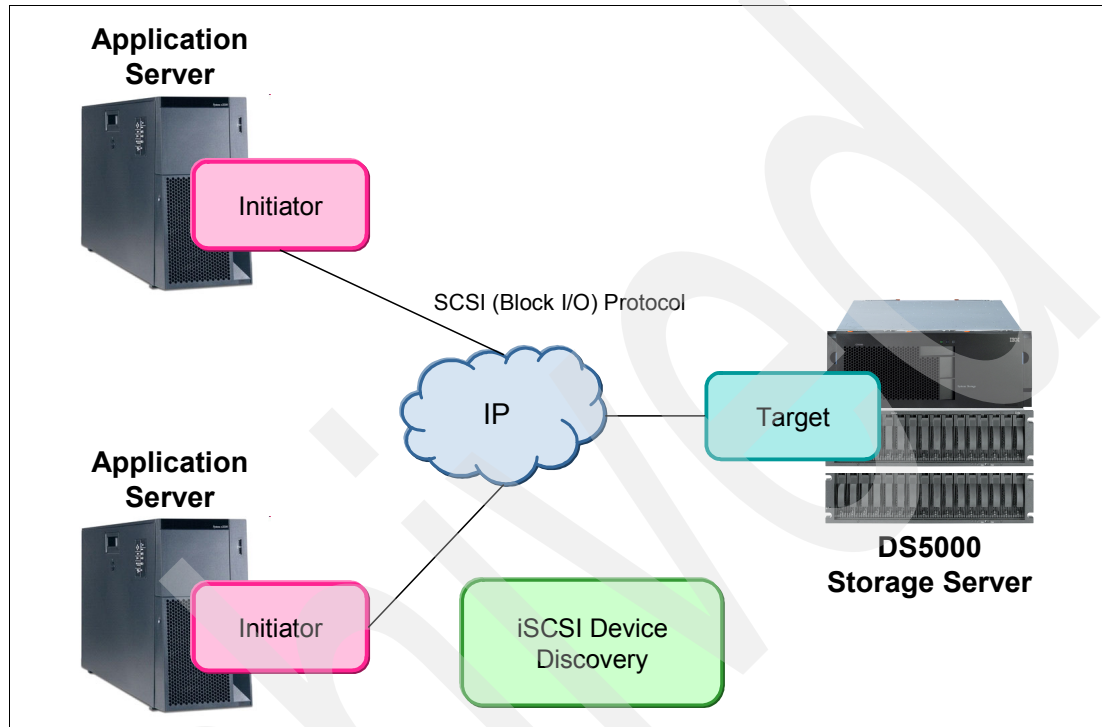


Figure 1-5 iSCSI components

iSCSI SANs might not be a solution for everyone, given that iSCSI SANs might not deliver the same performance that Fibre Channel SANs can deliver. However, there are ways to address iSCSI performance:

- ▶ While the host servers can use almost any Ethernet network interface card for iSCSI traffic, this does mean that the CPUs on the host server have to run the iSCSI stack (to perform encapsulation of SCSI commands and data). This causes increased CPU and memory processing, which can impact performance. For increased performance, it is better to use dedicated iSCSI HBAs to process the TCP/IP stack. This technology is known as TCP Offload Engine (TOE). TOE technology relieves the CPUs on the host server from having to process the SCSI encapsulation, which can lead to better performance.
- ▶ To avoid excessive traffic, but mainly to ensure that only valid initiators connect to storage servers, it is better to run iSCSI over dedicated network segments or virtual LANs (VLAN).

1.7 Fibre Channel direct/switch basics

Fibre Channel (FC) is a high-speed disk attachment technology, designed to connect a large number of storage devices to a number of host servers across a Storage Area Network (SAN). Fibre Channel Protocol (FCP) transfers SCSI commands and data across physical FC links.

FC supports a much higher number of devices and much longer cable lengths than SCSI. It has become the preferred disk attachment technology in midrange and large scale data center solutions.

At the time of writing, the IBM Midrange System Storage series' maximum FC throughput is 8 Gbps; 10 Gbps links can be used today, but only for SAN switch interconnection.

Host servers contain one or more FC Host Bus Adapters (HBA). The HBAs provide connectivity to the storage devices using FC cabling and SAN Switch.

For more information about Fibre Channel and SANs, see *Introduction to Storage Area Networks*, SG24-5470.

FC topologies

FC supports several connectivity topologies:

- ▶ Point-to-point

This is the simplest topology and provides a direct link between an FC HBA inside a host server and a storage device.

- ▶ Arbitrated loop

This topology can be used to interconnect several FC devices. A typical example is to attach a certain number of host servers to an FC storage subsystem. A loop can consist of up to 126 devices.

Devices on the loop use one-way ring communication. In any given moment, only two devices on the loop can communicate. This means the devices share bandwidth, so the arbitrated loop topology is not suitable for high performance requirements.

Arbitrated loops were commonly implemented with the use of an FC hub. Even though this is physically a star topology, logically it will be a loop. Alternatively, devices can be connected in a daisy chain manner. Arbitrated loops are rarely seen these days, as switched fabrics have become the norm.

- ▶ Switched fabric

The most commonly used topology in a typical SAN today is switched fabric. SAN switches are used to provide FC connectivity between the host servers and storage devices. Switched fabrics can become very complex in large scenarios, connecting hundreds of host servers to a very large number of storage subsystems.

SAN switches provide optimized traffic flow and increased performance by allowing concurrent data transfers between many connected hosts and storage devices. Switched fabrics can provide dedicated bandwidth, as opposed to arbitrated loop technology, where the bandwidth is shared among all the devices in the loop.

FC protocol layers

The FC protocol is split into five layers, named FC0 through FC4. Let us look briefly at them:

- ▶ FC0 is the physical layer, which describes cabling, connectors, signalling, and so on. This layer defines the physical media implementation.
- ▶ FC1 is the data link layer. This layer contains the 8b/10b encoding and decoding of signals for transmission across the physical media.
- ▶ FC2 is the network layer and defines the main FC protocols. This layer defines how the frames are transferred.
- ▶ FC3 is the common services layer. This layer provides services such as multi-casting and striping.

- ▶ FC4 is the application protocol mapping layer. In storage connectivity applications, FCP protocol is used to encapsulate SCSI data into FC frames.

FC cable types

FC implementations can utilize either single-mode or multi-mode FC cables. The name multi-mode fiber indicates that multiple modes, or rays of light, can travel through the cable core simultaneously. The multi-mode fiber cable uses a larger diameter core, which makes it easier to couple than the single-mode fibre cable. With a throughput of 8 Gbps, the length of the cable can be up to 150 m.

Single-mode fiber transfers a single ray of light. The core diameter is much smaller than the core of multi-mode cable. Therefore, coupling is much more demanding and tolerances for single-mode connectors and splices are very low. However, single-mode fiber cables can be much longer. The cable length can exceed 50 km.

Multi-mode cabling is much more common, as it is easier to work with and meets the requirements of most customer scenarios. However, in situations where very long cable lengths are needed, single-mode cabling will be required.

FC world wide names (WWN)

FC devices are presented with a unique identifier called world wide name (WWN). The WWNs are somewhat similar to the MAC addresses in Ethernet terms. For example, each FC HBA has its own WWN, which is hardcoded (or burned-in) during manufacturing. The HBA will be uniquely identified by the storage subsystem using its WWN.



New features

This chapter provides a brief description about new features that are part of controller firmware version (CFW) 7.60 and the IBM System Storage DS5020 announcement.

2.1 Full Disk Encryption capable disk drive modules (DDM)

Disk Security is a new feature introduced with the DS5000 that complements the newly available Full Disk Encryption (FDE) drives. It is supported by the latest level of DS5000 firmware (Version 7.60) and Storage Manager V10.60. This new feature can add a greater level of security to data that resides on disk.

The Full Disk Encryption (FDE) disk is used in conjunction with IBM Disk Encryption Storage Manager on the DS5000 storage subsystem. IBM Disk Encryption Storage Manager will generate encryption and decryption keys that are used to lock each FDE drive so that all data that resides on a disk is fully encrypted.

See Chapter 5, “Disk Security with Full Disk Encryption drives” on page 259 for details about how FDE works and how to set it up.

2.2 Solid State Drive (SSD) module

The need for higher IOPS in enterprise storage led to a new storage type known as Solid State Drive (SSD). SSDs have existed for some time, and were built in the past by using volatile DRAM memory that was supplemented with battery backup. However, these were enormously expensive and were often up to 1,000 times the cost of a high-performance disk drive with an equivalent capacity.

Some applications require very high IOPS rates, and there are specific military or industrial use cases that benefit from insensitivity to shock and vibration. But the vast majority of applications cannot justify the extra cost, so SSDs have remained a rarity.

A new storage device type based on non-volatile flash memory is now available for enterprise workloads that require high IOPS. Flash memory is less than 10 percent of the cost of DRAM, and expected innovations should lower the cost by another order of magnitude in the next 2-3 years. By itself, however, flash memory is too slow to perform WRITE operations and has an unacceptably short life. Clever engineering has been applied to use flash memory in combination with DRAM and embedded software to create a device for enterprise applications, one that is interface and function compatible with an HDD, but has substantially higher IOPS performance. These devices made their appearance in 2008, and will be widely available as optional devices in disk arrays. Flash-based SSDs are still more costly than HDDs of equivalent capacity, but for high IOPS workloads, they can cost less.

The availability of a new type of storage device raises questions, including:

- ▶ What is the right use of the new devices?
- ▶ Are additional changes to the server, storage, and software stack required to see the full value at the application level?
- ▶ How do we place the right data on these (expensive) devices at the right time to achieve performance objectives without creating an undue burden of management?

IBM DS5000 offers a 76 GB SSD drive running on 4 Gbps Fibre Channel (FC) that is compatible with the EXP5000.

You can read more about SSD drives in Appendix C, “Solid State Drives on the IBM System Storage DS5000 series” on page 527.

2.3 600 GB FC disk drive module

The DS5000 offers a 600 GB FC enhanced disk drive module (E-DDM). The dual port FC DDM operates at 4 Gbps FC speed. It is designed to fit into the EXP810, EXP5000, and EXP520. However, the drive requires a certain enclosure service module (ESM) firmware and controller firmware (CFW) version to be recognized by the DS5000 subsystem. See the latest ESM readme to verify compatibility with your subsystem.

The drive is also available as an FDE Drive.

The DDM characteristics are:

- ▶ 600 GB capacity
- ▶ 15,000 RPM (15K)
- ▶ Hot swappable
- ▶ Dual FC interface to EXP

2.4 1 TB SATA enhanced disk drive module

The DS5000 offers a 1000 GB (1 TB) SATA enhanced disk drive module (E-DDM). The disk drive operates at 3 Gbps SATA speed. The ATA interposer card on the DDM converts the ATA protocol of the drive to the 2 or 4 Gbps Fibre Channel protocol, so that the DDM module fits in every EXP810, EXP5000, and EXP520. However, the drive requires a certain enclosure service module (ESM) firmware and controller firmware version (CFW) to be recognized by the DS5000 subsystem. See the latest ESM readme to verify the compatibility with your subsystem.

The DDM characteristics are:

- ▶ 1000 GB capacity
- ▶ 7200 RPM
- ▶ Hot swappable
- ▶ Dual FC interface to EXP
- ▶ 2 and 4 Gbps FC speed

2.5 IBM System Storage DS5020 storage subsystem and EXP520

The IBM System Storage DS5020 storage system (Machine Type 1814-20A) is designed to provide solutions to meet the needs of midrange/departmental storage requirements, delivering high performance, advanced function, high availability, modular and scalable storage capacity. It also provides SAN-attached 8 Gbps Fibre Channel (FC) and 1 Gbps iSCSI connectivity, and support for RAID levels 0, 1, 3, 5, and 6 up to over 49 terabytes (TB) when using 450 GB Fibre Channel hard drives and up to 112 TB when using 1 TB Serial Advanced Technology Attachment (SATA) Enhanced Disk Drive Modules (E-DDMs).

A 3U rack-mountable enclosure houses the DS5020 redundant, dual-active RAID controllers with either four Fibre Channel ports or two Fibre Channel and two iSCSI ports per controller. The DS5020 can be configured for the attachment of host servers and EXP520 and EXP810 storage expansion enclosures and up to 16 4 Gbps Fibre Channel or SATA E-DDMs. The base DS5020 storage subsystem controllers each have four Fibre Channel ports.

The DS5020 Machine Type is 1814-20A.

The IBM System Storage EXP520 storage expansion enclosure (Machine Type 1814, Model 52A) provides high-capacity, Fibre Channel and SATA disk storage for the DS5020 storage subsystem. Up to six EXP520 expansion units can be attached to the DS5020 disk subsystem, providing modular storage solutions for up to 112 TB physical storage capacity.

The EXP520 machine type is 1812-52A.

2.5.1 Highlights

Some highlights of this system are:

- ▶ The IBM System Storage DS5020 disk system is designed to deliver high performance, advanced function, high availability, and modular and scalable storage capacity.
- ▶ It supports RAID 0, 1, 3, 5, 6, and 10.
- ▶ It provides SAN-attached 8 Gbps Fibre Channel (FC) host connectivity, as well as optional 1 GbE iSCSI host connectivity.
- ▶ It accommodates up to 16 disk drives installed within the DS5020 enclosure with attachment support for up to six EXP520 expansion units, providing modular and highly scalable storage solutions that range up to 112 TB physical storage capacity.
- ▶ It supports an intermix of SATA drives, FC drives, and encryption-capable FC drives within the DS5020 and EXP520 enclosures.

See Chapter 3, “IBM System Storage DS4000 and DS5000 hardware” on page 23 for more details about the DS5020 and EXP520.

2.6 EXP5060 expansion enclosure

The IBM System Storage EXP5060 storage expansion enclosure provides high-capacity SATA disk storage for the DS5100 and DS5300 storage subsystems. The storage expansion enclosure provides continuous and reliable service by using hot-swap technology for easy replacement, without shutting down the system, and supports redundant, dual-loop configurations. External cables and Small Form-Factor Pluggable (SFP) modules connect the DS5100 or DS5300 storage subsystem to the EXP5060 storage expansion enclosure. The EXP5060 uses redundant 4 Gbps Fibre Channel to make connections to the DS5100 or DS5300 storage subsystem and another EXP5060 storage expansion enclosure in a cascading cabling configuration, offering reliability and performance.

Note: A maximum of eight EXP5060 storage expansion enclosures (with 480 hard drives) can be attached only to the DS5100 and DS5300 storage subsystems and only SATA disks are supported in the EXP5060 expansion.

The EXP5060 is a 4U rack-mountable enclosure that supports up to 60 SATA Disk Drive Modules (DDMs), offering up to 60 TB of SATA disk space per enclosure using 1 TB SATA DDMs. The expansion enclosure contains 60 drive bays (arranged on five stacked drawers with twelve drives for each drawer), dual-switched 4 Gbps ESMs, and dual power supplies and cooling components. Coupled with a storage subsystem (DS5100 or DS5300), you can configure RAID-protected storage solutions of up to 480 TB when using 1 TB SATA DDMs and eight EXP5060 storage expansion enclosures, providing economical and scalable storage for your rapidly growing application needs. The Attach up to 8 EXP5060s feature pack must be purchased for the DS5100/DS5300 storage subsystem to enable it to be connected to up to eight EXP5060 storage expansion enclosures.

2.7 iSCSI host interface

The DS5000 storage subsystems support 1 Gbps iSCSI connectivity. You must determine how the host systems will connect to the storage subsystem.

The iSCSI ports support IPv4 and IPv6 TCP/IP addresses, CHAP, and iSNS. Use either Cat5E or Cat6 Ethernet cable types for iSCSI port connections. A Cat6 Ethernet cable provides optimal performance.

iSCSI host interface cards (HICs) are only an option when initially purchasing the IBM DS5020. iSCSI HICs for the DS53100 and DS5300 are available as field replaceable upgrades (MES) that will be installed by trained IBM service personnel.

See Chapter 3, “IBM System Storage DS4000 and DS5000 hardware” on page 23 for specific information about iSCSI on a particular DS5000 subsystem.

2.8 8 Gbps FC host interface

The DS5000 offers 8 Gbps Fibre Channel (FC) host attachment. Upon purchase of the DS5000, you must determine how many host ports you will need.

The DS5100 and DS5300 storage subsystems are field upgradable to 8 Gbps by installing the quad-port HIC.

There are up to four (two per controller) quad-port 8 Gbps FC host interface cards (HICs) available per DS5100 and DS5300 storage subsystem. They will be installed in pairs of two (one in each controller).

A DS5020 storage subsystem is not upgradable. Host attachment must be ordered with the initial purchase of the DS5020.

2.9 16 port host port upgrade for the DS5100

Initially, the DS5100 storage subsystem supports only one host interface card (HIC) per controller. That limits the number of host attachment ports to a maximum of eight. Starting with firmware 7.50, the DS5100 allows a second HIC card to be installed into each controller to extend the number of host ports to 16 (eight per controller). No premium feature is required.

2.10 64 GB cache upgrade for the DS5100 and DS5300

With controller firmware version (CFW) 7.60, the DS5300 allows a cache size of 32 GB per controller. There are field upgrade options in different configurations available.

These options are available:

- ▶ 8 GB to 16 GB
- ▶ 8 GB to 32 GB
- ▶ 8 GB to 64 GB
- ▶ 16 GB to 32 GB
- ▶ 16 GB to 64 GB
- ▶ 32 GB to 64 GB

2.11 Remote Support Manager Hardware Model RS2

IBM System Storage DS Remote Support Manager for Storage (DS-RSM) Model RS2 is designed to manage and analyze alert notifications with accuracy, efficiency, and cost-effectiveness. DS-RSM Model RS2 replaces the DS-RSM Model RS1 (1818-RS1) and has the following characteristics:

- ▶ Enhanced with a processor and memory technology upgrade
- ▶ Designed to provide fast response time to alerts
- ▶ Offers detailed information about each alert for error analysis
- ▶ Designed to manage up to 50 DS5000/DS4000/DS3000 storage systems per implementation
- ▶ Sends log and status information along with the alert to IBM for problem resolution
- ▶ Enables the IBM service support center to dispatch inquiries to determine problem and speed problem resolution

For additional information about the DS-RSM, go to the following address:

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

See Chapter 6, “IBM Remote Support Manager for Storage” on page 285 for implementation details.

2.12 Increased drive support for the DS5100 and DS5300

Beginning with firmware 7.40, 448 drives are supported for a redundant drive channel pair on DS5100 and DS5300 storage systems.

With the introduction of the EXP5060 storage expansion enclosure, DS5100 and DS5300 storage subsystems can now be configured with eight EXP5060s that support a maximum of 480 SATA hard drives.

2.13 New Storage Manager Subsystem Management window (SMW) design

Beginning with Storage Manager Version 10.50, the Subsystem Management window (SMW) has a different design. A new Summary tab has been added to view a subsystem summary. The Logical/Physical view has been replaced by separate Logical and Physical view tabs, and a Support tab has been added to combine all the support activities together. Figure 2-1 gives an overview of this new design. See Chapter 4, “IBM System Storage DS planning and configuration” on page 103 for details about the new functions.

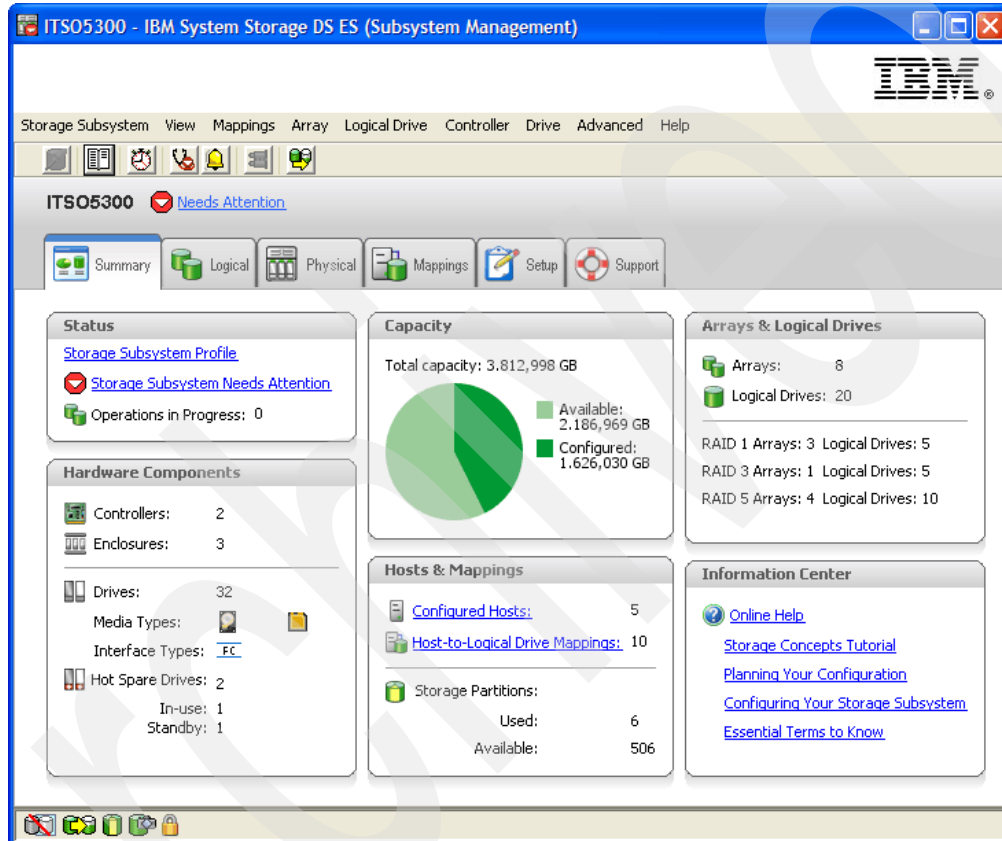


Figure 2-1 New Storage Manager SMW design

Archived

IBM System Storage DS4000 and DS5000 hardware

This chapter describes the hardware and features of the current IBM System Storage DS4000 and the DS5000 family, including the new IBM System Storage DS5020 Express. This section details the DS5000 products, which include the DS5020, DS5100, and the DS5300. The DS4000 product covered in this section includes only the DS4700 model; for more details about the older DS4000 models, see *IBM System Storage DS4000 and Storage Manager V10.30*, SG24-7010.

The chapter also includes a description of additional hardware components that are essential for a complete storage solution. These include the new EXP5000 and the EXP520, with cabling considerations.

The IBM System Storage DS storage subsystem is designed for the rigors of consolidation. The architecture allows for linear scalability, and it is easy to customize a balanced workload to fulfill high IOPS and MBps requirements. Replaceable host cards enable modular adoption for different host interface techniques, such as 4 Gbps FC, 8 Gbps FC, iSCSI, and so on.

3.1 IBM System Storage DS4700 Express

The IBM System Storage DS4700 Express storage subsystem (Figure 3-1) uses 4 Gbps technology (4 Gbps capable drives and the HBAs required to achieve 4 Gbps throughput speeds). It is designed to handle the most rigorous customer workloads and can expand from workgroup to enterprise-wide capability with the attachment of six DS4000 EXP810 disk enclosures. The DS4700 includes 16 disks and offers a fully switched drive-side architecture. Its RAID controllers use a 667 MHz xScale processor with an embedded high-speed XOR engine that generates RAID parity with no performance penalty.



Figure 3-1 IBM System Storage DS4700 Express storage subsystem

Designed to excel at input/output operations per second (IOPS) and MBps, the DS4700 Express can work well for both transaction-oriented and bandwidth-intensive applications. Additionally, the DS4700 Express is a good choice for environments with replication requirements, because it is designed to handle the performance demands of FlashCopy, VolumeCopy, and Enhanced Remote Mirroring.

One of the key advantages of the DS4700 Express 4 Gbps technology is that it is backwards compatible with 2 Gbps and even 1 Gbps technologies. This means that you can avoid replacing your entire SAN with 4 Gbps technology. You can add new technology incrementally. The 4 Gbps products will slow down to the 2 Gbps or 1 Gbps speed if they are connected, but zoning allows you to implement a rolling upgrade strategy with minimal disruption.

The DS4700 Express Disk System and the EXP810 Storage Expansion Unit offer new models designed to be powered from a - 48 V dc Telco industry standard power source and are NEBS-3 compliant. They also include an air filter option with the optional front bezel.

3.1.1 DS4700 features

The available DS4700 models include:

- ▶ DS4700 Express Model 70 (1814-70A)
 - Contains 2 GB of cache memory (1 GB per controller),
 - Four 4 Gbps FC host ports (two ports per controller),
- ▶ DS4700 Express Model 72 (1814-72A)
 - Contains 4 GB of cache memory (2 GB per controller),
 - Eight 4 Gbps FC host ports (four ports per controller),

Attention: New enhancements for the DS4700 Express models are now available as Models 70-DC and 72-DC. These new models are designed to be powered from a - 48 V dc Telco industry standard power source and are NEBS-3 compliant. It also includes an air filter option with the optional front bezel.

The key DS4700 features are:

- ▶ Compact, 3U rack-mountable enclosures containing dual high-performance intelligent RAID controller cards, accommodating 16 SATA or FC E-DDMs.
- ▶ DS4700 supports up to 12 TB with 750 GB SATA disks, up to 16 TB with 1 TB SATA disks, and 4.8 TB with 300 GB FC disks (7.2 TB with 450 GB FC disks). Adding the EXP810 or EXP710 storage enclosure raises the size of the storage up to 84 TB (using 750 GB SATA), 112 TB (using 1 TB SATA) using the EXP810, and to 30 TB using the EXP710.
- ▶ Controller firmware (CFW) Version 6.23 and later (V7.10 and V7.30) allows the intermixing of expansion units (EXP710 and EXP810) and also the intermixing of FC and SATA disks in the same enclosure (only for EXP810) or storage subsystem.
- ▶ Utilizes 667 Mhz xScale processor with embedded XOR engines on RAID controllers.
- ▶ Dual, redundant 4 Gbps RAID controllers with up to 4 GB of cache memory (2 GB per RAID controller).
- ▶ New lithium-ion battery backup protects data cache in each controller for at least 72 hours.
- ▶ Hot-swappable cache backup battery.
- ▶ Redundant, hot-swappable power supplies and fans.
- ▶ Supports RAID 0, RAID 1, RAID 3, RAID 5, RAID 6, and RAID 10 (RAID 1+0).
- ▶ Switched loop architecture that supports two dual redundant FC disk loops.
- ▶ Auto-negotiates 1, 2, or 4 Gbps host connection speeds.
- ▶ Supports Global Hot Spare.
- ▶ Host-side connections support Fibre Channel switched, loop, or direct connections.
- ▶ Redundant drive-side connections designed to avoid any single point of failure and maintain high availability.
- ▶ Supports up to six EXP810 (each EXP810 has 16 disk bays) to attach up to 112 FC disks (additional licenses required).
- ▶ The following disks are supported:
 - 2 Gbps FC: 10K rpm, 300 GB/146 GB/73 GB/36 GB
 - 2 Gbps FC: 15K rpm, 300 GB/146 GB/73 GB/36 GB
 - 4 Gbps FC: 15K rpm, 300 GB/146 GB/73 GB/36 GB
 - 4 Gbps FC: 15K rpm, 600GB/450 GB
 - 4 Gbps SATA: 7200 rpm, 500 GB E-DDM, 750 GB E-DDM, 1 TB E-DDM (SATA and FC cannot be intermixed in the same enclosure.)
- ▶ Supports shortwave Fibre Channel 4 Gbps host attachment (on initial release).
- ▶ Supports multiple heterogeneous server and operating system (host kits required).
- ▶ Both models 70 and 72 have selectable storage partitions up to 128.
- ▶ Powerful on demand functions: Dynamic volume expansion, dynamic capacity expansion, and dynamic RAID level migration, which allow users to modify storage configurations on-the-fly without incurring any downtime.
- ▶ New Remote Support Manager (RSM) support to notify IBM (Call Home) in case of problems or failures. See Chapter 6, “IBM Remote Support Manager for Storage” on page 285 for more details about RSM.
- ▶ New dual 10/100 Ethernet for out-of-band management to separate out-of-band management from service diagnostics for each controller.
- ▶ FlashCopy, VolumeCopy, and Enhanced Remote Mirroring (premium features).

- ▶ DS4000 Storage Manger: SM Version 9.23 until the newest of 10.60
- ▶ Controller firmware version (CFW) 6.23 until the newest of 7.60.

3.2 DS4700 model comparison

Figure 3-2 summarize the characteristics of the DS4700 models available today.

Feature/Function	DS4700-72A	DS4700-70A
Machine type-model	1814-72A	1814-70A
Architecture		
Host ports – max	8 FCP	4 FCP
Native host link speed	4Gbps	4Gbps
Supported host link speeds	1, 2, 4Gbps	1, 2, 4Gbps
Drive ports / loop pairs or loops	4 / 2	4 / 2
Processors	2	2
Processor type	Intel xScale 667 MHz	Intel xScale 667 MHz
Cache per subsystem	4 GB	2 GB
Disk Storage		
Maximum drives	112 (EXP810) 100 (EXP710)	112 (EXP810) 100 (EXP710)
Drives supported	FC, SATA	FC, SATA
RAID levels supported	0, 1, 3, 5, 6, 10	0, 1, 3, 5, 6, 10
Intermix FC and SATA	Yes	Yes
Availability Features		
FlashCopy	Yes	Yes
Volume Copy	Yes	Yes
Remote mirror, synchronous copy, to 10km	Metro Mirror	Metro Mirror
Remote mirror, asynchronous copy, extended distance	Global Copy	Global Copy
Remote mirror, asynch copy with write consistency, extended dist.	Global Mirror	Global Mirror
Call Home support	RSM for Storage	RSM for Storage
Non-call Home support	email alerts	email alerts
Hot swappable disks/components	Yes	Yes
Concurrent firmware upgrade	Yes	Yes

Figure 3-2 DS4700-72 and DS4700-70 characteristics

3.3 DS4000 series expansion enclosures

There are various models of expansion enclosures available for the DS4000 series. We provide a short description for the latest models; for more details, see *IBM System Storage DS4000 and Storage Manager V10.30*, SG24-7010.

The models are:

- ▶ DS4000 EXP710 FC Expansion Enclosure with 2 Gbps Switched Disk Technology (This technology is also referred to as a Switched Bunch Of Disks (SBOD).)
- ▶ DS4000 EXP810 FC Expansion Enclosure with 4 Gbps capability

In this section, we discuss the features of the DS4000 expansion units (EXP) and how they can be attached and combined to expand the storage capacity of the DS4000 storage servers. In particular, we look at the possibilities and limitations of intermixing FC and SATA enclosures, as that is now supported with Storage Manager V9.x and later.

For additional information about drive migration and special handling procedures for expansion enclosures, see the *IBM System Storage DS4000/DS5000 Hard Drive and Storage Expansion Enclosure Installation and Migration Guide*, GC53-1139.

3.3.1 IBM Total Storage DS4000 EXP710 expansion enclosure

The EXP710 (Figure 3-3) is a rack-mountable storage expansion enclosure that contains 14 bays for slim-line hot-swappable Fibre Channel disk drives. It occupies 3U inside a rack and features hot-pluggable and redundant power supplies and fans.

The redesigned Environmental Service Module board (two ESMs in an EXP710) provides a redundant switched connection to each drive in EXP710, eliminating the arbitrated loop within the disk enclosure that existed in the previous EXP700 model.

This feature allows not only better performance in a large configuration, resulting from reduced loop latency, but also improves diagnostics and troubleshooting tasks. Indeed, the point-to-point topology provides a full isolation of the drives, eliminating the risk of a single drive disrupting the loop and causing other drives to fail on the loop (fault propagation). This also enables selective diagnosis of FC errors.

For more information about DS4000 EXP710 Expansion Enclosure, go to the following address:

<https://www-947.ibm.com/systems/support/supportsite.wss/brandmain?brandind=5345868>



Figure 3-3 EXP710 front view

3.3.2 IBM System Storage EXP810 expansion enclosure

The IBM System Storage EXP810 Storage Expansion Unit (1812-81A) expands the sixth-generation architecture of the DS4000 series disk storage systems with increased, built-in capability for 4 Gbps Fibre Channel operation and 16-bay disk capacity. The Expansion Unit is rack-mountable, occupies 3U rack space, and comes standard with two 4 Gbps shortwave small form-factor pluggable fiber optic transceivers for connection to selected DS4000 midrange disk systems. The EXP810 attaches to the DS4300, DS4500, DS4700, DS4800, DS5100, and DS5300 storage subsystems.

Further information can be found at the following address:

<http://www-03.ibm.com/servers/storage/disk/ds4000/exp810/>



Figure 3-4 EXP810 front view

Note: There is also an EXP810 model designed to be powered from a - 48 V dc Telco industry standard power source and it is NEBS-3 compliant. It includes an air filter option with the optional front bezel.

3.3.3 Intermixing EXP810 and EXP710

When intermixing the EXP810 and EXP710, even though the enclosure grouping by type is not a requirement, IBM recommends that you group them by enclosure type in a redundant drive channel pair to simplify maintenance and troubleshooting. Also, the EXP810 link rate switch must be set at 2 Gbps if you use them in the same drive channel.

There is a limit for the number of maximum configurable disks in redundant drive channel pairs. This limit determines the fixed number of drive enclosure combinations (Table 3-1). The total number of configurable drives is 112.

Table 3-1 Total number of attached expansion units

Enclosure	Maximum number of configurable expansion units in a drive channel pair
EXP710	8
EXP810	7
Intermixed	7

IBM does not recommend the intermixing of expansion units in the same redundant drive channel pair. IBM recommends connecting all of the EXP710s to a redundant drive channel pair and all of the EXP810s to the other (second) redundant drive channel pair. In this case, you can operate the EXP810 at a higher speed and avoid the subsequent separation of intermixed drive channel pairs.

For detailed instructions for any intermix configuration, see the *IBM System Storage DS4000/DS5000 Hard Drive and Storage Expansion Enclosure Installation and Migration Guide*, GC53-1139.

IBM System Storage DS4000 and Storage Manager V10.30, SG24-7010 also has information about intermixing different expansion enclosures behind the DS4000 storage systems.

3.4 DS5100 and DS5300 storage subsystems

The first models of the DS5000 family that have been released offer increased performance compared to the family's DS4000 predecessors. With the ability to hold up to 448 TB of SATA disk physical capacity or more than 268 TB of FC disk physical capacity, the DS5000 can be customized for both performance-oriented and capacity-oriented storage solutions.

Figure 3-5 shows a DS5300 storage subsystem with attached EXP5000 expansion modules.



Figure 3-5 DS5300 storage subsystem

The following two models are available:

- ▶ DS5100 (1818-51A)
- ▶ DS5300 (1818-53A)

The two models differ in the number of host ports and amount of cache for data and performance.

Model DS5100 initially comes with eight host channels, 4 Gbps FC, and 8 GB cache for data.

Model DS5300 comes with eight or sixteen host channels, 4 Gbps FC, and 8 or 16 GB cache for data and full performance activation.

Models DS5100 (1818-51A) and DS5300 (1818-53A) have these features:

- ▶ Compact 4U rack-mountable enclosure.
- ▶ Utilizes new seventh generation dedicated ZIP ASIC engines on RAID controllers.
- ▶ Features Intel® Xeon 2.8 GHz processor.
- ▶ Dual, redundant controllers.
- ▶ PCI Express x8 bus technology.
- ▶ Dedicated cache for data (base model has 8 GB cache) with enhanced diagnostics. The architecture is designed to support a 32 GB cache per controller.
- ▶ Dedicated processor memory of 2 GB per controller.
- ▶ Hot-swappable lithium-ion battery for backup and destaging data from cache.
- ▶ New flash memory to store dirty data from cache during power outage.
- ▶ Two dedicated PCI Express buses for cache mirroring.
- ▶ Redundant, hot-swappable power supplies and fans.
- ▶ Hot-swappable interconnect module acts as midplane.

- ▶ Supports RAID 0, 1, 3, 5, 6, and 10 (RAID 1+0).
- ▶ Supports RAID 1 and 10 dedicated mirrored drive pairs configurations.
- ▶ The ability to create a RAID 10 or 0 group on all available drives to maximize performance for a LUN.
- ▶ Flexible host interface modules can be added, changed, or mixed as the infrastructure changes. The quad-4 Gbps Fibre Channel (FC) Host Interface Cards (HIC) can be replaced now with quad-8 Gbps FC HIC or dual-1 Gbps iSCSI Host Interface Card. Field replacements (MES) are available.
- ▶ Supports an unlimited number of Global Hot Spare drives with the ability to enable/disable the copy back function (important for SATA drives).
- ▶ Supports eight host-side connections (two HICs) per controllers on both models.
- ▶ Host-side connections support
 - Fibre Channel Switched Fabric
 - Arbitrated Loop and Fibre Channel Direct Connections
 - Ethernet Direct Connection and Ethernet Switched Network (with iSCSI HIC)
- ▶ Supports sixteen 4 Gbps drive-side connections for both controllers. This allows a total of eight dual-redundant drive channel pairs to be implemented to support expansion enclosure additions.
- ▶ Redundant drive-side connections are designed to avoid any single-point of failure and maintain high availability.
- ▶ Supports up to 28 EXP5000s (or a mix of EXP5000s and EXP810s for migration purposes) for a total of 448 disks. This will allow you to install up to 448 TB of raw capacity with 1 TB SATA disks or 268 TB of raw capacity with FC drives or mixed.
- ▶ Supports a maximum of 20 Solid State Drives (SSDs) within EXP5000.
- ▶ Fully supports Fibre Channel/SATA intermix (premium feature) by allowing the simultaneous usage of SATA and Fibre Channel behind one DS5000 controller, allowing user flexibility and increased storage capacity utilization. It is also possible to mix disks of different size and technology inside one enclosure.
- ▶ Supports up to 512 host storage partitions that isolate LUNs for different servers or groups of servers.
- ▶ Supports up to 2048 volumes.
- ▶ Supports up to 2048 host logins.
- ▶ Supports 4096 command queue depth (maximum drive queue depth is 16).
- ▶ Supports logical volumes greater than 2 TB (when required and supported by the operating system).
- ▶ Supports shortwave Fibre Channel 4 and 8 Gbps host attachment.
- ▶ Supports 1 Gbps copper iSCSI host attachment.
- ▶ Multiple heterogeneous server and operating system support (host kits required).
- ▶ Powerful On Demand functions: Dynamic Volume Expansion, Dynamic Capacity Expansion, and Dynamic RAID Level Migration. Dynamic Segment Size allows users to modify storage configurations on-the-fly without incurring any downtime.
- ▶ Remote Service Manager notifies IBM if there is an issue. see Chapter 6, “IBM Remote Support Manager for Storage” on page 285 for more details.
- ▶ New dual 10/100/1000 Ethernet for out-of-band management to separate out-of-band management from service diagnostics for each controller.

- ▶ FlashCopy (premium feature) for up to 16 copies per base volume. There are two FlashCopies per default (without premium feature).
- ▶ VolumeCopy (premium feature).
- ▶ Remote Volume Mirroring: Metro Mirror, Global Mirror, and Global Copy (premium features) for up to 128 pairs.
- ▶ Standard DB-9 serial connection for service purposes.

Additionally, the DS5100 offers further value in flexibility, scalability, and investment protection by providing the ability to upgrade to a DS5300. The DS5100 provides excellent performance for dynamic storage consolidation solutions in the deployment stage, with the full assurance that this model can be upgraded to a DS5300 performance level if required.

3.4.1 DS5100 and DS5300 controller architecture

The DS5100 uses the same hardware as the DS5300, although the internal bus speed is reduced by the controller code (controller firmware). Both use two controllers connected by the interconnect module. It does not have a backplane board.

This section describes the architecture of a single controller. The DS5300 controller uses a seventh generation dedicated ZIP application-specific integrated circuit (ASIC) chip, which is designed and customized for a particular use, rather than intended for general purpose use. The new ZIP ASIC is designed for supporting I/O operations, especially for calculating redundancy information for RAID 3, RAID 5, and RAID 6. The parity can be calculated much faster by dedicated hardware than by software, which results in good performance, especially for writing data that is protected by one of the mentioned RAID levels.

The ZIP ASIC chip calculates data in cache, so it uses a fast 17 Gbps bus to cache memory dedicated for data. The DS5300 controller is based on a PCI Express x8 2 Gbps architecture. With the ZIP ASIC, there is also an Intel Xeon® 2.8 GHz processor with dedicated 2 GB of memory for I/O control, management, and other general purposes. The ZIP ASIC chip is one of the most important components in the architecture, so all disk and host chips are directly connected using fast PCI Express buses to achieve high performance.

There are two quad FC 4 Gb chips dedicated to disk channels connected to the ASIC through two PCI Express buses. The first chip is connected to eight external FC 4 Gbit ports through internal switches in the same controller. The second one is connected to eight external FC 4 Gb ports in the second controller through interconnect modules. Internal switches are used to have a non-blocking connection to each disk channel port from both controllers.

The ZIP ASIC is connected through two PCI Express buses to two Host Interface Cards. Each card has four FC 4 Gb ports.

DS5300 uses two dedicated PCI Express buses for mirroring the cache. The buses connect directly to the ASICs on both controllers through the interconnect module.

Figure 3-6 shows the main components of the DS5000 controller.

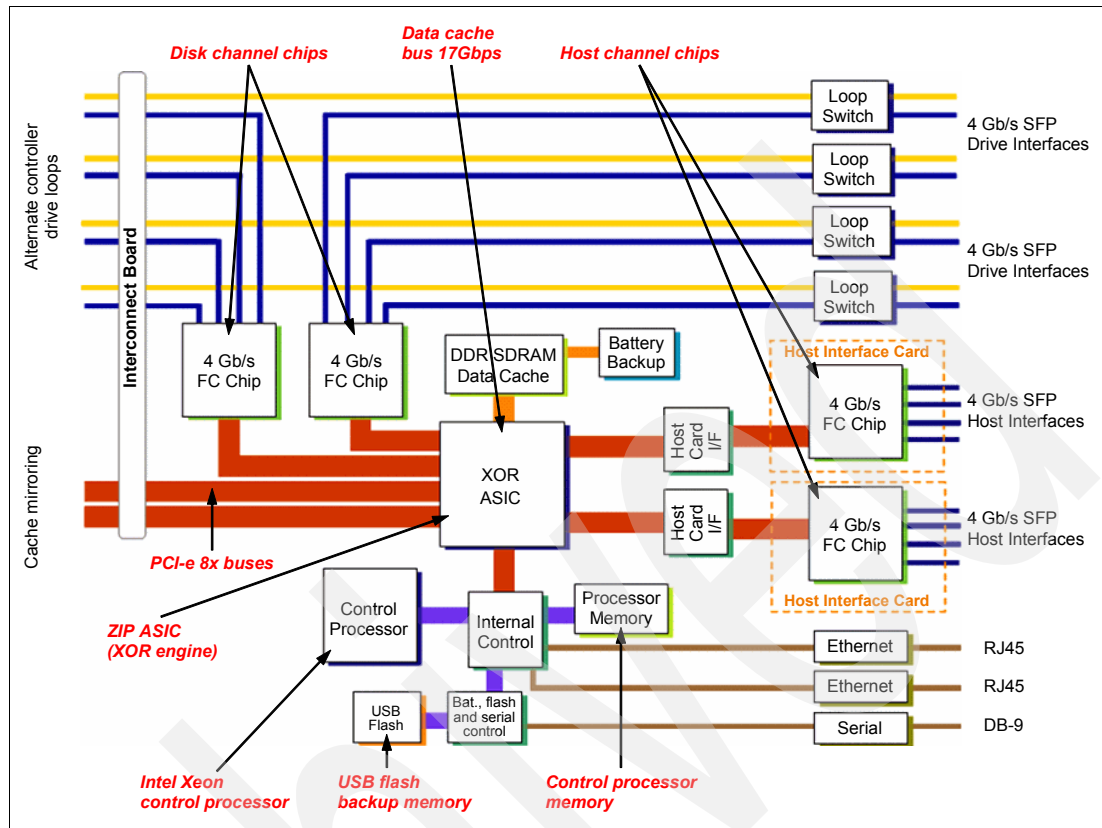


Figure 3-6 DS5300 controller's main components

Figure 3-7 shows the DS5300's controller B with the top lid removed. Model 5100 may have only one (left) Host Interface Card and has less data cache memory.

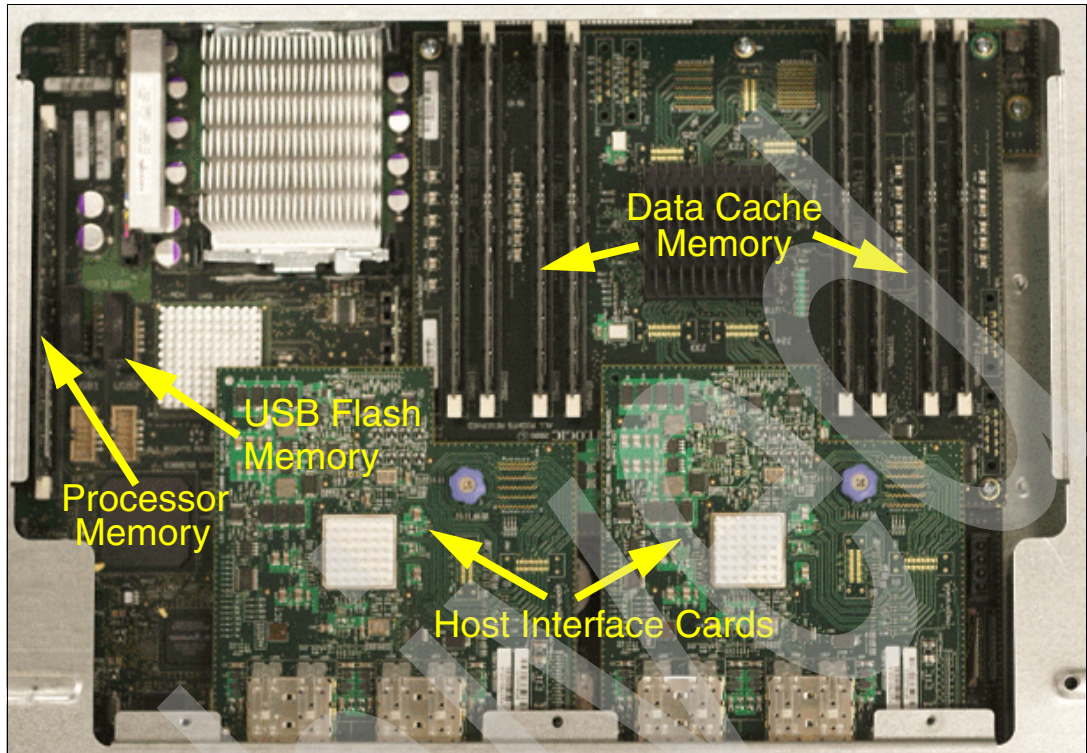


Figure 3-7 DS5300 controller

Figure 3-8 shows a DS5300 controller without the memory board and Host Interface Cards.

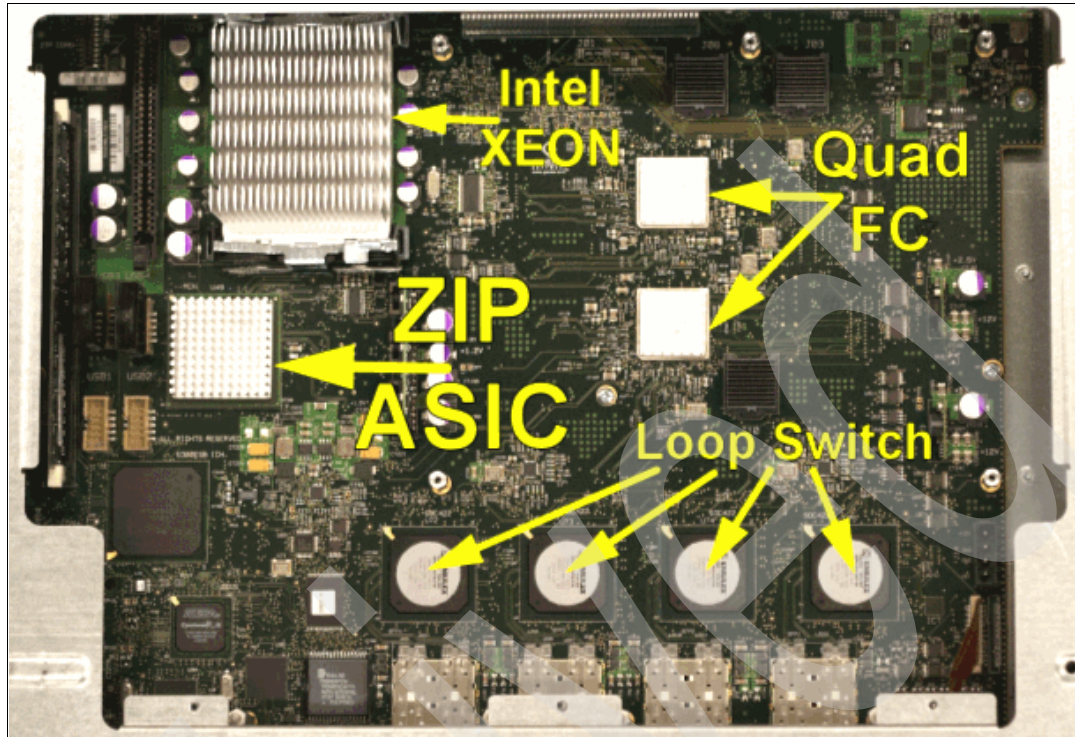


Figure 3-8 DS5000 controller without memory board and Host Interface Cards

Data cache memory

The controller cache has a large amount of physical memory dedicated to I/O operations between the controller and hosts and between the controller and disk drives. It is logically and physically separate from the controller processor's memory and can participate in Direct Memory Access (DMA) operations with both host side and drive side physical channel adapters, as shown in Figure 3-9 on page 35. Consequently, the controller's processor is not required to execute the data movement in and out of the cache memory. The controller cache is a significant contributor to the overall performance of the storage array. The use of the cache increases controller performance in several ways:

- ▶ The cache acts as a buffer so that host and drive data transfers do not need to be synchronized.
- ▶ If write-back caching is used, the host can send subsequent write commands before the data from a previous write operation has been written to a drive.
- ▶ The data for a read or write operation from the host may already be in the cache from a previous operation, thus eliminating the need to access the drive. This is referred to as "Reach Cache".
- ▶ If cache pre-fetch is enabled, sequential read access is optimized as cache pre-fetch, which makes it much more likely that a read operation will find its data in cache rather than have to read it from a disk drive. This is also referred to as "Read Ahead".

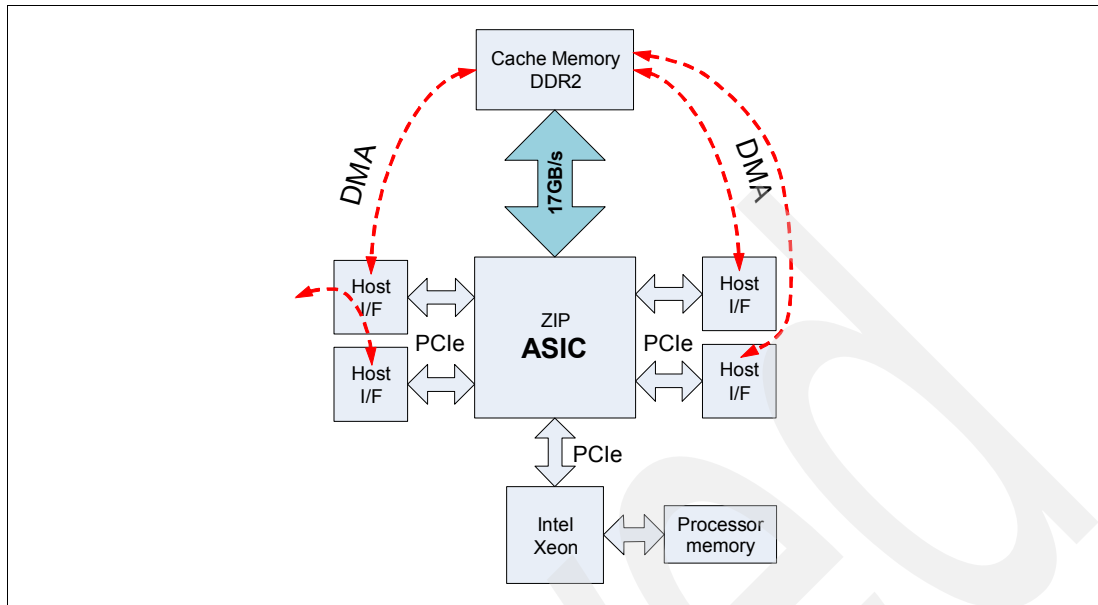


Figure 3-9 Cache memory access

Dirty data: A cache block that contains new data that is not consistent with data on disk is dirty. The cache block holding the write data is dirty until the data is written to disk and the cache block becomes clean.

The cache may have dirty data in it. The cache memory is RAM, so if there is a power failure, the contents of the cache will be lost, including all the dirty data. There is no way for an application using the storage subsystem to tell exactly how many writes were lost in a case like this, so recovery can be very difficult or impossible. This severely limits the applicability of write caching to non-essential data like scratch data, temporary files, and so on. To circumvent this limitation, the controller can make the cache contents persist through power failures by using persistent cache backup devices.

Persistent cache backup devices is a new method implemented in the DS5000 family. The controller has persistent cache backup devices (USB flash memory) into which cache contents can be stored for an indefinite period of time. In this case, the controller also has a battery with enough capacity to let it write the full contents of the cache memory to the persistent cache backup devices in case of a power failure. The controller firmware turns off the batteries when the backup has been completed.

The state of the cache persistence mechanism is tied to the cache management subsystem and affects write caching directly. If the cache persistence mechanism fails for some reason, for example, the battery is missing in a battery-backed cache memory controller, write caching will be disabled for all volumes, except those that have the “cache without batteries” attribute set to true (see 4.10.6, “Cache parameters” on page 249 for more information).

When the controller loses power and dirty data is in the cache, the backup process starts (copying data from cache to flash memory). If power is restored by the time the backup is complete, normal operations continue. If power is *not* restored by the time the backup is complete, the controller disables the batteries and powers off. When it powers on, any dirty data in flash memory will be copied to cache. Host I/O is supported while backup data is being restored. Normal power up continues as well. Restored data is immediately scheduled to be flushed to disks.

Note: Both data cache as well as flash memory are CRC protected.

Data flow

Figure 3-10 shows a simplified data flow from host port to disk drive (not covering the data cache flow). This diagram shows the configuration where an array is built on one disk enclosure and two disks ports are used for I/O operations to one volume. If there are two volumes on the same array, the second one can be managed by controller B. For each volume you set, the default controller will manage it, which allows you to evenly load both controllers and all ports. In this example, there is a RAID 5 array group defined on five drives and one volume with the default controller A (preferred ownership) and a segment size of 256 KB. There is a 1 MB write request. Drives 1, 3, and p (parity) use port 8 in controller A and drives 2 and 4 use port 1 in controller B.

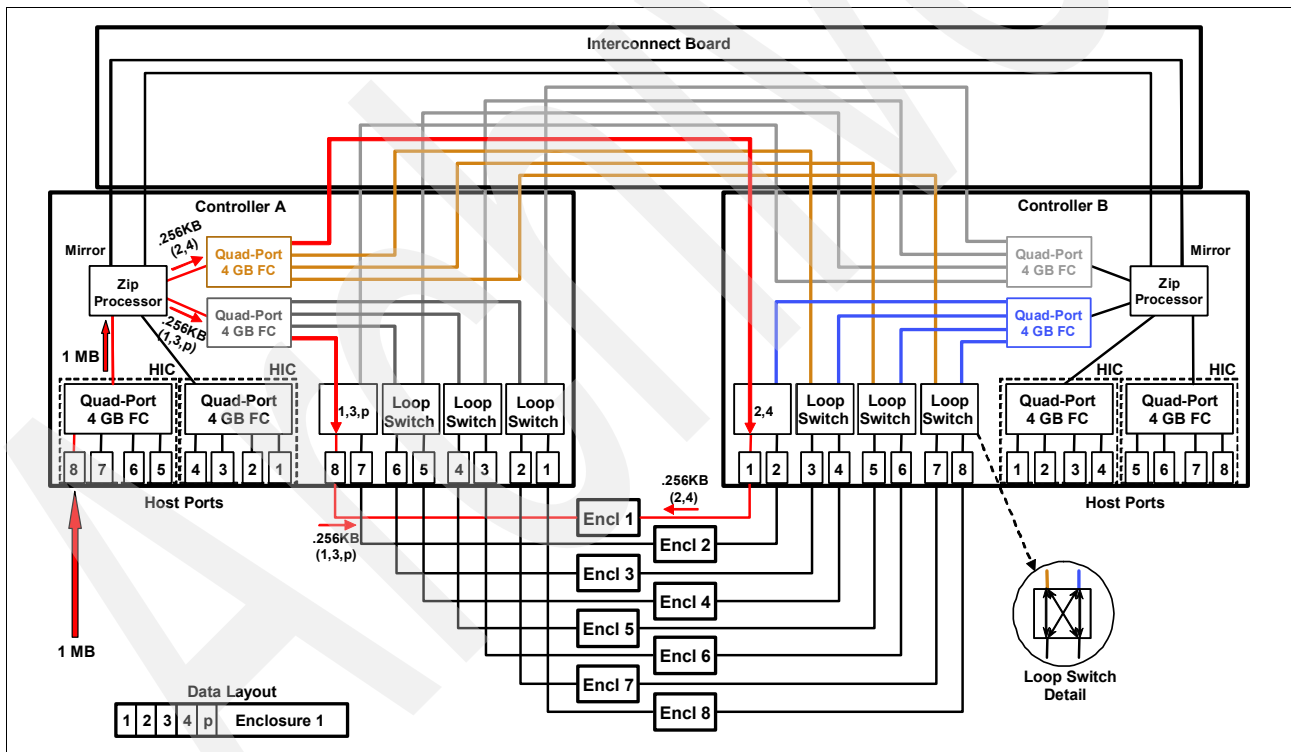


Figure 3-10 Data flow with one enclosure

In the example shown in Figure 3-11, the array is built on five enclosures and uses five disk ports to communicate. Note that disk number 2 is placed in the second enclosure in the second position in the enclosure, and so on (“barber pole” volume). If all drives are placed vertically (on many enclosures), but in the same position, for example, the first slot, five disk ports are used, but all in the same controller, using one (of two) quad disk port chips.

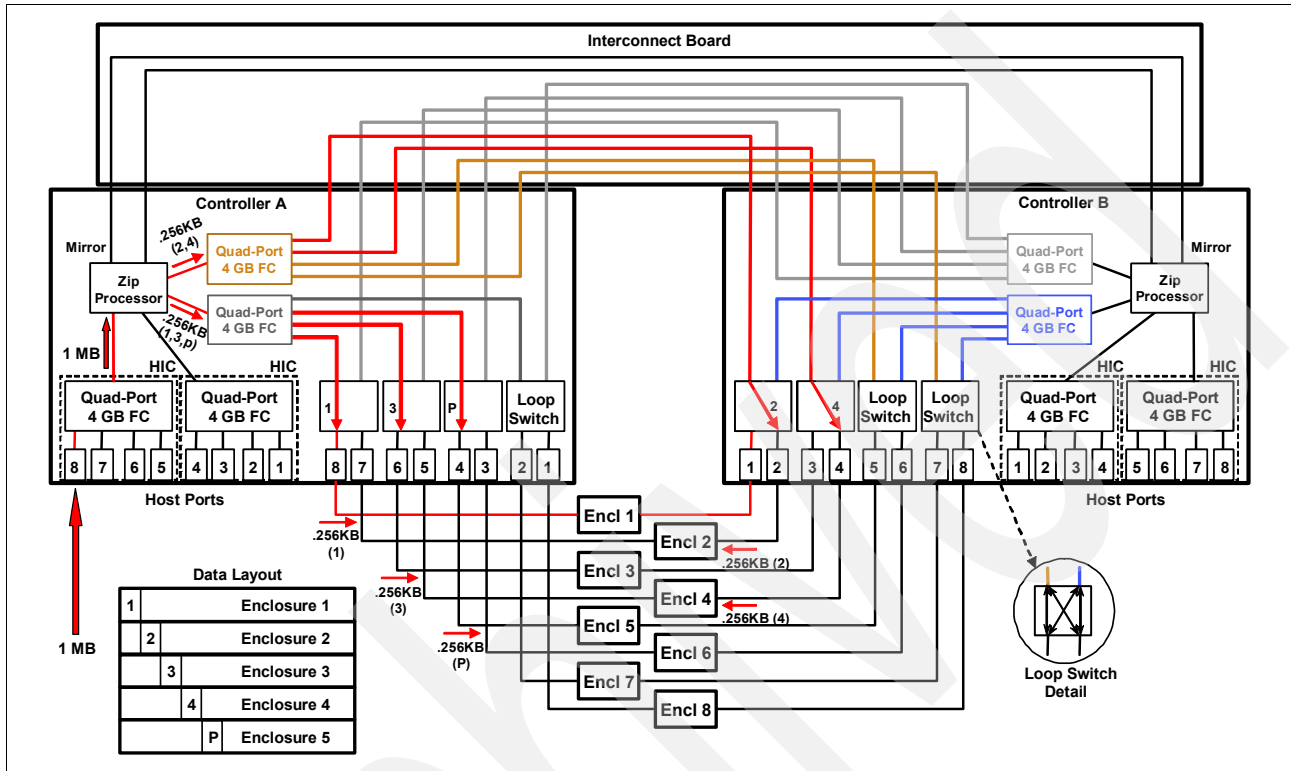


Figure 3-11 Data flow with more expansions

Note: Normally, when both controllers are running, disk loop ports in controller A access odd drives in a disk enclosure. Disk loop ports in controller B access even drives in a disk enclosure. Each controller has connections to all disk ports in both controllers. When the controller fails, the running one uses its ports to communicate to all drives.

Write operation steps

Here are the write operation steps:

1. Data comes through the host port of the preferred controller for the volume.
2. Data is written to cache.
 - a. If cache mirroring is enabled, cache without battery is disabled, the battery state is okay (or just cache without battery is enabled), and data is copied to the second controller.
 - b. If Enhanced Remote Mirroring is established for the volume, the remote copy procedure starts. If synchronous mode is used, data is copied to a remote DS5000.
 - c. If write caching is enabled, the host receives the message I/O completed.
3. Data is written to disks of the array.
4. Depending on the number of disks, the local or second controller's ports are used for writing.

5. The cache block status is set to Clean.
6. The host receives the message I/O completed (if it did not receive this message before).

Read operation steps

The read operation steps are:

1. A host requests data.
2. If data is in the cache, the data is sent to the host and the I/O is completed.
3. Data is read from disk.
4. Data is copied into cache.
5. The host receives the required data and the I/O is completed.
6. If sequential reading is discovered, the next *n* segments are read and copied into the cache.

Cache block states

The cache block states are:

- ▶ **Empty:** During the start-of-day processing, the controller firmware will partition the cache memory into cache blocks (except for the part of cache memory reserved for various cache metadata). All these cache blocks will be put on the Free list and their state will be Empty.
- ▶ **Clean:** A cache block that contains data that is consistent with the data on disk is Clean.
- ▶ **Dirty Write-Through (Dirty WT):** A cache block that contains new data that is not consistent with data on disk is Dirty. In the write-through cache mode, status is not returned to the host until the data has been written to disk. The cache block holding the write data is Dirty Write-Through until the data is written to disk and the cache block becomes Clean.
- ▶ **Dirty Write-Back (Dirty WB):** In the write-back cache mode, the status is returned to the host as soon as the data has been inserted into the cache. The cache block holding the write data is Dirty Write-Back until the data is written to disk and the cache block becomes Clean.
- ▶ **Dirty Write-Back Mirrored (Dirty WBM):** If cache mirroring is enabled for the volume, the dirty cache block will be replicated to the cache memory in the other controller, and when the replication is completed, the cache block transitions from Dirty Write-Back to Dirty Write-Back Mirrored.

Figure 3-12 shows a simplified cache data flow for a DS storage subsystem. It does not cover remote mirror operations and parity calculations for RAID 3, 5, or 6. In the example, I/O size is equal or smaller than segment size.

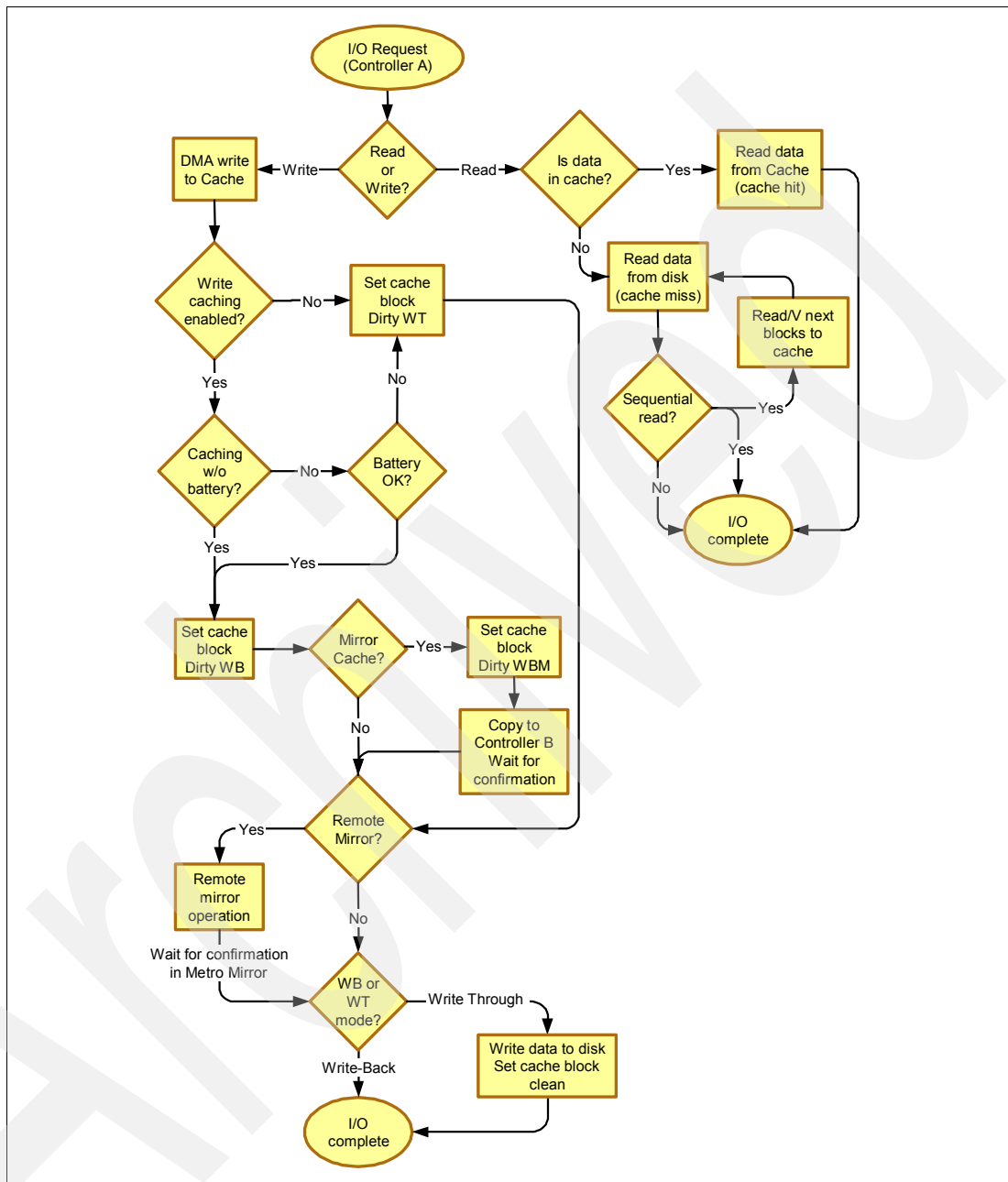


Figure 3-12 Simplified data cache flow diagram

3.4.2 DS5000 storage subsystem chassis design

In Figure 3-13, we show a diagram of the DS5000 modules.

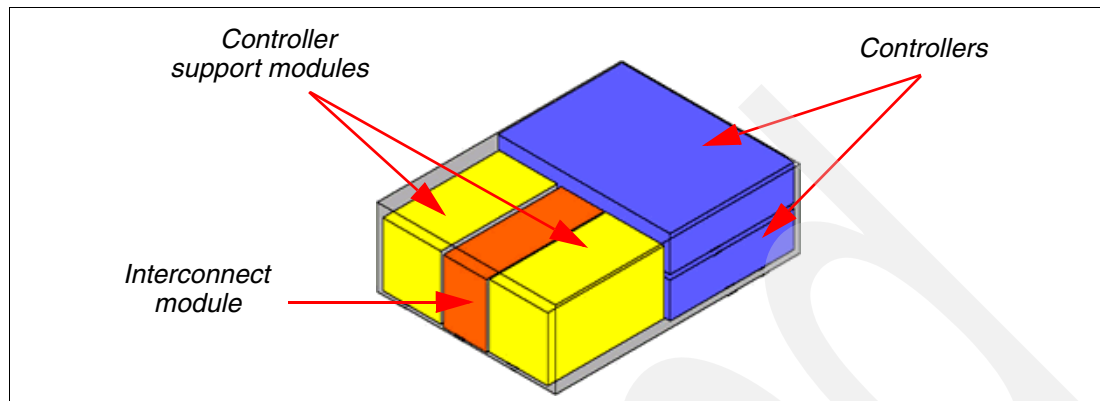


Figure 3-13 DS5000 modules

The DS5000 base controller unit is broken down into five primary Field Replacement Units (FRUs). These components are two controller modules, two controller support modules, and one interconnect module:

- ▶ The controller modules contain the ASIC engines, the processors, cache memory, processor memory, flash memory for cache backup, and additional electronics that allow the DS5000 to process I/O.
- ▶ The controller support modules contain the power supplies, battery charging, and fans.
- ▶ The interconnect module holds the batteries and functions as a hot-swappable midplane. Controllers use this module to mirror the data cache and share disk loops ports.

All of the five FRUs and their individual components are hot-swappable and can be interchanged while the system is online, allowing DS5000 users to maximize their uptime. See the corresponding service guide for specific replacement instructions that take place for the different FRUs. The latest service guide can be found at:

<http://www-947.ibm.com/systems/support/supportsite.wss/selectproduct?taskind=7&brandind=5000028&familyind=5368957&typeind=0&modelind=0&osind=0&psid=sr&continue.x=1>

Note: A DS5000 storage subsystem has no single point of failure; each module, including the midplane, is hot-swappable. When any single module fails, I/O operations continue.

3.4.3 DS5000 storage subsystem front view

The front section of the DS5000 storage subsystem contains the two controller support modules and the interconnect module.

The controller support modules are the units (FRUs) on the left and right. They each house a fan and a 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 new hot-swappable midplane. When the interconnect module is removed, the DS5000 storage subsystem automatically suspends controller B, fails over all the LUNs to controller A, and continues to operate. The midplane has batteries, so write caching is disabled for the default settings of the LUNs. If you still want to use cache for writing, open Storage Manager, select a volume, and select **Logical Drive** → **Change** → **Cache settings** → **Enable write caching without batteries**. You can find more details about this topic in Chapter 4, “IBM System Storage DS planning and configuration” on page 103.

Possible loss of data: When you use the feature Enable write caching without batteries, you can lose data in a cache when the batteries are removed or discharged and power fails.

When the interconnect module is put back in place, the DS5000 storage subsystem can revert back to normal operation. If the operating system uses non-AVT path failover drivers, you have to manually redistribute the LUNs to their respective owning controller (A or B). All the components depicted in Figure 3-14 are hot-swappable and can be replaced on-the-fly to ensure that users do not incur any downtime.

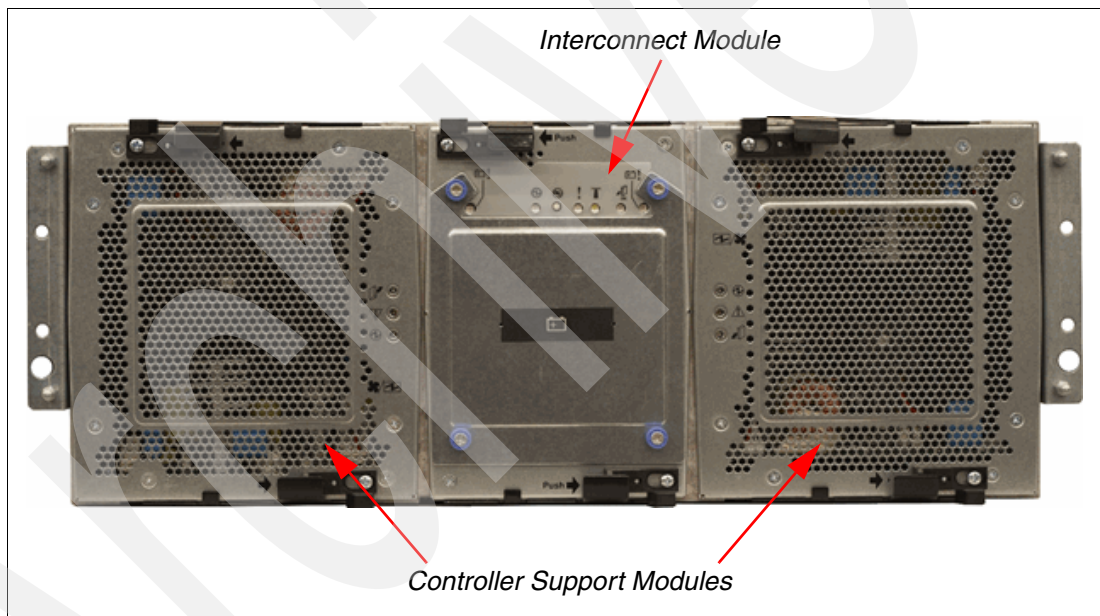


Figure 3-14 Front view of the DS5000 storage subsystem

3.4.4 Interconnect module and battery packs

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 (Figure 3-15). An interconnect module works as a midplane, so it also connects disk loops and buses to mirror cache data between controllers.

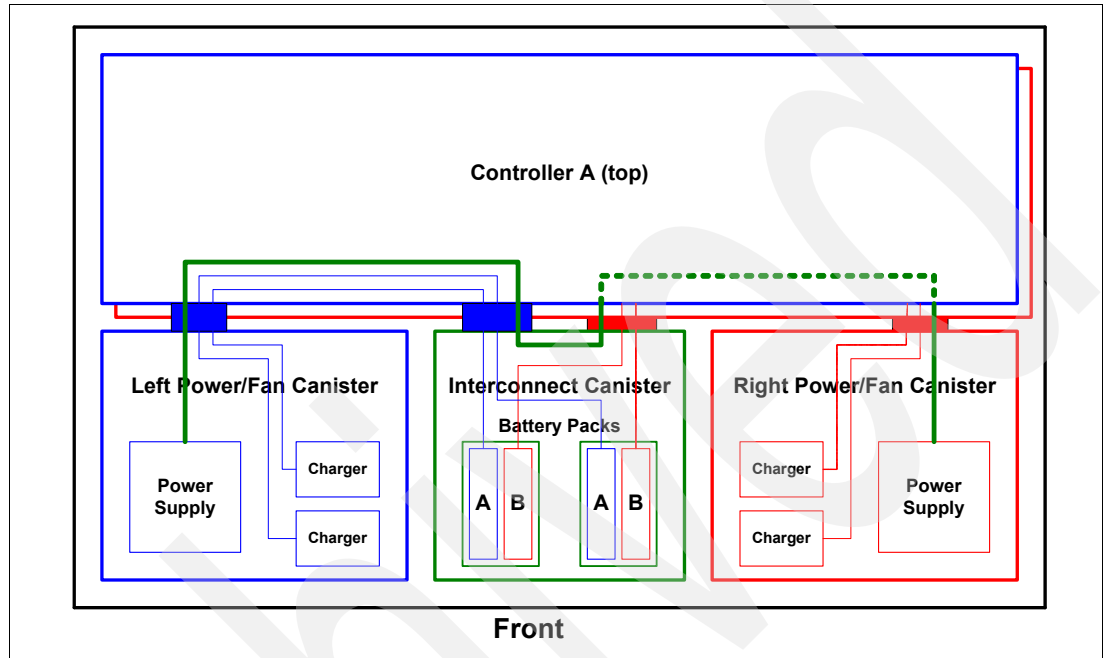


Figure 3-15 Power distribution

The DS5000 storage subsystem battery packs do not use the expiration dates. Only replace a battery pack when the LEDs have indicated that they have failed (see “Interconnect module LEDs” on page 54). You only have to replace the battery pack that failed, not both battery packs.

Because 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 impact due to the disabling of the write-caching function.

Figure 3-16 shows a DS5000 storage subsystem without a chassis (controller A is removed).

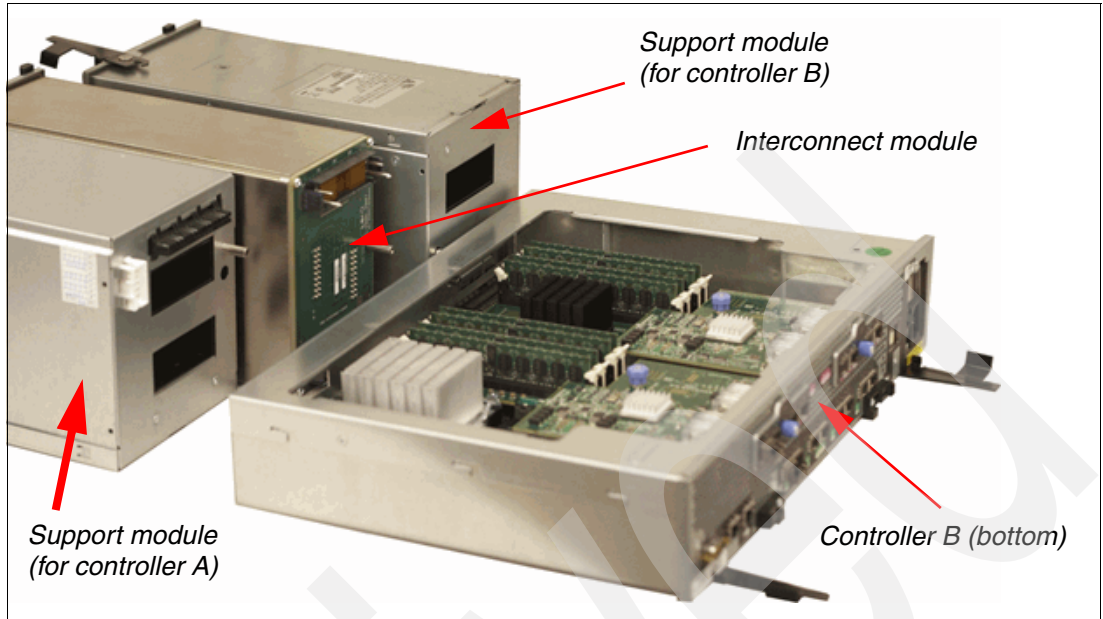


Figure 3-16 DS5000 storage subsystem without a chassis

Figure 3-17 shows the interconnect module with the removed battery.

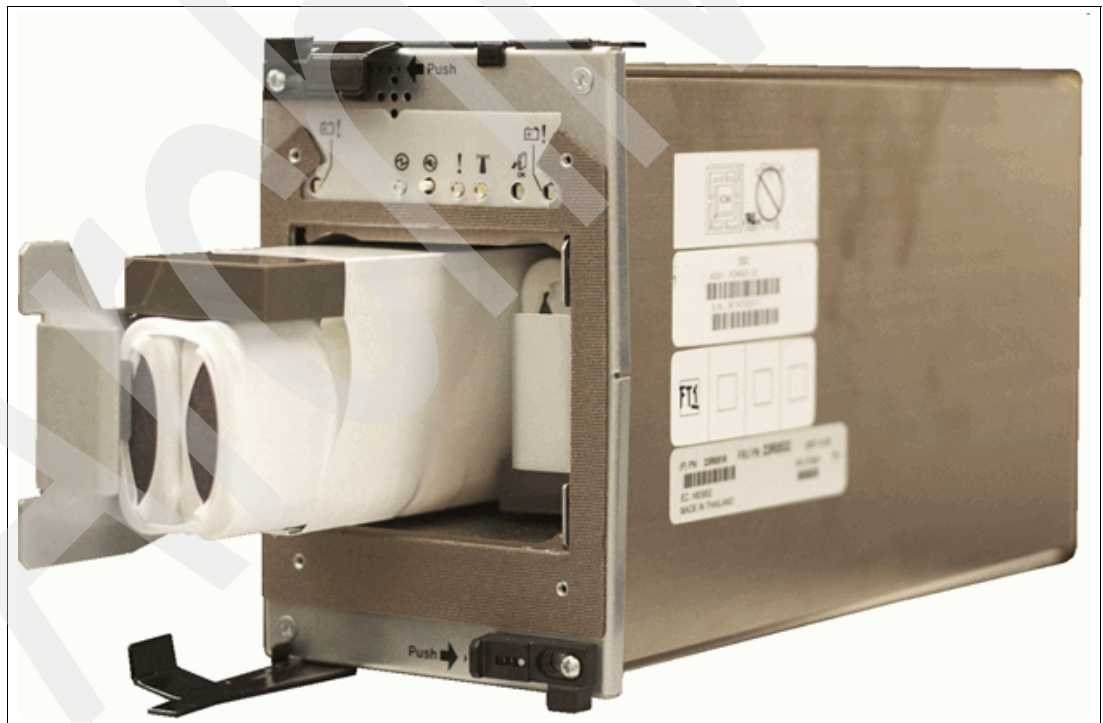


Figure 3-17 Interconnect module with removed battery pack

Smart Battery Backup Units (BBUs)

The Smart Battery Management feature allows the management facilities of the storage subsystem to take advantage of the capabilities provided by a “smart” battery unit. This feature provides a mechanism that enables the controller firmware to accurately determine how much capacity is left in the battery in a battery backup unit (BBU) and to access the battery’s state of health (SOH). This information allows the controller firmware to determine exactly how long the BBU can hold the cache up and take action as needed, for example, disable write-back caching if the battery backup facilities are insufficient.

Batteries are used to hold controller power to destage data from cache to flash memory. A large capacity allows a load of 150W for a minimum of 30 minutes.

The battery can have three states:

- ▶ Full Charge: The battery is fully charged. The cache is set to Write Caching mode.
- ▶ Maintenance Charge: The battery is “trickle charging”. The cache state is unchanged.
- ▶ Learn Cycle (for Smart Battery): The battery is testing itself for capacity. The battery will retain enough charge to flush the cache. The cache state is unchanged.

Learn cycles

To properly calibrate the battery gas gauge, a fully charged smart battery unit must periodically be taken through a controlled discharge into the discharge load. The battery gas gauge is properly calibrated when the charge level decreases to a predetermined threshold. The threshold varies with the specific hardware design of the smart battery unit. Once the battery is discharged to the predetermined threshold, it is then fully recharged, following any required rest period. This controlled discharge, followed by a rest period, followed by a charge, is referred to as a *learn cycle*.

Learn cycles occur automatically at scheduled intervals. The time between learn cycles is based on the start time of each learn cycle so that the period remains constant, regardless of the length of each cycle. The learn cycle interval is scheduled in weeks (the default is 8 weeks), so that the start time for each learn cycle will occur on the same day of the week, at the same time of day.

Each controller receives backup power from one or more battery components. The controllers execute learn cycles on their respective set of battery components concurrently. In other words, controller A discharges and recharges its battery components at the same time that controller B discharges and recharges its battery components. If a controller receives backup power from more than one battery component or set of battery cells, the learn cycles for the battery components associated with that controller are executed sequentially. In other words, if there are two sets of battery cells that supply backup power to controller A, one set of cells for that controller is discharged/recharged before the second set of cells is discharged/recharged for that controller. In a duplex configuration, controller B is performing sequential discharges and recharges on its battery components at the same time that controller A is performing sequential discharges and recharges on its battery components.

The batteries are not totally discharged during a learn cycle. There is always power left to supply backup power for destaging data from cache to flash memory.

A controlled discharge begins once the following conditions are true:

- ▶ All batteries are fully charged.
- ▶ No batteries are overheated.

3.4.5 DS5000 storage subsystem rear view

The rear of the DS5000 shows the two controllers stacked in a horizontal fashion. Controller A is located on the top and controller B is located on the bottom. The DS5100 and DS5300 system can have two Host Interface Cards in total. Both controllers are hot-swappable (Figure 3-18 and Figure 3-19 on page 46).

The rear view of the subsystem varies depending on the version of HIC cards you have installed. Host connection is described in more detail in 3.4.7, “DS5000 storage subsystem host-side connections” on page 55.

You can have a:

- ▶ Fibre Channel version, which has one or two FC HICs installed per controller (Figure 3-18).
- ▶ iSCSI Version, which has one or two iSCSI HICs installed per controller (Figure 3-19 on page 46).
- ▶ Mixed version, which has one FC and one iSCSI HIC installed.

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.

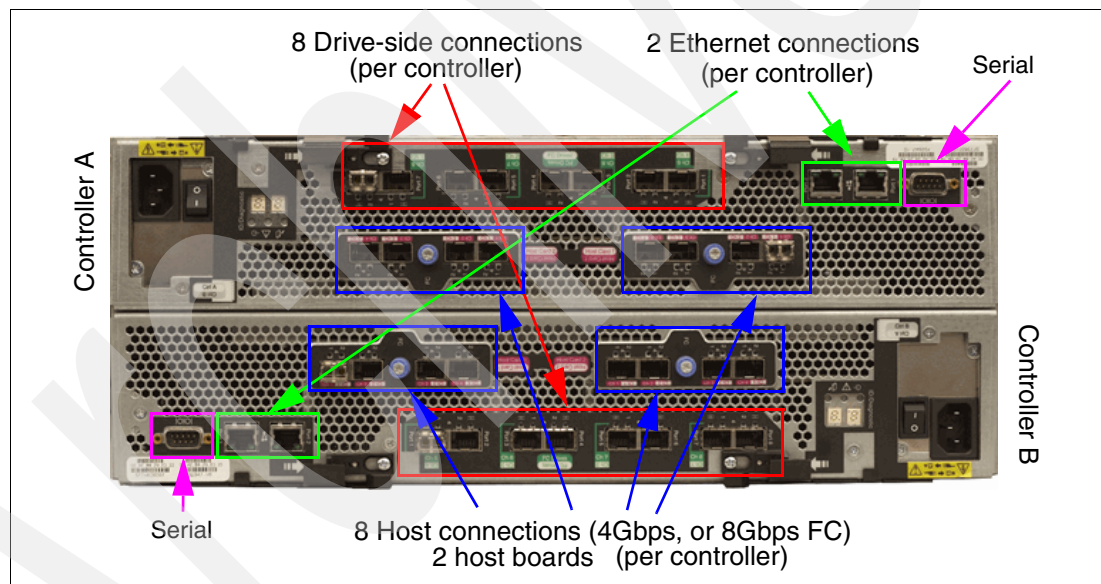


Figure 3-18 Rear view of the DS5300 storage subsystem: FC version

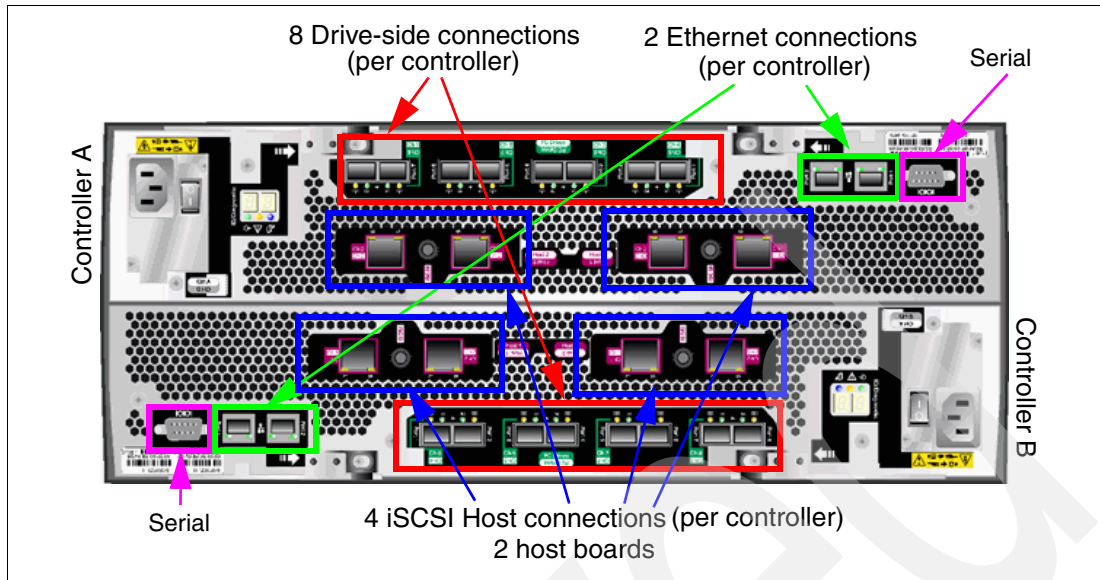


Figure 3-19 Rear view of the DS5300 storage subsystem: iSCSI version

Each controller is also equipped with two Ethernet RJ45 connectors and one serial port DB-9. These controllers are discussed in 3.4.10, “DS5100 and DS5300 storage subsystem additional connections” on page 65.

3.4.6 DS5000 storage subsystem LED indicator lights

LED indicator lights allow the DS5000 to communicate with the user. There are four main components with LEDs: front bezel panel, RAID controllers, controller support modules, and the interconnect module.

Front bezel LEDs

Figure 3-20 shows the front bezel panel of the DS5000 storage subsystem and its LED indicator lights.

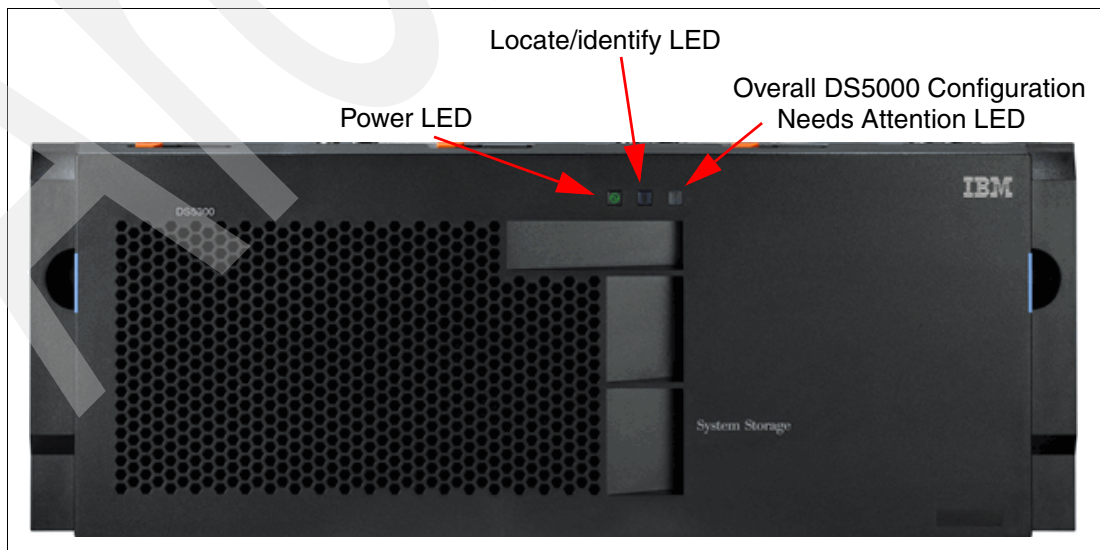


Figure 3-20 Front bezel LEDs

Important: The displayed order of the Overall Configuration Needs Attention and Locate/Identify LEDs on the interconnect-battery unit are reversed when the bezel is removed. See “Interconnect module LEDs” on page 54 for more information.

The LEDs are:

- ▶ Power LED (green):
 - On: Storage subsystem is powered on.
 - Off: Storage subsystem is powered off.
- ▶ Locate/identify (blue):
 - Off: Normal status.
 - On: Storage subsystem locate.

Note: This LED is shown as white (and displayed in a different order) on the interconnect-battery unit when the DS5000 storage subsystem bezel is removed.

- ▶ Overall DS5000 storage subsystem configuration needs attention (amber):
 - Off: Normal status.
 - On: One or more failures exist either in the storage system chassis or expansion enclosures.

RAID controller LEDs

The LEDs on the RAID controllers serve as indicators of key information (Figure 3-21).

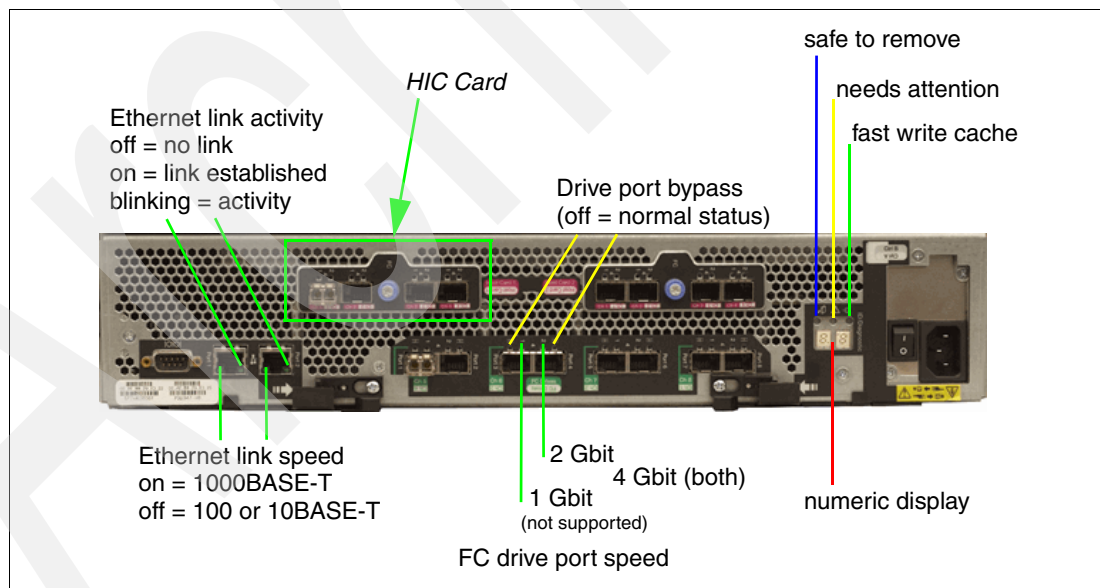


Figure 3-21 RAID controller LEDs (controller B): FC HIC version

Figure 3-25 on page 49 labels each LED (see the descriptions of each LED number after Figure 3-25 on page 49).

There are three different HIC cards available. The following figures illustrate the LEDs and their usage for each host port:

- ▶ Figure 3-22 describes the LEDs for a 4 GBps FC HIC.
- ▶ Figure 3-23 describes the LEDs for a 8 Gbps HIC.
- ▶ Figure 3-24 on page 49 describes the LEDs for a iSCSI HIC.

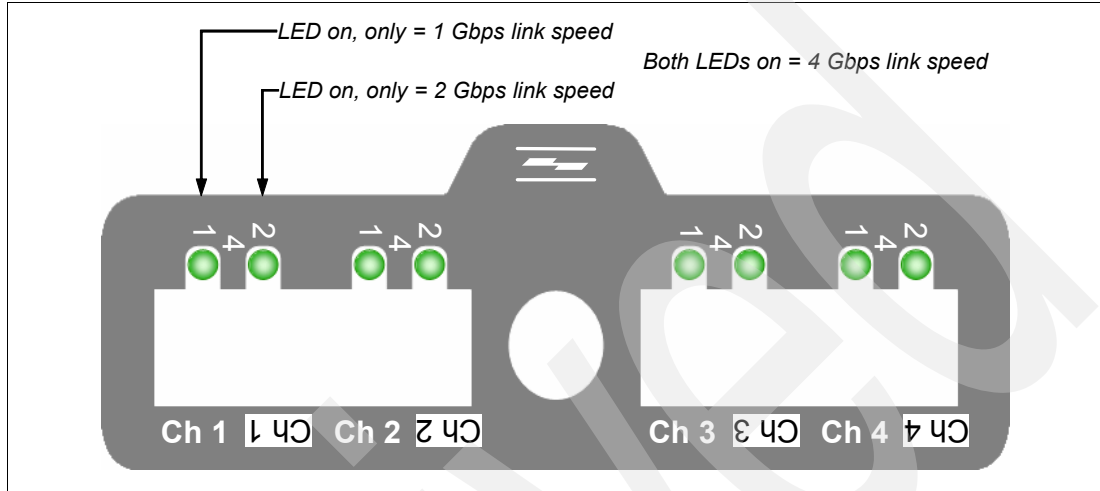


Figure 3-22 4 Gbps HIC faceplate design

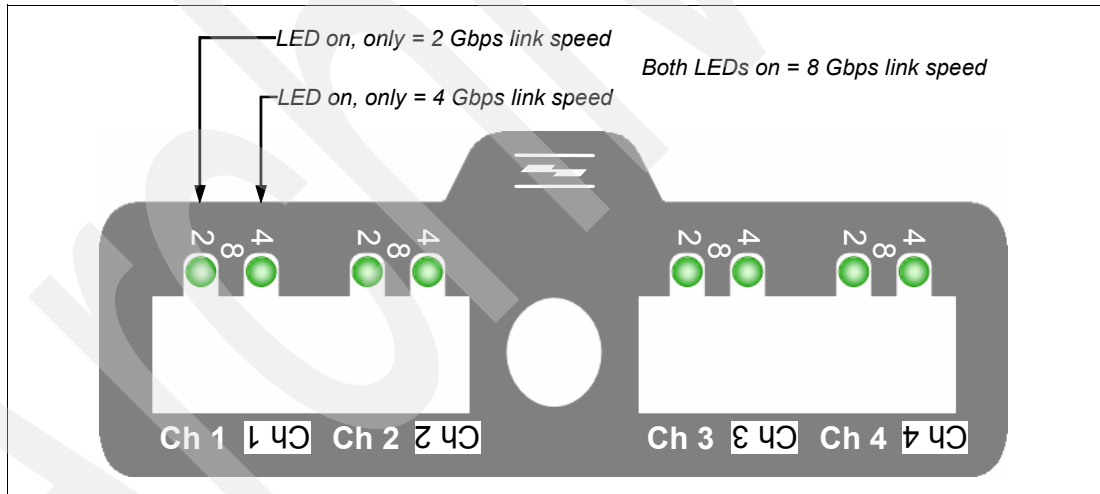


Figure 3-23 8 Gbps faceplate design

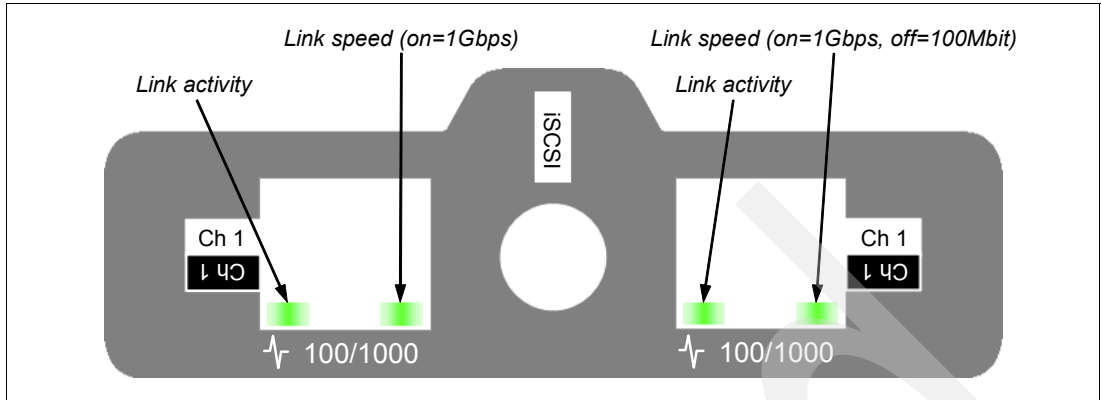


Figure 3-24 iSCSI faceplate design

Figure 3-25 shows details about the controller LED and its status.

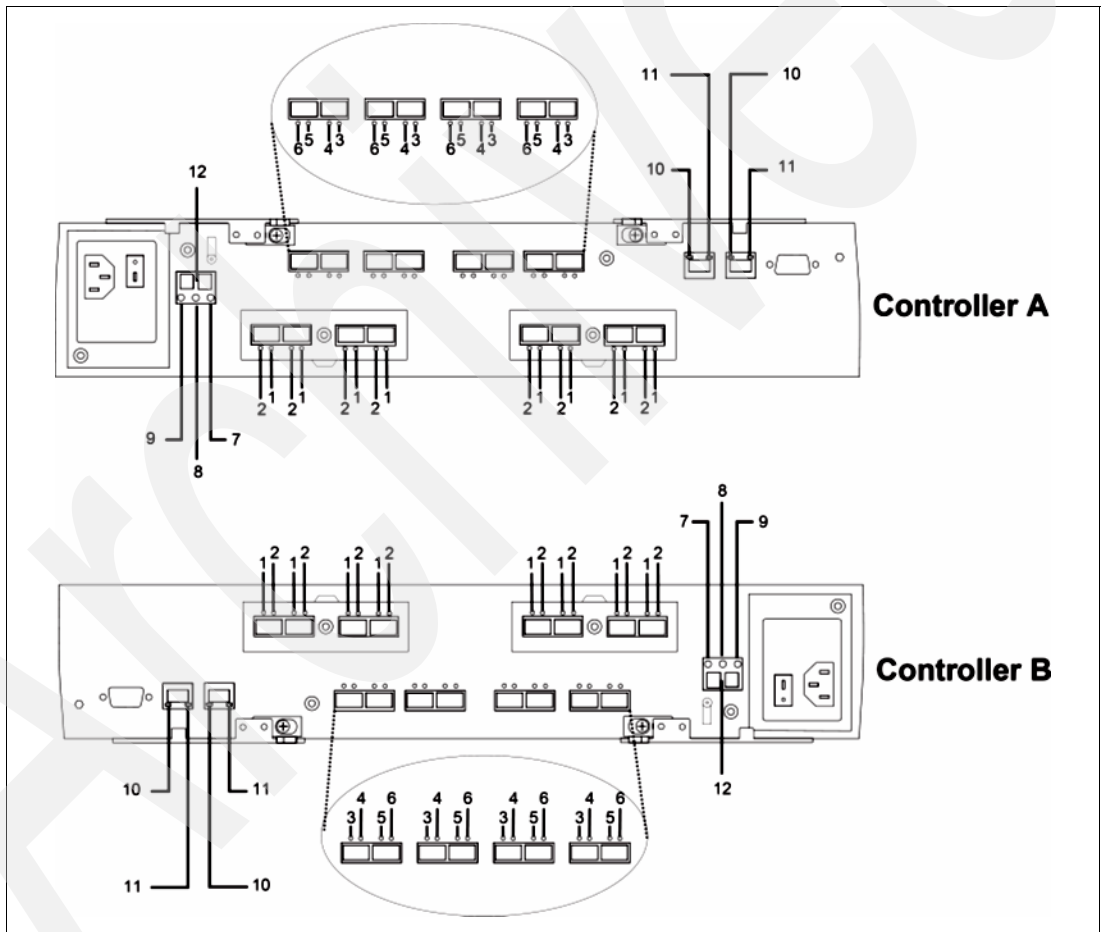


Figure 3-25 DS5000 storage subsystem RAID controller rear LEDs

The LEDs are:

- ▶ LED #1 and LED #2. See Table 3-2 on page 51.
- ▶ LED #3 (amber): Drive port bypass:
 - Off: Normal status.
 - On: Drive port bypass problem.
- ▶ LED #4 (green): Drive channel speed L1:
 - Speed L1. See Table 3-2 on page 51.
- ▶ LED #5 (green): Drive channel speed L2:
 - Speed L2. See Table 3-2 on page 51.
- ▶ LED #6 (amber): Drive port bypass:
 - Off: Normal status.
 - On: Drive port bypass problem.
- ▶ LED #7 (blue): Service action allowed:
 - Off: Normal status.
 - On: Safe to remove.
- ▶ LED #8 (amber): Needs attention:
 - Off: Normal status.
 - On: Controller needs attention (controller fault or controller is offline).
- ▶ LED #9 (green): Cache active:
 - On: Data in cache.
 - Off: No data in cache.
- ▶ LED #10 (green): Ethernet link speed:
 - Off: 100BASE-T or 10BASE-T.
 - On: 1000BASE-T.
- ▶ LED #11 (green): Ethernet link activity:
 - Off: No link established.
 - On: Link established.
 - Blinking: Activity.
- ▶ LED #12 (green/yellow): Numeric display (enclosure id/diagnostic display):
 - Diagnostic LED: On = Diagnostic code is displayed.
 - Diagnostic LED: Flashing = Controller enclosure ID is displayed.

See “Numeric display LEDs” on page 51 for more information.

Table 3-2 Host and drive channel LED definition

HIC version	L1 (Label)	L2 (Label)	Definition
4 Gbps FC	Off (1 Gbps)	Off (2 Gbps)	When both LEDs for a host or drive channel are off, this indicates one or more of the following conditions: <ul style="list-style-type: none"> ▶ The host or drive channel ports are bad. ▶ An SFP module is inserted with no Fibre Channel cable attached. ▶ No SFP module is inserted in one or both of the host or drive ports in the channel.
	On (1 Gbps)	Off (2 Gbps)	The host channel is operating at 1 Gbps.
	Off (1 Gbps)	On (2 Gbps)	The host or drive channel is operating at 2 Gbps.
	On (1 Gbps)	On (2 Gbps)	The host or drive channel is operating at 4 Gbps.
8 Gbps FC	Off (2 Gbps)	Off (4 Gbps)	When both LEDs for a host or drive channel are off, this indicates one or more of the following conditions: <ul style="list-style-type: none"> ▶ The host or drive channel ports are bad. ▶ An SFP module is inserted with no Fibre Channel cable attached. ▶ No SFP module is inserted in one or both of the host or drive ports in the channel.
	On (2 Gbps)	Off (4 Gbps)	The host channel is operating at 2 Gbps.
	Off (2 Gbps)	On (4 Gbps)	The host or drive channel is operating at 4 Gbps.
	On (2 Gbps)	On (4 Gbps)	The host or drive channel is operating at 8 Gbps.
iSCSI	Off (Link)	Off (Speed)	When both LEDs for a host channel are off, this indicates one or more of the following conditions: <ul style="list-style-type: none"> ▶ The host channel ports are bad. ▶ No ethernet connectivity.
	On/blink (Link)	Off (Speed)	100 Mbps Ethernet connection established, transfer data (blinking).
	On/blink (Link)	On (Speed)	1000 Mbps ethernet connection established, transfer data (blinking).

Numeric display LEDs

When the storage subsystem is operating normally, the numeric display shows the enclosure identification (enclosure ID) of the storage subsystem and the diagnostic LED flashes once every two seconds. The storage subsystem tray ID is normally set at the factory to either values 85 or 00. Verify that the attached storage expansion enclosures are not set to either of these enclosure IDs.

If an error has occurred and the controller Needs Attention LED is on, the numeric display shows diagnostic information. The numeric display indicates the information is diagnostic by illuminating an LED that appears as a decimal point between the display numbers. The diagnostic LED turns off when the numeric display shows the storage subsystem enclosure ID. The numeric display shows various diagnostic codes as the controllers perform the startup process after each power cycle or reset. After diagnostics are complete, the current storage subsystem enclosure ID is displayed.

Diagnostic codes in the form of Lx, where x is a hexadecimal digit, indicate controller state information. In general, these codes are displayed only when the controller is in a non-operational state. The controller might be non-operational due to a configuration problem (such as mismatched controller types), or it might be non-operational due to hardware faults.

If the controller is non-operational due to system configuration, the controller Needs Attention LED is off. If the controller is non-operational due to a hardware fault, the controller Needs Attention LED is on. The definitions for Lx diagnostic codes are listed in Table 3-3.

Table 3-3 Numeric display diagnostic codes

Value	Controller state	Description	Storage Manager view
L0	Suspend	Mismatched controller types.	Needs Attention condition for board type mismatch
L1	Suspend	Missing interconnect-battery unit.	Needs Attention condition for missing interconnect-battery unit
L2	Suspend	Persistent memory errors.	Needs Attention condition for offline controller
L3	Suspend	Persistent hardware errors.	Needs Attention condition for offline controller
L4	Suspend	Persistent data protection errors.	Needs Attention condition for offline controller
L5	Offline	The alternate controller has incompatible firmware, but the automatic controller firmware synchronization (ACS) cannot be performed.	Needs Attention condition for offline controller
L7	Suspend/Offline	A controller with a different controller submodel ID is inserted.	Needs Attention for offline controller
L8	Offline	Unsupported memory is present or memory is not populated in the correct memory slots.	Need Attention because the controller is in offline/failed state
88	Reset	A controller is held in reset by an alternate controller.	N/A

Controller support module LEDs

The controller support modules are located on the left side and right side in the front section of the DS5000 storage subsystem, behind the front bezel panel. The two modules each contain the power supplies and fans that are needed to operate the DS5000 storage subsystem. The LED positions on the right and left power supply and fan units are in mirrored positions. The LED indicator lights are shown in Figure 3-26.

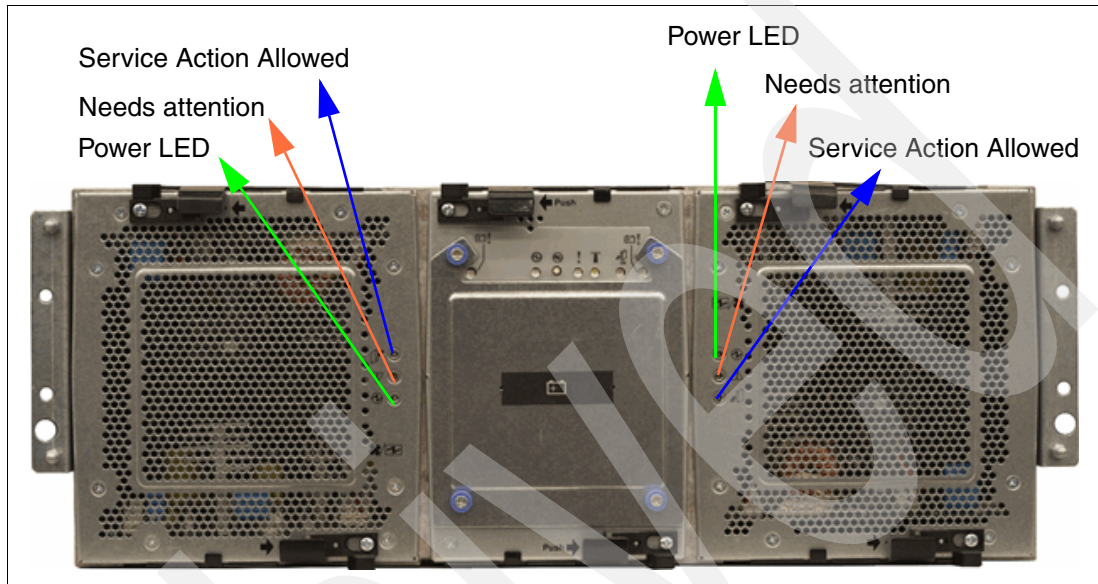


Figure 3-26 Controller support module (power supply/fan) LEDs

Note: The right power supply and fan unit is linked with RAID controller A. The left power supply and fan unit is linked with RAID controller B in the DS5000 storage subsystem.

The LEDs are:

- ▶ Power LED (green):
 - On: Power supply and fan unit is providing power.
 - Off: Power supply and fan unit is not providing power.
- ▶ Needs attention (amber):
 - Off: Normal status.
 - On: Power supply and fan unit needs attention.
- ▶ Service action allowed (blue):
 - Off: Normal status.
 - On: Safe to remove.

Interconnect module LEDs

The interconnect module is located in the forward section of the DS5000 storage subsystem, behind the front bezel panel. It is located in between the two controller support modules. It holds the cache batteries and the removable midplane. The LED indicator lights are shown in Figure 3-27.

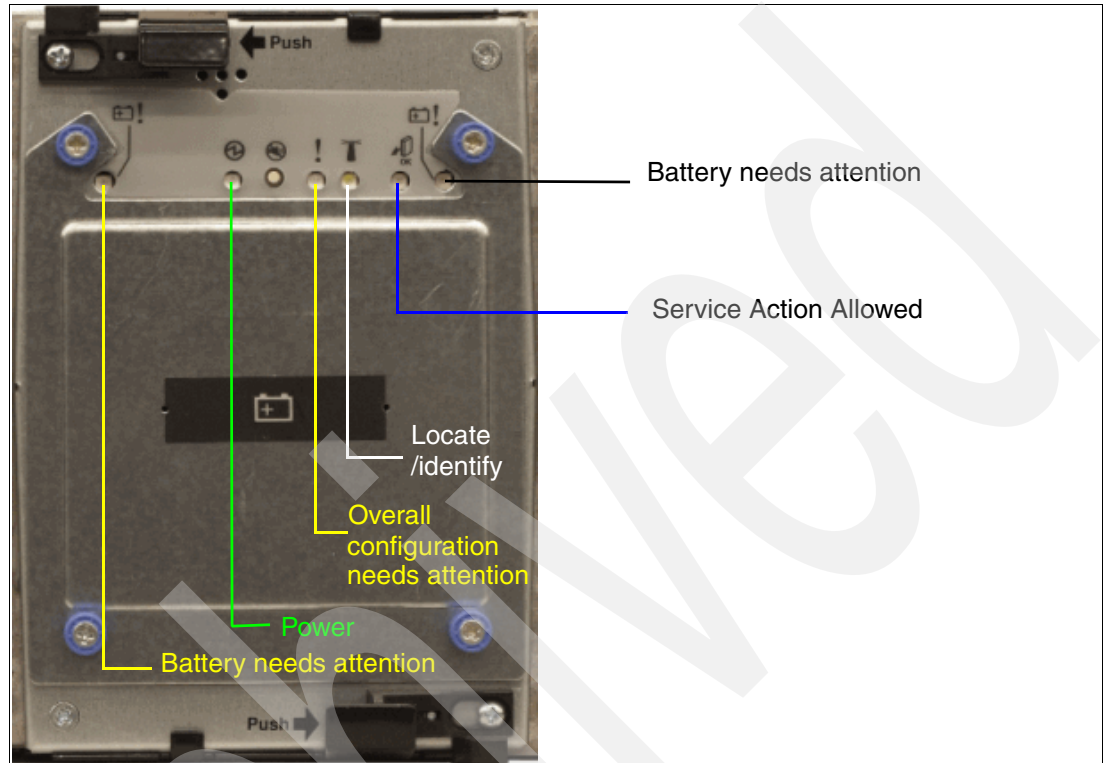


Figure 3-27 Interconnect module LEDs

The LEDs are:

- ▶ Battery needs attention (amber):
 - Off: Normal status.
 - On: Battery failed.
- ▶ Power LED (green):
 - On: Storage subsystem is powered on.
 - Off: Storage subsystem is powered off.
- ▶ Overall DS5000 storage subsystem configuration requires attention (amber):
 - Off: Normal status.
 - On: A Component in the storage system has developed a fault.
- ▶ Locate/Identify (white, appears as blue when front bezel is installed):
 - Off: Normal status.
 - On: Storage subsystem locate.
- ▶ Service action allowed (blue):
 - Off: Normal status.
 - On: Safe to remove.

3.4.7 DS5000 storage subsystem host-side connections

The DS5000 storage subsystem integrates the host-side and drive-side connections into the controller itself. DS4000 models use built-in host ports while DS5000 models use flexible Host Interface Cards that can be replaced by IBM service personnel only. Each DS5000 controller holds up to two Host Interface Cards (HICs) (see Figure 3-28).

These HICs are currently available:

- ▶ 4 Gbps FC HIC (four ports per HIC)
- ▶ 8 Gbps FC HIC (four ports per HIC)
- ▶ 1 Gbps iSCSI HIC (two ports per HIC)

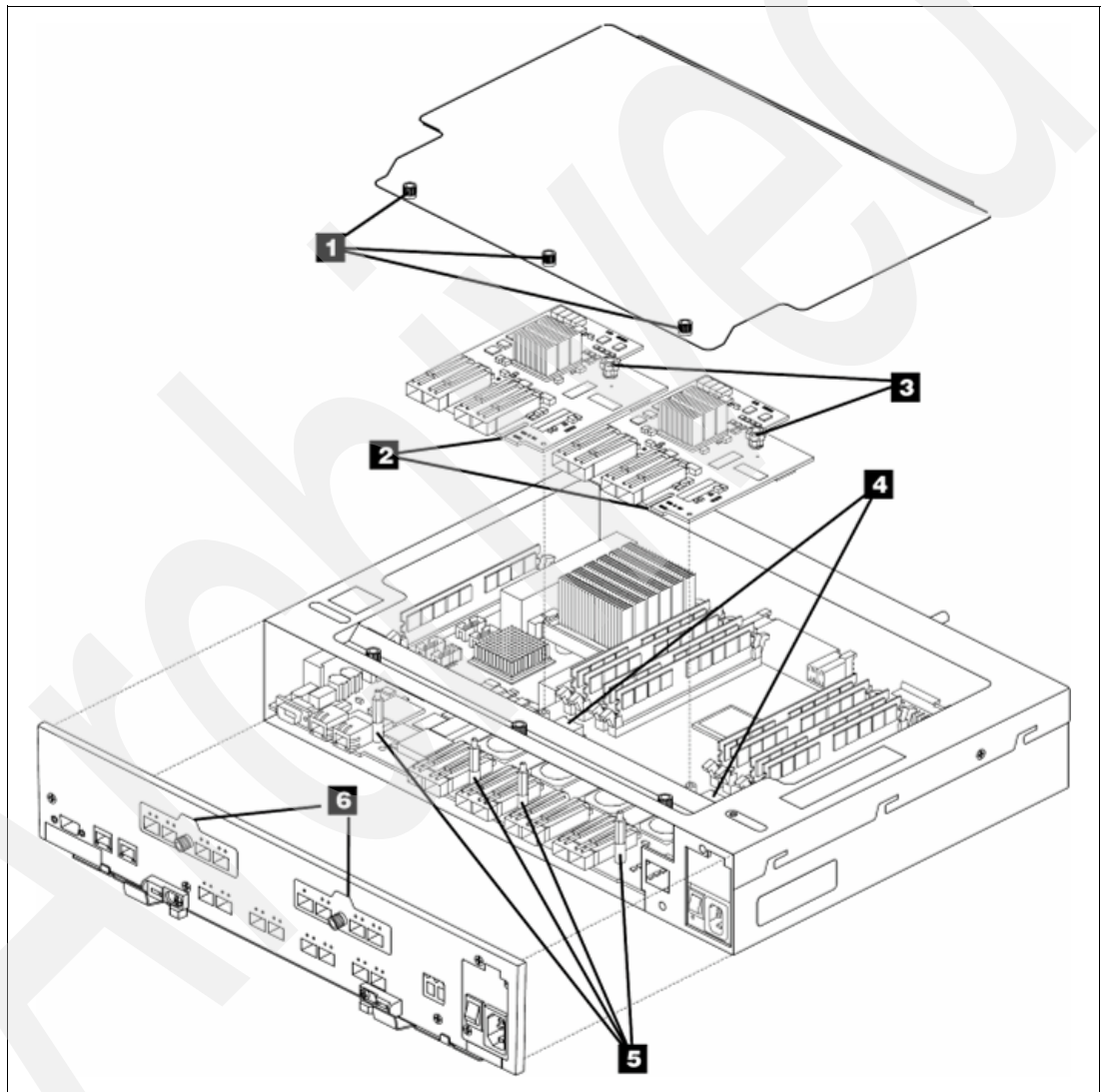


Figure 3-28 Controller exposed

The HICs noted as (2) are shown in Figure 3-28. Thumb screws (3) support easy replacement. Different faceplates (6) are used to describe the LED status of the host port LEDs. The faceplates are described in Figure 3-22 on page 48, Figure 3-23 on page 48, and Figure 3-24 on page 49.

The DS5100 and DS5300 holds up to sixteen host connections.

Host connections support Fibre Channel or iSCSI attachment (depending on the options ordered) through switches and direct connections. The host ports are labeled sequentially from 1 through 4, from left to right, on controller B (bottom) for both host cards. Conversely, they are labeled in reverse order, from 4 to 1, from the left to the right on controller A (top). As previously indicated, this is because controller A is installed “upside-down” relative to controller B.

Having sixteen independent host ports allows us to establish fully redundant direct connections to up to eight hosts.

It is important to match up host or fabric connections to the DS5000 storage subsystem by attaching one connection to each controller. In doing so, you take advantage of the DS5000 storage subsystem’s ability to fail over and distribute the workload among the two controllers. For any given host, make sure to connect to the same host port number in the same host card on each controller. Remember that the right most host port on controller A and the left most host port on controller B are both host port #1, as shown in Figure 3-29 on page 57.

Important: The DS5000 storage subsystem does not support a direct connection to a host if it is only connected to one of the two controllers.

Host ports channels are numbered in the following way (see Figure 3-29 on page 57 and Figure 3-30 on page 57):

- ▶ FC HIC 1 ports 1-4 have channel numbers 1-4.
- ▶ FC HIC 2 ports 1-4 have channel numbers 5-8 (in FC only version).
- ▶ FC HIC 2 ports 1-4 have channel numbers 3-6 (in mixed host type version).
- ▶ iSCSI HIC 1 ports 1 and 2 have channel numbers 1 and 2.
- ▶ iSCSI HIC 2 ports 1 and 2 have channel numbers 3 and 4 (in iSCSI only version).
- ▶ iSCSI HIC 2 ports 1 and 2 have channel numbers 5 and 6 (in mixed host type version).

According to Figure 3-29, host ports in controller B have numbers 1-8 from left to right, while the ports in controller A have them from right to left.

The DS5000 storage subsystem also fully supports Fibre Channel switched connections. Figure 3-29 shows how the DS5000 would be connected into dual-redundant fabrics.

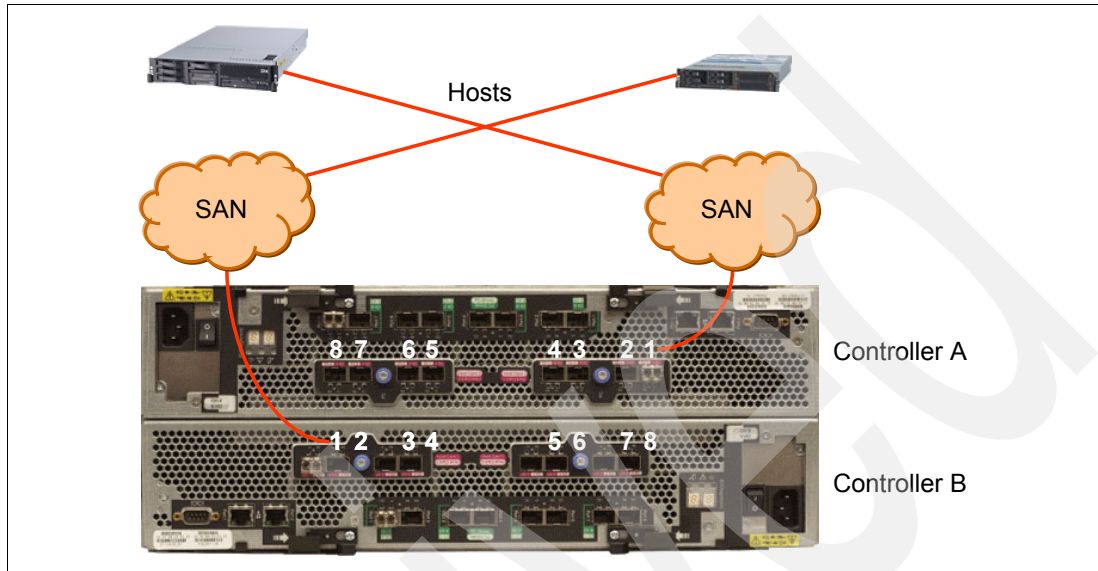


Figure 3-29 SAN connected hosts to DS5000 storage subsystem

According to Figure 3-30, host ports in controller B have numbers 1-6 from left to right, while the ports in controller A have them from right to left.

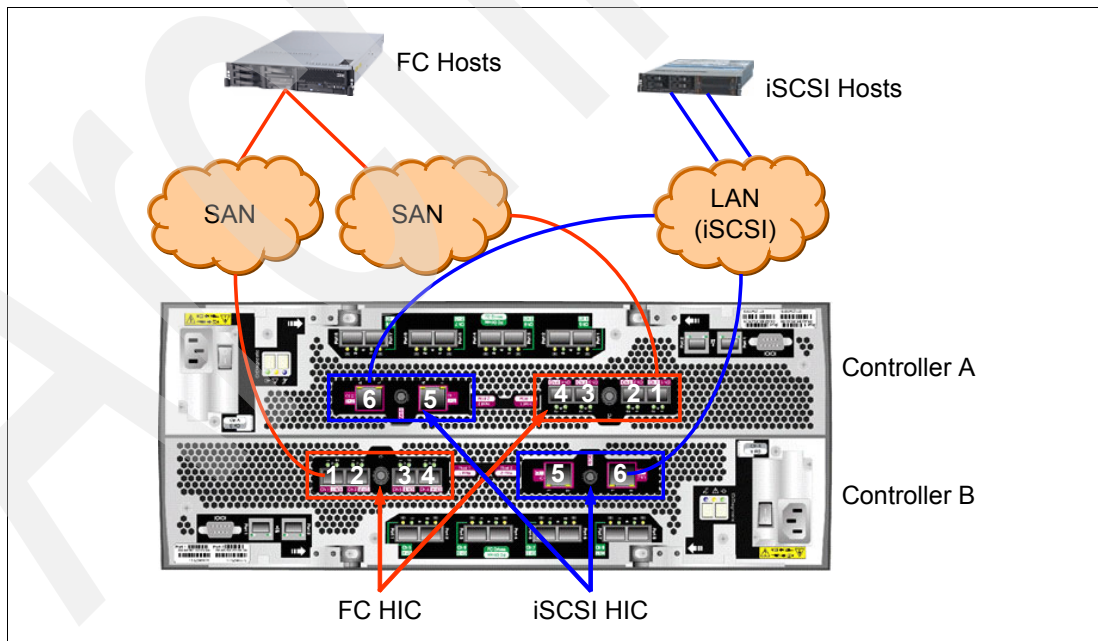


Figure 3-30 Mixed host type HICs

Figure 3-31 shows that the host ports are connected to eight sets of FC host bus adapters (HBAs). You can replace one or all of these sets of FC HBAs with FC switches as required. In our example, HBA 1 of the host AIX is connected to port 1 in the host card 1 in RAID controller A. HBA 2 is connected to the same port and card number in controller B.

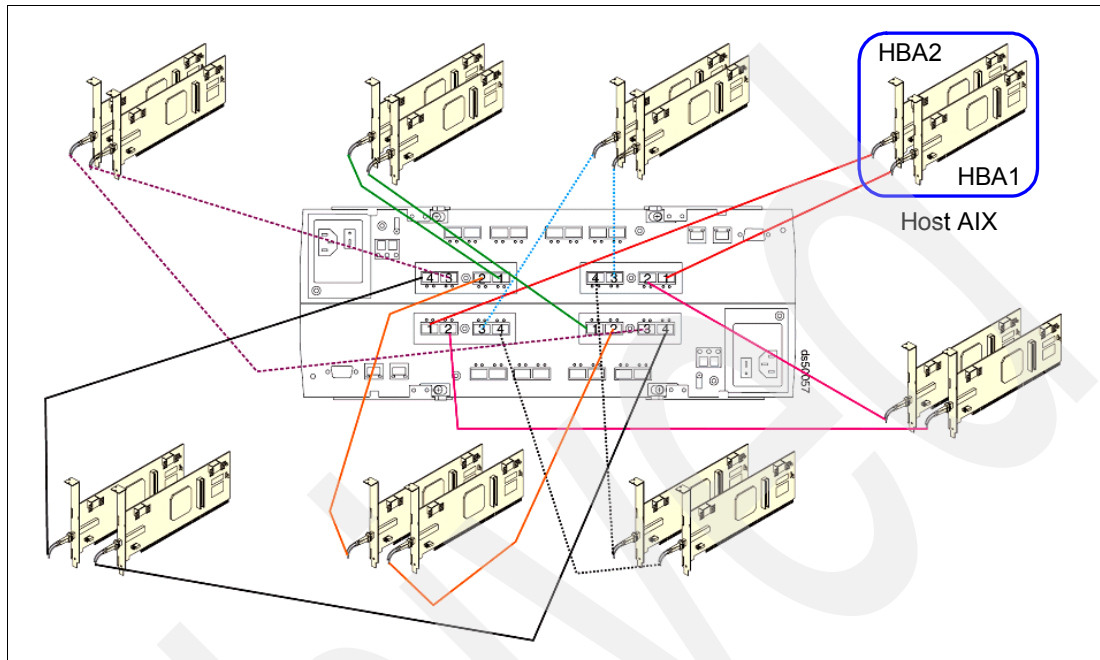


Figure 3-31 Directly connected hosts to the DS5000 storage subsystem

In practice, there is no difference about which ports in which order (in one controller) will be connected to hosts or switches. However, we recommend connecting hosts or switches in the following order:

Models 5100 and 5300 with eight host ports (single Host Interface Card per controller):

1. Controller A, host card 1, port 1 <-> controller B, host card 1, port 1
2. Controller A, host card 1, port 3<-> controller B, host card 1, port 3
3. Controller A, host card 1, port 2<-> controller B, host card 1, port 2
4. Controller A, host card 1, port 4<-> controller B, host card 1, port 4

Models 5100 and 5300 with 16 host ports (two Host Interface Cards per controller):

1. Controller A, host card 1, port 1 <-> controller B, host card 1, port 1
2. Controller A, host card 2, port 1 <-> controller B, host card 2, port 1
3. Controller A, host card 1, port 3<-> controller B, host card 1, port 3
4. Controller A, host card 2, port 3<-> controller B, host card 2, port 3
5. Controller A, host card 1, port 2<-> controller B, host card 1, port 2
6. Controller A, host card 2, port 2<-> controller B, host card 2, port 2
7. Controller A, host card 1, port 4<-> controller B, host card 1, port 4
8. Controller A, host card 2, port 4<-> controller B, host card 2, port 4

In this situation, ports dedicated for Enhance Remote Mirroring (ERM) (refer to *IBM Midrange System Storage Copy Services Guide*, SG24-7822 for more details) are used last. Another reason to connect the ports in the suggested order is easier manageability and higher availability (in case of a failure of a host card, you only lose half of the connections to the controller instead of all of them).

Note: When ERM is enabled, the following ports are dedicated for replication (which stops the host I/O):

- ▶ Models with eight FC host ports: FC HIC, FC port 4 in both controllers
- ▶ Model with 16 FC host ports: Host card 2 and port 4 in both controllers

iSCSI ports are not supported for ERM.

Rule of Thumb: The latest FC host port is always dedicated for mirror traffic if ERM is enabled.

3.4.8 DS5000 storage subsystem drive-side connections

The drive-side connections operate at up to 4 Gbps (2 Gbps or 4 Gbps) and allow connection of disk expansion enclosures to the base controller unit.

There are 16 total drive-side connections (eight on each controller). The numbering scheme for the drive connections is structured like the host connections. Because controller A is upside-down, the left-to-right numbering of the drive connection ports is reversed. This means controller A is numbered left to right, 8 through 1, in reverse sequential order. Controller B is numbered left to right, 1 through 8, in forward sequential order (see Figure 3-32).

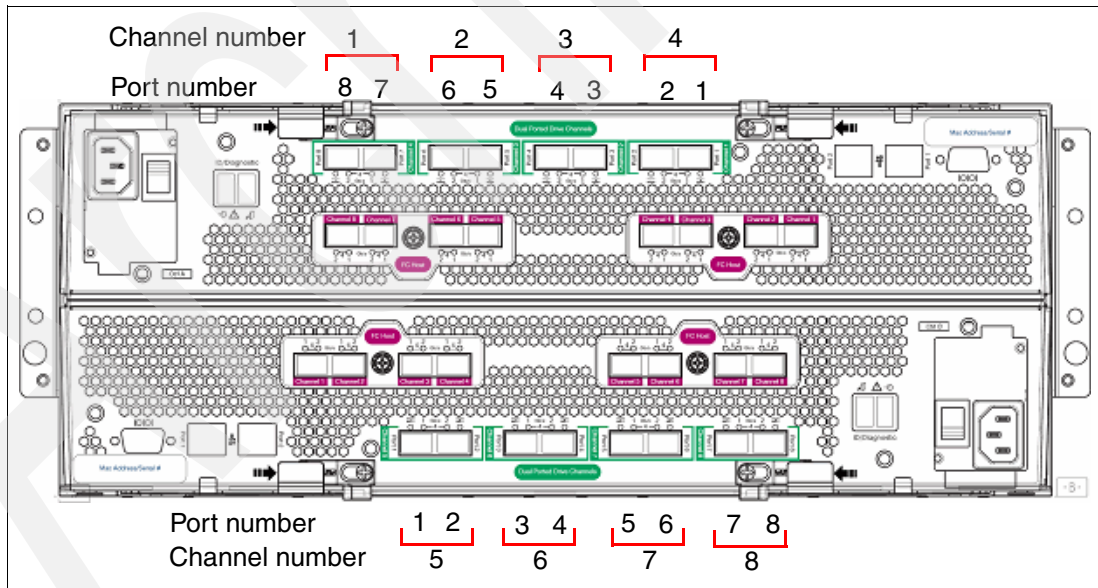


Figure 3-32 Disk drive channel and port numbering

The DS5000 storage subsystem supports eight redundant drive channel pairs on which to place expansion enclosures. Having many drive channel pairs allows you to achieve high bandwidth and linear scalability. The FC-AL standard allows you to connect 127 devices together into a loop. It is important not to have a fully utilized FC-AL to fulfill the linear performance scalability requirement. The DS5000 storage subsystem allows attachment of 448 drives. Many disk loops bring the DS5000 system to an unprecedented level of availability, with the ability to spread out disk enclosures over eight back-end drive channels. For 128 and fewer drives, you can have a dedicated disk loop for each disk enclosure.

Ports 1 and 2 on each controller are grouped together in one drive channel group. Similarly, ports 3 and 4, 5 and 6, and 7 and 8 are grouped together in other drive channel groups. If you look at the rear of a properly installed DS5000 storage subsystem, you will see them clearly labeled. In this case:

- ▶ Ports 8 and 7 on controller A are channel group 1.
- ▶ Ports 6 and 5 on controller A are channel group 2.
- ▶ Ports 4 and 3 on controller A are channel group 3.
- ▶ Ports 2 and 1 on controller A are channel group 4.
- ▶ Ports 1 and 2 on controller B are channel group 5.
- ▶ Ports 3 and 4 on controller B are channel group 6.
- ▶ Ports 5 and 6 on controller B are channel group 7.
- ▶ Ports 7 and 8 on controller B are channel group 8.

The two ports on each drive channel group must run at the same speed. There is no blocking between the two adjacent ports at the drive channel group level. It is best to spread out the drive-side channel pairs among the channel groups to ensure maximum availability. Each channel (two ports) is internally connected to a 4 Gbit FC port in Controller A and Controller B. Each controller has two quad ports disk chips for disk connections. One chip is dedicated for local ports, with one link per channel. The second one for remote ports (in the second controller) has one link per channel. Refer to Figure 3-6 on page 32 and Figure 3-10 on page 36 for architectural details.

A drive-side channel pair is made up of one port from each controller, going left to right. For example, drive channel pair 1 is composed of controller A, port 8, and controller B, port 1. Drive channel pair 2 is composed of controller A, port 7, and controller B, port 2, and so on.

The recommended pairing is driven by the connections between FC loop switches within controllers A and B. For attaching the first four enclosures, we recommend using one port of each channel. Because of easier management, performance, and availability, we propose connecting channel pairs in the following order (refer to Figure 3-33 on page 61):

1. Disk port 8 controller A <-> disk port 1 controller B
2. Disk port 6 controller A <-> disk port 3 controller B
3. Disk port 4 controller A <-> disk port 5 controller B
4. Disk port 2 controller A <-> disk port 7 controller B
5. Disk port 7 controller A <-> disk port 2 controller B
6. Disk port 5 controller A <-> disk port 4 controller B
7. Disk port 3 controller A <-> disk port 6 controller B
8. Disk port 1 controller A <-> disk port 8 controller B

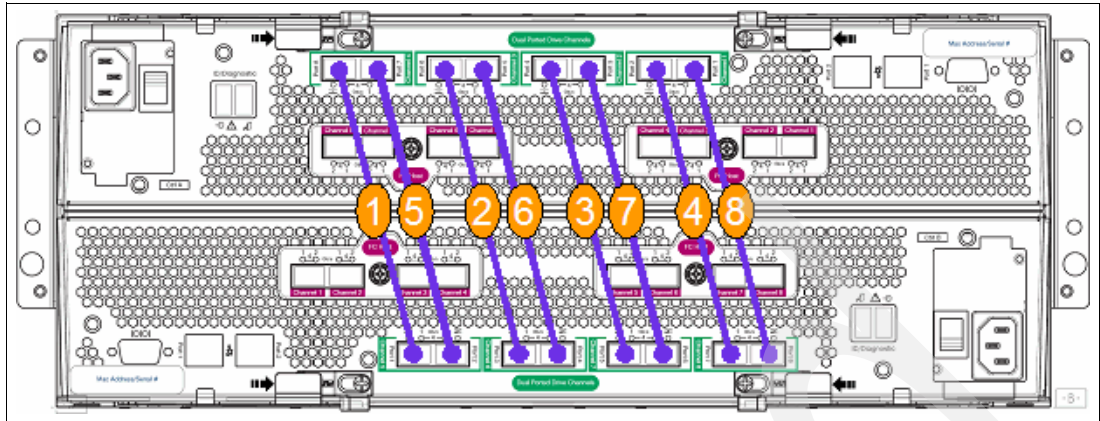


Figure 3-33 Enclosure connections order

Note: The numbering in Figure 3-33 is to help with the growth of a subsystem by adding enclosures. It is not indicative of channel or port numbers.

3.4.9 DS5100 and DS5300 storage subsystem drive-side cabling

The following rules apply while cabling the DS5000 with an expansion enclosure:

- ▶ The maximum number of expansion enclosures (EXP5000 or EXP810) in a drive loop is limited to seven, as the maximum 112 disks limitation is reached with seven enclosures configured with 16 drives each.
- ▶ Each FC-drive and expansion unit (EXP) in the drive channel must operate at the same Fibre Channel speed (either 2 Gbps or 4 Gbps). SATA drives autonegotiate to the speed of the EXP.
- ▶ The DS5000 controller drive port must always be connected to the EXP5000 port labelled 1B. Because the left and right EXP5000 ESMs (ESMs A and B) are inserted in the ESM bays in different orientations, ensure that you use the port labeled 1B, as shown in Figure 3-34. This applies to the previous DS4800 and DS4700 with its expansion units in the same way. Port 1A is used to daisy-chain the next expansion units, which we refer to later in this section

Note: To achieve the best performance, spread all your expansion units among all port channels.

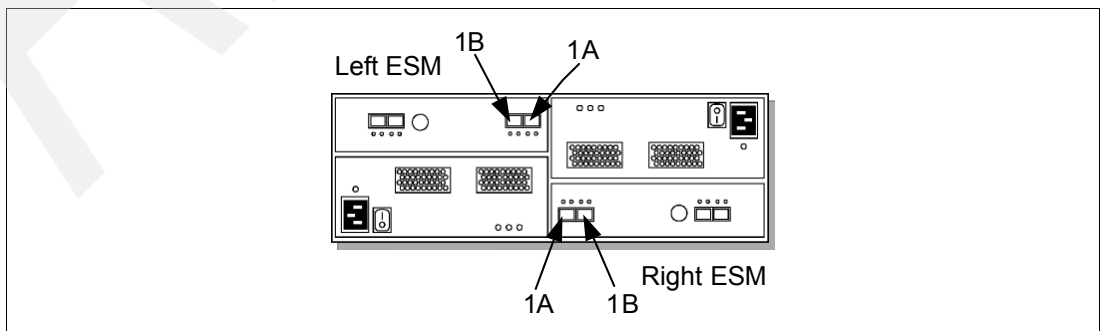


Figure 3-34 Port labels on EXP5000

The DS5000 storage subsystem can attach up to 28 EXP5000s or EXP810s or a intermix of the two for migration purposes. It also fully supports an intermix of the FC and SATA drives inside enclosures to allow users maximum flexibility for customizing storage solutions. It is generally best to spread enclosures evenly among the eights drive channel pairs as you scale up your DS5000 storage subsystem's storage capacity. This allows you to fully utilize the maximum drive-side bandwidth. For attaching the first four enclosures, use all the disk channels, with one port per channel, as shown in Figure 3-35.

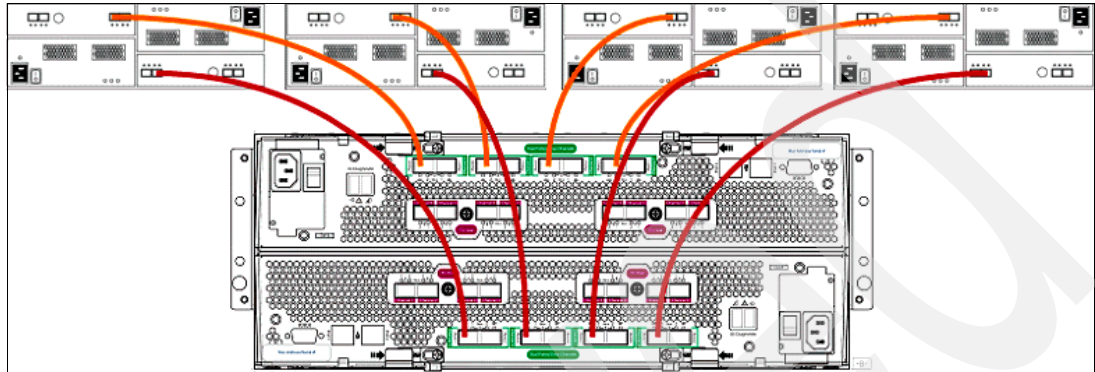


Figure 3-35 DS5000 with four disk enclosures

When attaching the next four enclosures, use the next ports in each channel, as shown in Figure 3-36. For an eight enclosure configuration, each enclosure is attached to a dedicated port pair.

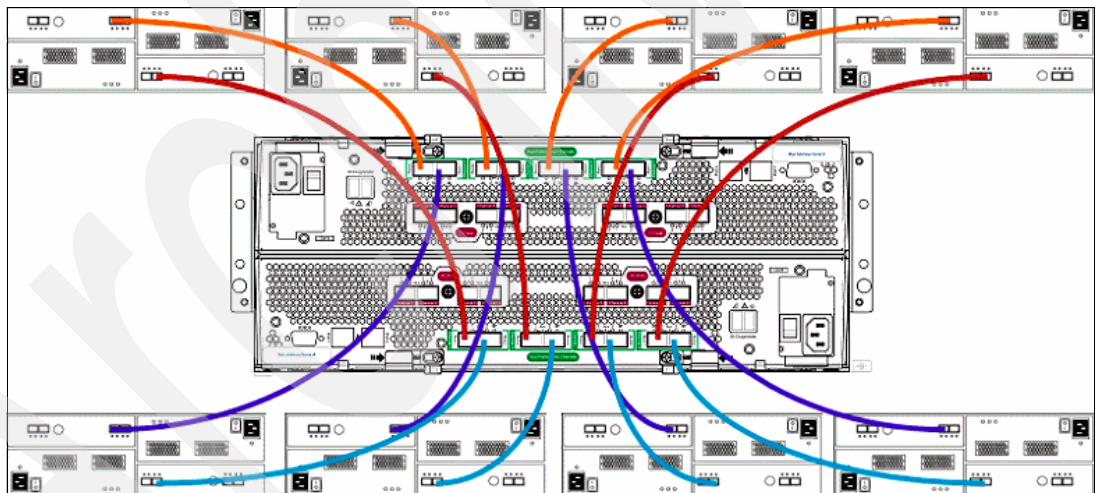


Figure 3-36 DS5000 with eight disk enclosures

When attaching the ninth or more enclosures, attach them to existing disk loops by daisy-chaining them. A 256 disk configuration should have two expansion enclosures on each drive-side channel pair and uses 16 EXPs in summary, as shown in Figure 3-37.

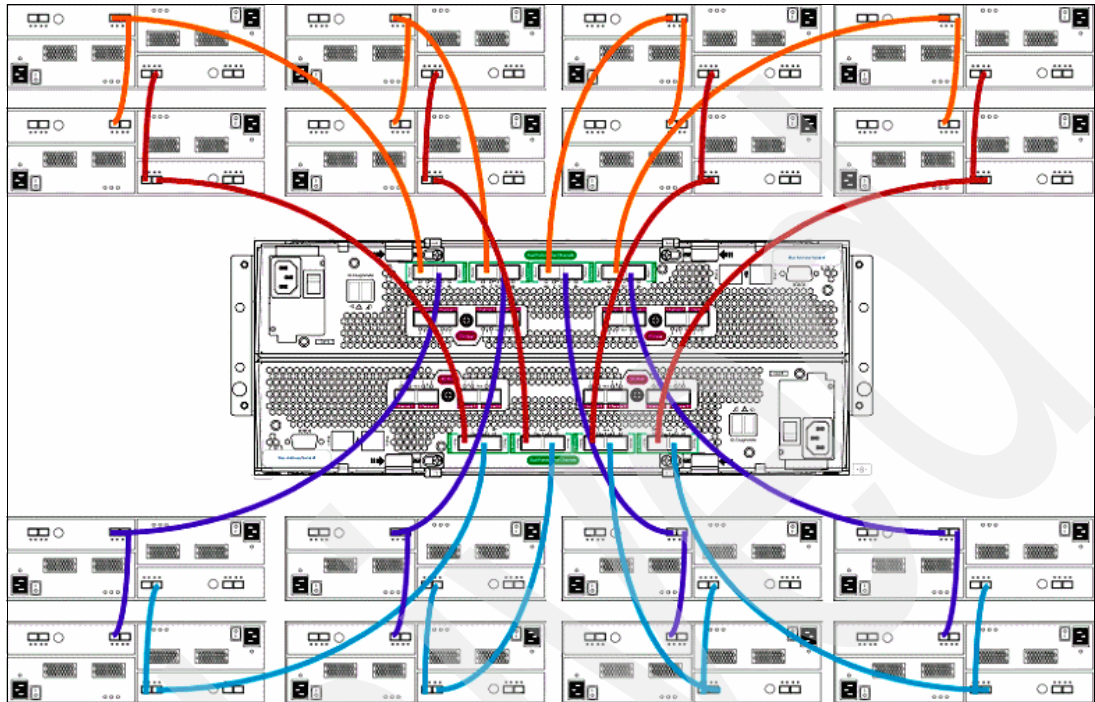


Figure 3-37 DS5000 with 16 disk enclosures

Starting with controller firmware Version 7.50, the DS5000 supports the 448 drive configuration. It requires you to attach seven expansion units (EXP5000) to one redundant port channel pair. As a result, you will have four EXPs connected to one drive port pair and three EXPs to the other drive port pair of the same channel, as shown in Figure 3-38.

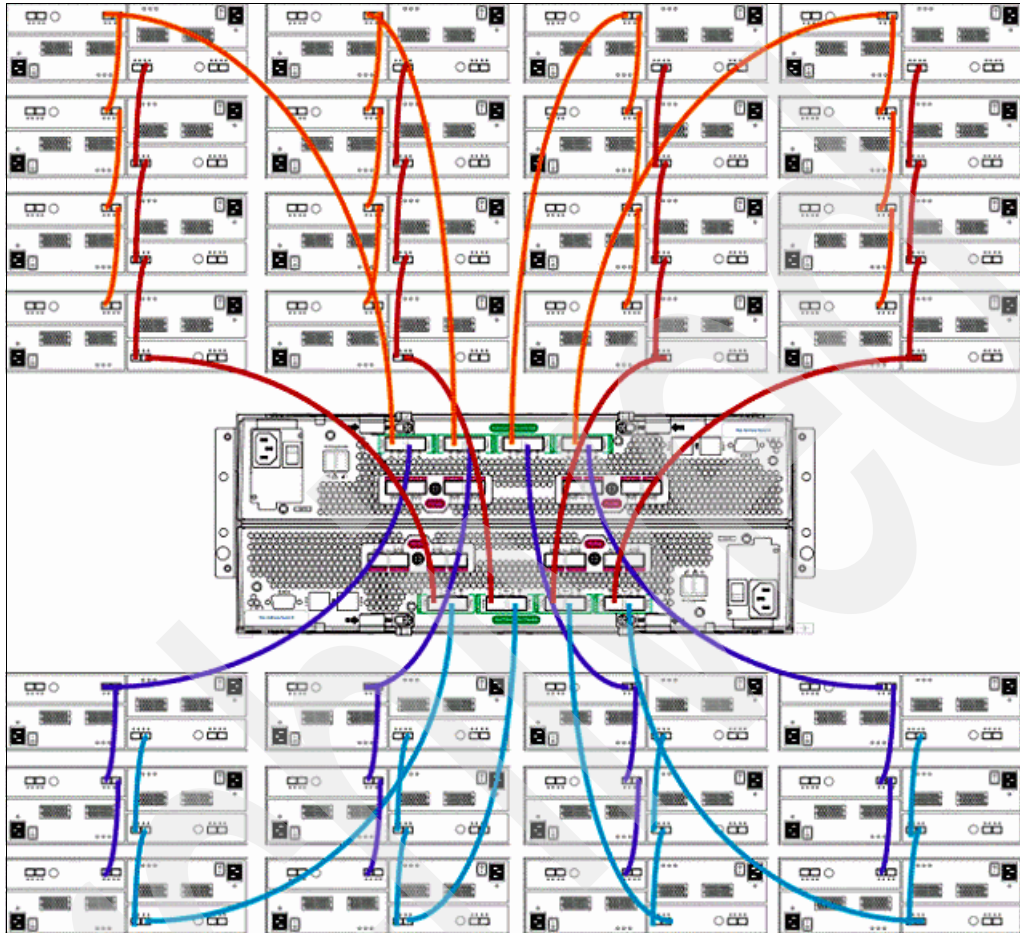


Figure 3-38 DS5000 with 28 expansion units

Note: Make sure that when you mix different speed EXPs that both disk ports in a disk channel operate at the same speed. If you attach a disk enclosure that operates at 2 Gbps, both ports in the controller's disk channel and all drives will work at 2 Gbps.

Remember to connect a 2 Gbps and 4 Gbps enclosure to different channels, as shown in Figure 3-39 on page 65.

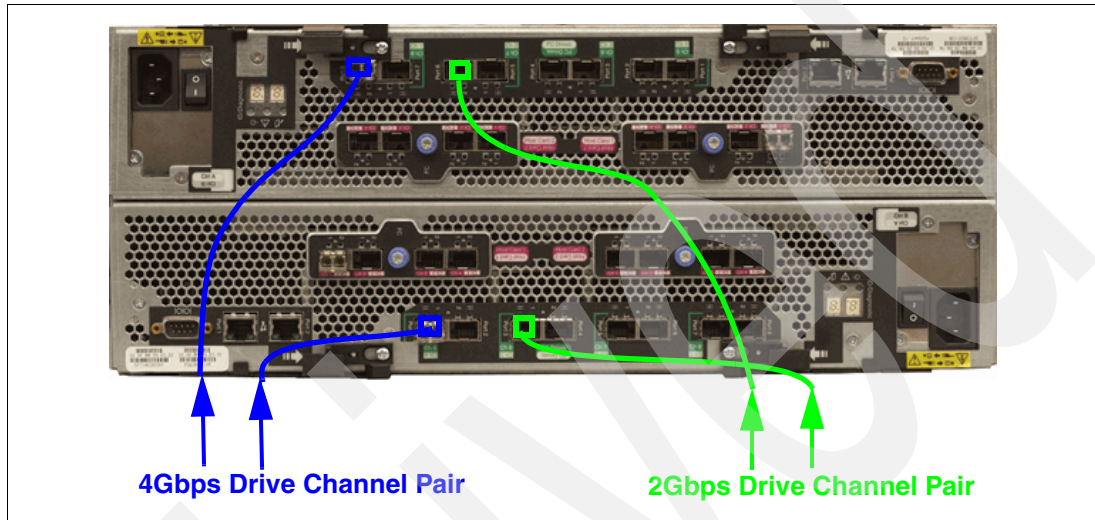


Figure 3-39 Example of a mixed drive-side environment (2 or 4 Gbps)

Additional materials regarding cabling can be found in the *IBM System Storage DS4000/DS5000 Hard Drive and Storage Expansion Enclosure Installation and Migration Guide*, GC53-1139.

3.4.10 DS5100 and DS5300 storage subsystem additional connections

In addition to host and drive connections, the DS5000 storage subsystem's rear also has connections for Ethernet and serial ports. You can manage and service the DS5000 storage subsystem through these ports. The ports are:

- ▶ RJ-45 Ethernet connector

This connector is for an RJ-45 10/100/1000BASE-Tx Ethernet connection. There are two connections per controller. One port is designed for out-of-band management and the other port is meant for serviceability. The logic behind adding an extra port was to introduce additional isolation and to separate management and service traffic from one another. Because of the extra port, it is preferable to have two IP addresses per controller in order to manage and service the DS5000 storage subsystem appropriately. However, you cannot attach this port to a routed network, because you can not set up a gateway for it. You can still operate the DS5000 storage subsystem with only one IP port active per controller. The best practice is to set Port 1 into customer network for out-of-band management and leave the Port 2 as the default in order to let IBM service personnel connect using the default IP addresses.

The default IP addresses for the controllers are shown in the Table 3-4. The default subnet mask for all four Ethernet ports is 255.255.255.0.

Table 3-4 Default IP addresses for the controllers

	Controller A	Controller B
Port 1	192.168.128.101	192.168.128.102
Port 2	192.168.129.101	192.168.129.102

► Serial 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 maximum baud rate is 115,200 bps. The default baud rate setting from the factory is 38,400 bps, N-8-1, with no flow control.

Attention: Managing the DS5000 storage subsystem through the serial interface has potential risks. Using certain commands, you can initialize the RAID controller, and therefore lose all your data. You should only use this interface when instructed to do so by IBM Support.

3.5 DS5020 storage subsystem

The IBM System Storage DS5020 disk system, shown in Figure 3-40, is the newest midrange disk storage offering from IBM. It is designed to deliver high performance, advanced functionality, high availability, and modular and scalable storage capacity.

The DS5020 disk system is designed to deliver up to a two-fold IOP performance increase over the current-generation DS4700 offerings and represents the eighth generation architecture within the midrange disk family.

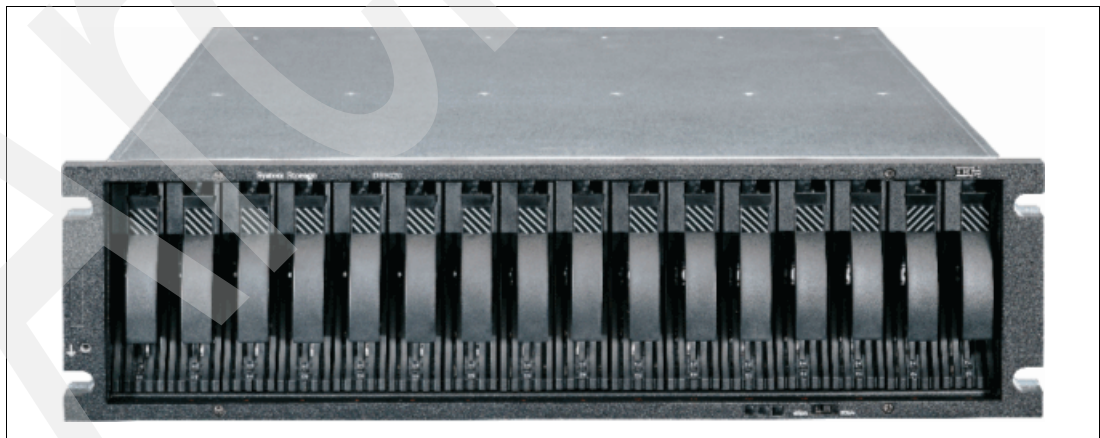


Figure 3-40 IBM System Storage DS5020 storage subsystem

DS5020 disk system members

- DS5020 Disk Controller (1814-20A)
- EXP520 Storage Expansion Unit (1814-52A)

Disk drives

- ▶ Up to 112 TB physical storage capacity.
- ▶ Accommodates up to 16 disk drives installed within the DS5020 enclosure.
- ▶ Attachment support for up to six EXP520 expansion enclosures.
- ▶ Attachment support for EXP810 with the *Attach EXP810 to DS5020 Activation feature*.
- ▶ Supports an intermix of SATA drives, FC drives, and encryption-capable FC drives (FDE) within the DS5020 and EXP520 enclosures.

Host attachment

- ▶ Provides SAN-attached 8 Gbps Fibre Channel (FC) host connectivity, as well as optional 1GbE iSCSI host connectivity.

All DS5020s have four 8 Gbps FC ports (two per controller).

Additionally you may order initially either:

- 2-Dual 8 Gbps Host Interface Cards (HIC)
- 2-Dual 1 Gbps iSCSI HIC

System cache

- ▶ DS5020 comes with 2 GB cache memory (1 GB per internal RAID controller).
- ▶ The 4 GB cache memory (2 GB per RAID controller) feature is available as an initial plant order feature.
- ▶ There are no cache memory upgrades available as field (MES) features for the DS5020.

Drive options

- ▶ FC disks without encryption:
 - 146.8 GB/15K 4 Gbps FC DDM
 - 300 GB/15K 4 Gbps FC DDM
 - 450 GB/15K 4 Gbps FC DDM
 - 600 GB/15K 4 Gbps FC DDM
- ▶ FC disk *with encryption*:
 - 146.8 GB/15K 4 Gbps FC encryption-capable DDM
 - 300 GB/15K 4 Gbps FC encryption-capable DDM
 - 450 GB/15K 4 Gbps FC DDM encryption-capable DDM
 - 600 GB/15k 4 Gbps FC DDM encryption-capable DDM
- ▶ SATA disks:
 - 750 GB/7.2K SATA DDM
 - 1000 GB/7.2K SATA DDM
- ▶ The DS5020 supports RAID 0, 1, 3, 5, 6, and 10.

3.5.1 DS5020 controller architecture

The DS5020 storage subsystem uses almost the same controller architecture as the DS4700 does (see Figure 3-41) with a few minor differences.

- ▶ Uses a faster XOR engine (1.2 GHz vs. 667 MHz on DS4700).
- ▶ One optional host interface (either dual port iSCSI or dual port 8 Gbps FC).
- ▶ The internal bus is PCI-E x8 (PCI-X on DS4700).
- ▶ Flash drives are used to destage dirty data from the cache in case of a power loss.

The heart of the controller is the 1.2 GHz XOR raid processor, which has improved speed compared to the DS4700 667 Mhz Xscale processor. The standard 8 Gbps host port and the optional Host Interface Cards in Figure 3-42 on page 69 and Figure 3-43 on page 69 share the same PCI-E x8 bus to transfer their data to the XOR chip. The main FC switch on the controller mainboard processes the FC traffic for the drive channels as well as for the two standard host channels.

The backup battery unit provides power to back up the cache memory of each controller onto flash drives in the event of a power failure.

To distribute the data to the drives, a chip called “switch-on-chip” (SOC) is used to switch the data directly to their internal destination drive or the external drive channel.

Both controllers connect indirectly via the drive channel.

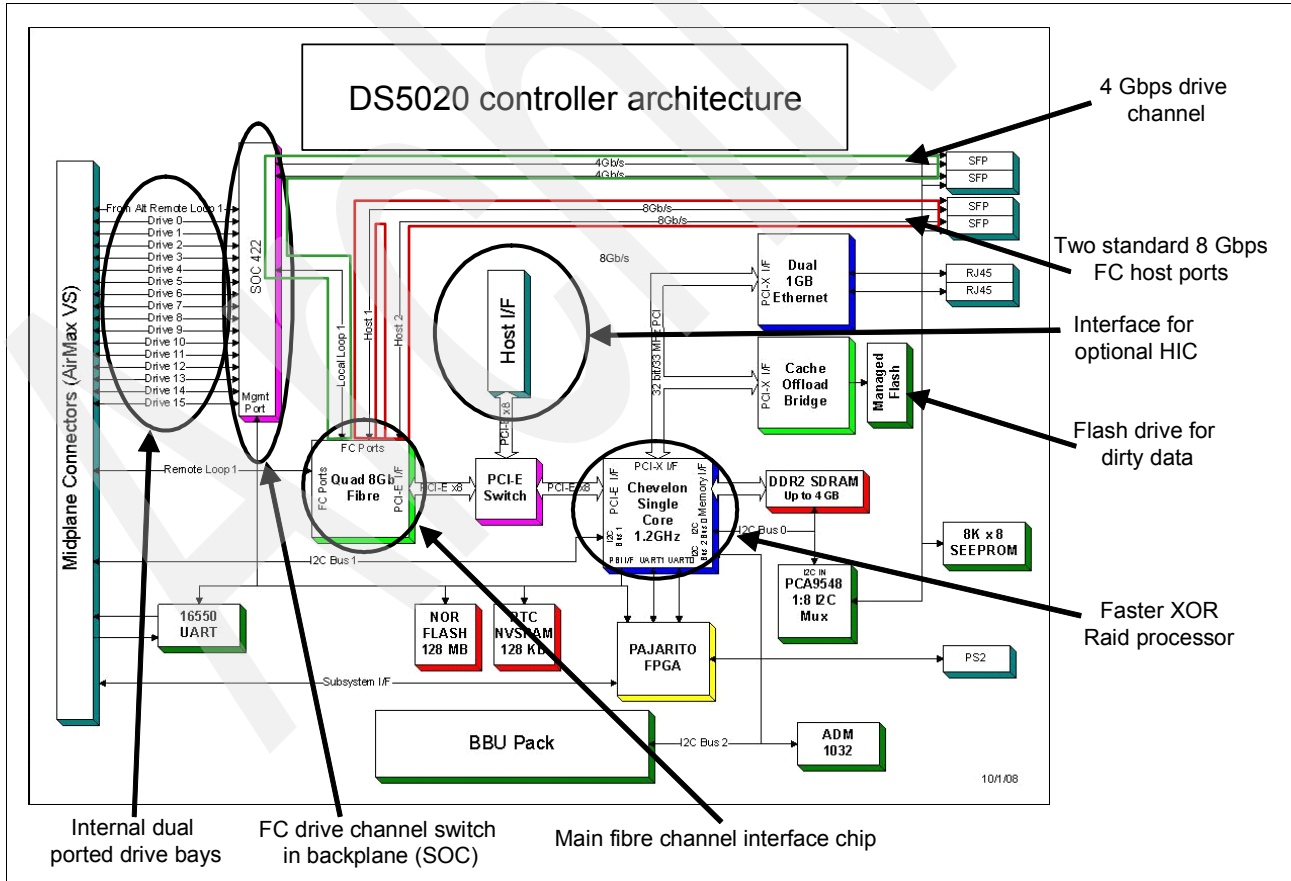


Figure 3-41 DS5020 controller architecture breakout

The optional dual-ported 8 Gbps Host Interface Card (HIC), shown in Figure 3-42, uses an two port FC chip. This chip connects via PCI-E x8 to the XOR engine.

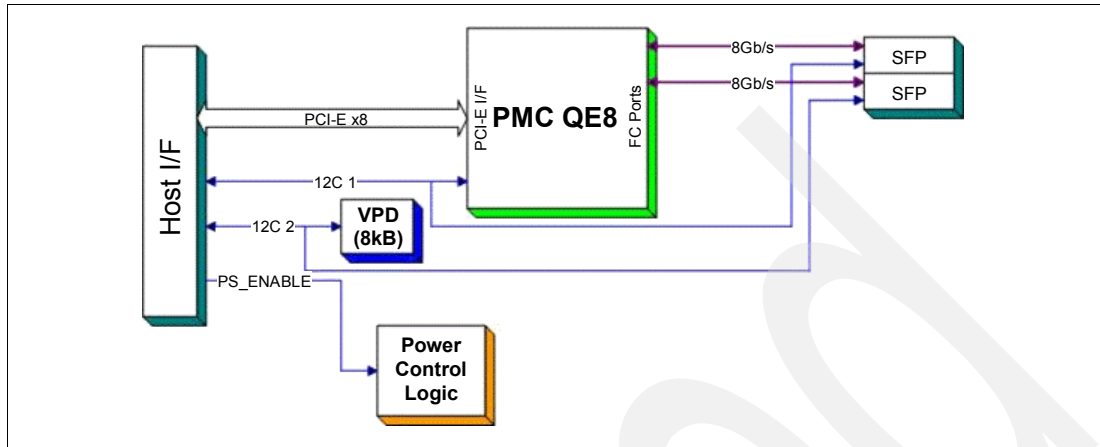


Figure 3-42 DS5020 Dual-Port 8 Gbps HIC architecture

If you have the iSCSI option, as shown in Figure 3-43, there is an QLogic top-off-engine (TOE) built into the DS5020. It will be built in as a daughter card of the controller mainboard. It reduces the workload of the controller CPU by calculating the whole Ethernet traffic and transfers just the SCSI packet through it.

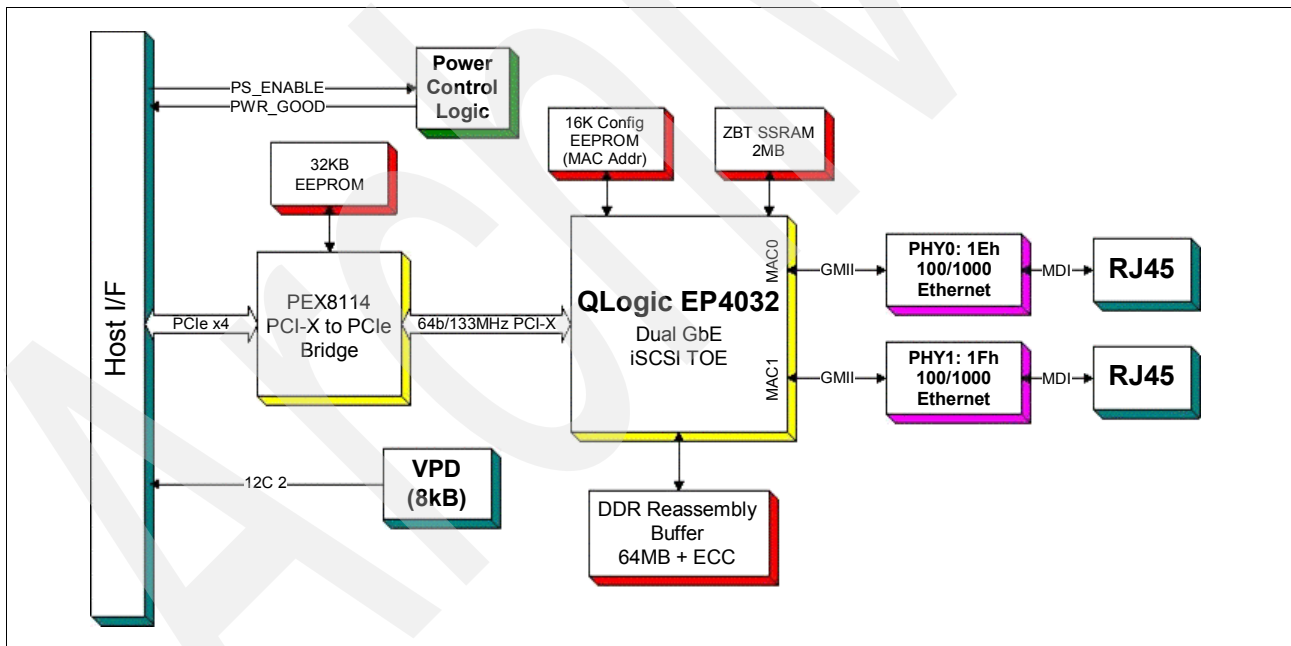


Figure 3-43 DS5020 iSCSI HIC architecture

Because of the similarities between the DS5000 controller models, we refer to 3.4.1, “DS5100 and DS5300 controller architecture” on page 31 for details about write cache, cache handling, write operations, and cache block flow.

3.5.2 DS5020 components

Figure 3-44 shows the subcomponents of the DS5020 storage subsystem.

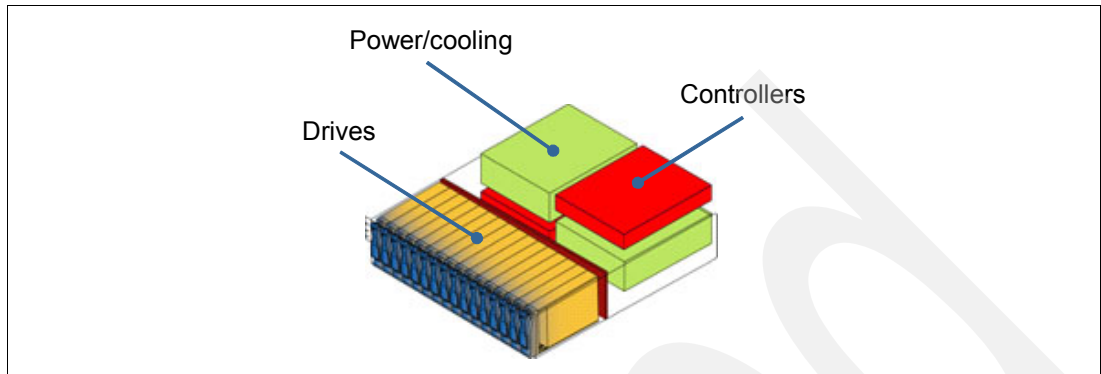


Figure 3-44 DS5020 hardware components

AC power supply and fan unit

Figure 3-45 shows the LEDs on the AC power supply and fan unit.

The LEDs display the status of the storage subsystem and components. The color of the LED is important:

- ▶ Green LEDs indicate a normal operating status.
- ▶ Amber LEDs (Needs Attention) indicate a possible failure.
- ▶ A blue LEDs on a CRU indicates that it is safe to remove the component.

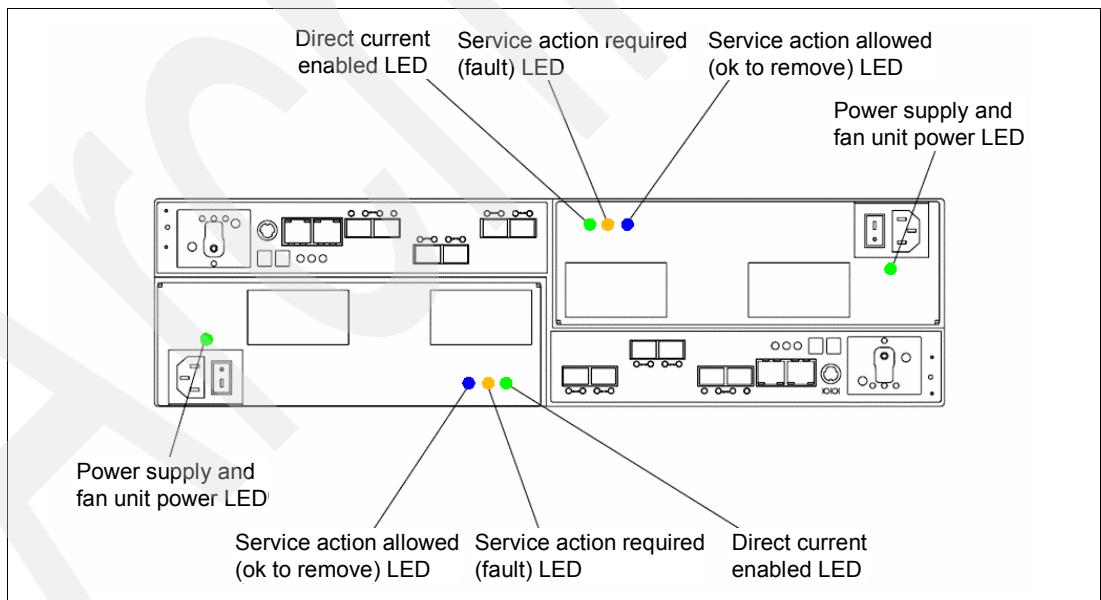


Figure 3-45 DS5020 power supply and fan unit LEDs

In normal operation, only the green LEDs (power LED and DC enabled LED) are on.

Controller

The controllers contain the storage subsystem control logic, interface ports, and LEDs. Depending on the DS5020 configuration you purchased, your controllers are one of the following types:

- ▶ Controllers with 1 GB memory and two standard 8 Gbps FC host ports
- ▶ Controllers with 1 GB memory, two standard 8 Gbps FC host ports, and one optional 2-port 8 Gbps FC host card
- ▶ Controllers with 1 GB memory, two standard 8 Gbps FC host ports, and one optional 2-port 1 GB iSCSI host card
- ▶ Controllers with 2 GB memory and two standard 8 Gbps FC host ports
- ▶ Controllers with 2 GB memory, two standard 8 Gbps FC host ports, and one optional 2-port 8 Gbps FC host card
- ▶ Controllers with 2 GB memory, two standard 8 Gbps FC host ports, and one optional 2-port 1 GB iSCSI host card

The controllers vary only in the size of the cache (either 1 or 2 GB) and the type of the optional Host Interface Card (either none, FC, or iSCSI).

Figure 3-46, Figure 3-47 on page 72, and Figure 3-48 on page 72 show the different DS5020 controller host interface configurations that are available. Figure 3-46 shows the base DS5020 storage subsystem, with two Fibre Channel host ports only. A field upgrade is not possible.

