

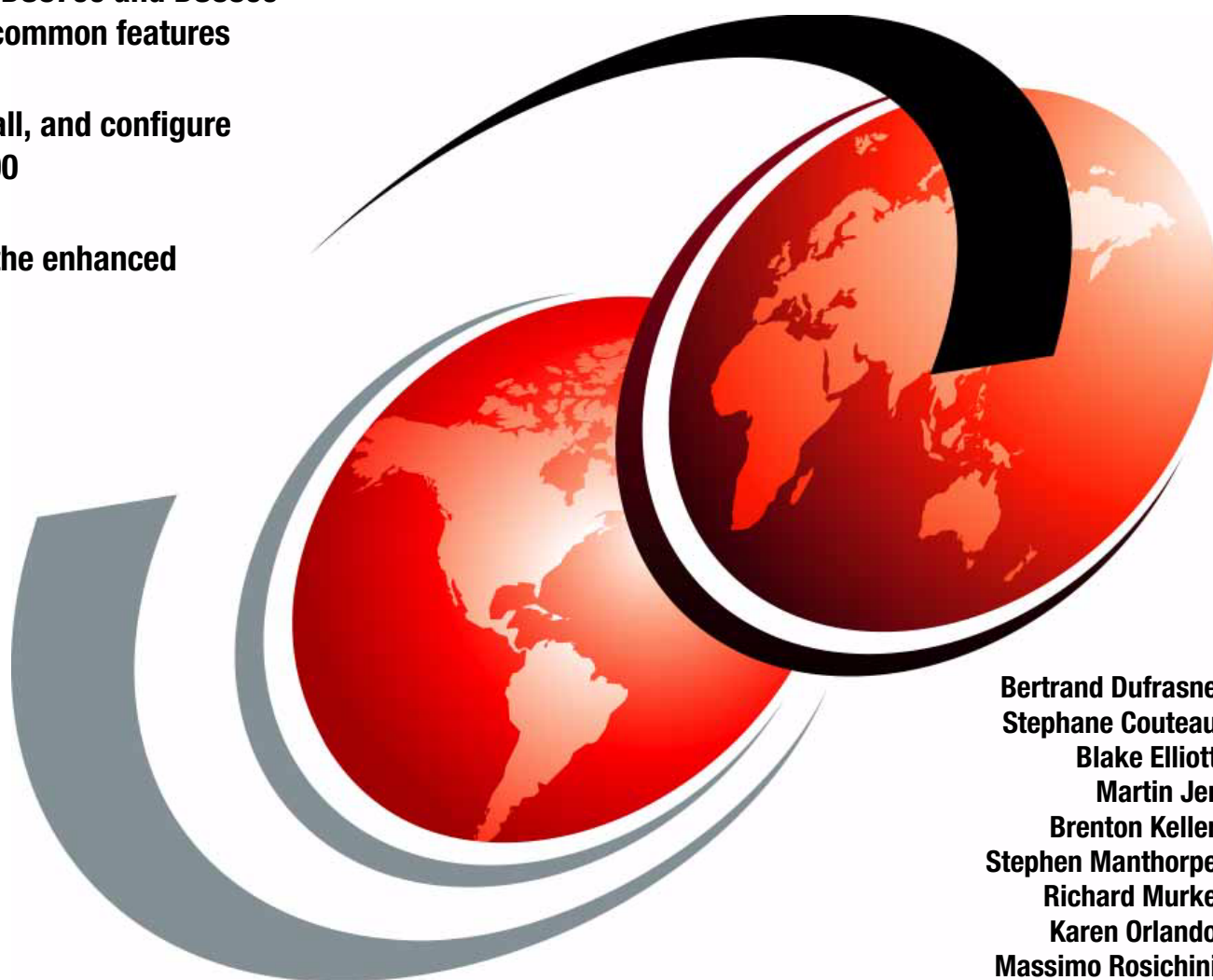
IBM System Storage DS8000

Architecture and Implementation

Learn the DS8700 and DS8800
new and common features

Plan, install, and configure
the DS8000

Discover the enhanced
DS GUI



Bertrand Dufrasne
Stephane Couteau
Blake Elliott
Martin Jer
Brenton Keller
Stephen Manthorpe
Richard Murke
Karen Orlando
Massimo Rosichini

Redbooks



International Technical Support Organization

**IBM System Storage DS8000: Architecture and
Implementation**

August 2011

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

Second Edition (August 2011)

This edition applies to the IBM System Storage DS8700 with DS8000 Licensed Machine Code (LMC) level 6.6.1.xx and the IBM System Storage DS8800 with DS8000 Licensed Machine Code (LMC) level 7.6.1.xx.

© Copyright International Business Machines Corporation 2011. 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	xiii
Trademarks	xiv
Preface	xv
The team who wrote this book	xv
Now you can become a published author, too!	xviii
Comments welcome	xviii
Stay connected to IBM Redbooks	xviii
Part 1. Concepts and architecture	1
Chapter 1. Introduction to the IBM System Storage DS8000 series	3
1.1 Introduction to the DS8700 and DS8800	4
1.1.1 Shared features of the DS8700 and DS8800	5
1.2 The DS8700: A member of the DS family	7
1.3 The DS8800: The premier member of the DS family	8
1.4 DS8000 architecture and functions overview	8
1.4.1 Overall architecture and components	9
1.4.2 Storage capacity	11
1.4.3 Supported environments	12
1.4.4 Easy Tier	12
1.4.5 I/O Priority Manager	13
1.4.6 Configuration flexibility	13
1.4.7 Copy Services functions	14
1.4.8 Resource Groups for copy services scope limiting	17
1.4.9 Service and setup	17
1.4.10 IBM Certified Secure Data Overwrite	17
1.5 Performance features	18
1.5.1 Sophisticated caching algorithms	18
1.5.2 Solid State Disk drives (SSD drives)	19
1.5.3 Multipath Subsystem Device Driver	19
1.5.4 Performance for System z	19
1.5.5 Performance enhancements for IBM Power Systems	20
1.5.6 Performance enhancements for z/OS Global Mirror	20
Chapter 2. IBM System Storage DS8000 models	21
2.1 DS8700 Model overview	22
2.1.1 DS8700 Model 941 overview	23
2.2 DS8800 model overview	28
2.2.1 DS8800 Model 951 overview	29
Chapter 3. Hardware components and architecture	35
3.1 Frames: DS8700	36
3.1.1 Base frame: DS8700	36
3.1.2 Expansion frames	37
3.2 Frames: DS8800	38
3.2.1 Base frame: DS8800	38
3.2.2 Expansion frames	40
3.2.3 Rack operator panel	40

3.3 DS8000 architecture	42
3.3.1 POWER6 and POWER6+ processor	42
3.3.2 Server-based SMP design	43
3.3.3 Peripheral Component Interconnect Express (PCI Express)	43
3.3.4 Storage facility architecture.	45
3.4 Storage facility processor complex (CEC).	47
3.4.1 Processor memory and cache management	49
3.4.2 Flexible service processor and system power control network.	50
3.4.3 RIO-G	50
3.5 I/O enclosures	50
3.5.1 DS8700 I/O enclosures.	51
3.5.2 DS8800 I/O enclosures.	52
3.5.3 Host adapters	52
3.5.4 Device adapters	57
3.6 Disk subsystem	58
3.6.1 Disk enclosures.	58
3.6.2 Disk drives	64
3.7 Power and cooling.	65
3.8 Management console network	67
3.8.1 Ethernet switches	67
3.9 System Storage Productivity Center	68
3.10 Isolated Tivoli Key Lifecycle Manager server	69
Chapter 4. RAS on IBM System Storage DS8000.	71
4.1 Names and terms for the DS8000 storage system	72
4.2 RAS features of DS8000 central electrical complex	74
4.2.1 POWER6 Hypervisor	74
4.2.2 POWER6 processor	74
4.2.3 AIX operating system	77
4.2.4 Central electrical complex dual hard drive rebuild	77
4.2.5 RIO-G interconnect.	78
4.2.6 Environmental monitoring	78
4.2.7 Resource deallocation	79
4.3 Central electrical complex failover and failback	79
4.3.1 Dual operational	80
4.3.2 Failover	81
4.3.3 Failback.	82
4.3.4 NVS and power outages.	83
4.4 Data flow in DS8000	84
4.4.1 I/O enclosures.	84
4.4.2 Host connections	84
4.4.3 Metadata checks.	88
4.5 RAS on the HMC.	88
4.5.1 Hardware	89
4.5.2 Microcode updates	89
4.5.3 Call Home and Remote Support.	89
4.6 RAS on the disk subsystem	89
4.6.1 RAID configurations	89
4.6.2 Disk path redundancy	90
4.6.3 Predictive Failure Analysis	91
4.6.4 Disk scrubbing	91
4.6.5 RAID 5 overview	91
4.6.6 RAID 6 overview	92

4.6.7 RAID 10 overview	93
4.6.8 Spare creation	94
4.7 RAS on the power subsystem	96
4.7.1 Components	96
4.7.2 Line power loss	97
4.7.3 Line power fluctuation	97
4.7.4 Power control	98
4.7.5 Emergency power off	98
4.8 RAS and Full Disk Encryption	99
4.8.1 Deadlock recovery	99
4.8.2 Dual platform TKLM servers	100
4.9 Other features	101
4.9.1 Internal network	101
4.9.2 Remote support	101
4.9.3 Earthquake resistance	101
Chapter 5. Virtualization concepts	103
5.1 Virtualization definition	104
5.2 The abstraction layers for disk virtualization	104
5.2.1 Array sites	105
5.2.2 Arrays	106
5.2.3 Ranks	107
5.2.4 Extent Pools	108
5.2.5 Logical volumes	112
5.2.6 Space Efficient volumes	114
5.2.7 Allocation, deletion, and modification of LUNs/CKD volumes	117
5.2.8 Logical subsystem	123
5.2.9 Volume access	125
5.2.10 Virtualization hierarchy summary	127
5.3 Benefits of virtualization	128
5.4 zDAC - z/OS FICON discovery and Auto-Configuration	129
5.5 EAV V2 - Extended Adresse Volumes (CKD)	131
Chapter 6. IBM System Storage DS8000 Copy Services overview	137
6.1 Copy Services	138
6.2 FlashCopy and FlashCopy SE	139
6.2.1 Basic concepts	139
6.2.2 Benefits and use	141
6.2.3 FlashCopy options	141
6.2.4 FlashCopy SE-specific options	143
6.2.5 Remote Pair FlashCopy	143
6.3 Remote Mirror and Copy	145
6.3.1 Metro Mirror	145
6.3.2 Global Copy	146
6.3.3 Global Mirror	146
6.3.4 Metro/Global Mirror	148
6.3.5 Multiple Global Mirror sessions	149
6.3.6 z/OS Global Mirror	152
6.3.7 z/OS Metro/Global Mirror	153
6.3.8 Summary of Remote Mirror and Copy function characteristics	154
6.4 Resource Groups for copy services	155
Chapter 7. Performance	157
7.1 DS8700 hardware: Performance characteristics	158

7.1.1	Fibre Channel switched disk interconnection at the back end	158
7.1.2	DS8700 Fibre Channel device adapter	161
7.2	DS8800 hardware: Performance characteristics	162
7.2.1	DS8800 Fibre Channel switched interconnection at the back-end	162
7.2.2	Fibre Channel device adapter	163
7.2.3	Eight-port and four-port host adapters	165
7.2.4	IBM System p POWER6: Heart of the DS8000 dual cluster design	166
7.2.5	Vertical growth and scalability	168
7.3	Software performance: Synergy items	168
7.3.1	End-to-end I/O priority: Synergy with AIX and DB2 on System p	169
7.3.2	Cooperative caching: Synergy with AIX and DB2 on System p	169
7.3.3	Long busy wait host tolerance: Synergy with AIX on System p	169
7.3.4	PowerHA Extended distance extensions: Synergy with AIX on System p	169
7.4	Performance considerations for disk drives	170
7.5	DS8000 superior caching algorithms	172
7.5.1	Sequential Adaptive Replacement Cache	173
7.5.2	Adaptive Multi-stream Prefetching	174
7.5.3	Intelligent Write Caching	175
7.6	Performance considerations for logical configuration	177
7.6.1	Workload characteristics	177
7.6.2	Data placement in the DS8000	177
7.6.3	Data placement	178
7.7	Performance and sizing considerations for open systems	183
7.7.1	Determining the number of paths to a LUN	183
7.7.2	Dynamic I/O load-balancing: Subsystem Device Driver	183
7.7.3	Automatic port queues	183
7.7.4	Determining where to attach the host	184
7.7.5	I/O Priority Manager	185
7.8	Performance and sizing considerations for System z	185
7.8.1	Host connections to System z servers	185
7.8.2	Parallel Access Volume	186
7.8.3	z/OS Workload Manager: Dynamic PAV tuning	189
7.8.4	HyperPAV	190
7.8.5	PAV in z/VM environments	192
7.8.6	Multiple Allegiance	193
7.8.7	I/O priority queuing	194
7.8.8	Performance considerations on Extended Distance FICON	195
7.8.9	High Performance FICON for z	197
7.8.10	Extended distance High Performance FICON	198

Part 2. Planning and installation 199

Chapter 8. Physical planning and installation	201
8.1 Considerations prior to installation	202
8.1.1 Who should be involved	203
8.1.2 What information is required	203
8.2 Planning for the physical installation	204
8.2.1 Delivery and staging area	204
8.2.2 Floor type and loading	205
8.2.3 Room space and service clearance	207
8.2.4 Power requirements and operating environment	208
8.2.5 Host interface and cables	210
8.3 Network connectivity planning	212

8.3.1	Hardware Management Console and network access	212
8.3.2	IBM Tivoli Storage Productivity Center	213
8.3.3	System Storage Productivity Center and network access	213
8.3.4	DS command-line interface DSCLI	214
8.3.5	Remote support connection	215
8.3.6	Remote power control	216
8.3.7	Storage Area Network connection	216
8.3.8	Tivoli Key Lifecycle Manager server for encryption	216
8.3.9	Lightweight Directory Access Protocol server for single sign-on	218
8.4	Remote mirror and copy connectivity	219
8.5	Disk capacity considerations	219
8.5.1	Disk sparing	219
8.5.2	Disk capacity	220
8.5.3	DS8000 Solid State Drive (SSD) considerations	222
8.6	Planning for growth	223
Chapter 9	DS8000 HMC planning and setup	225
9.1	Hardware Management Console overview	226
9.1.1	Storage Hardware Management Console hardware	226
9.1.2	Private Ethernet networks	228
9.2	Hardware Management Console software	228
9.2.1	DS Storage Manager GUI	229
9.2.2	Command-line interface	230
9.2.3	DS Open Application Programming Interface	230
9.2.4	Web-based user interface	230
9.3	HMC activities	232
9.3.1	HMC planning tasks	232
9.3.2	Planning for microcode upgrades	233
9.3.3	Time synchronization	233
9.3.4	Monitoring DS8000 with the HMC	234
9.3.5	Call Home and remote support	234
9.4	HMC and IPv6	234
9.5	HMC user management	237
9.5.1	User management using the DS CLI	239
9.5.2	User management using the DS GUI	241
9.6	External HMC	245
9.6.1	External HMC benefits	246
9.6.2	Configuring DS CLI to use a second HMC	246
9.7	Configuration worksheets	247
9.8	Configuration flow	248
Chapter 10	IBM System Storage DS8000 features and license keys	251
10.1	IBM System Storage DS8000 licensed functions	252
10.2	Activation of licensed functions	254
10.2.1	Obtaining DS8000 machine information	255
10.2.2	Obtaining activation codes	257
10.2.3	Applying activation codes using the GUI	262
10.2.4	Applying activation codes using the DS CLI	265
10.3	Licensed scope considerations	266
10.3.1	Why you get a choice	267
10.3.2	Using a feature for which you are not licensed	267
10.3.3	Changing the scope to All	268
10.3.4	Changing the scope from All to FB	269

10.3.5 Applying an insufficient license feature key	270
10.3.6 Calculating how much capacity is used for CKD or FB.	270
Part 3. Storage configuration.	273
Chapter 11. Configuration flow	275
11.1 Configuration worksheets	276
11.2 Configuration flow	276
Chapter 12. System Storage Productivity Center	279
12.1 System Storage Productivity Center overview	280
12.1.1 SSPC components	280
12.1.2 SSPC capabilities	282
12.1.3 SSPC upgrade options	282
12.2 SSPC setup and configuration	283
12.2.1 DS8000 Storage Management	284
12.2.2 Configuring SSPC for DS8000 remote GUI access	284
12.2.3 Preparing your browser.	284
12.2.4 Accessing the TPC on SSPC	285
12.2.5 Manage embedded CIMOM on DS8000.	288
12.3 Configuring Tivoli Productivity Center For DS8000.	292
12.3.1 Before you start.	292
12.3.2 Adding a DS8000 server by using the Configure Devices wizard.	292
12.3.3 Running a probe job to collect DS8000 data	299
12.3.4 Configuring the TPC Element Manager to access the DS8000 GUI.	301
12.4 Working with a DS8000 system in TPC-BE	304
12.4.1 Display disks and volumes of DS8000 Extent Pools.	304
12.4.2 Display the physical paths between systems	306
12.4.3 Storage health management.	308
12.4.4 Display host volumes through SVC to the assigned DS8000 volume.	308
Chapter 13. Configuration using the DS Storage Manager GUI	309
13.1 DS Storage Manager GUI overview	310
13.1.1 Accessing the DS GUI	310
13.1.2 DS GUI Overview window.	314
13.2 Logical configuration process	317
13.3 Examples of configuring DS8000 storage.	317
13.3.1 Define storage complex	318
13.3.2 Create arrays	321
13.3.3 Create ranks	326
13.3.4 Create Extent Pools	329
13.3.5 Configure I/O ports	334
13.3.6 Configure logical host systems	335
13.3.7 Create fixed block volumes.	340
13.3.8 Create volume groups.	344
13.3.9 Create LCUs and CKD volumes	348
13.3.10 Additional actions on LCUs and CKD volumes.	355
13.4 Other DS GUI functions.	356
13.4.1 Check the status of the DS8000.	356
13.4.2 Explore the DS8000 hardware	358
Chapter 14. Configuration with the DS Command-Line Interface	363
14.1 DS Command-Line Interface overview	364
14.1.1 Supported operating systems for the DS CLI	364

14.1.2	User accounts	364
14.1.3	DS CLI profile	365
14.1.4	Command structure	367
14.1.5	Using the DS CLI application	367
14.1.6	Return codes	370
14.1.7	User assistance	370
14.2	Configuring the I/O ports	371
14.3	Monitoring the I/O ports	372
14.4	Configuring the DS8000 storage for FB volumes	373
14.4.1	Create arrays	373
14.4.2	Create ranks	374
14.4.3	Create Extent Pools	375
14.4.4	Creating FB volumes	377
14.4.5	Creating volume groups	382
14.4.6	Creating host connections	384
14.4.7	Mapping open systems host disks to storage unit volumes	386
14.5	Configuring DS8000 Storage for Count Key Data Volumes	388
14.5.1	Create arrays	388
14.5.2	Ranks and Extent Pool creation	388
14.5.3	Logical control unit creation	390
14.5.4	Create CKD volumes	391
14.5.5	Resource Groups	397
14.5.6	Performance I/O Priority Manager	397
Part 4.	Maintenance and upgrades	399
Chapter 15.	Licensed machine code	401
15.1	How new microcode is released	402
15.2	Bundle installation	403
15.3	Concurrent and non-concurrent updates	405
15.4	Code updates	405
15.5	Host adapter firmware updates	405
15.6	Loading the code bundle	406
15.7	Post-installation activities	406
15.8	Summary	406
Chapter 16.	Monitoring with Simple Network Management Protocol	407
16.1	Simple Network Management Protocol overview	408
16.1.1	SNMP agent	408
16.1.2	SNMP manager	409
16.1.3	SNMP trap	409
16.1.4	SNMP communication	409
16.1.5	SNMP Requirements	410
16.1.6	Generic SNMP security	410
16.1.7	Message Information Base	411
16.1.8	SNMP trap request	411
16.1.9	DS8000 SNMP configuration	411
16.2	SNMP notifications	412
16.2.1	Serviceable event using specific trap 3	412
16.2.2	Copy Services event traps	412
16.2.3	I/O Priority Manager SNMP	419
16.2.4	Thin Provisioning SNMP	420
16.3	SNMP configuration	420
16.3.1	SNMP preparation	420

16.3.2	SNMP configuration with the HMC using WUI Menu	421
16.3.3	SNMP configuration with the DS CLI	423
Chapter 17.	Remote support	425
17.1	Introduction to remote support	426
17.1.1	Suggested reading	426
17.1.2	Organization of this chapter	426
17.1.3	Terminology and definitions	427
17.2	IBM policies for remote support	428
17.3	VPN rationale and advantages	428
17.4	Remote connection types	429
17.4.1	Asynchronous modem	430
17.4.2	IP network	431
17.4.3	IP network with traditional VPN	432
17.4.4	IP network with Business-to-Business VPN	432
17.5	DS8000 support tasks	433
17.5.1	Call Home and heartbeat (outbound)	433
17.5.2	Data offload (outbound)	434
17.5.3	Code download (inbound)	436
17.5.4	Remote support (inbound and two-way)	437
17.6	Remote connection scenarios	437
17.6.1	No connections	437
17.6.2	Modem only	438
17.6.3	VPN only	439
17.6.4	Modem and network with no VPN	439
17.6.5	Modem and traditional VPN	440
17.6.6	Modem and Business-to-Business VPN	441
17.7	Audit logging	442
Chapter 18.	Capacity upgrades and CoD	445
18.1	Installing capacity upgrades	446
18.1.1	Installation order of upgrades	447
18.1.2	Checking how much total capacity is installed	448
18.2	Using Capacity on Demand (CoD)	449
18.2.1	What is Capacity on Demand	449
18.2.2	Determining if a DS8800 has CoD	450
18.2.3	Using the CoD storage	453
Appendix A.	Tools and service offerings	455
	Capacity Magic	456
	Disk Magic	457
	HyperPAV analysis	458
	FLASHDA	458
	IBM i SSD Analyzer Tool	460
	IBM Tivoli Storage Productivity Center	460
	IBM Certified Secure Data Overwrite	461
	IBM Global Technology Services: service offerings	463
	IBM STG Lab Services: Service offerings	463
Abbreviations and acronyms		465
Related publications		469
	IBM Redbooks publications	469
	Other publications	469

Online resources	470
How to get IBM Redbooks publications	470
Help from IBM	470
Index	471

Notices

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

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

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

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

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

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

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

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

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

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

COPYRIGHT LICENSE:

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

Trademarks

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

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

AIX 5L™	Power Architecture®	System p®
AIX®	Power Systems™	System Storage DS®
DB2®	POWER5™	System Storage®
DS4000®	POWER5+™	System x®
DS6000™	POWER6+™	System z10®
DS8000®	POWER6®	System z®
Enterprise Storage Server®	PowerHA™	TDMF®
ESCON®	PowerPC®	Tivoli®
FICON®	POWER®	TotalStorage®
FlashCopy®	pSeries®	WebSphere®
GDPS®	Redbooks®	XIV®
HACMP™	Redpapers™	z/OS®
HyperSwap®	Redbooks (logo)  ®	z/VM®
i5/OS®	RMF™	z10™
IBM®	S/390®	z9®
NetView®	System i®	

The following terms are trademarks of other companies:

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

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

Java, and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

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

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

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

Preface

This IBM® Redbooks® publication describes the concepts, architecture, and implementation of the IBM System Storage® DS8700 and DS8800 storage systems. The book provides reference information to assist readers who need to plan for, install, and configure the DS8700 and DS8800.

The DS8700 includes IBM POWER6®-based controllers. The IBM System Storage DS8800 is the most advanced model in the IBM DS8000® lineup and is equipped with IBM POWER6+™ based controllers. Both systems feature a dual 2-way or dual 4-way processor complex implementation. They also feature enhanced 8 Gpbs device adapters and host adapters. Their extended connectivity, with up to 128 Fibre Channel/FICON® ports for host connections, makes them suitable for multiple server environments in both open systems and IBM System z® environments. Both systems now support thin provisioning. They also support the Full Disk Encryption (FDE) feature. If desired, they can be integrated in an LDAP infrastructure.

The DS8800 is equipped with high-density storage enclosures populated with 24 small-form-factor SAS-2 drives. The DS8700 and DS8800 storage subsystems can be equipped with Solid-State Drives (SSDs).

The DS8700 and DS8800 can automatically optimize the use of SSD drives through the Easy Tier feature, which is available for no extra fee. For details about Easy Tier, refer to *IBM System Storage DS8000: Easy Tier Concepts and Usage*, REDP-4667.

Host attachment and interoperability topics for the DS8000 series including the DS8800 are now covered in *IBM System Storage DS8000: Host Attachment and Interoperability*, SG24-8887.

For information related to specific features, refer to *IBM System Storage DS8000: Easy Tier Concepts and Usage*, REDP-4667, *IBM System Storage DS8000: Priority Manager*, REDP-4760, *IBM System Storage DS8000: Copy Services Resource Groups*, REDP-4758, *IBM System Storage DS8700 Disk Encryption Implementation and Usage Guidelines*, REDP-4500, and *IBM System Storage DS8000: LDAP Authentication*, REDP-4505.

To read about DS8000 Copy Services functions, refer to *IBM System Storage DS8000: Copy Services for Open Environments*, SG24-6788, and *IBM System Storage DS8000: Copy Services for IBM System z*, SG24-6787.

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.

Bertrand Dufrasne is an IBM Certified Consulting IT Specialist and Project Leader for IBM System Storage disk products at the International Technical Support Organization, San Jose Center. He has worked at IBM in various IT areas. Bertrand has written many IBM Redbooks publications, and has also developed and taught technical workshops. Before joining the ITSO, he worked for IBM Global Services as an Application Architect in the retail, banking, telecommunication, and health care industries. He holds a Masters degree in Electrical Engineering.

Stephane Couteau is an IT Specialist and has been with IBM France for 12 years. His areas of expertise include System p®, AIX®, Linux, DS8000, SVC, XiV, DS5000, and N series. He has been involved in various projects that typically include server and storage solutions implementation and migration, and more recently storage benchmarking. Stephane holds a degree in Mathematics from University of Metz, and has more than 20 years of experience in computer industry.

Blake Elliott is a Field Technical Sales Specialist based in Houston, Texas. He is a graduate of the IBM Summit Program. He uses his knowledge of Enterprise Disk, SAN Volume Controller, and Scale Out Network Attached Storage systems in the health care, and oil and gas industries. He started his IBM career in 2010. He holds a Bachelor of Science in Economics from Berry College and a Masters in Business Administration with a focus on International Business from Liberty University.

Martin Jer is a Certified IT Storage Specialist in New York, working as a pre-sale advisor for enterprise storage topics. He has more than 30 years of experience with IBM large systems, storage systems, tape systems, and virtualization systems. His areas of expertise include IBM high-end disk and tape storage subsystems and disaster recovery solutions using the capabilities and features of IBM storage products. He has contributed to storage-related IBM Redbooks publications in the past.

Brenton Keller is Storage Field Technical Sales Specialist based in Chicago, Illinois. He is a graduate of the IBM Summit Program and holds a degree in electrical and computer engineering from Washington University in St. Louis.

Stephen Manthorpe worked for IBM Australia for 20 years as Large Systems CE working across most platforms. He then assumed the positions of country support, High End Disk for Australia and New Zealand and regional Support for SE Asia. He moved to IBM USA in 2007 for STG/DS8000 Development as Functional Verification Test Team Lead. His focus is on RAS Functional Test, and Serviceability, New Function, and New HW.

Richard Murke is a Senior IT Specialist in Germany. He has 28 years of experience with S/390® Hardware and Software on S/390 and z/OS®, including 11 years in high end disk storage. He works as a Field Technical Sales Support Specialist for high end storage and SAN environments. His area of expertise includes performance analysis and implementations for DS8000 within FICON infrastructures, IBM disk storage servers, and SAN directors from multiple vendors.

Karen Orlando is a Project Leader at the International Technical Support Organization, Tucson Arizona Center. Karen has over 25 years in the IT industry with extensive experience in open systems, product test, and development of IBM hardware and software storage. She holds a degree in Business Information Systems from the University of Phoenix and has been Project Management Professional (PMP) certified since 2005.

Massimo Rosichini is an IBM Certified in Supporting IT Solutions and Country Specialist in the GTS System Storage Technical Support Group in Italy. He has extensive experience in support and delivery of Professional Services System Storage solutions in the IMT Italy. He is an DS8000 Top Gun Specialist, and has co-authored several IBM Redbooks publications related to Storage High End Copy Services for Open System and SAN FS Environment.



Figure 1 The team: Stephane, Massimo, Martin, Elliot, Richard, Karen, Stephen, Brenton, Bertrand

Many thanks to the following people who helped with equipment provisioning and preparation:

Douglas Acuff
 Robert Brown
 Joshua Crawford
 Michael Lopez
 Rachel Mikolajewski
 Addie Richards
 Teena Worley
 IBM Storage Test Lab, Tucson, Arizona

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

Dale Anderson
 Stephen Blinick
 John Bynum
 Larry Chiu
 James Davison
 John Elliott
 Craig Gordon
 Peter Kimmel
 Lee LaFrese
 Michael Lopez
 Denise Luzar
 Rosemary Mccutchen
 Paul Muench
 Brian Rinaldi
 Rick Ripberger
 Kavilah Shah
 Brian Sherman
 Jeff Steffan

Sonny Williams
Jens Wissenbach
Allen Wright
Yan Xu
IBM

Now you can become a published author, too!

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

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

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

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

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

ibm.com/redbooks

- Send your comments in an email to:

redbooks@us.ibm.com

- Mail your comments to:

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

Stay connected to IBM Redbooks

- Find us on Facebook:

<http://www.facebook.com/IBMRedbooks>

- Follow us on Twitter:

<http://twitter.com/ibmredbooks>

- Look for us on LinkedIn:

<http://www.linkedin.com/groups?home=&gid=2130806>

- ▶ Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks weekly newsletter:
<https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm>
- ▶ Stay current on recent Redbooks publications with RSS Feeds:
<http://www.redbooks.ibm.com/rss.html>

Concepts and architecture

This part gives an overview of the IBM System Storage DS8000 concepts and architecture. The topics covered include:

- ▶ Introduction to the IBM System Storage DS8000 series
- ▶ IBM System Storage DS8000 models
- ▶ Hardware components and architecture
- ▶ RAS on IBM System Storage DS8000
- ▶ Virtualization concepts
- ▶ IBM System Storage DS8000 Copy Services overview



Introduction to the IBM System Storage DS8000 series

This chapter introduces the features, functions, and benefits of the IBM System Storage DS8000 series. The functions and features covered in this chapter apply to the DS8700 and DS8800 models.

More detailed information about functions and features is provided in subsequent chapters. The topics presented here are:

- ▶ Shared functions of the DS8700 and the DS8800
- ▶ DS8700 only functions
- ▶ DS8800 only functions
- ▶ DS8000 architecture and functions overview
- ▶ Performance features

1.1 Introduction to the DS8700 and DS8800

IBM has a wide range of product offerings that are based on open standards and share a common set of tools, interfaces, and innovative features. The System Storage DS8000 family is designed as a high performance, high capacity, and resilient series of disk storage systems. It offers high availability, multiplatform support, and simplified management tools to help provide a cost-effective path to an on-demand world.

The DS8700 and DS8800 (Figure 1-1) are IBM third- and fourth-generation high-end disk systems in the DS8000 series. Both are designed to support the most demanding business applications with their exceptional all-around performance and data throughput. Combined with the world-class business resiliency and encryption features, both machines, provide a unique combination of high availability, performance, and security. Both are tremendously scalable, have broad server support, and virtualization capabilities. These features can help simplify the storage environment by consolidating multiple storage systems onto a single machine. High density storage enclosures offer a considerable reduction in footprint and energy consumption, making them the most space and energy-efficient models in the DS8000 series.

Compared with their predecessors, the IBM System Storage DS8100 and IBM System Storage DS8300, the DS8700 and DS8800 introduce new functional capabilities, allowing you to choose the combination that is right for your application needs.



Figure 1-1 DS8700 (left) and DS8800 (right)

1.1.1 Shared features of the DS8700 and DS8800

The DS8700 and DS8800 with Release 6.1 of the firmware (that is, Licensed Machine Code level 6.6.1.xx for the DS8700 Licensed Machine Code (LMC) level 7.6.1.xx for the DS8800), offer the following shared features:

- ▶ Storage virtualization offered by the DS8000 series allows organizations to allocate system resources more effectively and better control application quality of service. The DS8000 series improves the cost structure of operations and lowers energy consumption through a tiered storage environment.
- ▶ Storage Pool Striping (*rotate extents*) is now the default when creating new volumes and not explicitly specifying an extent allocation method (EAM). Storage Pool Striping helps maximize performance without special tuning and greatly reduces “hot spots” in ranks.
- ▶ Easy Tier: This feature enables automatic dynamic data relocation capabilities. Configuration flexibility and overall storage cost-performance can greatly benefit from the exploitation of this feature. In Release 6.1, Easy Tier can now support any two tiers of storage (Nearline and Enterprise, Nearline and Solid State Disk, or Enterprise and Solid State Disk). Easy Tier also allows several manual data relocation capabilities (extent pools merge, rank depopulation, volume migration). Refer to *IBM System Storage DS8000 Easy Tier*, REDP-4667 for more information.
- ▶ Storage Tier Advisor Tool: This tool is used in conjunction with the Easy Tier facility to help clients understand their current disk system workloads and provide guidance on how much of their existing data would be better suited for the various drive types (spinning or solid state).
- ▶ Resource Groups: This new feature is a policy based resource scope limiting function that enables the secure use of Copy Services functions by multiple users on a DS8000 series storage subsystem. Resource Groups are used to define an aggregation of resources and policies for configuration and management of those resources. The scope of the aggregated resources can be tailored to meet each hosted customers’ Copy Services requirements for any given operating system platform supported by the DS8000 series. Refer to *IBM System Storage DS8000 Resource Groups*, REDP-4758 for more information.
- ▶ I/O Priority Manager: A new feature that provides application level Quality of Service (QoS). It provides a way to manage quality of service for I/O operations associated with critical workloads and give them priority over other I/O operations associated with non-critical workloads. Refer to *IBM System Storage DS8000 I/O Priority Manager*, REDP-4760 for more information.
- ▶ Large FB LUNs: With release 6.1, supported LUNs have increased from 2 TB up to 16 TB. This helps to alleviate address constraints to support large storage capacity needs.
- ▶ Active Volume Protection: This is a feature that prevents the deletion of volumes still in use.
- ▶ The Dynamic Volume Expansion simplifies management by enabling easier, online volume expansion to support application data growth, and to support data center migration and consolidation to larger volumes to ease addressing constraints.
- ▶ Thin Provisioning: This feature allows the creation of over-provisioned devices for more efficient usage of the storage capacity.
- ▶ Quick Initialization for open system (FB) volumes: This feature provides volume initialization that is up to 2.6 times faster, and therefore allows the creation of devices and making them available as soon as the command completes.
- ▶ Peripheral Component Interconnect Express (PCI Express Generation 2) IO enclosures: To improve I/O Operations Per Second (IOPS) and sequential read/write throughput, the

new I/O enclosures are directly connected to the internal servers with point-to-point PCI Express cables. I/O enclosures no longer share common loops. They connect directly to each internal server through separate cables and link cards.

- ▶ 8 Gbps host adapters (HAs): The DS8000 model offers enhanced connectivity with 4-port Fibre Channel/FICON host adapters located in the I/O enclosures that are directly connected to the internal processor complexes. The DS8000s 8 Gbps Fibre Channel/FICON host adapter supports FICON attachment to FICON Express8 on zEnterprise 196 (z196) and System z10® (and later). The DS8000 8 Gbps Fibre Channel/FICON host adapter also provides support for FICON Express2-attached and FICON Express4-attached systems.
- ▶ Optional Solid-State Drives (SSDs) provide extremely fast access to data, energy efficiency, and higher system availability.
- ▶ Processor memory offerings: The DS8000, which has a 2-way configuration, offers up to 128 GB processor memory. With a 4-way configuration the DS8000 offers up to 384 GB of processor memory. Non-volatile Storage (NVS) scales with the processor memory size on a 1/32 scale (minimum of 1 GB).
- ▶ Adaptive Multi-stream Prefetching (AMP) caching algorithm can dramatically improve sequential performance, thereby reducing times for backup, processing for business intelligence, and streaming media. Intelligent Write Caching (IWC) improves the Cache Algorithm for random writes.
- ▶ The Full Disk Encryption (FDE) feature can protect business-sensitive data by providing disk-based hardware encryption combined with a sophisticated key management software (Tivoli® Key Lifecycle Manager). The Full Disk Encryption support feature is available only as a plant order. Refer to *IBM System Storage DS8700 Disk Encryption Implementation and Usage Guidelines*, REDP-4500, for more information about this topic.
- ▶ Disk encryption key management that helps address Payment Card Industry Data Security Standard (PCI-DSS) requirements:
 - Encryption deadlock recovery key option: When enabled, this option allows the user to restore access to a DS8000 when the encryption key for the storage is unavailable due to an encryption deadlock scenario.
 - Dual platform key server support: DS8000 requires an isolated key server in encryption configurations. The isolated key server currently defined is an IBM System x® server. Dual platform key server support allows two server platforms to host the key manager with either platform operating in either *clear key* or *secure key* mode.
 - Recovery key Enabling/Disabling and Rekey data key option for the Full Disk Encryption (FDE) feature: Both of these enhancements can help clients satisfy Payment Card Industry (PCI) security standards
- ▶ High Performance FICON for System z (zHPF) Extended Distance capability: This feature enhances zHPF write performance by supporting the zHPF “Disable Transfer Ready” protocol.
- ▶ The FlashCopy® SE capability enables more space efficient utilization of capacity for copies, enabling improved cost effectiveness.
- ▶ Remote Pair FlashCopy: This allows you to establish a FlashCopy relationship where the target is a remote mirror Metro Mirror primary volume keeping the pair in the *full duplex* state.
- ▶ System Storage Productivity Center (SSPC) single pane control and management integrates the power of the Tivoli Storage Productivity Center (TPC) and the DS Storage Manager user interfaces into a single view.

- ▶ An improved DS GUI management interface with views that show the mappings of elements of the logical configuration to physical hardware components.
- ▶ LDAP authentication support, which allows single sign-on functionality, can simplify user management by allowing both the DS8700 and DS8800 to rely on a centralized LDAP directory rather than a local user repository. Refer to *IBM System Storage DS8000: LDAP Authentication*, REDP-4505 for more information
- ▶ The DS8000 series has been certified as meeting the requirements of the IPv6 Read Logo program, indicating its implementation of IPv6 mandatory core protocols and the ability to interoperate with other IPv6 implementations. The IBM DS8000 can be configured in native IPv6 environments. The logo program provides conformance and interoperability test specifications based on open standards to support IPv6 deployment globally.
- ▶ Value based pricing/licensing: The Operating Environment License is now priced based on the performance, capacity, speed, and other characteristics that provide value in customer environments.
- ▶ Data Protection: The DS8000 series is designed for the most demanding, mission-critical environments requiring extremely high availability. It is designed to avoid single points of failure. With the advanced Copy Services functions the DS8000 series integrates, data availability can be enhanced even further. FlashCopy and FlashCopy SE allow production workloads to continue execution concurrently with data backups.

Metro Mirror, Global Copy, Global Mirror, Metro/Global Mirror, z/OS Global Mirror, and z/OS Metro/Global Mirror business continuity solutions are designed to provide the advanced functionality and flexibility needed to tailor a business continuity environment for almost any recovery point or recovery time objective.

The DS8000 also offers three-site solutions with Metro/Global Mirror and z/OS Metro/Global Mirror for additional high availability and disaster protection. Another important feature for z/OS Global Mirror (2-site) and z/OS Metro/Global Mirror (3-site) is Extended Distance FICON, which can help reduce the need for channel extenders configurations by increasing the number of read commands in flight.

The Copy Services can be managed and automated with IBM Tivoli Storage Productivity Center for Replication (TPC-R).

1.2 The DS8700: A member of the DS family

The IBM System Storage DS8700 adds Models 941 (base frame) and 94E (expansion unit) to the 242x machine type family. Compared with its predecessors, IBM System Storage DS8100 and DS8300, the DS8700 is designed to provide capabilities for the combination of price and efficiency.

Functions include:

- ▶ IBM POWER6 processor technology: The DS8700 features the IBM POWER6 server technology to help support high performance. Compared to the POWER5+™ processor in previous models, the POWER6 processor can deliver more than a 50% performance improvement in I/O operations per second (IOPS) in transaction processing workload environments. Additionally, sequential workloads can receive as much as 150% bandwidth improvement. The DS8700 offers either a dual 2-way processor complex or a dual 4-way processor complex.
- ▶ A nondisruptive upgrade path for the DS8700 Model 941 and additional Model 94E expansion frames allows processor, cache, and storage enhancement to be performed concurrently without disrupting applications.

1.3 The DS8800: The premier member of the DS family

The IBM System Storage DS8800 adds Models 951 (base frame) and 95E (expansion unit) to the 242x machine type family, delivering cutting edge technology, improved space, improved energy efficiency, and increased performance.

Functions include:

- ▶ IBM POWER6+ processor technology

The DS8800 features the IBM POWER6+ server technology to help support high performance. Compared to the performance of the DS8700 (POWER6), the processor aids the DS8800 in achieving sequential read throughput performance improvement up to 20% and sequential write throughput performance improvement up to 40%. The DS8800 offers either a dual 2-way processor complex or a dual 4-way processor complex.

- ▶ Improved configuration options

The DS8800 standard cabling is optimized for performance and highly scalable configurations with capacity for large long-term growth. The DS8800 with standard cabling allows for up to three frames and up to sixteen 8-port host adapters, or up to sixteen 4-port host adapters, providing a high performance and scalable storage environment.

The DS8800 also provides a business class configuration option. The business class option allows a system to be configured with more drives per device adapter, thereby helping to reduce configuration cost and increasing adapter utilization.

- ▶ A nondisruptive upgrade path for DS8800 Model 951 (standard cabling), and additional Model 95E expansion frames allows processor, cache, and storage enhancements to be performed concurrently without disrupting applications.

The DS8800 provides a nondisruptive upgrade path for the DS8800 Model 951 (standard cabling), and additional Model 95E expansion frames allowing processor, cache, and storage enhancements to be performed concurrently without disrupting applications.

- ▶ High density storage enclosures

The DS8800 provides storage enclosure support for 24 small form factor (SFF, 2.5-inch) drives in 2U of rack space. This option helps improve the storage density for disk drive modules (DDMs) as compared to previous enclosures.

- ▶ Improved high density frame design

The DS8800 can support a total of 1056 drives in a smaller footprint (three frames) than previous generations, supporting higher density and helping to preserve valuable raised floor space in data center environments. DS8800 is designed for hot and cold aisle data center design, drawing air for cooling from the front of the system and exhausting hot air at the rear.

Coupled with this improved cooling implementation, the reduced system footprint, and small form factor Enterprise-2 drives, a fully configured DS8800 consumes up to 40% less power than previous generations of DS8000. The DS8800 base frame supports up to 240 drives, with the first expansion frame supporting up to 336 drives and second expansion frame supporting up to 480 drives.

1.4 DS8000 architecture and functions overview

The IBM System Storage DS8000s highlights include:

- ▶ Robust, flexible, Enterprise class, and cost-effective disk storage
- ▶ Exceptionally high system availability for continuous operations

- ▶ IBM POWER6 family
- ▶ Capacities from 600 MB to 2048 TB on the DS8700
- ▶ Capacities from 2.3 TB to 633 TB on the DS8800
- ▶ Point-in-time copy function with FlashCopy and FlashCopy SE
- ▶ Remote Mirror and Copy functions with Metro Mirror, Global Copy, Global Mirror, Metro/Global Mirror, z/OS Global Mirror, and z/OS Metro/Global Mirror with Incremental Resync capability
- ▶ Support for a wide variety and intermix of operating systems, including IBM i and System z
- ▶ Designed to increase storage efficiency and utilization, ideal for green data centers

1.4.1 Overall architecture and components

From an architectural point of view, the DS8700 and DS8800 offer continuity with respect to the fundamental architecture of their predecessors the DS8100 and DS8300 models. This ensures that both the DS8700 and the DS8800 can use a stable and well-proven operating environment, offering optimal availability. The hardware is optimized to provide higher performance, connectivity, and reliability.