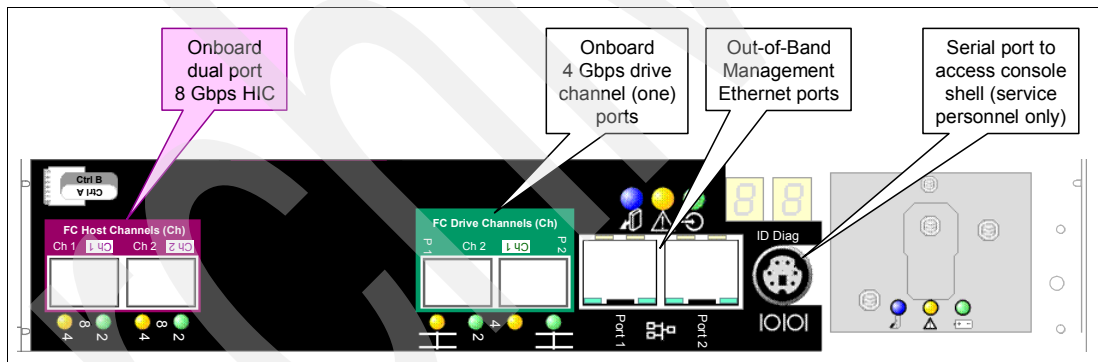


Figure 3-46 DS5020 controller with standard dual 8 Gbps FC host ports: Base model

Figure 3-47 shows the controller with the additional Dual-Port 8 Gbps FC Host Interface Card (HIC). The daughter card will be factory installed only. No field change can be done to this controller.

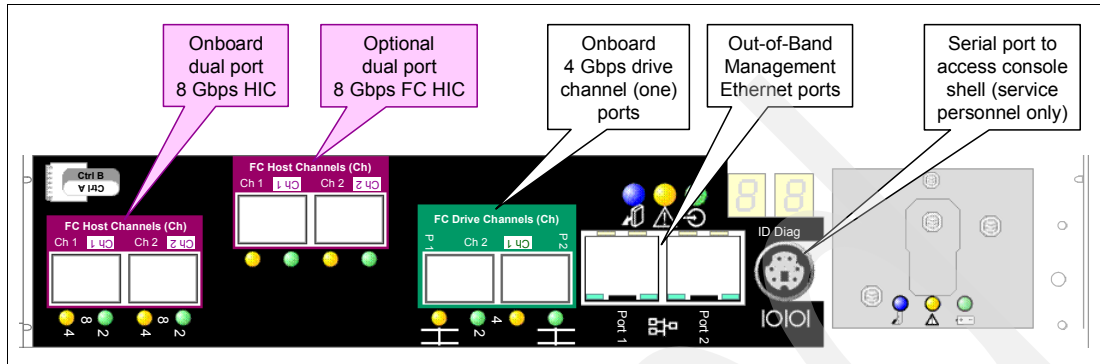


Figure 3-47 DS5020 controller: Additional 8 Gbps daughter card upgrade

The third configuration, shown in Figure 3-48, include a Dual-Port 1 Gbps iSCSI HIC card that will be factory installed as well. There is no field upgrade/change for the HIC cards.

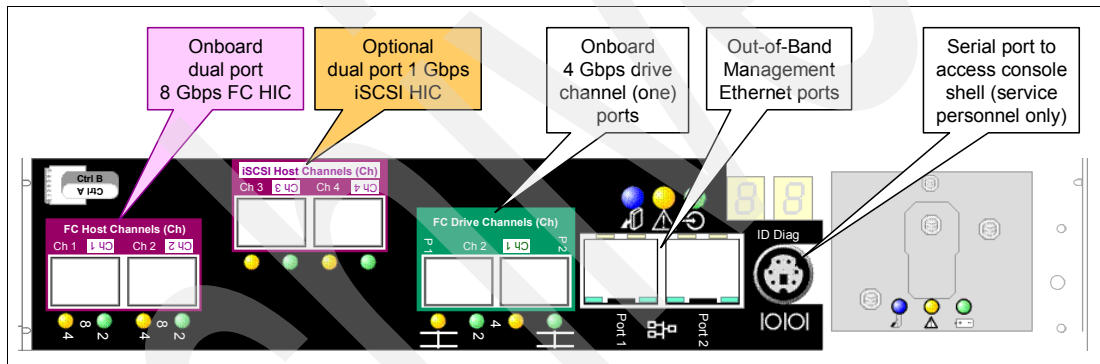


Figure 3-48 DS5020 controller with iSCSI host ports

The DS5020 comes with all SFPs pre-installed.

Enhanced Disk Drive Modules (E-DDMs)

The hot-swap drive bays that are accessible from the front of your storage subsystem are shown in Figure 3-49.

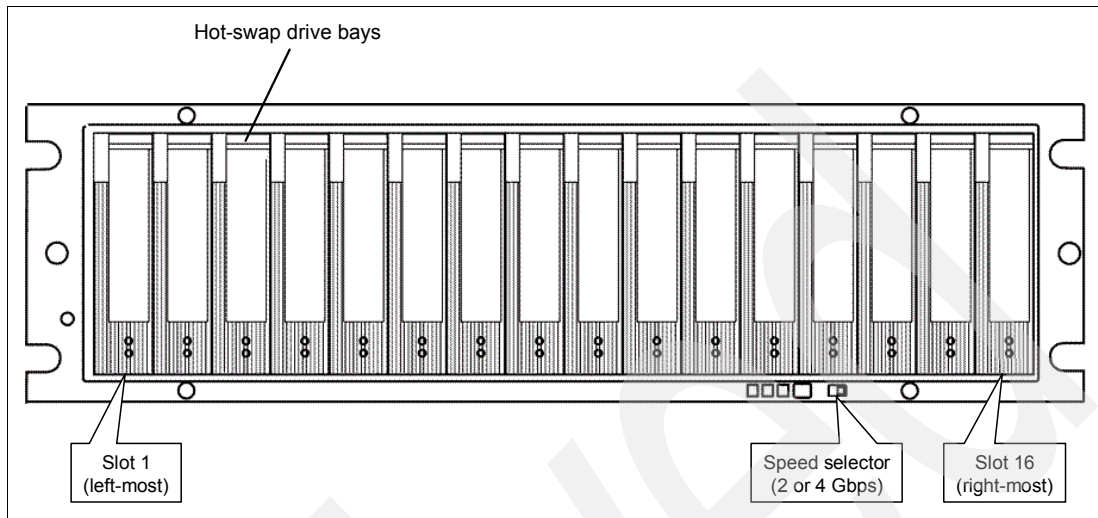


Figure 3-49 DS5020 hot-swap drive bays

The DS5020 supports both Fibre Channel (FC) and SATA E-DDMs intermixed in the storage subsystem drive chassis. The DS5020 supports up to sixteen 4 Gbps FC or 3 Gbps SATA E-DDMs.

SATA E-DDMs have an ATA translator card (interposer card) that converts the Fibre Channel protocol interface of the DS5020 drive channel or loop into the hard drive SATA protocol interface. It also provides dual paths to the SATA drive for drive CRU path redundancy. Each drive, ATA translator card, and carrier assembly is called SATA E-DDM CRUs. The Fibre Channel E-DDMs consist of the Fibre Channel and the carrier assembly (drive tray).

Install E-DDM CRUs in the 16 drive bays on the front of the storage subsystem from the leftmost slot (slot 1) to the rightmost slot (slot 16). When an E-DDM is installed, the drive and tray slot designation is set automatically. The hardware addresses are based on the enclosure ID, which is set by the controller software, and on the E-DDM physical location in the storage subsystem.

The DS5020 storage subsystem drive channel operates at a 4 Gbps Fibre Channel interface speed. Even the 3 Gbps SATA E-DDMs operate at 4 Gbps Fibre Channel speed.

Note: Even though the DS5020 has a 2 or 4 Gbps Fibre Channel Link Rate switch that can be used to set the drive channel speed at 2 Gbps, the link rate speed must be set to 4 Gbps. The DS5020 supports only 4 Gbps FC speed in the drive channel.

The Link Rate switch on the DS5020 storage subsystem and all storage expansion enclosures connected to it must have the same setting.

There are no serviceable parts in an E-DDM CRU. If it fails, it must be replaced in its entirety (E-DDM, ATA translator card, bezel, and tray). The DS5020 drive tray is not interchangeable with the drive tray of other DS4000 storage subsystems, such as DS4100 or DS4300 storage subsystems. The DS5020 E-DDM option CRUs are not interchangeable with those of the DS4200 Express and EXP420. When replacing an E-DDM CRU, be sure to order and install the correct E-DDM CRU. Using non-supported E-DDM options or FRUs will result in the E-DDM being locked out by the DS5020 controller firmware and might also damage the drive connector in the enclosure midplane.

The following precautions must be taken while replacing the E-DDM CRU.

- ▶ After you remove an E-DDM CRU, wait 70 seconds before replacing or reseating the E-DDM CRU to allow it to properly spin down. Failure to do so might cause undesired events.
- ▶ Never hot-swap an E-DDM CRU when its associated green Activity LED is flashing. Hot-swap an E-DDM CRU only when its associated amber Fault LED lights is not flashing or when the E-DDM is inactive and its associated green Activity LED lights is not flashing.
- ▶ If the E-DDM you want to remove is not in a failed or bypass state, always use the Storage Manager client program either to place it in a failed state or to place the array that is associated with the E-DDM (or E-DDMs) in an offline state before you remove it from the enclosure.

The IBM System Storage DS5020 storage subsystem (Machine Type 1814-20A) supports RAID levels 0, 1, 3, 5, and 6 up to over 67.2 TB when using 600 GB Fibre Channel hard drives and up to 112 TB when using 1 TB Serial Advanced Technology Attachment (SATA) Enhanced Disk Drive Modules (E-DDMs).

The DS5020 supports configurations of FC disks with or without Full Disk Encryption (FDE), or SATA disks, or a mix of disk drives. To install FDE disks in a DS5020, you must purchase the Full Disk Encryption option.

3.5.3 DS5020 Storage Subsystem front view

Figure 3-50 shows the DS5020 from the front side.

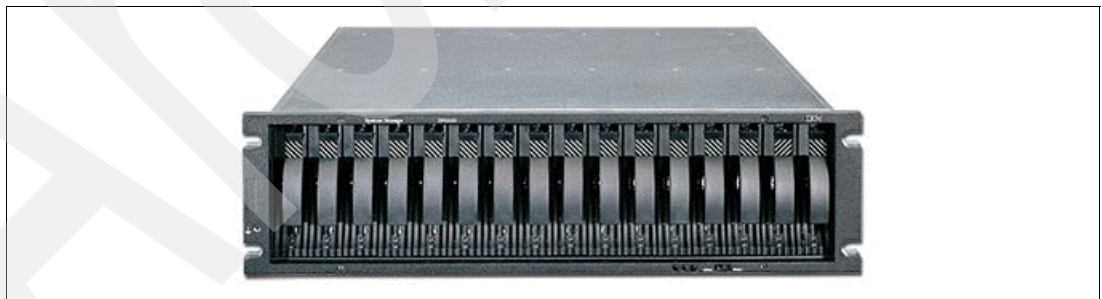


Figure 3-50 DS5020 storage subsystem front view

The hot-swap features of the DS5020 enable you to remove and replace the 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.

The DS5020 Express supports up to 16 disk drive modules (DDMs). Each drive bay also provides dual paths to each drive for path redundancy. The drives are customer replacement units (CRUs).

Several LED indicators and the FC Link speed selector are also visible from the front of the storage unit. Refer to “Front panel LEDs and FC link speed selector” on page 79 for details about the LEDs.

Attention: Never hot-swap an E-DDM CRU when its associated green activity LED is flashing. Hot-swap a drive CRU only when its associated amber fault LED light is not flashing or when the drive is inactive and its associated green activity LED light is not flashing. Wait 70 seconds before inserting the drive back into the bay.

3.5.4 DS5020 storage subsystem rear view

The rear of the DS5020 appears in different versions depending on what host port configuration has been ordered for the system (Figure 3-51, Figure 3-52 on page 76, and Figure 3-53 on page 76).

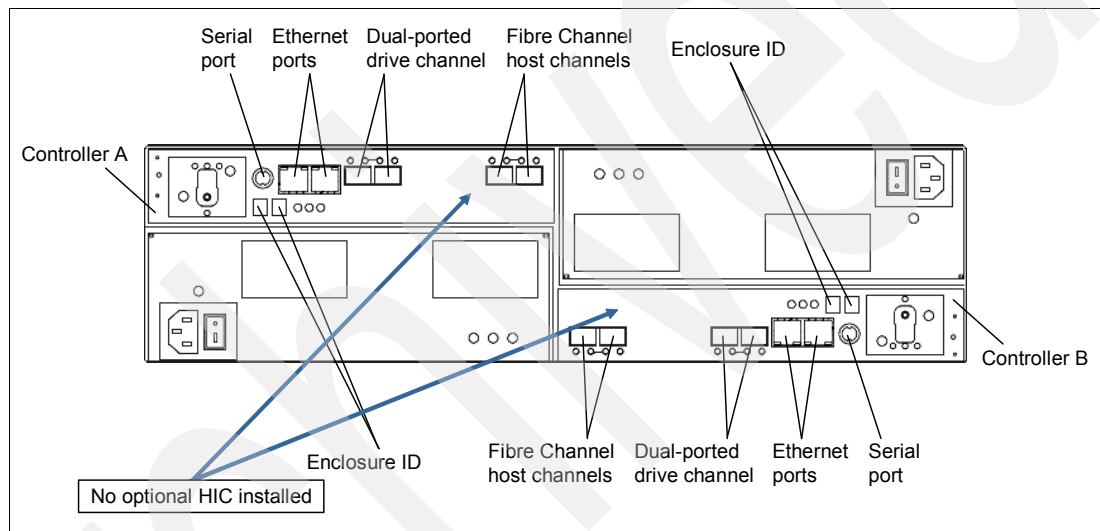


Figure 3-51 DS5020 rear view: Base model

They vary only in host port configurations. The base model does not have the optional Host Interface Cards (HIC). It comes with four 8 Gbps Fibre Channel host ports.

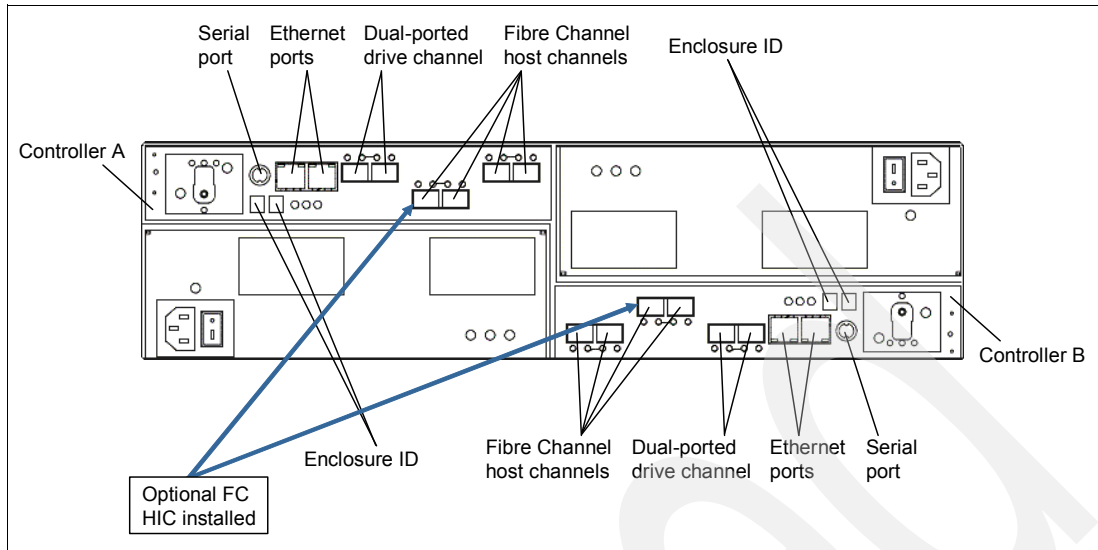


Figure 3-52 DS5020 rear view: 8 FC Port model

Another configuration of the DS5020 has additional four 8 Gbps FC host ports (Figure 3-52) or additional four 1 Gbps iSCSI host ports instead (Figure 3-53).

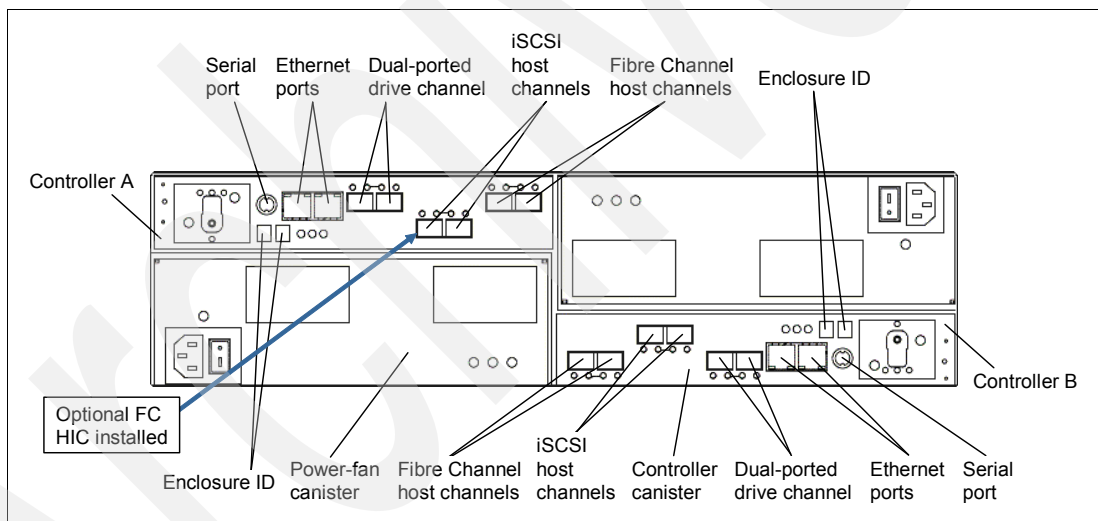


Figure 3-53 DS5020 rear view: 4 FC and 4 iSCSI port model

The DS5020 rear view shows four hot swappable parts:

- ▶ The two controllers with the Backup Battery Unit (BBU)
- ▶ The two Power Supply and Fan Units

The two controllers hold host and drive interfaces as well as the batteries. The left controller is controller A and the right controller is controller B. Note that controller A is upside-down relative to controller B. The same configuration applies to the power supply and fan unit. It is important to keep this information in mind when connecting the back-end ports to hosts and drive-side expansion enclosures. Refer to 3.5.7, “DS5020 storage subsystem drive-side connections” on page 85 for more details.

Figure 3-54 shows a closer view of a DS5020 controller.

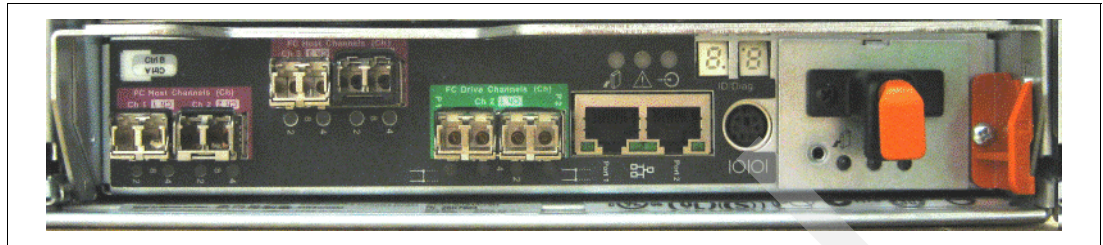


Figure 3-54 DS5020 controller photo rear view

SFP modules

The storage subsystem supports a fiber-optic interface for host and storage expansion enclosure connections. You must install a Small Form-factor Pluggable (SFP) module in each interface connector on the controller where a fiber-optic cable is to be installed.

Note: Do not install an SFP in any port that will not have a fiber-optic cable attached.

Remove any SFP from any port that does not have fiber-optic cables attached.

The DS5020 storage subsystem host ports support 2, 4, and 8 Gbps Fibre Channel speeds. The DS5020 storage subsystem drive ports support only 4 Gbps Fibre Channel speeds.

The maximum operating speed of the Fibre Channel port is determined by two factors:

- ▶ The speed of the SFP module that is installed
- ▶ The speed of the Fibre Channel connection

For example, a 4 Gbps SFP that is plugged into a 8 Gbps-capable port will limit the speed of that port to a maximum of 4 Gbps. Conversely, an 8 Gbps SFP that is plugged into a 4 Gbps-capable port will limit the speed of the port to a maximum of 4 Gbps. Carefully check the SFP IBM part number, option number, and FRU part number to identify its speed. There are no physical features that distinguish an 8 Gbps SFP from a 4 Gbps SFP.

Figure 3-55 shows an example of SFP module with fiber-optic cable.

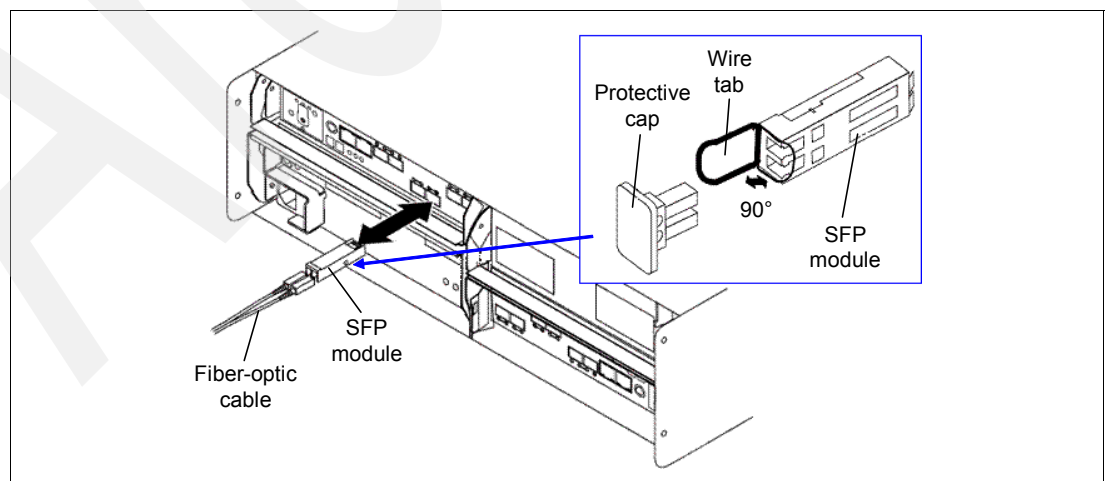


Figure 3-55 DS5020 SFP Module

SFP for all FC ports are shipped with the system.

Backup Battery Unit (BBU)

The Backup Battery Unit provides power to back up the cache memory of each controller onto flash drives in the event of a power failure. Each battery unit contains a sealed, rechargeable SMART lithium ion battery. The battery unit contains enough charge to back up the cached data in each controller to a flash drive in the event of a power failure.

When the unit is powered on the first time or whenever the battery is replaced, the battery chargers will charge the battery to the programmed level. Then, the controller will start a battery learning cycle to determine whether the battery current capacity is sufficient.

Note: Data caching starts after the battery is charged to the programmed level.

During the battery learn cycle, the cache will be active if the battery is in good condition. If the battery fails the learn cycle, it is marked as failed. The battery learning cycle lasts up to three hours. After the first battery learn cycle, the controller will perform a learn cycle every 8 weeks to re-calibrate the battery-charging level. The battery unit is hot-swappable. You can remove the battery unit for servicing and then reinsert it while the DS5020 continues to perform I/O operations. The battery should be removed using the Storage Manager: Prepare for Removal procedure (see Chapter 7, “Advanced maintenance, troubleshooting, and diagnostics” on page 327). However, write I/O caching is disabled when the battery is in a failed state or removed from the controller chassis. Replace the failed battery as soon as possible to minimize the time that the write IO caching is disabled. Information about the condition of the battery unit is conveyed by indicator LEDs on the front of battery unit.

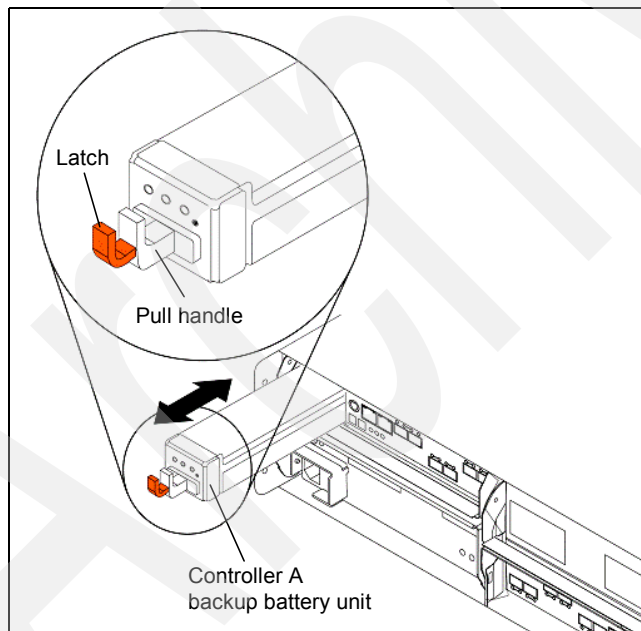


Figure 3-56 Replaceable Backup Battery Unit (BBU)

To physically remove the BBU, just push the latch, as shown in Figure 3-56, and pull out the battery unit.

Important: Unlike the batteries for DS4000 storage subsystems, the DS5020 storage subsystem battery units do not use the expiration dates given in Storage Manager. Do not replace these batteries after a certain usage period.

3.5.5 DS5020 storage subsystem LED indicator lights

LED indicator lights allow the DS5020 to communicate with the user. There are four main components with LEDs:

- ▶ Front panel
- ▶ RAID controllers
- ▶ Battery
- ▶ Power supply fans

Front panel LEDs and FC link speed selector

Figure 3-57 shows the DS5020 front panel and its LED indicator lights.

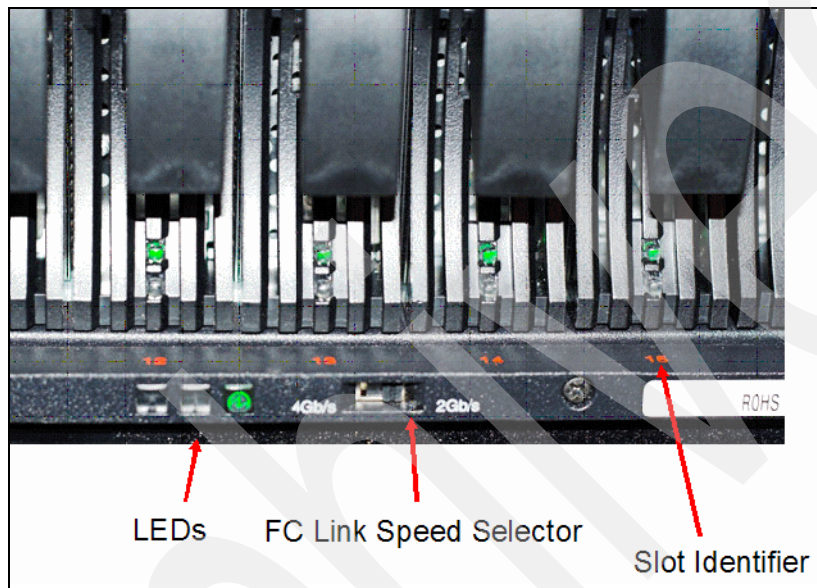


Figure 3-57 DS5020 front panel LEDs

Link speed selector

The FC link speed selector is a physical switch that must be used to set the enclosure speed. The DS5020 Express storage system drive channel operates at only at 4 Gbps Fibre Channel interface speed.

Note: The DS5020 storage system SATA E-DDM CRUs have an ATA translator card that converts E-DDM 3 Gbps SATA drive interface protocol to 4 Gbps Fibre Channel interface speed.

Front LEDs

These are:

- ▶ Locate LED (blue)
 - On: This indicates storage subsystem locate.
 - Off: This is the normal status.
- ▶ Service action required LED (amber)
 - On: There is a corresponding needs attention condition flagged by the controller firmware. Some of these conditions might not be hardware related.
 - Off: This is the normal status.

- ▶ Power LED (green)
 - On: The subsystem is powered on.
 - Off: The subsystem is powered off.

RAID controller LEDs

There are LEDs on the RAID controllers that serve as indicators of key information (refer to Figure 3-58, Table 3-5, and Table 3-6 for LED status details).

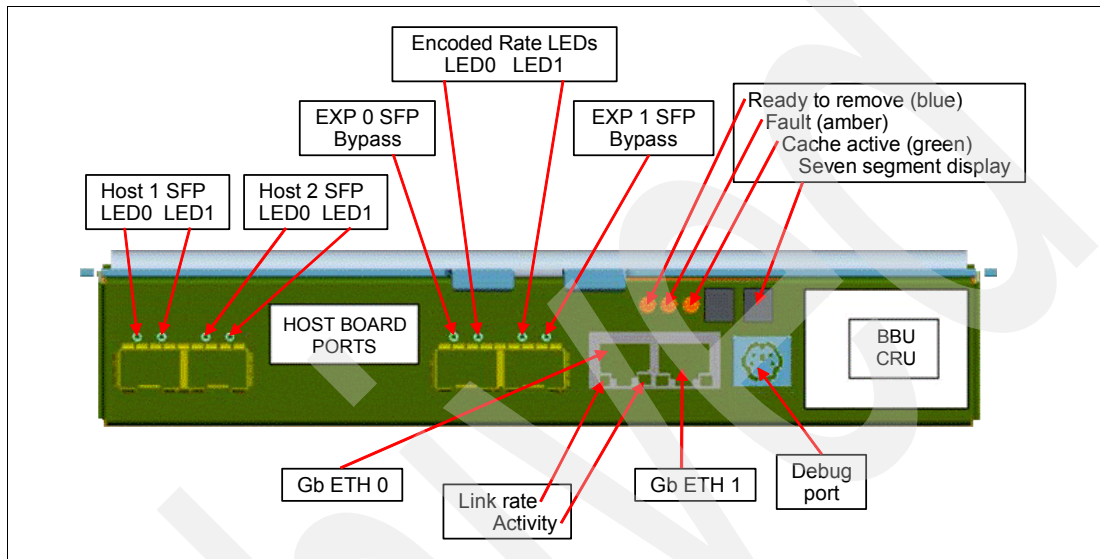


Figure 3-58 DS5020 RAID controller LEDs

Host SFP LEDs

- ▶ LED #0 (green): Host channel speed L1
- ▶ LED #1 (green): Host channel speed L2

Table 3-5 FC host SFP status LED definitions

LED#0	LED#1	Port status
OFF	OFF	Link down
ON	OFF	Link rate 2 Gbps
OFF	ON	Link rate 4 Gbps
ON	ON	Link rate 8 Gbps

Disk Channel SFPs LEDs

Table 3-6 FC disk expansion port SFP LED definitions

LED#0	LED#1	Port status
OFF	OFF	Link down
ON	OFF	Reserved
OFF	ON	Link rate 2 Gbps
ON	ON	Link rate 4 Gbps

- ▶ Drive channel bypass / EXP bypass (amber)
 - Off: Normal status
 - On: Drive port bypass problem

Other LEDs

- ▶ Serviced action allowed / Ready to remove (blue) (see “Enclosure ID” on page 82 for details)
 - Off: Normal status
 - On: Safe to remove
- ▶ Need attention (amber)
 - Off: Normal status
 - On: Controller needs attention (controller fault or controller is offline)
- ▶ Caching active (green)
 - On: Data in cache
 - Off: No data in cache
- ▶ Ethernet link rate
 - Off: 100 Mbps
 - On: 1 Gbps BASE-T
- ▶ Ethernet link activity
 - Off: No link established
 - On: Link established (blinks off with transmit or receive activity)

Enclosure ID

The enclosure ID, comprised of two seven-segment numbers, is located on the back of each controller next to the indicator lights, as shown in Figure 3-59. It provides a unique identifier for each enclosure in the DS5020 storage subsystem configuration.

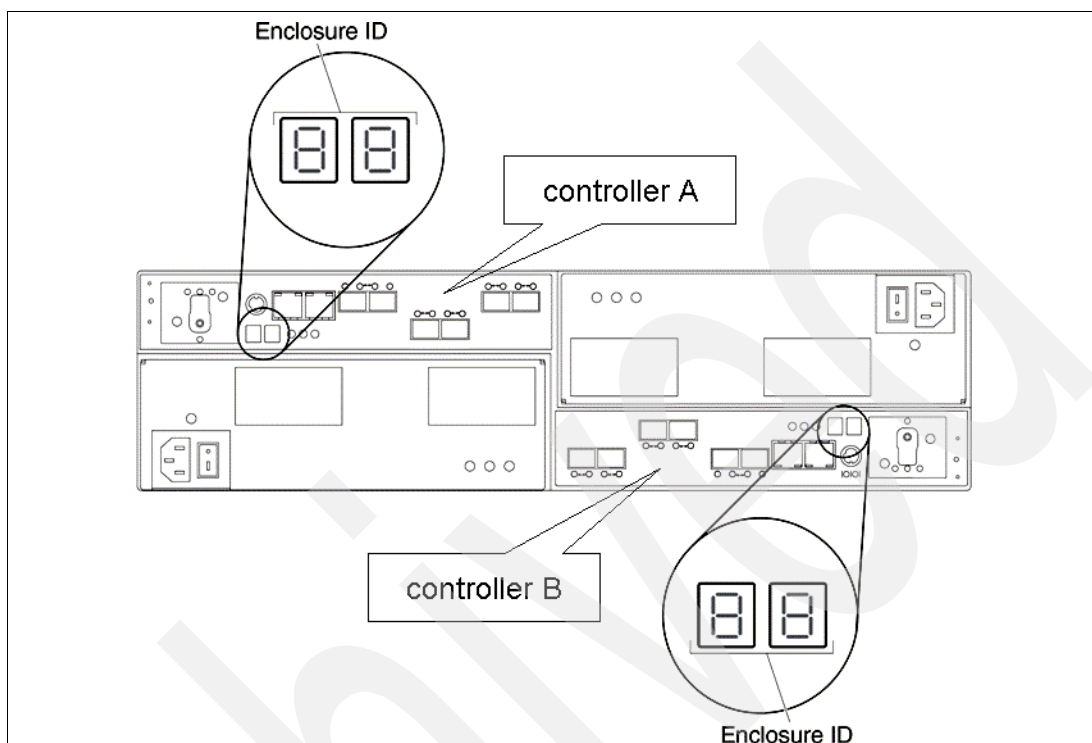


Figure 3-59 DS5020 controller enclosure ID display

The controller automatically sets the enclosure ID number. You can change the setting through the storage management software if necessary. Both controller enclosure ID numbers are identical under normal operating conditions. Each storage expansion enclosure (including the DS5020 storage subsystem itself) in the DS5020 configuration must have a unique storage enclosure ID.

In addition, the single digits (x1) of the enclosure IDs of all storage expansion enclosures and the DS5020 storage subsystem in the redundant drive channel/loop pair must be unique. Because the DS5020 has only one drive channel, all expansion enclosures will have the same single digit (x1).

Although the allowable ranges for enclosure ID settings are 0-99, do not set the enclosure ID to 00 or any number less than 80. The DS5020 enclosure ID is usually set to a value of 85 before it is shipped.

Service action allowed (SAA) LEDs

There are few things to talk about regarding the SAA LEDs (Figure 3-60 on page 83).

Each controller, power supply and fan unit, and battery unit has a blue Service Action Allowed status (SAA) LED. The purpose of the SAA LED is to help make sure that a component is not removed before it is safe to do so. Do not remove any storage subsystem component unless the Service Action Allowed status LED for that component is lit.

Attention: If you do a controller replacement, make sure that the controller that you want to replace is in offline mode or failed before you physically remove it. Using only the Prepare for Removal function will not change the controller state.

Use the Prepare for Removal function in the DS Storage Manager Subsystem Management window or refer to the applicable component replacement instructions for this case. Refer to Chapter 7, “Advanced maintenance, troubleshooting, and diagnostics” on page 327 for detailed information.

Note: Wait at least two minutes after you replace each component for the controller to recognize the new component and update the LED status. If the SAA LED is turned on, you must remove and replace the component in order to get the LED turned off.

In most cases when a single component fails, the Service Action Allowed status LED turns on steadily when the Needs Attention status LED is turned on for the component.

Battery LEDs

Each DS5020 RAID controller has its own battery. There are three LED indicator lights on each battery:

- ▶ Service action allowed (blue) (see “Enclosure ID” on page 82 for details)
 - Off: Normal status.
 - On: Safe to remove.
- ▶ Battery charging (green)
 - On: Battery charged and ready.
 - Blinking: Battery is charging.
 - Off: Battery is faulted, discharged, or missing.
- ▶ Needs attention or service action required (amber)
 - Off: Normal status.
 - On: Controller firmware or hardware requires attention.

Figure 3-60 shows the battery LEDs.

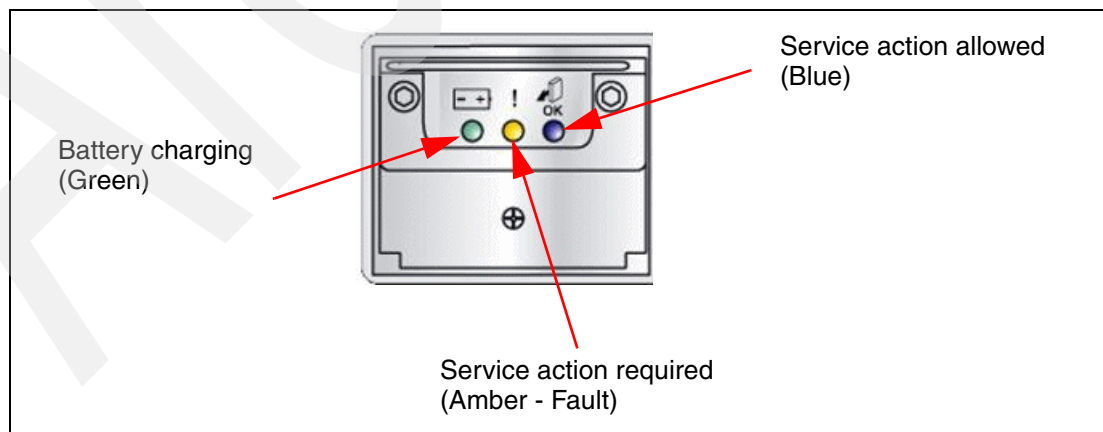


Figure 3-60 Battery LEDs

Power supply and fan unit LEDs

Each power supply fan (Figure 3-61) contains one power supply and two fans.

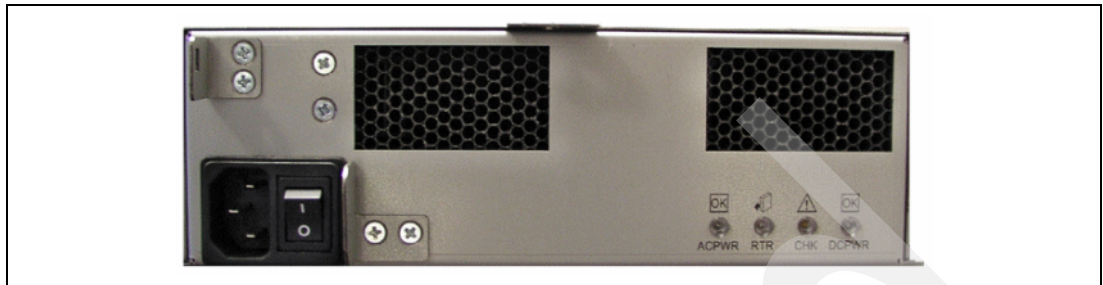


Figure 3-61 Power supply and fan unit LEDs

The LEDs are:

- ▶ Power supply fan LED (AC power) (green)
 - Off: Power supply fan is not providing AC power.
 - On: Power supply fan is providing AC power.
- ▶ Serviced action allowed (blue) (see “Enclosure ID” on page 82 for details)
 - On: Safe to remove.
 - Off: Normal status.
- ▶ Needs attention (amber)
 - Off: Normal status.
 - On: Power supply fan requires attention.
- ▶ Power supply fan Direct Current Enabled (DC power) (green)
 - Off: Power supply fan is not providing DC power.
 - On: Power supply fan is providing DC power.

3.5.6 DS5020 storage subsystem host-side connections

The DS5020 integrates the host-side and drive-side connections into the controller itself. The DS5020 has two 8 Gbps host connections by default. They are mounted on the mainboard of the controller. Another two host ports per controller can be added when ordering the DS5020. They can be installed only by factory. No MES field upgrade will be possible. Currently, you can order 1 Gbps iSCSI or 8 Gbps FC Host ports. Refer to “Controller” on page 71 for the configuration options. The FC Host connections support Fibre Channel attachment through SAN switches or direct connections. The iSCSI Host connections support 100 Mbps or 1 Gbps switched Ethernet or iSCSI network and direct connection.

- ▶ 8 Gbps FC host ports

These ports autonegotiate with 2 Gbps, 4 Gbps, and 8 Gbps Fibre Channel speed if an 8 Gbps SFP is installed. 1 Gbps FC speed will not be supported.

- ▶ 4 Gbps FC drive ports

The controller has two drive ports that belong to the same drive channel. Both ports must run at 4 Gbps speed because only EXPs and drives running a 4 Gbps are supported to be attached to the DS5020.

► 1 Gbps iSCSI host Ports

The iSCSI ports support both IPv4 and IPv6 TCP/IP addresses, CHAP, and iSNS. Use either Cat5E or Cat6 Ethernet cable types for iSCSI port connections. A Cat6 Ethernet cable provides optimal performance. The setup of the host ports will be done in the Storage Manager. By default, the iSCSI ports autonegotiate between 100 and 1000 Mbps Ethernet speed.

It is important to match up host or fabric connections to the DS5020 by attaching one connection to each controller. In doing so, you take advantage of the DS5020's ability to fail over and distribute the workload among the two controllers. For any given host, make sure to connect to the same host port number on each controller. The host port layout is shown in Figure 3-68 on page 91.

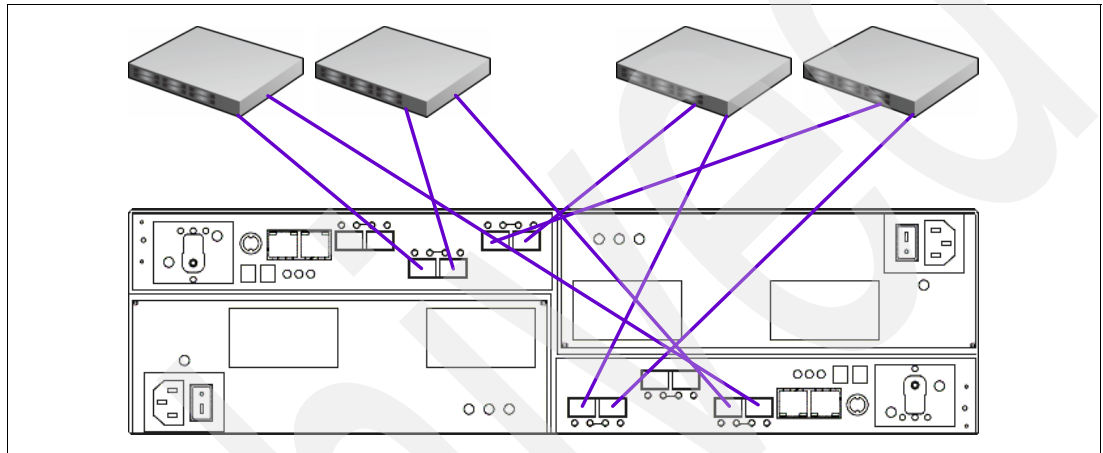


Figure 3-62 DS5020 mixed host connection layout

3.5.7 DS5020 storage subsystem drive-side connections

The DS5020 can attach up to six expansion enclosures. The regular expansion enclosure is the EXP520. However, you can buy a feature that enables you to attach EXP810 as well (refer to “Disk drives” on page 67). Only these two expansion enclosures are currently supported. It is generally best to spread the enclosures evenly between the two drive channel pairs as you scale up the DS5020 in storage capacity. A fully configured DS5020 should have three expansion enclosures on each drive-side channel pair.

Both drive ports on the DS5020 controller belong to the same drive channel, which means that both drive ports must operate at the same Fibre Channel speed. However, there are only 4 Gbps expansion units supported.

Note: There are three rules for the expansion cabling:

- ▶ With the DS5020, you should only connect a maximum of three enclosures per controller drive port.
- ▶ The DS5020 controller drive port must always be connected to the EXP520 or EXP810 port labelled 1B. Because the left (ESM A) and right (ESM B) enclosure service modules (ESM) are inserted in different orientations, ensure that you use the port labeled 1B before making the Fibre Channel connection to the DS5020 storage subsystem, as shown in Figure 3-63.
- ▶ Spread expansion enclosures among the two drive channel pairs. For example, if you attach four enclosures, it is better to have two enclosures behind each drive port rather than three and one.

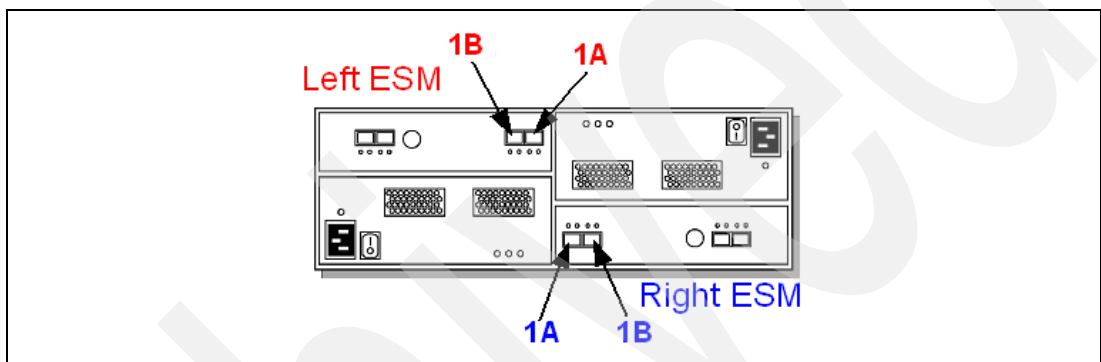


Figure 3-63 Port labels on EXP520 or EXP810

Refer to “Disk Channel SFPs LEDs” on page 80 for more details about the LED status on the disk channels.

3.5.8 DS5020 storage subsystem drive-side cabling

The drive-side cabling for the DS5020 depends on how many expansion units you must attach:

- ▶ If you attach only one enclosure, make sure that you have one connection to each of the controllers, thus using one of the two ports on each controller (controller A port 2 and controller B port 1), as shown in Figure 3-64.

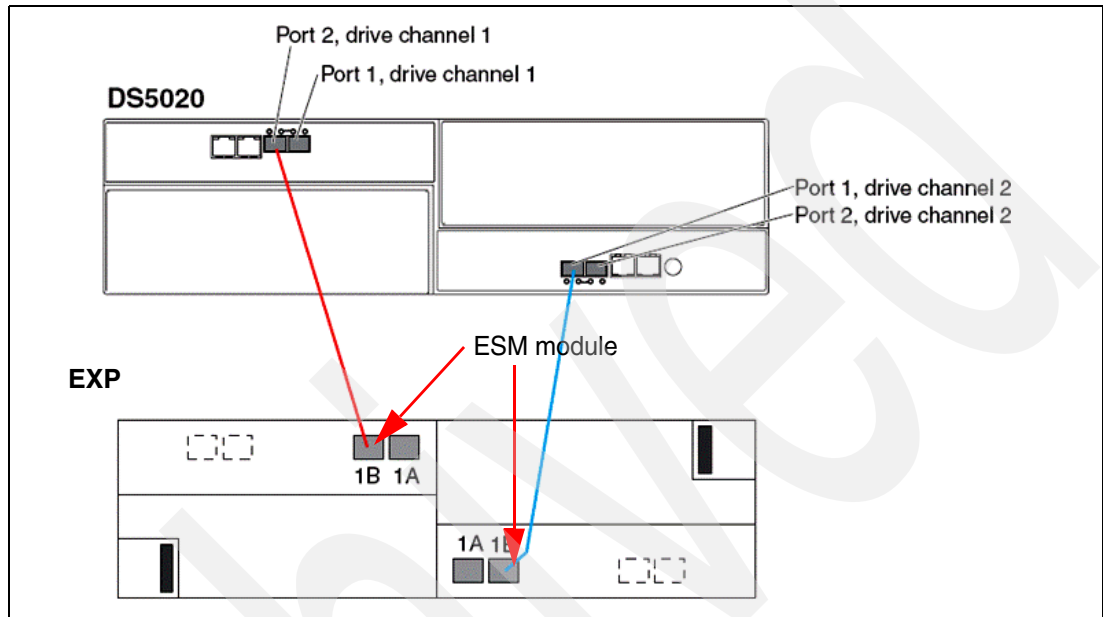


Figure 3-64 DS5020 drive cabling with one EXP

- ▶ If you attach a second expansion unit, connect it by using the second port on the controller (controller A port 1 and controller B port 2), as shown in Figure 3-65.

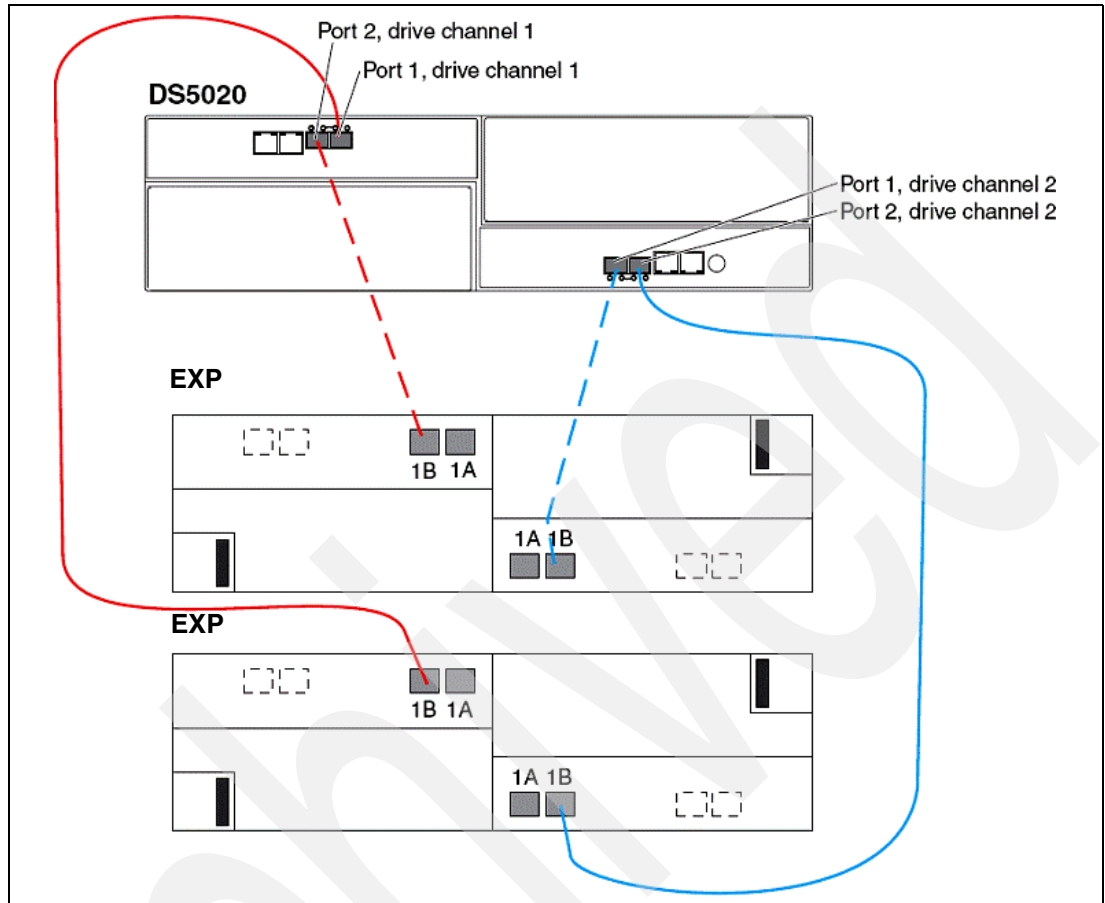


Figure 3-65 DS5020 drive cabling with two EXPs

- Beyond two enclosures (up to a maximum of six), make sure that you equally distribute the enclosures among the redundant drive channel pairs (Figure 3-66 and Figure 3-67 on page 90).

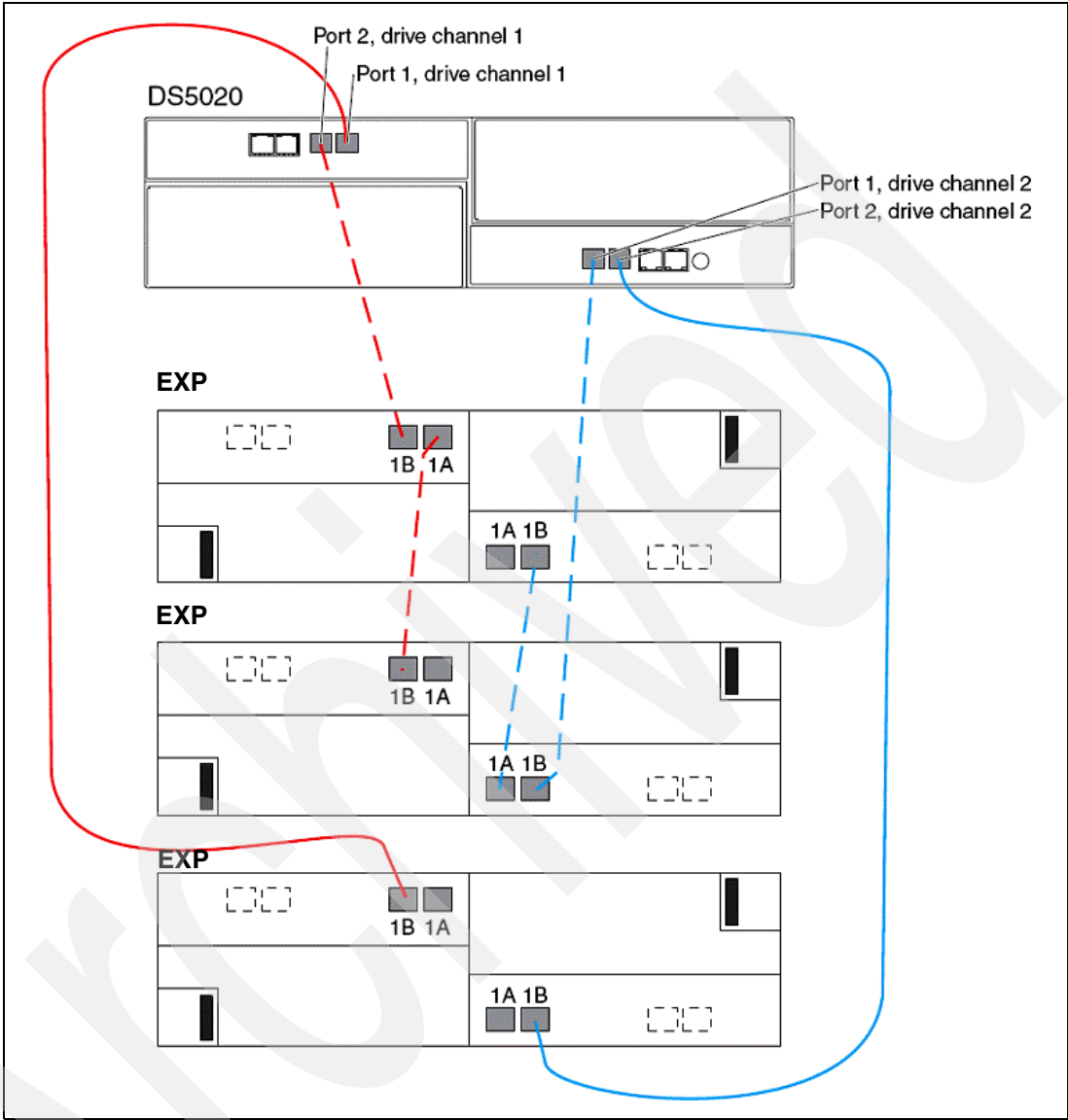


Figure 3-66 DS5020 drive cabling with three EXPs

When six enclosures are required, the same method is employed again, maintaining drive channel redundancy and using both controllers, as shown in Figure 3-67.

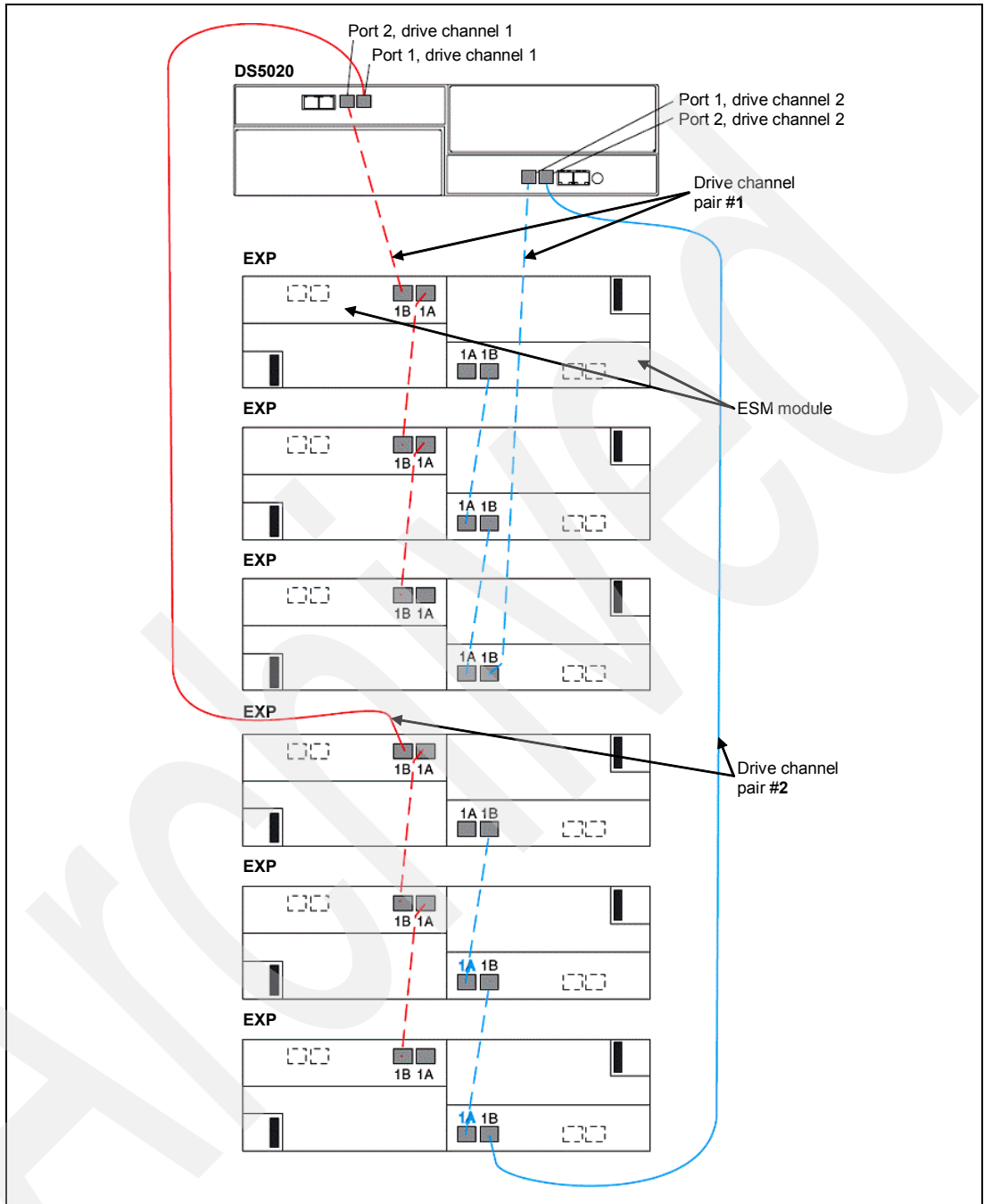


Figure 3-67 DS5020 drive-side cabling with six EXPs

Drive-side cabling example

As shown in Figure 3-67, the DS5020 is cabled using all two-drive channel pairs, assuming that there are six expansion enclosures (EXP) evenly spread out across the drive channel pairs (three enclosures each).

Each enclosure service module (ESM) only has one pair of ports, labelled 1A and 1B, that can be used to connect FC cables. The other pair of ports is reserved for future use. Proceed as follows:

1. Start with the first expansion enclosure, which we attach to drive channel pair #1. Cable controller A, drive port 2 to the port (1B) on the left ESM of the first EXP.
2. Cable the port (1A) of the left ESM on the first EXP to the port (1B) on the left ESM of the second EXP.
3. Cable the port (1A) on the left ESM of the second EXP to the port (1B) on the left ESM of the third EXP.
4. Cable the port (1B) on the right ESM of the first EXP to the port (1A) on the right ESM of the second EXP.
5. Cable the port (1B) on the right ESM of the second EXP to the port (1A) on the right ESM of the third EXP.
6. Cable controller B, drive port 1 to the port (1B) on the right ESM of the third EXP located on the first drive channel pair. This is the last step of the first drive channel pair.

Repeat steps 1–6 (using the next drive-side channel pair ports) for the second drive channel pairs (three EXPs each).

3.5.9 DS5020 storage subsystem additional connections

The DS5020 storage subsystem has various kinds of connectors, as shown in Figure 3-68.

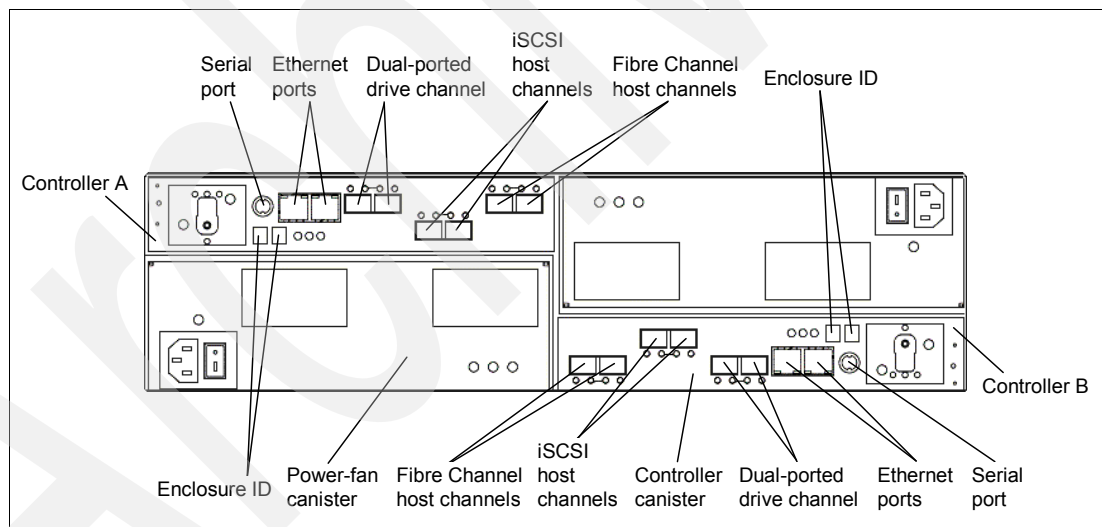


Figure 3-68 DS5020: All connectors

These connectors are:

► Management Ethernet connectors

This connector is for an RJ-45 10/100/1000 BASE-Tx Ethernet connection. There are two connections per controller. One port is designed for out-of-band management and the other port is meant for serviceability. The logic behind adding an extra port was to introduce additional isolation and to separate management and service traffic from one another. Because of the extra port, it is preferable to have two IP addresses per controller in order to manage and service the DS5020 storage subsystem appropriately. However, you cannot attach this port to a routed network, because you cannot set up a gateway for it. You will still operate the DS5020 with only one IP port active per controller. The best practice is to set port 1 in the customer network for out-of-band management and leave the Port 2 as the default in order to let service personnel to connect using the default IP addresses.

The default IP addresses for the controllers are shown in the Table 3-7. The default subnet mask for all four Ethernet ports is 255.255.255.0.

Table 3-7 Default IP addresses for management ports

	Controller A	Controller B
Port 1	192.168.128.101	192.168.128.102
Port 2	192.168.129.101	192.168.129.102

► Serial 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.

Note: We do not recommend the terminal program PuTTY, because certain versions of PuTTY send characters to the controller that cause the controller to reboot.

The maximum baud rate is 115,200 bps. The default baud rate setting from the factory is 38,400 bps, N-8-1, with no flow control.

Attention: Managing the DS5000 storage subsystem through the serial interface has potential risks. Using certain commands, you can initialize the RAID controller, and therefore lose all your data. You should only use this interface when instructed to do so by IBM Support.

3.6 DS5000 series product comparison

This section will compare the DS5000 products.

3.6.1 DS5020 product comparison

Take a quick look at the specifications of the DS4700 and DS5020 (Figure 3-69), You see significant improvements in many areas.

<i>Dual-controller system (unless noted)</i>	DS4700	DS5020
Host interfaces options	Four 4 Gb/s FC; or Eight 4 Gb/s FC	Four 8 Gb/s FC; or Eight 8 Gb/s FC; or Four 8 Gb/s FC & four 1 Gb/s iSCSI
Redundant drive channels	Four 4 Gb/s	Four 4 Gb/s
Max drives	112 FC/SATA	112 FC/FDE/SATA
RAID Processor	Intel xScale667MHz	Intel xScale1.2 GHz
Bus Technology	PCI-X (1GB/s)	PCI-Express x8 (4 GB/s)
Internal controller bandwidth	2 GB/s	8 GB/s
Cache memory (min/max)	2 GB / 4 GB	2 GB / 4 GB
Cache IOPS	120,000 IOPS	200,000 IOPS
Disk Read IOPS	40,000 IOPS	50,000 IOPS
Disk Read MB/s	980 MB/s	1,600 MB/s

Figure 3-69 DS5020 / DS4700 specification comparison

While the max host ports (8) and drives (112) are the same, the DS5020's 8 Gbps interfaces and support for 1 Gbps iSCSI give it the connectivity advantage. The DS5020's support for self-encrypting drives (also referred to as FDE for full disk encryption) provide an clear advantage.

The controller architecture is where we get the performance gains. It has a 80% faster processor and 300% faster busses. The two items are key contributors to the DS5020's performance advantages, which are up to 67% higher.

A vendor neutral company called Storage Performance Council has done a performance test on the DS5020. The results of this test can be found at the following address:

<http://www.storageperformance.org/results>

3.6.2 DS5300 product comparison

Figure 3-70 compares the DS5100 and DS5300. There is no difference between these models except for the internal front side bus speed. In earlier versions of code, there were restrictions on cache upgrade and number host channels, which no longer apply with CFW V7.60.

The DS5100 can be upgraded to the DS5300 by using the Performance Upgrade premium feature.

<i>Dual-controller system (unless noted)</i>	DS5100	DS5300
Number of host channels	Two (iSCSI only), Eight, or Sixteen	
Host channel topologies	4 Gbps FC, 8 Gbps FC, 1 Gbps iSCSI	
Redundant drive channels	Sixteen 4 Gbps	
Max drives	448 FC/SATA	
Processor	Intel Xeon 2.8 GHz	
Processor memory (per controller)	2 GB	
XOR technology	Dedicated ZIP ASIC	
Internal Frontside Bus Speed	Reduced clock speed through internal code	Full clock speed
Internal controller bandwidth (per controller)	4 GB/s	
Data cache (min/max)	8 / 16 / 32 / 64 GB	
Cache Holdup	Permanent	
Cache Mirroring	Two dedicated busses	
Cache bandwidth (single controller)	17 GB/s	

Figure 3-70 DS5100 / DS5300 comparison chart

Figure 3-71 compares the DS5100 and DS5300 with their predecessor, the DS4800, and points out the main differences.

	DS4800		DS5000
Host interfaces	Eight 4 Gbps FC	200%	Eight or Sixteen 4 Gbps FC, 8 Gbps FC, 1 Gbps iSCSI
Max drives	224	200%	448
Data cache	4, 8 or 16 GB	400%	8 to 64 GB
Cache protection	Battery backed		Battery-backed and destaged to disk (permanent)
Cache mirroring	Across backend drive loops		Dedicated PCIeExpress busses
Internal bandwidth (single controller)	1 Gbps on single PCIeX buss	400%	4 Gbps on dual PCIeExpress busses
Cache bandwidth (ASIC to Cache)	3.2 Gbps	>500%	17 Gbps

Figure 3-71 Key improvements of the DS5000 over the DS4800

A vendor neutral company called Storage Performance Council has done a performance test on the DS5300, which we refer to for performance details. The results of this test can be found at the following address:

<http://www.storageperformance.org/results>

3.7 DS5000 series physical specifications

This section contains the physical specifications for the DS5020, DS5100, and DS5300 systems and the expansion units.

IBM System Storage DS5020 storage subsystem (1814-20A)

- ▶ Height: 129.5 mm (5.1 in.)
- ▶ Width: 482.6 mm (19.0 in.)
- ▶ Depth: 571.5 mm (22.5 in.)
- ▶ Weight
 - Drive-ready (without drive modules installed): 27.67 kg (61 lbs.)
 - Fully configured (16 drive modules installed): 39.92 kg (88 lbs.)

EXP520 Expansion Unit (1814-52A)

- ▶ Height: 129.5 mm (5.1 in.)
- ▶ Width: 482.6 mm (19.0 in.)
- ▶ Depth: 571.5 mm (22.5 in.)
- ▶ Weight:
 - Drive-ready (without drive modules installed): 26.31 kg (58 lbs.)
 - Fully configured (16 drive modules installed): 38.56 kg (84 lbs.)

IBM System Storage DS5100 (1818-51A) and DS5300 (1818-53A) storage subsystems

- ▶ Height: 174.50 mm (6.87 in.)
- ▶ Width: 481.75 mm (18.97 in.)
- ▶ Depth: 634.92 mm (25.0 in.)
- ▶ Weight: 40.90 kg (90.0 lbs.)

IBM System Storage EXP5000 Expansion Unit (1818-D1A)

- ▶ Height: 132.10 mm (5.20 in.)
- ▶ Width: 482.60 mm (19.00 in.)
- ▶ Depth: 558.80 mm (22.0 in.)
- ▶ Weight:
 - Drive-ready (without drive modules installed): 31.30 kg (69 lbs.)
 - Fully configured (16 drive modules installed): 38.56 kg (85 lbs.)

Operating environment

This section will cover operating environment specifications.

IBM System Storage DS5020 storage subsystem (1814-20A)

- ▶ Temperature (operating):
 - 10 to 35° C (50 to 95° F) at 0 to 914 m (0-3,000 ft.)
 - 10 to 32° C (50 to 90° F) at 914 to 2,133 m (3,000-7,000 ft.)
- ▶ Relative humidity (operating): 8% to 80%
- ▶ Relative humidity (storage): 5% to 80%
- ▶ Electrical power (per power supply, system rating):
 - Voltage range: 100-240V ac
 - Operating current: 6.0 - 2.5 amperes
 - Power: 600 watts
 - Frequency: 50/60 Hz
- ▶ Heat dissipation: 1529 BTU per hour
- ▶ Noise level (normal operation): 6.4 bels

IBM System Storage EXP520 (1814-52A) Expansion Unit

- ▶ Temperature (operating):
 - 10 to 35° C (50 to 95° F) at 0 to 914 m (0-3,000 ft.)
 - 10 to 32° C (50 to 90° F) at 914 to 2,133 m (3,000-7,000 ft.)
- ▶ Relative humidity (operating): 8% to 80%
- ▶ Relative humidity (storage): 5% to 80%
- ▶ Electrical power (per power supply, system rating):
 - Voltage range: 100-240V ac
 - Operating current: 6.0 - 2.5 amperes
 - Power: 600 watts
 - Frequency: 50/60 Hz
- ▶ Heat dissipation: 1516 BTU per hour (fully configured)
- ▶ Noise level (normal operation): 6.5 bels

IBM System Storage DS5100 (1818-51A) and DS5300 (1818-53A) storage subsystems

- ▶ Temperature (operating):
 - 10 to 35 degrees C (50 to 95 degrees F) at 0 to 914 m (0-3,000 ft.)
 - 10 to 32 degrees C (50 to 90 degrees F) at 914 to 2,133 m (3,000-7,000 ft.)
- ▶ Relative humidity (operating): 8% to 80%
- ▶ Relative humidity (storage): 5% to 80%
- ▶ Electrical power (per power supply, system rating):
 - Voltage range: 100-240V ac
 - Operating current: 5.4 - 2.25 amperes
 - Power: 580 watts
 - Frequency: 50/60 Hz
- ▶ Heat dissipation: 804 BTU per hour
- ▶ Noise level (normal operation): 6.75 bels

IBM System Storage EXP5000 (1818-D1A) Expansion Unit

- ▶ Temperature (operating):
 - 10 to 35 degrees C (50 to 95 degrees F) at 0 to 914 m (0-3,000 ft.)
 - 10 to 32 degrees C (50 to 90 degrees F) at 914 to 2,133 m (3,000-7,000 ft.)
- ▶ Relative humidity (operating): 8% to 80%
- ▶ Relative humidity (storage): 5% to 80%
- ▶ Electrical power (per power supply, system rating):
 - Voltage range: 100-240V ac
 - Operating current: 5.4 - 2.25 amperes
 - Power: 580 watts
 - Frequency: 50/60 Hz
- ▶ Heat dissipation: 1570 BTU per hour (fully configured)

- ▶ Noise level (normal operation): 6.75 bels

3.8 DS5000 supported operating systems

The intent of this section is to list the most popular supported operating system platforms for the DS5000 series. For a complete and up-to-date list, go to the following address:

<http://www-01.ibm.com/systems/support/storage/config/ssic/index.jsp>

The supported systems are:

- ▶ IBM Power Systems™
 - AIX 5L™ V5.3 and V6.1 TL2
 - Red Hat Enterprise Linux up to V5.3
 - Novell SUSE up to SLES 11
- ▶ IBM BladeCenter®
 - Microsoft Windows 2003 and 2008
 - Red Hat Enterprise Linux up to V5.3 (on Intel and Power)
 - Novell SUSE up to SLES 11 (on Intel and Power)
 - AIX 5L V5.3, and V6.1 TL2 (on Power)
- ▶ IBM System x, Intel, and AMD
 - Microsoft Windows 2003 and 2008
 - Red Hat Enterprise Linux up to V5.3
 - Novell SUSE up to SLES 11
- ▶ VMware V3.5 and V4.0
- ▶ Hewlett-Packard: HP-UX 11iV2 and 11iV3
- ▶ Solaris: Sun Solaris 9 U8 and 10 U7 (all architectures)

3.9 DS5000 storage subsystem disk enclosures

The EXP5000 is the expansion enclosure available for the DS5000 series, except for the DS5020 storage subsystem. The DS5020 has its own expansion enclosure called EXP520. The EXP810 can be attached to the DS5000 storage systems for purposes of migrating data from the EXP810 to the DS5000 series. However, a license is required to attach the EXP810 to a DS5020 storage subsystem.

- ▶ EXP5000 Storage Expansion Unit (1818-D1A)
- ▶ DS4000 EXP810 Storage Expansion Unit (1812-81A) (refer to “IBM System Storage EXP810 expansion enclosure” on page 27 for details)
- ▶ EXP520 Storage Expansion Unit (1814-52A)

In this section, we discuss the features of the DS5000 expansion units (EXPs) and how they can be attached and combined to expand the storage capacity of the DS5000 storage systems. In particular, we look at the possibilities and limitations of intermixing FC and SATA drives. Connectors, LED indicators, and switches located at the rear of the expansion units are similar on all models.

As part of the base machine functionality, the DS5000 controllers and EXP5000 are designed to support integration and use of up to 448 high-performance Fibre Channel or high-capacity SATA disk drive modules (DDMs).

When configuring more than 48 SATA DDMs on the DS5000 controllers and EXP5000s, one or more SATA Disk Drive Attachment features need to be selected to maintain entitlement and support. The SATA Disk Drive Attachment features have been withdrawn from market as of January 2009. Customers are now entitled to attach as many SATA drives as they want, up to the maximum of 448 drives.

3.9.1 EXP5000 and EXP520 Storage Expansion Unit

The EXP5000 (1818-D1A) and EXP520 (1814-52A) Storage Expansion Units are each packaged in a 3U rack-mountable, high-capacity 16-drive bay enclosure containing dual switched 4 Gbps ESMs, dual power supplies, and redundant cooling. EXP5000s connect to DS5100 and DS5300 controllers through high-speed 4 Gbps FC disk expansion ports and have a physical storage capacity of up to 16 TB per enclosure, using 1000 GB SATA DDMs (disk drive modules) and up to 9.6 TB per enclosure, using 600 GB FC DDMs. It is also possible to mix SATA and FC drives in the same enclosure. The EXP520 uses the same disk types and sizes, but only connects to DS5020 storage controllers.

The EXP5000 Expansion Unit (1818-D1A) and EXP520 (1814-52A) base model includes a 3U, rack-mount 16-bay disk enclosure, two software SFP transceivers, dual power supplies, redundant cooling, rack mounting rails, softcopy documentation, and two rack PDU power cords (36L8886). Any country-specific wall outlet power cord is obtained through feature number 98xx.

Both expansion units support the following drives:

- ▶ 146.8 GB, FC 4 Gbps, 15000 rpm (FDE and non-FDE)
- ▶ 300 GB, FC 4 Gbps, 15000 rpm (FDE and non-FDE)
- ▶ 450 GB, FC 4 Gbps, 15000 rpm (FDE and non-FDE)
- ▶ 600 GB, FC 4Gbps, 15000 rpm (FDE and non-FDE)
- ▶ 750 GB, SATA, 7200 rpm
- ▶ 1000 GB, SATA, 7200 rpm
- ▶ 76 GB, FC, solid state drive (only in EXP5000)

Note: The usable disk space is less than the overall disk capacity. Consider these usable capacity amounts for storage capacity calculation issues:

- ▶ 73.4 GB formatted capacity is equal to 67.85 GB usable capacity.
- ▶ 146.8 GB formatted capacity is equal to 136.21 GB usable capacity.
- ▶ 300 GB formatted capacity is equal to 278.88 GB usable capacity.
- ▶ 450 GB formatted capacity is equal to 415.58 GB usable capacity.
- ▶ 600 GB formatted capacity is equal to 558.28 GB usable capacity.
- ▶ 750 GB formatted capacity is equal to 697.98 GB usable capacity.
- ▶ 1000 GB formatted capacity is equal to 930.81 GB usable capacity.

The usable capacities are what the SMclient will report as storage that can be used by the hosts. We arrive at this number through the following steps:

1. Take the listed raw disk amount (listed in decimal, per the storage industry standard), multiply by 1000^3 , and divide by 1024^3 to get a raw binary capacity (1 decimal GB = 1,000,000,000 bytes. 1 binary GB = 1,073,741,824 bytes (2^{30} bytes)).
2. Subtract the 512 MB configuration database, that is, the DACstore (the region that holds controllers and configuration information) after converting to binary.

This gives you the usable binary capacity that can be utilized by hosts and is what the SMclient will report to you as usable capacity.

Front view

The front of the subsystem provides access to the sixteen portable canister contained drives. There are three summary LEDs at the bottom of the front panel, as shown in Figure 3-72:

- ▶ **White:** Locate (number 3)
This White Locator LED aids in module identification.
- ▶ **Amber:** Summary Fault (number 4)
When there is a fault condition, this LED glows amber.
- ▶ **Green:** Power (number 5)
This LED glows green when at least one power supply is operational.

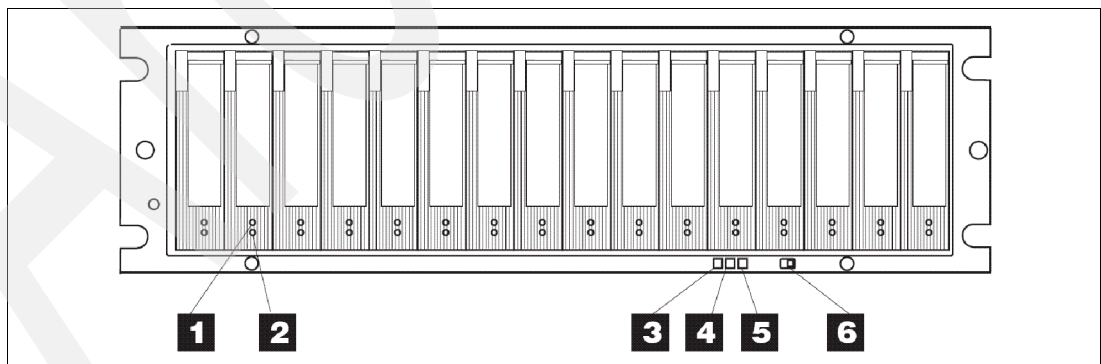


Figure 3-72 EXP5000 and EXP520 front panel

There are two LEDs on each drive tray (see Figure 3-72 on page 100):

- ▶ Drive Activity (number 1)
 - On (not blinking): No data is being processed.
 - Blinking: Data is being processed.
- ▶ Drive Fault (number 2)
 - Blinking: Drive, volume, or storage array locate function.
 - On: A problem has occurred.

The Link Rate switch (number 6) should be set to 4 Gbps (depending on your configuration).

Rear view

This section describes the primary LEDs, controls, and connectors on the rear of the storage expansion enclosure for all models. The back view in Figure 3-73 shows the following components:

- ▶ Fans and power supplies: Two removable power supply and fan unit FRUs, each containing one power supply and two fans
- ▶ ESMs: Two removable environmental services monitors (ESMs)

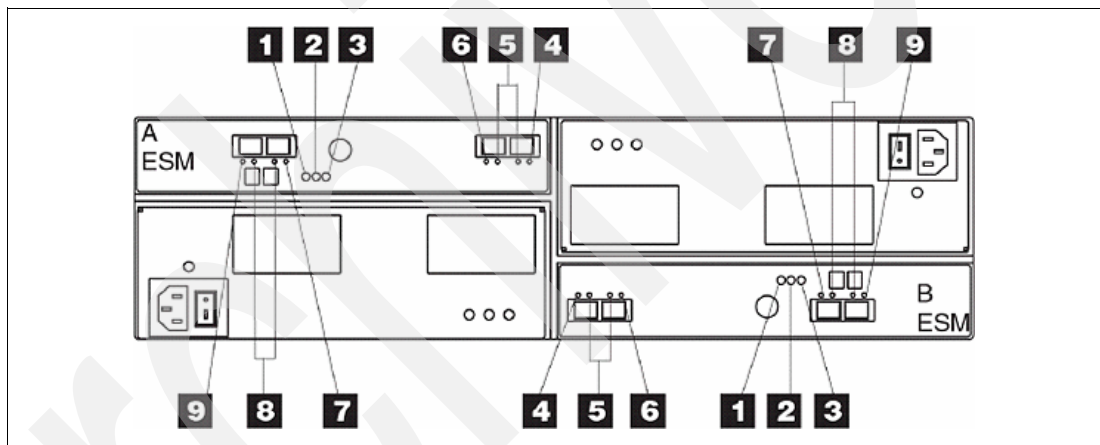


Figure 3-73 EXP5000 and EXP520 rear view

In Figure 3-73, you see the following LEDs and connectors:

- ▶ ESM Power LED (number 1): Normally ON.
- ▶ ESM Fault LED (number 2): ON when failures occur.
- ▶ ESM Service Action Allowed (number 3): When ON, can be removed.
- ▶ ESM Port 1 in Bypass (port is labeled 1A) (LED number 4): Normal status OFF when a cable is connected and ON if no cable is connected. Problem status ON when a cable is connected and a failure occurs.
- ▶ Data Rate LEDs (number 5): 2 Gbps - one LED is lit. 4 Gbps - two LEDs are lit.
- ▶ ESM Port 2 in Bypass (port is labeled 1B) (LED number 6): Normal status OFF when a cable is connected and ON if no cable is connected. Problem status ON when a cable is connected and a failure occurs.
- ▶ ESM ports 3 and 4 (LEDs number 7, 8, and 9) are reserved for future use.

EXP5000 and EXP520 ESM board

The ESMs contain the storage expansion enclosure control logic, interface ports, and LEDs. Each ESM has four SFP module ports that you could use to connect the storage expansion enclosure to the storage subsystem. However, only the two ESM SFP ports (labeled 1A and 1B) near the center of the storage expansion enclosure are used. The SFP ports labeled 2A and 2B are reserved for future use. Refer to Figure 3-74.

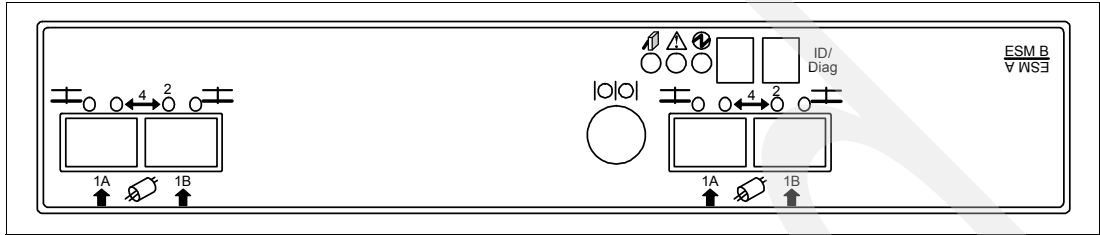


Figure 3-74 EXP5000 and EXP520 ESM ports

The EXP5000 and EXP520 ESM board allows for host based monitoring and control of the subsystem through SCSI-3 Enclosure Services (SES) over the FC link.

The EXP5000 and EXP520 supports two ESM CRUs to help provide system redundancy. Each ESM connects a separate Fibre Channel loop to one of the two Fibre Channel ports on the disk drives, which is designed to provide both loops access to all drives. It is designed to maintain access to all drives through the second loop and ESM. In the event of an FC ESM failure or inoperable loop, access to all drives is still available through the second loop and ESM.

There are switches inside ESM that allow direct access (1 hop) to a required disk from ESM, as shown in Figure 3-75.

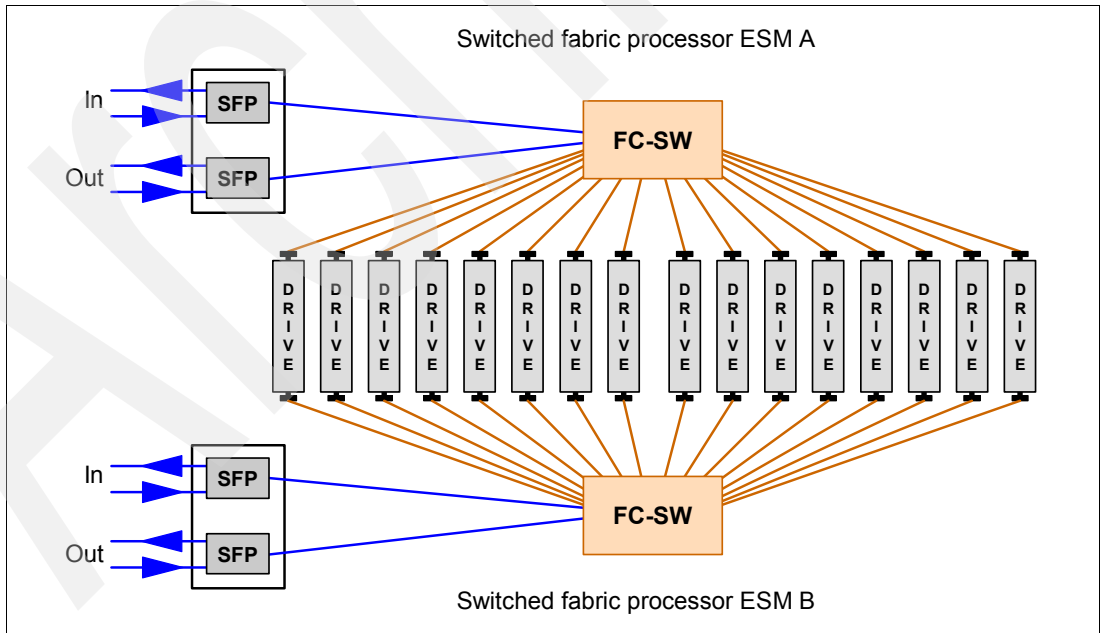


Figure 3-75 EXP5000 and EXP520 switched architecture

IBM System Storage DS planning and configuration

Careful planning is essential to any new storage installation. This chapter provides guidelines to help you with the planning process, and then allow a correct implementation.

Choosing the right equipment and software, and also knowing what the right settings are for a particular installation, can be challenging. Every installation has to answer these questions and accommodate specific requirements, and there can be many variations in the solution.

Having a well, thought-out design and plan prior to the implementation will help you get the most out of your investment for the present and protect it for the future.

During the planning process, you need to answer numerous questions about your environment:

- ▶ What are my host connection requirements?
- ▶ What additional hardware do I need?
- ▶ What reliability do I require?
- ▶ What redundancy do I need? (For example, do I need off-site mirroring?)
- ▶ What compatibility issues do I need to address?
- ▶ Will I use any storage virtualization product such as IBM SAN Volume controller?
- ▶ Will I use any unified storage product like the IBM System Storage N series?
- ▶ What operating system am I going to use (existing or new installation)?
- ▶ What applications will access the storage subsystem?
- ▶ What are the hardware and software requirements of these applications?
- ▶ What will be the physical layout of the installation? Only local site, or remote sites as well?
- ▶ What level of performance do I need?

This list of questions is not exhaustive, and as you can see, some go beyond simply configuring the DS5000 storage subsystem.

4.1 Planning your DS storage structure

Planning the details of the configuration of your DS Storage System is the key to a successful implementation. We cover in this section the different alternatives available to configure your system.

4.1.1 DS5000 arrays and RAID levels

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

When defining arrays, you often have to compromise among capacity, performance, and redundancy.

RAID levels

We go through the different RAID levels and explain why we choose a particular setting in a particular situation, and then you can draw your own conclusions.

RAID 0: For performance, but generally not recommended

RAID 0 (Figure 4-1) 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 0 for applications that will benefit from the increased performance capabilities of this RAID level. Never use this level for critical applications that require high availability.

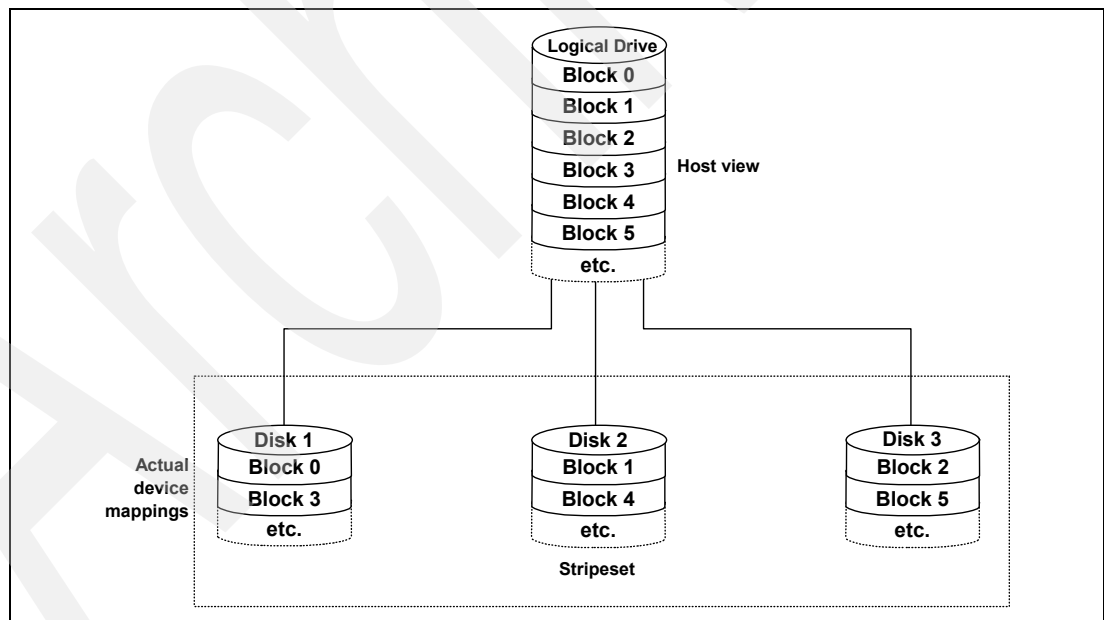


Figure 4-1 RAID 0

RAID 1: For availability and good read response time

RAID 1 (Figure 4-2) 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 cost is not a major concern.

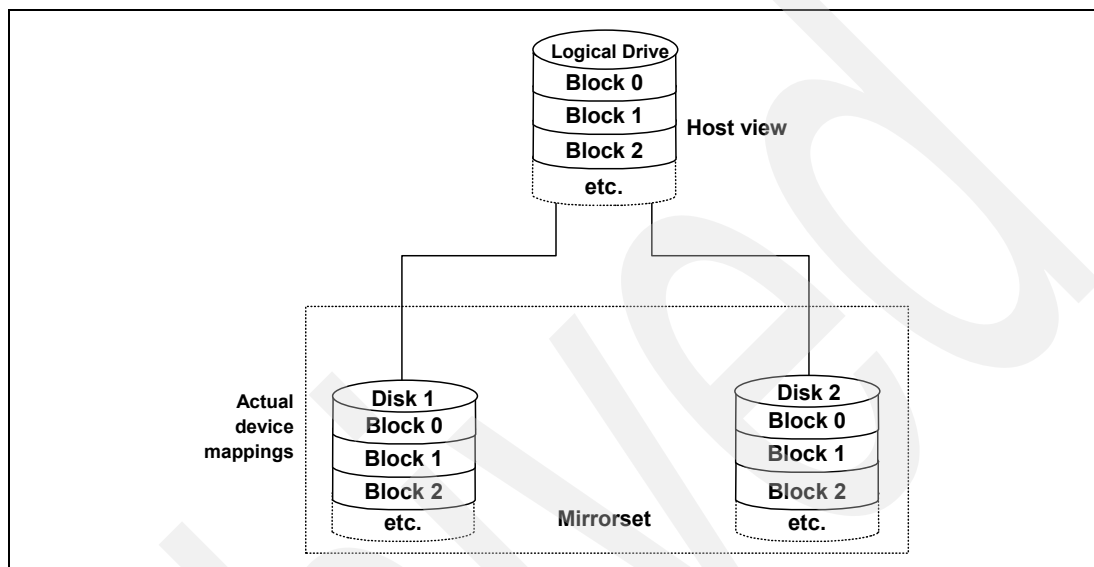


Figure 4-2 RAID 1

Because the data is mirrored, the capacity of the logical drive, when using RAID 1, is 50% of the array capacity.

Here are some recommendations when using RAID 1:

- ▶ 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 by using a write cache. When using a write cache, be sure it is battery-backed up.

Note: RAID 1 is actually implemented only as RAID 10 (see “RAID 10: Higher performance than RAID 1” on page 108) on the DS5000 storage subsystem products.

RAID 3: Sequential access to large files

RAID 3 is a parallel process array mechanism, where all drives in the array operate in unison. Similar to data striping, information that will 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.

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 operations. It 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 impact incurred by mirroring in RAID 1.

RAID 5: High availability and fewer writes than reads

RAID 5 (Figure 4-3) stripes data and parity across all drives in the array. RAID 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.

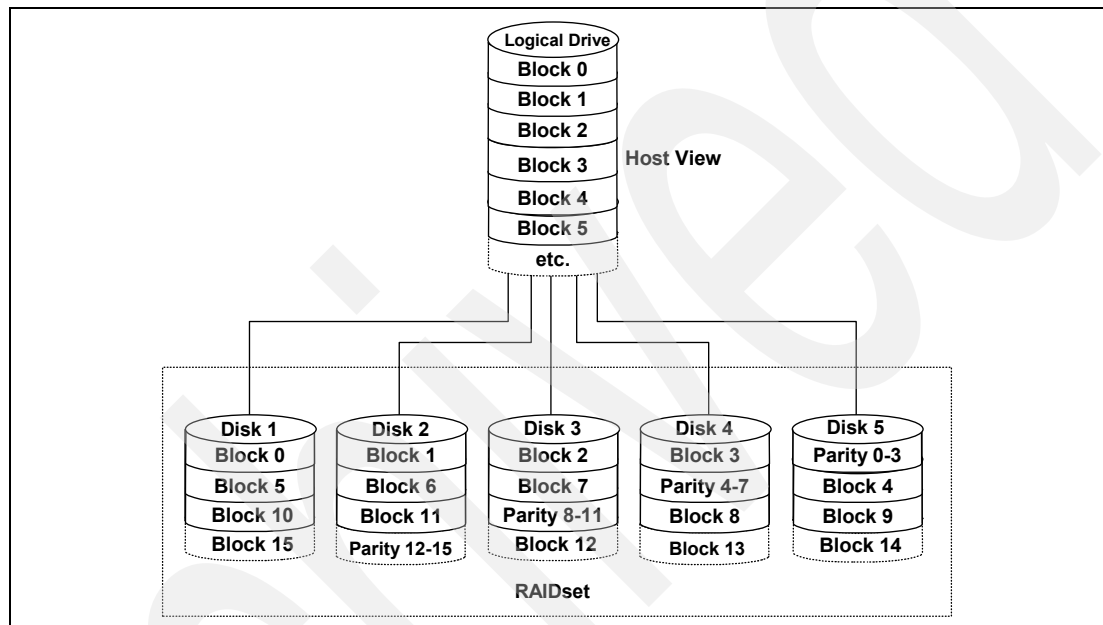


Figure 4-3 RAID 5

RAID 5 is best used in environments requiring high availability and 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 (random write-intensive) do not perform as well on RAID 5 logical drives because of the way 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; however, if the write activity consists of sequential, large block I/O, then RAID 5 is the best option.

Use write caching on RAID 5 arrays, because RAID 5 writes will not be completed until at least two reads and two writes have occurred. The response time of writes will be improved through the use of write cache (be sure it is battery-backed up). RAID 5 arrays with caching can give as good as performance as any other RAID level, and with some workloads, the striping effect gives better performance than RAID 1. For better cache usage, you can double the cache memory space by disabling the write cache mirror, so the controllers will not have a copy of the other cache data. Do that action only when data can be rewritten if it is lost, because if there is a controller failure, the cache data is lost.

RAID 6: High availability with additional fault tolerance

RAID 6 (see Figure 4-4) is a RAID level that employs $n+2$ drives, which can survive the failure of any two drives. 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 DS5000 storage subsystem, this calculation is made by the hardware and thus more performant 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 impact of more parity calculations, in terms of writing data, RAID 6 is slower than RAID 5, but might be faster in random reads thanks to the spreading of data over one more disks.

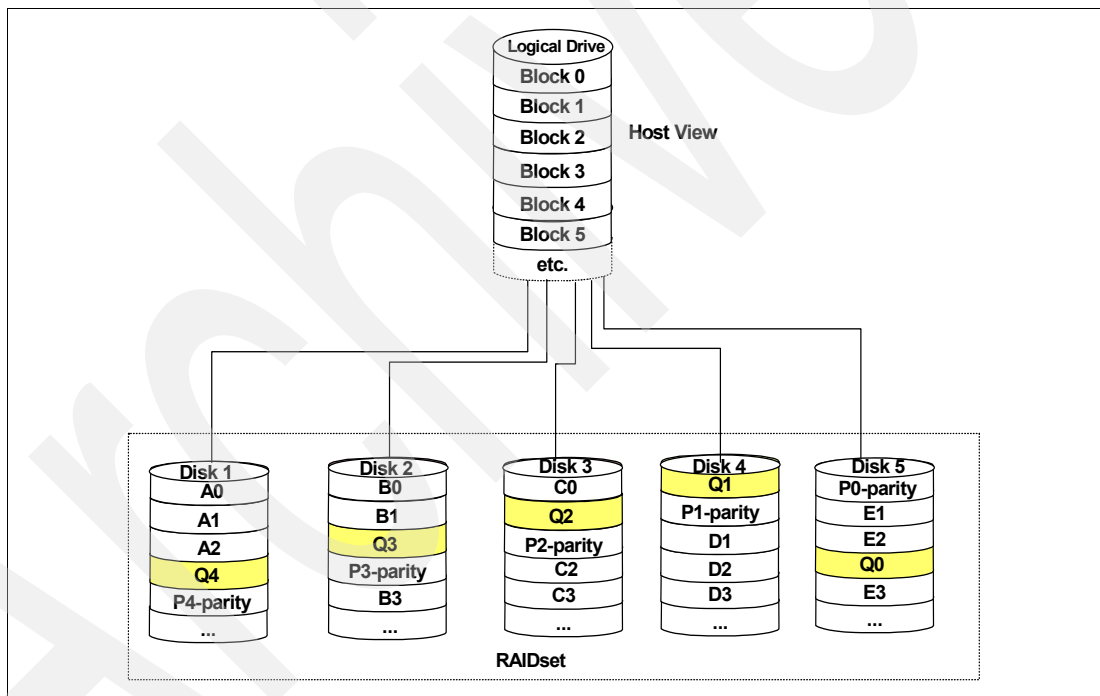


Figure 4-4 RAID 6

RAID 6 is striping with Dual Rotational Parity.

Table 4-1 shows the RAID 6 details.

Table 4-1 RAID 6 details

Feature	Description
Definition	Distributed parity: Disk striping and two independent parity blocks per stripe. Can survive the loss of two disks without losing data.
Benefits	Data redundancy, high read rates, and good performance.
Considerations	<ul style="list-style-type: none"> ▶ Requires two sets of parity data for each write operation, resulting in a significant decrease in write performance. ▶ There are additional costs because of the extra capacity required by using two parity blocks per stripe.
Uses	<ul style="list-style-type: none"> ▶ Any application that has high read request rates and average write request rates. ▶ Transaction servers, Web servers, data mining applications, and Exchange servers.
Drives	Minimum of three.
Fault Tolerance	Yes.

RAID 10: Higher performance than RAID 1

RAID 10 (Figure 4-5), 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.

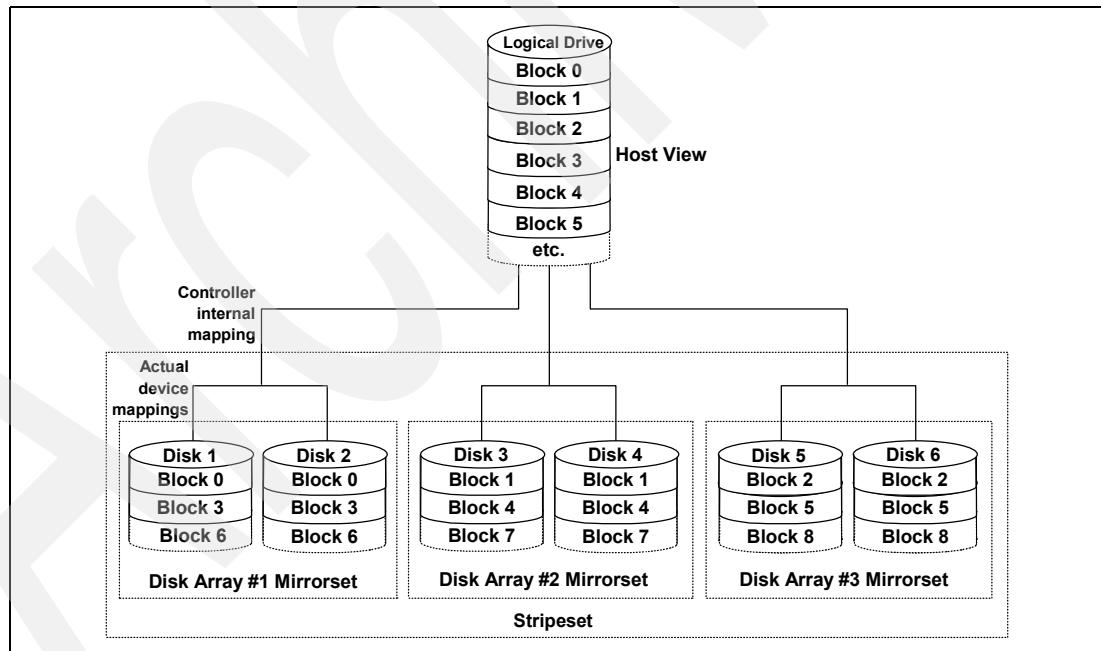


Figure 4-5 RAID 10

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.

The recommendations for using RAID 10 are as follows:

- ▶ 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 a RAID 10 write will not be completed until both writes have been done, write performance can be improved through the use of a write cache (be sure 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.

The following note and Table 4-2 summarize this information.

Summary: 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 offers high availability with performance slightly lower than RAID 5.
- ▶ RAID 10 offers higher performance than RAID 1 and more reliability than RAID 5 and is the best option for random write performance.

Table 4-2 RAID levels comparison

RAID	Description	Application	Advantage	Disadvantage
0	Stripes data across multiple drives.	IOPS Mbps	Performance, due to parallel operation of the access.	No redundancy. If one drive fails, the data is lost.
1	The 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 the same number of disks.	IOPS	Performance, as multiple requests can be fulfilled simultaneously. This is the best option for random write performance. Most reliable RAID level.	Storage costs are doubled.

RAID	Description	Application	Advantage	Disadvantage
3	Drives operate 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, and graphics).	Degraded performance with 8-9 I/O threads, random IOPS, and smaller, more numerous IOPS.
5	Drives operate 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. Best for sequential writes, throughput intensive applications. Best option for mixed workload and space efficiency.	Random 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 than RAID 5 is required.	Slower in random writing data, complex RAID controller architecture.

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 would appear 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 rebuild the stripe is not available. The end effect of this issue will depend on the type of data and how sensitive it is to corruption. While most storage subsystems (including the DS5000 storage subsystem) 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 faultless, 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 the appropriate software, according to business needs.

Array configuration

Before you can start using the physical disk space, you must configure it. You divide your (physical) disk drives into arrays and create one or more logical drives inside each array.

In simple configurations, you can use all of your drive capacity with just one array and create all of your logical drives in that unique array. However, this presents the following drawbacks:

- ▶ If you experience a (physical) drive failure, the rebuild process affects all logical drives, and the overall system performance goes down.
- ▶ Read/write operations to different logical drives are still being made to the same set of physical hard drives.

The array configuration is crucial to performance. You must take into account all the logical drives inside the array, as all logical drives inside the array will impact the same physical disks. If you have two logical drives inside an array and they both are high throughput, then there might be contention for access to the physical drives as large read or write requests are serviced. It is crucial to know the type of data that each logical drive is used for and try to balance the load so contention for the physical drives is minimized. Contention is impossible to eliminate unless the array only contains one logical drive.

Drive types

There are three different disk drive types available for the DS Storage Systems:

- ▶ Fibre Channel (FC) disks (Encryption capable FDE or not)
- ▶ Solid State Drives (SSD)
- ▶ Serial ATA disks (SATA)

RAID arrays can only be created by using the same disk types. Depending the usage planned for a specific array, you might want to choose the right disk type to optimize the overall performance.

If your application demands high levels of throughput, then both SATA and FC disk types have similar performance values.

SSD drives, on the other hand, are much better for I/O intensive applications, but not as good as FC or SATA for throughput. If your application is critical for I/O operations, then your best disk selection is SSD, then FC; avoid using SATA drives.

Tip: Select the best drive types for your array, depending on your application needs:

- ▶ For I/O demanding applications, use as first choice SSD, then FC, and then SATA.
- ▶ For throughput demanding applications, use FC or SATA as your first options.

Number of drives

The more physical drives you have per array, the shorter the access time for read and write I/O operations is. So having many disk drives per array will benefit the transactional environment, where a high number of I/O operations per second are needed. However, this is not the same number of drives that the sequential access applications need, so the number of drives to select per array have to be considered according to the application environment.

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 transactional 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. However, pay attention to the protection level, because with larger disk drives, the rebuild time might expose you to a second failure and lead to an offline array.

Enclosure layout and loss protection planning

Depending on the DS storage system and expansions in your configuration, you need to plan in advance how these expansions will be best interconnected with your DS system controller. In order to optimize performance from each controller to each expansion, you need to determine the best layout of your cabling configuration.

Especially in large configurations with multiple expansion enclosures, you can provide better controller access to the drives, and thus optimize performance, by selecting array members from different enclosures. Other considerations to optimize the DS storage subsystem is to evenly spread disk array members from odd and even slots, because each controller has a preferred path to odd or even slots.

Enclosure loss protection is a good way to make your system more resilient against hardware failures. Enclosure loss protection means that you spread your protection arrays across multiple enclosures rather than in one enclosure so that a failure of a single enclosure does not take a whole array offline.

By default, the automatic configuration is enabled. However, this is not a best practice for the method of creating arrays. Instead, use the manual method, as this allows for more configuration options to be available at creation time.

Best practice: Manual array configuration allows for greater control over the creation of arrays because it allows you to specify your planned optimal configuration options.

Figure 4-6 shows an example of enclosure loss protection. If enclosure number 2 fails, the array with the enclosure loss protection would still function (in a degraded state), as the other drives are not affected by the failure.

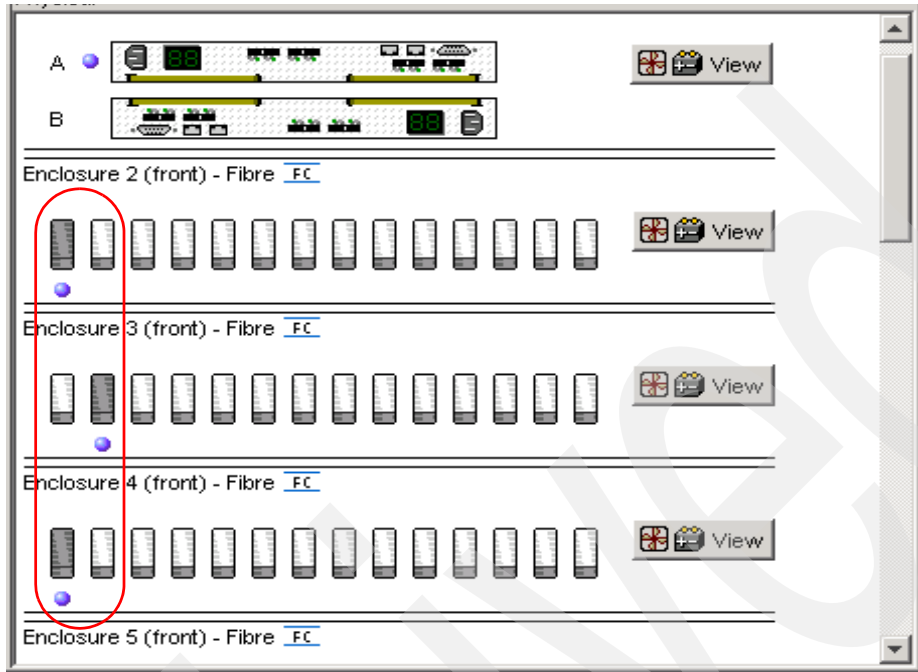


Figure 4-6 Enclosure loss protection

In the example shown in Figure 4-7, without enclosure loss protection, if enclosure number 2 fails, the entire array becomes inaccessible.

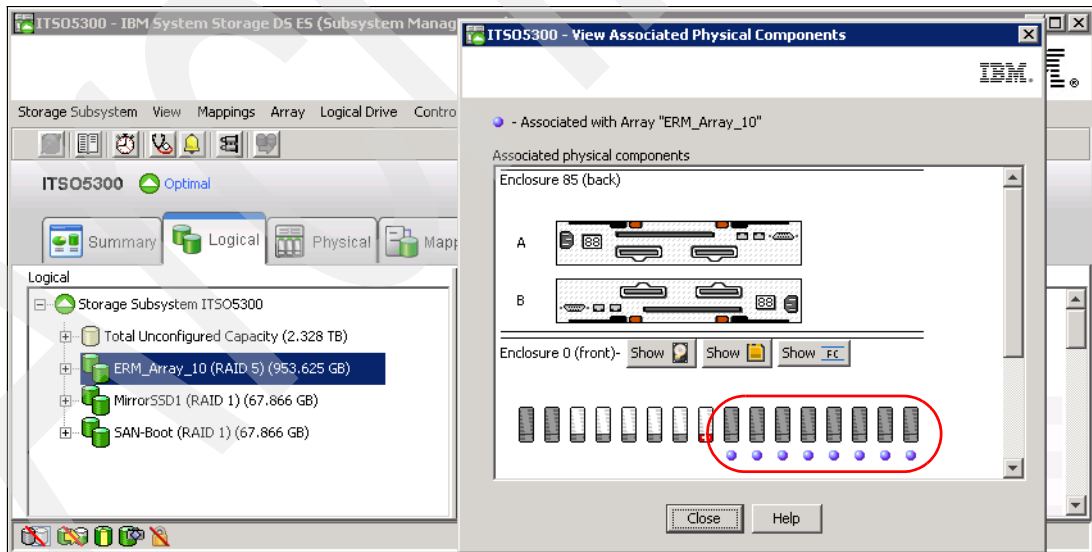


Figure 4-7 Array without enclosure loss protection

Best practice: Plan to use enclosure loss protection for your arrays.

4.1.2 Logical drives and controller ownership

Logical drives, sometimes simply referred to as volumes or LUNs (LUN stands for Logical Unit Number and represents the number a host uses to access the logical drive), are the logical segmentation of arrays. A logical drive is a logical structure you create on a storage subsystem for data storage. A logical drive is defined over a set of drives (called an array) and has a defined RAID level and capacity (see 4.1.1, “DS5000 arrays and RAID levels” on page 104). The drive boundaries of the array are hidden from the host computer.

The IBM System Storage DS5000 storage subsystem provides great flexibility in terms of configuring arrays and logical drives. However, when assigning logical volumes to the systems, it is very important to remember that the DS5000 storage subsystem uses a preferred controller ownership approach for communicating with LUNs. This means that every LUN is owned by only one controller. It is, therefore, important at the system level to make sure that traffic is correctly balanced among controllers. This is a fundamental principle for a correct setting of the storage system.

Balancing traffic is unfortunately not always a trivial task. For example, if an application requires a large disk space to be located and accessed in one chunk, it becomes harder to balance traffic by spreading the smaller volumes among controllers.

In addition, typically, the load across controllers and logical drives is constantly changing. The logical drives and data accessed at any given time depend on which applications and users are active during that time period, which is why it is important to monitor the system.

Best practice: Here are some guidelines for LUN assignment and storage partitioning:

- ▶ Assign LUNs across all controllers to balance controller utilization.
- ▶ Use the manual method of creating logical drives. This allows greater flexibility for configuration settings, such as enclosure loss protection and utilizing both drive loops.
- ▶ If you have highly used LUNs, where possible, move them away from other LUNs and put them on their own separate array. This will reduce disk contention for that array.

Enhanced Remote Mirror considerations

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 subsystem, controller A owns the associated secondary logical drive in the secondary storage subsystem. If controller ownership changes on the primary logical drive, then this will cause a corresponding controller ownership change of the secondary logical drive.

4.1.3 Hot spare drive

A hot spare drive is like a replacement drive installed in advance. Hot spare disk drives provide additional protection that might prove to be essential if there is a disk drive failure in a fault tolerant array.

Because a DS5000 storage subsystem configuration might have different disk types (FC, SATA, FDE, and SSD) with different capacities and speeds, you have to plan in advance how to provide the best hot spare protection for your storage subsystem.

Follow these guidelines to plan your hot spare coverage properly:

- ▶ Hot spare disk drives must be of the same media type and interface type as the disk drives that they are protecting.
- ▶ Similarly, hard disk drives can only be hot spares for other hard disk drives, not Solid State Drives (SSDs).
- ▶ Hot spare disk drives must have capacities equal to or larger than the used capacity on the disk drives that they are protecting. The DS5000 storage subsystem can use a larger drive to recover a smaller failed drive to it. It will not use smaller drives to recover a larger failed drive. If a larger drive is used, the remaining excess capacity is blocked from use.
- ▶ FDE disk drives provide coverage for both security capable and non-security capable disk drives. Non-security capable disk drives can provide coverage only for other non-security capable disk drives.
 - For an array that has secured FDE drives, the hot-spare drive should be an unsecured FDE drive of the same or greater capacity.
 - For an array that has FDE drives that are not secured, the hot-spare drive can be either an unsecured FDE drive or a non-FDE drive.

In a mixed disk environment that includes non-security capable SATA drives, non-security-capable Fibre Channel drives, and FDE Fibre Channel drives (with security enabled or not enabled), use at least one type of global hot-spare drive (FDE Fibre Channel and a SATA drive) at the largest capacity within the array. If a secure-capable FDE Fibre Channel and SATA hot-spare drive are included, all arrays are protected.

Important: When assigning disks as hot spares, make sure they have enough storage capacity. If the failed disk used capacity is larger than the hot spare, reconstruction is not possible. Ensure that you have at least one of each size or all larger drives configured as hot spares.

Hot spares locations

Distribute the hot spare drives evenly across the different expansions of your storage subsystem, but avoid having multiple ones in a single enclosure. Because hot spare drives are in standby, without traffic or I/O until a drive fails, then you want to maximize the overall performance of your system by evenly distributing your production drives across the different expansions. At the same time, this avoids the risk of a single disk drive channel, or expansion enclosure failure, causing loss of access to all hot spare drives in the storage subsystem.

However, in some configurations, for example, a DS5000 storage subsystem with five expansions evenly distributed across the four different drive channels to maximize the traffic to each enclosure, you can choose to maximize performance over availability by having all the spares defined in the fifth expansion. This way, the channel with two expansions will not be penalized for excessive traffic, because the spares expansion will not contribute to the traffic load in that channel.

Important: Distribute your spare drives evenly across the different expansion to avoid the possibility of a general enclosure failure.

Quantity and type of hot-spares drives

There is no fixed rule about the quantity of disk drives to assign as hot spares, but as a general rule about disk usage and availability, we recommend defining one of every 30 drives of a particular media type and interface type, or one for every two fully populated enclosures. Because of disk sizes, and especially in large configurations with arrays containing numerous drives, the reconstruction of a failed drive to a hot spare drive can take a long time, proportional to the quantity of drives in the array and the size of the disks. If in addition to that time, you have to wait to have a new disk available onsite to replace the failed drive, then the probability of having another disk failure increases. Having multiple spare drives will not mitigate the reconstruction time, but at least an array will be prepared sooner for a second disk failure.

Note: There is no definitive recommendation about how many hot spares you should install, but it is common practice to use a ratio of one hot spare for about 30 drives.

4.1.4 Storage partitioning

Storage partitioning adds a high level of flexibility to the DS5000 storage subsystem. It enables you to connect multiple and heterogeneous host systems to the same storage subsystem, 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 they access.

Without storage partitioning, the logical drives configured on a DS5000 storage subsystem can only be accessed by a single host system or by a single cluster. This can lead to inefficient use of storage subsystem hardware unless the use of the DS5000 storage subsystem is dedicated to a single host (for example, SVC attachment, where it is seen as a single host).

Storage partitioning, on the other hand, allows the creation of “sets”, containing the hosts with their host bus adapters and the logical drives. We call these sets *storage partitions*. 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 allows mapping and masks the logical device or LUN (that is why it is also referred to as LUN masking). This means that after the LUN is assigned to a host, it is hidden from all other hosts connected to the same storage subsystem. Therefore, access to that LUN is exclusively reserved for that host.

It is a good practice to configure storage partitioning prior to connecting multiple hosts. Operating systems such as AIX or Windows will write their signatures to any device it can access.

Heterogeneous host support means that the host systems can run different operating systems. But be aware that all host systems within a particular storage partition have unlimited access to all logical drives in this partition. Therefore, file systems on these logical drives must be compatible with host systems. To ensure that this is true, it is best to run the same operating system on all hosts within the same partition. Some operating systems might be able to mount foreign file systems.

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. To map a logical drive or LUN to a specific host server or group of hosts, each component of the storage partition must be defined.

A storage partition contains several components:

- ▶ Host 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.

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

A *host port* is the identifier of a port on a Fibre Channel host bus adapter (HBA), identified by its world-wide name (WWN), or the identifier of the initiator name running the iSCSI protocol. A single host can contain more than one host port. If the servers are attached in a redundant way (highly recommended), each server will have two host bus adapters, that is, it needs two host ports within the same host system. It is possible to have a host with a single HBA, but for redundancy, it should be able to access both DS5000 controllers. When using FC, this redundancy can be achieved using SAN zoning, and when using iSCSI, it is provided by the Ethernet switch.

The DS5000 storage subsystem only communicates through the use of this host port's identifiers. 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 reflect a logical view of the physical connections to the storage, 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 no host (or host port, to be precise) can access these logical drives until they are mapped.

With Storage Manager, it is possible to have up to 512 storage partitions on a DS5000. This allows the storage subsystem to have storage capacity to a greater amount of heterogeneous hosts, allowing for greater flexibility and scalability.

For the number of maximum storage partitions for a specific model, see Table 4-3. Note that on DS5000 models that the number of partitions also depends on the premium feature licenses that have been purchased.

Table 4-3 LUN restrictions

Operating system	Maximum number of LUNs supported per partition
Windows Server 2003	255
Windows Server 2008	255
HP-UX	127
AIX	255
Linux	255

Every mapping of a logical drive to a new host or host group creates a new storage partition. If additional logical drives are required for an existing host or host group, a new storage partition is not required. For example, a cluster with two nodes with redundant I/O paths are configured as one host group with two hosts. Each host has two host ports for redundancy. Several logical drives are mapped to this host group. All these components represent one storage partition. If another single host system is attached to the same storage system and different logical drives mapped to that host, another storage partition must be created for it. If a new logical drive is created and mapped it to either the cluster or the single host, it uses an existing storage partition.

For a step-by-step guide, see 4.1.4, “Storage partitioning” on page 116.

Note: There are limitations as to how many logical drives you can map per host. DS5000 series storage subsystems will support up to 256 LUNs (including the access LUN) per partition (although there are also restrictions, depending on the host operating system) and a maximum of two partitions per host. Keep these limitations in mind when planning the DS5000 storage subsystem installation.

Storage partitioning considerations

In a heterogeneous environment, the “access” logical drive is mapped only to the default host group. If in-band management is being used, then the access logical drive must be mapped to the storage partition for the managing host.

In a security-sensitive environment, you can also assign the access logical drive to a particular storage partition and ensure host-based management access only through the servers in this storage partition. In this environment, you can assign a password to the storage system as well.

Note: Each host with uniquely assigned storage will use a storage partition. Each host group with uniquely assigned storage will also use an additional storage partition.

For Fibre Channel attachments, in order to do the storage partitioning correctly, you need the WWN of your HBAs. Mapping is done on a WWN basis. Depending on your HBA, you can obtain the WWN either from the BIOS or QLogic SANsurfer tool if you have QLogic cards. Emulex adapters and IBM adapters for IBM System p and IBM System i® servers have a sticker on the back of the card. The WWN is also usually printed on the adapter itself or the box in which the adapter was shipped.

If you are connected to a hub or switch, check the Name Server Table of the hub or switch to identify the WWN of the HBAs.

For iSCSI attachment, you need to know the host port identifier, or iSCSI initiator name of the cards running iSCSI protocol. The iSCSI initiator can be implemented by specific host bus adapter hardware called *iSCSI cards*, or by software that runs on the host and provides connectivity through an Ethernet network interface card (NIC). iSCSI names, as the WWPN in FC, have a global scope, are independent of address or location, and are persistent and globally unique. Names must be extensible and scalable with the use of naming authorities. Changing this name might cause conflicts within the system. The standard supported for the DS5000 reading iSCSI naming convention is the or iSCSI qualified name (IQN) standard.

When planning your partitioning, keep in mind that:

- ▶ In a cluster environment, you need to use host groups.
- ▶ You can optionally purchase partitions up to the limit of the specific DS model.
- ▶ A given partition should not contain hosts that are FC-based as well as hosts that are iSCSI-based.
- ▶ A single host should not be configured for both iSCSI connections and FC connections to the storage system.

When planning for your storage partitioning, you should create a table of planned partitions and groups so that you can clearly map out and define your environment.

Best practice: If you have a single server in a host group that has one or more LUNs assigned to it, we recommend that you do the mapping to the host and not the host group. All servers having the same host type (for example, Windows servers) can be in the same group if you want, but by mapping the storage at the host level, you can define what specific server accesses which specific LUN.

However, if you have a cluster, it is good practice to assign the LUNs at the host group, so that all of the servers on the host group have access to all the LUNs.

Table 4-4 shows an example of a storage partitioning plan. This clearly shows the host groups, hosts, port names, WWN of the ports, and the operating systems used in that environment. Other columns can be added to the table for future references, such as HBA BIOS levels, driver revisions, and switch ports used, all of which can then form the basis of a change control log.

Table 4-4 Sample plan for storage partitioning

Host group	Host name	Port alias	Host port ID	OS type
Windows 2008	Mail	MailAdp_A	200000E08B28773C	Windows 2008 Non-Clustered
		MailAdp_B	200000E08B08773C	
Linux	Lin_Host	LinAdp_A	200100E08B27986D	Linux
		LinAdp_B	200000E08B07986D	
System p	AIX1	AIX1Adp_A	20000000C926B6D2	AIX
		AIX1Adp_B	20000000C926B08	
iSCSI Windows	TSM	TSM_iP0	iSCSI IP + IQN ^a	Windows 2008 Non-Clustered
		TSM_iP1	iSCSI IP + IQN ^a	

a. *IQN or iSCSI qualified name

Heterogeneous hosts

When implementing a DS5000 storage subsystem solution, a mixture of host servers with different operating systems can be used (clustered and non-clustered variants of the same operating systems). However, all logical drives in a single storage partition must be configured for the same operating system. Also, all hosts in that same storage partition must run the same defined operating system.

Important: Heterogeneous hosts are only supported with storage partitioning enabled.

Access logical drive

The DS5000 storage subsystem will automatically create a LUN 31 for each host attached. This is used for in-band management, so if you do not plan to manage the DS5000 storage subsystem from that host, you can delete LUN 31, which will give you one more LUN to use per host.

You can always recreate it if later you decide to use in-band management.

4.1.5 Segment size

A segment, in a logical drive, is the amount of data, in kilobytes, that the controller writes on a single physical drive before writing data on the next physical drive. Depending on your data structure, you can optimize the creation of your logical drives for this value.

When you create a logical drive, the default segment size choose is 128 KB, unless you are creating logical drives under a RAID 3, or choosing the multimedia type for the logical drive, in which case a segment size of 256 KB is used.

The choice of a segment size can have a major influence on performance in both IOPS and throughput. You can choose the defaults, which is a good choice for general usage, or consider the following to properly select a the logical drive segment size customized for your needs:

- ▶ The characteristics of your application, IOPS demanding, or Throughput demanding, whether random or sequential.
- ▶ The I/O block size of the host that will use the logical drive.
 - Small IO sizes allow for greater transaction (IOPS) rates.
 - Large IO sizes allows a better transfer rate in MBps.
- ▶ The RAID protection and number of disk drives that are participating in the logical drive array.

To calculate the best segment size in this environment, consider the RAID protection and number of drives participating in the array.

Segment Size > I/O block size request

This is the best option for high IOPs and random requests.

- ▶ Several transactions are satisfied in a single operation to a disk.
- ▶ There is higher IOPS.
- ▶ It is ideal for random I/O requests, such as the database file system.
- ▶ The best option is start with Segment size $\geq 2 \times$ Block Size for high IOPS.

Segment Size = I/O block size request

It is very difficult to align segment size with block size, so this option is not practical to implement, so use the other two scenarios.

- ▶ Every request transaction uses exactly one disk operation.
- ▶ There is high IOPS.
- ▶ This option is ideal for random I/O requests, such as the database file system.

Segment Size < I/O block size request

This is the best option to obtain MBps performance and low IOPs, which is the norm for multimedia applications.

- ▶ More disks are used or requested.
- ▶ There is higher throughput (Mbps).
- ▶ It is ideal for sequential writes, such as a large multimedia application.
- ▶ It is optimized when a single I/O request can be serviced with a single or exact multiple data stripes (the segment size multiplied by the number of drives in the array that are used for I/O). In this case, multiple disks are used for the same request, but each disk is only accessed once, or the exact number of times if the block size is too big and you have to use multiple stripes.

To calculate the best segment size in this environment, consider the RAID protection and number of drives participating in the array by using the following formula:

$$\text{Segment size} = (\text{Block size} / X)$$

Where X is the number of drives used in the array for I/O and where:

- RAID 0=Number of drives
- RAID 1 or 10=Number of drives / 2
- RAID 5= Number of drives - 1
- RAID 6= Number of drives - 2

If the resulting Segment Size is >512 KB, then divide it by an integer number starting with 2 to obtain multiple data stripes of each block requested.

Tips: The possible segment sizes available are 8 KB, 16 KB, 32 KB, 64 KB, 128 KB, 256 KB, and 512 KB.

- ▶ Storage Manager sets a default block size of 128 KB for every logical volume, except for RAID 3 volumes, which are set to 256 KB.
- ▶ For database applications, block sizes between 32–128 KB have shown to be more effective.
- ▶ In a large file environment, such as media streaming or CAD, 128 KB or more are recommended.
- ▶ For a Web server or file and print server, the range should be between 16–64 KB.

The performance monitor can be used to evaluate how a given segment size affects the workload.

Note: A performance testing schedule should be undertaken in the environment before going into production with a given segment size. Segment size can be dynamically changed, but only by rewriting the data, which consumes bandwidth and impacts performance. Plan this configuration carefully to avoid having to reconfigure the options chosen.

4.1.6 Cache parameters

Cache memory is an area of temporary volatile storage (RAM) on the controller that has a faster access time than the drive media. This cache memory is shared for read and write operations.

Efficient use of the RAID controller cache is essential for good performance of the DS5000 storage subsystem.

Figure 4-8 shows a schematic model of the major elements of a disk storage system, which are elements through which data moves (as opposed to other elements, such as power supplies). In the model, these elements are organized into eight vertical layers: four layers of electronic components shown inside the dotted ovals and four layers of paths (that is, wires) connecting adjacent layers of components to each other. Starting at the top in this model, there are some number of host servers (not shown) that connect (over some number of paths) to host adapters. The host adapters connect to cache components. The cache components, in turn, connect to disk adapters that, in turn, connect to disk drives.

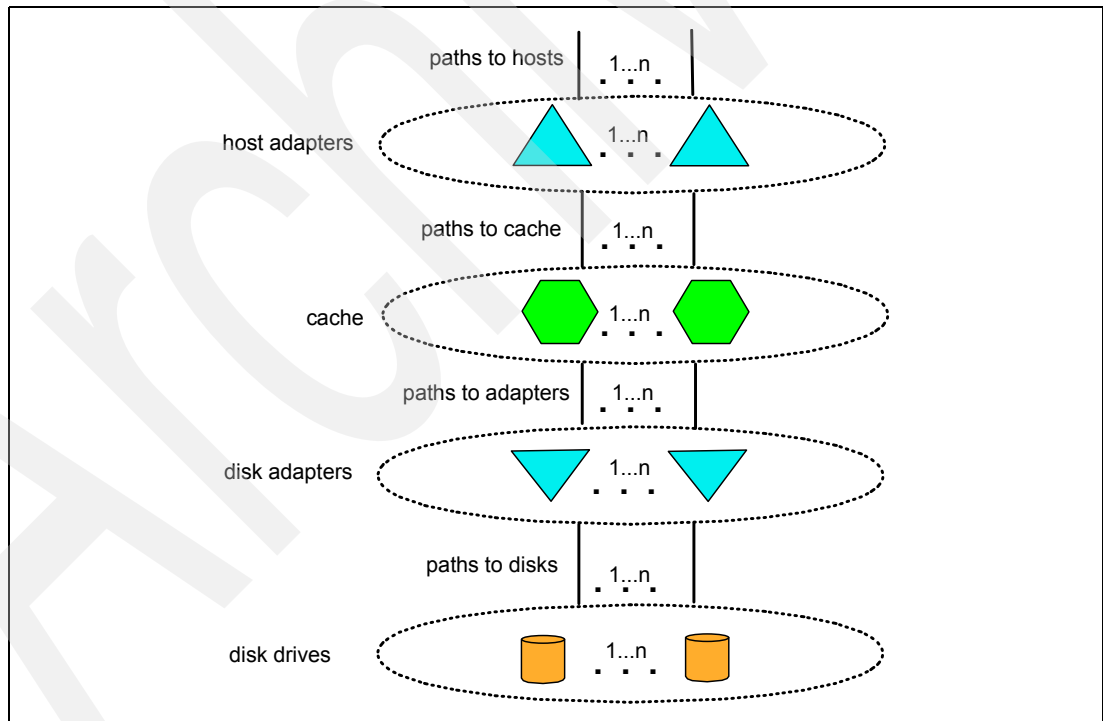


Figure 4-8 Conceptual model of disk caching

In this model, a read I/O request is handled where a host issues a read I/O request that is sent over a path (such as a Fibre Channel) to the disk system. The request is received by a disk system host adapter, which checks whether the requested data is already in cache. In such a case, it is immediately sent back to the host. If the data is not in cache, the request is forwarded to a disk adapter that reads the data from the appropriate disk and copies the data into cache. The host adapter sends the data from cache to the requesting host.

Most (hardware) RAID controllers have some form of read or write caching. These caching capabilities can be used, as they enhance the effective I/O capacity of the disk server. The principle of these controller-based caching mechanisms is to gather smaller and potentially nonsequential I/O requests coming in from the host server (for example, SQL Server) and try to batch them with other I/O requests. Consequently, the I/O requests are sent as larger (32 KB to 128 KB), and possibly sequential, requests to the hard disk drives. The RAID controller cache arranges incoming I/O requests by making the best use of the hard disk's underlying I/O processing ability. This increases the disk I/O throughput.

There are many different settings (related to caching) that come into play. The IBM System Storage DS Storage Manager utility enables the following settings to be configured:

- ▶ DS5000 system wide settings:
 - Start and stop cache flushing levels (this setting will affect all arrays and logical drives created on the system)
 - Cache Block size
- ▶ Logical drive specific settings:
 - Read caching
 - Dynamic read cache prefetch
 - Write caching or write-through mode (write caching disabled)
 - Enable or disable write cache mirroring

The default parameters configured in the Storage Manager are generally good for normal installations. Select the defaults, or see 4.10.6, “Cache parameters” on page 249 for specific configuration details.

When implementing a DS5000 storage subsystem as part of a whole solution, performance testing and monitoring to adjust the settings should be planned over at least a one week period.

4.1.7 Security planning: Full Disk Encryption

Businesses must comply with a growing number of corporate standards and government regulations, Drive Security is one tool that can enhance security by complying with these new standards and regulations.

Full Disk Encryption

This is a new premium feature where Full Disk Encryption (FDE) protects the data on the disks only when the drives are removed from storage subsystem enclosures. Drive Security and FDE requires security capable drives and provide access to data only through a controller that has the correct security key.

FDE does not prevent someone from copying the data in the storage subsystems through Fibre Channel host port connections when the drives are unlocked and operating. FDE also does not prevent unauthorized management access. A security capable drive encrypts data during writes and decrypts data during reads.

The security enabled drives can be used as normal drives and intermixed in an array with drives of equal type and capacity when this feature is not enabled. This new feature is detailed in Chapter 5, “Disk Security with Full Disk Encryption drives” on page 259.

Securing the DS5000 storage subsystem

Besides this new level of protection within the drives, it is important to plan ahead for other more trivial, but not less important, security issues, such as:

- ▶ Password protection: Storage Manager has the facility to set passwords to allow only authorized users access.
- ▶ Virtual Private Network (VPN) connection: Permits only authorized users access to your internal network.
- ▶ Remote Desktop Users/Passwords: Permits only authorized users to access your Storage Manager Client workstation.
- ▶ Plan redundancy for Storage Manager stations and networks:
 - One network card for the DS5000 Management LAN
 - One network card for the customer WAN
 - Physically separates the DS5000 management network from the rest of the enterprise network

4.2 Planning for premium features

Premium features are optional features that are activated, after being purchased, by registering a key using the Storage Manager software. When planning for any of the premium features, it is a good idea to document what are the goals and rationale for purchasing the feature. This clearly defines from the outset what you want to achieve and why. To learn the actual procedure to enable the premium features, see 7.2, “Handling premium features” on page 362

The following are the premium features that are available, depending your DS5000 storage subsystem:

- ▶ Disk Encryption or FDE
- ▶ Storage Partitioning
- ▶ Drive Slot Limit
- ▶ FlashCopy Logical Drives
- ▶ VolumeCopy
- ▶ Remote Logical Drive Mirroring
- ▶ FC/SATA Intermix
- ▶ Storage Partitioning

Feature packs

In addition to the premium features, which are normally available for all the DS5000 storage subsystems for one specific function, you can receive a feature pack activation, which is a bundle of optional features packed together in a single activation file. The basic feature pack shipped depends on the model of the DS5000 storage subsystem.

Here we describe the planning requirements of each feature pack. All of the needs and requirements should be documented.

4.2.1 Disk Encryption or FDE

This feature provides the licensing to enable the disk encryption functions. This feature is for activation only; you need to have also Encryption capable FC disk drives installed to actually make use of the encryption function. With only the encryption capable drives, denominated Full Disk Encryption (FDE) drives, or the activation feature, you cannot enable drive encryption; you must have both.

For more information about FDE, see 4.1.7, “Security planning: Full Disk Encryption” on page 123

4.2.2 Storage partitioning

Depending on the DS5000 storage subsystem, you can implement from 1 up to 512 partitions per storage subsystem.

Depending on the quantity of hosts you plan to attach to your DS5000 storage subsystem, you can request incremental licenses to activate the necessary partitions. Upon activation of the optional Storage Partition feature, the DS Storage Manager enables management of a single midrange disk storage system to function as multiple, logical storage systems to meet the needs of storage consolidation in heterogeneous environments.

4.2.3 Drive slot limit

Depending on the DS5000 storage subsystem, the basic limit of the number of drives allowed can be expanded by purchasing an additional feature and activating it in your server.

For all DS5000 models, the base number of disk drives supported can be upgraded up to 448 drives.

4.2.4 FlashCopy

A FlashCopy logical drive is a point-in-time image of a logical drive. It is the logical equivalent of a complete physical copy, but you create it much more quickly than a physical copy. Additionally, it requires less disk space. In Storage Manager, the logical drive from which you are basing the FlashCopy, called the base logical drive, must be a standard logical drive in the storage subsystem. Typically, you create a FlashCopy so that an application (for example, an application to take backups) can access the FlashCopy and read the data while the base logical drive remains online and user accessible.

FlashCopy takes only a small amount of space compared to the base image.

4.2.5 VolumeCopy

The VolumeCopy feature is a firmware-based mechanism for replicating logical drive data within a storage subsystem. This feature is designed as a system management tool for tasks, such as relocating data to other drives for hardware upgrades or performance management, data backup, and restoring Snapshot™ logical drive data.

A VolumeCopy creates a complete physical replication of one logical drive (source) to another (target) within the same storage subsystem. The target logical drive is an exact copy or clone of the source logical drive. VolumeCopy can be used to clone logical drives to other arrays inside the DS5000 storage subsystem. Careful planning should be considered with regard to the space available to make a FlashCopy of a logical drive.

The VolumeCopy premium feature must be enabled by purchasing a feature key. For efficient use of VolumeCopy, FlashCopy must be installed as well.

4.2.6 Enhanced Remote Mirroring (ERM)

The Enhanced Remote Mirroring option is a premium feature that comes with the Storage Manager software and is enabled by purchasing a premium feature key. The Enhanced Remote Mirroring option is used for online, real-time replication of data between two DS5000 storage subsystems over a remote distance.

The Enhanced Remote Mirroring is a redesign of the former Remote Volume Mirroring and offers both synchronous mirroring mode or asynchronous mirroring mode. Synchronous mode is also called Metro Mirror. Asynchronous mirroring with the consistency group option is referred to as Global Mirror, and asynchronous mirroring without the consistency group option is referred to as Global Copy.

Here we explain each method:

- ▶ **Metro Mirror**

Metro Mirror is a synchronous mirroring mode. Any host write request is written to the primary (local) storage subsystem and then copied to the secondary (remote) storage subsystem. The host application must wait until receiving acknowledgement that the write request has been executed on both (local and remote) storage controllers.

In the event of a communication failure, synchronous write mode offers the best chance of full data recovery from the secondary logical drive, so it is the recommended write mode when the maximum distance between the DS5000 storage subsystems in an Enhanced Remote Mirroring configuration is up to 10 km, although because it performs synchronous writes, host I/O performance is slower than it is in asynchronous write mode.

- ▶ **Global Copy**

Global Copy is an asynchronous write mode. All write requests from a host are written to the primary (local) storage subsystem and immediately reported as completed to the host system. The secondary (target) is written to at a later time independently. This capability is designed to prevent primary performance from being affected by wait time from writes on the secondary system. Therefore, the primary and secondary copies can be separated by long distances. This function is appropriate for remote data migration, offsite backups, and transmission of inactive database logs at virtually unlimited distances.

- ▶ **Global Mirror**

This mode of asynchronous copy must be used when the host application has dependent data spread across several logical drives to ensure that the dependent write requests are carried out in the same order at the remote site. This is done through the use of a consistency group. Global Mirror ensures that write requests to multiple primary logical drives are carried out in the same order on the secondary logical drives as they are on the primary logical drives, preventing data on the secondary logical drives from becoming inconsistent with the data on the primary logical drives

Asynchronous write mode, both Global Mirror or Global Copy, offers faster host I/O performance than synchronous write mode, but does not guarantee that the copy has been successfully completed before processing the next write request. This is the recommended mode when the maximum distance between the DS4000 or DS5000 storage subsystems in an Enhanced Remote Mirroring configuration is more than 10 km.

Enhanced Remote Mirroring has also been equipped with new functions for better business continuance solution design and maintenance tasks.

A minimum of two storage subsystems is required. One storage subsystem can have primary volumes being mirrored to arrays on other storage subsystems and hold secondary volumes from other storage subsystems. Also note that because replication is managed on a per-logical drive basis, you can mirror individual logical drives in a primary storage subsystem to appropriate secondary logical drives in several different remote storage subsystems.

Planning considerations for Enhanced Remote Mirroring

Here are some planning considerations:

- ▶ DS5000 storage subsystems (minimum of two)
- ▶ Fiber links between sites
- ▶ Distances between sites (ensure that it is supported)
- ▶ Switches or directors used
- ▶ Redundancy
- ▶ Additional storage space requirements

Note: ERM requires a dedicated *switched fabric* connection per controller to be attached to host port 4 on both A and B controllers.

This dedication is required at both ends of the ERM solution.

4.2.7 FC/SATA intermix

For some specific DS5000 storage subsystems, this intermix premium feature supports the concurrent attachment of Fibre Channel and SATA storage expansion enclosures for a single DS4000 or DS5000 controller configuration. With this premium feature, you can create and manage distinct arrays or logical drives that are built from either Fibre Channel disks or SATA disks in a DS4000 or DS5000 storage subsystem, and allocate the drives to the appropriate applications in the attached host servers.

When planning your type of drives, consider these three classes of storage:

- ▶ *Online (or primary) storage* is the storage for applications that require immediate access to data, such as databases and frequently accessed user data. Primary storage holds business-critical information and data with the highest value and importance. This storage requires high performance and high availability technologies such as Fibre Channel technology.
- ▶ *Near-line (or secondary) storage* is used for applications that do not require immediate access but still require the performance, availability, and flexibility of disk storage. It can also be used to cache online storage to reduce the time required for data backups. Secondary storage represents a large percentage of a company's data and is an ideal fit for Serial Advanced Technology Attachment (SATA).
- ▶ *Offline (archival) storage* is used for backup or long-term storage. For this type of storage, tape remains the most economical solution.

Now that we have identified that SATA technology is best suited to near-line storage usage, we have the following considerations regarding its use in the same storage subsystem as online storage:

- ▶ *Performance:* Instead of having a storage subsystem for online storage and another for near-line storage, both types of storage can be combined with one storage subsystem, and use higher performing controllers.
- ▶ *Scalability:* The total SATA disk capacity is far greater than the Fibre Channel offering. The new EV-DDM and E-DDM drives are 1000 GB, as compared to the largest FC drive being 450 GB.

- ▶ *Cost:* Consider an existing Fibre Channel environment where you want to implement SATA technology. This consideration is further enhanced with Fibre Channel and SATA drive intermixing within the same enclosures. The cost considerations of intermixing versus implementing a separate SATA storage subsystem are:
 - Implementation costs of a new SATA storage subsystem versus intermixing with the existing Fibre Channel enclosures.
 - SAN infrastructure costs for more ports or hardware.
 - Administration costs (effort) depend on the number of controllers managed. Intermixing eliminates controller pairs by leveraging existing controllers.

After these factors have been considered and the criteria for the environment has been identified, the choice to intermix or not can be determined.

4.3 Planning your host attachment method

In this section, we review the different attachments methods available for the IBM Midrange System Storage DS5000 storage subsystem so you can evaluate the requirements needed in each case. The options to attach your hosts to the DS5000 storage subsystems are:

- ▶ Fibre Channel
 - Direct Attach
 - Storage Area Network attached
- ▶ iSCSI attach

4.3.1 Fibre Channel: SAN or Direct Attach

When planning the setup of a Storage Area Network (SAN), you want the solution to answer your current requirements and fulfill your future needs.

First, the SAN should be able to accommodate a growing demand in storage (it is estimated that storage needs double every two years). Second, the SAN must be able to keep up with the constant evolution of technology and resulting hardware upgrades and improvements. It is estimated that a storage installation needs to be upgraded every 2 to 3 years.

Ensuring compatibility among different pieces of equipment is crucial when planning the installation. The important question is what device works with what, and also who has tested and certified that equipment.

When designing a SAN storage solution, it is a best practice to complete the following steps:

1. Produce a statement outlining the solution requirements that can be used to determine the type of configuration you need. It should also be used to cross-check that the solution design delivers the basic requirements. The statement should have easily defined bullet points covering the requirements, for example:
 - New installation or upgrade of an existing infrastructure.
 - Host Bus Adapter (HBA) selection (1 Gb, 2 Gb, 4 Gb, or 8 Gb speed).
 - HBA driver type selection: SCSIPort or StorPort.
 - Multipath Driver selection (RDAC, MPIO, or SDDPCM).
 - Types of applications accessing the SAN. (Are the applications I/O intensive or high throughput?)

- Required capacity.
 - Required redundancy levels.
 - Type of data protection needed.
 - Current data growth patterns for your environment.
 - Is the current data more read or write based?
 - Backup strategies in use (Network, LAN-free, or Server-less).
 - Premium features required (FC/SATA Intermix, Partitioning, FlashCopy, Volume Copy, Enhanced Remote Mirroring, and Drive Encryption).
 - Number of host connections required.
 - Types of hosts and operating systems that will connect to the SAN.
 - What zoning is required?
 - Distances between equipment and sites (if there is there more than one site).
2. Produce a hardware checklist. It should cover such items that require you to:
 - Make an inventory of existing hardware infrastructure. Ensure that any existing hardware meets the minimum hardware requirements and is supported with the DS5000 storage subsystem.
 - Make a complete list of the planned hardware requirements.
 - Ensure that you have enough rack space for future capacity expansion.
 - Ensure that the power and environmental requirements are met.
 - Ensure that your existing Fibre Channel switches and cables are properly configured.
 3. Produce a software checklist to cover all the required items that need to be certified and checked. It should include such items that require you to:
 - Ensure that the existing versions of firmware and storage management software are up to date.
 - Ensure host operating systems are supported with the DS5000 storage subsystem. Check the IBM System Storage DS5000 interoperability matrix available at this Web site for more information:
<http://www.ibm.com/servers/storage/disk>

This list is not exhaustive, but the creation of the statements is an exercise in information gathering and planning; it gives you a greater understanding of what your needs are in your current environment and creates a clear picture of your future requirements. The goal should be quality rather than quantity of information.

Use this chapter as a reference to help you gather the information for the statements.

Understanding the applications is another important consideration in planning for your DS5000 storage subsystem setup. Applications can typically either be I/O intensive (high number of I/O per second (IOPS)), or characterized by large I/O requests (that is, high throughput or MBps).

- ▶ Typical examples of high IOPS environments are Online Transaction Processing (OLTP), databases, and Microsoft Exchange servers. These have random writes and fewer reads.
- ▶ Typical examples of high throughput applications are data mining, imaging, and backup storage pools. These have large sequential reads and writes.

By understanding your data and applications, you can also better understand growth patterns. Being able to estimate an expected growth is vital for the capacity planning of your DS5000 storage subsystem installation. Clearly indicate the expected growth in the planning documents: The actual patterns might differ from the plan according to the dynamics of your environment.

Selecting the right DS5000 storage subsystem model for your current and perceived future needs is one of the most crucial decisions you will make. The good side, however, is that the DS5000 offers scalability and expansion flexibility. Premium features can be purchased and installed at a later time to add functionality to the storage subsystem.

In any case, it is perhaps better to purchase a higher model than one strictly dictated by your current requirements and expectations. This will allow for greater performance and scalability as your needs and data grow.

SAN zoning

Zoning is an important part of integrating a DS5000 storage subsystem into a SAN. When done correctly, it can eliminate many common problems.

Depending on your host operating system and the device driver you are planning to use, you have to zone your SAN differently. In AIX or Solaris using old RDAC, only one path per controller is supported, so you have to create a zone for the connection between the host bus adapter (HBA1) and controller A and a separate zone that contains the other HBA2 to controller B.

If using a multipath driver as MPIO, or MPP in Linux, then you can connect and zone your SAN to have multiple paths to the same controller, and then create additional zones for access to other resources. This isolates each zone down to its simplest form.

Best practices:

- ▶ Connect your storage and create zones by considering the number of paths allowed to the same controller by your host operative system multipath driver.
- ▶ Make sure to have only one initiator per zone.
- ▶ Disk and tape should be on separate HBAs, which follows the best practice for zoning.

Enhanced Remote Mirroring considerations

When using Enhanced Remote Mirroring (ERM), you must create two additional zones:

- ▶ The first zone contains the ERM source DS5000 controller A and ERM target DS5000 controller A.
- ▶ The second zone contains the ERM source DS5000 controller B and ERM target DS5000 controller B.

4.3.2 Planning for iSCSI attachment

In addition to using Fibre Channel as the interface connection method, iSCSI host interfaces allows servers with specialized hardware iSCSI cards, or with software running over regular Ethernet cards, run the iSCSI protocol as the connection method to attach the DS5000 storage subsystems. As with FC, before beginning an iSCSI deployment, plan in advance and understand and document the network topology that will be used.

Hardware initiators

As of the writing of this book, only the following hardware initiators are supported:

- ▶ IBM iSCSI Server TX Adapter
- ▶ IBM iSCSI Server SX Adapter
- ▶ QLogic iSCSI Single-Port PCIe HBA for IBM System x
- ▶ QLogic iSCSI Dual-Port PCIe HBA for IBM System x

All of the hardware initiators that are supported use the same base firmware code and the SANsurfer management application. Before you install and configure these adapters, make sure that you have installed the latest management application and the latest firmware code.

Software initiators

The native MPIO that is provided with the Microsoft iSCSI Software Initiator (Version 2.03 or later) is not supported. You must use the DSM that is provided with the Storage Manager software to make sure that failover and I/O access are correct. If the native MPIO from the Microsoft iSCSI Software Initiator is used, it might cause unwanted effects.

Partitioning

As described in 4.1.4, “Storage partitioning” on page 116, because the FC and iSCSI protocols provide radically different latency and throughput capabilities, and this mixture within a server might cause failover driver conflict, performance degradation, or potential data loss, you have to plan ahead by using the following recommendations:

- ▶ Define separated partitions for FC-based hosts from iSCSI-based hosts, avoid mixing them in same storage partition.
- ▶ A single host should not be configured for both iSCSI connections and FC connections to the storage system.

In order to define hosts and partitions in iSCSI, remember that you have to plan for iSCSI addresses for the host ports, and use the iSCSI qualified name (IQN) of the host you want to map.

Network settings

Unlike traditional Fibre Channel, which requires special cabling and SAN switches, iSCSI can be run over an existing network infrastructure. However, in complex network environments, and in order to protect the integrity of the data in your DS5000 storage subsystem, and its continuous access, we suggest that, whenever possible, to try to isolate the iSCSI traffic in a dedicated network. The iSCSI multipathing architecture provides failover to the alternate controller in the event of an outage situation. Also with MPIO, IBM provides DSM, which also offers load-balancing algorithms.

For better redundancy, you can increase the availability of your connections using redundant networks, so a failure in one does not interrupt the remaining redundant connection.

Aside from the basic iSCSI connectivity parameters, such as IP address per target Ethernet port and associated iSCSI qualified names, you have to plan in advance several optional configuration parameters, including enablement of jumbo frames, configuration of a VLAN, and setting a specific Ethernet priority:

- ▶ Jumbo frames are created when the MTU is adjusted above 1500 bytes per frame, and they are set by port. The frame sizes supported are of 1500 and 9000 bytes. When using jumbo frames, ensure that all of the devices on your iSCSI network, including switches, initiators, and targets, are configured to use the same maximum jumbo frame size.

- ▶ VLAN: As previously mentioned, we suggest, for performance and availability reasons, having separated networks for redundant interfaces. If it is not possible to segregate an iSCSI storage system onto a physically separate LAN, with the IBM DS5000 storage subsystems that are connected by iSCSI, you can use VLANs to maximize the potential performance.
- ▶ Ethernet priority: Ethernet priority, sometimes referred to as quality of service or class of service, is supported in the DS5000 series of storage systems. You can set the Ethernet priority of the target iSCSI interfaces to increase the class of service received within the network itself.

Security

Unlike FC SANs, Ethernet networks can be more open, so in order to provide additional security, you can configure the following additional authentication protocols on the DS5000 storage subsystems:

- ▶ The Internet Storage Name Service (iSNS) protocol allows for automated discovery, management, and configuration of iSCSI devices on a TCP/IP network. iSNS servers offer additional security services through explicitly defined initiator-to-target mappings and simplified asset locators, similar to that provided by DNS and WINS for IP address lookup facilities
- ▶ Challenge Handshake Authentication Protocol (CHAP) provides an additional security layer within the iSCSI SAN on the IBM Storage System DS5000 series.

4.4 Host support and multipathing

The intent of this section is to list the most popular supported operating system platforms and topologies used on the DS5000 storage subsystem. For a complete and up-to-date list, see the DS5000 series interoperability matrix, available at the following address:

<http://www-01.ibm.com/systems/support/storage/config/ssic/index.jsp>

In 4.4.4, “Multipathing” on page 133, we discuss the available multipathing drivers as well as their supported operating systems.

4.4.1 Supported server platforms

The supported server platforms are:

- ▶ IBM System x
- ▶ IBM System p
- ▶ IBM System i
- ▶ IBM Power Systems
- ▶ IBM BladeCenter
- ▶ HP Compatible servers
- ▶ AMD and Intel Compatible servers
- ▶ Sun Compatible servers

4.4.2 Supported operating systems

The following are supported operating systems at the time of publication:

- ▶ Microsoft Windows Server 2003 SP2
- ▶ Microsoft Windows Server 2008 SP1
- ▶ Red Hat Enterprise Linux 4.6
- ▶ Red Hat Enterprise Linux 5.1
- ▶ Novell SUSE SLES 9 SP4
- ▶ Novell SUSE SLES 10 SP1
- ▶ VMware ESX 3.5 U2 (restrictions apply)
- ▶ AIX 5L V5.1, V5.2, and V5.3, and V6.1
- ▶ IBM i
 - Via VIOS guest client: DS4700, DS5020, DS5100, and DS5300
 - Native IBM i attach: DS5100 and DS5300
- ▶ VIOS 1.5.2 and 2.1.1
- ▶ HP-UX 11iv2 (11.23) and HP-UX 11iv3 (11.31)

Note: Make sure to check the System Storage Interoperation Center found at the following Web site for the current supported operating systems and clustering environments:

<http://www.ibm.com/systems/support/storage/config/ssic/index.jsp>

4.4.3 Clustering support

Supported clustering services are:

- ▶ IBM PowerHA™ (formerly HACMP)
- ▶ Microsoft Cluster Services
- ▶ Novell Cluster Services
- ▶ Linux RHEL Native Cluster
- ▶ Sun Cluster
- ▶ Steeleye: Lifekeeper 5.3, 6.0 U1
- ▶ HP MC/Service Guard 11.18.00
- ▶ Symantec VERITAS Cluster
- ▶ VMware Cluster Service

4.4.4 Multipathing

IBM offers different multipath drivers that you can use with your DS5000 storage subsystem. Only one of these drivers is required. Each driver offers multipath support, I/O load balancing, and automatic path failover.

The multipath driver is a proxy for the real, physical-level HBA drivers. Each multipath driver hides from the application the fact that there are redundant connections by creating a virtual device. The application uses this virtual device, and the multipath driver will connect the application to the correct physical path.

When you create a logical drive, you assign one of the two active controllers to own the logical drive (called *preferred controller ownership*, as described in 4.1.2, “Logical drives and controller ownership” on page 114) and to control the I/O between the logical drive and the application host along the I/O path. The preferred controller normally receives the I/O requests from the logical drive. If a problem along the data path (such as a component failure) causes an I/O to fail, the multipath driver issues the I/O to the alternate controller.

A multipath device driver is not required when the host operating system has its own mechanism to handle multiple I/O paths, but if not, you have to install the supplied multipath device driver, even if your server does not have multiple paths.

Table 4-5 shows the DS5000 storage subsystem interoperability matrix of multipathing drivers.

Table 4-5 DS5000 storage subsystem interoperability matrix of multipathing drivers

	Windows 2003/2008	RHEL	SLES	AIX 5L V5.3	AIX V6.1	HP-UX 11.23	HP-UX 11.31
DS5000 RDAC		X	X				
Win MPIO	X						
AIX MPIO				X ^a	X ^a		
AIX SDDPCM				X	X		
LVM (HP-UX)						X	X
SDD for HP-UX						X	

a. Included in the operating system

Load Balancing Policy

When you have multiple paths to the DS5000 storage subsystem, the driver can select between different options about how to manage the traffic between the host and storage. The options available for the different operative systems are shown in Table 4-6.

Table 4-6 Load Balancing Policies per OS

Operating system	Multipath driver	Load Balancing Policy
AIX	MPIO	Round robin Selectable path priority
Red Hat Enterprise Linux 4 Update 7	MPP	Round robin Least queue depth
Solaris	MPxIO	Round robin
SUSE Linux Enterprise 9 Service Pack 4	MPP	Round robin Least queue depth
Windows	MPIO	Round robin Least queue depth Least path weight

One path per controller: Failover

Failover is not a balancing algorithm, but we include it here because it is an option that is presented with the actual balancing policies in some operating systems, such as Windows. The driver use failover only when there is one path to each controller of your system. Using failover, one of the device paths is active, the one corresponding to the controller who owns the logical volume, and the other is in a standby state. This is the default for RDAC drivers (AIX and Solaris), because it only supports two paths, that is, one to each controller. Make sure to zone accordingly when working with RDAC so that there is not more than one path per controller.

Multiple paths per controller

The multi-path driver transparently balances I/O workload without administrator intervention, across multiple paths to the same controller, but not across both controllers. To make use of the feature, you need to cable and zone your SAN properly. The load balancing policy uses one of three following algorithms:

- ▶ Round robin with subset

The round robin with subset I/O load balance policy routes I/O requests, in rotation, to each available data path to the controller that owns the volumes. This policy treats all paths to the controller that owns the volume equally for I/O activity. Paths to the secondary controller are ignored until ownership changes. The basic assumption for the round robin policy is that the data paths are equal. With mixed host support, the data paths might have different bandwidths or different data transfer speeds.

- ▶ Least queue depth with subset

The least queue depth with subset policy is also known as the least I/Os or least requests policy. This policy routes the next I/O request to a data path that has the least outstanding I/O requests queued. For this policy, an I/O request is simply a command in the queue. The type of command or the number of blocks that are associated with the command are not considered. The least queue depth with subset policy treats large block requests and small block requests equally. The data path selected is one of the paths in the path group of the controller that owns the volume.

- ▶ Least path weight with subset

The least path weight with subset policy assigns a weight factor to each data path to a volume. An I/O request is routed to the path with the lowest weight value to the controller that owns the volume. If more than one data path to the volume has the same weight value, the round-robin with subset path selection policy is used to route I/O requests between the paths with the same weight value.

4.4.5 Microsoft Windows MPIO

This section discusses the available Windows multipath options.

MPIO is a Driver Development Kit (DDK) from Microsoft for developing code that manages multipath devices. It contains a core set of binary drivers, which are installed with the DS5000 Device Specific Module (DSM) to provide a transparent system architecture that relies on Microsoft Plug and Play to provide LUN multipath functionality while maintaining compatibility with existing Microsoft Windows device driver stacks.

Note: The MPIO Driver is included in the Storage Manager software package for Windows and supports Microsoft Windows 2003 and 2008 on 32-bit and x64 systems. In Windows 2008, MPIO is already part of the operating system.

The MPIO driver performs the following tasks:

- ▶ Detects and claims the physical disk devices presented by the DS5000 storage subsystems based on vendor/product ID strings and manages the logical paths to the physical devices.
- ▶ Presents a single instance of each LUN to the rest of the Windows operating system.
- ▶ Provides an optional interface through WMI for use by user-mode applications.
- ▶ Relies on the vendor's (IBM) customized Device Specific Module (DSM) for information about the behavior of storage subsystem devices for the following items:
 - I/O routing information
 - Conditions requiring a request to be retried, failed, failed over, or failed back (for example, vendor-unique errors)
 - Handles miscellaneous functions, such as release/reservation commands
- ▶ Multiple Device Specific Modules (DSMs) for different disk storage subsystems can be installed in the same host server.

See 4.4.10, "Auto Logical Drive Transfer (ADT) feature" on page 138 for more details about multipathing and failover considerations.

For compatibility information, always see the current DS5000 interoperability matrix or the readme file for your HBA driver. The matrix can be found at:

<http://www-01.ibm.com/systems/support/storage/config/ssic/index.jsp>

4.4.6 AIX MPIO

With Multiple Path I/O (MPIO), a device can be uniquely detected through one or more physical connections, or paths. A path-control module (PCM) provides the path management functions.

An MPIO-capable device driver can control more than one type of target device. A PCM can support one or more specific devices. Therefore, one device driver can be interfaced to multiple PCMs that control the I/O across the paths to each of the target devices.

The AIX PCM has a health-check capability that can be used to do the following actions:

- ▶ Check the paths and determine which paths are currently usable for sending I/O.
- ▶ Enable a path that was previously marked failed because of a temporary path fault (for example, when a cable to a device was removed and then reconnected).
- ▶ Check currently unused paths that will be used if a failover occurred (for example, when the algorithm attribute value is failover, the health check can test the alternate paths).

MPIO is part of the AIX operating system and does not need to be installed separately. You can find more information about MPIO in *IBM Midrange System Storage Implementation and Best Practices Guide*, SG24-6363.

4.4.7 AIX Subsystem Device Driver Path Control Module (SDDPCM)

SDDPCM is a loadable path control module for supported storage devices to supply path management functions and error recovery algorithms. When the supported storage devices are configured as Multipath I/O (MPIO) devices, SDDPCM is loaded as part of the AIX MPIO Fibre Channel Protocol (FCP) device driver during the configuration. The AIX MPIO-capable device driver with the supported storage devices SDDPCM module enhances data availability and I/O load balancing.

4.4.8 Linux: RHEL/SLES RDAC

The Redundant Disk Array Controller (RDAC), also known as Multi-Path Proxy (MPP), is the recommended multipathing driver for Linux based operating systems like Red Hat Enterprise Linux (RHEL) or SUSE Linux Enterprise Server (SLES).

The current RDAC driver implementation performs the following tasks:

- ▶ Detects and claims the physical devices (LUNs) presented from the DS5000 storage subsystems (*hides* them) based on vendor/product ID strings and manages all of the paths to the physical devices.
- ▶ Presents a single instance of each LUN to the rest of the Linux operating system components.
- ▶ Manages all of the plug and play interactions.
- ▶ Provides I/O routing information.
- ▶ Identifies conditions requiring a request to be retried, failed, or failed over.
- ▶ Automatically fails over the LUNs to their alternate controller when detecting problems in sending I/Os to the LUNs in their preferred controller, and fails back the LUNs to their preferred controller when detecting the problems in the preferred path fixed.
- ▶ Handles miscellaneous functions, such as persistent reservation translation.
- ▶ Uses a round robin (load distribution or load balancing) model.

RDAC is implemented between the HBA driver and the operating system disk driver, operating as a low-level filter driver. It has the following advantages:

- ▶ It is much more transparent to the OS and applications.
- ▶ I/O controls at the HBA driver level are not as tightly coupled to the OS as those at the disk driver level. Consequently, it is easier to implement I/O control functionality in the MPP-based RDAC driver for routing functions.

As the driver is positioned at the HBA level (see Figure 4-9), it has access to the SCSI command and sense data interface of the HBA driver and therefore can make more informed decisions about what to do in the case of path failures.

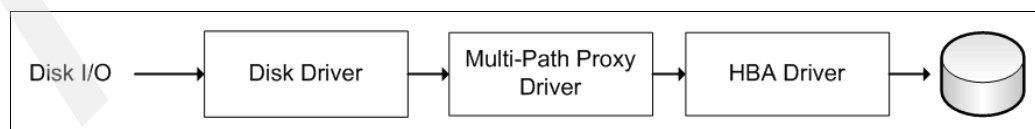


Figure 4-9 Linux RDAC/MPP driver

Note: In Linux, RDAC cannot be installed with the Installation Wizard. If you need RDAC, you have to download and install it separately.

4.4.9 Virtualization

With the growth and popularity of storage area networks, storage environments are getting more and more complex. Storage virtualization reduces the complexity and costs of managing storage environments and optimizes storage utilization in a heterogeneous environment.

The IBM Storage Area Network Volume Controller (SVC) and IBM TotalStorage Productivity Center products address some of these needs.

IBM System Storage SAN Volume Controller overview

The IBM System Storage SAN Volume Controller is a scalable hardware and software solution to allow aggregation of storage from different disk servers. It provides storage virtualization and a consistent view of storage across a Storage Area Network (SAN).

The SAN Volume Controller provides in-band storage virtualization by creating a pool of managed disks from attached back-end disk storage systems. These managed disks are then mapped to a set of virtual disks for use by various host systems.

In conjunction with the DS5000 storage subsystem family, the SAN Volume Controller (SVC) can increase the storage copy services functionality and also the flexibility of SAN-based storage. The SVC is very flexible in its use. It can manage all host storage requirements or just part of them. A DS5000 storage system can still be used to allocate storage to hosts or use the SVC. This is dependant upon various needs and requirements.

The SVC also offers an alternative to FlashCopy, VolumeCopy, and Enhanced Remote Mirroring for disaster recovery, high availability, and maintenance. If you plan to use SVC with a DS5000 storage subsystem, then these premium features will not be required.

The SVC might also reduce the requirement for additional partitions. The SVC only consumes one storage partition. If you plan to use the SVC for all of your hosts, then a storage partition upgrade might not be required.

SVC is licensed by the capacity that is being managed. This capacity also includes the capacity used by the copy services.

For detailed information about SVC implementation, see *Implementing the IBM System Storage SAN Volume Controller V4.3*, SG24-6423.

For more information about Storage Virtualization, see the IBM TotalStorage Virtualization Web site at the following address:

<http://www.ibm.com/servers/storage/software/virtualization/index.html>

For more information about SAN Volume Controller, see its home page at the following address:

<http://www.ibm.com/servers/storage/software/virtualization/svc/index.html>

4.4.10 Auto Logical Drive Transfer (ADT) feature

In a DS5000 storage subsystem, you can provide redundant I/O paths with the host systems. There are two different components that provide this redundancy:

- ▶ A multipath driver
- ▶ Auto Logical Drive Transfer (ADT)

Auto-Logical Drive Transfer (ADT) is a built-in feature of controller firmware that allows logical drive-level failover rather than controller-level failover. Depending on the attached host, the feature will allow auto-volume transfer based on the characteristics of the operative system and the driver. You can also enable or disable this feature.

Note: ADT is not a failover driver. ADT provides storage systems with the flexibility to work with some third-party failover software.

The two modes are:

- ▶ ADT-disabled failover

The multi-path software will send a SCSI Mode Select command to cause a change in volume ownership before using the alternate path. All logical drives on the preferred controller are transferred to the alternate controller. This is the configuration setting for Microsoft Windows, IBM AIX, and Linux (when using the RDAC or SDD driver and non-failover Fibre Channel HBA driver) systems. When ADT is disabled, the I/O data path is still protected as long as you use a multi-path driver. After the I/O data path problem is corrected, the preferred controller does not automatically reestablish ownership of the logical drive. You must open a storage management window, select **Redistribute Logical Drives** from the Advanced menu, and perform the Redistribute Logical Drives task.

Figure 4-10 shows the ADT-disabled failover mode phases.

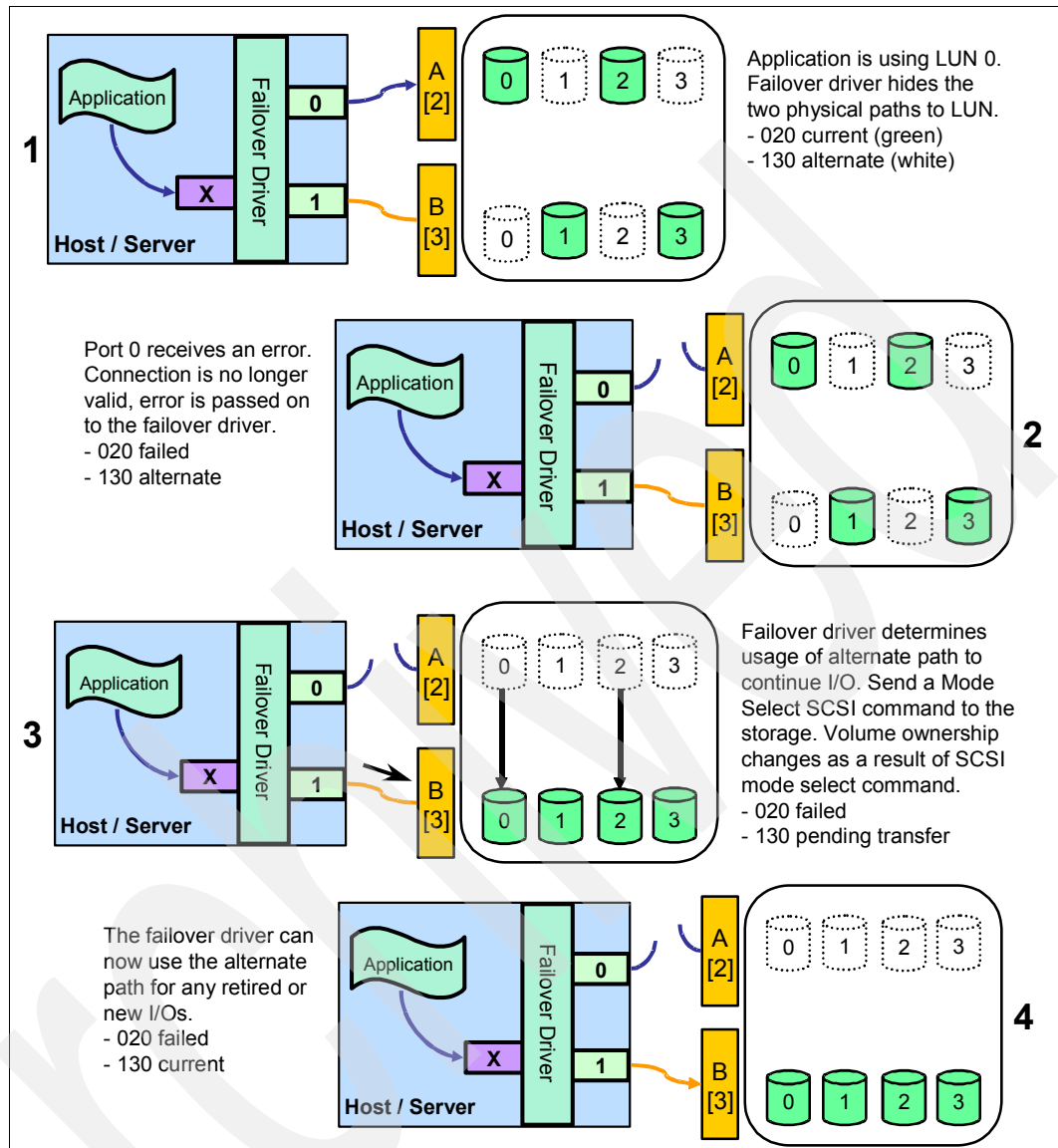


Figure 4-10 ADT-disabled mode path failover

Note: In ADT-disabled mode, you are required to issue a redistribution command manually to balance the LUNs across the controllers.

► ADT-enabled failover

The multi-path driver starts using the alternate path by sending the I/O down the path it chooses and lets the ADT react. This is the normal configuration setting for Novell NetWare, Linux (when using the FC HBA failover driver instead of RDAC), and HP-UX systems. After the I/O data path problem is corrected, the preferred controller automatically reestablishes ownership of the logical drive as soon as the multipath driver detects that the path is normal again.

Figure 4-11 shows the phases of failover in the AVT-enabled case.

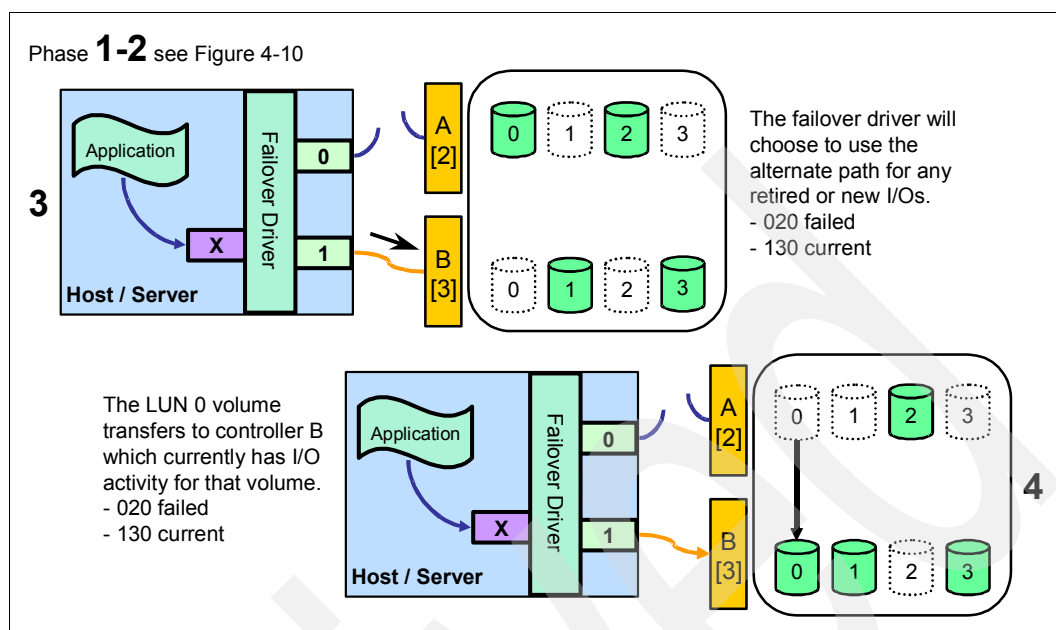


Figure 4-11 AVT-enabled mode path failover

Note: In AVT mode, RDAC automatically redistributes the LUNs to their preferred path after the failed path is operational again.

4.5 Operating system restrictions

In your planning, you should also consider the restrictions of your operating system. In this section, we cover the maximum file system size ordered by operating system and their commonly used file systems, the maximum number of LUNs that can be assigned to one host, and the maximum logical volume size supported by various operating systems.

4.5.1 Maximum file system size

Table 4-7 shows the maximum capacity a logical drive can have and be formatted with a specific file system.

Table 4-7 Maximum file system size supported by various operating systems

Operating system	File system	Size
Linux	ext2	32 TB
	ext3	32 TB
AIX	JFS2	64 PB
Windows	NTFS	16 EB
	FAT32	8 TB
HP-UX 11i v2	HFS	128 GB
	JFS 3.5	2 TB / 32 TB ^a

a. VxVM 3.5 only and OnlineJFS license for greater than 2 TB file systems (only HP-UX 11iV2 B. 11.23)

4.5.2 Maximum number of LUNs per host

The maximum number of LUNs that can be assigned to a host depends on the host operating system type and mapping limit within a storage partition on the DS5000 storage subsystem. The DS5000 storage subsystem supports 256 LUNs per single storage partition. You can find more information about storage partitioning in 4.1.4, “Storage partitioning” on page 116. Table 4-8 lists the operating system limits as well as the limits of the DS5000 storage subsystem.

Table 4-8 Maximum number of LUNs per host

Host operating system type	Maximum number of LUNs per host (theoretical OS limit)	Maximum number of LUNs per host (DS5000 storage subsystem limit)
AIX	64000	256
HP-UX	16000	256
Linux	256	256
Windows	256	256
ESX Server 3.5 Update 2	256	256

Note: Keep in mind that the total number of LUNs on a DS5000 storage subsystem cannot exceed 2048.

4.6 Storage Manager software

The IBM System Storage DS Storage Manager software (also referred to as Storage Manager or SM) is used to configure arrays and logical drives, assign logical drives into storage partitions, replace and rebuild failed disk drives, expand the size of arrays and logical drives, and convert from one RAID level to another. Storage Manager also allows the user to perform troubleshooting and management tasks, such as checking the status of the storage subsystem components, updating the firmware of controllers, and similar tasks.

Advanced functions, such as FlashCopy, VolumeCopy, and Enhanced Remote Mirroring, are also configured using Storage Manager.

The IBM System Storage DS Storage Manager software is packaged as two separate groups of software, as shown in Figure 4-12.

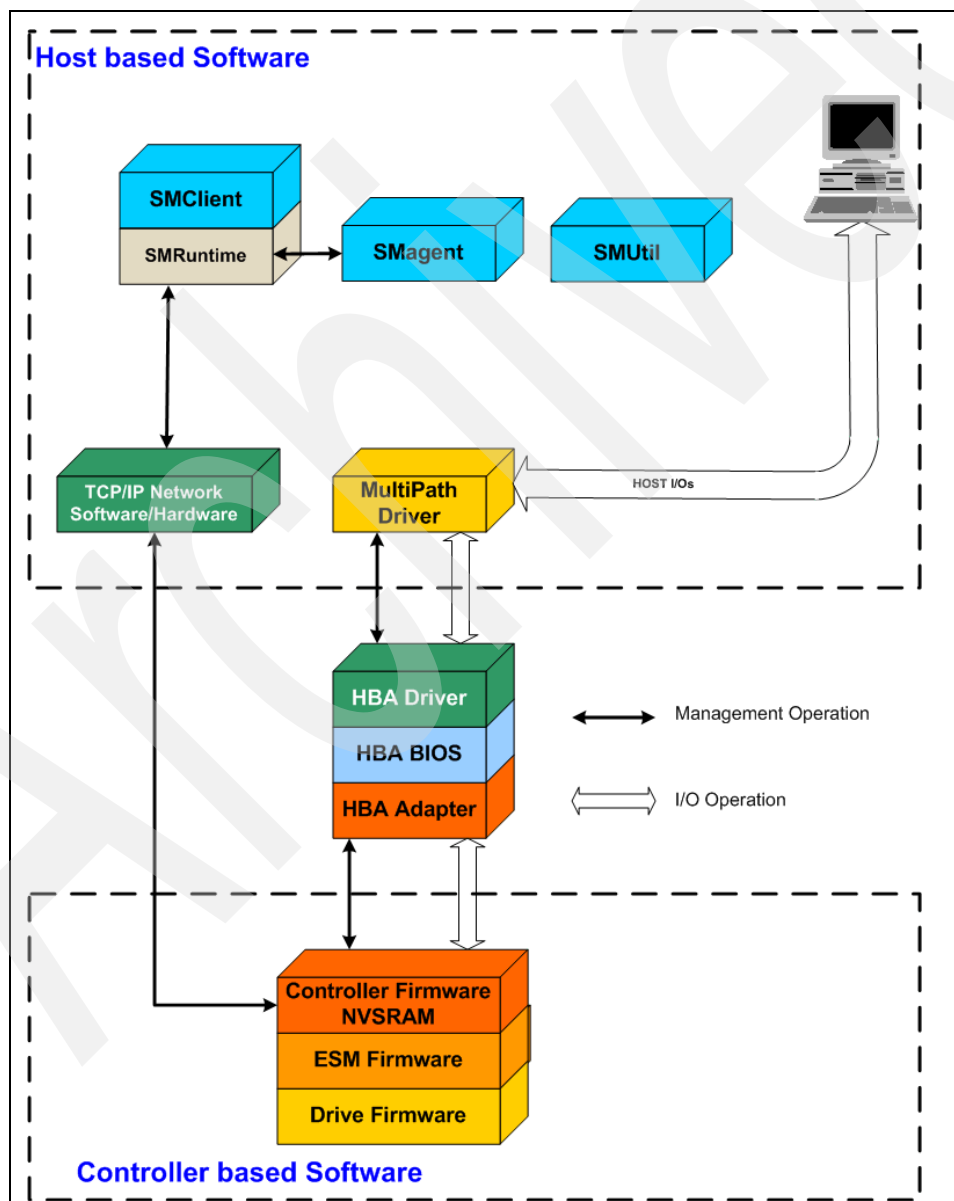


Figure 4-12 Storage Manager software components

The *host-based software* consists of all the software components that must be installed on a host system using the storage subsystem or a management station controlling the server. The *controller-based software* includes everything required for the components of the storage subsystem to be managed through a remotely LAN/SAN attached host server or PC.

Note: Not all software components are available for all platforms.

Host-based software

The host-based software includes the following software components:

▶ **Storage Manager Client (SMclient):**

The SMclient component provides a graphical user interface (GUI) for managing storage systems through the Ethernet network or from the host server. It has two main components:

- **Enterprise Management:** The client can manage multiple storage subsystems in a storage domain, enabling the user to add, remove, and monitor each storage subsystem.
- **Subsystem Management:** The client can manage individual storage subsystems.

The SMclient provides an interface for storage management, based on information supplied by the storage system controllers. When the SMclient is installed on a management station, commands are sent to the storage subsystem controllers. The controller firmware contains the necessary logic to carry out the storage management requests. The controller is responsible for validating and executing the commands and providing the status and configuration information that is sent back to the SMclient.

The SMclient is bundled with an Event Monitor that can run in the background and send alert notifications when critical events occur. The Event Monitor service handles notification functions (e-mail and SNMP traps) and can monitor storage systems whenever the Enterprise Management window is not open.

The command-line interface (SMcli) is also packaged with the SMclient and provides command-line access to perform all management functions.

See 4.6.2, “Storage Manager client” on page 148 for additional details.

▶ **Storage Manager Runtime (SMruntime):**

The SMruntime is a Java runtime environment that is required for the SMclient to function. It is not available on every platform as a separate package, but in those cases, it has been bundled into the SMclient package.

▶ **Storage Manager Agent (SMagent):**

The SMagent package is an optional component that allows in-band management of the DS5000 storage subsystem. This agent allows management of the storage subsystem using the same path as the I/O requests coming from the system. The host agent software receives requests from a management station that is connected to a host server through a network connection and passes the requests to the storage system controllers through the Fibre Channel I/O path.

The host agent, along with the network connection on the host server, provides an in-band host agent type network management connection to the storage system instead of the out-of-band direct network management connection through the individual Ethernet connections on each controller. The management station can communicate with a storage system through any host server that has host agent management software installed. The host agent receives requests from the management station through the network connection to the host server and sends them to the controllers in the storage system through the Fibre Channel or iSCSI paths.

- ▶ **Storage Manager Utilities (SMutil):**

SMutil can be used to register and map new logical drives to the operating system and to verify mapping. It can be installed on all UNIX-based host operating systems, including AIX, HP-UX, Solaris, and Linux host servers that are attached to a storage subsystem through Fibre Channel.

The Storage Manager Utilities package contains a number of command-line tools. See 4.6.4, “Storage Manager utilities” on page 152 for more details.

Controller-based software

The controller-based software consists of:

- ▶ **DS5000 storage subsystem controller firmware and NVSRAM**

The controller firmware and NVSRAM are always installed as a pair and provide the “brains” of the DS5000 storage subsystem. All commands from the SMclient and SMcli come through this code to cause actions to be performed on the storage devices. The controller firmware can be thought of as the operating system and application code for the storage subsystem, and the NVSRAM can be thought of as the configuration of the application.

- ▶ **Environmental Service Module (SMesm software)**

The ESM canister is a component in a storage expansion enclosure that monitors the environmental condition of the components in that enclosure. The ESM software is required for automatic ESM firmware synchronization. The ESM firmware controls the interface between the controller and the drives.

- ▶ **DS5000 storage subsystem drive firmware**

The drive firmware is the software that tells the Fibre Channel (FC) drives how to behave on the FC loop.

4.6.1 Storage subsystem management methods

The storage management software provides two methods for managing storage systems:

- ▶ Host agent (in-band) management method
- ▶ Direct (out-of-band) management method

Depending on specific storage system configurations and host systems, either or both methods can be employed. The management method selected will determine which software components need to be installed.

Host agent (in-band) management method

When the host agent (in-band) management method is used, the storage systems are managed through the Fibre Channel or iSCSI I/O path from the host server. This requires installation of the Storage Manager agent (SMagent) package. The management information can either be processed on the host server or passed to the management station through the network connection, as shown in Figure 4-13.

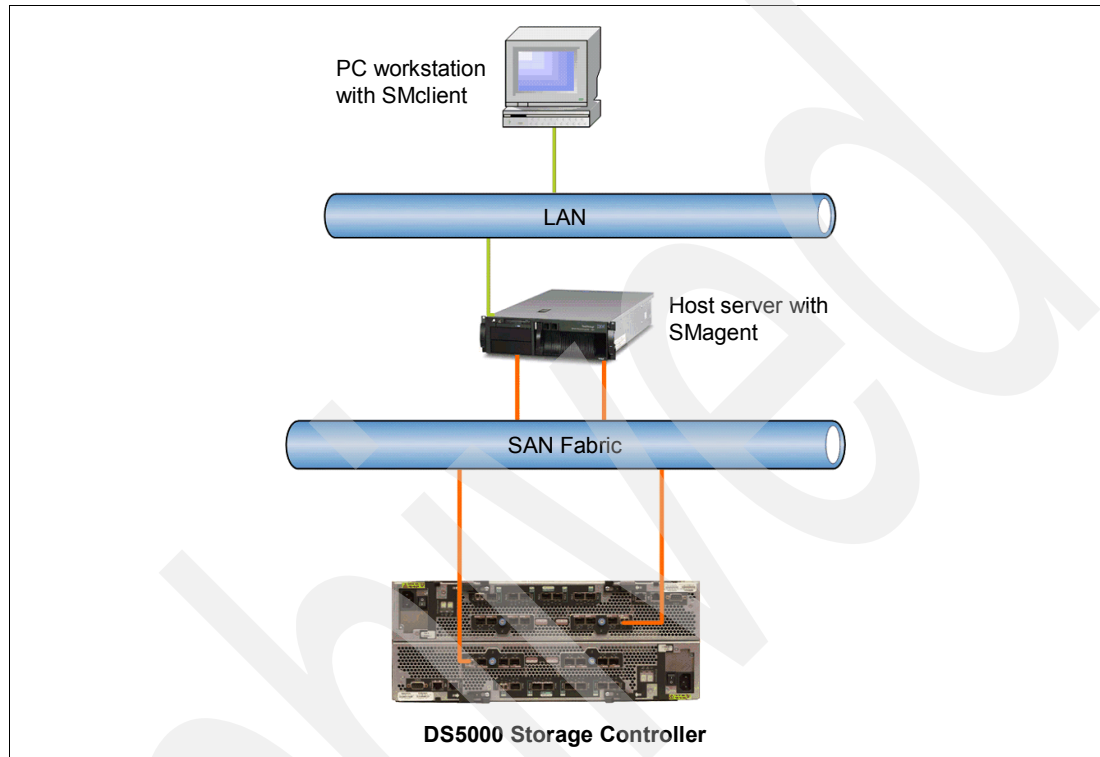


Figure 4-13 In-band management

Managing storage systems through the host agent has the following advantages:

- ▶ Ethernet cables do not need to be run to the controllers' management ports.
- ▶ A host name or IP address must only be specified for the host instead of for the individual controllers in a storage system. Storage systems that are attached to the host can be automatically discovered.

Managing storage systems through the host agent has the following disadvantages:

- ▶ The host agent requires a special logical drive, called the *access logical drive*, to communicate with the controllers in the storage system. Therefore, the host server is limited to one less logical drive than the maximum number that is allowed by the operating system and the host adapter that is being used. Not all operating systems support the *access logical drive*. In-band management is not supported on these systems.
- ▶ If the connection through the Fibre Channel or iSCSI path is lost between the host and the server, the server cannot be managed or monitored.

Important: If a host already has the maximum number of logical drives configured, either use the direct management method or give up a logical drive for use as the access logical drive.

Direct (out-of-band) management method

When the direct (out-of-band) management method is used, storage systems are managed directly over the network through a TCP/IP Ethernet connection to each controller. To manage the storage system through the Ethernet connections, the IP address and host name for each controller must be defined or the DS5000 storage subsystem must be set to use DHCP/BOOTP settings. The controllers must be attached to a Local Area Network (LAN) through Ethernet, as shown in Figure 4-14.

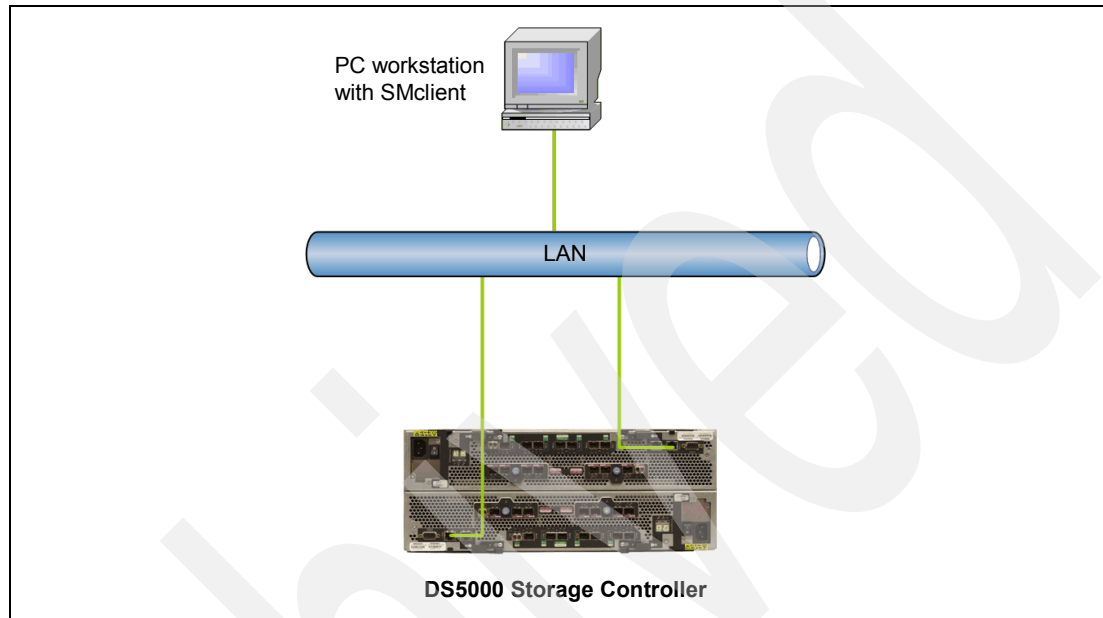


Figure 4-14 Out-of-band management

Managing storage systems using the direct (out-of-band) management method has the following advantages:

- ▶ The Ethernet connections to the controllers enable a management station running SMclient to manage storage systems that are connected to a host running any of the operating systems that are supported by the current level of Storage Manager.
- ▶ An access logical drive is not needed to communicate with the controllers. The maximum number of logical drives that are supported by the host operating system can be configured and used.
- ▶ The storage system can be managed when there are problems with the Fibre Channel links.
- ▶ Security is enhanced when management LANs/VLANs and more advanced solutions, such as VPN, can be used to manage the system remotely.
- ▶ More DS5000 storage subsystems in the network can be managed through one Storage Manager interface.

Managing storage systems using the direct (out-of-band) management method has the following disadvantages:

- ▶ Two Ethernet cables are required to connect the storage system controllers to a network.
- ▶ When adding devices, the IP address or host name for each controller must be provided.
- ▶ DHCP/BOOTP server and network preparation tasks are required. This can be avoided by assigning static IP addresses to the controller, or by using the default IP address.

To assign static IP addresses, see 4.8.3, “Configuring IP addresses of the controllers” on page 161.

Tip: To manage storage subsystems through a firewall, configure the firewall to open port 2463 for TCP and UDP data.

4.6.2 Storage Manager client

This section continues our overview of the Storage Manager client, reviewing the different screens and information available.

We know already that the Storage Manager client can be used for either in-band or out-of-band management of the storage subsystem. In-band management uses the Fibre Channel network to communicate with the IBM System Storage DS5000 storage subsystem, and out-of-band management uses the TCP/IP network. On host platforms that support both methods, it is possible to use them on the same machine if you have a TCP/IP connection and also a Fibre Channel connection to the DS5000 storage subsystem.

When you install SMclient and SMagent on a stand-alone host and manage the storage subsystem through the Fibre Channel I/O path (rather than through the Ethernet network), you should still install the TCP/IP software on the host and assign an IP address to it. This way permits other workstations in the network with the SMclient installed to manage the DS5000 storage subsystem by connecting through the SMagent.

Next, we review the different host systems where you can install the Storage Manager client and the different windows presented.

Supported host systems

The Storage Manager client is a Java-based GUI utility that is available for various operating systems. For up-to-date information about support for specific DS5000 models and operating systems, check the DS5000 storage subsystem support Web site at the following address:

<http://www.ibm.com/servers/storage/support/disk>

The Storage Manager client uses two main windows that gives control over the storage system:

- ▶ The Enterprise Management window
- ▶ The Subsystem Management window

The Enterprise Management window

The Enterprise Management window is the first window that opens when the storage management software is started. The Enterprise Management window has two different tabs:

- ▶ Devices
- ▶ Setup

Enterprise Management Setup tab

The Setup tab of the Enterprise Management window contains different functions that you can use to perform the initial setup tasks for your storage subsystem, as shown in Figure 4-15.



Figure 4-15 Enterprise Management Setup tab

From here you can add your storage subsystem to the management software, start your configuration, and perform other various tasks, such as providing proper names to each subsystem, configuring alert notification destinations, and updating firmware or selecting a particular storage subsystem (or subsystem) to be managed.

Enterprise Management Devices tab

The Devices view of the Enterprise Management window presents the different storage subsystems that the client can access either directly or through the host agents. If a certain storage subsystem can be accessed in both ways, and possibly through several host agents, it will be listed several times in the Enterprise Management window.

The name, status, and management type (through Ethernet or through host agent) are shown for each listed storage subsystem, as shown in Figure 4-16. In our illustration, we assume out-of-band management from a machine that only has a TCP/IP connection to the DS5000 storage subsystem.

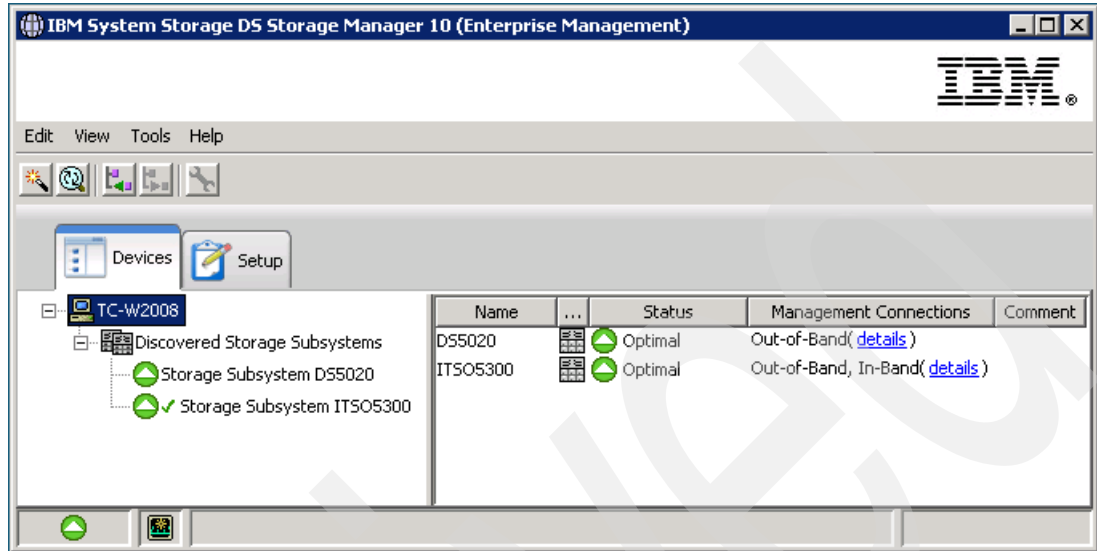


Figure 4-16 Enterprise Management Device view

Note: Although a single storage subsystem can be listed several times in the left pane when it is accessed by various host agents or directly attached, it only appears once in the right pane.

The Subsystem Management window

Once a system to be managed is selected in the Enterprise Management window, clicking it opens the Subsystem Management window for that particular system. In Figure 4-17, there is a sample window showing the summary view of the Subsystem Management window. Note the multiple tabs that are available, each presenting a different view of the subsystem. This multiple tabs view is available with firmware version 10.50 and later. If you are managing a storage subsystem with an earlier firmware version, you will see only two tabs in this view.

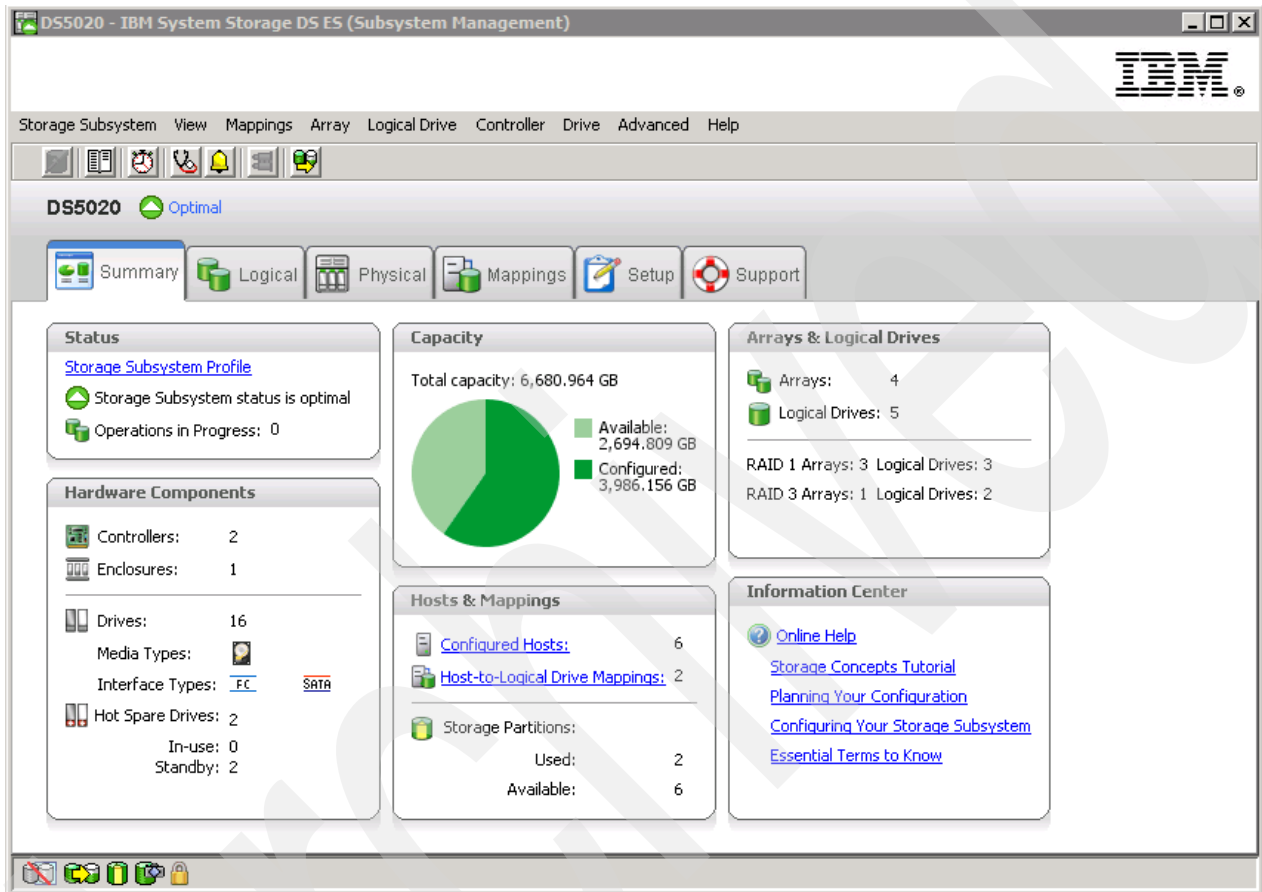


Figure 4-17 Subsystem Management Summary view

We cover the different options available in the Subsystem Management windows later in 4.7.1, "Installing Storage Manager in a Windows Server 2008 host" on page 153.

4.6.3 Event Monitor service

The Event Monitor service handles notification functions (e-mail and SNMP traps) and monitors storage systems whenever the Enterprise Management window is not open.

The Event Monitor is a separate program bundled with the Storage Manager client software (the Event Monitor cannot be installed without the client.) The Event Monitor can only be installed on a management station or host server connected to the storage systems. For continuous monitoring, the Event Monitor must be installed on a host server that runs 24 hours a day. Once installed, the Event Monitor runs in the background and checks for possible critical problems. If it detects a problem, it notifies a remote system through e-mail, Simple Network Management Protocol (SNMP), or both.

In order to use SNMP notification, you need an SNMP machine listening for the notifications sent by the Event Monitor.

For additional information about how to set up the Event Monitor, see 4.9.7, “Monitoring and alerting” on page 223.

4.6.4 Storage Manager utilities

Storage Manager comes with command-line utilities that are installed separately from the other components. These vary by operating system type, but generally include the following utilities:

- ▶ **hot_add:** This utility is used to scan for new disks available to the operating system after they are defined and mapped in Storage Manager. This is especially useful for operating systems that normally have to be re-booted.
- ▶ **SMdevices:** This utility lists all logical drives available to the host, including target ID and logical drive name (as defined in the Storage Manager). This is useful if there are several logical drives of the same size defined for a given host, because it is able to identify which logical drive is which before mounting and formatting them under the operating system.
- ▶ **SMrepassist:** This utility is a host-based utility for Windows platforms that performs certain functions needed to make the subsystem hardware FlashCopy work smoothly. It is run against a specific drive or mount point and causes the buffers to be flushed to disk.
- ▶ **mppUtil:** This utility is used in conjunction with the Linux driver to configure and troubleshoot the driver. This utility can display information about the RDAC driver itself, assist in debugging errors, and can manually initiate one of the RDAC driver’s scan tasks. This utility is installed with the RDAC driver itself.

For information about how to install Storage Manager utilities, see 4.7, “Installing IBM System Storage DS Storage Manager” on page 152.

4.7 Installing IBM System Storage DS Storage Manager

The IBM System Storage DS Storage Manager consists of a set of client and host server software packages that enable you to manage and connect to the DS5000 storage subsystem.

The IBM System Storage DS5000 storage subsystem comes with a CD that provides the following packages for the various supported operating systems:

- ▶ **Client Software package**
 - **SMruntime software:** Provides the Java runtime environment required to run the SMclient.
 - **SMclient software:** Java-based graphical user interface.
 - **SMesm software:** Provides functions for the Environmental Service Module, ESM, and firmware delivery package.
- ▶ **Host Software package**
 - **SMagent software:** Provides management capabilities through the host I/O path.
 - **SMutil software:** An utility to register and map new logical drives to the operating system.

Note: Before installing your DS5000 storage subsystem hardware and software, be sure to meet all the requirements for your DS5000 model: HBAs, firmware version, SAN, and OS level, as specified at the IBM Support Web site.

Go to the Web page for the System Storage Interoperation Center to get the latest DS5000 storage subsystem compatibility information at the following address:

<http://www-01.ibm.com/systems/support/storage/config/ssic/index.jsp>

4.7.1 Installing Storage Manager in a Windows Server 2008 host

This section covers a guided installation of the Storage Manager V10.60 software in a Windows Server 2008 host environment. You can use it as a reference also for installing on other environments as well.

The host software for Windows includes the following components:

- ▶ SMclient
- ▶ Multipath driver
- ▶ SMagent
- ▶ SMutil

Perform the following steps:

1. Because you are installing new software, including new drivers, you need to log on with administrator rights.
2. Locate and run the installation executable file, either in the appropriate CD-ROM directory, or the file that you have downloaded from the IBM support Web site. Once executed, select your language of choice. After the introduction and copyright statement windows, you are asked to accept the terms of the license agreement. This is required to proceed with the installation.
3. The next step is to select the installation target directory. The default installation path is C:\Program Files\IBM_DS, but you can select another directory.

4. Now you need to select the installation type, as shown in Figure 4-18.



Figure 4-18 *InstallAnywhere: Select Installation Type*

- The installation type you select defines the components that will be installed. For example, if you select Management Station, then the multipath driver and Agent components will not be installed, because they are not required on the management computer. In most cases, you select either the Management Station or Host installation type.

- Because having only these two options can be a limitation, two additional choices are offered: Typical (Full Installation) and Custom. As the name suggests, Typical (Full Installation) installs all components and Custom installation lets you choose the components, as you can see in Figure 4-19.

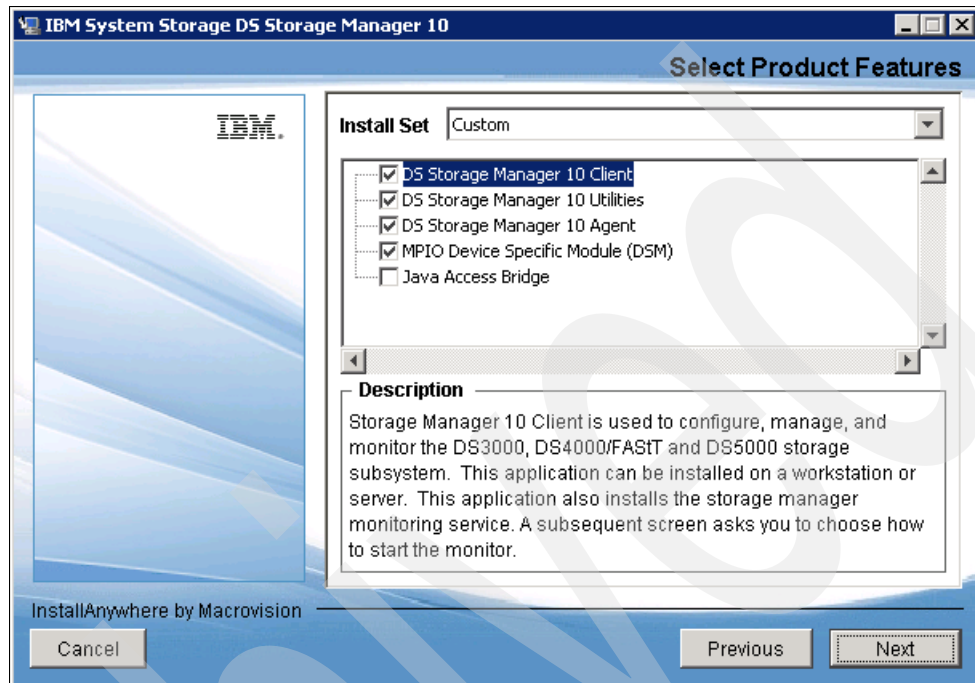


Figure 4-19 InstallAnywhere: Select Storage Manager components

Note: Remember that the SM Failover drivers (MPIO/DSM) are only supported for connection to DS5000 storage subsystems with controller firmware V6.19 and later.

Besides the usual Storage Manager components, you can choose to install Java Access Bridge. This selection enables support for the window reader (like JAWS from Freedom Scientific, Inc.) for blind or visually impaired users.

5. The next installation window asks you whether you want to automatically start the Storage Manager Event Monitor, as shown in Figure 4-20. This depends on your particular management setup. In case there are several management machines, the Event Monitor should only run on one. If you want to use the Event Monitor with SNMP, you have to install the Microsoft SNMP service first, because the Event Monitor uses its functionality.

Note: The Event Monitor needs to be enabled for both the automatic ESM synchronization, as well as the automatic support bundle collection on critical events.

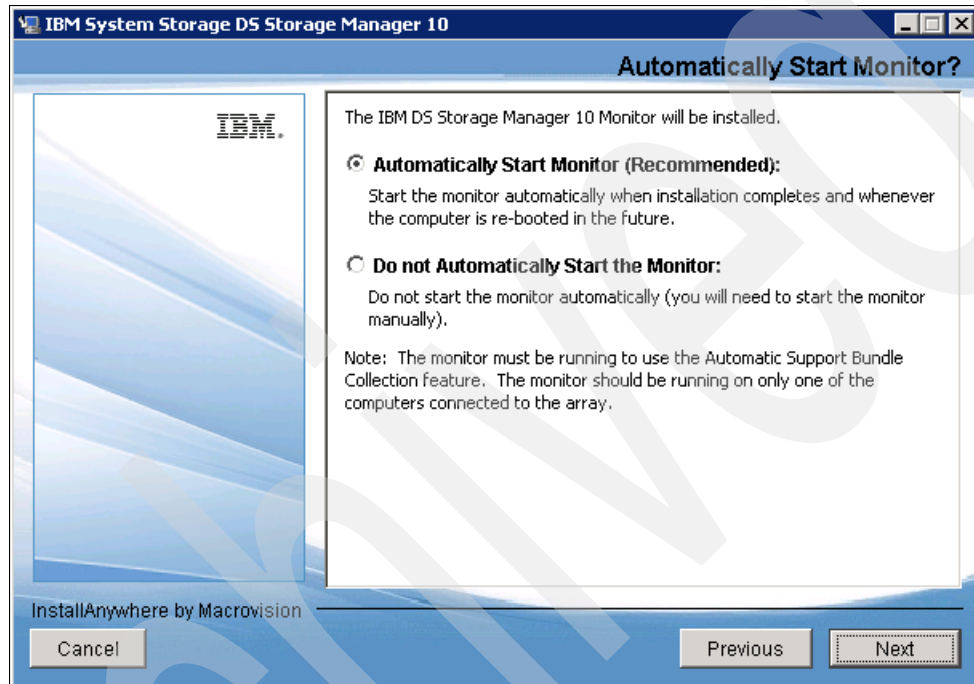


Figure 4-20 InstallAnywhere: Automatically Start Monitor

6. Finally, you are presented with the Pre-Installation Summary window, just to verify that you have selected the correct installation options, as shown in Figure 4-21. Click the **Install** button and the actual installation process starts.



Figure 4-21 *InstallAnywhere: Pre-Installation Summary window*

Verifying the SM installation

This section provides instructions about how to verify that you have installed the SMagent correctly in Windows Server 2008.

To verify the correct installation of the SM, perform the following steps:

1. Look in your Programs folder for a new program entry called IBM DS Storage Manager 10 Client. This is the name of the program created after a successful installation. It also generates a log file with the details of the installation process and options selected, and places it into the installation directory. The file name is `IBM_System_Storage_DS_Storage_Manager_10_InstallLog.log`. In case of problems during the installation, have a look at the file for a possible hint about what could be wrong.
1. Select **Start** → **Administrative Tools** → **Services**. The Services window opens.

2. Scroll through the list of services until you find IBM DS Storage Manager 10 Agent, as shown in Figure 4-22.

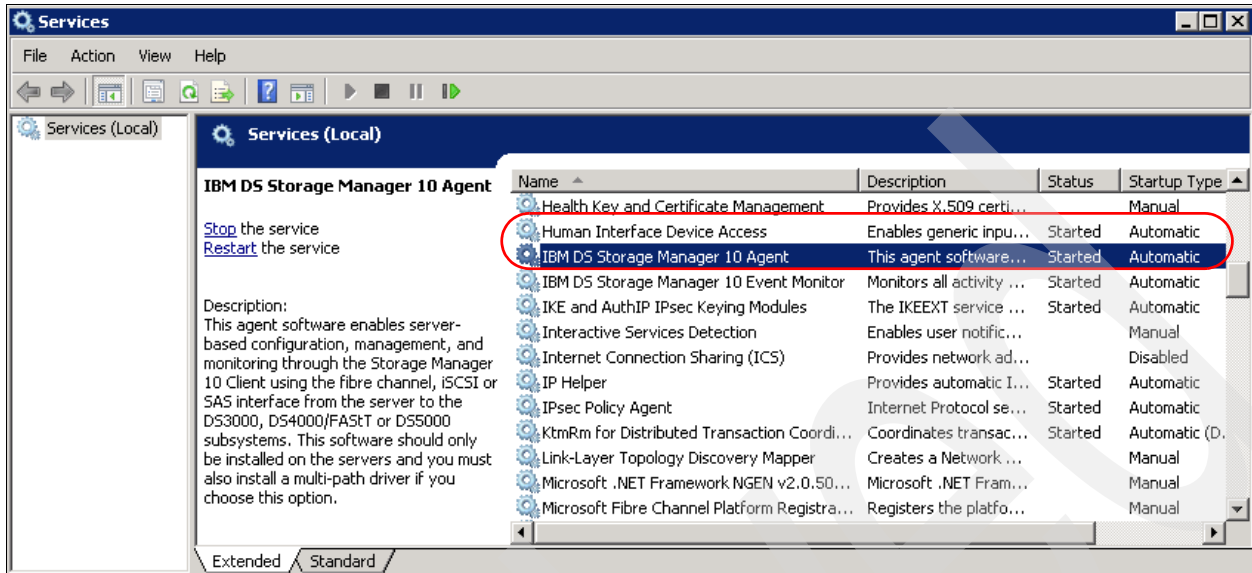


Figure 4-22 Verifying the SMagent installation

3. The installation program creates both the IBM DS Storage Manager 10 Agent and the Event Monitor services and starts both of them by default. If this service is not started, right-click it and select **Start**. Make sure the Startup Type is set to **Automatic**.

If you are installing the host server and do not plan to use the host-agent software to manage one or more storage systems, you can set the Startup Type to **Manual**.

Verifying the SUtil installation

To verify that you have installed the SUtil correctly on Windows operating systems, perform the following steps:

1. Go to the *installation_directory*\Util directory. This is typically be C:\Program Files\IBM_DS\Util.
2. Verify that the directory contains the following files:
 - hot_add.exe
 - SMdevices.bat
 - SMrepassist.exe

Verifying the SM Failover driver (MPIO/DSM) installation

There are different types of MPIO/DSM drivers supported for Windows 2008. Follow these steps to verify the installation:

1. Check the install directory for the SM Failover driver. The default install directory is C:\Program Files\DSMDrivers\ds4dsm.

2. Open the Device Manager in Computer Management. There should be a Multi-Path Support entry under the SCSI and RAID Controller folder, as shown in Figure 4-23.

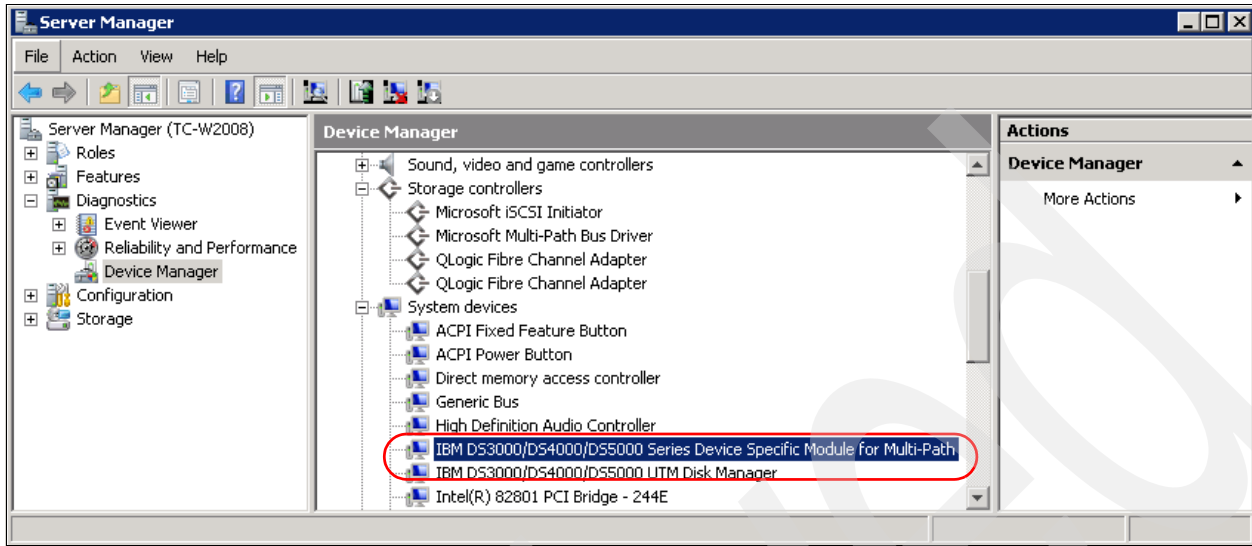


Figure 4-23 SM Failover (MPIO/DSM) Multi-Path Support

4.8 Preparing the DS5000 storage subsystem

This section explains how to configure the DS5000 storage subsystem. This includes:

- ▶ Physical installation
- ▶ Powering on the storage system
- ▶ Configuring IP addresses on the DS5000 storage subsystem
- ▶ Initializing the Storage Manager client
- ▶ Updating the controller microcode
- ▶ Updating the drive and ESM microcode

4.8.1 Physical installation

You should have already completed the physical installation of your DS5000 storage subsystem hardware. See the installation, user's, and maintenance guides for your specific IBM System Storage DS5000 for specific instructions.

Take note of the order in which you install the expansion enclosures. Once the Storage Manager software is installed, it is important that the representation in the software interface matches the current disposition of the expansions in the rack.

Follow the recommendations on cabling the expansion enclosures to the storage subsystem controllers, to allow an optimization of the drive channel paths, and consider any plans for future expansions.

4.8.2 Powering on the storage subsystem

Before powering on your storage subsystem, make sure you have your hardware installed correctly. In particular, check whether your system has a speed switch located on the front of the system. This should be set to 2 GB or 4 GB and match the speed of the drives installed in that enclosure. You might also want to check, if you have multiple old expansion units, that all of them have a different enclosure ID assigned. New expansion enclosures do not require a switch setting for the enclosure ID, because they are automatically assigned at first power on. For specific details, see 3.4, “DS5100 and DS5300 storage subsystems” on page 29.

When turning power on or off, a certain order must be followed on the DS5000 storage subsystem.

The drives attached to the storage subsystem must be available when the storage subsystem is powered up. In other words, when expansion enclosures are attached to a storage subsystem, the expansions must be powered up first.

Important: Always power up the expansion enclosures first. The controllers in the storage subsystem might not recognize the correct configuration if the drives are powered up after the storage subsystem.

The Fibre Channel and Ethernet hubs and switches, if installed, must be powered on before the host to ensure proper loop initialization. When connecting your DS5000 storage subsystem ports to a SAN switch, or directly to a host, match whenever possible the maximum speed of the ports, that is, avoid connecting an 8 Gb SFP to a 2 Gb switch or host bus adapter, for example.

The power up procedure includes the following steps:

1. Turn on the expansion enclosures.
2. After the last enclosure is powered on, wait for all the disks to be ready (steady green LED for each disk).
3. Turn on the SAN or Ethernet switches (if attached).
4. Turn on the DS5000 storage subsystem.
5. Turn on the host application server.

The power-down procedure includes the following steps:

1. Turn off the host application server.
2. Turn off the DS5000 storage subsystem.
3. Turn off hubs/switches (if attached).
4. Turn off the expansion enclosures.

Important: Ensure that your system is in an optimal state before you shut it down. Never turn the power off if any fault light is lit. Be sure to resolve any error conditions before you shut down the system.

4.8.3 Configuring IP addresses of the controllers

By default, the DS5000 storage subsystem will try the BOOTP/DHCP service to request an IP address. If no BOOTP/DHCP service is found in the network, the controllers revert to fixed IP addresses. By default, the fixed addresses are:

- ▶ Controller A: 192.168.128.101
- ▶ Controller B: 192.168.128.102

The DS5000 storage subsystem has two Ethernet ports per controller. Use one Ethernet port for daily management of your DS5000 storage subsystem. Reserve the other port for use by service personnel or for subsystem monitoring hardware that might be available in the future.

The default IP addresses of the additional Ethernet ports are:

- ▶ Controller A: 192.168.129.101
- ▶ Controller B: 192.168.129.102

See the specific Ethernet port locations with their default IP addresses in Figure 4-24.

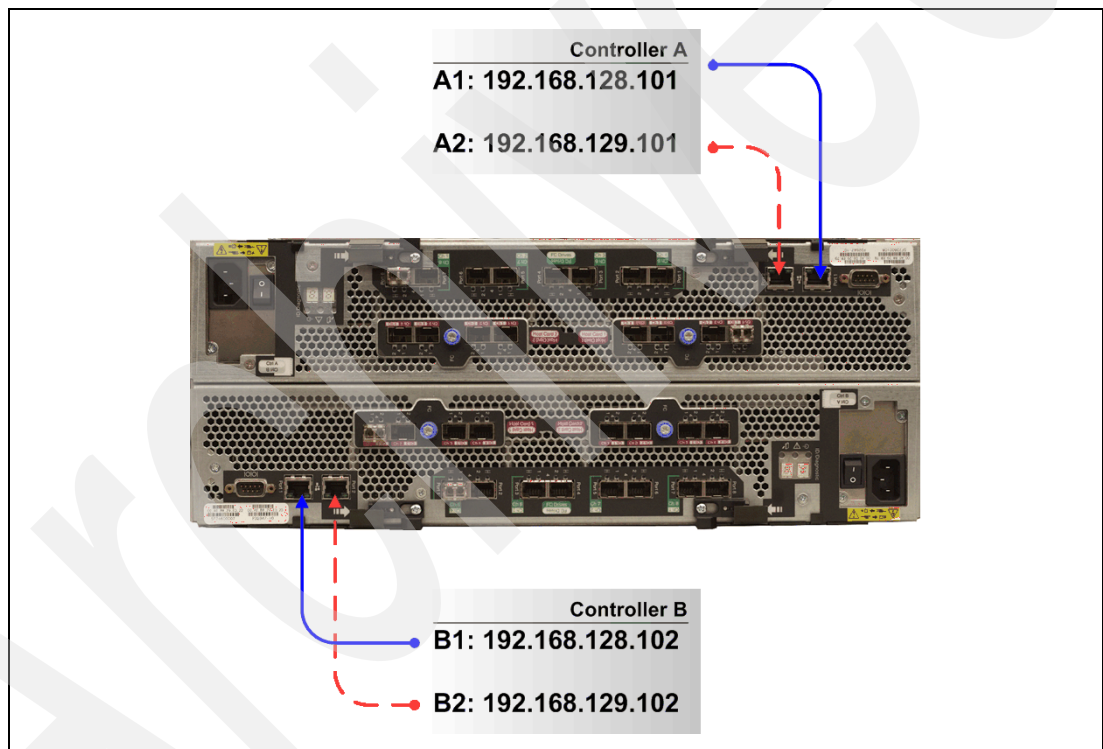


Figure 4-24 DS5000 storage subsystem default IP addresses

Given the importance of the storage subsystems in any environment, we recommend that an IP address be assigned to them rather than relying on a BOOTP/DHCP dynamic address. There are two ways to change the IP addresses of the controllers by changing them with the Storage Manager utility:

- ▶ Using out-of-band management (recommended)
- ▶ Using in-band management

We describe both methods in the following sections.

Changing the IP addresses with the Storage Manager utility

In order to change the default IP addresses, first we must connect to the DS5000 storage subsystem. The procedure described in this section uses the Ethernet management capability, because the in-band management method needs to configure first the hosts HBAs, which requires specific details, depending on the host platform.

Perform the following steps:

1. Connect the Ethernet ports to an isolated Ethernet switch.
2. Set the IP address in your management machine to match the network address of the default IPs on the DS5000 storage subsystem.
3. Connect the management machine to the same Ethernet switch.
4. Use the Storage Manager client interface to change the IP addresses on both controllers:
 - a. Start the Storage Manager client interface.

If this is the first time you are using the Storage Manager software, after it is launched, the program presents a window to start automatic discovery of the attached devices.

- b. Select **OK** to initiate an automatic discovery.

If the DS5000 storage subsystem does not appear after this action, check the network connectivity to both default IP addresses of the storage subsystem using the **ping** command. Consider that each time a network cable is plugged to an Ethernet port of a controller, it detects the linkup and initiates a request for a dynamic address. If did not find one, it assigns the default IP address. If you try to connect to the storage during that time, you receive an error, so we recommend waiting at least 5 minutes after plugging in the cable to attempt either the automatic discovery or to manually add the controller IP addresses.

Tip: Before starting the automatic discovery or adding the DS5000 storage subsystem manually, wait for the controller to finish its boot process and then another 5 minutes after connecting the network cable to allow for the DHCP process to complete.

Alternatively, you can add your storage manually from the Enterprise Management selecting the **Setup** tab, selecting **Add Storage Subsystem**, and then completing the fields with the default IP addresses, as shown in Figure 4-25.

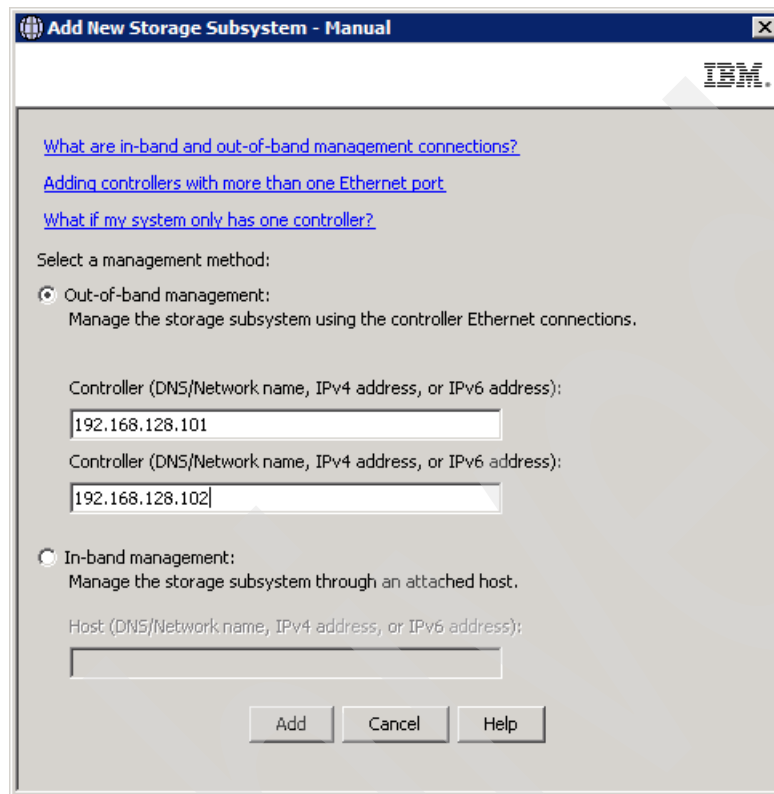


Figure 4-25 Manually adding a storage system

- c. A confirmation message appears and prompts you for additional storage subsystems. Click **No** to finish adding storage subsystems.
- d. After successfully adding a storage subsystem, the Devices tab shows the new system. Note that you can check the IP address of the managed system by selecting **Details** under Management Connections.

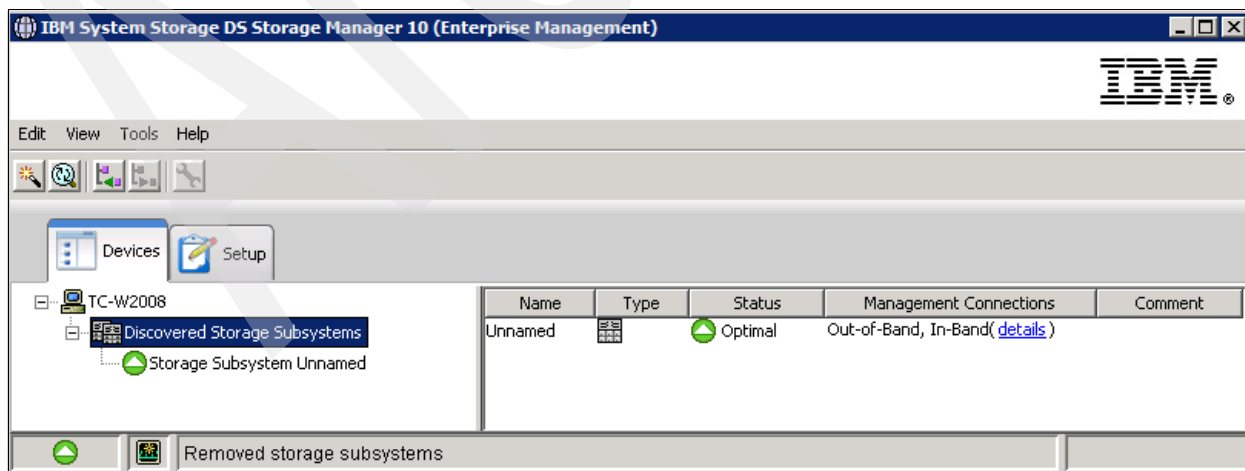


Figure 4-26 Managing a storage system

Select the **Setup** tab to view the Initial Setup window. Right-click the storage subsystem or double-click it to open the Subsystem Management window. You are prompted to set a password for your storage subsystem, as shown in Figure 4-27.

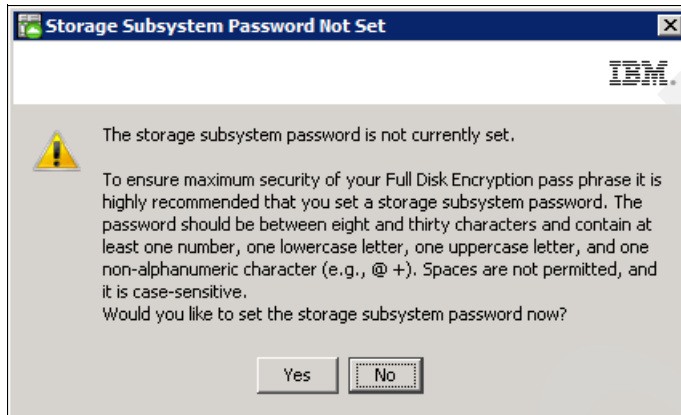


Figure 4-27 Password protection

Note: Make sure to take advantage of the security features of the Storage Manager by setting a password for each of your DS5000 storage subsystems.

- e. The storage subsystem synchronizes with the management station clock when they become out of synchronization. Click **OK** to accept synchronization if prompted to do so.

- f. Now we are connected to the DS5000 storage subsystem, but with the default IP addresses. We need to change the address to your specific ones. In the Subsystem Management window, select the **Setup** tab, and then scroll down to select **Configure Ethernet Management Ports**, as shown in Figure 4-28.

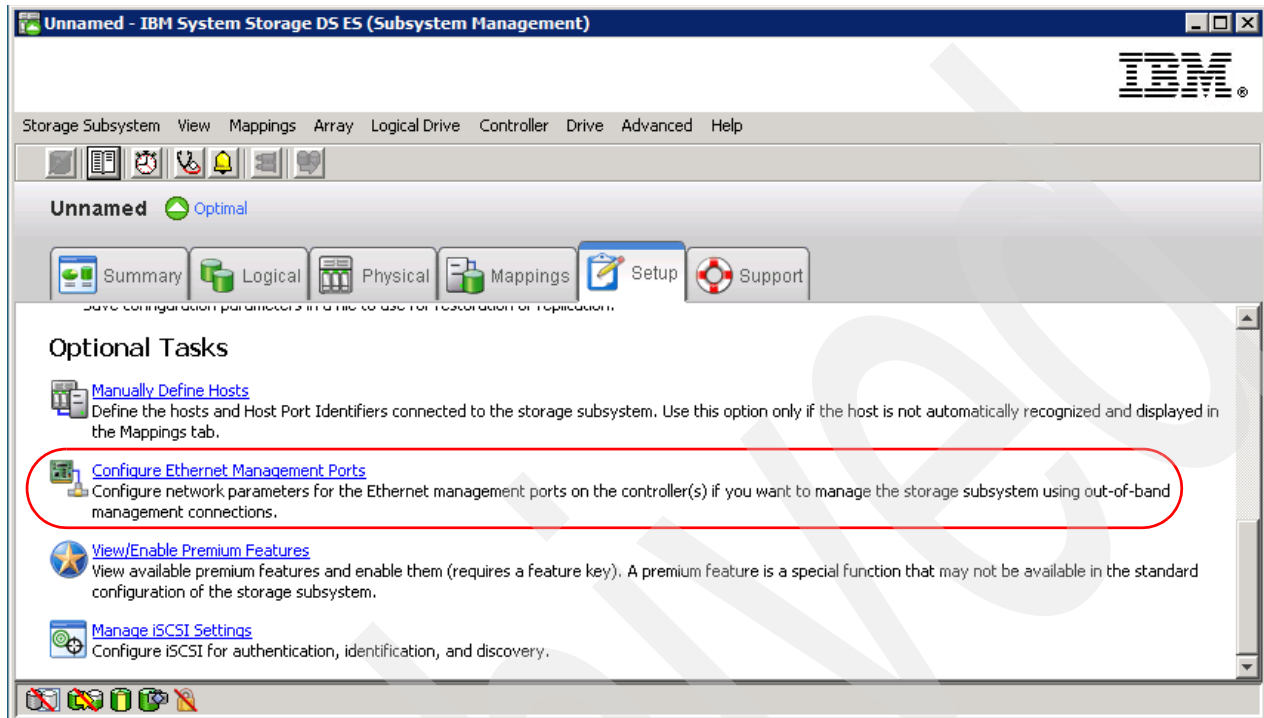


Figure 4-28 Changing the network configuration in the Subsystem Management window

- g. Input the network information and click **OK**, as shown in Figure 4-29. You can configure both controllers from this window by selecting the Ethernet port. Make sure to click **OK** after configuring both of them.

The DS5020 has two Ethernet ports per controller. Make sure to select the port of each controller you already cabled. If you are planning to have the system connected to only one network, select port 1, leaving port 2 for service. See Figure 4-24 to be sure of where to connect the LAN cables.

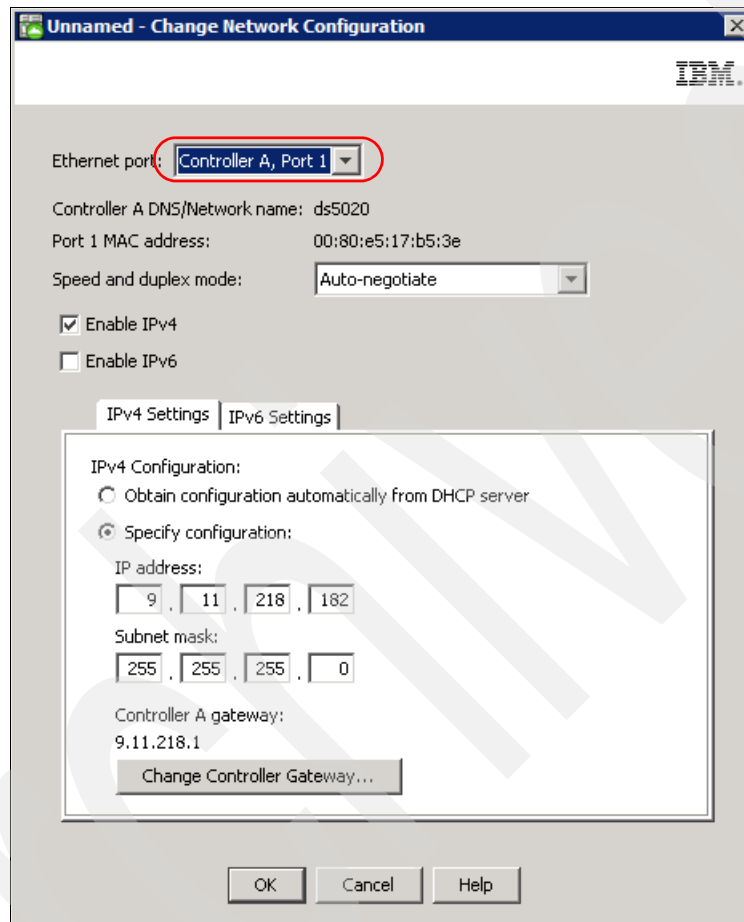


Figure 4-29 DS5000 Controller Network Configuration change

After changing the IP addresses, close the Storage Subsystem window, and remove the added device from the Enterprise Management window. Then add the subsystem, repeating step b, but this time, with your newly assigned IP addresses.

Note: To manage storage subsystems through a firewall using out-of-band management, configure the firewall to open port 2463 to TCP data.

4.8.4 Using and configuring the DS Storage Manager client

We already know by now that the Storage Manager GUI uses two main windows:

- ▶ Enterprise Management window
- ▶ Subsystem Management window

In the following sections, we review the different tasks that you should implement before the configuring the arrays and data storage.

When you start the Storage Manager client, the Enterprise Management window opens, either showing the Device Management section or Setup section, as shown in Figure 4-30.



Figure 4-30 Enterprise Management Initial Setup Tasks window

Adding storage systems

Before you can manage a DS5000 storage subsystem, you have to add it to the Enterprise Management window. The first time the DS Storage Manager is used, it will prompt you to add a storage subsystem, either automatically or manually. Using the Enterprise Management setup view, you can also invoke the Add Storage Subsystem option at a later time as well as any of the other initial setup tasks shown. You can select either the automatic or the manual addition method, as shown in Figure 4-31.

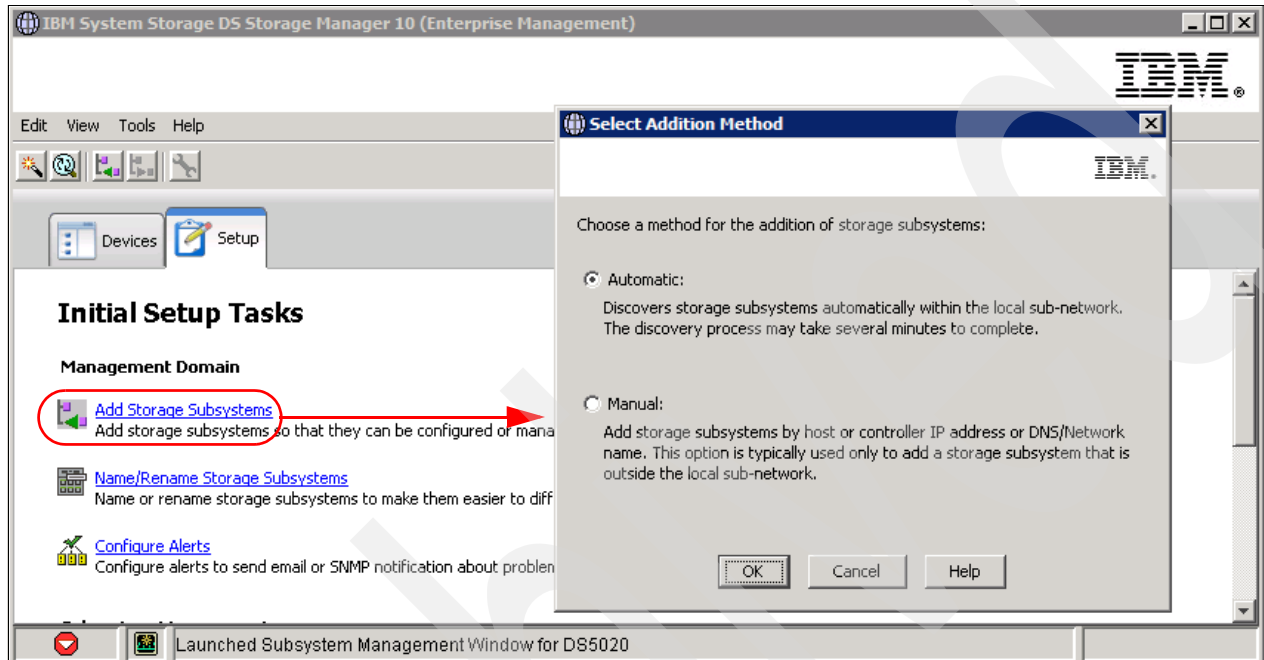


Figure 4-31 Add Storage Subsystems setup task

The automatic discovery process sends out broadcasts through Fibre Channel (if SMagent is installed and started) and the IP network. If it finds directly attached storage systems or hosts running the Storage Manager agent (with an attached storage system), and it adds these storage systems into the Enterprise Management window.

The manual addition method requires that you provide the IP addresses (or host names) of both controllers for out-of-band management. If one of the controllers is not connected or is not reachable, then some of the management functions cannot be performed (except in cases where you manage a single-controller storage subsystem).

If the DS5000 storage subsystem is managed through the FC path (in-band management), and you want to manage it using a workstation with SMclient installed instead of using the in-band management, you have to specify the IP address of the host attached to the storage subsystem. The Enterprise Management window with different DS5000 storage subsystems is shown in Figure 4-32.

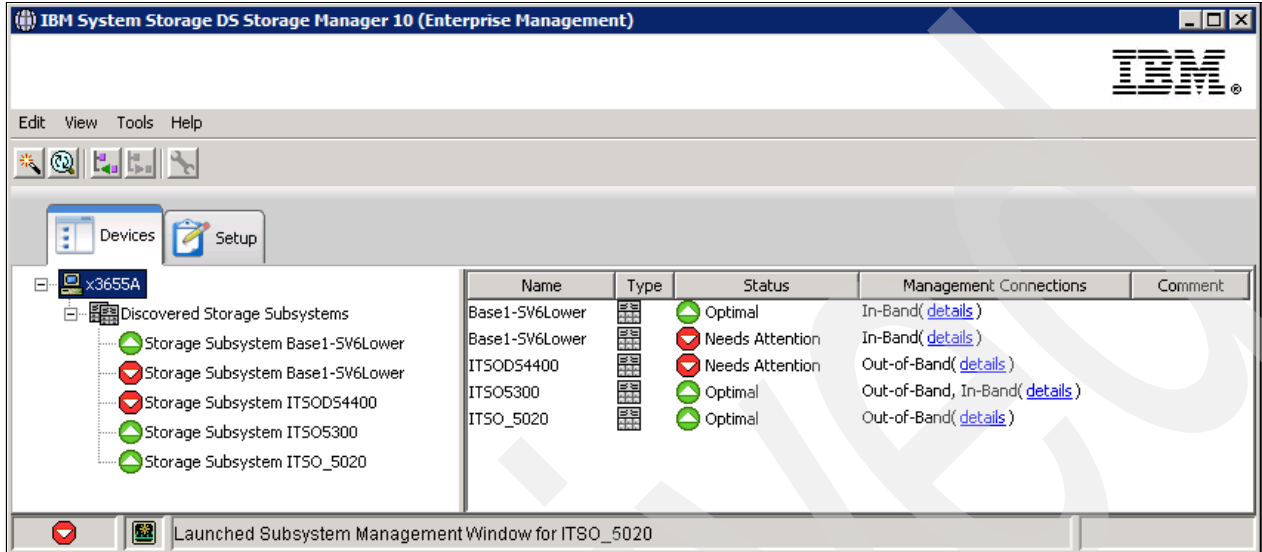


Figure 4-32 Enterprise Management window

You can see all detected DS5000 storage subsystems and how they are managed, either direct (out-of-band) or host-agent attached (in-band), or through both connections. There is also a status column. Usually, the status is Optimal with a green icon next to it, but if there are any problems, the status changes to Needs Attention and a red icon is displayed.

Naming or renaming a storage subsystem

If you installed multiple DS5000 systems, or plan to install more than one, it is important to give each a unique and meaningful name so that you can differentiate it easily from others in the Enterprise Management window.

To accomplish this task, perform these steps:

1. To rename the DS5000 storage subsystem, from the Devices view of the Enterprise Management window, right-click the subsystem to rename and select **Storage Subsystem** → **Rename**.

The other option is to select the **Name/Rename Storage Subsystem** option in the Setup view of the Enterprise Management window, as shown in Figure 4-33.

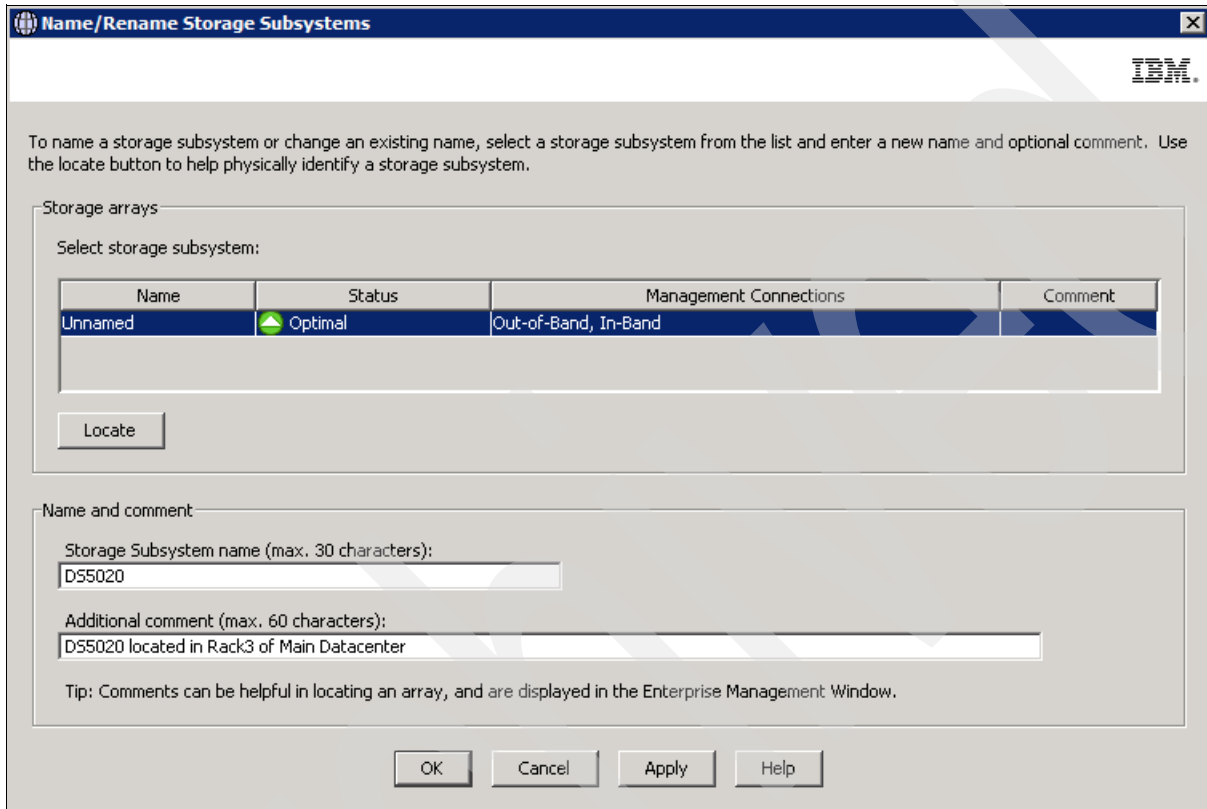


Figure 4-33 Initial Setup Tasks: Name or rename a storage system

2. Type in the name you want to assign to your storage subsystem (there is a 30-character limit). All leading and trailing spaces are deleted from the name. Use a unique, meaningful naming scheme that is easy to understand and remember. You can also assign a comment to facilitate its identification.
3. Click **OK** to finish the name assignment.

Setting the controller clocks

Because the DS5000 storage subsystem stores its own event log, you need to synchronize the controller clocks with all the host systems accessing it. This improves problem determination procedures. If you have not already set the clocks on the storage subsystems, set them now. Be sure that your local system is using the correct time, and then select **Storage Subsystem** → **Set Controller Clock** (Figure 4-34).

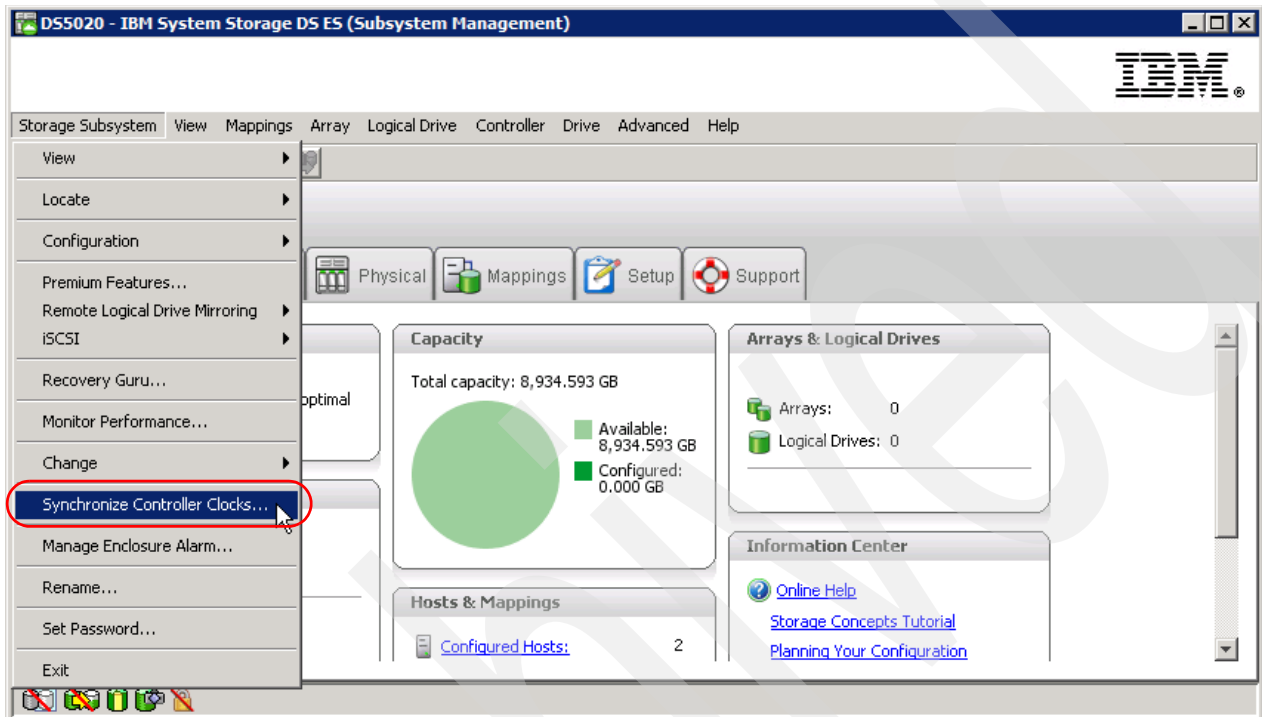


Figure 4-34 Setting the controller clock

Setting a security password

For security reasons, especially if the DS5000 storage subsystem is directly attached to the network, you should set a password. This password is required for all actions on the DS5000 storage subsystem that change or update the configuration in any way.

To set a password, highlight the storage system, right-click it, and select **Storage Subsystem** → **Set Password** (see Figure 4-35). This password is then stored on the DS5000 storage subsystem. It is used if you connect through another SMclient, no matter whether you are using in-band or out-of-band management.



Figure 4-35 Setting the password

Note: Setting a password provides additional protection for a DS5000 storage subsystem in a public LAN. However, you gain higher protection by connecting the DS5000 storage subsystem to a private network reserved for administration.

Be aware that after the password is set, no modification commands will be allowed without it, so do not forget the password.

Configuring alerts

This option allows you to set up the alert structure. If there are any problems with any of the storage subsystems, an e-mail or an SNMP notification can be sent. The Initial Setup Tasks view of the Enterprise Management window allows you to configure alerts for all storage subsystems, for a group of them, or for a single one.

We provide a detailed description of the Configure Alerts option in 4.9.7, “Monitoring and alerting” on page 223.

Setting Enclosure Order

Verify that the enclosures in the right half of the window reflect your actual physical layout. This ensures that you do not perform maintenance activities in the wrong enclosure. If the enclosures are listed in an incorrect order, select **Storage Subsystem** → **Change** → **Enclosure Order**, as shown in Figure 4-36.

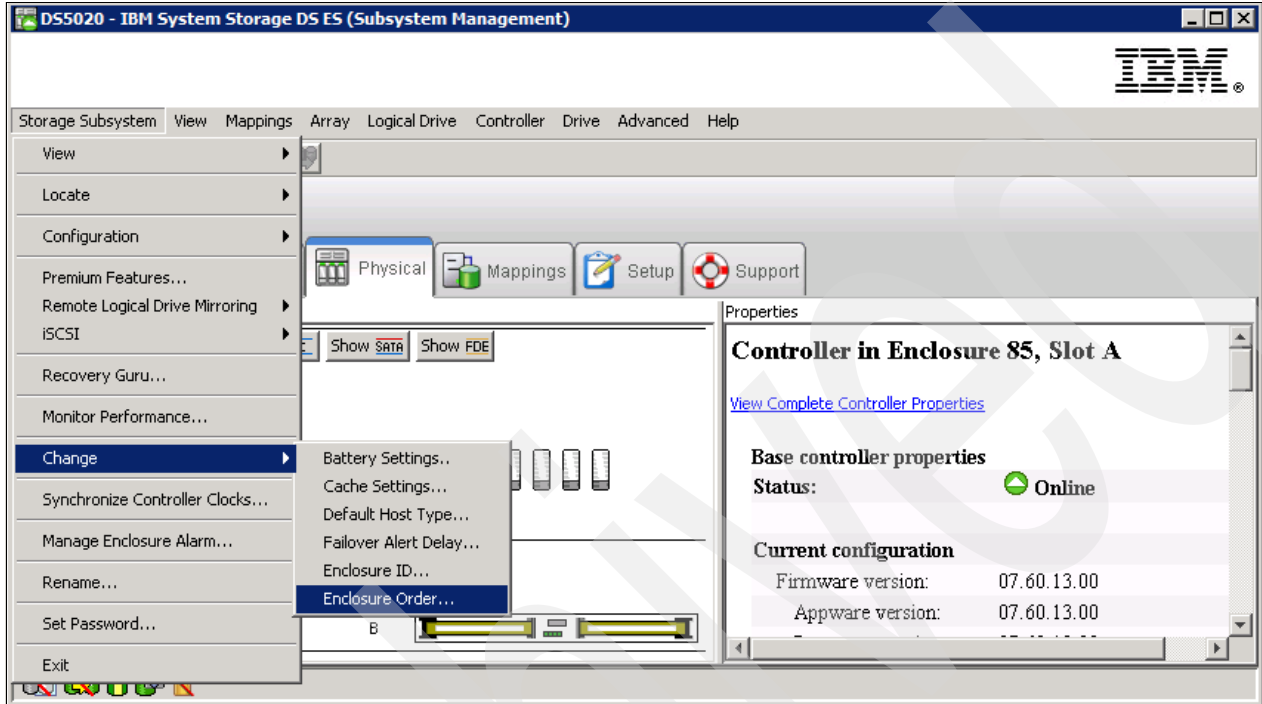


Figure 4-36 Change the enclosure order

Now you can sort the enclosures according to your site setup, as shown in Figure 4-37.

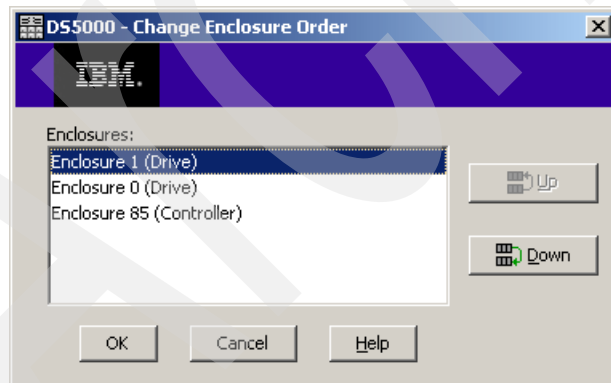


Figure 4-37 Changing the enclosure order

Manage Enclosure Alarm

Enclosure alarms can be managed from the Storage Manager subsystem window. This is done by selecting **Storage Subsystem** → **Manage Enclosure Alarms**.

Note: This option is not available with all storage subsystem models.

Manage Enclosure Alarms is used to respond to an alarm on the storage subsystem. If a critical failure of the storage subsystem or the enclosures occurs, then an audible alarm will sound if it is enabled.

The audible alarm is just one of three ways in which Storage Manager will inform you of a fault. The other two are that the Alarm button becomes animated and the Recovery Guru button in the toolbar appears.

By default, the alarm is disabled, but this can be changed by the Storage Manager client. This can be done on all enclosures or individual enclosures and controllers (Figure 4-38). The two options are to always sound an alarm or never sound an alarm.

The lower part of the window shows the alarms that are already sounding, which can be silenced by clicking the **Silence Alarm(s)** button.

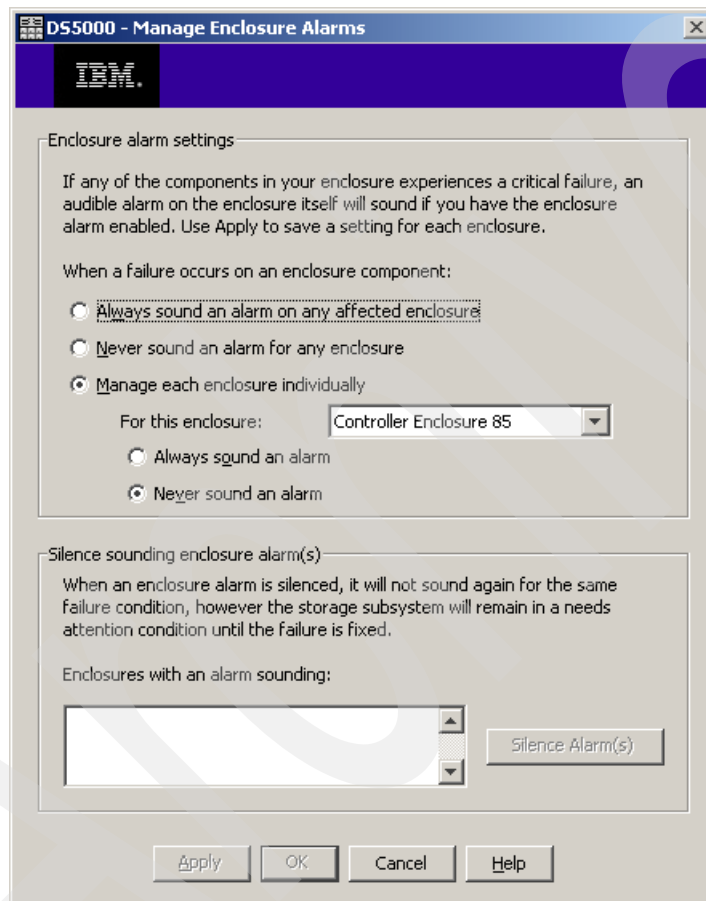


Figure 4-38 Manage Enclosure Alarms

4.8.5 Updating the controller microcode

Before you use your DS5000 storage subsystem to process production data, update the firmware (or microcode) of your DS5000 storage subsystem to prevent any unnecessary, known failures, and make use of the latest improvements. New firmware might be required when installing a new version of the Storage Manager software. Check the IBM support Web site for available updates at the following address:

<http://www.ibm.com/servers/storage/support/disk>

Because firmware updates are a customer responsibility, it is a recommended practice to periodically review this Web site for newer versions in order to keep your DS storage subsystem updated with the latest fixes and enhancements.

To facilitate this task, you can subscribe to a service at the IBM My Support Web site that will send you automatic notifications about new firmware. You will receive an e-mail when new firmware levels have been updated and are available for download and installation. To register for My Support, visit the following address:

<http://www.ibm.com/support/mysupport/us/en>

From the Web site, select your specific DS5000 model and click **Download**. The available firmware packages and versions for your product will be shown.

The DS5000 storage subsystem components that can be updated are:

- ▶ Controller firmware
- ▶ Non-Volatile Static Random Access Memory (NVSRAM)
- ▶ Environmental services monitor, ESM, and canister
- ▶ Disk drives

For complete instructions, see the readme file of the specific version you want apply, which is available at the IBM support Web site. You can find a detailed procedure in 7.1.5, “Updating controller firmware” on page 333.

Always check the IBM System Storage Interoperation Center (SSIC) Web site for the latest supported combinations of firmware in your specific environment at the following address:

<http://www.ibm.com/systems/support/storage/config/ssic/index.jsp>

Note: Use the IBM System Interoperation Center Web site to check the latest compatibility information before making any changes in your storage subsystem environment

4.9 Step-by-step configuration

This section describes the major steps in the configuration of a DS5000 storage subsystem:

1. Configuration planning
2. Enabling the premium features
3. Creating arrays and logical drives
 - a. Using the Automatic Configuration Wizard
 - b. Using the manual procedure
4. Configuring storage partitioning
5. Configuring logical drives from Windows
6. Monitoring and alert options
7. Protecting the configuration

4.9.1 Configuration planning

A configuration of a DS5000 storage subsystem can be complex, especially when different operating systems and storage partitioning are involved. Therefore, you should plan the configuration you need to apply in advance.

On one hand, you need to define the arrays and the logical drives you need, including considerations such as number of drives in the arrays, size of the logical drives, RAID level, and segment size. To plan the disk layout, you need to know the attached hosts, their operating system, and the applications using the DS5000 storage subsystem. You also need to consider the performance aspects and requirements.

On the other hand, you also need to define the layout of the attached hosts with their host bus adapters and the mappings of the logical drives to specific host groups or hosts.

Before starting to configure your DS5000 storage subsystem, you should draw up a plan. This will allow you implement your storage requirements in the best way, minimizing any problems with a future desired reconfiguration. Keep all the necessary information regarding the different configuration options in your plan in order to configure your storage.

4.9.2 Enabling the premium features

In order to activate your purchased premium features, perform these steps:

1. Obtain the Feature Enable Identifier for your Storage Subsystem.
2. Use the Web to generate the activation file.
3. Install the activation file using the DS Storage Manager.

Obtain the Feature Enable Identifier for your storage subsystem

From the Subsystem Management window of your DS5000 storage subsystem, select **Storage Subsystem** → **Premium Features**. This opens a window that shows the current activation status of the premium features in your subsystem, as shown in Figure 4-39.

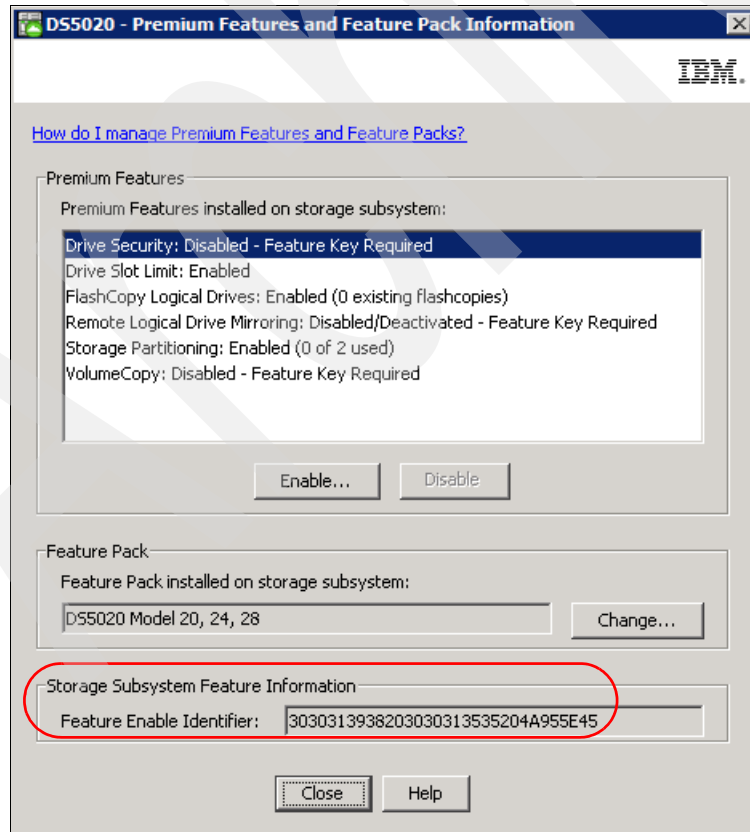


Figure 4-39 Premium Features and Feature Pack Information

Record the Feature Enable Identifier and continue with the next section to generate your activation file. Later we use the same options and window to activate the purchased premium features.

Use the Web to generate the activation file

With the Feature Enable Identifier and the registration card provided with the machine for the purchased premium features, go to the following Web site to generate the activation file:

<https://www-912.ibm.com/PremiumFeatures/>

In the Web site, select **Activate a Premium Feature**, read the requirements, and then click **Continue**. Complete the fields presented in the window shown in Figure 4-40 by entering the feature activation code received in the card shipped with the DS5000 storage subsystem, the feature enable identifier obtained in “Obtain the Feature Enable Identifier for your storage subsystem” on page 176, and your specific machine type, model number, and serial number.

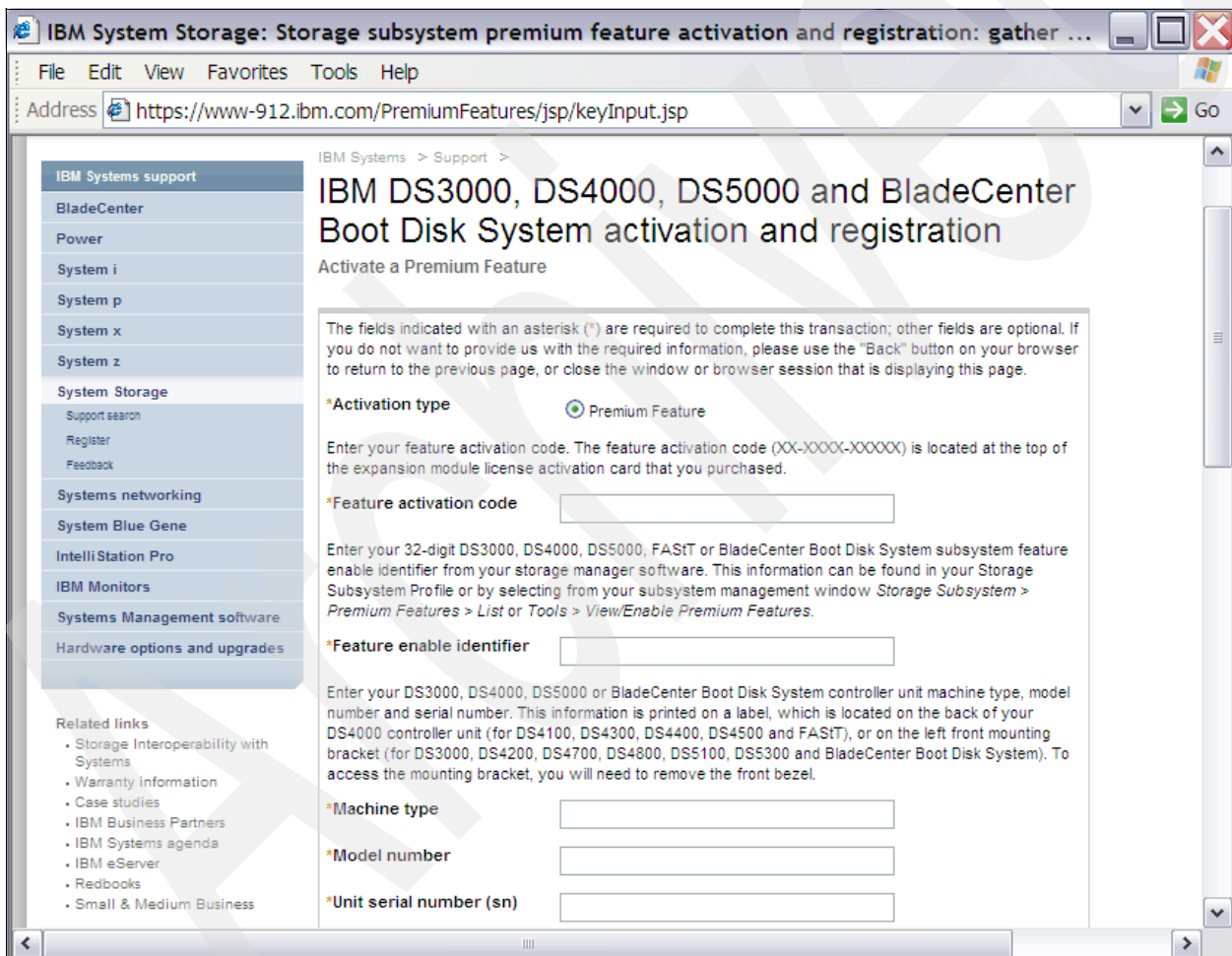


Figure 4-40 Premium features registration

Scroll down, complete the remaining fields by entering your e-mail address, and submit the information by clicking **Continue**. The activation key file is then e-mailed to you. Save the received file in your folder, and continue with next step.

Install the activation file using the DS Storage Manager

From the Subsystem Management window of your DS5000 storage subsystem, select **Storage Subsystem** → **Premium Features**. This opens a window that shows the current activation status of the premium features in your subsystem.

Click **Enable** and, as shown in Figure 4-41, select the key file that you received by e-mail.

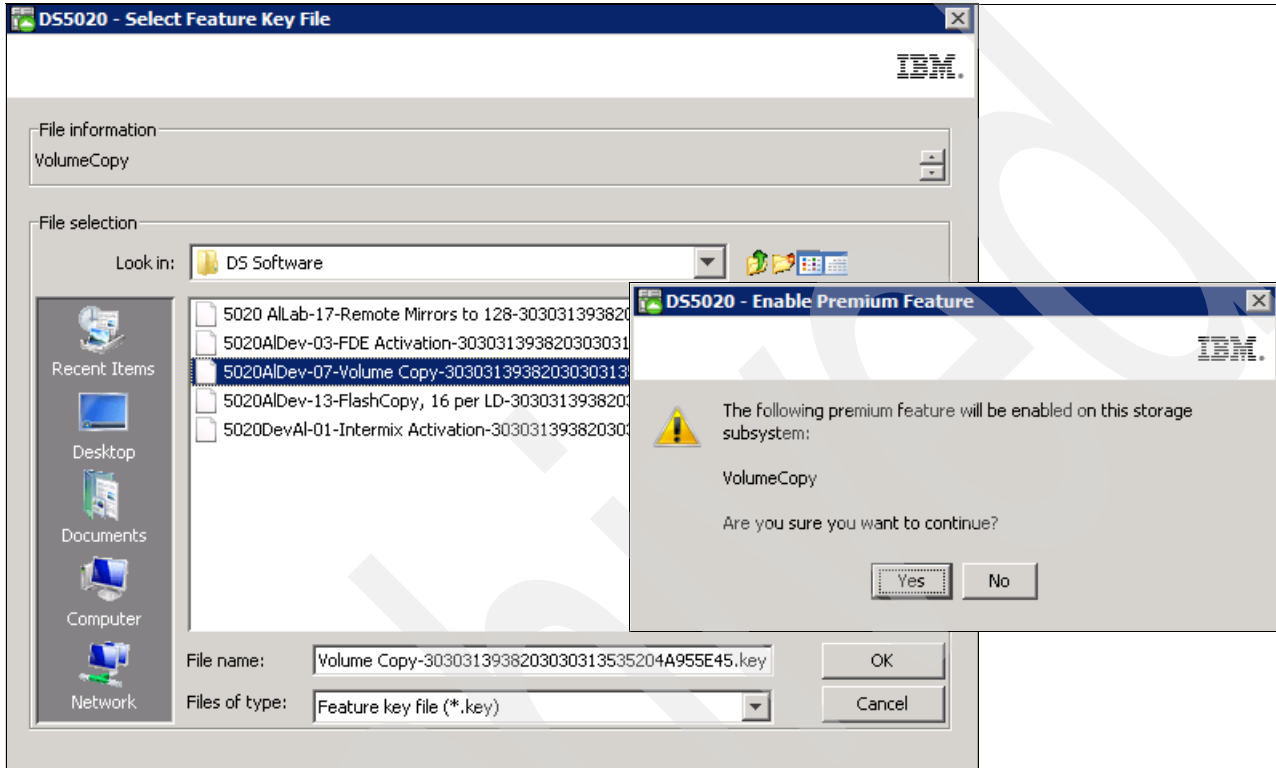


Figure 4-41 Selecting key file for premium features activation

Confirm the Enable Premium Feature action by clicking the **Yes** button. Then select **Storage Subsystem** → **Premium Features**, or select the **Setup** view of the Subsystem Management window and then **View/enable Premium Features**. The premium feature activated shows as **Enabled**, as shown in Figure 4-42.

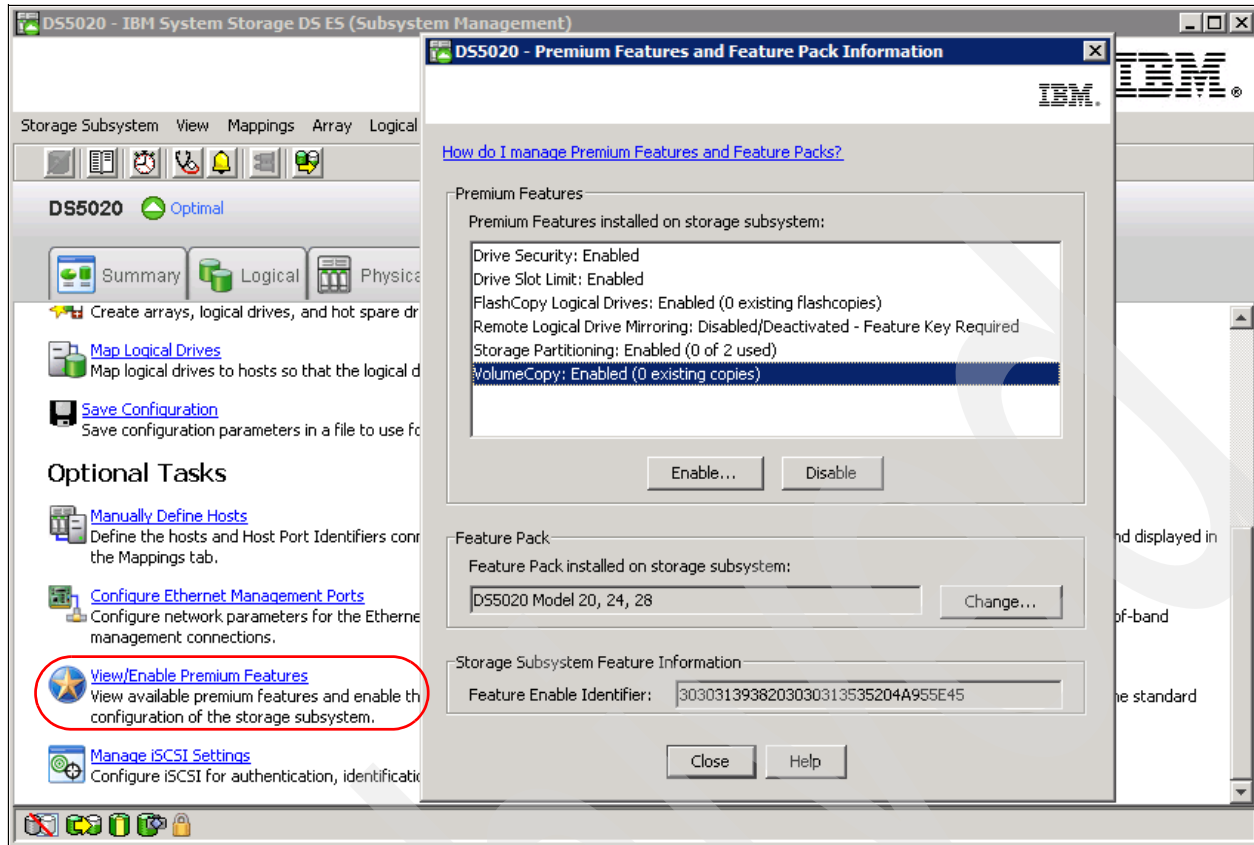


Figure 4-42 Premium feature activated

Repeat the previous steps for any purchased premium features that have not yet been activated.

4.9.3 Automatic configuration

In this section, we cover the necessary steps to configure unconfigured storage subsystem capacity into logical drives using the Automatic Configuration wizard integrated into the DS Storage Manager software.

The Automatic Configuration wizard can be used to create multiple arrays with logical drives, and hot spare drives using the same attributes for all arrays, such as RAID level, number of drives per array, number of logical drives, and I/O type. The wizard configures all the non-configured disk drives in the system with minimal interaction.

If you need to define space that has not been used in already defined arrays, or configure capacity with some specific parameters, you will find the manual configuration method more useful. It is described in 4.9.4, “Manual configuration” on page 185, and allows you to select each of the parameters for the creation of both arrays and logical drives.

To access the Automatic Configuration wizard, perform these steps:

1. From the Enterprise Management window, select the DS5000 storage subsystem you want to configure. Double-click it or right-click it and select **Manage Storage Subsystem**.
2. From the Subsystem Management interface, select **Storage Subsystem** → **Configuration** → **Automatic**, or from the Setup view of the Subsystem Management interface, select **Configure Storage Subsystem** → **Automatic Configuration** → **OK**.

Here is an example showing the different steps:

1. Select **Configure Storage Subsystem** → **Automatic Configuration** → **OK**, as shown in Figure 4-43.

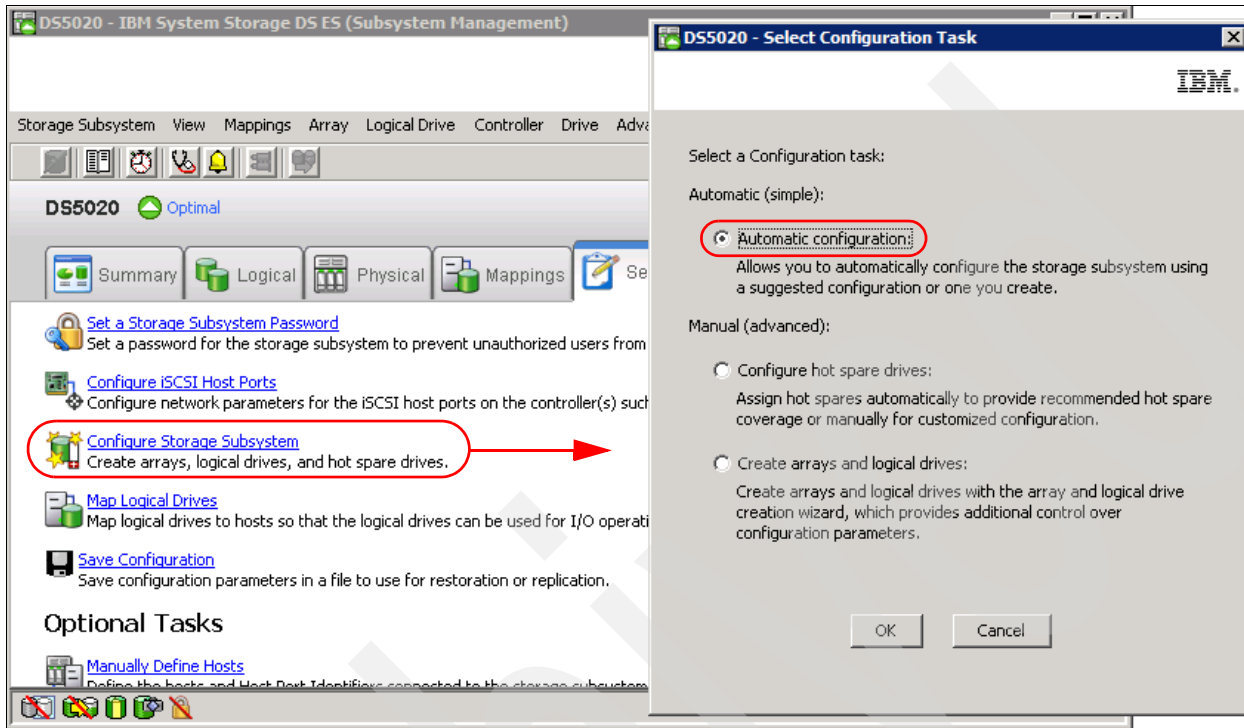


Figure 4-43 Automatic configuration

2. Read the introduction window. It reminds you to quit the wizard and start a manual configuration process if your requirements need different RAID levels, volume sizes, and so on. If this is not the case, click **Next** to continue.

- In the window that follows, choose between **Choose a suggested configuration**, selecting the RAID level desired, or **Create your own configuration**, as shown in Figure 4-44.

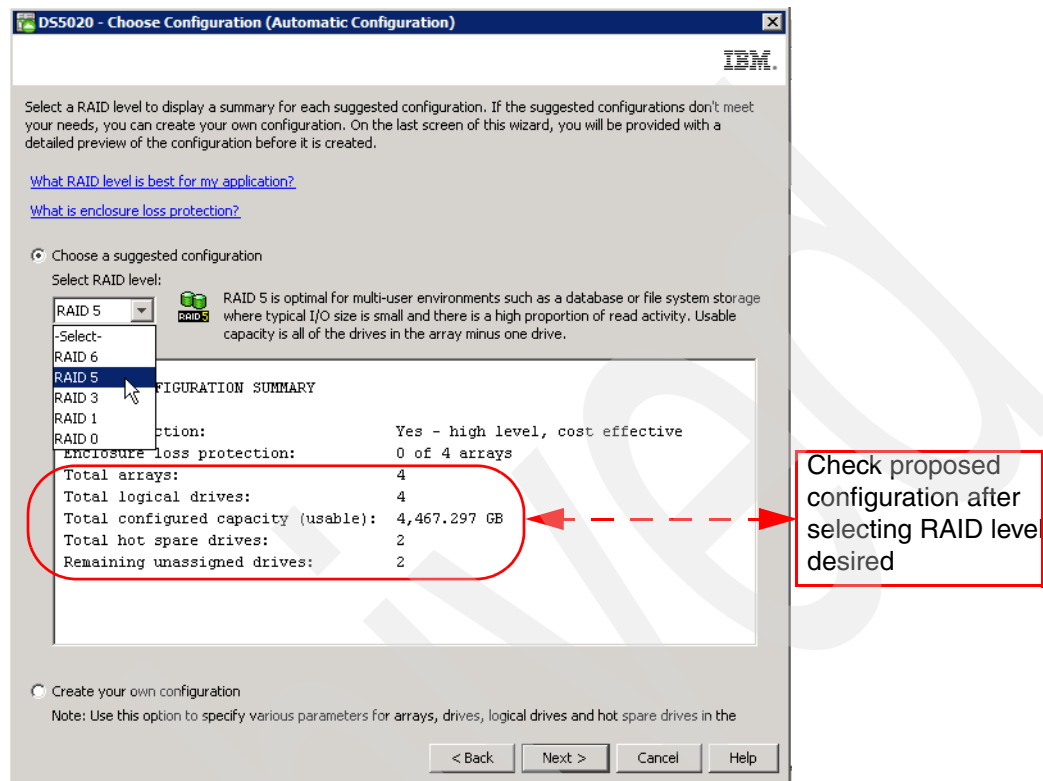


Figure 4-44 Choosing the RAID level

- Suggested configuration:** This is the default option, where you only have to select the RAID level to be applied to all unconfigured disks. Once you select the RAID level, a summary is presented based on the resources available in your configuration. Check this summary configuration. You can also click **Next** to see a preview of the resulting configuration, based on the RAID level selected and the quantity of free drives in the system.

Review the information carefully, and if it is correct, click **Finish** to proceed.

- Create your own configuration: This allows you to customize the configuration by selecting the additional parameters shown in Figure 4-45.

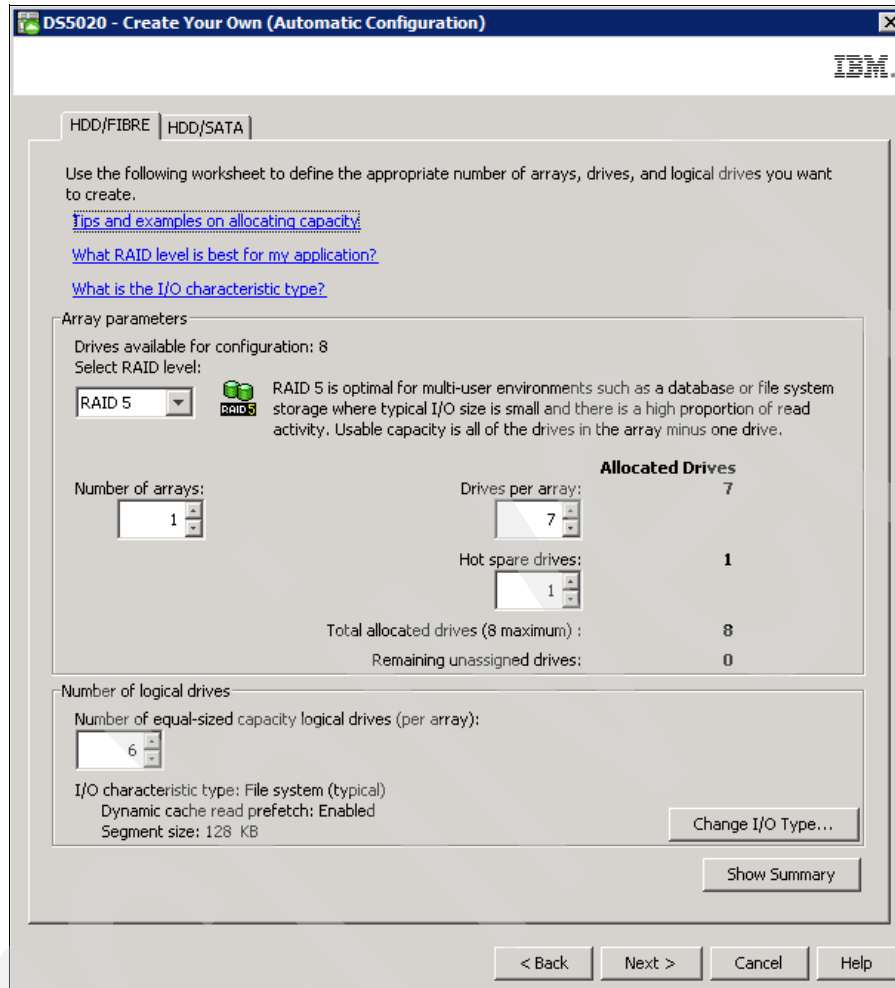


Figure 4-45 Customized automatic configuration

The options for both Fibre Channel or SATA drives are:

- RAID level: 0, 1, 3, 5, or 6 (or RAID 10 by selecting multiple drives in RAID 1)
- Number of logical drive groups or arrays
- Number of drives per array
- Number of hot spares
- Number of logical drives per array
- I/O type characteristics

The window that appears lets you modify each of the parameters, as though you were using a worksheet, and gives you the option to have a red warning appear if your selections exceed the available resources. If you have both FC and SATA drives, you have to complete two separate worksheet windows to set your desired configuration.

You can select the online help for an explanation of each of the parameters. For additional information, see 4.1.1, “DS5000 arrays and RAID levels” on page 104.

Click **Change I/O type** if you want to set up your logical drives for other types of access (this is different than the preset defaults for regular file systems, as shown in Figure 4-46).

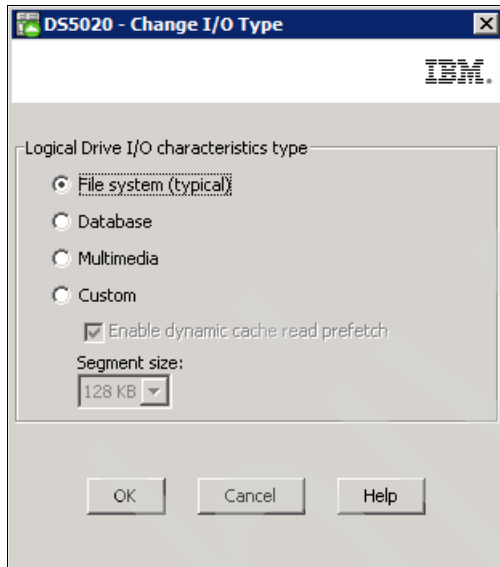


Figure 4-46 Changing I/O characteristics

For additional information about these values, see 4.1.5, “Segment size” on page 120.

Once you have specified all of your desired values, click **OK** to open the window shown in Figure 4-47.

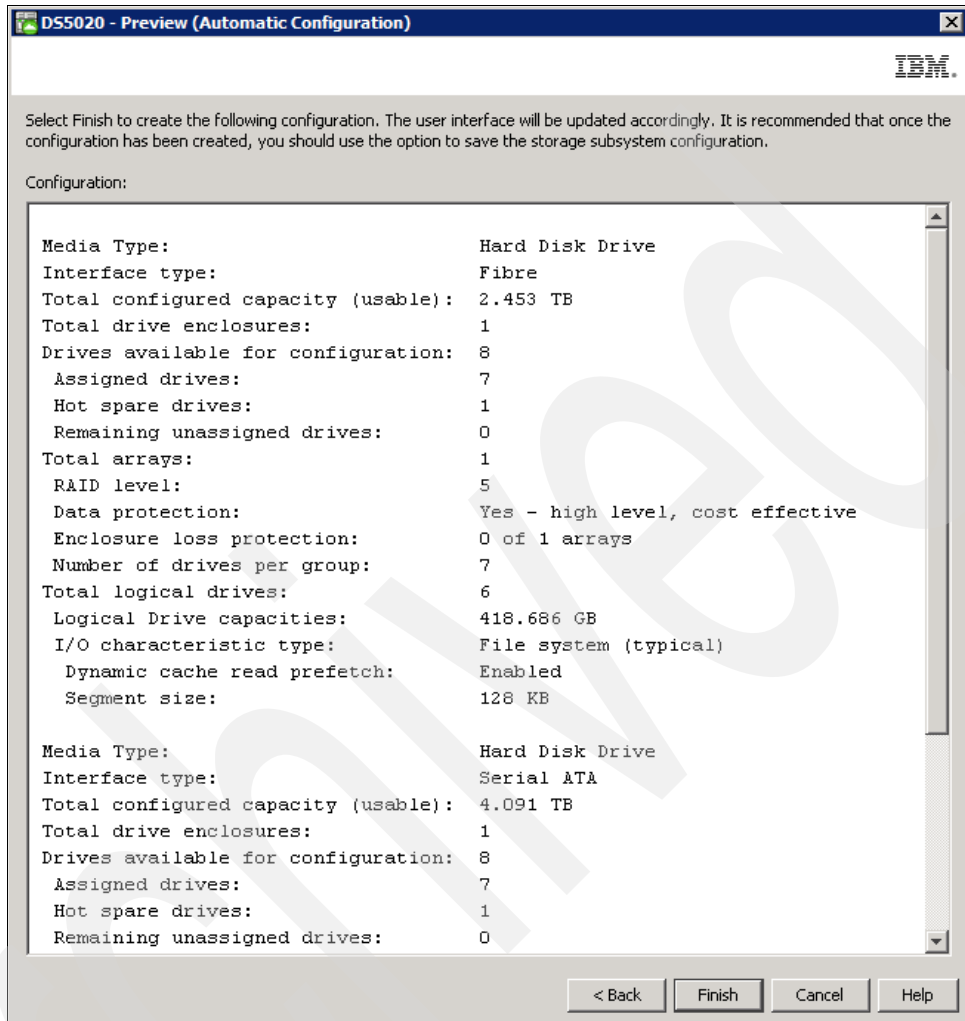


Figure 4-47 Preview automatic configuration

4. Review the configuration that will be created based on your input, and click **Finish** to proceed. You see a confirmation message informing you that, depending on the number of drives and volumes, it might take some time for all the changes to appear in the Management window. Do not submit another request until this configuration is complete. You can check the event log to see whether the operation was successful.

- After completion, the Storage Manager shows the arrays, logical drives, and hot spares created automatically based on your input, as shown in Figure 4-48.

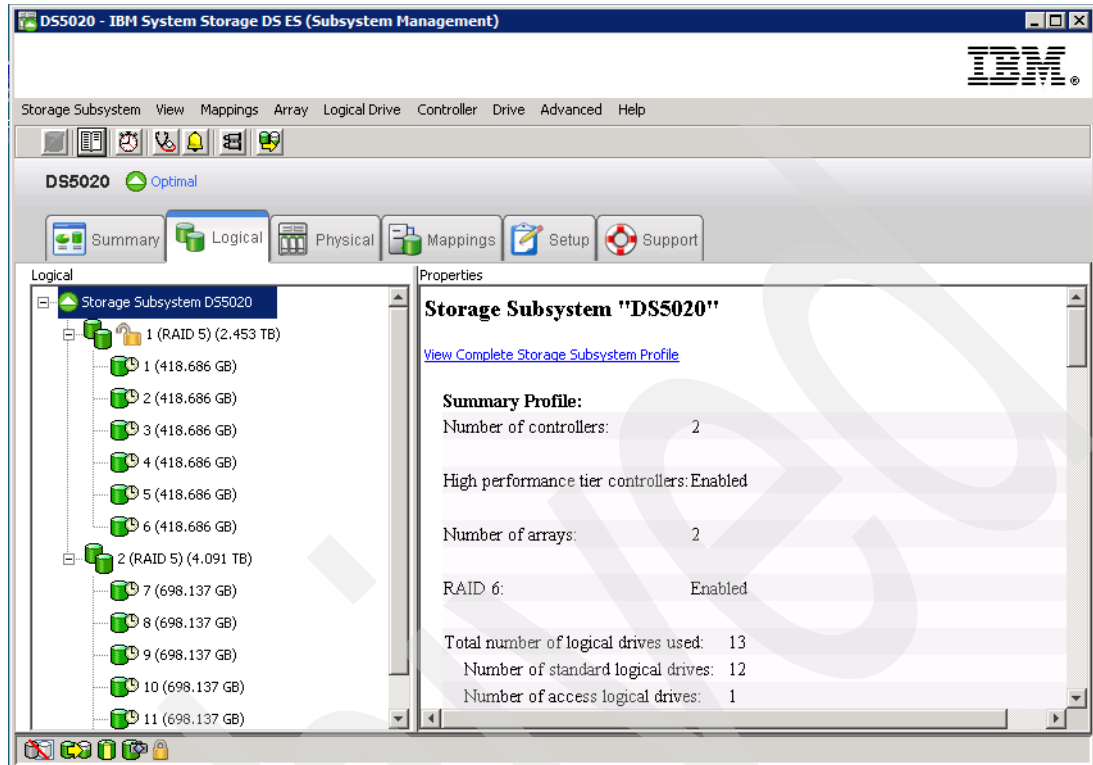


Figure 4-48 Automatic configuration results

4.9.4 Manual configuration

In this section, we cover the necessary steps to configure available storage subsystem capacity into logical drives using the Storage Manager interface. We already covered the automatic procedure, where the options are limited. Now we cover the additional parameters that you can select to better configure the storage for your environment.

We split the tasks into two sections, starting with the creation of global hot spares. We recommend performing this task first to ensure that you have enough spare drives before creating the logical drives.

Defining hot spare drives

The concept of a hot spare drive is explained in 4.1.3, "Hot spare drive" on page 114. Here we cover the available methods of defining hot spare drives with Storage Manager V10.60. Because we plan for this configuration, we start assigning hot spare drives and then continue with the arrays.

To start configuring hot spares, from the Setup view of the your Subsystem management window, select **Configure Storage Subsystem**, as shown in Figure 4-49.

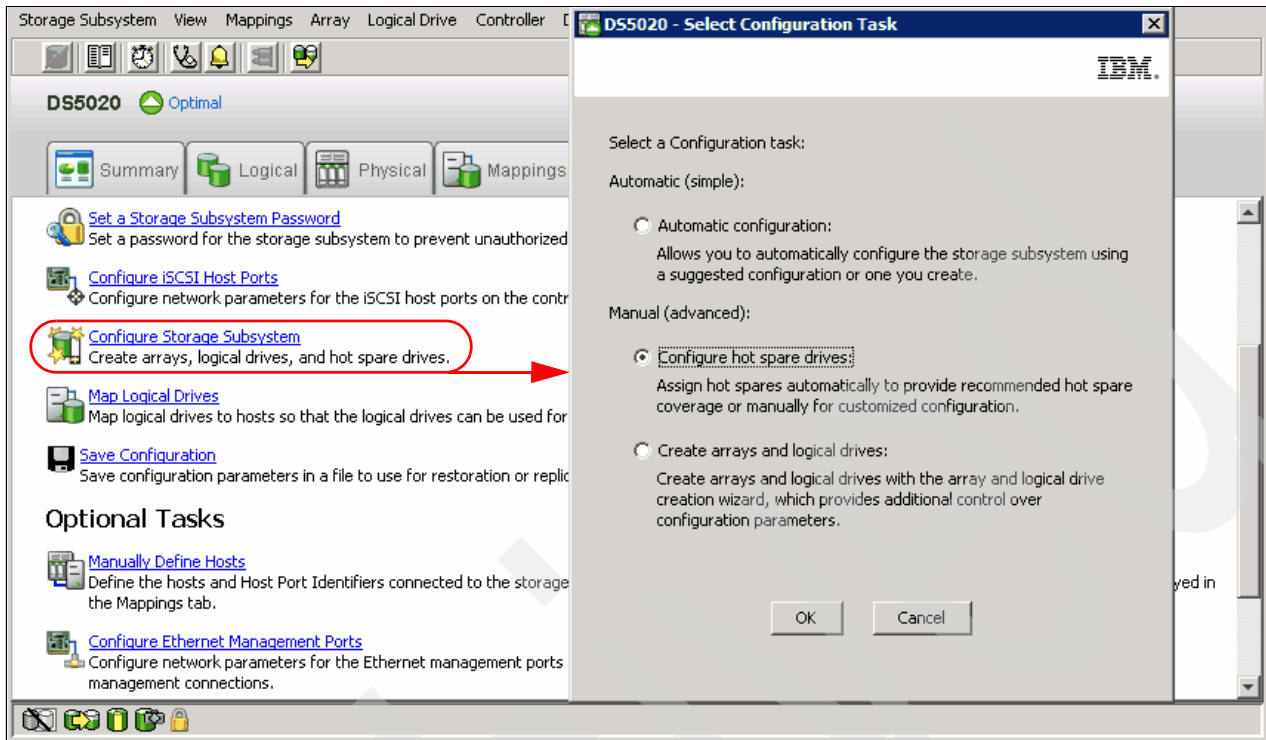


Figure 4-49 Configuring hot spares

Select **Configure hot spare drives**. The window shown in Figure 4-50 opens.

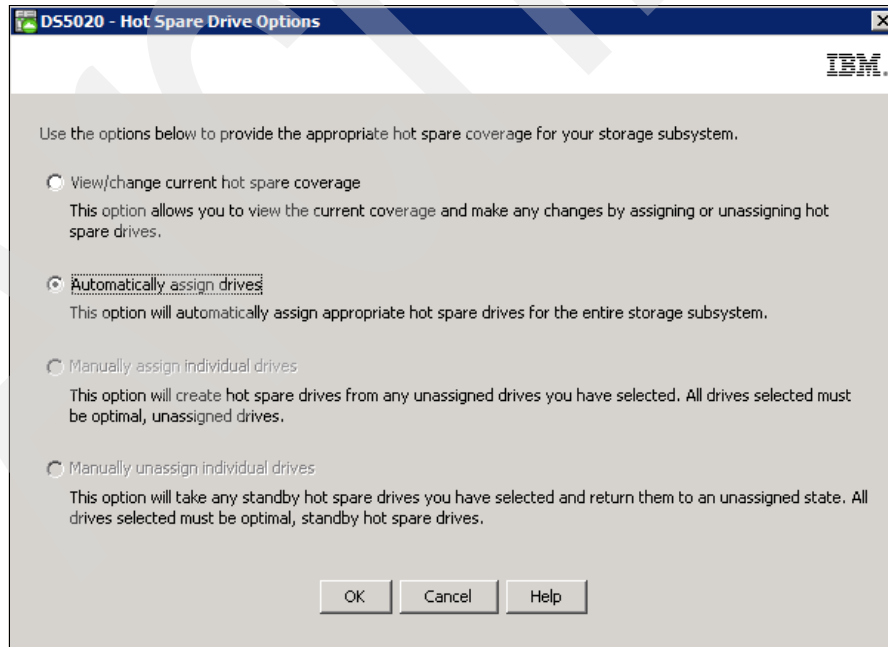


Figure 4-50 Hot Spare Drive Options

You have two methods to define hot spare drives:

- ▶ Automatic assignment
- ▶ Manual assignment

Tip: Select drives of equal or greater size than the total capacity of the largest disk in the storage subsystem.

Especially in large configurations with arrays containing many drives, it might be necessary to define multiple hot spares, because the reconstruction of a failed drive to a hot spare can take a long time.

Consider having spares for each type of disk drives, that is, FC, SATA, and security enabled FDE.

For an array that has FDE drives that are not secured, the hot-spare drive can be either an unsecured FDE drive or a non-FDE drive.

For an array that has secured FDE drives, the hot-spare drive should be an unsecured FDE drive of the same or greater capacity. After the unsecured FDE hot-spare drive is used as a spare for a failed drive in the secured RAID array, it is security enabled.

Automatic assignment

For automatic hot spare assignment, follow these steps:

1. To automatically create the hot spare coverage using the drives that are available, select **Automatically assign drives**, as shown in Figure 4-50. The recommended quantity of spares drives needed for your configuration are created automatically.
2. Select the **Physical** tab to view the results of the automatic hot spare creation. You see the drives assigned in Figure 4-51.

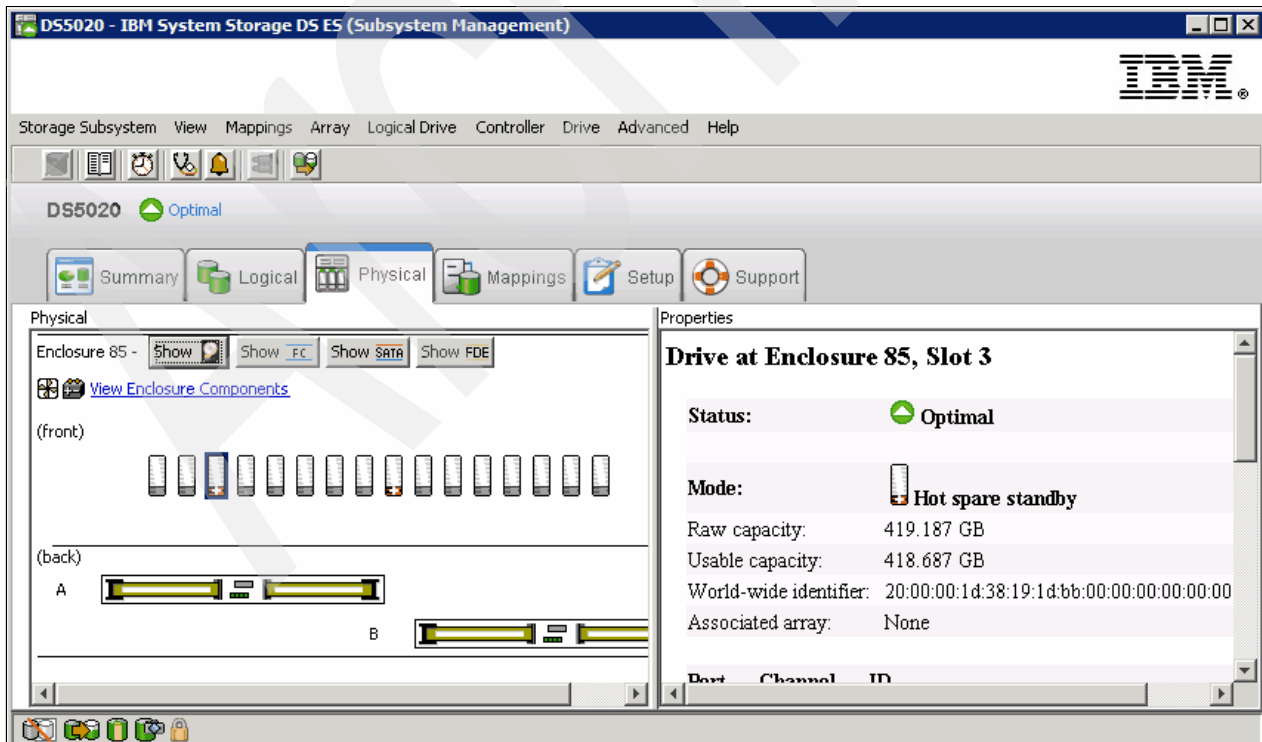


Figure 4-51 Automatic hot spare creation

This automatic hot spare creation function created one hot spare drive for every 30 disk drives of the same type (FDE/FC/SATA).

Manual assignment

To perform manual hot spare assignment, follow these steps:

1. To manually define a hot spare drive, select, from the Setup view, **Configure Storage Subsystem** → **Configure hot spare drives** → **View/change current hot spare coverage**.

You can also select, from the Physical view, the specific non-assigned drive you want to set up as a hot spare, right-click it, and select **Hot Spare Coverage**, or, from the top menu, select **Drive** → **Hot Spare Coverage**, as shown in Figure 4-52.

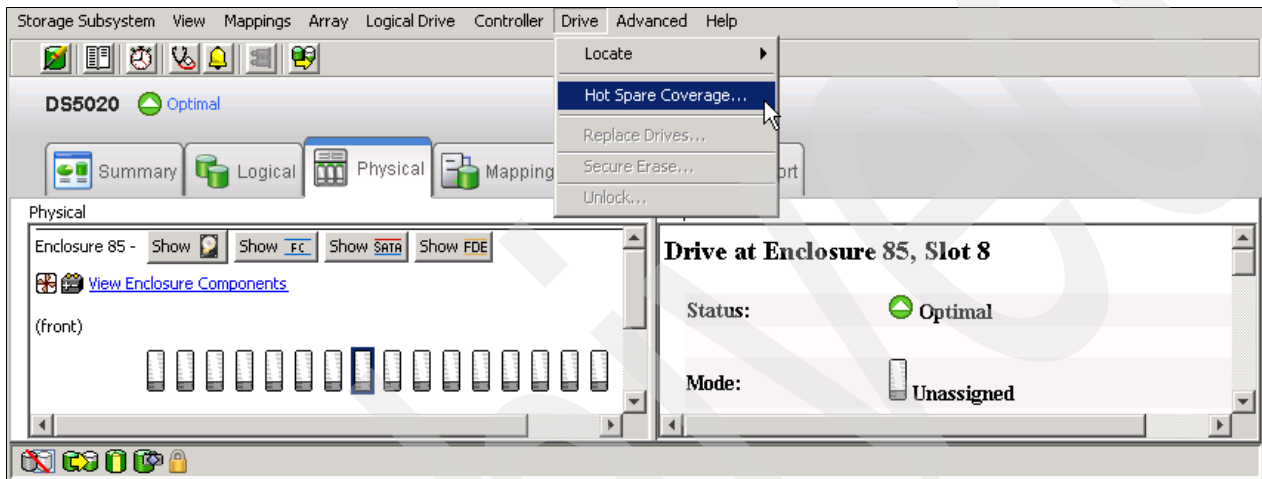


Figure 4-52 Hot Spare Coverage

2. This opens the Hot Spare Drive Options window, as shown in Figure 4-50, but this time with the option **Manually assign individual drives** selected by default. Click **OK** and the unassigned drive is defined as a hot spare.

If any array on the DS5000 storage subsystem contains larger drives than the drive you have chosen, a warning message appears notifying you that not all arrays are protected by the hot spare drive.

Remember to create hot spares for each type of disk drive. Use the push buttons in the Physical view to show each type of drives, as shown in Figure 4-53.

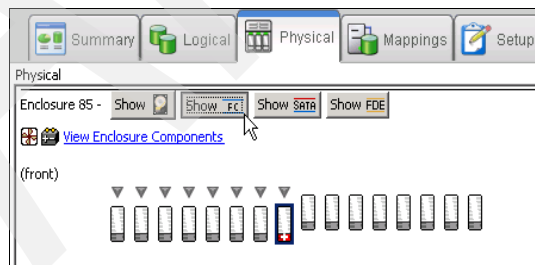


Figure 4-53 FC drives protected with hot spare

3. Because in this example we have more than one type, we repeat the previous steps to create a second hot spare drive for the SATA disk drives, which are still unprotected. The operation finishes and the window shown in Figure 4-54, in the physical view, opens. Click the **Show SATA** button.

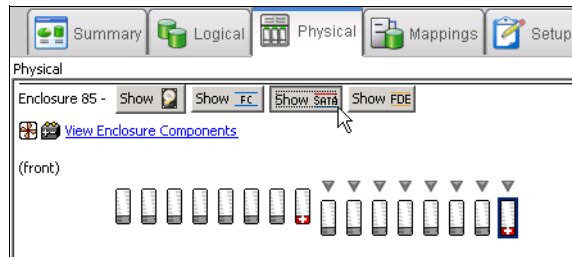


Figure 4-54 SATA drives protected with hot spare

4. To unassign a hot spare drive and have it available again as a free drive, highlight it, select **Drive** → **Hot Spare Coverage**, and then select **Manually unassign individual drives**.

View/change hot spare coverage

After you have configured your spare drives, you can review or modify your settings using this option. Perform the following actions:

1. To start the option from the Storage Manager Client, select **Drive** → **Hot Spare Coverage**.

- The Hot Spare Drive Options window opens, as shown in Figure 4-50. Select **View/Change current hot spare coverage** and click **OK**. This opens the window shown in Figure 4-55.

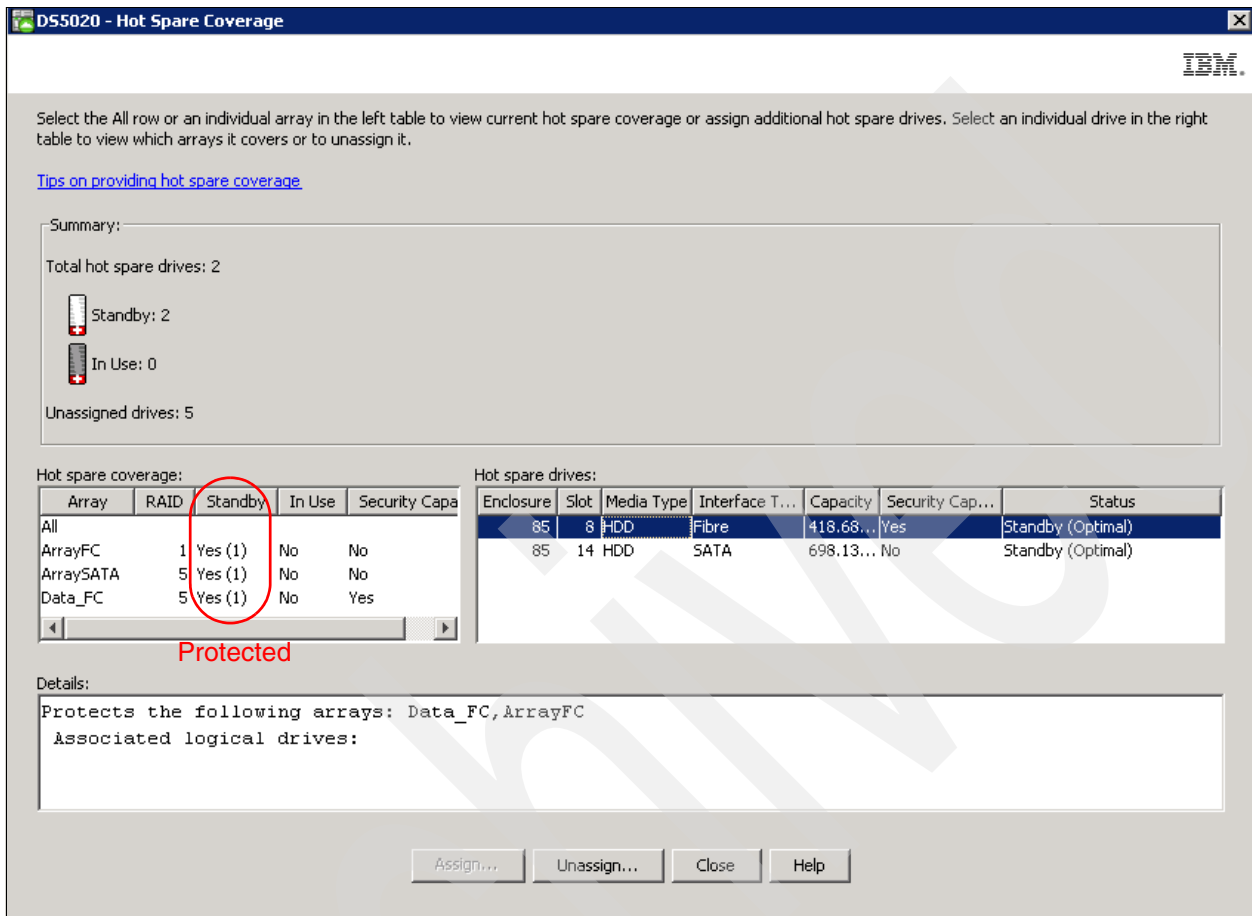


Figure 4-55 View/Change hot spare coverage

Make sure to use this window to check that all your arrays are protected. If they are not protected, implement complete hot spare protection for all the arrays, as shown in Figure 4-56.

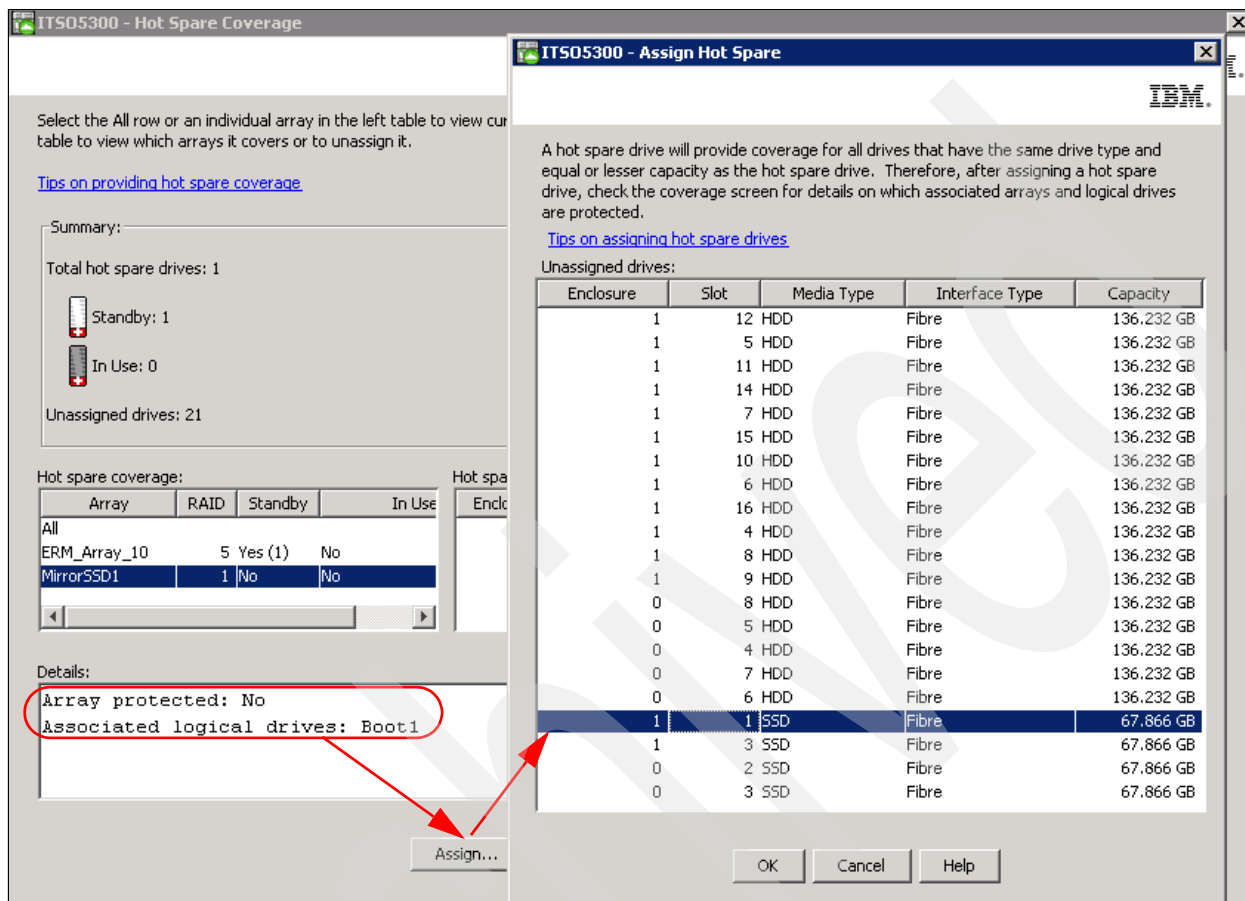


Figure 4-56 Assigning hot spare protection for an array

Notice in this example that there is only one spare drive assigned, but because there are arrays defined with different types of disk, Array MirrorSSD1 is not protected by a hot spare drive. You can run into this situation if the arrays are created before assigning hot spare drives, or by adding drives of a different type after the storage is defined. Let us assign an additional hot spare drive using this interface to solve this issue

- To resolve this exposure, click the **Assign** option button. The Assign Hot Spare window opens, as shown in Figure 4-56.
- It is clear that there are different disks in the storage subsystem, both hard disk drives (HDDs) and Solid State Drives (SSDs). Select a disk of the same type as the ones in the unprotected array (SSD in this case), and click **OK** to assign it. Remember to select a drive type of equal or greater capacity than the disks to protect.
- Finally, review the results window, checking that all arrays are protected, and click **Close**.

Creating arrays and logical drives

At this stage of the process, the storage system is installed, upgraded to the newest microcode level, and at least one hot spare drive is defined for each drive type. Arrays and logical drives can now be configured.

You can define logical drives from unconfigured capacity or from free capacity already available in previously defined arrays on the storage system:

- ▶ When you create a logical drive from unconfigured capacity, you create an array by first choosing the RAID type, and so on, and then the needed logical drives.
- ▶ If an array is not created, you can still create a logical drive, but the Storage Manager will direct you to create the array first.
- ▶ If there is free capacity on a previously defined array, you can create additional logical drives using that free capacity.

The example that follows assumes that there is unconfigured capacity on the DS5000 storage subsystem. Note that the unconfigured capacity for Fibre Channel and SATA disks are grouped separately. This procedure illustrates the most common steps to follow in setting up a logical drive from unconfigured drives:

1. In the Subsystem Management window (Figure 4-57), right-click the unconfigured capacity for the selected drive type. Because we have selected unconfigured capacity, we first create an array option. Choosing **Create Logical Drive** causes a window to open that requests that you create the array first. Select **Create Array**.

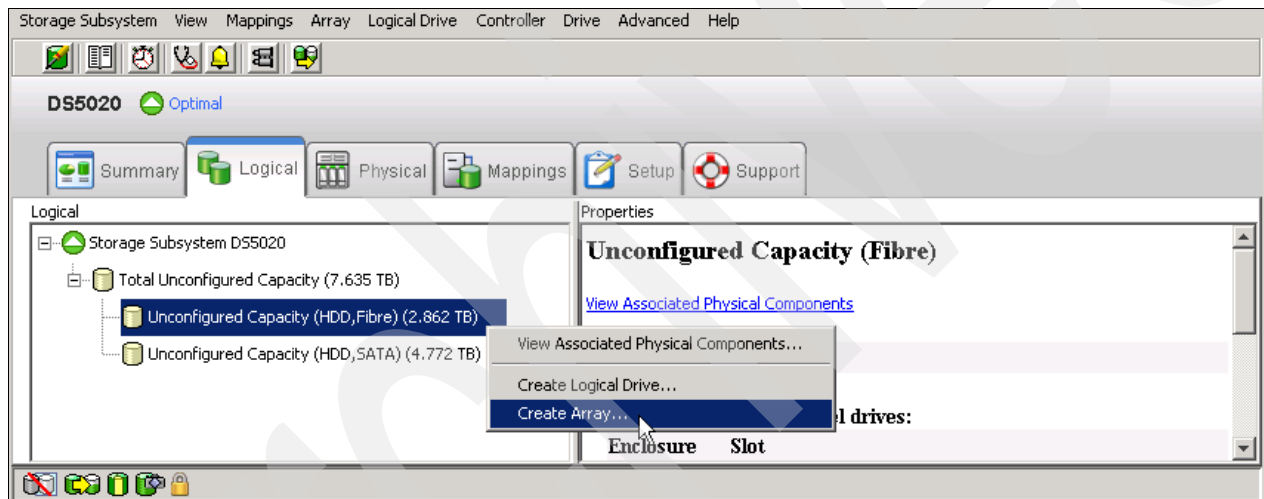


Figure 4-57 Create logical drive from unconfigured capacity

This action starts the wizard for creating the array first and then the logical drives. The first window of the wizard is an introduction to the process. It displays the available unconfigured capacity for the type of disks selected. Read the introduction and then click **Next** in order to proceed.

2. In the Create Array window, type a meaningful array name that describes your set of disks, and then select either Automatic mode or Manual mode. Automatic mode is the default option, as shown in Figure 4-58. By selecting the Automatic option, the Storage Manager software selects a combination of available drives to optimize performance and availability, and attempts to select physical drives from different enclosures in order to provide enclosure protection whenever possible. Click **Next**.

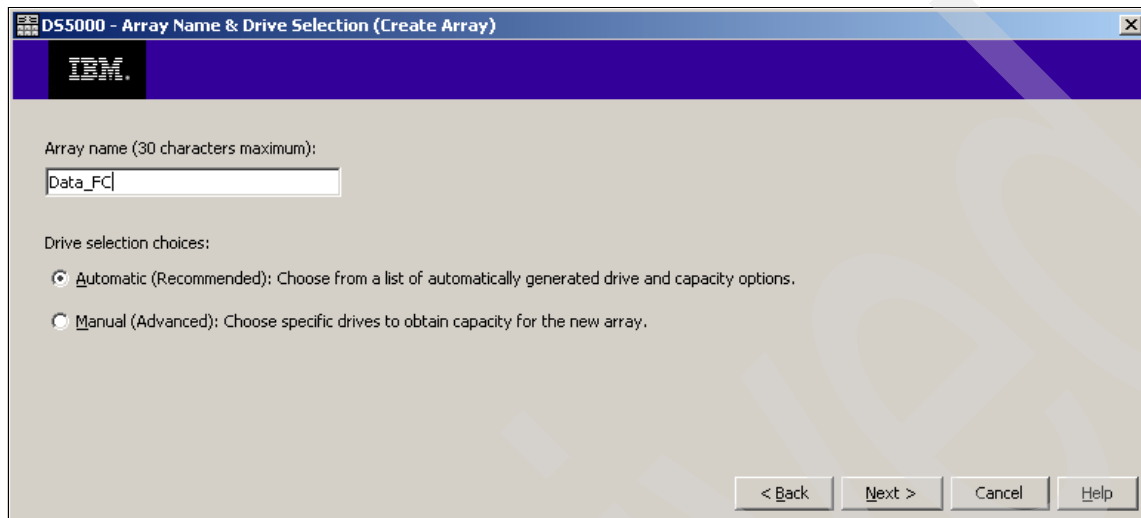


Figure 4-58 Create array and assign a name and mode

The automatic creation of the array selects the drives in the following order:

- Same capacity-same speed enclosure redundancy
- Same capacity-mixed speed enclosure redundancy
- Same capacity-mixed speed no enclosure redundancy
- Mixed capacity-same or mixed speed-no enclosure redundancy

In manual mode, you have to select all the drives individually. Make sure that you select them to maximize performance and availability.

3. Click **Next** in order to select the RAID level.
 - Select the desired RAID level. The window now displays the different capacity options depending on the unconfigured drives available in your configuration. If you have different disk sizes, you have more than one option for the same number of disks. See 4.1.1, “DS5000 arrays and RAID levels” on page 104 to determine the best RAID level for your specific environment and application.
 - Select the total capacity required and click **Finish**.

In our example (Figure 4-59), we created a RAID 0 array with three drives.

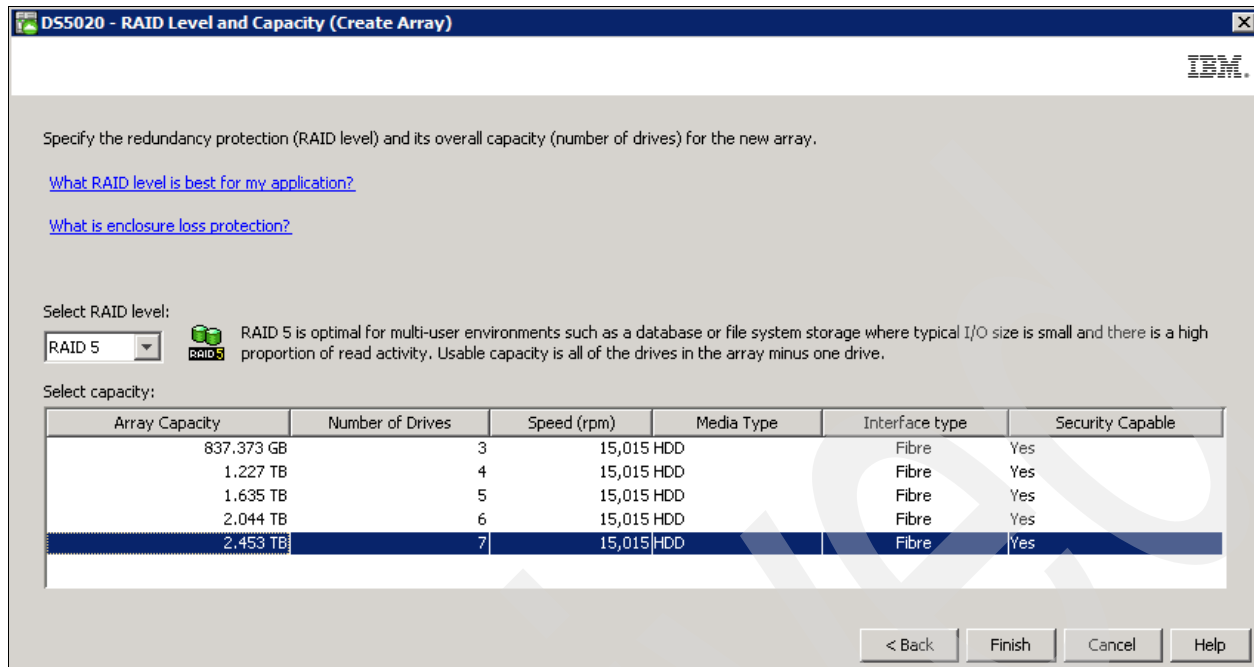


Figure 4-59 RAID level and capacity

4. The Array Success dialog box appears and confirms that the array is now created. Click **Yes** to continue with the creation of a logical drive.
5. In the Create Logical Drive wizard window, click **Next** to start configuring the logical drive.
6. In the Specify Capacity/Name dialog window:
 - a. If you want to define more than one logical drive in the array, enter the desired size capacity below the limit of the array capacity.
 - b. Assign a name to the logical drive.
 - c. If you want to change advanced logical drive settings, such as the segment size or cache settings, select the **Customize settings** option and click **Next**.

Note: The recommended settings values for creating volumes (Figure 4-60) are dynamic cache read prefetch enabled for all RAID types, and the segment size 128 KB for all but RAID 3, which is set to 256 KB.

DS5020 - Specify Capacity/Name (Create Logical Drive)

IBM.

On this screen, you specify the capacity and unique name for an individual logical drive. You must indicate exactly how much of the array's available capacity you want to allocate for an individual logical drive.

NOTE: Make sure to leave some free capacity if you want to create more logical drives on the same array.

Logical Drive parameters

Array name: Data_FC
Array RAID level: RAID 5
Free capacity: 2,512.113 (GB)

New logical drive capacity: 200 Units: GB

Logical Drive name (30 characters maximum): LogDrive1

Advanced logical drive parameters:

Use recommended settings
 Customize settings (I/O characteristics and controller ownership)

< Back Next > Cancel Help

Figure 4-60 Specifying logical drive capacity

7. The Customize Advanced Logical Drive Parameters window opens (Figure 4-61). You can set your new logical drive using any of the predefined I/O types listed, or manually set the cache read ahead multiplier, segment size, and controller ownership.
 - a. For logical drive I/O characteristics, you can specify file system, database, or multimedia defaults. The Custom option allows you to disable or enable the dynamic cache read prefetch and the segment size. Table 4-9 shows the defaults predefined for each I/O type.

Table 4-9 Logical drive defaults I/O characteristics

I/O type	File system	Database	Multimedia
Segment size	128 K	128 K	256 K
Modification priority	High	High	High
Read cache	Enable	Enable	Enable
Write cache	Enable	Enable	Enable
Write cache without batteries	Disable	Disable	Disable
Write cache with mirroring	Enable	Enable	Enable
Flush write cache after	10 seconds	10 seconds	10 seconds
Dynamic cache prefetch	Enable	Enabled	Enable
Enable background media scan	Enable	Enable	Enable
Media scan with redundancy check	Disable	Disable	Disable
Pre-read redundancy check	Disabled	Disabled	Disabled

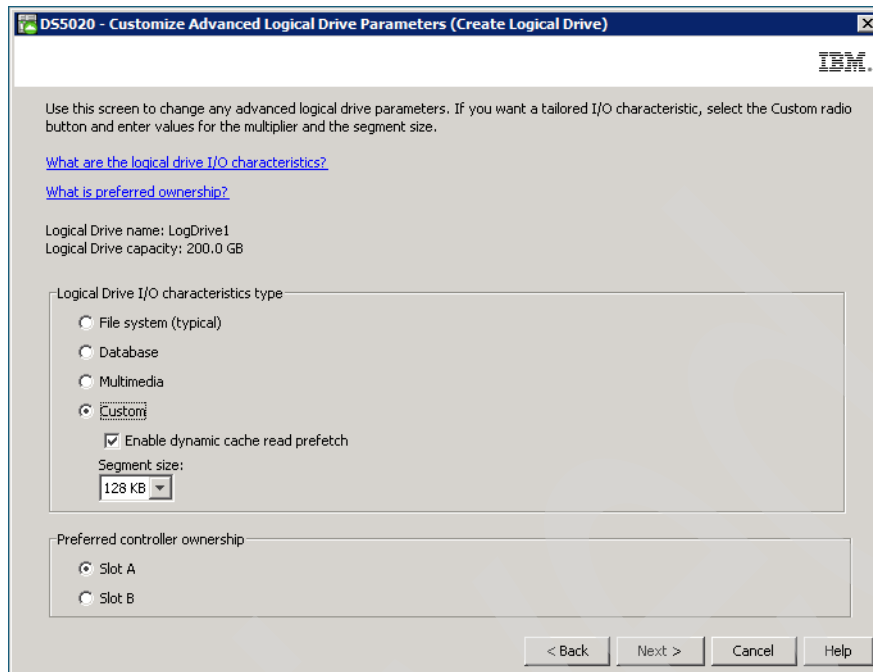


Figure 4-61 Customize logical drive parameters

- b. The segment size is chosen according to the usage pattern. For custom settings, you can directly define the segment size.
- c. Dynamic cache read prefetch is important for multimedia applications that use sequential I/O, by allowing the controller, while it is reading and copying host-requested data blocks from disk, to copy additional data blocks into the cache.
- d. The preferred controller handles the logical drive normally if both controllers and I/O paths are online. You can distribute your logical drives between both controllers to provide better load balancing between them. The default is to alternate the logical drives on the two controllers.

It is better to spread the logical drives by the load that they cause on the controller. If you do not know the expected access pattern for each logical drive, you can evaluate it by using the performance monitor option integrated with the Storage Manager client. Based on data gathered from the Performance Monitor, move some logical drives to the other preferred controller to balance the load if required (see 4.9.7, “Monitoring and alerting” on page 223, and 4.1.2, “Logical drives and controller ownership” on page 114 for more information).

- The Specify Logical Drive-to-LUN Mapping window opens (Figure 4-62). This window allows you to choose between mapping your created logical drive to the default group or to “Map later using the Mappings View.” If you choose the Default mapping, then the physical volume is mapped to the default host group and is available to any host zoned to the DS5000 storage subsystem, so it is not the recommended choice if your DS5000 storage subsystem supports more than a single partition.

Important: Manually mapping logical drives to hosts prevents unwanted mapping and is always the recommended choice.

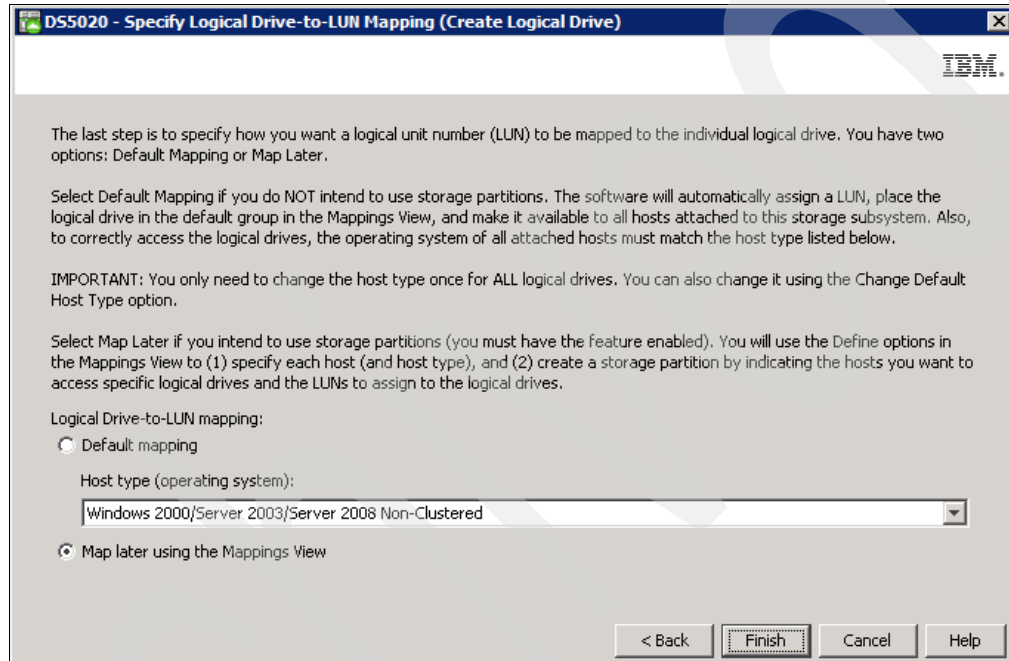


Figure 4-62 Logical drive mapping

- Click **Finish** and the Creation Successful window opens. You are prompted about whether you want to create another logical drive. Click **No** if you have finished creating logical drives or want to finish at a later time.
- The Completed window opens. The Mapping section presents the logical volume to a desired host. Click **OK** to finish.
- Repeat the same process for creating other arrays and logical drives.

If you left unconfigured capacity inside the array, you can later define another logical drive in this array. Simply highlight this capacity, right-click, and choose **Create Logical Drive**. Follow the steps that we previously outlined in this section, except for the selection of drives and RAID level (because the array is already defined).

Displaying arrays and logical drives

After creating the arrays and logical drives, the Logical view of the Subsystem Management window shows the status of the current configuration for arrays and logical drives, as shown in Figure 4-63.

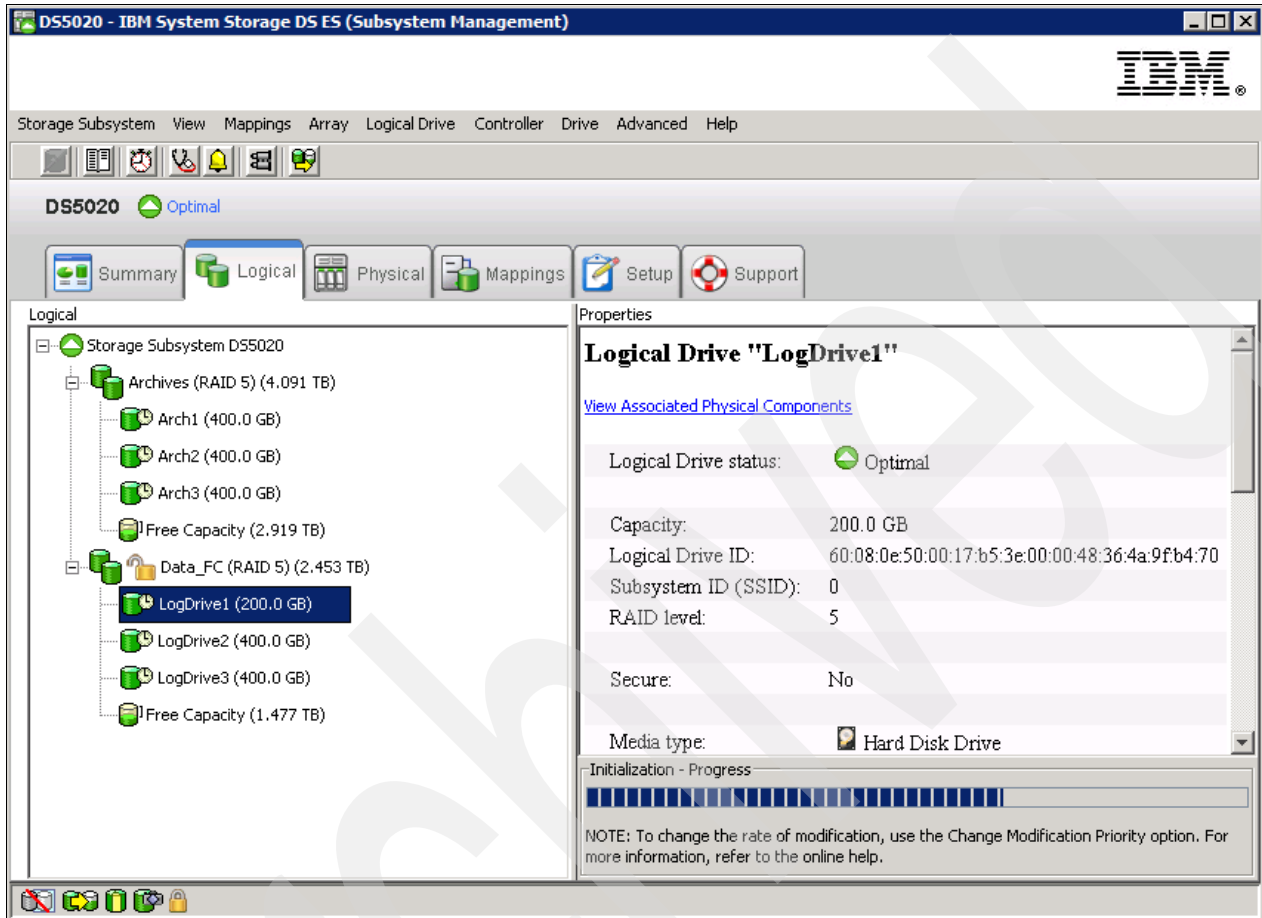


Figure 4-63 Logical Drive initialization progress

Notice that after creating a logical drive that the disk space is initialized, and shows a clock icon to the right of its name. The right frame shows the properties for the selected logical drive or array, with details about the initialization's progress in the lower part of the window. Be aware that even during the initialization process, the logical drive is immediately available for access if mapped.

To attach the arrays or logical drives to the physical disks, right-click one of them and select **View associated physical components**. The window that opens shows a blue dot under each physical disk where the array or logical drive selected belongs, as shown in Figure 4-64.

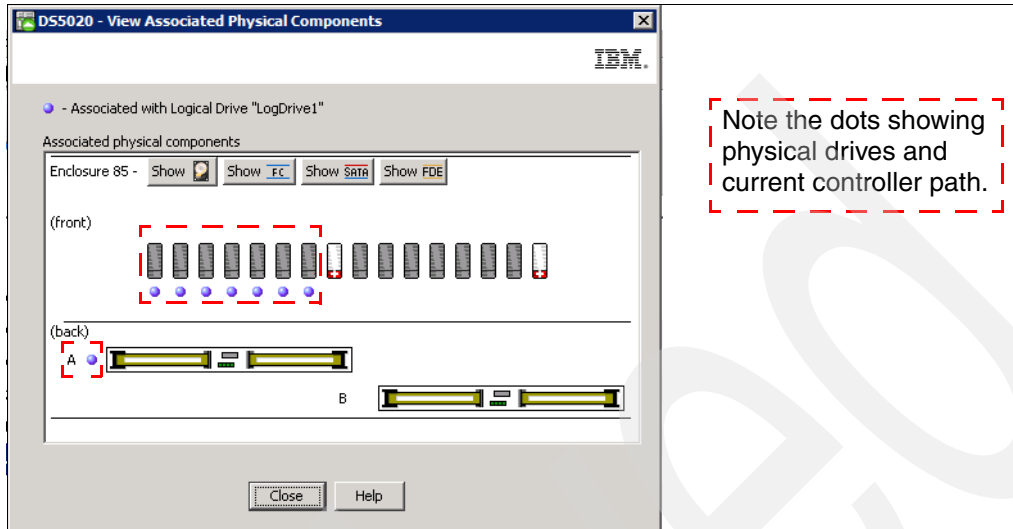


Figure 4-64 Logical/Physical relationship

The current path of the logical drive is represented by the blue dot at the right side of the controller. In this case, the logical drive named LogDrive1 is using controller A.

You can also use the Physical view of the Subsystem window and click one disk at a time to display that disk's related array, as shown in Figure 4-65.

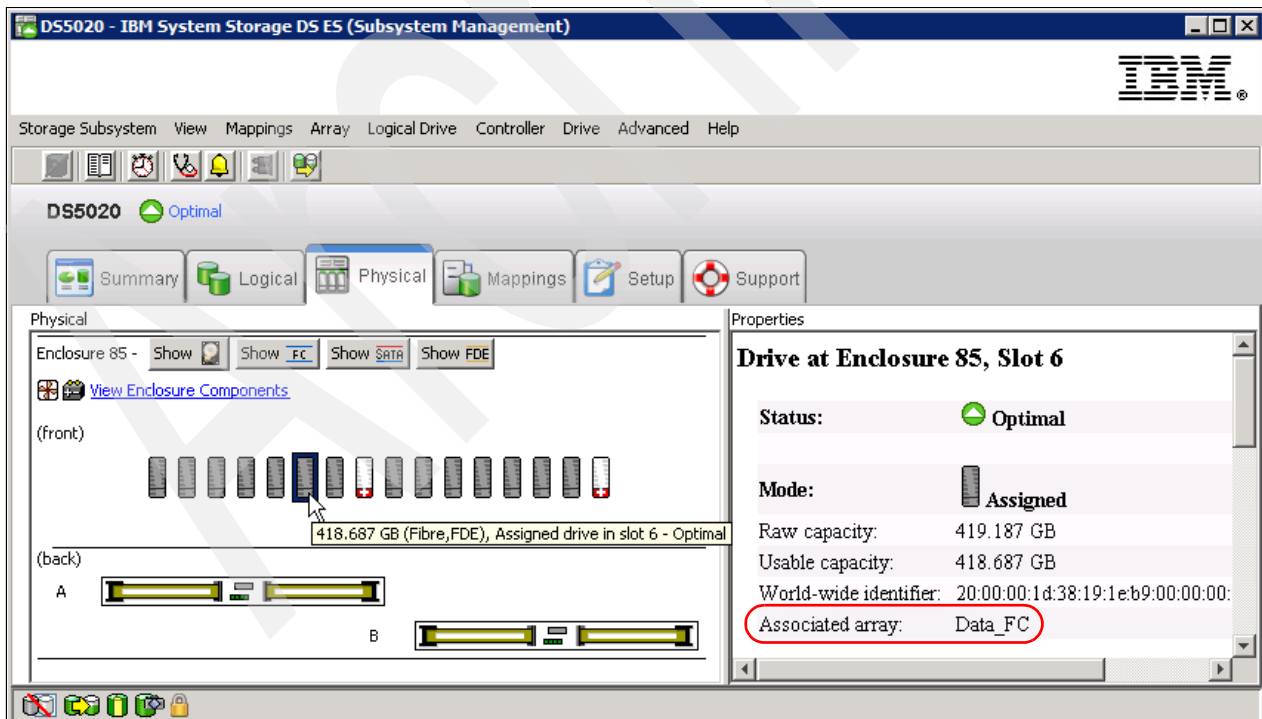


Figure 4-65 Physical/Logical relationship

4.9.5 Configuring storage partitioning

We explain the concept of storage partitioning in 4.1.4, “Storage partitioning” on page 116. Here we show an example of configuring partitions for a FC host. If you need a specific procedure for iSCSI attachment, see *IBM Midrange System Storage Implementation and Best Practices Guide*, SG24-6363.

Because heterogeneous hosts can be attached to the DS5000 storage subsystem, you need to configure storage partitioning for two reasons:

- ▶ Each host operating system requires slightly different settings on the DS5000 storage subsystem. You need to tell the storage system the host type to which it is attached.
- ▶ There is interference between the hosts if every host has access to every logical drive. By using storage partitioning and LUN masking, you ensure that each host or host group only has access to its assigned logical drives. You can have up to 256 LUNs assigned to a single storage partition. You might have a maximum of 2048 LUNs configured per DS5000 storage subsystem.

The overall process of defining the storage partitions is as follows:

1. Define host groups.
2. Define hosts.
3. Define host ports for each host.
4. Define storage partitions by assigning logical drives to the hosts or host groups.

The Subsystem Management has an specific view called *Mappings*. From this view, you can display the current status, create your hosts and hosts groups, and map the logical drives to them.

Selecting the **Mappings** view, if you have not defined any storage partitions, opens the Mapping Start-Up Help window (shown in Figure 4-66). The information in the window advises you to only create host groups if your plan includes sharing logical volumes across different hosts, normally a cluster.

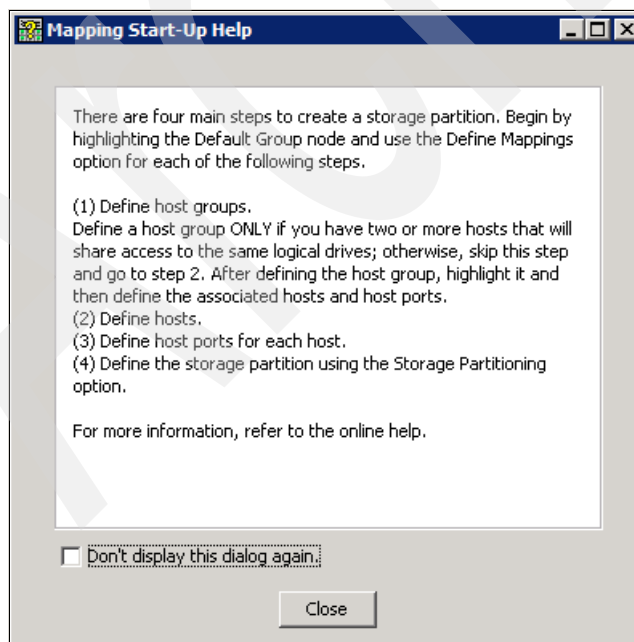


Figure 4-66 Mappings Start-Up Help

The Setup view has the option Map Logical Drives, which you can use for mapping your logical drives to the Default group, or to previously defined hosts or groups.