The DS8700 and the DS8800 are available with separate configurations, which are discussed in detail in Chapter 2, “IBM System Storage DS8000 models” on page 21.

IBM POWER 6 processor technology

Both the DS8700 and DS8800 use IBM POWER® 6 processor technology. The Symmetric Multiprocessing (SMP) system features 2-way or 4-way, copper-based, silicon-on-insulator (SOI)-based POWER technology. The DS8700's P6 processor runs at 4.7 GHz, whereas the DS8800's P6+ processor runs at 5.0 GHz.

Both the DS8700 and the DS8800 offer either a dual 2-way processor complex or a dual 4-way processor complex. A processor complex is also referred to as a storage server or central electronics complex. For more information, see Chapter 4, “RAS on IBM System Storage DS8000” on page 71.

Internal PCIe-based fabric

The DS8700 and DS8800 use direct point-to-point high speed PCI Express (PCIe) connections to the I/O enclosures to communicate with the device adaptors and host adapters. Each single PCIe connection operates at a speed of 2 GBps in each direction. There are up to 16 PCIe connections from the processor complexes to the I/O enclosures. For more information, see Chapter 3, “Hardware components and architecture” on page 35 or at: <http://www.redbooks.ibm.com/Redbooks.nsf/RedbookAbstracts/tips0456.html?open>

Device adapters

The DS8700 and DS8800 offer 4-port Fibre Channel Arbitrated Loop (FC-AL) Device Adapters (DA). All adapters provide improved IOPS, throughput, and scalability over previous DS8000s. They are optimized for SSD technology and architected for long-term support for scalability growth. These capabilities complement the POWER server family to provide significant performance enhancements allowing up to a 400% improvement in performance over previous generations. For more information, see Chapter 3, “Hardware components and architecture” on page 35.

Switched Fibre Channel arbitrated loop

The DS8000 uses a switched Fibre Channel arbitrated loop (FC-AL) architecture as the backend for its disk interconnection. The DAs connect to the controller cards in the storage enclosures using FC-AL with optical short wave multi-mode interconnection. The Fibre Channel Interface Controller cards (FCIC) offer a point-to-point connection to each drive and device adapter, so that there are four paths available from the DS8000 processor complexes to each disk drive. For more information, see Chapter 3, “Hardware components and architecture” on page 35.

Disk drives

Both the DS8700 and DS8800 offer a variety of disk drives to meet the requirements of various workload and security configurations. For more information, see Chapter 8, “Physical planning and installation” on page 201.

► DS8700

300, 450, and 600 GB (15K RPM) Enterprise disk drives can currently be installed in the DS8700. The DS8700 also supports 300, 450, and 600 GB (15K RPM) Full Disk Encryption (FDE) disk drives.

146 GB (15K RPM) is also supported but has been withdrawn from marketing as of the Release 6.1 microcode.

Additionally, 2 TB (7200K RPM) Nearline (SATA) disk drives are supported.

The DS8700 also supports 600 GB Solid State Disk drives (SSDs). 73 GB and 146 GB SSDs are also supported, but have been withdrawn from marketing as of this Microcode release.

► DS8800

The DS8800 supports 146 GB (15 K RPM), 450 GB (10K RPM), 600 GB (10K RPM), Enterprise (SAS) disk drives. 450 GB (10K RPM) and 600 GB (10K RPM) Full Disk Encryption (FDE) Enterprise (SAS) drives are also supported.

The DS8800 supports Solid-State Drives (SSDs) with a capacity of 300 GB.

Solid State Disk drives (SSDs)

Solid State Disk drives (SSD) are the best choice for I/O-intensive workloads. They provide up to 100 times the throughput and 10 times lower response time than 15K rpm spinning disks. They also consume less power than traditional spinning disks. For more information, see Chapter 8, “Physical planning and installation” on page 201.

Host adapters

Each DS8700 Fibre Channel adapter offers four 4 Gbps or 8 Gbps Fibre Channel ports. Each 4 Gbps port independently auto-negotiates to either 1, 2, or 4 Gbps link speed. Each 8 Gbps port independently auto-negotiates to either 2, 4, or 8 Gbps link speed. Each of the four ports on an DS8700 adapter can also independently be either Fibre Channel protocol (FCP) or FICON.

Each DS8800 Fibre Channel adapter offers four or eight 8 Gbps Fibre Channel ports. Each 8 Gbps port independently auto-negotiates to either 2, 4, or 8 Gbps link speed. Each of the ports on an DS8800 host adapter can also independently be either Fibre Channel protocol (FCP) or FICON. For more information, see Chapter 3, “Hardware components and architecture” on page 35

IBM System Storage Productivity Center management console

The DS8000 leverages the IBM System Storage Productivity Center (SSPC), an advanced optional management console that can provide a view of both IBM and non-IBM storage environments. The SSPC can enable a greater degree of simplification for organizations confronted with the growing number of element managers in their environment. The SSPC is an external System x server with preinstalled software, including IBM Tivoli Storage Productivity Center Basic Edition (TPC-BE). TPC-BE is required for launching the DS GUI.

Utilizing TPC-BE, SSPC extends the capabilities available through the IBM DS Storage Manager. SSPC offers the unique capability to manage a variety of storage devices connected across the Storage Area Network (SAN). This allows the administrator to explore the health of the environment at an aggregate or in-depth view. The TPC-BE, which is pre-installed on the SSPC, can be optionally upgraded to Tivoli Storage Productivity Center Standard Edition (TPC-SE), which includes enhanced functionality:

- ▶ Monitoring and reporting capabilities that can be used to enable more in-depth performance reporting
- ▶ Asset and capacity reporting
- ▶ Automation for the DS8000
- ▶ Management of other resources including storage devices, server file systems, tape drives, tape libraries, and SAN environments.

For more information, see Chapter 12, “System Storage Productivity Center” on page 279.

Storage Hardware Management Console for the DS8000

The Hardware Management Console (HMC) is the focal point for maintenance activities. The HMC is a dedicated workstation that is physically located inside the DS8000 and can proactively monitor the state of your system, notifying you and IBM when service is required. It can also be connected to your network to enable centralized management of your system using the IBM System Storage DS® Command-Line Interface (DSCLI). The HMC supports the IPv4 and IPv6 standards. For more information, see Chapter 9, “DS8000 HMC planning and setup” on page 225.

An external management console is available as an optional feature and can be used as a redundant management console for environments with high availability requirements.

Tivoli Key Lifecycle Manager isolated key server

The Tivoli Key Lifecycle Manager (TKLM) software performs key management tasks for IBM encryption-enabled hardware, such as the IBM System Storage DS8000 Series family, by providing, protecting, storing, and maintaining encryption keys that are used to encrypt information being written to, and decrypt information being read from, encryption-enabled disks.

For DS8000 storage systems shipped with Full Disk Encryption (FDE) drives, two TKLM key servers are required. An Isolated Key Server (IKS) with dedicated hardware and non-encrypted storage resources is required and can be ordered from IBM. For more information, see Chapter 3, “Hardware components and architecture” on page 35

1.4.2 Storage capacity

The physical storage capacity, for both the DS8700 and DS8800, is contained in the disk drive sets. A disk drive set contains 16 Disk Drive Modules (DDMs), which have the same capacity and the same revolutions per minute (rpm). In addition, Solid State Disk drives (SSDs) are available in half sets (8) or full sets (16) of disk drives or DDMs. The available drive options

provide industry class capacity and price/performance to address enterprise application and business requirements. DS8000 storage capacity can be configured as RAID 5, RAID 6, RAID 10, or as a combination (restrictions apply for Full Disk Encryption and Solid-State Drives).

The DS8700 can have up to 1024 DDMs installed. The total storage capacity is up to 2 TBs. The 941 base frame installs with up to 128 drives. The first, second, and third expansion frame (94E) can have up to 256 drives, and the fourth expansion frame (94E) can have 128.

The DS8800 can have up to 1056 DDMs installed, with a total physical storage capacity up to 633 TBs. The 951 base frame installs with up to 240 drives. The first expansion frame (95E) can have up to 336 drives and the second expansion frame (95E) can have up to 480 drives.

IBM Standby Capacity on Demand offering for the DS8000

Standby Capacity on Demand (CoD) provides standby on demand storage for the DS8000 that allows you to access the extra storage capacity whenever the need arises. With CoD, IBM installs additional CoD disk drive sets in your DS8000. At any time, you can logically configure your CoD drives, concurrently with production, and you are automatically charged for the additional capacity. The models have the following capacities:

- ▶ DS8700 can have up to four Standby CoD drive sets (64 drives).
- ▶ DS8800 can have up to six Standby CoD drive sets (96 drives).

1.4.3 Supported environments

The DS8000 offers connectivity support across a broad range of server environments, including IBM Power Systems™, System z, System p, System i®, and System x servers, servers from Sun and Hewlett-Packard, non-IBM Intel, and AMD-based servers.

The DS8000 supports over 90 platforms. For the most current list of supported platforms, see the DS8000 System Storage Interoperation Center at:

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

This rich support of heterogeneous environments and attachments, along with the flexibility to easily partition the DS8000 storage capacity among the attached environments, can help support storage consolidation requirements and dynamic environments.

1.4.4 Easy Tier

Easy Tier is a DS8000 built-in dynamic data relocation feature that allows a host-transparent movement of data among the storage subsystem resources. This significantly improves the configuration flexibility, the performance tuning, and planning.

Tip: Easy Tier is available for DS8000 for no extra fee. It is not supported on systems with the Full Disk Encryption (FDE) feature.

Easy Tier can operate in two modes:

- ▶ Easy Tier Automatic Mode

Easy Tier Automatic Mode autonomically manages the capacity allocated in a DS8000 Extent Pool containing mixed disk technology (Enterprise, Solid State Disk, or Nearline) to place the most demanding pieces of data (hot data) on the appropriate storage media. The data relocation is at the extent level. This significantly improves the overall storage cost-performance ratio and simplifies the performance tuning and management.

► Easy Tier Manual Mode

Easy Tier manual mode allows a set of manual initiated actions to relocate data among the storage subsystem resources in dynamic fashion without any disruption of the host operations. The Easy Tier Manual Mode capabilities are *Dynamic Volume Relocation* and *Dynamic Extent Pool Merge*. Dynamic Volume Relocation allows a DS8000 volume to be migrated to the same or a separate Extent Pool. Dynamic Extent Pool Merge allows an Extent Pool to be merged to another Extent Pool. By combining these two capabilities, we can greatly improve the configuration flexibility of the DS8000.

Tip: Dynamic Volume Relocation within the same extent pool is not supported in hybrid pools.

Tip: Easy Tier is a DS8000 firmware function available with LMC level 6.1.xx or later. An additional LIC feature must be order and installed at no extra fee.

Easy Tier provides a performance monitoring capability whether or not the licensed feature is activated. This monitoring capability enables workload data collection to be off-loaded and further processed with the *Storage Tiering Advisor Tool*. Providing a graphical representation of hot data distribution at the volume level, this powerful tool allows you to analyze the workload characteristics and evaluate the benefits of the higher performance possible with the Solid State Drive technology.

Refer to *IBM System Storage DS8000 Easy Tier*, REDP-4667-01 for more information.

1.4.5 I/O Priority Manager

The I/O Priority Manager can help you effectively manage quality of service levels for each application running on your system. This feature aligns distinct service levels to separate workloads in the system to help maintain the efficient performance of each DS8000 volume. The I/O Priority Manager detects when a higher-priority application is hindered by a lower-priority application that is competing for the same system resources. This might occur when multiple applications request data from the same disk drives. When I/O Priority Manager encounters this situation, it reduces the lower-priority I/O streams to assist the more critical I/O streams in meeting their performance targets.

For more details, refer to *IBM System Storage DS8000: I/O Priority Manager*, REDP-4760.

1.4.6 Configuration flexibility

The DS8000 series uses virtualization techniques to separate the logical view of hosts onto Logical Unit Numbers (LUNs) from the underlying physical layer, thus providing high configuration flexibility. Virtualization is discussed in Chapter 5, “Virtualization concepts” on page 103.

Dynamic LUN/volume creation, deletion, and expansion

The DS8000 gives a high degree of flexibility in managing storage, allowing LUNs to be created and deleted nondisruptively. When a LUN is deleted, the freed capacity can be used with other free space to form a LUN of a different size. A LUN can also be dynamically increased in size.

Large LUN and large count key data (CKD) volume support

You can configure LUNs and volumes to span arrays, allowing for larger LUN sizes of up to 16 TB. Note that Copy Services is not supported for LUN sizes greater than 2 TBs. The maximum Count Key Data (CKD) volume size is 1,182,006 cylinders (1 TB), greatly reducing the number of volumes managed and creating a new volume type on z/OS called 3390 Model A. This capability is referred to as Extended Address Volumes (EAV) and requires z/OS 1.12 or later.

Flexible LUN-to-LSS association

With no predefined association of arrays to LSSs on the DS8000 series, users are able to put LUNs or CKD volumes into Logical Subsystems (LSSs) and make best use of the 256 address range, particularly for System z.

Simplified LUN masking

The implementation of volume group-based LUN masking simplifies storage management by grouping some or all World Wide Port Names (WWPNs) of a host into a Host Attachment. Associating the Host Attachment to a Volume Group allows all adapters within it access to all of the storage in the Volume Group.

Thin provisioning features

The DS8000 provides two types of space efficient volumes: Track Space Efficient volumes and Extent Space Efficient volumes. Both these features enable over-provisioning capabilities that provide more efficient usage of the storage capacity and reduced storage management requirements.

Logical definitions: Maximum values

The current DS8000 maximum values for the major logical definitions are:

- ▶ Up to 255 logical Subsystems (LSS)
- ▶ Up to 65280 logical devices
- ▶ Up to 16TB Logical Unit Numbers (LUNs)
- ▶ Up to 1,182,006 cylinders (1 TB) Count Key Data (CKD) volumes
- ▶ Up to 130560 Fibre Connection (FICON) logical paths (512 logical paths per control unit image) on the DS8000
- ▶ Up to 1280 logical paths per Fibre Channel (FC) port
- ▶ Up to 8192 process logins (509 per SCSI-FCP port)

1.4.7 Copy Services functions

For IT environments that cannot afford to stop their systems for backups, the DS8000 provides a fast replication technique that can provide a point-in-time copy of the data in a few seconds or even less. This function is called *FlashCopy*.

For data protection and availability needs, the DS8000 provides Metro Mirror, Global Mirror, Global Copy, Metro/Global Mirror, and z/OS Global Mirror, which are Remote Mirror and Copy functions. These functions are also available and are fully interoperable with previous models of the DS8000 family. These functions provide storage mirroring and copying over large distances for disaster recovery or availability purposes.

We briefly discuss Copy Services in Chapter 6, “IBM System Storage DS8000 Copy Services overview” on page 137. For detailed information about Copy Services, see the Redbooks *IBM*

System Storage DS8000: Copy Services for Open Systems, SG24-6788, and IBM System Storage DS8000: Copy Services for IBM System z, SG24-6787.

FlashCopy

The primary objective of FlashCopy is to quickly create a point-in-time copy of a source volume on a target volume. The benefits of FlashCopy are that the point-in-time target copy is immediately available for use for backups or testing, and that the source volume is immediately released so that applications can continue processing with minimal application downtime. The target volume can be either a logical or physical copy of the data, with the physical copy copying the data as a background process. In a z/OS environment, FlashCopy can also operate at a data set level.

The following sections summarize the options available with FlashCopy.

Multiple Relationship FlashCopy

Multiple Relationship FlashCopy allows a source to have FlashCopy relationships with up to 12 targets simultaneously.

Incremental FlashCopy

Incremental FlashCopy provides the capability to *refresh* a LUN or volume involved in a FlashCopy relationship. When a subsequent FlashCopy is initiated, only the data required to make the target current with the source's newly established point-in-time is copied.

Remote Pair FlashCopy

Remote Pair FlashCopy provides improvement to resiliency solutions by ensuring data synchronization when a FlashCopy target is also a Metro Mirror source. This keeps the local and remote site consistent which facilitates recovery, supports HyperSwap®, and reduces link bandwidth utilization.

Remote Mirror Primary FlashCopy

Remote Mirror primary FlashCopy allows a FlashCopy relationship to be established where the target is also a remote mirror primary volume. This enables a full or incremental point-in-time copy to be created at a local site, and then uses remote mirroring commands to copy the data to the remote site. While the background copy task is copying data from the source to the target, the remote mirror pair goes into a *copy pending* state.

Consistency Groups

FlashCopy Consistency Groups can be used to maintain a consistent point-in-time copy across multiple LUNs or volumes, or even multiple DS8000 systems.

Inband commands over remote mirror link

In a remote mirror environment, inband FlashCopy allows commands to be issued from the local or intermediate site and transmitted over the remote mirror Fibre Channel links for execution on the remote DS8000. This eliminates the need for a network connection to the remote site solely for the management of FlashCopy.

IBM FlashCopy SE

The IBM FlashCopy SE feature provides a “space efficient” copy capability that can greatly reduce the storage capacity needed for point-in-time copies. Only the capacity needed to save pre-change images of the source data is allocated in a copy repository. This enables more space efficient utilization than is possible with the standard FlashCopy function. Furthermore, less capacity can mean fewer disk drives and lower power and cooling requirements, which can help reduce costs and complexity. FlashCopy SE can be especially useful in the creation of temporary copies for tape backup, online application checkpoints, or

copies for disaster recovery testing. For more information about FlashCopy SE, refer to *IBM System Storage DS8000 Series: IBM FlashCopy SE*, REDP-4368.

Remote Mirror and Copy functions

The Remote Mirror and Copy functions include Metro Mirror, Global Copy, Global Mirror, and Metro/Global Mirror. There is also z/OS Global Mirror for the System z environments. As with FlashCopy, Remote Mirror and Copy functions can also be established between DS8000 systems.

The following sections summarize the Remote Mirror and Copy options available with the DS8000 series.

Metro Mirror

Metro Mirror, previously called Peer-to-Peer Remote Copy (PPRC), provides a synchronous mirror copy of LUNs or volumes at a remote site within 300 km. Metro Mirror Consistency Groups, when used with a supporting application, can be used to maintain data and transaction consistency across multiple LUNs or volumes, or even multiple DS8000 systems.

Global Copy

Global Copy, previously called Extended Distance Peer-to-Peer Remote Copy (PPRC-XD), is a non-synchronous, long-distance copy option for data migration and backup.

Global Mirror

Global Mirror provides an asynchronous mirror copy of LUNs or volumes over virtually unlimited distances. The distance is typically limited only by the capabilities of the network and channel extension technology being used. A Global Mirror Consistency Group is used to maintain data consistency across multiple LUNs or volumes, or even multiple DS8000 systems.

Metro/Global Mirror

Metro/Global Mirror is a three-site data replication solution for both Open Systems and the System z environments. Local site (Site A) to intermediate site (Site B) provides high availability replication using synchronous Metro Mirror, and intermediate site (Site B) to remote site (Site C) provides long distance disaster recovery replication using asynchronous Global Mirror.

Tip: Up to 32 Metro/Global Mirror sessions can be run in release 6.1.

z/OS Global Mirror

z/OS Global Mirror, previously called Extended Remote Copy (XRC), provides an asynchronous mirror copy of volumes over virtually unlimited distances for the System z. It now provides increased parallelism through multiple SDM readers (Multiple Reader capability).

z/OS Metro/Global Mirror

This is a combination of Copy Services for System z environments that uses z/OS Global Mirror to mirror primary site data to a remote location that is at a long distance, and Metro Mirror to mirror the primary site data to a location within the metropolitan area. This enables a z/OS three-site high availability and disaster recovery solution.

z/OS Global Mirror also offers Incremental Resync, which can significantly reduce the time needed to restore a Disaster Recovery (DR) environment after a HyperSwap in a three-site

z/OS Metro/Global Mirror configuration. After the Incremental Resync, you can change the copy target destination of a copy relation without requiring a full copy of the data.

1.4.8 Resource Groups for copy services scope limiting

Copy services scope limiting is the ability to specify policy-based limitations on copy services requests. With the combination of policy-based limitations and other inherent volume-addressing limitations, you can control which volumes can be in a copy services relationship, which network users or host LPARs issue copy services requests on which resources, and other copy services operations. Use these capabilities to separate and protect volumes in a copy services relationship from each other. This can assist you with multi-tenancy support by assigning specific resources to specific tenants, limiting copy services relationships so that they exist only between resources within each tenant's scope of resources, and limiting a tenant's copy services operators to an “operator only” role. When managing a single-tenant installation, the partitioning capability of resource groups can be used to isolate various subsets of the environment as though they were separate tenants. For example, to separate mainframes from open servers, Windows from UNIX, or accounting departments from telemarketing.

For more information, refer to *IBM System Storage DS8000: Resource Groups*, REDP-4758.

1.4.9 Service and setup

The installation of the DS8000 is performed by IBM in accordance with the installation procedure for this machine. The client's responsibility is the installation planning, retrieval and installation of feature activation codes, logical configuration, and execution.

For maintenance and service operations, the Storage Hardware Management Console (HMC) is the focal point. The management console is a dedicated workstation that is physically located inside the DS8000 storage system where it can automatically monitor the state of your system. It will notify you and IBM when service is required. Generally, use a dual HMC configuration, particularly when using Full Disk Encryption.

The HMC is also the interface for remote services (Call Home and Remote Support), which can be configured to meet client requirements. It is possible to allow one or more of the following:

- ▶ Call home on error (machine-detected)
- ▶ Connection for a few days (client-initiated)
- ▶ Remote error investigation (service-initiated)

The remote connection between the management console and the IBM Service organization is done using a Virtual Private Network (VPN) point-to-point connection over the internet or modem. A new secure SSL connection protocol option is available for call home support and additional audit logging.

The DS8000 storage system can be ordered with an outstanding four-year warranty, an industry first, on both hardware and software.

1.4.10 IBM Certified Secure Data Overwrite

Sometimes regulations and business prudence require that the data actually be removed when the media is no longer needed.

An STG Lab Services Offering for the DS8000 series includes the following services:

- ▶ Multi-pass overwrite of the data disks in the storage system

- Purging of client data from the server and HMC disks

Tip: The secure overwrite functionality is offered as a service exclusively and is not intended for use by clients, IBM Business Partners, or IBM field support personnel.

For more information about the IBM Certified Secure Data Overwrite service offerings, contact your IBM sales representative or IBM Business Partner.

1.5 Performance features

The IBM System Storage DS8000 offers optimally balanced performance. This is possible because the DS8000 incorporates many performance enhancements, such as the dual 2-way and dual 4-way POWER6 and POWER6+ processor complex implementation, fast 8 Gbps Fibre Channel/FICON host adapter cards, Solid State Disk drives, and the high bandwidth, fault-tolerant point-to-point PCI Express internal interconnections.

With all these components, the DS8000 is positioned at the top of the high performance category.

1.5.1 Sophisticated caching algorithms

IBM Research conducts extensive investigations into improved algorithms for cache management and overall system performance improvements.

Sequential Prefetching in Adaptive Replacement Cache

One of the performance enhancers of the DS8000 is its self-learning cache algorithm, which improves cache efficiency and enhances cache hit ratios. This algorithm, which is used in the DS8000 series, is called Sequential Prefetching in Adaptive Replacement Cache (SARC).

SARC provides the following abilities:

- Sophisticated algorithms to determine what data should be stored in cache based on recent access and frequency needs of the hosts
- Pre-fetching, which anticipates data prior to a host request and loads it into cache
- Self-learning algorithms to adaptively and dynamically learn what data should be stored in cache based upon the frequency needs of the hosts

Adaptive Multi-stream Prefetching

Adaptive Multi-stream Prefetching (AMP) is a breakthrough caching technology that improves performance for common sequential and batch processing workloads on the DS8000. AMP optimizes cache efficiency by incorporating an autonomic, workload-responsive, and self-optimizing prefetching technology.

Intelligent Write Caching

Intelligent Write Caching (IWC) improves performance through better write cache management and destaging order of writes. It can double the throughput for random write workloads. Specifically, database workloads benefit from this new IWC Cache algorithm.

SARC, AMP, and IWC play complementary roles. While SARC is carefully dividing the cache between the RANDOM and the SEQ lists to maximize the overall hit ratio, AMP is managing the contents of the SEQ list to maximize the throughput obtained for the sequential workloads. IWC manages the write cache and decides what order and rate to destage to disk.

1.5.2 Solid State Disk drives (SSD drives)

To improve data transfer rate (IOPS) and response time, the DS8000 series provides support for Solid State Disk drives (SSDs). SSDs have improved I/O transaction-based performance over traditional spinning drives. The DS8700 is available with 600 GB. The DS8800 is available with 300 GB.

SSDs are high-IOPS class enterprise storage devices targeted at Tier 0, I/O-intensive workload applications that can use a high level of fast-access storage. SSDs offer a number of potential benefits over Hard Disk Drives, including better IOPS, lower power consumption, less heat generation, and lower acoustical noise. For more information, see Chapter 8, “Physical planning and installation” on page 201.

1.5.3 Multipath Subsystem Device Driver

The Multipath Subsystem Device Driver (SDD) is a pseudo-device driver on the host system designed to support the multipath configuration environments in IBM products. It provides load balancing and enhanced data availability capability. By distributing the I/O workload over multiple active paths, SDD provides dynamic load balancing and eliminates data-flow bottlenecks. SDD helps eliminate a potential single point of failure by automatically rerouting I/O operations when a path failure occurs.

SDD is provided with the DS8000 series at no additional charge. Fibre Channel (SCSI-FCP) attachment configurations are supported in the AIX, HP-UX, Linux, Windows, Novell NetWare, and Oracle Solaris environments.

Tip: Support for multipath is included in an IBM i server as part of Licensed Internal Code and the IBM i operating system (including i5/OS®).

For more information about SDD, refer to *IBM System Storage DS8000: Host Attachment and Interoperability*, SG24-8887.

1.5.4 Performance for System z

The DS8000 series supports the following IBM performance enhancements for System z environments:

- ▶ *Parallel Access Volumes (PAVs)* enable a single System z server to simultaneously process multiple I/O operations to the same logical volume, which can help to significantly reduce device queue delays. This is achieved by defining multiple addresses per volume. With Dynamic PAV, the assignment of addresses to volumes can be automatically managed to help the workload meet its performance objectives and reduce overall queuing. PAV is an optional feature on the DS8000 series.
- ▶ *HyperPAV* is designed to enable applications to achieve equal or better performance than with PAV alone, while also using fewer Unit Control Blocks (UCBs) and eliminating the latency in targeting an alias to a base. With HyperPAV, the system can react immediately to changing I/O workloads.
- ▶ *Multiple Allegiance* expands the simultaneous logical volume access capability across multiple System z servers. This function, along with PAV, enables the DS8000 series to process more I/Os in parallel, helping to improve performance and enabling greater use of large volumes.
- ▶ *I/O priority queuing* allows the DS8000 series to use I/O priority information provided by the z/OS Workload Manager to manage the processing sequence of I/O operations.

- ▶ *High Performance FICON for z (zHPF)* reduces the impact associated with supported commands on current adapter hardware, thereby improving FICON throughput on the DS8000 I/O ports. The DS8000s also supports the new zHPF I/O commands for multi-track I/O operations.

Chapter 7, “Performance” on page 157, gives you more information about the performance aspects of the DS8000 family.

1.5.5 Performance enhancements for IBM Power Systems

Many IBM Power Systems users can benefit from the following DS8000 features:

- ▶ End-to-end I/O priorities
- ▶ Cooperative caching
- ▶ Long busy wait host tolerance
- ▶ Automatic Port Queues

Chapter 7, “Performance” on page 157, gives you more information about these performance enhancements.

1.5.6 Performance enhancements for z/OS Global Mirror

Many users of z/OS Global Mirror, which is the System z-based asynchronous disk mirroring capability, will benefit from the DS8000 enhancement “z/OS Global Mirror suspend instead of long busy option”. In the event of high workload peaks, which can temporarily overload the z/OS Global Mirror configuration bandwidth, the DS8000 can initiate a z/OS Global Mirror SUSPEND, preserving primary site application performance, which is an improvement over the previous LONG BUSY status.

Consider the following points:

- ▶ All users of z/OS Global Mirror benefit from the DS8000’s “z/OS Global Mirror Multiple Reader” support. This recent enhancement spreads the z/OS Global Mirror workload across more than a single reader. In the event of high workload peaks restricted to a few volumes, which can mean restricted to a single reader, the peak demand can now be balanced across a set of up to 16 readers. This enhancement provides more efficient use of the site-to-site network capacity (a higher single volume throughput capability) and an environment that can effectively use Parallel Access Volumes (PAVs).
- ▶ Extended Distance FICON is a capability that can help reduce the need for channel extenders in z/OS Global Mirror configurations by increasing the numbers of read commands in flight.

See *IBM System Storage DS8000: Copy Services for IBM System z*, SG24-6787, for a detailed discussion of z/OS Global Mirror and related enhancements.



IBM System Storage DS8000 models

This chapter provides an overview of the DS8700 and DS8800 storage subsystems, describes the different models, and how well they scale regarding capacity and performance.

2.1 DS8700 Model overview

The DS8000 series includes the DS8700 Model 941 base frame and the associated DS8700 expansion frame model 94E.

► DS8700 Model 941

This model is available as either a dual 2-way processor complex with enclosures for 128 DDMs and 16 FC host adapter cards, or a dual 4-way processor complex with enclosures for 128 DDMs and 16 FC host adapter cards. Host adapter cards can be either 8 Gbps (certain restrictions apply) or 4 Gbps.

Tip: Model 941 supports nondisruptive upgrades from dual 2-way to dual 4-way.

► DS8700 Model 94E

This expansion frame for the 941 model includes enclosures for additional DDMs and additional FC host adapter cards to allow a maximum configuration of 32 FC host adapter cards. The expansion frame 94E can only be attached to the 941 4-way processor complex. Up to four expansion frames can be attached to a Model 941. Additional FC host adapter cards can only be installed in the first expansion frame.

Tip: A Model 941 supports nondisruptive upgrades from an eight drive install to a full system with four expansion frames.

Table 2-1 provides a comparison of the DS8700 model 941 and its available combination of resources.

Table 2-1 DS8700 series Model 941 comparison and additional resources

Base model	Expansion model	Processor type	Max DDMs	Max processor memory	Max host adapters
941	None	2-way 4.7 GHz	128	128 GB	16
941	None	4-way 4.7 GHz	128	384 GB	16
	1 x 94E		384		32
	2 x 94E		640		
	3 x 94E		896		
	4 x 94E		1024		

Each Fibre Channel/FICON host adapter has four Fibre Channel ports, providing up to 128 Fibre Channel ports for a maximum configuration.

Machine type 242x

DS8700 series models, like DS8800 models, are associated with machine type 242x. This machine type corresponds to the length of warranty offer that allows a 1 year, 2 year, 3 year, or 4 year warranty period (x=1, 2, 3, or 4, respectively). The 94E expansion frame has the same 242x machine type as the base frame.

2.1.1 DS8700 Model 941 overview

The DS8700 Model 941, shown in Figure 2-1, has the following features:

- ▶ Two processor complexes, each with a IBM System p POWER6 4.7 GHz 2-way or 4-way central electronic complex.
- ▶ A 2-way configuration requires two battery packs. A 4-way configuration requires three battery packs.
- ▶ A base frame with up to 128 DDMs for a maximum disk storage capacity of 256 TB using 2 TB nearline DDMs.
- ▶ Up to 128 GB (2-way) or 384 GB (4-way) of processor memory, also referred to as the *cache*. Note that the DS8700 supports concurrent cache upgrades.
- ▶ 4 Gbps or 8 Gbps Fibre Channel/FICON host adapters (HAs). A total of 16 HAs can be installed, eight of which can be 8 Gbps HAs. Each port can be independently configured as either:
 - FCP port to open systems hosts attachment
 - FCP port for Metro Mirror, Global Copy, Global Mirror, and Metro/Global Mirror connectivity
 - FICON port to connect to System z hosts
 - FICON port for z/OS Global Mirror connectivity
 - This totals up to 64 ports with any mix of FCP and FICON ports
- ▶ The DS8700 Model 941 can connect up to four expansion frames (Model 94E). Figure 2-1 displays a front view of a DS8700 Model 941 and 94E with the covers off.

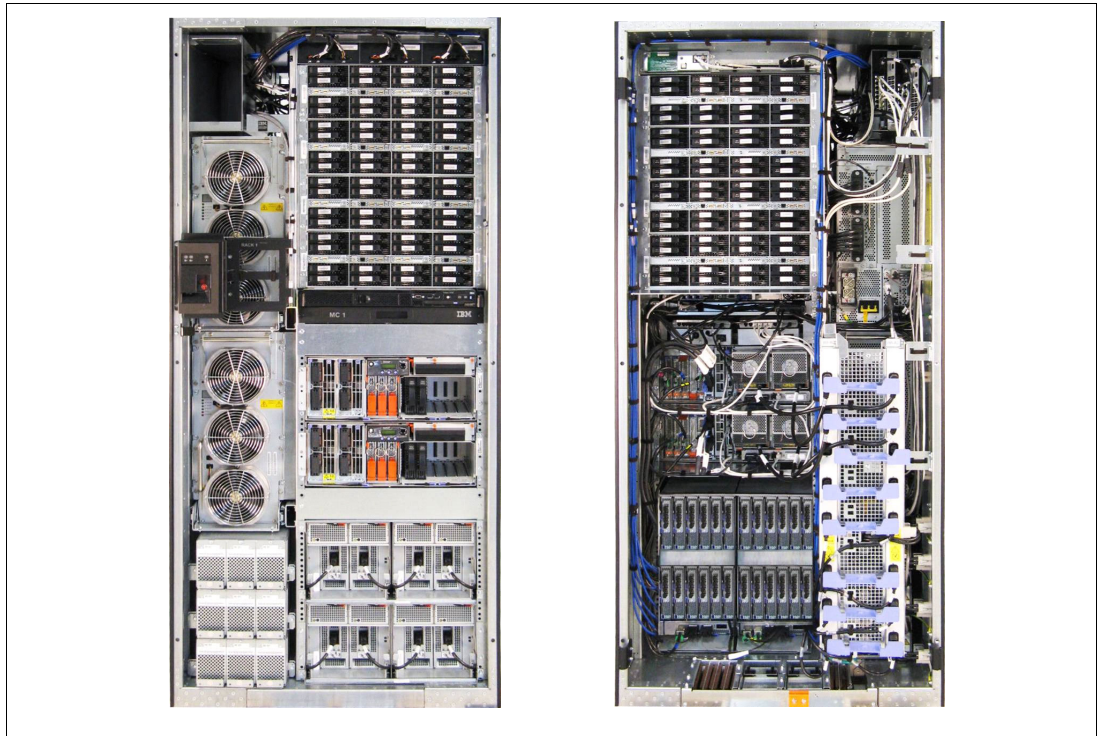


Figure 2-1 DS8700 base frame with covers removed: Front and rear

Figure 2-2 shows the maximum configuration for a DS8700 Model 941 base frame with one 94E expansion frame. It shows the placement of the hardware components within the frames.

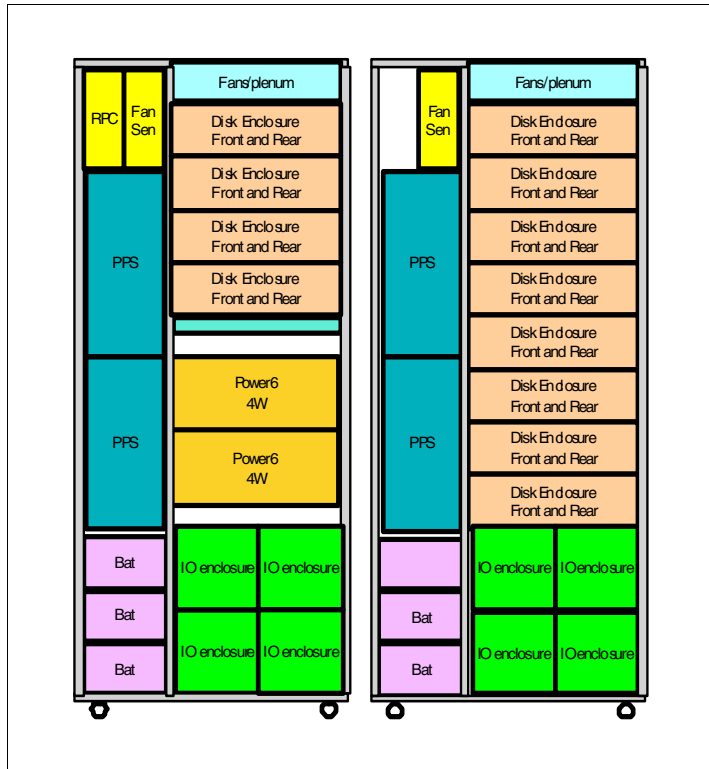


Figure 2-2 DS8700 configuration: 941 base unit with one 94E expansion frame

There are no I/O enclosures installed for the second, third, and fourth expansion frames. The result of installing all possible 1024 DDMs is that they will be distributed evenly over all the device adapter (DA) pairs. For an explanation of DA pairs, refer to 3.5.4, “Device adapters” on page 57. Figure 2-3 shows a maximum DS8700 configuration.

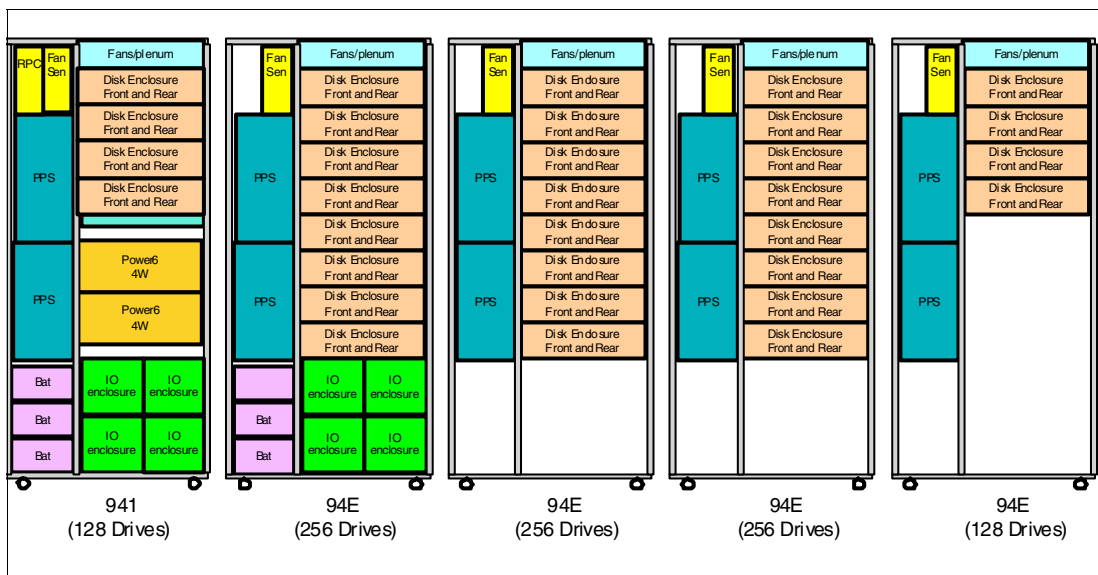


Figure 2-3 DS8700 models 941/94E maximum configuration with 1024 disk drives

The DS8700 can contain 300, 450, and 600 GB (15K RPM) enterprise disk drives and 2 TB (7,200 RPM) nearline (SATA) disk drives.

Besides enterprise and nearline hard disk drives (HDDs), it is also possible to install 73 GB, 146 GB, and 600 GB Solid-State Drives (SSDs) in the DS8700. However, the 73 and 146 GB SSDs cannot be ordered anymore. SSD drives can be ordered in 16 drive install groups (disk drive set) like HDD drives, or in eight drive install groups (half disk drive set). The suggested configuration of SSDs, for optimum price performance, is 16 drives per DA pair. For additional information about SSD configurations, see 8.5.3, “DS8000 Solid State Drive (SSD) considerations” on page 222.

Tip: Intermix of drives of different capacity and speed is supported on a DA pair, but not within a storage enclosure pair.

The DS8700 can be ordered with Full Disk Encryption (FDE) drives, with a choice of 300 GB (15K RPM), 450 GB (15K RPM), and 600 GB (15K RPM) enterprise drives. You cannot intermix FDE drives with other drives in a DS8700 system. For additional information about FDE drives, refer to *IBM System Storage DS8700 Disk Encryption Implementation and Usage Guidelines*, REDP-4500.

- In the 2-way configuration, the DS8700 model 941 can have up to 128 DDMs and 16 FC adapter cards with up to eight 8 Gbps host adapter cards.
- In the 4-way configuration, the DS8700 model 941 can have up to 1024 DDMs and 32 FC adapter cards with up to sixteen 8 Gbps host adapter cards.

Tip: The placement of 8 Gbps HAs is restricted within an I/O enclosure. For more information, see Chapter 3, “Hardware components and architecture” on page 35

Table 2-2 summarizes the capacity characteristics of the DS8700. The minimum capacity is achieved by installing one half drive group of eight 73 GB SSD drives (note that the 73 GB drives have now been withdrawn from marketing).

Table 2-2 Capacity comparison of device adapters, DDMs, and storage capacity

Component	2-way base frame with one I/O enclosure pair	2-way and 4-way base with two I/O enclosure pairs	4-way (one expansion frame)	4-way (four expansion frames)
DA pairs	1	1 to 2	1 to 8	1 to 8
HDDs	Up to 64 increments of 16	Up to 128 increments of 16	Up to 384 increments of 16	Up to 1024 increments of 16
SSDs	Up to 32 increments of 8	Up to 64 increments of 8	Up to 192 increments of 8	Up to 256 increments of 8
Physical capacity	0.6 to 128 TB	0.6 to 256 TB	0.6 to 768 TB	0.6 to 2048 TB

Adding DDMs and Capacity on Demand

The DS8700 has a linear capacity growth up to 2048 TB.

A significant benefit of the DS8700 is the ability to add DDMs without disruption. IBM offers capacity on demand solutions that are designed to meet the changing storage needs of rapidly growing e-business. The Standby Capacity on Demand (CoD) offering is designed to

provide you with the ability to tap into additional storage. This is particularly useful if you have rapid or unpredictable storage growth. Up to four standby CoD disk drive sets (64 disk drives) can be concurrently field-installed into your system. To activate, simply logically configure the disk drives for use, which is a nondisruptive activity that does not require intervention from IBM.

Tip: A significant benefit of the DS8700 is the ability to add DDMs without disruption.

Upon activation of any portion of a standby CoD disk drive set, you must place an order with IBM to initiate billing for the activated set. At that time, you can also order replacement standby CoD disk drive sets. For more information about the standby CoD offering, refer to the DS8700 series announcement letter, which can be found at the following address:

<http://www.ibm.com/common/ssi/index.wss>

Device Adapters and performance

By default, the DS8700 comes with a new pair of Device Adapters per 64 DDMs. If you order a system with, for example, 128 drives, you will get two Device Adapter (DA) pairs. When ordering 512 disk drives, you get eight DA pairs, which is the maximum number of DA pairs. Adding more drives will not add DA pairs. Having enough DA pairs is important to achieve the high throughput level required by certain sequential workloads, such as data warehouse installations.

It is possible that your sequential throughput requirements will be high, but your capacity requirements are low. For example, you might have capacity requirements for 256 disks only, but still want the full sequential throughput potential of all DAs. For such situations, IBM offers the *Performance Accelerator* feature (PAF, FC 1980). It is a plant-only feature, meaning that it can only be installed on new machines. After the feature is enabled, you will get one new DA pair for every 32 DDMs. The feature is supported on machines with at most two expansion units, but the second expansion unit can only hold up to 64 drives. With this feature, the maximum configurations are:

- ▶ Base frame only: Two DA pairs and 64 drives
- ▶ One expansion unit: Six DA pairs and 192 disk drives
- ▶ Two expansion units: Eight DA pairs and 256 drives.