Figure 4-67 shows an example of the Mappings view. The right side of the window lists all mappings that are owned by the object you select on the left side. If you highlight the storage system, you see a list of all defined mappings. If you highlight a specific host group or host, its mappings are listed.

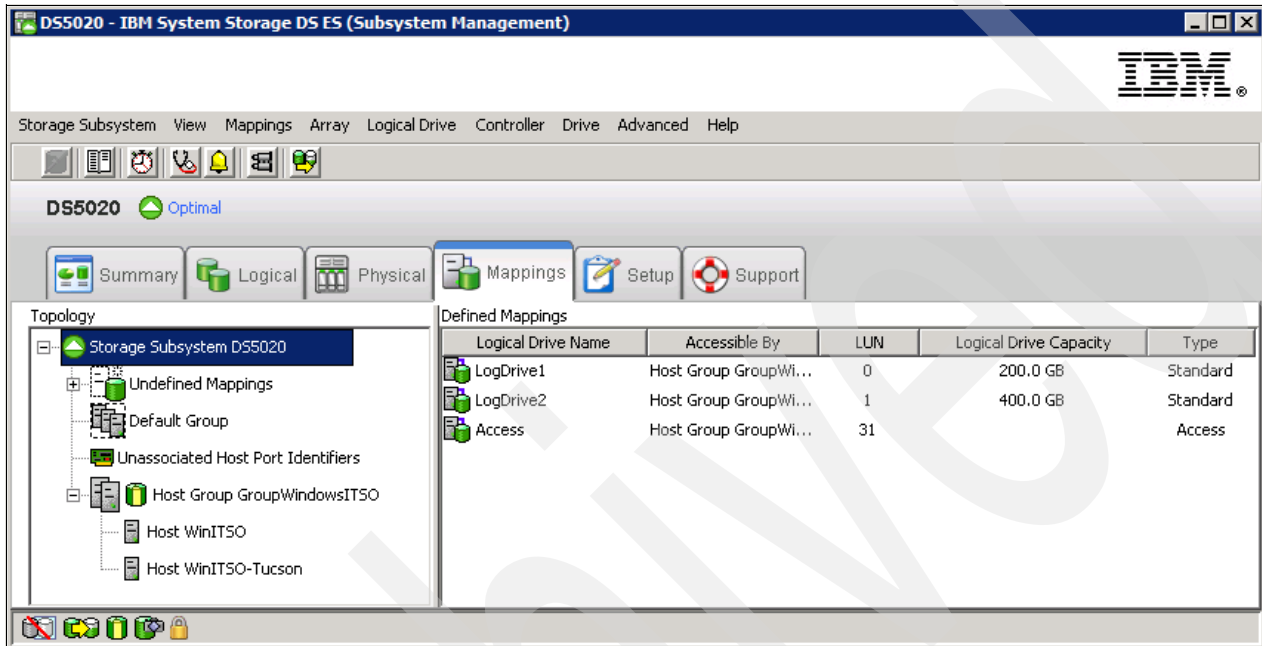


Figure 4-67 Mappings view in Subsystem Management window

In our example, we show how to configure the following storage partitioning scenario:

- ▶ We have a host with Windows Server 2008. We want that host to have access to the logical drives LogDrive1 and LogDrive2 of the storage subsystem.
- ▶ This host contains two FC HBAs.
- ▶ At some time, the host becomes a part of the Windows 2008 cluster, so we have to put it into a host group. The host group name will be WinGroupITSO.
- ▶ We map the two logical drives (LogDrive1 and LogDrive2) to that host group.
- ▶ We show how to modify the mapping to integrate it later into a cluster.

Note: Before configuring the partitions, if you have your storage connected to a FC switch, make sure to first define the zones appropriately.

Next, perform a hardware rescan on each host that is to be mapped to the DS5000 storage subsystem to reconfigure the FC devices and allow the WWN to be presented in the Storage Manager.

It is a best practice to configure only one host HBA per zone, together with one DS5000 storage subsystem controller.

Defining hosts

Perform the following steps:

1. Right-click your storage subsystem and select **Define** → **Host**, as shown in Figure 4-68. Even with no host group created, it is easier to create the hosts first and then the host groups.

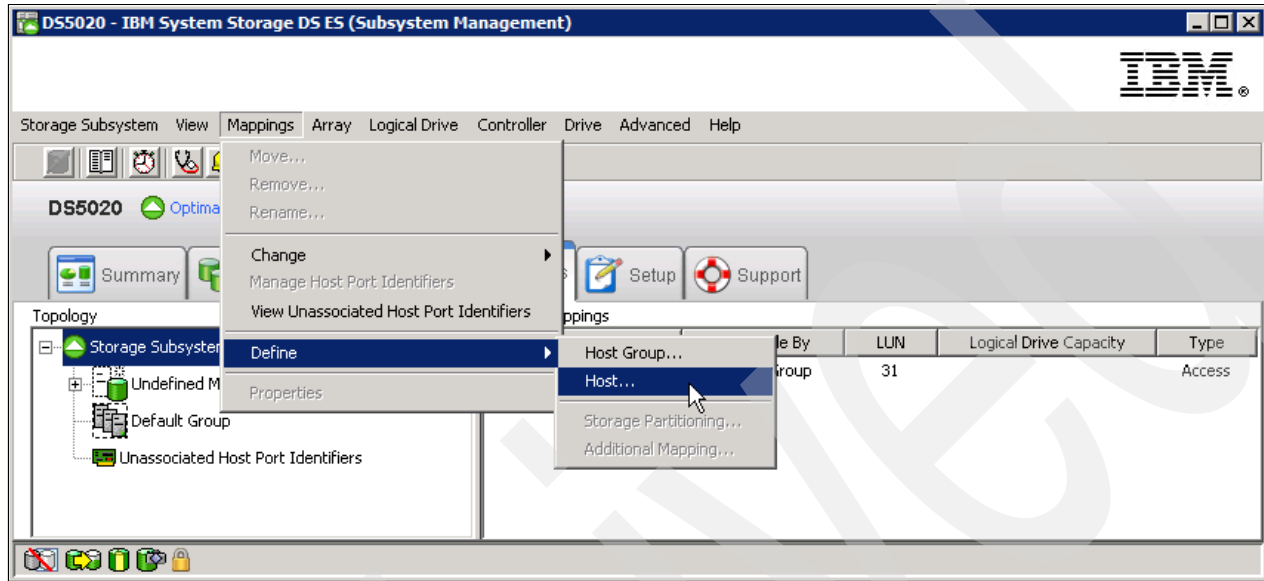


Figure 4-68 Selecting Define Host

This launches the wizard for defining a new host. You have to provide the following information during the process:

- Name of the new host.
- Protocol attachment method:
 - FC
 - iSCSI
- Host port identifiers with their Alias or Labels.
- Host type (the operating system that runs on the host).
- Whether the host is going to participate in a cluster.
- Host Group name if the host is defined as clustered.

- The first window of the wizard is an introduction (Figure 4-69). You are asked to assign a name to the host being defined, and whether you plan to use storage partitioning.

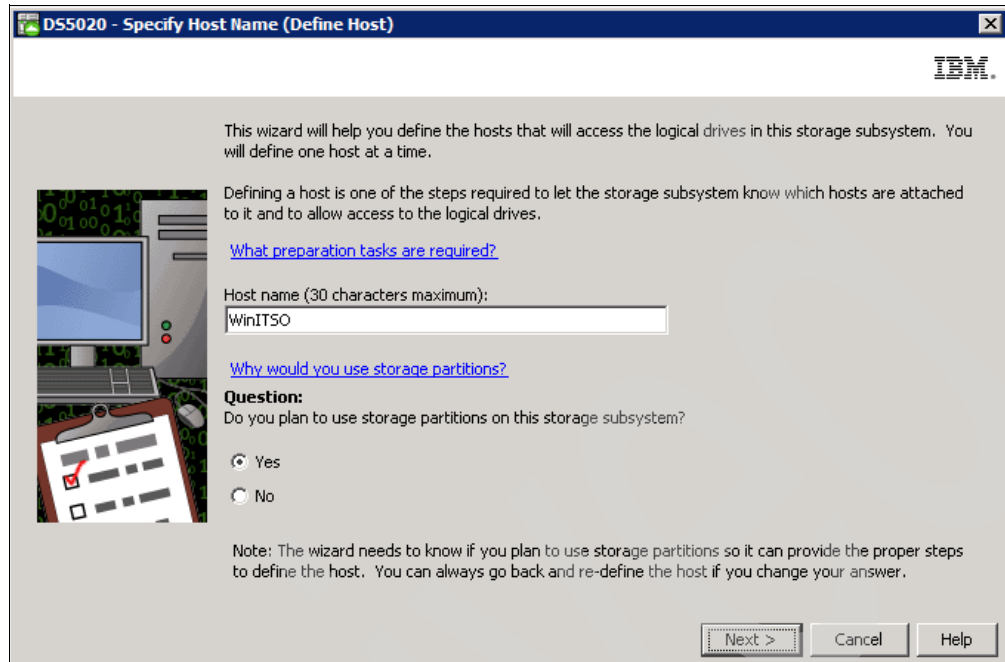


Figure 4-69 Define Host wizard

- If you only plan to attach a single host server, then you do not need storage partitioning. But if you plan group different host environments together, or if clusters sharing logical drives is recommended, click **Yes** to create different hosts groups at a later time.
- In the next window, you need the attachment protocol to specify the host interface type, FC or iSCSI, and the HBA host port information (host port identifier and alias). Remember that the HBA host port identifier is the world-wide port name of the particular HBA. To make sure that you know what your server's HBA WWPN is, use the SANsurfer management tool (which is discussed in 7.11.3, "Qlogic HBAs and SANsurfer (Windows/Linux)" on page 443), as shown in Figure 4-70.

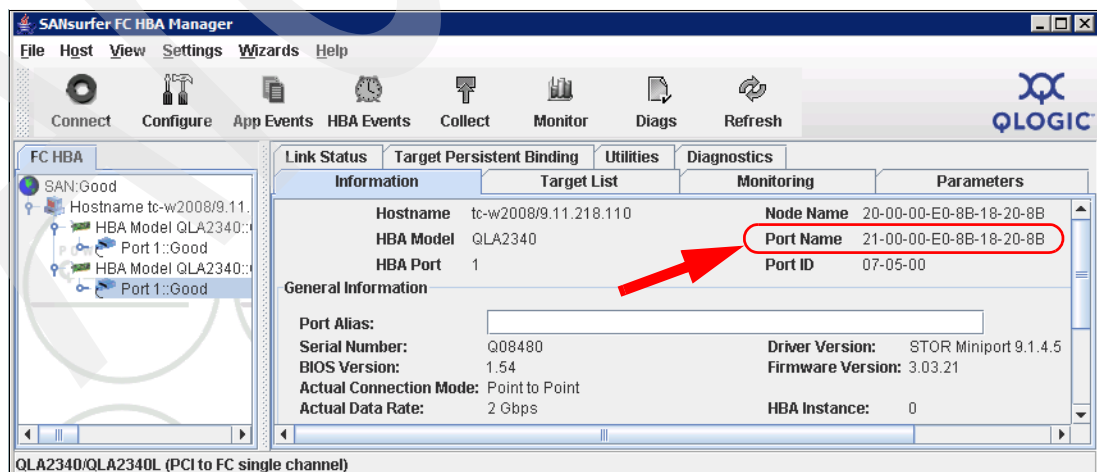


Figure 4-70 Displaying WWPN with SANsurfer

Note: The *IBM QLogic SANsurfer User's Guide* and the QLogic SANsurfer program are located on the IBM DS Storage Manager CD.

5. Enter the data from step 4 on page 204 into the window shown in Figure 4-71.

DS5020 - Specify Host Port Identifiers (Define Host)

The host communicates with the storage subsystem through its host bus adapters (HBAs) or its iSCSI initiators where each physical port has a unique host port identifier. In this step, select or create an identifier, give it an alias or user label, then add it to the list to be associated with host WinIT50.

[How do I match a host port identifier to a host?](#)

Choose a host interface type:
FC

Choose a method for adding a host port identifier to a host:

Add by selecting a known unassociated host port identifier

Known unassociated host port identifier:
21:00:00:e0:8b:89:2c:c0 Refresh

Add by creating a new host port identifier

New host port identifier (16 characters required):

Alias (30 characters maximum):

Add ? Remove ?

Host port identifiers to be associated with the host:

Host Port Identifier	Alias / User Label
21:00:00:e0:8b:18:20:8b	HBA-1
21:00:00:e0:8b:89:2c:c0	HBA-2

< Back Next > Cancel Help

Figure 4-71 Define Host: Specifying the host name and HBA

- Select your interface type (FC in our example) and an active port already detected by the DS subsystem, or type its world-wide port name directly into the field.
 - Type an alias for the specified host port identifier and click **Add**.
 - Repeat the same process until you define all the HBAs. Remember that if you define only one HBA, the host can lose access if there is a problem. Click **Next** after defining all the HBAs.
6. The next window requires you to specify the host type. This is basically the operating system running on the host. It is vital that you select the appropriate host type because the RDAC and ADT settings rely on this setting. In addition, this is the part of the configuration where you configure the heterogeneous host support. Each operating system expects slightly different settings and handles SCSI commands differently. If you make a wrong choice, your operating system might not boot anymore or path failover cannot be used.

We show an example of this window in Figure 4-72. In our particular case, we selected **Windows 2000/Server 2003/Server 2008 Non-Clustered**.

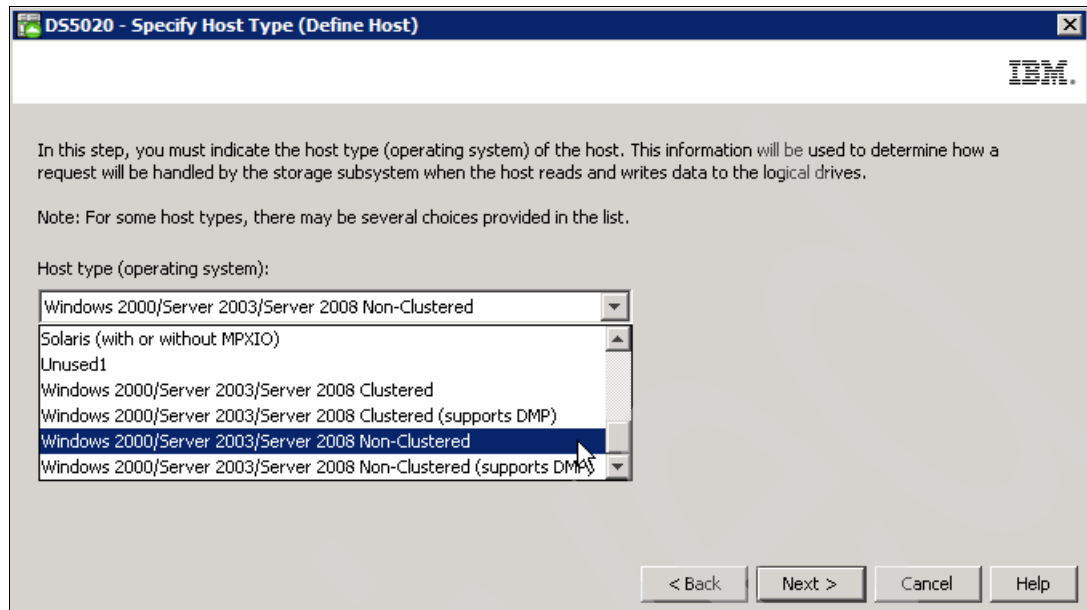


Figure 4-72 Define Host: Specifying the host type

7. In the next step, you are asked whether the host is a part of a cluster. Click **Yes** or **No**, depending your configuration, and **Next** to continue.

8. If the answer is **Yes**, as in our example, then you need to specify a host group. The host group can be either a new or an existing one, as shown in Figure 4-73.

DS5020 - Specify Host Group (Define Host)

IBM

[What is a host group?](#)

Because you specified on the previous screen that the host you are defining will share access to logical drives with one or more other hosts, you must indicate the name of the host group that this host will be associated with.

You can either (1) manually enter a new host group name or (2) select an existing host group. If you select an existing one, you will be shown the hosts currently associated with it.

Enter name (30 characters maximum)

GroupWindowsIT50

Select existing host group

-Select from list-

Associated hosts in host group:

Name	Host Type
------	-----------

< Back Next > Cancel Help

Figure 4-73 Define Host: Specifying a host group

9. Finally, you have the chance to preview the new host definition (Figure 4-74). If all the selections are correct, click **Finish** to define the new host.

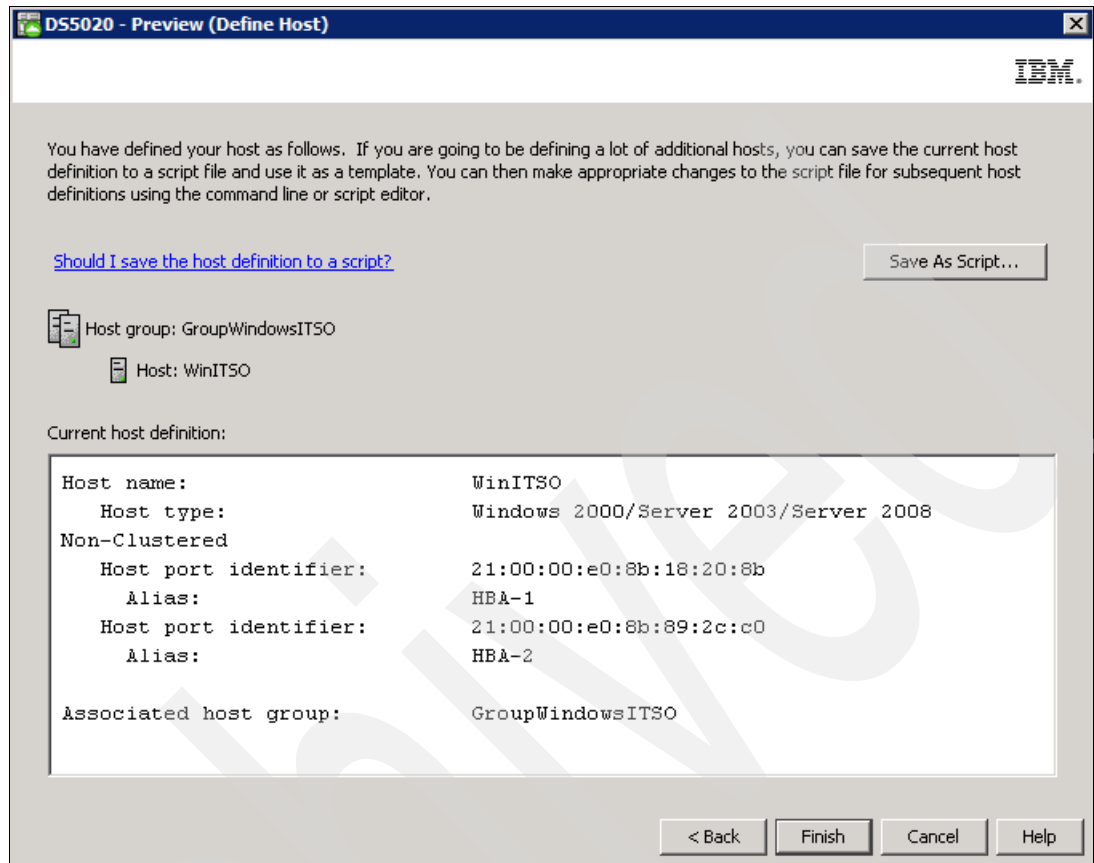


Figure 4-74 Define Host: Preview

10. Repeat the same steps to create more hosts and host groups. Once finished, the new host (and the host group, if it was also defined) is placed in the default group. It will stay there until you actually create a storage partition by assigning the logical drives to that host (or group). Figure 4-75 shows an example.

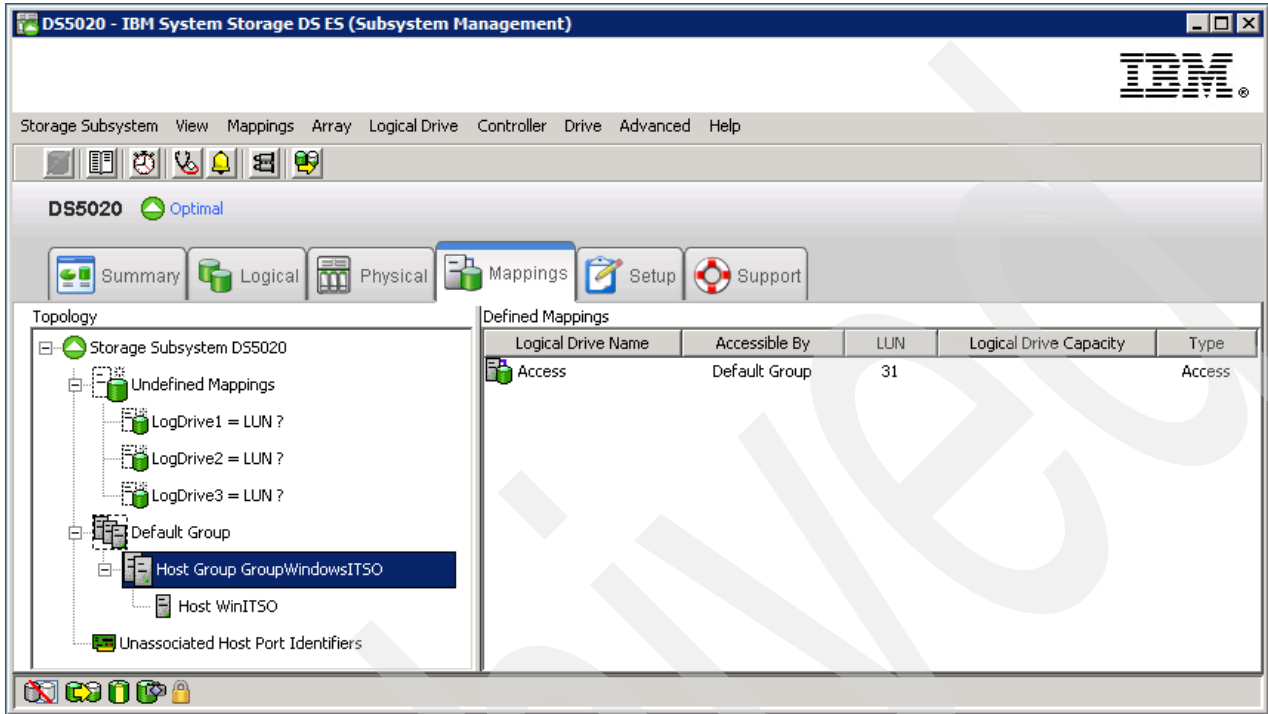


Figure 4-75 New host and host group placed in the default group

Defining storage partitioning

Next, we define storage partitioning:

1. We start by creating a storage partition by assigning the logical drives to the hosts or host groups. The Storage Partitioning wizard leads you through the process, and you initiate it by right-clicking **Default Group** and selecting **Define** → **Storage Partitioning**. We show an example in Figure 4-76.

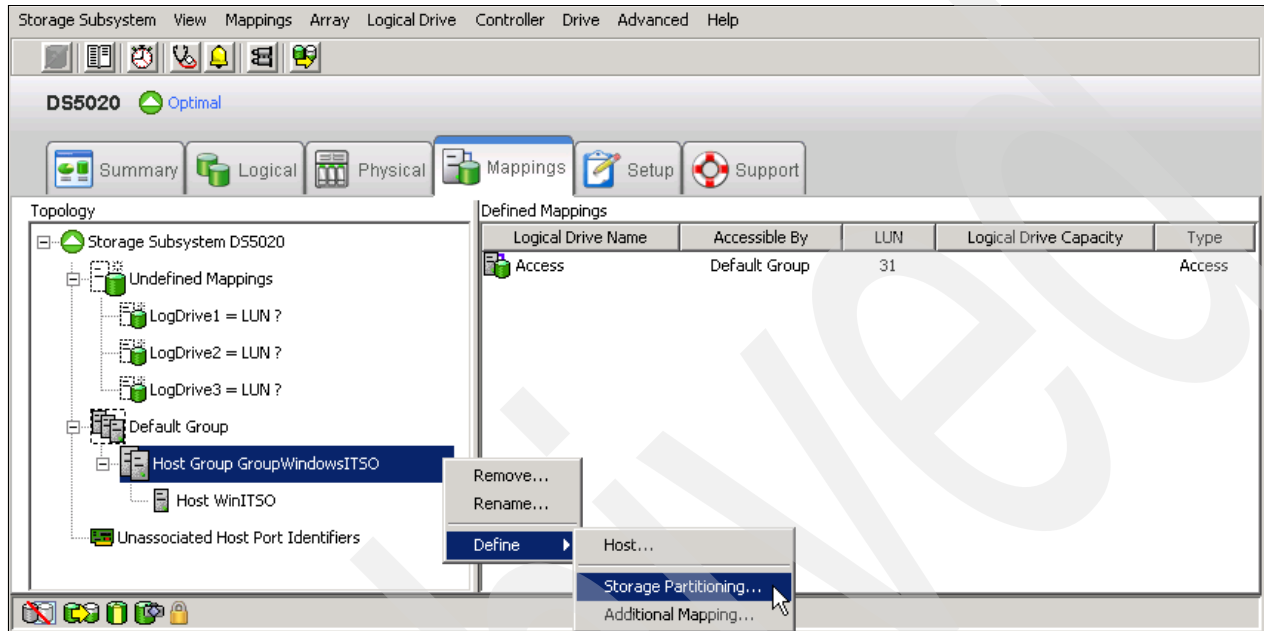


Figure 4-76 Define storage partitioning

2. After the introductory window, the wizard asks you to select either a host or a group of hosts. If you are creating a storage partition for clustered host servers, you need to specify the appropriate group; otherwise, you can select an individual host.

- The next window allows you to select the logical drives that are going to be mapped to the host or the group. You also have to specify a LUN for each logical drive. In our example, shown in Figure 4-77, we selected WindowsLD1 and WindowsLD2. We assigned the LUNs 0 and 1, respectively, to these two logical drives.

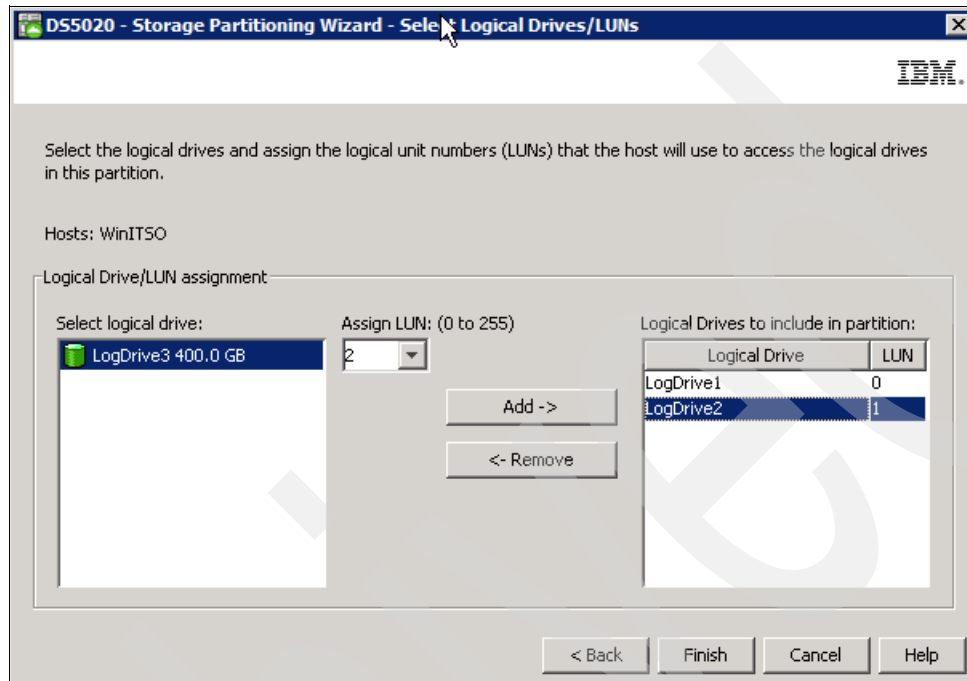


Figure 4-77 Storage Partitioning wizard: Selecting logical drives/LUNs

- Click **Finish** when you are done with selecting the logical drives and assigning the LUNs.
- Display your newly defined host groups, host, and mappings by selecting the **Mappings** view in the Subsystem management windows, as shown in Figure 4-77.

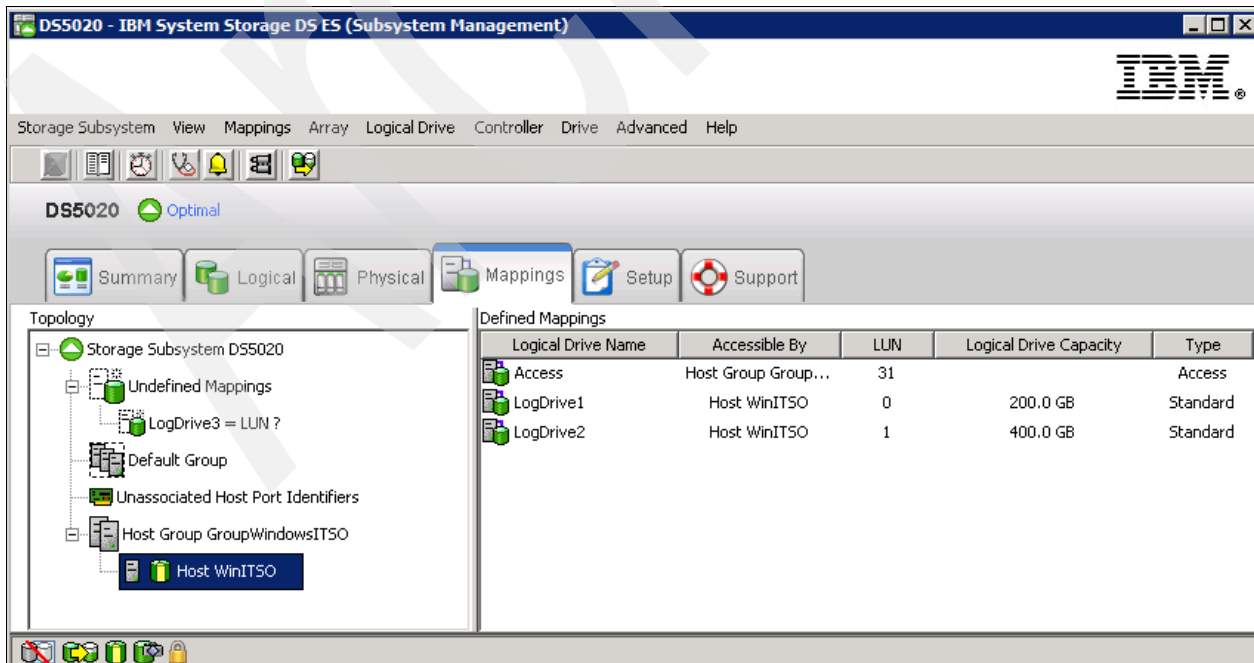


Figure 4-78 Displaying created mappings

Notice that even if the host WinITSO is defined under a host group, the logical drives are mapped exclusively to the host WinITSO. If later you decide to incorporate another host, and make it work together with the first host as a cluster, perform the following steps:

1. Add the second host by performing the steps shown in “Defining hosts” on page 203. Specify the host type and select **Will participate in a cluster**, and incorporate it into the existing host group where your other host is already defined.
2. Change the previous host type from Windows 2008 Nonclustered to Clustered.
3. Reassign the mappings of logical drives. Now we want to map them to the host group, not to an specific host, so all the hosts belonging to the group can reach all the logical volumes.

After performing these steps, the Mappings view appears as shown in Figure 4-79, showing the correct configuration for a cluster environment.

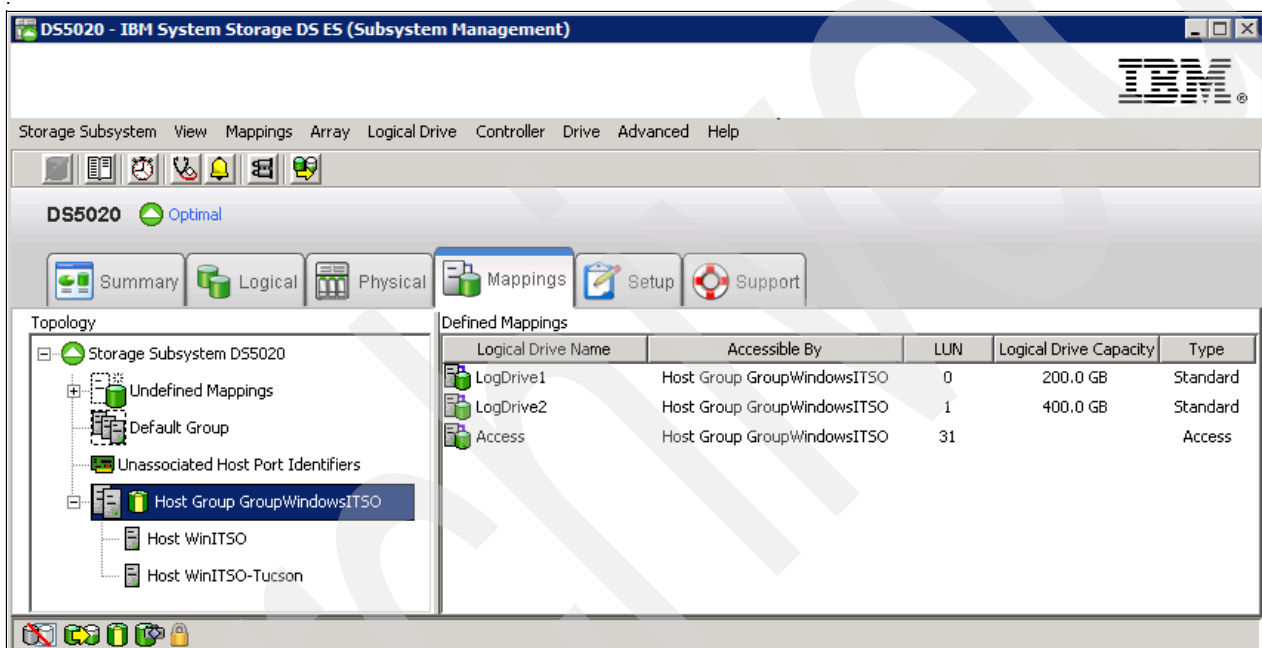


Figure 4-79 Host group mappings for clustering

You can continue creating other partitions or host groups, you should know that even though you can create multiple partitions, you can only map logical drives to partitions up to the maximum number allowed by your specific system and premium feature configuration. Check your storage subsystem for the allowed number of partitions by selecting, in the Subsystem Management view, **Storage Subsystem** → **Premium Features**, and review the information presented, as shown in Figure 4-80.

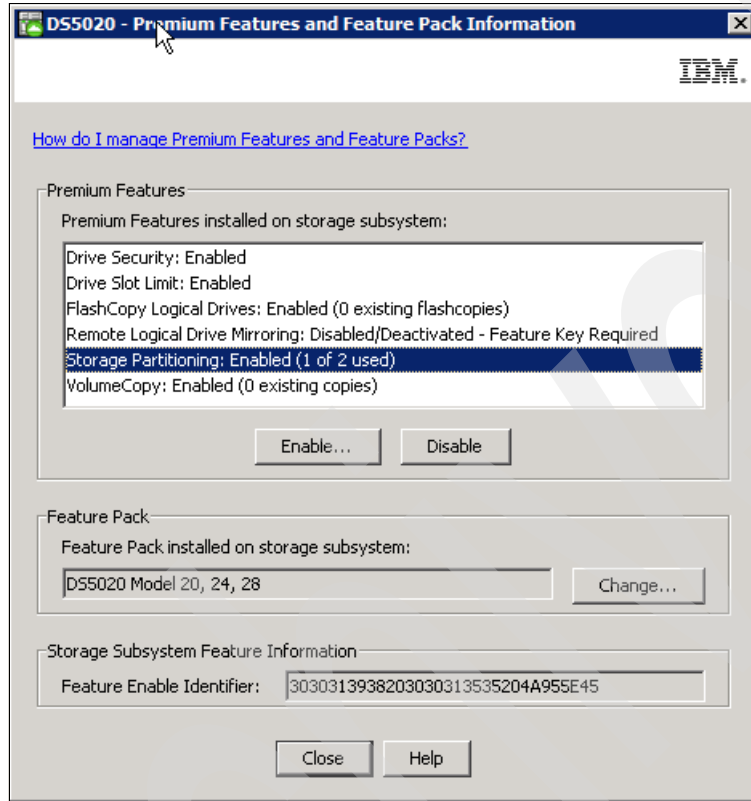


Figure 4-80 Partitions allowed

Define Host Group option

Because the host groups can be created with the Define Host wizard, there is usually no need to define the groups from outside the wizard. But you still have the option to do so by performing the following steps:

1. Right-click **Default Group** and select **Define** → **Host Group**, as shown in Figure 4-81.

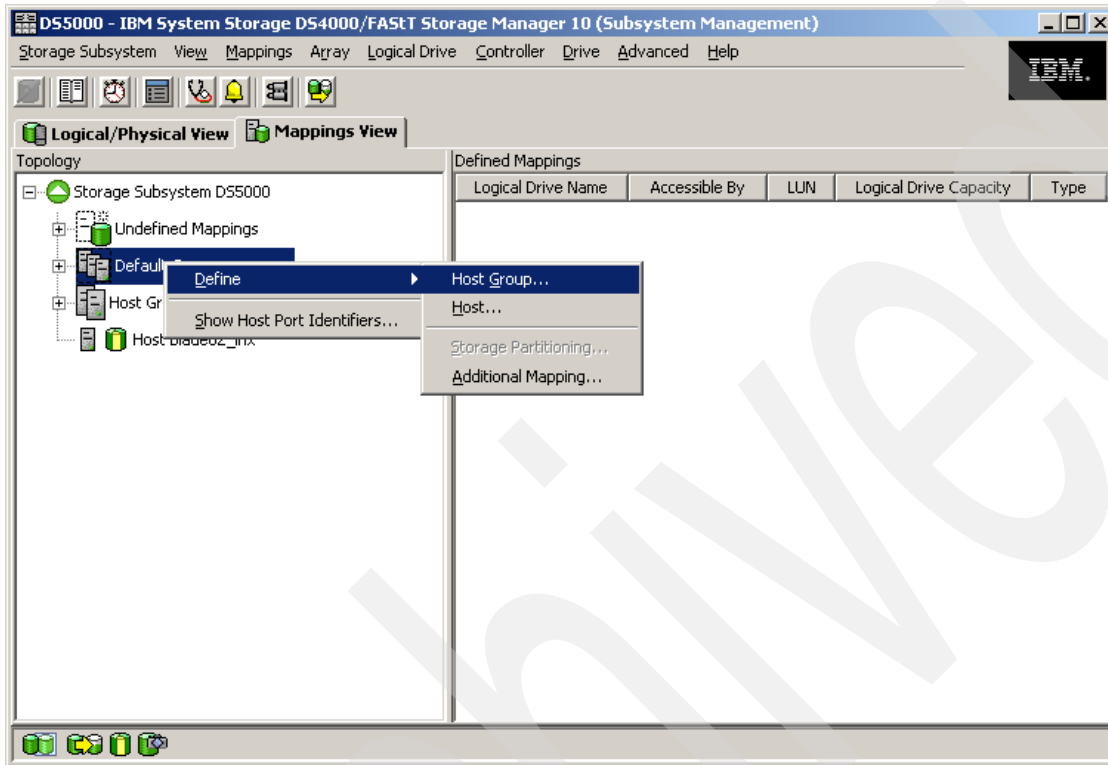


Figure 4-81 Define Host Group

2. The Define Host Group window (shown in Figure 4-82) opens. Enter the name of the host group you want to add. Select every host you want to add to the group and click **Add**. When you are done, you can exit the dialog by clicking **OK**.

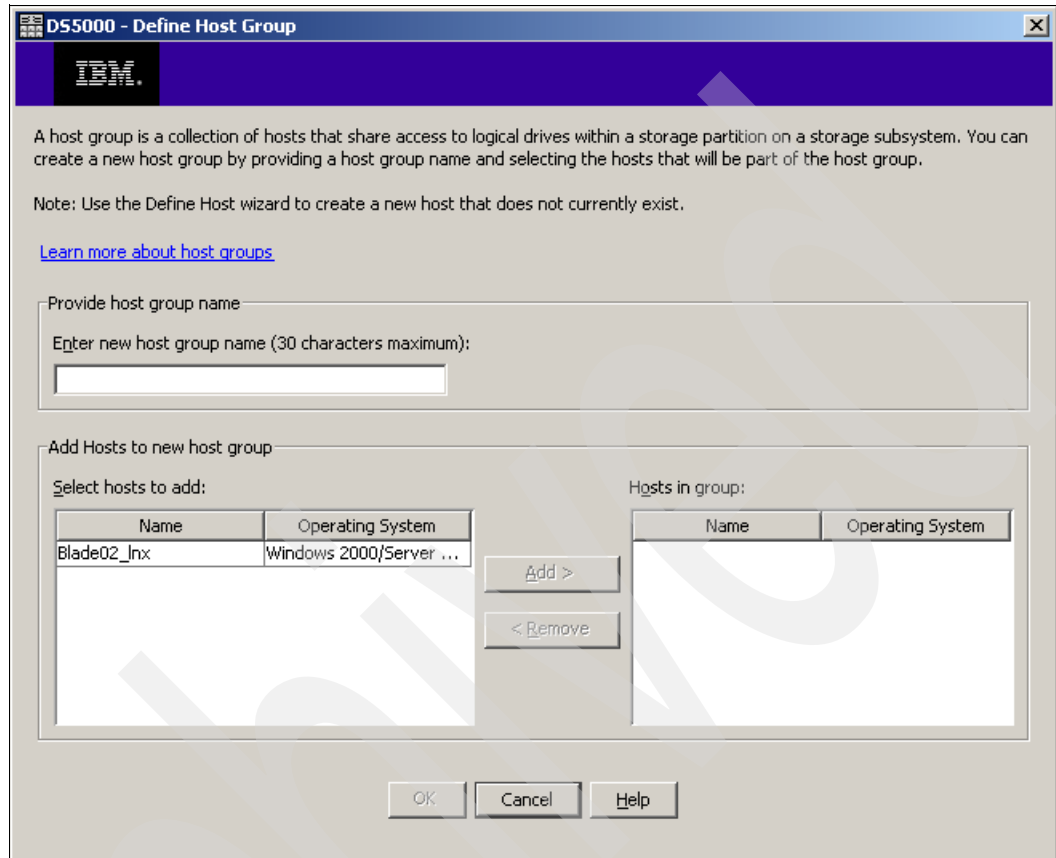


Figure 4-82 Define Host Group name

3. If you accidentally assign a host to the wrong host group, you can move the host to another group. Simply right-click the host and select **Move**. A window opens and prompts you to specify the host group name.
4. Because storage partitioning of the DS5000 storage subsystem is based on the world wide names of the host ports, the definitions for the host groups and the hosts only represent a view of the physical and logical setup of your fabric. When this structure is available, it is much easier to identify which host ports are allowed to see the same logical drives and which are in different storage partitions.

Manage Host Port Identifiers

For environments with multiple hosts and attachment types, you can use this option to have a single source of information about the different available host ports.

From the Subsystem Management window, select **Mappings** → **Manage Host Port Identifiers**. This option opens a window similar to Figure 4-83.

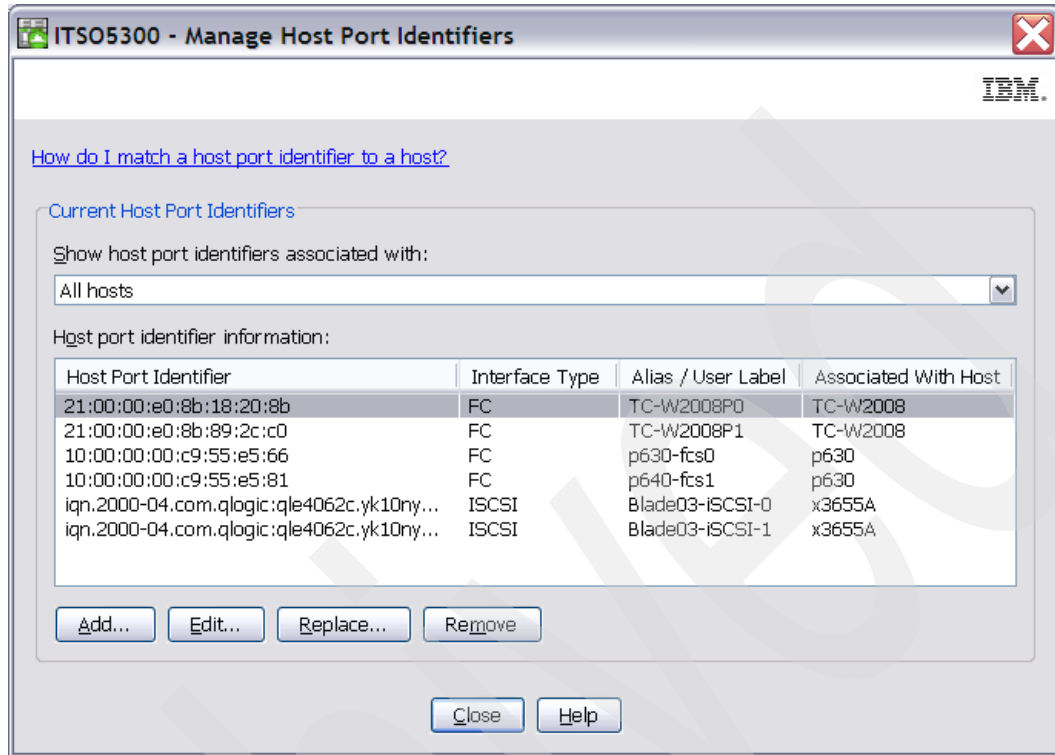


Figure 4-83 Manage Host Port Identifiers

Select this option if you need to review your port configuration assignment, add, remove, and change port settings, and whenever you need to replace a host interface card after a hardware replacement to continue presenting the logical volumes to the new host port identifier, WWPN, or iSCSI Initiator.

Define Additional Mapping option

Suppose that a particular host (or a host group) is already a part of a certain storage partition. This means the logical drives are already mapped to that host or group. If you need to map additional logical drives to the same host or group, use the Define Additional Mapping option:

1. Right-click the host or group to which you want to map a new logical drive. Select **Define** → **Additional Mapping**, as shown in Figure 4-84.

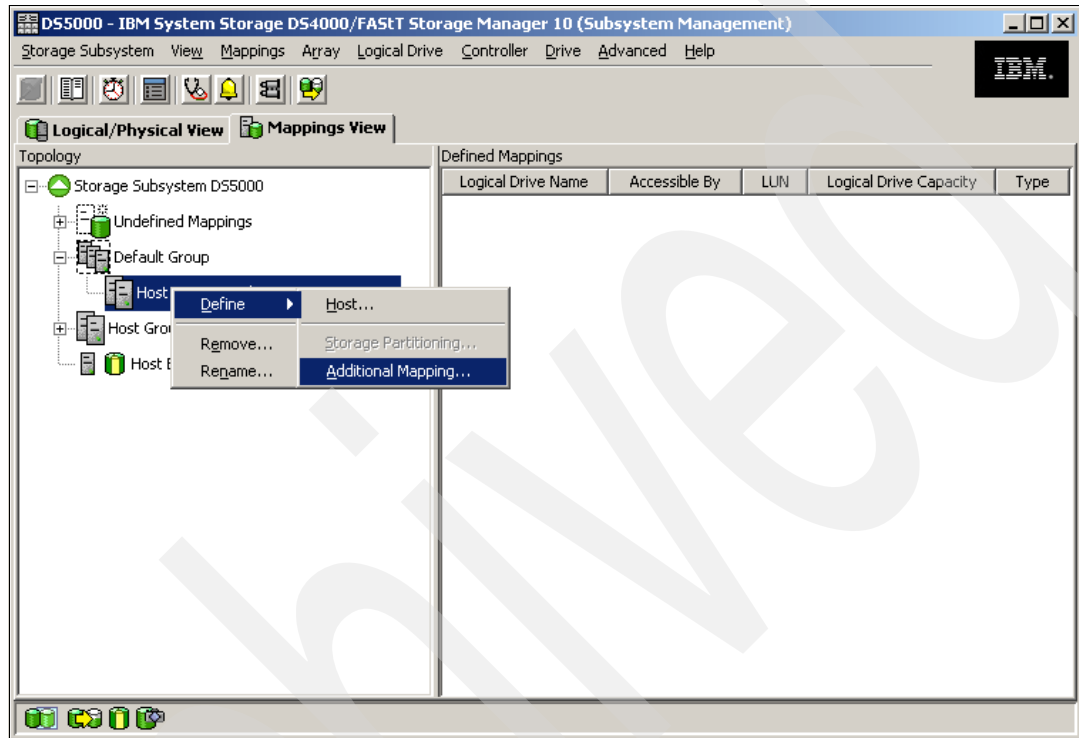


Figure 4-84 Define Additional Mapping

2. In the Define Additional Mapping window, select the logical drive you want to map to this host group or host and assign the correct LUN number, as shown in Figure 4-85.

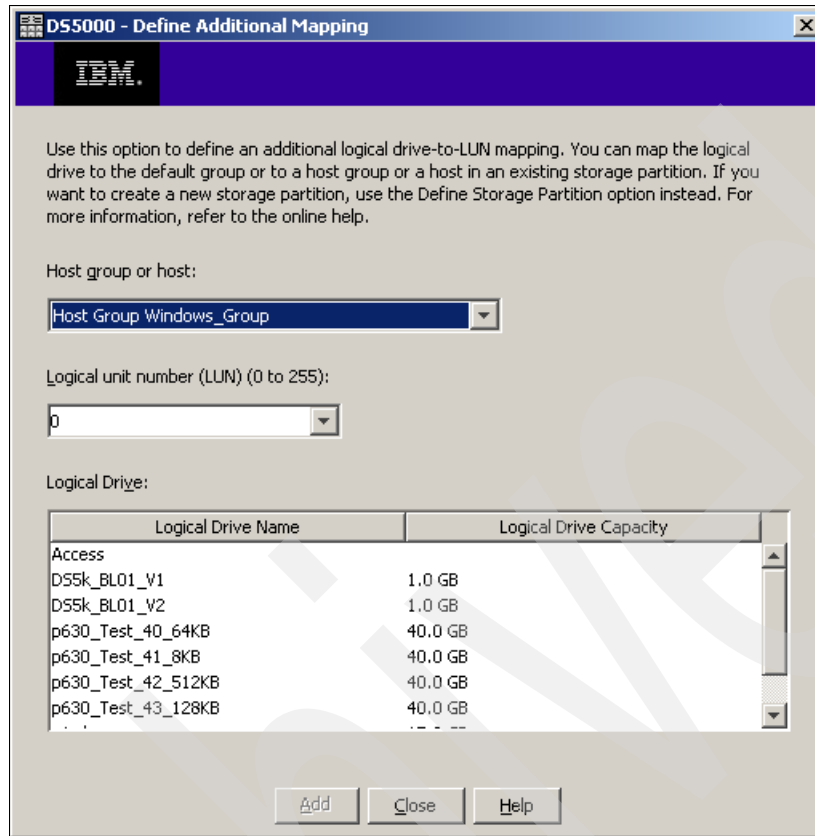


Figure 4-85 Define Additional Mapping

Note: If you change an existing mapping of a logical drive, the change is effective immediately. Therefore, make sure that this logical drive is not in use or even assigned by any of the machines attached to the storage system.

3. To make the logical drives available to the host systems without rebooting, the Storage Manager Utilities package provides the **hot_add** command-line tool (for some operating systems). You simply run **hot_add**, all host bus adapters are rescanned for new devices, and the devices are assigned within the operating system.
 - You might have to take appropriate steps to enable the use of the storage inside the operating system, such as formatting the disks with a file system and mounting them.
 - If you attached a Linux system to the DS5000 storage subsystem, you need to remove the mapping of the access logical drive. Highlight the host or host group containing the Linux system in the Mappings View. In the right part of the window, you see the list of all logical drives mapped to this host or host group. To remove the mapping of the access logical drive, right-click it and choose **Remove Mapping**. The mapping of the access logical drive is removed immediately.

4.9.6 Configuring mapped drives from Windows

To check the new mapped logical drives from the Windows Server 2008 used in this example, start with the SANsurfer management tool. Refresh it so that the HBAs search for all the new mapped drives. After that action completes, you see the information shown in Figure 4-86.

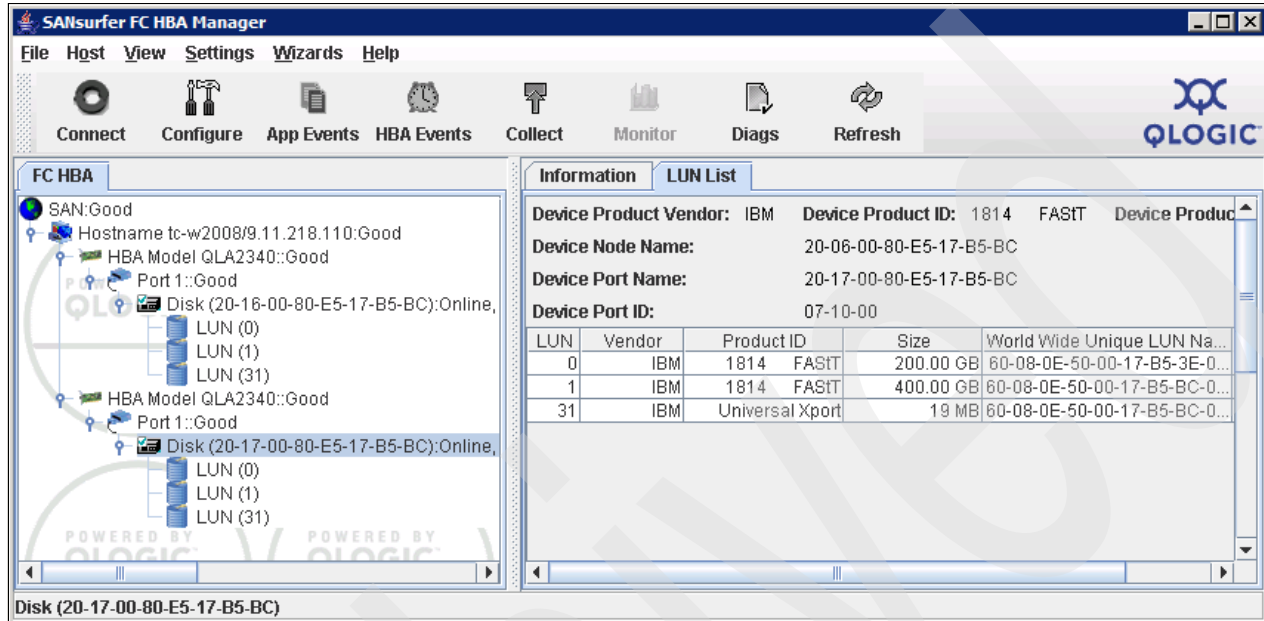


Figure 4-86 Displaying mapped logical drives

The verified HBAs show all the mapped LUNs, so you can use the Windows Disk Administrator utility to scan for new storage, create and format partitions, and begin using the new disk space.

First, use the Windows Device Manager to scan for new hardware changes. Once refreshed, you see the newly mapped logical drives from your DS5000 storage subsystem under the Disk drives folder of the server.

Each logical drive is presented as IBM xxx Multipath Disk Device under the Disk drives folder, where xxx is the product ID of the DS Storage subsystem, as shown in Figure 4-87.

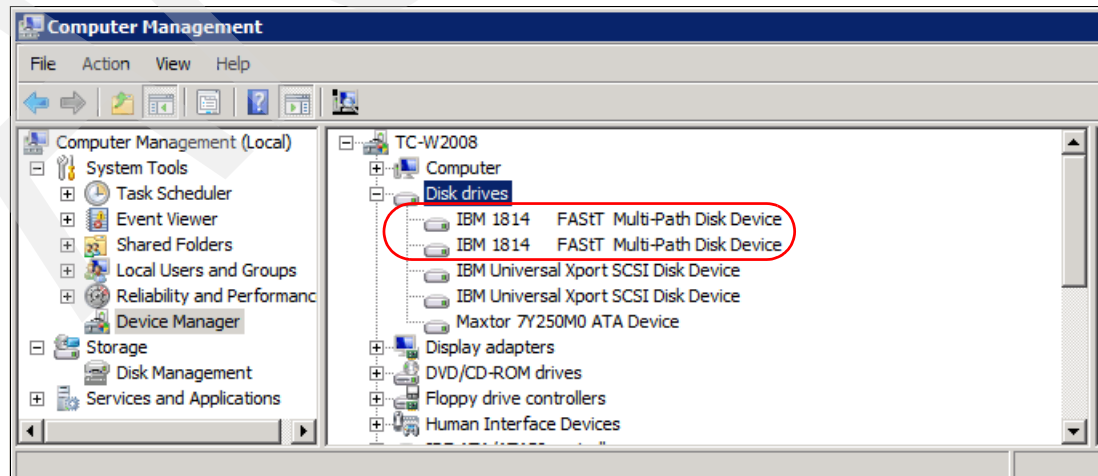


Figure 4-87 Verifying new disks in Device Manager

With MPIO, each device listed here represents a logical drive. In order to display the different paths of each one, select it and then choose **Properties**. Select the **MPIO** tab to display the different paths to the DS5000 storage subsystem. From the same Properties window, you can set the different Load Balance policies for the MPIO driver. Depending on the current physical path state, and the policy selected, the paths are both Active (round robin, weighted paths), or one Active and the other in Standby (failover, Least queue depth), as shown in Figure 4-88.

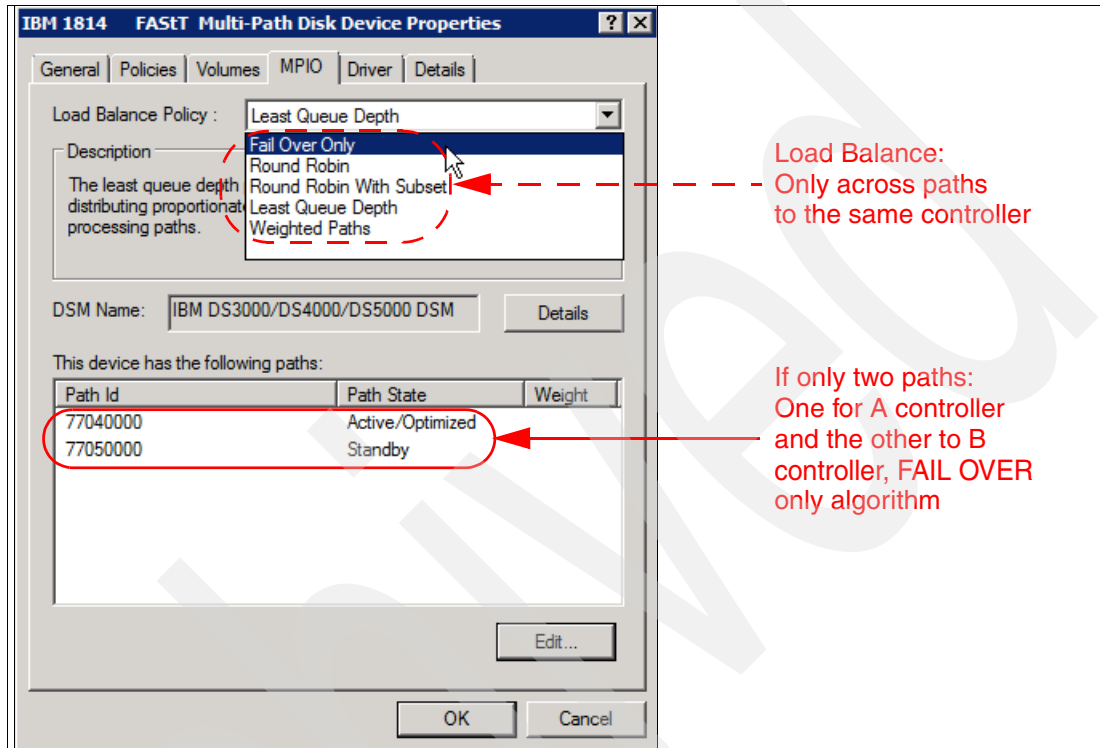


Figure 4-88 MPIO load balancing policy

Important: The Load Balance algorithms only work between multiple paths to the same controller. If you only have one path, you can only have failover capability.

For a correct data assignment to your defined logical volumes, you need to determine which of the disks drives in the Windows Device Manager or Disk Manager relates to the previously defined logical volumes of your DS5000 storage subsystem.

Select the disk drive, right-click it, and select **Properties**. Use the LUN number to isolate each logical drive, as shown in Figure 4-89.

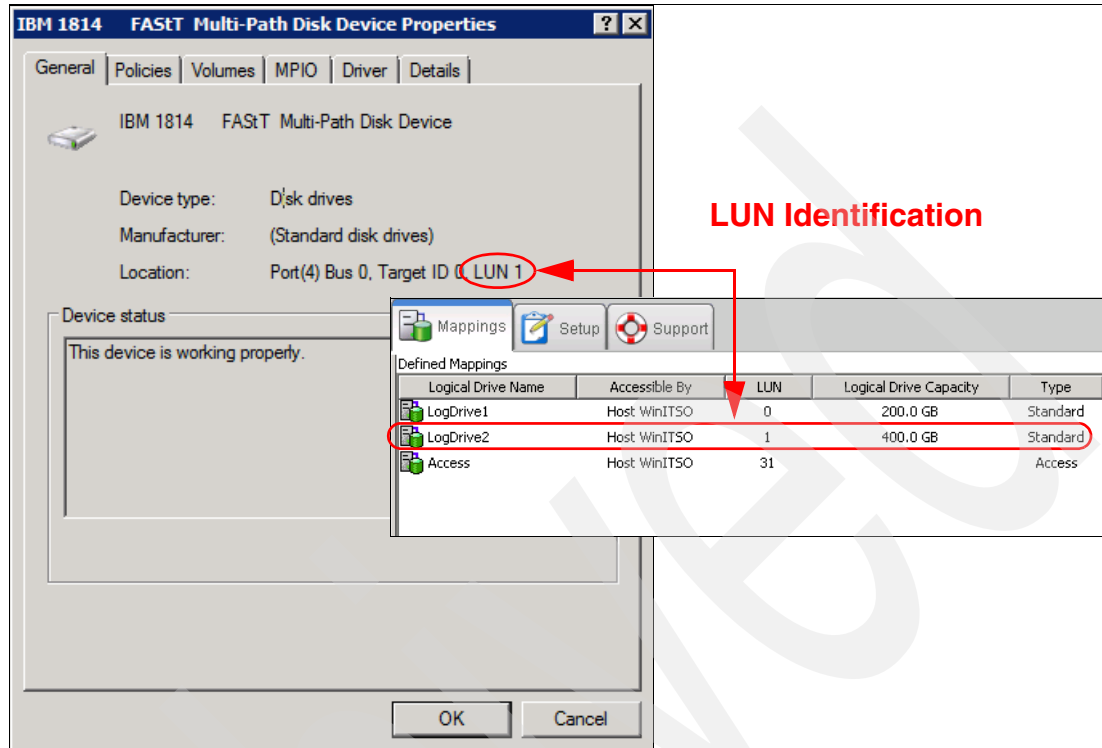


Figure 4-89 Matching LUNs between a Windows host and SM

Finally, use the Windows Disk Management to start using your new mapped DS5000 storage subsystem disks, as shown in Figure 4-90. You can get the information shown in Figure 4-89 by clicking each of the drives and selecting **Properties**.

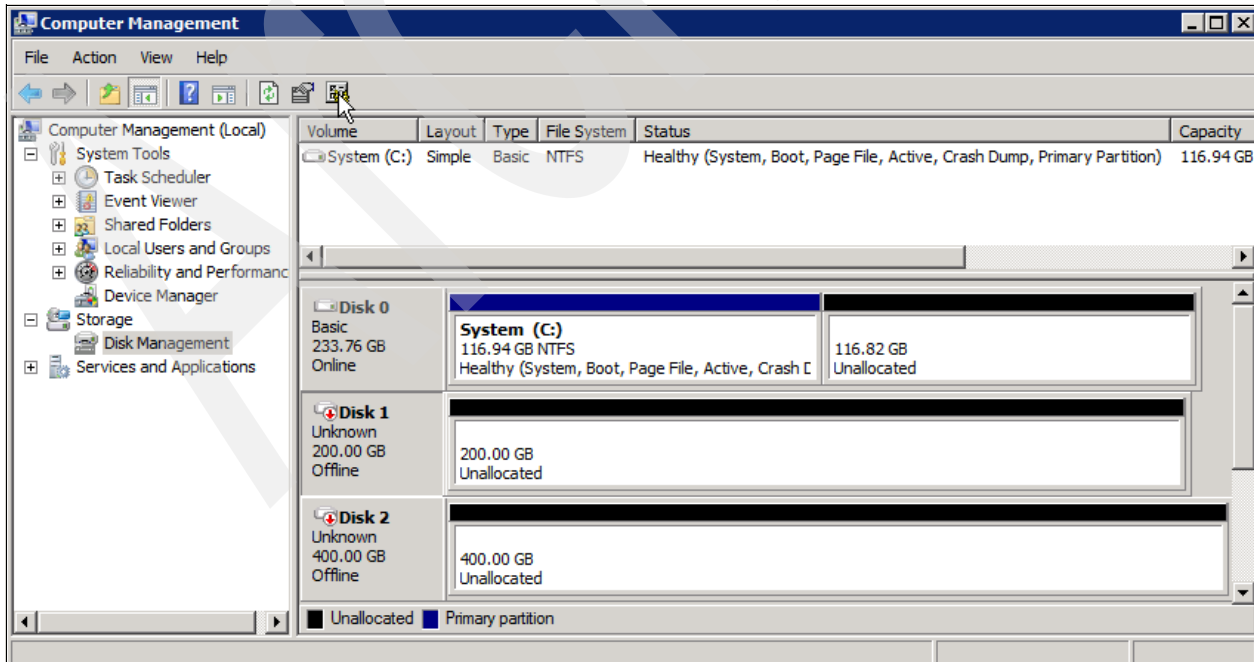


Figure 4-90 Windows Disk Management

The disks drives are shown as Offline. To start using one, right-click it, select **Online**, and then **Initialize**. After that, click the **Unallocated** space to partition the disk and assign a drive letter.

Storage Manager provides another utility to display the logical drives mapped. You can access this utility by opening your Storage Manager installation directory and run the SMdevices utility from a command prompt. Once you run the utility, you will see the output shown in Example 4-1.

Example 4-1 SMdevices output

```
C:\Program Files\IBM_DS\util>SMdevices
DS Storage Manager Utilities Version 10.01.35.01
Built Fri Jun 19 06:12:18 CDT 2009
IBM System Storage DS Storage Manager(Enterprise Management)
(C) Copyright International Business Machines Corporation, 2003-2009 Licensed Ma
terial - Program Property of IBM.
All rights reserved.
US Government Users Restricted Rights - Use, Duplication, or disclosure restrict
ed by GSA ADP Schedule Contract with IBM Corp.
IBM Support URL=http://www.ibm.com/servers/storage/support/disk/

  \\.\PHYSICALDRIVE1 [Storage Subsystem DS5020, Logical Drive LogDrive1, LUN 0,
Logical Drive ID <60080e500017b53e000048364a9fb470>, Preferred Path (Controller-
A): In Use]
  \\.\PHYSICALDRIVE2 [Storage Subsystem DS5020, Logical Drive LogDrive2, LUN 1,
Logical Drive ID <60080e500017b5bc000044e34a9fb474>, Preferred Path (Controller-
B): In Use]
  \\.\SYMsmUTMLun0 [Storage Subsystem DS5020, Logical Drive Access, LUN 31, Logi
cal Drive ID <60080e500017b5bc000043834a955f8a>]
  \\.\SYMsmUTMLun1 [Storage Subsystem DS5020, Logical Drive Access, LUN 31, Logi
cal Drive ID <60080e500017b5bc000043834a955f8a>]
```

Note that so far we updated the storage system to the latest level, configured logical drives, set up storage partitioning, and configured you host system. The next step is to define the alerting methods to be used in case of failures.

For more information about drive mapping in other operating systems, see *IBM Midrange System Storage Implementation and Best Practices Guide*, SG24-6363.

4.9.7 Monitoring and alerting

The Event Monitor program is included in the Storage Manager Client package. It enables the host running this monitor to send out alerts (through e-mail (SMTP) or traps (SNMP)) about any of the DS5000 storage subsystems in your environment.

Earlier versions of Storage Manager include the Event Monitor, where the Enterprise Management window had to remain open to monitor the storage subsystems and receive alerts. Now you will receive an alert either from the SMclient, if it is opened, or from the Event Monitor program.

For continuous monitoring, install the Event Monitor on a host computer that runs 24 hours a day. If you choose not to install the Event Monitor, you should still configure alerts on the host computer where the client software is installed, but it will only work when the SMclient is opened. The installed server should be capable of out-of-band and in-band management. This ensures proper alerting, even if one server is down, or a connection type has failed.

Important: The DS5000 storage subsystem does not send the e-mail or SNMP trap itself. The management station running the event monitor service sends the e-mail or SNMP trap on behalf of the DS5000 storage subsystem. If the management station is down, or if the event monitor process is not running, no notifications will be sent.

The Enterprise Storage Management Task Assistant lets you configure alerts with the following options:

- ▶ All storage subsystems in the management domain
- ▶ An individual storage subsystems
- ▶ All storage subsystems managed through an specific host

The steps are:

1. Make sure the SMclient package has been installed in the workstation you configure to send alerts.
2. Decide which storage systems you want to monitor. You can set up alert-notification destination addresses where you will be notified. Use the Enterprise Task or the following options to configure the alerts.

3. If you right-click your local system in the Enterprise Management window (at the top of the tree) and choose **Configure Alerts**, this applies to all storage systems listed in the Enterprise Management window. If you want to monitor only one subsystem, then click that particular one and select **Edit** → **Configure Alerts**. You can also select the **Setup** tab of the Enterprise Management window, and then select the **Configure Alerts** option. You will be prompted about whether you want to configure alerts for all storage subsystems or any one in particular, as shown in Figure 4-91.

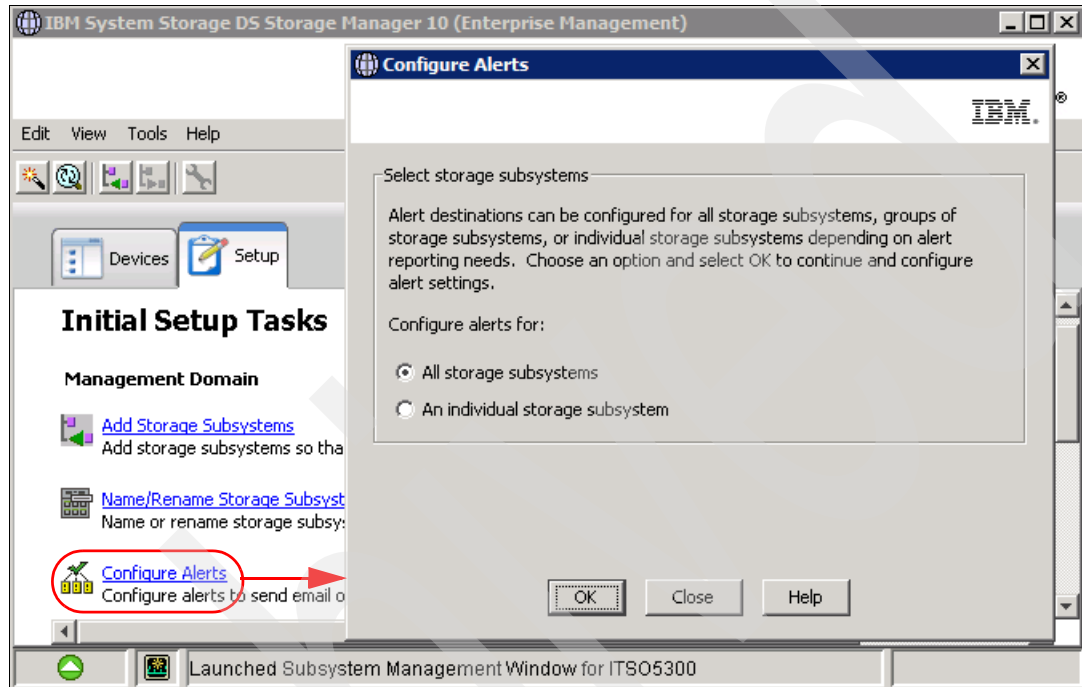
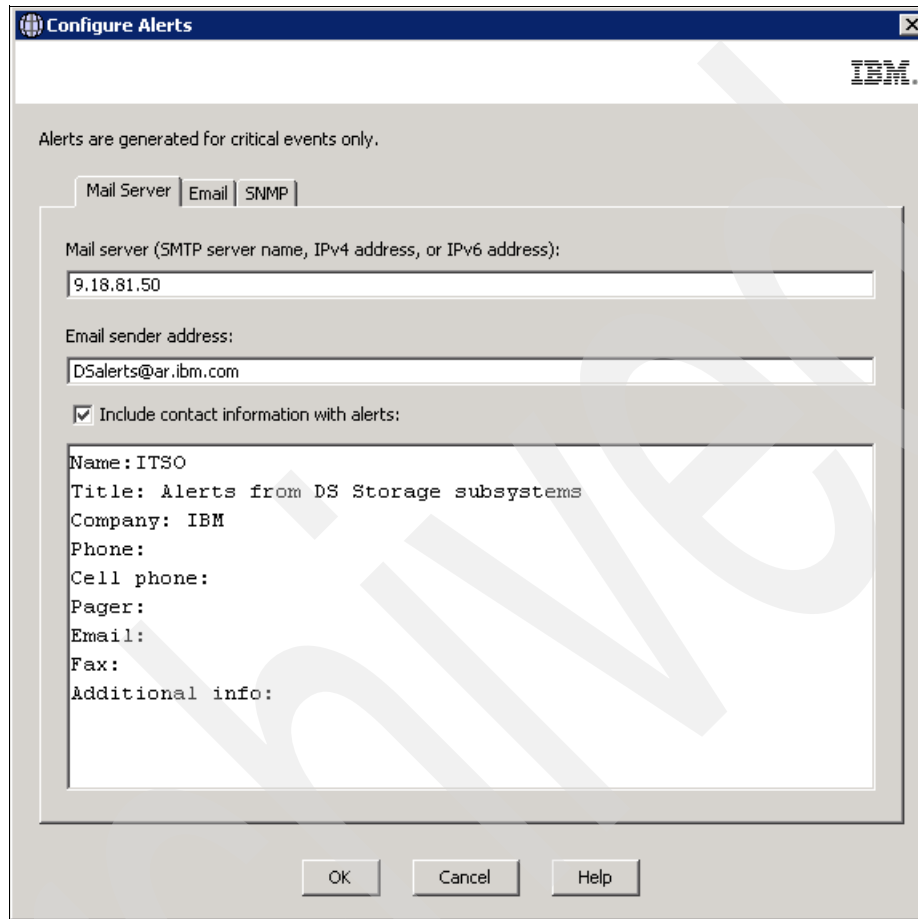


Figure 4-91 Selecting the Configure Alerts option

4. If you want to send e-mail alerts, you have to define an SMTP server first. Enter the Mail SMTP Server name or IP address of your mail server and the e-mail that will show as the origin of the message, as shown in Figure 4-92. You can also complete the contact information, if desired.



The screenshot shows a window titled "Configure Alerts" with an IBM logo in the top right corner. Below the title bar, there is a message: "Alerts are generated for critical events only." Below this, there are three tabs: "Mail Server", "Email", and "SNMP". The "Mail Server" tab is selected. Inside this tab, there are three text input fields: "Mail server (SMTP server name, IPv4 address, or IPv6 address):" containing "9.18.81.50", "Email sender address:" containing "DSalerts@ar.ibm.com", and a checked checkbox labeled "Include contact information with alerts:". Below the checkbox is a text area containing the following text: "Name: ITSO", "Title: Alerts from DS Storage subsystems", "Company: IBM", "Phone:", "Cell phone:", "Pager:", "Email:", "Fax:", and "Additional info:". At the bottom of the dialog box are three buttons: "OK", "Cancel", and "Help".

Figure 4-92 Configure Alerts: Defining the SMTP server

5. Select the **Email** tab. This tab allows you to configure the e-mail addresses to which the alerts are sent. Enter the e-mail address and press **Add** to append it to the list of notifications.

6. The default notification sends only the event when it occurs.

Note that you can also modify the Information to Send field to forward profile data or support data, either when an event occurs or at regular intervals. Make use of this feature whenever possible, as shown in Figure 4-93.

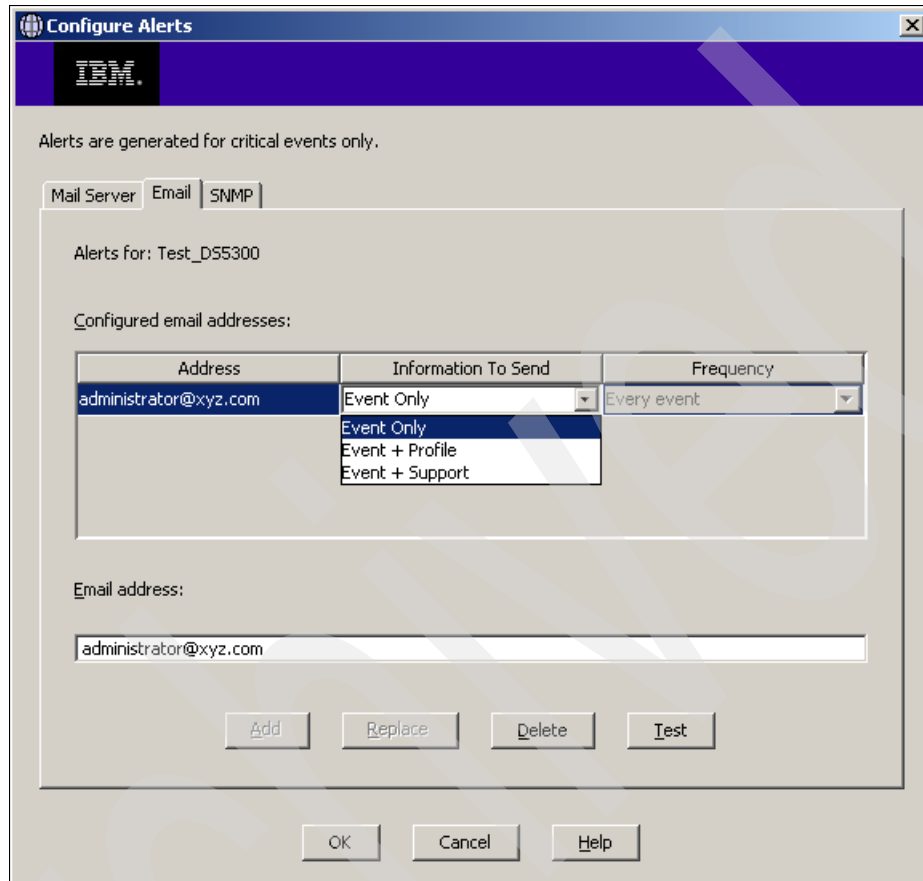


Figure 4-93 E-mail notification setup

Be aware that sending the support data at regular frequencies might impact the capacity of the destination e-mail address box.

Note: Besides the e-mail notification with error capture attached, there is another option to collect diagnostic data when an error is detected, that is, by enabling automatic collection of support data. See “Gathering support information” on page 231 for more information.

7. When you have finished adding e-mail addresses to notify, test your specified SMTP and e-mail addresses by clicking the **Test** button.
8. If you have a Network Management Station (NMS) in your network collecting traps, select the **SNMP** tab to define the settings for SNMP alerts. Type the IP address of your SNMP console and the community name. As with the e-mail addresses, you can define several trap destinations.

The NMS can decode the received traps using the MIB file included in the Storage Manager software CD, which should be compiled into the NMS console to allow proper display of the traps. To set up alert notification to an NMS using SNMP traps, perform the following steps:

- Insert the IBM DS Storage Manager CD into the CD-ROM drive on an NMS. You need to set up the designated management station only once.
- Copy the SMxx.x.MIB file from the SMxxMIB directory to the NMS.
- Follow the steps required by your NMS to compile the management information base (MIB) file. (For details, contact your network administrator or see the documentation specific to your particular storage management product.)

After configuring the storage subsystem for alerts, check the results in the Enterprise Management window, as shown in Figure 4-94.

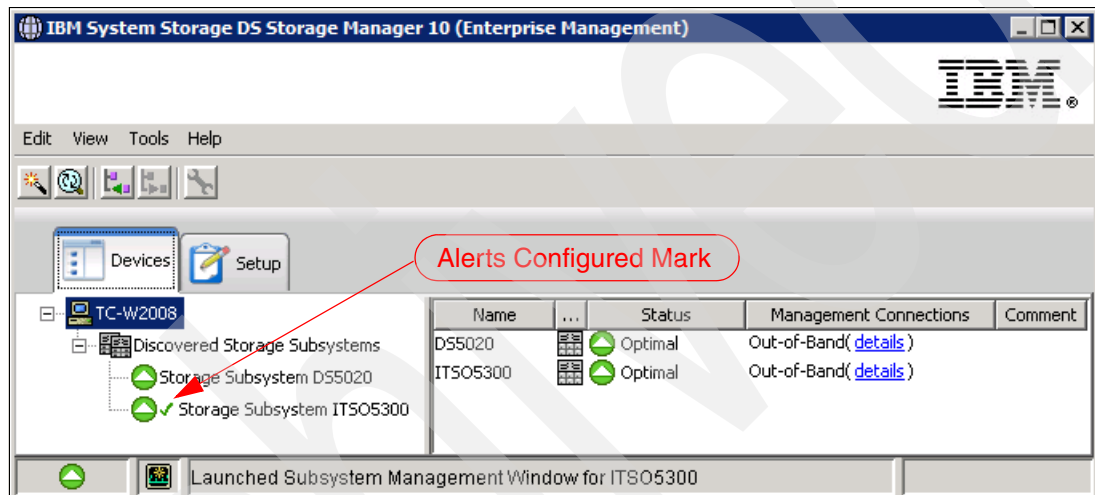


Figure 4-94 Alerts configured for one specific DS5000 storage subsystem

Check for the mark besides each representation of the storage subsystem configured for alerts.

- ▶ Only the storage subsystems with a Alerts Configured mark will be monitored.
- ▶ Event Monitor will send only one alert per event, even if the events are monitored through different hosts or out-of-band. However, if you configure the Event Monitor in more than one management station, you will receive duplicate messages.

See Chapter 6, “IBM Remote Support Manager for Storage” on page 285 if you want IBM to receive these notifications.

4.9.8 Saving the configuration

Once your DS5000 storage subsystem is configured and running, you should save this configuration. This allows you to replicate the configuration already performed in another DS5000 with identical physical resources, or in the same system in case you need it. It can also be used to recover part of the configuration in case of problems.

These are the different type of information to save:

- ▶ Save Configuration option
- ▶ Storage Profile option
- ▶ Support Data

We show how to save each of the options available in the following sections. If you need assistance to recover part of the configuration, contact your IBM Support representative.

Save Configuration

The Save Configuration option includes information for the arrays and logical drive configuration, the name of the subsystem, its cache settings, and other parameters, including the storage partitioning configuration.

The saved file can be used to restore the configuration data to the same DS5000 storage subsystem, or also to other DS5000 storage subsystems in case you want to set up multiple storage systems with the same configuration. To allow that action, 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 select **Storage Subsystem** → **Configuration** → **Save** (Figure 4-95).

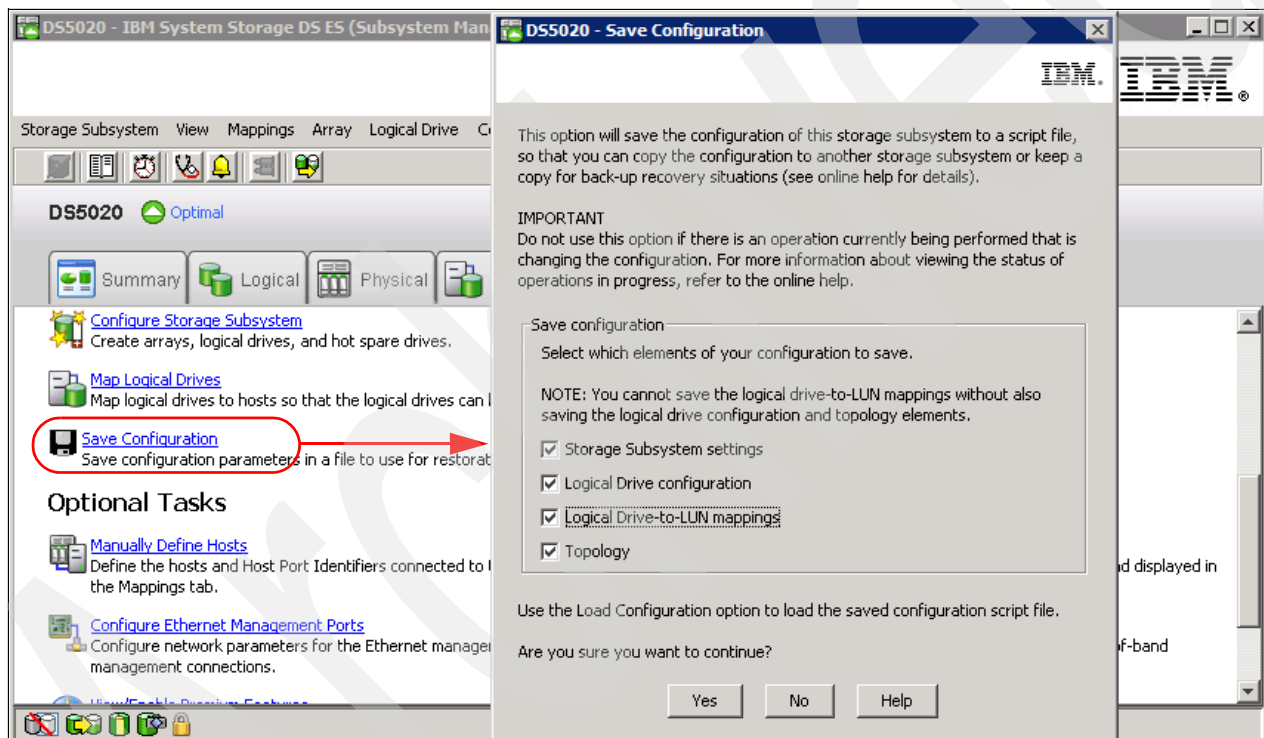


Figure 4-95 Saving the DS5000 storage subsystem configuration

We can choose to save specific elements of the configuration.

Select the desired configuration elements, click **Yes**, and select a file name and destination folder in which to save the file. Make sure not to use a directory located on a DS5000 storage subsystem disk, or you might not be able to access it when needed.

The script created can be used to replicate the configuration of the DS5000 storage subsystem. You can apply the configuration to the destination subsystem for all the saved elements, or any particular element.

Remember that this save option does not include the data resident on the logical volumes, only configuration data. Make sure to make periodic backups of your data to avoid exposures.

Storage subsystem profile

Configuring a DS5000 storage subsystem is a complex task and it is therefore essential to document the configuration. The profile data contains information that could help recover part or all of the configuration of the DS5000 storage subsystem, and can also be saved in a file known as the *subsystem profile*. This profile stores information about the controllers, attached drives, and enclosures, and their microcode levels, arrays, logical drives, and storage partitioning.

Tip: You should save a new profile every time you change the configuration of the DS5000 storage system, even for minor changes. The profile is stored in a location where it is available even after a complete configuration loss, for example, after a site loss.

1. To obtain the profile, open the Subsystem Management window and select **Storage Subsystem** → **View** → **Profile**, as shown in Figure 4-96.

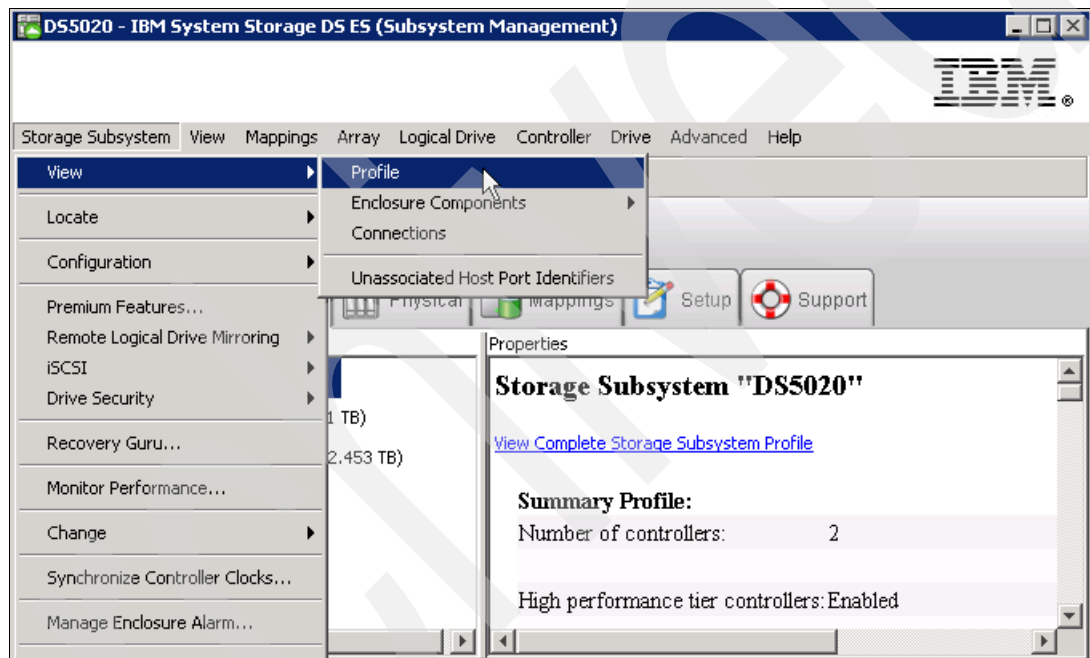


Figure 4-96 Viewing the storage subsystem profile

2. The information is gathered from various components when you request the profile. The profile can be saved locally and included in the documentation to maintain a change history for the storage system.

We recommend that you save a new version of the profile and store it securely whenever a configuration change takes place. Even in the case of a complete configuration loss, you can restore the arrays and logical drives configuration as well as the mappings for the storage partitioning using the profile information.

A sample profile window is shown in Figure 4-97.

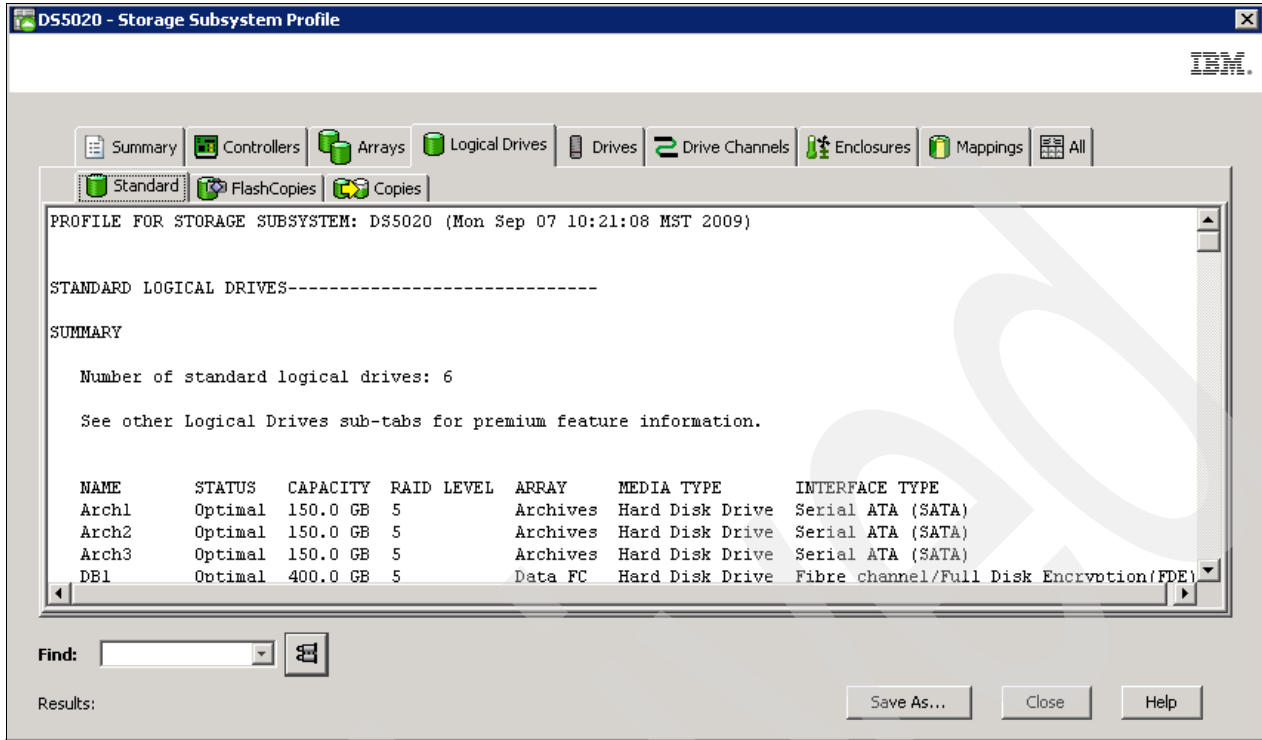


Figure 4-97 Storage Subsystem Profile

3. Select the **Save As** option to keep a copy of the current profile configuration. In the next window, select **All Sections** and specify a directory and file name where you will save a copy, as shown in Figure 4-98.

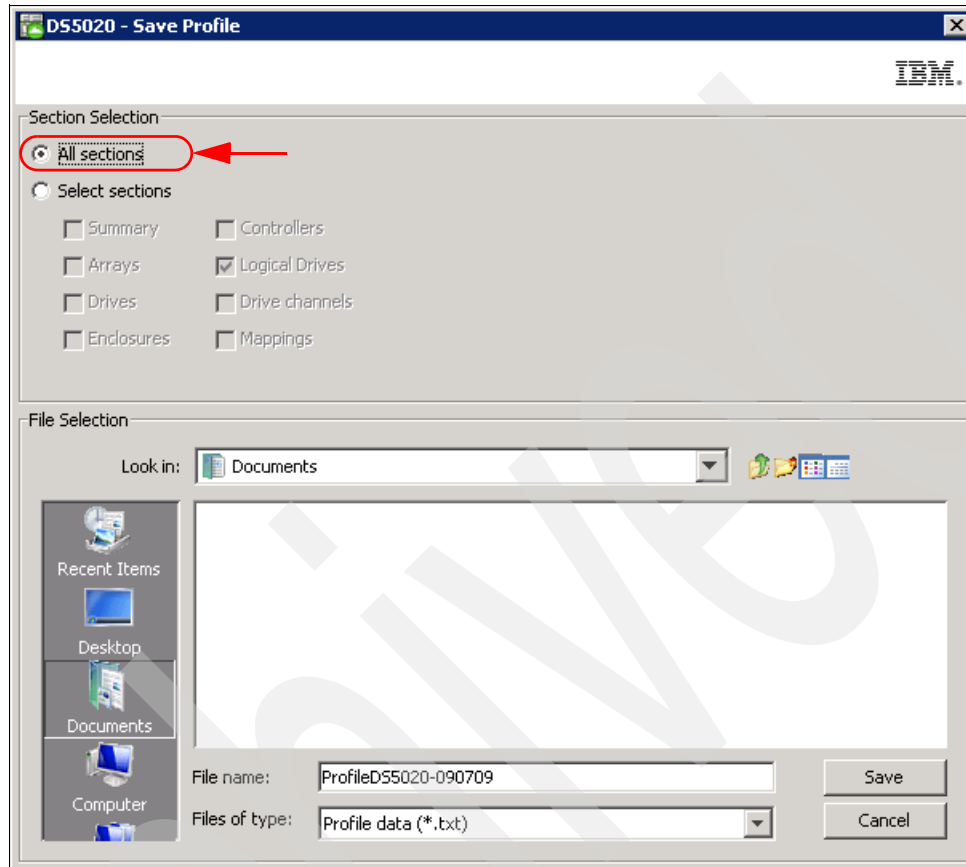


Figure 4-98 Save Profile: All sections

Gathering support information

Support Data is an option of the Storage Manager that lets you collect all the internal information of the DS5000 storage subsystem for review by the support organization. This includes the storage Subsystems Profile, majorEventLog, driveDiagnosticData, NVSRAM data, readLinkStatus, performanceStatistics, and many others.

This information can be collected manually at any time, but if there is a critical problem, it is collected automatically and saved to the folder `\client\data\monitor\` under the installation path of the Storage Manager. In case of problems, your IBM support representative might need this data to identify the source of the problem and make the necessary corrections.

Make sure to check that the automatic collection of support data is not disabled, select **Advanced** → **Troubleshooting** → **Support Data** → **Support Data** → **Automatic Settings**, as shown in Figure 4-99.

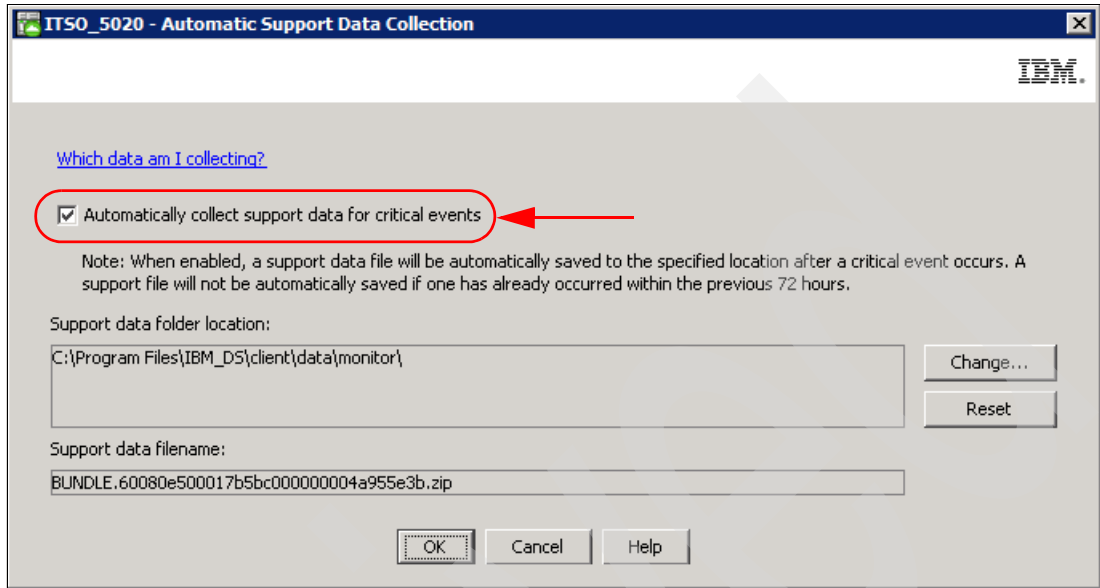


Figure 4-99 Automatic Support Data Collection

Collect data before and after any major changes by selecting **Gather support information** from the Support view of the Subsystem Management window, as shown in Figure 4-100.

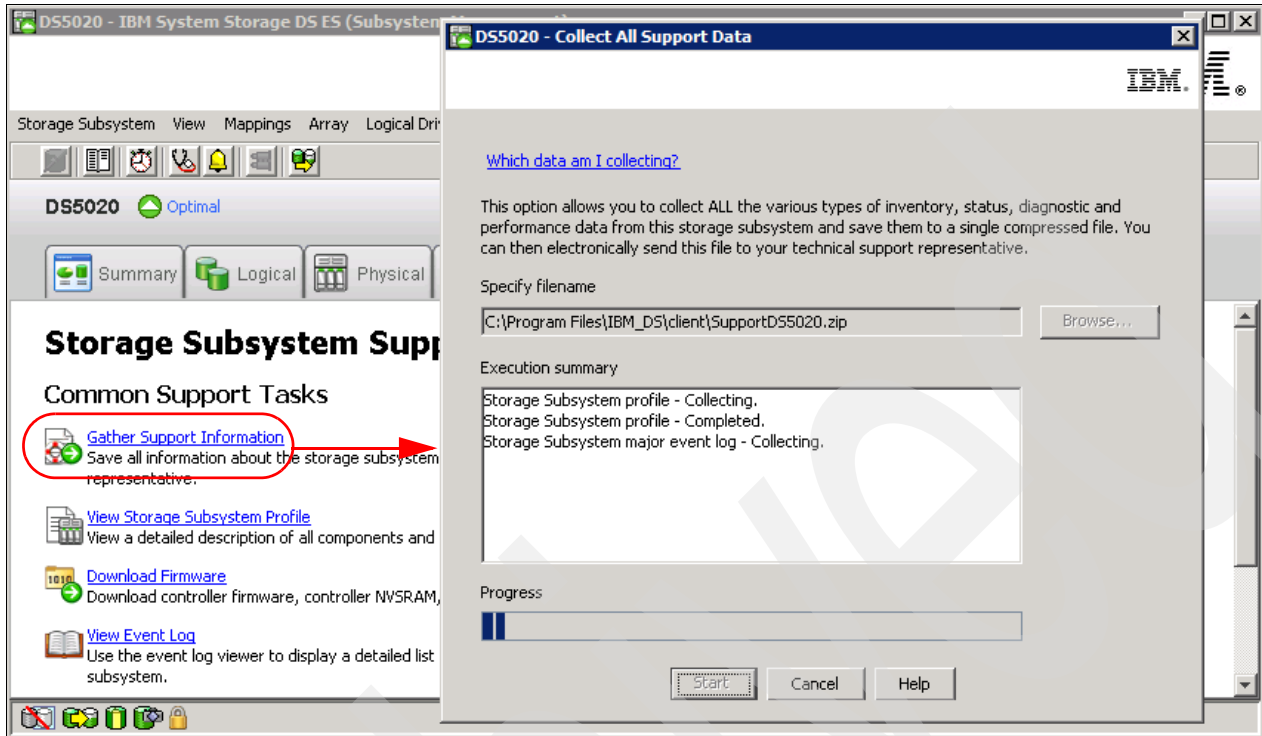


Figure 4-100 Gather support information

Important: The profile and support data files provides configuration information that can be used in case of a failure to recover the configuration. Remember to save a copy after making changes in the configuration, and keep it outside of the DS5000 storage subsystem disks.

4.10 Advanced functions

This section introduces some of the advanced features of the Storage Manager.

4.10.1 Expanding arrays

The ability to increase the available free capacity in an array (*Dynamic Capacity Expansion (DCE)*) without needing to restart the host system is a very important feature. In today's IT environment, the need for storage space grows constantly. Many customers exhaust their existing space sooner or later and have to expand their storage capacity. It is essential that this process be nondisruptive and not cause any downtime.

With Storage Manager, it is possible to add new disk drives to the storage subsystem and start the expansion procedure while the system remains fully operational. Once the procedure starts, it cannot be stopped. This procedure might have a performance impact, because the expansion process competes with normal disk access. We recommend that, where possible, that this type of activity be performed when I/O activity is at a minimum. The new free capacity can be used to create additional logical drives. Existing logical drives in the array do not increase in size as a result of this operation.

Note: Storage Manager supports RAID 0 and 1 arrays with more than 30 drives. In certain DS5000 storage subsystem configurations, this can improve performance, provided that the system is optimally tuned. It also improves the data capacities of these arrays. RAID 1 or 10 requires an even number of disk drives.

Attention: It is still not possible to use more than 30 drives in RAID 3, 5, and 6 arrays. Once the maximum number of drives is reached, you obviously cannot add new drives anymore.

To add new drives to an array, highlight the array, right-click, and select **Add free Capacity (Drives)**. In the Add Drives window (Figure 4-101), choose one or two drives to be added to the array, depending on whether RAID 1 or 10 is being used by the array. The controller firmware imposes a maximum of two drives to be added *at one time*, although this operation can be repeated to add more than two drives to an array.

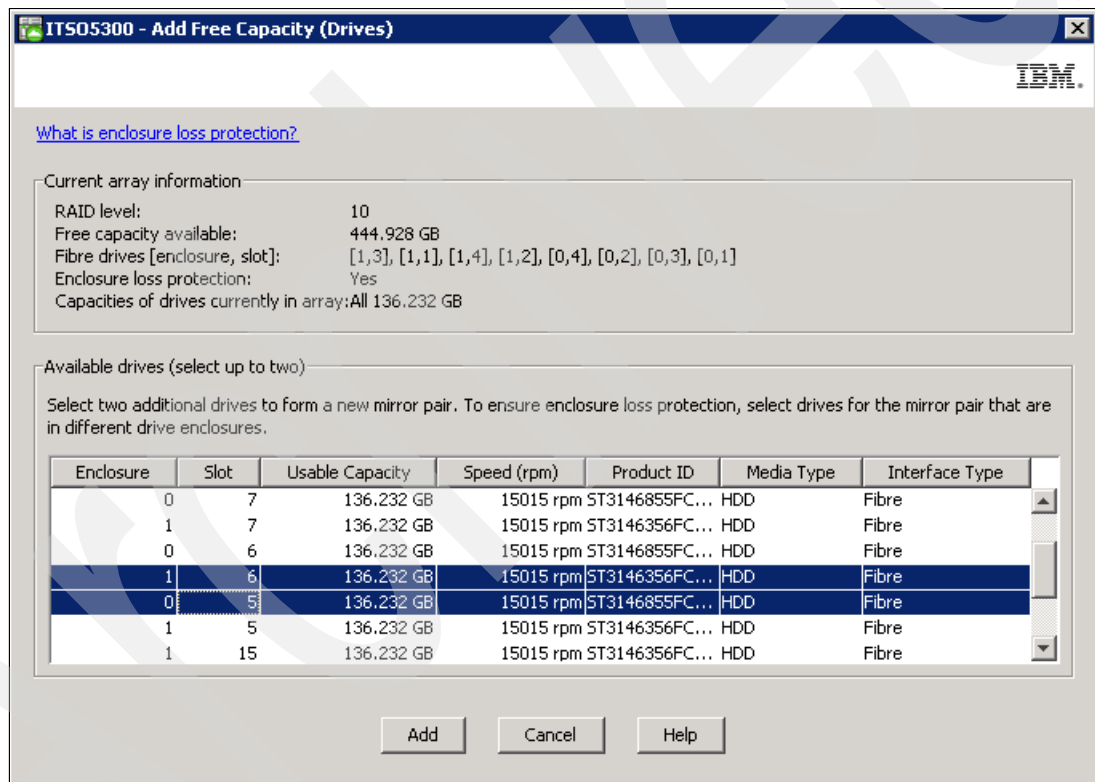


Figure 4-101 Adding new drives to an array

For RAID levels 3, 5, and 6, select one drive, and for RAID levels 1 and 10, two drives must be selected, with the following considerations:

- ▶ Only the type of disks are listed as candidates to add (FC/SATA).
- ▶ Select drives that have a capacity equal to the current drive capacities in the array.

Note: Drives larger than the other drives participating in the array can be added, but we do not recommend it, because their usable capacity will be reduced so that they match the current drives capacities in the array.

Once the procedure is started, it cannot be stopped, as the subsystem needs to redistribute the data contained in the array to all drives, including the new ones. There is a performance impact during this operation, but the logical drives of the array remain available to the host systems.

4.10.2 Defragment an 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.

New logical drives cannot spread across several free space nodes, so 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 to one free space node for the array. Then, all new logical drives can use the whole available free space.

To accomplish this task, open the Subsystem Management window, highlight the array to defragment, and select **Advanced** → **Recovery** → **Defragment Array** to start the procedure, as shown in Figure 4-102. The defragmentation can run concurrently with normal I/O, but it impacts performance because the 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.

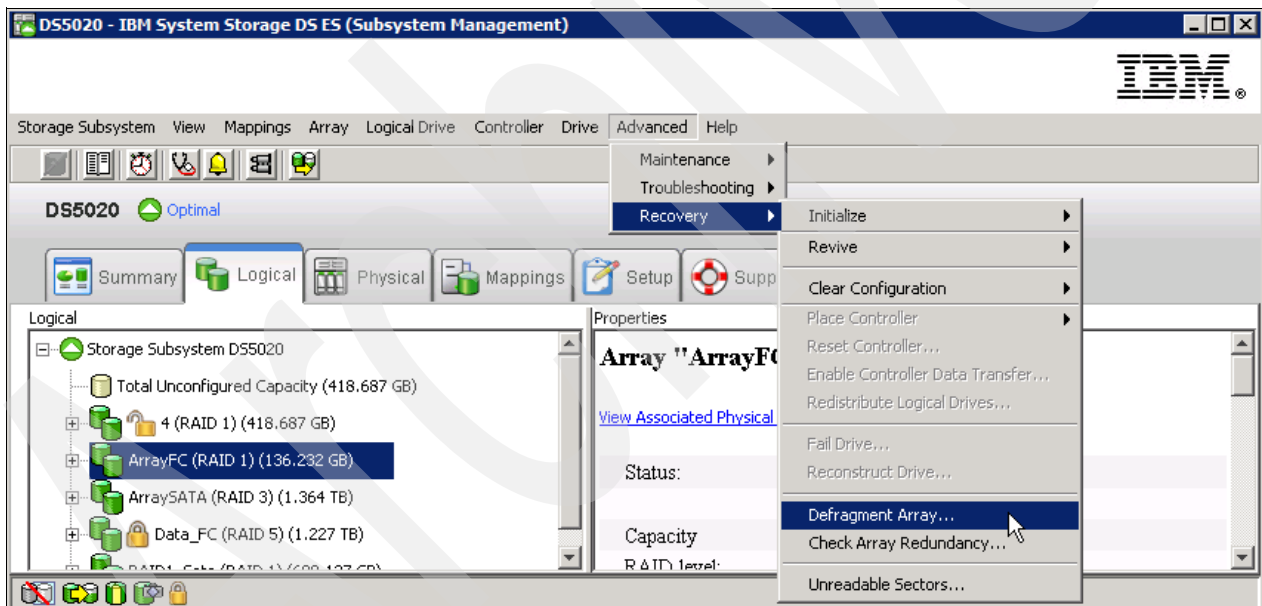


Figure 4-102 Defragment an array

The defragmentation done on the DS5000 storage subsystem 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.

4.10.3 Changing the RAID array level

Changing the RAID level of an array is performed in a nondisruptive manner. The system remains fully operational while the process takes place. A few possible reasons why customers might want to do this operation are:

- ▶ The storage requirements changed over time and existing RAID levels are no longer optimal for a particular environment.
- ▶ The performance tuning process indicates that a different RAID level is more appropriate than the existing one.

It is possible to change any RAID level to any other one. There are some restrictions that apply to the new arrays:

- ▶ RAID 1 or 10 requires an even number of disk drives.
- ▶ RAID 3 and 5 require at least three drives.
- ▶ RAID 6 requires at least five drives.
- ▶ There is a limit of 30 drives per array for RAID 3, 5, and 6 arrays.

There are limitations if there is not enough free space in the array. For example, a RAID 5 array of four disk drives with no free space cannot be migrated directly to RAID 1. If this migration is attempted, an error message will be displayed stating that there is not enough free space. There must be enough free capacity to change the RAID level. Also, if the array has an odd number of drives and a migration to RAID 1 is required, a disk must be added to the array prior to performing the procedure.

When changing from RAID 1 to RAID 5, free space in the array can be gained, which can be used to define new logical drives or expand existing ones.

When the procedure starts, it reorganizes the data segments in the array according to the new RAID level, and a large amount of I/O happens, so there is an impact on performance while the migration lasts. The performance impact can be controlled to a certain extent by changing the value of the modification priority. This parameter is set on a logical drive basis, which is where it should be changed for all logical drives in the array.

Changing the modification priority to a low value during the migration process minimizes performance degradation. When the migration finishes, the value can be increased to reduce the time for a rebuild in case of a drive failure. This minimizes the critical time of non-redundant operation caused by the disk drive fault.

Attention: Once the migration starts, it cannot be stopped.

Note: Even though RAID migration is a nondisruptive process, we recommend carrying out this migration when I/O activity is at a minimum.

Even though the DS5000 storage subsystem always tries to optimize the layout of the disk arrays, some settings might require a change to optimize the disk usage or the performance.

To change an array's RAID level, from the Storage Manager window, select the array, right-click it, select **Change** → **RAID Level**, and then the desired RAID level, as shown in Figure 4-103.

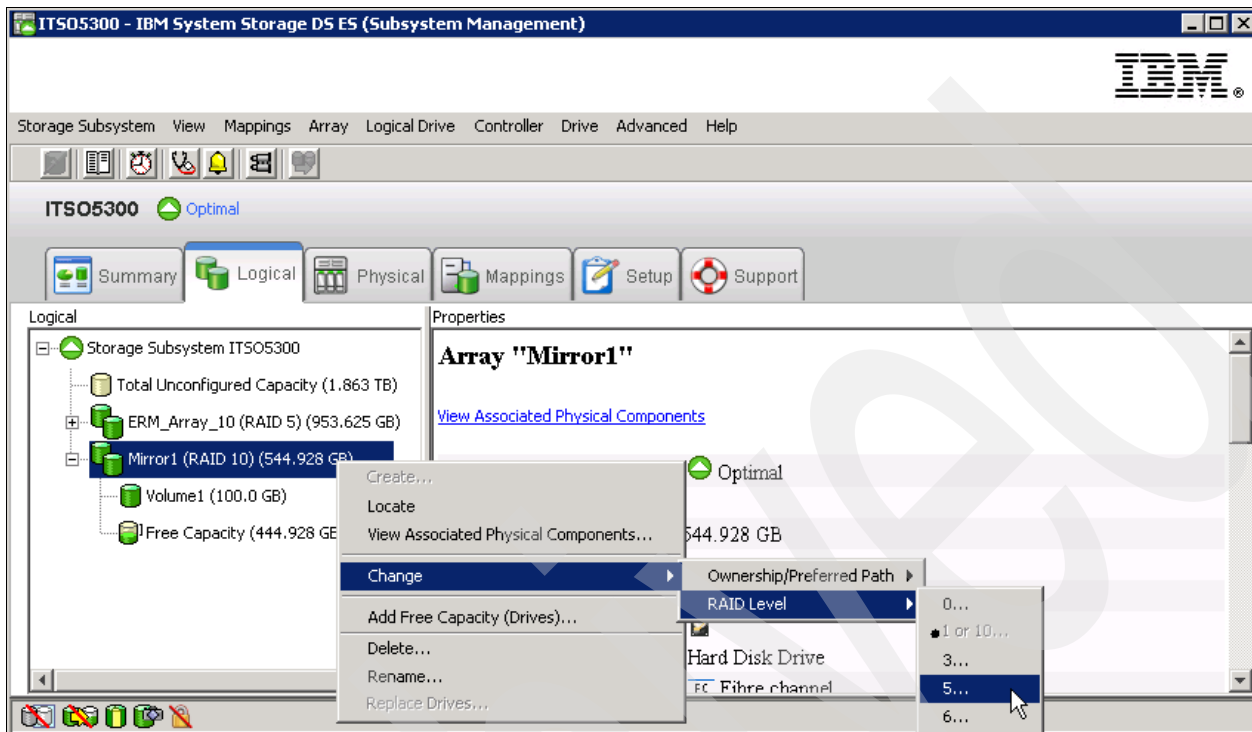


Figure 4-103 Changing RAID level

A confirmation message appears and informs you that the operation cannot be stopped until it is complete. The data remains accessible during this operation, which might take a long time. To check the progress status, select the logical drives in the array being modified, right-click it, and then select **Properties** to display a progress bar for the operation.

4.10.4 Unconfiguring a storage subsystem and arrays

Storage Manager allows the storage subsystem or previously created arrays to be cleared if required.

Clearing the storage subsystem completely removes the complete configuration of the storage subsystem and data, bringing it back to the state when it was initially installed. Information defining all arrays, logical drives, and hot spares are deleted. This feature can be used to create a new configuration on a storage subsystem that already has a configuration defined that is now redundant.

Clearing an array configuration clears logical drives and group configuration in *all* arrays and leaves the remaining subsystem configuration intact.

Warning: In both cases (clearing the storage subsystem or array configuration), data loss occurs. Ensure that a backup of the storage subsystem data as well as the storage subsystem configuration profile is made before attempting these operations.

4.10.5 Performing advanced functions on logical drives (LUNs)

This section details the advanced functions that can be performed on logical drives.

Pre-read redundancy check

RAID arrays are created to provide tolerance against disk failure and errors. To allow recovery from failures, the RAID array stores additional data called redundancy data. This redundancy data can potentially become inconsistent from time to time. Redundancy data is “consistent” if every portion of the redundancy group can be reliably and consistently reconstructed in the event of a failure to the array, and “inconsistent” if it cannot be read or the consistency cannot be verified.

With Storage Manager, a pre-read redundancy check feature is available. The pre-read redundancy feature allows the consistency of redundancy data of a RAID array to be optionally checked prior to returning data for host read requests. The consistency check cannot determine whether the data itself is correct. It does, however, determine that a redundancy consistency error has occurred within the extent of the read command. Once an inconsistency is discovered, an error is reported in the storage subsystem, and no read data is returned, depending on the data inconsistency error.

This feature can be activated on a logical drive basis, which is created as a RAID array that supports redundancy information. The feature allows data verification in environments where data consistency is a key requirement. This feature cannot be enabled for logical drives that do not have any redundancy data.

When a read request is issued to the storage subsystem, the controller verifies that the redundancy group is consistent for the extent of the data specified in the read request. If the controller found that the redundancy group data is in a consistent state, the read request is returned to the host successfully. If the redundancy group data is found to be in an inconsistent state, the read request is returned to the host with a check condition status. The consistency check is only performed for data not already in cache.

In order to verify the consistency of the redundancy group, the entire redundancy group must be read from disk into cache. Should no inconsistency be found, all user data that comprises the redundancy group is left in cache in order to potentially satisfy subsequent cache hits. The data left in cache might include data that is outside the extent of the original read request. Such data does not require a consistency check, because the data is already in cache.

Note: Should a drive associated with an array become degraded, the pre-check feature is disabled.

During the process of verifying the consistency of the redundancy group, a media error or unreadable sector might be encountered within the extent of the redundancy group. When such an error occurs, the RAID controller attempts to perform a reconstruction of the data, using the redundancy information of the data. If this operation is successful, the consistency check indicates that the redundancy group is consistent and the read data is returned to the host.

During the process of verifying the consistency of the redundancy group, an unrecoverable read error might be encountered. When such an error occurs, the consistency check fails. However, the original read request continues to be processed. If the entire extent of the request is still readable, the read data is returned to the host. If some portion of the original read request is unavailable, the read command fails.

In most situations, recovery from an inconsistent redundancy group involves restoring the volume from a backup source.

Note: In a FlashCopy or Enhanced Remote Mirror copy pair relationship, the pre-read redundancy feature is only supported on the source logical drives.

When activating the pre-read redundancy check feature, the media scan feature no longer automatically corrects discovered redundancy group inconsistencies.

Attention: Take care when enabling this feature, as it could have an impact on I/O performance.

To enable this feature, perform these steps:

1. Select the logical drive where the pre-read redundancy check is to be enabled.
2. Select **Logical Drive** → **Change** → **Pre-Read redundancy check**, as shown in Figure 4-104.

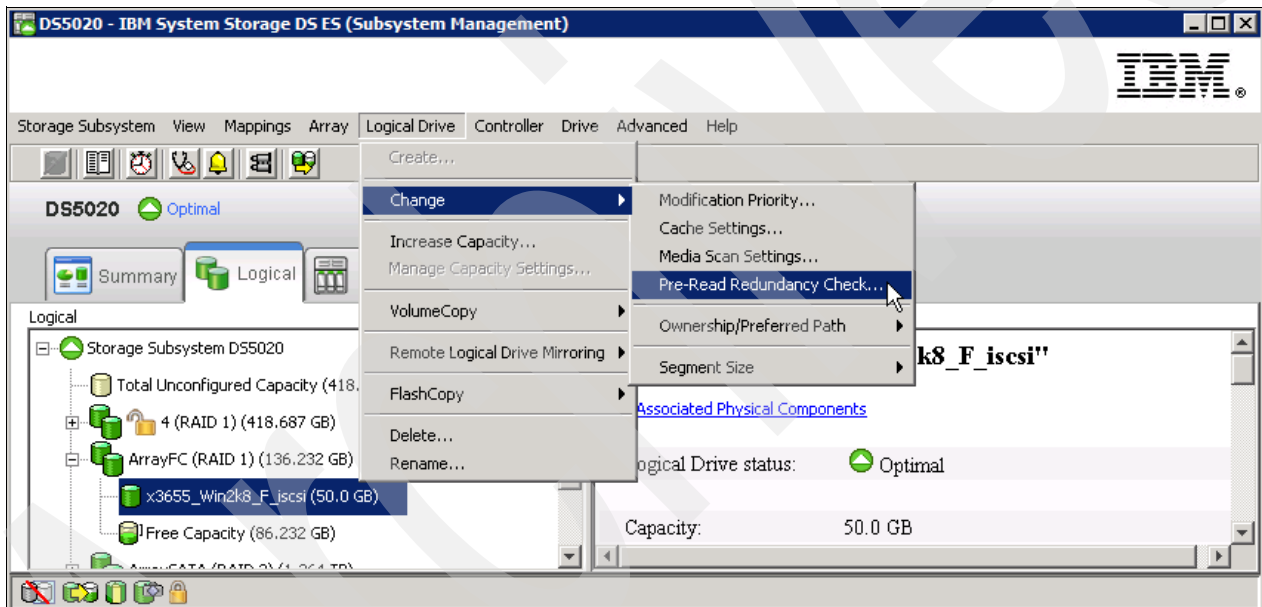


Figure 4-104 Pre-read redundancy check

3. Make sure that the correct logical volume is selected, and click the lower box to enable the value, as shown in Figure 4-105.

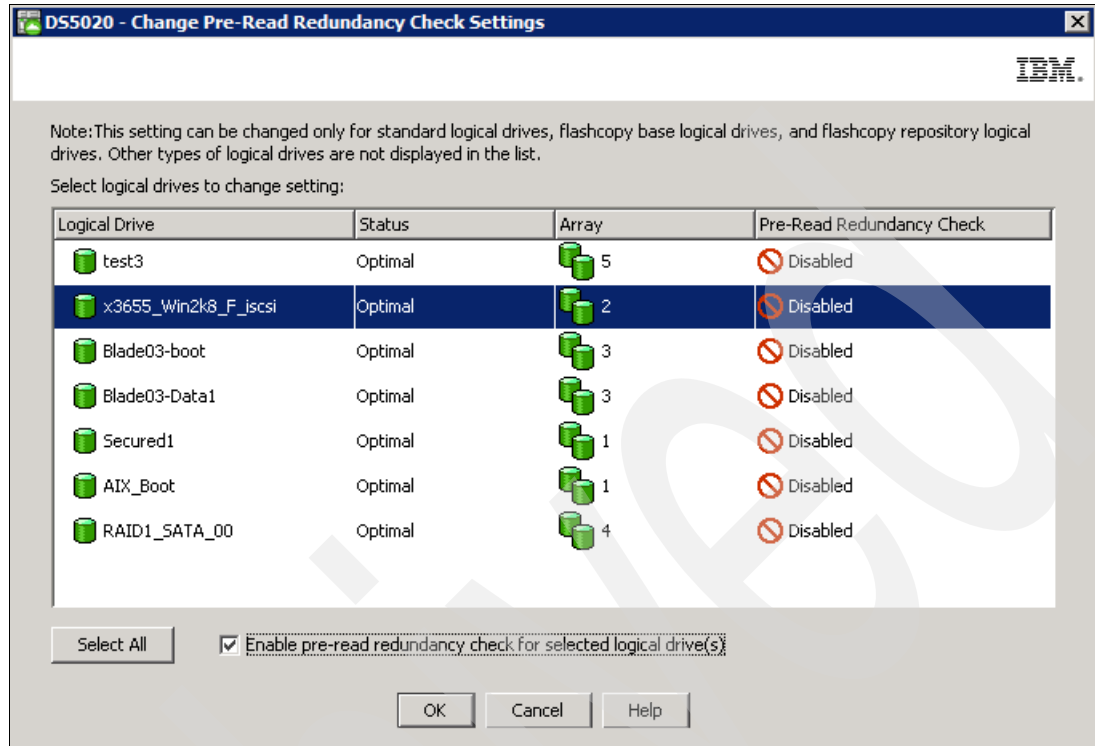


Figure 4-105 Enabling pre-read redundancy check

Expanding logical drives

It is also possible to increase the size of logical drives. This action is called *Dynamic Volume Expansion* (DVE). The capacity of standard logical drives and FlashCopy repository logical drives might be increased using one or both of the following capacities:

- ▶ Free capacity available on the array of the standard or FlashCopy repository logical drive
- ▶ Unconfigured capacity (in the form of unused drives) on the array of the standard or FlashCopy repository logical drive

Increasing the capacity of a FlashCopy repository logical drive does not increase the capacity of the associated FlashCopy logical drive. The FlashCopy logical drive's capacity is always based on the capacity of the base logical drive at the time the FlashCopy is created.

Note: Storage Manager provides support for logical drives greater than 2 TB. This improves the capacity requirement for applications that requires large capacity logical drives.

Note: Increasing the capacity of a standard logical drive is only supported on certain operating systems. If the logical drive capacity is increased on a host operating system that is not supported, the expanded capacity will be unusable and the original logical drive capacity will not be able to be restored.

The operating systems that support a dynamic increase of capacity in a mapped logical drive are:

- ▶ AIX
- ▶ Linux
- ▶ NetWare
- ▶ Windows Dynamic Disks
- ▶ Windows Basic Disks

Tip: If a logical drive-to-LUN mapping has not yet been defined, it is possible to increase the capacity for a standard logical drive on any host operating system, that is, before host system data is placed on the drive.

The storage capacity of a standard logical drive cannot be increased if:

- ▶ One or more hot spare drives are in use in the array that it resides on.
- ▶ The logical drive has *Non-Optimal* status.
- ▶ Any logical drive in the array is in any state of modification.
- ▶ The controller that owns this logical drive is in the process of adding capacity to another logical drive (each controller can add capacity to only one logical drive at a time).
- ▶ No free capacity exists in the array and no unconfigured capacity (in the form of drives) is available to be added to the array.

To increase the logical drive capacity, on the SM Logical/Physical view, highlight the logical drive to be expanded, right-click it, and select **Increase Capacity**. In the Increase Logical Drive Capacity window (Figure 4-106), enter the amount of space by which the logical drive will be enlarged.

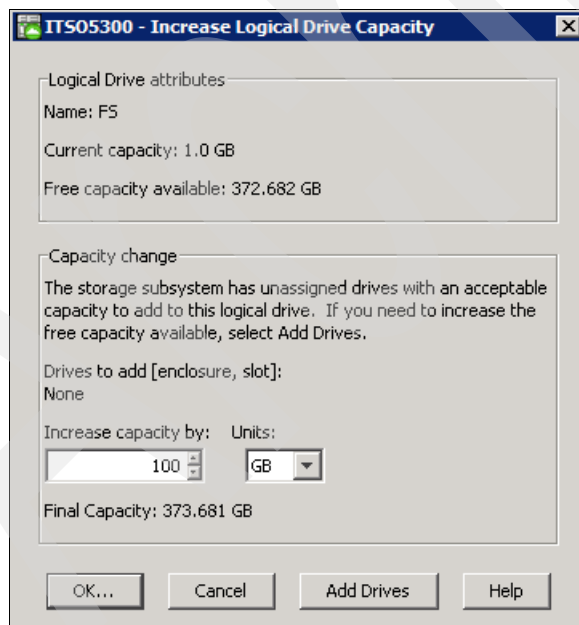


Figure 4-106 Dynamic logical drive expansion

In the top part of the window shown in Figure 4-106, the current size of the logical drive and the available free capacity in the array is displayed. If no free configured space is available, but there are unassigned drives, these can be added from this same window by clicking **Add Drives** before proceeding to enlarge the logical drive. This will perform the same process described in 4.10.1, “Expanding arrays” on page 233.

Note: If the RAID level of the array is 3, 5, or 6, and the drive enclosure has enclosure loss protection, the Add Drives option displays only drives that ensure enclosure loss protection. If the RAID level is 1 or 10, a minimum of two drives must be added.

Click **OK** after selecting the capacity to add. A warning message appears indicating that this operation cannot be stopped after it is started and that it might take a long time to complete. However, the data on the selected logical drive and all other logical drives on this array (if new drives have been added) remains accessible during this time. As with all operations requiring a redistribution of the data on the physical disks, the procedure might affect the performance. From the host operating system, the administrator will then have to perform a procedure in order to utilize the newly allocated space.

Extending a basic disk on a Windows platform

In the following example, we have a Windows 2008 system with a basic disk partition of 1 GB. The partition has data on it, the partition is disk 6, and its drive letter is J. We have used DS5000 Dynamic Volume Expansion (DVE) to expand the logical drive by adding 100 GB. This leaves the operating system with a disk of 101 GB, with a partition of 1 GB and free space of 100 GB, as shown in Figure 4-107.

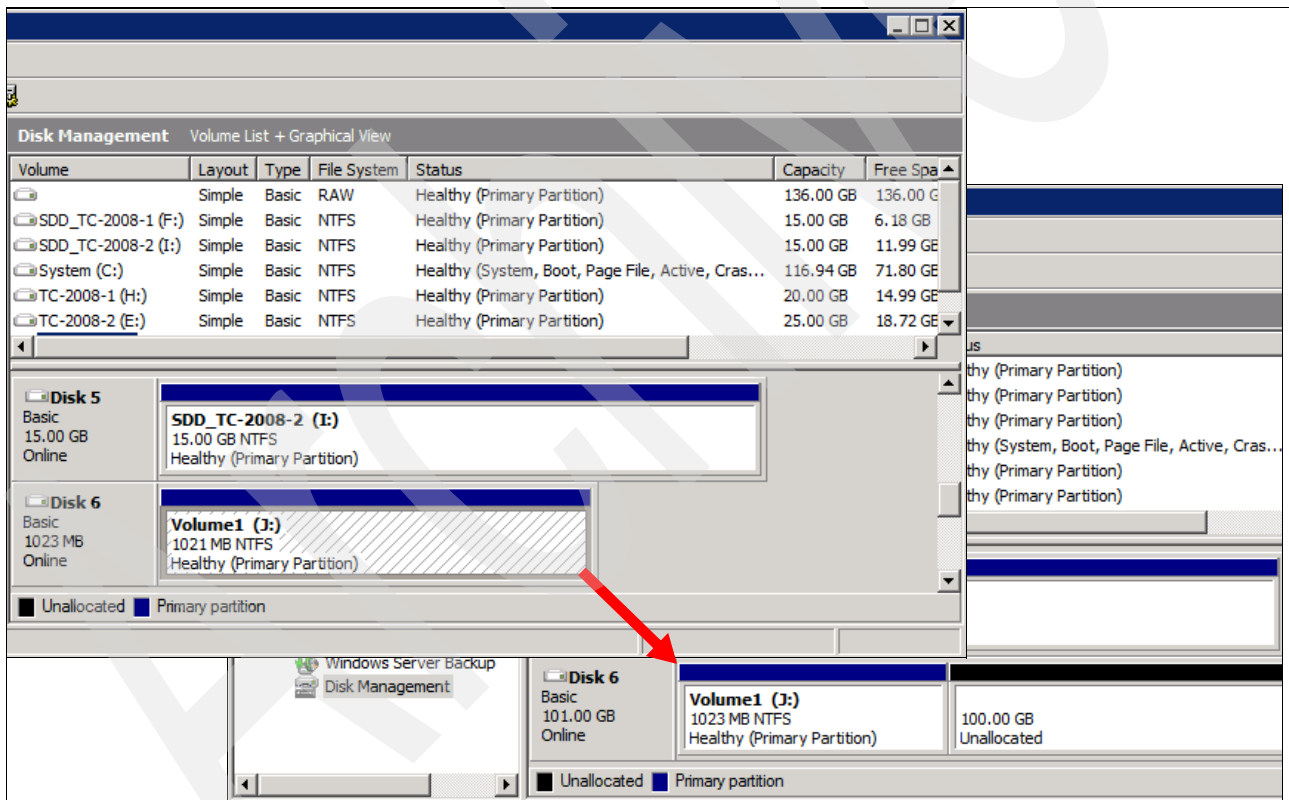


Figure 4-107 The Windows 2008 basic disk with free space

We use the Windows 2008 command-line utility diskpart.exe to extend the original 1021 MB partition to the full size of the disk, as shown in Example 4-2.

Example 4-2 The diskpart utility extends the basic disk in a command window

```
CMicrosoft DiskPart version 6.0.6002
Copyright (C) 1999-2007 Microsoft Corporation.
On computer: TC-W2008
```

```
DISKPART> list volume
```

Volume ###	Ltr	Label	Fs	Type	Size	Status	Info
Volume 0	D			DVD-ROM	0 B	No Media	
Volume 1	G			DVD-ROM	0 B	No Media	
Volume 2	C	System	NTFS	Partition	117 GB	Healthy	System
Volume 3	H	TC-2008-1	NTFS	Partition	20 GB	Healthy	
Volume 4	E	TC-2008-2	NTFS	Partition	25 GB	Healthy	
Volume 5			RAW	Partition	136 GB	Healthy	
Volume 6	F	SDD_TC-2008	NTFS	Partition	15 GB	Healthy	
Volume 7	I	SDD_TC-2008	NTFS	Partition	15 GB	Healthy	
Volume 8	J	Volume1	NTFS	Partition	1021 MB	Healthy	

```
DISKPART>
```

```
DISKPART> select volume 8
```

```
Volume 8 is the selected volume.
```

```
DISKPART> extend
```

```
DiskPart successfully extended the volume.
```

```
DISKPART> list volume
```

Volume ###	Ltr	Label	Fs	Type	Size	Status	Info
Volume 0	D			DVD-ROM	0 B	No Media	
Volume 1	G			DVD-ROM	0 B	No Media	
Volume 2	C	System	NTFS	Partition	117 GB	Healthy	System
Volume 3	H	TC-2008-1	NTFS	Partition	20 GB	Healthy	
Volume 4	E	TC-2008-2	NTFS	Partition	25 GB	Healthy	
Volume 5			RAW	Partition	136 GB	Healthy	
Volume 6	F	SDD_TC-2008	NTFS	Partition	15 GB	Healthy	
Volume 7	I	SDD_TC-2008	NTFS	Partition	15 GB	Healthy	
* Volume 8	J	Volume1	NTFS	Partition	101 GB	Healthy	

```
DISKPART>exit
```

After diskpart.exe has extended the disk, the partition is now 101 GB. All the data is still intact and usable, as shown in Figure 4-108.

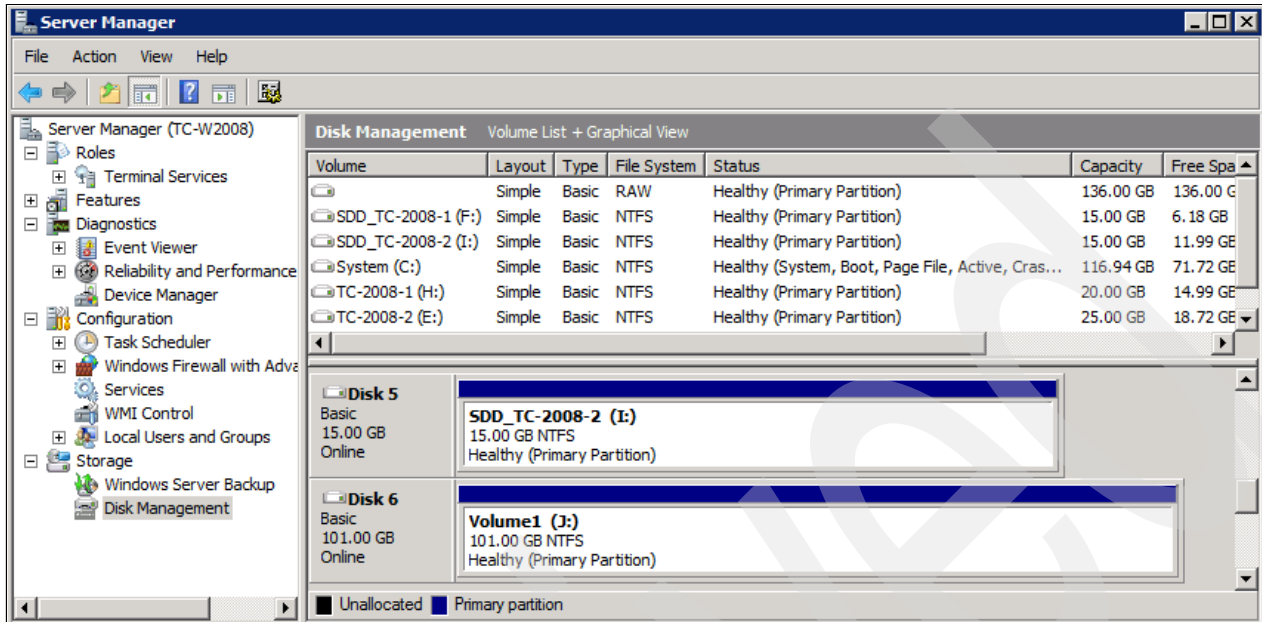


Figure 4-108 Disk Management after diskpart has extended the partition

Notes:

- ▶ The **extend** operation is dynamic.
- ▶ The **extend** command only works on NTFS formatted volumes.
- ▶ Officially, you do not need to stop I/O operations to the disk before you extend. However, keeping I/O operations at a minimum is a best practice.

With dynamic disks, the Disk Management GUI utility can be used to expand logical drives.

Modification priority

The modification priority defines how much processing time is allocated for operations modifying the logical drive relative to the system performance. Operations that cause a logical drive modification are:

- ▶ Initializing a logical drive
- ▶ Reconstructing after a disk failure
- ▶ Copying back from a hot spare drive
- ▶ Changing the segment size of a logical drive
- ▶ Expanding a dynamic logical drive
- ▶ Adding free capacity to an array
- ▶ Defragmenting an array
- ▶ Changing the RAID level of an array

If the logical drive contains critical data, you might prefer a high modification priority to keep the time of a critical state (for example, after losing a disk) as short as possible, even if this affects the system performance during the modification process.

The following modification priority rates are available:

- ▶ Lowest
- ▶ Low
- ▶ Medium
- ▶ High
- ▶ Highest

Note: The lowest priority rate favors system performance, but the modification operation takes longer. The highest priority rate favors the modification operation, but system performance might be compromised.

The progress bar at the bottom of the Logical Drive Properties window displays the progress of a modification operation.

When a storage system logical drive is a primary logical drive and a full synchronization is necessary, the controller owner performs the full synchronization in the background while processing local I/O writes to the primary logical drive and associated remote writes to the secondary logical drive. The full synchronization diverts controller processing resources from I/O activity, where it can impact performance on the host application. The synchronization priority defines how much processing time is allocated for synchronization activities relative to system performance.

The guidelines in Table 4-10 can help determine how long a synchronization can take and how much various synchronization priorities can affect system performance.

Table 4-10 Impact of modification priority on relative time for full synchronization

Modification priority	Relative time
Highest	Fastest possible time
High	Two times longer than fastest possible time
Medium	Three and a half times longer than fastest possible time
Low	Six times longer than fastest possible time
Lowest	Eight times longer than fastest possible time

To change the modification priority, perform the following steps:

1. Select a logical drive, right-click it, and select **Change** → **Modification Priority**, as shown in Figure 4-109.

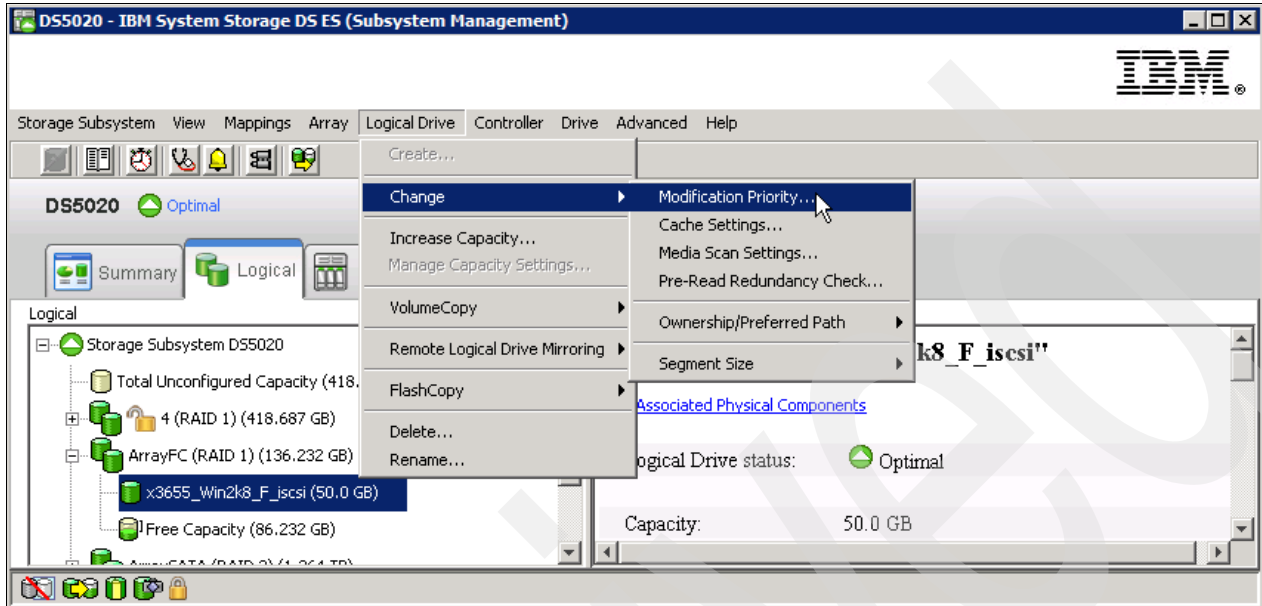


Figure 4-109 Changing the modification priority

2. Make sure that the correct logical drive is selected, and set the new Modification Priority value, as shown in Figure 4-110.

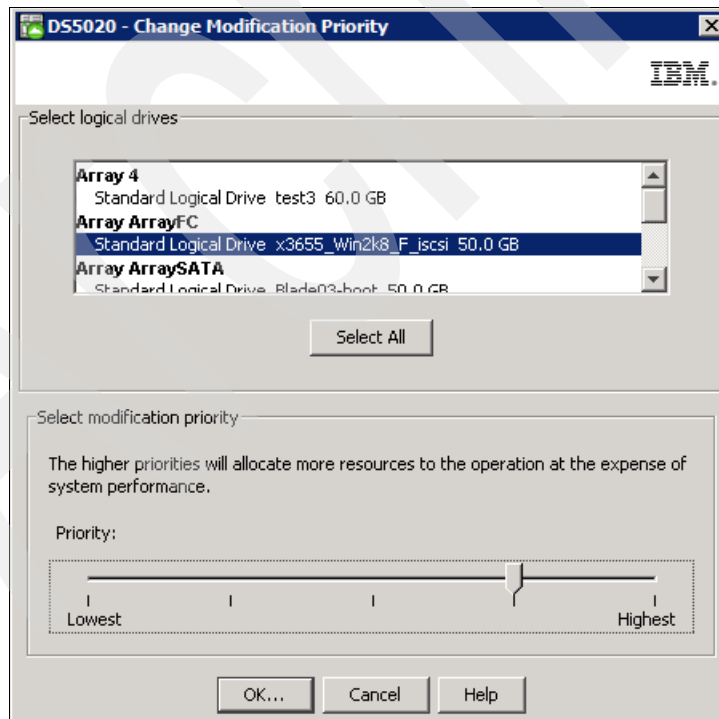


Figure 4-110 Modification priority for a logical drive

If a logical drive modification is in progress, a status bar appears at the bottom of the window, as shown in Figure 4-111.

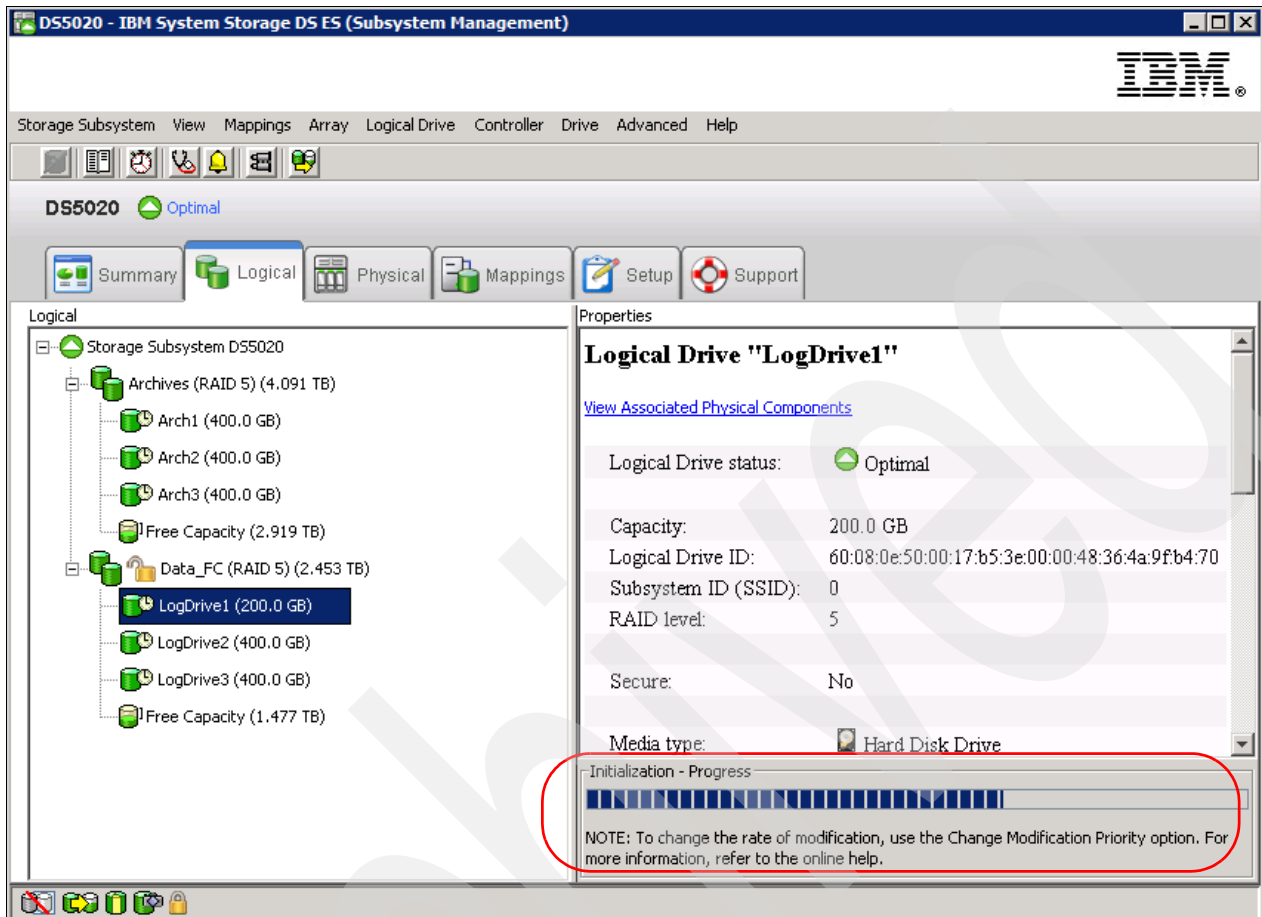


Figure 4-111 Modification progress

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 controller. The alternate controller only takes over and handles the I/O requests in case of a failure along the I/O path, for example, a defective host bus adapter or switch. When defining logical drives, the system normally alternates ownership between the two controllers as they are defined.

Situations can occur when all heavily stressed logical drives can reside on only one controller and the other one handles only a small amount of all I/O requests. To balance the workload between the controllers, the preferred ownership of a logical drive can be changed to the other controller.

Important: Be sure that the operating system using the logical drive uses a multipath I/O driver; otherwise, it loses access to the logical drive.

Balancing traffic is unfortunately not always a trivial task. For example, if an application requires large disk space to be located and accessed in one chunk, it becomes harder to balance traffic by spreading the smaller volumes among controllers.

In addition, typically, the load across controllers and logical drives are constantly changing. The logical drives and data can be accessed at any given time depending on which applications and users are active during that time period, which is why monitoring the system is important.

The Performance Monitor provides data that 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. Identify actual I/O patterns in the individual logical drives and compare them with the expectations for an application. If a particular controller has considerably more I/O activity, consider moving logical drives to the other controller in the storage system.

A disparity in the total I/Os (workload) of controllers might be noticed. For example, the workload of one controller is heavy or is increasing over time, and that of the other controller is lighter or more stable. In this case, consider changing the controller ownership of one or more logical drives to the controller with the lighter workload.

Tip: Here are some guidelines for logical drives assignment and storage partitioning:

- ▶ Assign defined logical drives evenly across all controllers to balance controller utilization.
- ▶ Use the manual method of creating logical drives. This allows for greater flexibility of configuration settings, such as enclosure loss protection and utilizing both drive loops.
- ▶ If some logical drives are highly utilized, where possible, separate them by putting them on their own or another array. This will reduce disk contention for that 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 has to be replaced, the controller should be disabled 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.

Consideration: 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. Changing the controller ownership of the primary logical drive causes a corresponding controller ownership change of the secondary logical drive.

To change the preferred ownership from one controller to the other, highlight the logical drive, right-click, and select **Change** → **Ownership/Preferred Path**. Then select the controller to which the logical drive is to be moved. Depending on the current workload, the operation can take a while to finish.

4.10.6 Cache parameters

The Storage Manager utility enables various cache settings to be configured:

- ▶ Specify the DS5000 system wide settings:
 - Start and stop cache flushing levels (this setting will affect all arrays and logical drives created on the system)
 - Cache Block size
- ▶ Specify settings per logical drive:
 - Read caching
 - Cache read-ahead multiplier
 - Write caching or write-through mode (write caching disabled)
 - Enable or disable write cache mirroring

Figure 4-112 shows the typical values when using the Create Logical Drive Wizard. With the Storage Manager, cache settings can be specified for each logical drive independently, giving greater flexibility.

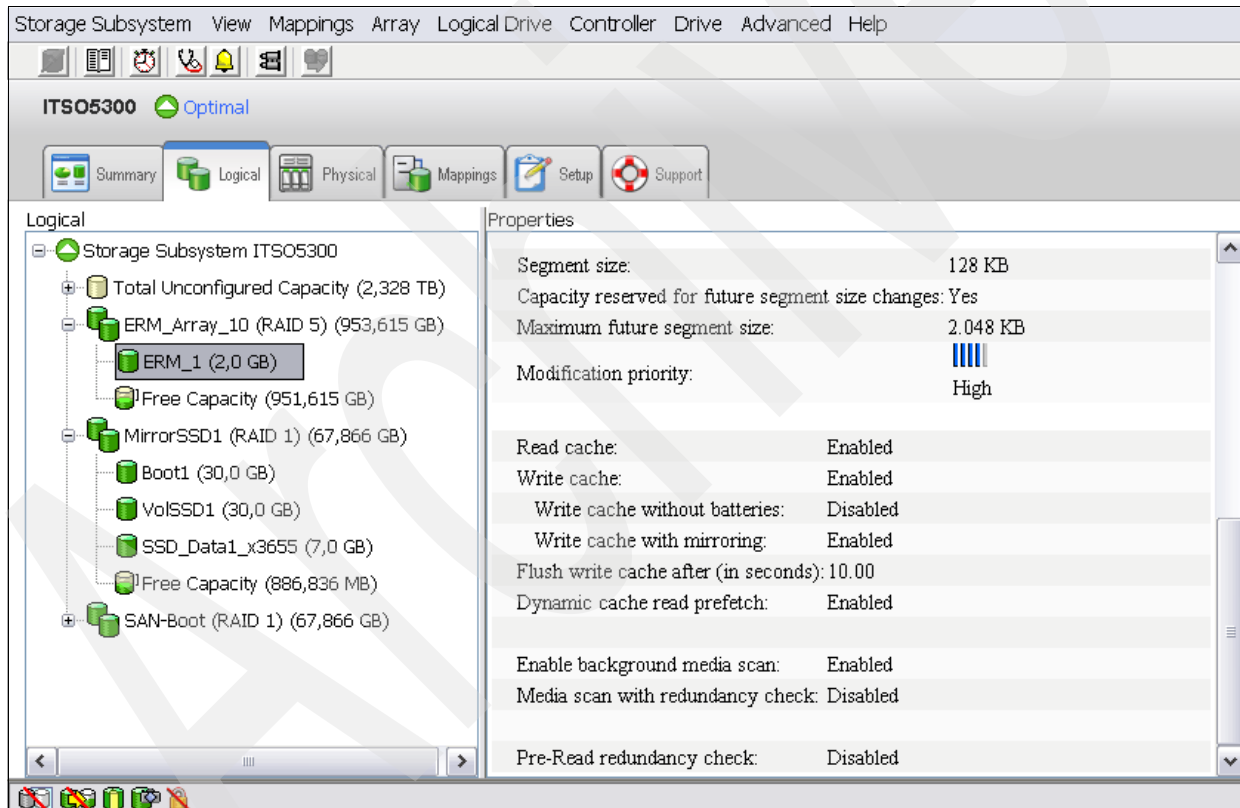


Figure 4-112 Default values used by the Create Logical Drive Wizard

Note: We recommend that the values are manually set during creation to suit the performance needs of the logical drive. These settings can be changed after logical drive creation for tuning purposes.

Changing cache settings

This can be achieved at logical drive and subsystem levels.

Logical drive

To locate this setting, highlight the logical drive in the Storage Manager menu and select **Logical Drive** → **Change** → **Cache Settings....** The window shown in Figure 4-113 will appear.

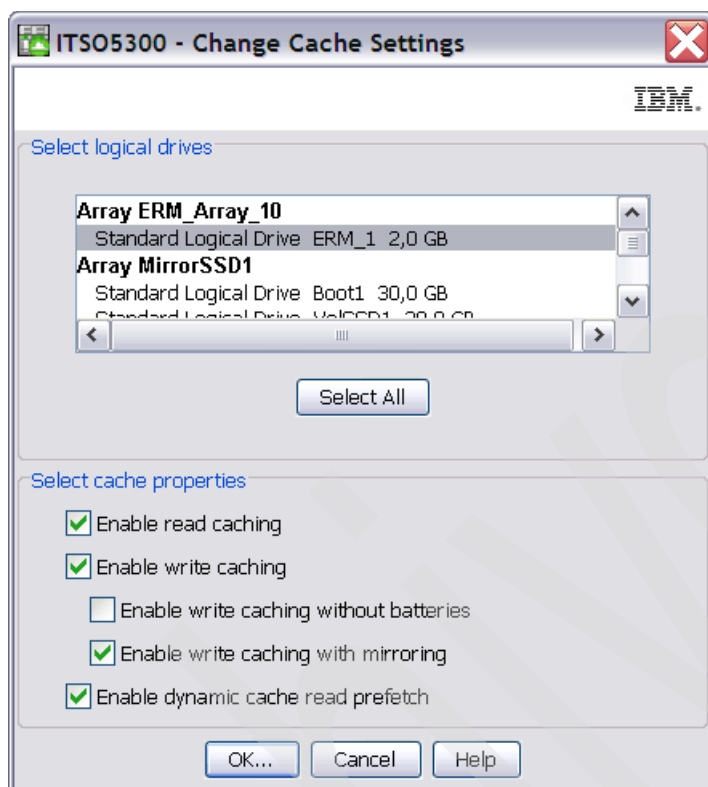


Figure 4-113 Logical drive changes to cache settings

These settings have a large impact on the performance of the DS5000 storage subsystem and on the availability of data. Be aware that performance and availability often conflict with each other. If maximum performance is required, in most cases availability might have to be compromised and vice versa.

The default settings are read and write cache for all logical drives, with cache mirroring to the alternate controller for all write data. The write cache is only used if the battery for the controller is fully charged. Read ahead is not normally used on the logical drives.

Read caching

Read caching allows read operations from the host to be stored in controller cache memory. If a host requests data that is not in the cache, the controller reads the needed data blocks from the disk and places them in the cache. Until the cache is flushed, any other requests for this data are fulfilled with the cache data instead of initiating another read operation to the disk.

Write caching

The write caching parameter enables the storage subsystem to cache write data instead of writing it directly to the disks. This can improve performance significantly, especially for environments with random writes, such as databases. For sequential writes, the performance gain varies with the size of the data written. If the logical drive is only used for read access, it might improve overall performance to disable the write cache for this logical drive, and no cache memory is reserved for the logical drive.

Write cache mirroring

The DS5000 storage subsystem write cache mirroring provides the integrity of cached data if a RAID controller fails. This is excellent from a high availability perspective, but it decreases performance. The data is mirrored between controllers across dedicated PCI Express buses. We recommend that the controller write cache mirroring be left enabled for data integrity reasons in case of a controller failure.

By default, a write cache is always mirrored to the other controller to ensure proper contents, even if the logical drive moves to the other controller. Otherwise, the data of the logical drive can be corrupted if the logical drive is shifted to the other controller and the cache still contains unwritten data. If you turn off this parameter, you risk data loss in the case of a controller failover, which might also be caused by a path failure in your fabric.

The cache of the DS5000 storage subsystem is protected by a battery against power loss. If the batteries are not fully charged, for example, just after powering on, the controllers automatically disable the write cache. If you enable the parameter, the write cache is used, even if no battery backup is available, resulting in a higher risk of data loss.

Write caching or write-through

Write-through means that writing operations do not use cache at all. The data is always going to be written directly to the disk drives. Disabling write caching frees up cache for reading (because the cache is shared for read and write operations).

Write caching can increase the performance of write operations. The data is not written straight to the disk drives; it is only written to the cache. From an application perspective, this is much faster than waiting for the disk write operation to complete. Therefore, a significant gain in application writing performance can be expected. It is the responsibility of the cache controller to eventually flush the unwritten cache entries to the disk drives.

Write cache mode appears to be faster than write-through mode, because it increases the performance of both reads and writes. But this is not always true, because it depends on the disk access pattern and workload.

A lightly loaded disk subsystem usually works faster in write-cache mode, but when the workload is high, the write cache can become inefficient. As soon as the data is written to the cache, it has to be flushed to the disks in order to make room for new data arriving into cache. The controller performs faster if the data goes directly to the disks. In this case, writing data to the cache is an unnecessary step that decreases throughput.

Dynamic cache read prefetch

Cache read-ahead, or “prefetch,” allows the controller, while it is reading and copying host-requested data blocks from disk into the cache, to copy additional data blocks into the cache. This increases the chance that a future request for data will be fulfilled from the cache. Cache read-ahead is important for multimedia applications that use sequential I/O.

This feature uses an automatic pre-fetching multiplier to maximize its cache hit efficiency and system performance. This will turn on monitoring of the I/O to the logical drive and enable the new algorithm to dynamically choose how much to read ahead. This simplifies the process for the administrator, as there is no need to manually set a specific value for the read ahead multiplier. The system will tune itself depending on the I/O characteristics. When sequential access is detected, the controller will automatically start using read ahead buffering. When random or non-sequential I/O is used, then it will stop using the read ahead buffer. To disable this feature, simply uncheck the **Dynamic Cache Read Prefetch** check box for the relevant logical drive.

Storage subsystem cache flushing levels

These two settings affect the way the cache controller handles unwritten cache entries. They are only effective when the write-back cache policy is configured. Writing the unwritten cache entries to the disk drives is called *flushing*. The start and stop flushing level values can be configured, as shown in Figure 4-114. They are expressed as percentages of the entire cache capacity. When the number of unwritten cache entries reaches the start flushing value, the controller begins to flush the cache (write the entries to the disk drives). The flushing stops when the number of unwritten entries drops below the stop flush value. The controller always flushes the oldest cache entries first. Unwritten cache entries older than 20 seconds are flushed automatically.

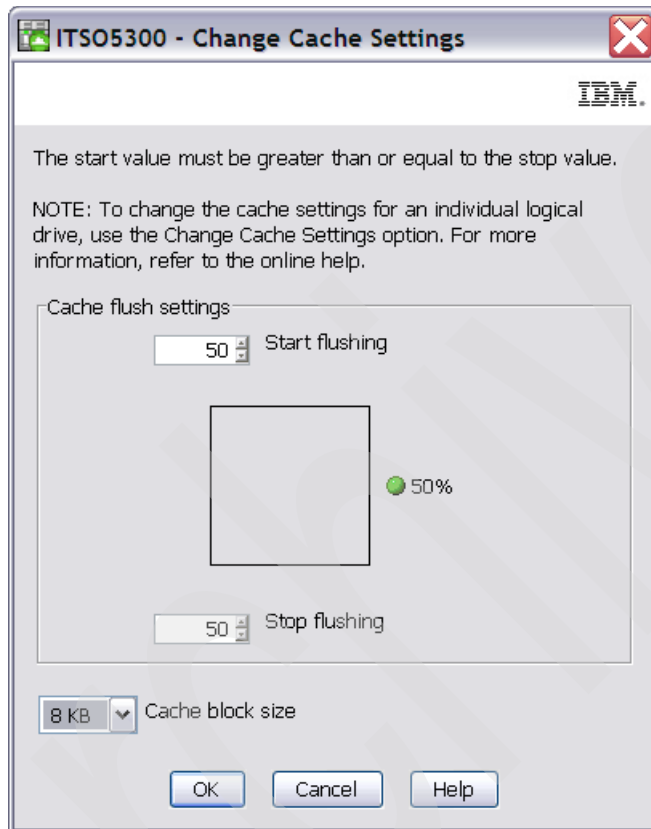


Figure 4-114 Storage subsystem cache settings

The default is the start flushing level and the stop flushing level is set to 80%. This means that the cache controller does not allow more than 80% of the entire cache size for write-back cache, but it also tries to keep as much of it as possible for this purpose. If such settings are used, a high number of unwritten entries in the cache is expected. This is good for writing performance, but be aware that it offers less data protection.

Performance tests have shown that it is a good idea to use similar values for start and stop flushing levels. If the stop level value is significantly lower than the start value, this causes a high amount of disk traffic when flushing the cache. If the values are similar, then the controller only flushes the amount needed to stay within limits.

Note: In our experience, a start flushing level of 50% and a stop flushing level of 50% is best for customer environments.

To locate this setting, highlight the logical drive in the Storage Manager menu and select **Storage Subsystem** → **Change** → **Cache Settings....**

Cache block size

This is the size of the cache memory allocation unit and can be either 4 K, 8 K, 16 K, or 32 K. This provides more flexibility in managing and customizing your cache setup on the DS5000 storage subsystem, based on specific requirements. By selecting the proper value for a particular situation, improvements can be seen in caching efficiency and performance. For example, if applications mostly access the data in small blocks up to 8 K, but a 16 K cache block size is used, then each cache entry block is only partially populated. Blocks will always occupy 16 K in cache to store 8 K (or less) of data. This means only up to 50% of cache capacity is effectively used to store the data. This inefficiency lowers the performance. For random workloads and small data transfer sizes, 4 K is better.

Alternatively, if the workload is sequential and a large segment size is used, as is the case with multimedia applications, it is better to use the larger cache block sizes of 16 or 32 K. A larger block size means a lower number of cache blocks and reduces cache delays. In addition, a larger cache block size requires fewer cache data transfers to handle the same amount of data.

4.10.7 Media scan

Media scan is a background process enabled by default that checks logical drives over hard disk drives for defects by reading the raw data from the disk and writing it back. This detects possible problems caused by bad sectors of the physical disks before they could eventually disrupt normal data reads or writes. This process is sometimes known as *data scrubbing*.

Important: Media scan is an option available for logical drive space configured on hard disk drives, not over Solid State Drives. Unused hard disks or hot spares are not scanned.

The media scan runs on all logical drives in the storage subsystem that meet the following conditions:

1. The logical drive is in an optimal status.
2. The logical drive is not defined over Solid State Drives.
3. There are no modification operations in progress.
4. The Media Scan parameter is enabled.

The media scan continuously runs in the background, using spare cycles to complete its work. The default media scan is for a scan every 30 days, that is, the maximum time the media scan has to complete the task. During the scan process, the DS5000 storage subsystem calculates how much longer the scan process will take to complete, and adjusts the priority of the scan to ensure that the scan completes within the time setting allocated. Once the media scan has completed, it starts over again and resets its time for completion to the current setting. This media scan setting can be reduced. However, if the setting is too low, priority is given to the media scan over host activity to ensure that the scan completes in the allocated time. This scan can impact performance, but will improve data integrity in the long term.

The media scan is enabled for the entire storage subsystem. The system-wide enabling specifies the duration over which the media scan runs, which by default is 30 days. By default, the media scan process runs without checking redundancy data. You can optionally specify whether to do a redundancy check or to stop media scan.

A media scan can be considered a surface scan of the hard drives, and a redundancy check scans the blocks of a RAID 3, 5, or 6 logical drive and compares it against the redundancy data. In the case of a RAID 1 logical drive, the redundancy scan compares blocks between copies on mirrored drives.

Note: A media scan is only capable of resolving media errors and data or parity mismatches. A media scan does not attempt to resolve any other sort of error occurring during I/O operations.

We have seen no effect on I/O when we use a 30-day setting unless the processor is utilized in excess of 95%. The length of time that it takes to scan the logical drives depends on the capacity of all the logical drives on the system and the utilization of the controller.

Important: The media scan must be enabled for the entire storage subsystem and enabled on each logical drive within the storage subsystem to protect the logical drive from failure due to media errors. This is the default and recommended configuration.

Table 4-11 shows the errors and describes several of the actions that the DS5000 storage subsystem takes as a result of a media scan and redundancy check operations.

Table 4-11 Media scan errors

Reported error	Description	Result
Unrecovered media error	The data cannot be read on its first attempt, or on any subsequent retries.	With redundancy check: Data is reconstructed and scanned again. Without redundancy check: No error correction.
Recovered media error	The drive cannot read the requested data on its first attempt, but succeeded on a subsequent attempt.	Data is written to drive and verified.
Redundancy mismatches	Redundancy errors are found.	The first 10 redundancy mismatches found on a logical drive are reported. Operating system data checking operations should be executed.
Unfixable error	The data cannot be read, and parity or redundancy information cannot be used to regenerate it.	An error is reported.

To locate this setting, highlight the logical drive in the Storage Manager menu and select **Storage Subsystem** → **Change** → **Media Scan Settings**. The window shown in Figure 4-115 opens.

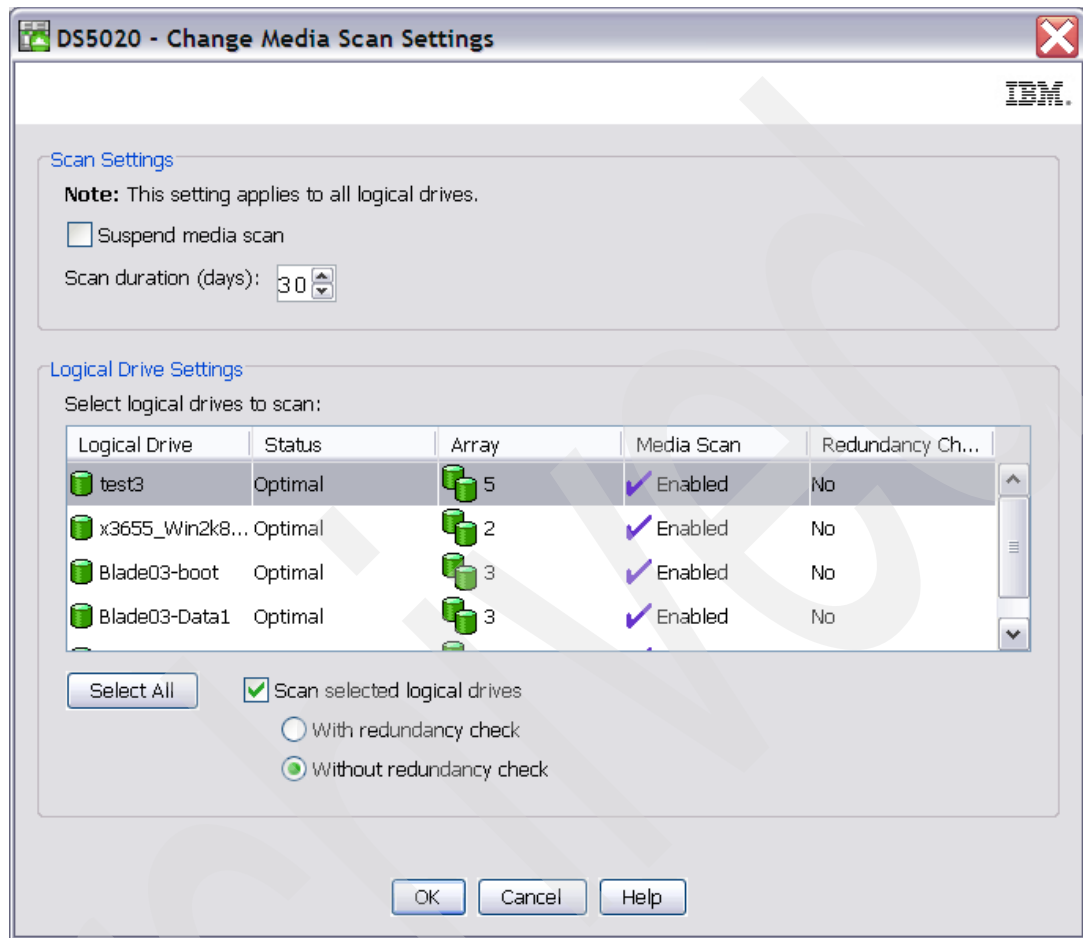


Figure 4-115 Media scan settings for storage subsystem

4.10.8 Failover alert delay

Storage Manager provides alert notification on ADT-induced logical drive ownership changes. The logical drive transfer alert notification is issued for any instance of a logical drive owned by a non-preferred controller, whether ADT is enabled or not, and is in addition to any informational or critical event already logged within the ADT or RDAC context.

A failover alert delay can be specified that lets you delay the logging of a critical event if the multipath driver transfers logical drives to the non-preferred controller. If the multipath driver transfers the logical drives back to the preferred controller within the specified delay period, no critical event is logged. If the transfer exceeds this delay period, then a logical drive-not-on-preferred-path alert is issued as a critical event. This option also can be used to minimize multiple alerts when many logical drives fail over because of a system error, such as a failed host adapter.

Attention: Whenever a logical drive not-on-preferred-path condition occurs, only the alert notification is delayed. A needs attention condition is raised immediately.

To make the best use of this feature, set the failover alert delay period such that the host driver failback monitor runs at least once during the alert delay period. Note that a logical drive ownership change might persist through the alert delay period, but correct itself before you can inspect the situation. In such a case, a logical drive-not-on-preferred-path alert is issued as a critical event, but the array will no longer be in a needs-attention state.

Important:

- ▶ The failover alert delay option operates at the storage system level, so one setting applies to all logical drives.
- ▶ The failover alert delay option is reported in minutes in the storage system profile as a storage system property.
- ▶ The default failover alert delay interval is five minutes. The delay period can be set within a range of 0 to 60 minutes. Setting the alert delay to a value of zero results in instant notification of a logical drive not on the preferred path. A value of zero does not mean alert notification is disabled.
- ▶ The failover alert delay is activated after controller start-of-day completes to determine if all logical drives were restored during the start-of-day operation. Thus, the earliest that the not-on-preferred path alert will be generated is after boot up and the configured failover alert delay time.

To change this setting, from the Subsystem Management window, select **Storage Subsystem** → **Change** → **Failover Alert Delay**. The window shown in Figure 4-116 opens.

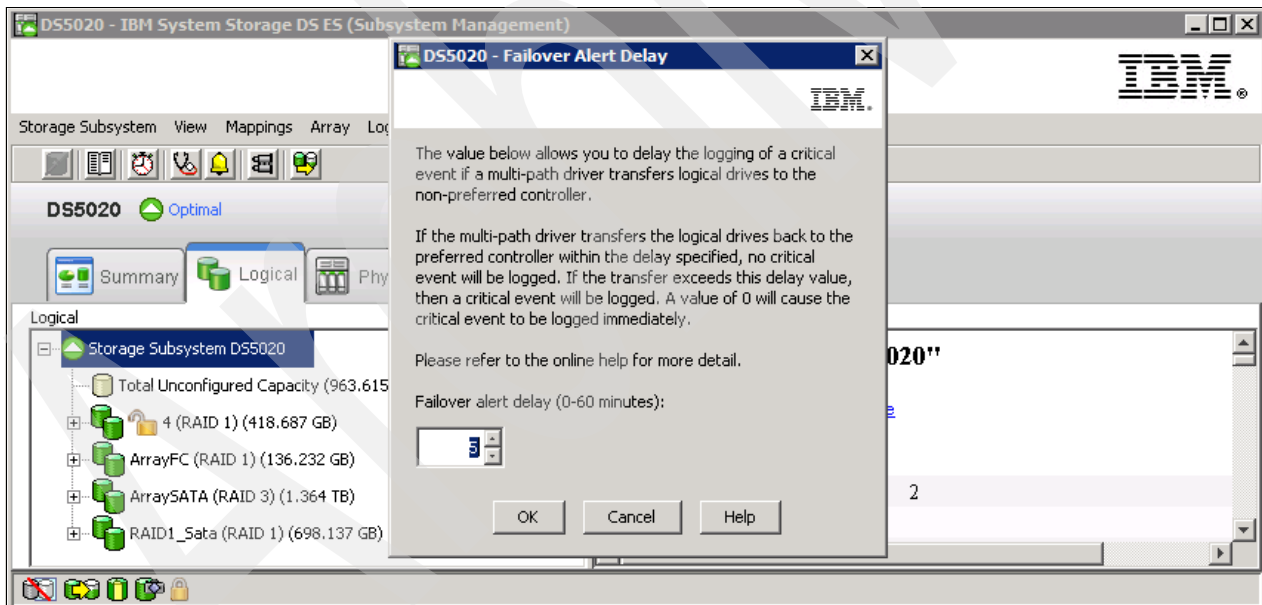


Figure 4-116 Storage subsystem failover alert delay

4.10.9 Persistent reservations

Persistent reservations are a SCSI-3 feature for restricting access to storage media, based on the concept of reservations that the host can establish and manipulate. Earlier versions of SCSI provide a simple reservation capability through the RESERVE and RELEASE commands. SCSI-3 persistent reservations provide a significant superset of the earlier capability. Improvements that come with persistent reservations include:

- ▶ A well-defined model for reserving across multiple host and target ports
- ▶ Levels of access control, for example, shared reads, exclusive writes, exclusive reads, and writes
- ▶ Ability to query the storage system about registered ports and reservations
- ▶ Provisions for persistence of reservations through power loss at the storage system

A logical drive reservation is a feature of the cluster software (the actual reservation of the logical drive is handled by the host application) that allows one or more host ports to reserve a logical drive, thus preventing other host ports from accessing the same logical drive.

Unlike other types of reservations, a persistent reservation reserves across multiple host ports, provides various levels of access control, offers the ability to query the storage system about registered ports and reservations, and, optionally, provides for persistence of reservations in the event of a storage system power loss.

The benefits of the persistent reservations feature is that it allows the DS5000 storage subsystem to integrate with cluster solutions that use shared logical drives for increased availability, scalability, and performance.

The Persistent Reservation options provided with the DS Storage Manager enables you to view and clear logical volumes reservations and associated reservations.

To locate this setting, highlight the logical drive in the Storage Manager menu and select **Advanced > Maintenance > Persistent Reservations**. The window shown in Figure 4-117 opens.

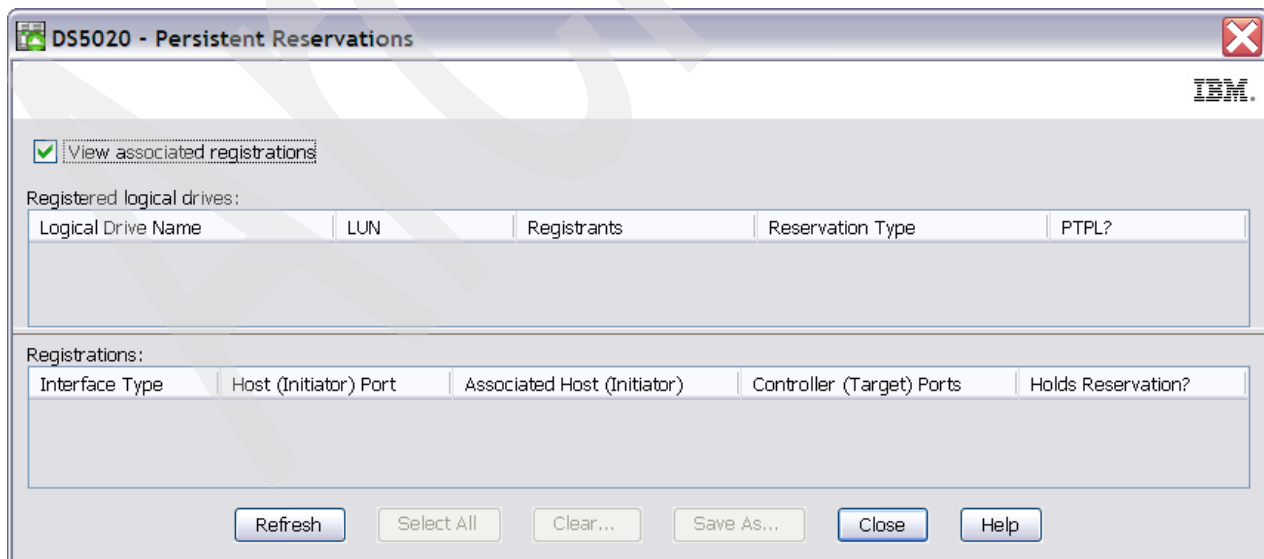


Figure 4-117 View of persistent reservations

4.10.10 Automatic firmware synchronization

The Automatic Code Synchronization (ACS) feature ensures that both controllers within a storage subsystem are executing the same version of the controller firmware. The primary purpose of this feature is to reconcile the possible difference in controller firmware version present on a replacement controller. ACS is also used to address inconsistent versions of firmware that can occur without controller replacement if a firmware upgrade is interrupted.

The replacement of both controllers is not a recommended operation, because ACS is intended to use the firmware image on the remaining native controller to resolve the inconsistency. ACS does address the replacement of both controllers. In this case, ACS synchronizes to the newer of the two firmwares.

A failure to successfully transfer the incumbent image from the native to the foreign controller results in the native controller holding the alternate controller in reset.

The ACS feature behavior is based on three key persistent data representations:

- ▶ The serial number of each controller.
- ▶ The serial numbers of the last known native controllers. The identities of the last known controllers are contained within the metadata on the native drive set.
- ▶ The firmware version number of the incumbent firmware. This is the firmware associated with the metadata on the native drive set. The version of the incumbent firmware is stored within the drive metadata.

The controller firmware uses this information to determine what ACS action is required. If the serial number of a controller does not match the serial number for the slot in which it resides, this controller is considered to be *foreign* for the purposes of ACS, and a firmware synchronization occurs.



Disk Security with Full Disk Encryption drives

Disk Security is a new feature introduced with the IBM System Storage DS5000 storage subsystem that uses the newly available Full Disk Encryption (FDE) disk drives. It is supported by the latest level of the DS5000 firmware (Version 7.60) and Storage Manager V10.60. This chapter discusses how this new feature can add a greater level of security while your data resides on disk, what it does, the various components of the feature, and how to implement it.

The Disk Security premium feature requires security capable drives. A security capable drive encrypts data during writes and decrypts data during reads. Each security capable drive has a unique drive encryption key. When a security capable drive has the security enabled, the drive requires the correct security key from the DS5000 for authentication before allowing reading or writing the data. This is managed on each of the DS5000 controllers by the IBM Disk Encryption Storage Manager. All of the drives in the DS5000 share the same security key and the shared security key provides read and write access to the drives, and the drive encryption key on each drive is used to encrypt the data.

5.1 The need for encryption

Data security breaches are becoming increasingly common, and the threat of unauthorized access to sensitive data and intellectual property is growing. Data security is now a common component of the corporate landscape, mostly driven by regulatory compliance, and security efforts can often be fragmented.

At some point, all drives are out of an organization's control and vulnerable to a breach. For example, with "re-purposing", decommissioning, or disposal of individual disks, common methods of security are often insufficient when:

- ▶ Deleted files can be reinstated.
- ▶ Drive disposal methods are subject to human error.
- ▶ Password protected data can still be accessed by a skilled individual.
- ▶ Disk reformatting erases data on disks but information can still be recovered.¹

In each case, a risk is introduced where legible data might be recovered from disk. This can be made significantly more difficult if Disk Security on a DS5000 storage subsystem is employed.

Important: FDE and drive-level encryption technology is a new and additional level of data protection, but it does not replace the access controls and security processes that exist; rather, it complements them.

5.1.1 Encryption method used

The FDE drives have encryption hardware, and can perform symmetric encryption and decrypting of data at full disk speed with no impact on performance. The disk encryption hardware is used in conjunction with IBM Disk Encryption Storage Manager on the DS5000 storage subsystem. It uses asymmetric encryption to encrypt and decrypt the data key. IBM Disk Encryption Storage Manager will generate encryption and decryption keys that are used to lock each FDE drive.

Without these IBM Disk Encryption Storage Manager managed keys, the user (authorized or unauthorized) can no longer decrypt the data on disk.

Important: Should these keys and all copies be lost, then the encrypted data on disk cannot be decrypted and is therefore considered lost.

¹ Some utilities used to "erase" data from disks and arrays are not fully successful when such data can still be recovered using forensic techniques

Figure 5-1 shows the relationship of the IBM Disk Encryption Manager and an individual FDE drive with encryption enabled.

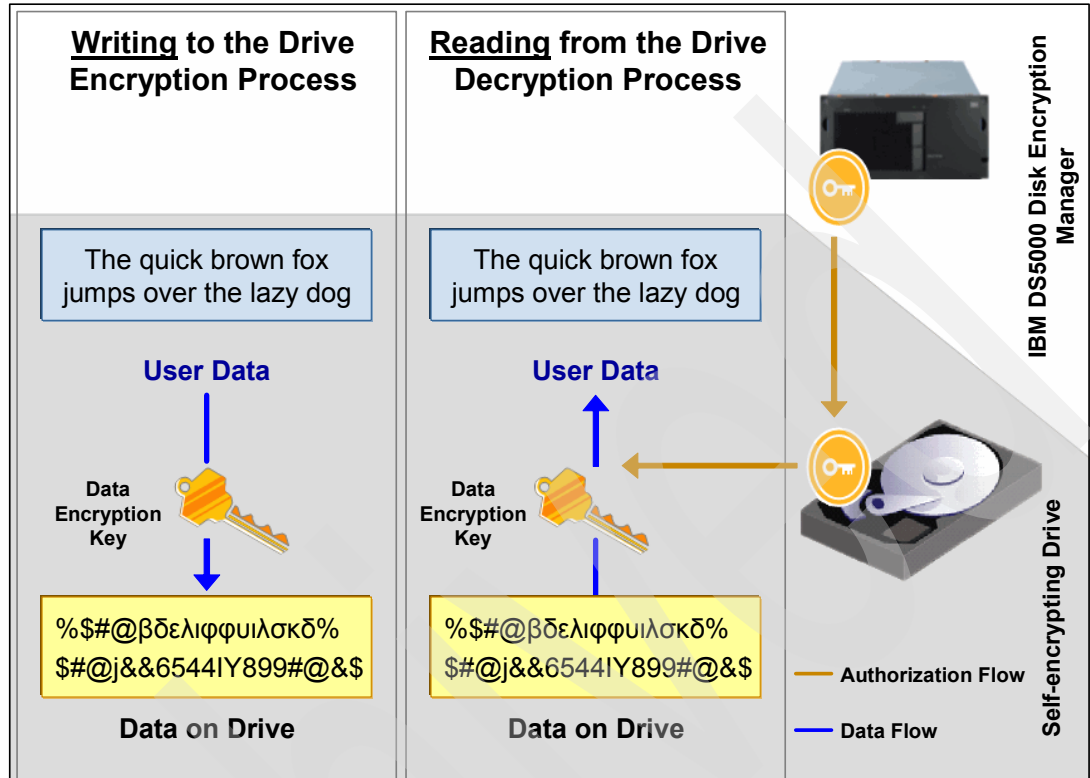


Figure 5-1 DS5000 Disk Encryption Manager using self encrypting FDE drives

With this relationship, the correct keys, and authentication, the FDE drive will encrypt data written and decrypt data read from it. But if the disk is removed and data on the disk is attempted to be read, as shown in Figure 5-2, the user will not have the appropriate authorizations, as data cannot be read from or written to the drive without authenticating with the DS5000 Disk Encryption Manager, which will unlock the drive.

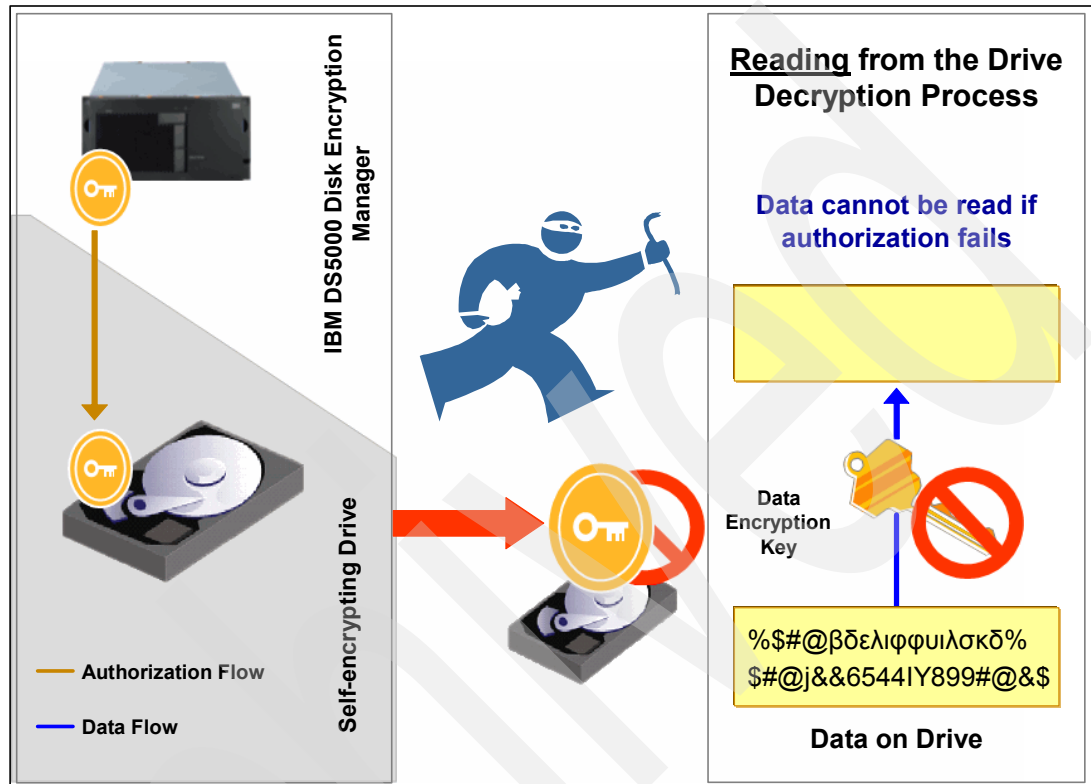


Figure 5-2 Unauthorized access to the drive results in the data remaining encrypted

5.2 Disk Security components

There are a number of new components to this new feature that are detailed in this section. All of these features are managed by the Storage Manager (V10.6.x and higher).

5.2.1 DS5000 Disk Encryption Manager

The Disk Encryption Manager on the DS5000 system maintains and controls the key linkage and communications with FDE drives. It is included with the firmware (V7.60) and Storage Manager (V10.6). It:

- ▶ Provides all the management tools necessary to quickly and simply enable and secure FDE drives.
- ▶ Establishes and manages a single authorization scheme for all the FDE drives in a DS5000 storage subsystem.
 - Places FDE drives in a secured state.
 - Defines secure arrays.
 - Supports the decommissioning or re-purposing of drives with Instant Secure Erase.

With this function you can record both the security key ID, pass phrase, and the secure file location in a safe place.

- ▶ Using the FDE drive, it generates and encrypts a security key:
 - Creates a unique security key ID that is paired with the security key.
 - Adds a randomly generated number.
 - The security key ID is saved. This folder location will be needed whenever a security operation requires the key ID (for example, when a drive powers up).
 - Creates a backup of the security key and the security key identifier.
 - A secure backup is provided in which the security key and the security key identifier are encrypted utilizing a user-selected pass phrase.

5.2.2 Full Data Encryption (FDE) disks

FDE drives are required to enable Disk Security. Currently, you must use Fibre Channel (FC) disks with a speed of 15,000 rpm. These disks include:

- ▶ Encryption Capable 4 GBps FC, 146.8 GB/15K: 41Y8461 146 GB ST3146356FC 9CG004-039
- ▶ Encryption Capable 4 GBps FC, 300 GB/15K: 41Y8462 300 GB ST3300056FC 9CK004-039
- ▶ Encryption Capable 4 GBps FC, 450 GB/15K: 41Y8463 450 GB ST3450056FC 9CN004-039

5.2.3 Premium feature license

The DS5000 requires that the Drive Security premium feature be installed and enabled for Disk Security to function. See 4.2, “Planning for premium features” on page 124 for details about this topic.

5.2.4 Keys

There are two types of keys that are used with Drive Security and FDE drives:

- ▶ The *encryption key* is generated by the drive and never leaves the drive, so it always stays secure. It is stored in encrypted form and performs symmetric encryption and decryption of data at full disk speed with no impact on disk performance. Each FDE drive uses its own unique encryption key that is generated when the disk is manufactured and regenerated when required by the storage administrator using the DS5000 Disk Encryption Manager.
- ▶ The *lock key* or *security key* is a 32 byte random number that authenticates the drive with the DS5000 Disk Encryption Manager using asymmetric encryption for authentication. When the FDE drive is secure “enabled”, it has to authenticate with the Disk Encryption Manager or it will not return any data and remains locked. After the drive has been authenticated, access to the drive operates like any other disk drive. One security key is created for all FDE drives on the DS5000 storage subsystem, where it is generated, encrypted, and hidden in the subsystem (NVS RAM). The authentication only occurs typically after the FDE has powered up, where it will be in a “locked” state.

If the lock key is not initially established between the DS5000 Disk Encryption Manager and the disk, then the disk is considered unlocked with access unlimited, as per a non-FDE drive.

5.2.5 Security key identifier

For additional protection, the security key that is used to unlock FDE drives is not visible to the user. The security key identifier is used to refer to a security key instead. You can see the security key identifier during operations that involve the drive security key backup file, such as creating or changing the security key. The security key identifier is stored in a special area of the disk; it can always be read from the disk and can be written to the disk only if security has been enabled and the drive is unlocked.

The security key identifier field in the FDE Drive Properties window, shown in Figure 5-3, includes a random number that is generated by the controller when you create or change the security key. One security key is created for all FDE drives on the storage subsystem.

Note that the Security Capable and Secure fields in the Drive Properties window show whether the drive is secure capable and whether it is in Secure (Yes) or Unsecured (No) state. The example shows that the drive is both capable (FDE) and enabled.

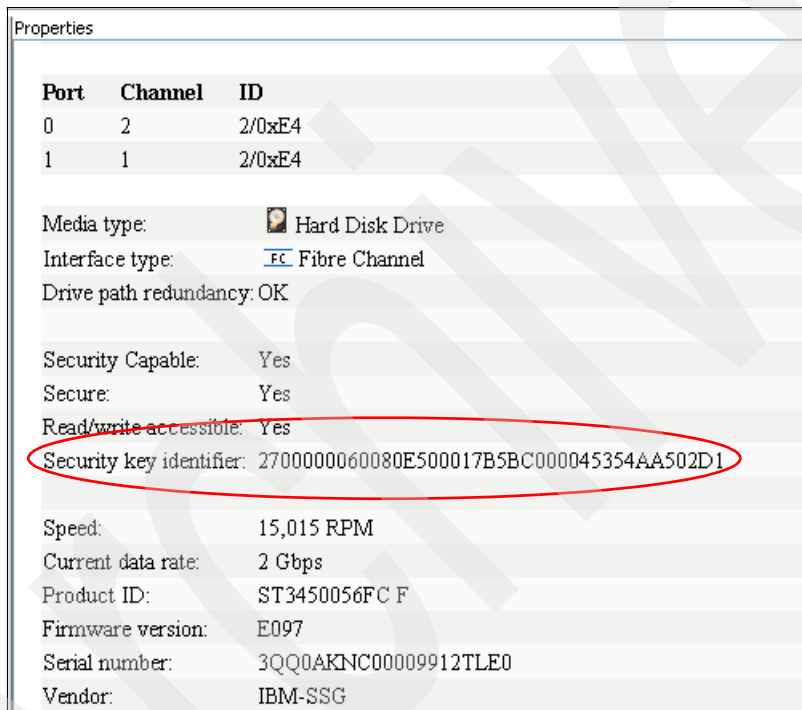


Figure 5-3 FDE drive properties showing security ID and status

5.2.6 Passwords

For Disk Security to be enabled, the DS5000 has to have the administration pass phrase or password set. The password must be “strong” and not easy to guess. A check is made on the password and if the system does not consider it to be strong enough when you log in or are prompted for the password, the message shown in Figure 5-6 on page 266 will appear. It will include suggestions about how the password can be made stronger.

The security key and the security key identifier are encrypted using a different password or pass phrase when the key is created or changed (see 5.3.2, “Secure key creation” on page 266 and 5.4.1, “Changing the security key” on page 270). The array then returns a file that is called a *blob*, or key backup. If the array needs that key later, you give the blob and pass phrase to the GUI, which sends it down to the array where the original key is decrypted.

The user-specified alphanumeric character string is not stored anywhere on the DS5000 or in the security key backup file.

5.3 Setting up and enabling a secure disk

This section shows a step-by-step process to create a key and file on the IBM Disk Encryption Storage Manager of the DS5000. It will then show how to enable a previously configured array that has FDE drives.

5.3.1 FDE and premium feature check

There are a number of checks to make prior to key creation. First, you must check that the premium feature key has been applied to the system. To do this task, from the Storage Manager window, select **Storage Subsystem** → **Premium Features**.

Figure 5-4 shows that the Drive Security premium feature key has been obtained and successfully installed. This premium feature key is installed the same as any other premium feature key (see 4.2, “Planning for premium features” on page 124 for details about how to install premium feature keys).

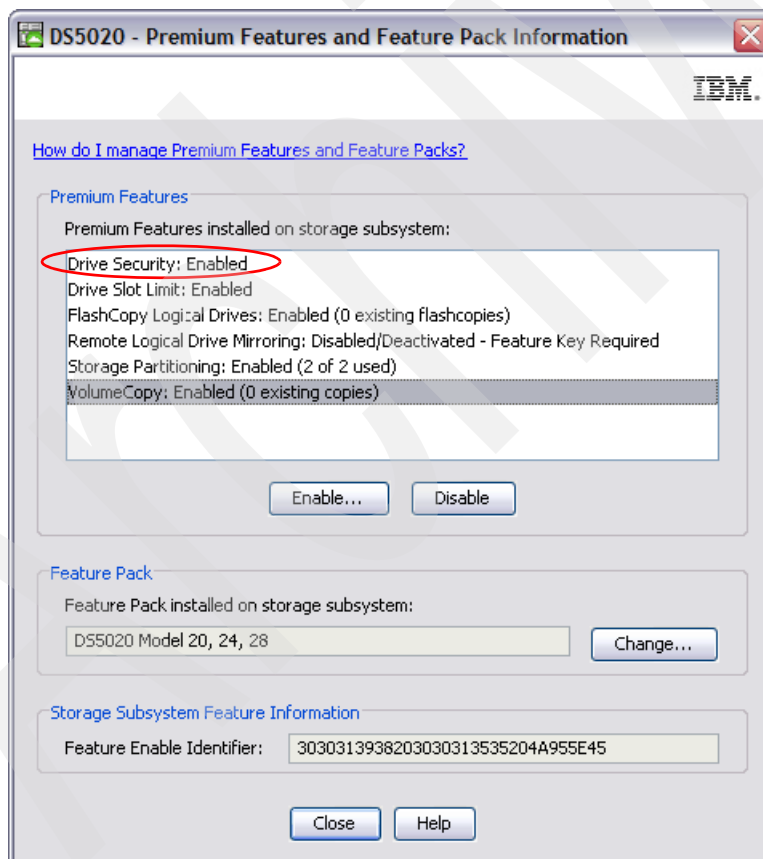


Figure 5-4 The correct premium feature has been obtained and installed

The next check is to ensure that the physical disks are FDE drives and are secure capable. In Figure 5-5, you can see that an array has been created where disk security is not enabled, as indicated by the unlocked padlock. If you click the array and then right-click and select **View Associated Physical Components**, you can view the disks. If you click the **Show FDE** button, it will display all the FDE drives and confirm that all the drives in the array are all secure capable.

The array is already being used and has several logical drives configured and consists of all FDE drives; therefore, it can be enabled.

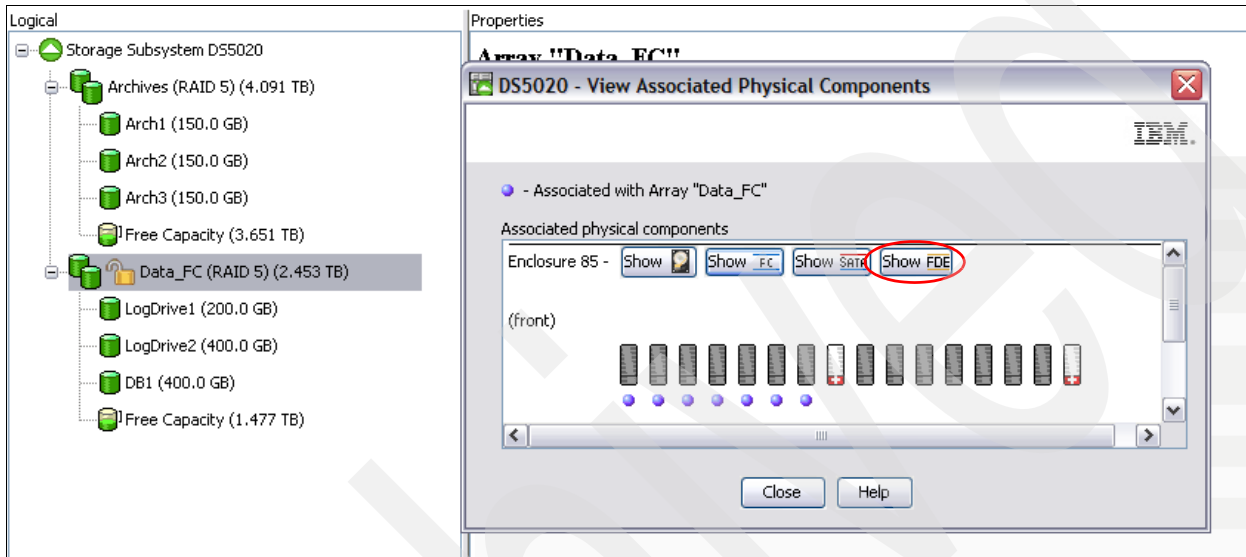


Figure 5-5 Array created with FDE drives with Disk Security disabled

5.3.2 Secure key creation

To create the secure key, select, in the top left corner of the IBM System Storage DS ES window, **Storage Subsystem** → **Drive Security** → **Create Security Key**.

You must have your DS5000 subsystem password set to proceed, and you might see the message shown in Figure 5-6 if the existing password be considered too weak.

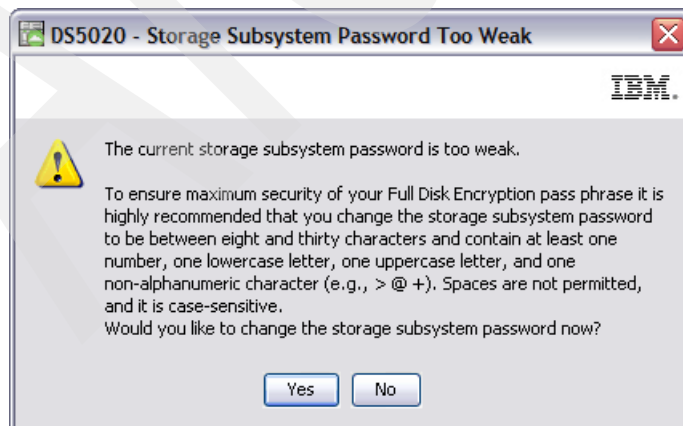


Figure 5-6 Warning regarding the weak subsystem password

The window shown in Figure 5-7 opens, where you need to complete the fields.

DS5020 - Create Security Key

IBM

Security key identifier

The security key identifier is paired with the security key to help you remember which key to use for secured operations. The security key identifier can be left blank or you may type up to 189 alphanumeric characters. The system will add the storage subsystem world-wide identifier and a randomly generated number to what you type in the field. You will have the opportunity to record the final security key identifier later.

Security key identifier:

Usethisone7_9_2009

File save location

File name:

T:\ST-9D24\Keys D5\SecKey_DS5020_7-9-2009.slk

Pass phrase

The pass phrase is required to perform security operations. It must be between eight and thirty-two characters and contain at least one number, one lowercase letter, one uppercase letter, and one non-alphanumeric character (e.g., < > @ +). Spaces are not permitted, and it is case-sensitive.

Pass phrase:

Confirm pass phrase:

Figure 5-7 Requirements displayed for the security key creation

The key location default is in the user's local PC directory. We strongly advise that the key be copied and kept in a safe location.

Tip: The best practice is to store the security key file with your key management policies along with the pass phrase. It is important to record and remember where this file is stored because the security key file is required along with pass phrase when a drive is moved from one storage subsystem to another or when both controllers in a storage subsystem are replaced at the same time.

When the process is complete, the key is created and located as specified; also, as shown in Figure 5-8, the security identifier is displayed. There are three items that must be kept secure in order to manage any changes to the FDE drives when they are encryption enabled.

- ▶ Security key file created from the process
- ▶ Security key identifier created from the process
- ▶ Password used



Figure 5-8 Security key creation process complete

The key authorizations now generated are synchronized between both controllers in the DS5000 storage subsystem. With these authorizations in place, arrays on the FDE drives in the storage subsystem can be secured.

5.3.3 Enable Disk Security on array

Now that the keys have been created, Disk Security can now be enabled on the array. Right-click the array and select **Secure Drives**, as shown in Figure 5-9.

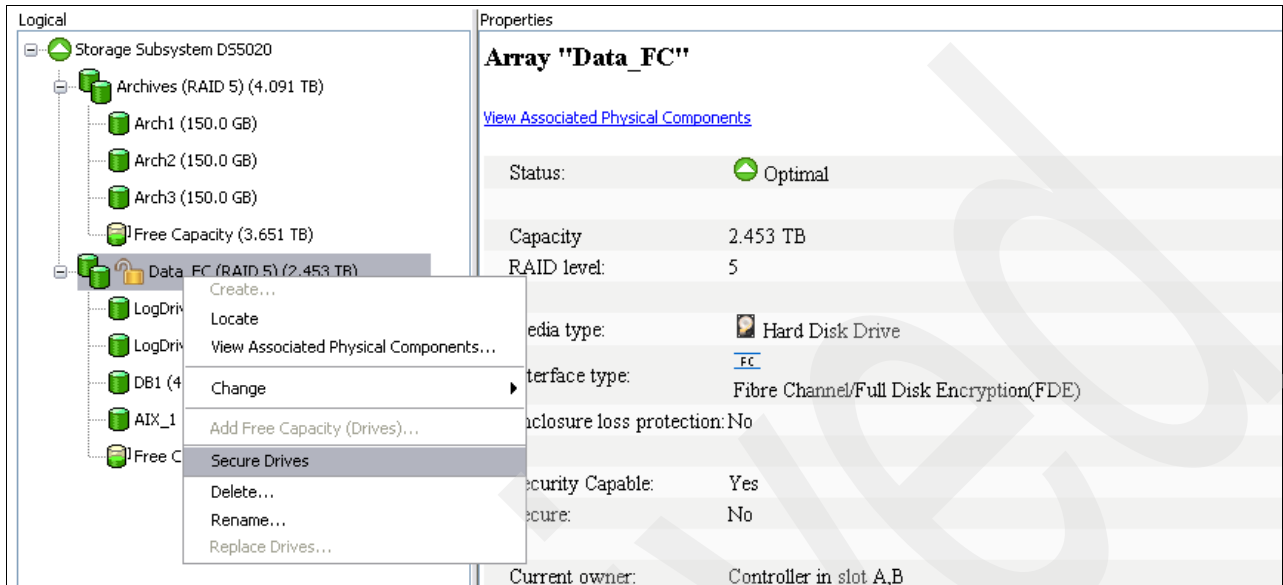


Figure 5-9 Secure all drives in the array

You will then be prompted to confirm the Secure Drives on the array, as shown in Figure 5-10.

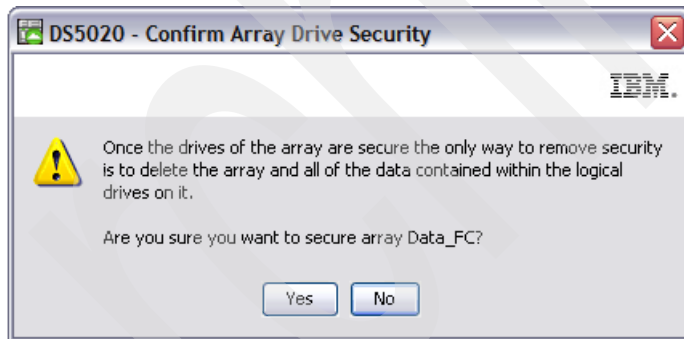


Figure 5-10 Confirm Array Drive Security

The array is now secured, as indicated by the padlock in a locked position, as shown in Figure 5-11.

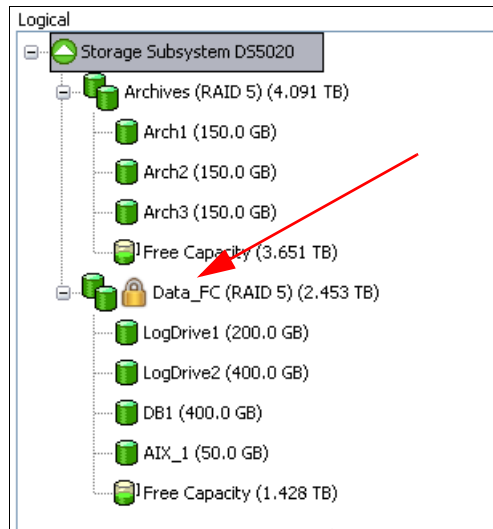


Figure 5-11 Array is now secured with Disk Security enabled

5.4 Additional secure disk functions

In the following sections, we discuss the following functions:

- ▶ Changing the security key
- ▶ Saving the security key file
- ▶ Secure disk erase
- ▶ FDE drive status
- ▶ Hot spare drives

5.4.1 Changing the security key

The security key can be changed if the details of the existing key be corrupted or the pass phrase forgotten, provided that there are no outstanding Secure Disk communications between the FDE drives and Disk Encryption Manager (for example, if a disk is in a “locked” state). Because the disk encryption key never leaves the disk, you might want to periodically change the encryption key, the way a user might periodically change the administrative password to an operating system. This depends on the organization’s security guidelines.

The process to change the security key is very similar to that of creating it initially. To change the key, select, in the top left hand corner of the Storage Manager menu, **Storage Subsystem** → **Drive Security** → **Change Security Key**.

The window shown in Figure 5-12 opens and you are prompted to add a security identifier (optional), a location to store the key file in, and a pass phrase.



Figure 5-12 Change Security Key options

The new security key is generated by the controller firmware and is hidden in the storage subsystem. The new security key replaces the previous key that was used to unlock the security-enabled FDE drives in the storage subsystem. The controller negotiates with all of the security-enabled FDE drives for the new key.

The original security key is also stored in the storage subsystem for protection in case something prevents the controllers from completing the negotiation of the new security key with the security-enabled FDE drives (for example, loss of storage subsystem power during the key change process). If this happens, you must change the security key so that only one version of the security key is used to unlock all drives in a storage subsystem. The original key is stored in the storage subsystem only. It cannot be changed directly or exported to a security key backup file.

When the security key has been successfully changed, a confirmation window opens, as shown in Figure 5-13, where the new key file location and security key identifier are shown.



Figure 5-13 Change Security Key Complete confirmation window

5.4.2 Save security key file

This action will save a backup of the security key file and will require the original pass phrase in order to copy it. It can therefore also be used to verify that the pass phrase stored is correct. To save the security key file, select, from the top left hand corner of the Storage Manager menu, **Storage Subsystem** → **Drive Security** → **Save Security Key File**.

You will be prompted for the location to store the file and the pass phrase used to create or change the existing security key file, as shown in Figure 5-14. The DS5000 Disk Encryption Manager uses the pass phrase to encrypt the security key before it exports the security key to the security key backup file.



Figure 5-14 Save Security Key File window

5.4.3 Secure erase

Secure erase provides a higher level of data erasure than other traditional methods. When you initiate secure erase with the DS5000 Disk Encryption Manager, a command is sent to the FDE drive to perform a “cryptographic erase”. This erases the existing data encryption key and then generates a new encryption key inside the drive, making it impossible to decrypt the data. Drive security becomes disabled and must be re-enabled if it is required again.

The secure erase process is shown in Figure 5-15.

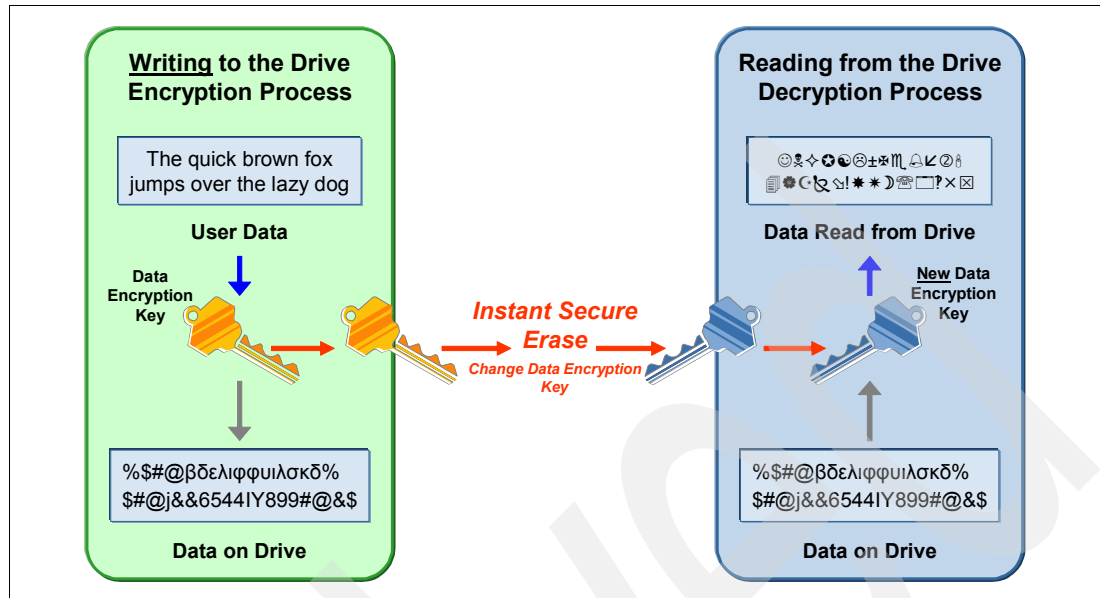


Figure 5-15 Secure erase process

Warning: All data on the disk will be permanently and irrevocably erased when the secure erase operation is completed for a security-enabled FDE drive. Do not perform this action unless you are sure that you want to erase the data, as there is no recovery.

Secure erase can only be performed on drives that are not allocated to an array. The process is also referred to as re-provisioning, where:

- ▶ The FDE drive becomes fully reusable.
- ▶ The drive can be reused in secure or non-secure applications.
- ▶ Previous data and keys are not accessible.
- ▶ It executes in less than a second.
- ▶ It returns the drive to the original factory state.

5.4.4 FDE drive status

The FDE drives have a status indicating whether the disk can be accessed. The statuses are:

- ▶ Locked
 - The drive is security capable.
 - The drive has security enabled.
 - The lock key has not been supplied to the drive.
 - Data cannot be read or written from drive.
- ▶ Unlocked
 - The drive is security capable.
 - The drive has security enabled.
 - The lock key has been supplied to the drive.
 - Data can be read or written from drive.

The locked state will rarely be seen, that is, only when the array containing the disks have been moved to another DS5000 or controllers have been replaced. The drive becomes locked whenever the disk is powered down. The drive will remain unlocked during firmware upgrades or while other components are being replaced. When the drive is powered on, the status will be locked. If it detects a security key identifier, then it will remain locked until it has successfully authenticated with the DS5000 Disk Encryption Manager.

5.4.5 Hot spare drive

If a disk drive fails in the DS5000 storage subsystem, the controller uses redundant data to reconstruct the data on the failed drive on a global hot-spare drive. The global hot-spare drive is automatically substituted for the failed drive without intervention. When the failed drive is eventually replaced, the data from the hot-spare drive is copied back to the replacement drive.

Hot-spare drives must meet the array hot-spare requirements. The following drive types are required for hot-spare drives when secure-capable arrays are configured. If a drive does fail, the Storage Manager automatically determines which hot-spare drive to substitute according to the type of the failed drive:

- ▶ For an array that has secured FDE drives, the hot-spare drive must be an unsecured FDE drive of the same or greater capacity. After the unsecured FDE hot-spare drive is used as a spare for a failed drive in the secured RAID array, it is security enabled.
- ▶ For an array that has FDE drives that are not secured, the hot-spare drive can be either an unsecured FDE drive or a non-FDE drive.
- ▶ An unconfigured secured FDE drive cannot be used as a global hot-spare drive. If a global hot spare is a secured FDE drive, it can be used as a spare drive only in secured arrays.

5.5 Migrating secure disk arrays

This section will detail how to migrate an array with Disk Security enabled to another DS5000. The process consists of exporting the array on the source DS5000 and importing the array on the target DS5000 after the disk have been physically moved. User data remains intact on the disks because configuration metadata is stored on every drive in the DS5000. The export process is the same as an ordinary array, but there are some extra steps to be carried out for the locked drives of the secure array when importing.

5.5.1 Planning checklist

This list is displayed in a window of the migration wizard (Figure 5-17 on page 277). It is important to follow this list when planning an export and import of an array.

On the source DS5000:

- ▶ Save the DS5000 configuration by selecting **Storage Subsystem** → **Configuration** → **Save**.
- ▶ Back up all data at the host level on logical drives of the array.
- ▶ Ensure that all I/O has been stopped to the array, and unmap any logical drives that are mapped on the array, ensuring that each host no longer requires the disk presented to it by the array to be migrated.
- ▶ Locate the array disks on the DS5000, noting the enclosure ID and location slot of each disk member.

- ▶ Save the security key file, as discussed in 5.4.2, “Save security key file” on page 272.
- ▶ Obtain blank canisters or new drives to be inserted in the DS5000 when the drives are removed.

On the target DS5000:

- ▶ Verify that there are available drive slots to host the disk drives.
- ▶ Check that the drives that are being moved are supported. The Drive Security premium feature should be already enabled, as shown in Figure 5-4 on page 265.
- ▶ Verify that the firmware is the same on the target DS5000 as the source DS5000, and is up to date.
- ▶ Unlock drives before importing the array.

5.5.2 Export the array

From the Storage Manager window, in the Logical tab, select the array that you want to export and then select **Advanced** → **Maintenance** → **Export Array**.

The Export Array wizard will guide you through the export.

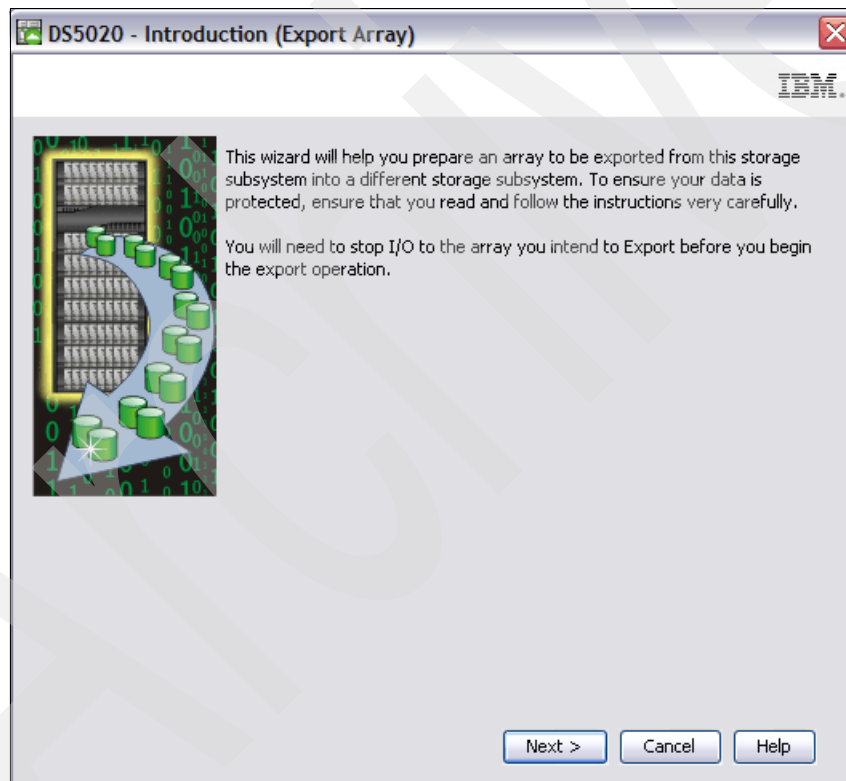


Figure 5-16 Export Array wizard

The window with the preparation checklist opens, as shown in Figure 5-17 on page 277. Check that no planning steps have been missed.

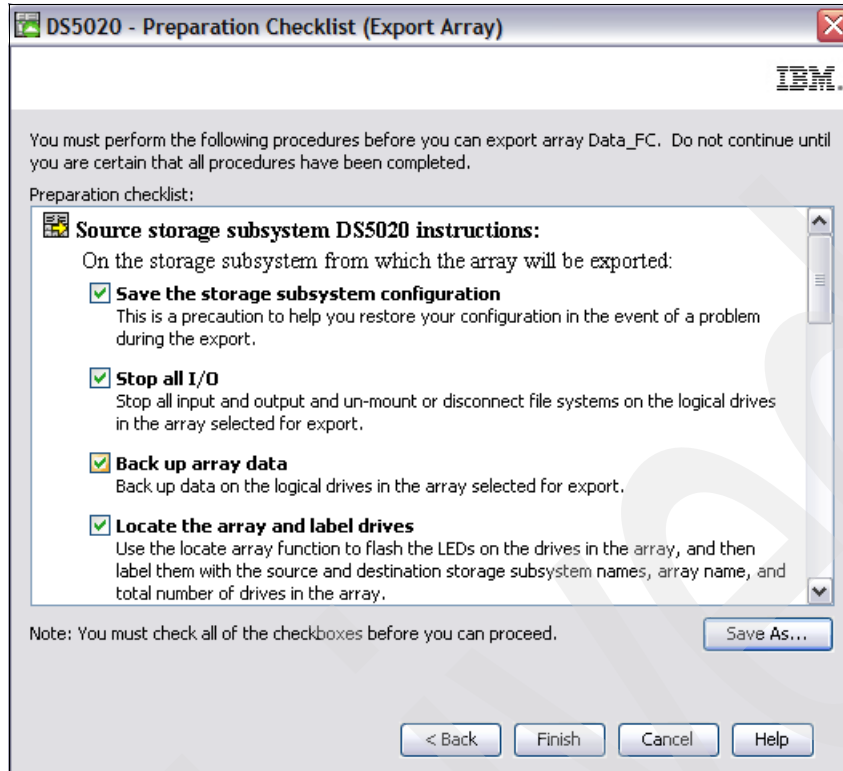


Figure 5-17 Preparation checklist

When the process has completed, the wizard will indicate that the drives can be removed, as shown in Figure 5-18.

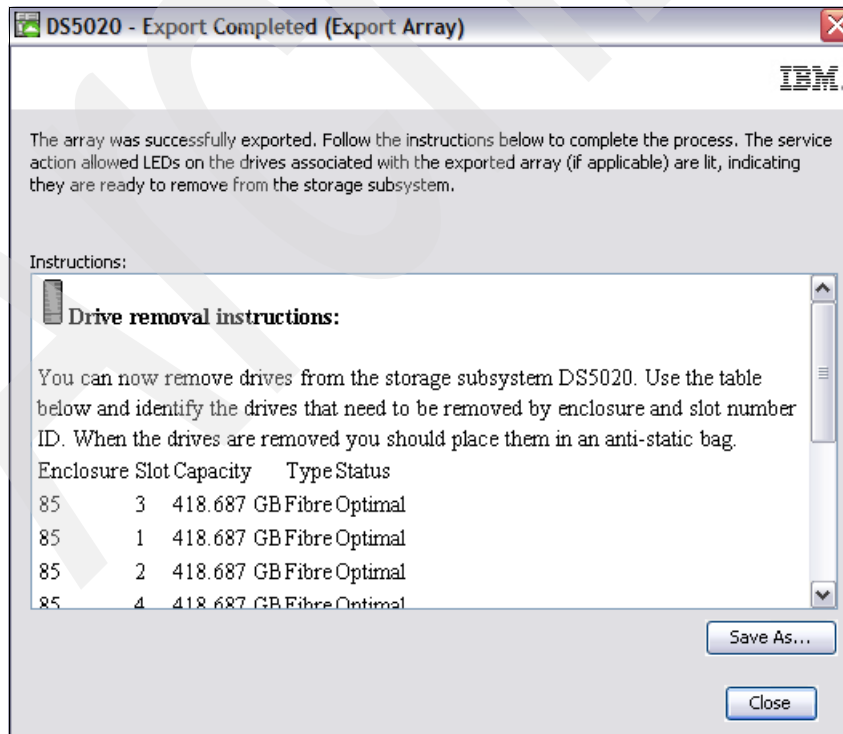


Figure 5-18 Export completed

The Storage Manager will now indicate that the array has been exported and the physical disks are offline. The windows shown in Figure 5-19 and Figure 5-20 confirm this information.

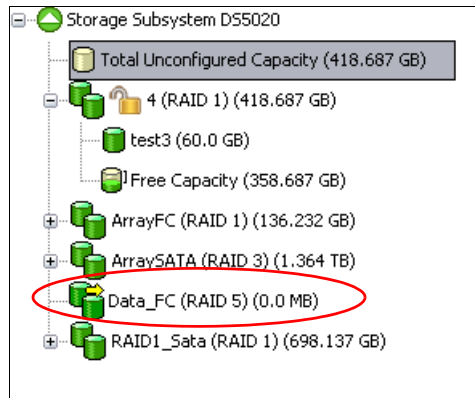


Figure 5-19 Array now exported

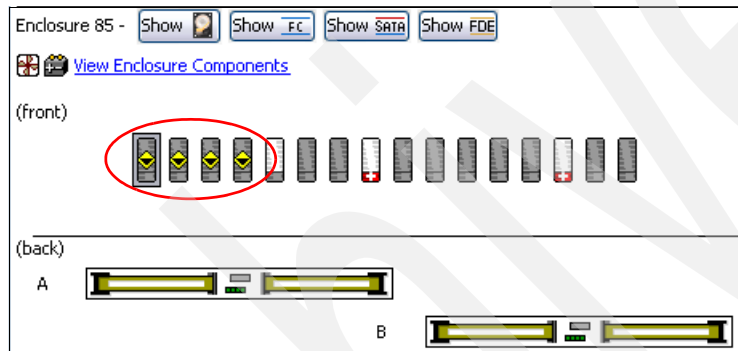


Figure 5-20 Array disks now offline

5.6 Import secure drive array

Once all the disks have been physically moved and are seated correctly in the target DS5000 enclosure, the drives can be unlocked and then the array can be imported. It is necessary to ensure that all the drives are seen by the DS5000 and that there is not a problem with them, as any problems will affect the array after it has been imported.

All drives will be in a locked state, as shown by the event viewer and the individual drive properties, as shown in Figure 5-21 on page 279. Access is not allowed due to invalid key details.

Port	Channel	ID
0	2	1/0xE8
1	1	1/0xE8

Media type:	Hard Disk Drive
Interface type:	Fibre Channel
Drive path redundancy:	OK
Security Capable:	Yes
Secure:	Yes
Read/write accessible:	No, invalid security key
Security key identifier:	2700000060080E500017B53E00004B974AA917CA
Speed:	15,015 RPM
Current data rate:	2 Gbps
Product ID:	ST3450056FC F

Figure 5-21 Drive properties show invalid key details and secure enabled

The DS5000 storage subsystem will also indicate that it needs attention due to the locked state of the drives. You can see the details about this situation in the Recovery Guru, as shown in Figure 5-22.

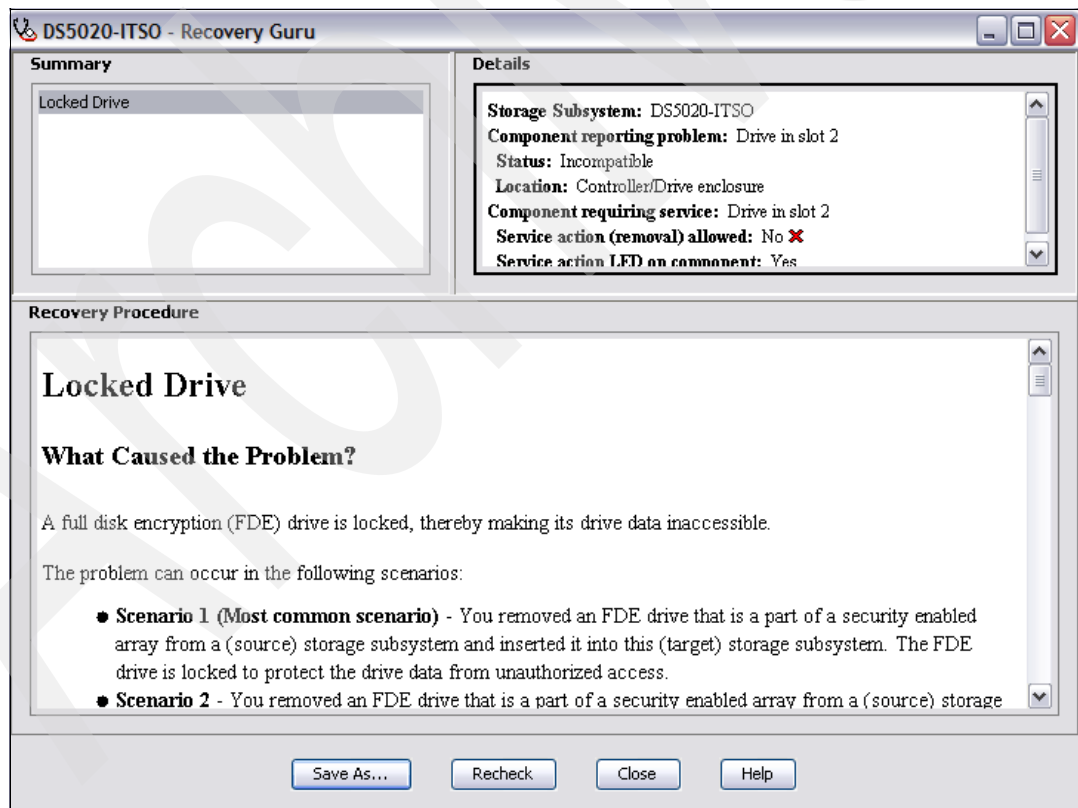


Figure 5-22 Drive locked message in the Recovery Guru

5.6.1 Unlock drives

Prior to importing the array, all drives must be unlocked. This procedure is performed on one disk, but will then be applied to all disks.

From the Physical tab in the Storage Manager window, select one of the FDE drives that has just been installed (it will be marked as offline). Right-click the drive and select **Unlock**, as shown in Figure 5-23.

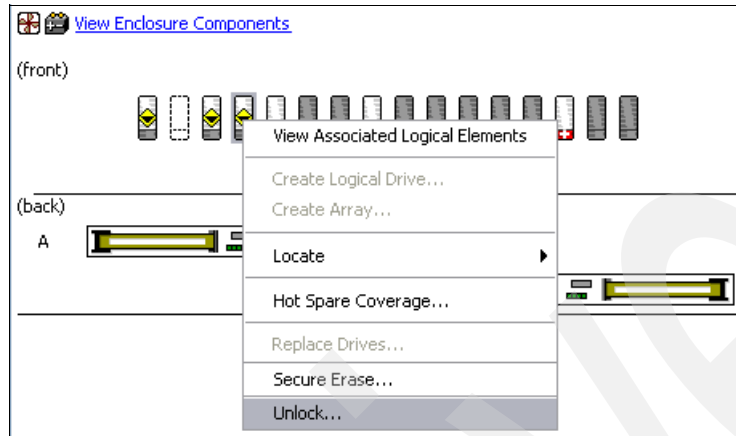


Figure 5-23 Select the Unlock option on the FDE drive

The DS5000 storage subsystem will recognize the FDE drives that are locked and will prompt for the key file and pass phrase, as shown in Figure 5-24.

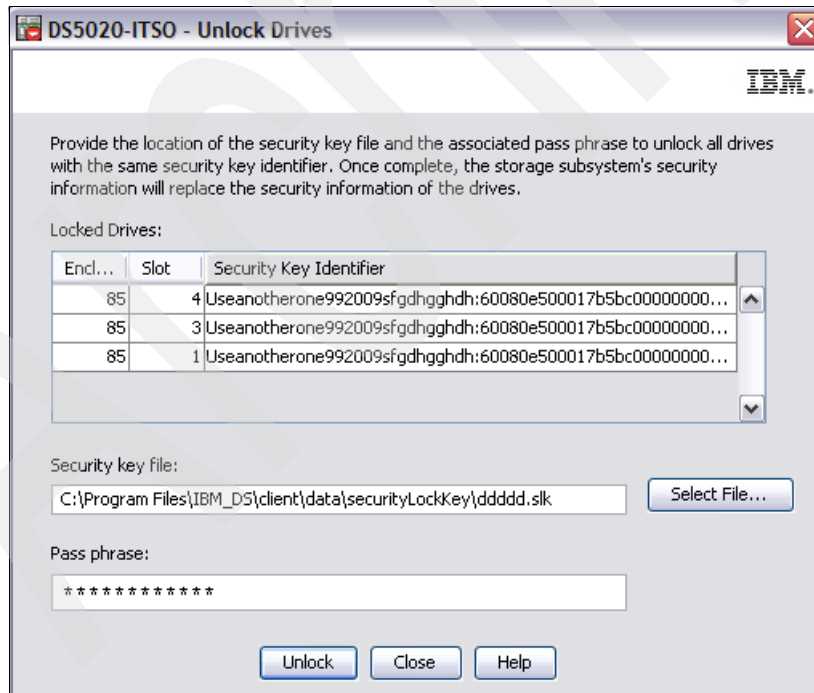


Figure 5-24 FDE drives locked with the key file and pass phrase to unlock

When the correct key file is selected and the pass phrase is entered, all the drives will be successfully unlocked, as shown in Figure 5-25.



Figure 5-25 FDE drive unlock completed successfully

5.6.2 Import array

The array will be displayed in the Logical tab of the Storage Manager window, and is ready to import, as shown in Figure 5-26.

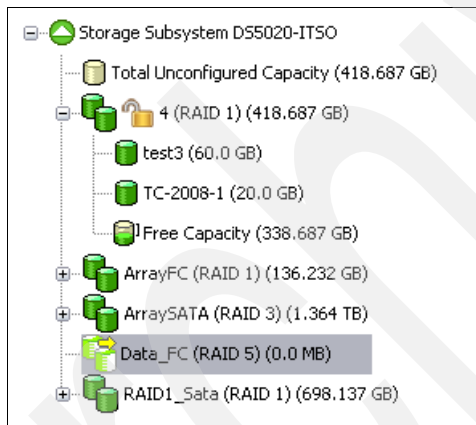


Figure 5-26 Array ready to be imported

Use the Import Array option to import an array that you have previously exported. From the Storage Manager window in the Logical view, select the array and then select **Advanced** → **Maintenance** → **Import Array**.

The import array, shown in Figure 5-27, gives details about the disks that will be imported. You need to check that all the disks that are displayed are correct and that none of them are missing.

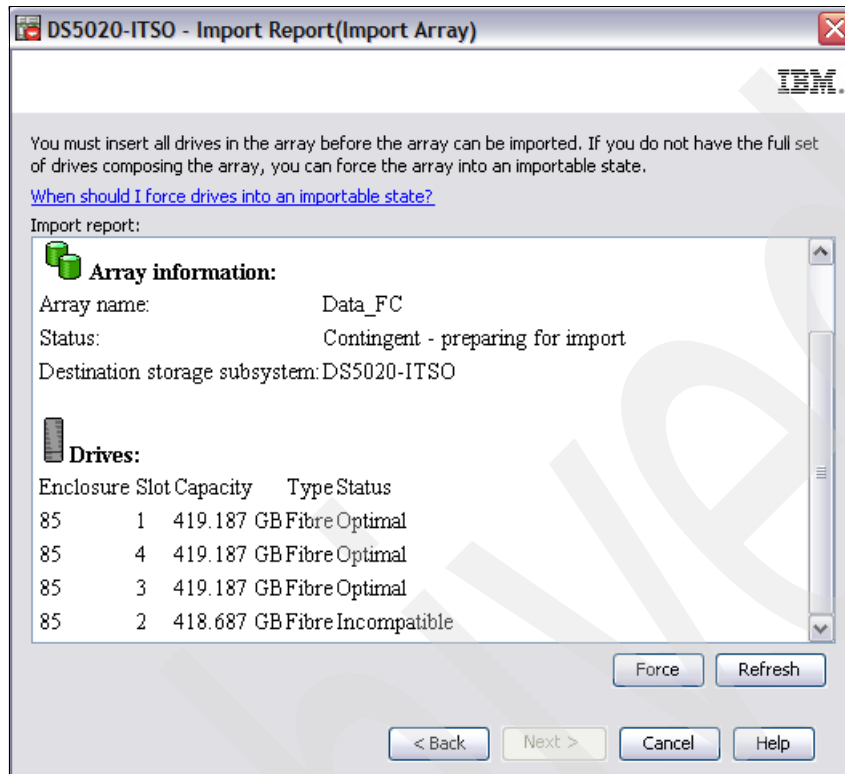


Figure 5-27 Import report indicating all that drives to be imported

Figure 5-28 confirms the importation of the drives.

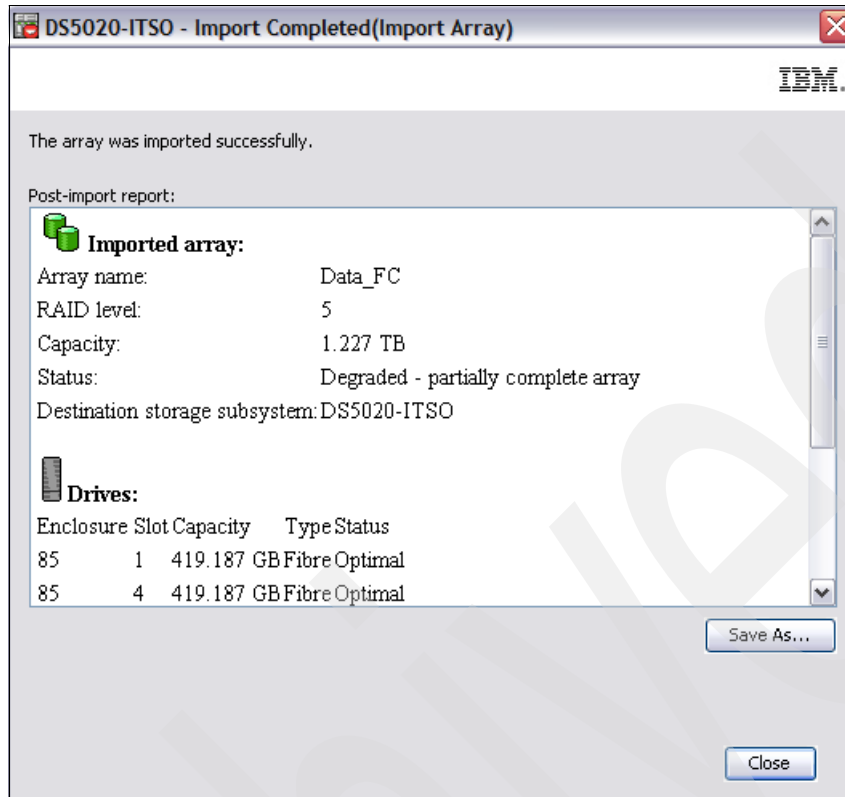


Figure 5-28 Import completed

The logical drives in the newly imported array are now ready to be mapped to hosts, where data can be accessed for read and write operations.

We recommend that you change the key after a secure array is imported to a DS5000 storage subsystem when it already has secure arrays. This will ensure that all secure enabled FDE drives will use a common key for authentication with the DS5000 storage subsystem when they are powered on.

Archived



IBM Remote Support Manager for Storage

In this chapter, we describe how to use IBM Remote Support Manager (RSM) for Storage with IBM Midrange System Storage storage subsystems.

6.1 IBM Remote Support Manager for Storage

The IBM Remote Support Manager for Storage (RSM for Storage) software is a no-charge software package that is installed on an IBM System x server running Novell SUSE Linux Enterprise Server 9, SUSE Linux Enterprise Server 10, Red Hat Enterprise Linux 4 Advanced Server, or Red Hat Enterprise Linux 5. It provides problem reporting and remote access for IBM Service for the IBM System Storage DS3000, DS4000, and DS5000 families.

The problem reporting utility provided by RSM for Storage automatically creates an entry in the IBM call management system for each subsystem that reports a problem. This is the equivalent of placing a voice call to IBM Service for a problem. Once in the system, problems are responded to with the priority specified by the maintenance agreement in place for the product.

RSM for Storage controls security for remote access by managing hardware and software components of the server on which it is installed. Once installed, the server should be considered a single purpose appliance for problem reporting and remote access support for your storage subsystems. Only applications approved by IBM and specified in this document should be installed. (Management of the internal firewall and other configuration changes made by the software might prevent other applications from working.) There is no guarantee that applications that work with the current version of RSM for Storage will continue to work with future releases.

Remote access to the RSM for Storage system by IBM Service is provided by either an external modem attached to the server or through an external SSH connection. This connection provides IBM Service with a command-line interface to the server. All bulk data transfers for logs and other problem determination files are sent to IBM through e-mail using the server's Ethernet interface, as shown in Figure 6-1. An internal firewall that is managed by the RSM for Storage software keeps remote and local users of the system from accessing other devices on your intranet. Local and remote IBM users of the system do not have the ability to change any security features of the software.

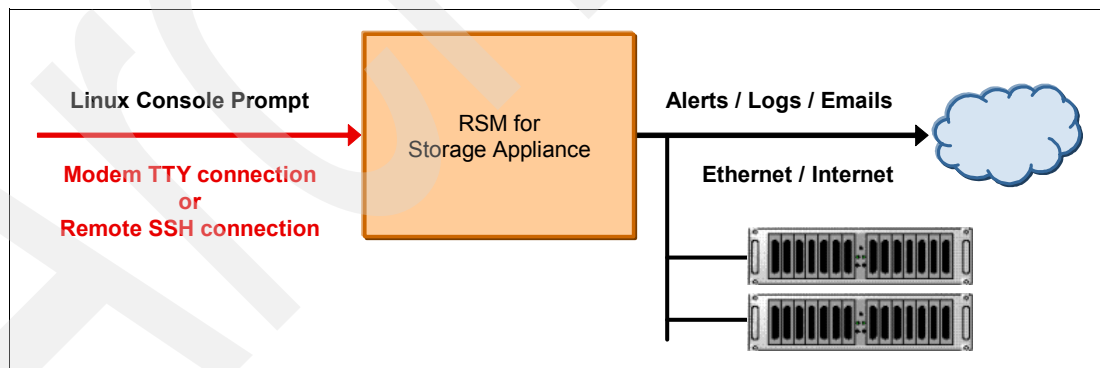


Figure 6-1 RSM layout

Your existing IBM DS Storage Manager software monitor the storage subsystems, and this software is configured to send SNMP traps to the Remote Support Manager when critical events are detected. Configuration of the management application is discussed in 6.2.4, “Configuring SNMP traps in Storage Manager” on page 314.

The RSM for Storage user interface allows you to control and view the status of four management areas:

- ▶ System configuration
- ▶ Reporting
- ▶ Remote access
- ▶ Internal firewall

Your contact person for RSM for Storage will also receive e-mails about status changes in these areas.

One RSM for Storage server can support up to 50 DS5000, DS4000, DS3000, FASiT 200, and FASiT 500 storage servers. Only IP connectivity to the Ethernet management ports of the subsystems is required; serial cables are not needed for access to the disk subsystems. The storage subsystems must be under a warranty or a current IBM maintenance agreement to be eligible to use the RSM for Storage software. There are no annual fees for RSM.

Note: Documentation and installation code is available at the following address:

<http://www.ibm.com/storage/disk/rsm>

6.1.1 Hardware and software requirements

RSM for Storage has the following hardware and software requirements.

Hardware requirements

IBM internal databases that support RSM for Storage use the IBM Machine Type and Serial Numbers of the servers, and therefore an IBM System x server must be used. This is required in order to properly register heartbeat records from the RSM for Storage system in IBM tracking databases. RSM for Storage can be run on a dedicated System x server or in a VMware client running on a System x server. See the *IBM RSM for Storage Compatibility Guide* (found at <http://www.ibm.com/support/docview.wss?uid=psg1MIGR-66062&rs=594>) for the minimum server requirements and a list of the specific servers that have been tested with the RSM software.

Recommended servers

The following servers have been tested with the RSM for Storage software, and installation instructions are available for configuring and using them with RSM for Storage software:

- ▶ IBM System x3250 4364
- ▶ IBM System x306m 8849
- ▶ IBM System x3550 7978
- ▶ IBM DS-RSM Model RS1 (1818-RS1)
- ▶ IBM DS-RSM Model RS2 (1818-RS2)

The *IBM Remote Support Manager for Storage: Installation Hints and Tips* document (found at <http://www.ibm.com/support/docview.wss?uid=psg1MIGR-66062&rs=594>) contains specific information about the BIOS configuration and device drivers that might be required to install the Linux OS on the above servers.

Other IBM servers

You may use other System x servers that meet the requirements listed below:

- ▶ 512 MB memory.
- ▶ 20 GB disk space.
- ▶ Serial port: The serial port must be on the server system board. Some servers have a build-in Remote Supervisor Adapter (RSA) that also includes a serial port. The serial port on the RSA cannot be accessed by the RSM for Storage software.
- ▶ 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.

Note: To use servers other than those specifically tested with RSM for Storage, you will need to see the System x server support Web site for technical information about BIOS configuration and device drivers that might be required to install the Linux OS or configure the server's internal RAID capability.

The *IBM RSM for Storage Compatibility Guide* also contains the setup required for a VMware client that will host the Linux operating system running RSM for Storage.

RSM for Storage in a VMware virtual client

RSM for Storage has been tested for operation in a VMware virtual client. See the *IBM RSM for Storage Compatibility Guide* for specific configurations.

In setting up the virtual client, allocate the following resources:

- ▶ Storage: 20 GB.
- ▶ Memory: 512 MB.
- ▶ Assign exclusive physical ownership of the host system's first serial port to the virtual client as `/dev/ttyS0`.

The client must be configured to automatically start when the host reboots.

RSM for Storage has been tested on both SLES 10 and RHEL 5 running on an System x server with:

- ▶ VMware Server 1.05
- ▶ VMware ESX Server 3

Software requirements

The RSM for Storage software requires the following prerequisite software to monitor an IBM Midrange System Storage storage subsystem:

- ▶ IBM Storage Manager V9.16 or later (the latest version is recommended) with Event Monitor installed on a management station in a different server.
- ▶ Storage subsystems with controller firmware supported by Storage Manager V9.16 or later. The latest supported firmware version is recommended.
- ▶ One of the following operating systems to install the RSM for Storage software:
 - Novell SUSE Linux Enterprise Server 9 (SP3 or SP4)
 - Novell SUSE Linux Enterprise Server 10 (Base, SP1 or SP2)
 - Red Hat Enterprise Linux 4 AS (Update 4, 5, or 5)
 - Red Hat Enterprise Linux 5 (Base, Update 1, or 2)

RSM for Storage software receives SNMP traps from the Event Monitor included with IBM DS Storage Manager. RSM for Storage software cannot be installed on the same system used to manage your storage network.

Note: See the *IBM RSM 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>

6.1.2 DS-RSM Model RS2

The IBM System Storage DS Remote Support Manager for Storage (DS-RSM) Model RS2 server is available to simplify the task of obtaining the hardware and software components required for running the RSM for Storage software. The DS-RSM Model RS2 (Figure 6-2) can be ordered as an appliance with your IBM DS5000 System Storage storage subsystem.

The DS-RSM Model RS2 is designed to help manage problem reporting and analyze alert notifications for IBM Midrange System Storage storage subsystems. The DS-RSM Model RS2 contains more advanced processor and memory technology for enhanced processing capabilities. The DS-RSM Model RS2 can attach to the LAN and collects alerts from IBM Midrange System Storage storage subsystems within the data center and expedite forwarding of the alerts to an IBM Service center.



Figure 6-2 DS-RSM Model RS2

Additionally, the DS-RSM Model RS2 is designed to enable:

- ▶ IBM to establish communication with the attached supported storage systems
- ▶ Faster response time on processing alerts
- ▶ The ability to transfer detailed information to assist error analysis
- ▶ Enhanced problem resolution

The DS-RSM Model RS2 can help improve and facilitate service and support for the DS storage subsystems. Additionally, IBM recommends that, for each data center with a DS storage subsystem present, that you install at least one DS-RSM Model RS2 that can interoperate with all of the DS5000/DS4000/DS3000 controllers located at the site. Each DS-RSM Model RS2 unit has the capability to manage up to 50 DS5000/DS4000/DS3000 storage systems.

DS-RSM Model RS2 is compatible with the following IBM Midrange System Storage storage subsystems:

- ▶ DS3000 series
- ▶ DS4000 series
- ▶ DS5000 series

The DS-RSM model RS2 is based on the System x3550 M2 Type 7946 server. The server is preloaded with Novell SLES 10 and RSM for Storage software. An internal modem is available as an option if you choose to use a modem to provide remote access for IBM Service. The internal modem might not be compatible in all countries or regions. In that case, you can provide an external modem appropriate for your phone system. As an alternative to using a modem, you can allow remote access using an external SSH connection. A keyboard, mouse, and monitor are required by IBM Service for installation and troubleshooting of the DS-RSM Model RS2 server. An optional, rack-mounted KVM drawer can be ordered with the DSM-RSM Model RS2 server in order to meet this requirement.

6.1.3 Installation choices for RSM for Storage

Various options exist for how RSM for Storage can be set up, as shown in Figure 6-3.

See the *IBM RSM for Storage Compatibility Guide* and the *IBM Remote Storage Manager for Storage Planning, Installation and Users Guide* found at the following address for the latest information related to these options:

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

Installation Choices				
Hardware Choices	Build your own	Use existing hardware	Use your own server and obtain a supported version of Linux. The system acts as a gateway for IBM Support with a limited scope of access to your network.	RSM documentation has step-by-step procedures for installing the Linux OS and the RSM for Storage software. Decide if you will allow IBM support to connect via a modem or SSH.
		Order a new System x server	Order a new System x server for use with RSM for Storage and get the Linux OS as a drop in the box option.	
		Run in a VMware guest OS	Install a supported OS distribution in a VMware system and reduce your hardware costs.	Consider hosting your SAN Management Station and the RSM for Storage on the same VMware platform.
	DS-RSM	Purchase a RSM ready appliance	DS-RSM is reloaded with SLES 10 and RSM software and an optional internal modem.	DS-RSM is installed and supported by IBM. You will need to complete the installation worksheet and have networking infrastructure in place in order for IBM to complete the configuration.
Remote Access	Problem reports are sent to IBM via email. You must choose a method of remote access.	SSH	Requires configuration of your external firewall.	The use of an authentication server adds another layer of access control.
		Modem	Requires an analog phone line and modem.	You can use most external modems or, if using the DS-RSM, order it with an internal modem adapter.
OS	RSM for Storage is a collection of software elements that integrates with a Linux OS.	SLES	9.3 & 9.4. 10.0, 10.1 & 10.2.	The DS-RSM is preloaded with Novell SLES 10.2.
		RHEL	4.4, 4.5 & 4.6. 5.0, 5.1 & 5.2.	RHEL is only available on "build your own" systems.
KVM	RSM requires an attached monitor, mouse and keyboard for some tasks		As the system can be anywhere in your network, you can connect to a KVM switch with other servers.	The DS-RSM can be ordered with a rack mounted KVM kit.

Figure 6-3 RSM installation choices

6.1.4 How RSM for Storage works

RSM for Storage uses an Ethernet connection for problem reporting and a modem (optional) for remote access by IBM Service, as shown in Figure 6-4. The storage system can be any supported IBM Midrange System Storage storage subsystem.

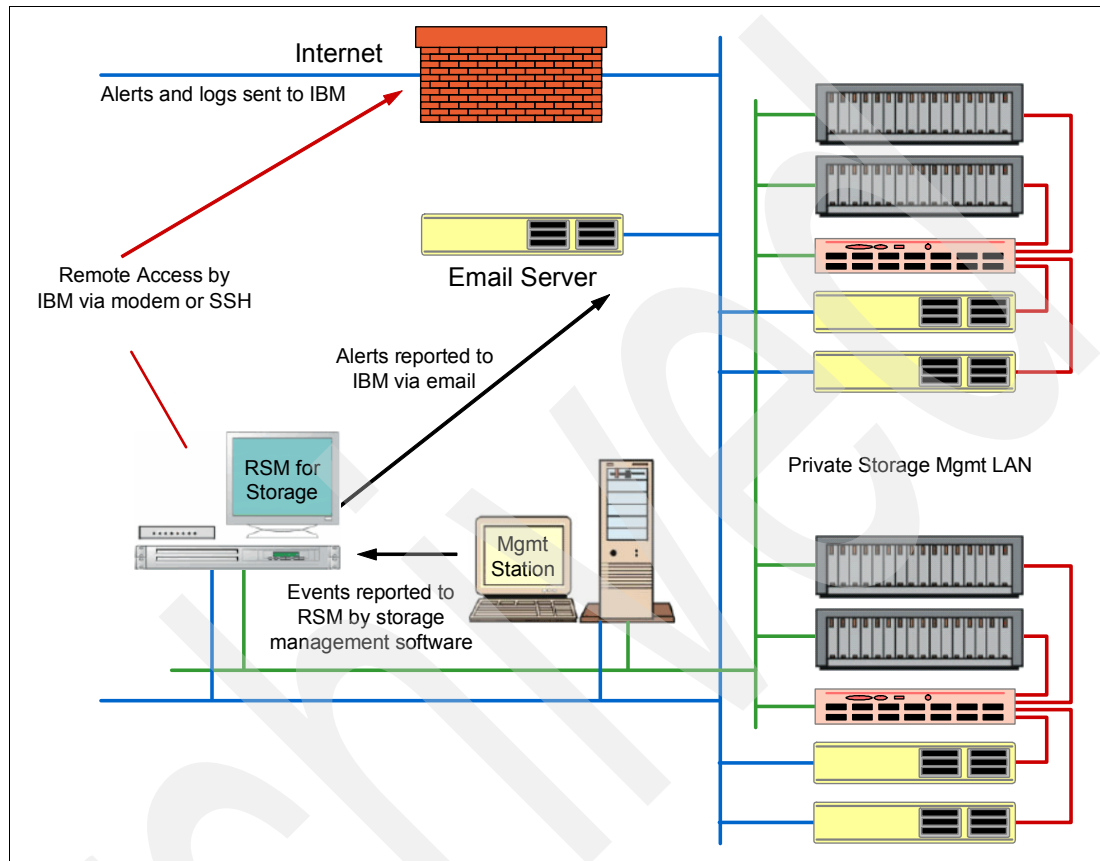


Figure 6-4 RSM for Storage connection

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 must be running IBM Storage Manager's Event Monitor. All the storage subsystems, the management station, the e-mail server, and Internet gateway must be 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 6-4).

Figure 6-4 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 in this instance needs at least two network adapters.

After RSM is installed, configured, and activated, here are the steps in a sample scenario for RSM. See Figure 6-4 on page 291 to understand the flow:

1. For example, an error occurs in one of the storage subsystems and a critical event is logged in the management station (running DS Storage Manager).
2. The management station reports the critical event to the RSM server through an SNMP trap. 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 the 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 by 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 might not always be running, we recommend that the Event Monitor be installed. See Storage Manager documentation to learn about the installation of Event Monitor.

See 6.1.5, "Notification e-mail and events filtering" on page 292 to configure the SNMP trap in Storage Manager.

3. IBM Service does problem determination based on information sent by the alert along with the problem determination data, such as MEL, RLS, and the subsystem profile. If the problem can be fixed with existing information, IBM Service contacts the customer either by phone or e-mail to resolve the problem. After the problem is solved, either IBM Service or the customer must indicate *Service Complete* for all alerts in the RSM. IBM Service can dial to the RSM modem or use an SSH connection to acknowledge and indicate *Service Complete* for all alerts for the subsystem.
4. If the problem cannot be fixed with existing information, IBM Service dials the RSM modem or uses an SSH connection, acknowledges the alert, and performs troubleshooting by connecting to the storage subsystem using the Storage Manager command-line interface (SMcli) or RLOGIN. IBM Service might need to contact the customer to obtain the password for the storage subsystem to use SMcli. IBM might also have to contact the customer to enable RLOGIN access. Indeed, we recommend normally disabling RLOGIN.

If IBM Service 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 or using SSH connection. Any data connected is 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).

5. After the problem is resolved, all alerts must be closed either by IBM Service 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.

6.1.5 Notification e-mail and events filtering

The RSM for Storage software sends e-mails to notify you about important changes in status.

Up to 20 contact people can be configured. One will be the primary contact for RSM related events and will receive all notifications about RSM system operation and problems. The primary contact will receive all notifications about RSM system operation and problems.

Additional contacts can be configured for subsystems located at other sites. They only receive notifications related to subsystem problems. See Figure 6-5 for information about the notifications.

Notification	Primary contact	Subsystem contacts
Remote Access is enabled or disabled	X	
Remote User logs into the RSM for Storage system or logs out	X	
Firewall is enabled or disabled	X	
Remote Access is about to expire or has expired	X	X
An alert occurs for a subsystem	X	X
Daily status updates at local noon.	X	
Daily status updates when a subsystem has an active alert	X	X
Daily phone line check fails	X	

Figure 6-5 E-mail notifications

There are several types of notifications sent by RSM to the primary contact or subsystem contact, as configured in the RSM:

- ▶ *Remote Access notifications* are sent when:
 - Remote access is enabled or disabled, as shown in Example 6-1.

Example 6-1 Remote access notification

Remote access to the RSM for Storage system has been enabled.

The remote access timeout for localhost is now 35:59 (HH:MM)

There are no remote users connected to localhost at this time.

```
#
# RSM system type:      7985PBF
# RSM system serial:   KQDBXV8
# RSM system name:     localhost.localdomain
# RSM modem Key:
K7EuTWmKlqoDuIsgP0xNYiVI13tDt1qoDuZbIvKXYI9pFXy5Le9qDMbeWKFxkn81093NAZxi96T3ZYr
b21zFSRUkaB00wEGcJ6NfmYvas12w
# RSM build:           T1.15.4
# RSM location:        Datacenter 1
# RSM address:         River Street
# RSM city:            River City
# RSM state:           NC
# RSM zip:             00000
# RSM country:         United States
# RSM contact name:    Max Musterman
# RSM contact phone:   11234567890
# RSM contact hours:   7am - 6pm
# RSM contact timezone: GMT - 6
# RSM contact alt. phone: 1 123 456 7891
# RSM contact alt. hours: 6pm - 7am
# RSM contact e-mail:  max.musterman@rivers.local
# RSM contact country: United States
```

- A remote user connects or disconnects from the system.

- The remote access automatic timeout is about to expire and the system has one or more active alerts. Timeout warnings are sent at 4 hours, 2 hours, and 1 hour before the timeout occurs.
- ▶ *Alerts Status notifications* are sent when an alert has been sent to IBM Service, as shown in Example 6-2.

Example 6-2 Alert Status notification

Notification from RSM for Storage system: localhost

An alert was sent to IBM Service at Thu Aug 28 22:26:58 2009 UTC.

Remote Access to the RSM for Storage system is enabled for the next 12 hours

Check the RSM for Storage console for additional information related to this problem.

```
#
# Machine type:          1818-53A
# Machine serial:       1234567
# Machine name:         ITS05300
# Machine location:     Dataenter 1
# Machine address:      River Street
# Machine city:         River City
# Machine state:        NC
# Machine zip:          00000
# Machine country:      United States
# Machine country code: 000
# Machine info:         None
# Machine info name:    None
# Machine software id:
# Machine contract:
#
# Contact name:         Max Musterman
# Contact phone:        11234567890
# Contact hours:        7am - 6pm
# Contact timeZone:     GMT - 6
# Contact alt. phone:   1 123 456 7891
# Contact alt. hours:   6pm - 7am
# Contact e-mail:       max.musterman@rivers.local
# Contact country:      United States
```

The event code was: 5005

The event message was:

```
Place controller offline.
Controller.
Controller in slot B
```

- *Daily Status e-mails* serve as a heartbeat notification that the RSM for Storage system is operational. This includes the summary status for the system and status for alerts that might be active for storage subsystems, as shown in Example 6-3.

Example 6-3 Daily Status e-mail

Status report for IBM Remote Support Manager running on: localhost.localdomain

Application has been running for: 0 days, 21 hours, 41 minutes

System status: OK

One or more subsystems has reported a problem.

The internal firewall is enabled.

Remote access is disabled.

1 storage subsystem has an active alert.

0 alerts have been acknowledged by IBM Service.

1 alert has been sent to IBM, but has not yet been acknowledged.

0 additional alerts are pending.

Subsystem: ITS05300, located at Datacenter 1

0 alerts acknowledged.

1 alert sent to IBM.

0 additional alerts pending.

Because IBM Service has already been alerted to a problem with this subsystem, no additional alerts will be sent to IBM for this subsystem until existing alerts are closed.

This problem has been open for 19 hours.

```
# Mini Inventory : Family      IBM MTM      IBM Serial Firmware  Hostname
# Mini Inventory : DS5000    1818-53A    1234567    07.60.13.00 ITS05300
#
# RSM system type:      7985PBF
# RSM system serial:    KQDBXV8
# RSM system name:      localhost.localdomain
# RSM modem Key:
K7EuTWmKlqoDuIsgP0xNYiVI13tDt1qoDuZbIvKXYI9pFXy5Le9qDMbeWKFxkn81093NAZxi96T3ZYr
b21zFSRUkaB00wEGcJ6NfmYvasl2w
# RSM build:            T1.15.4
# RSM location:         Datacenter 1
# RSM address:          River Street
# RSM city:             River City
# RSM state:            NC
# RSM zip:              00000
# RSM country:          United States
# RSM contact name:     Max Musterman
# RSM contact phone:    11234567890
# RSM contact hours:    7am - 6pm
# RSM contact timezone: GMT - 6
# RSM contact alt. phone: 1 123 456 7891
# RSM contact alt. hours: 6pm - 7am
# RSM contact e-mail:   max.musterman@rivers.local
# RSM contact country:  United States
```

- *Firewall Status notifications* are sent when the internal firewall is enabled or disabled by a root or admin user, as shown in Example 6-4.

Example 6-4 Firewall notification

RSM for Storage (localhost) internal firewall has been disabled.

```
#
# RSM system type:          7985PBF
# RSM system serial:       KQDBXV8
# RSM system name:         localhost.localdomain
# RSM modem Key:
K7EuTWmKlqoDuIsgP0xNYiVI13tDt1qoDuZbIvKXYI9pFXY5Le9qDMbeWKFxkn81093NAZxi96T3ZYr
b21zFSRUkaB00wEGcJ6NfmYvas12w
# RSM build:                T1.15.4
# RSM location:             Datacenter 1
# RSM address:              River Street
# RSM city:                 River City
# RSM state:                NC
# RSM zip:                  00000
# RSM country:              United States
# RSM contact name:         Max Musterman
# RSM contact phone:        11234567890
# RSM contact hours:        7am - 6pm
# RSM contact timezone:     GMT - 6
# RSM contact alt. phone:   1 123 456 7891
# RSM contact alt. hours:   6pm - 7am
# RSM contact e-mail:       max.musterman@rivers.local
# RSM contact country:      United States
```

- *Ping check notifications* are sent when two ping checks in a row fail. Ping checks of the configured SMTP sever and Management Station are performed each half hour.
- *Ignored Event notifications* are sent when an event is received that is configured to be ignored by the RSM for Storage system, and is therefore not reported to IBM Service. These are events for which a response by IBM Service is not usually required, as shown in Table 6-1.

Table 6-1 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	The logical drive is not on a preferred path due to ADT/RDAC.

- *Phone Line Check notification* is sent if the RSM for Storage system cannot verify that an active phone line is connected to the modem. The check is run daily after midnight local time.

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 *must be disabled* when RSM for Storage is installed, if 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.

To disable e-mail alerts in Storage Manager, do the following steps:

1. Right-click your management station in the Storage Manager Enterprise window and select **Configure Alerts** to select all storage subsystems.
2. On the Email 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 (see Figure 6-12 on page 304 for the Contact List), it is not necessary to delete them in Storage Manager.

6.1.6 Remote access methods

The required Remote Access for IBM Service can be provided by one or both of two methods. An external modem can be attached to the server's serial port, or remote access through an SSH client can be enabled.

Remote access by modem

The functional requirements for the modem are minimal; it is only used for remote access by IBM Service. Most "Hayes-compatible" external modems can be used. Any V.92 56K modem supporting the common AT command set will work.

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

- ▶ Multitech MultiModem II MT5600BA
- ▶ Multitech MultiModem ZBA MT5634ZBA

See the *IBM RSM for Storage Compatibility Guide* and *IBM Remote Support Manager for Storage: Installation Hints and Tips* for updated information about which modems are supported. You will need to contact your IT support staff for installation and problem resolutions related to the modem.

The optional internal modem that can be ordered with the DS- RSM Model RS2 server is supplied with an RJ-11 phone cord. You might need to supply an adapter in order to connect the RJ-11 phone cord to your phone system.

Remote access by SSH

Instead of using a modem for external access to the RSM for Storage system, you can allow remote access by an external SSH connection. To do this, you will need to map a port on your external firewall to the IP address and SSH port (22) on the RSM for Storage system. While the RSM for Storage system has several layers of login protection on the SSH port, you can also require authentication before the external firewall makes a connection to the RSM for Storage system.

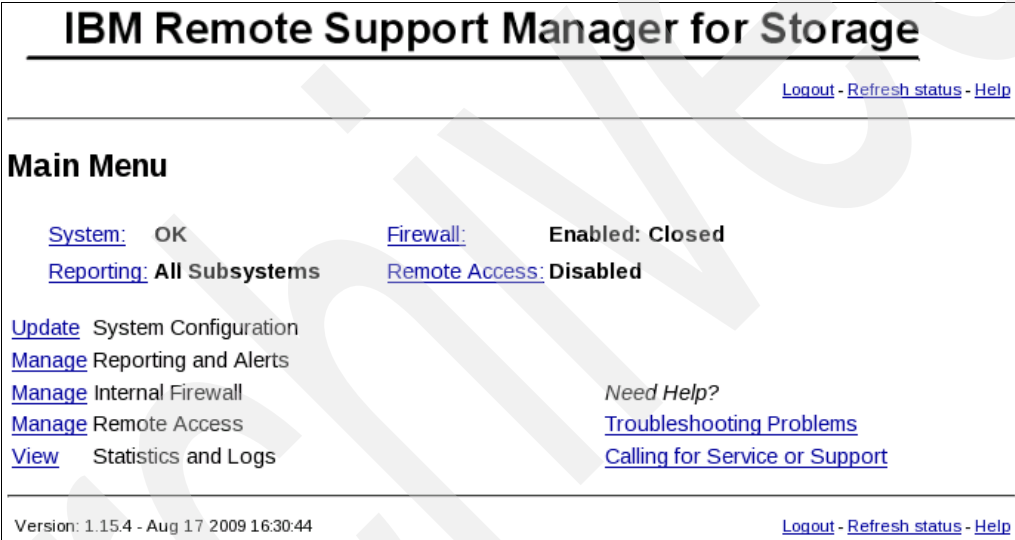
You can also choose to allow remote access by both methods. More information about setting up and using an SSH connection is available in the *IBM Remote Storage Manager for Storage Planing, Installation, and User's Guide*, GC26-7933, as well as the supplement *IBM Remote Support Manager for Storage: Installation Hints and Tips*, found at the following address:

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

6.1.7 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 Service, the interface is a command-line interface when connected to the RSM server through a modem or SSH connection.

Figure 6-6 shows an example of a System Configuration menu for an already configured and activated RSM system.



The screenshot displays the 'IBM Remote Support Manager for Storage' interface. At the top, the title is underlined, followed by links for 'Logout - Refresh status - Help'. Below this is a 'Main Menu' section. It lists system status: 'System: OK', 'Firewall: Enabled: Closed', 'Reporting: All Subsystems', and 'Remote Access: Disabled'. A list of actions includes 'Update System Configuration', 'Manage Reporting and Alerts', 'Manage Internal Firewall', 'Manage Remote Access', and 'View Statistics and Logs'. To the right, there are links for 'Need Help?', 'Troubleshooting Problems', and 'Calling for Service or Support'. At the bottom, the version '1.15.4 - Aug 17 2009 16:30:44' and another set of 'Logout - Refresh status - Help' links are visible.

Figure 6-6 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 might be displayed to the left of each link. The status values for each of these are as follows:

- ▶ System:
 - OK: Remote Support Manager is operating properly.
 - Incomplete: One or more required configuration settings are missing or incorrect.
 - Problem: There is a problem that is preventing correct operation.

Note: Reporting is disabled until all configuration problems are fixed.

- ▶ Reporting:
 - All Subsystems: Reporting is enabled for all configured subsystems.
 - Standby: Reporting has been disabled for all subsystems.
 - Partial: Reporting has been disabled for some but not all subsystems.
 - Suspended: Reporting is not performed while a configuration problem exists.
 - Storage Problem: A problem has been reported by one or more subsystems.
- ▶ Firewall:
 - Enabled:Closed: The firewall is enabled. Only connections required for reporting are allowed.
 - Enabled:Open: The firewall is enabled. Connections are open to one or more devices being serviced.
 - Enabled:Custom; The firewall is enabled. Only connections required for reporting and permitted by the custom rules defined in `/etc/rsm/rsm-firewall.conf` are allowed.
 - Disabled: The firewall is disabled. There are no restrictions on access to the networks connected to the RSM.
- ▶ Remote Access:
 - Disabled: Modem answer and remote user login is disabled.
 - Enabled: Modem is enabled and remote user login is allowed.
 - Active: A remote user is logged into the system.

6.1.8 RSM security considerations

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

Remote access to your system has the following four layers of control:

- ▶ The modem is configured to only answer when Remote Access is enabled by the RSM for Storage software. Likewise, the SSH daemon is only allowed to respond to connection attempts when Remote Access is enabled.

You can manually enable and disable remote access, or you can choose to have remote access automatically enabled when a storage subsystem reports a problem. When remote access is enabled, a timer is started that will automatically disable remote access when it expires. You do not have to remember to make the system secure after service has been completed.

The person identified as the primary contact for the RSM for Storage system is notified by e-mail whenever a change in the remote access settings occurs and all state changes are also written to the security log.

- ▶ The user ID reserved for remote access (*rservice*) is only valid when Remote Access is enabled. Attempts to log in using the *root*, *admin*, or *lservice* user IDs are rejected.

Note: For this reason, do not create additional users on this system.

- ▶ The initial login password is changed daily at midnight UTC. IBM Service has an internal tool that provides the current password for RSM for Storage systems.
- ▶ 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.

User ID

During installation, the 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 restrictions regarding the directories 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 ID *admin* and *lservice* for the RSM for Storage browser user interface can be changed by the Linux *root* user using the command `rsm-passwd admin` or `rsm-passwd lservice`. We recommend setting a different password for each user ID.

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

The Switch User (**su**) command is disabled to prevent a normal user from attempting to become “root” and have unrestricted access to the system. The RSM for Storage software makes other changes in program and directory permissions to limit what programs and files these users can access.

Internal firewall

RSM for Storage includes an internal firewall to limit the scope of access a remote user has to your network. Without an internal firewall, the remote user will have unrestricted access to your network. The RSM software configures an internal firewall on the RSM system to limit the scope of access that users of the RSM system have to your network, as shown in Figure 6-7 on page 301. When no alerts are active, the firewall only allows incoming SNMP traps and outbound SMTP email. When an alert occurs, a rule is automatically added to the firewall to allow access to the configured controllers for the storage subsystem reporting the problem. There may be times when you want to allow IBM Service to be able to access a device to troubleshoot a problem (such as a performance issue) for a subsystem that is not reporting a failure. You can manually enable “service access” for any configured storage subsystem. Service Access settings have a configurable timeout from 12 to 96 hours, after which the firewall rules for access are removed.

Firewall rules that are added for a device that is reporting an alert are removed when the alert is closed on the RSM for Storage system.

While the internal firewall effectively limits the scope of access for remote users, at the same time it also limits the access of any program running on the server. Because management applications such as IBM Storage Manager or IBM Director require access to all devices being managed, the presence of this internal firewall prevents the use of the RSM server as a management station. Therefore, the RSM for Storage system should be considered to be a single purpose appliance dedicated to problem reporting and remote access for IBM Service.

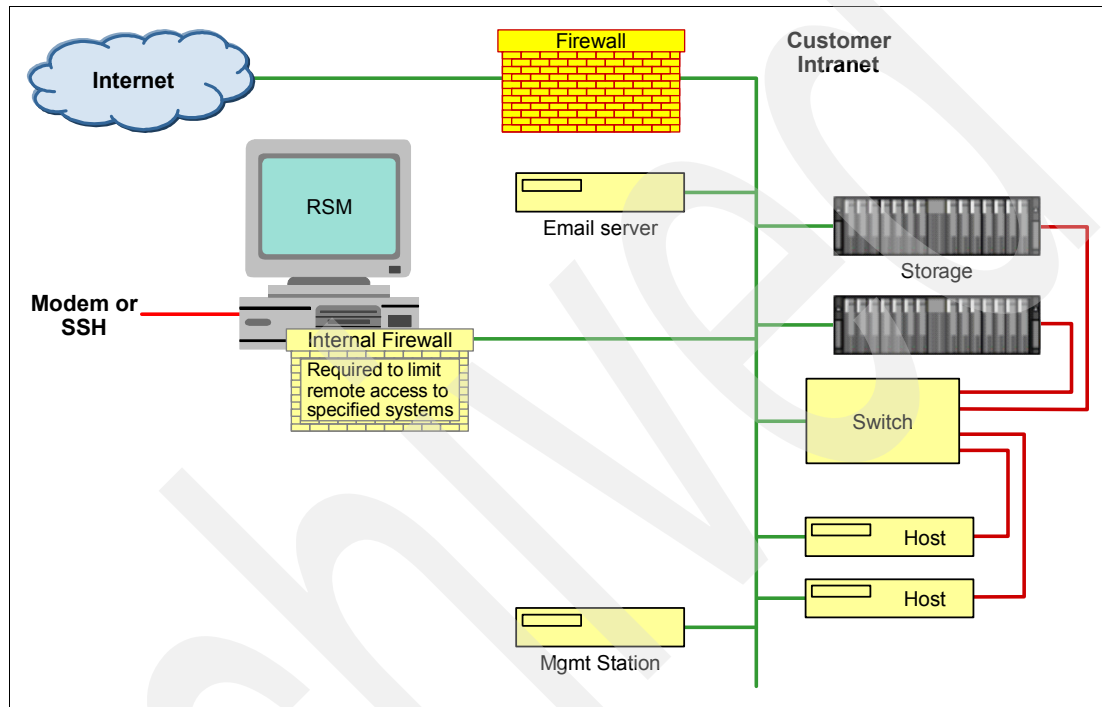


Figure 6-7 RSM internal firewall

6.2 Installing and setting up RSM

In this section, we show how to install and configure RSM. Before beginning this task, go to the RSM support Web page and carefully review the *IBM Remote Storage Manager for Storage* documents *Planning, Installation and Users Guide*, GC26-7933, and *Installation Hints and Tips*, found at the following address:

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

Tip: Do not use a remote connection when installing the RSM for Storage on the workstation. We recommend that you be logged on locally to the Graphical User Interface of the workstation, because RSM resets the firewall settings to prevent remote access to the Linux workstation and will be configured later.

6.2.1 Installing the host OS

There are various operating systems supported for the RSM host, as shown in “Software requirements” on page 288. In our example, we used a Red Hat Linux Enterprise License 5 as the host operating system on the RSM server. When installing RHEL5, we selected the following additional packages:

- ▶ expect
- ▶ mgetty
- ▶ Simple Web server (apache2)
- ▶ KDE desktop environment

See *IBM Remote Storage Manager for Storage: Planing, Installation, and User's Guide*, GC26-7933 for specific operating system installation instructions.

6.2.2 Installing RSM

The RSM software can be downloaded from:

<http://www.ibm.com/storage/disk/rsm>

We installed RSM according to the instructions in *IBM Remote Storage Manager for Storage Planing, Installation, and User's Guide*, GC26-7933. After the installation, you have to define the admin and lservice user IDs.

6.2.3 Setting up RSM

After the installation is complete, we have to set up RSM by performing the following steps:

1. On the Linux login window, perform the following steps:
 - a. Click **Session** and select **KDE**.
 - b. Log in as the admin user.
 - c. Click the Manage icon to open a Web browser that shows the main administration window, as shown in Figure 6-8. Click **Login**.

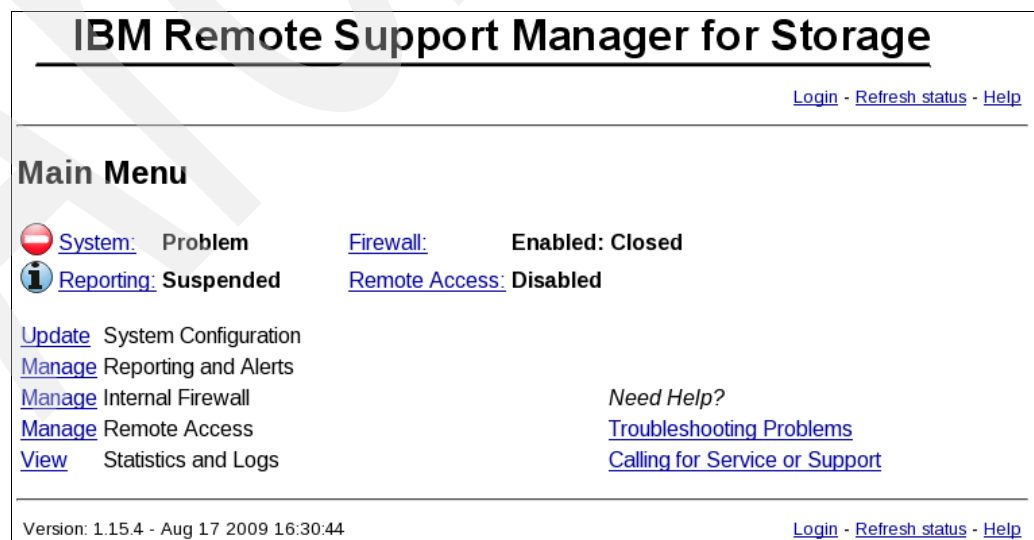


Figure 6-8 RSM administration window

2. Click **Login** and enter the user name and password of the RSM administrator. This account is named admin and the password was defined during the installation of RSM. See Figure 6-9.

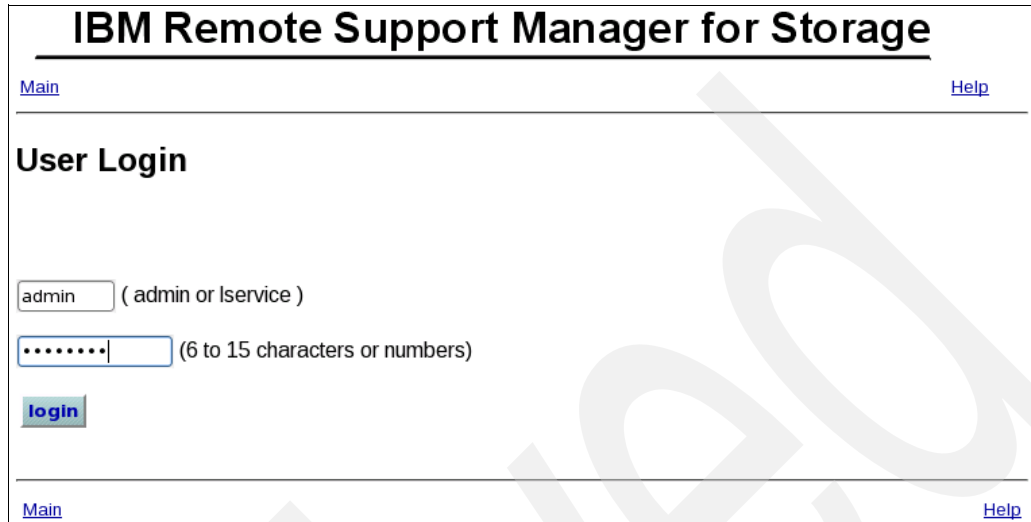


Figure 6-9 RSM logon window

3. You return to the Main Menu. The upper right menu contains a logout link. To start the setup, click **System**, as shown in Figure 6-10.

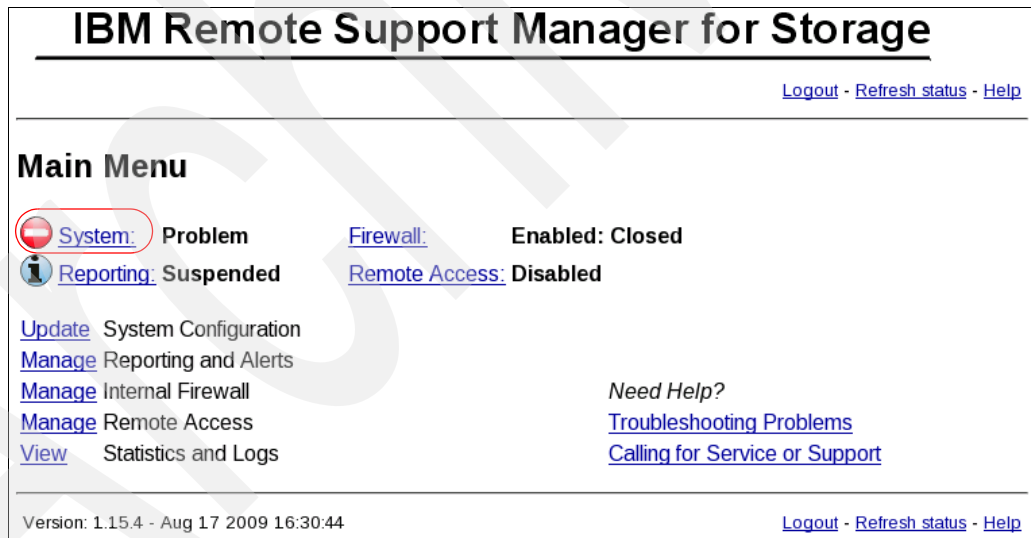


Figure 6-10 RSM main menu

- The System Configuration window shows incomplete tasks that need to be accomplished before the RSM system can be used (Figure 6-11).

IBM Remote Support Manager for Storage

[Main](#) [Logout](#) - [Refresh status](#) - [Help](#)

System Configuration

[System](#): **Problem** [Firewall](#): **Enabled: Closed**
[Reporting](#): **Suspended** [Remote Access](#): **Disabled**

View and define information for: **localhost.localdomain** Type: **7985-PBF** Serial: **KQDBXV8**

[Contact Information](#): **Incomplete**
[Company Information](#): **Incomplete**
[Connection Information](#): **Configuration Incomplete**
[Storage Subsystems](#): **Configuration Incomplete** (0 subsystems currently defined)
[Other SAN Devices](#): **OK** (0 switches currently defined)
[System Activated](#): **No**
[Options](#)

Configuration Test

All configuration incomplete errors must be fixed before the configuration test can be run.

Test has not been run since the last restart.

Figure 6-11 System Configuration incomplete

- Click **Contact Information** to access the contact list in order to add a new contact (Figure 6-12).

IBM Remote Support Manager for Storage

[Main](#) > [Configuration](#) [Logout](#) - [Refresh status](#) - [Help](#)

Contact List

[System](#): **Problem** [Firewall](#): **Enabled: Closed**
[Reporting](#): **Suspended** [Remote Access](#): **Disabled**

Who should IBM Support contact:

View/Configure
 1 [Select to add](#)

[Main](#) > [Configuration](#) [Logout](#) - [Refresh status](#) - [Help](#)



Figure 6-12 RSM contact list

- Click **Select to add**, fill out the form, and click **Update configuration**. This information is very important; when a DS Storage Subsystem reports a problem, IBM Service will contact the person specified here. See Figure 6-13 on page 305.

IBM Remote Support Manager for Storage

[Main](#) > [Configuration](#) > [Contact List](#)
[Logout](#) - [Refresh status](#) - [Help](#)

Contact Person Information

 [System](#): **Problem** [Firewall](#): **Enabled: Closed**
 [Reporting](#): **Suspended** [Remote Access](#): **Disabled**

Who should IBM Support contact:

* next to an entry indicates missing or incorrect information

<input type="text" value="Max Musterman"/>	* Contact person
<input type="text" value="max.musterman@rivers.local"/>	* E-mail address
<input type="text" value="+1 123 456 7890"/>	* Phone number
<input type="text" value="7am - 6pm"/>	* Hours to call
<input type="text" value="+1 123 456 7891"/>	Alternate phone number (optional)
<input type="text" value="6pm - 7am"/>	Hours to call alternate number (optional)
<input type="text" value="GMT - 5"/>	* Time Zone

Make this person the primary contact for the RSM for Storage system

* Country or region

[Update configuration](#)

To remove this contact from the configuration: [Delete this contact](#)



Figure 6-13 RSM contact person information

7. The specified contact will be added to the contact list. Multiple contacts can be defined that might be assigned with different storage subsystems. Click **Configuration** to return to System Configuration, as shown in Figure 6-14.

IBM Remote Support Manager for Storage

[Main](#) > [Configuration](#)
[Logout](#) - [Refresh status](#) - [Help](#)

Contact List

 [System](#): **Problem** [Firewall](#): **Enabled: Closed**
 [Reporting](#): **Suspended** [Remote Access](#): **Disabled**

Who should IBM Support contact:

View/Configure

1 [Max Musterman](#) United States (Primary contact for RSM for Storage)

2 [Select to add](#)

[Main](#) > [Configuration](#)
[Logout](#) - [Refresh status](#) - [Help](#)

Figure 6-14 RSM contact list with contacts

8. The task Contact Information in System Configuration will be marked OK. Click **Company Information**, as shown in Figure 6-11 on page 304.
9. Complete the form with the appropriate information and click **Update configuration**, as shown in Figure 6-15.

IBM Remote Support Manager for Storage

[Main > Configuration](#)
[Logout](#) - [Refresh status](#) - [Help](#)

Company Information

⊘ [System](#): **Problem** [Firewall](#): **Enabled: Closed**
i [Reporting](#): **Suspended** [Remote Access](#): **Disabled**

Company Information:

* next to an entry indicates missing or incorrect information

Rivers.local	* Company name
River Street	* Street Address 1
	Address 2
River City	* City
NC	* State or Province
00000	* Postal Code
<input type="text" value="United States"/>	* Country or region

[Main > Configuration](#)
[Logout](#) - [Refresh status](#) - [Help](#)

Figure 6-15 RSM company information

10. The task Company Information in System Configuration will be marked OK. Click **Connection Information**, as shown in Figure 6-11 on page 304.

11. Complete the form with the appropriate information and click **Update configuration**, as shown in Figure 6-16.

IBM Remote Support Manager for Storage

[Main > Configuration](#)
[Logout](#) - [Refresh status](#) - [Help](#)

Connection Information

⊘ **System:** Problem **Firewall:** Enabled: Closed
ⓘ **Reporting:** Suspended **Remote Access:** Disabled

This system is: **localhost.localdomain** Type: **7985-PBF** Serial: **KQDBXV8**

* next to an entry indicates missing or incorrect information

DIRECT	IP address of SMTP server, or DIRECT (See Help page)
9.11.218.110	IP address of Management Station (optional, see help)
Datacenter 1	* Location of RSM for Storage system
River Street	* Street Address
River City	* City
NC	* State or Province
00000	* Postal Code
United States	* Country or region

Remote Access Connections - you must configure one or both of the following:

For a remote modem connection:

NOMODEM * Modem phone number (0...9 and spaces) or "NOMODEM"

DISABLE * Phone Line Check number. (See Help page.)

For a remote SSH connection:

9.11.218.1 * External IP Address IBM should use (blank to disable)

22 * External Port IBM should use

External Firewall User ID (optional)

External Firewall Password (optional)

External Firewall Password Confirm

22 * SSH Port RSM will listen on for remote access (default: 22)

[Update configuration](#)

[Main > Configuration](#)
[Logout](#) - [Refresh status](#) - [Help](#)

Figure 6-16 RSM connection information

12. The task Connection Information in System Configuration will be marked OK. Click **Storage Subsystems**, as shown in Figure 6-11 on page 304.
13. Click **Select to add** to define a storage subsystem in RSM (Figure 6-17).

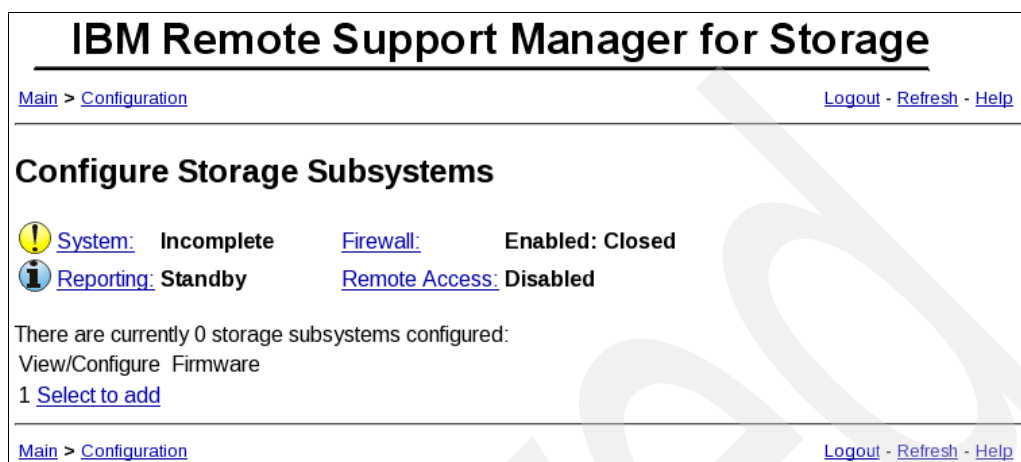


Figure 6-17 RSM configure storage subsystem

RSM for Storage reports DS Storage Systems Expansion units as well, and therefore the configuration of drive expansion units is also required. During the configuration test, the profile for each subsystem will be downloaded. This will verify connectivity to the subsystem, verify that this version of RSM for Storage software is compatible with the firmware on the subsystem, and determine if there are any drive expansion units attached to the controller. If any drive expansion units are detected, the configuration status for the subsystem will change to Configuration Incomplete and additional configuration fields will now be available for setting the IBM machine type and serial numbers of each detected drive expansion unit.

Complete the form with the following fields:



- a. The name of a storage subsystem, as shown in the DS Storage Manager.
- b. The country where the storage subsystem is located.
- c. A description about where to find the storage subsystem, for example, room number or rack number, in the location field.
- d. Management interface IP addresses of all controllers in the subsystem. If only one controller is installed, use only one IP address.
- e. Type, model, and the serial number. This information can be found on the top of the DS storage subsystem. Select an entry from the contact list.

Click **Update configuration**. Figure 6-18 shows the completed form with sample data.

IBM Remote Support Manager for Storage

[Main](#) > [Configuration](#) > [Storage Subsystems](#) [Logout](#) - [Help](#)

Storage Subsystem Information

 [System](#): **Incomplete** [Firewall](#): **Enabled: Closed**
 [Reporting](#): **All Subsystems** [Remote Access](#): **Disabled**

Storage Subsystem Information:

<input type="text" value="ITSO5300"/>	Name
<input type="text" value="United States"/>	Country or region
<input type="text" value="Datacenter 1"/>	Location: Room / Building
<input type="text" value="River Street"/>	Street Address
<input type="text" value="River City"/>	City
<input type="text" value="NC"/>	State or Province
<input type="text" value="00000"/>	Postal Code
<input type="text" value="9.11.218.172"/>	IP Address #1
<input type="text" value="9.11.218.173"/>	IP Address #2 (If present)
<input type="text" value="1818-53A"/>	IBM Product ID (TTTT-MMM) or MTM. Refer to the Help Page for this field.
<input type="text" value="1234567"/>	IBM Serial Number (7 characters)
<input type="text" value="Max Musterman"/>	Contact person for this subsystem

Part of an IBM Solution: **None** (Must be set by IBM Support. Refer to the Help for this page.)

[Update configuration](#)

To remove this subsystem from the configuration: [Delete this device](#)



Figure 6-18 RSM storage subsystem information

14. The storage subsystem will be added to the list of configured storage subsystems, as shown in Figure 6-19. Up to 50 storage subsystems and SAN switches can be added. Click **Configuration** to return to the System Configuration.

IBM Remote Support Manager for Storage

[Main](#) > [Configuration](#) [Logout](#) - [Refresh](#) - [Help](#)

Configure Storage Subsystems

 [System](#): **Incomplete** [Firewall](#): **Enabled: Closed**
 [Reporting](#): **All Subsystems** [Remote Access](#): **Disabled**

There are currently 1 storage subsystems configured:
View/Configure Firmware

1 [ITSO5300](#) **Configuration Incomplete** Not Available
2 [Select to add](#)

[Main](#) > [Configuration](#) [Logout](#) - [Refresh](#) - [Help](#)



Figure 6-19 RSM Configure Storage Subsystems

15. When all tasks are completed correctly, run the configuration test. Click **Run Configuration Test**, as shown in Figure 6-20.

IBM Remote Support Manager for Storage

[Main](#) [Logout](#) - [Refresh status](#) - [Help](#)

System Configuration

 [System](#): **Incomplete** [Firewall](#): **Enabled: Closed**
 [Reporting](#): **All Subsystems** [Remote Access](#): **Disabled**

View and define information for: **localhost.localdomain** Type: **7985-PBF** Serial: **KQDBXV8**

[Contact Information](#): **OK**
[Company Information](#): **OK**
[Connection Information](#): **OK**
[Storage Subsystems](#): **OK** (1 subsystems currently defined)
[Other SAN Devices](#): **OK** (0 switches currently defined)
[System Activated](#): **No**
[Options](#)

Configuration Test

A change has been made to the configuration. The Configuration Test should be re-run.
Test has not been run since the last restart.

[Run Configuration Test](#)

[Main](#) [Logout](#) - [Refresh status](#) - [Help](#)

Figure 6-20 RSM run configuration test

16. Click **Refresh status** to see the progress of the test (Figure 6-21).

The screenshot displays the IBM Remote Support Manager for Storage interface. At the top, the title "IBM Remote Support Manager for Storage" is centered. Below the title, there are links for "Main", "Logout", "Refresh status", and "Help". The main section is titled "System Configuration". It shows several status indicators: "System: Incomplete" (with a yellow warning icon), "Firewall: Enabled: Closed" (with an information icon), "Reporting: All Subsystems" (with an information icon), and "Remote Access: Disabled". Below this, it says "View and define information for: localhost.localdomain Type: 7985-PBF Serial: KQDBXV8". A list of configuration items follows: "Contact Information: OK", "Company Information: OK", "Connection Information: OK", "Storage Subsystems: OK (1 subsystems currently defined)", "Other SAN Devices: OK (0 switches currently defined)", and "System Activated: No". There is also an "Options" link. A red box highlights the "Configuration Test" section, which states "Test is currently running and is 93% complete. (Refresh status to see results)". Below this, a log window shows the following entries: "2009 09 04 18:28:55 UTC - Configuration Test: Checking connectivity to ITS05300", "2009 09 04 18:28:55 UTC - Configuration Test: Validating subsystem name for ITS05300", and "2009 09 04 18:28:55 UTC - Configuration Test: Waiting for results of SMcli: 'show storagesubsystem p'".

Figure 6-21 RSM configuration test

17. The results from this test are logged in the activity log. You can access the activity log by using an icon on the KDE desktop or you can use the command `tail -fn 10000 /var/log/rsm/activity.txt` to see the contents of the log file, as shown in Example 6-5.

Example 6-5 Activity log example

```
2009 09 04 18:28:55 UTC - Starting Configuration Test...
2009 09 04 18:28:55 UTC - Configuration Test: Checking for Modem
2009 09 04 18:28:55 UTC - Configuration Test: NOMODEM specified, skipping modem
checks
2009 09 04 18:28:55 UTC - Configuration Test: Setting all devices in Temporary
Service Access mode for duration of test.
2009 09 04 18:28:55 UTC - Configuration Test: Checking connectivity to
Management Station at 9.11.218.110
2009 09 04 18:28:55 UTC - Configuration Test: Checking connectivity to ITS05300
2009 09 04 18:28:55 UTC - Configuration Test: Validating subsystem name for
ITS05300
2009 09 04 18:28:55 UTC - Configuration Test: Waiting for results of SMcli:
"show storagesubsystem profile"
2009 09 04 18:29:03 UTC - Configuration Test: Checking results of SMcli command
2009 09 04 18:29:03 UTC - History file has been updated
2009 09 04 18:29:03 UTC - Found 2 drive expansion boxes for ITS05300
2009 09 04 18:29:03 UTC - Adding new expansion box (0) to configuration
2009 09 04 18:29:03 UTC - Adding new expansion box (1) to configuration
2009 09 04 18:29:03 UTC - Adding new expansion box (2) to configuration
2009 09 04 18:29:03 UTC - Configuration Test: No problems detected.
2009 09 04 18:29:03 UTC - Configuration Test: Restoring firewall to original
state.
```

18. Disk enclosures were found, and the storage subsystem has the status Configuration Incomplete in the System Configuration window. Go back to Storage Subsystem configuration and click the storage subsystem you just added. Information about disk enclosures has to be updated, as shown in Figure 6-22.

Storage Subsystem Information:

<input type="text" value="ITSO5300"/>	Name
<input type="text" value="United States"/> <input type="button" value="v"/>	Country or region
<input type="text" value="Datacenter 1"/>	Location: Room / Building
<input type="text" value="River Street"/>	Street Address
<input type="text" value="River City"/>	City
<input type="text" value="NC"/>	State or Province
<input type="text" value="00000"/>	Postal Code
<input type="text" value="9.11.218.172"/>	IP Address #1
<input type="text" value="9.11.218.173"/>	IP Address #2 (if present)
<input type="text" value="1818-53A"/>	IBM Product ID (TTTT-MMM) or MTM. Refer to the Help Page for this field.
<input type="text" value="1234567"/>	IBM Serial Number (7 characters)
<input type="text" value="Max Musterman"/> <input type="button" value="v"/>	Contact person for this subsystem

Part of an IBM Solution: **None** (Must be set by IBM Support. Refer to the Help for this page.)

Drive Expansion Units

Expansion Unit ID: 0

<input type="text"/>	* IBM Product ID (TTTT-MMM) or MTM.
<input type="text"/>	* IBM Serial Number (7 characters)

Expansion Unit ID: 1

<input type="text"/>	* IBM Product ID (TTTT-MMM) or MTM.
<input type="text"/>	* IBM Serial Number (7 characters)

To remove this subsystem from the configuration:

Figure 6-22 Adding disk enclosures to RSM

19. After the configuration is updated, go to the Main menu, select **System**, and run the configuration test. When the test is complete, the date and time of the last run is shown, as shown in Figure 6-23.

IBM Remote Support Manager for Storage

[Main](#) [Logout - Refresh status - Help](#)

System Configuration

! **System:** Incomplete [Firewall:](#) Enabled: Closed
i **Reporting:** All Subsystems [Remote Access:](#) Disabled

View and define information for: **localhost.localdomain** Type: **7985-PBF** Serial: **KQDBXV8**

[Contact Information:](#) OK
[Company Information:](#) OK
[Connection Information:](#) OK
[Storage Subsystems:](#) OK (1 subsystems currently defined)
[Other SAN Devices:](#) OK (0 switches currently defined)
[System Activated:](#) No
[Options](#)

Configuration Test

Test was last run at: Fri Sep 4, 2009 18:39:14 UTC

No problems were found.

[Run Configuration Test](#)

[Main](#) [Logout - Refresh status - Help](#)

Figure 6-23 RSM configuration test done

6.2.4 Configuring SNMP traps in Storage Manager

To allow the DS storage subsystem management workstation to send SNMP traps to the RSM server, set the RSM server as your SNMP traps destination in the Storage Manager client.

To configure the DS Storage Manager to send SNMP alerts for each defined storage subsystem in RSM to the RSM host, perform these steps:

1. Open the Enterprise management window of the Storage Manager.
2. Right-click a DS storage subsystem. In the Connect menu, click **Configure Alerts...**
3. In the Configure Alerts window, click the **SNMP** tab.
4. Enter the host name or the IP address of the RSM host in the trap destination field and click **Add** (Figure 6-24 on page 315). Do not change the SNMP community name (the default is *public*).

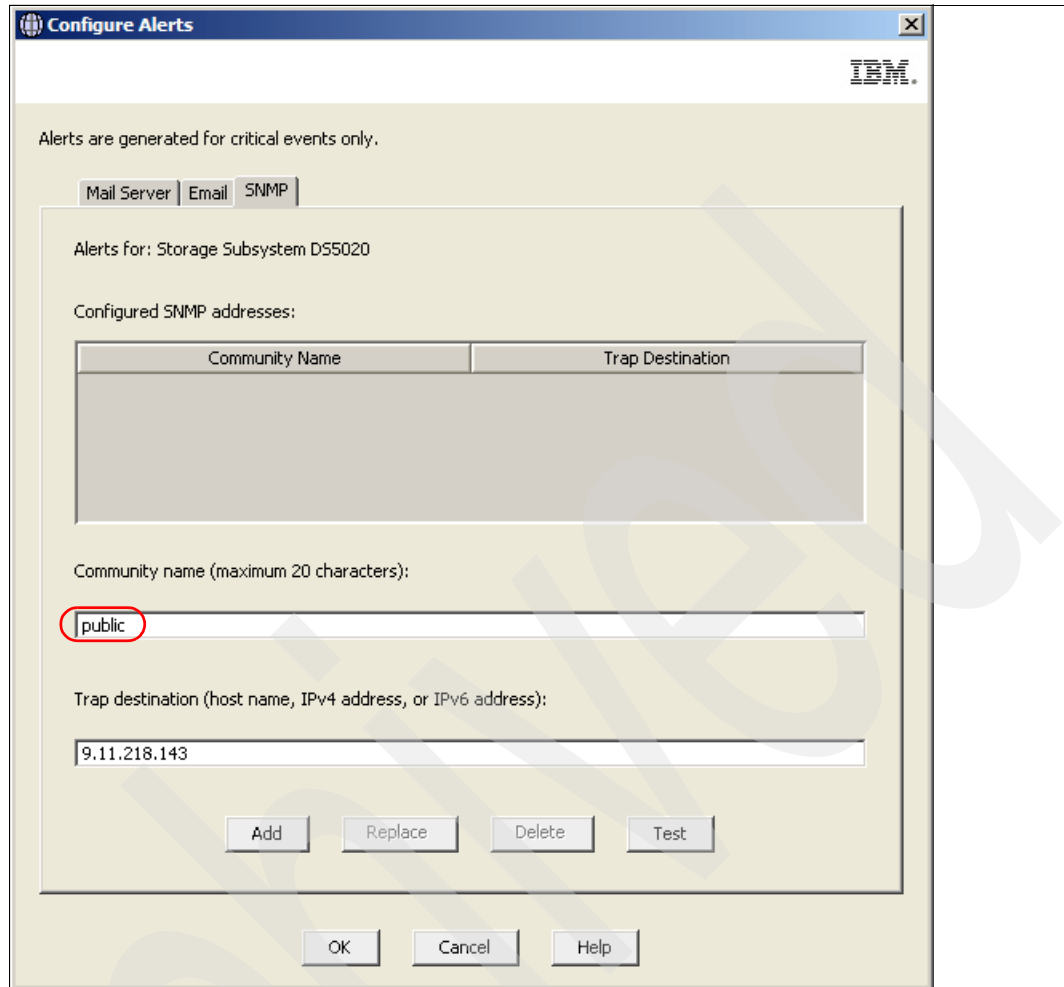


Figure 6-24 Add SNMP trap receiver

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.

Validate SNMP configuration

To validate the SNMP configuration, select the SNMP trap destination and click **Test** to send a test trap to the RSM host (Figure 6-25).

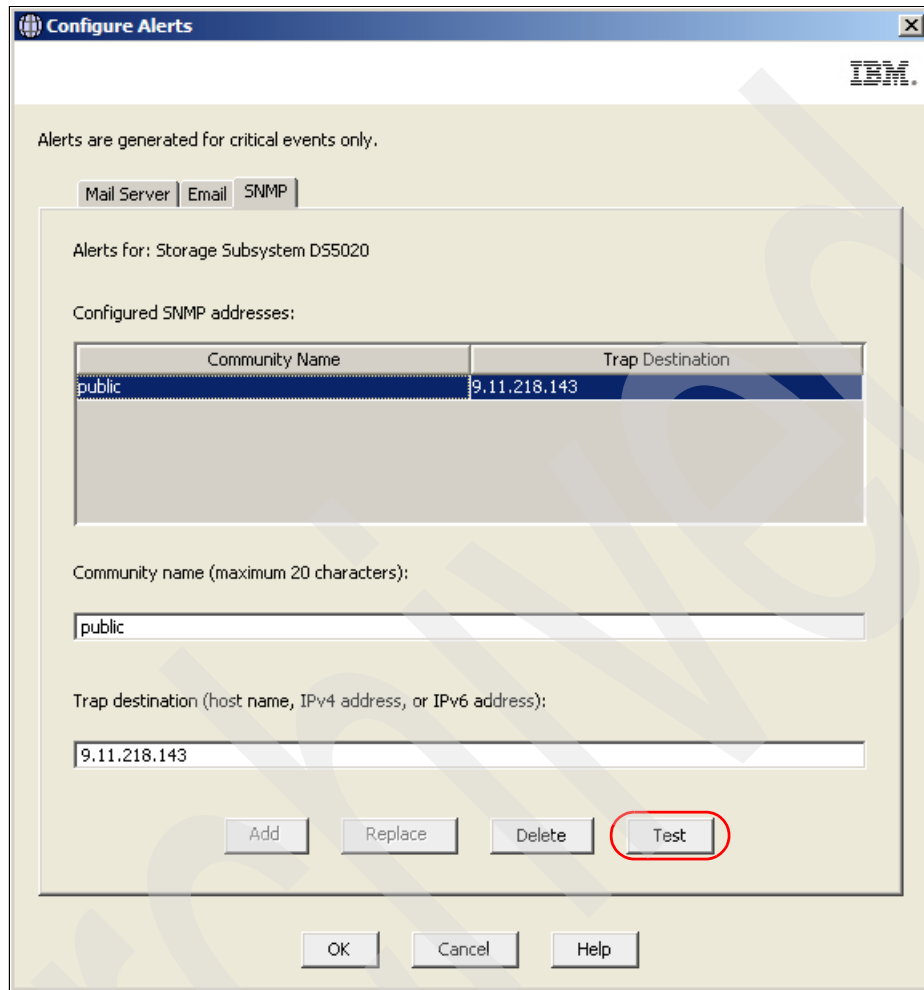


Figure 6-25 Send test trap

Check the activity log (as shown in Figure 6-23 on page 314) and verify that the trap was received. The activity log will contain an entry, as shown in Example 6-6.

Example 6-6 Test alert received

2009 09 04 22:05:08 UTC - Received a test alert for ITS053000 from 9.11.218.181

6.2.5 Activating RSM

The final step is to activate your system. Complete all the other configurations and run a successful Configuration Test before contacting IBM Service to activate RSM.

To activate RSM, perform these steps:

1. Make sure Remote Access is enabled. Click **Remote Access** in the main RSM window to activate it, if required (as shown in 6.2.6, “Remote access security” on page 317).
2. Call the number for IBM Service for your region and give the IBM Machine Type and Serial Number of one of the DS storage subsystems to be monitored by RSM for Storage. For support telephone numbers in your country or region, see the following Web site:

<http://www.ibm.com/planetwide>

3. Tell the disk support contact person that you are activating a RSM for Storage system.
4. Provide IBM Service with the phone number of the modem attached to the RSM for Storage system (if used). IBM Service will connect to the system, verify that the configuration is correct, send a test alert via e-mail, verify receipt of the alert and associated attachments, and then activate the system.

When the RSM for Storage system is ready to receive events, the System Status will be OK. See Figure 6-26.

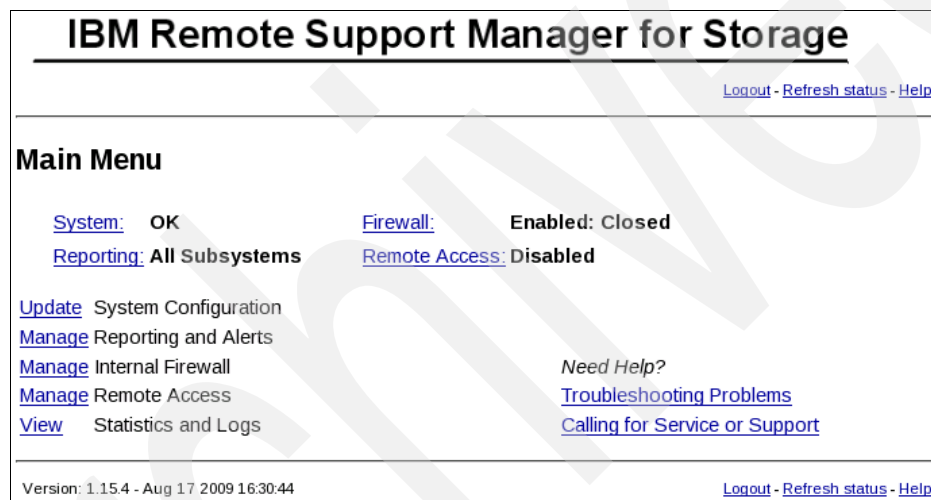


Figure 6-26 RSM main menu: Configured

Note: If you decide to install the RSM for Storage software on another server, you will need to contact IBM Service to obtain a new activation key for the new server.

6.2.6 Remote access security

This section will cover how to secure your IBM Midrange System Storage storage subsystem from unauthorized access.

SSH connectivity

In the RSM for Storage Remote Access window, click **Enable Remote Access**. This will reconfigure the RSM for Storage internal firewall to allow connections through SSH port 22.

Verify connectivity in the following sequence:

1. From inside your network, open an SSH client and connect to the RSM for Storage system on port 22. (Remember, if you perform these connectivity checks over several days, that the RSM for Storage Remote Access control has a timeout that might need to be reset.) Verify that you are able to obtain a login prompt from the RSM for Storage system.

2. From outside your network, open an SSH client and connect to your external IP address port that has been assigned (port mapped) to the RSM for Storage system.

You should be connected to the RSM for Storage system and receive a login prompt.

Note: You will not be able to complete the login. Authentication requires a special tool only available to IBM Remote Support.

If an authentication process has been put in place by your firewall administrator, verify that the user ID and password for the external firewall is specified in the RSM for Storage Connections configuration.

3. Verify that the SSH connection is correctly configured on the Connections Configuration page. This information will be encrypted and sent with each alert to IBM.

Modem connectivity

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

In RSM, the modem used 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.

On the RSM for Storage Remote Access window, click **Enable Remote Access**. This will enable the modem to answer when called. Verify modem connectivity by calling the modem phone number from a voice phone:

1. Most modems will either flash an LED or you might hear a sound when a call is being received. If there is no indication that a call is being received:
 - a. Plug the phone cable into an analog voice phone and verify that a dial tone is present.
 - b. If a dial tone is present, hang up and then call this phone number from another phone and verify that the phone you just connected rings.
 - c. Reconnect the phone cord to the modem connector labeled line #1.
2. Try again to verify the modem connectivity by calling the modem phone number from a voice phone.
3. Check the instructions that came with the modem and review any troubleshooting information.

To configure the Remote Access policy, click **Remote Access** from the Main Menu, as shown in Figure 6-26 on page 317.

In the Remote Access setting window, 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 6-27.

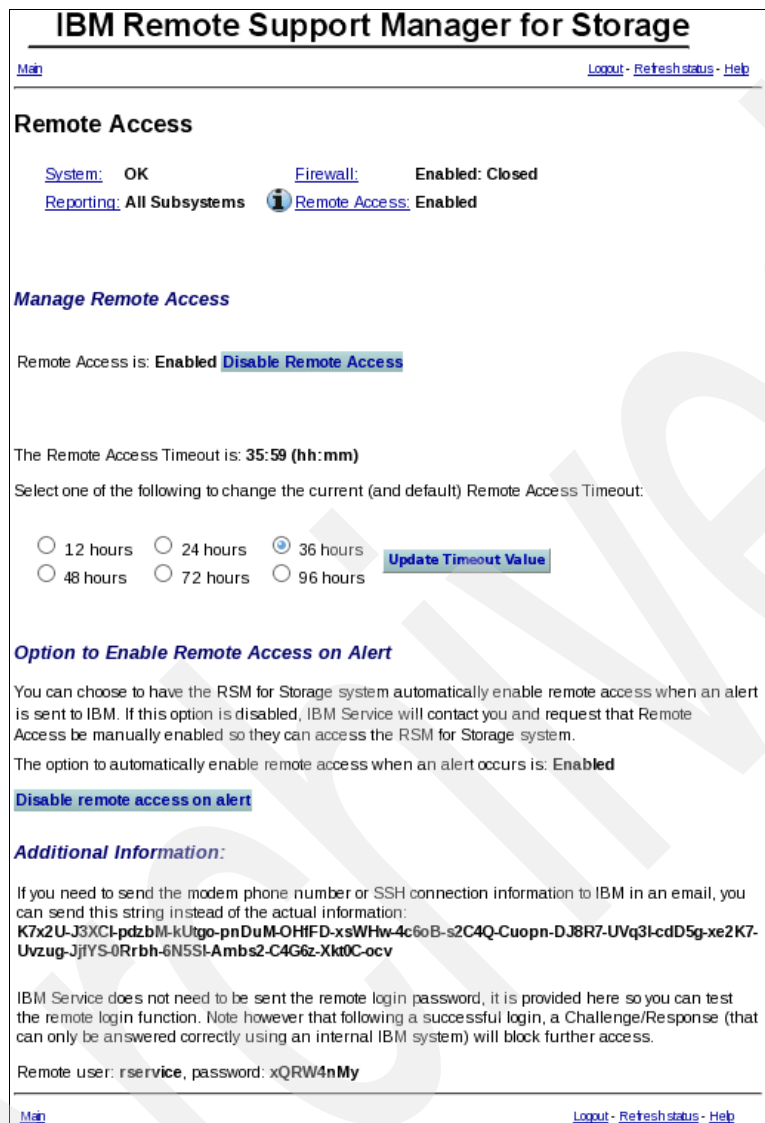


Figure 6-27 Remote Access settings

Remote Access also has a configurable timeout between 12 to 96 hours. You can manually disable remote access when the service is complete or allow it to time out. After the timeout period has elapsed, the system is guaranteed to return to a secure state without intervention.

To configure the timeout value, scroll down the Remote Access settings window, select the desired timeout value, and click **Update Timeout Value**, as shown in Figure 6-27.

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 software.

- ▶ The normal state for the firewall is Enabled:Closed, which means that the firewall is operational and configured to allow SNMP traps to be received and e-mails to be sent; however, access to other devices on your network is not allowed.
- ▶ The Enabled:Custom state indicates that one or more custom rules have been added to `/etc/rsm/rsmfirewall.conf`. These rules will be active any time the firewall is enabled.
- ▶ The Enabled:Open state means that access to one or more other devices has been enabled. The firewall allows access to any storage subsystem that has an active alert, and also storage subsystems and other SAN devices that have been placed in Service Access mode.

Service Access mode allows you to manually allow access to a device from the RSM for Storage system. You can select storage subsystems that you have previously configured.

Disabling the firewall allows unrestricted access from the RSM for Storage system to your network. To maintain the security of your network, disabling the firewall will also disable remote access. Likewise, enabling Remote Access will automatically enable the firewall.

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 the 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 6-26 on page 317.

In the Internal Firewall and Service Access window, you can change the internal firewall status and service access mode of your managed storage subsystems, as shown in Figure 6-28.

IBM Remote Support Manager for Storage

[Main](#) [Logout - Refresh status - Help](#)

Internal Firewall and Service Access

[System:](#) OK [Firewall:](#) Enabled: Closed
[Reporting:](#) All Subsystems [Remote Access:](#) Enabled

Manage Internal Firewall

Firewall is **Enabled** - All connections are blocked except those required for reporting.

Note: Disabling the firewall will also disable remote access.
[Disable Firewall](#)

RSM for Storage Firewall Status

Current Rules	Your Network	Internet
HTTPS ←	Systems Configured in RSM	IBM's WWW & FTP Sites
SNMP ←		
SSH ←		
IBM Sites →	Other Systems	

Grey No Access
 Yellow Possible Access
 Green Full Access

Manage Service Access

Any configured device may be placed in temporary in Service Access mode. This creates a rule for the internal firewall that will allow connections to that device from the RSM for Storage system. These temporary firewall rules will be removed when the Service Access Timeout expires.

[Manage](#) Service Access for storage subsystems. (0 - currently enabled)
[Manage](#) Service Access for other SAN devices (0 - currently enabled)
[Disable Service Access for all devices](#)

The Service Access Timeout is: **N/A**

Select one of the following to change the current (and default) Service Access Timeout:

12 hours
 24 hours
 36 hours
 48 hours
 72 hours
 96 hours
 [Update Timeout Value](#)

[Main](#) [Logout - Refresh status - Help](#)

Figure 6-28 Internal Firewall and Service Access window

Placing a device into Service Access mode will create a rule for the internal firewall that will allow 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 window, select the desired Service Access Timeout value, and click **Update Timeout Value**, as shown in Figure 6-28 on page 321.

IBM Midrange System Storage security

The IBM Midrange System Storage 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 IBM Midrange System Storage storage subsystems also have a controller shell environment that is accessible using a remote login (RLOGIN) client. IBM Midrange System Storage Storage Manager has an option to disable RLOGIN, and we normally recommend disabling RLOGIN.

See “Securing the DS5000 storage subsystem” on page 124 and 4.8.3, “Configuring IP addresses of the controllers” on page 161 for information about how to set the Storage Manager password and to disable remote login.

6.2.7 Managing alerts

To manage alerts, perform these steps:

1. When RSM receives SNMP alerts from one of the defined storage subsystems, an attention mark (!) is shown next to the reporting link, as shown in Figure 6-29. Click **Reporting** to see the alerts.

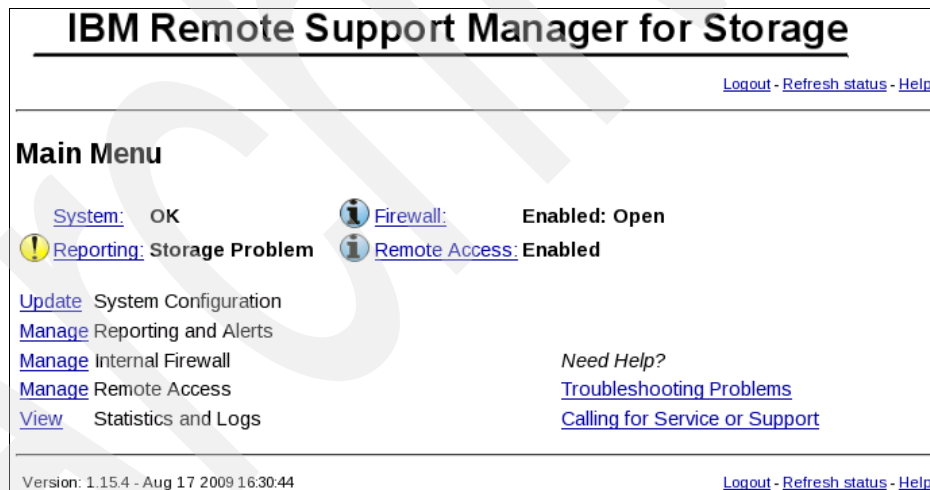





Figure 6-29 Main Menu: Storage Problem

2. The reporting and alert window shows all those subsystems that have sent alerts sent to RSM (Figure 6-30 on page 323). Click **View or Close Alerts**.

IBM Remote Support Manager for Storage

[Main](#) [Logout - Refresh status - Help](#)

Reporting and Alerts

[System](#): OK  [Firewall](#): Enabled: Open
 [Reporting](#): Storage Problem  [Remote Access](#): Enabled

Products with Active Alerts: **1** Alerts sent to IBM: **1**
 Products with Reporting Enabled: **1** Alerts Pending: **0**
 Products with Reporting Disabled: **0** Alerts Acknowledged: **0**

[View/Change](#) reporting state for each subsystem
 Events waiting to be processed: **1**
[View Closed Alerts](#) (**0**)

Active Alerts:	Sent	Acknowledged	Pending	Firmware
View or Close Alerts for: ITSO5300	1	0	0	07.60.13.04

[Main](#) [Logout - Refresh status - Help](#)




Figure 6-30 RSM storage subsystem with alerts

- The alert list for the selected storage subsystem (Figure 6-31) shows all alerts that were received by RSM. Click **View** to see the details of an alert.

IBM Remote Support Manager for Storage

[Main > Reporting and Alerts](#) [Logout - Refresh Status - Help](#)

Alert List for ITSO5300

[System](#): OK  [Firewall](#): Enabled: Open
 [Reporting](#): Storage Problem  [Remote Access](#): Enabled

Subsystem: **ITSO5300, 9.11.218.172** 9.11.218.173
 IBM Type/Serial: **1818-53A / 1234567**
 Firmware Version: **07.60.13.04** (as of Fri Sep 4, 2009 11:39:14)
 Location: Datacenter 1, River Street, River City, 00000, United States
 Contact: Max Musterman

Total alerts for this subsystem: **2**

NOTE: Do not close alerts unless IBM Service has contacted you about the problem. See the help page for more information.

View - Close	Date and Time	State	Duplicates	Event Code
	Tue Sep 8, 2009 20:27:46 UTC	Sent	0	226c
	Tue Sep 8, 2009 20:27:46 UTC	Getting Logs...	0	222d

[Close all active alerts for ITSO5300](#)

[Main > Reporting and Alerts](#) [Refresh Status - Help](#)

Figure 6-31 RSM list of alerts

4. The alert details and an error message are shown in Figure 6-32.

IBM Remote Support Manager for Storage

[Main](#) > [Reporting and Alerts](#) > [ITSO5300](#) [Logout](#) - [Refresh Status Help](#)

Alert Details

[System](#): OK  [Firewall](#): Enabled: Open
 [Reporting](#): Storage Problem  [Remote Access](#): Enabled

Subsystem Name: **ITSO5300**
Type and Model: **1812-81A**
Serial number: **1234568**
Alert Status: **Sent**
IBM Service logged in? **No**
Time of alert: **Tue Sep 8, 2009 20:27:46 UTC**
Sent to IBM at: **Tue Sep 8, 2009 20:27:58 UTC**
Duplicates: **0**
Time of last duplicate: **N/A**
Alert Event Code: **226c**
Alert Description: **Drive failure.
Drive.
Enclosure 0, Slot 1**
Log File Available: **Log files could not be downloaded**

NOTE: Do not close alerts unless you have talked with IBM Service
[Close this alert](#) (Service is complete)

[Main](#) > [Reporting and Alerts](#) > [ITSO5300](#) > [Refresh Status Help](#)

Figure 6-32 RSM alert details

5. The alert list of the selected storage subsystem shows the modified status of an alert, as shown in Figure 6-31 on page 323. When the problem is solved, click **Close**.
6. After the alert is closed, it disappears from the alert list.
7. The main menu status changes so that the attention mark disappears after all problems are solved (Figure 6-33).

IBM Remote Support Manager for Storage

[Main](#)
[Logout](#) - [Refresh status](#) - [Help](#)

Reporting and Alerts

[System:](#) **OK** [Firewall:](#) **Enabled: Closed**

[Reporting:](#) **All Subsystems** [Remote Access:](#) **Enabled**

Products with Active Alerts:	0	Alerts sent to IBM:	0
Products with Reporting Enabled:	1	Alerts Pending:	0
Products with Reporting Disabled:	0	Alerts Acknowledged:	0

[View/Change](#) reporting state for each subsystem

There are no active alerts.

[View Closed Alerts \(2 \)](#)

[Main](#)
[Logout](#) - [Refresh status](#) - [Help](#)

Figure 6-33 RSM main menu

Archived

Advanced maintenance, troubleshooting, and diagnostics

In this chapter, we explain the advanced Storage Manager functions. 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.

For some operating systems that can be used with the DS5000, we cover tools to manage your DS5000 logical disks, with commands' usage and examples.

We address advanced maintenance topics for the IBM System Storage DS5000 storage subsystem, such as:

- ▶ Upgrade firmware for the different DS5000 components
- ▶ Upgrade HBA firmware
- ▶ Other management and maintenance tasks, such as handling the premium features
- ▶ How to back up your configuration
- ▶ How to perform drive migration from between different DS storage subsystems using the new facilities for import-export volumes

7.1 Upgrades and maintenance

This section covers how to manage your DS5000 storage subsystem firmware, describing the steps that are required to upgrade the DS5000 storage subsystem code (this includes the Storage Manager client and firmware of all components) to the latest supported level.

7.1.1 Displaying installed firmware versions

You need to know the firmware level currently installed in your DS5000 storage subsystem. 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 DS5000 storage subsystem. You can view this date by performing the following steps:

1. Select **Storage Subsystem**, → **View** → **Profile** from the Subsystem Management window, or **Storage Subsystem** → **View Profile** if using older levels of the Storage Manager client.
2. Click the **Controller** tab to display a window, as shown in Figure 7-1. This shows the controller firmware (CFW) and NVSRAM version installed

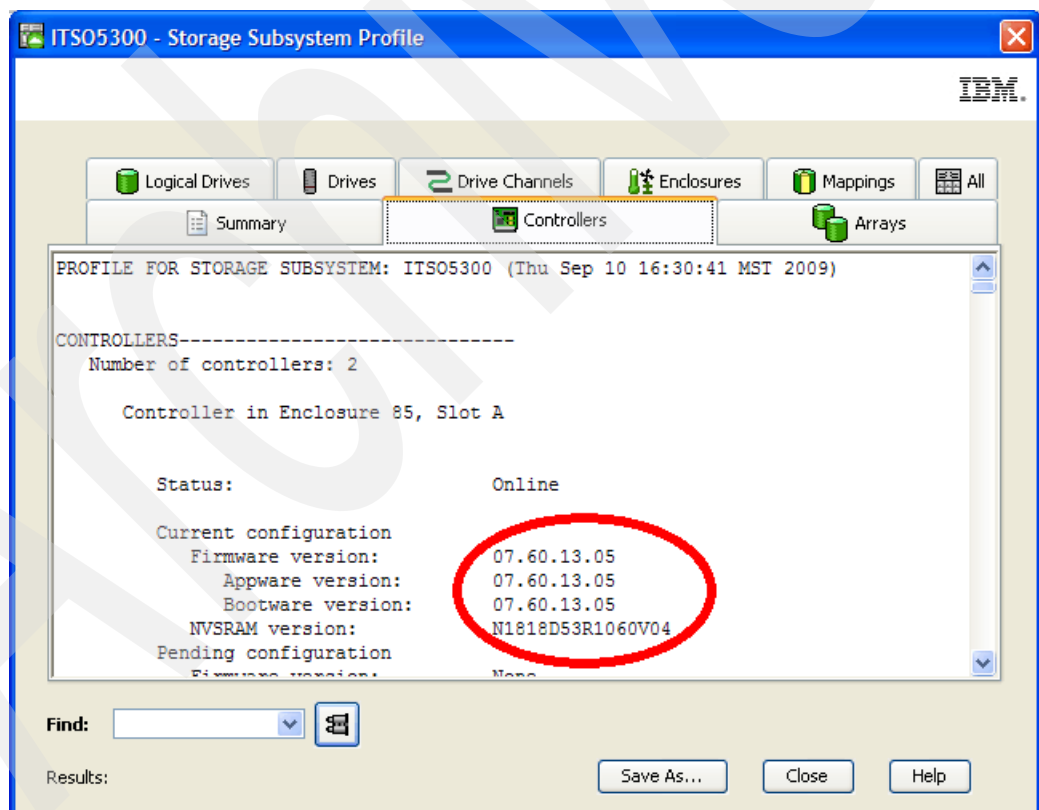


Figure 7-1 Controller and NVSRAM version in Storage Subsystem Profile

3. Click the **Enclosures** tab to see the ESM current firmware level. As shown in Figure 7-2, you have to use the scroll bar to show all of the ESM firmware (two per enclosure), as shown in Figure 7-2 by the pointers.

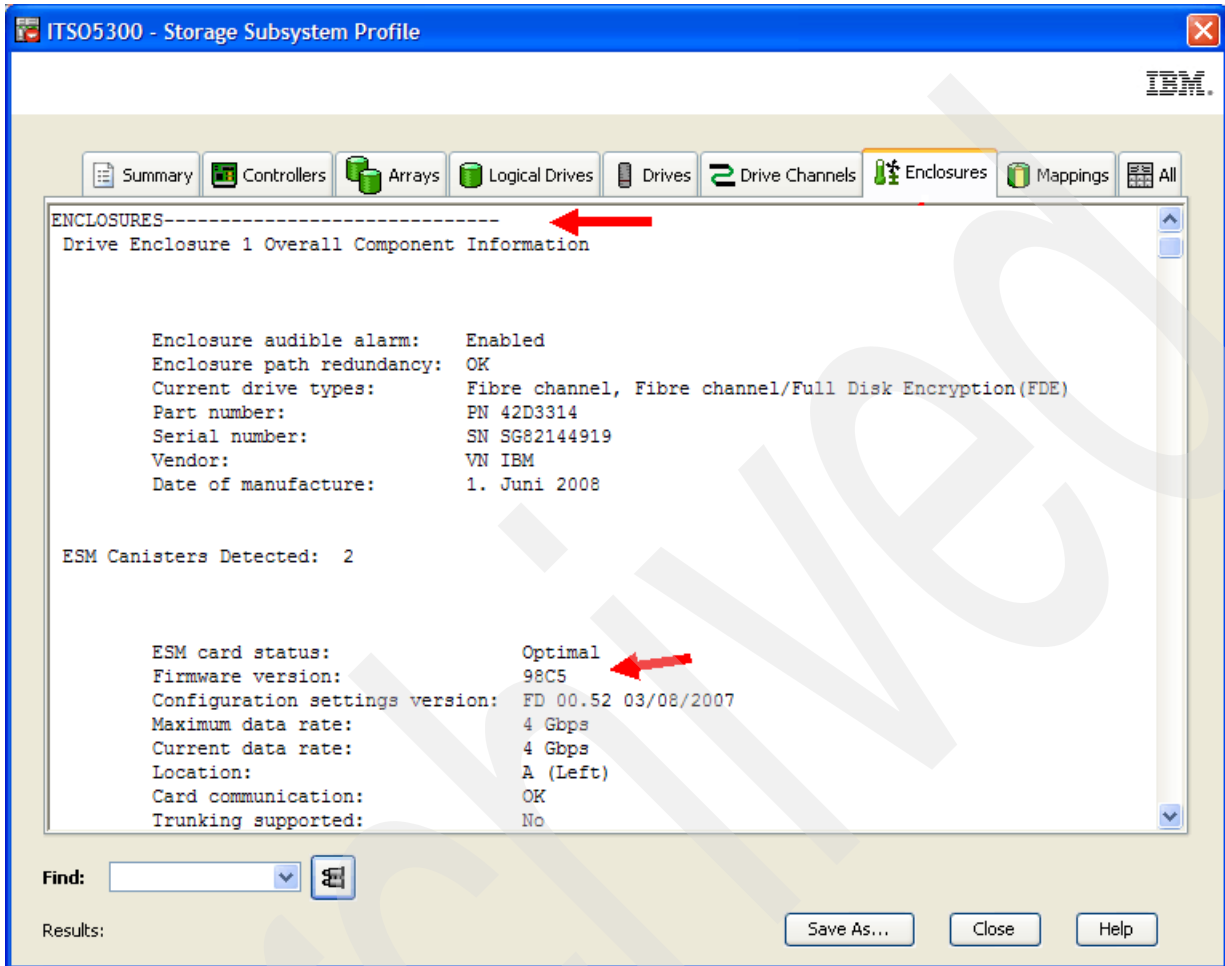


Figure 7-2 Viewing ESM firmware

- Click the **Drives** tab to see the drives' information. Scroll to the right side of the window until the column Firmware Version is visible. Be aware that you might have different drive types in your enclosures, so you can find multiple versions. See Figure 7-3.

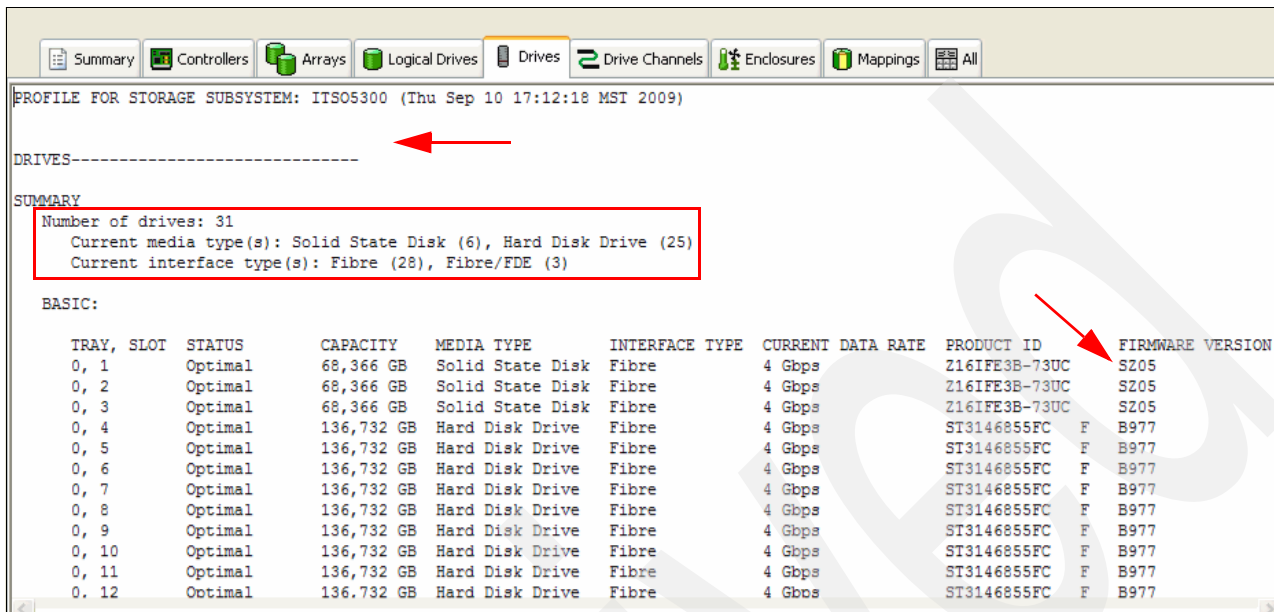


Figure 7-3 Viewing the drives' firmware

7.1.2 Obtaining updates

To download the latest firmware and drivers, and for updates, hints, and tips about known problems, fixes, technical notes, and product flashes (alerts), consult the IBM Support Web site at the following address:

<http://www.ibm.com/servers/storage/support/disk>

Select your specific DS5000 storage subsystem 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 DS5000 storage subsystem.

Info: You will be required to sign in with an IBM Identity in order to access the code.

If you do not already have one, click the **Register** link on the Sign in page to obtain one. IBM Support is instituting this new procedure so that those who download the code can be contacted if that becomes necessary.

Support Notifications subscription service

There is a Support Notifications subscription service that can be accessed at the IBM System Storage product support Web site. This service is designed to keep you informed of new or updated IBM System Storage support site information, without the need to periodically check the IBM Support Web site for updates. More information about the Support Notifications service can be found at the following address:

<http://www-304.ibm.com/systems/support/storage/subscribe/moreinfo.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 for the latest supported combinations at the following Web site:

<http://www.ibm.com/systems/support/storage/config/ssic/index.jsp>

This replaces the older compatibility matrix PFD files.

7.1.3 Planning for upgrades

Upgrading the firmware and management software for the DS5000 storage subsystem is a relatively simple procedure, but some preparation and caution are needed to ensure a smooth upgrade.

Important: Always check the readme files of the firmware you are going to install for prerequisites before proceed with the update.

Attention: Users upgrading from 6.xx.xx.xx controller firmware *must* upgrade to 7.1x.xx.xx or 7.36.xx.xx firmware first before upgrading to 7.50.13.00 or above.

Check the Dependencies section for further details. The components that typically need to be updated are:

- ▶ Storage Manager software
- ▶ Controller firmware
- ▶ NVSRAM firmware
- ▶ ESM firmware
- ▶ Disk drives firmware

Note: Always make a backup of data before you start an upgrade.

Upgrading large storage configurations can be time consuming. Time estimates for upgrading all the associated firmware and software are listed in Table 7-1.

Table 7-1 Upgrade times

Element being upgraded	Approximate time of upgrade
Storage Manager and associated drivers and utilities	35 minutes
ESM firmware	5 minutes per ESM
DS5000 storage subsystem firmware and NVSRAM	5 to 35 minutes
Hard drives	2 minutes per drive (but it is possible to do a parallel firmware upgrade, even for multiple drive types)

These times were observed in the lab. They are approximate and might vary from system to system. When activating controller firmware or NVSRAM, the maximum times are 15 minutes per controller.

It is critical that if you update one component that you also update other components to the matching levels (see the readme file for the release to which you are upgrading). You must *not* run a mismatched set.

Linux environments

For Linux environments, it is critical that you have the correct kernel version installed in order to install the RDAC driver. RDAC is still the preferred failover driver for Linux systems. The Storage Manager for Linux readme lists the supported kernel versions. See Figure 7-4 for an example with SM10.60.

OS	Version	RDAC
LoP-Redhat 4 update 7 (RHEL4-u7)	2.6.18-128.E1	Yes (note a) rdac-LINUX-09.03.0B05.0214
LoP-Redhat 5 update 3 (RHEL5-u3)	2.6.18-128.EL5	Yes (note b) rdac-LINUX-09.03.0C05.0214
LoP-SLES 9 Service Pack 4 (SLES9-SP4)	2.6.5-7.308-smp	Yes (note a) rdac-LINUX-09.03.0B05.0214
LoP-SLES 10 Service Pack 2 (SLES10-SP2)	2.6.16.60-smp	Yes (note a) rdac-LINUX-09.03.0C05.0214
LoP-SLES 11	2.6.27.19-5-default	Yes (note a) rdac-LINUX-09.03.0C05.0214

Notes:

- The redundant failover/failback capability is supported by the IBM DS Storage Manager Linux RDAC driver versions 9.03.0B05.0214 for SLES 9, RedHat 4 and SLES 10 (2.6 kernels), and 9.03.0C05.0214 for Redhat 5.x and SLES10 SP1 and greater and SLES 11.
- Redhat 5.0 is supported by the Storage Manager client only.

Figure 7-4 Linux kernel listing an example for SM V10.60

7.1.4 Updating the DS5000 storage subsystem host software

In this section, we discuss updating the DS5000 storage subsystem host software on a Windows and Linux host server, and we present a specific sequence to follow.

Important: Storage Manager host software has to be updated to V10.60 to manage a DS5000 storage subsystem with firmware V7.60. You must update the SM host software before performing the firmware upgrade.

Older storage subsystems running CFW 5.x are no longer manageable with Storage Manager Version 10.60.

Code update for the Windows environment

The IBM System Storage DS Storage Manager (using the Installation wizard InstallAnywhere) supports an installation upgrade. Launch the installation program and it automatically upgrades all the components that you previously installed. For details about the installation procedures, see 4.7, “Installing IBM System Storage DS Storage Manager” on page 152.

Code update for the Linux environment

You can update the host-based DS5000 storage subsystem software for Linux either with InstallAnywhere or manually. See the readme of the host software for a detailed procedure.

The steps below explain the manual update procedure for all hosts running with RDAC:

1. Uninstall the earlier version of Storage Manager components.
2. Install SMruntime.
3. Install SMclient.
4. Disable and enable the Event Monitor.
5. Install the IBM FC HBA non-failover version device driver for Linux.
6. Install Linux RDAC.
7. Install SMagent (optional).
8. Install SMutil.
9. Make sure that the correct host type is specified.

To update the Linux environment with IBM HBA failover driver as a multipath device driver, you must perform these steps:

1. Uninstall the earlier version of Storage Manager components.
2. Install SMruntime.
3. Install SMclient.
4. Disable and enable the Event Monitor.
5. Install SMutil.
6. Uninstall the existing IBM DS5000 Management Suite Java.
7. Install the IBM FC HBA failover version device driver for Linux.
8. Install the IBM DS5000 Management Suite Java and QLRemote Agent.
9. Configure path failover/failback. See the device driver readme file for detailed information.

7.1.5 Updating controller firmware

You can transfer the controller firmware and NVSRAM to a designated flash area on the DS5000 storage subsystem controllers and activate it at a later time, or transfer the files and apply them at the same moment.

The firmware is transferred to one of the controllers, which then copies the file to the other. The image is verified through a CRC check on both controllers. If the checksum is okay, the uploaded firmware is marked as ready and available for activation. If one of the two controllers fails to validate the CRC, the image is marked as invalid on both controllers and is not available for activation. An error is returned to the management station as well.

For information about how to use the IBM System Storage DS4000 Controller Firmware Upgrade Tool to upgrade your DS4700 Express controller firmware from 06.xx to 07.xx, see 7.1.6, “Controller Firmware Upgrade Tool” on page 339.

Important: Always check the readme files of the firmware you are going to install for prerequisites before proceed with the update.

Upgrading firmware and NVSRAM

The microcode of the DS5000 storage subsystem controllers consists of two packages:

- ▶ Controller firmware
- ▶ NVSRAM

The NVSRAM is similar to the settings in the BIOS of a host system. The firmware and the NVSRAM are closely tied to each other and are *not* independent. Be sure to install the correct combination of the two packages.

Important: Before upgrading the storage subsystem firmware and NVSRAM, make sure that the system is in an optimal state. Fix any problems before you proceed with the upgrade.

The upgrade procedure needs two independent connections to the DS5000 storage subsystem, one for each controller. It is not possible to perform a microcode update with only one controller connected. Therefore, both controllers must be accessible either through Fibre Channel or Ethernet. Both controllers must also be in the Optimal state.

Important: Make sure to install the firmware for each of the DS5000 storage subsystem components (ESM, drives, controller, and NVSRAM) in the sequence described in the readme file for that version.

Update the controller firmware and then the NVSRAM, or both at the same time.

Any power or network/SAN interruption during the update process might lead to configuration corruption or extended downtime. Therefore, do not power off the DS5000 storage subsystem or the management station during the update. If you are using in-band management and have Fibre Channel hubs or managed hubs, then make sure that no SAN-connected devices are powered up during the update. Otherwise, this can cause a loop initialization process and interrupt the process.

In general, the controller, NVSRAM, and ESM firmware upgrades can be done online during non-peak periods of traffic if your DS5000 storage subsystem model has redundant controllers, and if a redundant driver is installed in all the hosts being serviced by your DS5000 storage subsystem.

In the above conditions, after the upgrade is initiated, the files are transferred to one controller, which then copies them to the other.

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, and 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 DS5000 storage subsystem does not activate the new firmware; it is only activated after the user specifically chooses to activate the firmware.

To perform the firmware and NVSRAM update, perform these steps:

1. Open the Subsystem Management window for the DS5000 storage subsystem that you want to upgrade. To download the firmware, select **Advanced** → **Maintenance** → **Download** → **Controller Firmware**, as shown in Figure 7-5.

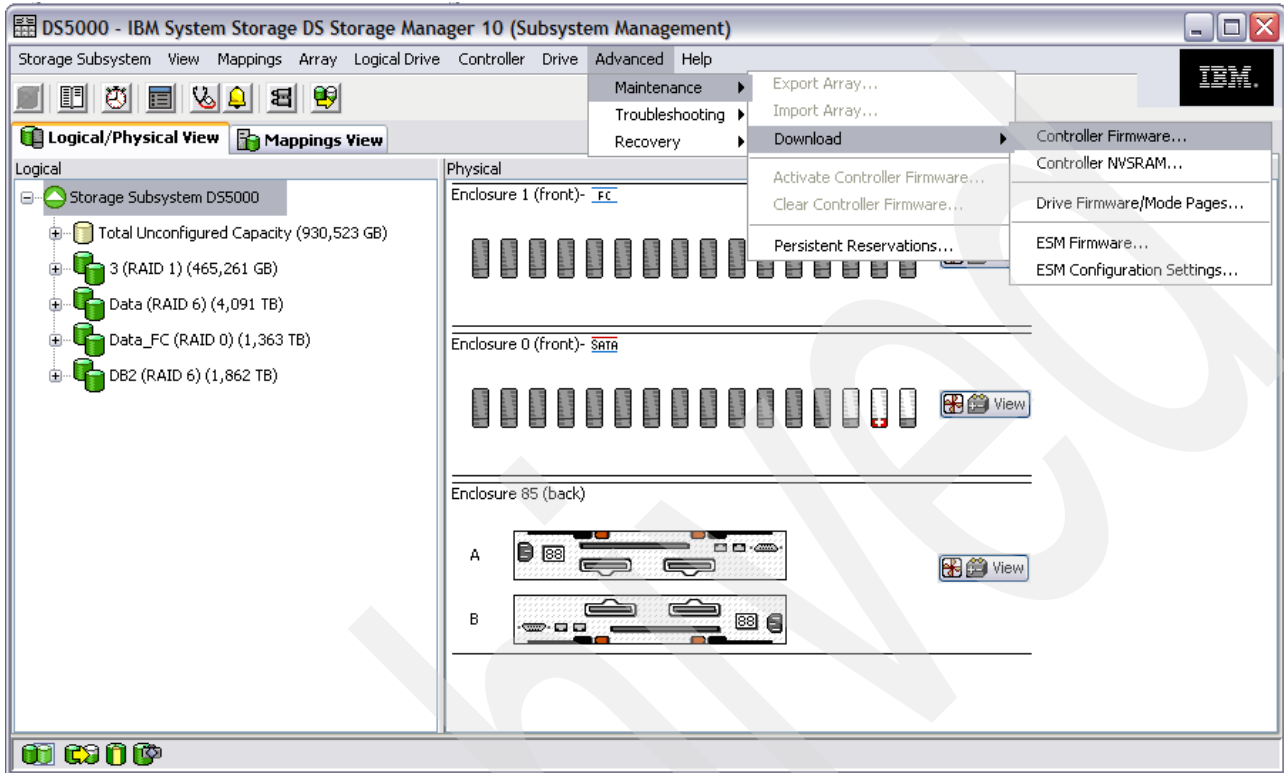


Figure 7-5 Subsystem Management window: Controller firmware update

2. The Download Firmware window opens, showing the current firmware and NVSRAM versions. Select the correct firmware and NVSRAM files, as shown in Figure 7-6. Check the check box to download the NVSRAM file as well.

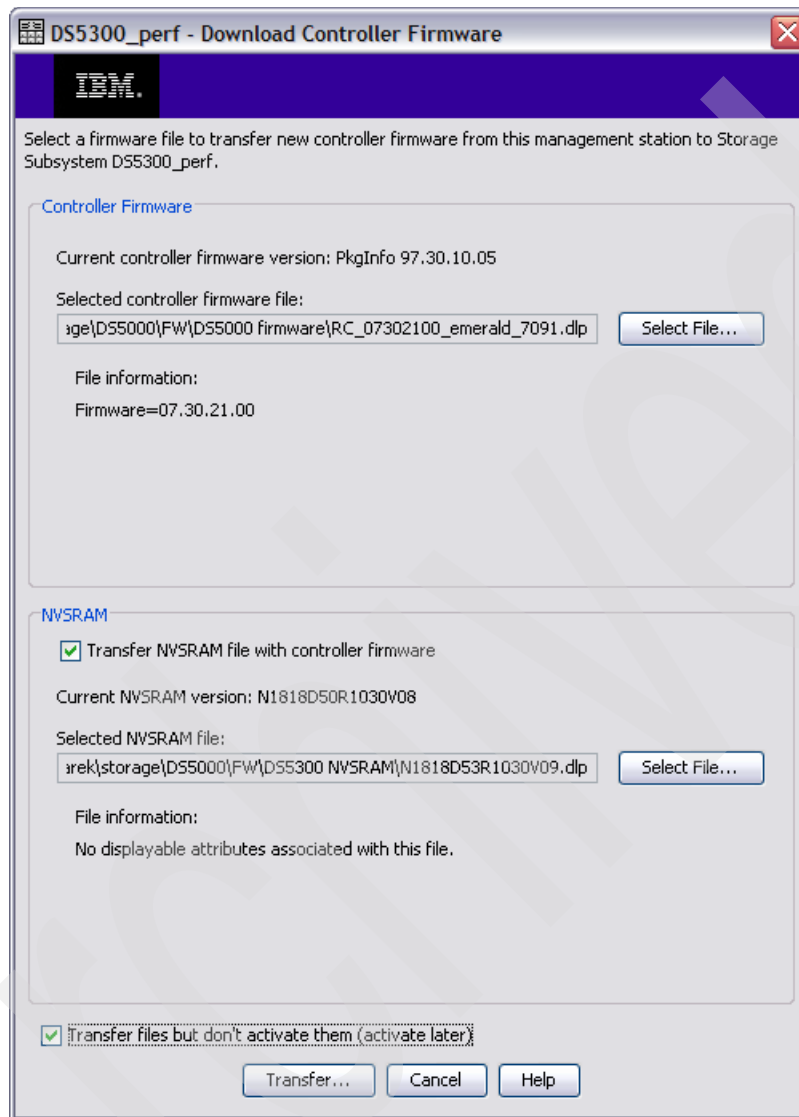


Figure 7-6 Download Firmware window

There is another check box at the bottom of the window (Transfer files but do not activate them (transfer later). Check this check box if you want to activate the new firmware at a later time. Then click **Transfer...** to continue. You will see the window shown in Figure 7-7.