Figure 2-4 on page 27 shows a Performance Accelerator configuration of 16 HDD disk drive sets. This is the maximum PAF configuration: no more drives can be added. The example assumes that all drives are of the same type and capacity.

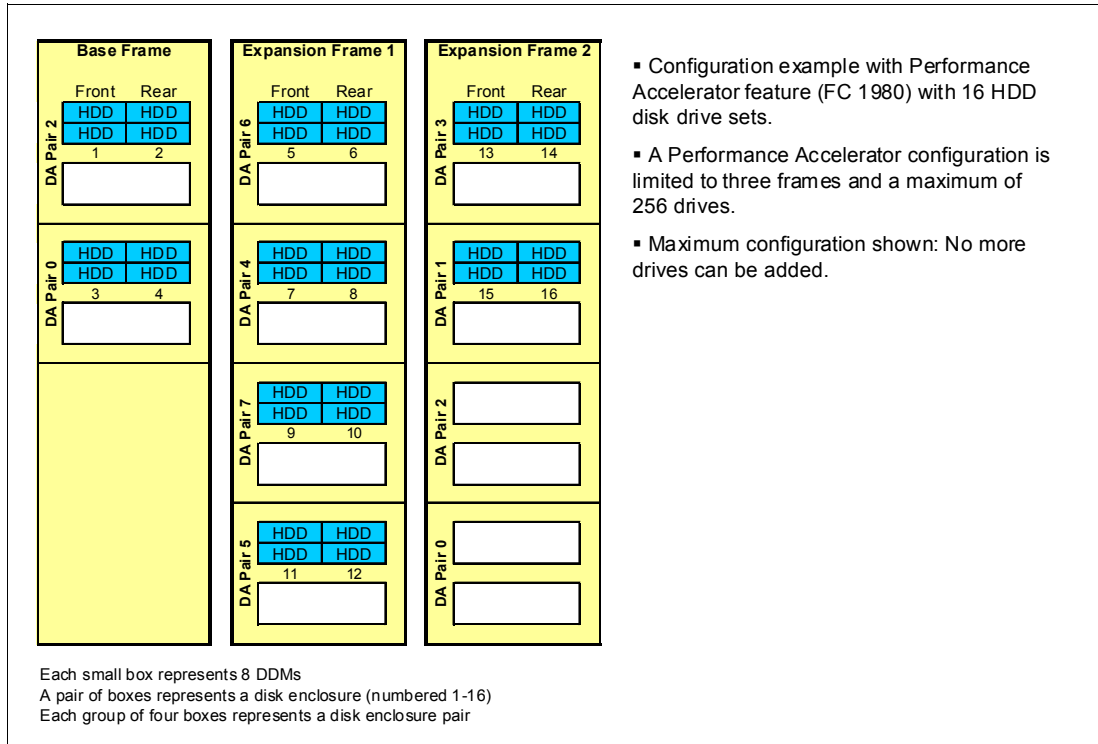


Figure 2-4 Configuration with Performance Accelerator feature

Scalable upgrades

With the DS8700, it is possible to start with a 2-way configuration with disk enclosures for 64 DDMs, and grow to a full scale, five frame configuration concurrently. See the upgrade path illustrated in Figure 2-5 for details.

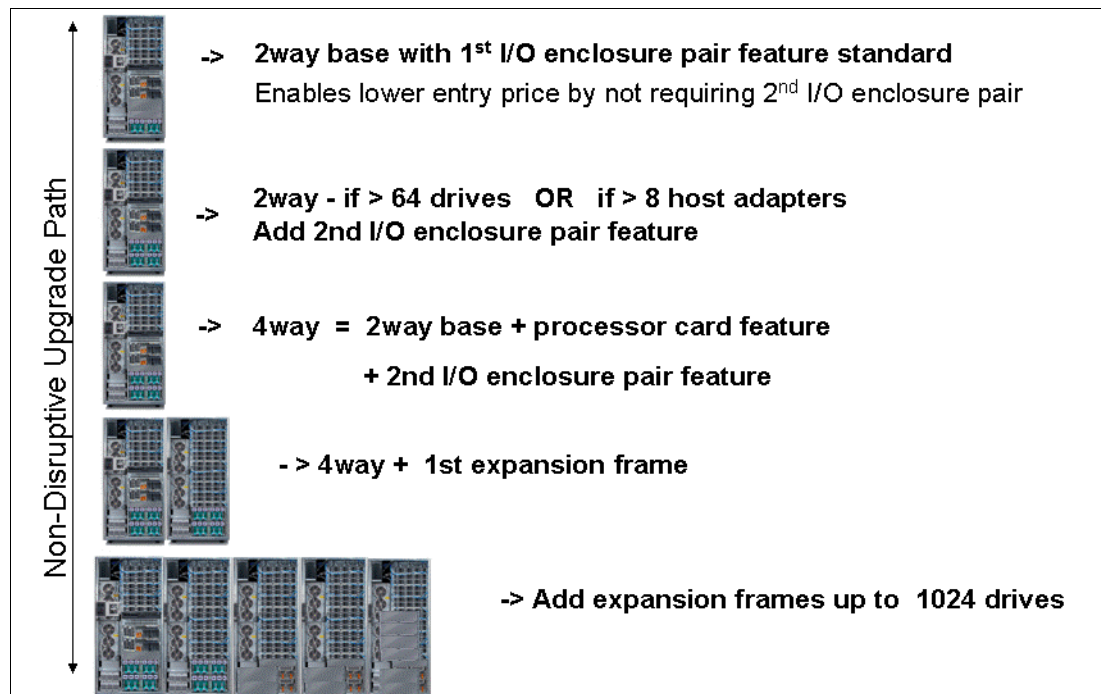


Figure 2-5 DS8700 concurrent upgrade path

2.2 DS8800 model overview

The DS8000 family includes the DS8800 Model 951 base frame and the associated DS8800 expansion frame 95E.

The DS8800 is available in either of the following configurations:

► DS8800 Model 951 standard cabling

This model is available as either a dual 2-way processor complex with storage enclosures for up to 144 DDMs and 4 FC host adapter cards, or as a dual 4-way processor complex with storage enclosures for up to 240 DDMs and 8 FC host adapter cards. Standard cabling is optimized for performance and highly scalable configurations, allowing large long-term growth. Model 951 standard cabling supports nondisruptive upgrades from dual 2-way to dual 4-way.

► DS8800 Model 951 Business Class cabling

This configuration of the Model 951 is available as a dual 2-way processor complex with storage enclosures for up to 240 DDMs and 4 FC host adapter cards. A business class system can now be configured with a minimum of 16 GB of cache. The business class option allows a system to be configured with more drives per device adapter, thus reducing configuration cost and increasing adapter utilization. Scalability is limited with this option.

Tip: Expansion frames cannot be added to the business class configuration.

► DS8800 Model 95E

This expansion frame for the 951 model includes enclosures for additional DDMs and additional FC adapter cards to allow a maximum configuration of 16 FC adapter cards. The expansion frame 95E can only be attached to the 951 4-way base frame. Up to two expansion frames can be attached to a model 951. FC adapter cards can only be installed in the first expansion frame.

Tip: Model 951 supports nondisruptive upgrades from an eight drive install to a full three frame system.

Table 2-3 provides a comparison of the DS8800 model 951 and its available combination of resources.

Table 2-3 DS8800 series model 951 comparison and additional resources

Base model	Cabling	Expansion model	Processor type	Max DDMs	Max processor memory	Max host adapters
951	Business ^a	None	2-way 5.0 GHz	240	64 GB	4
951	Standard	None	2-way 5.0 GHz	144	128 GB	4
951	Standard	None	4-way 5.0 GHz	240	384 GB	8
951	Standard	1 x 95E	4-way 5.0 GHz	576	384 GB	16

Base model	Cabling	Expansion model	Processor type	Max DDMs	Max processor memory	Max host adapters
951	Standard	2 x 95E	4-way 5.0 GHz	1056	384 GB	16

a. For copy services support, a business class system with feature code 4211 (16 GB memory) requires a memory upgrade to either 32 GB or 64 GB to increase the number of DDMs beyond 96. Business class with feature code 4211 does not support copy services, full disk encryption, or SSDs.

Each Fibre Channel/FICON host adapter has four or eight Fibre Channel ports, providing up to 128 Fibre Channel ports for a maximum configuration.

Machine type 242x

DS8800 models are associated to machine type 242x. This machine type corresponds to the length of warranty offer that allows a 1-year, 2-year, 3-year, or 4-year warranty period (x=1, 2, 3, or 4, respectively). The 95E expansion frame has the same 242x machine type as the base frame.

2.2.1 DS8800 Model 951 overview

The DS8800 Model 951, shown in Figure 2-6 on page 30, has the following features:

- ▶ A base frame with up to 240 DDMs for a maximum base frame disk storage capacity of 140 TB in high density storage enclosures. **1**
- ▶ Two processor complexes, each with a IBM System p POWER6+ 5.0 GHz, 2-way or 4-way central electronic complex. **2**
- ▶ Up to 128 GB (2-way) or 384 GB (4-way) of processor memory, also referred to as the *cache*. Note that the DS8800 supports concurrent cache upgrades.
- ▶ Up to eight 4-port or 8-port Fibre Channel/FICON host adapters (HAs) of 8 Gbps. **3** Each port can be independently configured as either:
 - FCP port to open systems hosts attachment
 - FCP port for Metro Mirror, Global Copy, Global Mirror, and Metro/Global Mirror connectivity
 - FICON port to connect to System z hosts
 - FICON port for z/OS Global Mirror connectivity

This totals up to 64 ports with any mix of FCP and FICON ports.

- ▶ Both 2-way and 4-way configurations require two battery packs. **4**
- ▶ The DS8800 has redundant primary power supplies (PPS) **5**. They provide a redundant 208 VDC power distribution to the rack. The processor complex, I/O enclosures, and storage enclosures have dual power supplies that are connected to the rack power distribution units (PDUs). **6**

- The DS8800 model 951 can connect up to two expansion frames (model 95E). Figure 2-6 displays a front and rear view of a DS8800 model 951 with the covers off, displaying the indicated components.

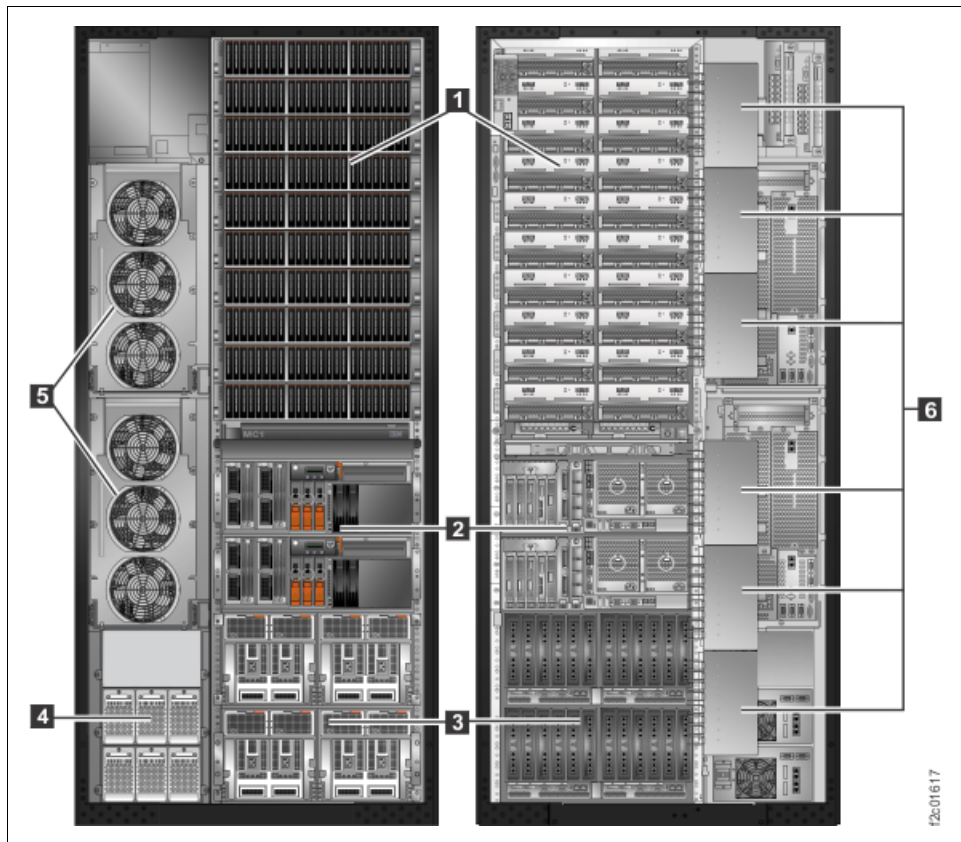


Figure 2-6 DS8800 base frame with covers removed: front and rear

Figure 2-7 shows the maximum configuration for a DS8800 model 951 base frame with one 95E expansion frame.

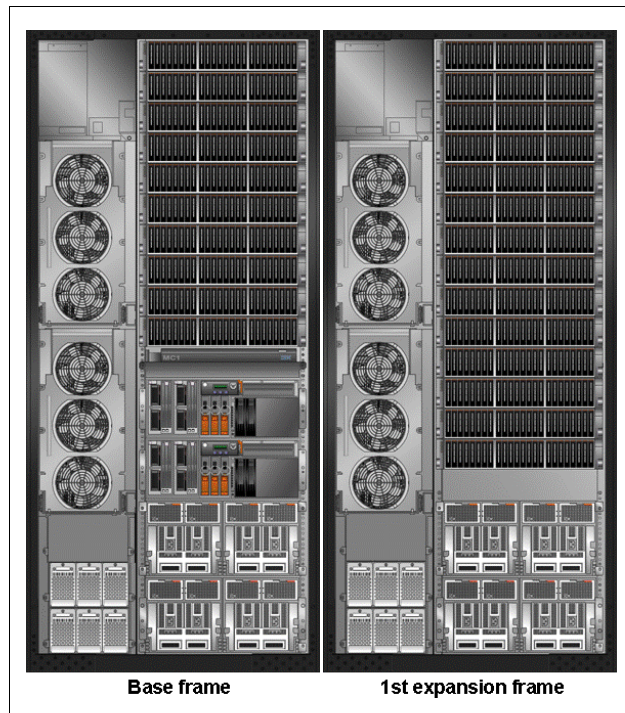


Figure 2-7 DS8800 configuration: 951 base unit with one 95E expansion frame: Front

There are no I/O enclosures installed in the second expansion frame. In a full three frame installation, the 1056 drives are distributed over all the device adapter (DA) pairs. For an explanation of DA pairs, see 3.5.4, “Device adapters” on page 57. The second 95E expansion frame is displayed in Figure 2-8.

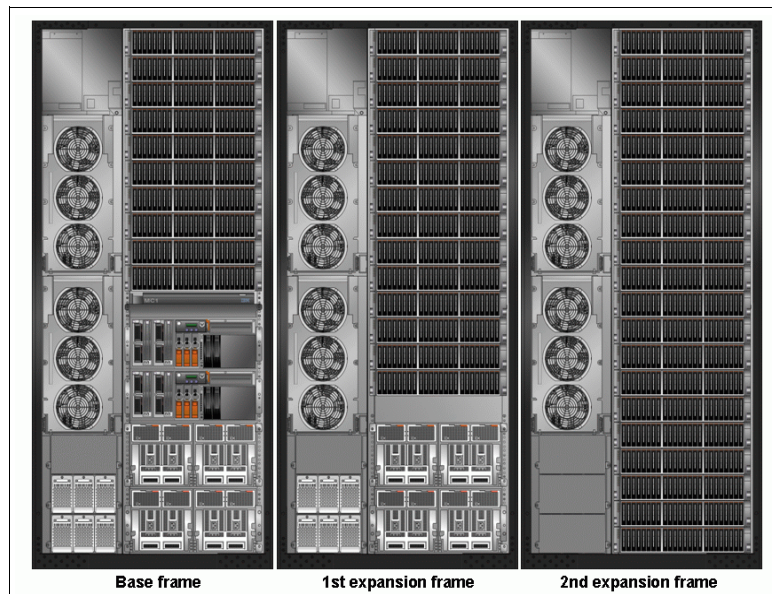


Figure 2-8 DS8800 models 951/95E maximum configuration with 1056 disk drives: Front

The DS8800 business class cabling option of the Model 951 is available as a dual 2-way processor complex with storage enclosures for up to 240 DDMs and 4 FC adapter cards. Figure 2-9 shows the maximum configuration of 2-way standard versus 2-way business class cabling.

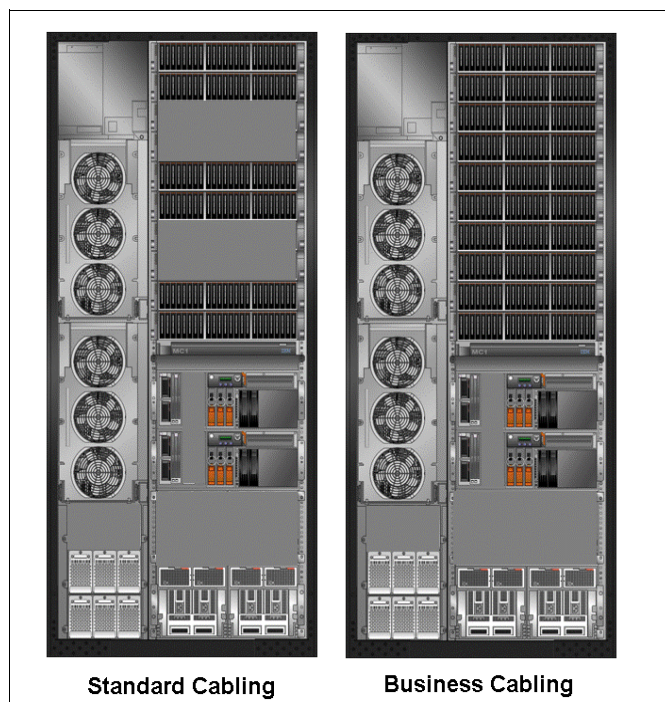


Figure 2-9 DS8800 2-way processor standard cabling versus business cabling: Front

The DS8800 offers enterprise class drives that feature a 6 Gbps interface. These enterprise class drives, using a 2.5 inch form factor, provide increased density and thus increased performance per frame. The drives are available in 146 GB (15K RPM), 450 GB (10K RPM), and 600 GB (10K RPM) capacities.

Besides enterprise class hard disk drives (HDDs), it is also possible to install 300 GB Solid-State Drives (SSDs) in the DS8800. SSD drives can be ordered in 16 drive install groups (disk drive set), like HDD drives, or in eight drive install groups (half disk drive set). The suggested configuration of SSDs for optimum price to performance ratio is 16 drives per DA pair. For additional information about SSD configurations, see 8.5.3, “DS8000 Solid State Drive (SSD) considerations” on page 222.

Tip: Intermix of drives of different capacity and speed is supported on a DA pair, but not within a storage enclosure pair.

The DS8800 can be ordered with Full Disk Encryption (FDE) drives, with a choice of 450 GB (10K RPM) and 600 GB (10K RPM) enterprise drives. You cannot intermix FDE drives with other drives in a DS8800 system. For additional information about FDE drives, see IBM Encrypted Storage Overview and Customer Requirements at the following URL:

<http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP101479>

- ▶ The DS8800 model 951 can have up to 144 DDMs and 4 FC adapter cards in the 2-way standard configuration.
- ▶ The DS8800 model 951 can have up to 240 DDMs and 4 FC adapter cards in the 2-way business class configuration.

- The DS8800 model 951 can have up to 240 DDMs and 8 FC adapter cards in the 4-way configuration in the base frame. The DS8800 model 951 4-way configuration supports up to 1056 DDMs and 16 FC adapter cards with two expansion frames.

A summary of the capacity characteristics is listed in Table 2-4. The minimum capacity is achieved by installing one 16 drive group of 146 GB 15K enterprise drives.

Table 2-4 Capacity comparison of device adapters, DDMs, and storage capacity.

Component	2-way base frame business cabling	2-way base frame standard cabling	4-way base frame	4-way (one expansion frame)	4-way (two expansion frames)
DA pairs	1 or 2	1 or 2	1 to 4	1 to 8	1 to 8
HDDs	Up to 240 increments of 16	Up to 144 increments of 16	Up to 240 increments of 16	Up to 576 increments of 16	Up to 1056 increments of 16
SSDs	N/A	Up to 96 increments of 8	Up to 192 increments of 8	Up to 384 increments of 8	Up to 384 increments of 8
Physical capacity	2.3 to 144 TB	2.3 to 86 TB	2.3 to 144 TB	2.3 to 346 TB	2.3 to 633 TB

Adding DDMs and Capacity on Demand

The DS8800 series has a linear capacity growth up to 633 TB.

A significant benefit of the DS8800 series is the ability to add DDMs without disruption. IBM offers capacity on demand solutions that are designed to meet the changing storage needs of rapidly growing e-business. The Standby Capacity on Demand (CoD) offering is designed to provide you with the ability to tap into additional storage and is particularly attractive if you have rapid or unpredictable storage growth.

Up to six standby CoD disk drive sets (96 disk drives) can be concurrently field-installed into your system. To activate, you simply logically configure the disk drives for use, which is a nondisruptive activity that does not require intervention from IBM.

Upon activation of any portion of a standby CoD disk drive set, you must place an order with IBM to initiate billing for the activated set. At that time, you can also order replacement standby CoD disk drive sets. For more information about the standby CoD offering, refer to the DS8800 series announcement letter, which can be found at the following URL:

<http://www.ibm.com/common/ssi/index.wss>

Device Adapters and performance

By default, the DS8800 comes with a pair of Device Adapters per 48 DDMs. If you order a system with, for example, 96 drives, you will get two Device Adapter (DA) pairs.

When ordering 432 disk drives, you get eight DA pairs, which is the maximum number of DA pairs. Adding more drives will not add DA pairs.

Having enough DA pairs is important to achieve the high throughput level required by certain sequential workloads, such as data warehouse installations.

Scalable Upgrades

With the DS8800, it is now possible to start with a 2-way configuration with disk enclosures for 48 DDMs, and grow to a full scale, 1056 drive 3-frame configuration concurrently.

- ▶ 2-way base with one I/O enclosure pair
 - Enables lower entry price by not requiring second I/O enclosure pair
- ▶ 4-way = 2-way base + processor card feature + second I/O enclosure pair feature
 - Enables improved performance on base rack
- ▶ 4-way base + first expansion frame
 - Enables 4 I/O enclosure pairs and 16 host adapters and 8 device adapter pairs
- ▶ 4-way base with first expansion frame + second expansion frame
 - Enables up to 1056 drives



Hardware components and architecture

This chapter describes the hardware components of the IBM System Storage DS8700 and DS8800. It provides readers with more insight into individual components and the architecture that holds them together. Although functionally the DS8700 and DS8800 are similar, many of their hardware components are different. Where there is a significant differences, we will describe the two systems separately.

The following topics are covered in this chapter:

- ▶ Frames: DS8700
- ▶ Frames - DS8800
- ▶ DS8000 architecture
- ▶ Storage facility processor complex (CEC)
- ▶ Disk subsystem
- ▶ Host adapters
- ▶ Power and cooling
- ▶ Management console network
- ▶ System Storage Productivity Center
- ▶ Isolated Tivoli Key Lifecycle Manager (TKLM) server

3.1 Frames: DS8700

The DS8700 is designed for modular expansion. From a high-level view, there appear to be three types of frames available for the DS8700. However, on closer inspection, the frames themselves are almost identical. The only variations are the combinations of processors, I/O enclosures, batteries, and disks that the frames contain.

Figure 3-1 is an attempt to show frame variations that are possible with the DS8700. The left frame is a base frame that contains the processors. In this example, it is two 4-way IBM System p POWER6 servers, as only the 4-way systems can have expansion frames. The center frame is the first expansion frame that contains additional I/O enclosures but no additional processors. The right frame is the second expansion frame that contains just disks and no processors, I/O enclosures, or batteries (without extended power line disturbance feature (ePLD)). Each frame contains a power area with power supplies and other power-related hardware. A DS8700 can consist of up to five frames. The third expansion frame is identical to the second expansion frame. The fourth expansion frame is also identical to the second expansion frame but can only be half full of disk enclosures,

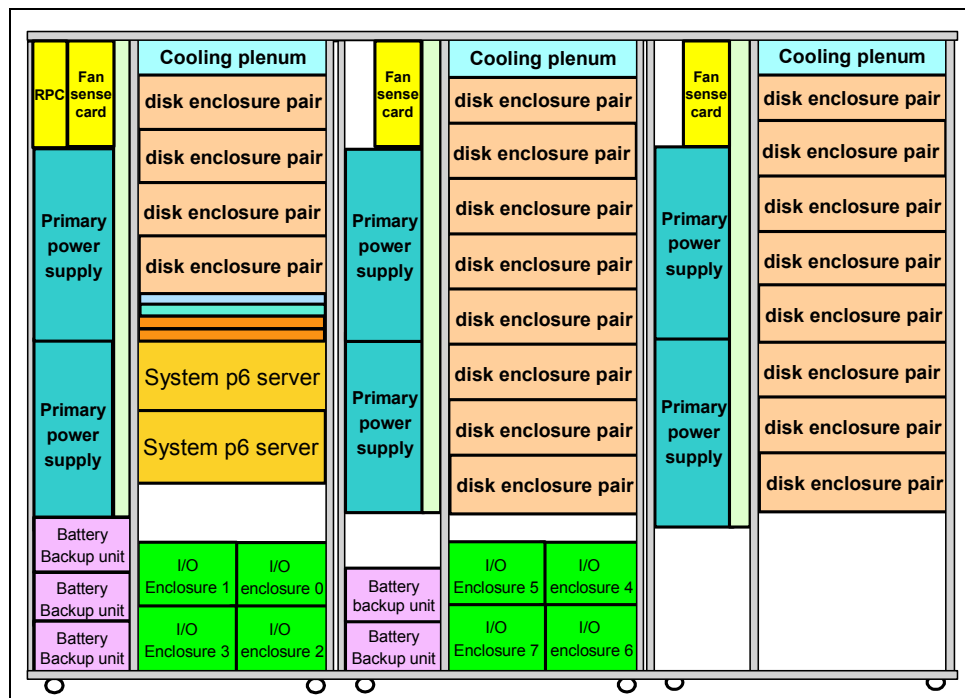


Figure 3-1 DS8700 frame types (front view)

3.1.1 Base frame: DS8700

The left side of the base frame, viewed from the front of the machine, is the frame power area. Only the base frame contains rack power control cards (RPC) to control power sequencing for the storage unit. It also contains two fan sense cards that provide rack ID/Pack ID location information and fan monitoring path for that frame. The base frame contains two primary power supplies (PPSs) to convert input AC into DC power. The power area also contains two (2-way system) or three (4-way system) battery backup units (BBUs).

The base frame can contain up to eight disk enclosures (installed in pairs), each of which can contain up to 16 disk drives. In a maximum configuration, the base frame can hold 128 disk

drives. Disk drives are either hard disk drives (HDD) with real spinning disks or Solid State Disk drives (SSD).

A disk enclosure pair can contain either HDDs or SSDs. Intermixing HDDs and SSDs in the same disk enclosure pair is not supported. HDDs are installed in groups of 16. SSDs can be installed in groups of 16 (full disk set) or 8 (half disk set). These groups are installed evenly across the disk enclosure pair. For more information about the disk subsystem, see 3.6, “Disk subsystem” on page 58.

Above the disk enclosures are cooling fans located in a cooling plenum. Between the disk enclosures and the processor complexes are two Ethernet switches and a Storage Hardware Management Console (HMC).

The base frame contains two processor complexes (CECs). These System p POWER6 servers contain the processor and memory that drive all functions within the DS8700.

Finally, the base frame contains two or four I/O enclosures. These I/O enclosures provide connectivity between the adapters and the processors. The adapters contained in the I/O enclosures can be either device adapters (DAs), host adapters (HAs), or both.

The communication path used for adapter to processor complex communication in the DS8700 consists of four lane (x4) PCI Express Generation 2 connections, providing a bandwidth of 2 GBps for each connection.

A RIO-G loop is still employed for the inter-processor complex communication, as in previous models of the DS8000 family. However, this RIO-G loop is no longer used to communicate to the I/O enclosures.

3.1.2 Expansion frames

In the DS8700, only a 4-way system can have expansion frames. There are two types of expansion frames. The first expansion frame always contains I/O enclosures. The second, third, and fourth expansion frames have no I/O enclosures.

The left side of each expansion frame, viewed from the front of the machine, is the frame power area. The expansion frames do not contain rack power control cards: These cards are only present in the base frame. The expansion frames also contain two fan sense cards that provide rack ID/Pack ID location information and fan monitoring path for that frame. Each expansion frame contains two primary power supplies (PPSs) to convert the AC input into DC power. Finally, the power area contains two battery backup units (BBUs) in the first expansion frame. The second through fourth expansion frames have no BBUs unless the extended Power Line Disturbance (ePLD) ePLD feature is installed. In that case, each of those frames will have two BBUs.

If the optional ePLD feature is installed, battery booster modules must be installed on the base frame and all expansion frames in the System. The ePLD feature is useful as an additional safeguard against environmental power fluctuations.

Expansion frames one through three can each hold up to 16 disk enclosures (installed in pairs), which contain the disk drives. In a maximum configuration, these expansion frames can hold 256 disk drives. Expansion frame four can only contain up to four disk enclosure pairs for a maximum total of 128 disk drives in this frame. A fully populated DS8700 five-frame system will contain a total of 1024 disk drives. Disk drives can be either HDDs or SSDs. SSDs are not supported in 3rd or 4th expansion frame.

3.2 Frames: DS8800

The DS8800 is designed for modular expansion. From a high-level view, there appear to be three types of frames available for the DS8800. However, on closer inspection, the frames themselves are almost identical. The only variations are the combinations of processors, I/O enclosures, storage enclosures (and cabling), batteries, and disks that the frames contain.

Figure 3-2 shows frame variations that are possible with the DS8800. The left frame is a base frame that contains the processors. In this example, it has two 4-way IBM System p POWER6+ servers. Only 4-way systems can have expansion frames. The center frame is an expansion frame that contains additional I/O enclosures but no additional processors. The right frame is an expansion frame that contains simply disks and no processors, I/O enclosures, or batteries (without ePLD). Each frame contains a power area with power supplies and other power-related hardware. A DS8800 can consist of up to three frames.

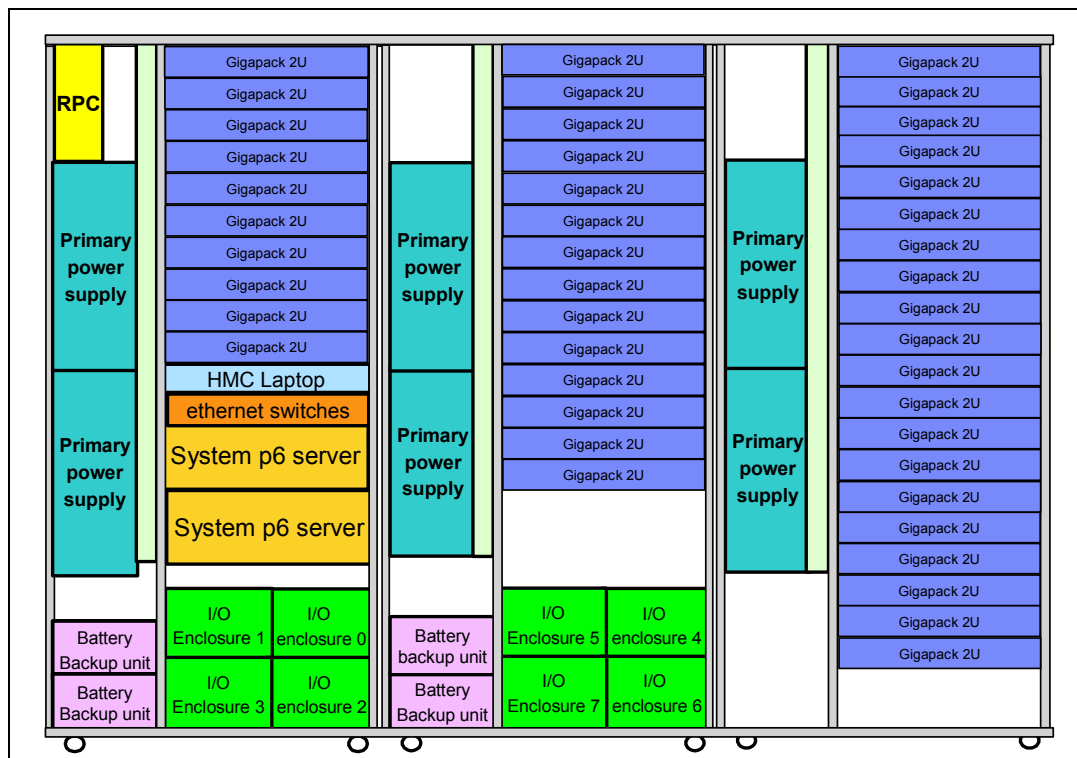


Figure 3-2 DS8800 frame types (front view)

3.2.1 Base frame: DS8800

The left side of the base frame, viewed from the front of the machine, is the frame power area. Only the base frame contains rack power control cards (RPCs) to control power sequencing for the storage unit. The base frame contains two primary power supplies (PPSs) to convert input AC into DC power. The power area also contains two battery backup units (BBUs). This is true whether it is a 2 or 4-way system, and if the system has optional ePLD (extended power line disturbance) feature or not.

The base frame can contain up to ten disk enclosures (installed in pairs), each of which can contain up to 24 disk drives. In a maximum configuration, the base frame can hold 240 disk drives. Disk drives are either hard disk drives (HDD) with real spinning disks or Solid State Disk drives (SSD).

A disk enclosure pair can contain either HDDs or SSDs. Intermixing HDDs and SSDs in the same disk enclosure pair is not supported. HDDs are installed in groups of 16. SSDs can be installed in groups of 16 (full disk set) or 8 (half disk set). These groups are installed evenly across the disk enclosure pair.

Inside the disk enclosures are cooling fans located in the storage enclosure power supply units. These fans pull cool air from the front of the frame and exhaust to the rear of the frame.

Between the disk enclosures and the processor complexes are two Ethernet switches and a Storage Hardware Management Console (HMC).

The base frame contains two processor complexes (CECs). These System p POWER6+ servers contain the processor and memory that drive all functions within the DS8800.

The base frame also contains I/O enclosures (installed in pairs). Two I/O enclosures for 2-way systems (including business cabled systems), and four I/O enclosures in 4-way systems. These I/O enclosures provide connectivity between the adapters and the processors. Each I/O enclosure can contain up to two device adapters and two host adapters.

The communication path used for adapter-to-processor complex communication in the DS8800 consists of PCI Express Generation 2 connections, providing a bandwidth of 2 GBps for each connection.

The interprocessor complex communication still utilizes the RIO-G loop as in previous models of the DS8000 family. However, this RIO-G loop no longer has to handle data traffic, which greatly improves performance.

The base frame can be configured using either standard or business class cabling. Standard cabling is optimized for performance and allows for highly scalable configurations with large long-term growth. The business class option allows a system to be configured with more drives per device adapter, thereby reducing configuration cost and increasing adapter utilization. This configuration option is intended for configurations where capacity and high resource utilization is of the most importance. Scalability is limited in the business class option.

Standard cabling supports either 2-way processors with one I/O enclosure pair or 4-way processors with two I/O enclosure pairs. Standard cabling with one I/O enclosure pair supports up to two DA pairs and six storage enclosures (144 DDMs). Standard cabling with two I/O enclosure pairs supports up to four DA pairs and ten storage enclosures (240 DDMs).

Business class cabling utilizes two-way processors and one I/O enclosure pair. Business class cabling supports two DA pairs and up to ten storage enclosures (240 DDMs).

Notes:

- ▶ Business class cabling is available as an initial order only. A business class cabling configuration can only be ordered as a base frame with no expansion frames.
- ▶ DS8800 does not support model conversion, that is, business class and standard class cabling conversions are not supported. Re-cabling is available as an RPQ only and is disruptive.

3.2.2 Expansion frames

In the DS8800, only a 4-way system can have expansion frames. The expansion frame model is called 95E. There are two types of expansion frames. The first expansion frame will always contain four I/O enclosures (two pair). The second expansion frame will have no I/O enclosures. The I/O enclosures provide connectivity between the adapters and the processors. The adapters contained in the I/O enclosures can be either device adapters or host adapters, or both. You cannot use expansion frames from previous DS8000 models as expansion frames for a DS8800 storage system.

The left side of each expansion frame, viewed from the front of the machine, is the frame power area. The expansion frames do not contain rack power control (RPC) cards. RPC cards are only present in the base frame. Each expansion frame contains two primary power supplies (PPSs) to convert the AC input into DC power.

The power area can contain zero or two battery backup units (BBUs), depending on the configuration. The first expansion rack requires two BBUs (with or without ePLD). The second expansion rack requires two BBUs (with ePLD) and no BBUs (without ePLD).

If the optional ePLD feature is installed, battery booster modules must be installed on the base frame and all expansion frames in the system. The ePLD feature is useful as an additional safeguard against environmental power fluctuations.

The first expansion frame can contain up to 14 storage enclosures (installed in pairs). A storage enclosure can have up to 24 small form factor (SFF) disks installed. In a maximum configuration, the first expansion frame can contain up to 336 disk drives.

The second expansion frame can contain up to 20 storage enclosures (installed in pairs), which contain the disk drives. The second expansion frame can contain up to 480 disk drives.

A fully configured 3 frame DS800 system can contain a maximum of 1056 disks.

Disk drives are either hard disk drives (HDD) with real spinning disks or Solid State Disk drives (SSD).

For more information about SSDs see 8.5.3, “DS8000 Solid State Drive (SSD) considerations” on page 222

3.2.3 Rack operator panel

Each DS8700 or DS8800 frame features status indicators. The status indicators can be seen when the doors are closed. When the doors are open, the emergency power off switch (an EPO switch) is also accessible. Figure 3-3 on page 41 shows the operator panel for DS8700 and the EPO switch.

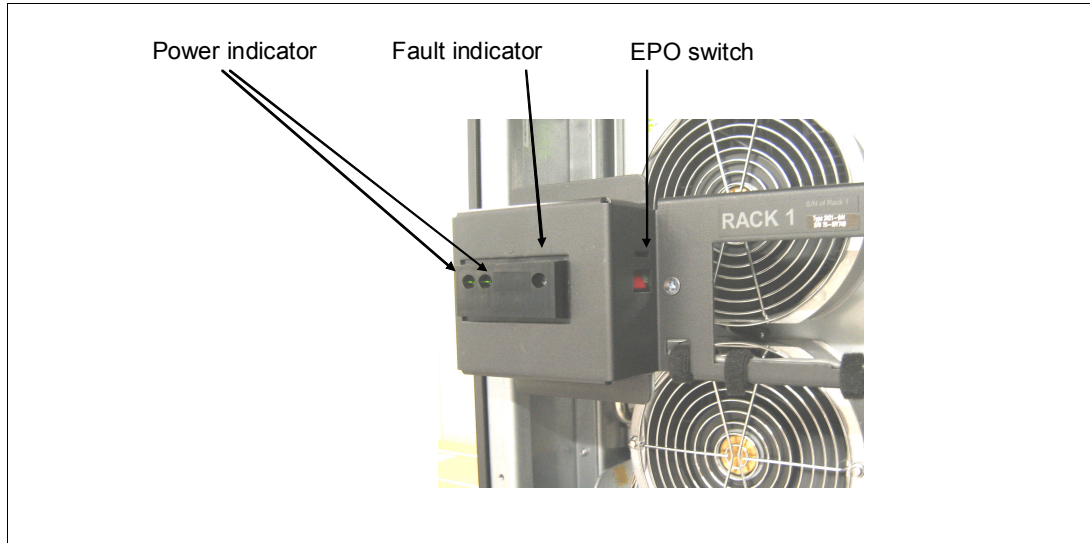


Figure 3-3 Rack operator window - DS8700

Figure 3-4 shows the operator panel for DS8800.

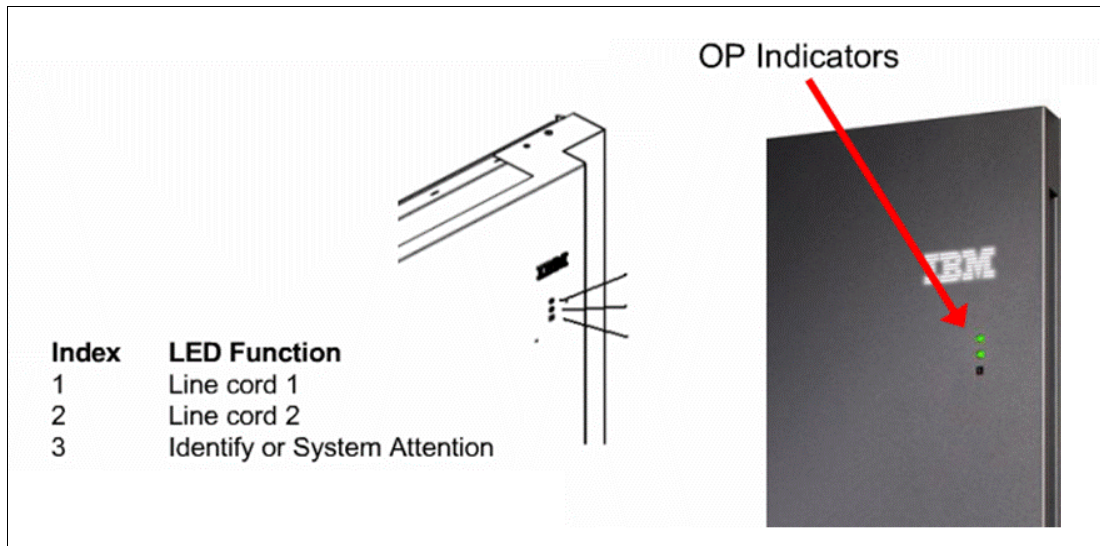


Figure 3-4 Rack operator window - DS8800

Each panel has two line cord indicators, one for each line cord. For normal operation, both of these indicators are illuminated green, if each line cord is supplying correct power to the frame. There is also a fault indicator. If this indicator is lit solid amber, use the DS Storage Manager GUI or the HMC **Manage Serviceable Events** menu to determine why the indicator is illuminated.

There is also an EPO switch to the right side of the operator panel on DS8700, or near the top of the PPS (Primary Power Supplies) on DS8800. This switch is only for emergencies. Tripping the EPO switch will bypass all power sequencing control and result in immediate removal of system power. Data in non-volatile storage (NVS) will *not* be destaged and will be lost. Do not trip this switch unless the DS8000 is creating a safety hazard or is placing human life at risk. Figure 3-5 on page 42 shows the EPO switch in the DS8800.

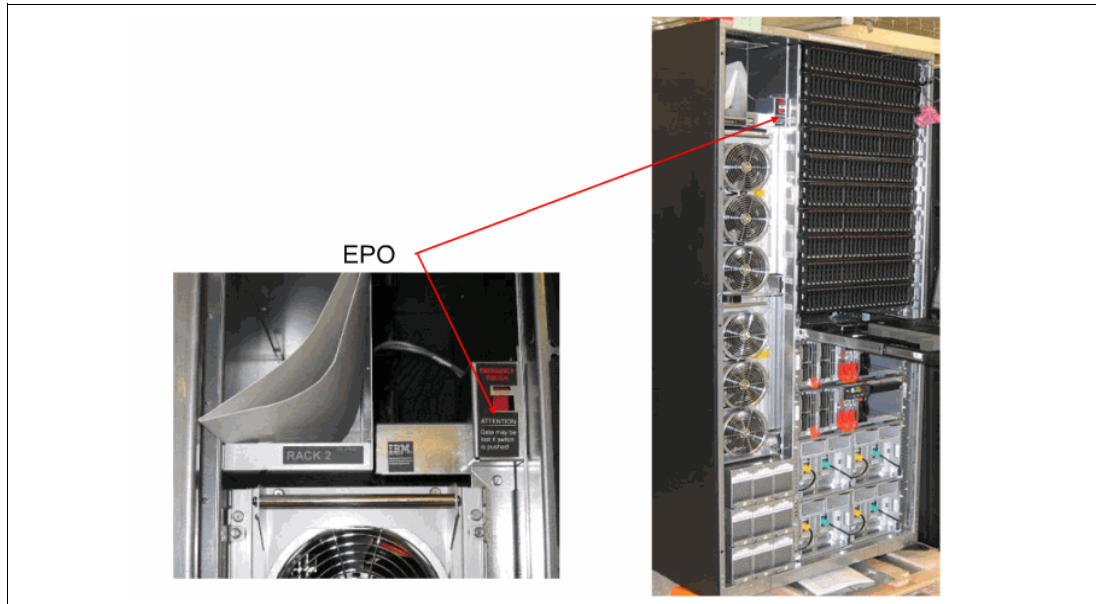


Figure 3-5 Emergency power off (EPO) switch - DS8800

There is no power on/off switch on the operator window because power sequencing is managed through the HMC. This ensures that all data in nonvolatile storage, known as *modified data*, is destaged properly to disk prior to power down. It is not possible to shut down or power off the DS8000 from the operator window, except in an emergency and using the EPO switch.

3.3 DS8000 architecture

Now that the frames have been described, the rest of this chapter explores the technical details of each component.

The DS8000 consists of two processor complexes (CECs). Each processor complex has access to multiple host adapters to connect to FC (Fibre Channel) or FICON hosts. DS8700 and DS8800 can have up to 128 FC or FICON host connections.

The installed storage is connected to the processors through internal switched FC fabrics.

3.3.1 POWER6 and POWER6+ processor

Both the DS8700 and the DS8800 use the POWER6 p570 based server technology. The 64-bit POWER6 processors in the p570 server are integrated into a dual-core single chip module or a dual-core dual chip module, with 32 MB of L3 cache, 8 MB of L2 cache, and 12 DDR2 memory DIMM slots. This enables operating at a high data rate for large memory configurations. Each new processor card can support up to 12 DDR2 DIMMs running at speeds of up to 667 MHz.

The Symmetric Multi-Processing (SMP) system features 2-way or 4-way, copper-based, Silicon-on Insulator-based (SOI-based) POWER6 microprocessors running at 4.7 GHz (DS8700) and POWER6+ microprocessors running at 5.0 GHz (DS8800).

Each POWER6 processor provides a GX+ bus that is used to connect to an I/O subsystem or fabric interface card. GX+ is a Host Channel Adapter used in POWER6 systems. For more information, see *IBM System p 570 Technical Overview and Introduction*, REDP-4405.

Also, see Chapter 4, “RAS on IBM System Storage DS8000” on page 71 and 7.2.4, “IBM System p POWER6: Heart of the DS8000 dual cluster design” on page 166 for additional information about the POWER6 processor.

3.3.2 Server-based SMP design

The DS8000 series, which includes the DS8700 and DS8800, benefits from a fully assembled, leading edge processor and memory system. The DS8000 systems use DDR2 memory DIMMs. Using SMPs as the primary processing engine sets the DS8000 systems apart from other disk storage systems on the market.

Additionally, the System p POWER6 and POWER6+ processors used in the DS8000 support the execution of two independent threads concurrently. This capability is referred to as *simultaneous multi-threading* (SMT). The two threads running on the single processor share a common L1 cache. The SMP/SMT design minimizes the likelihood of idle or overworked processors, whereas a distributed processor design is more susceptible to an unbalanced relationship of tasks to processors.

The design decision to use SMP memory as an I/O cache is a key element of the IBM storage architecture. Although a separate I/O cache could provide fast access, it cannot match the access speed of the SMP main memory.

All memory installed on any processor complex is accessible to all processors in that complex. The addresses assigned to the memory are common across all processors in the same complex. Alternatively, using the main memory of the SMP as the cache leads to a partitioned cache. Each processor has access to the processor complex's main memory, but not to that of the other complex. You should keep this in mind with respect to load balancing between processor complexes.

3.3.3 Peripheral Component Interconnect Express (PCI Express)

The DS8700 and DS8800 processor complex utilizes a PCI Express infrastructure to access the I/O subsystem, which provides a great improvement in performance over previous DS8000 models.

PCI Express was designed to replace the general-purpose PCI expansion bus, the high-end PCI-X bus, and the Accelerated Graphics Port (AGP) graphics card interface.

PCI Express is a serial I/O interconnect. Transfers are bidirectional, which means data can flow to and from a device simultaneously. The PCI Express infrastructure involves a switch so that more than one device can transfer data at the same time.

Unlike previous PCI-X interfaces, rather than being a bus, it is structured around point-to-point full duplex serial links called *lanes*. Lanes can be grouped by 1x, 4x, 8x, 16x, or 32x, and each lane is high speed, using an 8b/10b encoding that results in 2.5 Gbps = 250 MBps per lane in a generation 1 implementation. Bytes are distributed across the lanes to provide a high throughput (Figure 3-6 on page 44).

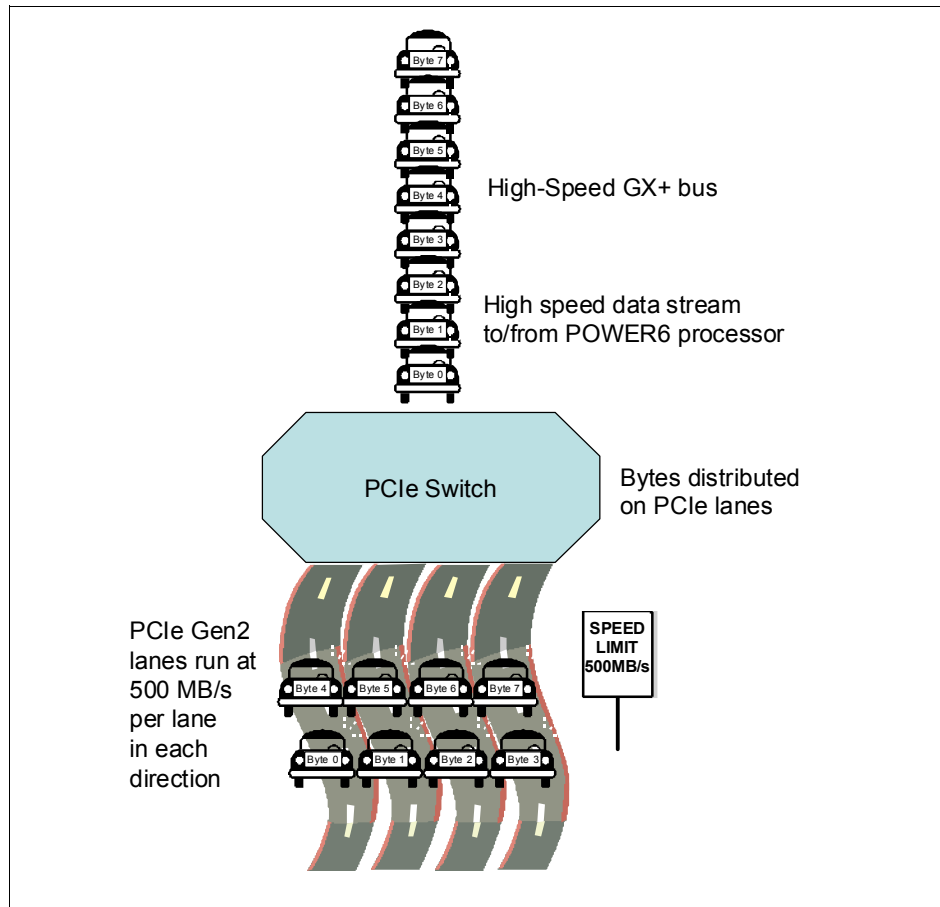


Figure 3-6 PCI Express architecture

There are two generations of PCI Express in use today:

- ▶ PCI Express 1.1 (Gen 1) = 250 MBps per lane (used in the DS8700 P6 and DS8800 P6+ processors)
- ▶ PCI Express 2.0 (Gen 2) = 500 MBps per lane (used in the DS8700 and DS8800 I/O enclosures)

You can learn more about PCI Express at the following site:

<http://www.redbooks.ibm.com/Redbooks.nsf/RedbookAbstracts/tips0456.html?Open>

The DS8000 processor complex (CEC) is equipped with two kinds of PCIe cards (Figure 3-7).

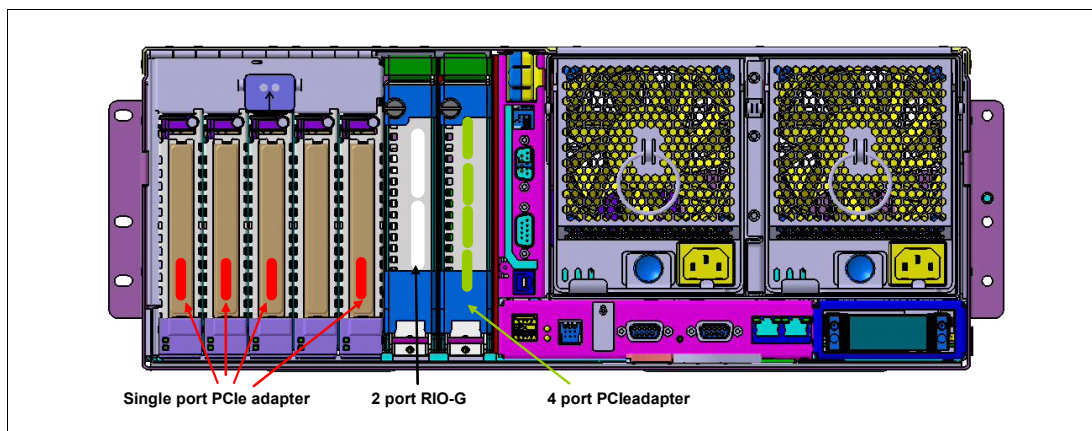


Figure 3-7 Processor complex - rear view

- Four half-high single port PCI Express cards (in slots 1, 2, 3, and 5). Each card converts a single PCIe x8 Gen1 bus into a PCIe x4 Gen2 external cable connection.
A bridge is used to translate the x8 Gen 1 lanes from the processor to x4 Gen 2 lanes used by the I/O enclosures.
- A four port PCI2 adapter that plugs into the CEC GX+ bus and has an onboard (P5IOC2) chip that supplies four PCIe x8 Gen1 busses that are converted into four PCIe x4 Gen2 external cable connections.

As shown in Figure 3-8, a bridge is used to translate the x8 Gen 1 lanes from the processor to the x4 Gen 2 lanes used by the I/O enclosures.

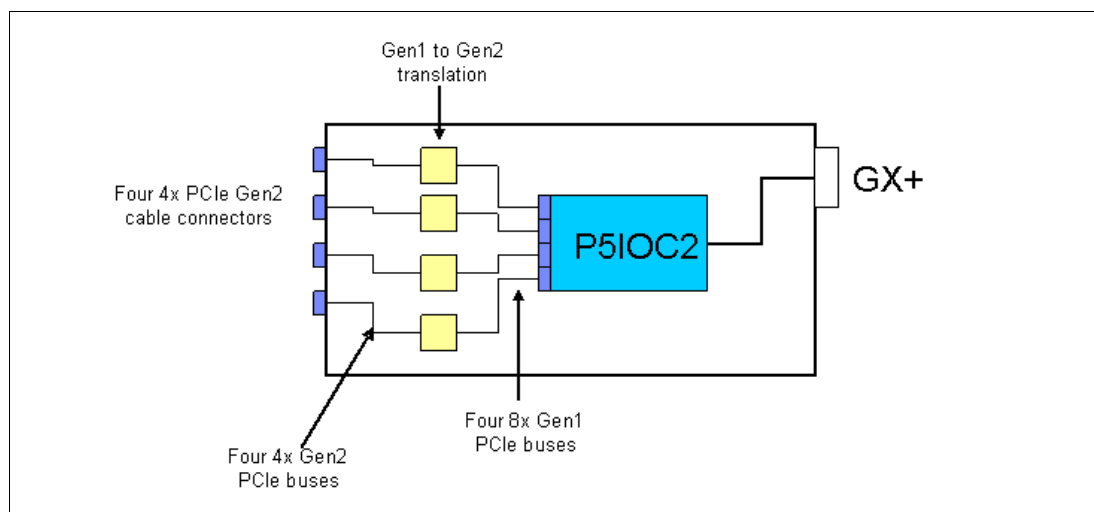


Figure 3-8 GX+ to PCI Express adapter

For more information, see 3.5, “I/O enclosures” on page 50.

3.3.4 Storage facility architecture

As already mentioned, the DS8700 and DS8800 storage facility consists of two POWER6 p570 servers. They form a processor complex that utilizes a RIO-G loop for processor

communication and a PCI Express infrastructure to communicate to the I/O subsystem (Figure 3-9).

When a host performs a read operation, the processor complexes (also called CECs) fetch the data from the disk arrays using the high-performance switched Fibre Channel architecture. The data is then cached in volatile memory in case it is required again. The servers attempt to anticipate future reads by an algorithm known as Sequential prefetching in Adaptive Replacement Cache (SARC). Data is held in cache as long as possible using this smart caching algorithm. If a cache hit occurs where requested data is already in cache, then the host does not have to wait for it to be fetched from the disks. The cache management has been enhanced by breakthrough caching technologies from IBM Research, such as the Adaptive Multi-stream Prefetching (AMP) and Intelligent Write Caching (IWC). See 7.5, “DS8000 superior caching algorithms” on page 172.

For DS8800, both the device and host adapters operate on high bandwidth fault-tolerant point-to-point 4 lane Generation 2 PCI Express interconnections. DS8800 device adapters feature an 8 Gb Fibre Channel interconnect speed with a 6 Gb SAS connection to the disk drives for each connection and direction. On a DS8800, as on a DS8700, the data traffic is isolated from the processor complex communication that utilizes the RIO-G loop

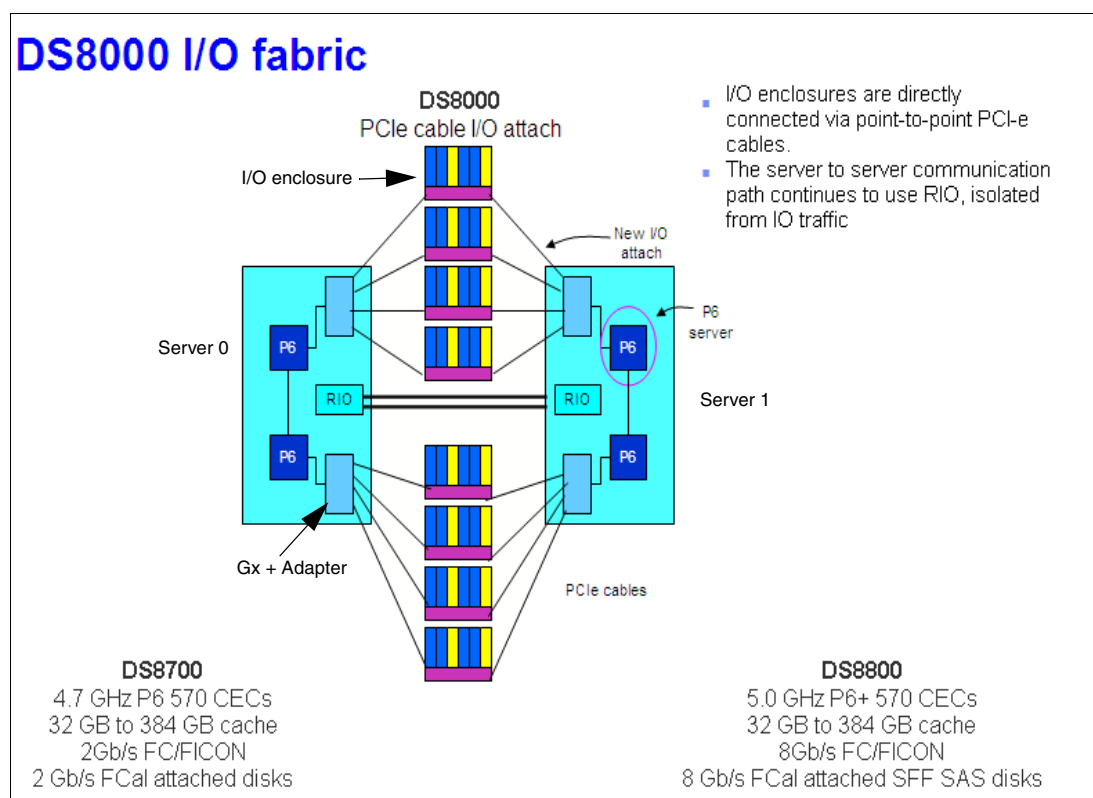


Figure 3-9 DS8000 series architecture

Figure 3-9 shows how the DS8800 hardware is shared between the servers. On the left side is one processor complex (CEC). The CEC uses the N-way symmetric multiprocessor (SMP) of the complex to perform its operations. It records its write data and caches its read data in the volatile memory of the left complex. For fast-write data, it has a persistent memory area on the right processor complex. To access the disk arrays under its management, it has its own affiliated device adapters. The server on the right operates in an identical fashion. The host adapters are shared between both servers.

3.4 Storage facility processor complex (CEC)

The DS8700 and DS8800 base frames contain two processor complexes. The 941 and 951 models can have the 2-way processor feature or the 4-way processor feature (*2-way* means that each processor complex has two CPUs, and *4-way* means that each processor complex has four CPUs).

Figure 3-10 shows a rear view of a DS8700 or DS8800 processor complex.

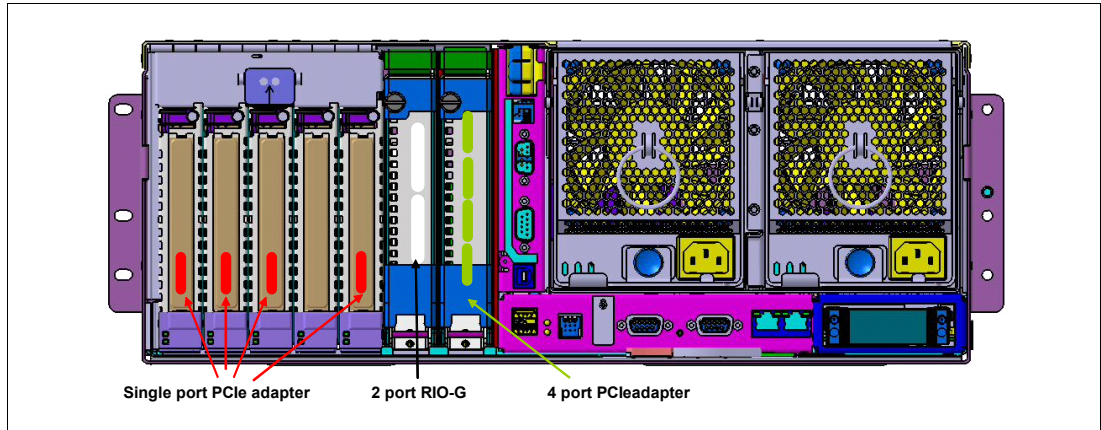


Figure 3-10 Processor complex - rear view

Figure 3-11 shows the DS8700 storage subsystem with the 2-way processor feature. There can be two or four I/O enclosures.

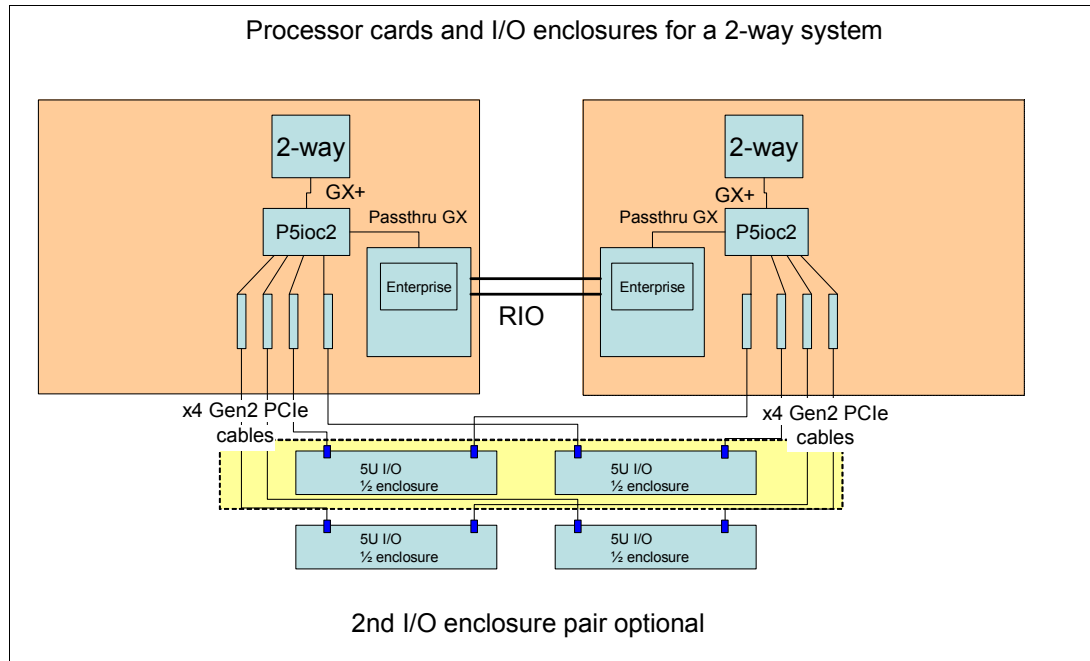


Figure 3-11 DS8700 2-way architecture

Figure 3-12 shows the DS8700 with the 4-way feature. Two I/O Enclosure pairs will be installed in the base frame. Two more I/O enclosure pairs will be in the first expansion frame (if installed).

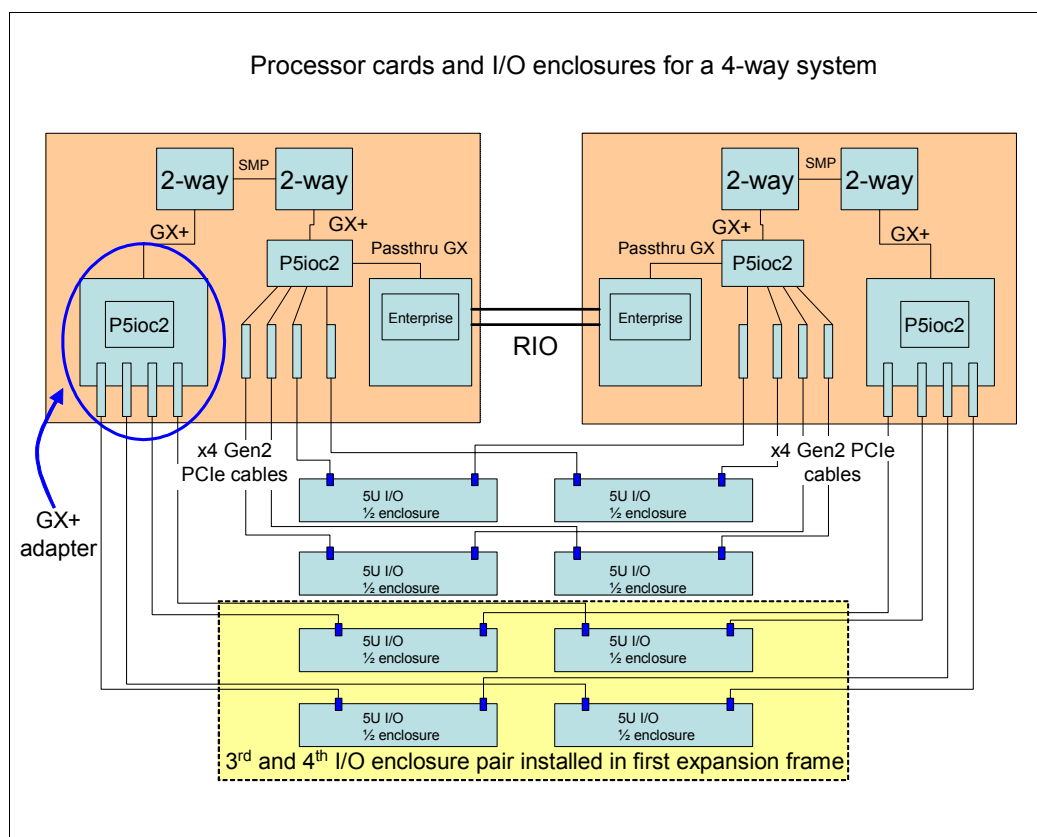


Figure 3-12 DS8700 4-way architecture

Figure 3-13 show the DS8800 storage subsystem with the 2-way processor feature. Only 1 I/O enclosure pair is supported.

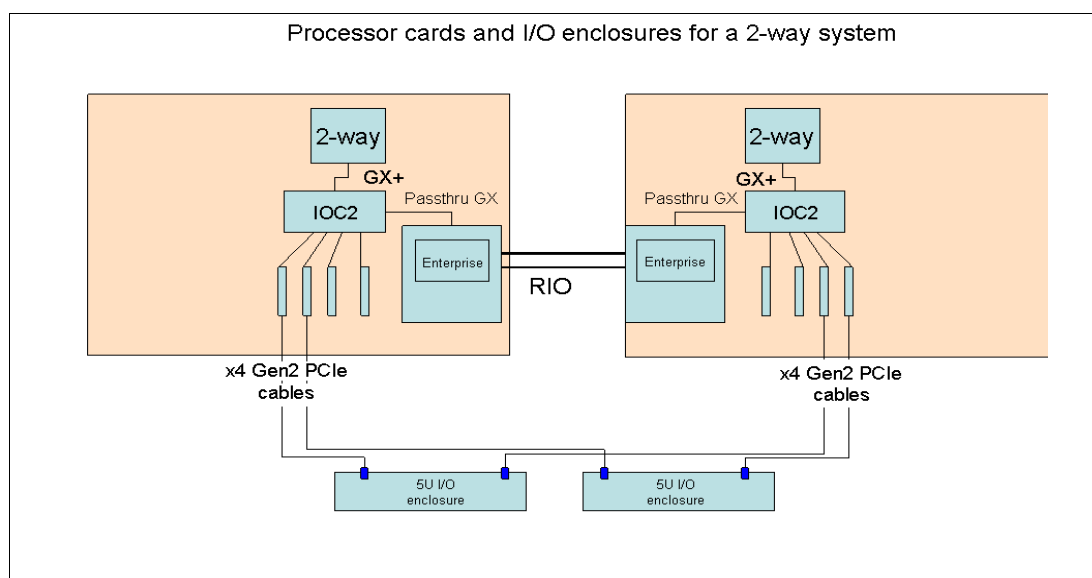


Figure 3-13 DS8800 2-way architecture

Figure 3-14 shows the DS8800 with the 4-way feature. Two I/O Enclosure pairs will be installed in the base frame. Two more I/O enclosure pairs will be in the first expansion frame (if installed).

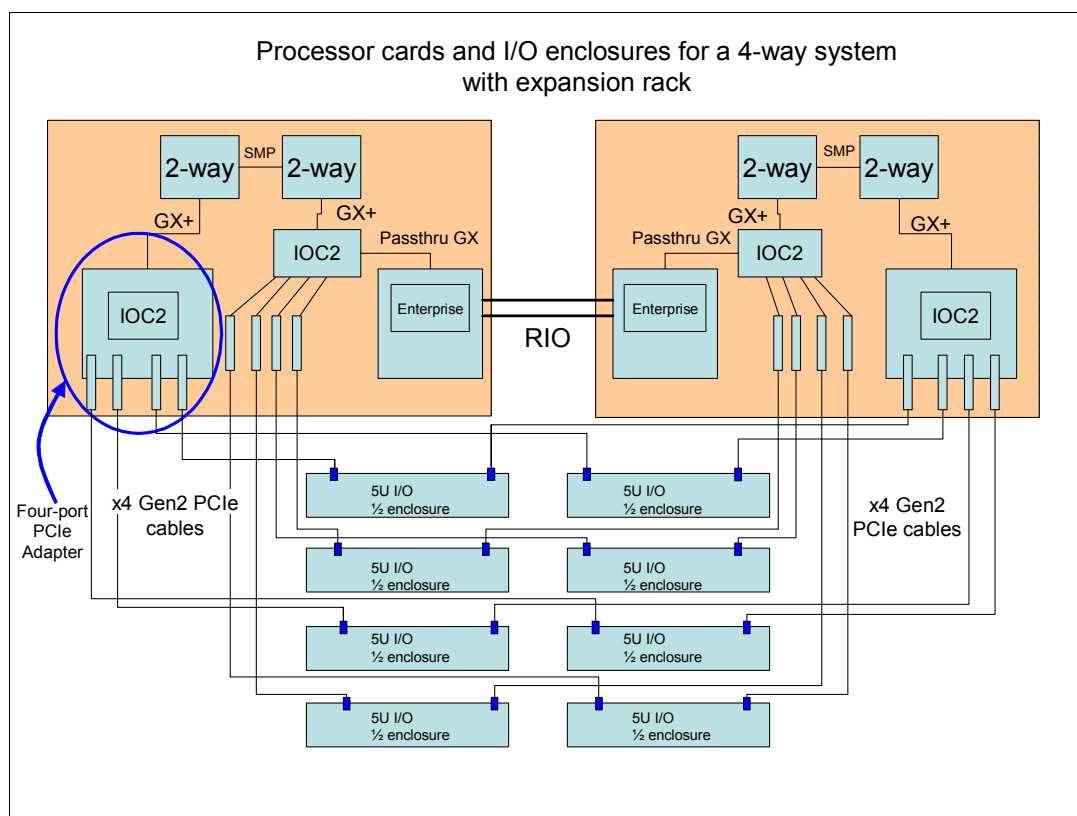


Figure 3-14 DS8800 with expansion frame and four I/O enclosure pairs

The DS8700 and DS8800 features IBM POWER6 server technology. Compared to the POWER5+ based processor models in the earlier DS8100 and DS8300, the POWER6 processor can achieve up to a 50% performance improvement in I/O operations per second in transaction processing workload environments and up to 150% throughput improvement for sequential workloads.

For details about the server hardware used in the DS8000, refer to *IBM System p 570 Technical Overview and Introduction*, REDP-4405, found at the following URL:

<http://www.redbooks.ibm.com/redpieces/pdfs/redp4405.pdf>

3.4.1 Processor memory and cache management

The DS8700 and DS8800 offer up to 384 GB of total processor memory. Each processor complex will have half of the total system memory. Caching is a fundamental technique for reducing I/O latency. Like other modern caches, DS8700 and DS8800 contain volatile memory used as a read and write cache and non-volatile memory used as a write cache. The non-volatile storage (NVS) scales to the processor memory size selected, which also helps to optimize performance.

The effectiveness of a read cache depends on the hit ratio, which is the fraction of requests that are served from the cache without necessitating a read from the disk (read miss).

To help achieve dramatically greater throughput and faster response times, the DS8000 uses Sequential-prefetching in Adaptive Replacement Cache (SARC). SARC is an efficient adaptive algorithm for managing read caches with both:

- ▶ Demand-paged data: It finds recently used data in the cache.
- ▶ Prefetched data: It copies data speculatively into the cache before it is even requested.

The decision of when and what to prefetch is made in accordance with the Adaptive Multi-stream Prefetching (AMP), a cache management algorithm.

The Intelligent Write Caching (IWC) manages the write cache and decides in what order and at what rate to destage.

For details about cache management, see 7.5, “DS8000 superior caching algorithms” on page 172.

3.4.2 Flexible service processor and system power control network

The flexible service processor (FSP) is an embedded controller that is based on a PowerPC® processor. The system power control network (SPCN) is used to control the power of the attached I/O subsystem. The SPCN control software and the FSP software are run on the same PowerPC processor.

The FSP performs predictive failure analysis based on any recoverable processor errors. The FSP can monitor the operation of the firmware during the boot process, and it can monitor the operating system for loss of control. This enables the FSP to take appropriate action.

The SPCN monitors environmentals such as power, fans, and temperature. Environmental critical and noncritical conditions can generate Early Power-Off Warning (EPOW) events. Critical events trigger appropriate signals from the hardware to the affected components to prevent any data loss without operating system or firmware involvement. Non-critical environmental events are also logged and reported.

3.4.3 RIO-G

In the DS8700 and DS8800, the RIO-G (remote I/O) loop is used for inter-processor communication only. The RIO-G has evolved from earlier versions of the RIO interconnect.

Each RIO-G port can operate at 1 GHz in bidirectional mode, and is capable of passing data in each direction on each cycle of the port. It is designed as a high performance, self-healing interconnect.

3.5 I/O enclosures

The DS8700 and DS8800 base frame, and expansion frame (if installed) both contain I/O enclosures. I/O enclosures are installed in pairs. There can be one or two I/O enclosure pairs installed in the base frame (depending on configuration - 2-way or 4-way). Two I/O enclosures are installed in the first expansion frame. Each I/O enclosure has six slots. Device adapters (DA) and host adapters (HA) are installed in the I/O enclosures. The I/O enclosures provide connectivity between the processor complexes and the HAs or DAs. The DS8700 can have up to 2 DAs and 4 HAs installed in an I/O enclosure. The DS8800 can have up to 2 DAs and 2 HAs installed in an I/O enclosure.

Each CEC has an onboard GX+ Bus to a P5IOC2 adapter. This P5IOC2 adapter drives four single port PCIe adapters which connect to four I/O enclosures. There is also a second GX+

bus, which is driven by the second CPU module, if installed (4-way feature). The second GX+ bus drives a 4 port PCIe adapter which connects to the other four I/O enclosures.

3.5.1 DS8700 I/O enclosures

Figure 3-15 shows the DS8700 CEC to I/O enclosure connectivity (4-way with first expansion frame). All I/O enclosures in the base frame will communicate using the first GX+ bus, through the 4 single port PCIe adapters. All I/O enclosures in the first expansion frame will communicate using the second GX+ Bus, through the 4 port PCIe adapter.

A 2-way configuration can have either one or two I/O enclosure pairs installed in the base frame.

A 4-way configuration will have two I/O enclosure pairs installed in the base frame and two I/O enclosures pairs in the first expansion frame (if installed). A 4-way configuration is required to support expansion frames.

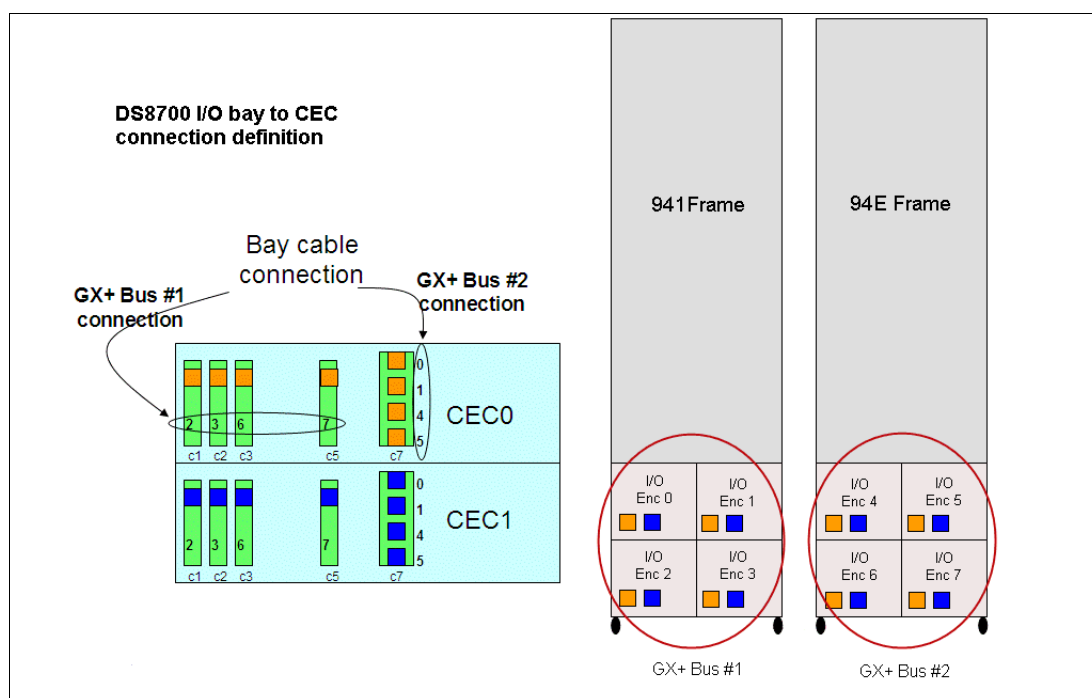


Figure 3-15 DS8700 I/O enclosure connections to CEC

3.5.2 DS8800 I/O enclosures

Figure 3-16 shows the DS8800 CEC to I/O enclosure connectivity (4-way with first expansion frame). The lower I/O enclosure pairs in the base and first expansion frames will communicate using the first GX+ bus, through the 4 single port PCIe adapters. The upper I/O enclosure pairs in the base and first expansion frame will communicate using the second GX+ bus through the 4 port PCIe adapter.

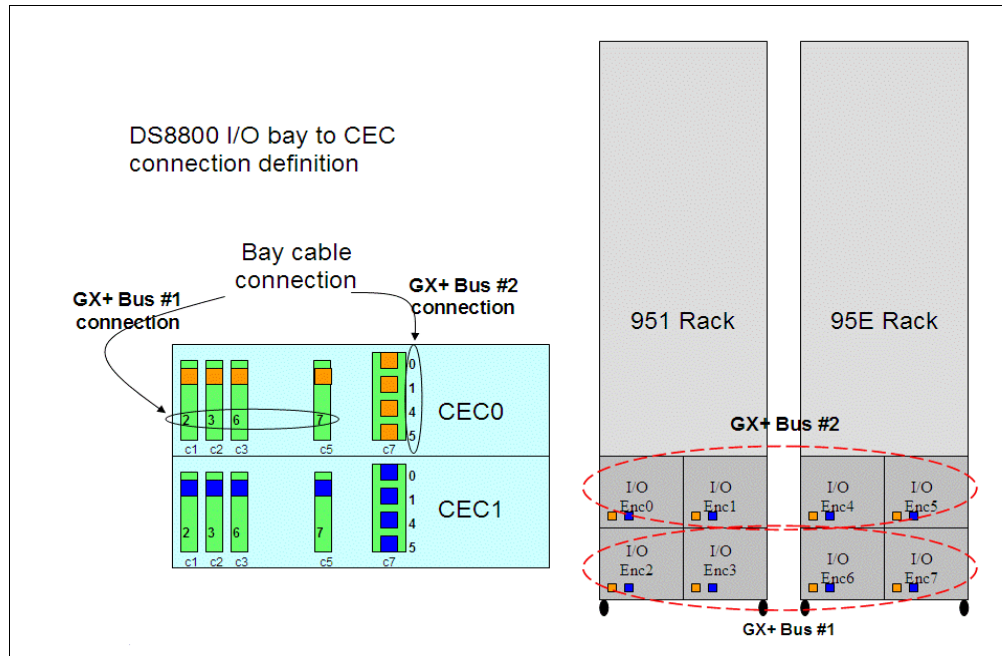


Figure 3-16 DS8800 I/O enclosure connections to CEC

A 2-way configuration would only use the lower I/O enclosure pair in the base frame.

A 4-way configuration will have two I/O enclosure pairs installed in the base frame and two I/O enclosures pairs in the first expansion frame (if installed). A 4-way configuration is required to support expansion frames.

Each I/O enclosure has the following attributes:

- ▶ 5U rack-mountable enclosure
- ▶ Six PCI Express slots
- ▶ Default redundant hot plug power and cooling devices

3.5.3 Host adapters

Attached host servers interact with software running on the complexes to access data on logical volumes. The servers manage all read and write requests to the logical volumes on the disk arrays. During write requests, the servers use fast-write, in which the data is written to volatile memory on one processor complex and preserved memory on the other processor complex. The server then reports the write as complete before it is written to disk. This provides much faster write performance than writing to the actual disk. Preserved memory is also called nonvolatile storage (NVS). Additional information about this topic is available.

DS8700 Host Adapters

The DS8700 supports up to four host adapters (HA) per I/O enclosure. Host Adapters on the DS8700 are available in either longwave or shortwave. Each port can be configured to operate as a Fibre Channel port or as a FICON port. Slots 1, 2, 4, and 5 support 4 Gbps 4 Port HAs. With Microcode release 6.1, DS8700 can also support up to two 8 Gbps 4 port HAs per I/O enclosure.

Tip: These 8 Gbps adapters are restricted to Slots 1 and 4 only

To add an 8 Gbps HA to an existing configuration, where slot 1 or 4 have an existing 4 Gbps adapter, remove the existing 4 Gbps adapter, and then reinstall it into an available slot location.

Figure 3-17 shows HA locations in the DS8700 I/O enclosure.

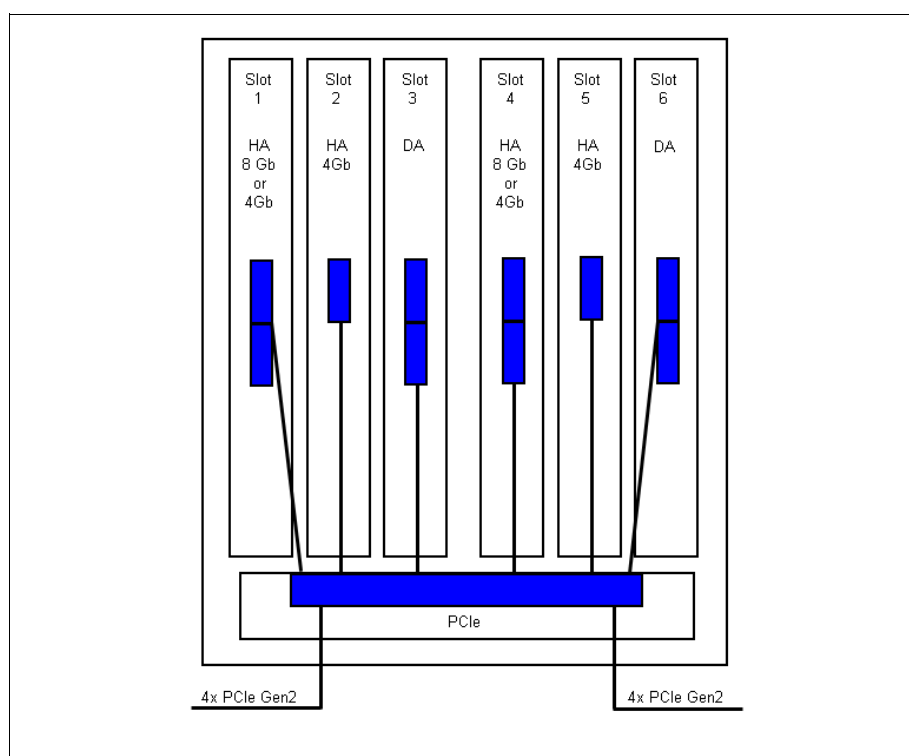


Figure 3-17 DS8700 I/O enclosure adapter layout

Figure 3-18 on page 54 illustrates the preferred HA plug order for DS8700. Host adapter positions and plugging order for the four I/O enclosures are the same for the base frame and the expansion frames with I/O enclosures. The chart shows the host adapter positions and plugging order for four I/O enclosures. The Install Sequence line indicates the order in which Host Adapter pairs are to be installed. To achieve optimum balance (and performance) they should be added in the following order.

Tip: If the DS8700 is a two I/O enclosures machine, use the plugging order for the XB3 and XB4 enclosure. Start card plugging with the left I/O enclosure, then the right enclosure following the plug order chart.