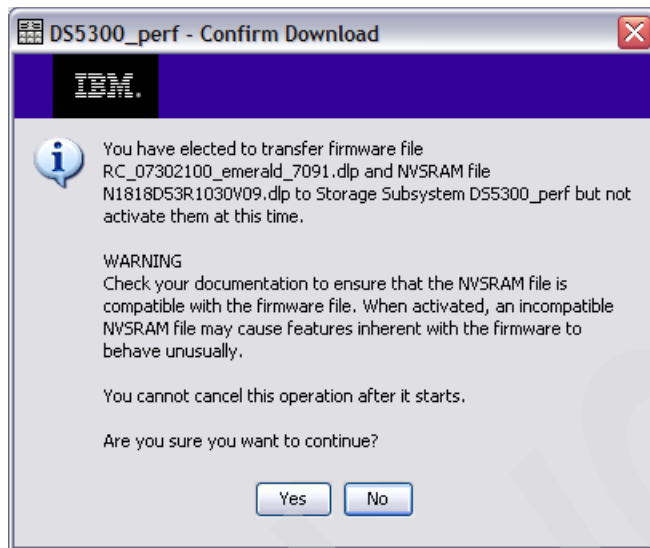


Figure 7-7 Firmware download confirmation

This window instructs you to confirm the firmware and NVSRAM download (because the process cannot be cancelled after 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 7-8. Click **Close**.

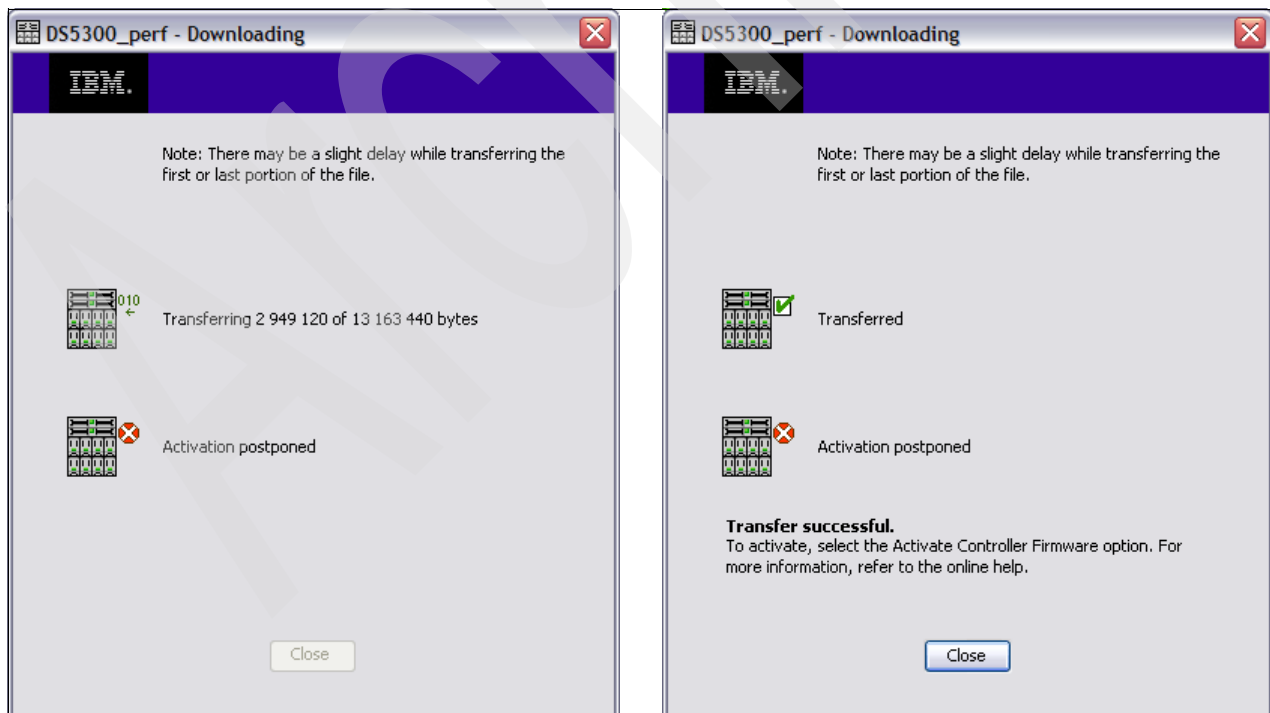


Figure 7-8 Firmware/NVSRAM download finished

3. After clicking **Close**, you are back in the Subsystem Management window.

- If you choose a staged firmware upgrade, the new firmware is now ready for activation. To activate the new firmware, select **Advanced** → **Maintenance** → **Activate Controller Firmware**, as shown in Figure 7-9.

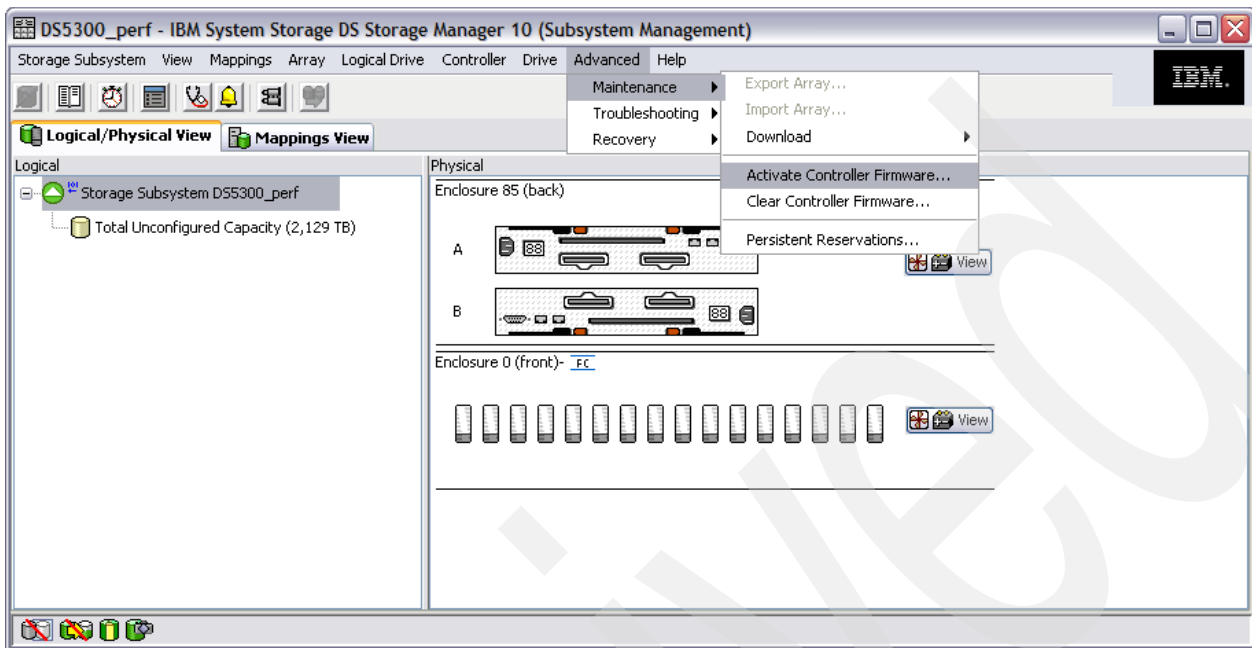


Figure 7-9 Subsystem Management window: Firmware activation

- The Activate Firmware window opens and asks you for confirmation to continue, as shown in Figure 7-10. Click **Yes**.

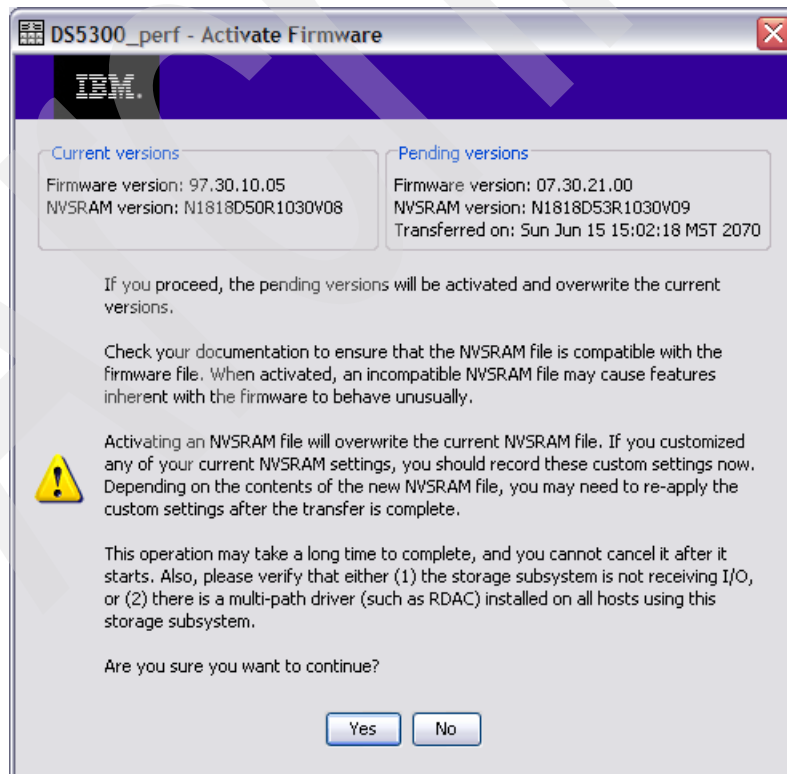


Figure 7-10 Activation firmware confirmation

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 more of a problem than with 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 7-11.

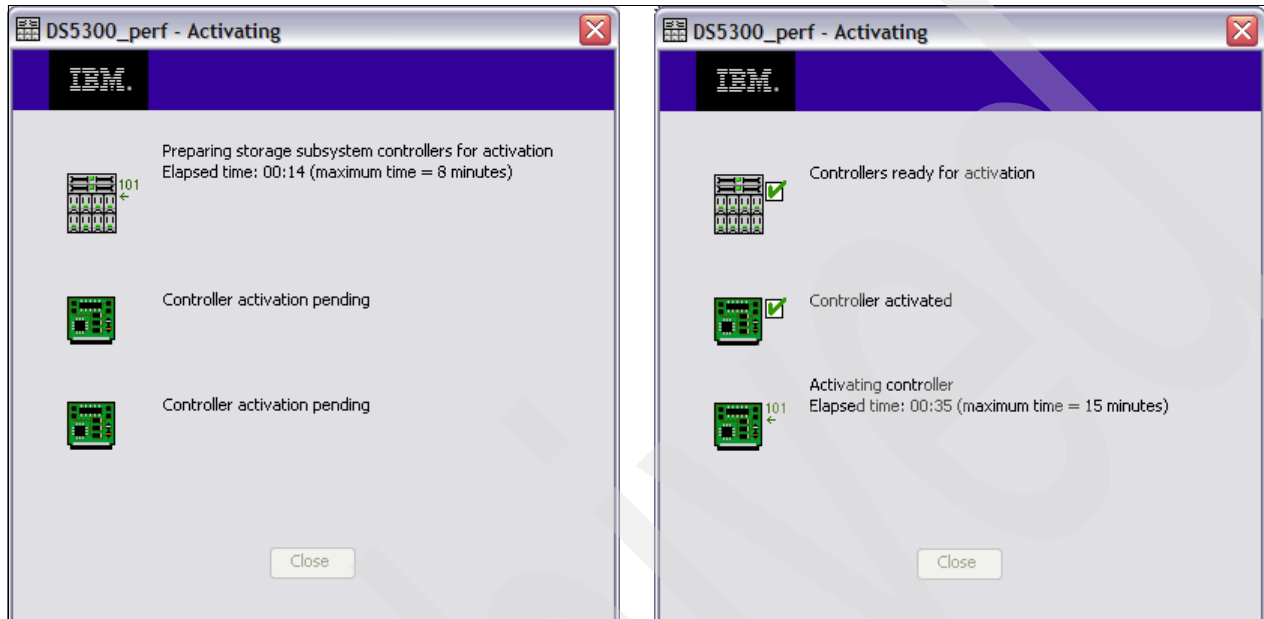


Figure 7-11 Activating the firmware

When the new firmware is activated on both controllers, you see the “Activation successful” message. Click **Close** to return to the Subsystem Management window.

7.1.6 Controller Firmware Upgrade Tool

The following information describes how to use the IBM System Storage DS4000/DS5000 Controller Firmware Upgrade Tool to upgrade your DS4800, DS4700, or DS4200 Express controller firmware from 06.xx to 07.xx.

Attention: Users upgrading from 6.xx.xx.xx controller firmware *must* upgrade to 7.1x.xx.xx or 7.36.xx.xx firmware first before upgrading to 7.50.13.00 or above.

Overview

With Storage Manager Version 10.50, the Controller Firmware Upgrade Tool has become part of the Enterprise Management window (EMW) and is no longer a separate tool. Be aware that:

- ▶ You can perform only a major release to major release (06.xx. to 07.xx) upgrade, using this tool.

Note: Do not attempt to perform this type of firmware upgrade using the Subsystem Management window (SMW). Once you are at the 07.xx firmware level, you do not need to use the Controller Firmware Upgrade Tool. Any future firmware upgrades can be performed by using the SMW.

- ▶ You must perform the upgrade offline.
- ▶ For most failover drivers to take effect, you must reboot the host.

Caution: Before using the IBM System Storage DS4000/DS5000 Controller Firmware Upgrade Tool, it is important that all data be completely backed up and that existing system configurations be saved. Once the tool has completed an upgrade, controllers cannot be returned to previous firmware version levels.

There is a potential loss of data access, so make sure the firmware you download is compatible with the Storage Manager software that is installed on your storage subsystem. If non-compatible firmware is downloaded, you might lose access to the drives in the storage subsystem, so upgrade Storage Manager first.

The Firmware Upgrade Tool will also automatically perform its own diagnostic check on these systems to determine if they are healthy enough to perform a controller firmware upgrade.

These are possible status conditions after the health check:

- ▶ **Optimal:** Every component in the device list is in the desired working condition.
- ▶ **Needs Attention:** A problem exists with the device that requires intervention to correct it.
- ▶ **Fixing:** A Needs Attention condition has been corrected, and the device is currently changing to an Optimal status.
- ▶ **Unresponsive:** The storage management station cannot communicate with the device, or one controller, or both controllers in the storage subsystem.
- ▶ **Contacting Device:** The management software is establishing contact with the device.
- ▶ **Needs Upgrade:** The storage subsystem is running a level of firmware that is no longer supported by the storage management software.
- ▶ **Refreshing:** The status is currently refreshing. This is a transition state.
- ▶ **Not Upgradable:** This subsystem does not fit the requirements to upgrade the firmware. Either it is already upgraded or it is not a DS4200, DS4700, or DS4800 storage subsystem.

Downloading the firmware

Only systems in optimal state can be upgraded. To download the firmware, perform these steps:

1. From the Enterprise Management window (EMW) menu bar, select **Tools** → **Upgrade Controller Firmware**, as shown in Figure 7-12. When the Firmware Upgrade window (Figure 7-13 on page 342) opens, any system listed in the EMW will also appear here.

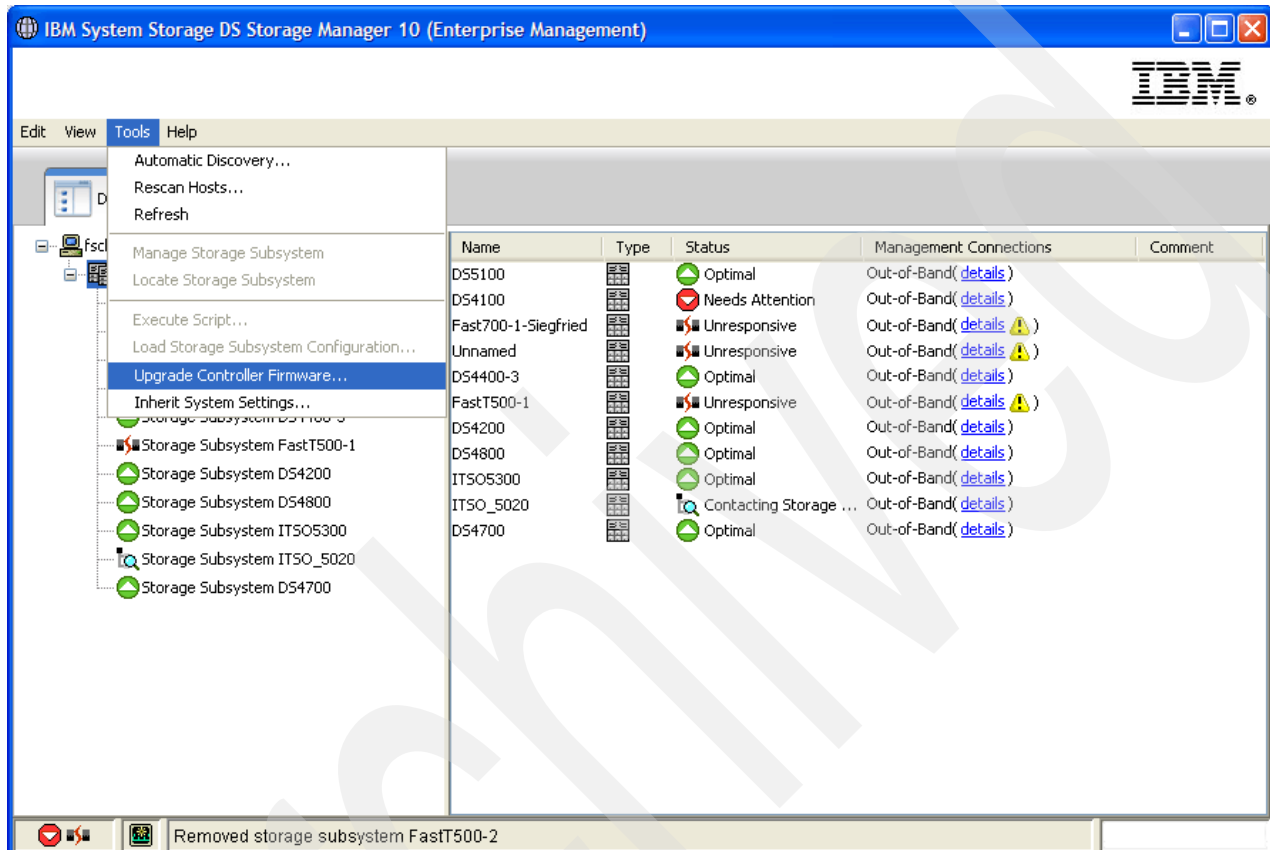


Figure 7-12 Access Controller Firmware Upgrade Tool

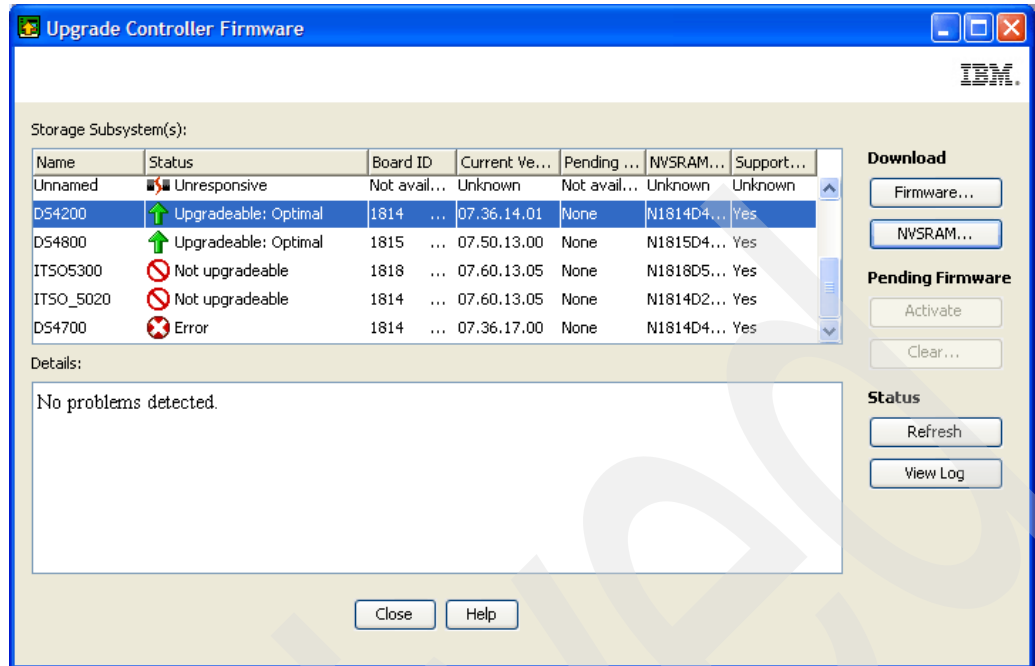


Figure 7-13 Upgrade Controller Firmware Tool

- From the Upgrade Controller Firmware tool, select the subsystem you want to upgrade and click **Firmware**.

Note: NVSRAM and Firmware are usually updated in one step. You do not need to come back and do the NVSRAM update separately.

- Select the controller firmware file that you want to download by clicking **Browse** for the Selected firmware file field (marked as “1” in Figure 7-14 on page 343) to choose the file from a directory on your computer.
- Select the NVSRAM file you want to download by clicking **Browse** for the Selected NVSRAM file field (marked as “2” in Figure 7-14 on page 343) to choose the file from a directory on your computer. Leave the check mark (marked in Figure 7-14 on page 343) checked to update both, NVSRAM and firmware in one step.
- Click **OK**. The firmware starts to download. A status bar appears in the Controller Firmware Upgrade window.

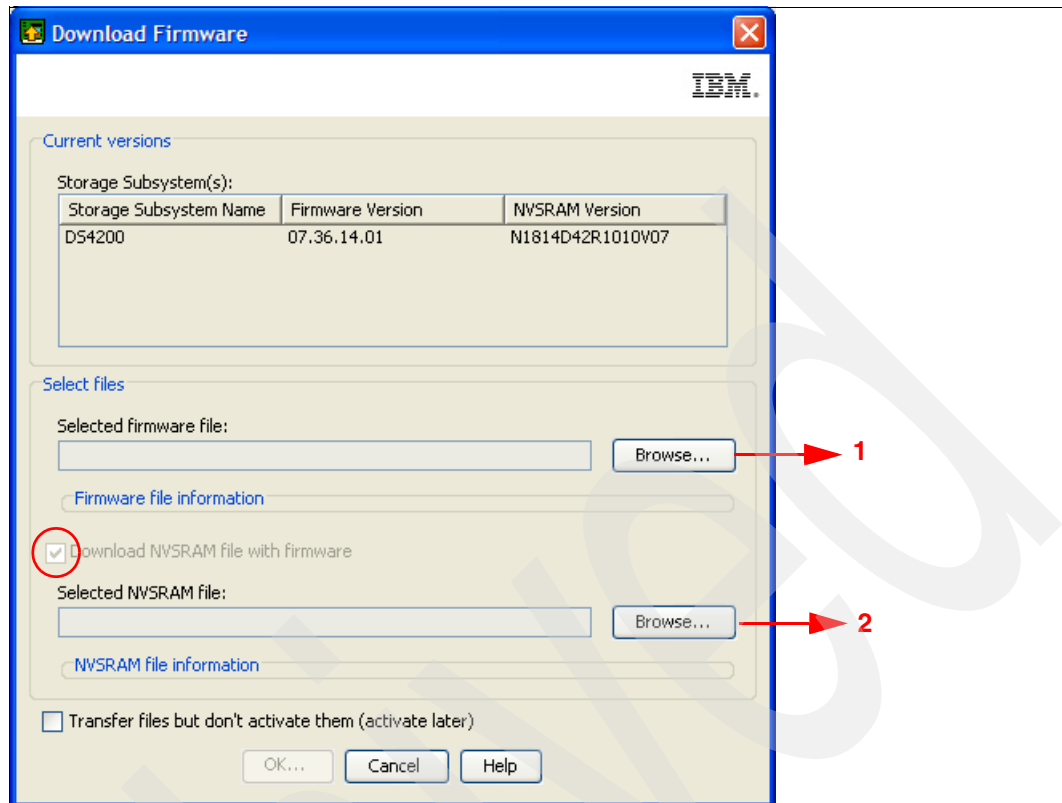


Figure 7-14 Download Firmware window

Do not make changes to your configuration or remove drives or enclosures during the upgrade process.

The update might take up to an hour. Do not interrupt the update process within this hour. Doing so might cause data loss. If you suspect that the update process has stopped, inspect the log file (select **View Log** in the window shown in Figure 7-13 on page 342) or consult IBM Support.

7.1.7 DbFix tool

Customers upgrading from controller code versions 07.36.08.00 or 07.36.12.00 need to take some additional interim steps prior to upgrading the firmware. These releases were affected by a code defect that could lead to recursive reboots when a controller is reset during the code upgrade. Therefore, a check utility was developed to ensure that the DS4000/DS5000 storage subsystem is not prone to these symptoms before proceeding.

The number of units running on these code versions is believed to be low. Nonetheless, customers that discover that if they have either 07.36.08.00 or 07.36.12.00 controller firmware still installed on the DS4000/DS5000 storage subsystem that they should:

- ▶ Avoid making any configuration changes
- ▶ Save a Support Data file from Storage Manager by selecting **Advanced** → **Troubleshooting** → **Support Data** → **Collect ...**
- ▶ Download the DbFix Utility from the firmware download page at the following address:
<http://www.storage.ibm.com/support>
- ▶ Check the online documentation for the latest recommendations and procedure

- ▶ Proceed with running the DbFix Utility

Installing the DbFix tool

First, extract DBFix_v24.zip into a temporary folder on a Windows host with network access to both controllers on the DS4000/DS5000 storage subsystem. The main utility is normally run on the workstation with Storage Manager already installed and configured to manage the DS4000/DS5000 storage subsystem that needs to be checked. However, there is also a separate stand-alone utility provided for those customers that normally run the Storage Manager client on a non-Windows host.

The DbFix utility can take several minutes to complete on larger configurations. It can be run concurrently, but might cause some performance degradation. Therefore, this should be scheduled to be run during off-peak hours.

The DbFix utility cannot be run through an in-band connection. An Ethernet connection to both controllers is required.

Any firewall between the DbFix utility and the DS4000/DS5000 controllers must allow both telnet and FTP traffic to pass.

Any FTP Server service running on the Windows workstation must be stopped before starting the DbFix procedure.

Running the DbFix tool

Use Windows Explorer to navigate into the DbFix subdirectory and perform the following steps:

1. Start the ftp server by double-clicking the startftpsrvr.bat icon. A command-line interface appears and displays the following message:

```
Using XML configuration file bin/arraycheck.xml...
FtpServer started
```

Leave this window running. It can be minimized, but do not close it until the DbFix utility has completed.

2. If you are using the Windows workstation with Storage Manager client already configured, then launch the DbFix utility by double-clicking the SMdbCk.bat icon; otherwise, select the SMdbCkStandalone.bat icon. The first window looks very similar to the normal Storage Manager Enterprise Management window, as shown in Figure 7-15.

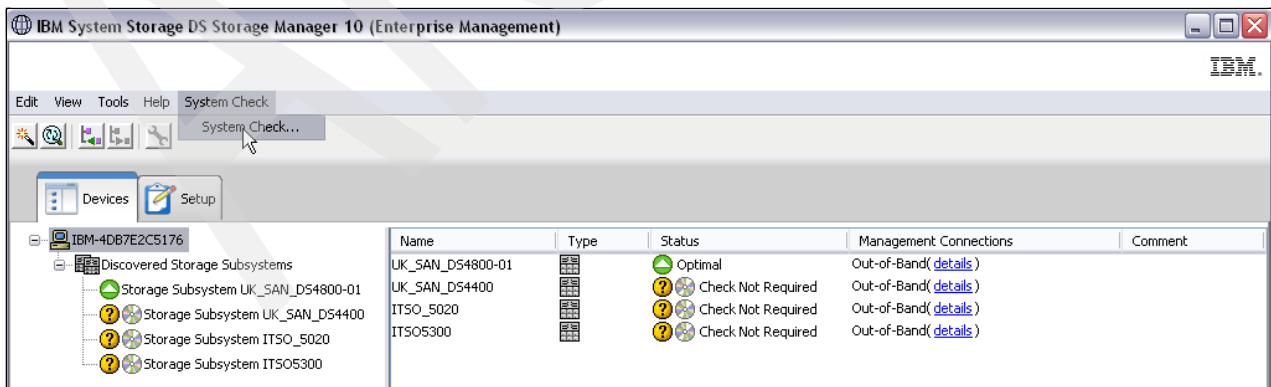


Figure 7-15 DbFix first window

- The utility automatically detects the predefined DS4000/DS5000 storage subsystems. The current status is only displayed for units running on code versions 07.15.xx or 07.36.xx. The status field indicates that a check is not required for code versions outside this range.
- Select **System Check** → **System Check**. A new window opens that shows just the DS4000/DS5000 storage subsystems that need checking, as shown in Figure 7-16.

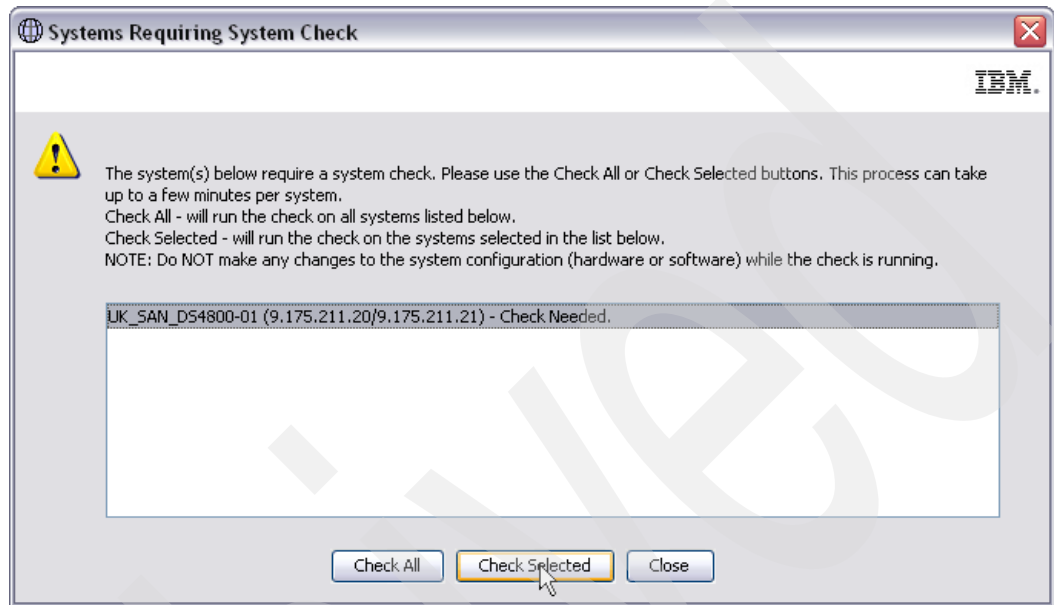


Figure 7-16 DbFix system check selection window

There are two modes of operation:

- You may check all controllers by clicking the **Check All** button.
- You may select individual controllers by selecting them and then clicking **Check Selected**.

- Click **Check All** or **Check Selected**, depending on which mode of operation you are using. A working window is displayed while the file system checks are running. Allow the operation to run to completion without interruption. Figure 7-17 shows the status window confirming that the DbFix utility did not detect any errors. Detailed logs are stored in the `system_check_logs` subdirectory.

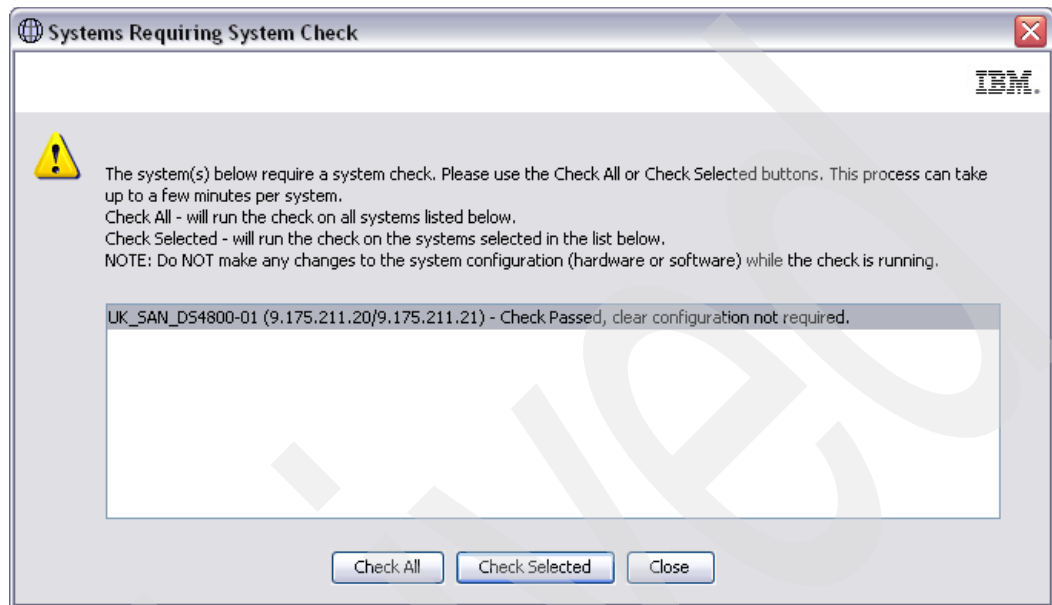


Figure 7-17 DbFix status notification

- If the utility returns a status of Check Passed, then it is safe to proceed with a firmware upgrade to 07.36.17.00. The utility must be rerun after the upgrade.
- If the utility returns a status of Check Failed or any other error, contact your local IBM Support Center for further assistance. Do not proceed with the firmware update unless you advised to do so by an IBM Support Representative.
- Stop the ftp server by closing the command-line window.
- Delete any 07.36.08.00 or 07.36.12.00 code image files from any Storage Manager workstations that might have been used for the code upgrades. These can be identified by the following file names:
 - FW_DS4200_07360800.d1p
 - FW_DS4200_07361200.d1p
 - FW_DS4700_07360800.d1p
 - FW_DS4700_07361200.d1p
 - FW_DS4800_07360800.d1p
 - FW_DS4800_07361200.d1p
 - FW_DS5100_07360800.d1p
 - FW_DS5100_07361200.d1p
 - FW_DS5300_07360800.d1p
 - FW_DS5300_07361200.d1p
- Uninstall the DbFix utility by deleting the temporary folder.

7.1.8 Updating the ESM board firmware

Before performing a firmware upgrade, read the specific firmware version readme file for details about the installation. Pay attention to any dependencies between controller and ESM, or drives, because there might be a specific sequence in which each of the components are updated.

ESM updates are necessary to allow new drive types in the expansion enclosure.

If an ESM is replaced in an EXP5000, EXP520, or EXP810, there is a firmware synchronization feature that ensures that a 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, if there is a failure that prevents the code from synchronizing, identify which ESM is the original and which is the replacement to synchronize.

For EXP5000, EXP520, 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 correct the problem before you proceed with the upgrade.

If you have new expansion enclosures firmware to update, perform the following steps to transfer a firmware file to the ESM boards:

1. From the Subsystem Management window, select **Advanced** → **Maintenance** → **Download** → **ESM Firmware**, as shown in Figure 7-18.

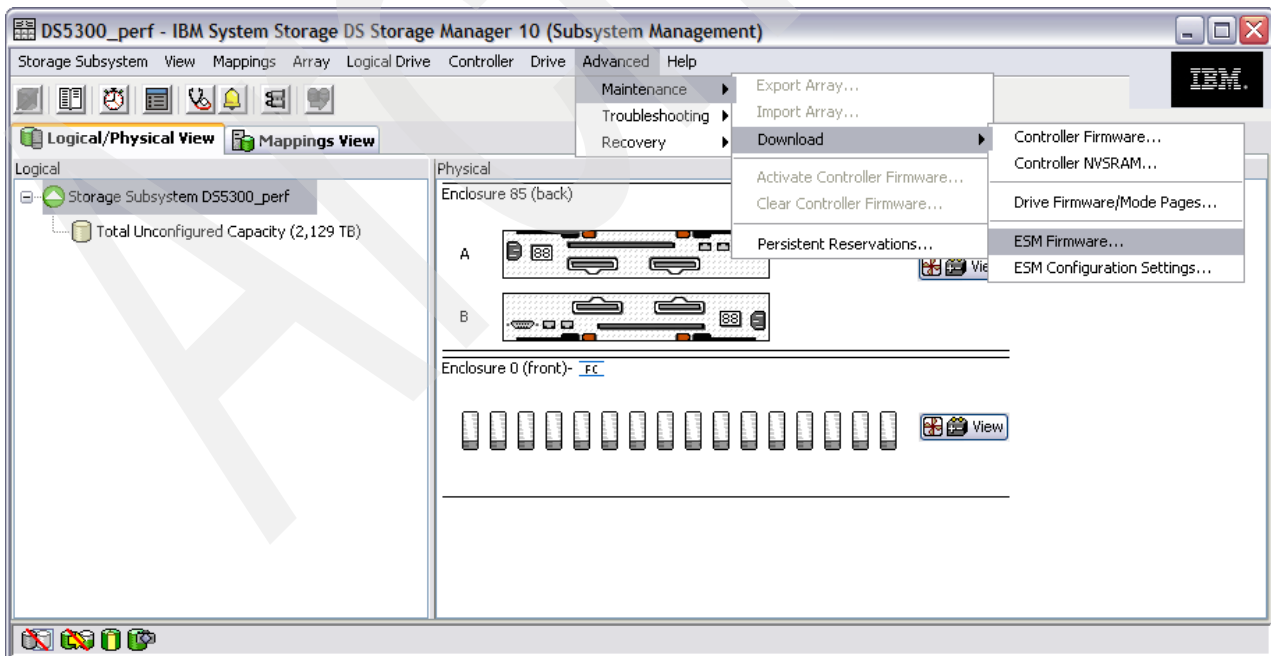


Figure 7-18 ESM firmware upgrade option

2. The Download Environmental (ESM) Card Firmware main window opens, as shown in Figure 7-19.
 - The Select enclosures table lists all the enclosures found attached to the storage array that contain 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 show up in the list (because of a loss of redundancy or some other problem), run the Recovery Guru to diagnose and correct the problem before continuing with the download to avoid losing both paths to the disks.

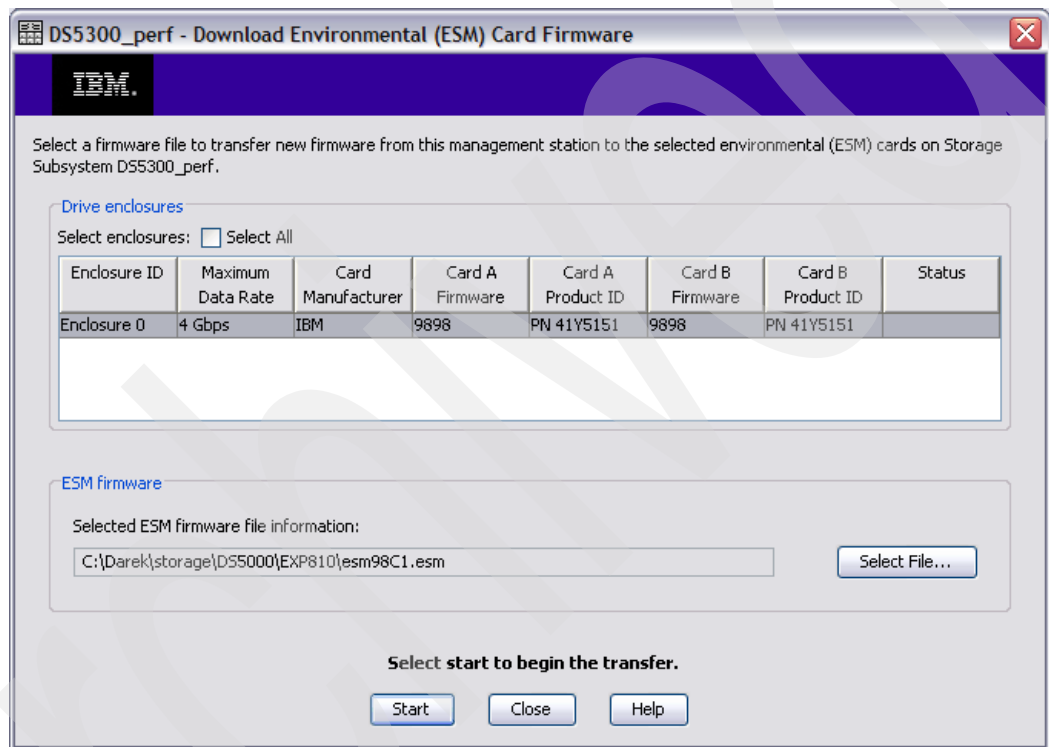


Figure 7-19 Download ESM firmware window

3. In the Select enclosures area, highlight each enclosure to which you want to download firmware or select the **Select All** button to highlight all drive enclosures in the storage subsystem (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 by 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. When the transfer is complete, you see the window shown in Figure 7-20.

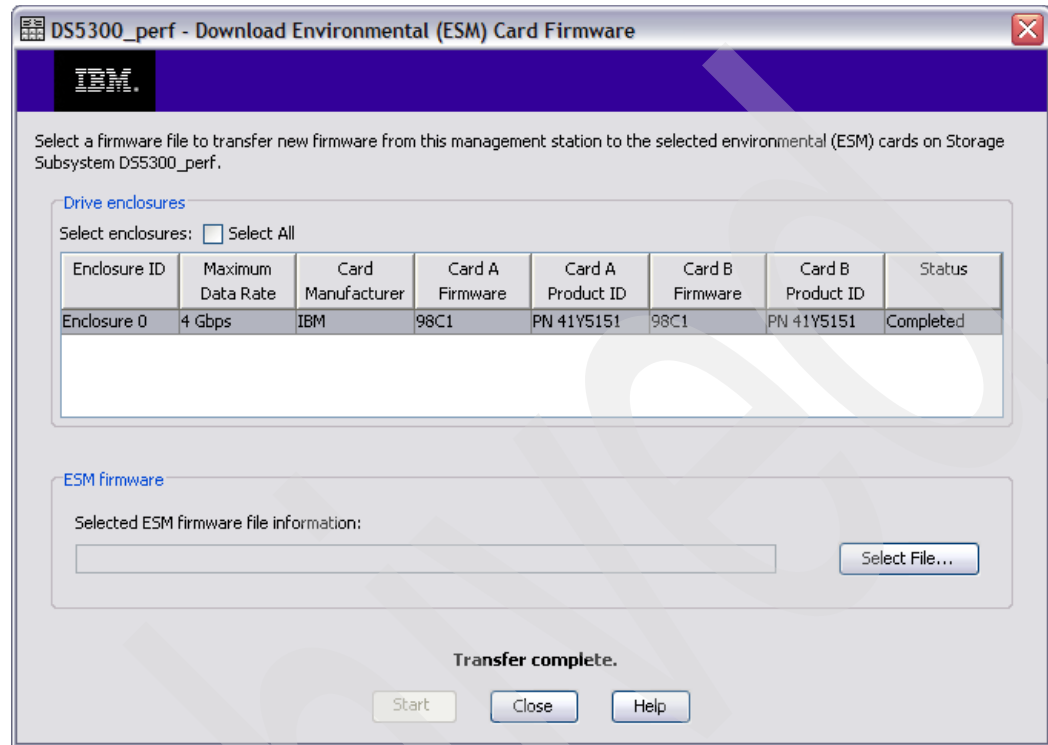


Figure 7-20 ESM firmware update completed

7.1.9 Updating the 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 you to download 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.

You can update up to four different drive types 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 DS5000 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 the following address:

<http://www-1.ibm.com/servers/storage/support/disk>

Perform these steps 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**, as shown in Figure 7-21.

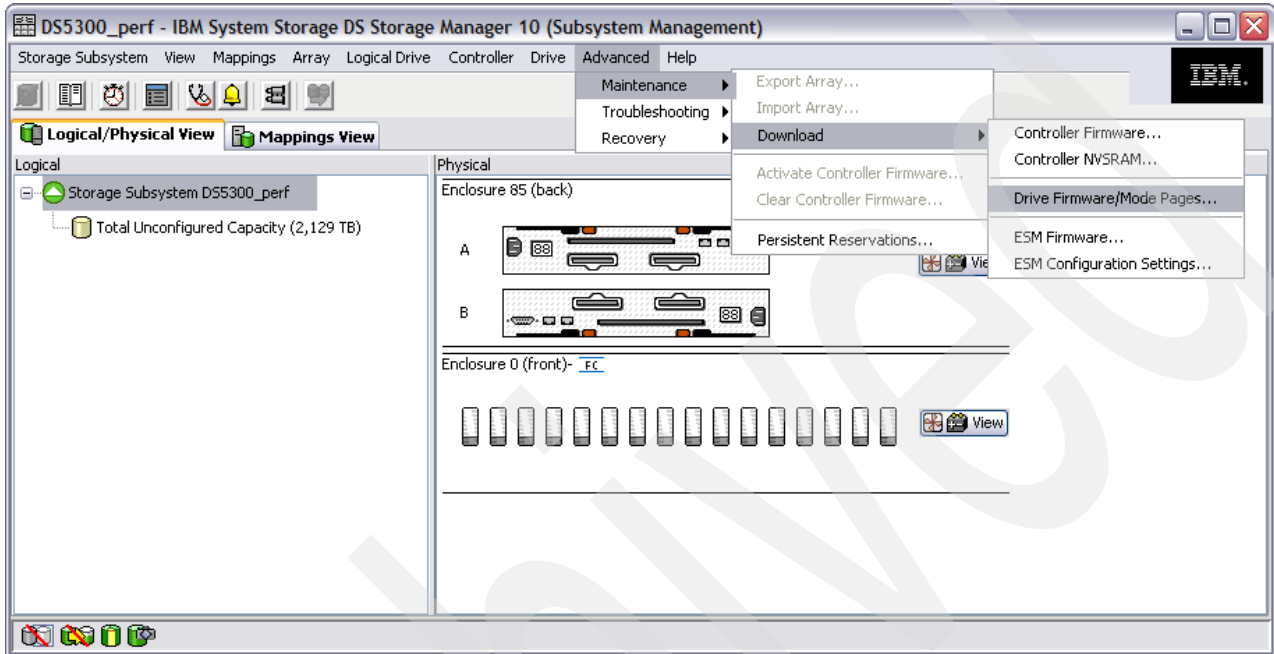


Figure 7-21 Subsystem Management window: Drive firmware update

2. Start the wizard and a window opens, as shown in Figure 7-22. Click **Next**.

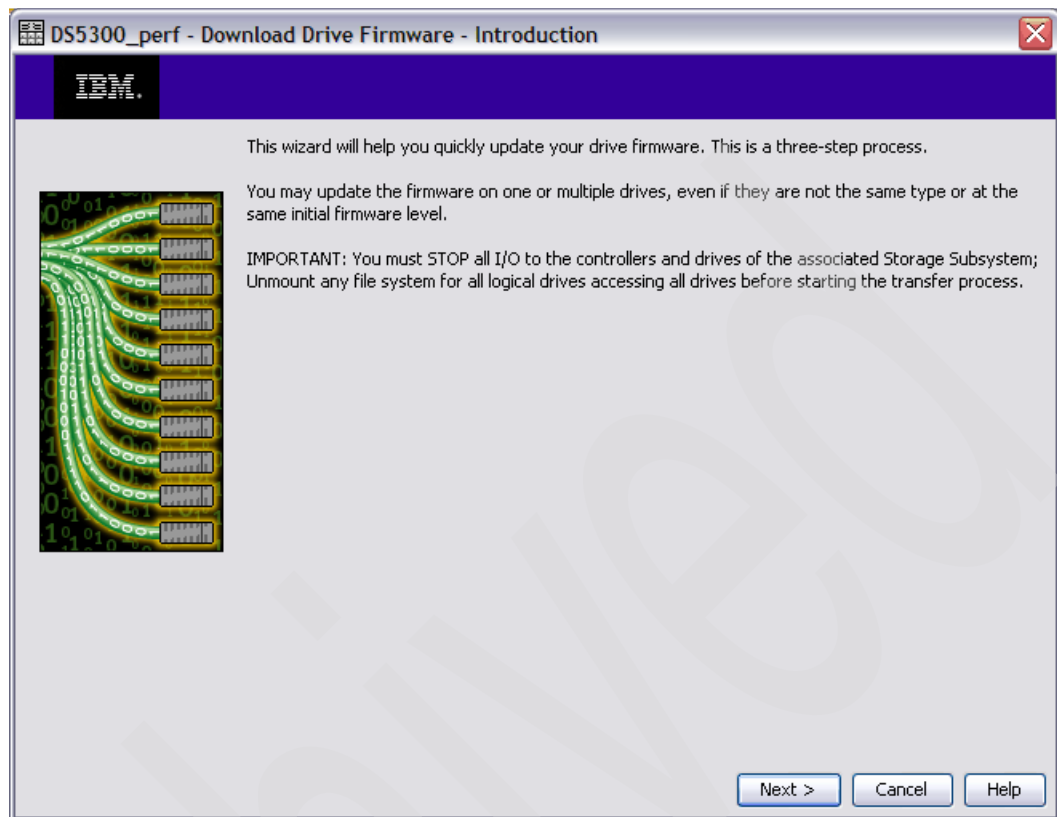


Figure 7-22 Drive upgrade wizard

3. The window shown in Figure 7-23 opens, which shows all the types of drives you have with the current firmware level. Click **Add** to choose the firmware package.

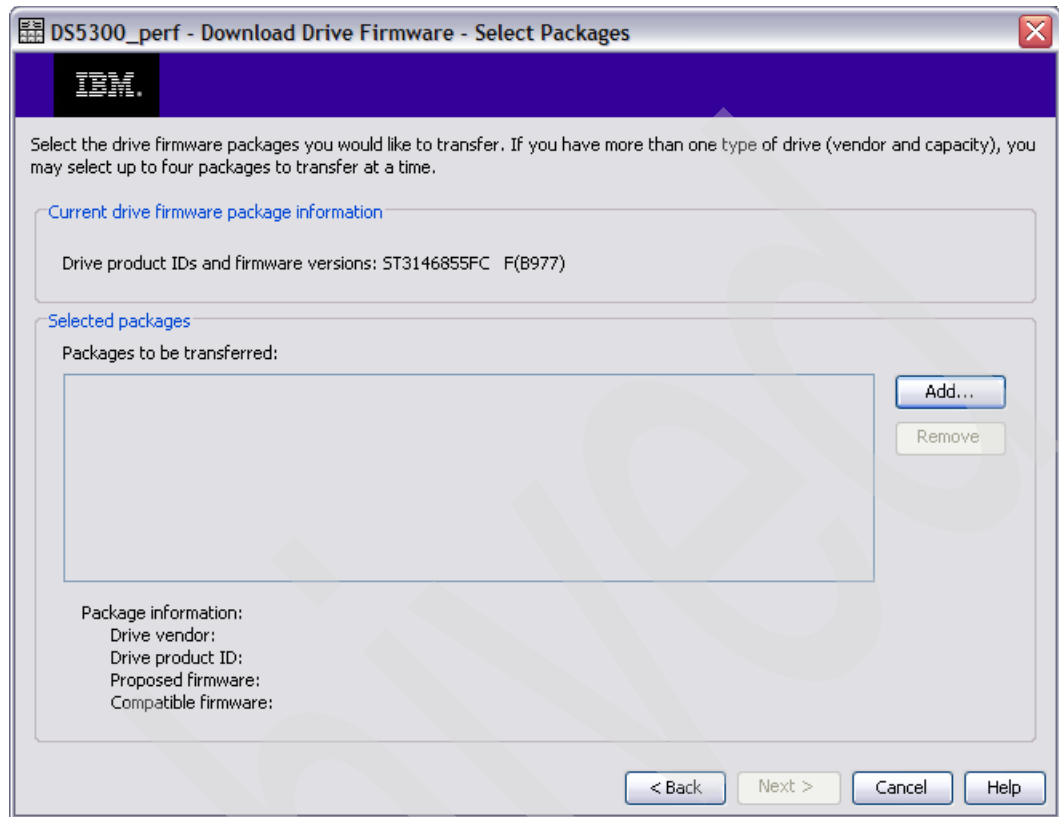


Figure 7-23 Selecting drive firmware packages

4. The window shown in Figure 7-24 opens. Here you can browse and choose drive firmware packages that you previously downloaded. This window shows if a firmware package is compatible in the File Information pane. Figure 7-24 shows compatible firmware and Figure 7-25 on page 354 shows incompatible firmware.

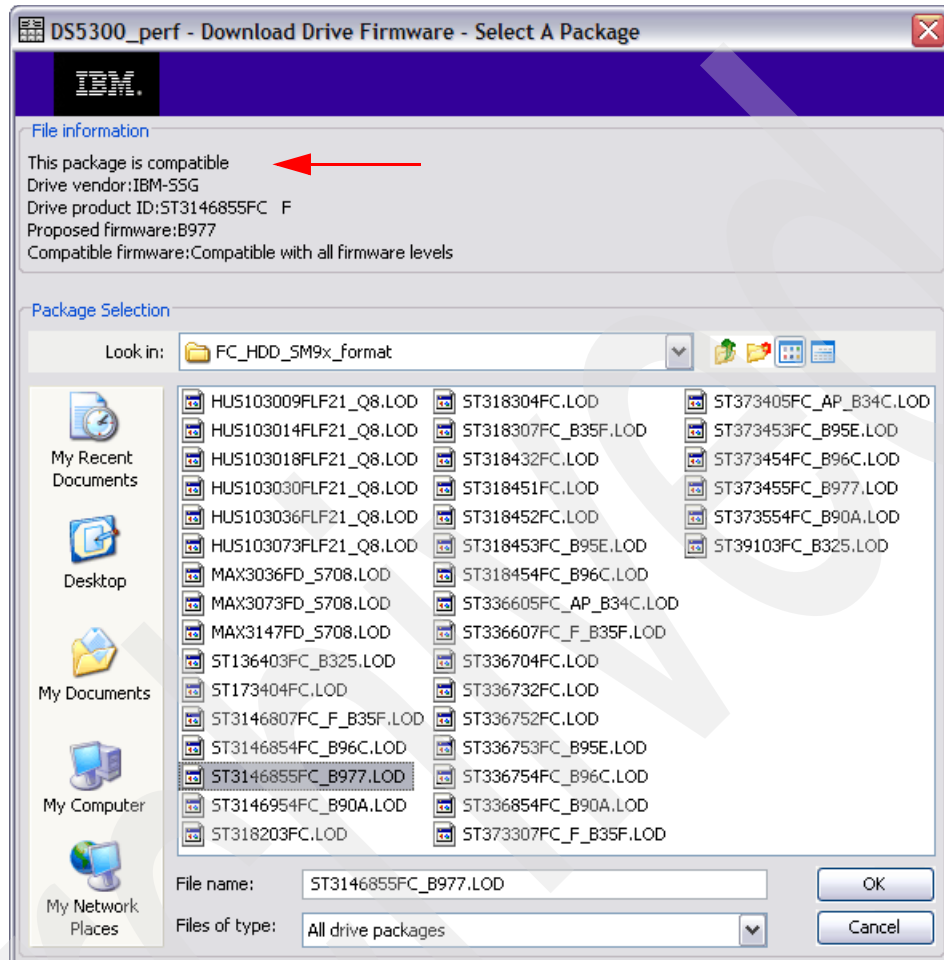


Figure 7-24 Drive firmware: Compatible

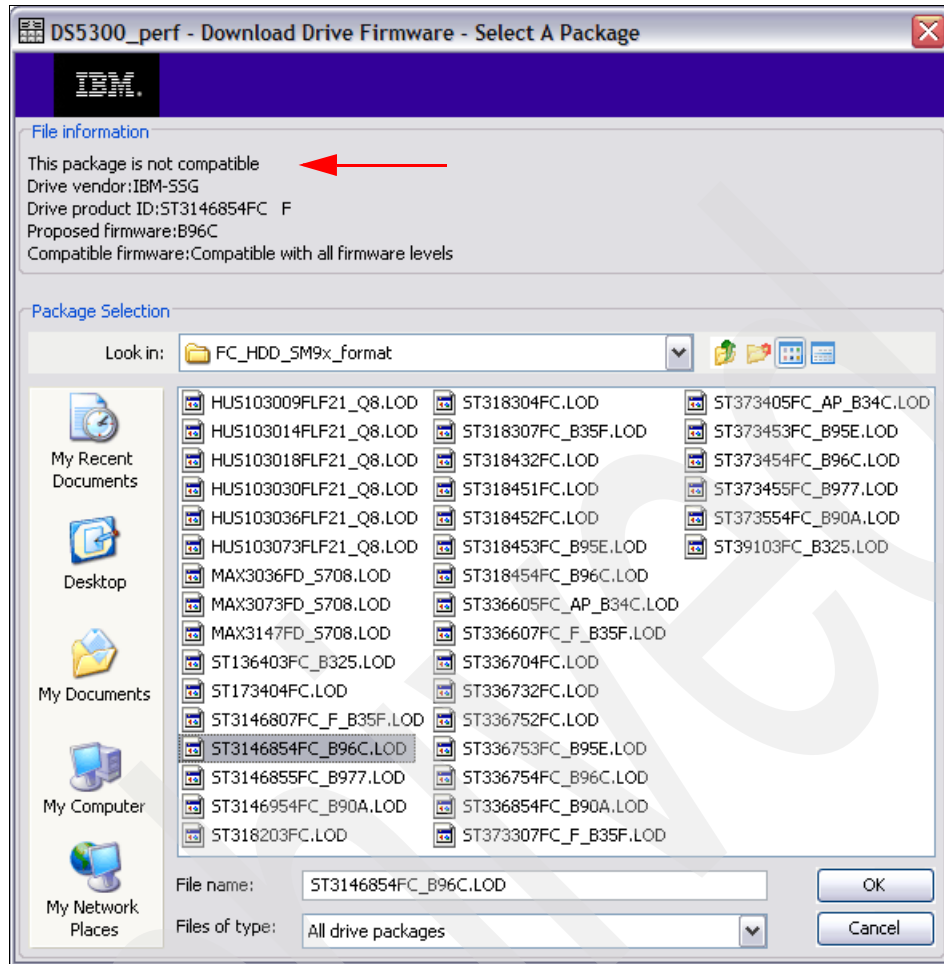


Figure 7-25 Drive firmware: Incompatible

5. You can select more drive firmware (see Figure 7-26) if you have different drives in enclosures.

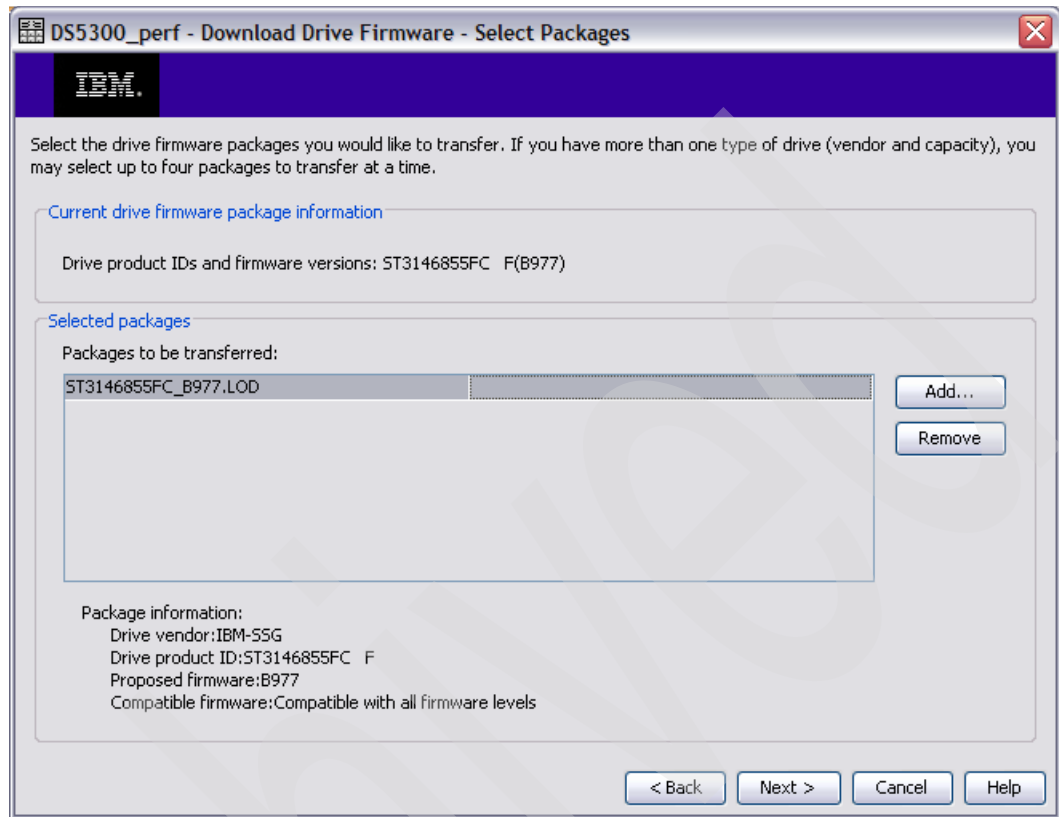


Figure 7-26 Adding drive firmware packages

- If you click **Next**, a new window opens showing all the drives with the current and proposed firmware. We had one type of drive that did not require an upgrade, because the firmware was on the same level, as shown in Figure 7-27. Select the drives you want to upgrade and click **Finish**.

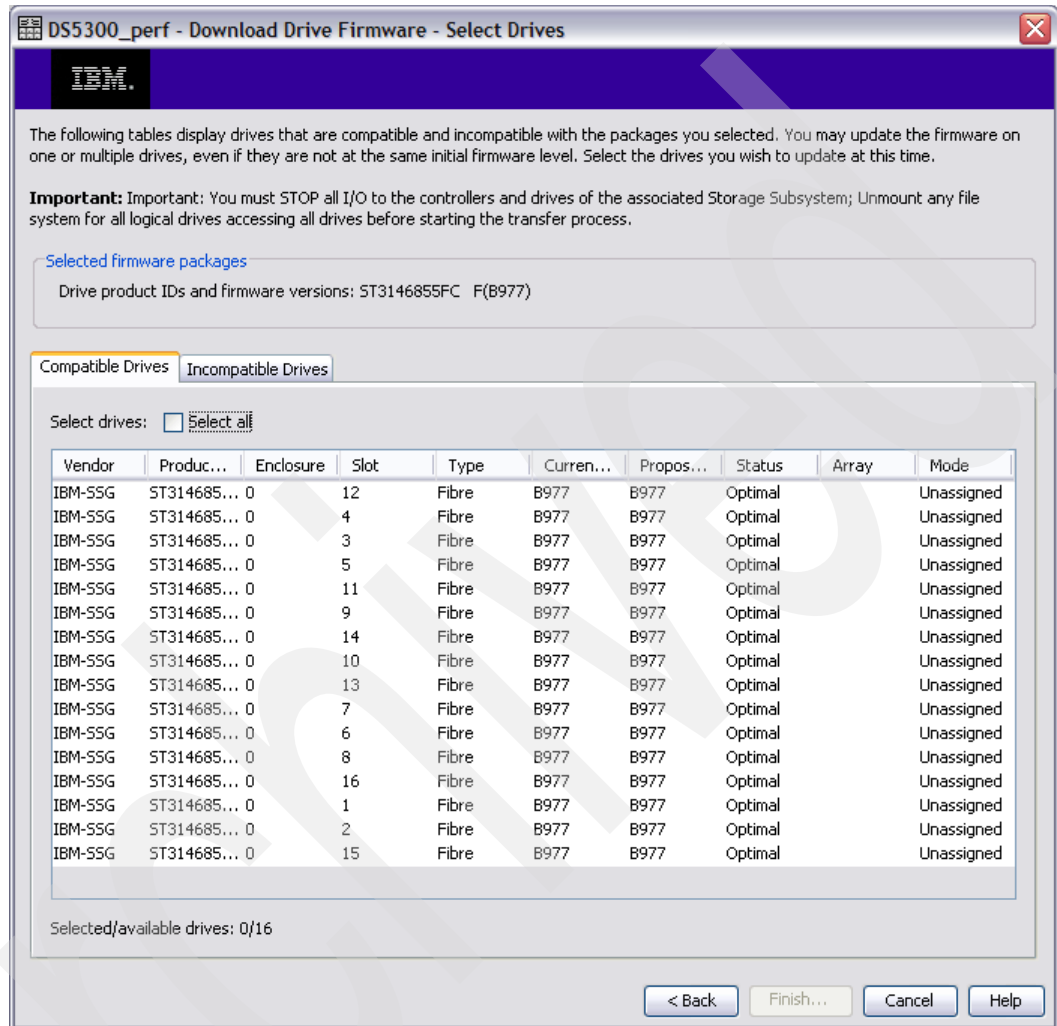


Figure 7-27 Selecting drives to upgrade

- In the window that appears, browse to the directory where you downloaded the drive firmware files. Select a package based on your disk types. You can use the top section while browsing through the different files. It shows whether each of them is its compatible with the disks that you have in your enclosures. Select a compatible file and click **OK**.

8. Once the drive firmware download procedure starts, you can monitor the progress, as shown in Figure 7-28. 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.

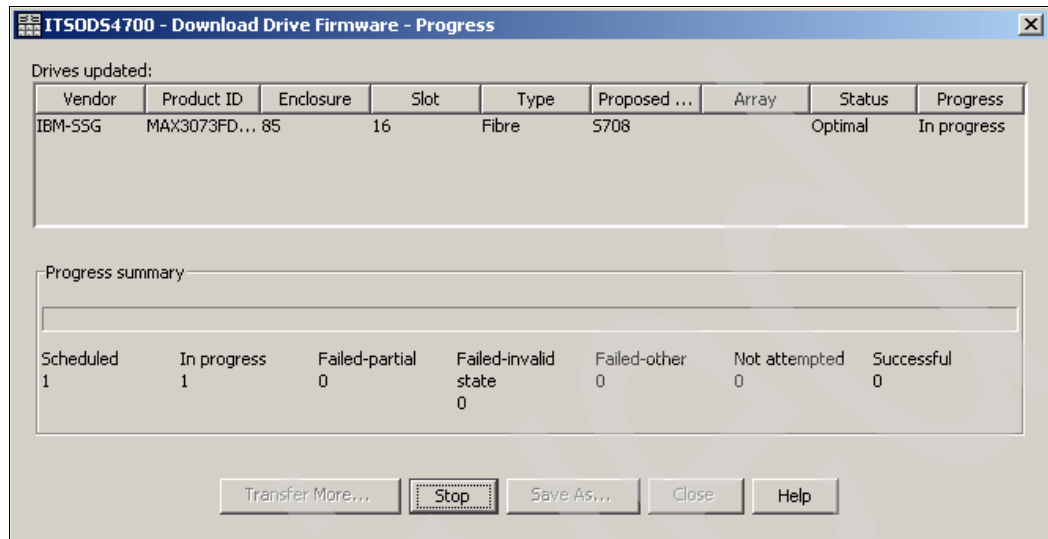


Figure 7-28 Drive firmware update: Download progress window

7.1.10 Updating host bus adapter (HBA) firmware

This section describes the procedure to update the HBA firmware in Windows, Linux (graphical based), and AIX environments. However, different HBA vendors offer different tools to manage the HBAs.

We describe the following procedures in this section:

- ▶ “Update Brocade HBA firmware using Brocade Host Connectivity Manager (HCM)” on page 358
For more details about Brocade HCM, see 7.11.1, “Brocade HBA and Brocade Host Configuration Manager (HCM)” on page 436
- ▶ “Update Emulex HBA firmware using Emulex HBAnyware” on page 359
For more details about Emulex HBAnyware, see 7.11.2, “Emulex HBA tools” on page 441
- ▶ “Update the HBA firmware using QLogic SANsurfer” on page 360
For more details about QLogic SANsurfer, see 7.11.3, “Qlogic HBAs and SANsurfer (Windows/Linux)” on page 443
- ▶ “Updating HBAs in AIX environments” on page 362

Download the HBA firmware

You can find the latest HBA firmware version at the IBM System Storage Support site at the following address:

<http://www-1.ibm.com/servers/storage/support/disk>

Select **System Storage**, then select **Disk systems** from the drop-down menu, and choose your IBM DS5000 storage subsystem model. Select the **Download** tab to access the supported firmware for your host bus adapter. You will find the appropriate files for your HBA depending on the vendor (either Brocade, Emulex, or Qlogic) in this table.

Note: When updating HBAs, make sure to install matching versions of BIOS and drivers for your adapters. Also make sure to install at least the minimum versions recommended on the IBM Storage support Web site. Do *not* download the firmware from the vendor Web site itself, as this is not IBM tested code.

Update Brocade HBA firmware using Brocade Host Connectivity Manager (HCM)

Brocade calls its HBA firmware *boot code* or *EFI*. There are different ways to update the boot code of the HBA. These are:

- ▶ Using a bootable live CD
- ▶ Using the HCM GUI (which is explained here)
- ▶ Using the BCU command line tools.

The Brocade HCM offers an easy way to update the HBA's firmware.

You can obtain the latest version of the firmware at the IBM storage support site at the following address:

<http://www-1.ibm.com/servers/storage/support/disk>

To update the firmware, perform the following steps:

1. Start HCM and log in to the agent. The HCM window opens, as shown in Figure 7-29.

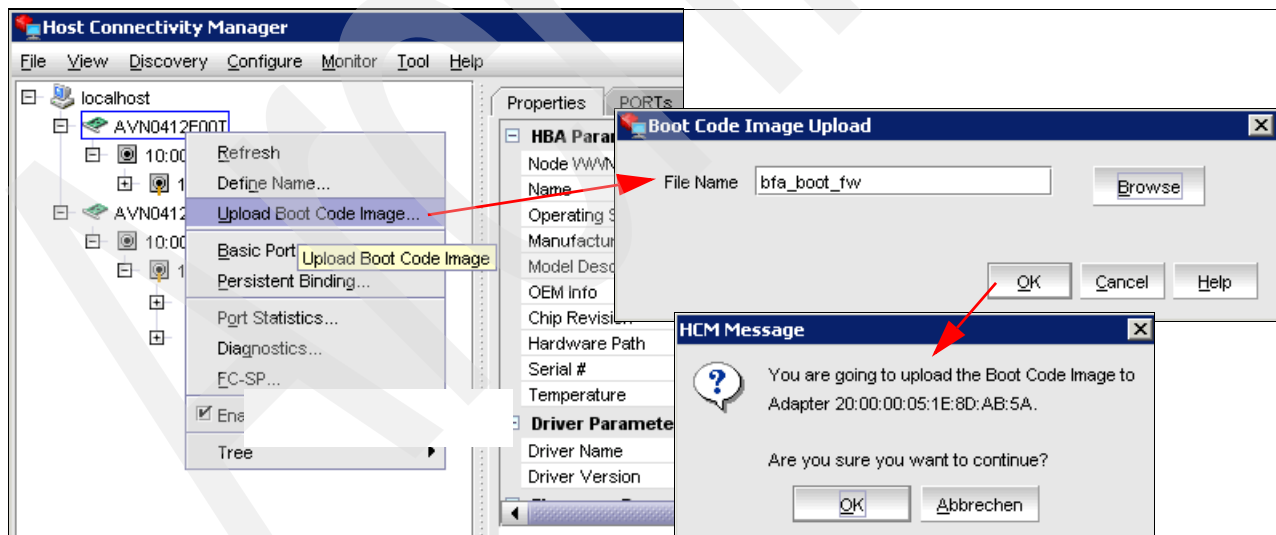


Figure 7-29 HCM update firmware

2. Right-click the HBA entry in the left pane and select **Upload Boot Code Image...** from the menu.
3. In the resulting Boot Code Image Upload window, click **Browse** to select the boot code image on the computer and then click **OK**.
4. Confirm your choice by clicking **OK** in the next window.

5. After a few seconds, a confirmation window appears.
6. Reboot the server in order to activate the new HBA BIOS.

Update Emulex HBA firmware using Emulex HBAnyware

You can update firmware or boot code with either HBAnyware or lputilnt.

- ▶ HBAnyware allows you to update firmware or boot code on remote and local HBAs.
- ▶ lputilnt allows you to update firmware or boot code on local HBAs only.

Prerequisites

You must meet the following prerequisites

- ▶ The SCSIport Miniport driver is properly installed.
- ▶ The HBAnyware utility is properly installed.
- ▶ The firmware or boot code file has been downloaded from the Emulex Web site and extracted to a directory on your local drive.

You can find the latest version at the IBM storage support site at the following address:

<http://www-1.ibm.com/servers/storage/support/disk>

Procedure

To update firmware or boot code using HBAnyware, perform the following steps:

1. Start the HBAnyware utility.
2. In the discovery-tree (left pane), click the HBA entry you want to update. In the menu bar, click **Batch** and select **Download Firmware** (see Figure 7-30).
3. A Select Firmware File window appears (Figure 7-30). Select the firmware you want to update and click **Open**.

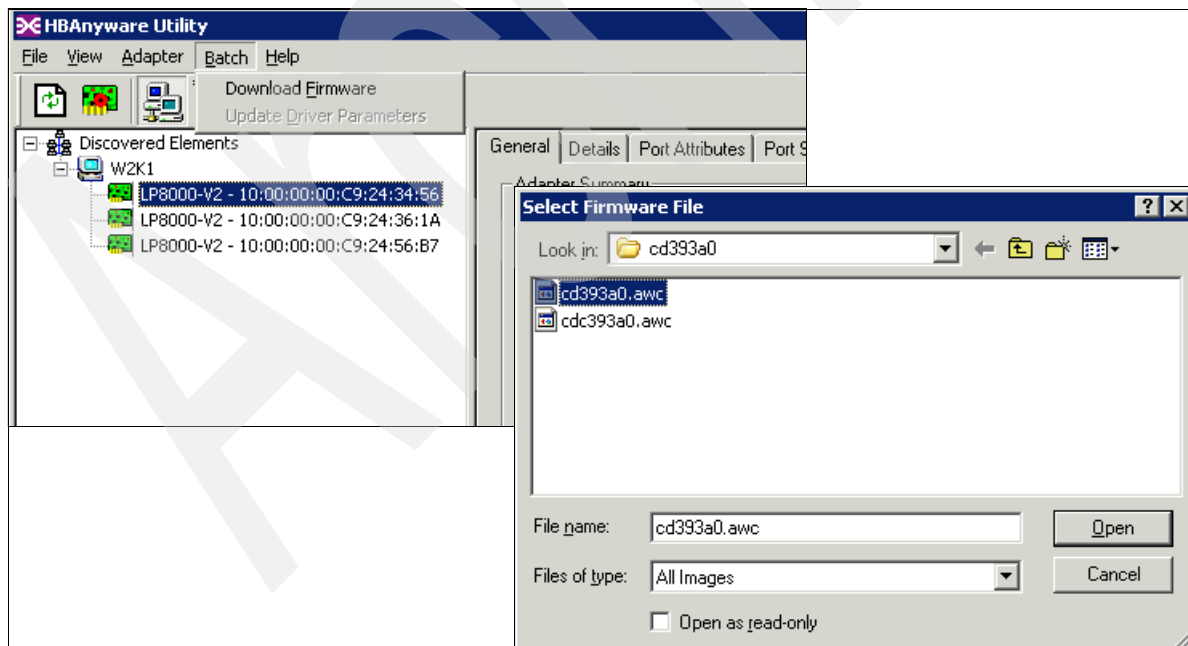


Figure 7-30 HBAnyware: Download firmware

4. The Firmware Download dialog box appears. Click **Yes**. Next, the adapter reset notice appears. Click **Yes**.

5. After a few seconds, the Firmware Upgrade Successful message appears. Click **OK**. The firmware update is done.
6. The server need to be rebooted in order to activate the new HBA firmware.

Update the HBA firmware using QLogic SANsurfer

The Qlogic SANsurfer is available for Windows and Linux hosts.

You can find the latest version at the IBM System Storage Support site at the following address:

<http://www-1.ibm.com/servers/storage/support/disk>

To update the HBA firmware using QLogic SANsurfer, perform these steps:

1. In order to update the BIOS version of the HBA, click the **Utilities** tab, and you will be presented with a window shown in Figure 7-31.

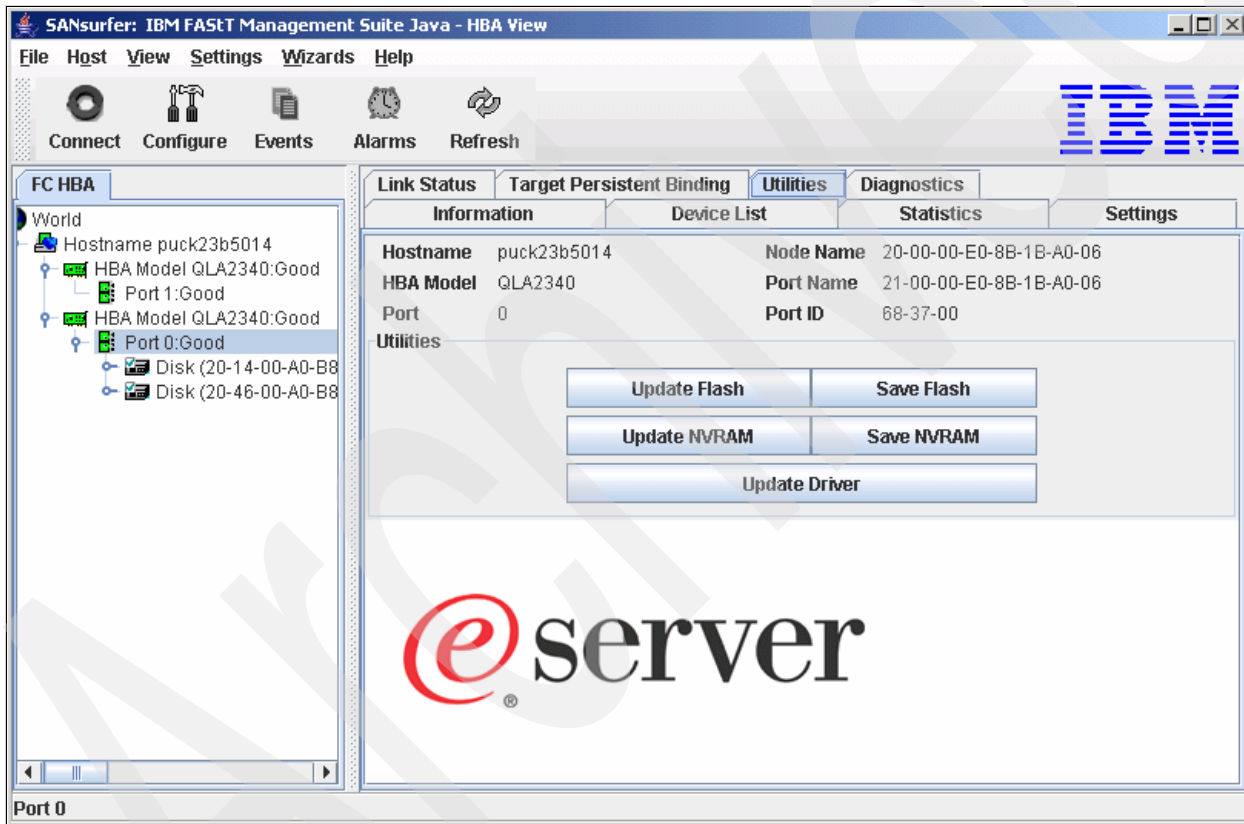


Figure 7-31 Qlogic SANsurfer Utilities tab

2. Notice that there are two options to update:

- Flash Update
- NVRAM

Click **Update Flash**, which corresponds to the BIOS version.

Note: The NVRAM usually does not need to be updated. It is used to save and recover HBA unique settings for this particular firmware and will be used in case your adapter NVRAM gets corrupt. For this function to work, NVRAM have been saved previously.

3. You are then presented with a warning, as shown in Figure 7-32. Make sure that you have suspended all I/O to the adapter before you proceed.

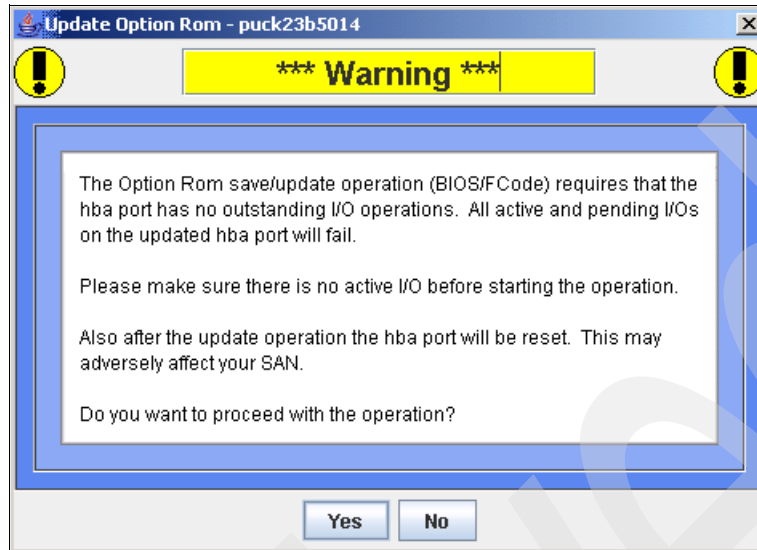


Figure 7-32 Qlogic SANsurfer BIOS update

4. Acknowledge the warning and click **Yes**. You must use the BIOS update file.
5. Select the corresponding file for your adapter. In Figure 7-33, we show the download for a Qlogic 234x adapter.

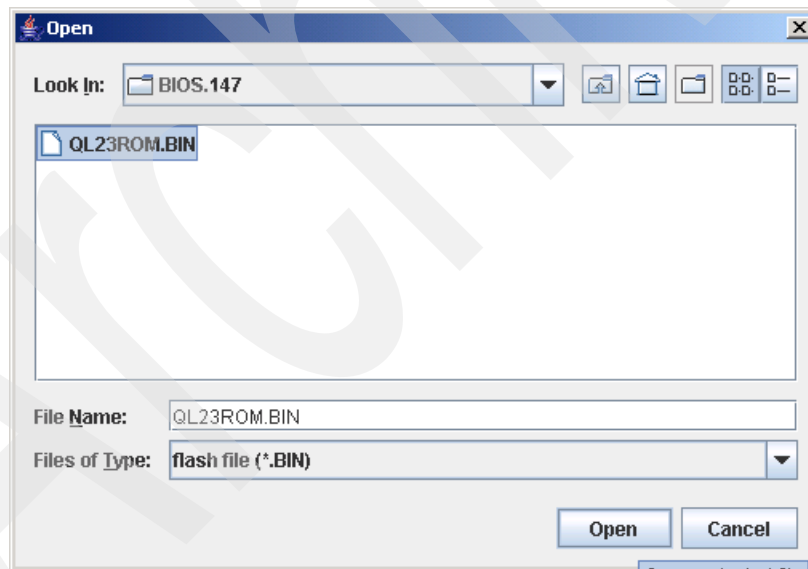


Figure 7-33 Qlogic SANsurfer HBA FC2-133 BIOS update file

Once you have selected the file, you will be prompted for the password. The default password for all adapters is *config*. Once the upload is finished, you will be prompted to reboot the server so the changes can take effect. You can upload BIOS for multiple adapters subsequently before rebooting the server. Remember, if you do not reboot the server, the changes will not take effect.

Updating HBAs in AIX environments

To update HBAs in AIX, perform these steps:

1. For IBM System p microcode, see the IBM support Web site at the following address:
<http://www-304.ibm.com/jct01004c/systems/support/>
2. In the Download microcode by device type section, select **Adapters**.
3. Search for your server-specific FC adapter type. Once found, the description provides detailed instructions about how to check the current version, install the update, and check the results.

Select the RPM file and click **Continue** at the bottom of the page to download the rpm file.

4. Follow the instructions to download the file and transfer it to `/etc/microcode`.
5. Unpack the transferred file in `/etc/microcode` by running the following command:

```
#cd /etc/microcode
#rpm -ihv --ignoreos pci.df1000f9-3-93a0.aix.noarch.rpm
```

Replace `pci.df1000f9-3-93a0.aix.noarch.rpm` with the name of the package downloaded for your particular HBA.

6. Flash the adapter microcode by running the following command:

```
#diag -d fcsX -T download
```

Where `X` is the number returned by the `lsdev -C | grep fcs` command. Self-explanatory menus take you through the microcode installation. Repeat this process for all the HBAs that need to be updated.

7. Verify the adapter microcode level by running the following command:

```
#lsmcode -d fcsX
```

Compare the obtained level with the readme file of the downloaded code.

7.2 Handling premium features

Depending on your specific IBM Midrange System Storage storage subsystem model, you might want to use a specific function or capability of your storage subsystem. Some of these capabilities, although included in the firmware, are not enabled by default. These are called premium features, and you have to 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.

Here we cover how to order, list, and install any of the premium features available for your IBM Midrange System Storage storage subsystem.

The available premium features are:

- ▶ Drive security
- ▶ Storage partitioning
- ▶ FlashCopy
- ▶ Enhanced Remote Mirroring (ERM)
- ▶ VolumeCopy
- ▶ FC/SATA Intermix

7.2.1 Listing premium features/feature enabler

From the Storage Manager client GUI, you can check what premium features are already activated in your DS storage subsystem.

Starting with Storage Manager V10 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** in Storage Manager V10.10 and higher.

To list the premium features, perform the following steps:

1. Select **Storage Subsystem** → **Premium Features** from the Subsystem Management window. The Premium Features window opens and shows a list of enabled premium features (see Figure 7-34).

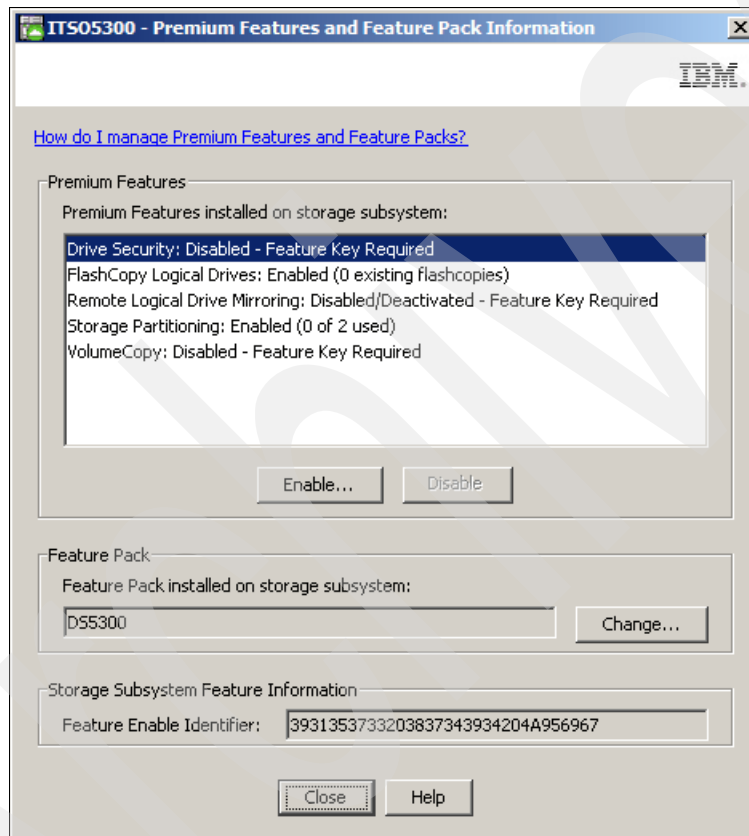


Figure 7-34 Listing premium features

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.

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 illustrated in Example 7-1.

Example 7-1 Collect support data featureBundle.txt file

```
FEATURE BUNDLE FOR STORAGE SUBSYSTEM:      ITS05300 (Mon Sep 07 18:46:44 CAT
2009)

Total logical drives allowed:                2048

Total logical drives allowed per partition:  256

Management application client:              1

Drives supported:                           B; J; S; E; BM

Drive enclosures supported:                  IBM,EXP5000,HUSKER;
IBM,EXP810,HUSKER; IBM,EXP5060,WEMBLEY-HUSKER

High Performance Tier:                       No data is available about
supported values of this feature

Storage Partitioning
  Enabled by default:                        Yes
  Default limit:                             2
  Enable/Upgrade via key:                   Yes
  Limit with feature key:                   512

FlashCopy Logical Drives
  Enabled by default:                        Yes
  Enable/Upgrade via key:                   Yes
  Total flashcopies allowed:                2

VolumeCopy
  Enabled by default:                        No
  Enable/Upgrade via key:                   Yes
  Total copies allowed:                     2048

Remote Logical Drive Mirroring
  Enabled by default:                        No
  Enable/Upgrade via key:                   Yes
  Total mirrors allowed:                    0

Mixed Drive Types
  Enabled by default:                        Yes
  Enable/Upgrade via key:                   No
```

3. Gather the following data along with the Feature Enable Identifier:
 - Machine type
 - Model
 - Serial number

Note: The machine type, model, and serial information is printed on a label on the back of your DS storage subsystem controller unit.

7.2.2 Enabling a premium feature

Obtaining a feature key depends upon the DS storage subsystem packaging procedures and time of order:

- ▶ If you bought any premium features together with the DS storage subsystem, you will receive the feature keys with the DS storage subsystem 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 DS storage subsystem, and you have to request the premium 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 about how to proceed to activate it.

Important: All the premium features are activated immediately and do not require a reboot of the controllers.

The following procedure will help you activate the new feature, or reactivate it if for any reason it is out of compliance:

1. On the card 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 **Storage Subsystem** → **Premium Features**.
3. Write down the 32-digit number next to the feature enable identifier, or 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 that was generated when you collected all the support data.

4. Go to the following Web site to personalize the received key to your specific system and generate the activation key file:

<http://www-912.ibm.com/PremiumFeatures/>

Select the option to activate a premium feature.

5. Complete the form in the Web page with the following information:
 - 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.

- Once you have received the key file, go to Storage Manager and select **Storage Subsystem** → **Premium Features** in the Subsystem Management window, as shown in Figure 7-35.

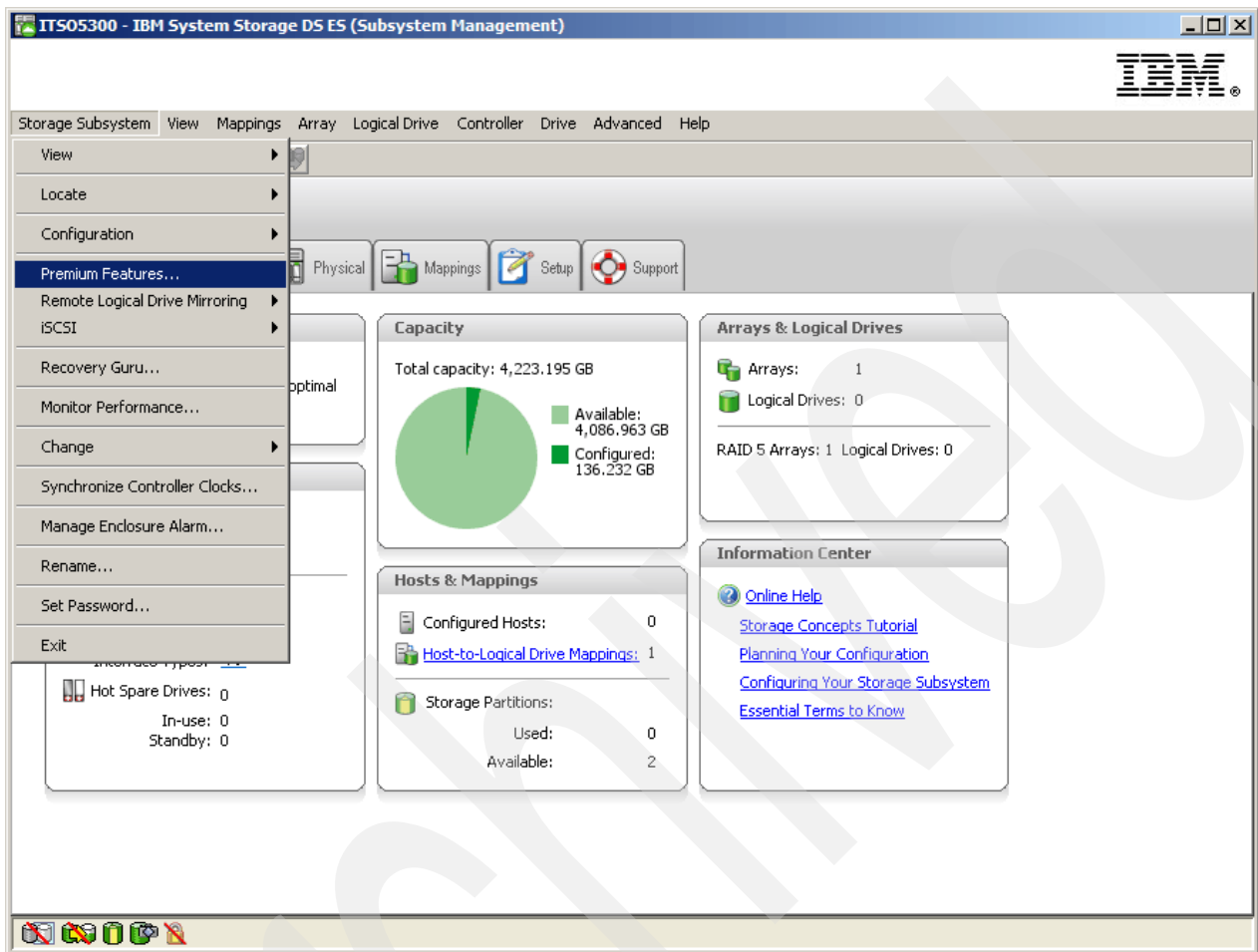


Figure 7-35 Premium Features menu

In the window, click **Enable** (as shown in Figure 7-36) and point to the location where the key file is stored. You need to confirm whether to enable the selected premium feature.

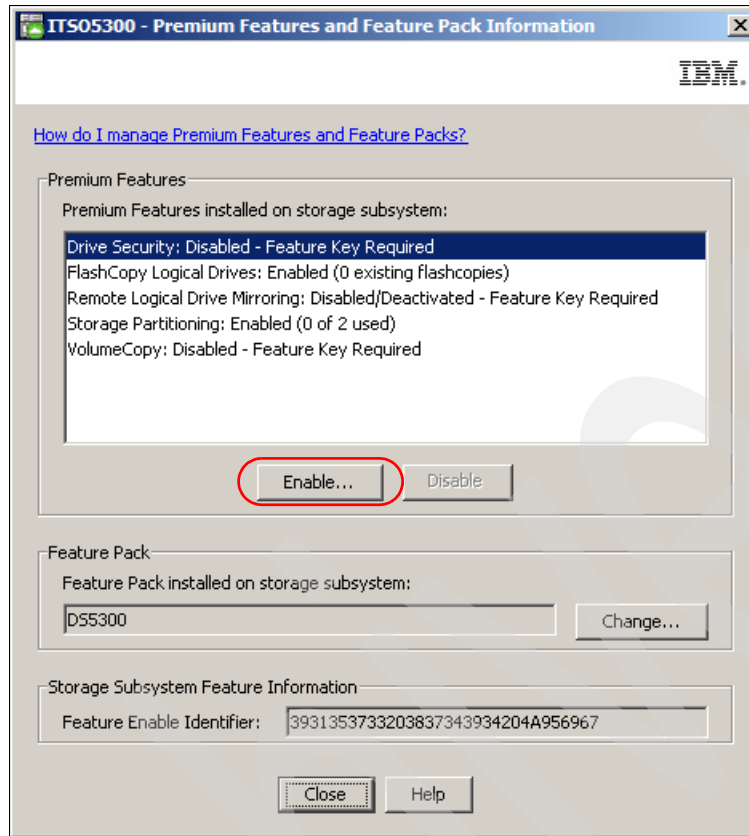


Figure 7-36 Enabling a premium feature

The DS storage subsystem validates the supplied code 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 is immediately available for use.

If the feature enable identifier does not match the DS storage subsystem, or it is already installed, you receive a notification and the key will not be installed.

7.2.3 Disabling a premium feature

To disable a premium feature, select **Storage Subsystem** → **Premium Features** in the Subsystem Management window. Choose the feature you want to disable from the list, as shown in Figure 7-37, and confirm.

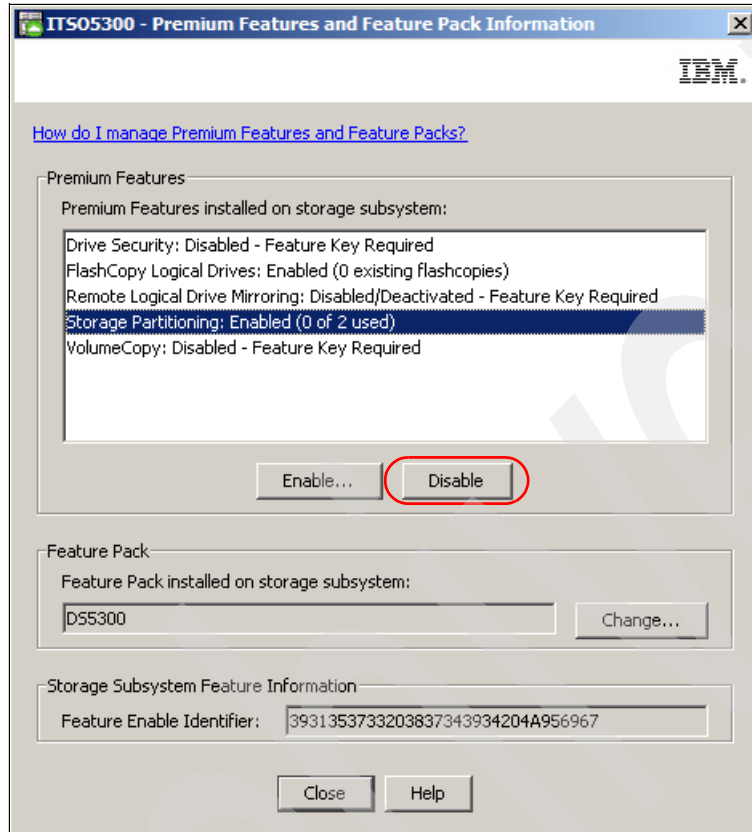


Figure 7-37 Disable premium feature

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.

7.3 Saving and loading the configuration

Once your DS storage subsystem is configured and running, save this configuration in order 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 parameters, including the storage partitioning configuration.

The saved file can be used to restore the configuration data to the same DS storage subsystem, or also to other DS storage subsystems in case you want to set up multiple storage subsystems with the same configuration. To allow that action, 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 select **Storage Subsystem** → **Configuration** → **Save** (Figure 7-38).

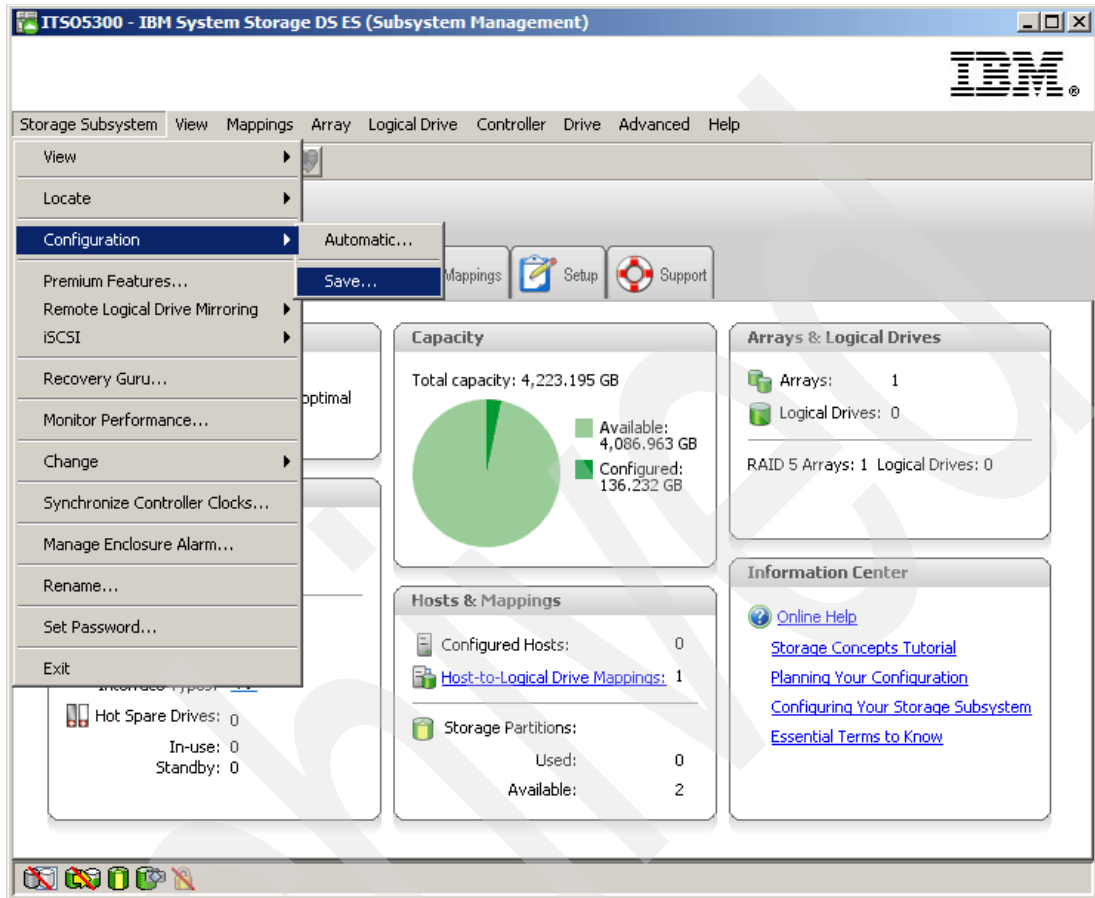


Figure 7-38 Saving the DS5000 storage subsystem configuration

We can choose to save specific elements of the configuration (Figure 7-39).

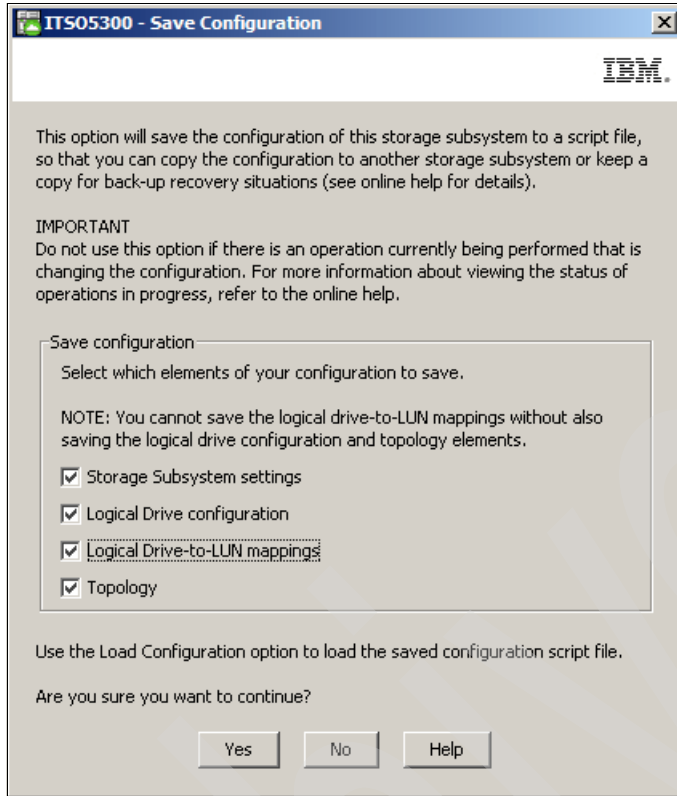


Figure 7-39 Saving configuration elements

Select the desired configuration elements, click **Yes**, and select a file name and destination folder in which to save the file. Make sure not to use a directory located on a DS storage subsystem disk, or you might not be able to access it when needed.

The script created can be used to replicate the configuration of the DS storage subsystem. 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

Attention: This procedure replaces any configuration on the storage subsystem. All data stored on the DS storage subsystem is lost because all logical drives are initialized. It is a good idea to save the configuration every time a change on the storage subsystem is made.

Do not attempt to load a saved configuration on the DS storage subsystem unless you fully understand the consequences.

To load the storage subsystem configuration, perform the following steps:

1. Open the Enterprise Management window and select the subsystem.
2. Select **Tools** → **Load Storage Subsystem Configuration** from the tool bar menu.
3. Point to the file containing the configuration and load it.
4. The Script Editor and a warning message appear. To load the configuration onto the DS storage subsystem, choose **Execute**. You can also edit the script before executing.

The procedure can take a long time, depending on the number of arrays and logical drives defined. When the procedure finishes, the storage 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 of the configuration. To read the current configuration, we use the View Profile option described in 7.3.1, “Storage subsystem profile” on page 371.

7.3.1 Storage subsystem profile

The storage subsystem profile is one of the most important items needed for IBM Support to help you solve whatever problem you have with your DS storage subsystem. 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 DS storage subsystem.

It can also be accessed through the Storage Manager interface in a readable format, as opposed to the saved configuration option (script), which is mostly used for backup and restore purposes.

Note: Always save your profile after you have changed the configuration of the DS storage subsystem. For example, if you create or delete logical drives, change the mapping, or add new disks or enclosures to the DS storage subsystem, save a new profile, as IBM Support might need it to help you in case of any problems.

To save the profile, select **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.

To see a more detailed procedure, including graphics, see “Storage subsystem profile” on page 229.

Reading the profile

Reading and interpreting the profile is a task usually done by IBM Support. There are some common information and failure situations that can be used, analyzed, and fixed easily by an administrator.

Controller and NVSRAM information

In Example 7-2, you can see the firmware and NVSRAM versions of controller A.

Example 7-2 Controller firmware and NVSRAM

Controller in Enclosure 85, Slot A

Status:	Online
Current configuration	
Firmware version:	07.60.13.04
Appware version:	07.60.13.04
Bootware version:	07.60.13.04
NVSRAM version:	N1818D53R1060V04
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 information shown in Example 7-3.

Example 7-3 Drive information

DRIVES-----

SUMMARY

Number of drives: 96
Supported drive types: Fibre (80), Serial ATA (SATA) (16)

TRAY, SLOT	STATUS	CAPACITY	TYPE	CURRENT DATA RATE	PRODUCT ID	FIRMWARE VERSION
0, 1	Optimal	698.638 GB	SATA	4 Gbps	HUA721075KLA330	43W9715 42C0417IBM GK80AB4A
0, 2	Optimal	698.638 GB	SATA	4 Gbps	HUA721075KLA330	43W9715 42C0417IBM GK80AB4A
0, 3	Optimal	698.638 GB	SATA	4 Gbps	HUA721075KLA330	43W9715 42C0417IBM GK80AB4A
0, 4	Optimal	698.638 GB	SATA	4 Gbps	HUA721075KLA330	43W9715 42C0417IBM GK80AB4A
0, 5	Optimal	698.638 GB	SATA	4 Gbps	HUA721075KLA330	43W9715 42C0417IBM GK80AB4A
0, 6	Optimal	698.638 GB	SATA	4 Gbps	HUA721075KLA330	43W9715 42C0417IBM GK80AB4A
0, 7	Optimal	698.638 GB	SATA	4 Gbps	HUA721075KLA330	43W9715 42C0417IBM GK80AB4A
0, 8	Optimal	698.638 GB	SATA	4 Gbps	HUA721075KLA330	43W9715 42C0417IBM GK80AB4A
0, 9	Optimal	465.762 GB	SATA	4 Gbps	ST3500641NS	39M4557 39M0181 IBM 3.AES
0,10	Optimal	465.762 GB	SATA	4 Gbps	ST3500641NS	39M4557 39M0181 IBM 3.AES
0,11	Optimal	465.762 GB	SATA	4 Gbps	ST3500641NS	39M4557 39M0181 IBM 3.AES
0,12	Optimal	465.762 GB	SATA	4 Gbps	ST3500641NS	39M4557 39M0181 IBM 3.AES
...						

Mappings information

Example 7-4 shows what is in the profile Mappings tab, which is similar to what is found in the mapping section of the Storage Manager client.

Example 7-4 Mappings information

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, select the **Arrays** tab to see the information shown in Example 7-5.

Example 7-5 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 from the disks in slots 1, 2, and 3 of enclosure 85. However, the drive in 85,1 seems to be failed and is using a hot spare, as location 85,4 is sparing for the drive in 85,1.

The status of the array is degraded, so the reconstruction is likely still in progress for the hot spare drive, but we have to check the logical drives (AIX1 and AIX2, in this case) that are part of this array, as shown in Example 7-6.

Example 7-6 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, because it still shows a Degraded status. We can cross check the information about the drives by using the Drives tab of the profile data (Example 7-7).

Example 7-7 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
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.

7.4 Migrating arrays between DS storage subsystems

The DS Storage Manager incorporates export and import options to safely move arrays between different DS storage subSystems without losing data.

This capability is very helpful when you have to upgrade or replace a DS storage subsystem with a new model or faster disks, but want to preserve the expansions and their data.

The export/import options check that all the conditions to support the disk migration are met before placing the disks offline and allowing removal of the disks.

Now, instead of using the option to place an array offline, as was the case with previous versions of Storage Manager, just select the **Export Array** option on the source machine. Select the **Import Array** option on the destination machine to accept the exported disks with their data.

- ▶ **Important:** We recommend upgrading both the source and the destination DS storage subsystems to the latest applicable firmware available before attempting any disk migration. Before you attempt to complete the drive migration procedure, review the information provided in the *IBM System Storage DS4000/DS5000 Hard Drive and Storage Expansion Enclosure Installation and Migration Guide*, GC53-1139.

7.4.1 Intermixing EXP810 and EXP5000 storage expansion enclosures

If you want to protect your current investment in IBM Midrange System Storage DS storage subsystems, you can migrate existing EXP810 expansion enclosures attached to an installed DS4700 or DS4800 to attach them to a DS5100 to DS5300. An RPQ approval from IBM is required for support of all migration configurations. With approved migration of EXP810 expansion enclosures to a DS5100 and DS5300, special consideration needs to be made about proper firmware levels and careful coordination needs to be made about differences in warranty and maintenance terms that will affect you. The purchase of new EXP810 expansion enclosures to attach to DS5100 and DS5300 will not be supported. When cabling EXP810 expansion enclosures behind a DS5100 and DS5300 storage subsystem, EXP810 expansion enclosures are cabled in the same manner as EXP5000 expansion enclosures. There are no special requirements to cable a mix of EXP810 and EXP5000 storage expansion enclosures behind a DS5100 and DS5300.

7.4.2 Intermixing EXP520 and EXP810 storage expansion enclosures

You can attach the EXP810 expansion enclosure to the DS5020 storage subsystem only after purchasing the Attach EXP810 to DS5020 Activation license option and activating it in the DS5020 storage subsystem. When cabling an EXP810 expansion enclosure behind a DS5020 storage subsystem, the EXP810 expansion enclosure is cabled in the same manner as an EXP520 expansion enclosure. There are no special requirements to cable a mix of EXP810 and EXP520 storage expansion enclosures behind a DS5020.

7.4.3 Migration prerequisites

To perform a successful export and import of arrays, the following conditions must be observed:

Attention: Failure to meet these conditions before you migrate hard disk drives might result in loss of data availability or loss of data.

- ▶ Run the Recovery Guru and make sure that the source and destination subsystems are in the Optimal state. Make sure to correct any problems before starting the migration. Check:
 - For failed drives
 - For hot spare drives in use (they all should be in Standby status)
 - That the entire array will be migrated (the disks in the array will be removed)
 - Missing logical drives
- ▶ The array being migrated must *not* have:
 - Logical drives with persistent reservations
 - Logical drives re-configuring or reconstructing
 - FlashCopy repository logical drive

- Volume copy source or target logical drive
- Mirror primary or secondary logical drive or mirror repository logical drive
- ▶ The array being migrated must have no mappings assigned.

We recommend installing the latest applicable levels of firmware in both the source and destination DS storage subsystems.

In addition:

- ▶ Check the *IBM System Storage DS4000/DS5000 Hard Drive and Storage Expansion Enclosure Installation and Migration Guide*, GC53-1139 for additional considerations.
- ▶ Do not migrate from a DS storage subsystem with a newer firmware level than the destination.
- ▶ Make sure that your source and destination DS storage subsystems have unique array names and logical drives names.

Important: Before you attempt to complete the drive migration procedure, review the information provided in the latest *IBM System Storage DS4000/DS5000 Hard Drive and Storage Expansion Enclosure Installation and Migration Guide*, GC53-1139.

In addition to these prerequisites, there are additional distinct considerations specifically for the source and for the destination DS storage subsystems.

Source DS storage subsystem prerequisites

On the source storage subsystem, perform the following actions before starting the export operation:

- ▶ Save the storage subsystem configuration and profile data. This is a precaution to help you restore your configuration in the event of a problem during the export. Make sure to save both the configuration and profile files outside of the DS storage subsystem.
- ▶ Stop all associated I/Os. Stop all input and output and un-mount or disconnect file systems on the logical drives in the array selected for export.
- ▶ Back up array data. Back up data on the logical drives in the array selected for export, and verify the backup. Make sure that you save the backup outside of the DS storage subsystem disks.
- ▶ Make sure that the hard disk drive firmware is the latest applicable level.
- ▶ Locate the array and label drives. Use the locate array function to flash the LEDs on the drives in the array, and then label them with the source and destination storage subsystem names, array name, and total number of drives in the array.
- ▶ If you continue using the source enclosure expansion, obtain blank drive canisters or new drives, because after you remove the drives in the array, you must replace them with blank drive canisters or new drives to maintain proper airflow within the enclosure.

Destination DS storage subsystem prerequisites

On the destination storage subsystem, be sure of the following items:

- ▶ Verify the available drive slots. Be sure that you have enough empty slots for the drives you will be importing.

- ▶ Check that the drives to be imported are supported by the storage subsystem.
 - The storage subsystem must support the type of drives that you are exporting (that is, an intermix of SATA and Fibre Channel drives).
 - You cannot exceed the maximum number of drives supported by the storage subsystem.
 - Make sure that the drives are compatible with the storage expansion enclosure. For example, insert a 4 GB drive into a storage expansion enclosure that supports 4 GB drives.
- ▶ You cannot exceed the maximum supported number of logical drives.
- ▶ Make sure that the storage subsystem has the latest applicable controller firmware, nonvolatile storage random access memory (NVSRAM), and ESM firmware. Also, make sure that the installed controller firmware in the storage subsystem supports the drives and expansion enclosures.

7.4.4 Migrating an array

The migration is started using the export option in the DS Storage Manager client, and it will perform a preliminary check to validate the migration.

Perform the following steps:

1. In the Subsystem Storage Manager window, select the array to be exported.

2. Select **Advanced** → **Maintenance** → **Export array** from the Subsystem Management window, as shown in Figure 7-40.

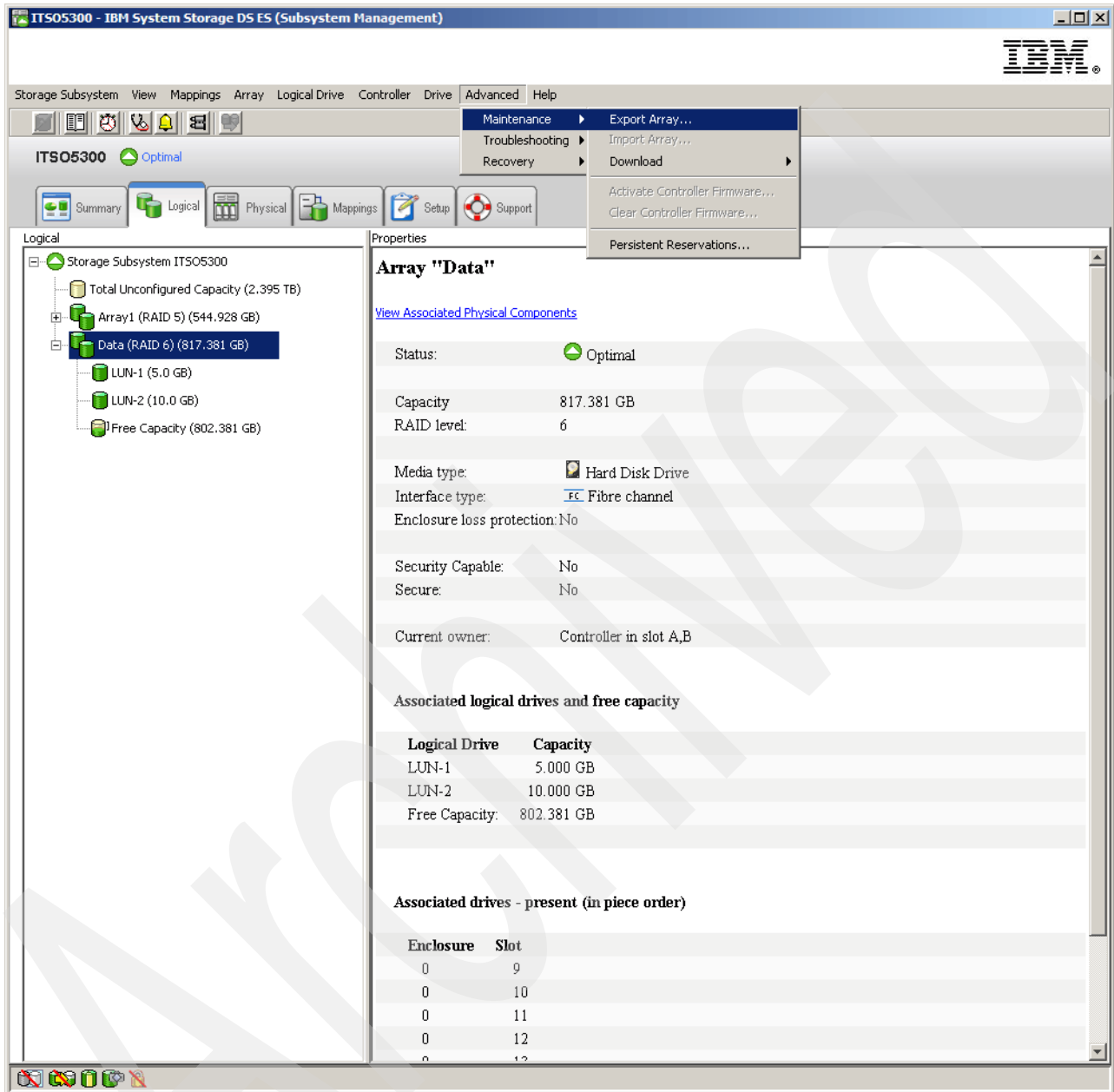


Figure 7-40 Exporting an array

3. The Export Array option opens a window detailing all the steps to cover before continuing the operation, as shown in Figure 7-41. The steps can be saved to a text file by clicking **Save As...**

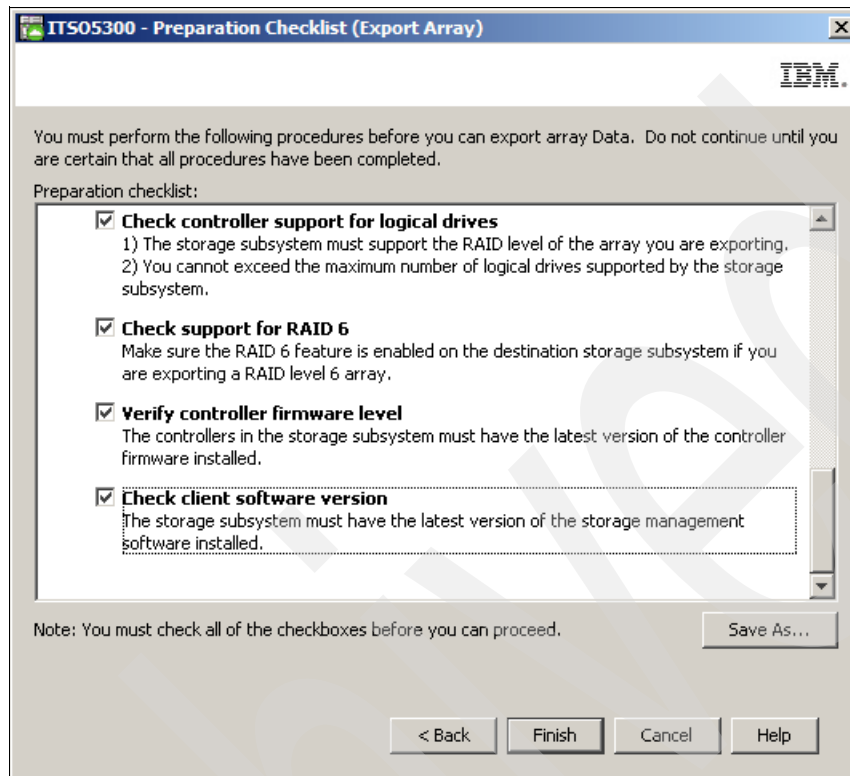


Figure 7-41 Exporting array instructions

The checklist reminds you to:

- For the source storage subsystem:
 - Save the storage subsystem configuration.
 - Stop all I/O on the logical drives and unmount or disconnect the file system for the logical drives contained in an array selected for export.
 - Back up data on the logical drives in the array selected for export.
 - Locate the array and label drives.
 - Obtain blank drive canisters or new drives to maintain airflow.
- For the target storage subsystem:
 - Verify the available drive slots.
 - Check support for drives.
 - Check controller support for logical drives.
 - Check support for RAID 6.
 - Verify the controller firmware level.
 - Check the client software version.

Once you have verified all the conditions listed, check the check box next to each item and click **Finish** to continue.

Note: Checking the box beside each task does not automatically cause the task to be completed. You must complete each task as you typically would. Checking the box simply helps you track the tasks you have completed and enables the Export button in the Export Array window.

4. Confirm the export operation in the confirmation window by typing Yes.

When finished, the window shown in Figure 7-42 opens and instructs you about how to remove the disks and transport them to the destination system. These instructions can be saved to a text file by clicking **Save As...** .

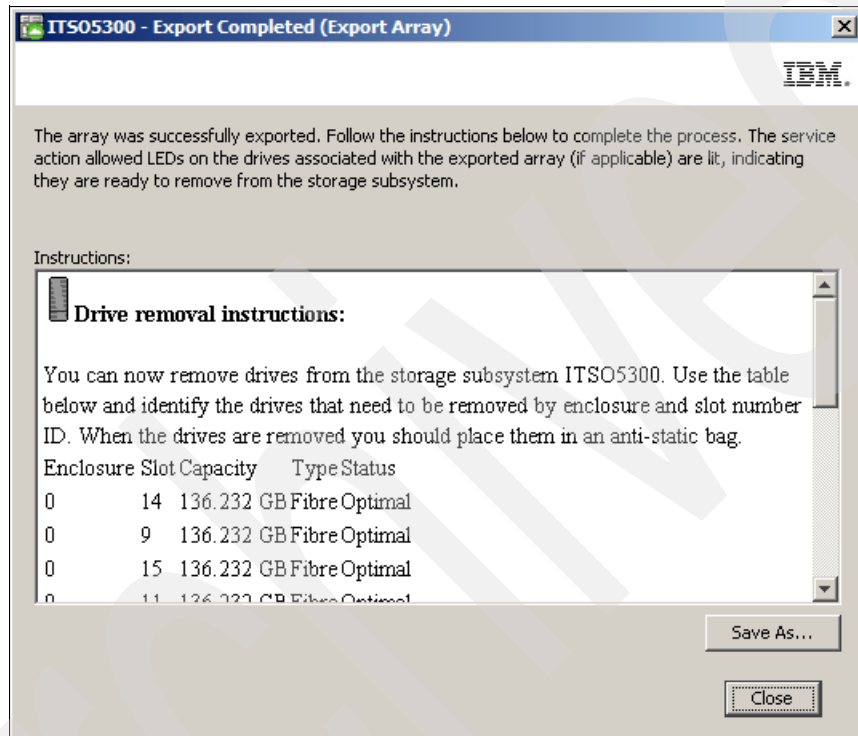


Figure 7-42 Exporting array: Disk removal instructions

- Before you start removing the disks, check the Storage Manager window to make sure that all the disks and the array are in an offline status, as shown in Figure 7-43.

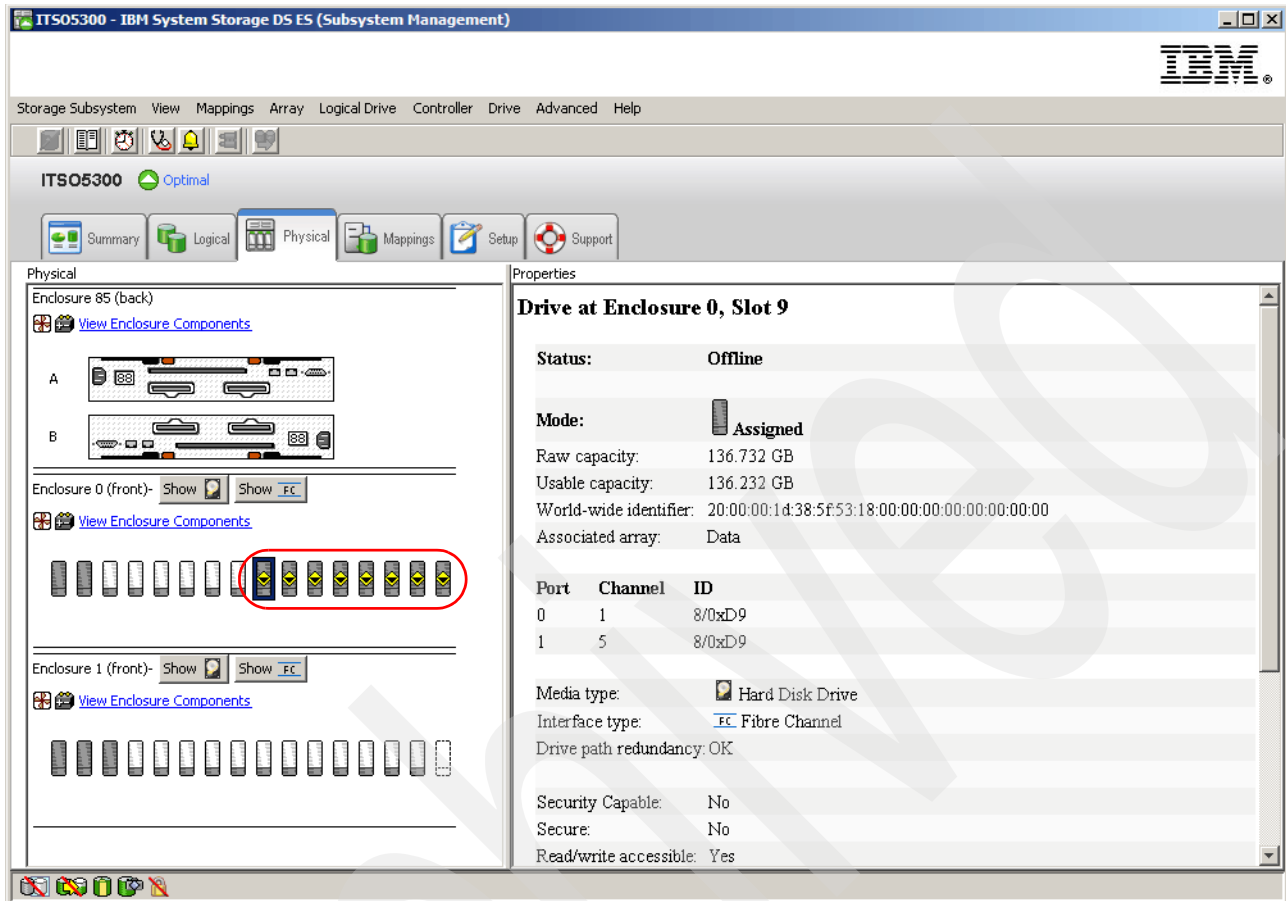


Figure 7-43 Export array: Disks offline

- You can now remove the drives (be sure to label them before removal). You have to remove *all* the of the exported array disks, and insert a dummy drive in each slot after removal to allow proper airflow. The dummy is needed only if the enclosure will remain operational.

Note: The export array operation sets all array disks in an offline status as part of the procedure. *Before* removing any disk, make sure that all the array disks are effectively marked offline, and all of them were appropriately labeled.

- Carefully package each drive and move it to the destination system.

Proceed with the import process described in the 7.4.5, “Importing an array” on page 382.

7.4.5 Importing an array

Once you have removed all the disks being migrated from the source DS storage subsystem, you can start importing the disks into the destination DS storage subsystem.

2. Click **OK** only if you are sure that all the drives making up the array are the ones that you really want to import (and that they were already correctly exported and removed from the source DS storage subsystem). The Import Array window is displayed, as shown in Figure 7-45.

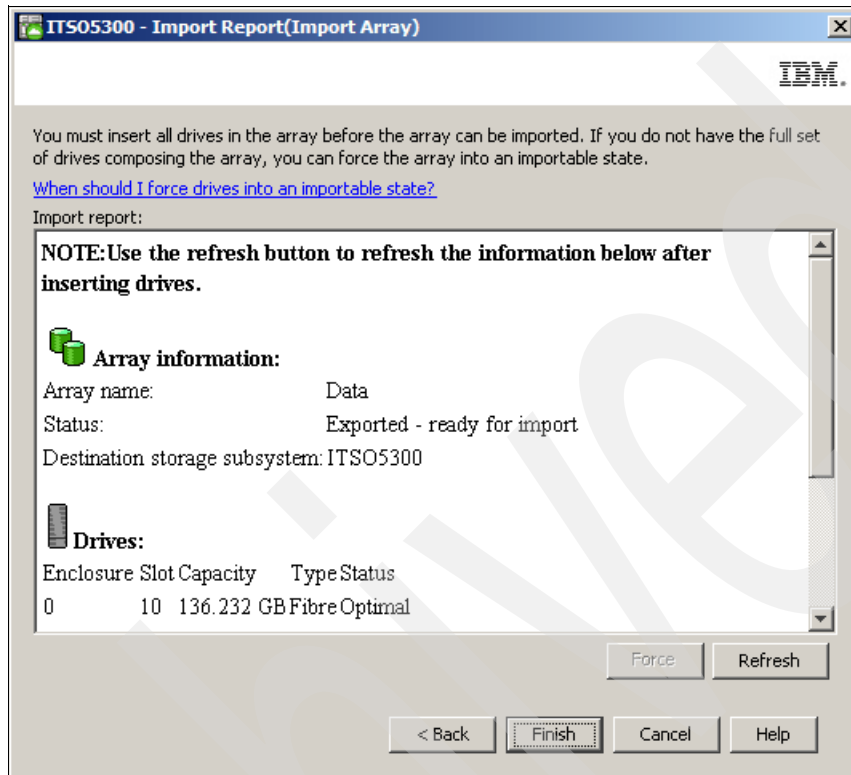


Figure 7-45 Import Array window

3. Click **Finish** if the configuration to import matches the exported configuration. If not, stop and go back to the original DS storage subsystem.
4. If the controller detects any condition that will interfere with the current configuration, it lets you know. If the condition detected is more critical, *stop here*. It is better to go back to the source DS storage subsystem to solve the issue.
5. A confirmation window displays. Enter Yes and click **OK**.

6. The next window (Figure 7-46) displays the result of the import process. Make sure that the operation finished successfully, and scroll down the window to see the logical drives imported with the array. Make sure that all of them are available and click **Close** to terminate the dialog.

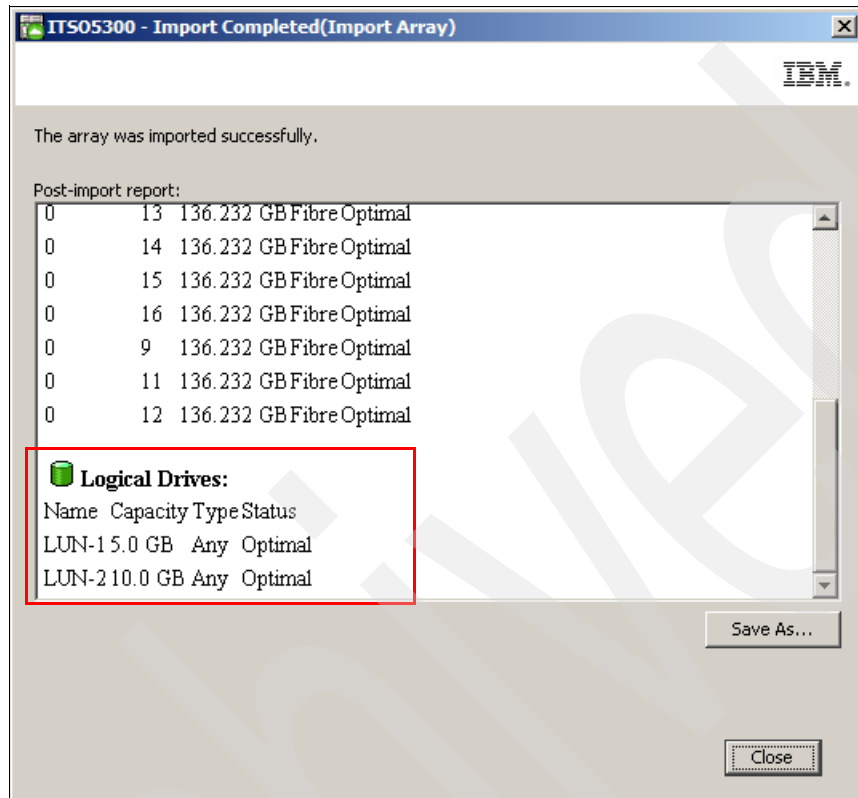


Figure 7-46 Import success

A successful import operation finishes with the Storage Manager recognizing the disks in a Optimal status, as well as the array and the logical drives as they were in the source DS Storage Subsystem. See Figure 7-47.

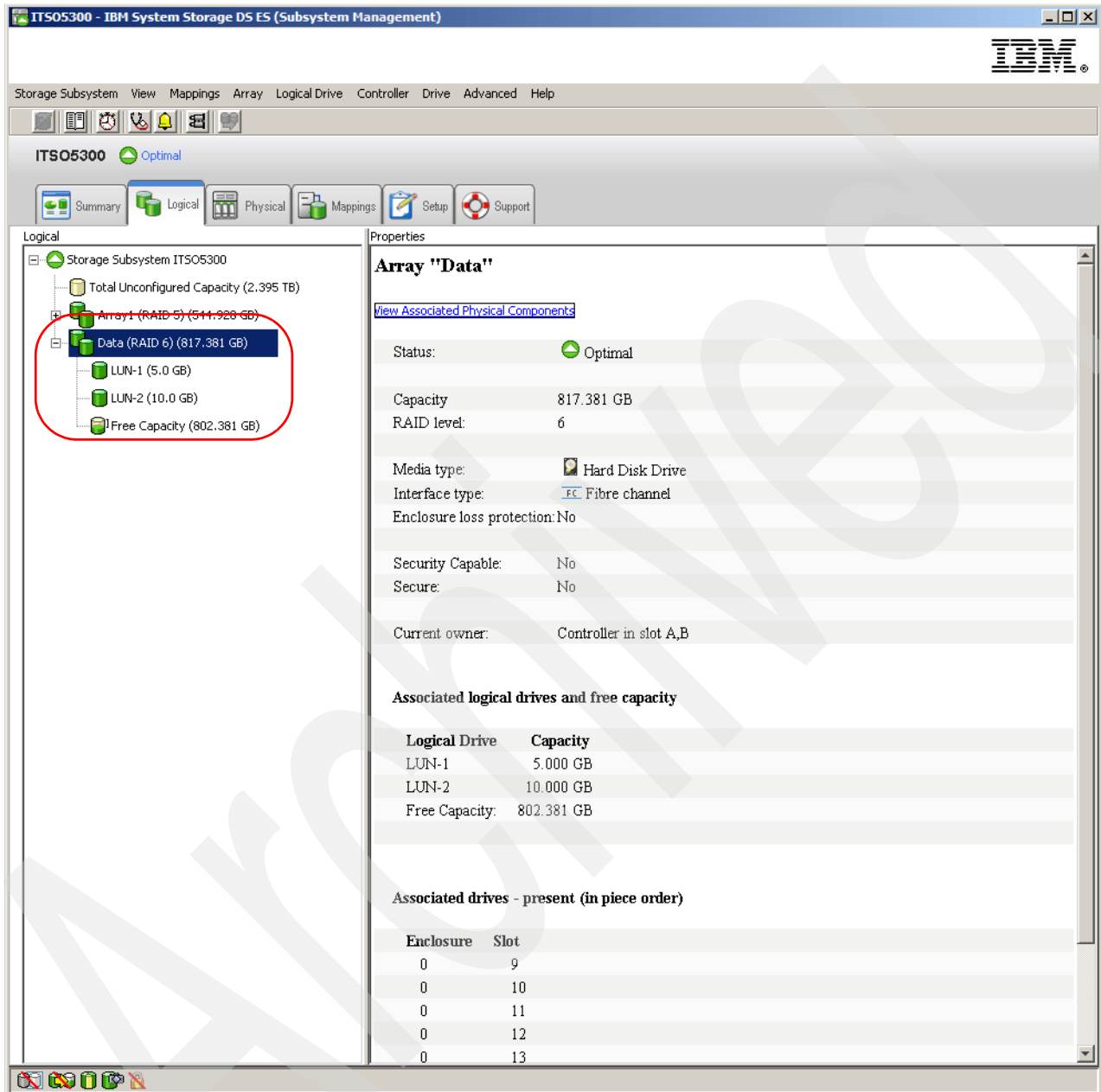


Figure 7-47 Finish import operation

A last step to perform before using the disks is to create the desired mappings. Once done, you will be able to access the migrated data, thus completing a successful migration.

7.5 Performing an upgrade from a DS4700 or DS4800 storage subsystem to a DS5000 storage subsystem

Important: Before you attempt the upgrade procedure, review the information provided in the *IBM System Storage DS4000/DS5000 Hard Drive and Storage Expansion Enclosure Installation and Migration Guide*, GC53-1139 for additional instructions.

To upgrade a DS4700 or DS4800 storage subsystem to a DS5000 storage subsystem, complete the following steps:

1. Make sure that the data in the existing configuration is backed up in a secure place before starting the upgrade procedure.
2. Consider the following items before you upgrade a functioning storage subsystem:
 - a. Host attachment and premium feature entitlements
 - i. To enable premium features in a new or replacement configuration that were enabled in an original configuration, you must purchase the applicable entitlements for the new or replacement storage subsystem, if that premium feature is not standard in the new or replacement storage subsystem. In addition, host attachment kits only assigned to specific storage subsystems. You must also purchase the applicable host attachment kits for new or replacement storage subsystems.
 - ii. After the upgrade, premium features that were previously enabled in the original storage subsystem along with the enable-by-default premium features in the new storage subsystem are automatically placed in the Out of Compliance state. You must generate new premium feature SAFE keys to re-enable premium features.
 - b. Storage firmware migration
 - i. You can migrate the drives and drive expansion enclosures in a functioning storage subsystem to a new controller enclosure only if the firmware in the controllers of the functioning storage subsystem and the new controller enclosure have the same major release version. Only major release Versions 6.xx and 7.xx are supported.
 - ii. To migrate from an original storage subsystem controller with firmware Version 6.xx to a new storage subsystem controller with firmware Version 6.xx, the original and the new storage subsystem controller firmware must be upgraded to the latest Version 6.xx before you perform the migration.
 - iii. To migrate from an original storage subsystem controller with firmware Version 7.xx to a new storage subsystem controller with firmware Version 7.xx, the controller firmware version in the original controller enclosure must be the same or earlier than the controller firmware version in the new controller. Otherwise, the controller firmware in the new controller enclosure is placed in the Controller Lock down state.
 - iv. To migrate from an original storage subsystem controller with firmware Version 6.xx or earlier to a new storage subsystem controller with firmware Version 7.xx or later, upgrade the original storage subsystem controller firmware to Version 7.xx or later before you perform the migration. Otherwise, the controller firmware in the new controller enclosure is placed in the Controller Lock down state.

- v. To migrate from an original storage subsystem controller with firmware Version 7.xx or later to a new storage subsystem controller with firmware Version 6.xx or earlier, the new storage subsystem controller firmware must be upgraded to Version 7.xx or later before you perform the migration. You must purchase two additional drives and a storage expansion enclosure to bring the new storage subsystem controller into the Optimal state with drives attached before you can update the controller firmware. If the upgrade of the new storage subsystem controller firmware to Version 7.xx or later is not possible, the drives that are configured in the storage subsystem with firmware Version 7.xx or later cannot be migrated into a storage subsystem with controller firmware Version 6.xx or earlier.
- c. Supported upgrades
 - i. Migrating from a configuration with an integrated drive/RAID controller DS4000 storage subsystem to one with the RAID controller only DS4000 storage subsystem requires an additional storage expansion enclosure for the drives that are installed in the integrated drive/RAID controller DS4000 storage subsystem chassis.
 - ii. Some storage subsystem models require that hard disk drives and drive expansion enclosures operate at a specified Fibre Channel speed. Make sure that the hard disk drive and drive expansion enclosures can operate at that speed before you begin the upgrade.

7.5.1 Planning the upgrade

To plan the upgrade, perform the following steps:

1. Migrating from a DS4700 or DS4800 storage subsystem to a DS5100 or DS5300 storage subsystem is supported. However, you must submit a RPQ with IBM to migrate any EXP810 drive expansion enclosures connected to the existing DS4700 or DS4800 controller. You can submit an RPQ to your IBM marketing representative or authorized reseller.

Note: Only EXP810 drive enclosures can be migrated from a DS4700 or DS4800 configuration into DS5100 or DS5300 configuration.

2. Purchase the premium feature entitlements that are enabled in the original storage subsystem for the new storage subsystem, if that premium feature is not standard in the new storage subsystem.
3. Purchase the host attachment entitlement kits for the new storage subsystem.
4. If you are migrating from a working DS4700 configuration, purchase an additional EXP5000 drive expansion enclosure to install the hard disk drives in the internal bays of the DS4700.
If there are enough empty drive bays in the existing EXP810 drive expansion enclosures, you can move the drives in the original DS4700 enclosure to the empty drive bays.
5. Lay out the drive expansion enclosure cabling to the new storage subsystem. See the documentation that comes with the new DS5000 storage subsystem for more information. See the *Installation, User's, and Maintenance Guide*, GC26-7798 for your storage subsystem for information about the storage expansion enclosures cabling rules.
6. Purchase any additional hardware that is required to cable the existing drive expansion enclosures to the new storage subsystem using the drive expansion enclosure cabling layout as a guide.
7. Make sure that the original subsystem is in the Optimal state.

8. Perform a full backup of the original storage subsystem and schedule it for down time.
9. Retrieve the proofs of purchase for both the original and new storage subsystems and for any additional premium feature entitlements on the new and original storage subsystems.
10. If there are any switch zoning definitions or applications that rely on the storage subsystem worldwide names, plan to update them to use the new storage subsystem worldwide names after the migration to the new storage subsystem is complete.

7.5.2 Preparing the new storage subsystem

To prepare the new storage subsystem for the upgrade, perform the following steps:

1. Unpack the new DS5000 storage subsystem and install it in a rack. Do not connect it to the drive expansion enclosures attached to the original DS4700 or DS4800 storage subsystem.
2. Connect the new storage subsystem to the systems-management network using the default IP addresses of the controllers and record the version of the controller firmware on the new storage subsystem.

The default TCP/IP address of controller A Ethernet port 1 is 192.168.128.101 and the default TCP/IP address of controller A Ethernet port 2 is 192.168.129.101.

The default TCP/IP address of controller B Ethernet port 1 is 192.168.128.102 and the default TCP/IP address of controller B Ethernet port 2 is 192.168.129.102.

7.5.3 Preparing the original storage subsystem

To prepare the original storage subsystem for the upgrade, perform the following steps:

1. If any long running tasks are processing in the original storage subsystem, make sure that they have completed. Examples of long running tasks are:
 - Dynamic logical drive capacity expansion
 - Dynamic volume expansion (DVE)
 - Dynamic capacity expansion (DCE)
 - Logical drive segment size modification
 - Array RAID-level modification
 - User-initiated array redundancy checking (on the Storage Subsystem Management window, select **Array** → **Check Redundancy**).
 - Remote mirror logical drive synchronization FlashCopy image or VolumeCopy image logical drive creation.
 - Logical drive reconstruction or copyback.
2. Save the storage subsystem profile in a safe location and not on the logical drives that are mapped from the original storage subsystem.
3. Record the version of the controller firmware that is on the storage subsystem.
4. Collect all support data of the original storage subsystem.
5. Stop all programs, services, and processes on the host servers that access the logical drives that are defined in the migrated hard disk drives. Also, make sure that there are no programs, services, or processes running in the background that write data to the logical drives.

6. Unmount the file systems to flush I/O from the server cache to disks.
 - If you are using a Windows operating system, remove the drive letter or the mount points of the drive-to-LUN map definitions, instead of unmounting the file systems.
 - See your operating-system documentation for information about the file system unmount procedure.
7. Perform an incremental backup of the data that was changed since the full backup that you made in 7.5.1, “Planning the upgrade” on page 388.
8. Make sure that the environmental service modules (ESMs) and hard disk drives in the original storage subsystem are updated to the latest firmware level. To download the latest firmware level, go to the following address:
<http://www.ibm.com/systems/support/storage/disk/>

7.5.4 Upgrading the controller firmware

To upgrade the controller firmware, perform the following steps:

1. Use the flow chart shown in Figure 7-48 on page 391 to determine the firmware version required in the new storage subsystem. To download the latest firmware level, go to the following address:
<http://www.ibm.com/systems/support/storage/disk/>
2. The controller firmware is normally listed as either xx.yy.zz.aa or xxyyzzaa, where xx.yy or xxyy is the controller firmware version used for compatibility checking. If the first x=0, it might not be identified. For example, 07.36.14.01 is the same as 7.36.14.01. The firmware version used for compatibility checking in this example is 7.36.

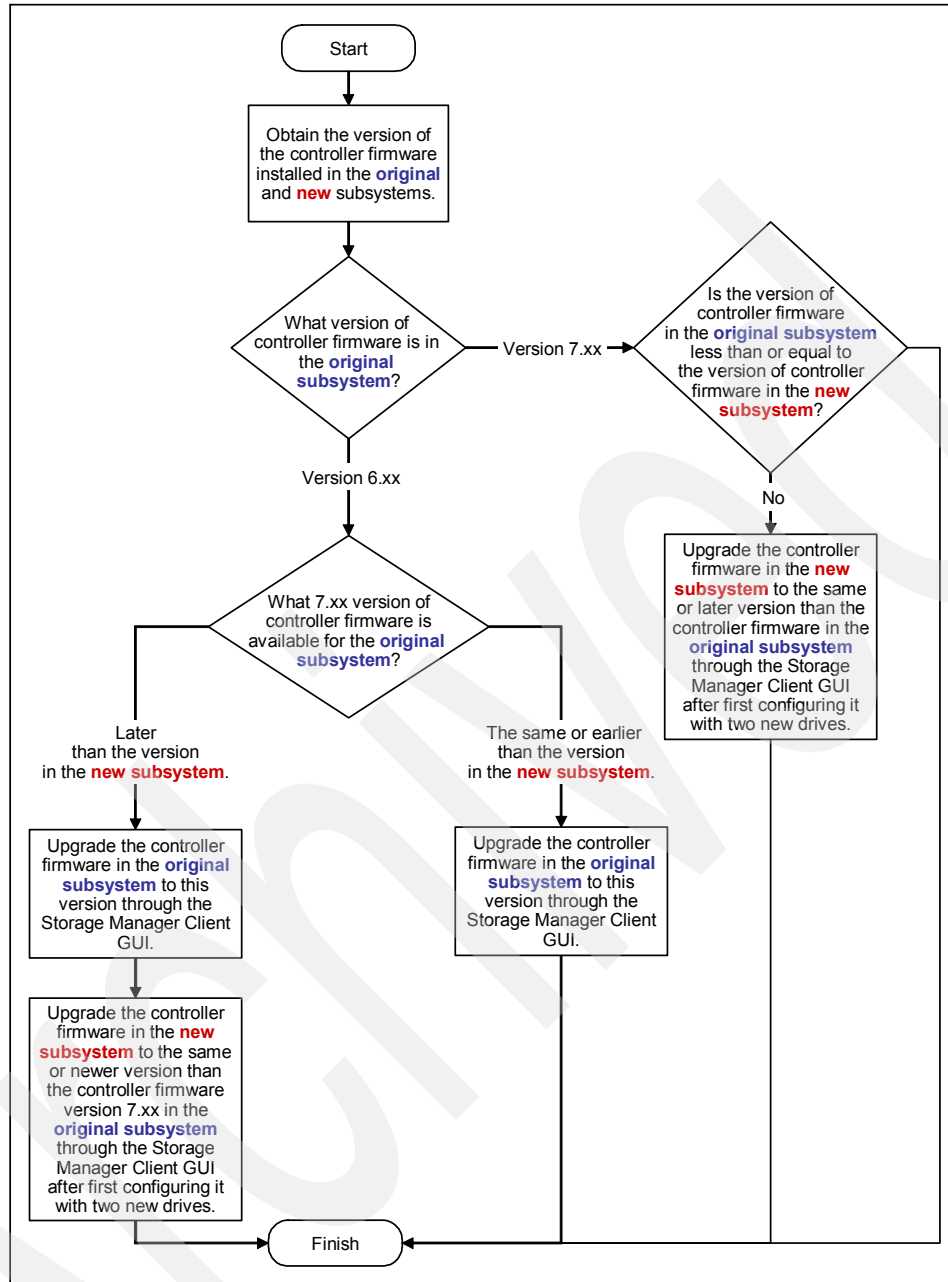


Figure 7-48 Firmware compatibility flow chart for a DS4700 or DS4800 to a DS5000 storage subsystem upgrade

3. If you configured the new storage subsystem with at least two new drives to place the storage subsystem into the Optimal state for updating the controller firmware, power off the new storage subsystem controller enclosure and remove the two drives (along with the additional expansion enclosure, if attached).
4. Save the full configuration of the original storage subsystem into a SMcli script file. Make sure that you check the check boxes for the storage subsystem settings, logical drive configurations, logical drive-to-LUN map definitions, and topology before proceeding with the configuration save. Make sure that the location you are saving to is not any of the logical drives that are mapped from the original storage subsystem.
5. Delete any FlashCopy images of the logical drives.

6. Delete any VolumeCopy logical drive pairs.
7. Delete any remote mirror relationships.
8. Delete any host-to-LUN mapping definitions in the original storage subsystem.

7.5.5 Switching from the original to the new storage subsystem

To switch from the original to the new storage subsystem, perform the following steps:

1. Perform one of the following tasks to export arrays, depending on the availability of hot spare drives and unconfigured drives in the storage subsystem:
 - If hot spare drives or unconfigured hard disk drives are installed in the original storage subsystem, export all of the defined arrays.
 - If no hot spare drives or unconfigured hard disk drives are installed in the original storage subsystem, keep one array on the original storage subsystem and export all the other arrays. One array must be in the Optimal state for the storage subsystem to be up and running.

Note: An error occurs if you try to export the last defined array in a storage subsystem configuration and no hot spare drives or unconfigured drives are installed.

To export an array using the Storage Subsystem Management window, right-click the name of the array and select **Advanced** → **Maintenance** → **Export Array**, and follow the instructions in the window that opens. You can also use the Start Array <array name> Export command in the IBM Storage Manager Client script window.

2. Power off the original controller enclosure first, and then power off the drive expansion enclosures. This is the best practice power-off sequence. See the documentation that comes with the storage subsystem for details about the power-off sequence.

Note: The controller enclosure must be powered off before the drive expansion enclosure.

3. Label all the cables connected to the original storage subsystem enclosure.
4. Wait until the LEDs on the storage subsystem chassis are off, and then disconnect all cables from the storage subsystem chassis.
5. Remove the original storage subsystem enclosure from the rack.
6. Install the new storage subsystem enclosure in the rack.
7. If the original storage subsystem is a DS4700 storage subsystem, install an EXP5000 drive expansion enclosure and move the hard disk drives from the original storage subsystem to the drive expansion enclosure.
8. Insert the SFPs into the new storage subsystem drive loop/channel port bays and cable the drive expansion enclosures to the new storage subsystem using the cabling layout you defined in 7.5.1, "Planning the upgrade" on page 388.
9. Insert the SFPs into the new storage subsystem host port bays and cable the host interface ports and the storage subsystem management ports of the new storage subsystem enclosure.
10. Make sure that all of the storage expansion enclosures are set to the same speed for each drive channel/loop.

7.5.6 Preparing the new storage subsystem for use

To prepare the new storage subsystem for use, perform the following steps:

1. If the controller TCP/IP addresses are assigned using DHCP, update the DHCP records with the new controller Ethernet port MAC addresses.

The controllers first check for a DHCP server during the boot process. If the controllers do not detect a DHCP server, they use either the static IP address (if defined) or the default IP addresses.

2. Power on the drive expansion enclosures if they are powered off. Do not power on the new storage subsystem controller enclosure. Check the drive expansion enclosure LEDs to make sure that the drive expansion enclosures are connected properly.
3. Power on the new storage subsystem controller enclosure.

If the TCP/IP addresses of the Ethernet management ports are statically defined for the original storage subsystem controllers, the TCP/IP addresses are used for the same Ethernet management ports in the new controllers.

4. Connect the new storage subsystem to the IBM DS Storage Manager Client either through the out-of-band method using the applicable TCP/IP addresses of the controller Ethernet management ports or through the in-band method through Fibre Channel connections.

Note: The new storage subsystem identifies itself as the machine type that it replaced until you download the applicable NVSRAM firmware for the new storage subsystem.

5. Make sure that the new storage subsystem configuration is in the Optimal state and that all of the drives are identified. Use the Recovery Guru in the DS Storage Manager Client Subsystem Management window to resolve any Needs Attention conditions.
6. Update the controller firmware of the new storage subsystem to the latest available version, if required.
7. Download the applicable NVSRAM firmware for the new storage subsystem.
8. Import all of the arrays that were exported in 7.5.5, “Switching from the original to the new storage subsystem” on page 392. Make sure that all of the arrays are online and in the Optimal state.
9. If there are any ghost hard disk drives, hard disk drives that are indicated as incompatible, or if any of the following conditions persist, contact IBM Support for assistance:
 - The empty drive bay icon is displayed for the drive bay into which you inserted the migrating drive.
 - The Failed unconfigured drive icon or the Failed configured drive icon is displayed for the drive bay into which you inserted the migrating drive.
 - Array configuration data on the drives you have added is incomplete.
 - You cannot bring the array online (controller firmware V6.xx.xx.xx or earlier) or import the array (controller firmware V7.xx.xx.xx or later).
10. Use the Enable Identifier storage subsystem premium feature to generate and apply premium features keys to remove Out of Compliance errors on enabled premium features from the original storage subsystem. See 7.2, “Handling premium features” on page 362 or the instructions that come with the Enable Identifier premium feature for information about generating the premium feature keys.

11. Extract the applicable SMCLI commands in the configuration script file that you saved in 7.5.4, “Upgrading the controller firmware” on page 390 to recreate the FlashCopy images, VolumeCopy images, remote mirror relationships, and host-to-LUNs map definitions, as required.
12. Make sure that the enclosure IDs in each drive loop/channel contain a unique first-position digit (x1). In addition, if the drive expansion enclosures are re-cabled behind the new storage subsystem controller enclosure, modify the second-position digit (x10) so that they have the same second-position digit for all the drive expansion enclosures in a drive channel/loop.
13. Update the switch zoning definitions and any applications that rely on the storage subsystem worldwide names to use the new storage subsystem worldwide names.

7.6 Securing the DS5000 storage subsystem client using remote management

The Remote Services feature of Microsoft Windows XP allows users of the DS5000 storage subsystem to remotely access the storage subsystem’s client, profile, information, and logs through a Virtual Private Network (VPN) connection and Microsoft’s Remote Desktop (RDP) function.

Remote access of the client workstation has some big advantages. It allows you to configure the DS5000 storage subsystem without having to be physically present at the client workstation. This will allow you to be more responsive to your storage environment at any time and from any place. However, using a secure connection is very important. If unauthorized users are able to gain access to your internal client workstation, they can launch the Storage Manager client and gain access to any of the DS5000 storage subsystem’s configuration information. Also, if you do not have a password set, they can even modify the configuration or potentially compromise data.

Note: The IBM System Storage DS Storage Manager client allows you to set a password to prevent unauthorized users from making configuration changes. However, the password does not prevent an unauthorized user from viewing the Storage Manager client if they can gain access to your client workstation.

Securing your internal Storage Manager Client workstation, while allowing authorized users remote access, can be best achieved by implementing the following:

- ▶ Virtual Private Network (VPN) connection: Permits only authorized users access to your internal network.
- ▶ Remote Desktop Users/Passwords: Permits only authorized users to access your Storage Manager Client workstation.
- ▶ “Dual-homed” Storage Manager Client workstation:
 - One network card for the DS5000 Management LAN
 - One network card for the customer WAN
 - Physically separates the DS5000 management network from the rest of the enterprise network
- ▶ Storage Manager Client Password: Permits only authorized users to make configuration changes to the DS5000 storage subsystem through the Storage Manager Client.

Figure 7-49 shows how to secure your environment for remote access. By introducing multiple layers of security, you will be best protected from unauthorized access to your DS5000 storage subsystem.

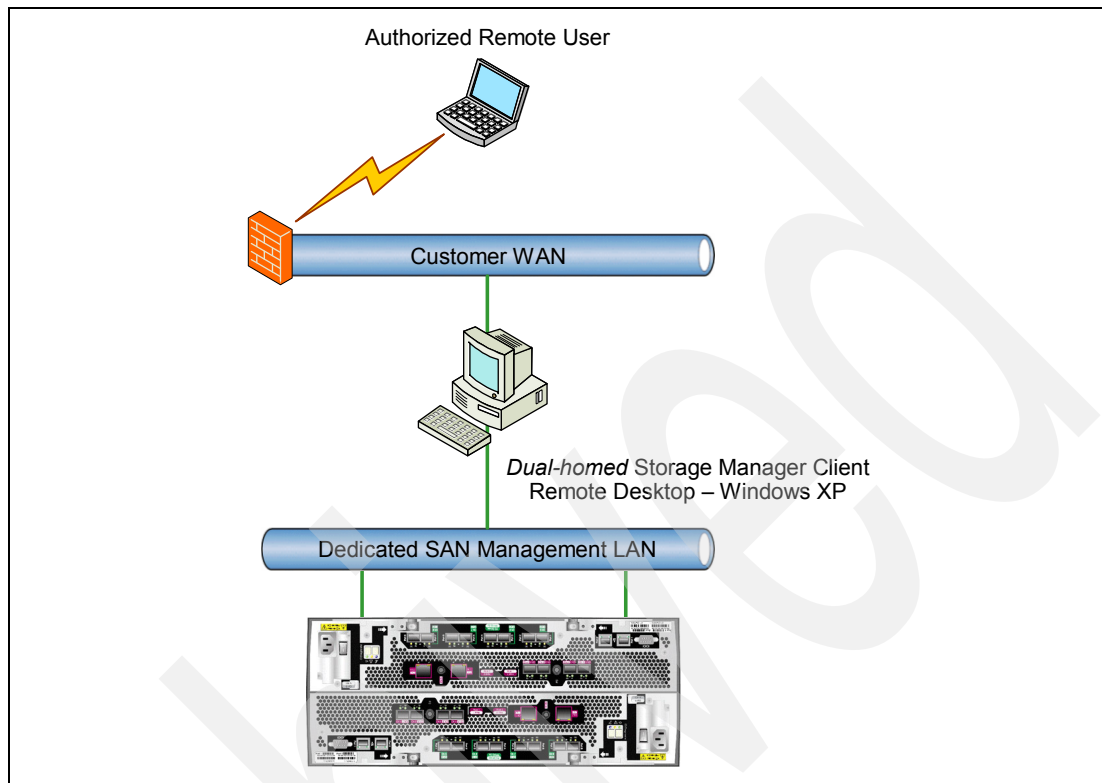


Figure 7-49 Securing remote access to the Storage Manager Client

In addition, Remote Desktop (RDP) is a high-encryption client that can run on any Windows server. It uses RSA Security's RC4 cipher and it has key strengths of 40-, 56-, or 128-bit.

Hardware requirements

In order to use Remote Desktop for the secure remote management of the DS5000 storage subsystem, you must have the following hardware/settings:

- ▶ Windows XP Professional on the client workstation
- ▶ Client workstation must be "dual-homed"
 - One network card for the dedicated DS5000 Management LAN (can be VLAN)
 - One network card for the customer WAN
- ▶ Accounts that will allow authorized users to remote control the desktop
- ▶ Client configured with Remote Desktop
- ▶ Client configured for Remote Assistance
- ▶ The client should have these DS5000 storage subsystem tools installed for management:
 - Storage Manager Client (latest version)
 - Storage Manager Utilities (latest version)
 - Qlogic SANsurfer (latest version)
 - PuTTY Version 0.6 or higher
- ▶ Dedicated DS5000 Management LAN (or VLAN)

- ▶ Customer firewall configuration
 - Open firewall for well-known TCP Port (3389)
 - Must be able to support dedicated VPN connection
 - Can be disabled (when not needed) for security precautions

The following Web sites offer more information about RDP and Remote Assistance:

- ▶ Microsoft information about RDP:
<http://www.microsoft.com/windowsxp/using/mobility/default.aspx>
- ▶ Microsoft RDP Frequently Asked Questions (FAQ):
<http://www.microsoft.com/windowsxp/using/mobility/rdfaqs.aspx>
- ▶ Microsoft Remote Assistance:
<http://www.microsoft.com/windowsxp/using/helpandsupport/learnmore/remotearr/intro.aspx>

7.7 Preventative maintenance and data collection

It is essential to regularly monitor the status of the DS4000 or DS5000 storage subsystem in order to identify potential problems promptly before they become more critical. The IBM System Storage DS Storage Manager V10.60 client software is ideally suited for this purpose. In this section, we cover some of the basic monitoring functions.

7.7.1 Storage Manager Enterprise Management window (EMW)

This window is opened when we first launch Storage Manager. It provides a clear indication of the status of all DS4000 and DS5000 storage subsystems managed from this workstation. These units may be accessed either out-of-band (via an Ethernet connection) or in-band (via a Fibre Channel or iSCSI connection).

Figure 7-50 shows a typical EMW window highlighting which units are in an Optimal state and which require attention. Double-clicking any one of the units will launch the Subsystem Management window (SMW).

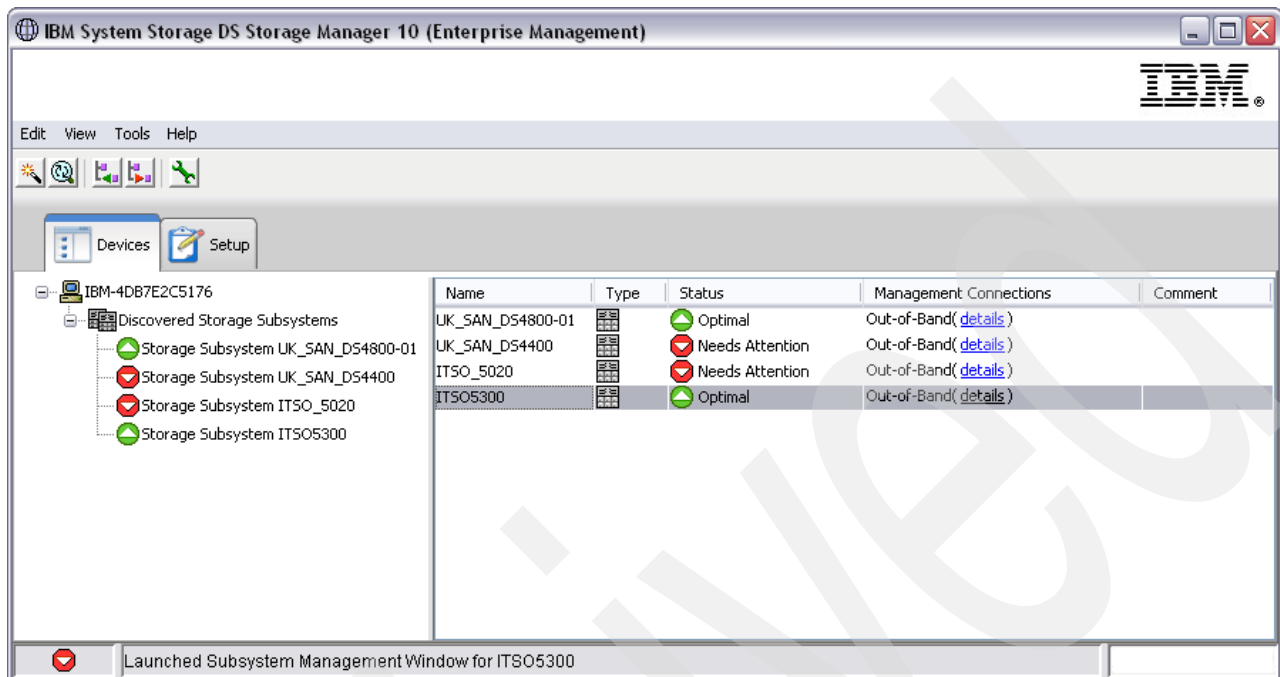


Figure 7-50 Storage Manager Enterprise Management window

Configuring e-mail or SNMP alert notifications in EMW

After you add devices to the management domain, you can set up alert notifications to report critical events on the storage subsystems. The following alert-notification options are available:

- ▶ Notification to a designated network management station (NMS) using Simple Network Management Protocol (SNMP) traps
- ▶ Notification to designated e-mail addresses

You can only monitor storage subsystems within the management domain. If you do not install the Event Monitor service, the Enterprise Management window must remain open. If you close the window, you will not receive any alert notifications from the managed storage subsystems. See the Enterprise Management window online help for additional information.

Alert notification with SNMP traps

To set up alert notification to a Network Management Station (NMS) using SNMP traps, perform the following steps:

1. Extract the downloaded Storage Manager image file onto an NMS. You need to set up the designated management station only once.
2. Copy the SMxx.x.MIB file from the SMxxMIB directory to the NMS.
3. Follow the steps required by your NMS to compile the management information base (MIB) file. (For details, contact your network administrator or see the documentation specific to your particular storage management product.)
4. Select **Storage subsystem** → **Edit** → **Configure alerts** from the Enterprise Management window and complete the entries in the SNMP tab.

Alert notification with e-mail

To set up alert notification without using SNMP traps, select **Storage subsystem** → **Edit** → **Configure alerts** from the Enterprise Management window.

First, select the **Mail Server** tab to define the SMTP server. Then select the **Email** tab to define the e-mail addresses in the notification list together with an option to select how much diagnostic information is to be sent in each alert notification e-mail.

7.7.2 Storage Manager Subsystem Management window (SMW)

The Storage Manager Subsystem Management interface includes a number of tools that can be used for monitoring, troubleshooting, and diagnostics:

- ▶ Subsystem Management GUI
- ▶ Storage Subsystem Profile
- ▶ Recovery Guru
- ▶ Major Event Log
- ▶ Collect all Support Data

Figure 7-51 shows some of the main visual problem indicators in the Subsystem Management window. Anyone using Storage Manager to monitor a DS4000 or DS5000 storage subsystem could immediately tell when the unit is in a non-optimal state.

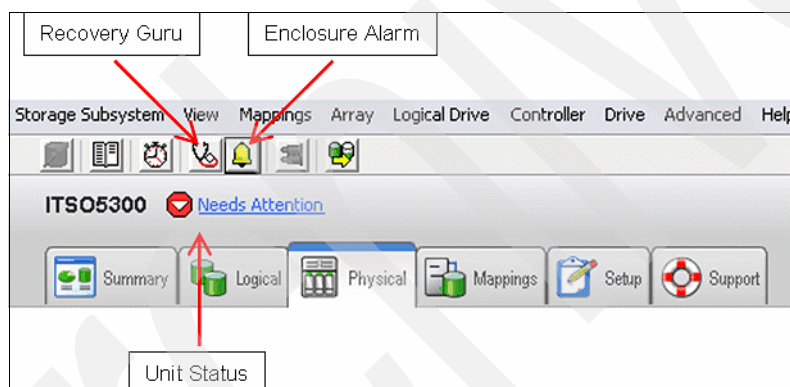


Figure 7-51 Storage Manager Subsystem Management window toolbar buttons

The Recovery Guru and Enclosure Alarm buttons blink when there is an error condition.

7.7.3 Storage subsystem profile

The storage subsystem profile provides a description of all of the components and properties of the storage subsystem. The profile viewer also allows the user to save the storage subsystem profile information to a text file. You might want to use the storage subsystem profile as an aid during recovery or as an overview of the current configuration of the storage subsystem. You might want to save a copy of the storage subsystem profile on the Storage Manager workstation and keep a hard copy with the DS4000 / DS5000 storage subsystem. Create a new copy of the storage subsystem profile if your configuration changes.

To open the storage subsystem profile, perform one of the following actions:

- ▶ Select **Storage Subsystem** → **View** → **Profile**.
- ▶ Select the **Summary** tab, and click **Storage Subsystem Profile** in the Status area.

The profile viewer provides tabs for navigating to a specific section of the profile. The information in the profile includes essential configuration and release code versions that might prove useful in troubleshooting.

The storage subsystem profile is included in the Collect all Support Data bundle.

7.7.4 Recovery Guru

The Recovery Guru is a component of the Subsystem Management window (SMW) that diagnoses problems and recommends recovery procedures to fix the problems. When a fault occurs, the user is notified with a flashing Recovery Guru button and the Unit Status changing to “Needs Attention”. By clicking either the “Needs Attention” or **Recovery Guru** button, we are presented with the Recovery Guru window (Figure 7-52) which shows:

- ▶ A summary view listing the faults detected
- ▶ A detailed view providing an expanded explanation of the fault
- ▶ A recovery procedure view that suggests some recovery actions

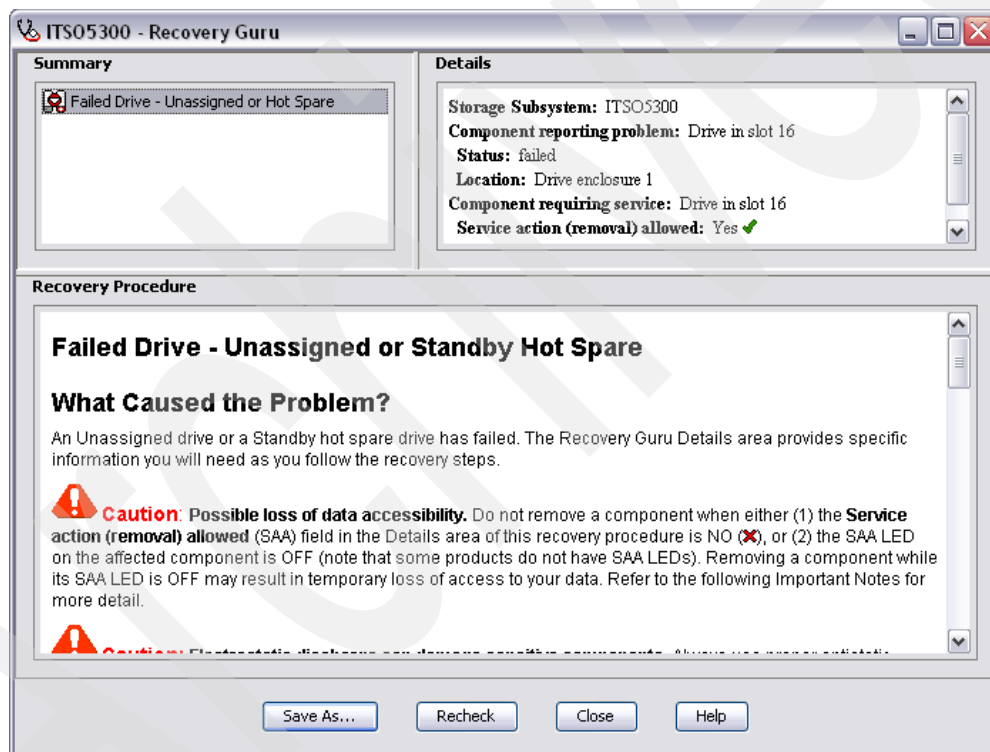


Figure 7-52 Recovery Guru window

In this example, the fault turned out to be a failed unassigned disk. The detailed view shows that the failed drive is in enclosure 1, slot 16 and that removal is allowed. The recovery procedure view provides a further explanation of the cause and the recommended recovery steps.

It is always advisable to review all the information in Recovery Guru before logging a support call.

The Recovery Guru status file is included in the Collect all Support Data bundle.

7.7.5 Major Event Log

The Major Event Log (MEL) is the primary source for troubleshooting a DS4000 or DS5000 storage subsystem. It provides a chronological trace of events logged from both controllers. To access the MEL, select **Advanced** → **Troubleshooting** → **View Event Log**. The MEL is shown in Figure 7-53.