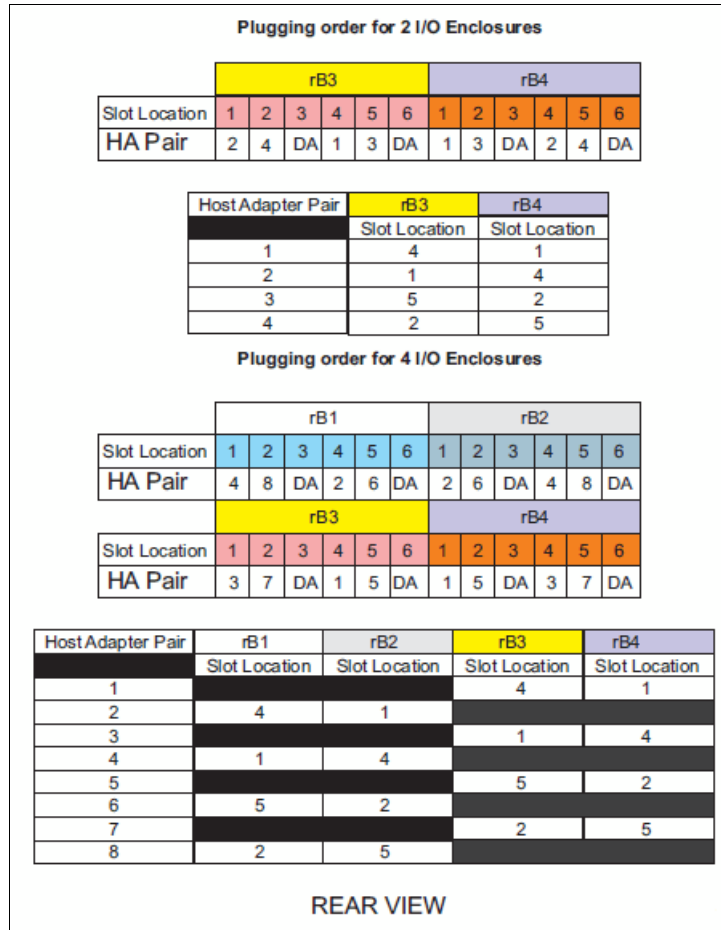


Figure 3-18 DS8700 HA plug order

Fibre Channel is a technology standard that allows data to be transferred from one node to another at high speeds and great distances (up to 10 km and beyond). The DS8700 uses the Fibre Channel protocol to transmit SCSI traffic inside Fibre Channel frames. It also uses Fibre Channel to transmit FICON traffic, which uses Fibre Channel frames to carry System z I/O.

Each DS8700 Fibre Channel adapter offers four 4 Gbps or 8 Gbps Fibre Channel ports. The cable connector required to attach to this adapter is an LC type. Each 4 Gbps port independently auto-negotiates to either 1, 2, or 4 Gbps link speed. Each 8 Gbps port independently auto-negotiates to either 2, 4, or 8 Gbps link speed. Each of the four ports on an DS8700 adapter can also independently be either Fibre Channel protocol (FCP) or FICON. The type of the port can be changed through the DS Storage Manager GUI or by using DSCLI commands. A port cannot be both FICON and FCP simultaneously, but it can be changed as required.

The card itself is PCI-X 64-bit 133 MHz. The card is driven by a new high function, that is, a high performance ASIC. To ensure maximum data integrity, it supports metadata creation and checking. Each Fibre Channel port supports a maximum of 509 host login IDs and 1,280 paths. This allows for the creation of large storage area networks (SANs)

DS8800 Host Adapters

The DS8800 supports up to two FC or FICON host adapters (HA) per I/O enclosure. Host Adapters on the DS8800 are available in either longwave or shortwave. Each port can be configured to operate as either a Fibre Channel port or a FICON port. Slots 1 and 4 support 8 Gbps 4 or 8 Port HAs. Slots 2 and 5 are reserved and cannot be used. Slots 3 and 6 support 8 Gbps 4 or 8 Port HAs.

Figure 3-19 shows HA locations in the DS8800 I/O enclosure.

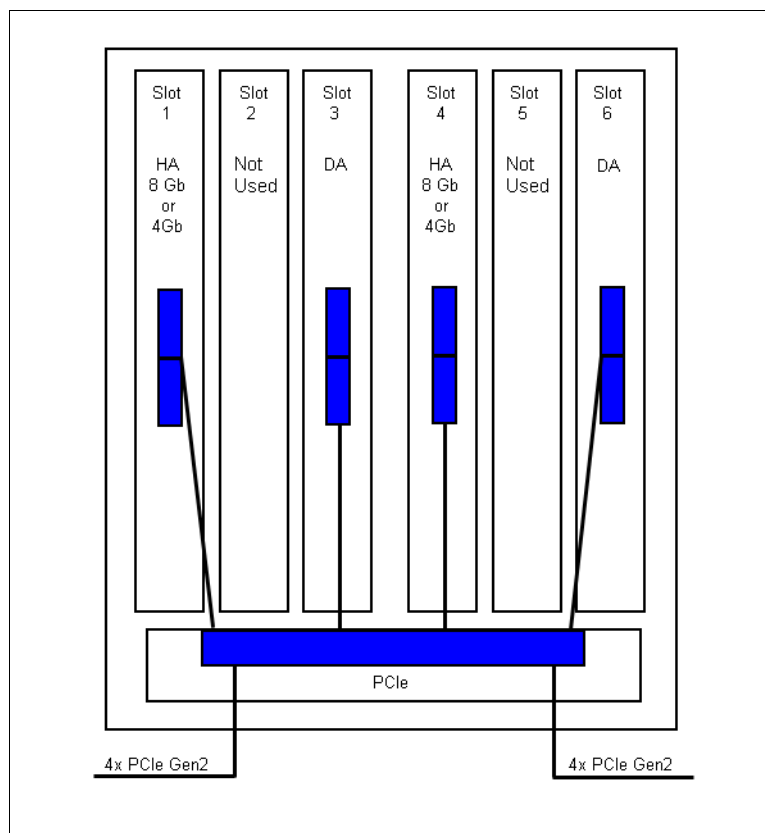


Figure 3-19 DS8800 I/O enclosure adapter layout

Tip: HA card slot locations 2 and 5 are not used in the DS8800.

Figure 3-20 on page 56 illustrates the preferred HA plug order for DS8800. Host adapter positions and plugging order for the four I/O enclosures are the same for the base frame and the expansion frames with I/O enclosures. The following chart shows the host adapter positions and plugging order for four I/O enclosures. The Install Sequence line indicates the order in which Host Adapter pairs are to be installed. To achieve optimum balance and performance, they should be added in the following order.

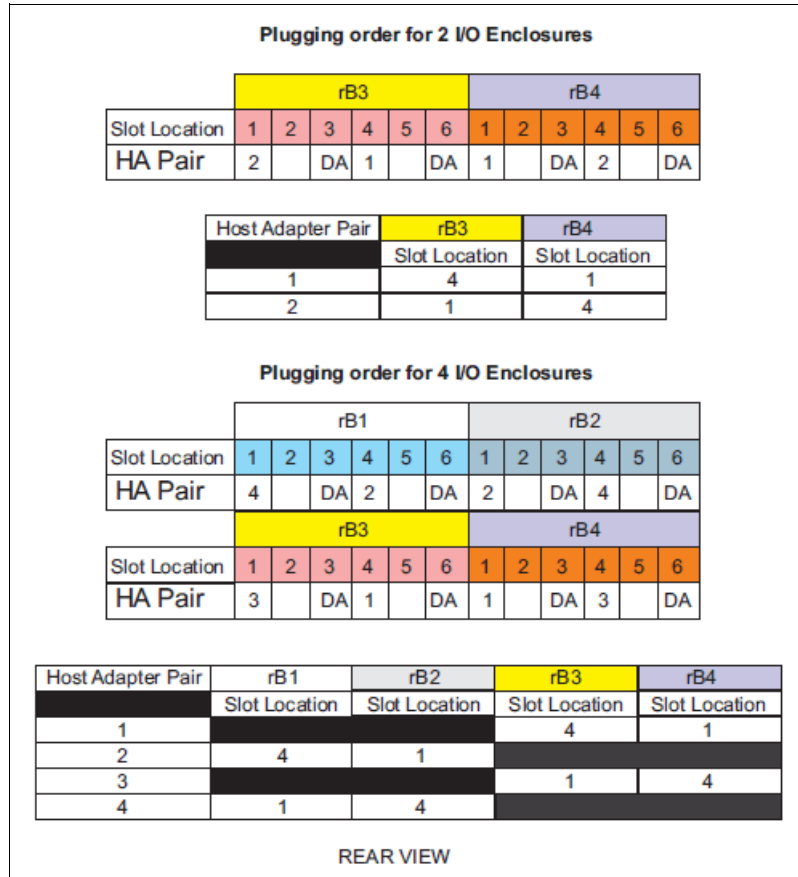


Figure 3-20 DS8800 HA plug order

Fibre Channel is a technology standard that allows data to be transferred from one node to another at high speeds and great distances (up to 10 km). The DS8800 uses the Fibre Channel protocol to transmit SCSI traffic inside Fibre Channel frames. It also uses Fibre Channel to transmit FICON traffic, which uses Fibre Channel frames to carry System z I/O.

Each DS8800 Fibre Channel adapter offers four or eight 8 Gbps Fibre Channel ports. The cable connector required to attach to this adapter is an LC type. Each 8 Gbps port independently auto-negotiates to either 2, 4, or 8 Gbps link speed. Each of the ports on an DS8800 host adapter can also independently be either Fibre Channel protocol (FCP) or FICON. The type of the port can be changed through the DS Storage Manager GUI or by using DSCSI commands. A port cannot be both FICON and FCP simultaneously, but it can be changed as required.

The card itself is PCIe Gen 2. The card is driven by a new high function, that is, a high performance ASIC. To ensure maximum data integrity, it supports metadata creation and checking. Each Fibre Channel port supports a maximum of 509 host login IDs and 1280 paths. This allows for the creation of large storage area networks (SANs).

Fibre Channel supported servers

The current list of servers supported by Fibre Channel attachment can be found at:

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

Consult these documents regularly because they contain the most current information about server attachment support.

Fibre Channel distances

There are two types of host adapter cards you can select:

- ▶ Longwave
- ▶ Shortwave

With longwave, you can connect nodes at distances of up to 10 km (non-repeated). With shortwave, you are limited to a distance of 500 meters (non-repeated). All ports on each card must be either longwave or shortwave. There can be no intermixing of the two types within a card.

3.5.4 Device adapters

Each processor complex accesses the disk subsystem by way of 4 port Fibre Channel arbitrated loop (FC-AL) device adapters (DAs). The DS8000 can have up to 16 of these adapters (installed in pairs).

Each DS8000 device adapter (DA) card offers four FC-AL ports. These ports are used to connect the processor complexes, through the I/O enclosures, to the disk enclosures. The adapter is responsible for managing, monitoring, and rebuilding the RAID arrays. The adapter provides remarkable performance thanks to a high function/high performance ASIC. To ensure maximum data integrity, it supports metadata creation and checking.

DS8700 Device adapters

Each adapter connects the processor complex to two separate switched Fibre Channel networks. Each switched network attaches disk enclosures that each contain up to 16 disks. Each enclosure contains two 20-port Fibre Channel switches. Of these 20 ports, 16 are used to attach to the 16 disks in the enclosure and the remaining four are used to either interconnect with other disk enclosures or to the device adapters. Each disk is attached to both switches. Whenever the device adapter connects to a disk, it uses a switched connection to transfer data. This means that all data travels through the shortest possible path.

DS8800 Device adapters

Each adapter connects the complex to two separate switched Fibre Channel networks. Each network attaches disk enclosures that each contain up to 24 disks. Each storage enclosure contains two 32-port bridges. Of these 32 ports, 24 are used to attach to the 24 disks in the enclosure, two are used to interconnect with other disk enclosures, and two interconnect to the device adapters. Each disk is attached to both switches. Whenever the device adapter connects to a disk, it uses a bridged connection to transfer data. This means that all data travels through the shortest possible path

In the DS8800, a faster application-specific integrated circuit (ASIC) and a faster processor is used on the device adapter cards compared to adapters of other members of the DS8000 family. This leads to higher throughput rates. The DS8800 replaces the PCI-X device and host adapters with native PCIe 8 Gbps FC adapters. This is an improvement from all previous DS8000 models (including the DS8700).

3.6 Disk subsystem

The disk subsystem consists of three components:

1. Device adapter pairs (installed in the I/O enclosures). Device adapters are RAID controllers that access the installed disk drives.
2. The device adapter pairs connect to Fibre Channel controller cards (FCIC) in the disk enclosures. This creates a switched Fibre Channel network to the installed disks.
3. The installed disks, commonly referred to as disk drive modules (DDMs).

We describe the disk subsystem components in the remainder of this section. See 4.6, “RAS on the disk subsystem” on page 89 for additional information.

3.6.1 Disk enclosures

The DS8000 data disks are installed in enclosures called disk enclosures or storage enclosures. These disk enclosures are installed in pairs. The DS8700 and DS8800 have different disk enclosure form factors, and are individually described in this section.

DS8700 disk enclosures

Each DS8700 frame contains a maximum of either 8 or 16 disk enclosures, depending on whether it is a base or expansion frame. Half of the disk enclosures are accessed from the front of the frame, and half from the rear. Each DS8700 disk enclosure contains a total of 16 DDMs or dummy carriers. A dummy carrier looks similar to a DDM in appearance, but contains no electronics. The enclosure is shown in Figure 3-21 on page 59.

Tip: If a DDM is not present, its slot must be occupied by a dummy carrier. This is because without a drive or a dummy, cooling air does not circulate properly.

The DS8700 also supports Solid-State Drives (SSDs). SSDs also come in disk enclosures which are either partially populated with 4 disks, 8 disks, or fully populated with 16 disks. They have the same form factor as DS8700 HDD disks. SSDs and other disks cannot be intermixed within the same enclosure pair.

Each DDM is an industry standard FC-AL or SATA disk. Each disk plugs into the disk enclosure backplane. The *backplane* is the electronic and physical backbone of the disk enclosure.

Each disk enclosure has a redundant pair of Fibre Channel interface control cards (FCIC) that provides the interconnect logic for the disk access and a SES processor for enclosure services. The interface control card has a 2 Gbps FC-AL switch with a Fibre Channel (FC) conversion logic on each disk port. The FC conversion function provides speed aggregation on the FC interconnection ports. The FC trunking connection provides full 2 Gbps transfer rates from a group of drives with lower interface speeds.

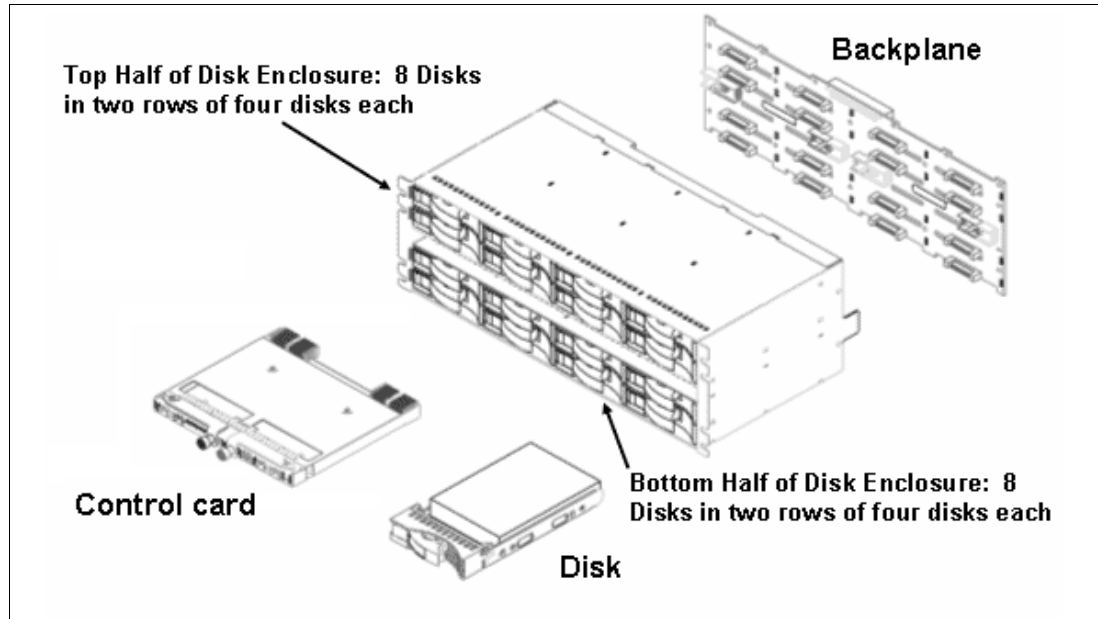


Figure 3-21 DS8700 disk enclosure

DS8800 disk enclosures

Each DS8800 frame contains a maximum of either 10, 14 or 20 disk enclosures, depending on whether it is a base or expansion frame. Each DS8800 disk enclosure contains a total of 24 small form factor (SFF) DDMs or dummy carriers. A dummy carrier looks similar to a DDM in appearance, but contains no electronics. The enclosure is shown in Figure 3-22 on page 60.

Tip: If a DDM is not present, its slot must be occupied by a dummy carrier. Without a drive or a dummy, cooling air does not circulate properly.

The DS8800 also supports Solid-State Drives (SSDs). SSDs also come in disk enclosures which are either partially populated with 4 disks, 8 disks, 16 disks, or fully populated with 24 disks. They have the same form factor as DS8800 HDDs disks. SSDs and HDDs cannot be intermixed within the same enclosure pair.

Each DDM is an industry standard Serial Attached SCSI (SAS) disk. The DDMs are 2.5-inch small form factor disks. This size allows 24 disk drives to be installed in each storage enclosure. Each disk plugs into the disk enclosure backplane. The *backplane* is the electronic and physical backbone of the disk enclosure.

Each disk enclosure has a redundant pair of Fibre Channel interface control cards (FCIC) that provides the interconnect logic for the disk access and a SES processor for enclosure services. The interface control card has an 8 Gbps FC-AL switch with a Fibre Channel (FC) to SAS conversion logic on each disk port. The FC and SAS conversion function provides speed aggregation on the FC interconnection ports. The FC trunking connection provides a full 8 Gbps transfer rates from a group of drives with lower interface speeds.

Enclosure Close-up

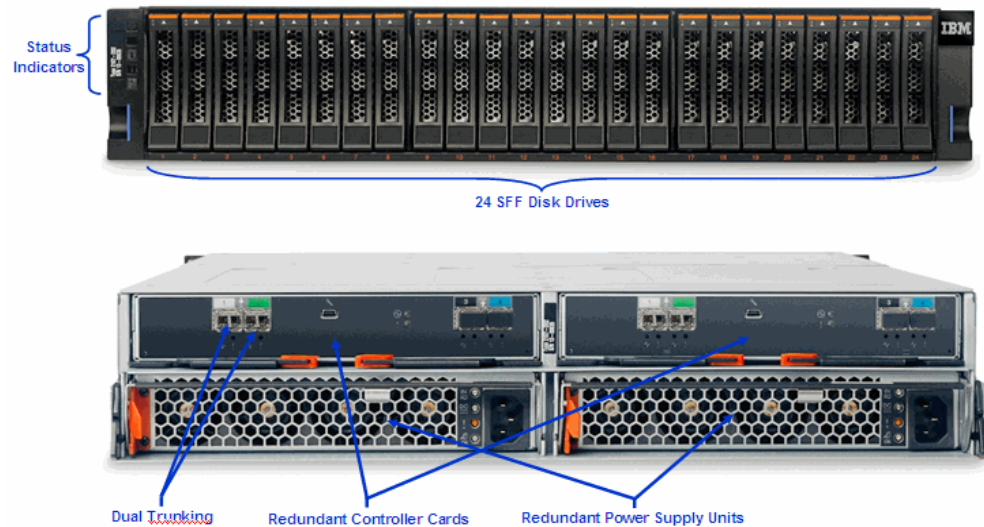


Figure 3-22 DS8800 disk enclosure

Switched FC-AL advantages

The DS8000 uses switched FC-AL technology to link the DA pairs and the DDMs. Switched FC-AL uses the standard FC-AL protocol, but the physical implementation is different. The key features of switched FC-AL technology are:

- ▶ Standard FC-AL communication protocol from DA to DDMs
- ▶ Direct point-to-point links are established between DA and DDM
- ▶ Isolation capabilities in case of DDM failures, providing easy problem determination
- ▶ Predictive failure statistics
- ▶ Simplified expansion, where no cable rerouting is required when adding another disk enclosure

The DS8000 architecture employs dual redundant switched FC-AL access to each of the disk enclosures. The key benefits of doing this are:

- ▶ Two independent networks to access the disk enclosures
- ▶ Four access paths to each DDM
- ▶ Each DA port operates independently
- ▶ Double the bandwidth over traditional FC-AL loop implementations

In Figure 3-23, each DDM is depicted as being attached to two separate Fibre Channel interface connectors (FCIC) with bridges to the disk drive. This means that with two DAs, we have four effective data paths to each disk. Each DA can support two switched FC networks.

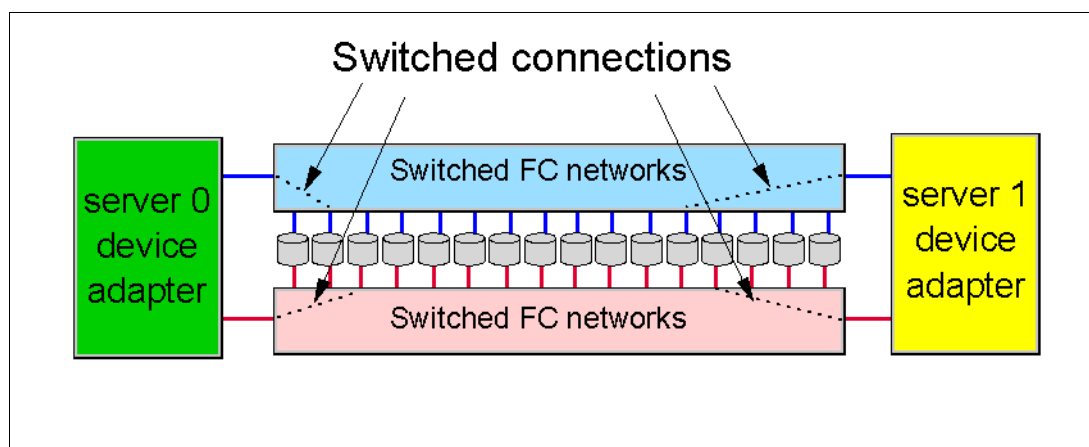


Figure 3-23 DS8000 Disk Enclosure (only 16 disks shown for simplicity)

When a connection is made between the device adapter and a disk, the storage enclosure uses backbone cabling at 8 Gbps, which is translated from Fibre Channel to SAS to the disk drives. This means that a mini-loop is created between the DA port and the disk.

DS8000 Series switched FC-AL implementation

Disk enclosures are installed in pairs.

- For DS8700 the disk enclosure pair is installed as one enclosure in the front of the frame, and one enclosure in the rear of the frame.
- For DS8800, all disk enclosure install from the front of the frame. The disk enclosure pair consists of two enclosures, one on top of the other.

For a more detailed look at how the switched disk architecture expands in the DS8800, see Figure 3-24 on page 62, which depicts how each DS8800 DA connects to two disk networks called *loops*.

Expansion is achieved by adding a disk enclosure pair to the expansion ports of the FCIC cards of the installed disk enclosures. Each loop can potentially have up to four enclosures.

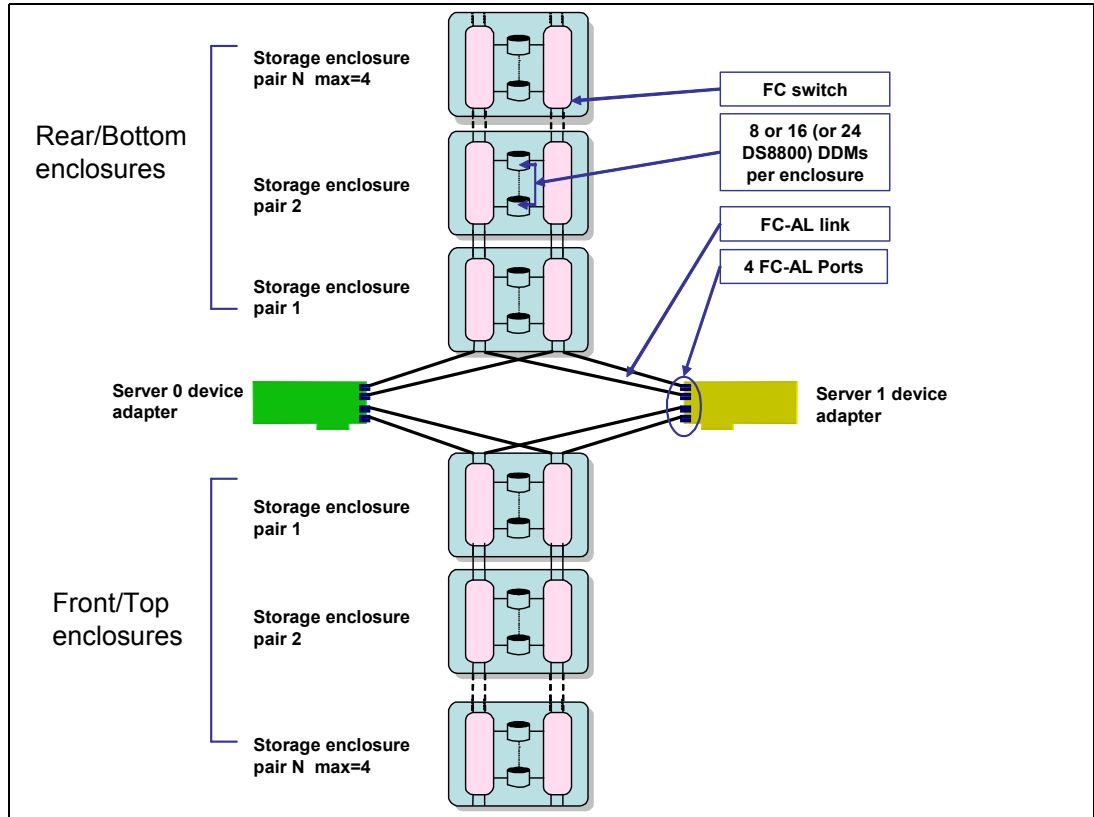


Figure 3-24 DS8000 switched disk expansion

Expansion

Disk enclosures are added in pairs and disks are added in groups of 16.

- For DS8700, it takes two orders of 16 DDMs to fully populate a disk enclosure pair (front and rear).

To provide an example, if a DS8700 had six disk enclosures total, it would have three at the front and three at the rear. If all the enclosures were fully populated with disks, and an additional order of 16 DDMs were purchased, then two new disk enclosures would be added, one at the front and one at the rear, as a pair.

- For DS8800, It takes three orders of 16 DDMs to fully populate a disk enclosure pair (top and bottom).

For example, if a DS8800 had six disk enclosures total and all the enclosures were fully populated with disks, there would be 144 DDMs in three enclosure pairs. If an additional order of 16 DDMs were purchased, then two new disk enclosures would be added as a pair.

In each case, the FC switched networks do not need to be broken to add the disk enclosures. They are simply added to the end of the *loop*; Eight DDMs will go in one disk enclosure of the pair and the remaining eight DDMs will go in the other disk enclosure. If an additional 16 DDMs gets ordered later, they will be used to fill up that pair of disk enclosures. These additional DDMs added have to be of the same capacity and speed as the 16 DDMs already residing in the enclosure pair.

Arrays and spares

Array sites, containing eight DDMs, are created as DDMs are installed. During the configuration, you have the choice of creating a RAID 5, RAID 6, or RAID 10 array by choosing one array site. Note that for SSDs, only RAID 5 is supported. The first four array sites created on a DA pair each contribute one DDM to be a spare.

The intention is to only have four spares per DA pair. However, this number can increase depending on DDM intermix. Four DDMs of the largest capacity and at least two DDMs of the fastest RPM are needed. If all DDMs are the same capacity and speed, four spares are sufficient.

Arrays across loops

Figure 3-25 is used to show the DA pair layout. One DA pair creates two switched loops.

- For DS8700, the front enclosure populates one loop, and the rear enclosures populates the other loop, in a disk enclosure pair. Each enclosure can hold up to 16 DDMs.
- For DS8800, the upper enclosure populates one loop, and the lower enclosure populates the other loop, in a disk enclosure pair. Each enclosure can hold up to 24 DDMs.

Each enclosure places two FC switches onto each loop. DDMs are purchased in groups of 16. Half of the new DDMs go into one disk enclosure and half go into the other disk enclosure of the pair.

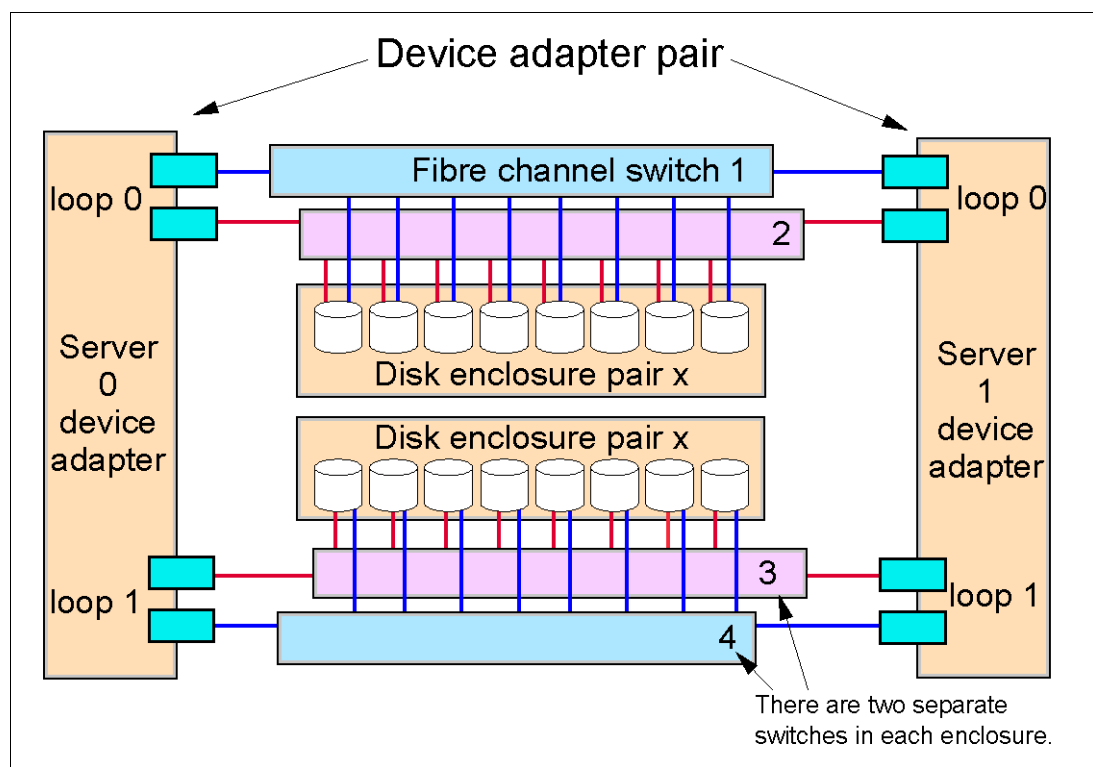


Figure 3-25 DS8000 switched loop layout (only 8 disks per enclosure are shown for simplicity)

An array site consists of eight DDMs. Four DDMs are taken from one enclosure in the disk enclosure pair, and four are taken from the other enclosure in the pair. This means that when a RAID array is created on the array site, half of the array is on each disk enclosure.

One disk enclosure of the pair is on one FC switched loop, and the other disk enclosure of the pair is on a second switched loop. This splits the array across two loops, known as *array*

across loops (AAL). To better understand AAL, see Figure 3-26. To make the diagram clearer, only 16 DDMs are shown, eight in each disk enclosure. When fully populated, there would be 16 or 24 DDMs in each enclosure (depending if this is a DS8700 or DS8800).

Having established the physical layout, the diagram reflects the layout of the array sites. Array site 1 in green (the darker disks) uses the four left DDMs in each enclosure. Array site 2 in yellow (the lighter disks), uses the four right DDMs in each enclosure. When an array is created on each array site, half of the array is placed on each loop. A fully populated disk enclosure pair would have four array sites (DS8700) or six array sites (DS8800).

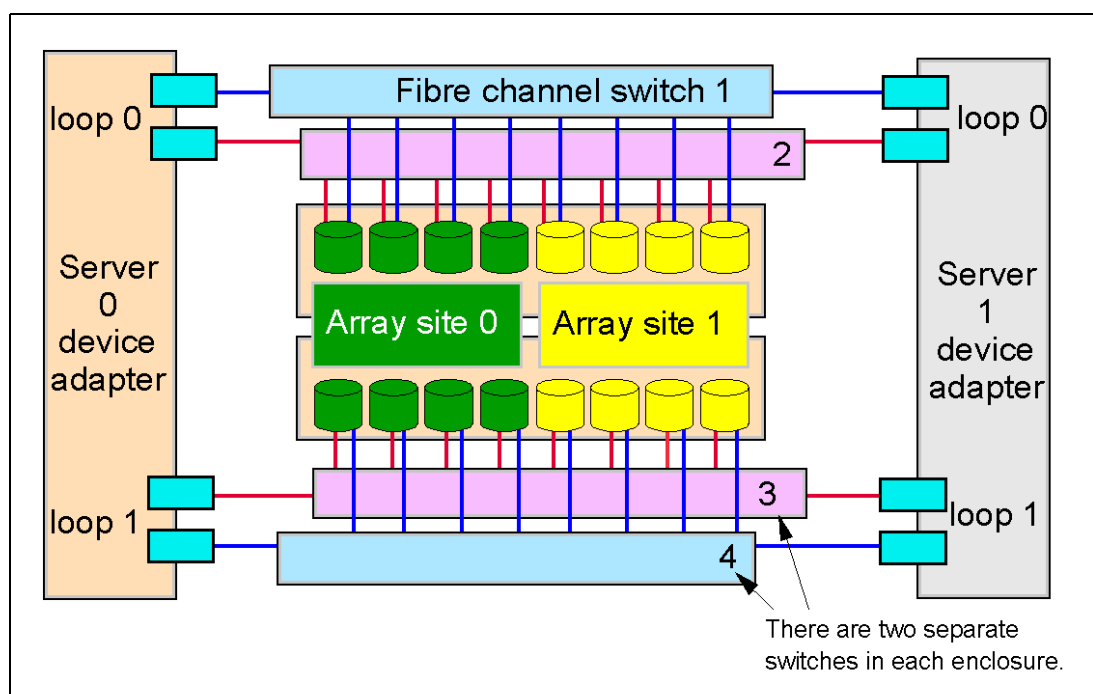


Figure 3-26 Array across loop

AAL benefits

AAL is used to increase performance. When the device adapter writes a stripe of data to a RAID 5 array, it sends half of the write to each switched loop. By splitting the workload in this manner, each loop is worked evenly. This aggregates the bandwidth of the two loops and improves performance. If RAID 10 is used, two RAID 0 arrays are created. Each loop hosts one RAID 0 array. When servicing read I/O, half of the reads can be sent to each loop, again improving performance by balancing workload across loops.

3.6.2 Disk drives

For the DS8700 and DS8800, each disk drive module (DDM) is hot pluggable and has two indicators. The green indicator shows disk activity, and the amber indicator is used with light path diagnostics to allow for easy identification and replacement of a failed DDM.

- The DS8700 supports 300, 450, and 600 GB (15K RPM) Enterprise disk drives. The DS8700 also supports 300, 450, and 600 GB (15K RPM) Full Drive Encryption (FDE) disk drives.

The 146 GB (15K RPM) is also supported but has been withdrawn from marketing as of this Microcode release 6.1 (LMC level 6.6.1.xx).

Additionally, 2TB (7200 RPM) Nearline (SATA) disk drives are supported.

The DS8700 supports 600 GB SSDs. 73 GB and 146 GB SSDs are also supported but have been withdrawn from marketing.

- The DS8800 supports 146 GB (15K RPM), 450 GB (10K RPM), and 600 GB (10K RPM) SAS Enterprise disk drives. The DS8800 also supports 450 GB (10K RPM) and 600 GB (10K RPM) Full Drive Encryption (FDE) SAS drives.

The DS8800 supports Solid State Disk drives (SSD) with a capacity of 300 GB.

For more information about SSDs, see 8.5.3, “DS8000 Solid State Drive (SSD) considerations” on page 222.

For information about encrypted drives and inherent restrictions, refer to *IBM System Storage DS8700 Disk Encryption Implementation and Usage Guidelines*, REDP-4500.

3.7 Power and cooling

The DS8000 series power and cooling system is highly redundant. The components are described in this section. See 4.7, “RAS on the power subsystem” on page 96, for more information about this topic.

Rack Power Control cards

The DS8000 has a pair of redundant Rack Power Control (RPC) cards that are used to control certain aspects of power sequencing throughout the DS8000. These cards are attached to the Flexible Service Processor (FSP) card in each processor complex, which allows them to communicate both with the Hardware Management Console (HMC) and the storage facility. The RPCs also communicate with each primary power supply (PPS).

Primary power supply

The DS8000 primary power supply (PPS) is a wide range power supply that converts AC input voltage into DC voltage. The line cord needs to be ordered specifically for the operating voltage to meet specific requirements. The line cord connector requirements vary widely throughout the world. The line cord might not come with the suitable connector for the country in which the system will be installed, in which case the connector will need to be replaced by an electrician after the machine is delivered.

There are two redundant PPSs in each frame of the DS8000. A single PPS is capable of powering the frame by itself. Each PPS has internal fans to supply cooling for that power supply.

There can also be an optional booster module that will allow the PPSs to temporarily run the disk enclosures off of the batteries, if the extended power line disturbance (ePLD) feature has been purchased (see Chapter 4, “RAS on IBM System Storage DS8000” on page 71 for a complete explanation of why this feature might be necessary for your installation).

- In the DS8700, the PPS creates 208V output power for the processor complexes and I/O enclosure power supplies. It also creates 5V and 12V DC power for the disk enclosures.
- For the DS8800, the PPS supplies 208V output power to six power distribution units (PDUs).

Each PDU is supplied from both PPSs in each frame for redundancy.

In the base frame, the PDUs supply power to the processor complexes, the I/O enclosures, and the disk enclosures. In the first expansion frame, the PDUs supply power to the I/O

enclosures and the disk enclosures. In the second expansion frame, the PDUs supply power to the disk enclosures, as there are no I/O enclosures or processor complexes in this frame.

Each disk enclosure has two power supply units (PSU). The disk enclosure PSUs are connected to two separate PDUs for redundancy.

Figure 3-27 shows the DS8800 base frame PDUs.

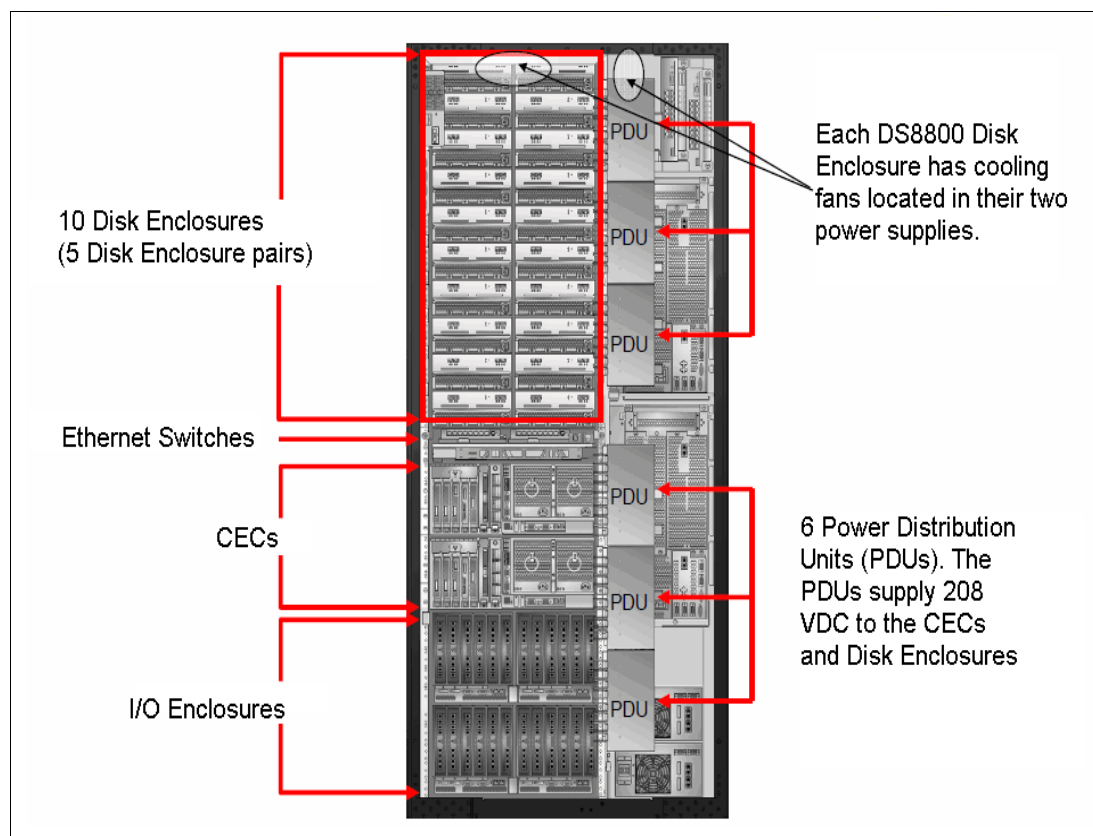


Figure 3-27 DS8800 Base frame power distribution units (PDUs)

Processor and I/O enclosure power supplies

Each processor complex and I/O enclosure have dual redundant power supplies to convert 208V DC into the required voltages for that enclosure or complex. Each enclosure also has its own cooling fans.

Disk enclosure power and cooling

For DS8700, the disk enclosures power directly from the PPSs. The disk enclosures are located in front and rear of each frame. They have cooling fans located in a plenum above the disk enclosures in each frame. They draw cooling air in through the front of each enclosure and exhaust air to the center plenum and then out the top of the frame.

For DS8800, the disk enclosures have two power supply units (PSU) for each disk enclosure. These PSUs draw power from the PPSs through the PDUs. There are cooling fans located in each PSU. These fans draw cooling air through the front of each disk enclosure and exhaust air out the rear of the frame. Figure 3-22 on page 60 shows the DS8800 disk enclosure PSUs.

Battery backup assemblies

The Battery BackUp (BBU) assemblies help protect data in the event of a loss of external power. In the event of a complete loss of AC input power, the battery assemblies are used to maintain power to the processor complexes and I/O enclosures for a sufficient period of time, to allow the contents of NVS memory (modified data not yet destaged to disk from cache) to be written to a number of disk drives internal to the processor complexes.

The DDMs are not protected from power loss unless the extended power line disturbance (ePLD) feature has been installed in the system.

3.8 Management console network

All base frames ship with one Hardware Management Console (HMC) and two Ethernet switches. A mobile computer HMC (Lenovo ThinkPad), shown in Figure 3-28, will be shipped with a DS8000.

DS8000 logical configuration creation and changes performed by the storage administrator using the GUI or DSCLI are passed to the storage system through the HMC.

More information about the HMC can be found in Chapter 9, “DS8000 HMC planning and setup” on page 225

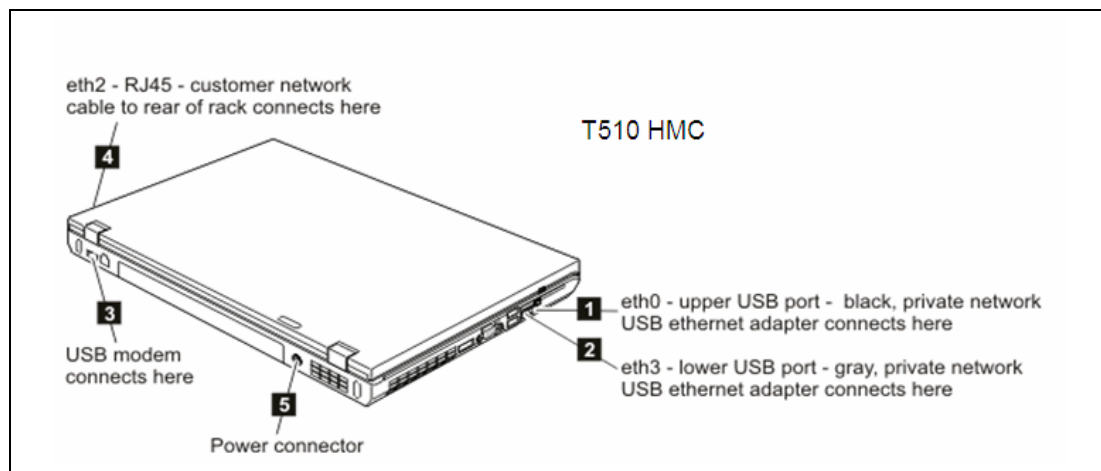


Figure 3-28 Mobile computer HMC

Tip: The DS8000 HMC supports IPv6, the next generation of the Internet Protocol. The HMC continues to support the IPv4 standard, and mixed IPV4 and IPv6 environments.