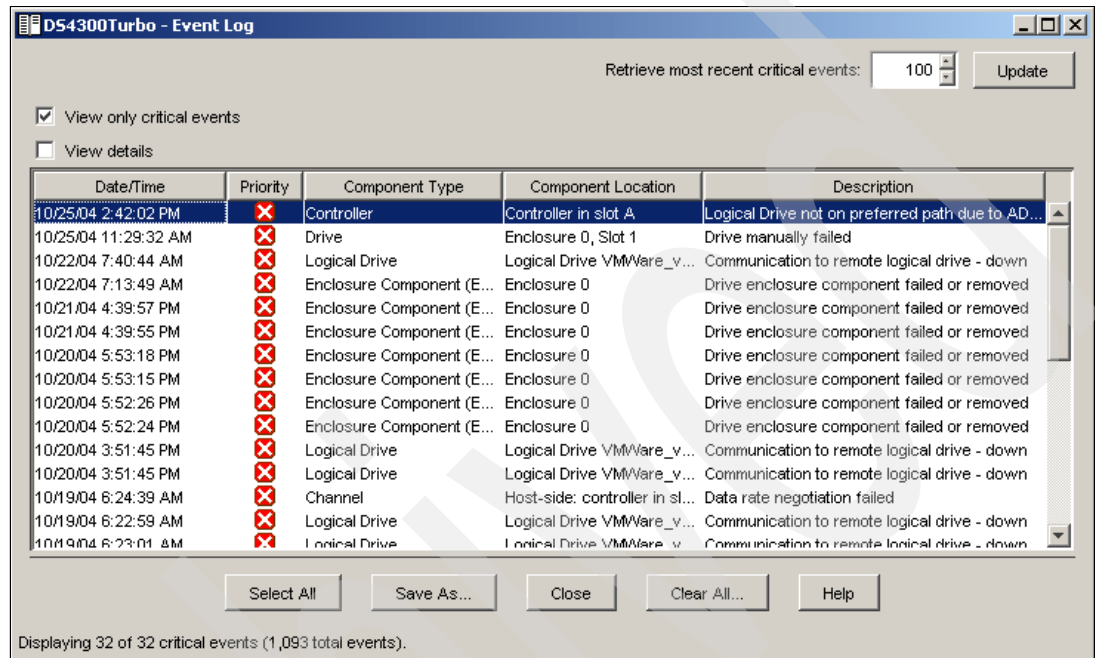


Figure 7-53 Major Event Log viewer

By default, only the 100 most recent critical events are shown, but this can be adjusted. The MEL holds up to 8192 events. It is best practice to capture a Collect all Support Data file as early as possible after an error occurs. Sometimes the error condition might generate a constant stream of events, so a delay could result in the MEL buffer wrapping around making it more difficult to determine the root cause of a fault. Enabling the Automatic Support Data Collection feature will ensure that this is captured whenever a critical event is logged.

There is a Clear All button to delete all the entries in the event log. Once a DS4000 or DS5000 storage subsystem is installed in a live environment, there is no reason to clear the MEL. The log will automatically wrap around when full. There is no benefit to reducing the number of events in MEL. However, clearing the MEL could result in the loss of vital information required in problem determination.

Important: Remember to configure the Event Monitor, and alerts will notify you in case of critical events as soon as they occur.

The MEL captures both critical errors and informational messages. Some of the informational messages, such as media scan start / stop events, will not normally be of any assistance during problem determination. The MEL viewer has a check box option for filtering the output:

View only critical events (default option) For a quick overview of all events that might affect the operational status of your DS4000 or DS5000. This is an effective way of filtering out only the critical events rather than looking through the entire log file.

View all events

Detailed information about all events logged by the controller. Sometimes this is useful when determining a time line of configuration changes or looking for marginal errors.

The MEL viewer also has a check box option called View details. When selected, this expands the window into two panes. The bottom pane will display additional data bytes associated with the event highlighted in the top pane. This information is also contained within the Collect all Support Data file, which might be required by the IBM Support representative during problem analysis.

7.7.6 Collect All Support Data option

Use the Collect All Support Data option to gather various types of inventory, status, and performance data that can help troubleshoot any problems with your storage subsystem. All of the files gathered are compressed into a single archive in a zipped-file format. Then, you can forward the archive file to an IBM Support representative for troubleshooting and further analysis. In most cases, this provides sufficient detail to help a hardware defect investigation, although sometimes switch logs, host logs, or cabling diagrams might also be required.

The data gathered in the Collect all Support Data bundle includes:

- ▶ The storage subsystem profile
- ▶ Major Event Log (MEL) information
- ▶ Read Link Status (RLS) data
- ▶ Nonvolatile static random access memory (NVSRAM) data
- ▶ Current problems and the associated Recovery Guru procedures
- ▶ Performance statistics for the entire storage subsystem
- ▶ Persistent registration and persistent reservation information
- ▶ Detailed information about the current status of the storage subsystem
- ▶ Disk drive diagnostic data
- ▶ A recovery profile for the storage subsystem
- ▶ Unreadable sectors detected on the storage subsystem
- ▶ Controller state capture data
- ▶ ESM state capture data
- ▶ The switch-on-a-chip (SOC) statistics
- ▶ Drive-side cabling connection summary

Automatic support data collection

Storage Manager incorporates an option to enable automatic support data collection. When enabled, a support data file is collected and transferred to a specified directory whenever a critical event occurs within a 72 hour period. This allows all information relevant for troubleshooting by your support representative to be preserved.

Tip: We recommend that you enable the Automatic Support Data Collection option in order to have a support data file automatically generated and saved to the specified location after the occurrence of a critical event. Make sure that:

- ▶ You specify a directory outside your DS4000 system to collect the information.
- ▶ The Event Monitor process is running on the workstation or host where you indicate to collect the logs.

To enable automatic captures, perform the following steps:

1. Select **Advanced** → **Troubleshooting** → **Support Data** → **Automatic Settings**.
2. Check the check mark to enable the automatic data collection and specify the destination folder, as shown in Figure 7-54.

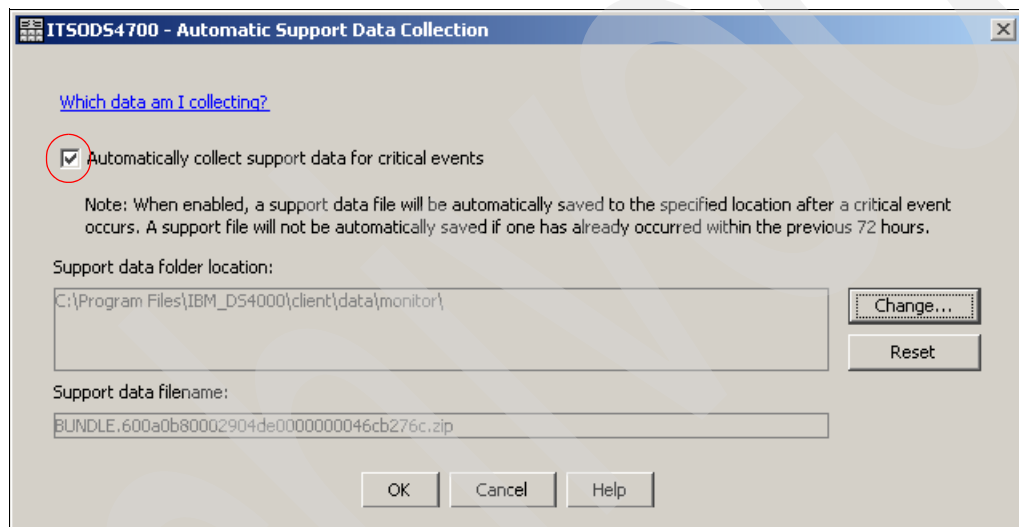


Figure 7-54 Enabling Automatic Data Collection

Running Collect All Support Data from Storage Manager GUI

In addition to automatically collecting support data, you can manually generate a new collection at any time from the Storage Manager by performing these steps:

1. Select **Advanced** → **Troubleshooting** → **Support Data** → **Collect**.

2. The Collect All Support Data window (Figure 7-55) opens. Specify the name and location of the zip file that you want to create and then click **Start** to begin the collection.

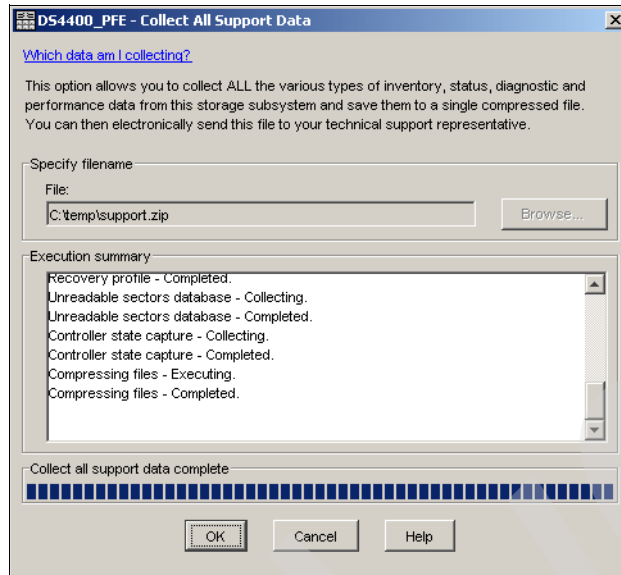


Figure 7-55 Collect All Support Data

A progress bar and execution summary is displayed throughout the data collection period. The compressed file size will typically be around 2-3 MB, depending on the configuration, making it convenient to send to IBM Support as an e-mail attachment.

Running collect all support data from script editor or SMcli

There is an alternate way to manually generate the Collect All Support Data bundle through the script editor or SMcli. The command to run is:

```
save storageSubsystem supportData file="c:\temp\filename.zip";
```

7.7.7 Media Scan

Media Scan is a background process that runs on all logical drives in the storage subsystem for which it has been enabled. A media scan provides error detection on the disk drive media. The Media Scan process scans all logical drive data to verify that it can be accessed. This is enabled by default when creating a new logical drive. There is also an option to scan the logical drive redundancy data during the Media Scan cycle.

Media Scan is not available for SSD drives. The menu option is grayed out for logical drives with SSD drives.

To modify the Media Scan settings, you need to be in the logical view in the Storage Manager Subsystem Management window. Then either right-click a logical drive or select **Logical Drive** → **Change** → **Media Scan Settings** to bring up the Change Media Scan Settings window shown in Figure 7-56.

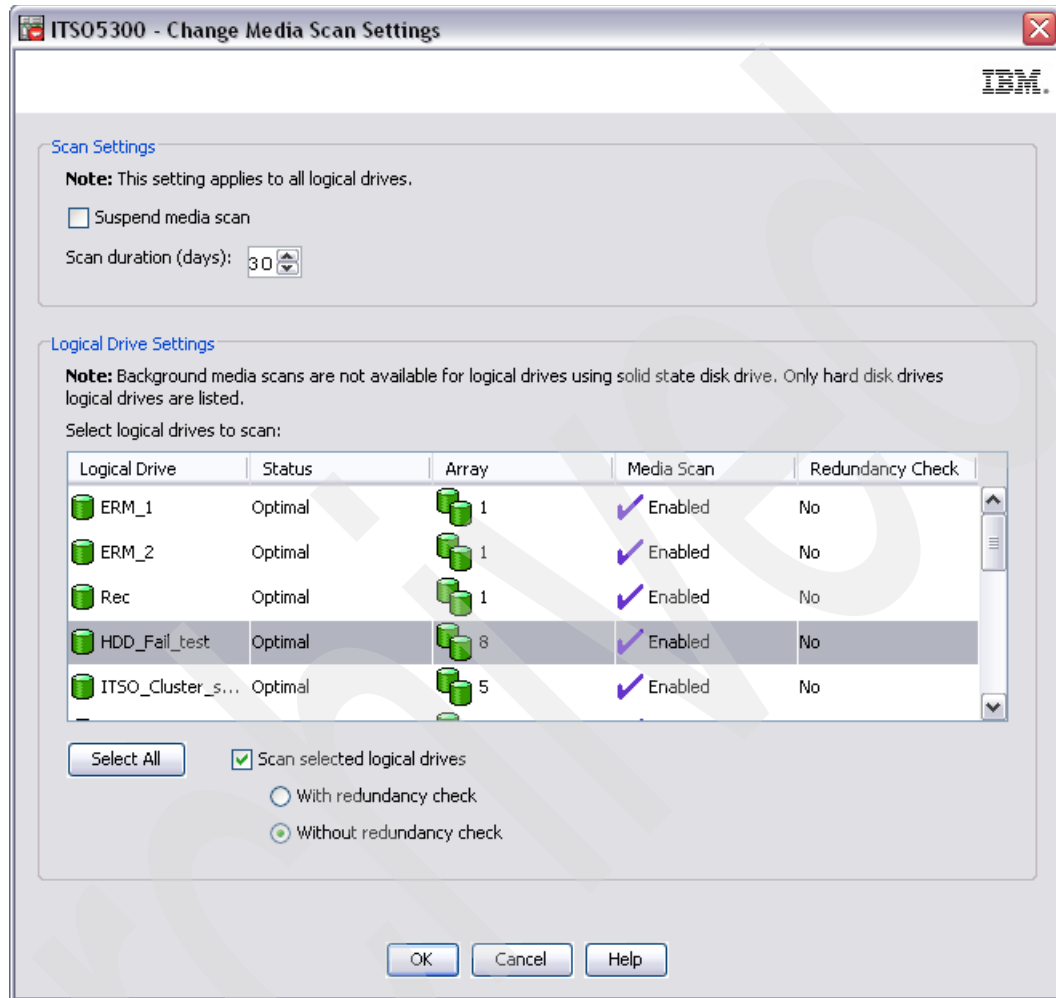


Figure 7-56 Media Scan settings

The options displayed at the top, Suspend media scan and Scan duration (days), are global settings. The logical drive settings at the bottom of the window apply to individual logical drives, although there is an option to set all logical drives identically by clicking the **Select All** button.

To specify the duration (in days) of the media scan, select a number in the Scan duration (days) box. The media scan duration specifies the number of days over which the media scan can run on the eligible logical drives. The controller uses the duration period, with its knowledge of which logical drives must be scanned, to determine a constant rate at which to perform media scan activities. This rate is maintained regardless of host I/O activity. Choosing a higher value for the duration will result in the media scan running for a longer period as a lower priority task.

A redundancy check scans the blocks in a RAID 3, 5, or 6 logical drive and checks the redundancy information for each block. A redundancy check compares data blocks on RAID 1 mirrored disk drives. RAID 0 logical drives have no data redundancy. There might be a performance impact when redundancy check is enabled.

7.7.8 Pre-read redundancy check

The pre-read redundancy check verifies the redundancy information during every read I/O. A logical drive that has this feature enabled returns read errors if the data is determined to be inconsistent by the controller. You can enable this option for logical drives that contain redundancy information, that is, RAID 1, 3, 5, and 6. However, there is a severe performance impact with this feature enabled, so it is only enabled in exceptional cases.

7.8 Problem determination

In this section, we explore some of the tools and methods used to diagnose:

- ▶ Drive-side problems
- ▶ Host-side problems
- ▶ Storage Manager communication problems

We only include techniques for determining faults using the Storage Manager and SMcli interface. Additional tools are available through the controller shell, although these are outside the scope of this book. A password is required for accessing the shell through either the serial port or a telnet session. It is only intended to be used by the IBM Support representative during problem determination.

However, there is a Capture State Information option in Storage Manager (also included within the Collect all Support Data bundle) that collects shell data from both controllers in the DS4000 or DS5000 storage subsystem. We will explore some of the included commands that might be relevant to fault finding.

With any suspected fault, it is essential to first rule out the possibility of a code mismatch or configuration error. This is particularly important when fault symptoms are detected following an installation or upgrade. Check that all prerequisites and configuration rules are adhered to before logging a hardware support call.

In many cases, Recovery Guru will assist in identifying cabling configuration errors with an explanation of the fault together with advice on correcting it, for example:

Channel miswired

Two or more drive channels are connected to the same Fibre Channel loop. The Recovery Guru Details area provides specific information that you will need as you follow the recovery steps.

Drive enclosures not cabled correctly

There are drive enclosures in the storage subsystem that are not cabled correctly because they have ESM canisters that must be cabled sequentially together. The Recovery Guru Details area provides specific information that you will need as you follow the recovery steps.

For limitations and version specific requirements, see the readme documents associated with the current code files, which can be found at the following address:

<http://www.storage.ibm.com/support>

For details of tested and supported configurations, see the System Storage Interoperation Center at the following address:

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

For cabling and configuration rules, see the following documents:

- ▶ *Installation, User's, and Maintenance Guide, GC26-7798*
- ▶ *Installation and Host Support Guide for DS Storage Manager v10.60, MIGR-5075652*
- ▶ *Installation and Migration Guide for Hard Drive and Storage Expansion Enclosure, MIGR-57818*

7.8.1 Diagnosing drive-side problems

In order to accurately diagnose drive-side problems, it is important to understand some of the differences between the various generations of DS4000 and DS5000 storage subsystems. Although the communication protocol is fundamentally unchanged in that they all use the Fibre Channel protocol for communication between the controllers and the back-end disks, the way it is implemented is different.

Disk speeds and capacity have increased, but they are all still arbitrated loop private devices. This protocol allows up to 127 devices to be attached on a single loop with each device being assigned a unique ALPA address. SATA disks include an interposer card to make them appear as dual channel Fibre Channel devices to the controllers.

The earlier models, such as the DS4300 storage subsystem and EXP-700 expansion enclosures, provide a physical connection between neighboring disks that daisy-chains them into a single loop. Fibre Channel frames are forwarded by each device in the chain until they reach their destination. This can result in bad frames being passed on to downstream devices by a faulty device. Therefore, we suspect the device immediately upstream of the first device to detect errors as the most likely culprit. The Read Link Status Diagnostics (RLSD) is the primary tool for analyzing drive-side errors in these configurations, as the error statistics are collected from the Fibre Channel interface on each disk, ESM, and controller.

The current generation of DS4000 and DS5000 storage subsystems and expansion enclosures include a switch-on-chip (SOC) built in to each ESM and controller. This allows the Fibre Channel frames to be passed between the controllers and ESM and then forwarded directly to the destination disk device via a point-to-point link. This improves performance, adds error diagnostic capability at the SOC, and eliminates the risk of a single drive disrupting the loop by causing downstream devices to fail. In this chapter, we only focus on diagnosing problems in the switched (SOC) environment.

Figure 7-57 illustrates the drive-side connections for one redundant channel pair in a 16 enclosure configuration. On the DS5300, each drive channel is associated with two ports. There are four drive channels and eight associated ports per controller. One channel from each controller combines to form a *redundant drive channel pair*.

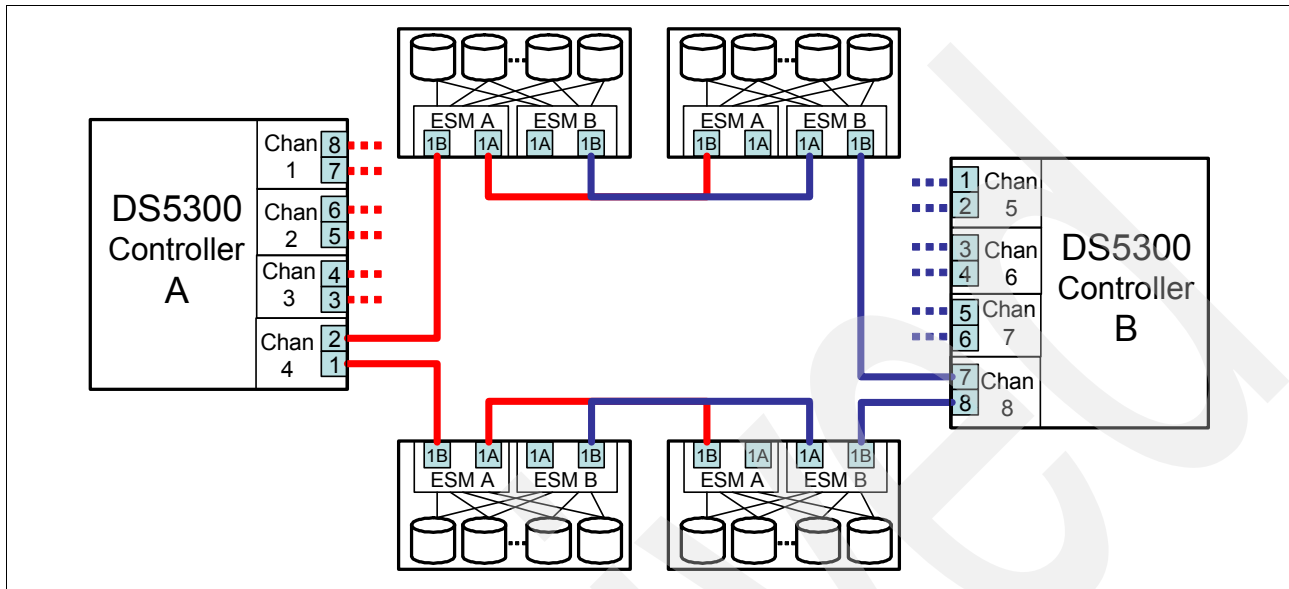


Figure 7-57 Drive-side connections

Degraded drive channel

The controller automatically fails a disk or ESM whenever an error threshold is exceeded. However, if there are excessive errors detected on a drive channel, then the controller will mark the channel as *degraded*. In this state, the channel is still available, but all I/Os are routed via the alternate channel in the redundant channel pair until the problem is identified and resolved. This is where we need to use available tools to determine which storage subsystem (enclosure) or component (cable, SFP, or ESM) on the channel is causing the excessive errors.

The main tools for diagnosing drive-side problems include:

- ▶ View Connections
- ▶ Storage Subsystem Profile
- ▶ Major Event Log (MEL)
- ▶ Read Link Status Diagnostics
- ▶ SOC Statistics
- ▶ Capture State Information

View Connections

Before starting any analysis of drive-side channel problems, it is essential to have a clear understanding of the cabling topology. By selecting **Storage Subsystem** → **View** → **Connections** from the Storage Manager Subsystem Management window, we are presented with a list of drive-side cable connections, as shown in Figure 7-58. This can be used to compile a topology diagram.

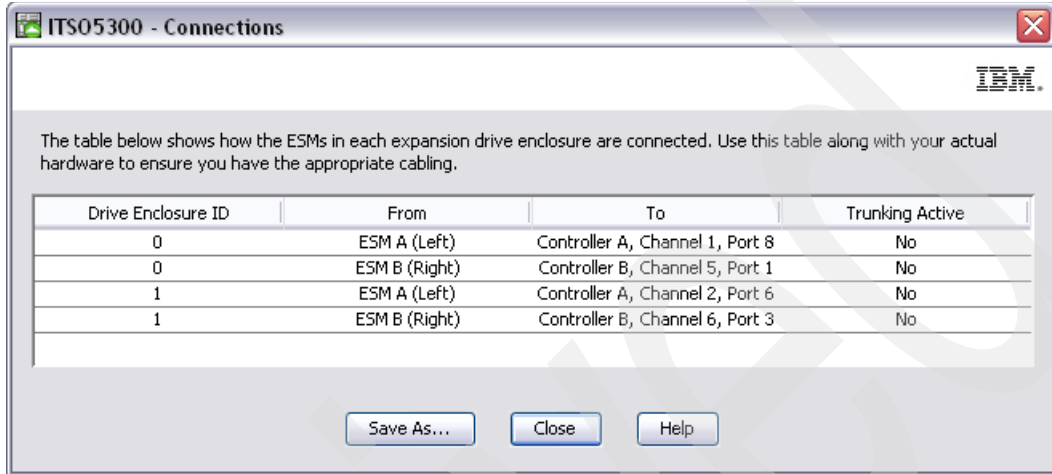


Figure 7-58 View Connections window

Storage Subsystem Profile

The Storage Subsystem Profile (Example 7-8) includes a section on drive channels. This provides a useful summary of status together with cumulative errors accumulated on each drive channel since the last controller reset (or since the counters were cleared).

Example 7-8 Storage Subsystem Profile

DRIVE CHANNEL 1

Port: 8, 7, ESM A 1B
Status: Degraded
Max. Rate: 4 Gbps
Current Rate: 4 Gbps
Rate Control: Auto
Controller A link status: Up
Controller B link status: Up
Trunking active: No

DRIVE COUNTS

Total # of attached drives: 16
Connected to: Controller A
Attached drives: 16
Drive enclosure: 0 (16 drives)

CUMULATIVE ERROR COUNTS

Controller A

Baseline time set: 23/09/09 15:38:31
Sample period (hh:mm:ss): 00:11:07
Controller detected errors: 0
Drive detected errors: 23160
Timeout errors: 0
Link down errors: N/A
Total I/O count: 6512

Controller B

Baseline time set: 10/09/09 12:12:52
Sample period (days, hh:mm:ss): 13 days, 03:37:27
Controller detected errors: 2
Drive detected errors: 1026598
Timeout errors: 0
Link down errors: N/A
Total I/O count: 65525656

The same information is available in the Storage Manager Subsystem Management window by selecting **Advanced** → **Troubleshooting** → **Drive Channels**, as shown in Figure 7-59.

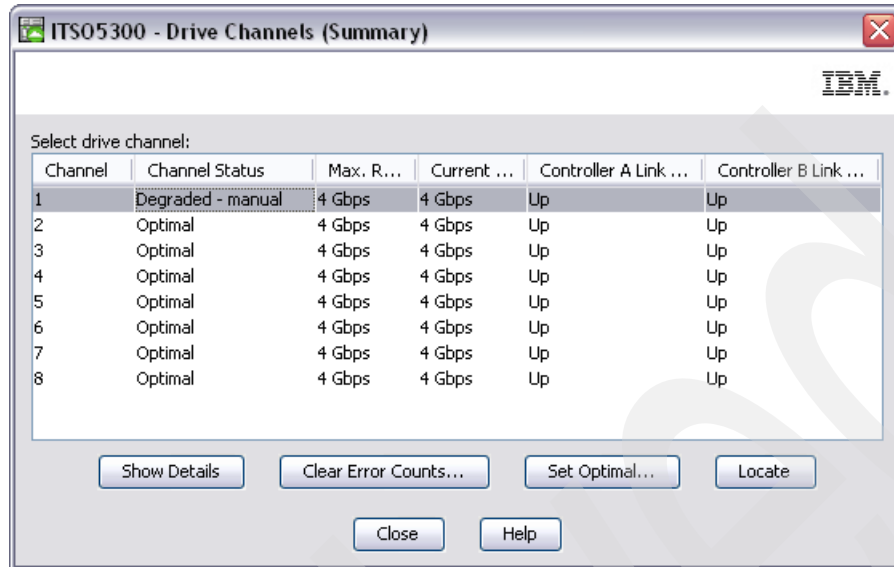


Figure 7-59 Drive Channels window

Clicking the **Show Details** button provides the cumulative error statistics for the selected channel. There is also an option to toggle the drive channel state between Optimal and Degraded. This action is not necessary, as the channel automatically returns to an Optimal state when the problem device is identified and excluded.

The Clear Error Counters option can be used to reset the counters. A sampling time of around 24 hours is normally sufficient to show any errors on a degraded drive-side loop. In some cases, the pattern of errors will begin to emerge within minutes of resetting the counters. If the counters have not been reset for an extended period, then the values might be historic rather than reflecting the current status.

Major Event Log (MEL)

The MEL is always likely to hold some clues to drive-side problems. Quite often, we see events logged against a faulty disk well before the error threshold is exceeded to mark it as failed. If there is a pattern of repeated errors against a single device, then it is worth removing it as a first step in diagnosing a degraded channel problem. It is also worth checking when the errors first started, as this might point to some other activity that was in progress at the time.

Read Link Status Diagnostics

The Read Link Status error counts refer to link errors that have been detected in the traffic flow of a Fibre Channel drive-side loop. The Read Link Status Diagnostics dialog retrieves the error counts and shows the controllers, disk drives, ESMs, and Fibre Channel ports in channel order.

By analyzing the error counts retrieved, you can determine the components on a drive-side channel that might be experiencing problems communicating with the other storage subsystems on the same channel. A high error count for a particular component might indicate that it is experiencing problems.

To run Read Link Status Diagnostics, select **Advanced** → **Troubleshooting** → **Run Diagnostics** → **Read Link Status**. The Read Link Status Diagnostics window will appear, as shown in Figure 7-60 on page 411.

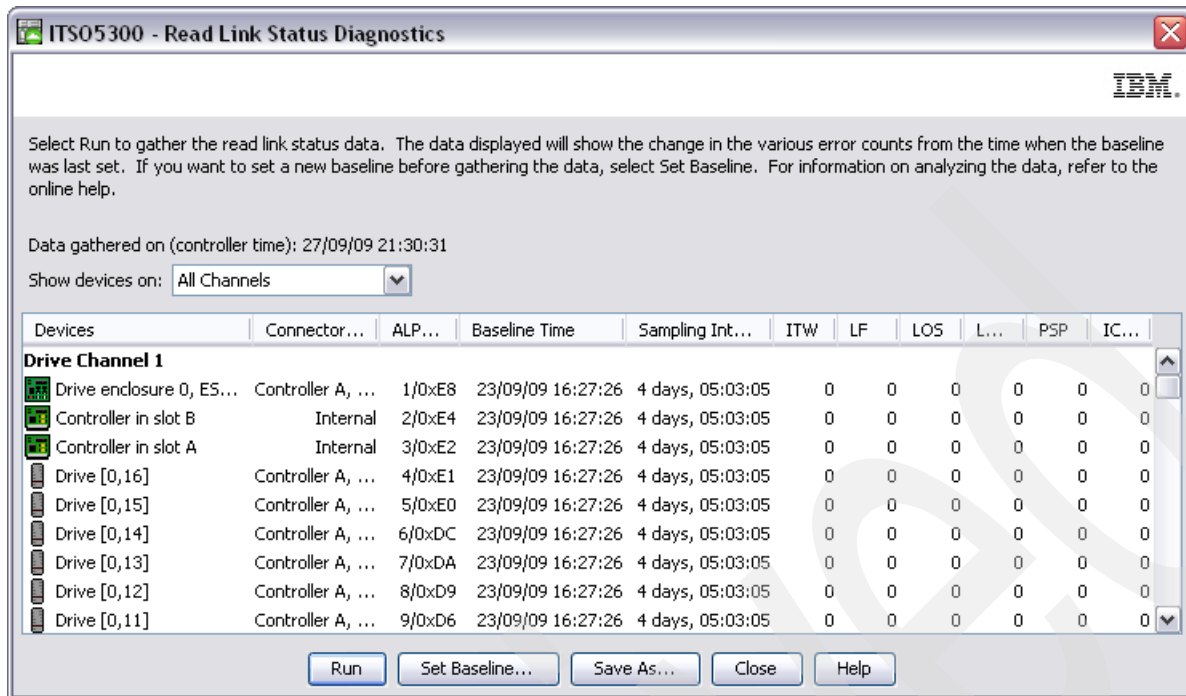


Figure 7-60 Read Link Status Diagnostics window

Error counts are calculated from a baseline. The baseline can be manually reset by pressing the **Set Baseline** button.

The columns displayed in the Read Link Status Diagnostics output are:

- Devices** A list of all of the storage subsystems on the Fibre Channel loop. The storage subsystems appear in channel order, and within each channel they are sorted according to the storage subsystem's position within the loop.
- Controller/Port** The controller ID or the port ID of the device.
- ALPA ID** The arbitrated loop physical address of the device.
- Baseline Time** The date and time of when the baseline was last set.
- Sampling Interval** 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 (ITW) errors detected on the Fibre Channel loop from the baseline time to the current date and time. ITW might be referred to as the "Received Bad Character Count."
- LF** The total number of link failure (LF) errors detected on the Fibre Channel loop from the baseline time to the current date and time.
- LOS** The total number of loss of synchronization (LOS) errors detected on the Fibre Channel loop from the baseline time to the current date and time.
- LOSG** The total number of loss of signal (LOSG) errors detected on the Fibre Channel loop from the baseline time to the current date and time.
- PSP** The total number of primitive sequence protocol (PSP) errors detected on the Fibre Channel loop from the baseline time to the current date and time.

ICRC The total number of invalid cyclic redundancy check (ICRC) errors detected on the Fibre Channel loop from the baseline time to the current date and time.

A sampling time of around 24 hours is normally sufficient to show any errors on a degraded drive-side loop. In some cases, the pattern of errors will begin to emerge within minutes of resetting the baseline. If the baseline has not been reset for an extended period, then the values might be historic rather than reflecting the current status.

ITW is the key error count to be used when analyzing the error count data. In a switched (SOC) ESM environment, a high ITW error count is likely to indicate a problem on the associated device.

The Read Link Status file can be saved as a comma separated file. It is included within the Collect all Support Data bundle.

SOC Statistics File (socStatistics.csv)

While RLSD error statistics are measured at the Fibre Channel port on the actual device, that is, the controller port, ESM port or disk, the SOC statistics are measured by the Switch-on-Chip (SOC) chip. Therefore, interpreting the differences between the two sets of output can be vital to accurately determine the root cause of a degraded drive-side channel.

The SOC statistics file can be generated with the following script command:

```
save storageSubsystem SOCCounts file="c:\socStatistics.csv";
```

It is also included within the Collect all Support Data bundle.

The columns displayed in the socStatistics file include:

OPM This is the Operating Port Mode (OPM). Valid states include non-cascade, tree, or string.

PS The state of the port (PS). Valid values could be inserted, loopback, unknown, or various bypassed states.

PIC Port Insertion Count (PIC) is the number of times the device has been inserted into this port.

LS Loop State (LS) is the condition of the loop between the SOC and component. Possible states include up, down, or various transition states.

LUC Loop Up Count (LUC) is the number of times the loop has changed from Down to Up.

CRCEC CRC Error Count (CRCEC) is the number of Cyclic Redundancy Check errors that are detected in frames.

RFDEA Relative Frequency Drift Error Average (RFDEA) is the difference between the port-received data rate and the internal clock of the SOC. A value in the 1,000s indicates a problem device.

LCC Loop Cycle Count (LCC) is the number of LIPs seen by the reference port.

OSEC Ordered Set Error Count (OSEC) is the number of invalid FC transmit words seen at the receiver of the port.

PCAC Port Connections Attempted Count (PCAC) is the number of times the port attempted to make a connection due to ARB connection requests.

PCHOC	Port Connections Held Off Count (PCHOC) is the number of times the port attempted to make a connection but it was held off by a busy port.
PUP	Port Utilization Percentage (PUP) is the percentage of time that frames are seen on the port, or percentage of time that a port is used if in switching mode.

During loop initialization (LIP), the SOC chips temporarily change to hub mode to allow all devices to see each other in an environment resembling a simple arbitrated loop topology. This skews some of the error statistics. Therefore, they can only be treated as reliable after a period of normal activity when no devices have been added, removed, or changed on the drive-side channels.

Capture State Information

This is a collection of shell commands executed on each controller. It can be started from the Storage Manager Subsystem Management window by selecting **Advanced** → **Troubleshooting** → **Capture State Information**. A large uncompressed text file is generated with a .dmp file name suffix.

This file is also included in the Collect all Support Data bundle.

Each command appears twice within the Capture State Information file, once on each controller. Each command is shown on a separate line with the following syntax:

Executing <command> on controller <A or B>

The file includes a few commands that might be useful for investigating drive-side problems:

fcDump	An incrementing number of fc exchange errors suggests that there is a problematic device on the loop. This is often a quick and easy check that can be carried out after excluding a suspect device from the loop.
chAll	Similar output to fcDump, just in a slightly different format.
luAll	This command provides a summary of logical unit information. For our purposes, the important information is in the ORP and Channels columns. ORP represents Operation, Redundancy and Performance. These normally are +++ for all disk devices. The Channels column displays a + for the preferred path to the device and * for the alternate path.
showEnclosuresPage81	This provides data similar to the SOC statistics file with a few additional items.

Clearing all error counters

We have already mentioned that there are options in Storage Manager to reset RLSD baseline and clear drive channel error counters. A neater way is to reset both of these and the SOC statistics with a simple script that can be executed from the Storage Manager Enterprise Management window (select **Tools** → **Execute Script** to open this window):

```
//Clear Drive Channel Statistics
clear allDriveChannels stats;
//
//Reset Storage Subsystem SOC Baseline
reset storageSubsystem SOCBaseline;
//
//Reset RLS baseline
reset storageSubsystem RLSBaseline;
```

The IBM Support representative handling your case might request that this script be run and that a new Collect all Support Data file be captured the next day.

Multiple drive failures

We highly recommend logging an IBM Service call whenever multiple drives fail simultaneously. Sometimes, the root cause is clearly understood and we can be reasonably confident that there are no underlying hardware defects. For example, if there was an unexpected loss of power to an expansion enclosure, then this could result in multiple disks remaining in a failed state. Those arrays that only lost RAID redundancy will remain online, but in a degraded state, and reconstruction or copyback will start automatically when power is restored to the enclosure. This can be observed in the Storage Manager Subsystem Management window physical view. The missing disks first re-appear in a replaced state. The associated logical drives return to an optimal state when reconstruction is complete without any intervention.

Any arrays and logical drives where the outage resulted in a failed array remain in a failed state after power is restored to the enclosure. This applies if two or more disks in the same RAID 5 array reside in the missing enclosure. Before taking any recovery action, it is important to understand the order in which the disks failed. Ideally, the failed disks should be revived in the opposite order in which they failed. With a power failure affecting a single enclosure, we can sometimes assume that all disks failed simultaneously. However, if a disk was in a failed state prior to the power outage, then it could contain stale data and therefore needs to be excluded from the array during recovery. Failing to do so might result in data corruption. If there is any doubt, then the IBM Support representative can determine the order in which the disks failed by reviewing the MEL and shell data.

Recovery actions: Multiple disks failures

To recover after multiple disks fail, perform these steps:

1. Determine the order in which the disks failed.
2. Unassign any standby hotspare drives.
3. Revive each disk starting with the drive that failed last until the associated logical drives change from a Failed to Degraded state. With a RAID 5 array, this means that one disk still remains in a failed state.

The Revive option is available through the Storage Manager Subsystem Management window by highlighting the drive and then selecting **Advanced** → **Recovery** → **Revive Drive**.

In this case, reviving a failed disk results in it being returned to an Optimal state.

4. Reboot the host(s) or rescan for the previously missing LUNs.

5. Check the data. If possible, avoid making any changes on the volume until it is clear that there is no data corruption, that is, mount the file system as read-only and perform **fsck** or **chkdsk** without attempting to fix errors. If data appears to be corrupt, then contact your IBM Support representative before taking any further actions.
6. After confirming data integrity, we can reconstruct the remaining failed disk(s) in the array. When complete, the associated logical drives return to an Optimal state.

The Reconstruct option is available through the Storage Manager Subsystem Management window by highlighting the drive and then selecting **Advanced** → **Recovery** → **Reconstruct Drive**.

Reconstructing a failed disk results in data being regenerated on it from the remaining disks in the RAID array.

7. Reassign the hotspare drives.

Note: Extreme care should be taken when selecting the Revive Drive option on an assigned disk. The function behaves differently depending on whether another drive is sparing for the drive being revived. If a hotspare has taken over, then reviving the drive will force a copyback to start. If there is no hotspare in use, then it will return the disk into an Optimal state as a member of the array, irrespective of whether it contains valid data or not.

Sometimes, following a power or channel incident, some disks return to a Replaced state even though there is no reconstruction or copyback in progress. In this case, the disk needs to be changed to a Failed state before it can be revived or reconstructed.

To fail a disk through the Storage Manager Subsystem Management window, highlight the drive and then select **Advanced** → **Recovery** → **Fail Drive**.

Checking RAID redundancy

We have already mentioned earlier in the Preventative Maintenance section that Media Scan includes an option to perform a RAID redundancy check. By default, the redundancy check is normally disabled for all logical drives. If enabled, Media Scan schedules to complete the check as a low priority task within the preset duration period (default of 30 days). However, there is also an option to perform an immediate data redundancy check on an array from Storage Manager.

This can be started from the Logical view in the Storage Manager Storage Subsystem window. First, highlight the array to be tested in the left-hand pane and select **Advanced** → **Recovery** → **Check Array Redundancy**. Figure 7-61 shows the Check Redundancy window. The array test starts after clicking the **Start** button.

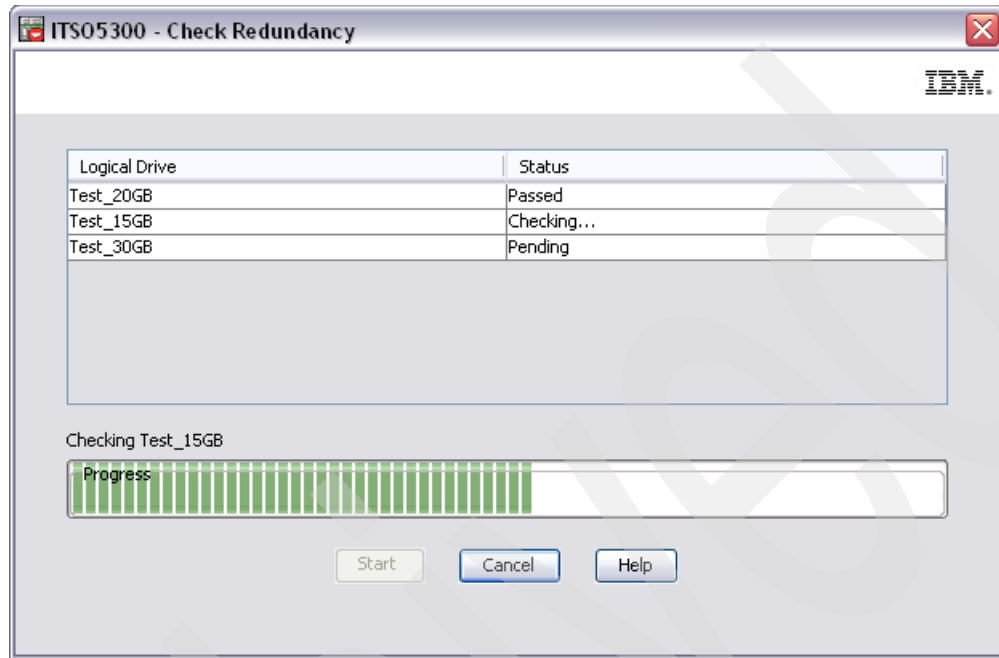


Figure 7-61 Check redundancy window

The utility will return a status for each logical drive associated with the array:

- Passed** The logical drive passed the redundancy check. No inconsistencies were detected in the redundancy information.
- Media error** The disk drive media is defective and is unreadable.
- Parity error** The parity is not what it should be for a given portion of the data.

If parity errors are detected, then it is important to determine the root cause. Contact your IBM Support representative for assistance.

A similar utility can also be run on an individual logical drive through the script editor in the Storage Manager Enterprise Management window. The command syntax is:

```
check logicalDrive [logicalDriveName] parity [parityErrorFile=filename]
[mediaErrorFile=filename] [priority=(highest | high | medium | low | lowest)]
[startingLBA=LBAvalue] [endingLBA=LBAvalue] [verbose=(TRUE | FALSE)];
```

Here is a typical usage example:

```
check logicalDrive ["LUN01"] parity parityErrorFile="c:\LUN01.parity.txt"
mediaErrorFile="c:\LUN01.media.txt" priority=high verbose=TRUE;
```

This runs the RAID redundancy check on the entire logical drive named LUN01 with parity errors being logged in the file c:\LUN01.parity.txt and any media errors in c:\LUN01.media.txt.

This command adds some flexibility to the redundancy check. It is never easy to predict how long the utility will take to complete and what impact it will have. Therefore, it is advisable to perform a quick test first by timing how long it takes to complete the check on the first 10 GB of data. The time to complete execution on the entire logical drive can then be estimated with some accuracy. If there are performance concerns, then the priority level can be reduced. The test command looks like this:

```
check logicalDrive ["LUN01"] parity parityErrorFile="c:\LUN01.parity.txt"
mediaErrorFile="c:\LUN01.media.txt" priority=low verbose=TRUE startingLBA=0
endingLBA=20000000;
```

Repairing RAID redundancy errors

If RAID parity errors are detected during execution of the logical drive parity check script, then there is another script command that can be run to fix them:

```
repair logicaldrive [logicalDriveName] parity parityErrorFile="filename"
[verbose=(TRUE | FALSE)];
```

So, our usage example looks like this:

```
repair logicalDrive ["LUN01"] parity parityErrorFile="c:\temp\LUN01.parity.txt"
verbose=TRUE;
```

This generates the following output in the bottom pane of the script editor while the sectors with invalid parity are being corrected:

```
Executing script...
Repair Logical Drive Parity called on logical drive: "LUN01" at LBA: 2,032
Repair Logical Drive Parity called on logical drive: "LUN01" at LBA: 2,033
Repair Logical Drive Parity called on logical drive: "LUN01" at LBA: 2,034
Repair Logical Drive Parity called on logical drive: "LUN01" at LBA: 2,035
Repair Logical Drive Parity called on logical drive: "LUN01" at LBA: 2,036
Repair Logical Drive Parity called on logical drive: "LUN01" at LBA: 2,037
Repair Logical Drive Parity called on logical drive: "LUN01" at LBA: 2,038
Repair Logical Drive Parity called on logical drive: "LUN01" at LBA: 2,039
Repair Logical Drive Parity called on logical drive: "LUN01" at LBA: 2,040
Script execution complete.
```

This command corrects the RAID parity information on the logical drive where errors were detected, but it has no way to determine which disk(s) have valid data and which are incorrect. Therefore, it is always safer to identify the root cause of the RAID parity corruption and repair it by forcing a reconstruction of data on the disk most likely to have contributed to the problem. The script command should only be used when the affected data area is known to occupy non-critical data or unused space.

Disk media errors

Over time, there will be a degree of deterioration on mechanical and Solid State Drive devices. The logic on the disks themselves will automatically re-map deteriorating sectors without the DS4000 or DS5000 controller even being aware. This extends the life and reliability of the media quite considerably. When enabled, the background Media Scan function forces a read of every sector in a logical drive over a preset period.

It is only when data on the sector becomes unreadable that the DS4000 or DS5000 controller must intercept by filling in the gaps using RAID redundancy data from the remaining disks in the array. If an error threshold is exceeded, then the disk is marked as failed and spun down awaiting replacement. At this point, if the array has RAID redundancy, there needs to be a reconstruction of data either to a standby hotspare drive, if available, or onto a replacement disk. With RAID 1 arrays, this involves a full copy of every sector for all logical drives from just the partner disk in the mirrored pair. With RAID 3 or 5 arrays, this involves a full copy of every sector for all logical drives from data and parity information about every remaining disk in the array.

Although extremely rare, it is nonetheless possible that an unreadable sector is detected on one of the source disks during this reconstruction. With RAID 3 and 5 arrays, the risk increases proportionally with the number of disks in the array. The type and capacity of the disks can also be a factor. RAID 6 arrays have two parity disks, so the potential exposure to this problem is statistically negligible.

When this occurs, there is no way to recover the data from the affected sector(s). The DS4000 or DS5000 controllers then remap the bad block without taking a prior copy of the data, while appending a log entry to the unreadable sector list. This allows the reconstruction process to complete, but data on the bad block is lost. Recovery Guru notifies the user of a non-optimal condition whenever there are any entries in the unreadable sector list. This list can be accessed from the Storage Manager Subsystem Management window by selecting **Advanced** → **Recovery** → **Unreadable Sectors**. By reviewing the logical LBAs in the list, it might be possible to determine whether the bad block(s) are in free space or a file. This could assist any decision on recovery action.

The DS4000 or DS5000 unit remains in a non-optimal condition until the Clear option is selected in the Unreadable Sectors window. However, make sure that a Collect All Support Data file is captured before doing so.

7.8.2 Diagnosing host-side problems

A significant proportion of support calls logged for host-side problems are found to be due to either misconfiguration or faults external to the DS4000 or DS5000. In this section, we look at some areas to check when attempting to identify host related problems.

Some of the more common host-side problems on the DS4000 or DS5000 include:

- ▶ Logical drive not on preferred path
- ▶ LUN bouncing
- ▶ Persistent reservations
- ▶ Target reset

Checking host configuration rules

The host configuration rules are described in detail in the *IBM System Storage DS Storage Manager Version 10 Installation and Host Support Guide*, GC53-1135. The latest operating system and code version specific updates are available in the readme files that accompany the Storage Manager download files. You can find these files at the following address:

<http://www.storage.ibm.com/support>

Supported configurations can be checked at the System Storage Interoperation Center at the following address:

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

Previously, there used to be a variety of host cable configuration rules and recommendations for the different implementations of the RDAC multipath driver. In most cases, these were restricted to a maximum of two paths. With MPIO and other native multipath drivers, there is a single highly resilient recommended configuration for Dual-HBA hosts attached through switched Fibre Channel SAN fabrics, as illustrated in Figure 7-62.

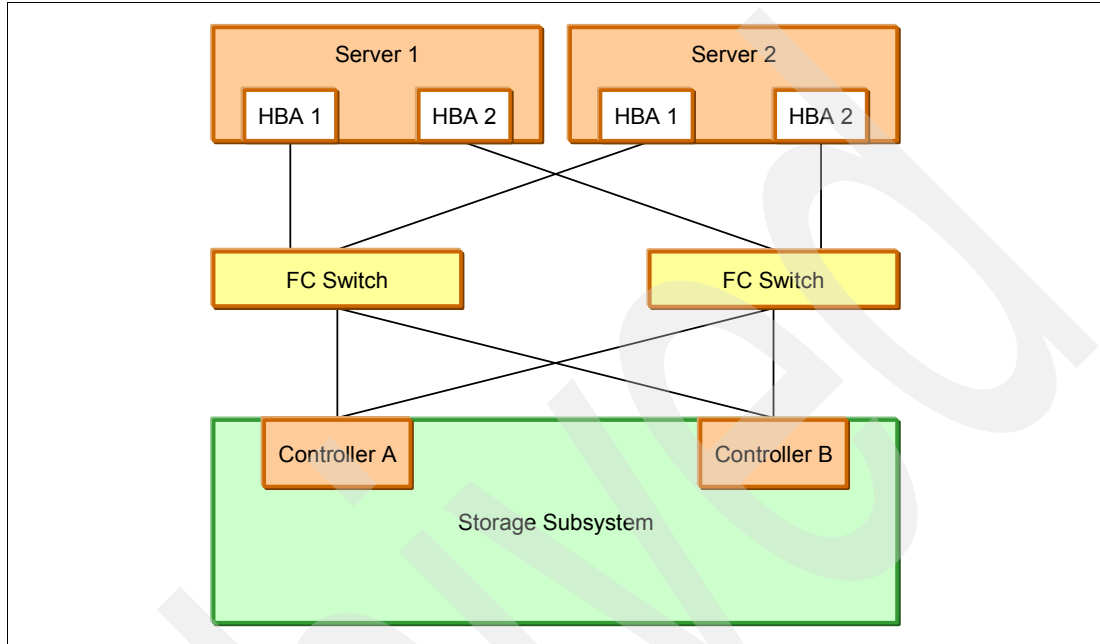


Figure 7-62 HBA to storage recommended configuration

Direct attached host configurations are permitted, although check the readme files for any operating system specific restrictions. Brocade HBAs do not support direct attached connections.

Single HBA configurations are permitted through a switch. The HBA must be zoned to have access to both controllers.

Logical drive not on preferred path

Each logical drive on the DS4000 or DS5000 storage subsystem is assigned to a preferred controller. The owning controller will service all I/O requests for this logical drive, while the alternate controller will act as a standby in case of failure. Whenever a logical drive failover occurs, a critical event is logged in MEL and Recovery Guru reports a non-optimal condition due to the logical drive not being on the preferred path. In the majority of cases, this does not indicate a fault on the DS4000 or DS5000 storage subsystem, but simply that the controllers behaved exactly as designed in transferring ownership when requested. Usually, the root cause of the failover is external to DS4000 or DS5000 storage subsystem.

Selecting the correct host type definition

The DS4000 or DS5000 storage subsystem automatically moves a logical drive onto the alternate controller for host types where Automatic Drive Transfer (ADT) is enabled whenever an I/O is received down the non-preferred path. Otherwise, it relies on the host multipath driver to issue a command (SCSI Mode Select 2C) to instruct the DS4000 or DS5000 storage subsystem when a failover is required. For these mechanisms to work, it is essential that the host type is set correctly to match the operating system and multipath driver on the host.

With controller firmware Version 07.60.xx, the host types shown in Table 7-2 are recognized.

Table 7-2 Host types

Host type	ADT status
AIX	Disabled
AIX-ADT/AVT	Enabled
DEFAULT	Disabled
HP-UX	Enabled
HPXTPGS	Disabled
IBM TS SAN VCE	Enabled
Irix	Disabled
LNXCLVMWARE	Disabled
Linux	Enabled
Linux_DMP	Disabled
NetWare Failover	Enabled
Solaris (with Veritas DMP)	Enabled
Solaris (with or without MPXIO)	Disabled
Unused1	Enabled
Windows 2000/Server 2003/Server 2008 Clustered	Disabled
Windows 2000/Server 2003/Server 2008 Clustered (supports DMP)	Enabled
Windows 2000/Server 2003/Server 2008 Non-Clustered	Disabled
Windows 2000/Server 2003/Server 2008 Non-Clustered (supports DMP)	Enabled

Note: There are two host type definitions for Windows 2000/Server 2003/Server 2008 Clustered and Windows 2000/Server 2003/Server 2008 Clustered Non-Clustered. The entries with the (supports DMP) suffix are only valid if VERITAS DMP is used as the multipath driver.

There are also scripts available to change the default ADT setting for specific host types. This might be required for SAN boot to work correctly. See the relevant documentation for usage instructions in *IBM Midrange System Storage Implementation and Best Practices Guide*, SG24-6363.

Missing path(s) to host

When multiple logical drives associated with different hosts (or host groups) suddenly failover onto their non-preferred controllers in the same direction, then the fault is likely to be closer to the DS4000 or DS5000 storage subsystem. In this case, the first place to check is the Recovery Guru summary of problems to see whether there are any other outstanding failures, such as failed SFP. If not, then take a look in the Major Event Log (MEL) for a possible controller reset or other host-side incidents at the time of failover. It is also worth checking the status and error count on the Fibre Channel switch ports that connect to the DS4000 or DS5000 storage subsystem. This could expose some marginal links to the host port(s) on the DS4000 or DS5000 controllers.

If logical drive(s) associated with a single host or host group fail over onto their non-preferred controllers while other logical drives remain unaffected, then the fault is likely to be closer to the host itself. In this instance, we need to focus just on the affected host(s) and paths to it.

If storage partitions are used, then the Storage Subsystem Profile and the host properties in the mapping section shows the WWPN of each HBA device in a host that the controllers are configured to see, as shown in Figure 7-63.

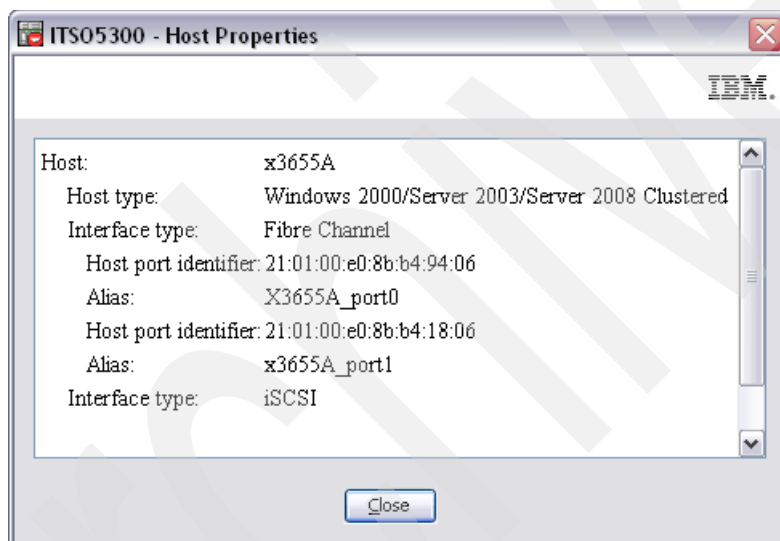


Figure 7-63 Host properties

Unfortunately, this does not provide a real-time view of the actual devices that are currently accessible. One way to check this is through the SAN switches. Here, we need to look at what devices are logged into the fabric and also what devices are zoned to see each other. This can sometimes be difficult with larger SAN configurations.

There is an alternate way to obtain a real-time view of devices accessible from each controller by analyzing the Capture State Information output. This is a text file containing low-level shell data from both controllers. The Capture State Information file can be generated from the Storage Manager Subsystem Management window by selecting **Advanced** → **Troubleshooting** → **Capture State Information**.

Use a text editor such as Notepad to view the Capture State Information file. The only command we need to check is **spmShow** from each controller. The sections we need to search for begin with:

Executing `spmShow(0,0,0,0,0,0,0,0,0,0)` on controller A

and

Executing `spmShow(0,0,0,0,0,0,0,0,0,0)` on controller B

From there, we need to page down to a subsection with the following heading:

---I-T-Nexus (PORT) Objects (ITN)---

This provides a list of initiator devices that are currently seen by this controller. The HBA ports can be recognized by either the alias name, which matches the alias defined in the Storage Manager mappings view or `Host_XXXXXXXXXXXXXXXX`, where `XXXXXXXXXXXXXXXX` represents the WWPN of the undefined HBA. For each Dual-HBA host configured in the recommended way shown in Figure 7-63 on page 421, we expect to see one entry for each HBA on each controller. The following output shows an example of how **spmShow** can be used to troubleshoot host-side problems:

Controller A `spmShow` output:

```
---I-T-Nexus (PORT) Objects (ITN)---
ITNID InitiatorPort                TargetPort                Online
-----
x0012 X3655A_port0                 FC_TargetPort_Ah_ch11    on
x0012 X3655A_port1                 FC_TargetPort_Ah_ch12    off
```

Controller B `spmShow` output:

```
---I-T-Nexus (PORT) Objects (ITN)---
ITNID InitiatorPort                TargetPort                Online
-----
x0012 X3655A_port0                 FC_TargetPort_Ah_ch11    on
x0012 X3655A_port1                 FC_TargetPort_Ah_ch12    off
```

The Online column shows a status of either on or off:

- on** Device is currently accessible from this controller.
- off** Device was recognized on this channel previously but is no longer accessible.

The output confirms that we need to focus on the HBA named `X3655A_port1` for the root cause of the logical drive moving onto the non-preferred path. This might be due to a zoning or configuration change, a hardware failure on the HBA card itself, or the fiber link to it. Checking the host logs and switch port status will reveal more information.

Note: This same technique can be used during initial configuration to confirm that all HBAs are zoned correctly. It is also good practice to simulate link failures during installation in order to verify that path failover is configured correctly on each host. Often, misconfiguration is only detected following a genuine failure when path failover does not behave as expected.

Managing a replaced HBA from Storage Manager

If an HBA is replaced in a host, then changes might need to be made in the switch zoning and Storage Manager host definition mappings. First, select the **Mappings** tab in the Storage Manager Subsystem Management window. Highlight the host with the new HBA in the left-hand pane and then select **Mappings** → **Manager Host Port Identifiers**. The Host Port Identifiers window opens, as shown in Figure 7-64.

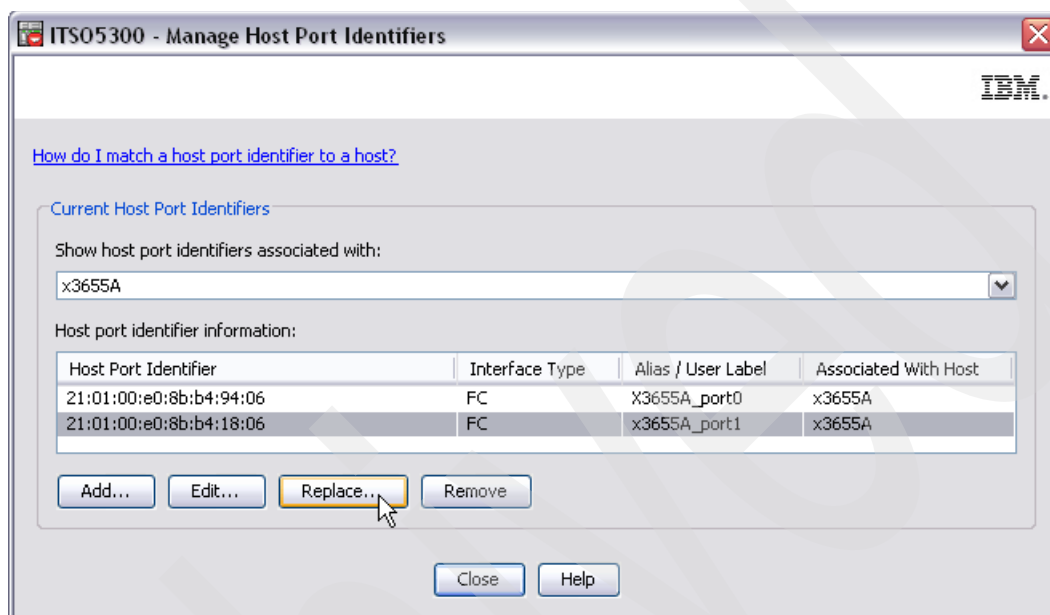


Figure 7-64 Host Port Identifiers window

By clicking the **Replace** button, we are given the option to select a new WWPN for this device from a drop-down list of accessible undefined devices.

Any logical drives might remain on the non-preferred controller while all available paths to the preferred controller are inaccessible.

Checking host configuration

If logical drives remain on the non-preferred path even though all paths have been confirmed to be accessible, then the focus must switch to the host itself. Here, we need to check that:

- ▶ Supported versions of the HBA driver, HBA BIOS, and multipath drivers are installed
- ▶ HBA parameters are set correctly
- ▶ Operating system parameters (that is, the Windows registry) are set correctly
- ▶ All operating system specific rules, limitations, and prerequisites are met

For hosts running SDDPCM/MPIO, the following commands are available for further troubleshooting:

```
pcmpath query adapter
pcmpath query adaptstats
pcmpath query device
pcmpath query devstats
pcmpath query essmap
pcmpath query portmap
pcmpath query wwpn
```

Returning logical drives back to their preferred paths

Once the underlying root cause of the logical drives switching to non-preferred path has been identified and resolved, then it is possible to return them back to their normal preferred paths. This can be done from the Storage Manager Subsystem Management window by selecting **Advanced** → **Recovery** → **Redistribute Logical Drives**. Figure 7-65 shows the Redistribute Logical Drives window with a progress bar that appears during the procedure.

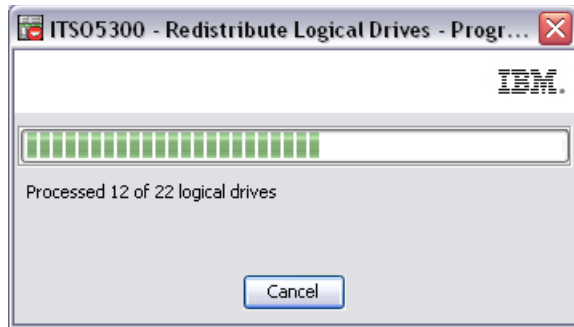


Figure 7-65 *Redistribute Logical Drives window*

When complete, the DS4000 or DS5000 unit returns to an Optimal state in Storage Manager. If any problems remain unresolved, the affected logical drives will fail back onto their non-preferred paths.

Logical drive bouncing

The constant transfer of logical drives back and forth between the two controllers can be caused by a number of reasons. It can result in a severe performance degradation and should be corrected. It is more likely to be a host configuration issue. In order to find evidence of logical drive bouncing, we need to check the Major Event Log (MEL), looking for a constant stream of non-critical event type 300D errors with a description of “Mode select for redundant controller page 2C received”, as shown in Figure 7-66.

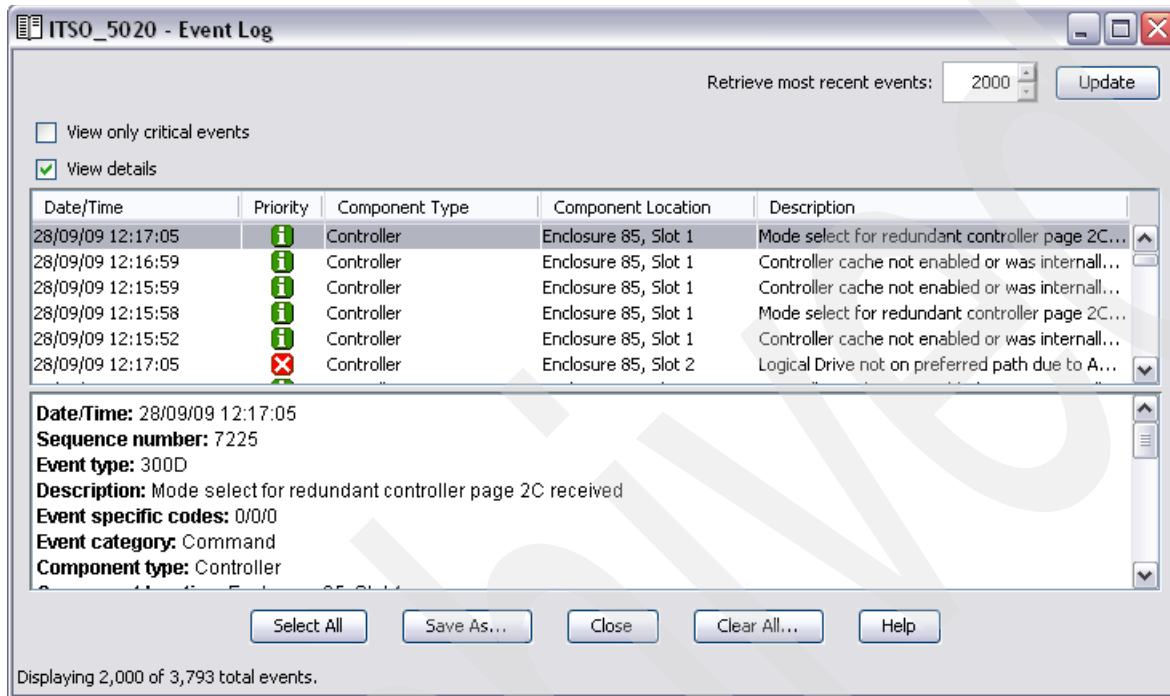


Figure 7-66 Mode Select 2C events

The DisableLunRebalance parameter

If the events appear exactly every 60 seconds, then it is possible that the DisableLunRebalance parameter is set incorrectly in an RDAC/MPP clustered environment. If a host in a cluster server configuration lost a physical path to a DS4000 or DS5000 storage subsystem controller, the logical drives that are mapped to the cluster group will periodically fail over and then fail back between cluster nodes until the failed path is restored. This behavior is the result of the automatic logical drive failback feature of the RDAC multipath driver. The cluster node with a failed path to a DS4000/DS5000 controller will issue a failover command of all logical drives that were mapped to the cluster group to the controller that it can access. After a programmed interval (normally 60 seconds), the nodes that did not have a failed path will issue **failback** command for the logical drives because they can access the logical drives both controllers, resulting in the cluster node with the failed path not being able to access certain logical drives. This cluster node will then issue a **failover** command for all logical drives, repeating the logical drives failover/failback cycle.

The workaround is to disable this automatic failback feature in all clustered configurations. In Windows, change the DisableLunRebalance registry setting of the [HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\rdacdisk\Parameters] registry key from 0 to 3 in each cluster node. Then, reboot each node for the changes to take effect.

Missing paths in a cluster environment

LUN bouncing can be caused by an incorrect configuration in which one cluster node can access the shared logical drives only through one controller and another node can only access the same shared logical drives through the alternate controller.

The missing paths can be identified using the same method described in “Missing path(s) to host” on page 421.

Persistent reservations

A host may issue a SCSI persistent reservation command to restrict which other HBA ports can access a particular LUN. These commands can be used by a server to prevent HBA ports in other servers from accessing the LUN and thereby prevent accidental data corruption caused by one server overwriting another server's data. “Reserve” and “Persistent Reserve” are often used by clustering software to control access to shared logical disks.

If a server is not shut down or removed from the server cluster in a controlled way, its reserves and persistent reserves can sometimes be left in place, preventing other servers from accessing data that is no longer in use by the server holding the reservation. In this situation, a storage administrator or server administrator might want to break the reservation and allow a new server to access the virtual disk.

In this situation, the safest thing to do is to have the server that owns the reservation explicitly release the reservation, as this ensures that the server concerned has flushed its caches and its software is aware that it has lost access to the disk. In circumstances where this is not possible, then most operating systems provide operating system specific tools to allow reservations to be removed. Consult the operating system documentation for details.

It is possible to view the Logical Drive Reservations and Logical Drive Registrations from the Storage Manager Subsystem Management window by selecting **Advanced** → **Maintenance** → **Persistent Reservations**. The dialog shows any logical drives in the storage subsystem that have registrations, with the first logical drive in the list highlighted by default. This table describes the information and buttons shown in the Persistent Reservation dialog:

Logical drive name	Shows the user label of the logical drive with persistent reservations. Logical drives are listed in alphabetical order. If a logical drive user label is not available, then its World Wide Identifier (WWID) appears.
LUN	Shows the assigned LUN number for the particular logical drive.
Registrations	Shows the number of registrations for the particular logical drive.
Reservation type	Shows an abbreviated form of the associated reservation type for the particular logical drive. Each addressable logical drive can have one reservation. Each reservation can grant access rights to one or more registrants, depending on the reservation type. You can reserve a logical drive for a specific access level by a group of registrants. All of the registrants within a group are restricted to the access level defined by the reservation type. This list shows reservation types as follows:
None	The logical drive has registrants, but it currently has no reservation.
WE	Write exclusive: Only the host port that reserved the logical drive may write to the logical drive. Reads shared: Any host port may read from the logical drive.
EA	Exclusive Access: Only the host port that reserved the logical drive may read from or write to the logical drive.

WE-RO	Write Exclusive - Registrants Only: Writes exclusive: All registered host ports may write to the logical drive. Reads shared: Any host port may read from the logical drive.
EA-RO	Exclusive Access - Registrants Only: Only registered host ports may read from or write to the logical drive.
WE-AR	Write Exclusive - All Registrants: Writes exclusive: All registered host ports may write to the logical drive. Reads shared: Any host port may read from the logical drive.
EA-AR	Exclusive Access - All Registrants: Only a host port may read from or write to the logical drive.
PTPL	Shows Yes if the registrations are set to persist through a power loss.

To view the registrations that are associated with the logical drive reservation, either double-click the desired logical drive, or highlight the logical drive and check the **View Associated Registrations** check box in the upper left of the dialog. The following information is then presented:

Interface type	Shows the type of interface: SATA, SAS, or Fibre.
Host (initiator) port	Shows the associated user label for the particular host port. If the host port alias has not been provided, then Not Available appears.
Associated host	Shows the host associated with the specific logical drive.
Controller (target) ports	Shows the port name of the target port.
Holds Reservation?	Shows either Yes or No, depending on whether the specific host port is the reservation holder.

Clearing persistent reservations

Sometimes it is not possible to clear the persistent reservation from the host. This is the case if a cluster host had been decommissioned without being shut down cleanly. Any persistent reservations prevent the associated logical drives from being deleted or changed in any way. There is an option to clear the logical drive reservations from the Storage Manager Subsystem Management window by selecting **Advanced** → **Maintenance** → **Persistent Reservations**.

1. In the upper-left corner of the Persistent Reservations window, make sure that the View Associated Registrations check box is cleared.
2. Click one or more desired logical drives. To select all of the logical drives, click **Select All**.
3. Click **Clear**.
4. In the text box in the Clear Registrations/Reservations dialog, type yes, and click **OK**. If you do not want to clear any reservations, click **Cancel** to return to the Persistent Reservations dialog.

The reservation and registrations that were associated with the logical drives that you highlighted in the upper pane are now cleared.

Target reset

It is normal to see TGT Reset events logged in the Major Event Log (MEL) during a host cluster failover. A support call will need to be logged for any unexplained target reset events.

7.8.3 Storage Manager communication problems

If Storage Manager shows the status of one of the DS4000 or DS5000 storage subsystems as unresponsive, then there are some recovery procedures to be followed.

In-band management

To recover the storage subsystems using in-band management, perform these steps:

1. Close all Storage Manager windows.
2. Check to make sure that the Access LUN is presented to the Storage Manager station. There is a utility provided called `SMDevices`, whose location and exact file name varies depending on the operating system. In Windows, the file is called `SMDevices.bat` and is located in the `Program Files\IBM DS4000\util` folder by default.

Use this utility to determine whether the Access LUN is visible. The default LUN number is 31. If the Access LUN is not mapped to the Storage Manager station, then in-band management will not be possible.

3. Restart the host agent software. On Windows 2003 and 2008 hosts, this can be done by performing these steps:
 - a. Select **Start** → **Administrative Tools** → **Services**. The Services window opens.
 - b. Right-click **IBM DS Storage Manager Agent**.
 - c. Click **Restart**. The IBM DS Storage Manager Agent stops and then starts again.
 - d. Close the Services window.
4. Restart the Storage Manager client.

Out-of-band management

To recover the storage subsystems using out-of-band management, perform these steps:

1. Ping both controllers from the Storage Manager station.
 - a. If the controllers respond to a ping, but Storage Manager remains unresponsive, then your firewall settings should be checked.
 - b. If the controllers do not respond to a ping, then there might be local network problems. Try using another mobile computer or host with a direct cable connection to one of the controllers. 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.
2. If the problem persists, then contact IBM Support for further assistance. A controller reset might be required.

7.9 Replacement and maintenance procedures

In this section, we discuss how to resolve some hardware failures on the DS4000 or DS5000 storage subsystem. See the *Installation, User's and Maintenance Guide* for your DS4000 or DS5000 storage subsystem for detailed parts replacement procedures. These publications can be downloaded from the Documents section on the IBM Storage Support Web site at the following address:

<http://www.storage.ibm.com/support>

7.9.1 Managing disk failures

The DS4000 and DS5000 controllers are constantly monitoring the status of the disk drives. Whenever an error threshold is exceeded, then the disk is marked as failed. This triggers the audible enclosure alarm to sound (unless disabled) and the subsystem appears in a non-optimal state. A critical event is logged in MEL and the Recovery Guru button starts flashing. All critical events are sent to the SNMP management console or to the e-mail recipient that you have configured to receive alert notifications by selecting **Edit** → **Configure Alerts** in the Enterprise Management window. The amber FAULT LED is illuminated on the faulty drive.

If the array has been configured with redundancy protection (RAID 1, 3, 5, or 6), then the drive failure will cause the array and associated logical drives to change to a degraded state. This indicates that the array has lost RAID redundancy. For RAID 1 or 6 arrays, this is only a partial loss of redundancy.

If a standby hotspare drive with the same (or greater) capacity and performance characteristics is available, then it takes over from the failed drive. Reconstruction of data onto the hotspare starts automatically. Once reconstruction of all associated logical drives is complete, the array returns to an Optimal state. At this point, the failed drive slot is still associated with the array. The hotspare drive remains assigned as a hotspare, but assumes a temporary association with the array.

Figure 7-67 shows both the physical and logical views in Storage Manager when a disk fails in a redundant array. There is a clock symbol next to one of the logical drives associated with the degraded array indicating that reconstruction is in progress on that logical drive. A progress bar is displayed when the logical drive is selected.

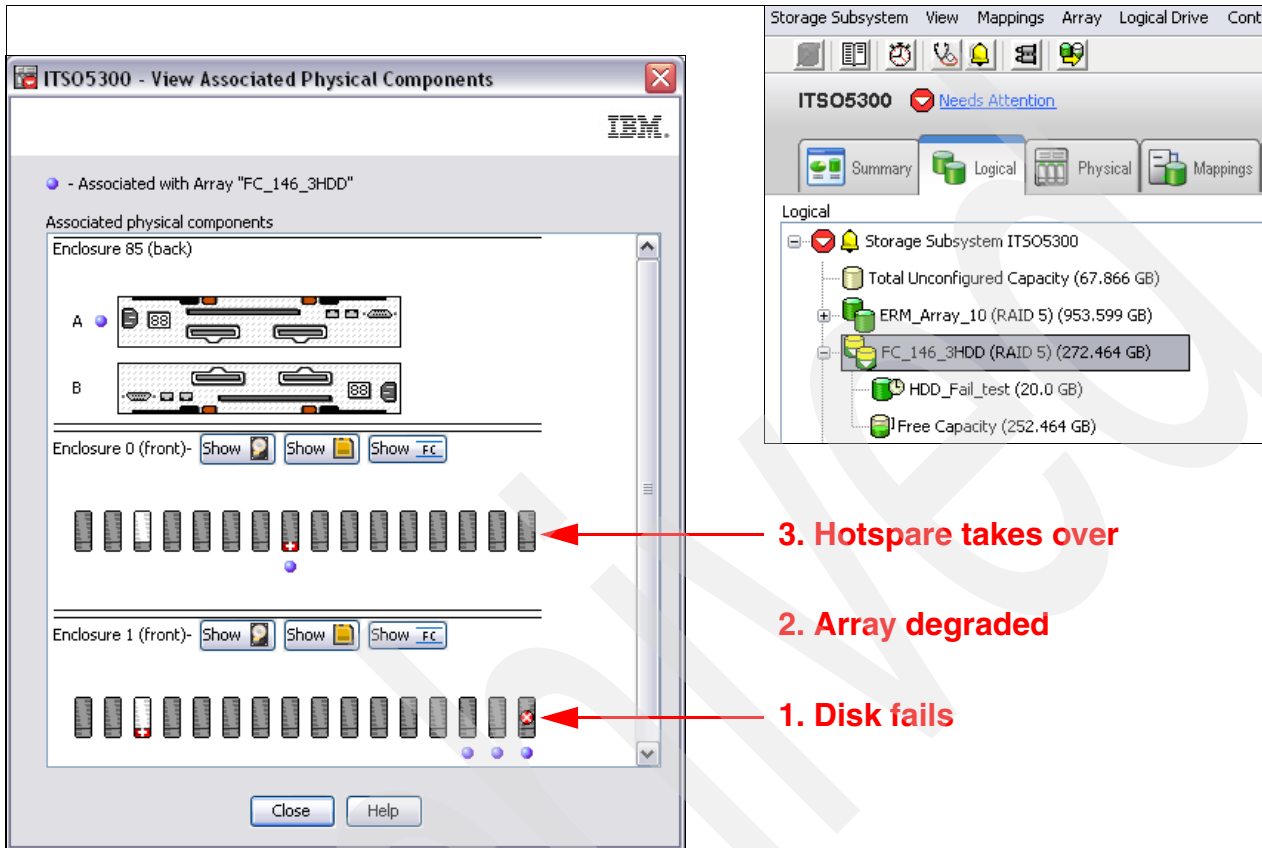


Figure 7-67 Storage Manager view after disk failure

At this point, we have a choice:

- ▶ The simplest and most common solution is to physically replace the failed disk. This results in an automatic copyback operation from the hotspare onto the new drive. When complete, the hotspare returns to an unassigned hotspare role.

- ▶ The alternate solution is to make another disk a permanent replacement for the failed drive. When we right-click the icon for the array with the failed disk in the logical view, we are given the Replace Drives option, as shown in Figure 7-68. This menu option is normally grayed out when all drives in the array are optimal.

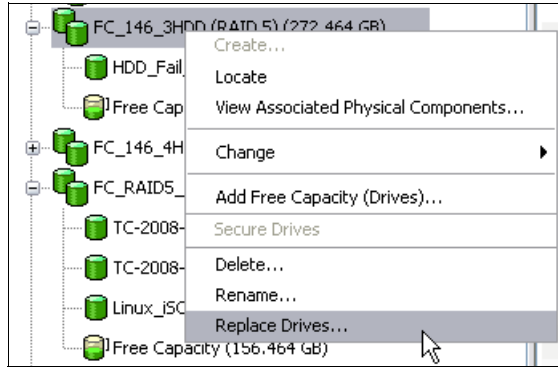


Figure 7-68 Replace Drives option in Storage Manager logical view

A new window appears showing the failed drive at the top and a list of potential replacement drives beneath, as shown in Figure 7-69. This allows us to make the hotspare drive or any other drive of equal capacity and type (FC, SATA, or SSD) a permanent replacement for the failed drive.

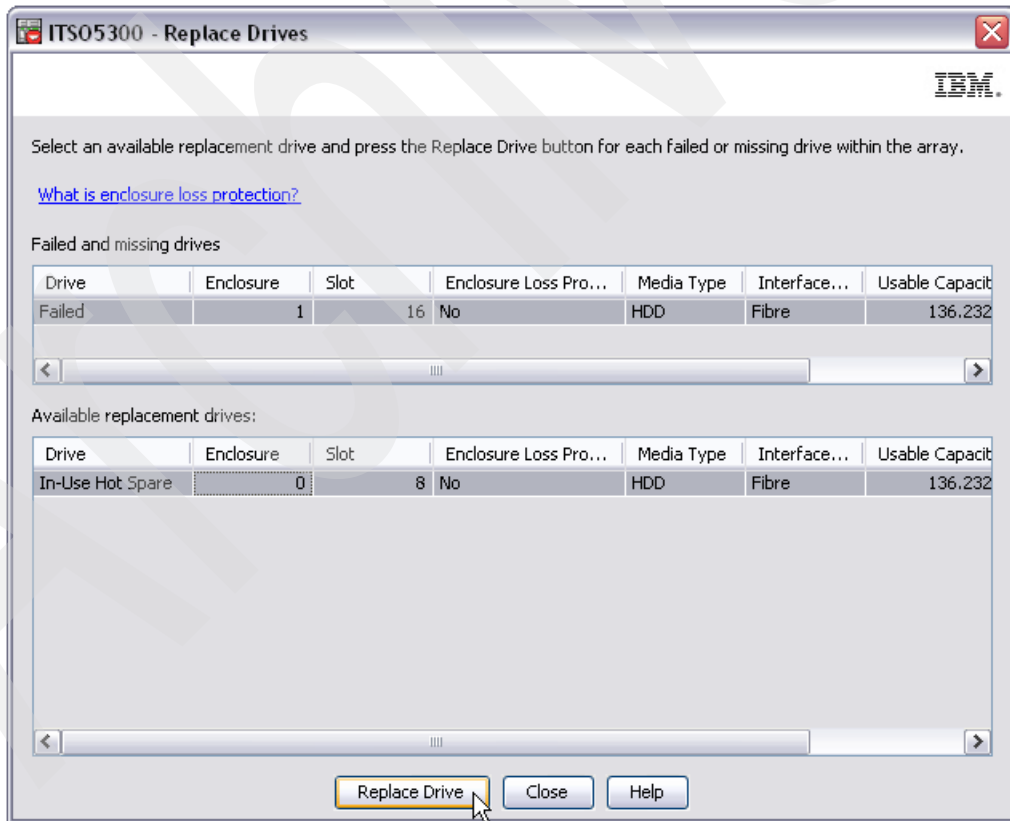


Figure 7-69 Replace Drives window

The straightforward option #1 is preferred in most cases, as all disks in the array remain in unchanged positions after the drive replacement. However, if drive positions were not an important factor during the initial planning, then option #2 might be considered.

The Replace Drives option can also be used as a concurrent method of migrating individual drives to new locations within the same storage subsystem. To do so, we need to ensure that the target drive is the only hotspare of the same capacity and type as the source. Then we can fail the source drive from Storage Manager by selecting it and navigating to **Advanced** → **Recovery** → **Fail drive**. When the target hotspare drive takes over, we can use the same procedure described in step #2 to make it the permanent replacement.

Disk replacement

The amber fault LED on the front of the drive indicates that it is in a powered down state and ready for replacement. To replace the disk, perform these steps:

1. Release the latch on the disk by pressing on the inside of the bottom of the tray handle.
2. Pull the tray handle out into the open position and slide the drive out.
3. Wait for at least 60 seconds before inserting the replacement drive. Gently push the new disk into the empty bay until the hinge of the tray handle latches beneath the storage subsystem enclosure bezel and then push the tray handle down into the closed (latched) position.

The amber fault LED on the front of the drive will flash while the drive is spinning up. When complete, the new drive appears in the Storage Manager physical view and copyback from hotspare starts automatically.

In cold climates, we recommend allowing the replacement disk to acclimatize within the drive slot for at least one hour before pushing it in fully. This reduces the risk of early life failures as the CRU drive is introduced into a controlled data center environment from a delivery vehicle. The sudden change in temperature and humidity can result in a buildup of condensation. The drive bays must never be left empty for an extended period, as this affects the internal airflow within the enclosure.

7.9.2 Managing disks with an impending drive failure error

The DS4000 and DS5000 controllers are constantly monitoring the status of the disk drives. A Predictive Failure Analysis (PFA) error is logged against the drive whenever a sufficient level of errors are detected and regarded as a concern yet the drive remains usable. This should be regarded as a warning that the drive is deteriorating and likely to fail in the near future. This triggers the audible enclosure alarm to sound (unless disabled) and the subsystem appears in a non-optimal state. A critical event is logged in MEL and the Recovery Guru button starts flashing. All critical events are sent to the SNMP management console or to the e-mail recipient that you have configured to receive alert notifications (you set these notifications by selecting **Edit** → **Configure Alerts** in the Enterprise Management window).

Recovery Guru reports three levels of impending drive failure:

- | | |
|------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Low risk PFA | This is when a PFA threshold is exceeded on an unassigned drive or standby hotspare drive. The suspect drive should be replaced whenever possible. |
| Medium risk PFA | When a PFA threshold is exceeded on a drive that is a member of a RAID 1, 3, 5, or 6 array. If the drive fails, then you might lose redundancy. The suspect drive should be replaced at the earliest opportunity. |

High risk PFA

When a PFA threshold is exceeded on a drive that is a member of an array where no more drives can fail without losing data. This is either a RAID 0 array or a degraded RAID 1, 3, 5, or 6 array. Immediate action should be taken to avoid data loss. We discuss some of the possible recovery actions later in this section.

When impending drive failure is detected, the affected disk remains powered and spinning.

With low or medium risk PFAs, the recovery actions are nondisruptive. The affected drive needs to be manually failed before it can be safely replaced. This is performed in the Storage Manager Subsystem Management window physical view by highlighting the affected drive and selecting **Advanced** → **Recovery** → **Fail drive**. Once in a failed state, the drive can be handled as a normal faulty drive with the procedure described in 7.9.1, “Managing disk failures” on page 429.

The recovery options for high risk PFAs are different. It is a good idea to back up all data on the affected logical drives and then proceed with the steps in either “PFA warning on a disk in a RAID 0 array” or “PFA warning on a disk in a degraded array” on page 434.

PFA warning on a disk in a RAID 0 array

An array is configured without redundancy (RAID 0) with the understanding that a single disk failure results in data loss. Only temporary or non-critical data should be stored on the associated logical drives. Therefore, the main PFA recovery action for RAID 0 arrays is a disruptive procedure with all associated LUNs being inaccessible while the affected drive is replaced and data restored.

Perform these steps:

1. Stop all I/O to the affected logical drives.
2. Volume Copy can be used as an alternative to tape backup and restore. This function is only available with the optional premium feature. If any of the affected logical drives are also source or target logical drives in a Volume Copy operation that is either Pending or In Progress, you must stop the copy operation before continuing. Go to the Copy Manager by selecting **Logical Drive** → **VolumeCopy** → **Copy Manager**, highlight each copy pair that contains an affected logical drive, and select **Copy** → **Stop**.
3. If you have FlashCopy logical drives associated with the affected logical drives, these FlashCopy logical drives will no longer be valid. Perform any necessary operations (such as backup) on the FlashCopy logical drives and then delete them.
4. Highlight the affected drive in the Physical View of the Subsystem Management window and select **Advanced** → **Recovery** → **Fail Drive**. The amber fault LED illuminates on the affected disk. The affected logical drives become Failed.
5. Replace the failed drive.
6. Highlight the array associated with the replaced drive in the Logical View of the Subsystem Management window and select **Advanced** → **Recovery** → **Initialize** → **Array**. The logical drives in the array are initialized, one at a time.

To monitor initialization progress for a logical drive, highlight the logical drive in the Logical View of the Subsystem Management window and select **Logical Drive** → **Properties**. Note that after the operation in progress has completed, the progress bar is no longer displayed in the Properties dialog.

When initialization is completed, all logical drives in the array have the Optimal status.

7. Use operating system tools to re-discover the initialized LUNs.

8. Restore data from backup or recreate any Volume Copy relationships by highlighting the copy pairs in the Copy Manager (select **Logical Drive** → **VolumeCopy** → **Copy Manager**) and selecting **Copy** → **Re-Copy**.

It might also be possible to add redundancy by changing the RAID level if sufficient spare drives are available. If successful, this alters the PFA risk level from high to medium, allowing the disk to be replaced without disruption. However, there will be data loss if the affected disk fails during this operation.

PFA warning on a disk in a degraded array

For the array to be in a degraded state, there must already be a failed disk when the PFA is detected on another disk in the same array. Reconstruction to a hotspare might already be in progress. Two replacement disks will be required. In this scenario, it is important to replace the failed disk as soon as possible by performing the following steps:

1. Although not required, I/O to the affected logical drives should be stopped to reduce the possibility of inducing a failure on the PFA disk before the failed disk is replaced.
2. If a standby hotspare drive is not available, replace the failed disk.
3. Monitor the progress of reconstruction on the affected logical drives or change the reconstruction rate by highlighting the logical drive in the Logical View of the Subsystem Management window and then selecting **Logical Drive** → **Properties**. Note that after the operation in progress has completed, the progress bar is no longer displayed in the Properties dialog.
4. When all affected logical drives have returned to the Optimal status, the PFA risk level reduces from high to medium. At this point, it is safe to resume I/O to the affected logical drives.
5. Highlight the PFA flagged drive in the Physical View of the Subsystem Management window and select **Advanced** → **Recovery** → **Fail Drive**. The amber fault LED for the affected disk illuminates. The affected logical drives become degraded until reconstruction is complete.
6. Replace the failed drive(s).

7.9.3 Monitoring Solid State Drives (SSD)

When we look at the SSD drive properties in the Storage Manager Subsystem Management physical view, we see some parameters that are unique to SSD drives (see Figure 7-70).



Media type:	 Solid State Disk
Interface type:	 Fibre Channel
Drive path redundancy:	OK
Wear life monitoring:	Enabled
Average erase count:	0%
Spare blocks remaining:	99%

Figure 7-70 SSD drive properties

A flash-based SSD has a limited wear life before individual memory locations can no longer reliably persist data. The disk drive continuously monitors itself and reports its wear life status to the controller. Two mechanisms exist to alert you that an SSD is nearing the end of its useful life: average erase count and spare blocks remaining. You can find these two pieces of information in the disk drive properties, which you can see in the storage management software by selecting a disk drive on the Physical tab.

The average erase count is reported as a percentage of the rated lifetime. When the average erase count reaches 80 percent, an informational event is logged to the Major Event Log (MEL). At this time, schedule the replacement of the SSD. When the average erase count reaches 90 percent, a critical event is logged, and a Needs Attention condition occurs. At this time, replace the SSD as soon as possible.

The spare blocks remaining are reported as a percentage of the total blocks. When the spare blocks remaining falls below 20 percent, an informational event is logged to the MEL. At this time, schedule the replacement of the SSD. When the spare blocks remaining falls below 10 percent, a critical event is logged, and a Needs Attention condition occurs. At this time, replace the SSD as soon as possible.

7.9.4 Managing battery issues

On DS5000 storage subsystems, the battery unit contains lithium-ion battery packs that can maintain power to the RAID controller caches for up to thirty minutes to flush cache memory to flash memory modules in the event of a power loss.

This battery unit provides backup power to each controller's cache memory. Each battery unit contains a sealed, rechargeable lithium-ion battery. The battery unit can maintain data in the cache for three days.

On both the DS4000 and DS5000 storage subsystems, the battery should not be replaced until it is marked failed by the controller. If the batteries are shown as expired, use the reset battery age function in the Storage Manager Subsystem Management window to reset the age.

7.10 Replacing adapters (HBA) and storage controllers

The logical volumes created in the DS5000 are mapped to the host's Fibre Channel adapter's worldwide name (WWN). Replacing an HBA affects the mappings in both the DS4000 and the SAN zoning configuration if the HBAs are not directly attached.

Consider the following items:

- ▶ Host adapter replacement

If for any reason an HBA is replaced and you are not using the default group to map your logical volumes in the DS5000, you will not see any disks through the new controller until the mappings are regenerated by replacing the old HBA WWPN with the new adapter WWPN.

Also, update your SAN zoning configuration after changing the HBA to allow the new WWPN to communicate with the target DS5000.

- ▶ DS5000 controller replacement

Under rare circumstances, it is possible that after replacing a failed DS5000 controller, the World Wide Port Names could be changed on the resident as well as the replaced controller. If a user has his Fibre Channel switch zoned by WWPN, this will cause a loss of access. The zoning configuration on Fibre Channel switches must be adjusted. It is also possible to see this exceptional condition after performing a DS5000 controller firmware upgrade.

There is a feature for automatic code synchronization of the controllers in case one is replaced. This ensures that both controllers execute the same level of firmware.

7.11 HBAs and operating system tools

This section provides practical information for problem determination on specific OS platforms and HBAs when you use Storage Manager utilities such as mppUtil and SMDevices, and other common operating system (OS) dependent commands used for reviewing disks status and data collection.

7.11.1 Brocade HBA and Brocade Host Configuration Manager (HCM)

This section gives a brief introduction to the Host Configuration Manager (HCM) for Brocade FC HBAs.

Brocade offers different software to manage their HBAs. A list of software bundles that are downloadable from the Web can be found at the following address:

http://www.brocade.com/sites/dotcom/services-support/drivers-downloads/HBA/HBA_IBM.page

The software bundles consist of the following items:

- ▶ Brocade Adapter Software installer
This package enables a single step installation for all software components, including Host Connectivity Manager (HCM), Drivers, Firmware, Agent, Brocade Command Utility (BCU), and APIs.
- ▶ Driver
This is a single driver package (per OS and server platform) for supporting all Brocade HBAs. The drivers will be packaged appropriately for each operating system.
- ▶ Firmware
The adapter firmware is bundled as part of the driver, and will be automatically updated after the driver is updated.
- ▶ Multi-boot code image
A multi-boot code image (BIOS/EFI) allows an adapter to be plugged into a server to support boot from SAN functionality for x86, IEM64T, AMD64, and IA64 server platforms. Server administrators can set up a bootable LUN in the SAN through an easy to use configuration utility menu.
- ▶ Driver Update Disks (DUDs)
DUDs (provided in ISO and zip format) are needed to install the drivers during an OS installation of a LUN attached to the SAN.

- ▶ LiveCD

The LiveCD can be used to boot up diskless or OS-less servers. The Brocade Command Utility (BCU) can then be used to update the boot code.
- ▶ Agent

The management agent is automatically installed as part of the driver installation process and can be started manually or automatically. This agent is required to manage Brocade HBAs through HCM.
- ▶ Host Connectivity Manager (HCM)

HCM is the Brocade Adapter management tool that has an intuitive and easy-to-use graphical user interface (GUI). This is a Java based application and can run on standard servers and workstations or a dedicated management server. HCM is used to install, configure, and manage local as well as remote adapters from a single interface. In addition, data center administrators can use the tool for detailed configuration tasks, driver and firmware upgrades, device level monitoring, and comprehensive diagnostics.
- ▶ Brocade Command Line Utility (BCU)

Brocade also includes the Brocade BCU, which is a command-line utility used to configure and manage local HBAs from the console. Many of the GUI configuration options are also available through the BCU. Run the `bcu -help` command to obtain more information about BCU options.
- ▶ APIs

SNIA HBA API V2.0 is supported. For both Windows and Solaris, the SNIA HBA API libraries are part of the OS. For Linux and VMware, separate API libraries are provided as part of the driver package.

Using HCM

The minimum set of requirements to support HCM include:

- ▶ Brocade FC HBAs (BR-415, BR-425, BR-815, and BR-825)
- ▶ A single-processor or multiprocessor server or workstation
- ▶ Pentium® III with 450 MHz (or equivalent) or greater for Windows, Red Hat, Novell, Solaris x86, and VMware.
- ▶ Sun Ultra 60 for Solaris SPARC.
- ▶ At least 256 MB of physical RAM; 512 MB is recommended.
- ▶ A video card capable of at least 256 colors and a screen resolution of 800 x 600 pixels.
- ▶ At least 150 MB of disk space.
- ▶ Internet Explorer (5.0 or later), Netscape Communicator (5.0 or later), or Firefox (1.0 or greater)
- ▶ TCP/IP protocol stack for communicating with the management agents.

After installing Brocade HCM, you can start it from the Windows Start menu.

HCM is secured by user and password, as shown in Figure 7-71. The default user is *administrator* with the default password of *password*. After a successful login, the window shown in Figure 7-72 opens, which is the HCM main window.

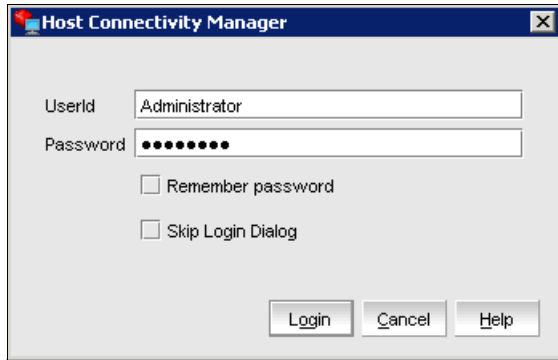


Figure 7-71 Agent login window

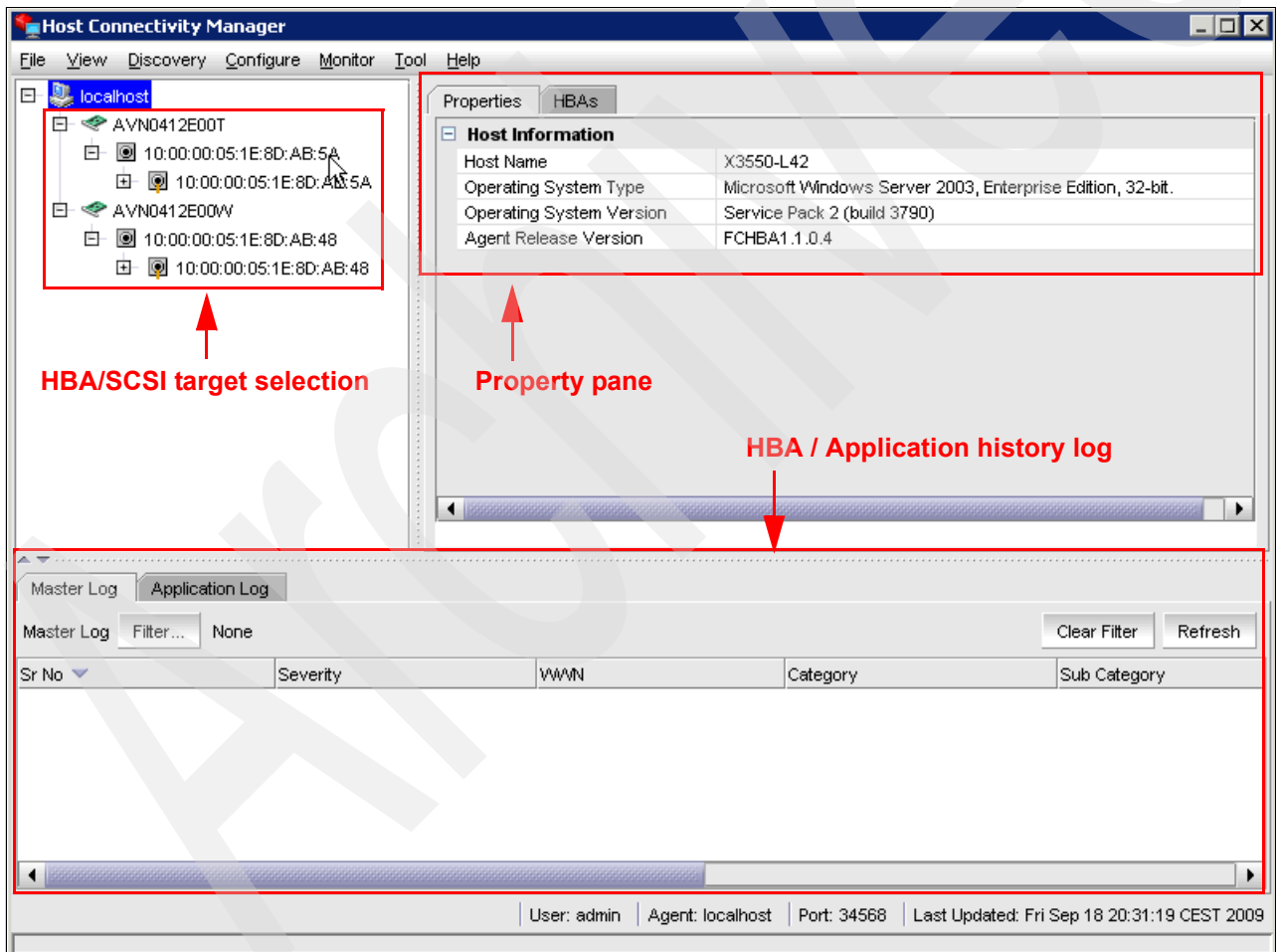


Figure 7-72 HCM main window

Figure 7-72 on page 438 shows, in the left pane, all the adapters and all ports, as well as all current connected targets with their available LUNs. Upon selecting a HBA, port, or target, the right pane will show the corresponding properties for that object. The bottom pane displays a change history of the configuration done by the user. By going through the menu items, you can perform various actions on the adapter and ports and retrieve much information from the targets.

With HCM, you can:

- ▶ Perform HBA firmware update (also known as Boot Code Image) (Figure 7-73)
- ▶ Perform data collection (Support Save) and HBA settings backup (Figure 7-74)
- ▶ Configure HBA/port settings, persistent binding, and diagnostics (Figure 7-75 on page 440)

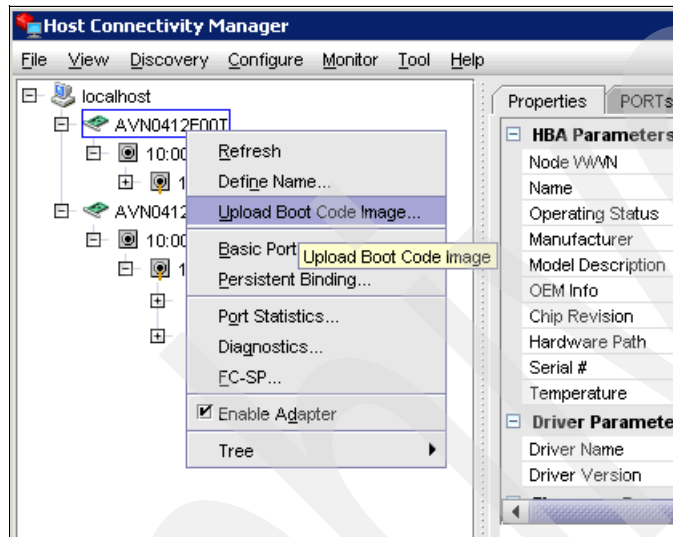


Figure 7-73 HCM Update HBA firmware

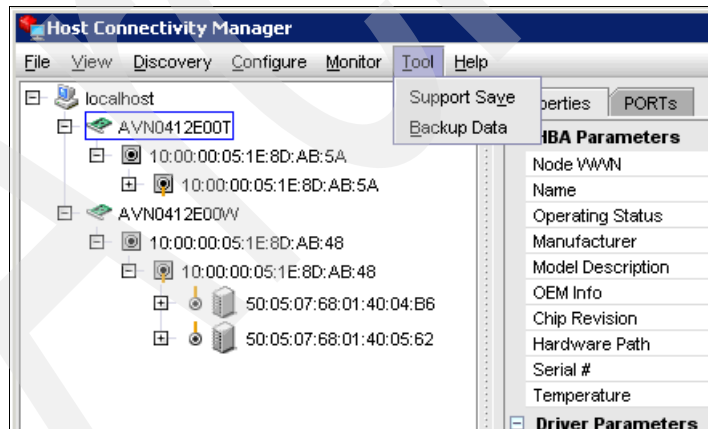


Figure 7-74 HCM Tool menu

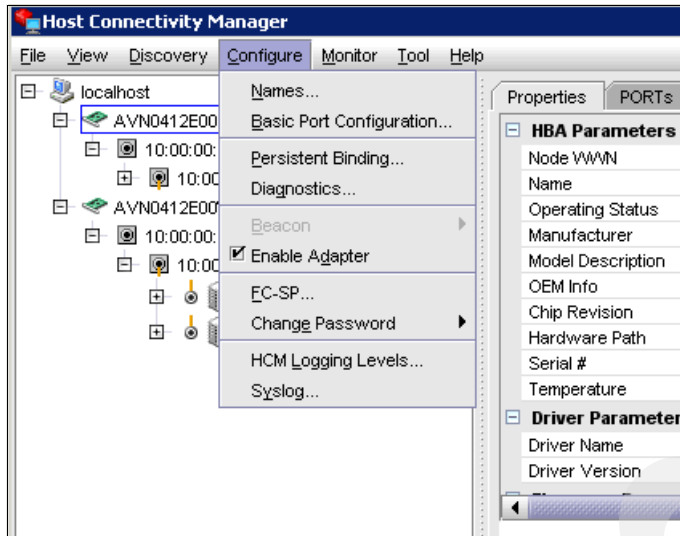


Figure 7-75 HCM Configure menu

Data collection

You can collect a variety of information about installed Brocade adapters, such as the firmware version installed, operational status, port speed, WWN, PCI data, configuration data, flash status, and other details, in order to troubleshoot the use of BCU commands, HCM menu options, and host operating system commands. This function is named Support Save.

Support Save

The Support Save feature is an important tool for collecting debug information from the driver, internal libraries, and firmware. You can save this information to the local file system and send it to support personnel for further investigation.

Use one of the following options to launch this feature:

- ▶ In HCM, launch Support Save from the Tools menu (see Figure 7-76).
- ▶ Using the Brocade Command Line Utility (BCU), run the `bfa_supportsave` command.
- ▶ Using your Internet browser (Internet Explorer 6 or later or Firefox 2.0 or later), you can collect `bfa_supportsave` output if you do not have root access, do not have access to file transfer methods such as FTP and SCP, or do not have access to the Host Configuration Manager (HCM).
- ▶ A `bfa_supportsave` collection can also occur automatically for a port crash event.

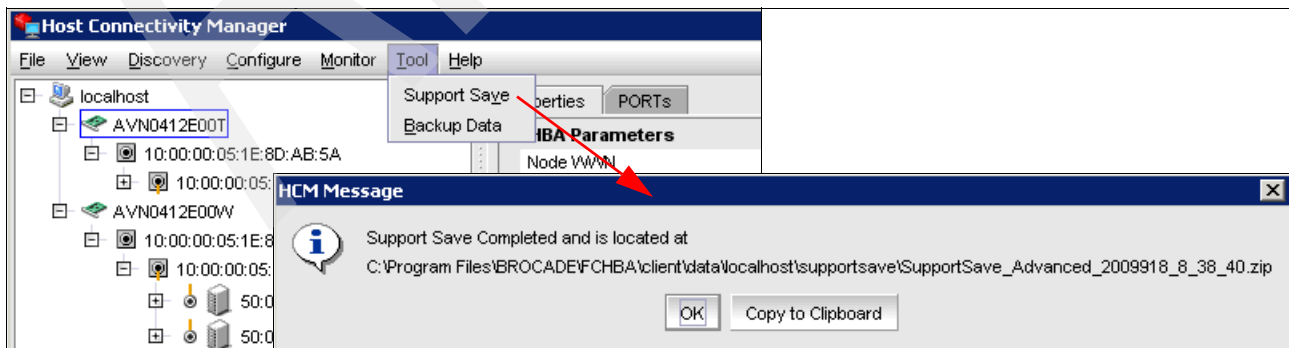


Figure 7-76 HCM Support Save

The Support Save feature saves the following information:

- ▶ Adapter model and serial number
- ▶ Adapter firmware version
- ▶ Host model and hardware revision
- ▶ All support information
- ▶ Adapter configuration data
- ▶ All operating system and adapter information needed to diagnose field issues
- ▶ Information about all adapters in the system
- ▶ Firmware and driver traces
- ▶ Syslog message logs
- ▶ Windows System Event log .evt file
- ▶ HCM GUI-related engineering logs
- ▶ Events
- ▶ Adapter configuration data
- ▶ Environment information
- ▶ Data.xml file
- ▶ Vital CPU, memory, and network resources
- ▶ HCM Agent (logs and configuration)
- ▶ Driver logs
- ▶ Install logs
- ▶ Core files
- ▶ Status and states of all adapter ports

7.11.2 Emulex HBA tools

This section briefly describes the tools used to maintain IBM branded Emulex HBAs.

Emulex offers this software to drive and manage the HBAs, and it includes the following tools:

- ▶ Driver
A host computer software component that controls the operation of peripheral controllers or HBAs attached to the host computer. Drivers manage communication and data transfer between applications and I/O devices, using HBAs as agents.
- ▶ The HBAnywareutility (HBAnyware)
This utility allows you to perform installation and configuration tasks on remote and local HBAs.
- ▶ Security Configurator
The HBAnyware security package allows you to control which HBAnyware systems can remotely access and manage HBAs on other systems in a Fibre Channel (FC) network.
- ▶ LightPulse utility (lputilnt)
This driver-specific utility for the Storport Miniport and SCSIport Miniport drivers provides a user-friendly interface that allows you to examine, manage, and configure installed HBAs. lputilnt is automatically installed when you install the HBAnyware utility.

To obtain the latest information for Emulex HBAs, go to the following address:

<http://www.ibm.com/servers/storage/support/disk/>

Installation instructions can be obtained from the Emulex Web site at the following address:

<http://www.emulex.com/downloads/ibm/fw-and-bootcode.html>

HBAware utility (HBAware)

The HBAware utility (HBAware) is a user-friendly graphical environment. Use HBAware to do any of the following actions:

- ▶ Discover local and remote hosts, host bus adapters (HBAs), targets, and LUNs.
- ▶ Reset HBAs.
- ▶ Set up persistent binding.
- ▶ Set HBA driver parameters.
- ▶ Set driver parameters simultaneously to multiple HBAs using Batch Update.
- ▶ Set global driver parameters to HBAs.
- ▶ Update firmware on the single HBA or multiple HBAs using Batch Update.
- ▶ Enable boot code.
- ▶ Run diagnostic tests on HBAs.
- ▶ Manage out-of-band HBAs.
- ▶ Manage local and in-band remote HBAs.
- ▶ Update EFIBoot (64-bit only).

To start HBAware for Windows, perform the following steps:

1. On the Windows desktop, select **Start** → **Programs** → **Emulex** → **HBAware**.
2. The initial discovery information for the host appears, as shown in Figure 7-77.

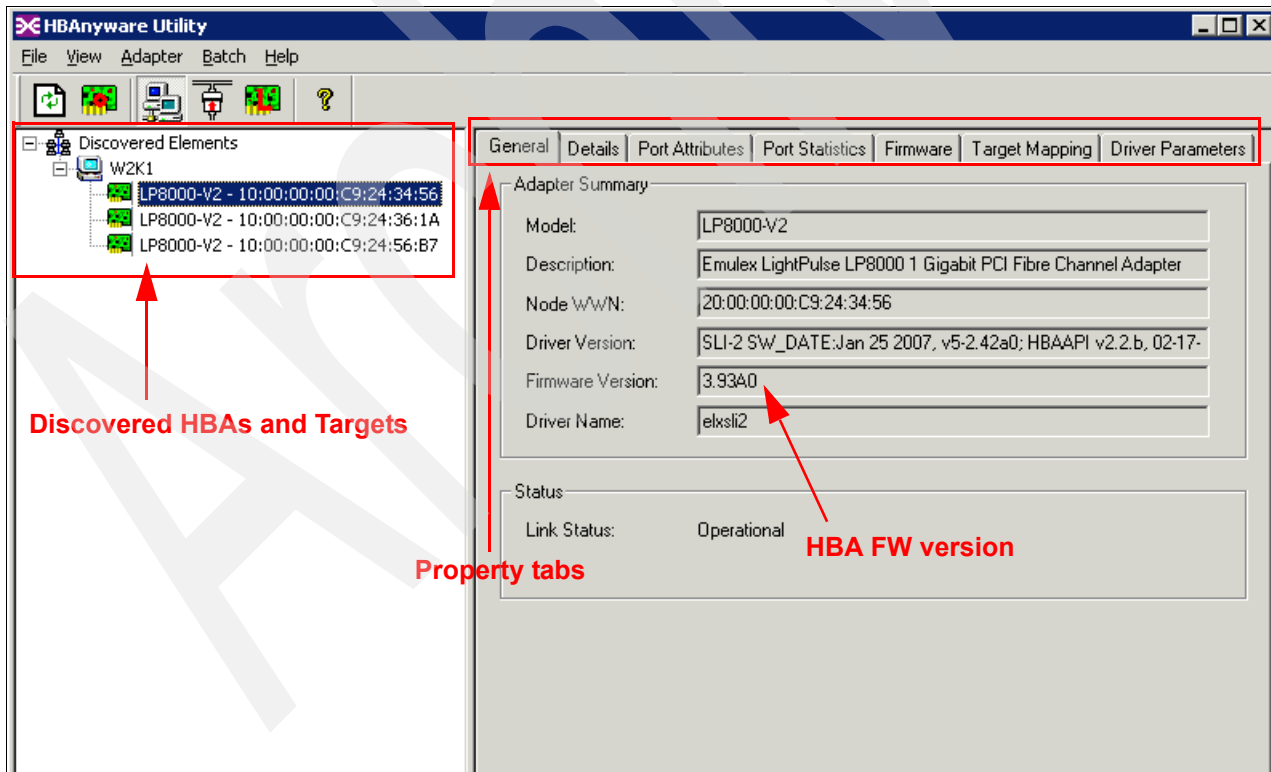


Figure 7-77 HBAware Utility main window

The HBAware utility contains five basic elements:

- ▶ Menu bar
- ▶ Toolbar
- ▶ Discovery tree
- ▶ Property tabs
- ▶ Status bar

7.11.3 Qlogic HBAs and SANsurfer (Windows/Linux)

The Qlogic SANsurfer is a graphical management interface program used to manage, configure, and diagnose Host Bus Adapters in servers running the Qlogic SANsurfer agent, which is available for Windows or Linux hosts. It is especially useful if you have hosts that are directly connected to the DS5000, as you can check for errors on the FC link of the HBA and troubleshoot path problems. Some of the features include:

- ▶ Timely and accurate detection of I/O failures
- ▶ Local and remote management of adapters
- ▶ Performance statistics
- ▶ Central control point in a network environment
- ▶ Diagnostics and utilities
- ▶ Firmware update

In this section, we provide an overview about how to use the tool with the DS5000 storage subsystem.

Setting the Qlogic SANsurfer client

You can find the latest version of the Qlogic SANsurfer at the IBM Disk Support Web site:

<http://www-1.ibm.com/servers/storage/support/disk>

Select your DS5000 model from the appropriate drop-down menu and select the **Download** tab on the right side of the page. Under the Drivers section for Qlogic FC2/4/8, you will find different packages for the Qlogic SANsurfer tool that correspond to different processor platforms with Linux or Windows.

Proceed with the installation. You have the option to install the agent and the manager separately. You can run the client from any workstation, while the agent must be installed on all the servers for which you want to manage the HBAs.

When you launch the SANsurfer application, you are prompted to specify a host to manage. Enter the IP or host name of the host running the agent, or leave the default as localhost if you are running the SANsurfer client from the server that you want to manage, as shown in Figure 7-78.

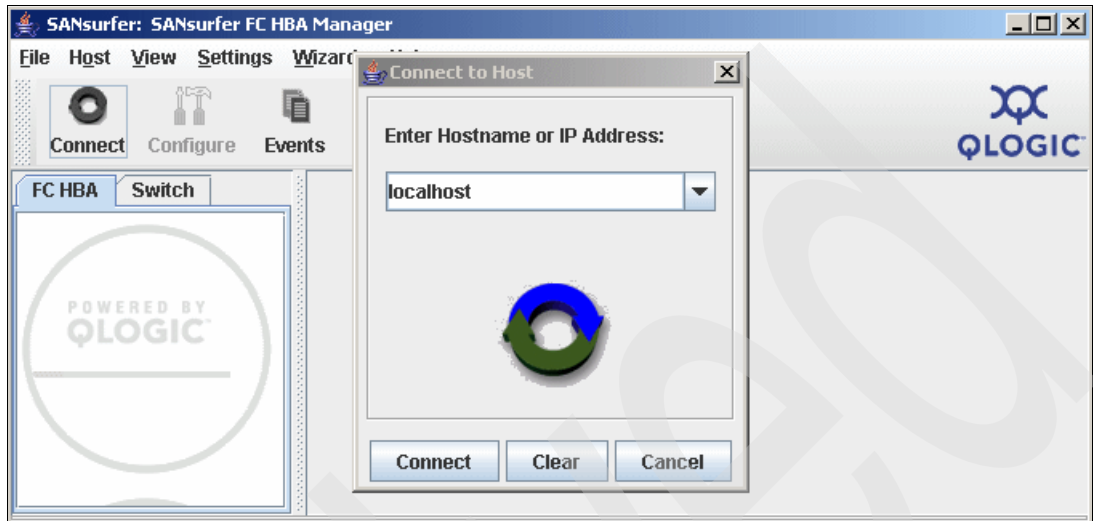


Figure 7-78 Qlogic SANsurfer FC HBA Manager view

Click **Connect** to start managing the HBAs on the specified host. At this point, you are returned to the HBA View window, and the host that you specified in the previous window is now displayed in the left pane (known as the HBA tree pane). The host bus adapters installed in the server appear below the host name (Figure 7-79).

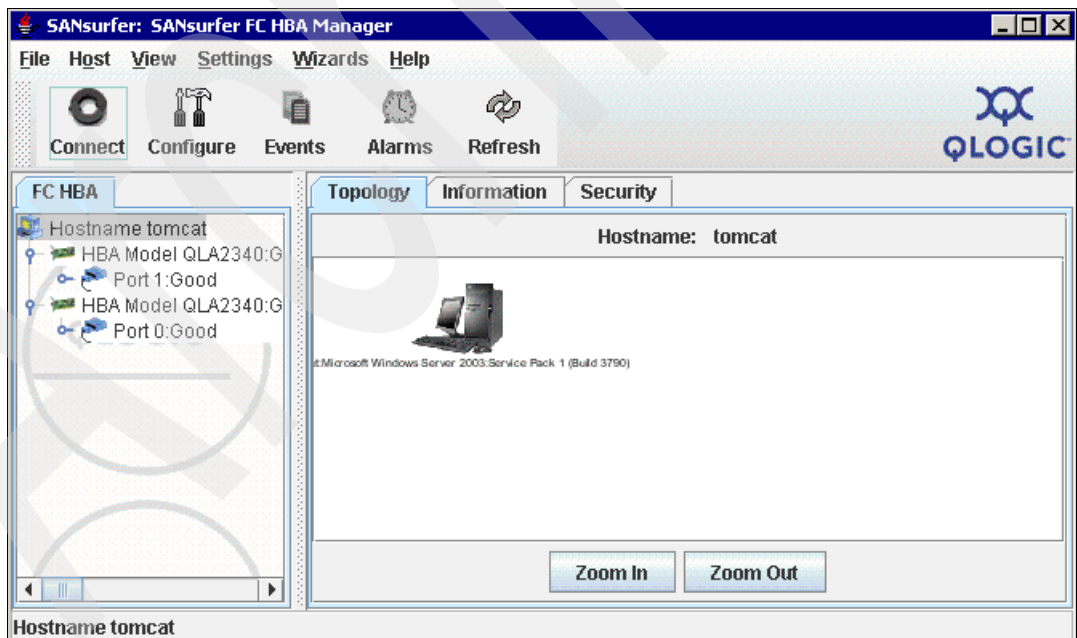


Figure 7-79 Qlogic SANsurfer Host view

When you click the host name in the left pane, three tabs are displayed in the right pane (known as the Tab pane), as shown in Figure 7-79 on page 444. They are:

- ▶ **Topology:** Contains a basic view of the topology of the currently connected server.
- ▶ **Information:** Contains basic information about the currently connected server, agent version running on the connected host, and OS version.
- ▶ **Security:** Contains security settings for the connected agent. This lets you set host security and application security.
 - Host access defines the authorized user with administrator or root privileges.
 - Application access specifies the password for changing settings on the HBAs (for firmware or NVSRAM download or BIOS settings), as shown in Figure 7-80.

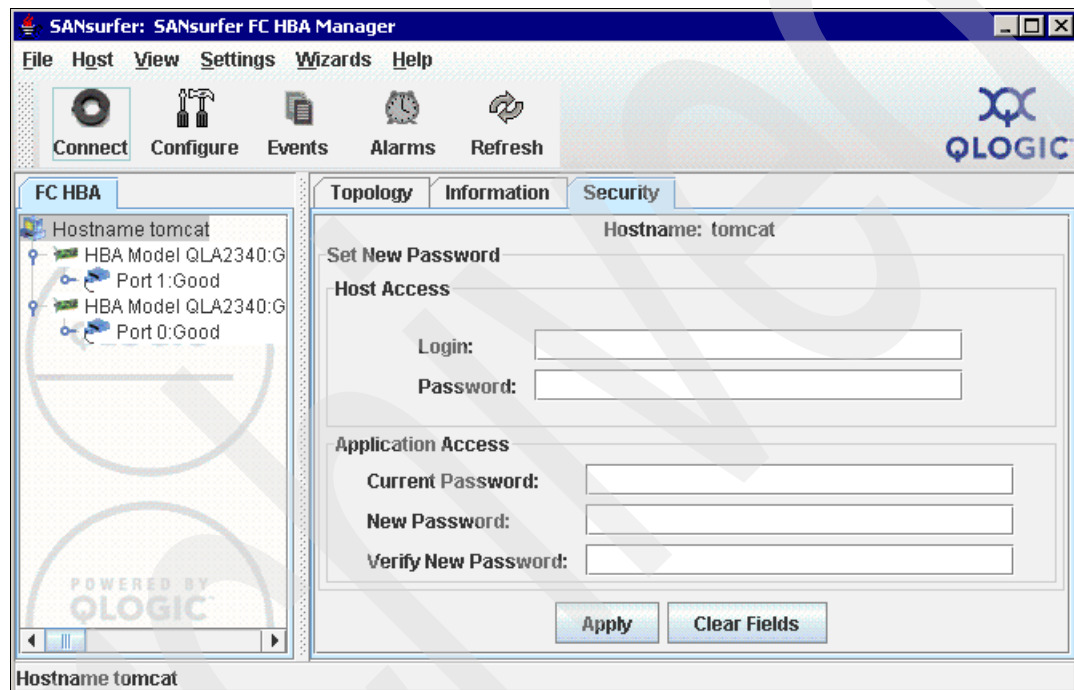


Figure 7-80 Qlogic SANsurfer Security tab

Viewing event and alarm logs

The Qlogic SANsurfer records an extensive amount of information to the event and alarm logs. The logs are saved as text files (`alarms.txt` and `events.txt`) in the folder where Qlogic SANsurfer is installed. Qlogic SANsurfer can parse and view these logs in a window. To view these logs, click **Event Log** or **Alarm Log** from the view menu, or click the appropriate button on the button bar.

Using the Qlogic SANsurfer tools

When you click one of the host bus adapter ports in the HBA tree pane, the tab pane displays eight tabs, as shown in Figure 7-81.

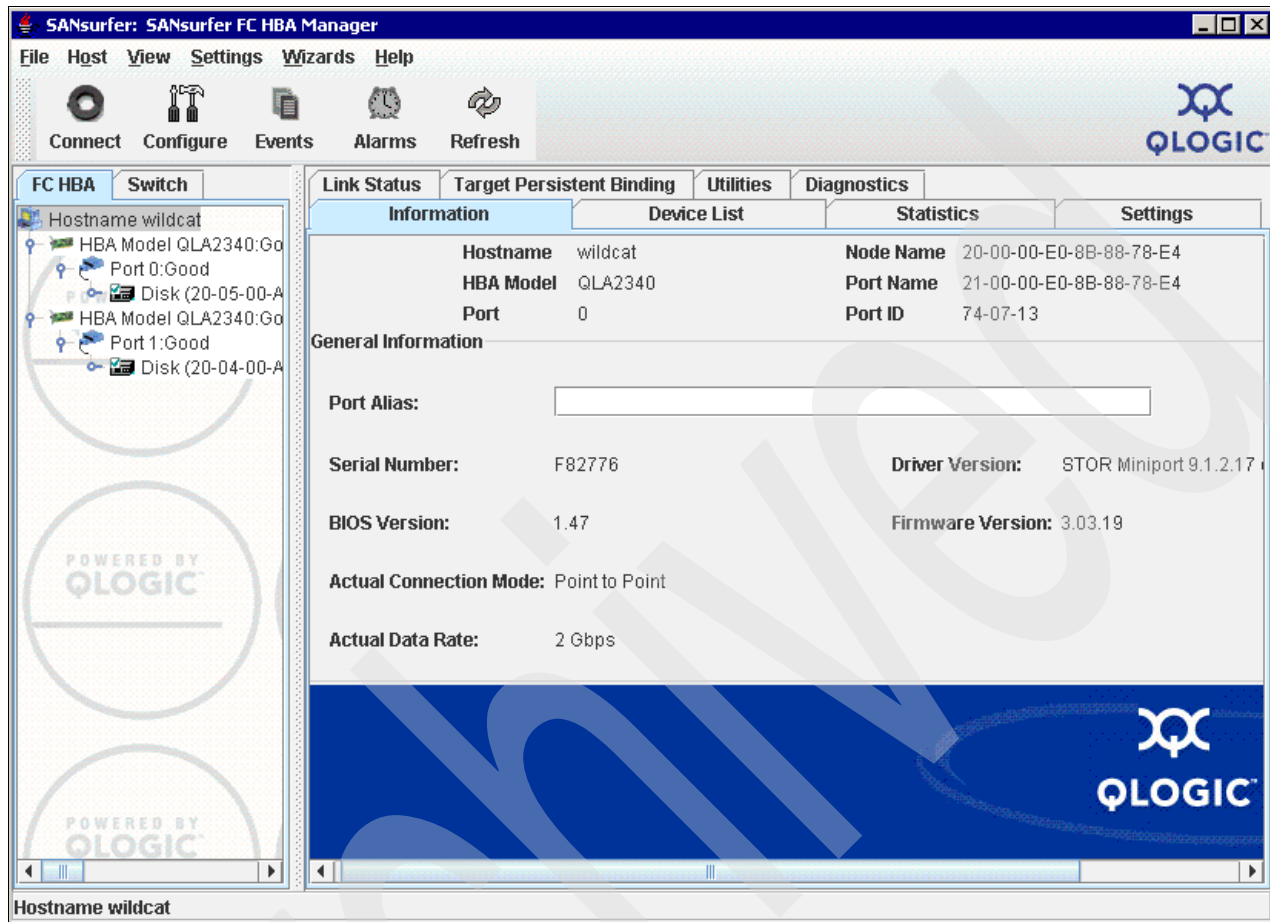


Figure 7-81 Qlogic SANsurfer FC HBA Manager view

The Qlogic SANsurfer has several tabs for the HBAs:

- ▶ **Information:** Displays general information about the server and Host Bus Adapters, such as worldwide name, BIOS, NVRAM, and driver version.
- ▶ **Device List:** Displays the devices currently available to the host bus adapter.
- ▶ **Statistics:** Displays a graph of the performance and errors on the host bus adapters over a period of time.
- ▶ **Settings:** Displays the current settings and allows you to make remote configuration changes to the NVSRAM of the adapters. All changes require a reboot of the server.
- ▶ **Link Status:** Displays link information for the devices attached to an adapter connected to a host.
- ▶ **Target Persistent Binding:** Allows you to bind a device to a specific LUN.
- ▶ **Utilities:** Allows you to update the flash memory and NVSRAM remotely.
- ▶ **Diagnostics:** Allows you to run diagnostic tests remotely.

Statistics

The Statistics tab (Figure 7-82) displays the following information:

- ▶ Adapter errors: The number of adapter errors reported by the adapter device driver (connection problem from or to switches or hubs).
- ▶ Device errors: The number of device errors reported by the adapter device driver (I/O problems to the storage subsystem, and so on). This item usually gives the first hint about what path to the storage subsystem controller has a problem.
- ▶ Reset: The number of LIP resets reported by the adapter's driver. If you get increasing numbers, there might be a communication problem between the HBAs and storage.
- ▶ I/O count: Total numbers of I/Os reported by the adapter's driver.
- ▶ IOPS (I/O per second): The current number of I/Os processed by the adapter.
- ▶ BPS (bytes per second): The current numbers of bytes processed by the adapter.

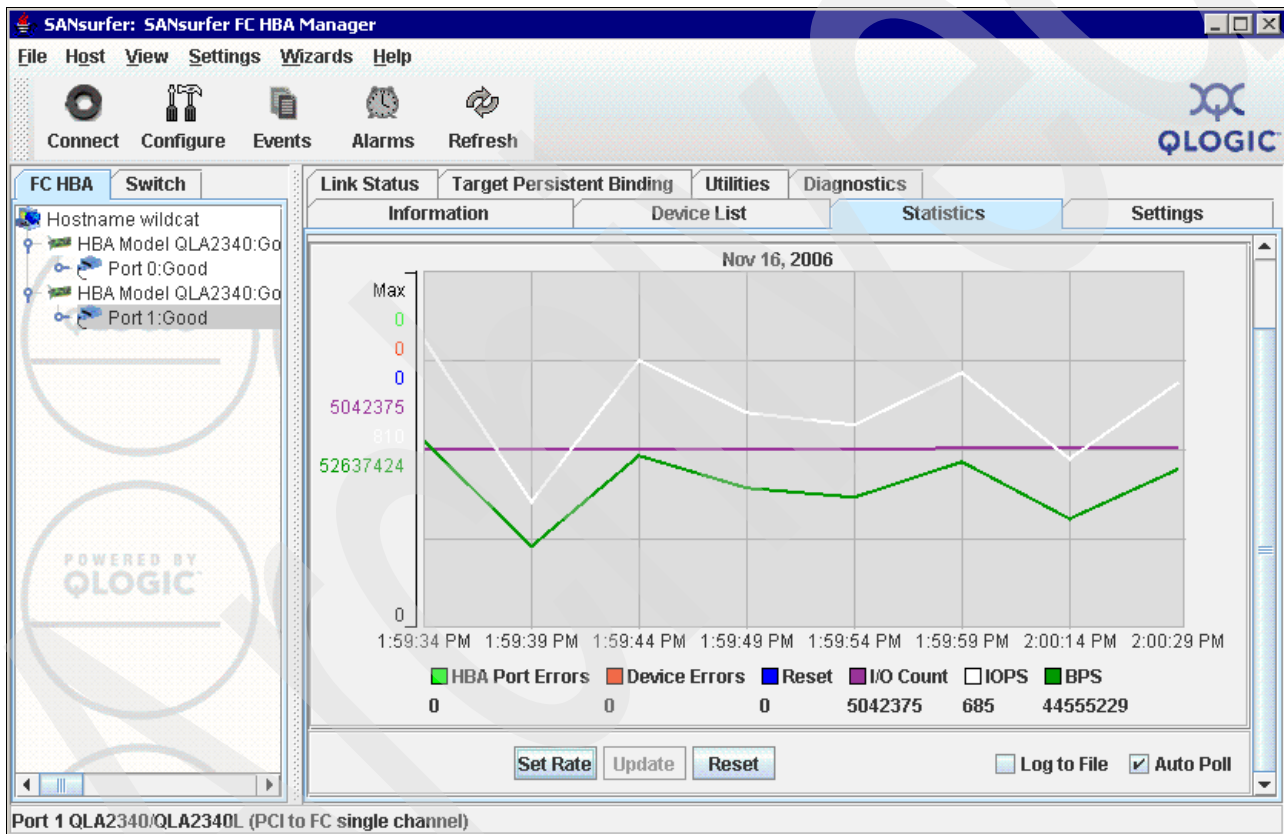


Figure 7-82 Statistics tab

Link status

If you experience problems with connectivity or performance or you see entries from RDAC or the HBA driver in Windows, use the Link Status tab (Figure 7-83) to narrow down the device causing problems (faulty cable, SFPs, and so on).

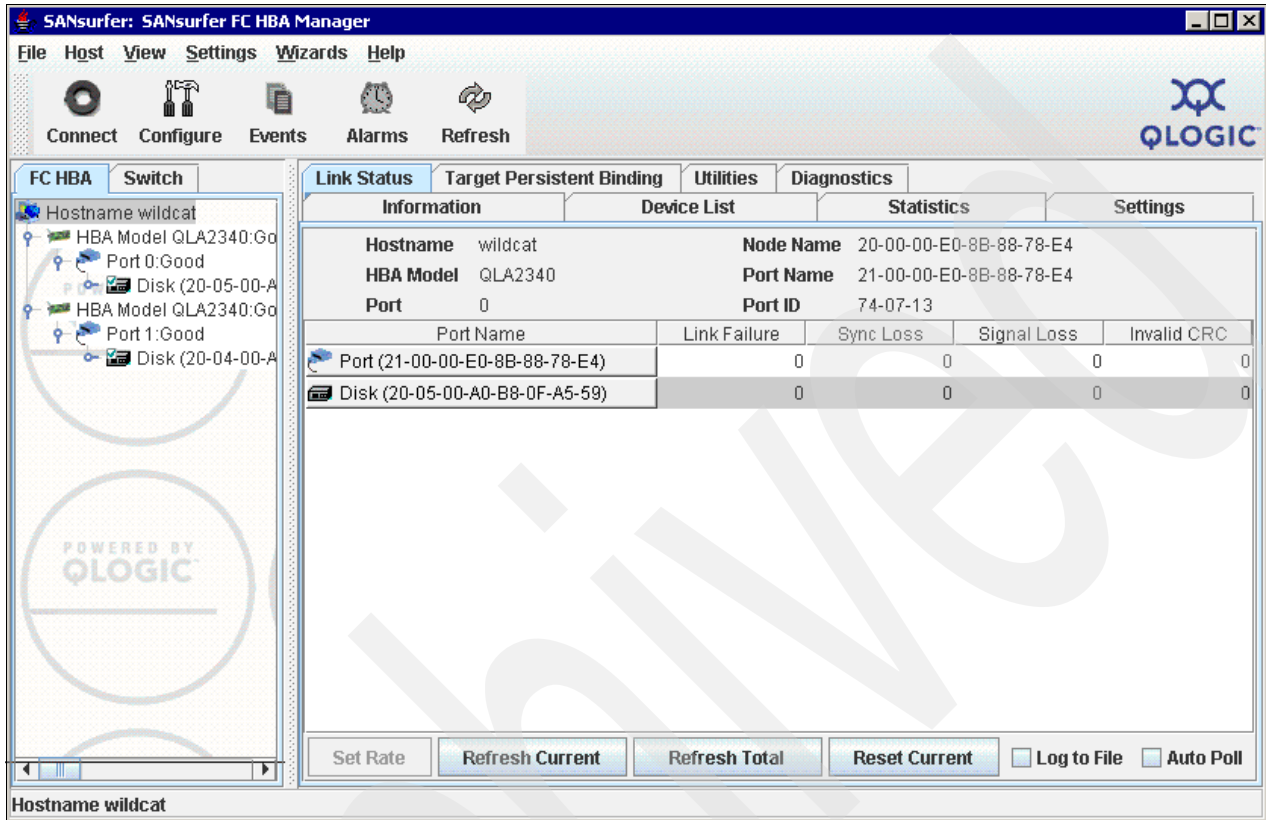


Figure 7-83 Link Status tab

The following information can be retrieved from the Link Status window:

- ▶ Link failure: The number of times that the link failed. A link failure is a possible cause for a timeout (see the Windows Event Log).
- ▶ Loss of sync: The number of times that the adapter had to re-synchronize the link.
- ▶ Loss signal: The number of times the signal was lost (dropped and re-connected).
- ▶ Invalid CRC: The number of cyclic redundancy check (CRC) errors that were detected by the device.

Diagnostics

You can use the Diagnostics tab (Figure 7-84 on page 449) to perform loopback and read/write buffer tests:

- ▶ The loopback test is internal to the adapter. The test evaluates the Fibre Channel loop stability and error rate. The test transmits and receives (loops back) the specified data and checks for frame CRC, disparity, and length errors.
- ▶ The read/write buffer test sends data through the SCSI Write Buffer command to a target device, reads the data back through the SCSI Read Buffer command, and compares the data for errors. The test also compares the link status of the device before and after the read/write buffer test. If errors occur, the test indicates a broken or unreliable link between the adapter and the device.

The Diagnostics tab has three main parts:

- ▶ Identifying Information: Displays information about the adapter being tested.
- ▶ Diagnostic Configuration: Contains testing options (like data patterns, number of tests, and test increments).
- ▶ Loopback Test Results: Displays the results of a test showing whether the test passed or failed and error counters.
 - For a loopback test, the test result includes the following information:
 - Test status
 - CRC error
 - Disparity error
 - Frame length error
 - For a read/write buffer test, the test result includes the following information:
 - Loop ID/status
 - Data mismatch
 - Link failure
 - Sync loss
 - Signal loss
 - Invalid CRC

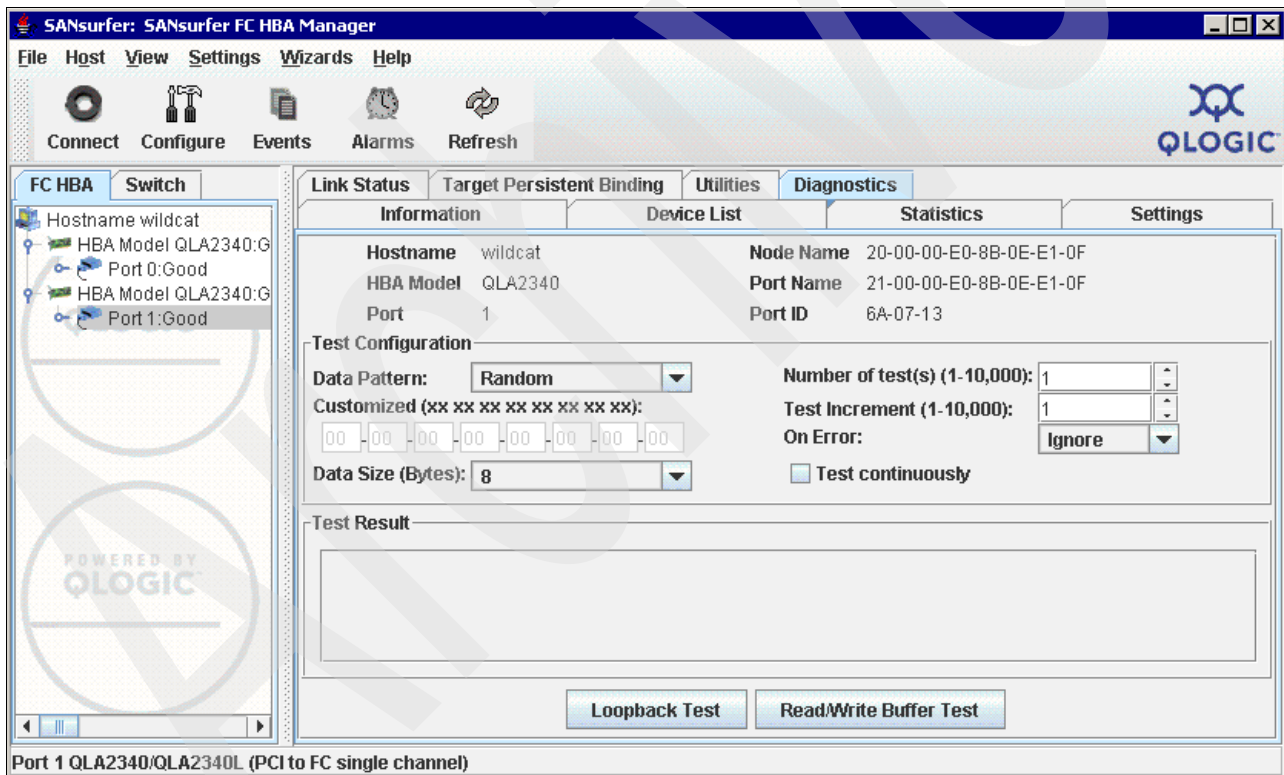


Figure 7-84 Qlogic SANsurfer Diagnostics tab

Utilities

Qlogic SANsurfer can also be used to update the BIOS, NVRAM, and drivers on the server HBAs, as shown in Figure 7-85. See “Update the HBA firmware using QLogic SANsurfer” on page 360 for instructions about updating the HBA firmware.

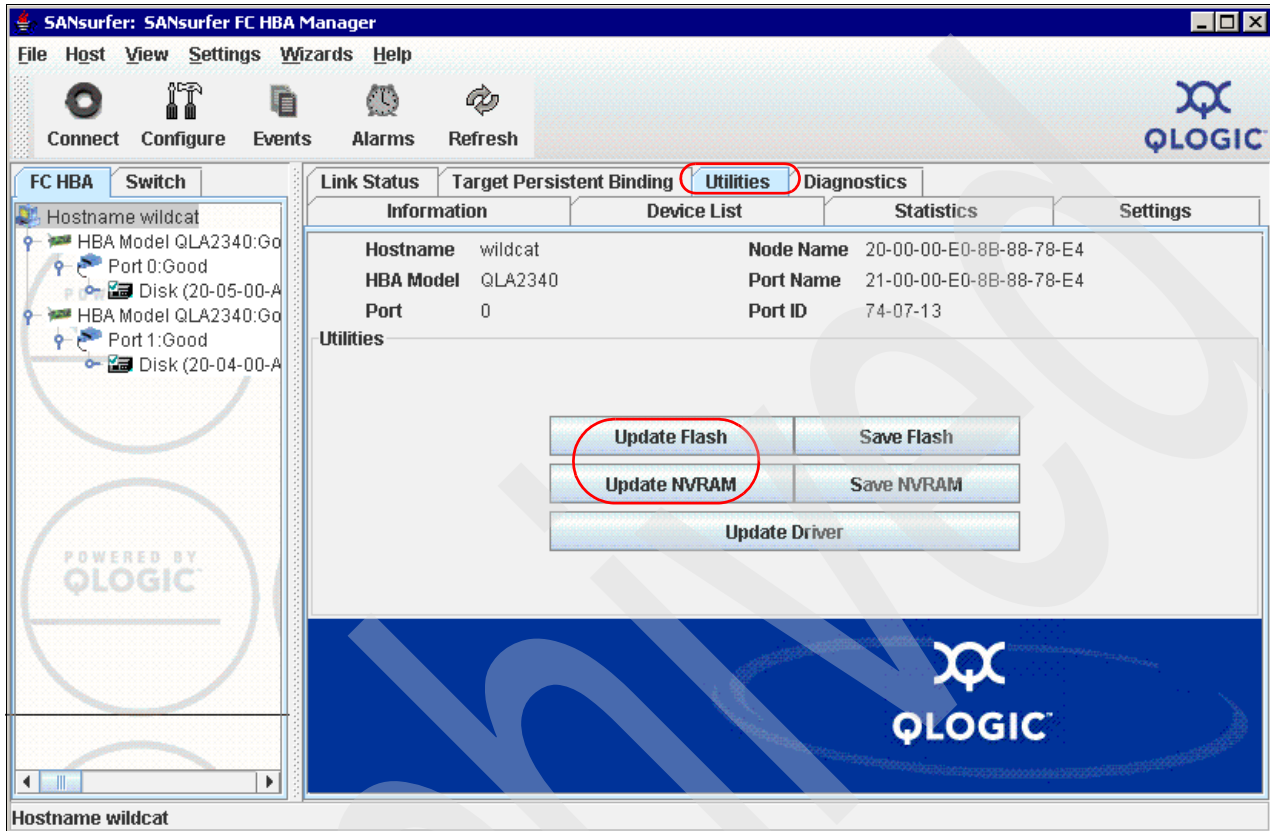


Figure 7-85 Qlogic SANsurfer utilities for updating the HBA BIOS and drivers

If you are installing your DS5000 for the first time, take time to update the HBA to the latest levels to minimize future exposures. Once your server is configured, this task will be more difficult to perform, because it requires a reboot of the server.

Displaying LUNs detected

The LUNs recognized by the HBAs are visible in the HBA tree view in the left pane, as shown in Figure 7-86. An alternate view can be accessed by going to the LUN List tab.

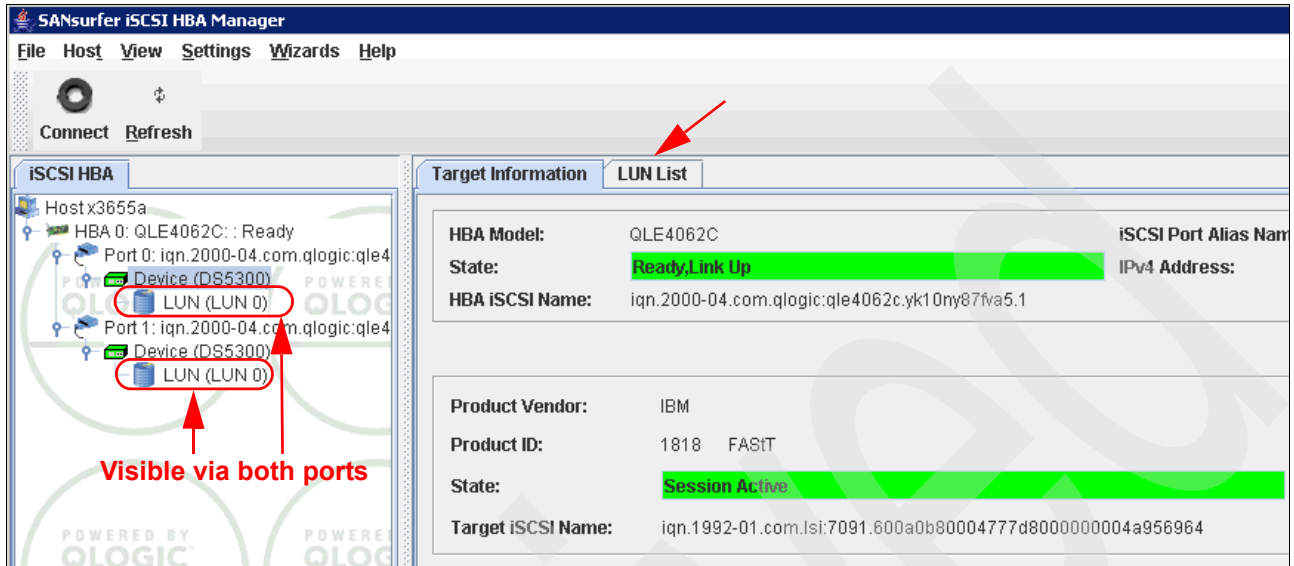


Figure 7-86 Displaying LUNs with SANsurfer

In a good configuration, you see the same LUN on both HBAs or HBA ports.

7.11.4 Windows Server 2008

The drivers used for failover in Windows environments are MPIO together with DSM installed by the Storage Manager installation package.

The former RDAC drivers for Windows are no longer supported, starting with IBM DS Storage Manager V10.10 (firmware V7.10).

Determining the driver and version in use

You can use the Windows Device Manager to check the driver and version being used:

- ▶ If you are running MPIO, as shown in Figure 7-87, there is a module under the Storage controllers tree that identifies the driver. Right-click it and select Properties to see the version of the driver installed under the Driver tab.
- ▶ If running the Storage Manager 10 Failover driver, you will find, under System devices, an entry named IBM DS3000/DS4000 series Device Specific Module, as shown in Figure 7-88 on page 453. This entry is used to set the failover policies.
- ▶ If RDAC is installed, the devices are presented differently, and there is an RDAC Multipath Pseudo-Bus entry under the Storage controllers folder of the Device Manager.

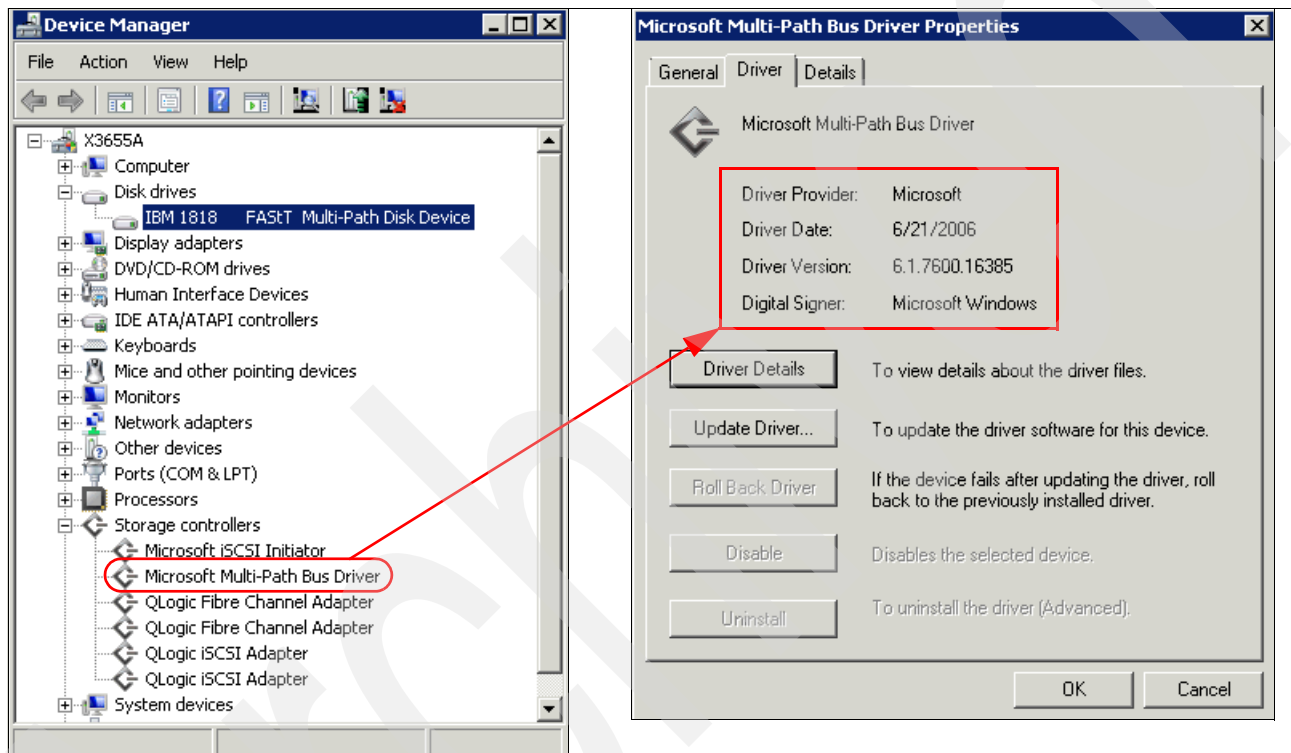


Figure 7-87 Determining the device driver in use

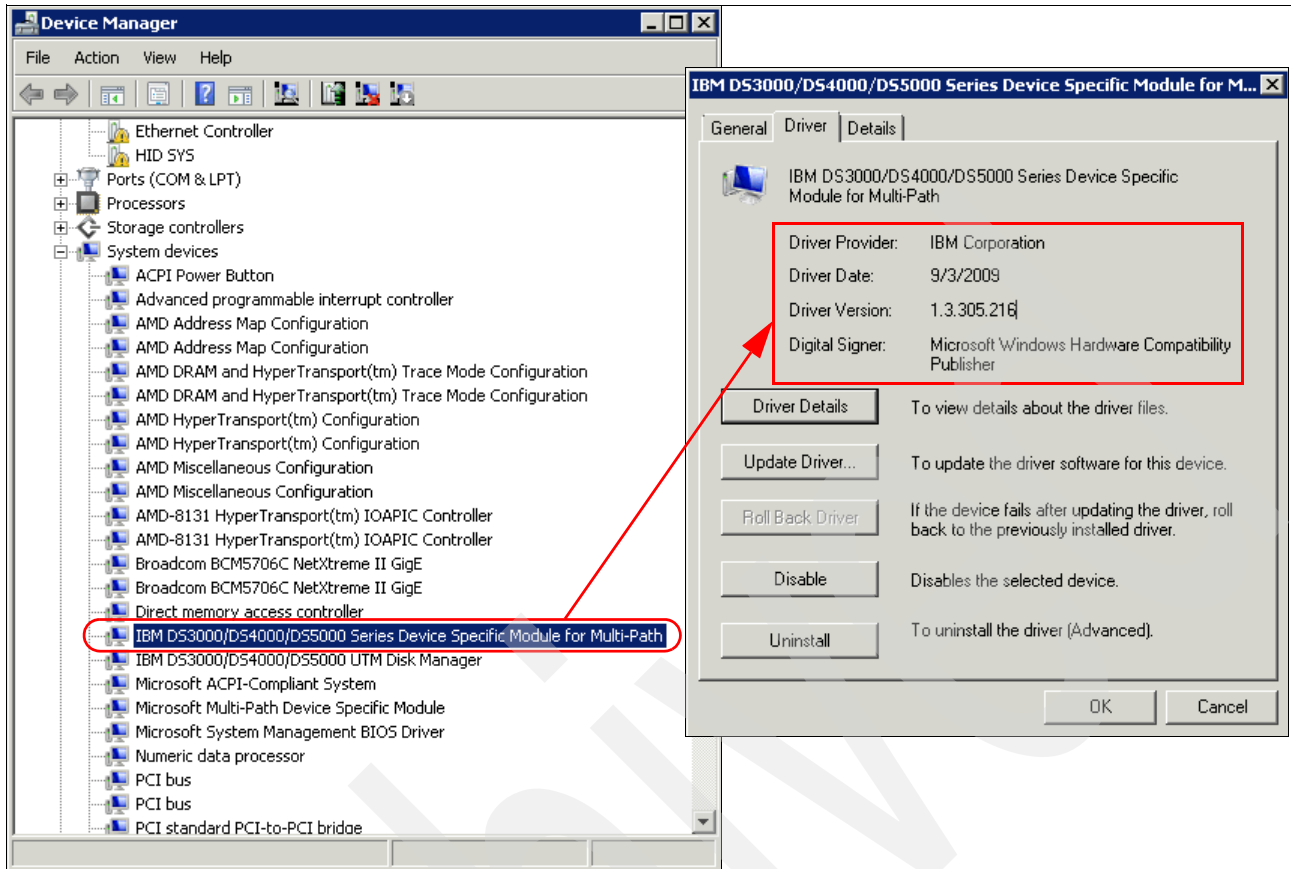


Figure 7-88 Windows Device Manager/DSM

DS5000 logical drive representation in Windows Server 2008

The DS5000 logical drives mapped to a Windows host are presented in the Windows Device Manager under the Disk drives section.

When using the device specific module (DSM) of the IBM Storage Manager together with Microsoft's MPIO architecture, the device is represented under the Device Manager as shown in Figure 7-89.

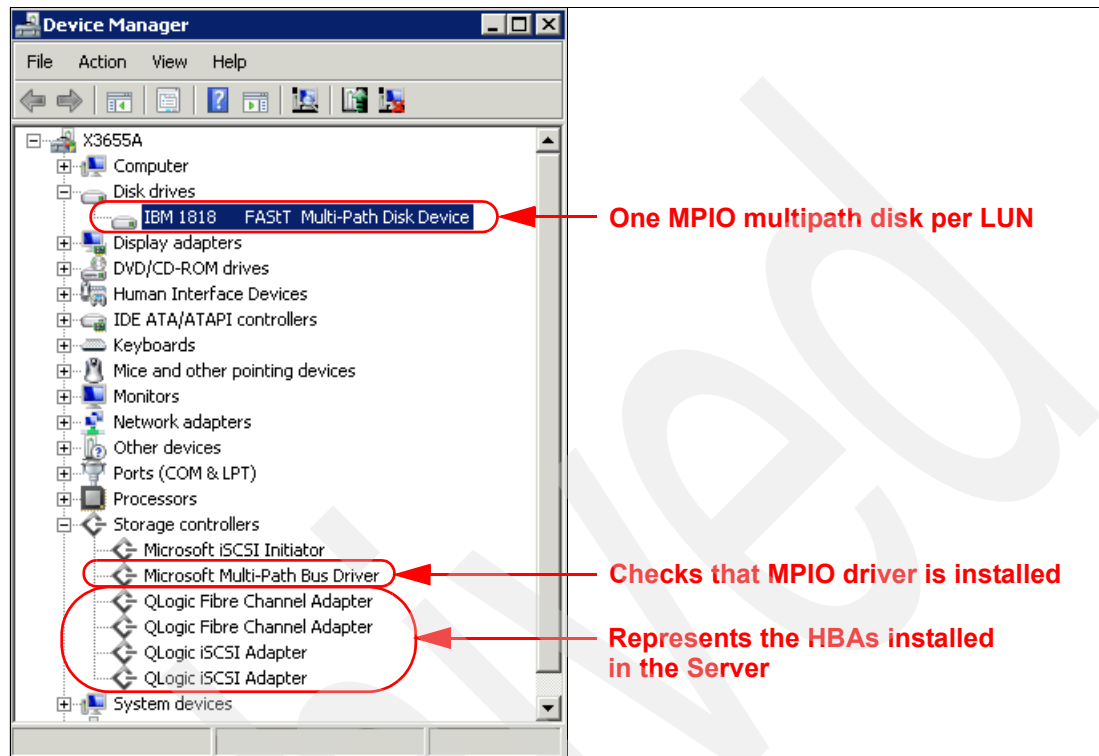


Figure 7-89 Device representation either using MPIO or SDDDSM

In the example presented in Figure 7-89, a DS5300 is configured with only one LUN. The Access LUN is not mapped to the server. Therefore, we see in the Disk drives list only one IBM 1818 Multi-path Disk Device that represents each logical drive mapped.

The physical paths are not represented, as was done with RDAC or Windows Server 2003 in the past. The best practise to verify the paths is to use the SANsurfer or other HBA utilities to verify the paths, as shown in Figure 7-86 on page 451.

However, there is also a way to gather this information in Windows. Figure 7-91 on page 455 shows the a way to get path details out of Windows tools. From the Windows Device Manager, under Disk devices, right-click the MPIO disk that represents the DS5000 LUN, as shown in Figure 7-90 on page 455.

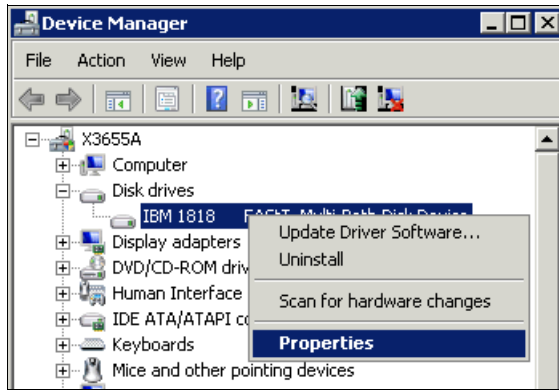


Figure 7-90 Disk drive properties

In our example in Figure 7-91, we see:

- ▶ A IBM 1818 Multi-Path Disk Device.
- ▶ MPIO counts two paths. One path is Active/Optimized, and the other one is Standby.
 - Active path on HBA port 4
 - Standby path on HBA port 5 (port number does not refer the hardware port number of the HBA)

Depending on the SAN/iSCSI topology, you might have more paths, but you usually have an equal number of active and standby paths.

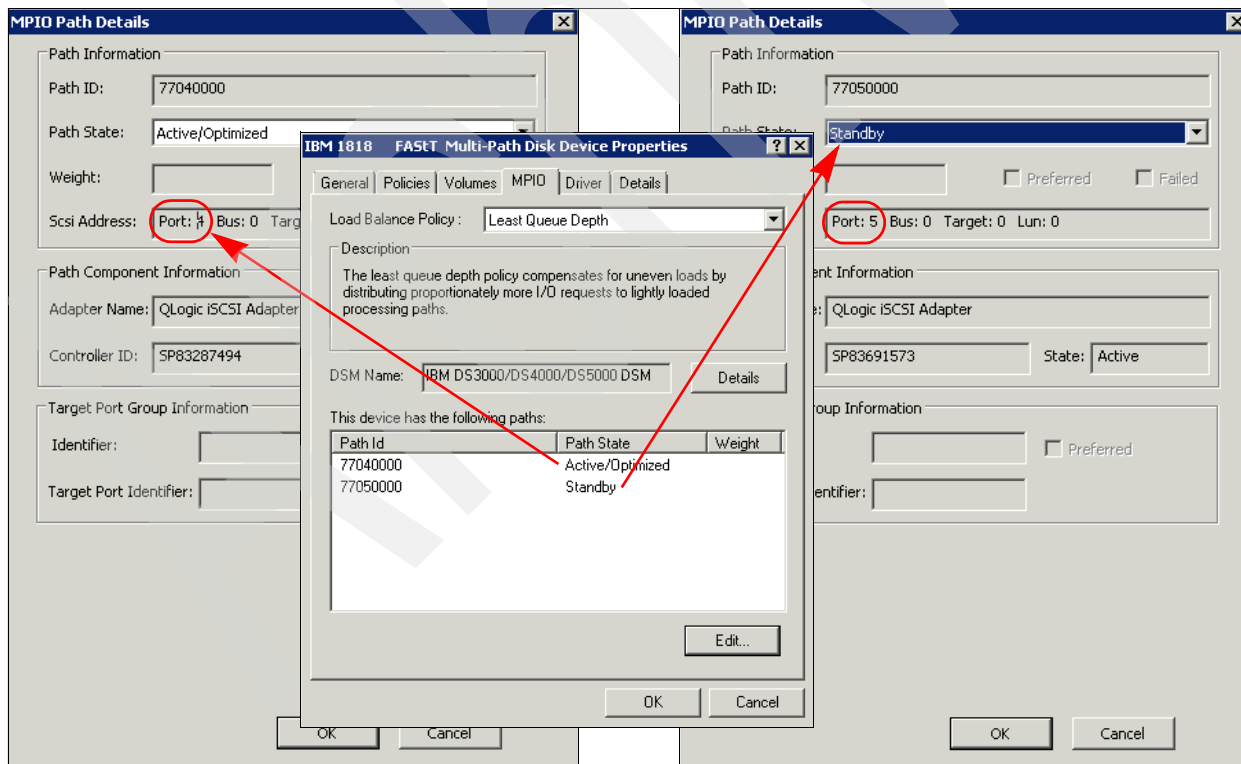


Figure 7-91 MPIO Path details

If the Access LUN is mapped to the server, an IBM Universal Xport SCSI Disk Device appears in the Disk drives list. There is one for each DS5000 connected to the server.

Matching DS5000 logical drives with Windows disks

You can cross-reference the disks represented in the Windows Device Manager list with the logical drives in the DS5000 Mapping view. Right-click the multi-path disk representation in the Device Manager and select **Properties** to display the information in the MPIO tab shown in Figure 7-92. If you select your host in the Mapping tab of SM, you see the matching LUN.

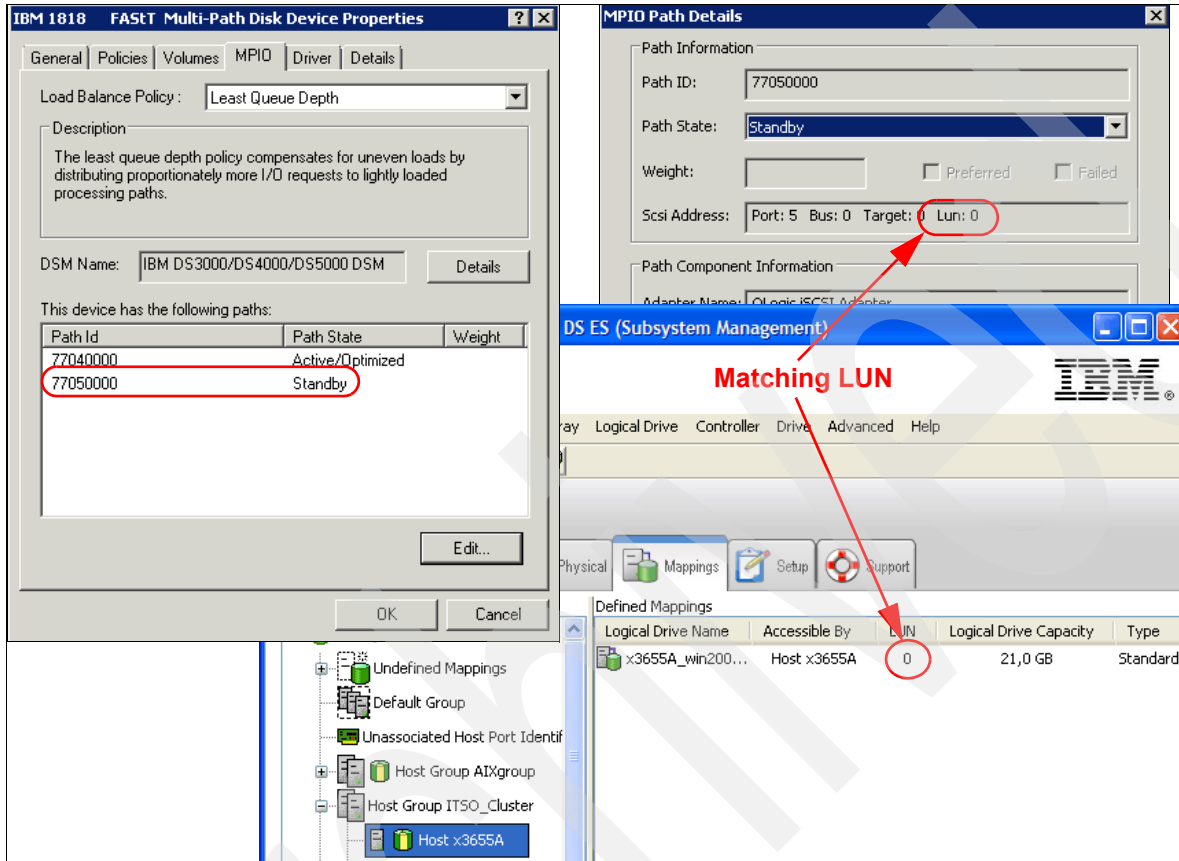


Figure 7-92 Matching disks in Windows and SM

You can access the disk device properties in Windows from either the Device Manager (Figure 7-90 on page 455) or, in Windows 2008, by selecting **Server Manager** → **Storage** → **Disk Management**. In the Disk Management window, select the disk to match from the lower right pane, then right-click and select **Properties**. The disk properties window opens, as shown in Figure 7-92 on page 456.

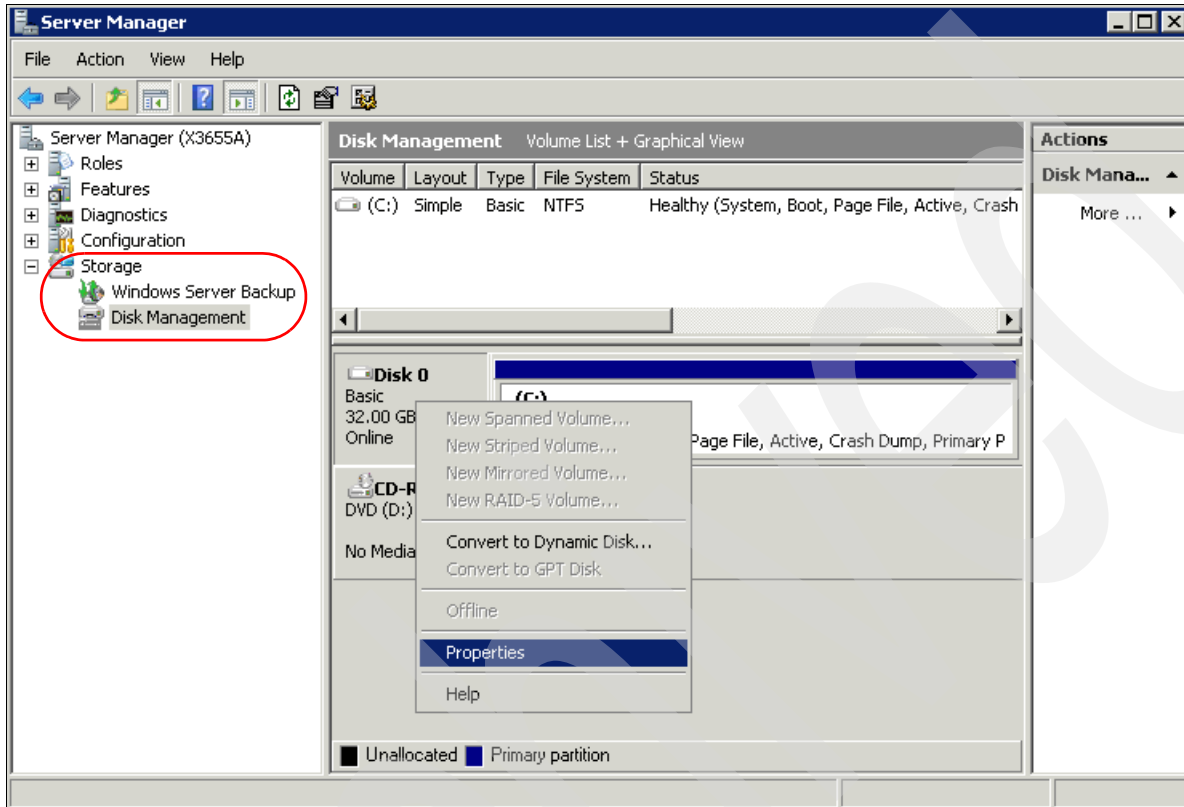


Figure 7-93 Disk Management

Using SM utilities to match disks

There are two commands provided by the Storage Manager to help you manage devices:

- ▶ **hot_add**: Invokes a hardware rescan to search for new DS5000 attached devices.
- ▶ **SMdevices**: Lists devices recognized by Windows, indicating DS5000 names and logical volume names as configured in the DS5000. It is very helpful to relate Windows devices with the logical drives set from Storage Manager.

SMdevices is installed under the util folder in the same directory where the SM Client is installed, and works with all DS5000 drivers (Figure 7-94).

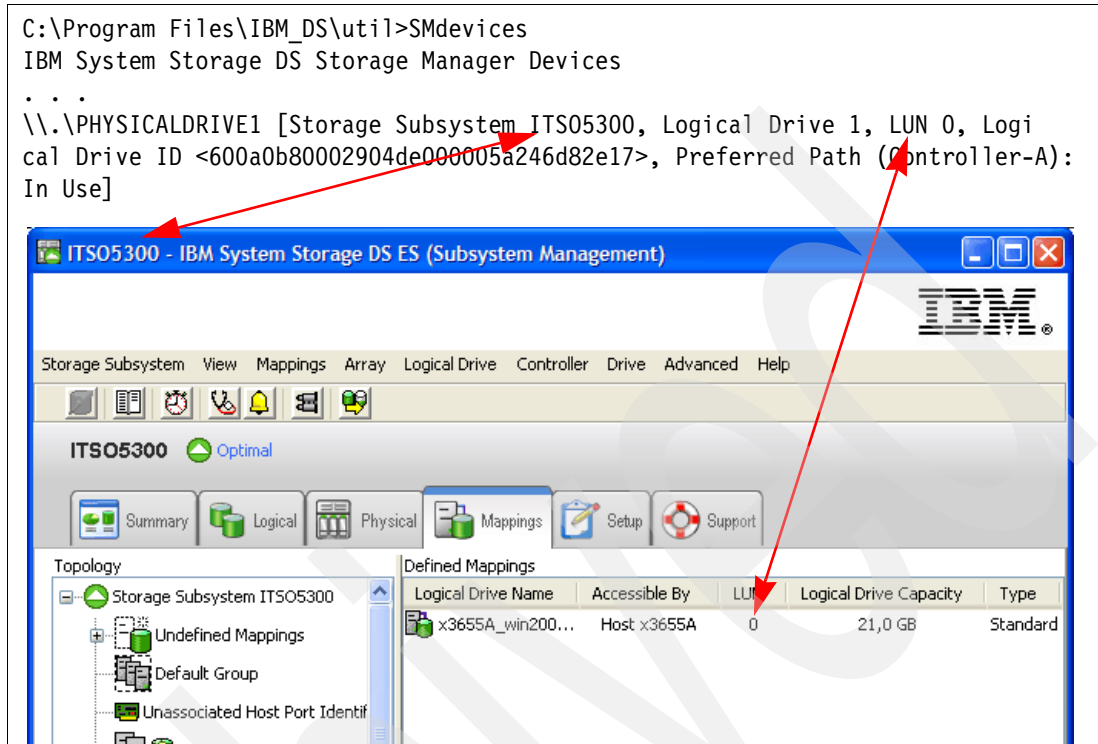


Figure 7-94 Using SMdevices utility

You can also use the drive ID string provided by the **SMdevices** output to identify your logical drive in the DS5000.

Collecting information

In addition to the information given in 7.7.6, “Collect All Support Data option” on page 401, also consider evaluating or sending your service provider the following information from your Windows host system to perform problem determination for certain failures:

- ▶ System event log: Right-click the **My Computer** icon on your desktop and select **Manage** to open the System Management window. From the System Tools option under Computer Management, select **Event Viewer** → **System**. Now click **Action** from the drop-down menu, and select **Save Log File as** to save the event view log file.
- ▶ Application log: Proceed as for the system event log, but this time select the application log.
- ▶ Dynamic system analysis (DSA log): IBM Dynamic System Analysis (DSA) collects and analyzes system information to aid in diagnosing system problems. This is a tool developed initially for IBM System x servers that not only collects the system event and application logs, but also information from your adapters, drivers, and configuration. This allows you to easily provide all the information that your support representative needs in one package. This tool is available at the following address:

<http://www-03.ibm.com/systems/management/dsa.html>

7.11.5 Linux

There are two available driver options to manage a DS5000 storage subsystem in Linux:

- ▶ The RDAC/MPP driver
- ▶ The HBA failover driver

To install RDAC/MPP, a specific kernel version is required. The RDAC driver readme has information about all the supported kernel versions. We currently recommend using RDAC as the failover driver.

RDAC provides additional functions and utilities to obtain much better information about the DS5000 from the operating system. Once the RDAC driver is installed, you have access to several commands that are useful for problem determination and correction. The utilities are:

- ▶ `mppUtil`: This utility is used with the Linux RDAC driver to perform the various functions provided by the driver.
- ▶ `hot_add`: This utility rescans your HBA to detect and configure new devices on SAN.
- ▶ `mppBusRescan`: This utility is equivalent to `hot_add` and also probes the physical bus hardware to discover any newly attached DS5000 devices. These can either be new LUNs on an existing system or a new storage subsystem. All the physical buses are scanned for new devices or LUNs. Any that are found are registered with the OS and attached to the RDAC driver. This utility should be run any time that new storage subsystem devices are connected or new LUNs are mapped to a system.
- ▶ `mppUpdate`: This utility uses the `mppUtil` utility to discover the RDAC-assigned virtual target IDs of the attached subsystems. It then updates the RDAC driver configuration file (`/var/mpp/devicemapping`) so that the entries will be persistent after a reboot. The first time that the RDAC driver sees a storage subsystem, it will arbitrarily assign a target ID for the virtual target that represents the storage subsystem. At this point, the target ID assignment is not persistent. It could change on a reboot. The `mppUpdate` utility updates the RDAC driver configuration files so that these target ID assignments are persistent and do not change across reboots. Once made persistent, the user does not have to worry about the possibility of device names changing and invalidating mount points.

`mppUpdate` must be executed after a DS5000 is added to or removed from the system. If a storage subsystem is added to the installation, `mppBusRescan` also must be run before `mppUpdate` is run. `mppUpdate` must also be executed whenever a user changes the RDAC configuration file `/etc/mpp.conf` to rebuild the new configuration file into the RAM Disk image that will be used for the next reboot.

Note: Linux RDAC supports rescan of a newly mapped LUN without rebooting the server. The utility program is packaged with the Linux RDAC driver. Rescan can be invoked by either the `hot_add` or the `mppBusRescan` command. `hot_add` is a symbolic link to `mppBusRescan`. There are man pages for both commands. However, the Linux RDAC driver does not support LUN deletion. You must reboot the server after deleting the mapped logical drives.

Listing DS5000 subsystems in Linux

If you want to scan for all the DS5000s that are attached to your Linux host, you must use the `mpputil -a` command, as shown in Example 7-9.

Example 7-9 `mppUtil -a`

```
[root@TC-2008 /]# mppUtil -a
Hostname = TC-2008
Domainname = (none)
Time = GMT 09/30/2009 04:34:27

-----
Info of Array Module's seen by this Host.
-----
ID          WWN          Type      Name
-----
0          60080e500017b5bc00000004a955e3b FC      ITS0_5020<--FC connected subsystem = ID #0
1          600a0b80004777d800000004a956964 FC      ITS05300
-----
```

In Example 7-9, two DS5000s are presented to the Linux host. Each one has an ID number assigned. You use that number to query for specific details of the subsystem.

Displaying DS storage devices from Linux

In addition to the RDAC utilities, installing the Storage Manager host software gives you access to the same utilities as in other operating systems, such as `SMdevices`.

In Example 7-10, we show the `SMdevices` output for the same system shown in Example 7-9.

Example 7-10 `SMdevices` output

```
[root@TC-2008 /]# SMdevices
DS Storage Manager Utilities Version 10.00.A5.13
...
..
 /dev/sda (/dev/sg0) [Storage Subsystem ITS0_5020, Logical Drive TC-2008-1, LUN 0, Logical
Drive ID <60080e500017b5bc000047d04aaa2559>, Preferred Path (Controller-A): In Use]
 /dev/sdb (/dev/sg1) [Storage Subsystem ITS0_5020, Logical Drive TC-2008-2, LUN 1, Logical
Drive ID <60080e500017b5bc000048444aafd60a>, Preferred Path (Controller-B): In Use]
 <n/a> (/dev/sg2) [Storage Subsystem ITS0_5020, Logical Drive Access, LUN 31, Logical
Drive ID <60080e500017b5bc000043834a955f8a>]
 <n/a> (/dev/sg3) [Storage Subsystem ITS0_5020, Logical Drive Access, LUN 31, Logical
Drive ID <60080e500017b5bc000043834a955f8a>]
 <n/a> (/dev/sg4) [Storage Subsystem ITS0_5020, Logical Drive Access, LUN 31, Logical
Drive ID <60080e500017b5bc000043834a955f8a>]
 <n/a> (/dev/sg5) [Storage Subsystem ITS0_5020, Logical Drive Access, LUN 31, Logical
Drive ID <60080e500017b5bc000043834a955f8a>]
 /dev/sdc (/dev/sg6) [Storage Subsystem ITS05300, Logical Drive TC-2008-Vol1, LUN 0,
Logical Drive ID <600a0b800047709c0000852c4abb84f4>, Preferred Path (Controller-A): In Use]
 /dev/sdd (/dev/sg7) [Storage Subsystem ITS05300, Logical Drive TC-2008-Vol2, LUN 1,
Logical Drive ID <600a0b80004777d800007e764abb86b6>, Preferred Path (Controller-B): In Use]
```

Notice the `SMdevices` output lists all the volumes found from any DS storage subsystem with LUNs mapped to this host.

Using the LUN Logical Drive ID

To display the Linux name assigned to each of the logical volumes, use the `lsvdev` utility, as shown in Example 7-11.

Example 7-11 Displaying Linux disk names with `lsvdev`

```
[root@TC-2008 mpp]# ./lsvdev
      Array Name      Lun      sd device
-----
      ITS0_5020       0      -> /dev/sda
      ITS0_5020       1      -> /dev/sdb
      ITS0_5020       2      -> /dev/sde
      ITS05300        0      -> /dev/sdc
      ITS05300        1      -> /dev/sdd
```

Use the `mppUtil` command, as shown in Example 7-12, to list the DS devices found together with their characteristics. Use the parameter `-g#`, where `#` specifies the DS Subsystem ID previously obtained by using the `mppUtil -a` command, as shown in Example 7-9 on page 460.

Example 7-12 `mppUtil -g0 - good condition`

```
[root@TC-2008 ~]# mppUtil -g0
Hostname      = TC-2008
Domainname    = (none)
Time          = GMT 10/01/2009 19:25:59

MPP Information:<-----DS5000 general information
-----
      ModuleName: ITS0_5020                               SingleController: N
VirtualTargetID: 0x000                                     ScanTriggered: N
      ObjectCount: 0x000                                   AVTEnabled: Y
      WWN: 60080e500017b5bc000000004a955e3b              RestoreCfg: N
      ModuleHandle: none                                   Page2CSubPage: Y
FirmwareVersion: 7.60.13.xx
ScanTaskState: 0x00000000
      LBPolicy: LeastQueueDepth

Controller 'A' Status:<----- Controller A breakdown view with path information
-----
ControllerHandle: none                                     ControllerPresent: Y
      UTMLunExists: N                                     Failed: N
      NumberOfPaths: 2                                    FailoverInProg: N
                                                         ServiceMode: N

      Path #1
      -----
DirectoryVertex: present                                  Present: Y
      PathState: OPTIMAL<-----first Path to CTRL A
      PathId: 77020000 (hostId: 2, channelId: 0, targetId: 0)

      Path #2
      -----
DirectoryVertex: present                                  Present: Y
      PathState: OPTIMAL<-----second Path to CTRL A
      PathId: 77030000 (hostId: 3, channelId: 0, targetId: 0)
```

```

Controller 'B' Status:<----- Controller Bbreakdown view with path information
-----
ControllerHandle: none                               ControllerPresent: Y
  UTMLunExists: N                                   Failed: N
  NumberOfPaths: 2                                  FailoverInProg: N
                                                    ServiceMode: N

  Path #1
  -----
  DirectoryVertex: present                           Present: Y
    PathState: OPTIMAL<-----first Path to CTRL A
    PathId: 77020001 (hostId: 2, channelId: 0, targetId: 1)

  Path #2
  -----
  DirectoryVertex: present                           Present: Y
    PathState: OPTIMAL<-----second Path to CTRL A
    PathId: 77030001 (hostId: 3, channelId: 0, targetId: 1)

Lun Information
-----
Lun #0 - WWN: 60080e500017b5bc00004be94ac48ff9<-----LUN0 breakdown
-----
  LunObject: present                                CurrentOwningPath: A<---
  RemoveEligible: N                                 BootOwningPath: A<---
  NotConfigured: N                                  PreferredPath: A<---
  DevState: OPTIMAL                                  ReportedPresent: Y
                                                    ReportedMissing: N
                                                    NeedsReservationCheck: N
                                                    TASBitSet: Y
                                                    NotReady: N
                                                    Busy: N
                                                    Quiescent: N

  Controller 'A' Path<-----LUN0 Path information for CTRL A
  -----
  NumLunObjects: 2                                  RoundRobinIndex: 0
    Path #1: LunPathDevice: present
              DevState: OPTIMAL
              RemoveState: 0x0 StartState: 0x1 PowerState: 0x0
    Path #2: LunPathDevice: present
              DevState: OPTIMAL
              RemoveState: 0x0 StartState: 0x1 PowerState: 0x0

  Controller 'B' Path<-----LUN0 Path information for CTRL B
  -----
  NumLunObjects: 2                                  RoundRobinIndex: 0
    Path #1: LunPathDevice: present
              DevState: OPTIMAL
              RemoveState: 0x0 StartState: 0x1 PowerState: 0x0
    Path #2: LunPathDevice: present
              DevState: OPTIMAL
              RemoveState: 0x0 StartState: 0x1 PowerState: 0x0

```

Example 7-12 on page 461 shows an example of Linux with one LUN connected to a DS5020 via Fibre Channel. Both HBAs see both controllers, which results in two paths per controller.

In Example 7-13, we see a bad condition example, where:

- ▶ Both paths to controller B failed.
- ▶ LUN0 is owned by controller A, which is in the Optimal state.
- ▶ LUN1 is owned by controller B, which is in the failed state. LUN1 has been taken over by controller A.

To correct these path issues, set the paths to controller B.

To relate the devices detected by Linux to the logical drives configured and mapped to your host in the DS5000, use the mppUtil utility. You can use the LUN number listed and the WWN string to determine the corresponding logical drive in your DS5000.

Example 7-13 mppUtil -g1 - controller failover

```
[root@TC-2008 ~]# mppUtil -g0
Hostname      = TC-2008
Domainname    = (none)
Time          = GMT 10/01/2009 19:27:28

MPP Information:
-----
      ModuleName: ITS0_5020                      SingleController: N
VirtualTargetID: 0x000                          ScanTriggered: N
      ObjectCount: 0x000                          AVTEnabled: Y
      WWN: 60080e500017b5bc000000004a955e3b      RestoreCfg: N
      ModuleHandle: none                          Page2CSubPage: Y
FirmwareVersion: 7.60.13.xx
ScanTaskState: 0x00000000
      LBPolicy: LeastQueueDepth

Controller 'A' Status:
-----
ControllerHandle: none                          ControllerPresent: Y
      UTMLunExists: N                              Failed: N
      NumberOfPaths: 2                            FailoverInProg: N
                                                    ServiceMode: N

      Path #1
      -----
DirectoryVertex: present                        Present: Y
      PathState: OPTIMAL
      PathId: 77020000 (hostId: 2, channelId: 0, targetId: 0)

      Path #2
      -----
DirectoryVertex: present                        Present: Y
      PathState: OPTIMAL
      PathId: 77030000 (hostId: 3, channelId: 0, targetId: 0)

Controller 'B' Status:
-----
ControllerHandle: none                          ControllerPresent: Y
      UTMLunExists: N                              Failed: Y<--indicates
HW problem
      NumberOfPaths: 2                            FailoverInProg: N
                                                    ServiceMode: N
```

Path #1

 DirectoryVertex: present Present: Y
 PathState: FAILED_NEED_CHECK<-----results out of the HW problem
 PathId: 77020001 (hostId: 2, channelId: 0, targetId: 1)

Path #2

 DirectoryVertex: present Present: Y
 PathState: FAILED_NEED_CHECK<-----results out of the HW problem
 PathId: 77030001 (hostId: 3, channelId: 0, targetId: 1)

Lun Information

 Lun #0 - WWN: 60080e500017b5bc00004be94ac48ff9<-----not affected, owning CTRL optimal

 LunObject: present CurrentOwningPath: A
 RemoveEligible: N BootOwningPath: A
 NotConfigured: N PreferredPath: A
 DevState: OPTIMAL ReportedPresent: Y
 ReportedMissing: N
 NeedsReservationCheck: N
 TASBitSet: Y
 NotReady: N
 Busy: N
 Quiescent: N

Controller 'A' Path

 NumLunObjects: 2 RoundRobinIndex: 0
 Path #1: LunPathDevice: present
 DevState: OPTIMAL
 RemoveState: 0x0 StartState: 0x1 PowerState: 0x0
 Path #2: LunPathDevice: present
 DevState: OPTIMAL
 RemoveState: 0x0 StartState: 0x1 PowerState: 0x0

Controller 'B' Path <-----recognise path failure

 NumLunObjects: 2 RoundRobinIndex: 0
 Path #1: LunPathDevice: present
 DevState: FAILED_NEED_CHECK<-----Failed
 RemoveState: 0x0 StartState: 0x1 PowerState: 0x0
 Path #2: LunPathDevice: present
 DevState: FAILED_NEED_CHECK<-----Failed
 RemoveState: 0x0 StartState: 0x1 PowerState: 0x0

Lun #1 - WWN: 60080e500017b5bc00004beb4ac4901c<---- affected LUN because CTRL B failed

 LunObject: present CurrentOwningPath: A<--failover
 RemoveEligible: N BootOwningPath: B
 NotConfigured: N PreferredPath: B
 DevState: OPTIMAL ReportedPresent: Y
 ReportedMissing: N
 NeedsReservationCheck: N
 TASBitSet: Y
 NotReady: N
 Busy: N
 Quiescent: N

```

Controller 'A' Path
-----
NumLunObjects: 2                                RoundRobinIndex: 0
  Path #1: LunPathDevice: present
            DevState: OPTIMAL
            RemoveState: 0x0 StartState: 0x1 PowerState: 0x0
  Path #2: LunPathDevice: present
            DevState: OPTIMAL
            RemoveState: 0x0 StartState: 0x1 PowerState: 0x0

Controller 'B' Path
-----
NumLunObjects: 2                                RoundRobinIndex: 0
  Path #1: LunPathDevice: present
            DevState: FAILED_NEED_CHECK<-----Failed
            RemoveState: 0x0 StartState: 0x1 PowerState: 0x0
  Path #2: LunPathDevice: present
            DevState: FAILED_NEED_CHECK<-----Failed
            RemoveState: 0x0 StartState: 0x1 PowerState: 0x0

```

Verify RDAC installation

To verify the installation of RDAC, you have many different commands available. Chose the right command that displays the output you want. You can use one of the following options:

- ▶ Type the following command to list the installed modules:

```
# lsmod
```

Verify that the module entries are included in the lsmod list, as follows:

- Module entries for SLES or RHEL:
 - mppVhba
 - mppUpper
 - lpfc for Emulex HBAs
 - qla2xxx for Qlogic HBAs
 - lpfcdfc (if ioctl module for Emulex HBAs is installed)
- mppVhba and MppUpper are RDAC modules.

Note: If you do not see the mpp_Vhba module, the likely cause is that the server was rebooted before the LUNs were assigned, so the mpp_Vhba module was not installed. If this is the case, assign the LUNs now, reboot the server, and repeat this step.

- ▶ Type the following command to verify the RDAC/MPP driver version:

```
# mppUtil -V
```

The Linux multipath driver version displays, as shown in Example 7-14.

Example 7-14 mppUtil -V

```
[root@TC-2008 ~]# mppUtil -V  
Linux MPP Driver Version: 09.03.0C05.0214
```

- ▶ Type the following command to verify that the devices are configured with the RDAC driver:

```
# ls -lR /proc/mpp
```

The output similar to the one shown in Example 7-15 displays:

Example 7-15 ls -lR /proc/mpp

```
/proc/mpp:  
total 0  
dr-xr-xr-x 4 root root 0 Oct 24 02:56 DS5020-sys1  
crwxrwxrwx 1 root root 254, 0 Oct 24 02:56 mppVBusNode  
/proc/mpp/ DS5020-sys1:  
total 0  
dr-xr-xr-x 3 root root 0 Oct 24 02:56 controllerA  
dr-xr-xr-x 3 root root 0 Oct 24 02:56 controllerB  
-rw-r--r-- 1 root root 0 Oct 24 02:56 virtualLun0  
/proc/mpp/ DS5020-sys1/controllerA:  
total 0  
dr-xr-xr-x 2 root root 0 Oct 24 02:56 lpfc_h6c0t2  
/proc/mpp/ DS5020-sys1/controllerA/lpfc_h6c0t2:  
total 0  
-rw-r--r-- 1 root root 0 Oct 24 02:56 LUN0  
/proc/mpp/ DS5020-sys1/controllerB:  
total 0  
dr-xr-xr-x 2 root root 0 Oct 24 02:56 lpfc_h5c0t0  
/proc/mpp/ DS5020-sys1/controllerB/lpfc_h5c0t0:  
total 0  
-rw-r--r-- 1 root root 0 Oct 24 02:56 LUN0
```

After you install the RDAC driver, the following commands and man pages are available:

- ▶ **mpptool**
- ▶ **mpptool**
- ▶ **mpptool**
- ▶ **RDAC**

HBA parameters in Linux

The 2.6.11 Linux kernel introduced certain changes to the lpfc (Emulex driver) and qla2xxx (Qlogic driver) Fibre Channel Host Bus Adapter (HBA) drivers; these changes removed the following entries from the proc pseudo-file system:

- ▶ /proc/scsi/qla2xxx
- ▶ /proc/scsi/lpfc

These entries provided a centralized repository of information about the drivers and connected hardware. After the changes, the drivers started storing all this information within the `/sys` file system. Since Red Hat Enterprise Linux 5 uses Version 2.6.18 of the Linux kernel, it is affected by this change.

Using the `/sys` file system means that all the Fibre Channel drivers now use a unified and consistent manner to report data. However, it also means that the data previously available in a single file is now scattered across a myriad of files in different parts of the `/sys` file system.

It is impractical to search through the `/sys` file system for the relevant files when there is a large variety of Fibre Channel-related information. Instead of using a manual search, use the **`systool`** command, which provides a simple but powerful means of examining and analyzing this information. We show several commands in this section that demonstrates what type of information that the **`systool`** command can be used to examine.

For example, to examine information about the Fibre Channel HBAs in a system, use the command **`systool -c fc_host -v`**, which shows:

- ▶ The WWN of a FC adapter
- ▶ The FC PortID
- ▶ The online state
- ▶ The HBA driver and firmware level

Figure 7-95 shows an example of this command.

```
[root@TC-2008 ~]# systool -c fc_host -v
Class = "fc_host"
Class Device = "host2" first fc adapter
Class Device path = "/sys/class/fc_host/host2"
  fabric_name      = "0x100500341e7096"
  issue_lip       = <store method only>
  node_name       = "0x200000e08b18208b"
  port_id         = "0x070100"
  port_name       = "0x210000e08b18208b"
  port_state      = "Online"
  port_type       = "NPort (fabric via point-to-point)"
  speed           = "2 Gbit"
  supported_classes = "Class 3"
  symbolic_name   = "QLA2340 FW:v3.03.26 DVR:v8.02.00.06.05.03-k"
  system_hostname = ""
  tgtid_bind_type = "wwpn (World Wide Port Name)"
  uevent          = <store method only>
[ content removed ]
Class Device = "host3" second fc adapter
Class Device path = "/sys/class/fc_host/host3"
  fabric_name      = "0x100500341e7096"
  issue_lip       = <store method only>
  node_name       = "0x200000e08b892cc0"
  port_id         = "0x070500"
  port_name       = "0x210000e08b892cc0"
  port_state      = "Online"
  port_type       = "NPort (fabric via point-to-point)"
  speed           = "2 Gbit"
  supported_classes = "Class 3"
  symbolic_name   = "QLA2340 FW:v3.03.26 DVR:v8.02.00.06.05.03-k"
  system_hostname = ""
  tgtid_bind_type = "wwpn (World Wide Port Name)"
  uevent          = <store method only>
[ content removed ]
```

Figure 7-95 HBA information in Linux: `systool -c`

Collecting information

In addition to the information mentioned in 7.7.6, "Collect All Support Data option" on page 401, you can also send additional information about the Linux host server to your service provider. This information can be collected by using the following commands:

- ▶ **SMdevices**: Lists the DS5000 recognized by your server (if SMutil is installed).
- ▶ **rpm -qa**: Lists installed software.
- ▶ **yum list installed**: Lists installed software if yum is installed.
- ▶ **mppUtil -V**: Lists the installed disk driver version.
- ▶ **ls -lR /proc/mpp**: Lists the devices recognized by the RDAC driver.
- ▶ **cat /proc/scsi/scsi**: Lists the LUNs recognized by the HBAs.
- ▶ **/opt/mpp/mppSupport**: Script provided by RDAC to collect information. Generates a compressed file in the /tmp folder.

Loss of Signal: 0
Primitive Seq Protocol Error Count: 0
Invalid Tx Word Count: 4
Invalid CRC Count: 0

IP over FC Adapter Driver Information
No DMA Resource Count: 0
No Adapter Elements Count: 0

FC SCSI Adapter Driver Information
No DMA Resource Count: 0
No Adapter Elements Count: 0
No Command Resource Count: 0

IP over FC Traffic Statistics
Input Requests: 0
Output Requests: 0
Control Requests: 0
Input Bytes: 0
Output Bytes: 0

FC SCSI Traffic Statistics
Input Requests: 28955
Output Requests: 22520
Control Requests: 172
Input Bytes: 2755219663

The text marked in red in Example 7-16 on page 469 indicates particularly important information, that is:

- ▶ HBA firmware level
- ▶ HBA WWNN and WWPN
- ▶ FCP information, such as FC attachment (Fabric or AL) and speed
- ▶ FC SCSI I/O statistics

Checking the installed AIX level

You must check that your AIX hosts meet the minimum recommended AIX OS level, maintenance level, and service pack level. You can check for this information by issuing the commands shown in Figure 7-96.

```
# oslevel  
6.1.0.0 ← AIX 6.1 installed  
  
# oslevel -r ← Technology level 3  
6100-03  
  
# oslevel -s  
6100-03-01-0921 ← Service pack 1
```

Figure 7-96 #oslevel

Collecting information

Execute the following sequence of commands to capture information to send to your IBM Support representative for problem analysis:

1. **snap -r**: This removes previous data collections from `/tmp/ibmsupt`.
2. **snap -gf1hc**: This collects information from your system and generates a compressed file in the `/tmp/ibmsupt` directory named `snap.pax.Z`. Send the compressed file output to your IBM Support representative for review.
3. **mpio_get_config -Av**: This command shows how your specific DS5000 is recognized by your MPIO AIX driver (this command does not work with SDDPCM).

Commands to display disk usage

The commands to display your disk usage are:

- ▶ **lspv**: Displays all disks and their usage per volume group (even volume groups that are not active).
- ▶ **lsvg -o**: Displays only online volume groups.
- ▶ **mount**: Displays file systems currently in use.
- ▶ **lsvg -l vgroupname**: Lists file systems for the *vgroupname* volume group (only if the volume group is active).
- ▶ **lspv -l hdisk***: Lists file systems in disk *hdisk**.

Other useful commands

Other useful commands are:

- ▶ **fcstat fcs***: Provides a complete listing of FC parameters of the HBA. Replaces **lsattr** and **lscfg**.
- ▶ **errpt -a**: Displays error log information.
- ▶ **lsdev -Cc disks**: Displays all disks recognized.
- ▶ **cfgmgr -v**: Invokes AIX scans for hardware changes.
- ▶ **lsslots -c pci**: Lists all PCI adapters in your system, and provides their physical locations.
- ▶ **lscfg -l hdisk***: Lists all recognized disks, and provides their physical locations.
- ▶ **lscfg -vl fcs***: Displays adapter *fcs**, which is used for the WWN, and the firmware level.
- ▶ **lsmcode -cd fcs***: Displays the adapter *fcs** microcode.
- ▶ **hot_add**: Similar to **cfgmgr**, but only for DS4000 devices (only if SMutil installed).
- ▶ **oslevel**: Displays the AIX version (**oslevel -s** also displays the maintenance level).
- ▶ **ls1pp -l**: Lists all installed file sets.
- ▶ **lspath**: Lists all paths and states for MPIO disks.
- ▶ **chpath -s enabled -l hdisk***: Enables paths for the *hdisk** device.

For additional information regarding AIX error messages, see your AIX software documentation or consult the Web site at the following address:

<http://publib.boulder.ibm.com/infocenter/pseries/>

Archived

Command-line interface and Script Editor

All Storage Manager functions available through the Subsystem Management window can also be performed and sent to an IBM System Storage DS storage subsystem using statements in scripts. The Script Editor can be used to create or edit a script file, save a script file to the Storage Manager station's local disk, or load a script file from disk. The command-line interface (CLI) can also be used to issue individual commands to the scripting engine from the host operating system command line or to invoke complete, pre-written scripts.

This chapter explains how to work with the command-line interface and the Script Editor. For detailed information about all the CLI parameters, consult the Command Line reference included in the SMclient online help or the guide *IBM System Storage DS3000, DS4000, and DS5000 Command Line Interface and Script Commands Programming Guide*, GC52-1275.

8.1 Command-line interface (CLI)

The Script Engine is built into the Storage Manager software and processes commands one at a time. It is invoked through one of three different methods:

- ▶ Interactive command-line interface
- ▶ Batch command-line interface
- ▶ Script Editor

Both the interactive and batch command-line interfaces use the **SMcli** command, which is packaged with the SMclient GUI. The *interactive command-line interface* connects to a storage subsystem and returns to a command prompt to await typed instructions. This interface is useful for basic storage subsystem configuration, command testing, and connectivity testing.

The *batch command-line interface* uses the **SMcli -f <scriptfile>** form of the command, where *scriptfile* is simply a series of commands that can be executed. This simple feature gives you the ability to create platform independent script files to execute on the DS5000 storage subsystem. The format of the commands that are put in this file are exactly as those entered on the interactive command-line interface.

The *Script Editor* is accessed from the GUI and provides a simple way to manage a single storage subsystem or test command syntax and execution. It is covered in more detail in 8.2, “Script Editor” on page 499.

Any of these methods provide an efficient way to edit, send, and execute Storage Manager commands on multiple network storage subsystems. The script engine makes it possible to automate many tasks, such as backup procedures or firmware upgrades on the storage subsystem.

The SMclient software must be installed on the station that will be used for command-line instructions. Detailed instructions about the installation of the SMclient can be found in 4.7, “Installing IBM System Storage DS Storage Manager” on page 152.

8.1.1 Using CLI commands

The basic syntax of a CLI command is:

```
command parameter;
```

The most important part of this syntax is the semicolon (;) at the end of the command. This terminates the command and causes the script engine to process it. A typical command looks like this:

```
show storagesubsystem;
```

Before a command can be executed, the user must connect SMcli to the storage subsystem. Both the interactive and batch command-line interfaces can address the target storage subsystem using a number of different identifiers, depending on the management method used:

- ▶ Directly managed (out-of-band): Use the *host name* or *IP address* of either or both controllers. The format is as follows:

```
SMcli <IP address of one controller> <IP address of second controller>
```

If the storage subsystem is configured in the Enterprise Management window, the storage subsystem can be specified by its user-supplied name only by using the `-n` option. The name must be unique to the Enterprise Management window, for example:

```
SMcli -n <name of the DS5000 storage subsystem>
```

If the name of the DS5000 storage subsystem consists of spaces and special characters, double quotes (") have to be used around the name, for example:

```
SMcli -n "Remote DS5100"
```

- ▶ **Host-agent managed (in-band):** Use the *host name* or *IP address* of the managing host (the host that is running the SMagent). For example:

```
SMcli <hostname of managing station>
```

The `-n` option must be used if more than one host-agent managed storage subsystem is connected to the host, for example:

```
SMcli <hostname of managing station> -n <name of the DS5000 storage subsystem>
```

If the world-wide name of the storage subsystem is specified, use the `-w` option instead of the `-n` option, for example:

```
SMcli -w 600A0B800029ED2200000000489C6F56
```

The interactive command-line mode is entered after executing one of the above commands, where one or more commands can be entered. **SMcli** verifies the existence of the specified storage subsystems and returns to a command prompt to await your instructions. Once you have entered the commands, press `Ctrl-c` to return to the operating system command prompt.

8.1.2 CLI parameters

Example 8-1 shows the help information that is displayed after executing `SMcli -?` in the command prompt window of a WIN platform; a similar output will be seen from a UNIX command line prompt.

Example 8-1 Help on SMcli

```
SMcli <DNS-network-name-or-IP-address>
    [<DNS-network-name-or-IP-address>]
    [-c "<command>;<command2>;..."]
    [-n <storage-array-name> | -w <WWID>]
    [-o <outputfile>] [-p <password>] [-e] [-S] [-quick]
SMcli <DNS-network-name-or-IP-address>
    [<DNS-network-name-or-IP-address>]
    [-f <scriptfile>]
    [-n <storage-array-name> | -w <WWID>]
    [-o <outputfile>] [-p <password>] [-e] [-S] [-quick]
SMcli {-n <storage-array-name> | -w <WWID>}
    [-c "<command>;<command2>;..."]
    [-o <outputfile>] [-p <password>] [-e] [-S] [-quick]
SMcli {-n <storage-array-name> | -w <WWID>}
    [-f <scriptfile>]
    [-o <outputfile>] [-p <password>] [-e] [-S] [-quick]
SMcli -d [-i] [-s] [-w] [-v] [-S]
SMcli -A [<DNS-network-name-or-IP-address1> [DNS-network-name-or-IP-address2]] [-S]
SMcli -X (-n <storage-array-name> | -w <WWID> | -h <hostName>)
SMcli -m <ip address> -F <e-mail address> [-g <contactInfoFile>] [-S]
SMcli -x email:<e-mail address>
    [<hostname or IP address1> [<hostname or IP address2>]]
    [-n <storage-array-name> | -w <WWID> | -h <hostName>]
    [-S]
SMcli -a email:<e-mail address>
    [<hostname or IP address1> [<hostname or IP address2>]]
    [-n <storage-array-name> | -w <WWID> | -h <hostName>]}
    [-I <informationToInclude>] [-q <frequency>] [-S]
SMcli {-a | -x} trap:<community>,<hostname or IP address>
    [<hostname or IP address1> [<hostname or IP address2>]]
    [-n <storage-array-name> | -w <WWID> | -h <hostName>]
    [-S]
SMcli -?
```

For additional information, refer to your Command Line Interface documentation

SMcli completed successfully.

Note: The method of accessing help on a CLI interface can vary depending on the operating systems. Enter `SMcli` without any parameters to display a short help message.

The command-line interface supports the command line parameters shown in Table 8-1.

Table 8-1 CLI parameters

Command-line parameter	Description
<IP address> or <hostname>	Specifies an IP address (xx.xx.xx.xx) or host name (of host-agent or controller) of a storage subsystem managed through the host-agent or directly managed method.
-a	<p>Adds a Simple Network Management Protocol (SNMP) trap destination or an e-mail address alert destination.</p> <p>When you add an SNMP trap destination, the SNMP community is automatically defined as the community name for the trap, and the host is the IP address or Domain Name Server (DNS) host name of the system to which the trap will be sent.</p> <p>When you add an e-mail address for an alert destination, the e-mail address is the e-mail address to which you want the alert message to be sent.</p>
-A	<p>Specifies a storage subsystem to add to the management domain. Specify an IP address (xx.xx.xx.xx) for each controller in the storage subsystem.</p> <p>Important: If you specify one IP address, the storage subsystem will be partially managed. If no IP address is specified, an automatic discovery is performed for storage subsystems attached to the local subnet.</p>
-c	<p>Specifies the list of commands to be performed on the specified storage subsystem.</p> <p>Important: Note the following usage requirements:</p> <ul style="list-style-type: none"> ▶ You cannot place multiple -c parameters on the same command line. However, you can include multiple commands after the -c parameter. ▶ Each command must be terminated with a semicolon (;). ▶ Windows: The entire command string must be enclosed in double quotes ("). Each command must be terminated with a semicolon (;). ▶ UNIX: The entire command string must be enclosed in single quotes ('). Each command must be terminated with a semicolon (;). <p>Note: Any errors encountered when executing the list of commands, by default, cause the execution to stop. Use the on error continue; command first in the list of commands to override this behavior.</p>
-d	<p>Displays the contents of the configuration file in the following format:</p> <pre><storageSubsystemName> <hostname> <hostname></pre> <p>The configuration file lists all known storage subsystems that are currently configured in the Enterprise Management window.</p>
-e	Executes the commands only, without performing a syntax check first.
-f	<p>Specifies the name of a file containing script engine commands to be performed on the specified storage subsystem. Use the -f parameter in place of the -c parameter.</p> <p>Note: Any errors encountered when executing the list of commands, by default, cause the execution to stop. Use the on error continue; command in the script file to override this behavior.</p>
-F	Specifies the e-mail address that sends the alerts.
-g	Specifies an ASCII file that contains e-mail sender contact information that will be included in all e-mail alert notifications. The CLI assumes that the ASCII file is text only, without delimiters or any expected format. Do not use this terminal if a userdata.txt file exists.
-h	Use this parameter to specify the host name that is running the SNMP agent to which the storage subsystem is connected. Use this parameter with the -a and -x parameters.
-l	<p>Specifies the type of information to be included in the e-mail alert notifications. You can select these values:</p> <ul style="list-style-type: none"> ▶ eventOnly ▶ profile ▶ supportBundle

Command-line parameter	Description
-i	Shows the IP address of the known storage subsystems. Use this terminal with the -d terminal. The file's contents use the format storage-system-name IP-address1 IPAddress2.
-m	Specifies the IP address or host name of the mail/SNMP server that will send the alerts.
-n	Specifies the storage subsystem name on which you want to perform the script commands. This name is optional when a <hostname> or <IP address> is used. However, if you are managing the storage subsystem using the host-agent management method, you must use the -n option if more than one storage subsystem is connected to the host at the specified address. This name is required when the <hostname> or <IP address> is not used. However, the storage subsystem name must be configured for use in the Enterprise Management window and must not be a duplicate of any other configured storage subsystem name.
-o	Specifies a file name for all output text from the script engine. If this parameter is not used, the output goes to stdout.
-p	Specifies the password for the storage subsystem on which you want to run commands. A password is not necessary under these conditions: <ul style="list-style-type: none"> ▶ A password has not been set on the storage subsystem. ▶ The password is specified in a script file that you are running. ▶ You specify the password by using the -c terminal and the set session password= password command.
-q	Specifies the frequency at which you want to include additional profile or support bundle information in the e-mail alert notifications. An e-mail alert notification containing at least the basic event information is always generated for every critical event. If you set the -l terminal to eventOnly, the only valid value for the -q terminal is everyEvent. If you set the -l terminal to either the profile value or the supportBundle value, this information is included with the e-mails with the frequency specified by the -q terminal. These values are valid frequency values: <ul style="list-style-type: none"> ▶ everyEvent: Information is returned with every e-mail alert notification. ▶ 2: Information is returned no more than once every two hours. ▶ 4: Information is returned no more than once every four hours. ▶ 8: Information is returned no more than once every eight hours. ▶ 12: Information is returned no more than once every 12 hours. ▶ 24: Information is returned no more than once every 24 hours.
-quick	Reduces the amount of time that is required to run a single-line operation. An example of a single-line operation is the recreate flashcopy logicalDrive command. This terminal reduces time by not running background processes for the duration of the command. Do not use this terminal for operations that involve more than one single-line operation. Extensive use of this command can overrun the controller with more commands than the controller can process, which causes operational failure. Also, status updates and configuration updates that are usually collected from background processes will not be available to the CLI. This terminal causes operations that depend on background information to fail.
-s	Displays the alert settings for the storage subsystems currently configured in the Enterprise Management window.
-S	Use this parameter to suppress informational messages describing command progress that appear when running script commands. (Suppressing informational messages is also called "silent mode.") This parameter suppresses the following messages: <ul style="list-style-type: none"> ▶ Performance syntax check. ▶ Syntax check complete. ▶ Executing script. ▶ Script execution complete. ▶ SMcli completed successfully.
-v	Use this parameter with the -d parameter to display the current global status of the known devices in a configuration file.

Command-line parameter	Description
-w	Specifies the storage subsystem, using its world-wide name (WWN), on which you want to perform the script commands. Note: The WWN is optional when a <hostname> is used or if the -n option is used to identify the storage subsystem with its <storagearrayname>. Use this option <i>instead</i> of the -n option.
-x	Delete an SNMP trap destination or e-mail alert destination. To delete an SNMP trap destination, enter: -x trap:Community, HOST Here COMMUNITY is the SNMP Community Name, and HOST is the IP address or the host name of a station running an SNMP service. To delete an e-mail alert destination, enter: -x email:MAILADDRESS Here MAILADDRESS is the fully qualified e-mail address to which the alert message will no longer be sent.
-X	Use this parameter to delete a storage subsystem from a configuration.
-?	Displays usage information.

Basic interactive command-line execution examples are shown in Example 8-2. In the example, the user runs the command **show storagesubsystem time;**. Note that the user presses Ctrl-c to exit the interactive mode and return to the command line prompt, which is the same for both UNIX and Windows.

Example 8-2 SMcli command in interactive mode

```
C:\temp>"C:\Program Files\IBM_DS\client\SMcli" 9.11.218.183 9.11.218.182 -p
xxxxxxx
Executing script...

show storagesubsystem time;
Controller in Slot A
Date/Time: Tue Sep 15 11:38:35 GMT-07:00 2009
Controller in Slot B
Date/Time: Tue Sep 15 11:38:32 GMT-07:00 2009
^C
Script execution complete.

C:\temp>
```

The interactive command-line method can also be invoked using the -c option, which allows a command to execute immediately. This method will execute the specified command and return to the operating system command prompt:

```
SMcli <IP address of one controller> <IP address of second controller> -c
"<command>;"
```

This executes the command that is found after the -c parameter. See Example 8-3 for more details.

Example 8-3 Non-interactive commands

```
C:\temp>"C:\Program Files\IBM_DS\client\SMcli" 9.11.218.183 9.11.218.182 -p xxxx  
xxxx -c "show storageSubsystem time;"  
Performing syntax check...
```

Syntax check complete.

Executing script...

```
Controller in Slot A  
Date/Time: Tue Sep 15 11:48:21 GMT-07:00 2009  
Controller in Slot B  
Date/Time: Tue Sep 15 11:48:17 GMT-07:00 2009  
Script execution complete.
```

SMcli completed successfully.

```
C:\temp>
```

If the batch command-line interface method is used, a script file can also be specified. The **SMcli** command will verify the file location and syntax of the commands. Each of the commands will be executed one at a time. When the commands are all executed, the **SMcli** command returns to the operating system command prompt. The command has the following format:

```
SMcli <IP address of one controller> <IP address of second controller> -f <script  
file name>
```

This executes the commands contained in the script file. If the script file is not in the **SMclient** directory or the path to the **SMclient** in the system environment settings, the full path to the script must be specified. See Example 8-4 on page 481, where the command **show StorageSubsystem time; show StorageSubsystem lunmappings;** is embedded in the pre-written script (script11.txt).

Example 8-4 Sending scripts to SMclient

```
C:\temp>"C:\Program Files\IBM_DS\client\SMcli" 9.11.218.183 9.11.218.182 -p xxx
xxxx -S -f "C:\temp\script11.txt"
Controller in Slot A
Date/Time: Tue Sep 15 15:11:46 GMT-07:00 2009
Controller in Slot B
Date/Time: Tue Sep 15 15:11:43 GMT-07:00 2009
MAPPINGS (Storage Partitioning - Enabled (3 of 8 used))-----
```

Logical Drive Name	LUN	Controller	Accessible by	Logical Drive status
AIX_Boot	3	A	Host Group AIX_ITS0	Optimal
Secured1	2	A	Host Group AIX_ITS0	Optimal
test3	4	A	Host Group AIX_ITS0	Optimal
RAID1_SATA_00	0	B	Host Group LAB_SVC_Cluster_1	Optimal
Access Logical Drive	31	A,B	Host WinITS0	Optimal
TC-2008-1	0	B	Host WinITS0	Optimal

```
C:\temp>
```

This method can be used, for example, to automatically create (by combining native operating system commands with CLI commands) a FlashCopy logical drive, mount it under the operating system, and create a backup of it.

For detailed information about the CLI parameters, consult the Command Line reference included in the SMclient online help or the guide *IBM System Storage DS3000, DS4000, and DS5000 Command Line Interface and Script Commands Programming Guide*, GC52-1275.

8.1.3 Syntax requirements

The **SMcli** command has some specialized syntax requirements, but most of them can be simplified by using the batch command-line interface. A script file simply contains the list of commands that are normally run from the command line, so no special formatting is required. The script can be used by any **SMcli** command on any platform without modification.

SMcli has the following usage and formatting requirements:

- ▶ Usage requirements that apply to all operating systems:
 - All statements must end with a semi-colon (;).
 - Separate each base command and any parameters with a space.
 - Separate each parameter and its parameter value with an equal sign.
 - The Script Editor and command-line interface are not case-sensitive. Any combination of upper- and lowercase letters can be entered.
 - Invoking **SMcli** with no arguments or with an unrecognized parameter will cause usage information to be displayed.
 - Arguments used after the -n, -o, -f, and -p options that contain a space, a number, or a special character (<, >, ', !, *, for example) need to be enclosed in single quotes (') or double quotes ("), depending on the operating system being used.
 - Arguments used after the -n, -o, -f, and -p options that contain a single quote character (') need to be enclosed in double quotes (").

- Invoking **SMcli** and specifying a storage subsystem, but not specifying the commands or script file to execute, will cause **SMcli** to run in interactive mode. Use Ctrl-c to stop **SMcli** execution.
- ▶ Usage requirements that apply to Windows operating systems only:
 - Insert a backslash (\) before each double quote character (") when the double quotes are used as part of a name or command syntax (for example, -c "set storageSubsystem userLabel=\"string\";").
 - Insert a backslash (\) before each quote around a user label that contains spaces (for example, -c "start logical driveCopy source=\"Mirror Repository 1\" target=trg9 priority=high;").
 - Insert three backslashes (\\) in front of the (") to display the backslash when used with the -n, -o, -f, or -p option (for example, -n "Jason\\\" to specify storage subsystem named Jason\).VolumeCopy
 - Insert five backslashes (\\\\) in front of the (") to use the backslash character as part of the literal command string (for example, -c "set storageSubsystem userLabel=\"Jason\\\\\";" to change the name of the storage subsystem to Jason\).
 - Insert a caret (^) before each special script character (^, &, |, <, >) when that character is used with the -n, -o, -f, and -p options (for example, -n "CLI^&CLIENT" to specify storage subsystem "CLI&CLIENT"). See the appropriate operating system scripting documentation for a list of special script characters.
 - Insert three carets (^) before each special script character when used within a literal script command string (for example, -c "set storageSubsystem userLabel=\"Finance^^&payroll\";" to change the name of the storage subsystem to Finance&Payroll).
- ▶ Usage requirements that apply to UNIX operating systems only:
 - The entire command string must be enclosed in single quotes ('), although some simple commands might also work with double quotes (").

Important: As a general recommendation, it is better to avoid using special characters in order to stick to a simple command syntax.

8.1.4 Error reporting

When the CLI encounters an error, it writes information describing the error directly to the command line and sets a return code. Depending on the return code, the CLI might also write additional information about which parameter caused the error. The CLI will also write information about what it was expecting in the command syntax to help you identify any syntax errors you might have entered.

When an exception occurs while a command is running, the CLI captures the error and, at the end of processing the command (after the command processing information has been written to the command line), the CLI automatically saves the error information to a file.

Special command-line options are not required to save the error data. Additionally, if the CLI must abnormally end CLI and script commands, error data is collected and saved before the CLI finishes.

The name of the file to which error information is saved is `excp rpt.txt`. It is located in Windows directory `C:\Program Files\IBM_DS\client\data`. The file name or location cannot be changed. The file is overwritten every time an exception occurs. To save the information in the file, copy it to a new file or directory.

8.1.5 Commands overview

This section shows all of the available commands, and are classified in the following categories:

- ▶ Storage subsystem
- ▶ Controller
- ▶ Physical disk drive
- ▶ Enclosure
- ▶ Array
- ▶ Logical drive
- ▶ Host topology
- ▶ FlashCopy
- ▶ Remote Mirror
- ▶ VolumeCopy
- ▶ Session
- ▶ Other

Throughout the command listings, we have highlighted the most useful commands in **bold**.

Please be advised that the list of commands you see in the following sections are an overview; for a complete list of the commands and syntax, see *IBM System Storage DS3000, DS4000, and DS5000 Command Line Interface and Script Commands Programming Guide*, MIGR-5076792.

Tip: While the Script Editor and command-line interface syntax have undergone some revisions, the former syntax is still supported. Any scripts that adhere to previous syntax rules will still pass the Script Editor syntax check and will execute.

Storage subsystem commands

Table 8-2 lists the storage subsystem commands.

Table 8-2 Storage subsystem commands

Command	Description
activate storageSubsystem firmware	This command activates firmware that you have previously downloaded to the pending configuration area on the controllers in the storage subsystem.
autoConfigure storageSubsystem	This command automatically configures a storage subsystem. Before entering the autoConfigure storageSubsystem command, enter the show storageSubsystem autoConfiguration command.
autoConfigure storageSubsystem hotSpares	This command automatically defines and configures the hot spares in a storage subsystem.
clear storageSubsystem configuration	This command clears the entire configuration from the controllers in a storage subsystem. There is potential storage subsystem configuration damage.

Command	Description
<code>clear storageSubsystem eventLog</code>	This command clears the storage subsystem event log by deleting the data in the event log buffer. There is potential storage subsystem configuration damage.
<code>clear storageSubsystem firmwarePendingArea</code>	This command deletes, from the pending area buffer, a firmware image or NVSRAM values you have previously downloaded. There is potential storage subsystem configuration damage.
<code>disable storageSubsystem featurePack</code>	This command disables a storage subsystem premium feature.
<code>download storageSubsystem driveFirmware</code>	This command downloads firmware images to all physical disks in the storage subsystem.
<code>download storageSubsystem firmware</code>	This command downloads firmware and, optionally, NVSRAM values for the storage subsystem controller.
<code>download storageSubsystem NVSRAM</code>	This command downloads NVSRAM values for the storage subsystem controller.
<code>enable storageSubsystem featureKey</code>	This command enables a feature using a feature key file.
<code>reset storageSubsystem batteryInstallDate</code>	This command resets the age of the batteries in a storage subsystem to zero days.
<code>reset storageSubsystem RLSBaseline</code>	This command resets the Read Link Status (RLS) baseline for all devices.
<code>reset storageSubsystem logicalDriveDistribution</code>	This command reassigns (moves) all logical drives to their preferred controller.
<code>save storageSubsystem configuration</code>	This command creates a script file that you can use to create the current storage subsystem logical drive configuration.
<code>save storageSubsystem allEvents</code>	This command saves events from the Major Event Log (MEL) to a file. You can save either all the events or only the critical events.
<code>save storageSubsystem performanceStats</code>	This command saves the performance statistics to a file. Before you use this command, issue the <code>set session performanceMonitorInterval</code> and <code>set session performanceMonitorIterations</code> commands to specify how often statistics are collected.
<code>save storageSubsystem RLSCounts</code>	This command saves the RLS counters to a file. Before using this command, issue the <code>reset storageSubsystem RLSBaseline</code> command to get current data.
<code>save storageSubsystem stateCapture</code>	This command saves the state capture to a file.

Command	Description
save storageSubsystem supportData	This command saves the support related information to a file. Support related information includes: <ul style="list-style-type: none"> ▶ Storage subsystem profile ▶ Major Event Log (MEL) information ▶ Read Link Status (RLS) data ▶ NVSRAM data ▶ Current problems and associated recovery information ▶ Performance statistics for the entire storage subsystem ▶ Persistent registration and reservation information ▶ Detailed information about the current status of the storage subsystem ▶ Physical disk diagnostic data ▶ A recovery profile for the storage subsystem ▶ Unreadable sectors detected on the storage subsystem ▶ State capture data
set storageSubsystem	This command defines the properties of the storage subsystem.
set storageSubsystem redundancyMode	This command sets the storage subsystem redundancy mode to either simplex or duplex. Use simplex mode when you have a single controller. Use duplex mode when you have two controllers.
set storageSubsystem time	This command sets the clocks on both controllers in a storage subsystem by synchronizing the controller clocks with the clock of the host from which you issue this command.
set storageSubsystem alarm	This command sets the alarm on or off.
set enclosure id	This command sets the enclosure ID of a controller module or a expansion drawer in a storage subsystem.
set storageSubsystem enclosurePositions	This command defines the position of all enclosures in the storage subsystem.
show storageSubsystem autoConfiguration	This command displays the default autoconfiguration that the storage subsystem will create if you issue the autoConfigure storageSubsystem command.
show storageSubsystem connections	This command lists where drive channel ports are located and where drive channels are connected.
show storageSubsystem	This command returns configuration information about the storage subsystem.
show storageSubsystem hostTopology	This command returns storage partition topology, host type labels, and host type index for the host storage subsystem.
show storageSubsystem lunMappings	This command returns information from the array profile about the storage subsystem LUN mappings.
show storageSubsystem unreadableSectors	This command returns a table of the addresses of all sectors in the storage subsystem that cannot be read.
start storageSubsystem locate	This command locates a storage subsystem by turning on the indicator lights for the storage subsystem.
stop storageSubsystem driveFirmwareDownload	This command stops a firmware download to the physical disks in a storage subsystem that was started with the download storageSubsystem driveFirmware command.
stop storageSubsystem locate	This command turns off the storage subsystem indicator lights that were turned on by the start storageSubsystem locate command.

Controller commands

Table 8-3 lists the controller commands.

Table 8-3 Controller commands

Command	Description
clear allDriveChannels stats	This command resets the statistics for all physical disk channels.
diagnose controller	This command runs diagnostic tests on the controller.
enable controller	This command revives a controller that has become quiesced while running diagnostics.
reset controller	This command resets a controller.
save controller NVSRAM	This command saves a copy of the controller NVSRAM values to a file.
set controller	This command defines the properties for the controllers.
set controller serviceAllowedIndicator	This command turns on or turns off the Service Action Allowed indicator light on a DS4800 82A/84A command module controller. This command is valid only for the DS4800 command modules.
set driveChannel status	This command defines how the physical disk channel performs.
set hostChannel	This command defines the loop ID for the host channel.
show controller	For each controller in a storage subsystem, this command returns: <ul style="list-style-type: none"> ▶ Status (online, offline) ▶ Current firmware and NVSRAM configuration ▶ Pending firmware and NVSRAM configuration configurations (if any) ▶ Board ID ▶ Product ID ▶ Product revision ▶ Serial Number ▶ Date of manufacture ▶ Cache/processor size ▶ Date/time to which the controller is set ▶ Associated logical drives (including preferred owner) ▶ Ethernet port ▶ Physical disk interface ▶ Host interface (this applies only to Fibre Channel host interfaces)
show controller NVSRAM	This command returns the NVSRAM bytes for the specified host type.
show driveChannel stats	This command displays cumulative physical disk channel data transfer and error information.
start driveChannel locate	This command identifies the physical disk enclosures connected to a specific physical disk channel by turning on the indicator lights for the physical disk enclosure connected.
stop driveChannel locate	This command turns off the physical disk enclosure indicator lights that were turned on by the start driveChannel locate command.

Physical disk drive commands

Table 8-4 lists these commands.

Table 8-4 Physical disk drive commands

Command	Description
download drive firmware	This command downloads a firmware image to a physical disk. There can be potential storage subsystem configuration damage.
revive drive	This command forces the specified physical disk to the Optimal state.
save allDrives logFile	This command saves the log sense data to a file. Log sense data is maintained by the storage subsystem for each physical disk.
set Drive hotSpare	This command assigns or unassigns one or more physical disks as a hot spare.
set Drive operationalState	This command sets a physical disk to the failed state.
show Drive	For each physical disk in the storage subsystem, this command returns: <ul style="list-style-type: none"> ▶ Total number of physical disks ▶ Type of physical disk (Fibre or SATA) ▶ Basic physical disk information: <ul style="list-style-type: none"> ▶ Enclosure and slot location ▶ Status ▶ Capacity ▶ Data transfer rate ▶ Product ID ▶ Firmware level ▶ Physical disk channel information <ul style="list-style-type: none"> ▶ Enclosure and slot location ▶ Preferred channel ▶ Redundant channel ▶ Hot spare coverage ▶ Details for each physical disk
show allDrives downloadProgress	This command returns the status of firmware downloads for the physical disks targeted by the download drive firmware or download storageSubsystem driveFirmware commands.
start Drive initialize	This command starts physical disk initialization. There can be potential storage subsystem configuration damage.
start Drive locate	This command locates a physical disk by turning on the physical disk indicator lights.
start Drive reconstruct	This command starts reconstructing a physical disk.
stop drive locate	This command turns off the physical disk indicator lights that were turned on by the start drive locate command.

Enclosure commands

Table 8-5 lists these commands.

Table 8-5 Enclosure commands

Command	Description
download enclosure firmware	This command downloads ESM firmware.
set enclosure id	This command sets the ID of a DS4800 82A/84A storage subsystem. The range of valid IDs is from 80 through 99. This range avoids conflicts with existing drive module IDs used for attached expansion enclosures.
set enclosure serviceAllowedIndicator	This command turns on or turns off the Service Action Allowed indicator light on the power-fan canister or interconnect module in a DS4800 82A/84A subsystem.
start enclosure locate	This command locates an enclosure by turning on the indicator lights.
stop enclosure locate	This command turns off the enclosure indicator lights that were turned on by the start enclosure locate command.

Array commands

Table 8-6 lists these commands.

Table 8-6 Array commands

Command	Description
delete array	This command deletes an entire array and its associated logical drives. There can be potential storage subsystem configuration damage.
revive array	This command forces the specified array and associated failed physical disks to the Optimal state.
set array	This command defines the properties for an array.
show array	This command returns the following information about an array: <ul style="list-style-type: none">▶ Status (online or offline)▶ Drive type (Fibre or SATA)▶ Enclosure loss protection (yes or no)▶ Current owner (controller slot A or slot B)▶ Associated logical drives and free capacity▶ Associated physical disks (drives)
start array defragment	This command starts a defragment operation on the specified array.

Logical drive commands

Table 8-7 lists these commands.

Table 8-7 Logical drive commands

Command	Description
check logicalDrive parity	This command checks a logical drive for parity and media errors, and writes the results of the check to a file.
clear logicalDrive reservations	This command clears persistent logical drive reservations.
clear logicalDrives unreadableSectors	This command clears unreadable sector information from one or more logical drives.
create logicalDrive driveCount= (Automatic Drive Select)	This command creates an array across the storage subsystem physical disks, and a new logical drive in the array. The storage subsystem controllers choose the physical disks to be included in the logical drive.
create logicalDrive drives= (Manual Drive Select)	This command creates a new array and logical drive, and enables you to specify the physical disks for the logical drive.
create logicalDrive array= (Free Capacity Base Select)	This command creates a logical drive in the free space of an array.
delete logicalDrive	This command deletes one or more standard logical drives or FlashCopy and FlashCopy repository logical drives. There can be potential storage subsystem configuration damage.
recover logicalDrive	This command creates a RAID logical drive with the given properties without initializing any of the user data areas on the disks.
remove logicalDrive lunMapping	This command removes the logical unit number mapping.
repair logicalDrive parity	This command repairs the parity errors on a logical drive.
set logicalDrive	This command defines the properties for a logical drive.

Command	Description
show logicalDrive	<p>For the logical drives in a storage subsystem, this command returns:</p> <ul style="list-style-type: none"> ▶ Number of logical drives ▶ Name ▶ Status ▶ Capacity ▶ RAID level ▶ Array where the logical drive is located ▶ Details ▶ Logical Drive ID ▶ Subsystem ID ▶ Physical disk type (Fibre or SATA) ▶ Enclosure loss protection ▶ Preferred owner ▶ Current owner ▶ Segment size ▶ Modification priority ▶ Read cache status (enabled or disabled) ▶ Write cache status (enabled or disabled) ▶ Write cache without batteries status (enabled or disabled) ▶ Write cache with mirroring status (enabled or disabled) ▶ Flush write cache after time ▶ Cache read ahead multiplier ▶ Enable background Media scan status (enabled or disabled) ▶ Media scan with redundancy check status (enabled or disabled) ▶ Pre-Read redundancy check (enabled or disabled) ▶ FlashCopy repository logical drives ▶ Mirror repository logical drives ▶ FlashCopy logical drives ▶ Copies
show logicalDrive actionProgress	<p>For a long-running operation that is currently running on a logical drive, this command returns information about the logical drive action and amount of the long-running operation completed. The amount of the long-running operation that is completed is shown as a percentage.</p>
show logicalDrive performanceStats	<p>This command returns information about the performance of the logical drives in a storage subsystem.</p>
show logicalDrive reservations	<p>This command returns information about the logical drives that have reservations.</p>
start logicalDrive initialize	<p>This command starts the formatting of a logical drive in a storage subsystem.</p>

Host topology commands

Table 8-8 lists these commands.

Table 8-8 Host topology commands

Command	Description
create host	This command creates a new host.
create hostGroup	This command creates a new host group.
create hostPort	This command creates a new host port.
delete host	This command deletes a host.
delete hostGroup	This command deletes a host group.
delete hostPort	This command deletes a host port.
set host	This command assigns a host to a host group or moves a host to a different host group.
set hostGroup	This command renames a host group.
set hostPort	This command changes the host type for a host port.
show allHostPorts	For all host ports connected to a storage subsystem, this command returns the following information: <ul style="list-style-type: none">▶ Host port identifier▶ Host port name▶ Host type

FlashCopy commands

Table 8-9 lists these commands.

Table 8-9 FlashCopy commands

Command	Description
create FlashCopyLogicalDrive	This command creates a FlashCopy logical drive.
recreate FlashCopy	This command starts a fresh copy-on-write operation using an existing FlashCopy logical drive.
set logicalDrive	This command defines the properties for a FlashCopy logical drive and enables you to rename a FlashCopy logical drive.
stop flashcopy	This command stops a copy-on-write operation.

Remote Mirror commands

Table 8-10 lists these commands.

Table 8-10 Remote Mirror commands

Command	Description
activate storageSubsystem feature=remoteMirror	This command creates the mirror repository logical drive and activates the Remote Mirror feature.
create remoteMirror	This command creates both the primary and secondary logical drives for a Remote Mirror.
deactivate storageSubsystem feature=remoteMirror	This command deactivates the Remote Mirror feature and tears down the mirror repository logical drive.
diagnose remoteMirror	This command tests the connection between the specified primary logical drives and mirror logical drives on a storage subsystem with the Remote Mirror feature installed.
recreate storageSubsystem mirrorRepository	This command creates a new Remote Mirror repository logical drive using the parameters defined for a previous Remote Mirror repository logical drive.
remove remoteMirror	This command removes the mirror relationship between the primary logical drive and secondary logical drive.
resume remoteMirror	This command resumes a suspended Remote Mirror operation.
set remoteMirror	This command defines the properties for a Remote Mirror pair.
show remoteMirror candidates	This command returns information about the candidate logical drives on the remote storage subsystem that you can use as secondary logical drives for a primary logical drive.
show remoteMirror localLogicalDrive synchronizationProgress	This command returns the progress of data synchronization between the primary logical drive and secondary logical drive in a Remote Mirror.
start remoteMirror primary synchronize	This command starts Remote Mirror synchronization.
suspend remoteMirror	This command suspends a Remote Mirror operation.

VolumeCopy commands

Table 8-11 lists these commands.

Table 8-11 Logical drive copy commands

Command	Description
<code>create volumeCopy</code>	This command creates a logical drive copy and starts the logical drive copy operation.
<code>recopy volumeCopy</code>	This command reinitiates a logical drive copy operation using an existing logical drive copy pair.
<code>remove volumeCopy</code>	This command removes a logical drive copy pair.
<code>set volumeCopy</code>	This command defines the properties for a logical drive copy pair.
<code>show volumeCopy</code>	This command returns information about logical drive copy operations. The information returned is: <ul style="list-style-type: none">▶ Copy status▶ Start time stamp▶ Completion time stamp▶ Copy priority▶ Source or target logical drive WWN▶ Target logical drive read-only attribute setting▶ Percent completed, if a logical drive copy operation is in progress
<code>show VolumeCopy sourceCandidates</code>	This command returns information about the candidate logical drives that you can use as the source for a logical drive copy operation.
<code>show volumeCopy targetCandidates</code>	This command returns information about the candidate logical drives that you can use as the target for a logical drive copy operation.
<code>stop volumeCopy</code>	This command stops a logical drive copy operation.

Note: Some earlier versions of Storage Manager used “LogicalDriveCopy” instead of the “VolumeCopy” command above. If you are updating to SM V10.6 and are using CLI scripting, we recommend that you check for usage in your scripts.

Other commands

Table 8-12 lists some other available commands.

Table 8-12 Other commands

Command	Description
<code>show “string”</code>	This command shows a string of text from a script file. This command is similar to the <code>echo</code> command in MS DOS and UNIX.
<code>set session</code>	This command defines how you want the current script engine session to run.

8.1.6 CLI examples

Here are examples of how CLI can be used to access and execute script engine commands. Note that the usage for the `-c` parameter varies depending on your operating system (enclosed in single quotation marks (') in UNIX or double quotation marks (") in Windows).

Upgrading the DS5000 controller, ESM, and disk drive firmware

The script in Example 8-5 can be used during a maintenance window to automate the upgrade of all components of the DS5000 storage subsystem. It should only be run in a maintenance window, because firmware updates require a quiescing of the I/O bus. This script is platform independent except for the specification of file names and assumes that the firmware has been downloaded to a local directory. The order for applying the updates is:

1. Controller
2. ESM
3. Drives

Example 8-5 SMcli script to download firmware updates to a DS5000 storage subsystem

On command line:

```
SMcli <AcontrollerIP> <BcontrollerIP> -f firmwareupgrade.scr
```

In the firmwareupgrade.scr file:

```
download storagesubsystem firmware,NVSRAM
file="C:\FW_06100600_06100100.dlp", "C:\N1742F700R910V03.dlp";

download alltrays firmware file="C:\esm9326.s3r";

download storagesubsystem drivefirmware file="C:\ST136403FC.LOD"
file="C:\ST173404FC.LOD" file="C:\ST3146807FC.LOD" file="C:\ST318203FC.LOD";
```

The script example assumes that all types of supported drives are present in the storage subsystem. In a specific environment, change the file names to only those updates that are required for that environment. Note that the command for parallel firmware download is different from the command for single drive firmware download.

The first command (**download storagesubsystem firmware, NVSRAM**) will download the firmware and NVSRAM to the controllers. The next command (**download all trays firmware**) downloads the firmware to the ESMs. The final command (**download storagesubsystem drive firmware**) uses the parallel drive download feature. There are 20 different drive types and associated file names, so this last command is be modified to suit the environment.

Staging an upgrade of DS5000 controller firmware and NVSRAM

The script in Example 8-6 stages an upgrade of the firmware for the controller and NVSRAM to be activated at another time.

Example 8-6 SMcli script to stage a controller and NVSRAM firmware update

On command line:

```
SMcli <AcontrollerIP> <BcontrollerIP> -f "c:\stagedfirmwareupgrade.scr"
```

In the c:\stagedfirmwareupgrade.scr file:

```
download storagesubsystem firmware,NVSRAM
file="C:\FW_06100600_06100100.dlp", "C:\N1742F700R910V03.dlp" activatenow=FALSE;
```

To activate the upgrade:

```
SMcli 9.11.218.163 9.11.218.164 -c "activate storagesubsystem firmware;"
```

Upgrading individual components in the DS5000 storage controller

The script in Example 8-7 can be useful for testing a new firmware level on a subset of the environment. It is included to show the slight differences in the commands that are required. In particular, the **drive update** command is different when you are not using the parallel upgrade functionality.

Example 8-7 SMcli script to update individual components in the DS4000 storage subsystem

To update the controller only:

```
SMcli 9.11.218.163 9.11.218.164 -c "download storagesubsystem firmware  
file=\"C:\FW_06100600_06100100.dlp\";"
```

To update the NVSRAM only:

```
SMcli 9.11.218.163 9.11.218.164 -c "download storagesubsystem NVSRAM  
file=\"C:\N1742F700R910V03.dlp\";"
```

To update a particular ESM:

```
SMcli 9.11.218.163 9.11.218.164 -c "download tray [1] firmware file=\"C:\esm9326.s3r\";"
```

To update a single drive:

```
SMcli 9.11.218.163 9.11.218.164 -c "download drive [2,3] firmware  
file=\"C:\ST336732FC.LOD\";"
```

Creating and mapping logical drives on AIX

The script in Example 8-8 on page 496 automates the creation of a logical drive and maps it to a storage partition for use by a specified hostgroup. The **SMcli** commands are platform independent, but the script example is from an AIX host and includes some recommended steps to make the storage available to the operating system. This example can be used if a large number of LUNs need to be created from predefined arrays and to existing hosts.

Example 8-8 Script to create and map a logical drives for AIX

```
#!/usr/bin/ksh
#
# -----
#
# Script:      Create LUN on DS5000
# Purpose:    Create several LUNs on DS5000 and allocate to a specific hostgroup.
#             If error message is detected, script will terminate
#
# -----
set -x
#
DATE=`date '+%y%m%d'`
FCMD=/usr/SMclient/SMcli
IP="9.11.218.183 9.11.218.182"
PWDD=xxxxxxx

Create_LUN()
{
LUNNO=$7
$FCMD $IP -p $PWDD -c "create logicaldrive array=$2 userlabel=\"\$1\" capacity=$3 GB
owner=$6 segmentsize=$4;"
error_chk $? Create_$1 quit

# Note we are attaching to a 'hostgroup', this can be changed to 'host' below
$FCMD $IP -p $PWDD -c "set logicaldrive [$1] logicalunitnumber=$7 hostgroup=\"\$5\";"
error_chk $? Allocate_$1 quit

}

error_chk()
{
TIME=`date '+%T'`
if [ $1 -eq 0 ]
then
echo "Process $2 completed Successfully at $TIME"
else
echo "Warning process $2 completed with code $1 at $TIME "
if [ "$3" = "quit" ]
then
echo " ##### ERROR
#####\n"
echo "Process $2 terminating at $TIME"
echo " ##### ERROR
#####\n"
exit 1
fi
}
# Passed variables used below
# LUN name=$1
# ARRAY name $2
# LUN size $3
# Segment size $4
# Host to map to $5
# Preferred Controlle a or b $6
# Host LUN number when mapped $7

Create_LUN AIX_A Data_FC 20 64 AIX_ITS0 a 10
```

```
Create_LUN AIX_B Data_FC 20 64 AIX_ITSO b 11
Create_LUN AIX_C Data_FC 20 128 AIX_ITSO a 12
Create_LUN AIX_D Data_FC 20 128 AIX_ITSO b 13

# Run cfgmgr and check final LUN configuration
/usr/sbin/cfgmgr
# For AIX 5.3 use /usr/bin/fget_config command
/usr/bin/mpio_get_config -Av
```

Renaming a DS5000 storage subsystem

Example 8-9, Example 8-10, and Example 8-11 demonstrate the differences between running a command with the `-c` and `-f` options. The `-c` option requires you to be aware of specialized parsing syntax for your particular platform. The `-f` option eliminates that need.

Example 8-9 Windows command line with -c option

```
C:\temp>"C:\Program Files\IBM_DS\client\SMcli.exe" 9.11.218.183 9.11.218.182 -p
xxxxxxx -c "set storagesubsystem userlabel=\"Tucson_5020\";"
Performing syntax check...
```

Syntax check complete.

Executing script...

Script execution complete.

SMcli completed successfully.

```
C:\temp>
```

Example 8-10 UNIX command line with -c option

```
/usr/SMclient/SMcli 9.11.218.183 9.11.218.182 -p xxxxxxx -c "set storagesubsystem
userlabel=\"ITS0_5020\";" <
Performing syntax check...
```

Syntax check complete.

Executing script...

Script execution complete.

SMcli completed successfully.

```
root@itsop630:/root
#
```

Example 8-11 Script file alternative on Windows or UNIX

```
In userlabel.scr:
set storageSubsystem userlabel="Test DS5020";
```

```
On Windows or UNIX:
SMcli 9.11.218.163 9.11.218.164 -f userlabel.scr
```

Deleting a logical drive

Example 8-12, Example 8-13, and Example 8-14 show the same set of commands being executed on Windows, UNIX, and a script file alternative. The tasks are:

- ▶ Remove mapping of the logical drive named AIX_D.
- ▶ Delete logical drive AIX_D.
- ▶ Show the health status of the storage subsystem, which is managed through the direct management method.

Example 8-12 Windows command line with -c option

```
C:\temp>"C:\Program Files\IBM_DS\client\SMcli.exe" 9.11.218.183 9.11.218.182 -p  
passw0rd -c "remove logicaldrive [\"AIX_B\"] lunmapping hostgroup=\"AIX_ITS0\"  
delete logicaldrive [\"AIX_B\"] ;"  
Performing syntax check...
```

Syntax check complete.

Executing script...

Script execution complete.

SMcli completed successfully.

```
C:\temp>
```

Example 8-13 UNIX command line with -c option deleting logical drive

```
# /usr/SMclient/SMcli 9.11.218.183 9.11.218.182 -p passw0rd -c "remove  
logicaldrive [\"AIX_B\"] lunmapping hostgroup=\"AIX_ITS0\";"  
Performing syntax check...
```

Syntax check complete.

Executing script...

Script execution complete.

SMcli completed successfully.

```
root@itsop630:/root
```

Example 8-14 Script file alternative on Windows or UNIX

```
In deletelogicaldrive.scr:  
remove logicaldrive ["AIX_A"] lunmapping hostgroup="AIX_ITS0";  
delete logicaldrive ["AIX_A"] ;  
show storageSubsystem healthStatus;
```

On Windows or UNIX:

```
SMcli 9.11.218.183 9.11.218.182 -p xxxxxxxx -f deletelogicaldrive.scr
```

Configuration Script example 1

Example 8-15 creates a new logical drive using the **create logicalDrive** command in the free space of a array.

Example 8-15 Creating logical drives

```
Show "Create RAID 5 Logical Drive WIN-1 on existing Array Data_FC";

//Create logicalDrive on array created by the create logicalDrive drives command

//Note: For arrays that use all available capacity, the last logicalDrive on the
// group is created using all remaining capacity by omitting the
capacity=logicalDrive
// creation parameter

create logicalDrive array=Data_FC RAIDLevel=5 userLabel="WIN-1" owner=A
segmentSize=16 capacity=20 GB;
show "Setting additional attributes for logicalDrive WIN-1";
//Configuration settings that cannot be set during logicalDrive creation
set logicalDrive["WIN-1"] cacheFlushModifier=10;
set logicalDrive["WIN-1"] cacheWithoutBatteryEnabled=false;
set logicalDrive["WIN-1"] mirrorEnabled=true; set logicalDrive["WIN-1"]
readCacheEnabled=true; set logicalDrive["WIN-1"] writeCacheEnabled=true; set logicalDrive["WIN-1"] mediaScanEnabled=false;
set logicalDrive["WIN-1"] redundancyCheckEnabled=false; set logicalDrive["WIN-1"]
modificationPriority=high;
```

The line beginning with `//Create` is a comment explaining that the purpose of this script file is to create a new logical drive using the **create logicalDrive** command on an existing array.

Tip: For an example of how to set up an entire storage subsystem and the commands required to do this task, look at the output of a saved configuration (select **Storage Subsystem** → **Configuration** → **Save...**) of a configured DS5000 storage subsystem. The output from this configuration will have all the commands and steps to duplicate the configuration or modified to create a different storage controller. It is an excellent source of example commands to complete individual tasks.

8.2 Script Editor

The Script Editor is a powerful tool to create and edit scripts. It can verify syntax, run a script, or both, and the scripts can also be saved and later executed to automate storage management procedures.

8.2.1 Using the Script Editor

To open the Script Editor, perform these steps:

1. Select a storage subsystem in the Device Tree View or Device Table from the Enterprise Management window in the Storage Manager Client.

2. Select **Tools** → **Execute Script**, as shown in Figure 8-1, or right-click and select **Execute Script**.

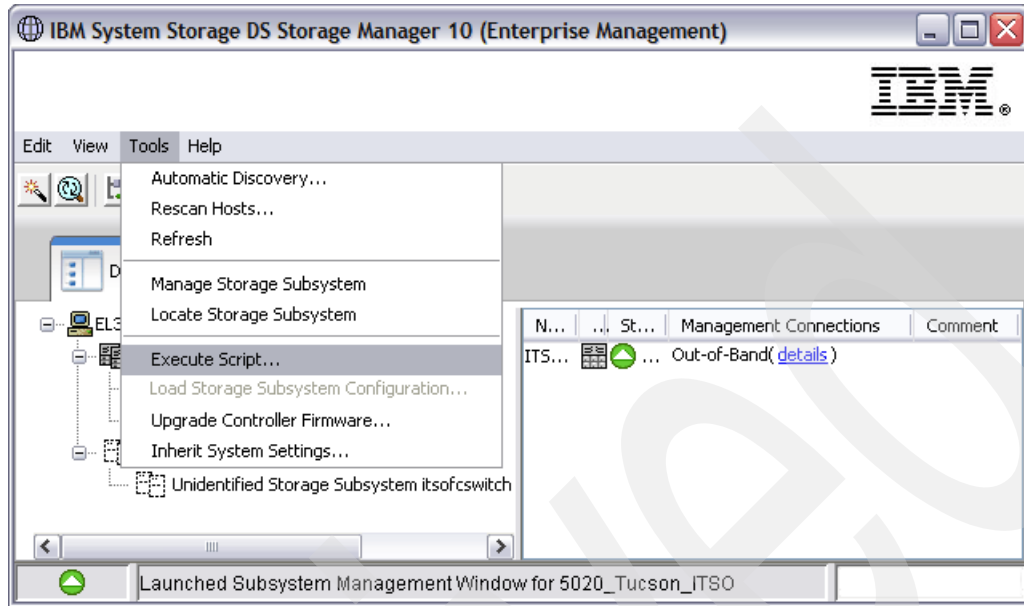


Figure 8-1 Starting the Script Editor

The Script Editor opens, as shown in Figure 8-2. There are two views in the window:

- ▶ Script view: Provides an area for inputting/editing script commands.
- ▶ Output view: Displays verification or execution results.

A splitter bar divides the window between Script View and Output View. You can use the splitter bar to resize the views.



Figure 8-2 The Script Editor

Usage guidelines

Follow these guidelines when using the Script Editor:

- ▶ All statements must end with a semicolon (;).
- ▶ Each base command and its associated primary and secondary parameters must be separated with a space.
- ▶ The Script Editor is not case-sensitive.
- ▶ Each statement must be on a separate line.
- ▶ Comments can be added to the scripts to make it easier for future reference and to understand the purpose of the command statements.

Adding comments to a script

The Script Editor supports the following comment formats:

- ▶ Text contained after two forward-slashes // until an enter character is reached.

For example, the comment The following command assigns hot spare drives is included for clarification and is not processed by the Script Editor:

```
//The following command assigns hot spare drives.  
set drives [1,2 1,3] hotspare=true;
```

Important: A comment beginning with // with must an with an end-of-line character, which is inserted by pressing the Enter key. If the script engine does not find an end-of-line character in the script after processing a comment, an error message is displayed and the script execution is terminated. This error commonly occurs when a comment is placed at the end of a script.

- ▶ Text contained between the characters /* and */.

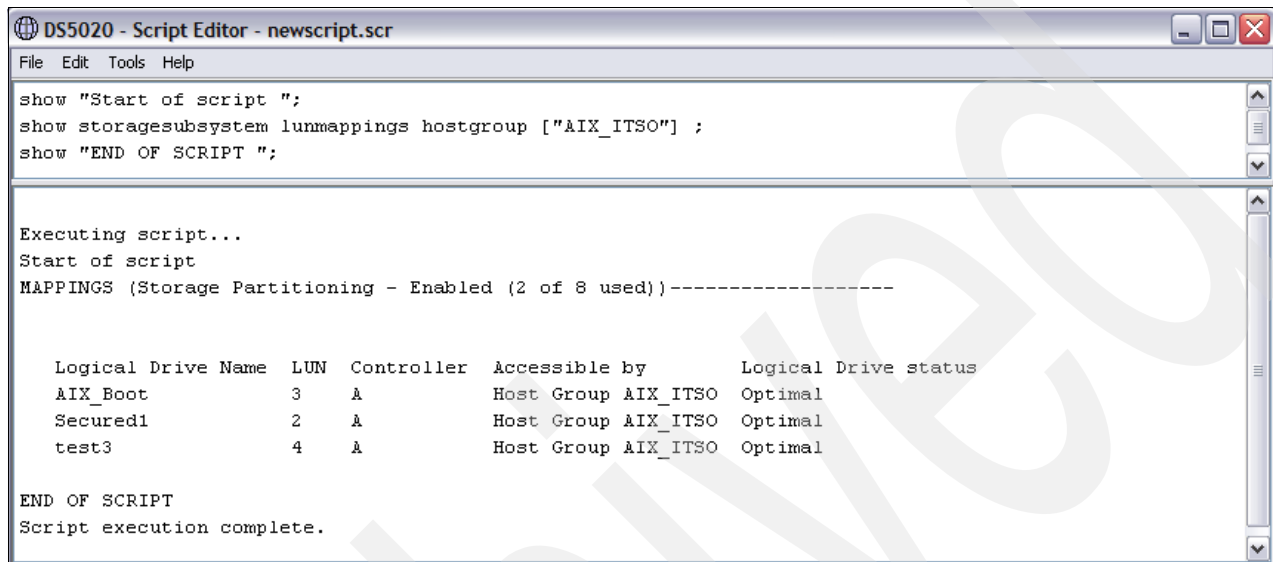
For example, the comment The following command assigns hot spare drives is included for clarification and is not processed by the Script Editor:

```
/* The following command assigns hot spare drives.*/  
set drives [1,2 1,3] hotspare=true;
```

Important: The comment must start with /* and end with */. If the script engine does not find both a beginning and ending comment notation, an error message is displayed and the script execution is terminated.

Using the show command

Use the **show** command with a string (enclosed in a double quotes) and no options to embed in your script comments that display in the output view during script execution. For example, including a **show "START of the script"** command in your script results in the display of START of the script in the output view when this line is processed during script execution. Including **show** statements can help you when writing longer scripts where intermediate steps can make troubleshooting easier, as shown in Figure 8-3.



The screenshot shows a window titled "DS5020 - Script Editor - newscript.scr". The menu bar includes "File", "Edit", "Tools", and "Help". The script content is as follows:

```
show "Start of script ";
show storagesubsystem lunmappings hostgroup ["AIX_ITSO"] ;
show "END OF SCRIPT ";
```

The output view shows the following execution results:

```
Executing script...
Start of script
MAPPINGS (Storage Partitioning - Enabled (2 of 8 used))-----

Logical Drive Name  LUN  Controller  Accessible by      Logical Drive status
AIX_Boot            3    A           Host Group AIX_ITSO  Optimal
Secured1            2    A           Host Group AIX_ITSO  Optimal
test3               4    A           Host Group AIX_ITSO  Optimal

END OF SCRIPT
Script execution complete.
```

Figure 8-3 Show command in Script Editor

Script Editor tools

The Script Editor offers the following tools to help you when writing scripts:

- ▶ Verify Syntax

To run this option, select **Tools** → **Verify Syntax** from the drop-down menu. The Script Editor engine parses the statements in the script file one line at a time and verifies that it has the correct syntax. Any syntax errors are displayed in the output view, reporting the line number of the error and a description of the error. If the Script Editor encounters a syntax error, no further syntax verification is performed on the script. Fix the syntax error and rerun the **Verify Syntax** command to validate the error correction and check the remainder of the statements in the script.

- ▶ Verify and Execute

To run this option, select the **Tools** → **Verify and Execute** option. The Script Editor engine parses the command statements in the script, interprets and converts the statements to the appropriate commands, and sends the commands to the storage subsystem.

If a *syntax error* is encountered, the execution stops and an error message is displayed. Fix the error, then use the Verify Syntax or Verify and Execute options to validate the error correction.

If an *execution error* occurs, the script might or might not continue to execute depending on the included On Error script statement. The On Error Stop statement stops the script if an execution error is encountered. The On Error Continue statement allows the script to continue even after an execution error is encountered. (This is the default.)

► **Execute Only**

To run this option, select the **Tools** → **Execute Only** option. The Script Editor engine executes a script. It displays an error message if a syntax error is encountered.

If an execution error occurs, the script might or might not continue to execute depending on the included On Error script statement. The On Error Stop statement stops the script if an execution error is encountered. The On Error Continue statement allows the script to continue even after an execution error is encountered. (This is the default.)

Note: Certain execution errors, including the inability to communicate with the storage subsystem, always cause the script execution to halt. In these cases, execution stops even if you have used the On Error Continue statement.

Interpreting the script execution results

During script execution, messages are displayed in the output view, beginning with:

Executing script...

After a successful script execution, you see the message:

Script execution complete.

If there is an error during the parse phase, an error indication is displayed in the Output View, giving the line and column number and a description of the syntax error.


If there is an error during execution, a message is displayed in the Output View stating that the command failed and reporting a description of the error.

8.2.2 Embedding commands in batch files

Due to the business demand for higher availability of data, the time window allocated to make backups is shrinking. Customers can no longer afford the daily downtime on their production server to perform backups. This becomes especially true as databases become larger and larger. Online backups improve the availability of the database, but cost CPU, disk, and network resources. Even with the use of incremental and differential backups, the time necessary to perform a backup can be significant.

On the other hand, a feature such FlashCopy enables customers to perform offline backups by creating a point in time copy of a logical device, which can then be presented to a host and backed up. A script can be written, such as the example in Example 8-8 on page 496, which will execute a FlashCopy procedure from the host at a chosen time of day using UNIX “cron” scheduling, for example. Thus, CLI and scripting can become an important tool in managing the DS5000 and host operations.

Archived



Overview of IBM System Storage DS5000 RAID types

In this appendix, we explain the various RAID levels that you can configure when setting up disk arrays on the DS5000 storage subsystems. RAID is an acronym for Redundant Array of Independent Disks. Therefore, the basic structure of RAID is an array of disk drives.

The array is a collection of drives that is configured, formatted, and managed in a particular way. The number of drives in the array, and the way that data is split between them, is what determines the RAID level, the capacity of the array, and its overall performance and data protection characteristics. Deciding what types of arrays to set up, and how to configure them, is the first thing you do when setting up a RAID implementation.

DS5000 arrays and RAID levels

An array is a set of drives that the system logically groups together to provide one or more logical drives to an application host or cluster. When defining arrays, you often have to compromise among capacity, performance, and redundancy. The DS5000 supports different number of RAID types.

Table A-1 shows a support matrix for DS4000 and DS5000 systems to the RAID levels.

Table A-1 RAID support matrix for DS4000 and DS5000 storage subsystems

Subsystem	RAID 0, 1, 10, 3, 5	RAID 6
DS4700 (CFW 6.xx)	Yes	No
DS4700 (CFW 7.xx)	Yes	Yes
DS5100/DS5300	Yes	Yes
DS5020	Yes	Yes

Generally, RAID 6 support is implemented in controller firmware version (CFW) 7.xx and above, but it also requires the DS4000 and DS5000 XOR chip to be able to handle the RAID 6 algorithm. DS4700 storage subsystem is the only DS4000 storage subsystems that is able to support RAID 6 with CFW 7.xx and higher.

The DS4000 and DS5000 are also able to dynamically change the RAID type without downtime. This feature is called Dynamic RAID Migration (DRM).

RAID levels

We go through the different RAID levels and explain why we choose a particular setting in a particular situation; you can draw your own conclusions.

RAID 0: For performance, but generally not recommended

RAID 0 (Figure A-1 on page 507) 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.

Since CFW 7.10, there is no limit for drives that belong to one RAID 0. CFW 6.xx limits the number of drives to a maximum of 30 drives.

There is no redundancy, so any disk failures require reloading from backups. Select RAID 0 for applications that will benefit from the increased performance capabilities of this RAID level, such as high performance computing (HPC) jobs. Never use this level for critical applications that require high availability.

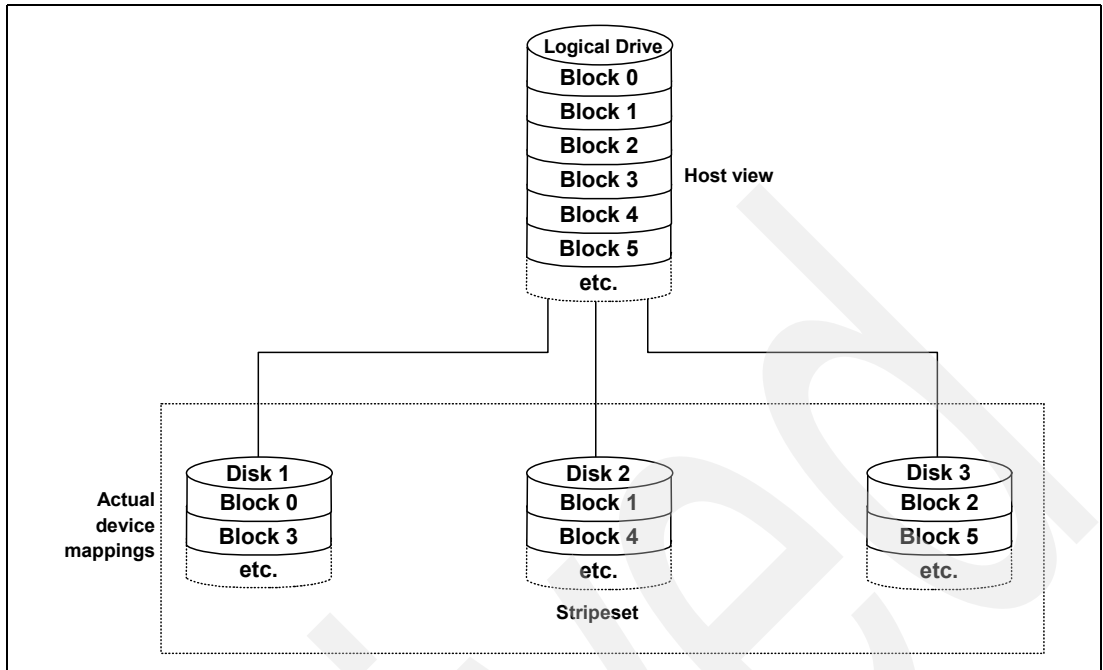


Figure A-1 RAID 0

RAID 1: For availability/good read response time

RAID 1 (Figure A-2) 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 cost is not the major concern.

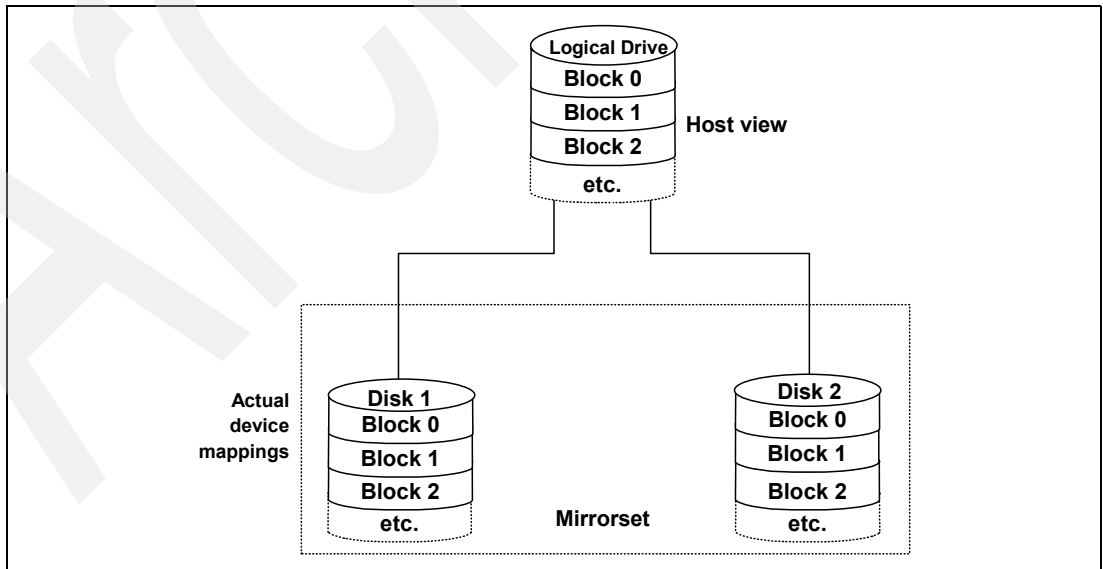


Figure A-2 RAID 1

Because the data is mirrored, the capacity of the logical drive when assigned RAID level 1 is 50% of the array capacity.

Here are some recommendations when using RAID 1:

- ▶ 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 completed on both disks, performance of writes can be improved through the use of a write cache. When using a write cache, be sure it is battery-backed up.

Note: RAID 1 is actually implemented only as RAID 10 (see “RAID 10: Higher performance than RAID 1” on page 511) on the DS5000 storage subsystem products.

RAID 3: Sequential access to large files

RAID 3 is a parallel process array mechanism, where all drives in the array operate in unison. Similar to data striping, information 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; with RAID 3, a dedicated disk is used, unlike RAID 5, where parity information is spread over all the drives in the array.

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. It 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 impact incurred by mirroring in RAID 1.

RAID 5: High availability and fewer writes than reads

RAID 5 (Figure A-3) stripes data and parity across all drives in the array. RAID 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.

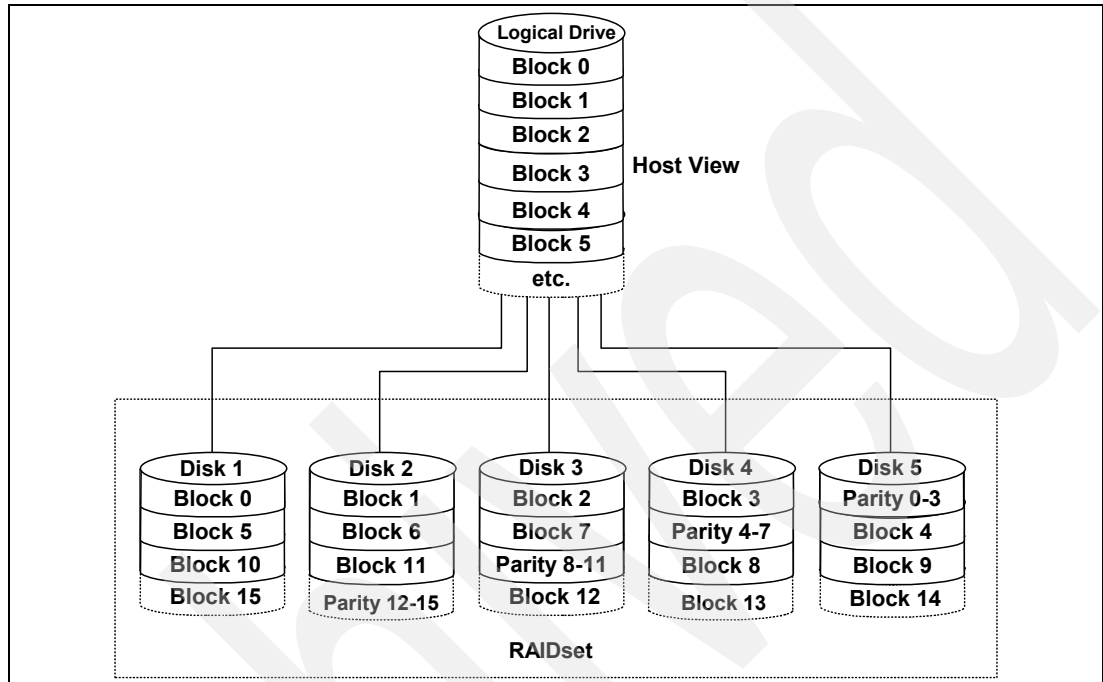


Figure A-3 RAID 5

RAID 5 is best used in environments requiring high availability and 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 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 have occurred. The response time of writes will be improved through the use of write cache (be sure it is battery-backed up). RAID 5 arrays with caching can give as good as 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

RAID 6 (see Figure A-4) is a RAID level employing $n+2$ drives, which can survive the simultaneous failure of any two drives. 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 DS5000 storage subsystem, this calculation is made by the hardware and thus there is more performance than the software-based implementation found in other storage subsystems.

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 impact of more parity calculations, in terms of writing data, RAID 6 is slower than RAID 5, but might be faster in random reads thanks to the spreading of data over one more disk.

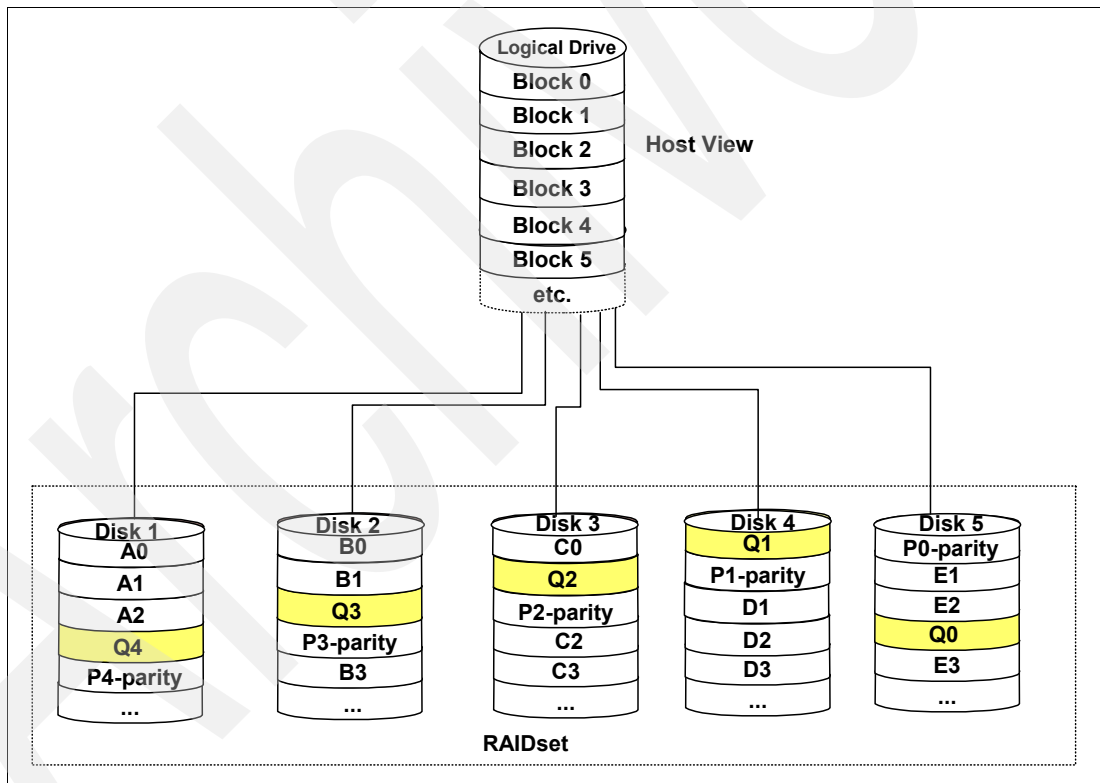


Figure A-4 RAID 6

RAID 6 is striping with Dual Rotational Parity but not “dual parity”.

Table A-2 shows the RAID 6 details.

Table A-2 RAID 6 details

Feature	Description
Definition	Distributed parity; Disk striping and two independent parity blocks per stripe Can survive the loss of two disks without losing data
Benefits	Data redundancy, high read rates, and good performance.
Considerations	Requires two sets of parity data for each write operation, resulting in a significant decrease in write performance. Additional costs because of the extra capacity required by using two parity blocks per stripe.
Uses	Any application that has high read request rates and average write request rates. Transaction servers, Web servers, data mining applications, and Exchange servers.
Drives	Minimum of three.
Fault Tolerance	Yes.

RAID 10: Higher performance than RAID 1

RAID 10 (Figure A-5), 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.

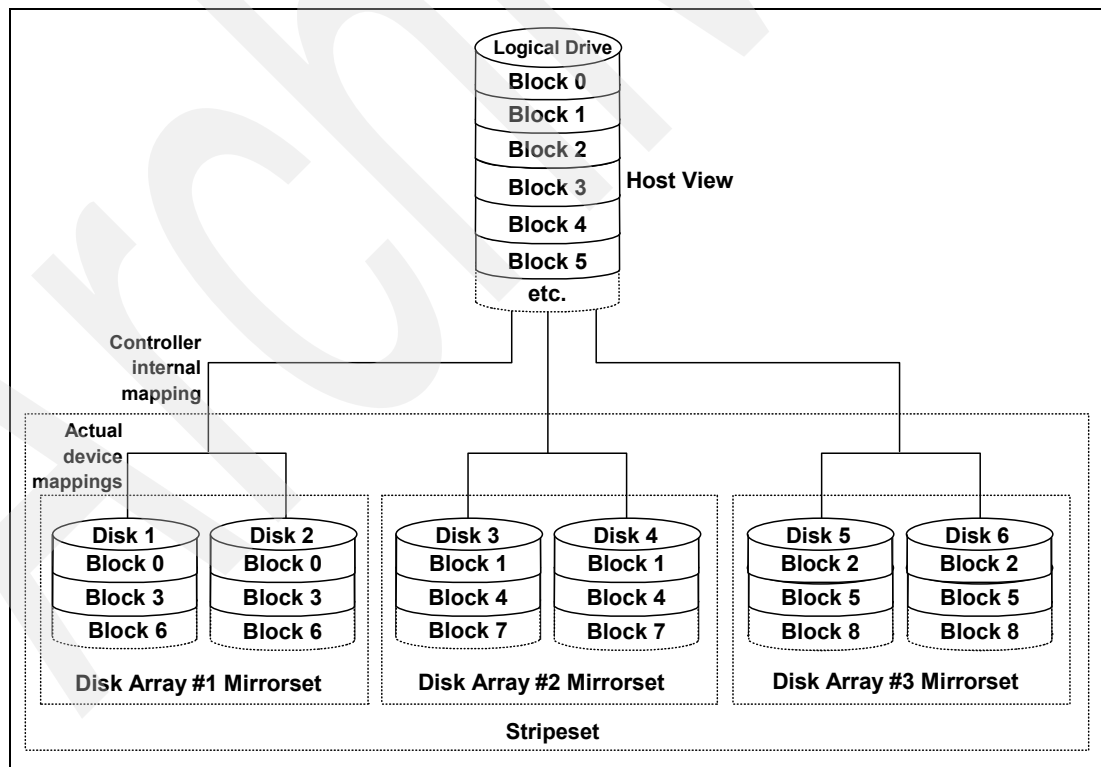


Figure A-5 RAID 10

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.

The recommendations for using RAID 10 are as follows:

- ▶ 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 a RAID 10 write will not be completed until both writes have been done, write performance can be improved through the use of a write cache (be sure 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.

RAID summary

The following note and Table A-3 on page 513 summarize this information.

Summary: 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 offers high availability with performance slightly lower than RAID 5.
- ▶ RAID 10 offers higher performance than RAID 1 and more reliability than RAID 5.

Table A-3 RAID levels comparison

RAID	Description	Application	Advantage	Disadvantage
0	Stripes data across multiple drives.	IOPS Mbps	Performance, due to parallel operation of the access.	No redundancy. If one drive fails, the data is lost.
1	The 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 the same number of disks.	IOPS	Performance, as multiple requests can be fulfilled simultaneously with better data protection.	Storage costs are doubled.
3	Drives operate 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, and graphics).	Degraded performance with 8-9 I/O threads, random IOPS, and smaller, more numerous IOPS.
5	Drives operate 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. This is the most reliable RAID level on the DS5000 storage subsystem	Slower in writing data, complex RAID controller architecture.

RAID reliability considerations

At first glance, both RAID 3 and RAID 5 appear to provide good protection against drive failure. With today's high-reliability drives, it appears unlikely that a second drive in an array will fail (causing data loss) before an initial failed drive can be replaced. But if you look at RAID 6 and calculate the possibility of data loss, the chance to lose data is theoretically much less than on RAID 3 and RAID 5.

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 rebuild 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 subsystems (including the DS5000 storage subsystem) 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 faultless, it will not protect your data from accidental corruption or deletion by program error or operator error. Therefore, all crucial data should be backed up by the appropriate software, according to business needs.

Deploying iSCSI host interface controllers on the IBM System Storage DS5000 series

This appendix reviews planning considerations when deploying IBM System Storage DS5000 storage subsystems in environments with mixed-host-interface requirements. The new Internet Small Computer Systems Interface (iSCSI) feature that has been added to the IBM System Storage DS5300 and IBM System Storage DS5100, and included in the IBM System Storage DS5020. Therefore, a number of new areas must be factored in when deploying the storage subsystems when you plan to use iSCSI. In the mixed interface environment (also referred to as SAN tiering), other iSCSI considerations include networking, performance, and security.

The DS5000 series of storage subsystems now support intermixed iSCSI and FC host interfaces. Using both interfaces provides a greater degree of flexibility in configuration and deployment of a consolidated storage solution to individual servers or hosts attached to a SAN.

iSCSI technology

iSCSI is a protocol that transmits SCSI commands and data over a standard Transmission Control Protocol/Internet Protocol (TCP/IP) based network. The advantage to using a TCP/IP-based network instead of an Fibre Channel based infrastructure is largely based on the cost associated with implementation and support of a Fibre Channel (FC) SAN. In many organizations, a TCP/IP based network already exists and is already being utilized for IP traffic over Local Area Network (LAN) and Wide Area Network (WAN). More recently, the TCP/IP network has also been used for voice traffic with the development of Voice over IP (VoIP). So for some organizations, further use of this existing resource with iSCSI makes sound economic sense.

The iSCSI protocol can be used to transmit data from the connected server to the storage subsystem and vice versa. Generically, the servers are called “initiators”, and the storage subsystem iSCSI ports, such as those available on the IBM DS5000 storage subsystems, are known as “targets”. The targets will listen for connection requests using TCP port 3260 (you can use another port if necessary). You may create other Ethernet ports to handle subsequent communications after a connection between initiator and target is established. An active iSCSI connection is called a *session*, and a session is required for actual data transfer to occur.

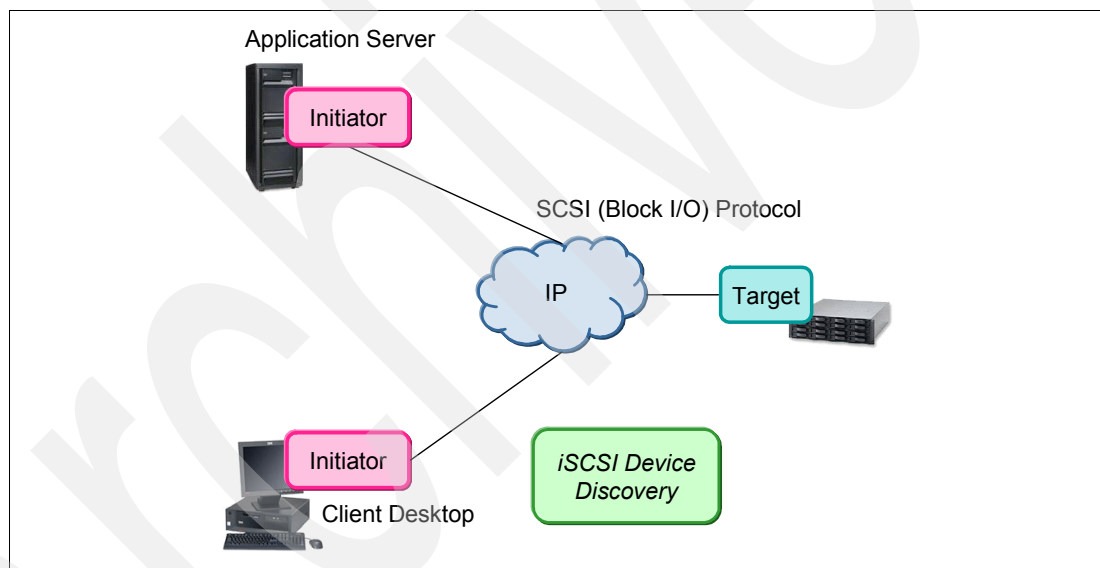


Figure B-1 iSCSI communication components

For the initiator to communicate with the target, you must assign a unique name to both, so that this communication layer can be established. Two types of naming conventions are associated with iSCSI: the IQN, or iSCSI qualified name, and the EUI, or the IEEE EUI-64 identifier.

Support is provided for IQN names on the IBM DS5000 series of storage subsystems.

iSCSI Qualified Name (IQN)

Four major sections comprise an IQN:

- ▶ Type
- ▶ Date
- ▶ Naming authority
- ▶ Unique string

The first section, the type, is the “IQN” string, which indicates that this is an iSCSI qualified name. A period delineates the next field, which is a date code. This date corresponds to the next field, which is a reversed domain name. The date should be the point in time at which the domain name was registered.

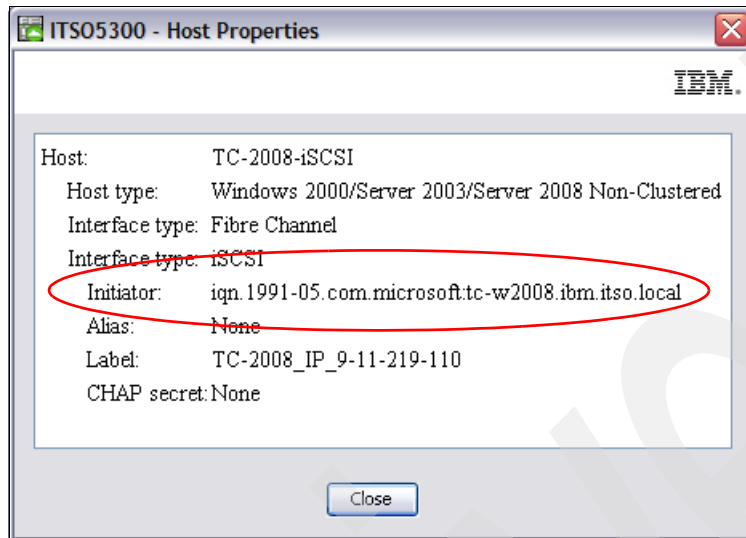


Figure B-2 Example of an IQN name in a defined host

The combination of the domain name and date indicate that the name is based on a registered authority. A colon separates the first three fields (the qualification) from the unique string that identifies the device. Many initiators and targets allow you to alter the iSCSI target name, although you cannot alter iSCSI target names in the IBM DS5000 platform. Note that even though iSCSI is in use, the underlying communication protocol is still TCP/IP, so you still need to assign valid IP addresses, along with normal associated IP configuration parameters to all of the appropriate iSCSI initiator and target ports. IQNs are used with iSCSI to allow for name translation services in the event that an underlying IP address were to change for a given initiator or target, for example, using Dynamic Host Control Protocol (DHCP).

After a session has been established between the target and the initiator, you can continue storage provisioning and usage in a manner similar to that of an FC-based storage subsystem.

iSCSI physical components

After planning the DS5000 as iSCSI target storage subsystem, there are several other components that you use to configure an iSCSI SAN. Of particular interest is the connectivity point for the iSCSI communications within the physical host server itself. You can use several different options to allow iSCSI network traffic to interface with the underlying server hardware. These eventually connect to the operating system and application by using hardware-accelerated host bus adapters (HBAs) and software initiators.

Generally, three types of options are available, as detailed in Table B-1.

Table B-1 Adapter technologies comparison

	NIC	TOE	iSCSI HBA
Description	Network Interface Card: Provides Ethernet connectivity.	TCP/IP Offload Engine: A specialized NIC that provides additional functionality.	An HBA that provides Ethernet connectivity and additional functionality.
Function	Physical and data link communications.	TCP/IP physical and data link communications.	iSCSI read/write processing, TCP/IP processing, and physical and data link communication.
Server CPU resource required	iSCSI protocol management and TCP/IP protocol management.	iSCSI protocol management.	None.
Requires Software based initiator	Yes.	Yes.	No.
Performance	Fair.	Good.	Best.

Network considerations

You can run the iSCSI protocol on the DS5000 series of storage subsystems, but you must consider a number of areas with regard to networked traffic. With normal FC SAN configurations, little planning is required because of the capacity and robustness of the switching and connectivity infrastructure or fabric. This is because Fibre Channel provides a very low-latency and high bandwidth communications medium between the initiator and the target. It has been specifically designed and implemented for high speed data (disk and tape) traffic. With iSCSI, these same statements might not be true and can be subject to external influences outside the traditional SAN itself, such as the existing aforementioned workload of the TCP/IP network.

Because iSCSI is hosted using a traditional Ethernet network, many of the same considerations that apply to file-level Networked Attached Storage (NAS) devices also will apply. You should analyze and understand the impact of the network into which an iSCSI target is to be deployed prior to the actual installation and configuration of a DS5000 storage subsystem.

Before beginning an iSCSI deployment, understand and document the network topology that will be used for the SAN. A poorly understood or inadequate networking infrastructure will inevitably lead to what is perceived as poor storage performance. Ensure that you connect a DS5000 storage subsystem with iSCSI host connectivity to the desired initiator using a dedicated 1 Gbps Ethernet network with as few “hops,” or switched and routed connections, as possible due to propagation loss and latency. Typically, approximately 1 ms of latency is added for every 100 miles of networked infrastructure, as well as additional latency of up to 1 ms for every routed connection. Thus, connectivity between an iSCSI initiator and an iSCSI target over a LAN is considered ideal and allows for adequate response time metrics between a given application and the underlying storage being used.

Several utilities, including ping and trace route, are available at the operating-system level that can be used to help determine network latency between two devices. In addition, you can use other more comprehensive utilities, such as Ethereal, to provide additional insight into the networking infrastructure. In particular, special attention must be made to any Address Resolution Protocol (ARP) packets being sent on the desired iSCSI SAN. ARP requests force a response from all of the devices located on the subnet from which they are broadcast and can have a serious impact on network performance and the underlying iSCSI-based storage subsystem.

iSCSI configurations on the DS5000 series

As well as the basic iSCSI connectivity parameters, such as IP address per target Ethernet port and associated IQN, as detailed in “iSCSI Qualified Name (IQN)” on page 516, you might modify several optional configuration parameters within the context of an iSCSI-enabled DS5000 storage subsystem, including enablement of jumbo frames, configuration of a VLAN, and setting a specific Ethernet priority.

Jumbo frames

Most enterprise gigabit Ethernet equipment provides some degree of support for jumbo frames, or frames larger than the standard 1500-byte payload limit. Enabling jumbo frames has the following effects:

- ▶ They can accelerate iSCSI performance by about 5 percent.
- ▶ They reduce server CPU utilization by 2 percent to 3 percent with standard NICs.

Since TOE cards or HBAs already perform off-loading, the CPU savings from jumbo frames is negligible when jumbo frames are used with a TOE or an HBA. However, jumbo frames can still accelerate performance. When using jumbo frames, ensure that all of the devices on your iSCSI network, including switches, initiators, and targets, are configured to use the same maximum jumbo frame size. Jumbo frame sizes are not standardized and can vary from 1501 bytes to 9000 bytes. If the frame size maximum differs between these components, a potentially serious I/O performance problem can be experienced by the host servers.

For example, if the servers, the storage subsystem, or both are set to a maximum frame size that is larger than the setting to which the switches are configured, a DS5000 storage subsystem might appear to be working perfectly. However, if you start performing large data transfers that exceed the switch's maximum frame size, I/O errors might result. Therefore, if jumbo frames are used, make sure to set up and configure jumbo frames across all devices with an agreed maximum size.

If jumbo frames have been enabled and performance appears to be substantially less than expected, disabling jumbo frames on the DS5000 storage subsystem is a simple and quick troubleshooting step, as shown in Figure B-3. We found that, based on empirical tests of several varieties of iSCSI HBAs, that using jumbo frames led to, on occasion, degraded performance.

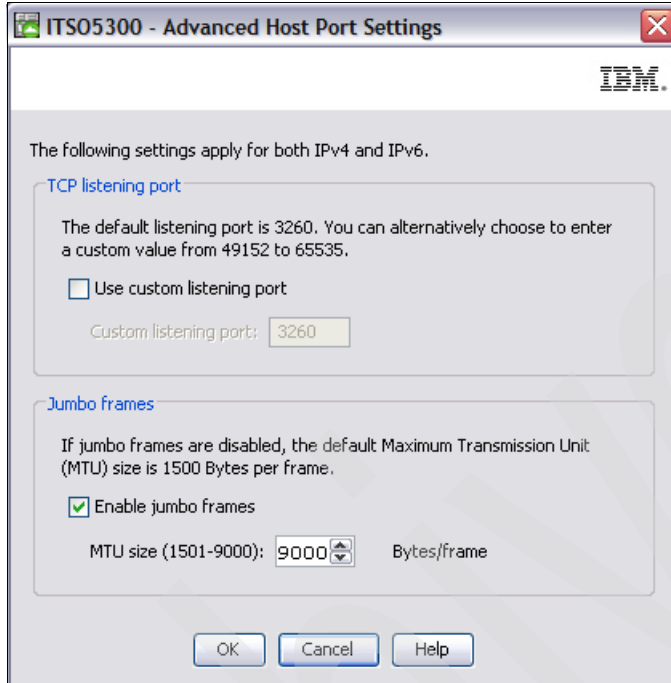


Figure B-3 Jumbo frames enabled on the iSCSI host port

Virtual Local Area Networks

Virtual Local Area Networks (VLANs) are a grouping of devices that communicate on the same broadcast network, even though they might be physically separate. A VLAN has the same basic characteristics of a normal LAN, but it allows for a greater degree of flexibility within a networked infrastructure, and it allows reconfiguration using software instead of physical movement of devices.

In the DS5000 storage subsystems that are connected using iSCSI, you can use an option that enables VLAN support and provides a valid VLAN ID, as shown in Figure B-4. If it is not possible to segregate an iSCSI storage subsystem onto a physically separate LAN, use a VLAN to maximize potential performance.

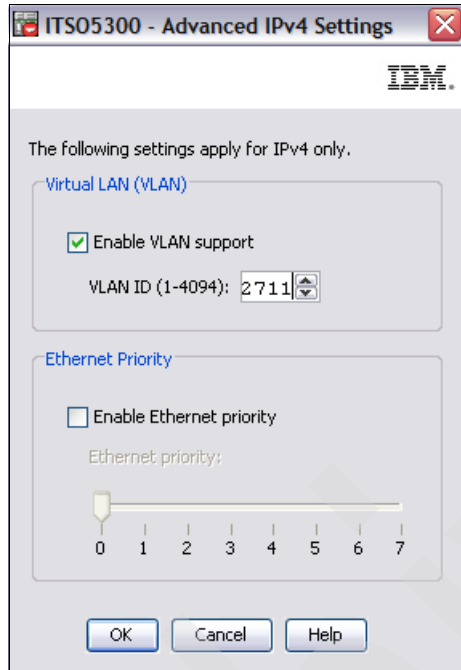


Figure B-4 Ethernet priority and VLAN support for iSCSI port

Ethernet priority

Ethernet priority, sometimes referred to as quality of service or class of service, is a relatively recent addition to the Ethernet specification. The 802.1 Ethernet standards working group has defined an extension to the Media Access Control (MAC) layer that can take into account a user-defined class of service. The 802.1p specification is a standard for traffic prioritization where network frames are tagged with one of eight priority levels using a 3-bit value added to the Tag Control Info (TCI) inside of a standard Ethernet frame, where 7 is high and 0 is low. Switches and routers that are 802.1p compliant can give traffic that is time-sensitive preferential treatment if the priority tag has been set to a higher value than other traffic. Generically, the seven levels are defined as shown in Figure B-4. The seven levels are:

0	Routine (default)
1	Priority
2	Immediate
3	Flash
4	Flash Override
5	Critical
6	Internetwork control
7	Network Control

For the DS5000 storage subsystem, you can modify the Ethernet priority of the target iSCSI interfaces to increase the class of service received within the network itself. Use Ethernet priority on isolated networks, both LANs and VLANs, only where multiple hosts and devices exist. Modifying this value can and will impact the performance of other devices located on the network. Avoid using a priority setting of 7, which is the highest priority.

Security

Several security mechanisms are provided by the DS5000 storage subsystem when an iSCSI Host Interface Card (HIC) is present that is not applicable to standard FC communications. Because the storage subsystems can now be employed on a relatively unsecure Ethernet network topology, additional security might be required to comply with organizational or corporate network security guidelines. You can configure both iSNS and CHAP authentication on the DS5000 storage subsystems.

Internet Storage Name Service

The Internet Storage Name Service (iSNS) protocol allows for automated discovery, management, and configuration of iSCSI devices on TCP/IP network. In a typical iSCSI-based storage subsystem without iSNS configured, any and all iSCSI session requests to the target may be allowed from devices on the iSCSI SAN subject to restrictions in place using the partitioning element of the DS5000 series of storage subsystems. The iSNS servers offer additional security services through explicitly defined initiator-to-target mappings and simplified asset locators, similar to that provided by DNS and WINS for IP address lookup facilities. Most modern operating systems allow iSNS services to be established. The DS5000 series of storage subsystems support this capability as well. Figure B-5 shows the configurations settings where the iSNS box has been checked, allowing the iSNS server to be defined.

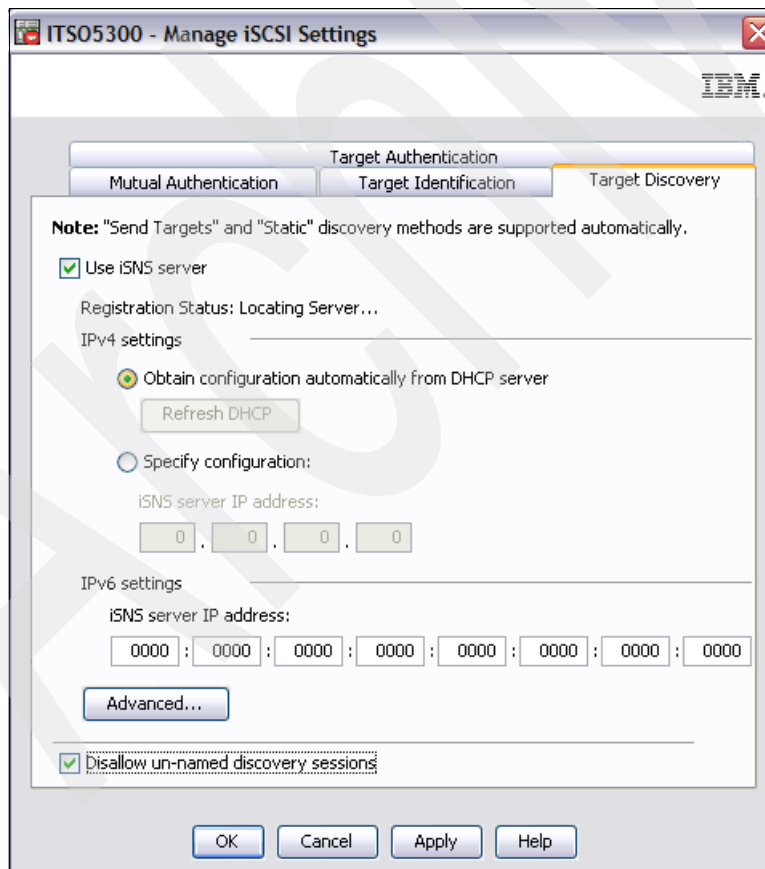


Figure B-5 iSNS settings

The iSNS server listening port is the TCP port number that the controller tries to connect to a server. This process lets the iSNS server register the storage subsystem with the iSCSI target and the iSCSI portals so that the host initiators can find the storage subsystem. The default value for this listening port is 3205.

Challenge Handshake Authentication Protocol

The Challenge Handshake Authentication Protocol (CHAP) provides an additional security layer within the iSCSI SAN on the DS5000 storage subsystems. When CHAP is enabled, the initiator sends the target a random value and an ID value. Both the initiator and the target share a predefined “secret,” or password. The peer then concatenates the random value, the ID, and the secret to calculate a one-way hash using the MD5 hash function. This hash value is then sent back to the initiator, which in turn builds the same string, calculates the MD5 sum, and compares the result with the value received by the target. If the values match, then the iSCSI session is considered established and future communication proceeds with subsequent increases of the ID value to prevent possible replay attacks. By default, CHAP is not enabled nor enforced on the DS5000 storage subsystem and this method is highly recommended for enhanced security. Figure B-6 shows the CHAP setup in the iSCSI settings window.

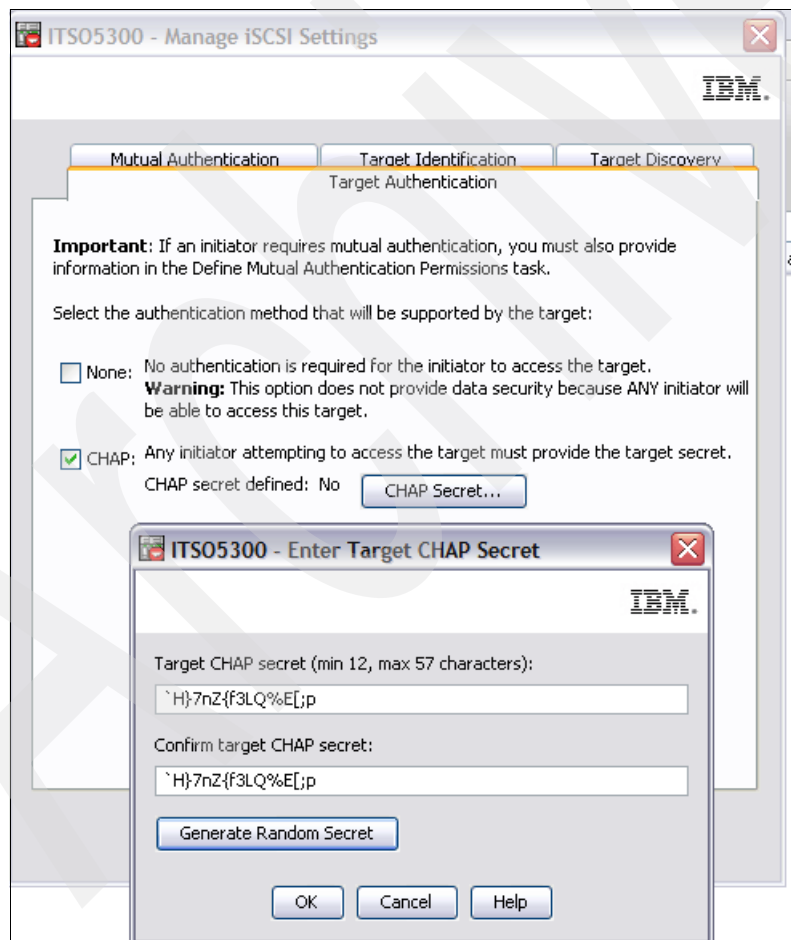


Figure B-6 Configuring CHAP and secret

The CHAP secret must be between 12 characters and 57 characters. The CHAP secret must use ASCII code characters with a decimal value between 32 and 126. In Figure B-6, the utility is used to generate a “secret”.

iSCSI performance considerations

A Fibre Channel SAN almost always yields higher performance than an iSCSI SAN. FC does not suffer from congestion and oversubscription problems as readily as iSCSI due to both the FC protocol and the general segregation of FC traffic itself. FC includes a mechanism for port-to-port throttling, not just end-to-end throttling, and can limit the amount of data that a connected system sends. This mechanism avoids situations in which information must be dropped and avoids degradation into a retransmission recovery scheme that can limit the storage network's performance. For this reason, and many others, including general protocol efficiency, FC is more capable of operating near-wire speed than iSCSI is likely to be.

From the host perspective, you must consider several elements before deploying an iSCSI SAN. These considerations include the number of physical NICs that will be installed in the hosts themselves, the types of NICs (see Table B-1 on page 518), whether a software initiator or a hardware initiator will be used, how many paths will exist between the initiator and the target, which failover, load-balancing driver, or both may be used and, most importantly, what speed will the infrastructure operate. Generically speaking, iSCSI solutions might be a good fit for those environments where the bulk of the data processing activity is considered to be of the random, small-block variety. Given the relatively modest throughput available within a current iSCSI SAN compared to 8 Gb FC fabrics, it is not likely that adequate throughput will be available within an Ethernet network or the specific target ports to handle workloads that consist primarily of large, sequential transactions. This behavior can be more easily seen using the following simplified equation:

Throughput = IOPS x I/O size

Thus, as the size of the inbound I/Os decrease and typically become more random in nature, as in some applications, databases and general file sharing, the amount of throughput required also is reduced. Conversely, as the size of the inbound I/Os increase and typically become more sequential in nature, such as video editing or disk-to-disk backup, the amount of throughput required greatly increases.

Initial testing of the DS5000 storage subsystems configured with 1 Gb iSCSI HICs has shown that a heavily configured storage subsystem will likely offer a greater proportion of the "small block random I/O" performance to an 8 Gb FC configuration. Alternatively, given the maximum of four 1 Gb iSCSI target ports currently available on the DS5000 series of storage subsystems, enough available bandwidth might not be available to provide high levels of "large block sequential" performance relative to what will potentially be available using multiple 8 Gb FC interfaces. The performance difference between the two interface types can become very large in this type of environment.

One of the most obvious conclusions given is that, typically, 1 Gb iSCSI is not a good fit for environments or applications that require high throughput.

After it has been determined whether a given application might not be a logical candidate for iSCSI connectivity based on throughput requirements, you need to consider whether you will use iSCSI HBAs or use on-board NICs with software initiators. The DS5000 storage subsystems support both methods, and you will make this determination on a cost to required performance basis. In principle, a software initiator performs nearly as well as a hardware-based initiator in terms of driving storage I/Os. The primary difference is largely in the amount of CPU cycles that are required to become devoted to handling iSCSI traffic in the absence of a hardware-based initiator with appropriate TOE support.

Multipathing iSCSI

As with Fibre Channel SANs, iSCSI SANs, in concert with the DS5000 series, offer the ability to provide failover to the alternate controller in the event of an outage situation. In addition, many failover drivers, such as the native MPIO driver in the Windows 2003 and Windows 2008 operating systems, when combined with the IBM-provided DSM, also offer load-balancing across available iSCSI routes between the target and the initiator as well. Therefore, you still need to configure iSCSI with multiple NIC interfaces for optimal performance and reliability for initiators connecting to a DS5300, DS5100, or DS5020 storage subsystem.

Other iSCSI performance considerations

You must consider several other factors when working with a mixed-host interface environment due to various limitations on the intermixing of the physical protocols at the controller level and the host level.

- ▶ A given host group must not contain hosts that are FC-based as well as hosts that are iSCSI-based.
- ▶ A single host must not be configured for both iSCSI connections and FC connections to the storage subsystem.
- ▶ The Remote Mirroring premium feature is only supported using FC connectivity.

Generically, even though both protocols are available within the storage subsystem itself, take care that at the host and host group level, these protocols remain independent and are used as a tiering strategy across the entire host pool rather than within a given host group or on a singular host itself.

As described previously, FC and iSCSI provide radically different latency and throughput capabilities, and this mixture within an initiator environment can be prone to failover driver conflict, performance degradation, and data loss.

Archived

Solid State Drives on the IBM System Storage DS5000 series

In this section, we discuss Solid State Drives (SSD), which is a new feature of the IBM System Storage DS series that was introduced with the DS5000 storage subsystems. We compare this technology to hard disk drives (HDD), and we base this comparison on tests we ran that compared SSDs to HDDs. Lab tests show a significant performance advantage with SSDs with a substantial reduction in the number of drives needed to meet the desired level of performance. Fewer drives translate into a smaller physical footprint, reduced energy consumption, and less hardware to maintain. Tests also showed better application response times using SSDs, which leads to increased productivity and higher customer satisfaction.

SSD technology was introduced more than two decades ago. Until recently, however, the high cost-per-gigabyte and limited capacity of SSDs restricted deployment of these drives to niche markets or military applications. Recent advances in SSD technology and economies of scale have driven down their cost, making them a viable storage option for many I/O intensive enterprise applications. While the cost of SSDs is trending downward, the dollar per gigabyte cost for SSDs is still substantially higher than that of HDDs. It is not cost-effective or necessary to replace all HDDs with SSDs. For example, infrequently accessed (cold) data can reside on lower cost HDDs, and frequently accessed (hot) data can be moved to SSDs for maximum performance. The appropriate mix of SSDs and HDDs can be used to strike a proper balance between performance and cost.

SSD technology

Solid State Drives are built mainly from flash memory and augmented with Dynamic Random Access Memory (DRAM) along with a sophisticated internal controller. Flash memory based on the NAND (the logical “Not And” operation) has been available for nearly two decades, and is used in several high-volume consumer electronics applications, including cell phones, PDAs, and MP3 players. Flash memory is non-volatile (it retains data without a power source), involves no mechanical parts, and can be manufactured as standard components in high volume.

Current flash technology manipulates a charge on the floating gate of specially designed transistors to allow representation of two (voltage) states, which translates to a single bit per cell, and is called single layer cell (SLC). SLC NAND-based flash has been dropping in price faster than DRAM and HDDs, and now sits between them in a price range approximately twenty to thirty times more expensive than HDDs but up to ten times cheaper than DRAM. At ten times the price and one hundred times the performance, the value proposition is very compelling for high IOPS applications. This explains why the technology has been around for some time but is just now being introduced in storage solutions, including the DS5000 series.

The flash technology is still evolving and we can expect to see a multiple layer cell or MLC in the future. MLC is designed to allow a more precise amount of charge on the floating gate of the transistor to represent four different states, so translating to two bits of information per cell. This higher density per cell and the potential to store three or more bits of information per cell provides for continuing cost improvement in the coming years. If MLC is successful, flash memory could reach down to only two to three times the price of high-performance HDDs (FC or SAS 15K RPM) in the foreseeable future, close enough to drive a significant substitution rate of low-capacity/high-performance HDDs.

Solid State Drives in tiered storage

Many storage environments have grown to support a diversity of needs and evolved into disparate technologies that have lead to storage sprawl. In a large-scale storage infrastructure, this yields a sub-optimal storage design that can be improved with a focus on data access characteristics analysis and management.

Tiered storage is an approach of utilizing different types of storage throughout the storage infrastructure. It is a mix of higher performing and higher cost storage with lower performing and lower cost storage and placing data accordingly based on specific characteristics, such as performance needs, age, and importance of data availability.

An example of an existing storage environment is shown in Figure C-1. The design results in a significantly increased cost associated with maintaining and supporting the infrastructure. In addition to the immediate effect associated with this balance, growth continues at an increased rate in the higher cost area of Tier 1. Thus, as the growth occurs, the distribution of data continues to grow in a non-optimal direction unless there is careful planning and discipline in deployment.

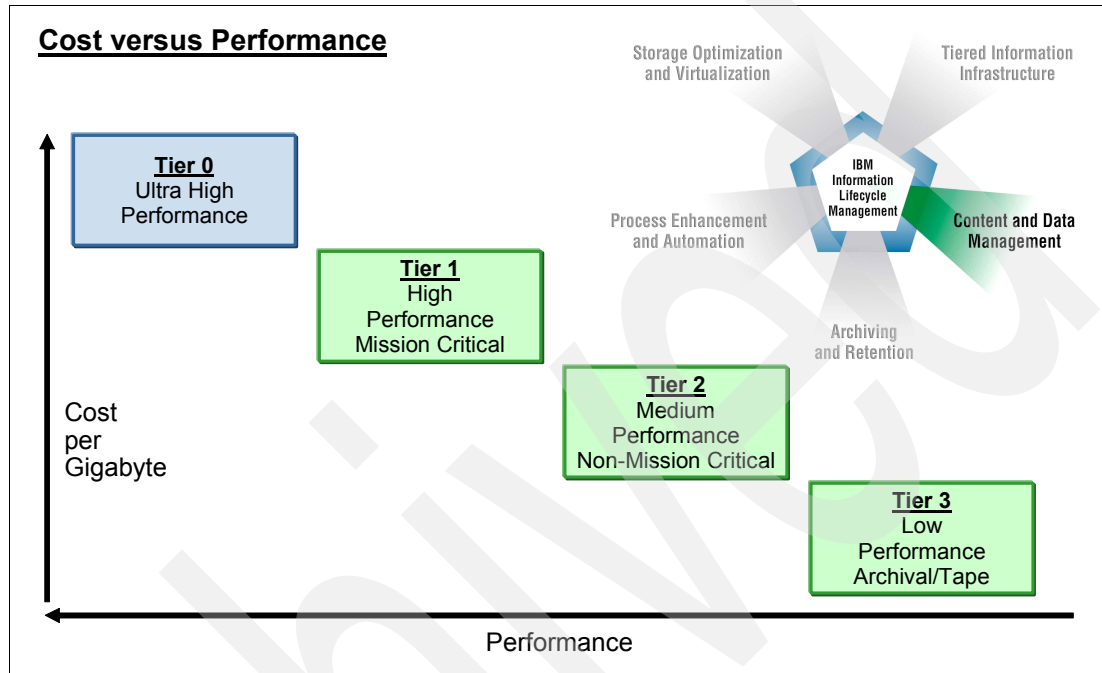


Figure C-1 Tiered Storage: Cost versus performance

Properly balancing these tiers leads to the minimal cost and best performance solution. Typically, an optimal design keeps the active operational data in Tier 0 and Tier 1 and uses Tiers 2 and 3 for less active data. An example is shown in Figure C-2. The benefits associated with a tiered storage approach are simple; it is all cost related. This approach will save significant cost associated with storage itself, as lower tiered storage is less expensive. Beyond that, there are the environmental savings, such as energy, footprint, and cooling reductions.

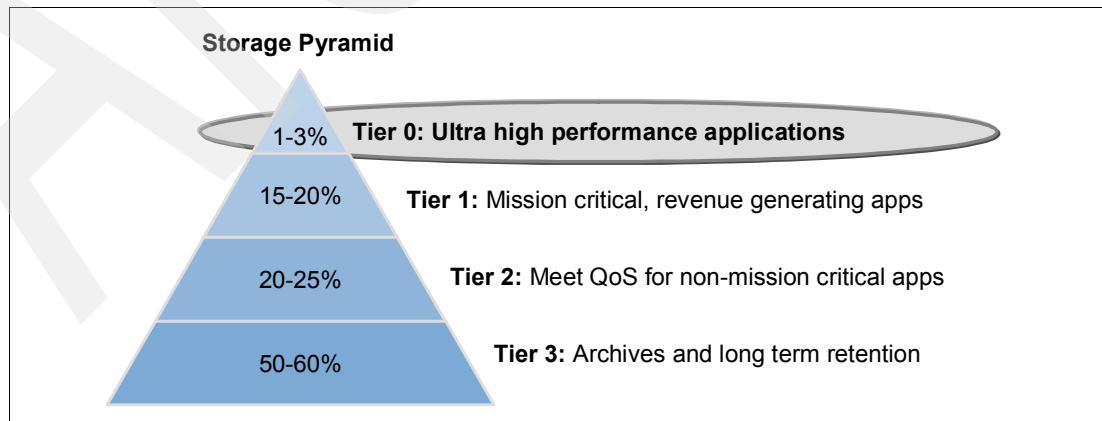


Figure C-2 Example of a tiered storage pyramid

On a DS5000 storage subsystem, delivering these tiers will be in the form of RAID level selection and disk types, so SDD will clearly be used for Tier 0 followed by FC 15k drives and SATA for the lower tiers. Longer term retentions may even be put on cheaper, external, or transportable media, such as tape, to be kept in libraries or a storage facility away from the data center.

Implementing tiered storage

There are three areas of interest critical to implementing, maintaining, and leveraging a tiered storage solution. These areas are:

- ▶ software tools: For identification and reporting of all components of the tiered storage solution.
- ▶ virtualization: To enable control and allocation of your solution.
- ▶ Offerings that are designed to provide alignment with your specific needs for IT governance.

One useful methodology is Information Lifecycle Management (ILM). ILM is the process of managing information, from creation, through its useful life, to its eventual destruction, in a manner that aligns storage costs with the changing business value of information. A more recent IBM offering is the Novus - Intelligent Storage Service Catalog (ISSC) offering, which is a single framework aimed at providing storage optimization through more efficient provisioning, better analytics of the storage environment, and proper alignment of data to storage tiers. You can obtain more information about this topic at the following address:

<http://www-935.ibm.com/services/us/index.wss/offering/its/a1031483>

Solid State Drives on a DS5000 storage subsystem

This new, supported feature on the DS5000 series is supported as a base product on DS5300, and DS5100 only with an EXP5000 expansion enclosure. To install SSDs on a DS5000 storage subsystem, there are some requirements and limitations.

The SSD emulates a conventional hard disk drive, thus easily replacing it in any application. SSDs are available with the same interfaces used by hard disk drives: Serial Advanced Technology Attachment (SATA) and Fibre Channel (FC).

The advantages of SSDs over hard disk drives in a DS5000 storage subsystem include:

- ▶ Faster start up (no spin up)
- ▶ Faster access to data (no rotational latency or seek time)
- ▶ Higher I/O operations per second (IOPS)
- ▶ Higher reliability
- ▶ Lower power usage
- ▶ Less heat produced and less cooling required

Your DS5000 storage subsystem must meet the following criteria to include SSDs:

- ▶ Firmware V7.60.
- ▶ Only twenty SSDs are supported in a DS5000 storage subsystem.
- ▶ Presently, only 73 GB SSDs are available for a DS5000 storage subsystem.

Identifying SSD in Storage Manager

You can identify SSDs in Storage Manager either by the label “SSD” or an icon, as shown in Figure C-3, in the row “Media Type”.

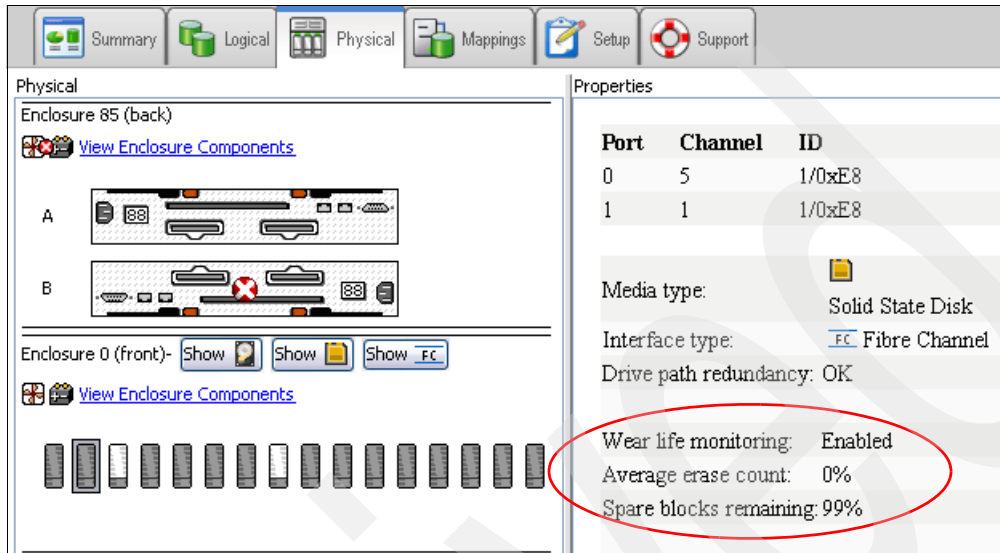


Figure C-3 SSD properties

SSD arrays

The same criteria for arrays apply to SSD as for other type disk drives. All of the disk drives in the SSD array must have the same media type (SSD) and the same interface type (FC only). Hot spare disk drives must also be SSD and be the same type as the disk drives they are protecting.

Wear life

The flash-based SSD has a limited wear life before individual memory locations can no longer reliably store data. The disk drive continuously monitors itself and reports its wear life status to the controller. Two mechanisms exist to alert you that an SSD is nearing the end of its useful life:

- ▶ Average erase count
- ▶ Spare blocks remaining

You can find these two pieces of information in the disk drive properties, which you can see in the storage management software by selecting a disk drive on the **Physical** tab. These are marked in Figure C-3.

The average erase count is reported as a percentage of the rated lifetime. When the average erase count reaches 80 percent, an informational event is logged to the Major Event Log (MEL). At this time, schedule the replacement of the SSD. When the average erase count reaches 90 percent, a critical event is logged, and a “Needs Attention” condition occurs on the DS5000. At this time, replace the SSD as soon as possible.

The spare blocks remaining are reported as a percentage of the total blocks. When the spare blocks remaining falls below 20 percent, an informational event is logged to the MEL. At this time, schedule the replacement of the SSD. When the spare blocks remaining falls below 10 percent, a critical event is logged, and a “Needs Attention” condition occurs. At this time, replace the SSD as soon as possible.

Write caching

Write caching will always be enabled for SSDs. Write caching improves performance and extends the life of the SSD.

Background media scan

You cannot enable background media scans on SSDs. Background media scans are not needed for SSDs because of the high reliability of SSDs. Furthermore, background media scans will be detrimental because they increase wear on the SSDs.

SSD performance on DS5000 storage subsystems

In addition to capacity growth, there is an increasing need to process data quickly. In some cases, the high volume of users accessing a database can result in the need for a high IOPS rate. In other cases, there is a need to quickly process data, and to index the incoming streams to allow high-speed searches and retrieval of needed records. Within the database or file system middleware, there are directories, lock managers, and access control records that must be accessed and updated at rates that scale with the size of the data or the number of clients being served.

A need for high performance disks

There are some applications that simply cannot be run fast enough to satisfy business needs. These include trading algorithms, complex simulations in aerospace or pharmaceutical design, and security video analysis. Online transaction processing (OLTP) systems are the classic example of these applications. Many of these usages of data create a need to operate at high speed, often on indices or subsets of larger collections of information. In cases where the IOPS performance of the storage is the system bottleneck, there is a high value in faster storage and using SSDs.

Over the years, HDDs have maintained a dramatic rate of improvement in terms of dollar per gigabyte, which has enabled data center administrators to keep up with storage capacity demand without greatly increasing expenses. HDDs have also performed relatively well in achieving improvements to sustained bandwidths (GBps) with recent 15K RPM drives advertising greater than 170 MBps speeds. But the rate of improvement in IOPS has lagged far behind other system elements. While the dollar to gigabyte ratio has improved at a rate of 50 percent per year or more over the last decade, IOPS has only increased at a rate of 5 percent per year, and has slowed even further in recent years.

This parameter of performance is dominated by the mechanical elements of the drive: the drive's rotational speed and the seek time of the arm. Substantial improvements for each drive form factor (for example, 3.5") are no longer attainable. Although some minor improvement are possible by using smaller drives (for example, 2.5" at 15K RPM), this comes at a higher dollar to gigabyte cost. This lack of improvement in IOPS means that HDDs are actually getting worse in access density as defined by IO per GB per second (I/O / GBps).

Storage management, performance, and cost are big issues in the database world. Database workloads, both transactional and data warehousing, typically require lots of HDDs for I/O performance, both IOPS and bandwidth. Traditional enterprise HDDs, including the 15K RPM HDDs, are limited by the rate of head movement and deliver random I/O performance of approximately 150 -175 IOPS with a latency of about 5 -7 ms and sequential scan bandwidth of about 30 - 60 MBps for most database workloads. Write-intensive batch jobs are under pressure to complete within the increasingly shrinking time-window, leading to reduced uptime for transactional database systems.

SSDs offer game-changing performance for database applications by removing the limitations traditional rotating disks impose on database design. This revolutionizes database architectural design by removing the traditional I/O bottleneck. SSDs eliminate the need to have a large number of underutilized (short-stroked) HDDs to meet the heavy I/O demands of database applications.

Initial lab tests of SSDs on a DS5000 storage subsystem

Lab tests with a DS5000 storage subsystem were made in order to illustrate a comparison between SSDs on a DS5000 storage subsystem and the traditional SATA and Fibre Channel (FC) disks. The tools used to create the I/O have been created purely for a lab environment, and the results shown below do not indicate any expectation to what a commercial application or database can expect to achieve on a DS5000 storage subsystem. The lab environment ensures that the tests performed are identical and a realistic comparison between the different disk drive results can be made.

Random I/O performance

The SSD limitation of twenty disks per subsystem are compared to SATA limitation of 256 disks and the FC limitation of 256 disks that can be used in a single DS5000 storage subsystem. The first set results, shown in Figure C-4, compare random I/O reads and writes for all three types of disk. The arrays were set up in the following configurations:

- ▶ SSD: 4 x (4+1) volumes in a RAID 5 configuration
- ▶ FC: 32 x (7+1) volumes in a RAID 5 configuration
- ▶ SATA: 32 x (7+1) volumes in a RAID 5 configuration
- ▶ Read cache disabled
- ▶ Write cache disabled

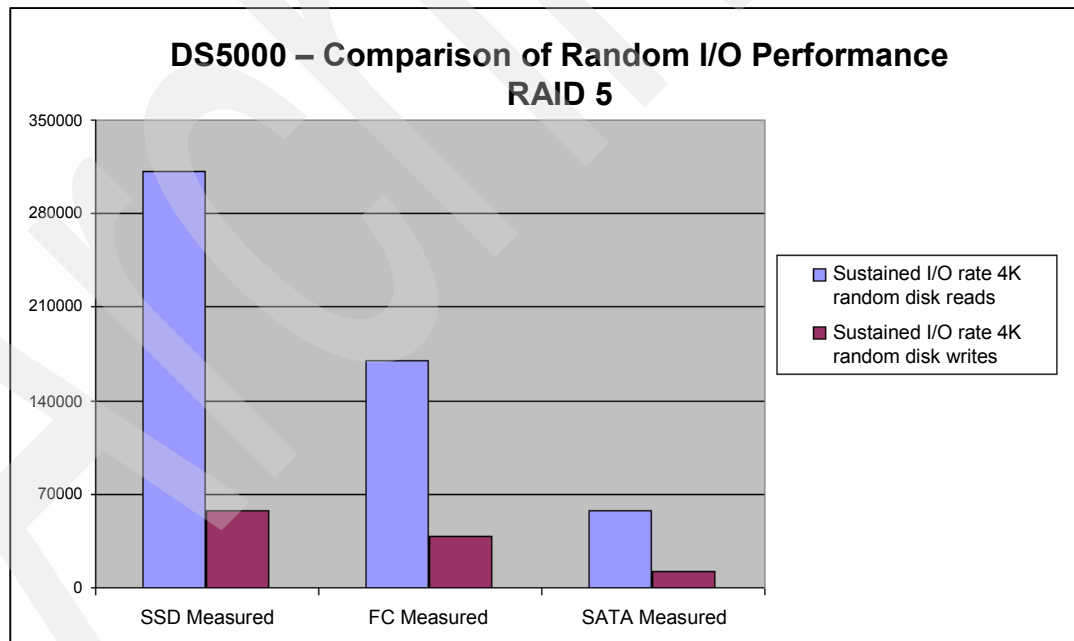


Figure C-4 Comparison of random I/O performance

Throughput performance

The second set of results show the throughput comparison. The arrays were set up in the following configurations:

- ▶ SSD: 4 x (4+1) volumes in a RAID 5 configuration
- ▶ FC: 16 x (8+1) volumes in a RAID 5 configuration
- ▶ SATA: 16 x (8+1) volumes in a RAID 5 configuration

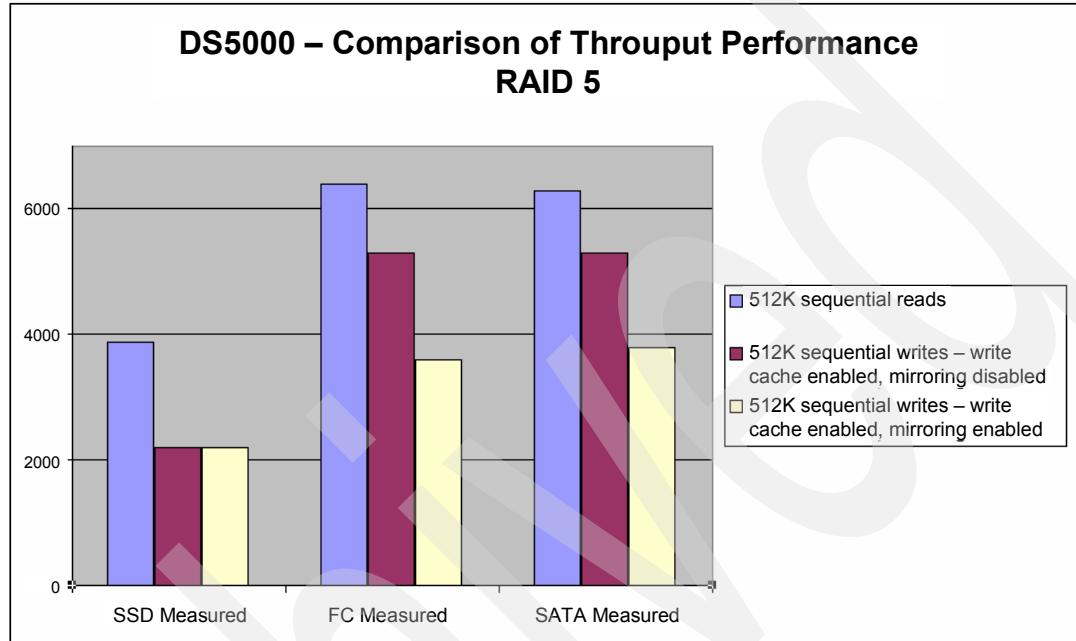


Figure C-5 Throughput performance comparison

SDD summary

From the comparisons given in Figure C-4 on page 533 and Figure C-5, SSD offers a greatly improved performance over HDDs, but SSD costs more. With such a performance, SDD usage can be increasingly justified for use in high performance tiered storage on a DS5000 storage subsystem. Also, SDD usage will be cost driven, and as the cost for SDD diminishes to only two to three times the price of high-performance HDDs, it will drive a significant substitution rate of low-capacity/high-performance HDDs.

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this book.

IBM Redbooks

For information about ordering these publications, see “How to get Redbooks” on page 536. Note that some of the documents referenced here may be available in softcopy only.

- ▶ *IBM Midrange System Storage Copy Services Guide*, SG24-7822
- ▶ *IBM Midrange System Storage Implementation and Best Practices Guide*, SG24-6363
- ▶ *IBM System Storage/Brocade Multiprotocol Routing: An Introduction and Implementation*, SG24-7544
- ▶ *IBM System Storage Business Continuity: Part 1 Planning Guide*, SG24-6547
- ▶ *IBM System Storage Business Continuity: Part 2 Solutions Guide*, SG24-6548
- ▶ *IBM System Storage Copy Services and IBM i: A Guide to Planning and Implementation*, SG24-7103
- ▶ *IBM System Storage DS3000: Introduction and Implementation Guide*, SG24-7065
- ▶ *IBM System Storage DS4000 and Storage Manager V10.30*, SG24-7010
- ▶ *Implementing an IBM/Brocade SAN with 8 Gbps Directors and Switches*, SG24-6116
- ▶ *Implementing an IBM/Cisco SAN*, SG24-7545

Other publications

These publications are also relevant as further information sources:

- ▶ *IBM System Storage DS3000, DS4000, and DS5000 Command Line Interface and Script Commands Programming Guide*, GC52-1275
- ▶ *IBM System Storage DS4000 Concepts Guide*, GC26-7734
- ▶ *IBM System Storage DS4000/DS5000 EXP810 Storage Expansion Enclosure Installation, User's and Maintenance Guide*, GC26-7798
- ▶ *IBM System Storage DS4000/DS5000 Fibre Channel and Serial ATA Intermix Premium Feature Installation Overview*, GC53-1137
- ▶ *IBM System Storage DS4000/DS5000 Hard Drive and Storage Expansion Enclosure Installation and Migration Guide*, GC53-1139
- ▶ *IBM System Storage DS4800 Storage Subsystem Installation, User's, and Maintenance Guide*, GC26-7845
- ▶ *IBM System Storage DS4800 Storage Subsystem Quick Start Guide*, GC27-2148
- ▶ *IBM System Storage DS5000 EXP5000 Storage Expansion Enclosure Installation, User's, and Maintenance Guide*, GC53-1141
- ▶ *IBM System Storage DS5100, DS5300, and EXP5000 Quick Start Guide*, GC53-1134

- ▶ *IBM System Storage DS5100 and DS5300 Storage Subsystems Installation, User's, and Maintenance Guide*, GC53-1140
- ▶ *IBM System Storage DS Storage Manager Version 10 Installation and Host Support Guide*, GC53-1135
- ▶ *IBM System Storage DS Storage Manager Version 10.50 Copy Services User's Guide*, GC53-1136
- ▶ *IBM System Storage DS Storage Manager Version 10.60 Copy Services User's Guide*, GC53-1136

Online resources

These Web sites are also relevant as further information sources:

- ▶ System Storage Interoperation Center (SSIC):
http://www-03.ibm.com/systems/support/storage/config/ssic/displayessearchwithoutjs.wss?start_over=yes
- ▶ Support for IBM Disk systems:
<http://www.ibm.com/systems/support/storage/disk>

How to get Redbooks

You can search for, view, or download Redbooks, Redpapers, 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

Numerics

- 1 Gbps iSCSI host ports 4
- 4 Gbps FC host ports 4
- 8 Gbps FC host interface 19
- 8 Gbps FC host ports 4

A

- access logical drive 146
- access LUN. See access logical drive
- ACS 258
- activation 334
- activation key 365
- Additional Mapping 217
- ADT 138, 205, 255
- airflow 377
- alarm 174
- alarm log 445
- alert 144, 172, 223
 - delay 255
 - notification 255
- Application Log 458
- arbitrated loop 13
- archival 127
- array 104, 114, 506
 - configuration 111–112
- array commands 488
- array configuration 111
- ATA translator card 79
- attached storage subsystem
 - DS4000 Storage Manager agent 168
- Auto-Logical Drive Transfer. See ADT
- Automatic Code Synchronization (ACS) 258
- automatic discovery 477
- Automatic Support Data Collection 401
- AVT 141
- AVT-disabled failover 139
- AVT-enabled failover 140

B

- Background Media Scan 532
- base logical drive 240
- basic disc 242
- battery 105, 109, 508, 512
- battery LEDs 83
- blocksize 121
- BOOTP 147, 161
- BOOTP/DHCP 161
- Brocade Host Connectivity Manager (HCM) 358

C

- cable 448
- cache 105, 109, 122, 508, 512

- block size 123, 249, 253
- flushing 123, 249, 252
- mirroring 123, 249, 251
- read-ahead multiplier 123, 249
- write-through 251
- cache block states 38
- cache memory 5, 250, 253
- cache read prefetch 195
- Capture State Information 413
- Challenge Handshake Authentication Protocol (CHAP) 523
- changing cache settings 249
- channel miswire 405
- checklist 129
 - hardware 129
 - software 129
- clearing all error counters 414
- CLI example 493
- clock 171
- clustering support 133
- Collect all Support Data 363
- command line
 - access 144
 - instruction 474
 - parameter 477
 - tool 145, 218
 - utility 152
- command-line interface (CLI) 473–474, 476, 481–482
- configuration wizard 179
- contention 111, 114, 248
- controller
 - ownership 114, 245, 247
- controller commands 486
- controller firmware 8, 144–145, 234, 333–335, 338, 349, 372, 390
- Controller Firmware Upgrade Tool 339
- controller firmware version (CFW) 7.60 15
- controller IP address 148
- controller support module 40
- copy
 - FlashCopy 129
- copy priority 493
- critical events 400
- Cyclic Redundancy Check (CRC) 448

D

- DACstore 100
- data block 250–251
- data cache memory 34
- data collection 440
- data flow 36
- data scrubbing 253
- data striping 104–105, 506, 508
- data verification 238

- DbFix tool 344
- DDK 135
- default group 117
- defragment an array 235
- degraded 238, 407
- degraded array 434
- Device Specific Module 136
- Device Tree View 499
- diagnostics 448
- dirty data 35
- disk
 - mirroring 105, 108, 507, 512
- disk drive 7, 105, 114–115, 143, 233–234, 236, 508
 - even number 234, 236
- disk encryption 125
- Disk Encryption Manager 262
- Disk Management 244
- disk media errors 417
- disk replacement 432
- Disk Security 259, 269
- Disk Security premium feature 259
- diskpart 243
- diskpart utility 243
- distribute 197
- Drive Channel 372
- drive firmware
 - download 357
- drive interfaces 5
- drive loop 114, 248
- drive port 5
- Drive Security 263
- drive type 488, 494
- drive types 111
- Driver Development Kit (DDK) 135
- drive-side cabling 87
- DSM 136
- DS-RSM Model RS2 20, 289
- dynamic cache read prefetch 251
- Dynamic Capacity Expansion (DCE) 233
- Dynamic Random Access Memory (DRAM) 528
- Dynamic System Analysis 458
- Dynamic Volume Expansion (DVE) 242

E

- E-DDM 6
- e-mail 223, 225–226
 - address 479
- e-mail alerts 297
- Emulex 118
- Emulex HBAnyware 359
- enclosure
 - loss protection 112
- enclosure alarm 174
- enclosure commands 488
- encryption 260
- Enhanced Disk Drive Modules (E-DDMs) 73
- Enhanced Remote Mirroring (ERM) 114, 126–127, 129–130
- Enterprise Management 148
- Enterprise Management window 144, 148, 151, 224,

- 371, 475, 477–478, 499
 - Device Table 499
- environmental requirements 129
- error reporting 482
- errpt 471
- ESM 27, 159, 329, 331, 494
- ESM board 27
 - firmware 347
- Ethernet 144, 147, 150, 334
- Ethernet network 144
- Ethernet port 161
- Ethernet priority 521
- event log 248, 327, 400, 445, 448
 - critical entry 248
 - informational entry 248
- Event Log menu 445
- Event Monitor 144, 151, 156, 333
 - service 151, 223
- Event Viewer 458
- EXP5000 5, 99
- EXP5000 Expansion Enclosure 6
- EXP5060 Expansion Enclosure 6
- EXP520 99
- EXP520 Expansion Enclosure 6
- EXP710 26–28
- EXP810 26–28
- EXP810 Expansion Enclosure 6
- expanding logical drives 240
- Expansion Enclosure 28, 347
 - special handling procedures 27

F

- failed drive 110, 116, 187, 513
- failover 8, 248
- failover alert delay 255
- failover driver 452
- FAStT
 - utilities 145
- FAStT configuration, saving 227, 368
- FAStT MSJ 118, 443, 445–446
- FAStT Storage Server 331, 334, 358
- FC 3, 8
- FC cable types 14
- FC drive 5
- FC protocol layers 13
- FC world wide names (WWN) 14
- FDE disk drives 10
- FDE Drive Status 274
- Feature Enable Identifier 364
- feature key 126
- Feature Packs 124
- featureBundle.txt 364
- Fibre Channel (FC) 3, 27–28, 129, 145–146, 192
 - connection 148
 - disk drive 27
 - host interface 486
 - hub 334
 - I/O path 144, 146, 148
 - link 147
 - loop 448

- network 148
- technology 127
- unconfigured capacity 192
- Fibre Channel Hard Disk Drives (FC HDD) 10
- Field Replacement Units 40
- file system 106, 109, 116, 196, 218, 235, 509, 512
- firewall 166, 320
- firmware 145, 174, 330
- firmware file 348, 494
- firmware level 328
- FlashCopy 7, 125, 129, 240, 483, 489–491
 - logical drive 240, 481
- flushing level 123, 249, 252
- Free Capacity
 - Base Select 489
- free capacity 233, 488
 - logical drive 192
- free space 236, 489
 - disk drive 236
 - logical drive 489
- Front LEDs 79
- FRU 40
- Full Disk Encryption (FDE) 123, 125
- Full Disk Encryption (FDE) drives 16, 259
- Full Disk Encryption Hard Disk Drives (FDE HDD) 10

G

- Global Copy 126
- Global Mirror 126
- graphical user interface (GUI) 144
- growth 130

H

- hard disk drive firmware 349
- hardware 23
- hardware initiators 131
- HBA firmware 357
- HBAnyware 442
- HBAs 445, 447
- heterogeneous hosts 120
- host 117, 119
 - host agent 144
 - storage subsystems 146
- host attachment 67
- Host Bus Adapter (HBA) 116–119, 130, 176, 218, 247, 443, 446
 - FC port 117
- host computer 138, 144, 151
 - Event Monitor 151
 - network connection 144
 - storage subsystem 144
- host connection 129
- host group 117, 119, 176, 198, 201, 215, 218, 371, 491
 - single server 119
- host name 146–147, 215, 444–445, 474, 477, 479
- host operating system 118, 201, 240
 - logical drive capacity 240
 - standard logical drive 241
- host port 117, 127, 215, 257, 371, 491

- host type 491
 - world wide name 215
- Host SFP LEDs 80
- host software 153
- host system 116, 144–145, 148, 171, 218, 233, 334
- host topology commands 491
- host type 119
- hot spare 185
- hot spare coverage 189
- hot spare drive 114, 116, 275
- hot_add 152, 218

I

- I/O activity 233, 236, 245, 248
 - full synchronization diverts controller processing resources 245
- I/O path 144, 146, 197, 247–248
- I/O per second (IOPS) 447
- I/O request 120–121, 123, 129, 134, 144, 247
- IBM Remote Support Manager (RSM) 285
- IBM System Storage Disk Systems family 1, 3
- IBM System Storage DS family 2
- IBM System Storage DS Remote Support Manager 20
- IBM System Storage DS3000 series 2
- IBM System Storage DS4000 series 8, 23, 26, 98, 103, 118, 130
 - logical drive 125
- IBM System Storage DS4000 Storage Manager
 - software 143
- IBM System Storage DS4200 storage subsystem
 - Battery LED 82
 - LED 79
- IBM System Storage DS4700 Express storage subsystem 24
- IBM System Storage DS4700 storage subsystem 4
 - chassis 24
 - RAID controller 26
- IBM System Storage DS4800 storage subsystem
 - battery pack 42
 - chassis 42
 - drive-side connections 59
 - front bezel LEDs 46
 - interconnect module 42
 - LED 53
 - RAID controller LED 47
- IBM System Storage DS5000 series 1–3, 6–8, 29
- IBM System Storage DS5020 Controller architecture 68
- IBM System Storage DS5020 storage subsystem 4, 15, 17
- IBM System Storage DS5100 storage subsystem 5
- IBM System Storage DS5300 storage subsystem 5
- IBM System Storage DS8000 series 2
- IBM System Storage EXP5060 18
- Import Array 281, 383
- Import Secure Drive Array 278
- in-band 144
- in-band management 118, 120, 144, 146, 148, 223, 334
- InstallAnywhere 333
- interconnect module 40
- intermixing 4

- Intermixing EXP520 and EXP810 376
- Intermixing EXP810 and EXP5000 376
- Internet Storage Name Service (iSNS) 522
- interoperability matrix 129
- IOPS 3, 109, 120, 129, 513
- IP address 146–148, 159, 161, 225–226, 474, 477, 479–480
 - controller 148
- iSCSI 130, 515–516
- iSCSI host interface 19
- iSCSI technology 516

J

- jumbo frames 519

K

- KB 121, 123
- keys 263

L

- LED indicator lights 79
- link failure (LF) 448–449
- link rate 28
- link speed selector 79
- Linux 119, 333
- lithium-ion battery 435
- load balancing 137
- Load Balancing Policy 134
- load distribution 137
- logical drive 105, 108, 111, 114, 116, 143, 146, 152, 175, 191, 197–198, 201, 215, 218, 222, 229, 235–236, 240–242, 244, 248, 251, 256, 371, 483–484, 486, 488–489, 492–493, 495–496, 498, 507, 512
 - actual reservation 257
 - command returns information 490
 - current owner 248
 - current size 241
 - existing mapping 218
 - free space 235
 - I/O requests 247
 - instant notification 256
 - maximum number 146
 - ownership 255–256
 - parity errors 489
 - physical disk 489
 - preferred controller ownership 248
 - preferred ownership 247
 - primary 114, 245, 248
 - secondary 114, 245, 248
 - segment size 244
 - specific assignments 117
- logical drive-to-LUN mapping 241
- logical unit number (LUN) 116, 198, 201, 446, 485
- logical view 117
- logical volume 114
- loopback 448
- lservice 300
- lspv 471

- lsvg 471
- LUN 114, 116, 119–120, 201
- LUN number 218

M

- Major Event Log (MEL) 292, 327, 400, 410, 484–485
- Manage Host Port Identifiers 215
- management interface 298
- management station 144, 146–147, 151, 223, 334
- managing alerts 322
- mapping 119, 202
 - logical drives to host systems 118
- Mappings View 117, 198, 218
 - Linux system 218
- media error 238
- media scan 196, 228, 253, 368, 371, 403, 490
- Metro Mirror 126
- MIB 227
- midplane 41
- midrange 2
- Migrating Secure Disk Arrays 275
- mirroring 105, 108, 507, 512
- missing paths 426
- modem connectivity 318
- modification operation 245
- modification priority 196, 236, 244, 370, 490
- Monitor process 402
- monitoring 114, 248
- MPIO 8, 128, 135
- mppBusRescan 459
- mppUtil 152, 436
- multipath driver 138
- Multi-Path Proxy (MPP) 137
- multipathing iSCSI 525
- Multiple Path I/O (MPIO) 136
- My Support 175

N

- near-line 127
- Network Management Station (NMS) 226
- network settings 131
- non-redundant operation 236
 - critical time 236
- NTFS 244
- NVSRAM 145, 334, 349, 446
- NVSRAM file 494
- NVSRAM information 372
- NVSRAM version 336, 372

O

- online transaction processing (OLTP) 129
- operating system (OS) 108, 110, 119, 148, 176, 218, 436, 459, 476, 481–482, 512–513
- Optimal state 487–488
- OS version 445
- out-of-band 144, 147
- ownership 114

P

- password 171, 264, 445
- performance 3, 104, 250, 506
- performance impact 235
- performance monitor 121, 248
- persistent reservation 137, 257–258, 426
- physical disk 242, 483–490
 - firmware downloads 487
 - storage subsystem 487
- physical disk drive commands 487
- physical drive 120
- planning 103, 128
- plug and play 137
- point-to-point 13
- power on 160
- power supply 27
 - fan unit LEDs 84
- Predictive Failure Analysis (PFA) 432
- preferred controller 114, 134, 247
- preferred owner 248, 490
- preferred path 256, 334, 419
- premium feature 124, 327, 362, 368
 - enabling 365
 - listing 363
- pre-read redundancy check 238, 405
- primary 245, 248
- primary contact 293
- priority rate 245
- profile 229, 256, 328, 365
- proxy 133
- PuTTY 395

Q

- QLogic 118
- Qlogic SANsurfer 360, 443

R

- RAID 3, 7, 104
 - comparison 109, 513
 - levels 104–106, 109, 114, 176, 198, 236, 370, 490, 505–507, 509, 512–513
 - reliability 110, 513
- RAID 0 104, 506
- RAID 1 105, 507
- RAID 10 511
- RAID 3 105, 508
- RAID 5 509
- RAID 6 510
- RAID controller LEDs 47, 80
- RAID redundancy 415
- RDAC 128, 205, 448
- RDAC driver 152–153, 332, 459
- RDAC installation 465
- RDP 394, 396
- read caching 250
- Read Link Status (RLS) 292, 327
- Read Link Status Diagnostics 410
- read-ahead multiplier 123, 249
- readme file 333

- Recovery Guru 347–348, 363, 399
- Redbooks Web site 536
 - Contact us xiv
- redundancy check 196, 371
 - media scan 196
- redundancy group 238
- redundant 3
- Redundant Disk Array Controller (RDAC) 137
- Remote Access 297
- Remote Assistance 396
- Remote Desktop 394–395
- Remote Mirror 248, 483, 492
 - secondary logical drive 248
- RJ-45 Ethernet connector 65
- RLOGIN 292, 322
- round-robin 137
- rservice 300
- RSM
 - change password 300
 - firewall 320
 - management interface 298
 - remote access security 318
 - security 299
 - user ID 300
- rsm-passwd 300

S

- SAN 3, 128
- SAN Volume Controller (SVC) 103
- SATA 3–4, 11
- Save Configuration 227
- script 370
- Script Editor 228, 369, 371, 473–474, 481, 483, 499, 501–502
 - tools 502
- script engine 501
- script execution 478, 501, 503
 - output view 502
- script file 473, 477, 480, 482, 484, 493, 498, 502
- SCSI Port 128
- SDDPCM 137
- secondary 127, 245
- secondary logical drive 492
- Secure Disk Erase 273
- security 132, 445
- Security Key Identifier 264
- segment 120
- segment size 120–122, 176, 195, 197, 253, 370, 490
- Serial ATA Hard Disk Drives (SATA HDD) 11
- serial number (SN) 486
- serial port 66
- Service Access Timeout 322
- set up alert notifications 397
- show storagesubsystem
 - battery command 480
- Simple Network Management Protocol (SNMP) 151, 156, 223, 226
- single layer cell (SLC) 528
- SM Failover driver 158
- SMagent 144, 153, 333, 475

- SMcli 474, 481
- SMclient 8, 144–145, 147–148, 153, 172, 474, 480
- SMdevices 152, 436, 468
- SMflashcopyassist 152
- SMrepassist 152
- SMruntime 8, 144
- SMTP server 225
- snap 471
- SNMP 223, 292
- SNMP traps 144, 151, 314, 397
- software Initiators 131
- Solid State Drive (SSD) 9, 16, 527
- spare 109, 512
- spare drive 185, 187–189, 244, 501
- SSH connectivity 317
- SSIC 331
- standard logical drive 240, 489
 - storage capacity 241
- step-by-step guide 118
- storage area network (SAN) 138
- storage capacity 27, 98, 233
- storage management
 - software 148, 474
 - station 473
- Storage Manager (SM) 8, 27, 121, 143–144, 146–147, 151–152, 159, 174, 185, 197, 218, 223, 227, 233, 255, 333, 396, 400, 499
 - controller based software 144
 - current level 147
 - host based software 144
 - software 1, 331
 - V9.10 client interface 162
- Storage Manager Agent 8
- Storage Manager Client 8
- Storage Manager Runtime 8
- Storage Manager Utilities 8
- Storage Manager V9.1
 - Runtime 333
- storage partition 116–118, 143, 201, 485, 495
 - logical drives 143
- storage partitioning 114, 116, 118, 125, 201, 210, 248
- Storage Profile 227
- storage subsystem 3, 98, 117, 138, 144–148, 151, 223, 228, 233, 245, 248, 256, 328, 334–335, 348, 363, 365, 368–369, 371, 400, 459, 473–474, 477, 479, 482–483, 486–492, 498–499, 502–503
 - application code 145
 - change history 229
 - command returns configuration information 485
 - commands 483
 - components 7
 - different components 144
 - drive enclosures 348
 - firmware upgrades 474
 - health status 498
 - in-band or out-of-band management 148
 - indicator lights 485
 - individual controllers 146
 - logical drives 114, 490
 - other controller 248

- physical disk 484, 487
- profile 371, 398
- recovery profile 485
- world wide name 475
- storageSubsystem userLabel 482
- StorPort 128
- stripe kill 110, 514
- subsystem contact 293
- Subsystem Management 151
- Subsystem Management window 148
- support
 - My Support 175
- Support Data 227, 402
- Support Notifications subscription service 330
- surface scan 254
- switch 247, 447
- switched bunch of disks (SBOD) 26
- switched fabric 13
- syntax check 477–478, 483
- syntax requirements 481
- system cache 67
- System Event Log 458
- System Storage Interoperation Center (SSIC) 153, 331

T

- tape 130
- throughput 112, 120, 123, 129, 251
- tiered storage 528
- transaction log 105, 508

U

- unconfigured capacity 179, 192, 198, 240–241
 - logical drive 192
- unlock drives 280
- unreadable sector 238, 485
- unrecoverable read error 238
- utilities 145

V

- Virtual Local Area Networks (VLANs) 520
- Virtual Private Network 124, 394
- virtualization 138
- volume 114
- VolumeCopy 125
- VolumeCopy commands 493
- VPN 394

W

- Windows 2000 119, 153, 451
 - IBM System Storage DS4000 series management software components 153
- Windows 2003 242–243
- WMI 136
- world wide name (WWN) 117–118, 479
- write cache mirroring 251
- write caching 250, 532
- write-back 251
- write-through 251

Z
zoning 129–130

Archived

Archived



Redbooks

IBM Midrange System Storage Hardware Guide

(1.0" spine)

0.875" x 1.498"

460 <-> 788 pages



IBM Midrange System Storage Hardware Guide



Redbooks®

DS4000 and DS5000 hardware planning and configuration

Remote Support Manager (RSM) configuration

Features of Storage Manager V10.60

This IBM Redbooks publication consolidates, in one document, detailed descriptions of the hardware configurations and options offered as part of the IBM Midrange System Storage servers, which include the IBM System Storage DS4000 and DS5000 families of products.

This edition covers updates and additional functions available with the IBM System Storage DS Storage Manager Version 10.60 (firmware level 7.60). This book presents the concepts and functions used in planning and managing the storage servers, such as multipathing and path failover. The book offers a step-by-step guide to using the Storage Manager to create arrays, logical drives, and other basic (as well as advanced) management tasks.

This publication also contains practical information about diagnostics and troubleshooting, and includes practical examples of how to use scripts and the command-line interface.

This publication is intended for customers, IBM Business Partners, and IBM technical professionals who want to learn more about the capabilities and advanced functions of the DS4000 series of storage servers with Storage Manager Software V10.60. It also targets those who have a DS4000 and DS5000 storage subsystem and need detailed advice about how to configure it.

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

SG24-7676-01

ISBN 0738434159