3.8.1 Ethernet switches

The DS8000 base frame has two 8-port Ethernet switches. Two switches are supplied to allow the creation of a fully redundant private management network. Each processor complex has connections to each switch to allow each server to access both private networks. These networks cannot be accessed externally, and no external connections are allowed. External Client network connection to the DS8000 system is through a separate patch panel connection. The switches get power from the internal power bus and thus do not require separate power outlets. The switches are shown in Figure 3-29 on page 68.

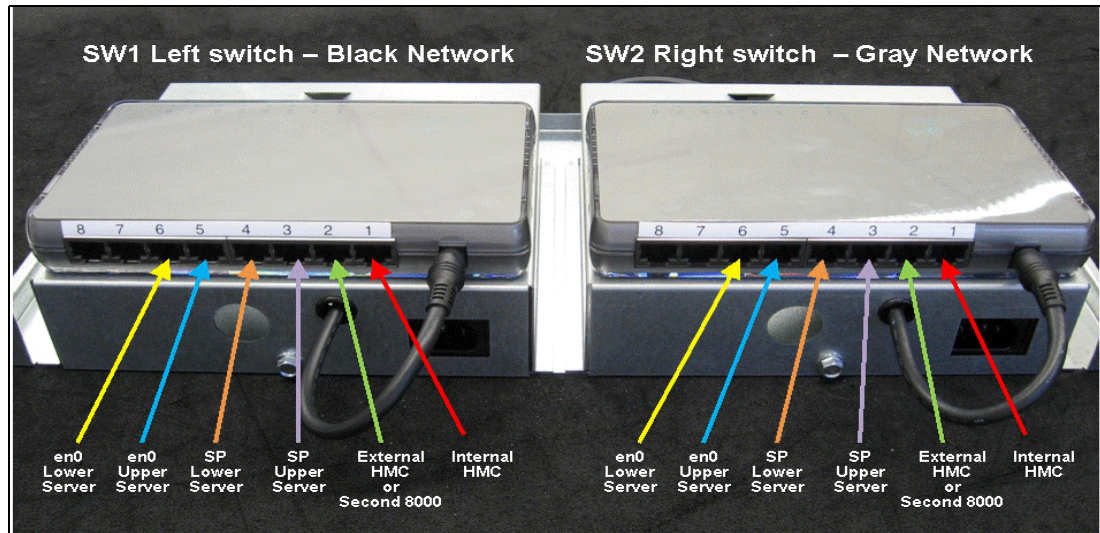


Figure 3-29 Ethernet switches

See 4.5, “RAS on the HMC” on page 88 for more information.

3.9 System Storage Productivity Center

The IBM System Storage Productivity Center (SSPC) is an optional hardware appliance with pre-installed software that can help you improve and centralize the management of your storage environment through the integration of products. It provides a single point of management integrating the functionality of the IBM Tivoli Storage Productivity Center with storage devices and element managers in an easy-to-use user interface for management.

With SSPC, it is possible to manage and fully configure multiple DS8000 storage systems from a single point of control.

The SSPC appliance consists of the following components:

- ▶ IBM Tivoli Storage Productivity Center licensed as TPC Basic Edition (includes the Tivoli Integrated Portal). A TPC upgrade requires that you purchase and add additional TPC licenses.
- ▶ DS CIM Agent Command-Line Interface (DSCIMCLI).
- ▶ IBM Tivoli Storage Productivity Center for Replication (TPC-R). To run TPC-R on SSPC, you must purchase and add TPC-R base license for Flashcopy.
- ▶ IBM DB2® Enterprise Server Edition 9.7 64-bit Enterprise.
- ▶ IBM JAVA 1.6 is preinstalled. You do not need to download Java from Sun Microsystems.

Optionally, the following components can be installed on the SSPC:

- ▶ DS8000 Command-Line Interface (DSCLI)
- ▶ Antivirus software

SSPC can be ordered as a software (SW) package to be installed on the client's hardware or can be ordered as Model 2805, which has the software preinstalled on an System x3550 with a Quad Core Intel processor (2.4 Ghz) with 8 GB of memory running Windows Server 2008 (Figure 3-30).



Figure 3-30 SSPC hardware

For more detailed information about SSPC, see Chapter 12, “System Storage Productivity Center” on page 279.

3.10 Isolated Tivoli Key Lifecycle Manager server

The Tivoli Key Lifecycle Manager (TKLM) software performs key management tasks for IBM encryption enabled hardware, such as the IBM System Storage DS8000 series and IBM encryption-enabled tape drives by providing, protecting, storing, and maintaining encryption keys that are used to encrypt information being written to, and decrypt information being read from, encryption enabled disks. TKLM operates on a variety of operating systems.

For DS8700 or DS8800 storage systems shipped with Full Disk Encryption (FDE) drives, two TKLM key servers are required. An Isolated Key Server (IKS) with dedicated hardware and non-encrypted storage resources is required.

The isolated TKLM key server can be ordered from IBM. It is the same hardware as is used for the SSPC. The following software is used on the isolated key server:

- ▶ Linux operating system
- ▶ Tivoli Key Lifecycle Manager V, which includes DB2 V9.1 FB4

No other hardware or software is allowed on the IKS. See 4.8, “RAS and Full Disk Encryption” on page 99 for more information.

For more information, refer to *IBM System Storage DS8700 Disk Encryption Implementation and Usage Guidelines*, REDP-4500.



RAS on IBM System Storage DS8000

This chapter describes the reliability, availability, and serviceability (RAS) characteristics of the IBM System Storage DS8700 and DS8800. The following topics are covered in this chapter:

- ▶ Names and terms for DS8000
- ▶ RAS features of the DS8000 central electronic complex
- ▶ Central electrical complex failover and failback
- ▶ Data flow in the DS8000
- ▶ RAS on the HMC
- ▶ RAS on the disk subsystem
- ▶ RAS on the power subsystem
- ▶ RAS and Full Disk Encryption
- ▶ Other features

4.1 Names and terms for the DS8000 storage system

It is important to understand the naming conventions used to describe DS8000 components and constructs to fully appreciate the discussion of RAS concepts. Although most terms have been introduced in previous chapters of this book, they are repeated and summarized here because the rest of this chapter will use these terms frequently.

Storage unit

The term *storage unit* describes a single DS8000 (base frame plus additional installed expansion frames). If your organization has one DS8000, then you have a single storage complex that contains a single storage unit.

Base frame

The DS8700 is available as a single model type (941) and the DS8800 is available as a single model type (951). Both include a complete storage unit contained in the base frame. To increase the storage capacity, expansion frames can be added. Expansion frames can only be added to 4-way systems. Up to four expansion frames can be added to the DS8700 base frame. Up to two expansion frames can be added to the DS8800 base frame.

2-way systems cannot have expansion frames (this includes business class cabled DS8800 systems).

A base frame contains the following components:

- ▶ Power and cooling components (power supplies, batteries, and fans)
- ▶ Power control cards: Rack Power Control (RPC) and System Power Control Network (SPCN)
- ▶ Two POWER6 (DS8700) or POWER6+ (DS8800) central electrical complexes
- ▶ Two or four I/O Enclosures for Host Adapters and Device Adapters
- ▶ 2 Gigabit Ethernet switches for the internal networks
- ▶ Hardware Management Console

For DS8700:

- ▶ Up to four pairs (eight total) of disk enclosures for storage disks:
 - Each disk enclosure can have up to 16 disk drive modules (DDM).
 - The base frame can have a maximum of 128 DDMs.

For DS8800:

- ▶ Up to five pairs (10 total) of disk enclosures for storage disks:
 - Each disk enclosure can have up to 24 disk drive modules (DDM).
 - The base frame can have a maximum of 240 DDMs.

Expansion frame

Expansion frames can be added one at a time to increase the overall capacity of the storage unit. All expansion frames contain the power and cooling components needed to run the frame. The first expansion frame contains storage disks and I/O enclosures. Subsequent expansion frames contain storage disks only. Adding an expansion frame is a concurrent operation for the DS8000.

For DS8700:

- ▶ Each Disk enclosure can have up to 16 disk drive modules (DDM).
- ▶ An expansion frame can have a maximum of 256 DDMs (the fourth expansion frame is limited to 128 DDMs).

For DS8800:

- ▶ Each disk enclosure can have up to 24 disk drive modules (DDM).
- ▶ The first expansion frame can have a maximum of 336 DDMs in 14 disk enclosures.
- ▶ The second expansion frame can have a maximum of 480 DDMs in 20 disk enclosures.

Storage complex

This term *storage complex* describes a group of DS8000s (that is, DS8700s or DS8800s) managed by a single management console. A storage complex can, and usually does, consist of simply a single DS8000 storage unit (base frame plus additional installed expansion frames).

Central Electronic Complex/processor complex/storage server

In the DS8000 (DS8700 or DS8800), a central electronic complex is an IBM System p server built on the POWER6 architecture. The central electrical complexes run the AIX V6.1 operating system and storage-specific microcode.

The DS8000 contains two central electrical complexes as a redundant pair so that if either fails, the DS8000 will fail over to the remaining central electrical complex and continue to run the storage unit. Each central electrical complex can have up to 192 GB of memory, and one or two POWER6 processor modules. In other models of the DS8000 family, a central electrical complex was also referred to as a *processor complex* or a *storage server*. The central electrical complexes are identified as CEC0 and CEC1. Some chapters and illustrations in this publication refer to *Server 0* and *Server 1*: These are the same as CEC0 and CEC1 for the DS8000.

HMC

The Hardware Management Console (HMC) is the management console for the DS8000 storage unit. With connectivity to the central electrical complexes, the client network, the SSPC, and other management systems, the HMC becomes the focal point for most operations on the DS8000. All storage configuration and service actions are managed through the HMC. Although many other IBM products also use an HMC, the DS8000 HMC is unique to the DS8000 family.

System Storage Productivity Center

The DS8000 can utilize the IBM System Storage Productivity Center (SSPC), which is a management system that integrates the power of the IBM Tivoli Storage Productivity Center (TPC) and the DS Storage Manager user interfaces (residing at the HMC) into a single view. The SSPC (machine type 2805-MC5) is an integrated hardware and software solution for centralized management of IBM storage products with IBM storage management software. The SSPC is described in detail in Chapter 12, “System Storage Productivity Center” on page 279.

Storage facility images and logical partitions

A logical partition (LPAR) is a virtual server within a physical processor complex. A storage facility image (SFI) consists of two logical partitions acting together as a virtual storage server. Earlier DS8000 models supported more than one SFI, meaning that 2 LPARs were implemented on a single central electrical complex. The DS8700 and DS8800 implement a single LPAR on a single central electrical complex. There is only one SFI, which owns 100% of the physical resources. So for the DS8700 and DS8800, the term *storage facility image* can be considered synonymous with storage unit.

4.2 RAS features of DS8000 central electrical complex

Reliability, availability, and serviceability (RAS) are important concepts in the design of the IBM System Storage DS8000. Hardware features, software features, design considerations, and operational guidelines all contribute to make the DS8000 extremely reliable. At the heart of the DS8000 is a pair of POWER6 based System p servers known as central electrical complexes. These two servers share the load of receiving and moving data between the attached hosts and the disk arrays. However, they are also redundant so that if either central electrical complex fails, the system will fail over to the remaining central electrical complex and continue to run the DS8000 without any host interruption. This section looks at the RAS features of the central electrical complexes, including the hardware, the operating system, and the interconnections.

4.2.1 POWER6 Hypervisor

The POWER6 Hypervisor (PHYP) is a component of system firmware that will always be installed and activated, regardless of the system configuration. It operates as a hidden partition, with no processor resources assigned to it.

The Hypervisor provides the following capabilities:

- ▶ Reserved memory partitions allow you to set aside a portion of memory to use as cache and a portion to use as NVS.
- ▶ Preserved memory support allows the contents of the NVS and cache memory areas to be protected in the event of a server reboot.
- ▶ I/O enclosure initialization control, so that when one server is being initialized, it does not initialize an I/O adapter that is in use by another server.
- ▶ Automatic reboot of a frozen partition or Hypervisor.

The AIX operating system uses PHYP services to manage the translation control entry (TCE) tables. The operating system communicates the desired I/O bus address to logical mapping, and the Hypervisor translates that into the I/O bus address to physical mapping within the specific TCE table. The Hypervisor needs a dedicated memory region for the TCE tables to translate the I/O address to the partition memory address, and then the Hypervisor can perform direct memory access (DMA) transfers to the PCI adapters.

4.2.2 POWER6 processor

IBM POWER6 systems have a number of new features that enable systems to dynamically adjust when issues arise that threaten availability. Most notably, POWER6 systems introduce the POWER6 Processor Instruction Retry suite of tools, which includes Processor Instruction Retry, Alternate Processor Recovery, Partition Availability Prioritization, and Single Processor Checkstop. Taken together, in many failure scenarios these features allow a POWER6 processor-based system to recover with no impact from the failing core. The DS8700 uses a POWER6 processor running at 4.7 GHz. The DS8800 uses a POWER6+ processor running at 5.0 GHz.

The POWER6 processor implements the 64-bit IBM Power Architecture® technology and capitalizes on all the enhancements brought by the POWER5™ processor. Each POWER6 chip incorporates two dual-threaded Simultaneous Multithreading processor cores, a private 4 MB level 2 cache (L2) for each processor, a 36 MB L3 cache controller shared by the two processors, integrated memory controller, and data interconnect switch. It is designed to provide an extensive set of RAS features that include improved fault isolation, recovery from

errors without stopping the processor complex, avoidance of recurring failures, and predictive failure analysis.

POWER6 RAS features

The following sections describe the RAS leadership features of IBM POWER6 systems in more detail.

POWER6 processor instruction retry

Soft failures in the processor core are transient errors. When an error is encountered in the core, the POWER6 processor will first automatically retry the instruction. If the source of the error was truly transient, the instruction will succeed and the system will continue as before. On predecessor IBM systems, this error would have caused a checkstop.

POWER6 alternate processor retry

Hard failures are more difficult, being true logical errors that will be replicated each time the instruction is repeated. Retrying the instruction will not help in this situation because the instruction will continue to fail. Systems with POWER6 processors introduce the ability to extract the failing instruction from the faulty core and retry it elsewhere in the system, after which the failing core is dynamically deconfigured and called out for replacement. The entire process is transparent to the partition owning the failing instruction. Systems with POWER6 processors are designed to avoid what would have been a full system outage.

POWER6 cache availability

In the event that an uncorrectable error occurs in L2 or L3 cache, the system will be able to dynamically remove the offending line of cache without requiring a reboot. In addition, POWER6+ utilizes an L1/L2 cache design and a write-through cache policy on all levels, helping to ensure that data is written to main memory as soon as possible.

POWER6 single processor checkstopping

Another major advancement in POWER6 processors is single processor checkstopping. A processor checkstop would result in a system checkstop. A new feature in System 570 is the ability to contain most processor checkstops to the partition that was using the processor at the time. This significantly reduces the probability of any one processor affecting total system availability.

POWER6 fault avoidance

POWER6 systems are built to keep errors from ever happening. This quality-based design includes such features as reduced power consumption and cooler operating temperatures for increased reliability, enabled by the use of copper chip circuitry, silicon on insulator (SOI), and dynamic clock-gating. It also uses mainframe-inspired components and technologies.

POWER6 First Failure Data Capture

If a problem should occur, the ability to diagnose it correctly is a fundamental requirement upon which improved availability is based. The POWER6 incorporates advanced capability in startup diagnostics and in runtime First Failure Data Capture (FFDC) based on strategic error checkers built into the chips. Any errors that are detected by the pervasive error checkers are captured into Fault Isolation Registers (FIRs), which can be interrogated by the service processor (SP). The SP has the capability to access system components using special-purpose service processor ports or by access to the error registers.

The FIRs are important because they enable an error to be uniquely identified, thus enabling the appropriate action to be taken. Appropriate actions might include such things as a bus retry, error checking and correction (ECC), or system firmware recovery routines. Recovery routines could include dynamic deallocation of potentially failing components.

Errors are logged into the system nonvolatile random access memory (NVRAM) and the SP event history log, along with a notification of the event to AIX for capture in the operating system error log. Diagnostic Error Log Analysis (DIAGELA) routines analyze the error log entries and invoke a suitable action, such as issuing a warning message. If the error can be recovered, or after suitable maintenance, the service processor resets the FIRs so that they can accurately record any future errors.

N+1 redundancy

High-opportunity components, or those that most affect system availability, are protected with redundancy and the ability to be repaired concurrently. The use of redundant parts allows the system to remain operational. Among them are:

- ▶ Redundant spare memory bits in cache, directories, and main memory
- ▶ Redundant and hot-swap cooling
- ▶ Redundant and hot-swap power supplies

Self-healing

For a system to be self-healing, it must be able to recover from a failing component by first detecting and isolating the failed component. It should then be able to take it offline, fix or isolate it, and then reintroduce the fixed or replaced component into service without any application disruption. Examples include:

- ▶ Bit steering to redundant memory in the event of a failed memory module to keep the server operational
- ▶ Bit scattering, thus allowing for error correction and continued operation in the presence of a complete chip failure (Chipkill recovery)
- ▶ Single-bit error correction using Error Checking and Correcting (ECC) without reaching error thresholds for main, L2, and L3 cache memory
- ▶ L3 cache line deletes extended from 2 to 10 for additional self-healing
- ▶ ECC extended to inter-chip connections on fabric and processor bus
- ▶ Memory scrubbing to help prevent soft-error memory faults
- ▶ Dynamic processor deallocation

Memory reliability, fault tolerance, and integrity

POWER6 uses Error Checking and Correcting (ECC) circuitry for system memory to correct single-bit memory failures and to detect double-bit memory failures. Detection of double-bit memory failures helps maintain data integrity. Furthermore, the memory chips are organized such that the failure of any specific memory module only affects a single bit within a four-bit ECC word (bit-scattering), thus allowing for error correction and continued operation in the presence of a complete chip failure (Chipkill recovery).

The memory DIMMs also utilize memory scrubbing and thresholding to determine when memory modules within each bank of memory should be used to replace ones that have exceeded their threshold of error count (dynamic bit-steering). Memory scrubbing is the process of reading the contents of the memory during idle time and checking and correcting any single-bit errors that have accumulated by passing the data through the ECC logic. This function is a hardware function on the memory controller chip and does not influence normal system memory performance.

Fault masking

If corrections and retries succeed and do not exceed threshold limits, the system remains operational with full resources, and neither you nor your IBM service representative need to intervene.

Mutual surveillance

The SP can monitor the operation of the firmware during the boot process, and it can monitor the operating system for loss of control. This enables the service processor to take appropriate action when it detects that the firmware or the operating system has lost control. Mutual surveillance also enables the operating system to monitor for service processor activity and can request a service processor repair action if necessary.

Additional memory keys in the POWER6+ (DS8800)

The DS8800 utilizes the POWER6+ processor, which delivers improved performance over the POWER6 processor. Additionally, the POWER6+ processor has 16 memory keys compared to eight memory keys in the POWER6 processor. This doubling of keys (eight for the kernel, seven for the user, and one for the Hypervisor) provides enhanced key resiliency that is important for virtualization environments.

4.2.3 AIX operating system

Each central electrical complex runs the IBM AIX Version 6.1 operating system. This is the latest generation of the IBM well-proven, scalable, and open standards-based UNIX-like operating system. This version of AIX includes support for Failure Recovery Routines (FRRs).

With AIX V6.1, the kernel has been enhanced with the ability to recover from unexpected errors. Kernel components and extensions can provide failure recovery routines to gather serviceability data, diagnose, repair, and recover from errors. In previous AIX versions, kernel errors always resulted in an unexpected system halt.

Refer to *IBM AIX Version 6.1 Differences Guide*, SG24-7559, for more information about how AIX V6.1 adds to the RAS features of AIX 5L™ V5.3.

You can also reference the IBM website for a more thorough review of the features of the IBM AIX operating system at:

<http://www.ibm.com/systems/power/software/aix/index.html>

4.2.4 Central electrical complex dual hard drive rebuild

If a simultaneous failure of the dual hard drives in a central electrical complex occurs, they need to be replaced and then have the AIX OS and DS8000 microcode reloaded. The DS8700 introduced a significant improvement in RAS for this process, known as a *rebuild*. Any fault that causes the central electrical complex to be unable to load the operating system from its internal hard drives would lead to this service action. This function is also supported on the DS8800.

For a rebuild on previous DS8000 models, the IBM service representative would have to load multiple CDs/DVDs directly onto the central electrical complex being serviced. For the DS8700 and DS8800, there are no optical drives on the central electrical complexes: Only the HMC has a DVD drive. For a central electrical complex dual hard drive rebuild, the service representative acquires the needed code bundles on the HMC, which then runs as a Network Installation Management on Linux (NIMoL) server. The HMC provides the operating system and microcode to the central electrical complex over the DS8000 internal network, which is much faster than reading and verifying from an optical disc.

All of the tasks and status updates for a central electrical complex dual hard drive rebuild are done from the HMC, which is also aware of the overall service action that necessitated the rebuild. If the rebuild fails, the HMC manages the errors, including error data, and allows the service representative to address the problem and restart the rebuild. When the rebuild

completes, the server is automatically brought up for the first time (IML). After the IML is successful, the service representative can resume operations on the central electrical complex.

Overall, the rebuild process on a DS8000 is more robust and straightforward, thereby reducing the time needed to perform this critical service action.

4.2.5 RIO-G interconnect

The RIO-G interconnect is a high speed loop between the two central electrical complexes. Each RIO-G port can operate at 1 GHz in bidirectional mode, and is capable of passing data in each direction on each cycle of the port. In previous generations of the DS8000, the I/O Enclosures were on the RIO-G loops between the two central electrical complexes. The RIO-G bus carried the CEC-to-DDM data (host I/O) and all CEC-to-CEC communications.

For the DS8700 and DS8800, the I/O Enclosures are wired point-to-point with each central electrical complex using a PCI Express architecture. This means that only the CEC-to-CEC (XC) communications are now carried on the RIO-G, and the RIO loop configuration is greatly simplified. Figure 4-1 shows the new fabric design of the DS8000.

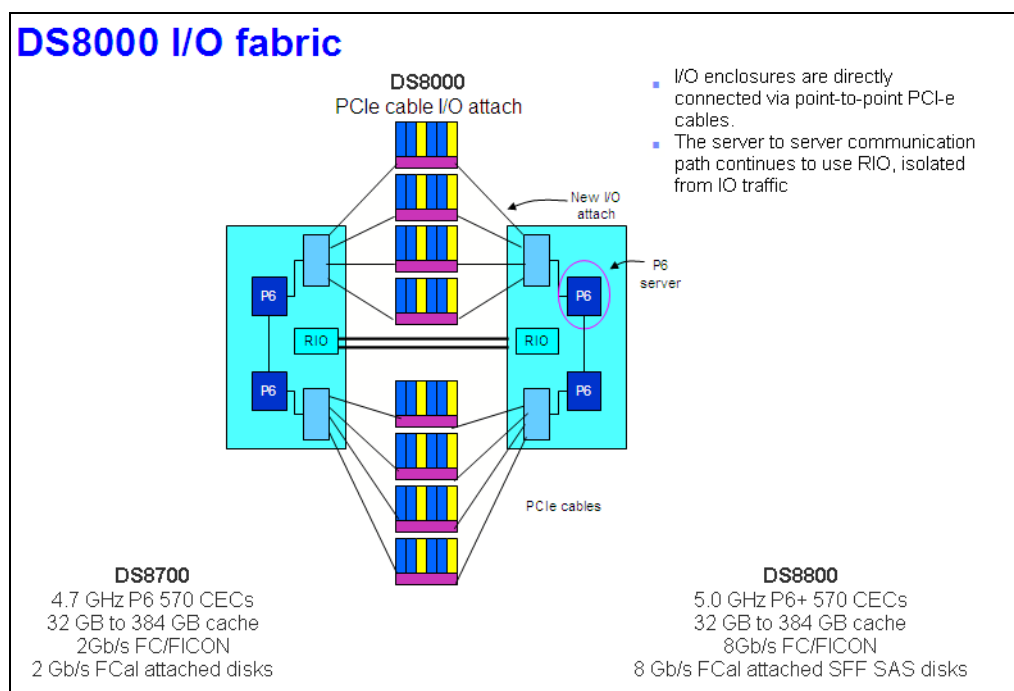


Figure 4-1 DS8000 design of RIO-G loop and I/O enclosures

4.2.6 Environmental monitoring

Environmental monitoring related to power, fans, and temperature is performed by the System Power Control Network (SPCN). Environmental critical and non-critical conditions generate Early Power-Off Warning (EPOW) events. Critical events (for example, a Class 5 AC power loss) trigger appropriate signals from hardware to the affected components to prevent any data loss without operating system or firmware involvement. Non-critical environmental events are logged and reported using Event Scan.

Temperature monitoring is also performed. If the ambient temperature goes above a preset operating range, the rotation speed of the cooling fans will be increased. Temperature

monitoring also warns the internal microcode of potential environment-related problems. An orderly system shutdown will occur when the operating temperature exceeds a critical level.

Voltage monitoring provides warning and an orderly system shutdown when the voltage is out of operational specification.

4.2.7 Resource deallocation

If recoverable errors exceed threshold limits, resources can be deallocated with the system remaining operational, allowing deferred maintenance at a convenient time. Dynamic deallocation of potentially failing components is nondisruptive, allowing the system to continue to run. Persistent deallocation occurs when a failed component is detected. It is then deactivated at a subsequent reboot.

Dynamic deallocation functions include the following components:

- ▶ Processor
- ▶ L3 cache lines
- ▶ Partial L2 cache deallocation
- ▶ PCIe bus and slots

Persistent deallocation functions include the following components:

- ▶ Processor
- ▶ Memory
- ▶ Deconfigure or bypass failing I/O adapters
- ▶ L2 cache

Following a hardware error that has been flagged by the service processor, the subsequent reboot of the server invokes extended diagnostics. If a processor or cache has been marked for deconfiguration by persistent processor deallocation, the boot process will attempt to proceed to completion with the faulty device automatically deconfigured. Failing I/O adapters will be deconfigured or bypassed during the boot process.

4.3 Central electrical complex failover and failback

To understand the process of central electrical complex failover and failback, we have to review the logical construction of the DS8000. For more complete explanations, see Chapter 5, “Virtualization concepts” on page 103.

Creating logical volumes on the DS8000 works through the following constructs:

- ▶ Storage DDMs are installed into predefined *array sites*.
- ▶ These array sites are used to form *arrays*, structured as RAID 5, RAID 6, or RAID 10 (restrictions apply for Solid-State Drives).
- ▶ These RAID arrays then become members of a *rank*.
- ▶ Each rank then becomes a member of an *Extent Pool*. Each Extent Pool has an affinity to either server 0 or server 1. Each Extent Pool is either open systems fixed block (FB) or System z count key data (CKD).
- ▶ Within each Extent Pool, we create *logical volumes*. For open systems, these are called LUNs. For System z, these are called *volumes*. LUN stands for *logical unit number*, which is used for SCSI addressing. Each logical volume belongs to a *logical subsystem* (LSS).

For open systems, the LSS membership is only significant for Copy Services. But for System z, the LSS is the logical control unit (LCU), which equates to a 3990 (a System z disk

controller which the DS8000 emulates). It is important to remember that LSSs that have an even identifying number have an affinity with CEC 0, and LSSs that have an odd identifying number have an affinity with CEC 1. When a host operating system issues a write to a logical volume, the DS8000 host adapter directs that write to the central electrical complex that *owns* the LSS of which that logical volume is a member.

4.3.1 Dual operational

One of the basic premises of RAS, in respect to processing host data, is that the DS8000 will always try to maintain two copies of the data while it is moving through the storage system. The central electrical complexes have two areas of their primary memory used for holding host data: cache memory and *non-volatile storage* (NVS). NVS is an area of the system RAM that is persistent across a server reboot.

Tip: For the previous generations of DS8000, the maximum available NVS was 4 GB per server. For the DS8700 and DS8800, that maximum has been increased to 6 GB per server.

When a write is issued to a volume and the central electrical complexes are both operational, this *write data* gets directed to the central electrical complex that owns this volume. The data flow begins with the write data being placed into the cache memory of the owning central electrical complex. The write data is also placed into the NVS of the other central electrical complex. The NVS copy of the write data is accessed only if a write failure should occur and the cache memory is empty or possibly invalid. Otherwise it will be discarded after the destaging is complete as shown in Figure 4-2.

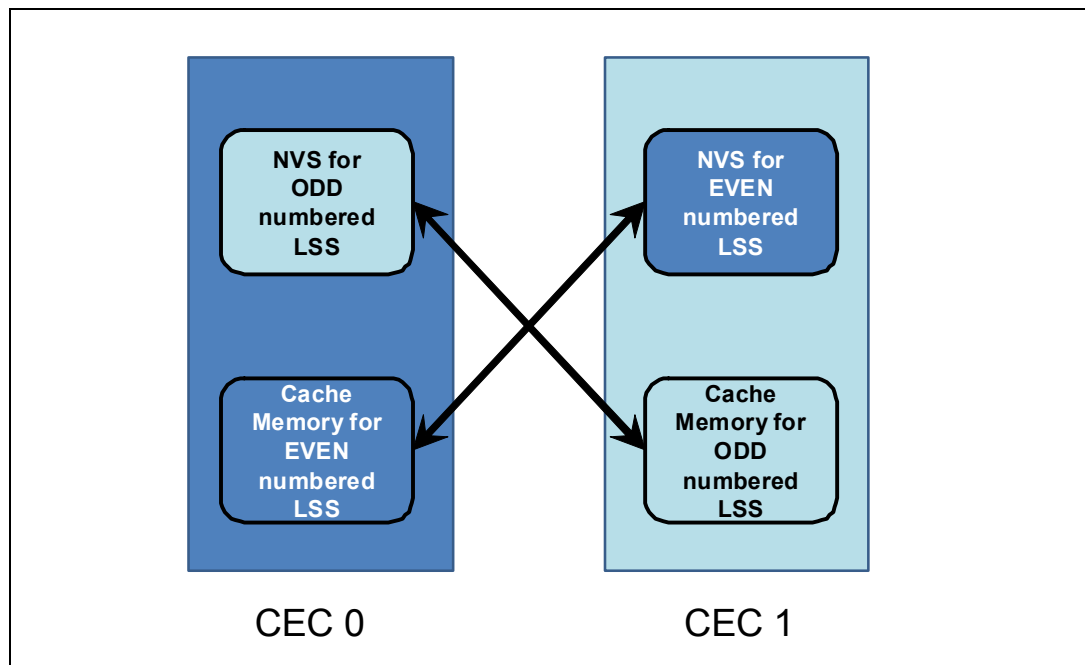


Figure 4-2 Write data when central electrical complexes are dual operational

Figure 4-2 shows how the cache memory of CEC 0 is used for all logical volumes that are members of the even LSSs. Likewise, the cache memory of CEC 1 supports all logical volumes that are members of odd LSSs. For every write that gets placed into cache, a copy gets placed into the NVS memory located in the alternate central electrical complex. Thus,

the normal flow of data for a write when both central electrical complexes are operational is as follows:

1. Data is written to cache memory in the owning central electrical complex. At the same time, data is written to NVS memory of the alternate central electrical complex.
2. The write operation is reported to the attached host as completed.
3. The write data is destaged from the cache memory to a disk array.
4. The write data is discarded from the NVS memory of the alternate central electrical complex.

Under normal operation, both DS8000 central electrical complexes are actively processing I/O requests. The following sections describe the failover and failback procedures that occur between the central electrical complexes when an abnormal condition has affected one of them.

4.3.2 Failover

In the example shown in Figure 4-3, CEC 0 has failed. CEC 1 needs to take over all of the CEC 0 functions. Because the RAID arrays are on Fibre Channel Loops that reach both central electrical complexes, they can still be accessed through the Device Adapters owned by CEC 1. See 4.6.1, “RAID configurations” on page 89 for more information about the Fibre Channel Loops.

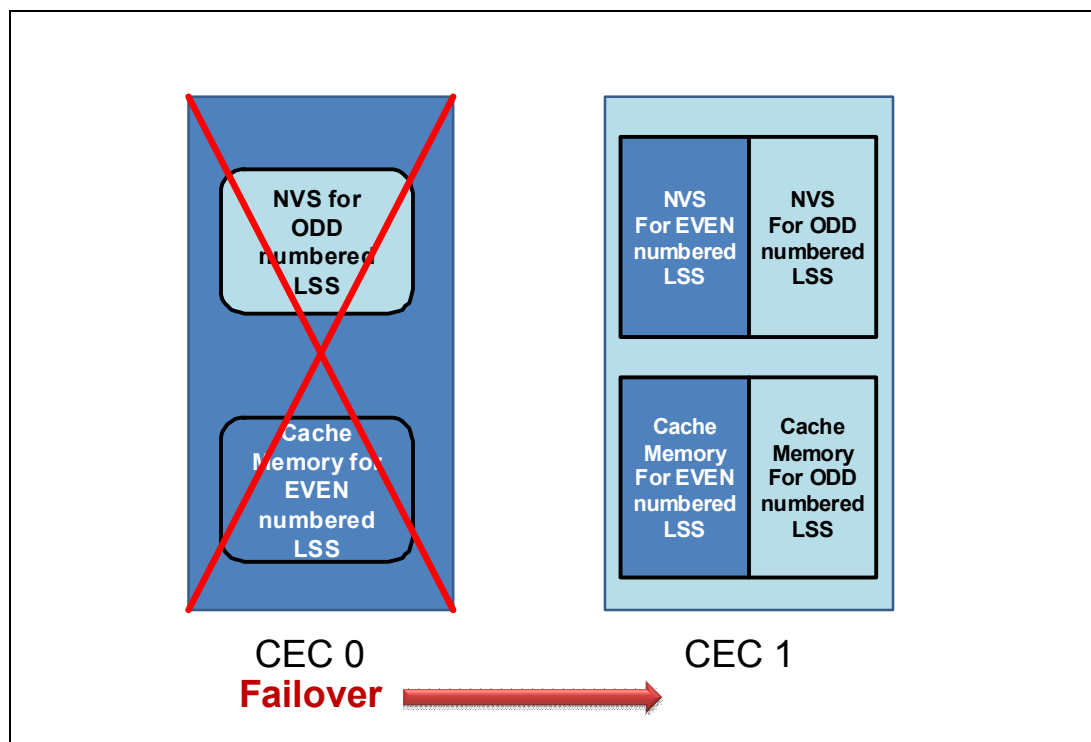


Figure 4-3 CEC 0 failover to CEC 1

At the moment of failure, CEC 1 has a backup copy of the CEC 0 write data in its own NVS. From a data integrity perspective, the concern is for the backup copy of the CEC 1 write data, which was in the NVS of CEC 0 when it failed. Because the DS8000 now has only one copy of that data (active in the cache memory of CEC 1), it will perform the following steps:

1. CEC 1 destages the contents of its NVS (the CEC 0 write data) to the disk subsystem. However, before the actual destage and at the beginning of the failover:
 - a. The working central electrical complex starts by preserving the data in cache that was backed by the failed central electrical complex NVS. If a reboot of the single working central electrical complex occurs before the cache data had been destaged, the write data remains available for subsequent destaging.
 - b. In addition, the existing data in cache (for which there is still only a single volatile copy) is added to the NVS so that it remains available if the attempt to destage fails or a server reboot occurs. This functionality is limited so that it cannot consume more than 85% of NVS space.
2. The NVS and cache of CEC 1 are divided in two, half for the odd LSSs and half for the even LSSs.
3. CEC 1 now begins processing the I/O for all the LSSs.

This entire process is known as a *failover*. After failover, the DS8000 now operates as shown in Figure 4-3 on page 81. CEC 1 now owns all the LSSs, which means all reads and writes will be serviced by CEC 1. The NVS inside CEC 1 is now used for both odd and even LSSs. The entire failover process should be transparent to the attached hosts.

The DS8000 can continue to operate in this state indefinitely. There has not been any loss of functionality, but there has been a loss of redundancy. Any critical failure in the working central electrical complex would render the DS8000 unable to serve I/O for the arrays, so IBM support should begin work right away to determine the scope of the failure and to build an action plan to restore the failed central electrical complex to an operational state.

4.3.3 Failback

The *failback* process always begins automatically as soon as the DS8000 microcode determines that the failed central electrical complex has been *resumed* to an operational state. If the failure was relatively minor and recoverable by the operating system or DS8000 microcode, then the resume action will be initiated by the software. If there was a service action with hardware components replaced, then the IBM service representative or remote support will resume the failed central electrical complex.

For this example where CEC 0 has failed, we should now assume that CEC 0 has been repaired and has been resumed. The failback begins with CEC 1 starting to use the NVS in CEC 0 again, and the ownership of the even LSSs being transferred back to CEC 0. Normal I/O processing with both central electrical complexes operational then resumes. Just like the failover process, the failback process is transparent to the attached hosts.

In general, recovery actions (failover/failback) on the DS8000 do not impact I/O operation latency by more than 15 seconds. With certain limitations on configurations and advanced functions, this impact to latency is often limited to just 8 seconds or less.

Important: On logical volumes that are not configured with RAID 10 storage, there are certain RAID related recoveries that cause latency impacts in excess of 15 seconds.

If you have real-time response requirements in this area, contact IBM to determine the latest information about how to manage your storage to meet your requirements.

4.3.4 NVS and power outages

During normal operation, the DS8000 preserves write data by storing a duplicate in the NVS of the alternate central electrical complex. To ensure that this write data is not lost due to a power event, the DS8000 contains battery backup units (BBUs). The single purpose of the BBUs is to preserve the NVS area of central electrical complex memory in the event of a complete loss of input power to the DS8000. The design is to not move the data from NVS to the disk arrays. Instead, each central electrical complex has dual internal disks that are available to store the contents of NVS.

Important: Unless the extended power line disturbance feature (ePLD) has been purchased, the BBUs do not keep the storage disks in operation. They keep the central electrical complexes and I/O enclosures operable long enough to write NVS contents to internal central electrical complex hard disks.

Should any frame lose AC input to both PPSs, the central electrical complexes would be informed that they are running on batteries and immediately begin a shutdown procedure. It is during this shutdown that the entire contents of NVS memory are written to the central electrical complex hard drives so that the data will be available for destaging after the central electrical complexes are operational again. If power is lost to a single PPS, the ability of the other power supply to keep all batteries charged is not impacted, so the central electrical complexes would remain online.

If all the batteries were to fail (which is extremely unlikely because the batteries are in an N+1 redundant configuration), the DS8000 would lose this NVS protection and consequently would take all central electrical complexes offline because reliability and availability of host data are compromised.

The following sections show the steps followed in the event of complete power interruption.

Power loss

When an on-battery condition shutdown begins, the following events occur:

1. All host adapter I/O is blocked.
2. Each central electrical complex begins copying its NVS data to internal disk (not the storage DDMs). For each central electrical complex, two copies are made of the NVS data.
3. When the copy process is complete, each central electrical complex shuts down.
4. When shutdown in each central electrical complex is complete (or a timer expires), the DS8000 is powered down.

Power restored

When power is restored to the DS8000, the following events occur:

1. The central electrical complexes power on and perform power on self tests and PHYP functions.
2. Each central electrical complex then begins boot up (IML).
3. At a certain stage in the boot process, the central electrical complex detects NVS data on its internal disks and begins to destage it to the storage DDMs.

4. When the battery units reach a certain level of charge, the central electrical complexes come online and begin to process host I/O.

Battery charging

In many cases, sufficient charging will occur during the power on self test, operating system boot, and microcode boot. However, if a complete discharge of the batteries has occurred, which can happen if multiple power outages occur in a short period of time, then recharging might take up to two hours.

Tip: The central electrical complexes will not come online (process host I/O) until the batteries are sufficiently charged to handle at least one outage.

4.4 Data flow in DS8000

One of the significant hardware changes for the DS8700 was the way in which host I/O was brought into the storage unit. The DS8800 continues this design for the I/O enclosures, which house the device adapters and host adapters. Connectivity between the central electrical complexes and the I/O enclosures was also improved. These changes use the many strengths of the PCI Express architecture.

See 3.3.3, “Peripheral Component Interconnect Express (PCI Express)” on page 43, for more information about this topic.

You can also discover more about PCI Express at the following URL:

<http://www.redbooks.ibm.com/Redbooks.nsf/RedbookAbstracts/tips0456.html?Open>

4.4.1 I/O enclosures

The DS8800 I/O enclosure is a design introduced in the DS8700. The older DS8000 I/O enclosure consisted of multiple parts that required removal of the bay and disassembly for service. In the DS8700 and DS8800, the switch card can be replaced without removing the I/O adapters, reducing time and effort in servicing the I/O enclosure. As shown in Figure 4-1 on page 78, each central electrical complex is connected to all four I/O enclosures (base frame) or all eight I/O enclosures (expansion frame installed) through PCI Express cables. This makes each I/O enclosure an extension of each server.

The DS8000 I/O enclosures use hot-swap adapters with PCI Express connections. These adapters are replaceable concurrently. Each slot can be independently powered off for concurrent replacement of a failed adapter, installation of a new adapter, or removal of an old one.

In addition, each I/O enclosure has N+1 power and cooling in the form of two power supplies with integrated fans. The power supplies can be concurrently replaced and a single power supply is capable of supplying DC power to the whole I/O enclosure.

4.4.2 Host connections

Each DS8000 Fibre Channel host adapter card provides four or eight ports (eight ports is DS8800 only) for connection either directly to a host, or to a Fibre Channel SAN switch.

Single or multiple path

In DS8000, the host adapters are shared between the central electrical complexes. To illustrate this concept, Figure 4-4 shows a potential machine configuration. In this example, two I/O enclosures are shown. Each I/O enclosure has a pair of Fibre Channel host adapters. If a host only has a single path to a DS8000, as shown in Figure 4-4, it would still be able to access volumes belonging to all LSSs because the host adapter will direct the I/O to the correct central electrical complex. However, if an error were to occur on the host adapter (HA), host port (HP), or I/O enclosure, or in the SAN, then all connectivity would be lost. Clearly, the host bus adapter (HBA) in the attached host is also a single point of failure.

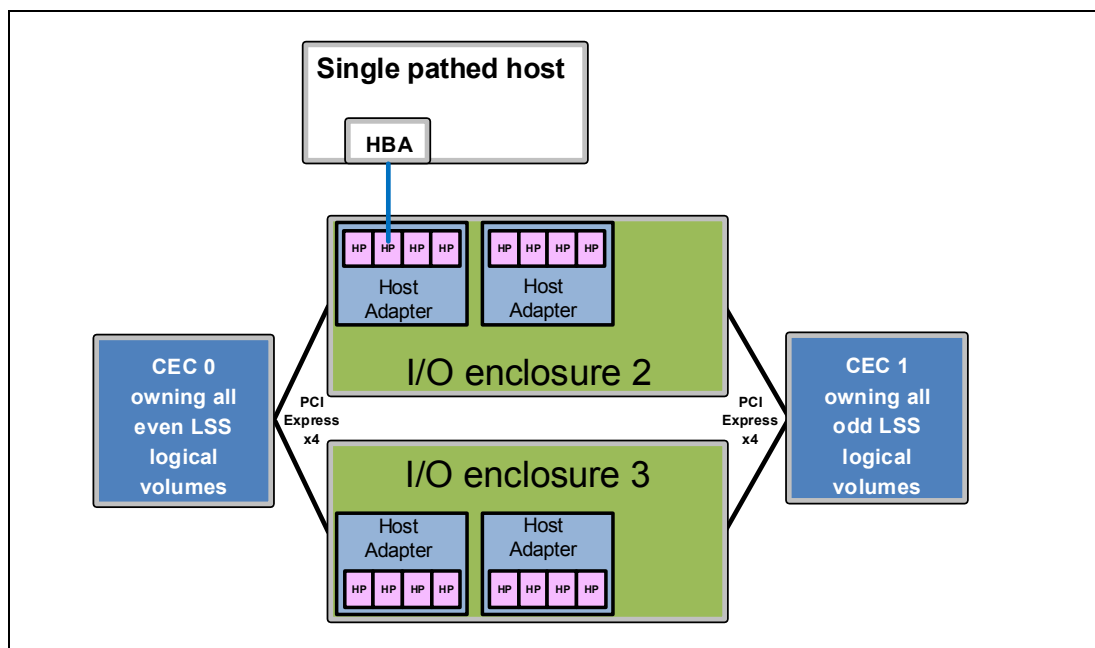


Figure 4-4 A single-path host connection

A more robust design is shown in Figure 4-5 where the host is attached to separate Fibre Channel host adapters in separate I/O enclosures. This is also important because during a microcode update, a host adapter port might need to be taken offline. This configuration allows the host to survive a hardware failure on any component on either path.

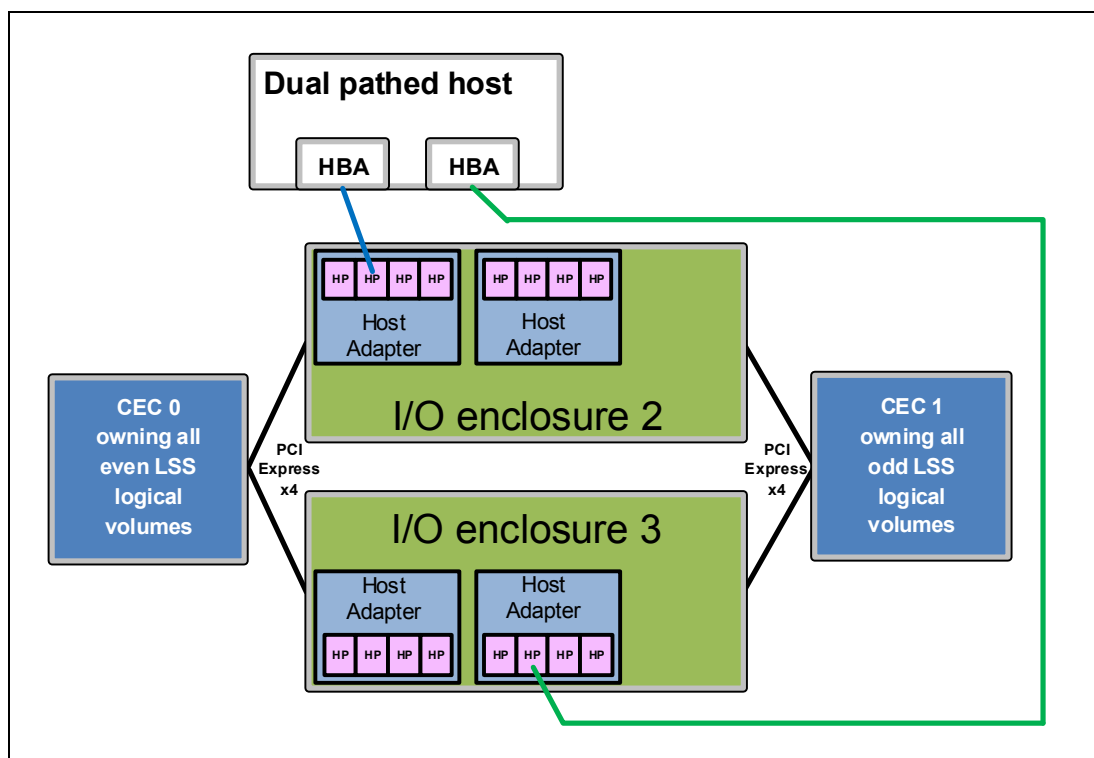


Figure 4-5 A dual-path host connection

Important: Best practice is that hosts accessing the DS8000 have at least two connections to host ports on separate host adapters in separate I/O enclosures.

SAN/FICON switches

Because a large number of hosts can be connected to the DS8000, each using multiple paths, the number of host adapter ports that are available in the DS8000 might not be sufficient to accommodate all the connections. The solution to this problem is the use of SAN switches or directors to switch logical connections from multiple hosts. In a System z environment, you will need to select a SAN switch or director that also supports FICON.

A logic or power failure in a switch or director can interrupt communication between hosts and the DS8000. Provide more than one switch or director to ensure continued availability. Ports from two separate host adapters in two separate I/O enclosures should be configured to go through each of two directors. The complete failure of either director leaves half the paths still operating.

Multipathing software

Each attached host operating system requires a mechanism to allow it to manage multiple paths to the same device, and to preferably load balance these requests. Also, when a failure occurs on one redundant path, then the attached host must have a mechanism to allow it to detect that one path is gone and route all I/O requests for those logical devices to an alternative path. Finally, it should be able to detect when the path has been restored so that the I/O can again be load-balanced. The mechanism that will be used varies by attached host operating system and environment, as detailed in the next two sections.

Open systems and SDD

In the majority of open systems environments, the Subsystem Device Driver (SDD) is useful to manage both path failover and preferred path determination. SDD is a software product that IBM supplies as an option with the DS8000 at no additional fee.

SDD provides availability through automatic I/O path failover. If a failure occurs in the data path between the host and the DS8000, SDD automatically switches the I/O to another path. SDD will also automatically set the failed path back online after a repair is made. SDD also improves performance by sharing I/O operations to a common disk over multiple active paths to distribute and balance the I/O workload.

SDD is not available for every supported operating system. Refer to *IBM System Storage DS8000 Host Systems Attachment Guide*, SC26-7917, and the interoperability website for guidance about which multipathing software might be required. The IBM System Storage Interoperability Center (SSIC), found at the following URL:

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

For more information about the SDD, refer to *IBM System Storage DS8000: Host attachment and Interoperability*, SG24-8887.

System z

In the System z environment, normal practice is to provide multiple paths from each host to a disk subsystem. Typically, four paths are installed. The channels in each host that can access each logical control unit (LCU) in the DS8000 are defined in the hardware configuration definition (HCD) or I/O configuration data set (IOCDS) for that host. Dynamic Path Selection (DPS) allows the channel subsystem to select any available (non-busy) path to initiate an operation to the disk subsystem. Dynamic Path Reconnect (DPR) allows the DS8000 to select any available path to a host to reconnect and resume a disconnected operation, for example, to transfer data after disconnection due to a cache miss.

These functions are part of the System z architecture and are managed by the channel subsystem on the host and the DS8000.

A physical FICON path is established when the DS8000 port sees light on the fiber (for example, a cable is plugged in to a DS8000 host adapter, a processor or the DS8000 is powered on, or a path is configured online by z/OS). At this time, logical paths are established through the port between the host, and some or all of the LCUs in the DS8000 controlled by the HCD definition for that host. This happens for each physical path between a System z CPU and the DS8000. There can be multiple system images in a CPU. Logical paths are established for each system image. The DS8000 then knows which paths can be used to communicate between each LCU and each host.

Control Unit Initiated Reconfiguration

Control Unit Initiated Reconfiguration (CUIR) prevents loss of access to volumes in System z environments due to incorrect path handling. This function automates channel path management in System z environments in support of selected DS8000 service actions.

CUIR is available for the DS8000 when operated in the z/OS and z/VM® environments. CUIR provides automatic channel path vary on and vary off actions to minimize manual operator intervention during selected DS8000 service actions.

CUIR also allows the DS8000 to request that all attached system images set all paths required for a particular service action to the offline state. System images with the appropriate level of software support respond to such requests by varying off the affected paths, and either notifying the DS8000 subsystem that the paths are offline, or that it cannot take the paths offline. CUIR reduces manual operator intervention and the possibility of human error during maintenance actions, while at the same time reducing the time required for the maintenance. This is particularly useful in environments where there are many z/OS or z/VM systems attached to a DS8000.

4.4.3 Metadata checks

When application data enters the DS8000, special codes or *metadata*, also known as *redundancy checks*, are appended to that data. This metadata remains associated with the application data as it is transferred throughout the DS8000. The metadata is checked by various internal components to validate the integrity of the data as it moves throughout the disk system. It is also checked by the DS8000 before the data is sent to the host in response to a read I/O request. Further, the metadata also contains information used as an additional level of verification to confirm that the data returned to the host is coming from the desired location on the disk.

4.5 RAS on the HMC

The HMC is used to perform configuration, management, and maintenance activities on the DS8000. It can be ordered to be located either physically inside the base frame (primary HMC) or external for mounting in a client-supplied rack (secondary HMC). The DS8000 HMC is able to work with IPv4, IPv6, or a combination of both IP standards. For further information, see 8.3, “Network connectivity planning” on page 212.

Important: The HMC described here is the Storage HMC, not to be confused with the SSPC console. SSPC is described in 3.9, “System Storage Productivity Center” on page 68.

If the HMC is not operational, then it is not possible to perform maintenance, power the DS8000 up or down, perform modifications to the logical configuration, or perform Copy Services tasks, such as the establishment of FlashCopies using the DSCLI or DS GUI. Generally, order two management consoles to act as a redundant pair. Alternatively, if Tivoli Storage Productivity Center for Replication (TPC-R) is used, Copy Services tasks can be managed by that tool if the HMC is unavailable.

Tip: The preceding alternative is only available if you have purchased and configured the TPC-R management solution.

4.5.1 Hardware

The DS8000 ships with a mobile computer HMC (Lenovo ThinkPad Model T510). Best practice is to order a second HMC to provide redundancy. The second HMC is external to the DS8000 rack(s). For more information about the HMC and network connections, see 9.1.1, “Storage Hardware Management Console hardware” on page 226.

4.5.2 Microcode updates

The DS8000 contains many discrete redundant components. Most of these components have firmware that can be updated. This includes the primary power supplies (PPS), Fibre Channel Interface Control cards (FCIC), device adapters, and host adapters. Both DS8000 central electrical complexes also have an operating system (AIX) and Licensed Machine Code (LMC) that can be updated. As IBM continues to develop and improve the DS8000, new releases of firmware and licensed machine code become available that offer improvements in both function and reliability.

For a detailed discussion about microcode updates, see Chapter 15, “Licensed machine code” on page 401.

Concurrent code updates

The architecture of the DS8000 allows for concurrent code updates. This is achieved by using the redundant design of the DS8000. In general, redundancy is lost for a short period as each component in a redundant pair is updated.

4.5.3 Call Home and Remote Support

Call Home is the capability of the HMC to contact IBM support services to report a problem. This is referred to as *Call Home for service*. The HMC will also provide machine-reported product data (MRPD) to IBM by way of the Call Home facility.

IBM Service personnel located outside of the client facility log in to the HMC to provide remote service and support. Remote support and the Call Home option are described in detail in Chapter 17, “Remote support” on page 425.

4.6 RAS on the disk subsystem

The reason for the DS8000's existence is to safely store and retrieve large amounts of data. Redundant Array of Independent Disks (RAID) is an industry-wide implementation of methods to store data on multiple physical disks to enhance the availability of that data. There are many variants of RAID in use today. The DS8000 supports RAID 5, RAID 6, and RAID 10. It does not support the non-RAID configuration of disks better known as JBOD (just a bunch of disks).

Important: Solid-State Drives support only RAID 5.

4.6.1 RAID configurations

The following RAID configurations are possible for the DS8000:

- ▶ 6+P RAID 5 configuration: The array consists of six data drives and one parity drive. The remaining drive on the array site is used as a spare.
- ▶ 7+P RAID 5 configuration: The array consists of seven data drives and one parity drive.

- 5+P+Q RAID 6 configuration: The array consists of five data drives and two parity drives. The remaining drive on the array site is used as a spare.
- 6+P+Q RAID 6 configuration: The array consists of six data drives and two parity drives.
- 3+3 RAID 10 configuration: The array consists of three data drives that are mirrored to three copy drives. Two drives on the array site are used as spares.
- 4+4 RAID 10 configuration: The array consists of four data drives that are mirrored to four copy drives.

For information regarding the effective capacity of these configurations, see Table 8-11 on page 221.

4.6.2 Disk path redundancy

Each DDM in the DS8000 is attached to two Fibre Channel switches. These switches are built into the disk enclosure controller cards. Figure 4-6 shows the redundancy features of the DS8000 switched Fibre Channel disk architecture.

Each disk has two separate connections to the backplane. This allows it to be simultaneously attached to both FC switches. If either disk enclosure controller card is removed from the enclosure, the switch that is included in that card is also removed. However, the FC switch in the remaining controller card retains the ability to communicate with all the disks and both device adapters (DAs) in a pair. Equally, each DA has a path to each switch, so it also can tolerate the loss of a single path. If both paths from one DA fail, it cannot access the switches. However, the partner DA retains connection.

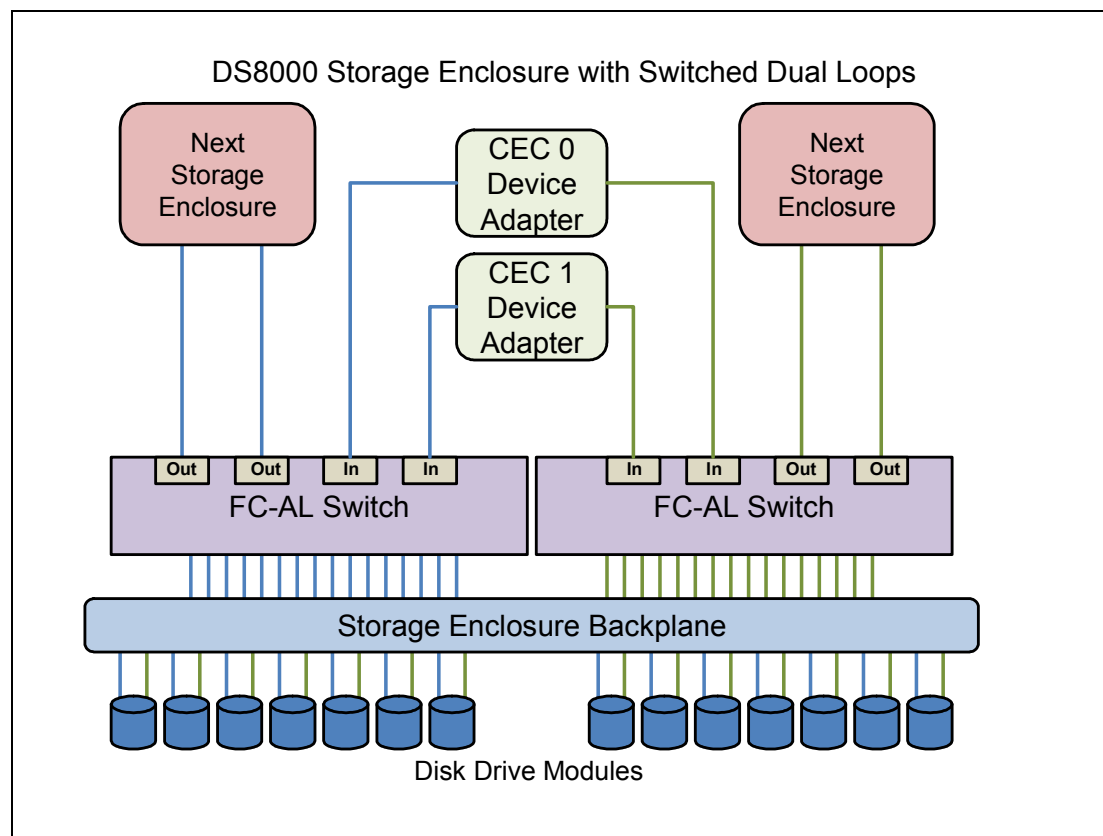


Figure 4-6 Switched disk path connections

Figure 4-6 on page 90 also shows the connection paths to the neighboring Storage Enclosures. Because expansion is done in this linear fashion, the addition of more enclosures is completely nondisruptive.

See 3.6, “Disk subsystem” on page 58 for more information about the disk subsystem of the DS8000.

4.6.3 Predictive Failure Analysis

The drives used in the DS8000 incorporate Predictive Failure Analysis (PFA) and can anticipate certain forms of failures by keeping internal statistics of read and write errors. If the error rates exceed predetermined threshold values, the drive will be nominated for replacement. Because the drive has not yet failed, data can be copied directly to a spare drive. This avoids using RAID recovery to reconstruct all of the data onto the spare drive.

4.6.4 Disk scrubbing

The DS8000 will periodically read all sectors on a disk. This is designed to occur without any interference with application performance. If error correcting code (ECC)-correctable bad bits are identified, the bits are corrected immediately by the DS8000. This reduces the possibility of multiple bad bits accumulating in a sector beyond the ability of ECC to correct them. If a sector contains data that is beyond ECC's ability to correct, RAID is used to regenerate the data and write a new copy onto a spare sector of the disk. This scrubbing process applies to both array members and spare DDMs.

4.6.5 RAID 5 overview

The DS8000 series supports RAID 5 arrays. RAID 5 is a method of spreading volume data plus parity data across multiple disk drives. RAID 5 provides faster performance by striping data across a defined set of DDMs. Data protection is provided by the generation of parity information for every stripe of data. If an array member fails, its contents can be regenerated using the parity data.

RAID 5 implementation in DS8000

In a DS8000, a RAID 5 array built on one array site will contain either seven or eight disks, depending on whether the array site is supplying a spare. A seven-disk array effectively uses one disk for parity, so it is referred to as a 6+P array (where the P stands for parity). The reason only seven disks are available to a 6+P array is that the eighth disk in the array site used to build the array was used as a spare. We refer to this as a 6+P+S array site (where the S stands for spare). An 8-disk array also effectively uses one disk for parity, so it is referred to as a 7+P array.

Drive failure with RAID 5

When a disk drive module fails in a RAID 5 array, the device adapter starts an operation to reconstruct the data that was on the failed drive onto one of the spare drives. The spare that is used will be chosen based on a smart algorithm that looks at the location of the spares and the size and location of the failed DDM. The rebuild is performed by reading the corresponding data and parity in each stripe from the remaining drives in the array, performing an exclusive-OR operation to recreate the data, and then writing this data to the spare drive.

While this data reconstruction is going on, the device adapter can still service read and write requests to the array from the hosts. There might be degradation in performance while the sparing operation is in progress because some DA and switched network resources are used to do the reconstruction. Due to the switch-based architecture, this effect will be minimal. Additionally, any read requests for data on the failed drive require data to be read from the other drives in the array, and then the DA performs an operation to reconstruct the data.

Performance of the RAID 5 array returns to normal when the data reconstruction onto the spare device completes. The time taken for sparing can vary, depending on the size of the failed DDM and the workload on the array, the switched network, and the DA. The use of arrays across loops (AAL) both speeds up rebuild time and decreases the impact of a rebuild.

4.6.6 RAID 6 overview

The DS8000 supports RAID 6 protection. RAID 6 presents an efficient method of data protection in case of double disk errors, such as two drive failures, two coincident medium errors, or a drive failure and a medium error. RAID 6 protection provides more fault tolerance than RAID 5 in the case of disk failures and uses less raw disk capacity than RAID 10.

RAID 6 allows for additional fault tolerance by using a second independent distributed parity scheme (dual parity). Data is striped on a block level across a set of drives, similar to RAID 5 configurations, and a second set of parity is calculated and written across all the drives, as shown in Figure 4-7.

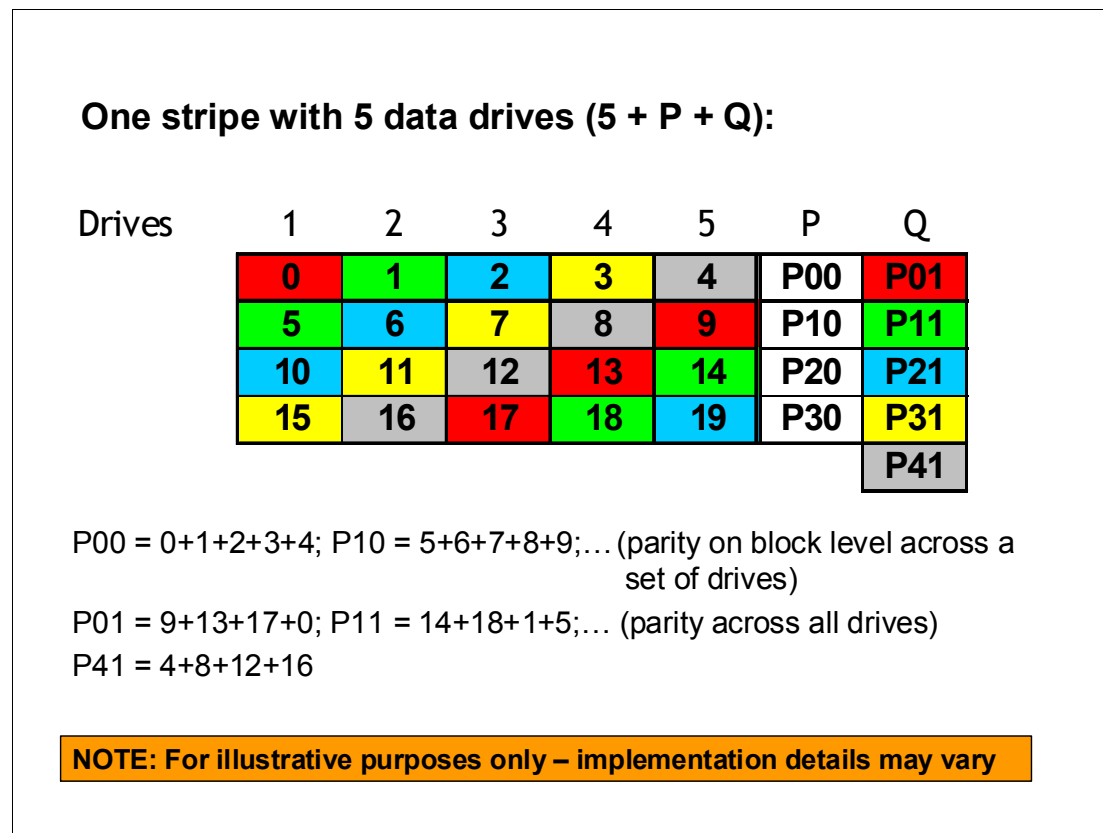


Figure 4-7 Illustration of one RAID 6 stripe

RAID 6 is best used in combination with large capacity disk drives, such as 2 TB SATA drives (DS8700), because they have a longer rebuild time. Comparing RAID 6 to RAID 5 performance gives about the same results on reads. For random writes, the throughput of a RAID 6 array is around only two thirds of a RAID 5, given the additional parity handling. Workload planning is especially important before implementing RAID 6 for write intensive applications, including copy services targets and FlashCopy SE repositories. Yet, when properly sized for the I/O demand, RAID 6 is a considerable reliability enhancement.

RAID 6 implementation in the DS8000

A RAID 6 array in one array site of a DS8000 can be built on either seven or eight disks:

- ▶ In a seven-disk array, two disks are always used for parity, and the eighth disk of the array site is needed as a spare. This kind of a RAID 6 array is hereafter referred to as a 5+P+Q+S array, where P and Q stand for parity and S stands for spare.
- ▶ A RAID 6 array, consisting of eight disks, will be built when all necessary spare drives are available. An eight-disk RAID 6 array also always uses two disks for parity, so it is referred to as a 6+P+Q array.

Drive failure with RAID 6

When a DDM fails in a RAID 6 array, the DA starts to reconstruct the data of the failing drive onto one of the available spare drives. A smart algorithm determines the location of the spare drive to be used, depending on the size and the location of the failed DDM. After the spare drive has replaced a failed one in a redundant array, the recalculation of the entire contents of the new drive is performed by reading the corresponding data and parity in each stripe from the remaining drives in the array and then writing this data to the spare drive.

During the rebuild of the data on the new drive, the DA can still handle I/O requests of the connected hosts to the affected array. Performance degradation could occur during the reconstruction because DAs and switched network resources are used to do the rebuild. Due to the switch-based architecture of the DS8000, this effect will be minimal. Additionally, any read requests for data on the failed drive require data to be read from the other drives in the array, and then the DA performs an operation to reconstruct the data. Any subsequent failure during the reconstruction within the same array (second drive failure, second coincident medium errors, or a drive failure and a medium error) can be recovered without loss of data.

Performance of the RAID 6 array returns to normal when the data reconstruction, on the spare device, has completed. The rebuild time will vary, depending on the size of the failed DDM and the workload on the array and the DA. The completion time is comparable to a RAID 5 rebuild, but slower than rebuilding a RAID 10 array in the case of a single drive failure.

4.6.7 RAID 10 overview

RAID 10 provides high availability by combining features of RAID 0 and RAID 1. RAID 0 optimizes performance by striping volume data across multiple disk drives at a time. RAID 1 provides disk mirroring, which duplicates data between two disk drives. By combining the features of RAID 0 and RAID 1, RAID 10 provides a second optimization for fault tolerance. Data is striped across half of the disk drives in the RAID 1 array. The same data is also striped across the other half of the array, creating a mirror. Access to data is preserved if one disk in each mirrored pair remains available. RAID 10 offers faster data reads and writes than RAID 5 because it does not need to manage parity. However, with half of the DDMs in the group used for data and the other half to mirror that data, RAID 10 disk groups have less capacity than RAID 5 disk groups.

RAID 10 is not as commonly used as RAID 5, mainly because more raw disk capacity is needed for every gigabyte of effective capacity. A typical area of operation for RAID 10 are workloads with a high random write ratio.

RAID 10 implementation in DS8000

In the DS8000, the RAID 10 implementation is achieved by using either six or eight DDMs. If spares need to be allocated on the array site, then six DDMs are used to make a three-disk RAID 0 array, which is then mirrored. If spares do not need to be allocated, then eight DDMs are used to make a four-disk RAID 0 array, which is then mirrored.

Drive failure with RAID 10

When a DDM fails in a RAID 10 array, the DA starts an operation to reconstruct the data from the failed drive onto one of the hot spare drives. The spare that is used will be chosen based on a smart algorithm that looks at the location of the spares and the size and location of the failed DDM. Remember a RAID 10 array is effectively a RAID 0 array that is mirrored. Thus, when a drive fails in one of the RAID 0 arrays, we can rebuild the failed drive by reading the data from the equivalent drive in the other RAID 0 array.

While this data reconstruction is going on, the DA can still service read and write requests to the array from the hosts. There might be degradation in performance while the sparing operation is in progress because DA and switched network resources are used to do the reconstruction. Due to the switch-based architecture of the DS8000, this effect will be minimal. Read requests for data on the failed drive should not be affected because they can all be directed to the good RAID 1 array.

Write operations will not be affected. Performance of the RAID 10 array returns to normal when the data reconstruction, onto the spare device, completes. The time taken for sparing can vary, depending on the size of the failed DDM and the workload on the array and the DA. In relation to a RAID 5, RAID 10 sparing completion time is a little faster. This is because rebuilding a RAID 5 6+P configuration requires six reads plus one parity operation for each write, whereas a RAID 10 3+3 configuration requires one read and one write (essentially a direct copy).

Arrays across loops and RAID 10

The DS8000 implements the concept of arrays across loops (AAL). With AAL, an array site is actually split into two halves. Half of the site is located on the first disk loop of a DA pair and the other half is located on the second disk loop of that DA pair. AAL is implemented primarily to maximize performance and it is used for all the RAID types in the DS8000. However, in RAID 10, we are able to take advantage of AAL to provide a higher level of redundancy. The DS8000 RAS code will deliberately ensure that one RAID 0 array is maintained on each of the two loops created by a DA pair. This means that in the extremely unlikely event of a complete loop outage, the DS8000 would not lose access to the RAID 10 array. This is because while one RAID 0 array is offline, the other remains available to service disk I/O. Figure 3-25 on page 63 shows a diagram of this strategy.

4.6.8 Spare creation

When the arrays are created on a DS8000, the microcode determines which array sites will contain spares. The first array sites on each DA pair that are assigned to arrays will contribute one or two spares (depending on the RAID option), until the DA pair has access to at least four spares, with two spares being placed on each loop.

A minimum of one spare is created for each array site assigned to an array until the following conditions are met:

- ▶ There are a minimum of four spares per DA pair.
- ▶ There are a minimum of four spares for the largest capacity array site on the DA pair.
- ▶ There are a minimum of two spares of capacity and RPM greater than or equal to the fastest array site of any given capacity on the DA pair.

Floating spares

The DS8000 implements a smart floating technique for spare DDMs. A *floating spare* is defined as follows: When a DDM fails and the data it contained is rebuilt onto a spare, then when the disk is replaced, the replacement disk becomes the spare. The data is not migrated to another DDM, such as the DDM in the original position the failed DDM occupied.

The DS8000 microcode takes this idea one step further. It might choose to allow the hot spare to remain where it has been *moved*, but it can instead choose to *migrate* the spare to a more optimum position. This will be done to better balance the spares across the DA pairs, the loops, and the disk enclosures. It might be preferable that a DDM that is currently in use as an array member is converted to a spare. In this case, the data on that DDM will be migrated in the background onto an existing spare. This process does not *fail* the disk that is being migrated, though it does reduce the number of available spares in the DS8000 until the migration process is complete.

The DS8000 uses this smart floating technique so that the larger or higher RPM DDMs are allocated as spares, which guarantees that a spare can provide at least the same capacity and performance as the replaced drive. If we were to rebuild the contents of a 450 GB DDM onto a 600 GB DDM, then approximately one-fourth of the 600 GB DDM will be wasted, because that space is not needed. When the failed 450 GB DDM is replaced with a new 450 GB DDM, the DS8000 microcode will most likely migrate the data back onto the recently replaced 450 GB DDM. When this process completes, the 450 GB DDM will rejoin the array and the 600 GB DDM will become the spare again.

Another example would be if we fail a 146 GB 15K RPM DDM onto a 600 GB 10K RPM DDM. The data has now moved to a slower DDM and is wasting a lot of space. This means the array will have a mix of RPMs, which is not desirable. When the failed disk is replaced, the replacement will be the same type as the failed 15K RPM disk. Again, a smart migration of the data will be performed after suitable spares have become available.

Hot pluggable DDMs

Replacement of a failed drive does not affect the operation of the DS8000 because the drives are fully hot pluggable. Each disk plugs into a switch, so there is no loop break associated with the removal or replacement of a disk. In addition, there is no potentially disruptive loop initialization process.

Overconfiguration of spares

The DDM sparing policies support the overconfiguration of spares. This possibility might be of interest to certain installations because it allows the repair of some DDM failures to be deferred until a later repair action is required.

4.7 RAS on the power subsystem

The DS8000 has completely redundant power and cooling. Every power supply and cooling fan in the DS8000 operates in what is known as N+1 mode. This means that there is always at least one more power supply, cooling fan, or battery than is required for normal operation.

4.7.1 Components

This section will review the power subsystem components of the DS8700 and DS8800.

Primary power supplies

Each frame has two primary power supplies (PPSs). Each PPS produces voltages for two separate areas of the machine.

For DS8700, 208V is produced to be supplied to each I/O enclosure and each processor complex. This voltage is distributed by each supply onto two redundant power buses.

12V and 5V are produced by modules in the PPS, which is then supplied to the disk enclosures.

For DS8800, 208V is produced to be supplied to each I/O enclosure, and each processor complex and to the disk enclosures. This voltage is distributed by each PPS to Power Distribution Units (PDUs). The PDUs distribute this voltage to the central electrical complexes, I/O enclosures, and the disk enclosures.

With the introduction of small form factor (SFF) disk enclosures on DS8800, certain changes have been made to the PPS to support the disk enclosure power requirements. The 5V/12V DDM power modules in the PPS (DS8700) is replaced with two 208V modules. The disk enclosures use 208V input and provides 5V/12V for the DDMs.

If either PPS fails, the other can continue to supply all required voltage to all power buses in that frame. The PPSs can be replaced concurrently.

Important: If you install the DS8000 so that both primary power supplies are attached to the same circuit breaker or the same switchboard, the DS8000 will not be well protected from external power failures. This is a common cause of unplanned outages.

Battery backup units

Each frame with I/O enclosures, or every frame if the extended power line disturbance (ePLD) feature is installed, will have battery backup units (BBUs). Each BBU can be replaced concurrently, if no more than one BBU is unavailable at any one time. The DS8000 BBUs have a planned working life of at least four years.

Power distribution unit (DS8800)

The power distribution units (PDUs) are used to distribute 208V from the PPS to the disk enclosures, central electrical complexes, and I/O enclosures. Each of the PDU modules can be replaced concurrently.

Disk enclosure power supply (DS8800)

The disk enclosure power supply unit provides 5V and 12V power for the DDMs, and houses the cooling fans for the disk enclosure. DDM cooling on the DS8800 is provided by these integrated fans in the disk enclosures. The fans draw air from the front of the frame, through the DDMs, and then move it out through the back of the frame. The entire rack cools from

front to back, enabling “hot and cold aisles”. There are redundant fans in each power supply unit and redundant power supply units in each disk enclosure. The disk enclosure power supply can be replaced concurrently.

Rack cooling fans (DS8700)

Each frame has a cooling fan plenum located above the disk enclosures. The fans in this plenum draw air from the front and rear of the frame, through the DDMs and then move it out through the top of the frame. There are multiple redundant fans in each enclosure. Each fan can be replaced concurrently.

Attention: Do not store any objects on top of a DS8700 frame that would block the airflow through the vent.

Rack Power Control card

The rack power control cards (RPCs) are part of the power management infrastructure of the DS8000. There are two RPC cards for redundancy. Each card can independently control power for the entire DS8000.

System Power Control Network

The System Power Control Network (SPCN) is used to control the power of the attached I/O subsystem. The SPCN monitors environmental components such as power, fans, and temperature. Environmental critical and noncritical conditions can generate Early Power-Off Warning (EPOW) events. Critical events trigger appropriate signals from the hardware to the affected components to prevent any data loss without operating system or firmware involvement. Non-critical environmental events are also logged and reported.

4.7.2 Line power loss

The DS8000 uses an area of server memory as nonvolatile storage (NVS). This area of memory is used to hold data that has not been written to the disk subsystem. If line power were to fail, where both PPS in a frame were to report a loss of AC input power, the DS8000 must take action to protect that data. See 4.3, “Central electrical complex failover and failback” on page 79 for a full explanation of the NVS Cache operation.

4.7.3 Line power fluctuation

The DS8000 primary frame contains battery backup units that are intended to protect modified data in the event of a complete power loss. If a power fluctuation occurs that causes a momentary interruption to power (often called a *brownout*), the DS8000 will tolerate this for approximately 30 ms. If the extended power line disturbance (ePLD) feature is not installed on the DS8000 system, then after that time, the DDMs will be powered off and the servers will begin copying the contents of NVS to the internal disks in the processor complexes. For many clients who use uninterruptible power supply (UPS) technology, this is not an issue. UPS-regulated power is in general reliable, so additional redundancy in the attached devices is often unnecessary.

Power line disturbance

If line power is not considered reliable, consider adding the extended power line disturbance (ePLD) feature. This feature adds two separate pieces of hardware to the DS8000:

- For each PPS in each frame of the DS8000, a booster module is added. When the BBUs supply power the primary power bus, this battery power is fed into the booster module, which then in turn keeps disk enclosure power present.

- Batteries will be added to expansion frames that did not already have them. The base frame and the first expansion frame will already have BBUs. Subsequent expansion frames do not get BBUs, until ePLD is installed.

With the addition of this hardware, the DS8000 will be able to run for 60 seconds on battery power before the central electrical complexes begin to copy NVS to internal disk and then shut down. This would allow for up to 60 second interruption to line power with no outage to the DS8000.

4.7.4 Power control

The DS8000 does not possess a white power switch to turn the storage unit off and on. All power sequencing is done using the Service Processor Control Network (SPCN) and RPCs. If you want to power the DS8000 off, you must do so by using the management tools provided by the HMC. If the HMC is not functional, it will not be possible to control the power sequencing of the DS8000 until the HMC function is restored. This is one of the benefits that is gained by purchasing a redundant HMC.

4.7.5 Emergency power off

Each DS8000 frame has an operator panel with three LEDs that show the line power status and the system fault indicator. The LEDs can be seen when the front door of the frame is closed. See Figure 4-8 on page 99 for an illustration of the DS8800 operator panel (DS8700 has a similar switch). On the side of the operator panel is an emergency power off (EPO) switch. This switch is red and is located inside the front door protecting the frame. It can only be seen when the front door is open. This switch is intended purely to remove power from the DS8000 in the following *extreme* cases:

- The DS8000 has developed a fault that is placing the environment at risk, such as a fire.
- The DS8000 is placing human life at risk, such as the electrocution of a person.

Apart from these two contingencies (which are highly unlikely), the EPO switch should never be used. The reason for this is that when the EPO switch is used, the battery protection for the NVS storage area is bypassed. Normally, if line power is lost, the DS8000 can use its internal batteries to destage the write data from NVS memory to persistent storage so that the data is preserved until power is restored. However, the EPO switch does not allow this destage process to happen and all NVS cache data is immediately lost. This will most likely result in data loss.



Figure 4-8 DS8000 EPO switch

If the DS8000 needs to be powered off for building maintenance or to relocate it, always use the HMC to shut it down properly.

4.8 RAS and Full Disk Encryption

The DS8700 and DS8800 can be ordered with DDMs that support Full Disk Encryption (FDE). FDE DDMs are available as 300 GB, 450 GB, and 600 GB (15K RPM) for DS8700 and capacities of 450 GB and 600 GB (10K RPM) for DS8800. The purpose of FDE drives is to encrypt all data at rest within the storage system for increased data integrity.

The DS8000 provides two important reliability, availability, and serviceability enhancements to Full Disk Encryption storage: deadlock recovery and support for dual-platform key servers.

For current considerations and best practices regarding DS8000 encryption, see *IBM Encrypted Storage Overview and Customer Requirements*, found at:

<http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP101479>

This link also includes the *IBM Notice for Storage Encryption*, which must be read by anyone acquiring an IBM storage device that includes encryption technology.

4.8.1 Deadlock recovery

The DS8000 family of storage servers with Full Disk Encryption drives can utilize a System z key server running the Tivoli Key Lifecycle Manager (TKLM) solution. A TKLM server provides a robust platform for managing the multiple levels of encryption keys needed for a secure storage operation. System z mainframes do not have local storage. Their operating system, applications, and application data are often stored on an enterprise-class storage server, such as a DS8000 storage subsystem.

Thus it becomes possible, due to a planning error or even the use of automatically-managed storage provisioning, for the System z TKLM server storage to end up residing on the DS8000 that is a client for encryption keys. After a power interruption event, the DS8000 becomes inoperable because it must retrieve the Data Key (DK) from the TKLM database on the

System z server. The TKLM database becomes inoperable because the System z server has its OS or application data on the DS8000. This represents a *deadlock* situation. Figure 4-9 depicts this scenario.

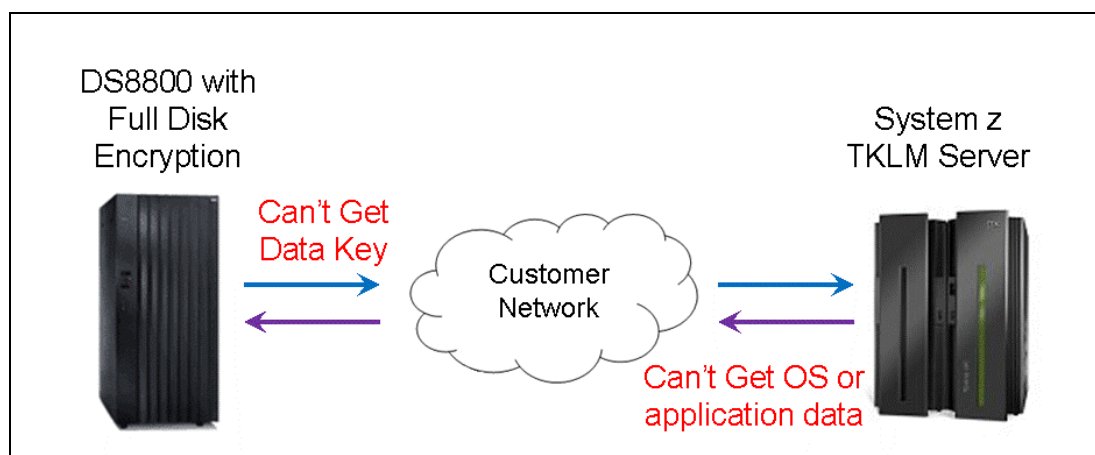


Figure 4-9 Deadlock scenario

The DS8000 mitigates this problem by implementing a *Recovery Key* (RK). The Recovery Key allows the DS8000 to decrypt the Group Key (GK) that it needs to come up to full operation. A new client role is defined in this process: the security administrator. The security administrator should be someone other than the storage administrator so that no single user can perform recovery key actions. Setting up the Recovery Key and using the Recovery Key to boot a DS8000 requires both people to take action. Use of a Recovery Key is entirely within your control. No IBM Service Representative needs to be involved. The DS8000 never stores a copy of the Recovery Key on the encrypted disks, and it is never included in any service data.

Refer to *IBM System Storage DS8700 Disk Encryption Implementation and Usage Guidelines*, REDP-4500, for a more complete review of the deadlock recovery process and further information about working with a Recovery Key.

Tip: Use the storage HMC to enter a Recovery Key. The Security Administrator and the Storage Administrator might need to be physically present at the DS8000 to perform the recovery.

4.8.2 Dual platform TKLM servers

The current DS8000 Full Disk Encryption solution requires the use of an IBM System x SUSE Linux-based Isolated Key Server (IKS), which operates in “clear key mode”. Clients have expressed a desire to run key servers that are hardware security module-based (HSM), which operate in “secure key mode”. Key servers like the IKS, which implement a clear key design, can import and export their public and private key pair to other key servers. Servers that implement secure key design can only import and export their public key to other key servers.

To meet this request, the DS8000 allows propagation of keys across two separate key server platforms. The current IKS is still supported to address the standing requirement for an isolated key server. Adding a z/OS Tivoli Key Lifecycle Manager (TKLM) Secure Key Mode server, which is common in Tape Storage environments, is currently supported by the DS8000.

After the key servers are set up, they will each have two public keys. They are each capable of generating and wrapping two symmetric keys for the DS8000. The DS8000 stores both

wrapped symmetric keys in the key repository. Now either key server is capable of unwrapping these keys upon a DS8000 retrieval exchange.

Refer to *IBM System Storage DS8700 Disk Encryption Implementation and Usage Guidelines*, REDP-4500, for more information regarding the dual-platform TKLM solution. Visit the following site for further information regarding planning and deployment of TKLM servers:

<http://www.ibm.com/developerworks/wikis/display/tivolidoccentral/Tivoli+Key+Lifecycle+Manager>

4.9 Other features

There are many more features of the DS8000 that enhance reliability, availability, and serviceability. Some of those are listed below.

4.9.1 Internal network

Each DS8000 base frame contains two Gigabit Ethernet switches to allow the creation of a fully redundant management (private) network. Each central electrical complex in the DS8000 has a connection to each switch. Each HMC also has a connection to each switch. This means that if a single Ethernet switch fails, then all traffic can successfully travel from the HMC(s) to other components in the storage unit using the alternate network.

There are also Ethernet connections for the FSP within each central electrical complex. If two DS8000 storage complexes are connected together, they will also use ports on the Ethernet switches. See 9.1.2, “Private Ethernet networks” on page 228 for more information about the DS8000 internal network.

Tip: Connections to the your network are made at the Ethernet patch panel at the rear of the machine. No network connection should ever be made to the DS8000 internal Ethernet switches.

4.9.2 Remote support

The DS8000 HMC has the ability to be accessed remotely by IBM Support personnel for many service actions. IBM Support can offload service data, change configuration settings, and initiate recovery actions over a remote connection. You decide which type of connection you want to allow for remote support. Options include:

- ▶ Modem only for access to the HMC command line
- ▶ VPN only for access to the HMC GUI (WebUI)
- ▶ Modem and VPN
- ▶ No access (secure account)

Remote support is a critical topic for Clients investing in the DS8000. See Chapter 17, “Remote support” on page 425 for a more thorough discussion of remote support operations. See Chapter 9, “DS8000 HMC planning and setup” on page 225 for more information about planning the connections needed for HMC installation.

4.9.3 Earthquake resistance

The Earthquake Resistance Kit is an optional seismic kit for stabilizing the storage unit rack, so that the rack complies with IBM earthquake resistance standards. It helps to prevent

human injury and increases the probability that the system will be available following the earthquake by limiting potential damage to critical system components, such as hard drives.

A storage unit frame with this optional seismic kit includes cross-braces on the front and rear of the frame that prevent the frame from twisting. Hardware at the bottom of the frame secures it to the floor. Depending on the flooring in your environment, specifically non-raised floors, installation of required floor mounting hardware might be disruptive.

This kit must be special ordered for the DS8000. Contact your IBM sales representative for further information.



Virtualization concepts

This chapter describes virtualization concepts as they apply to the IBM System Storage DS8000.

This chapter covers the following topics:

- ▶ Virtualization definition
- ▶ The abstraction layers for disk virtualization:
 - Array sites
 - Arrays
 - Ranks
 - Extent Pools
 - Dynamic Extent Pool merge
 - Track Space Efficient volumes
 - Logical subsystems (LSSs)
 - Volume access
 - Virtualization hierarchy summary
- ▶ Benefits of virtualization

5.1 Virtualization definition

In a fast-changing world, to react quickly to changing business conditions, IT infrastructure must allow for on-demand changes. Virtualization is key to an on-demand infrastructure. However, when talking about virtualization, many vendors are talking about different things.

For this chapter, the definition of *virtualization* is the abstraction process from the physical disk drives to a logical volume that is presented to hosts and servers in a way that they see it as though it were a physical disk.

5.2 The abstraction layers for disk virtualization

In this chapter, when talking about virtualization, we mean the process of preparing physical disk drives (DDMs) to become an entity that can be used by an operating system, which means we are talking about the creation of LUNs.

The DDMS are mounted in disk enclosures and connected in a switched FC topology, using a FC-AL protocol. The way DDMs are physically installed differs between DS8700 and DS8800:

- ▶ For the DS8700, DDMs are mounted in 16 DDM enclosures. You can order disk drives in groups of 8 or 16 drives of the same capacity and revolutions per minute (rpm). The options for 8-drive sets only apply for the 600 GB Solid-State Drives (SSDs).
- ▶ The DS8800 disks have a smaller form factor and are mounted in 24 DDM enclosures. Disk drives can be ordered in groups of 8 or 16 drives of the same capacity and rpm. The option for 8-drive sets only apply for the 300 GB Solid-State Drives (SSDs).

The disk drives can be accessed by a pair of device adapters. Each device adapter has four paths to the disk drives. One device interface from each device adapter is connected to a set of FC-AL devices so that either device adapter has access to any disk drive through two independent switched fabrics (the device adapters and switches are redundant).

Each device adapter has four ports, and because device adapters operate in pairs, there are eight ports or paths to the disk drives. All eight paths can operate concurrently and could access all disk drives on the attached fabric. In normal operation, however, disk drives are typically accessed by one device adapter. Which device adapter owns the disk is defined during the logical configuration process. This avoids any contention between the two device adapters for access to the disks.

Because of the switching design, each drive is in close reach of the device adapter, and some drives will require a few more hops through the Fibre Channel switch. Therefore, it is not really a loop but a switched FC-AL loop with the FC-AL addressing schema, that is, Arbitrated Loop Physical Addressing (AL-PA).

Figure 5-1 shows the physical layer on which virtualization is based.

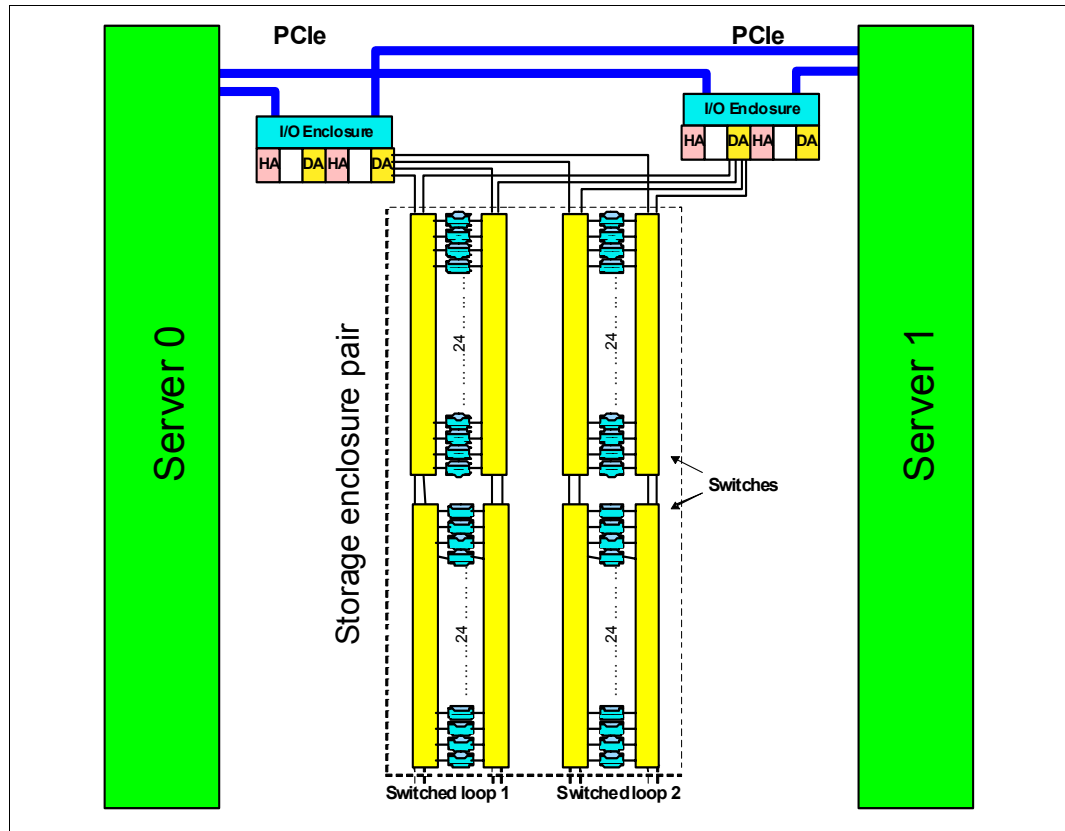


Figure 5-1 Physical layer as the base for virtualization

5.2.1 Array sites

An *array site* is a group of eight DDMs. Which DDMs are forming an array site is predetermined automatically by the DS8000. The DDMs selected can be from any location within the disk enclosures. Also note that there is no predetermined server affinity for array sites. The DDMs selected for an array site are chosen from two disk enclosures on different loops (Figure 5-2).

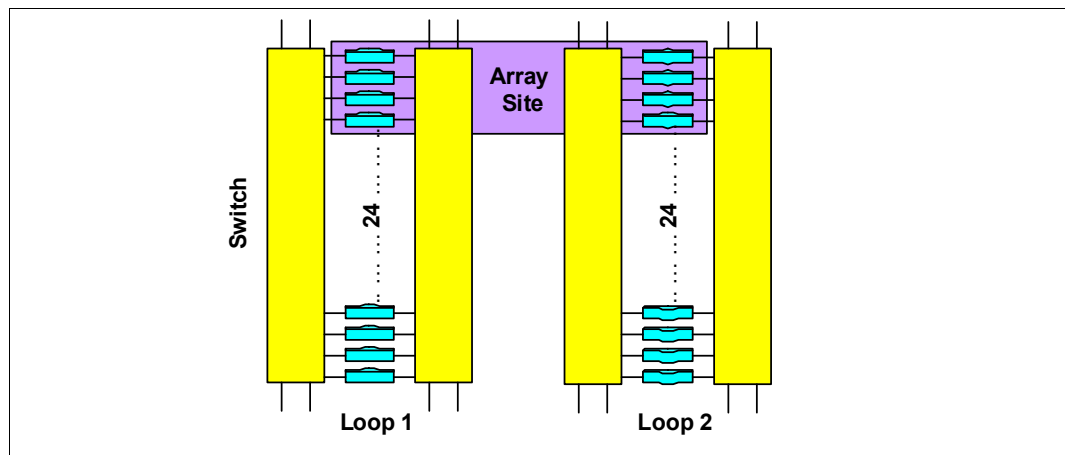


Figure 5-2 Array site

The DDMs in the array site are of the same DDM type, which means the same capacity and the same speed (rpm).

As you can see from Figure 5-2 on page 105, array sites span loops. Four DDMs are taken from loop 1 and another four DDMs from loop 2. Array sites are the building blocks used to define arrays.

5.2.2 Arrays

An *array* is created from one *array site*. Forming an array means defining it as a specific RAID type. The supported RAID types are RAID 5, RAID 6, and RAID 10 (see “RAID 5 implementation in DS8000” on page 91, “RAID 6 implementation in the DS8000” on page 93, and “RAID 10 implementation in DS8000” on page 94). For each array site, you can select a RAID type (remember that Solid-State Drives can only be configured as RAID 5). The process of selecting the RAID type for an array is also called defining an array.

Tip: In a DS8000 series implementation, one array is defined using one array site.

According to the DS8000 series sparing algorithm, from zero to two spares can be taken from the array site. This is discussed further in 4.6.8, “Spare creation” on page 94.

Figure 5-3 shows the creation of a RAID 5 array with one spare, also called a 6+P+S array (it has a capacity of 6 DDMs for data, capacity of one DDM for parity, and a spare drive). According to the RAID 5 rules, parity is distributed across all seven drives in this example.

On the right side in Figure 5-3, the terms D1, D2, D3, and so on stand for the set of data contained on one disk within a stripe on the array. If, for example, 1 GB of data is written, it is distributed across all the disks of the array.

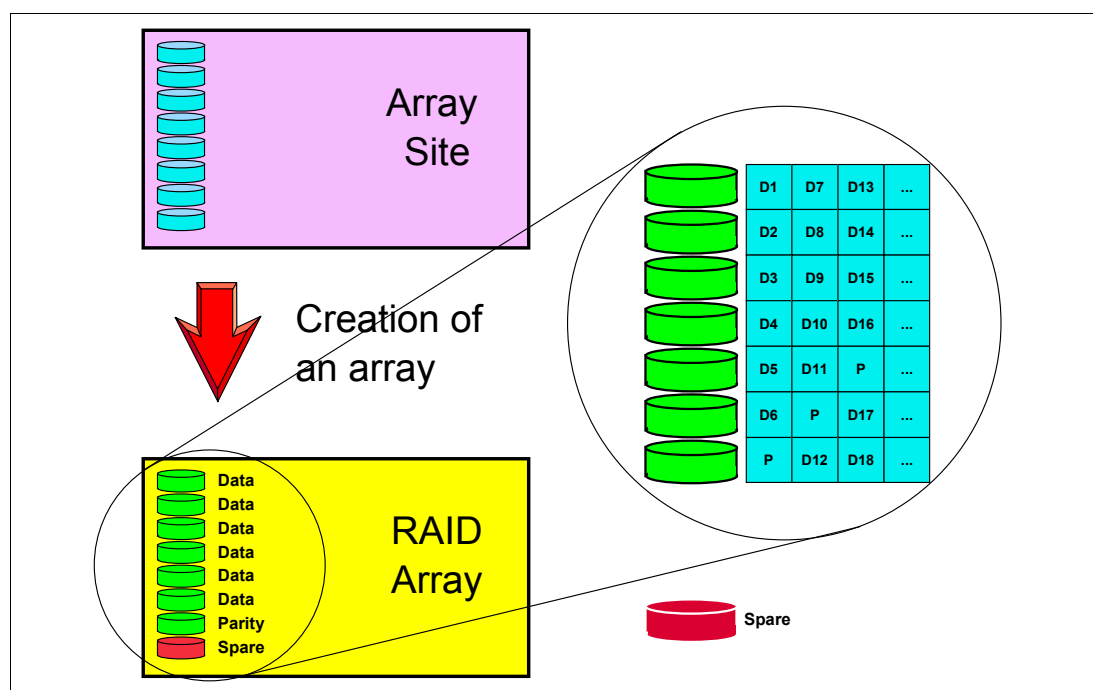


Figure 5-3 Creation of an array

So, an array is formed using one array site, and although the array could be accessed by each adapter of the device adapter pair, it is managed by one device adapter. You define which server is managing this array later on in the configuration path.

5.2.3 Ranks

In the DS8000 virtualization hierarchy, there is another logical construct called a *rank*. When defining a new rank, its name is chosen by the DS Storage Manager, for example, R1, R2, or R3, and so on. You have to add an array to a rank.

Tip: In the DS8000 implementation, a rank is built using just one array.

The available space on each rank will be divided into *extents*. The extents are the building blocks of the logical volumes. An extent is striped across all disks of an array as shown in Figure 5-4 on page 108 and indicated by the small squares in Figure 5-5 on page 110.

The process of forming a rank does two things:

- ▶ The array is formatted for either fixed block (FB) data for open systems or count key data (CKD) for System z data. This determines the size of the set of data contained on one disk within a stripe on the array.
- ▶ The capacity of the array is subdivided into equal-sized partitions, called *extents*. The extent size depends on the *extent type*, FB or CKD.

A FB rank has an extent size of 1 GB (more precisely, GiB, gibibyte, or binary gigabyte, being equal to 2^{30} bytes).

IBM System z users or administrators typically do not deal with gigabytes or gibibytes, and instead they think of storage in terms of the original 3390 volume sizes. A 3390 Model 3 is three times the size of a Model 1 and a Model 1 has 1113 cylinders, which is about 0.94 GB. The extent size of a CKD rank is one 3390 Model 1 or 1113 cylinders.

Figure 5-4 shows an example of an array that is formatted for FB data with 1 GB extents (the squares in the rank just indicate that the extent is composed of several blocks from separate DDMs).

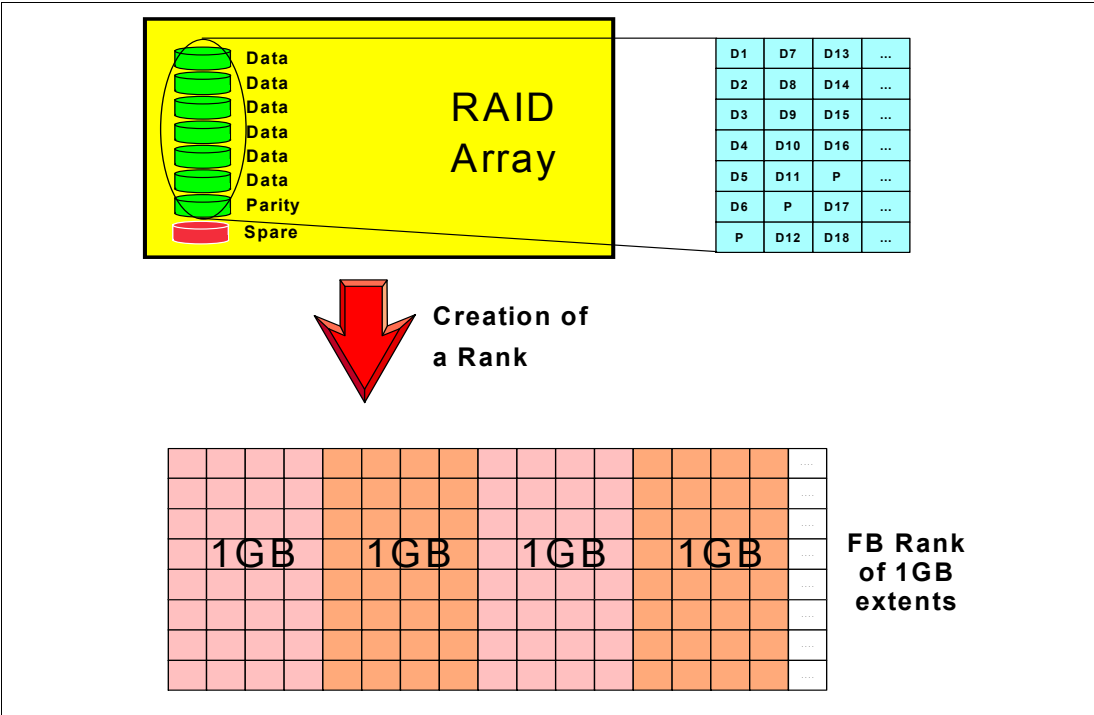


Figure 5-4 Forming an FB rank with 1 GB extents

It is still possible to define a CKD volume with a capacity that is an integral multiple of one cylinder or a fixed block LUN with a capacity that is an integral multiple of 128 logical blocks (64 KB). However, if the defined capacity is not an integral multiple of the capacity of one extent, the unused capacity in the last extent is wasted. For example, you could define a one cylinder CKD volume, but 1113 cylinders (1 extent) will be allocated and 1112 cylinders would be wasted.

Encryption group

A DS8000 series can be ordered with encryption capable disk drives. If you plan to use encryption before creating a rank, you must define an *encryption group*. For more information, refer to *IBM System Storage DS8700 Disk Encryption Implementation and Usage Guidelines*, REDP-4500. Currently, the DS8000 series supports only *one* encryption group. All ranks must be in this encryption group. The encryption group is an attribute of a rank. So, your choice is to encrypt everything or nothing. You can switch on (create an encryption group) encryption later, but then all ranks must be deleted and re-created, which means your data is also deleted.

5.2.4 Extent Pools

An Extent Pool is a logical construct to aggregate the extents from a set of ranks, forming a domain for extent allocation to a logical volume. Typically the set of ranks in the Extent Pool are to have the same RAID type and the same disk RPM characteristics so that the extents in the Extent Pool have homogeneous characteristics.

Important: Do not mix ranks with separate RAID types or disk rpm in an Extent Pool. Do not mix ranks of different *classes* (or *tiers*) of storage in the same Extent Pool, *unless* you want to enable the Easy Tier Automatic Mode facility.

There is no predefined affinity of ranks or arrays to a storage server. The affinity of the rank (and its associated array) to a given server is determined at the point it is assigned to an Extent Pool.

One or more ranks with the *same* extent type (FB or CKD) can be assigned to an Extent Pool. One rank can be assigned to only one Extent Pool. There can be as many Extent Pools as there are ranks.

There are considerations regarding how many ranks should be added in an Extent Pool. *Storage Pool Striping* allows you to create logical volumes striped across multiple ranks. This will typically enhance performance. To benefit from Storage Pool Striping (see “Storage Pool Striping: Extent rotation” on page 118), more than one rank in an Extent Pool is required.

Storage Pool Striping can enhance performance significantly, but when you lose one rank (in the unlikely event that a whole RAID array failed due to a scenario with multiple failures at the same time), not only is the data of this rank lost, but also all data in this Extent Pool because data is striped across all ranks. To avoid data loss, mirror your data to a remote DS8000.

The DS Storage Manager GUI prompts you to use the same RAID types in an Extent Pool. As such, when an Extent Pool is defined, it must be assigned with the following attributes:

- ▶ Server affinity
- ▶ Extent type
- ▶ RAID type
- ▶ Drive Class
- ▶ Encryption group

Just like ranks, Extent Pools also belong to an encryption group. When defining an Extent Pool, you have to specify an encryption group. Encryption group 0 means no encryption. Encryption group 1 means encryption. Currently, the DS8000 series supports only one encryption group and encryption is *on* for *all* Extent Pools or *off* for *all* Extent Pools.

The minimum number of Extent Pools is two, with one assigned to server 0 and the other to server 1 so that both servers are active. In an environment where FB and CKD are to go onto the DS8000 series storage system, four Extent Pools would provide one FB pool for each server, and one CKD pool for each server, to balance the capacity between the two servers. Figure 5-5 on page 110 is an example of a mixed environment with CKD and FB Extent Pools. Additional Extent Pools might also be desirable to segregate ranks with different DDM types. Extent Pools are expanded by adding more ranks to the pool. Ranks are organized in two *rank groups*: Rank group 0 is controlled by server 0 and rank group 1 is controlled by server 1.

Important: For best performance, balance capacity between the two servers and create at least two Extent Pools, with one per server.

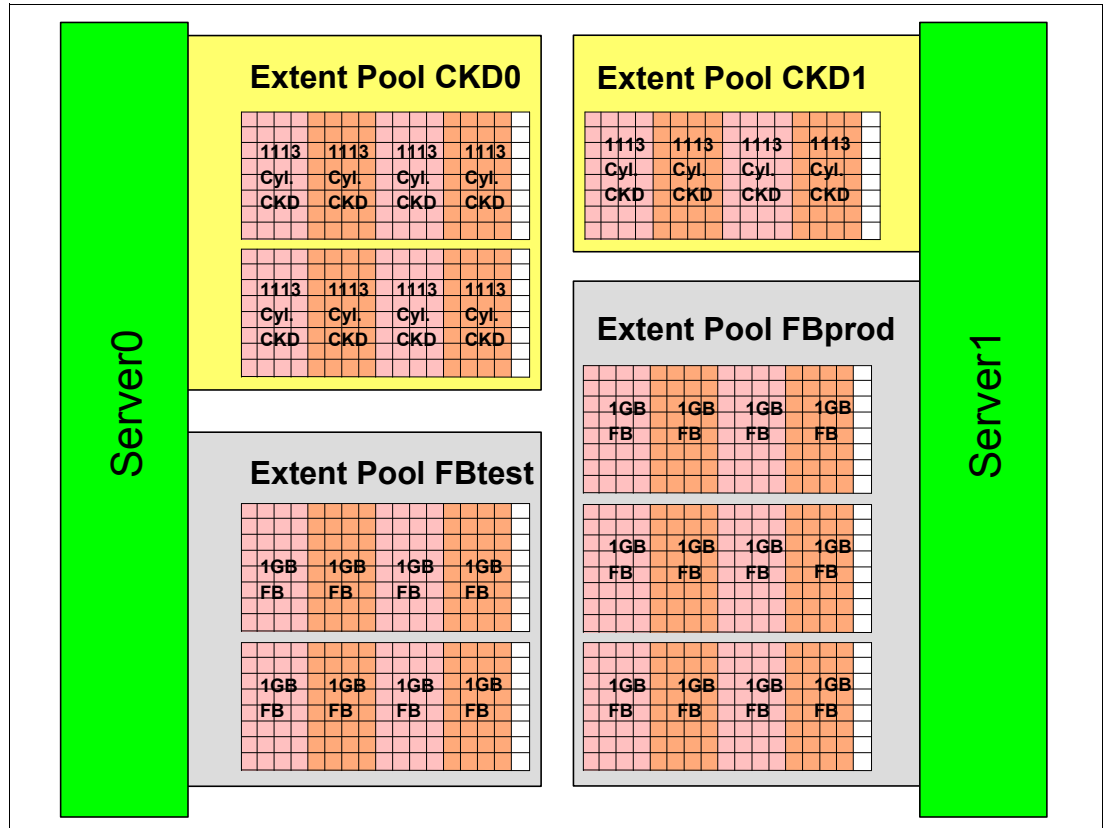


Figure 5-5 Extent Pools

Dynamic Extent Pool merge

Dynamic Extent Pool Merge is a capability provided by the Easy Tier manual mode facility.

Dynamic Extent Pool Merge allows one Extent Pool to be merged into another Extent Pool while the logical volumes in both Extent Pools remain accessible to the host servers.

Dynamic Extent Pool Merge can be used for the following reasons:

- For the consolidation of two smaller Extent Pools with equivalent storage type (that is, the same disk class, disk RPM, and RAID) into a larger Extent Pool. Creating a larger Extent Pool allows logical volumes to be distributed over a greater number of ranks, which improves overall performance in the presence of skewed workloads. Newly created volumes in the merged Extent Pool will allocate capacity as specified by the extent allocation algorithm selected. Logical volumes that existed in either the source or the target Extent Pool can be redistributed over the set of ranks in the merged Extent Pool using the Migrate Volume function.
- For consolidating two Extent Pools with exactly two different storage tiers to create a merged Extent Pool with a mix of storage technologies (either SSD + Enterprise Disk, SSD + Nearline Disk, or Enterprise and Nearline Disk). This is a prerequisite for using the Easy Tier automatic mode feature.

Figure 5-6 depicts the Easy Tier manual mode volume migration.

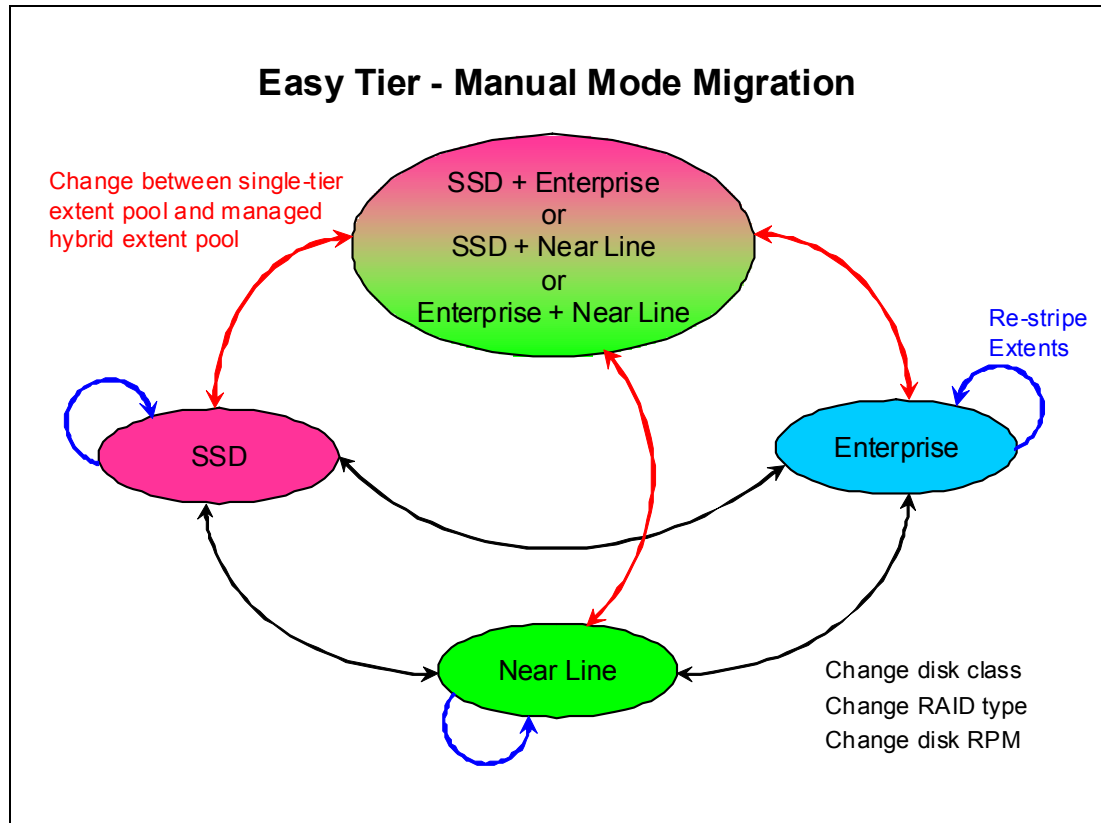


Figure 5-6 Easy Tier: Manual mode volume migration

Tip: Volume migration (or Dynamic Volume Relocation) within the same extent pool is not supported in hybrid (or multi-tiered) pools. The Easy Tier Automatic Mode automatically rebalances the volumes extents onto the ranks within the hybrid extent pool, based on the activity of the ranks.

An Extent Pool merge operation is not allowed under any of the following conditions:

- ▶ The source and target Extent Pools are not on the same storage server (server 0 or server 1). Both the source and target Extent Pools must have an even (or odd) Extent Pool number.
- ▶ The source and target Extent Pools both contain virtual capacity or both contain a space efficient repository.
- ▶ One Extent Pool is composed of SSD ranks and has either virtual capacity or a space efficient repository and the other Extent Pool contains at least one non-SSD rank.

Refer to *IBM System Storage DS8700 Easy Tier*, REDP-4667 for more information.

Tip: The Easy Tier function currently supports encryption-capable drives only when used in manual mode.

5.2.5 Logical volumes

A *logical volume* is composed of a set of extents from one Extent Pool.

On a DS8000, up to 65280 (we use the abbreviation 64 K in this discussion, even though it is actually 65536 - 256, which is not quite 64 K in binary) volumes can be created (either 64 K CKD, or 64 K FB volumes, or a mixture of both types with a maximum of 64 K volumes in total).

Fixed block LUNs

A logical volume composed of fixed block extents is called a LUN. A fixed block LUN is composed of one or more 1 GiB (2^{30} bytes) extents from one FB Extent Pool. A LUN cannot span multiple Extent Pools, but a LUN can have extents from separate ranks within the same Extent Pool. You can construct LUNs up to a size of 16 TiB (16×2^{40} bytes, or 2^{44} bytes).

Important: There is no Copy Services support for logical volumes larger than 2 TiB (2×2^{40} bytes). Do not create LUNs larger than 2 TiB if you want to use Copy Services for those LUNs, unless you want to integrate it as Managed Disks in a IBM Storage Volume Controller (SVC) with at least release 6.2 installed, and then use SVC Copy Services instead.

LUNs can be allocated in binary GiB (2^{30} bytes), decimal GB (10^9 bytes), or 512 or 520 byte blocks. However, the physical capacity that is allocated for a LUN is always a multiple of 1 GiB, so it is a good idea to have LUN sizes that are a multiple of a gibibyte. If you define a LUN with a LUN size that is not a multiple of 1 GiB, for example, 25.5 GiB, the LUN size is 25.5 GiB, but 26 GiB are physically allocated, of which 0.5 GiB of the physical storage remain unusable.

CKD volumes

A System z CKD volume is composed of one or more extents from one CKD Extent Pool. CKD extents are of the size of 3390 Model 1, which has 1113 cylinders. However, when you define a System z CKD volume, you do not specify the number of 3390 Model 1 extents but the number of cylinders you want for the volume.

Starting with the DS8000 microcode Release 4.1, you can define CKD volumes with up to 262,668 cylinders, which is about 223 GB. This volume capacity is called Extended Address Volume (EAV) and is supported by the 3390 Model A.

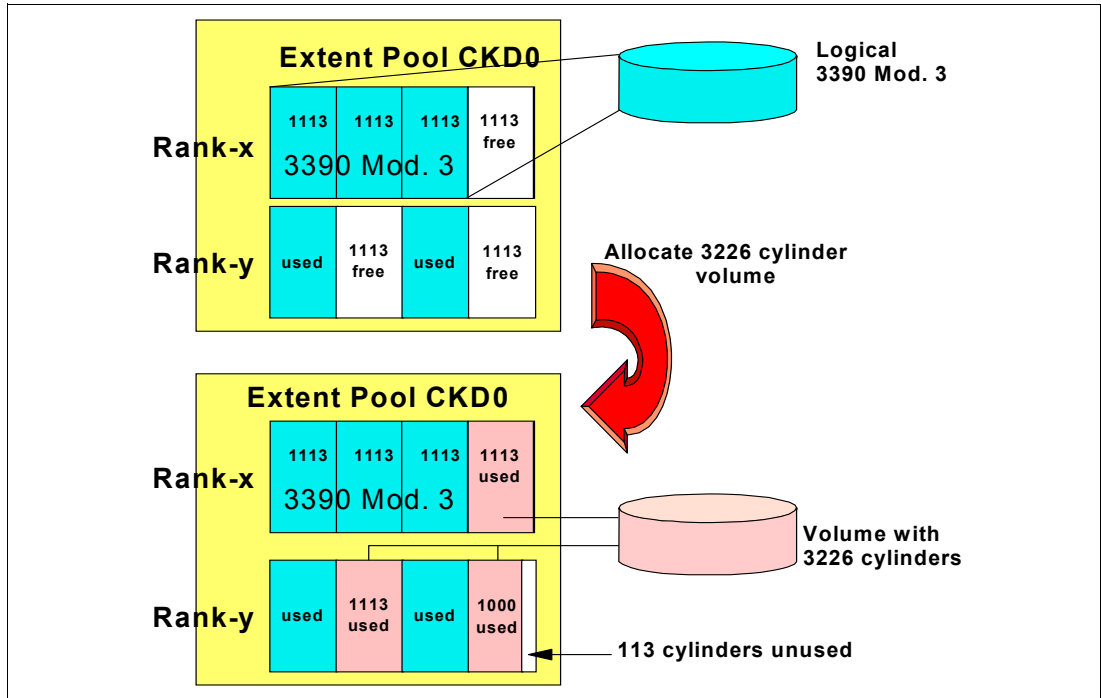


Figure 5-7 Allocation of a CKD logical volume

A CKD volume cannot span multiple Extent Pools, but a volume can have extents from different ranks in the same Extent Pool, or you can stripe a volume across the ranks (see “Storage Pool Striping: Extent rotation” on page 118). Figure 5-7 shows how a logical volume is allocated with a CKD volume as an example. The allocation process for FB volumes is similar and is shown in Figure 5-8.

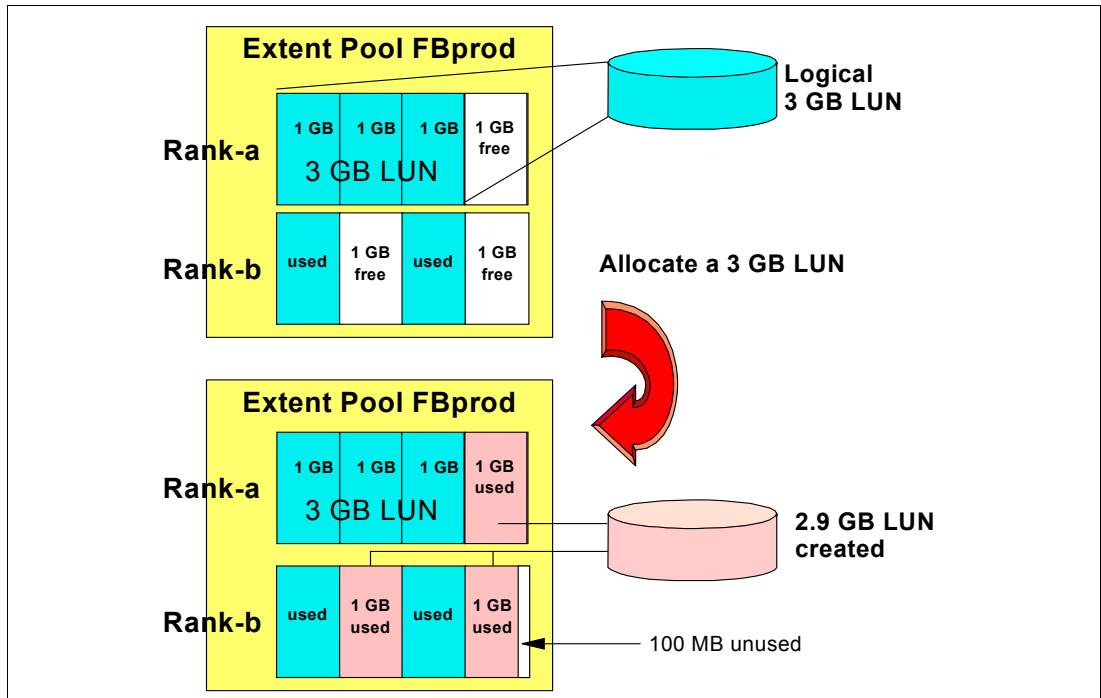


Figure 5-8 Creation of an FB LUN

IBM i LUNs

IBM i LUNs are also composed of fixed block 1 GiB extents. There are, however, special aspects with System i LUNs. LUNs created on a DS8000 are always RAID-protected. LUNs are based on RAID 5, RAID 6, or RAID 10 arrays. However, you might want to deceive i5/OS and tell it that the LUN is *not* RAID-protected. This causes the i5/OS to do its own mirroring. System i LUNs can have the attribute *unprotected*, in which case, the DS8000 will report that the LUN is not RAID-protected.

The i5/OS only supports certain fixed volume sizes, for example, model sizes of 8.5 GB, 17.5 GB, and 35.1 GB. These sizes are not multiples of 1 GB, and hence, depending on the model chosen, space is wasted. IBM i LUNs expose a 520 Byte block to the host. The operating system uses 8 of these Bytes so the usable space is still 512 Bytes like other SCSI LUNs. The capacities quoted for the IBM i LUNs are in terms of the 512 Byte block capacity and are expressed in GB (10^9). These capacities should be converted to GiB (2^{30}) when considering effective utilization of extents that are 1 GiB (2^{30}). For more information about this topic, refer to *IBM System Storage DS8000: Host Attachment and Interoperability*, SG24-8887.

5.2.6 Space Efficient volumes

When a standard FB LUN or CKD volume is created on the physical drive, it will occupy as many extents as necessary for the defined capacity.

For the DS8800 with Licensed Machine Code 7.6.1.xx or the DS8700 with Licensed Machine Code 6.6.1.xx, there are now two types of Space Efficient volumes that can be defined: *Extent Space Efficient Volumes* and *Track Space Efficient Volumes*. These two concepts are described in detail in *DS8000 Thin Provisioning*, REDP-4554.

A Space Efficient volume does not occupy physical capacity when it is created. Space gets allocated when data is actually written to the volume. The amount of space that gets physically allocated is a function of the amount of data written to or changes performed on the volume. The sum of capacities of all defined Space Efficient volumes can be larger than the physical capacity available. This function is also called *over-provisioning* or *thin provisioning*.

Space Efficient volumes for the DS8000 can be created when it has the IBM FlashCopy SE feature enabled (licensing is required).

The general idea behind Space Efficient volumes is to use or allocate physical storage when it is only potentially or temporarily needed.

Important: There is no Copy Services support for logical volumes larger than 2 TiB (2×2^{40} bytes).

Repository for Track Space Efficient volumes

The definition of Track Space Efficient (TSE) volumes begins at the Extent Pool level. TSE volumes are defined from *virtual space* in that the size of the TSE volume does not initially use physical storage. However, any data written to a TSE volume must have enough physical storage to contain this write activity. This physical storage is provided by the *repository*.

Tip: The TSE repository cannot be created on SATA Drives.

The repository is an object within an Extent Pool. In a certain sense it is similar to a volume within the Extent Pool. The repository has a physical size and a logical size. The physical size

of the repository is the amount of space that is allocated in the Extent Pool. It is the physical space that is available for all Space Efficient volumes in total in this Extent Pool. The repository is striped across all ranks within the Extent Pool. There can only be one repository per Extent Pool.

Important: The size of the repository and the virtual space it utilizes are part of the Extent Pool definition. Each Extent Pool can have a TSE volume repository, but this physical space cannot be shared between Extent Pools.

Virtual space in an Extent Pool is used for both TSE and ESE volumes, whereas the repository is only used for TSE volumes for FlashCopy SE. ESE volumes use available extents in the Extent Pool in a similar fashion as standard, fully provisioned volumes, but extents are only allocated as needed to write data to the ESE volume.

The logical size of the repository is limited by the available virtual capacity for Space Efficient volumes. As an example, there could be a repository of 100 GB reserved physical storage and you defined a virtual capacity of 200 GB. In this case, you could define 10 TSE-LUNs with 20 GB each. So the logical capacity can be larger than the physical capacity. Of course, you cannot fill all the volumes with data because the total physical capacity is limited by the repository size, which is 100 GB in this example.

Tip: In the current implementation of Track Space Efficient volumes, it is not possible to expand the physical size of the repository. Therefore, careful planning for the size of the repository is required before it is used. If a repository needs to be expanded, all Track Space Efficient volumes within this Extent Pool must be deleted, and then the repository must be deleted and recreated with the required size.

Space allocation

Space for a Space Efficient volume is allocated when a write occurs. More precisely, it is allocated when a destage from the cache occurs and there is not enough free space left on the currently allocated extent or track. The TSE allocation unit is a track (64 KB for open systems LUNs or 57 KB for CKD volumes).

Because space is allocated in extents or tracks, the system needs to maintain tables indicating their mapping to the logical volumes, so there is a performance impact involved with Space Efficient volumes. The smaller the allocation unit, the larger the tables and the impact.

Virtual space is created as part of the Extent Pool definition. This virtual space is mapped onto ESE volumes in the Extent Pool (physical space) and TSE volumes in the repository (physical space) as needed. Virtual space would equal the total space of the required ESE volumes and the TSE volumes for FlashCopy SE. No actual storage is allocated until write activity occurs to the ESE or TSE volumes.

Figure 5-9 illustrates the concept of Track Space Efficient volumes.

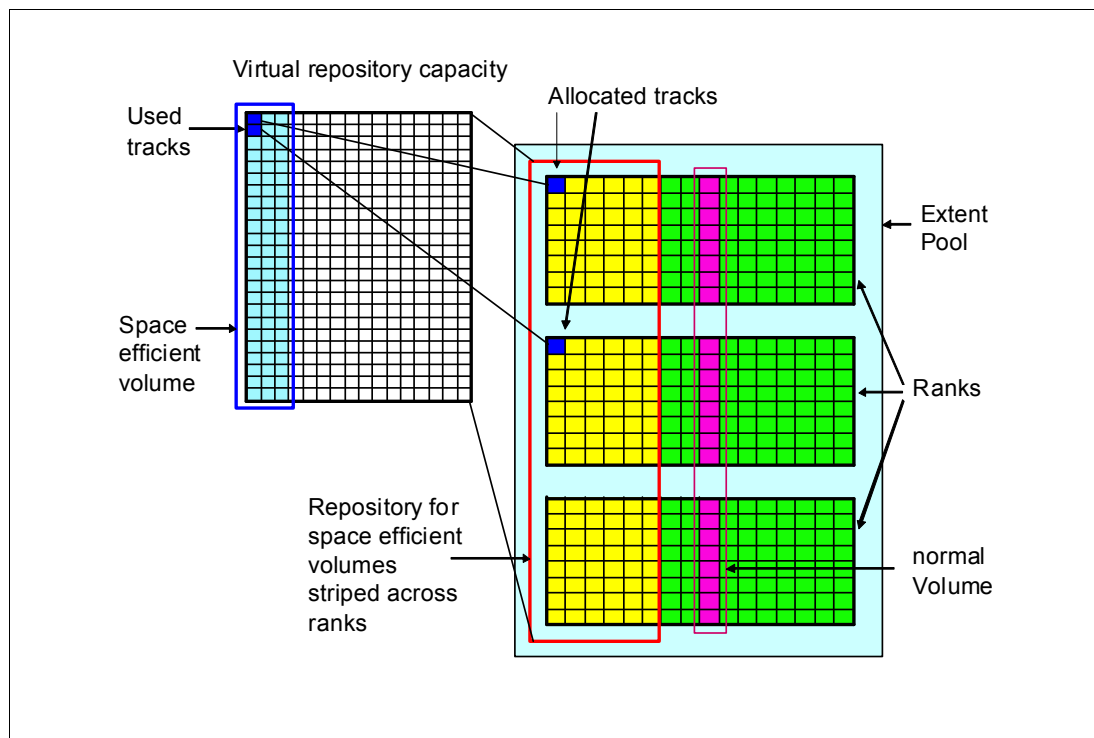


Figure 5-9 Concept of Track Space Efficient volumes for FlashCopy SE

The lifetime of data on Track Space Efficient volumes is expected to be short because they are used as FlashCopy targets only. Physical storage gets allocated when data is written to Track Space Efficient volumes, and we need a mechanism to free up physical space in the repository when the data is no longer needed.

The FlashCopy commands have options to release the space of Track Space Efficient volumes when the FlashCopy relationship is established or removed.

The CLI commands **initfbvol** and **initckdvol** can also release the space for Space Efficient volumes (ESE and TSE).

Figure 5-10 illustrates the concept of ESE logical volumes.

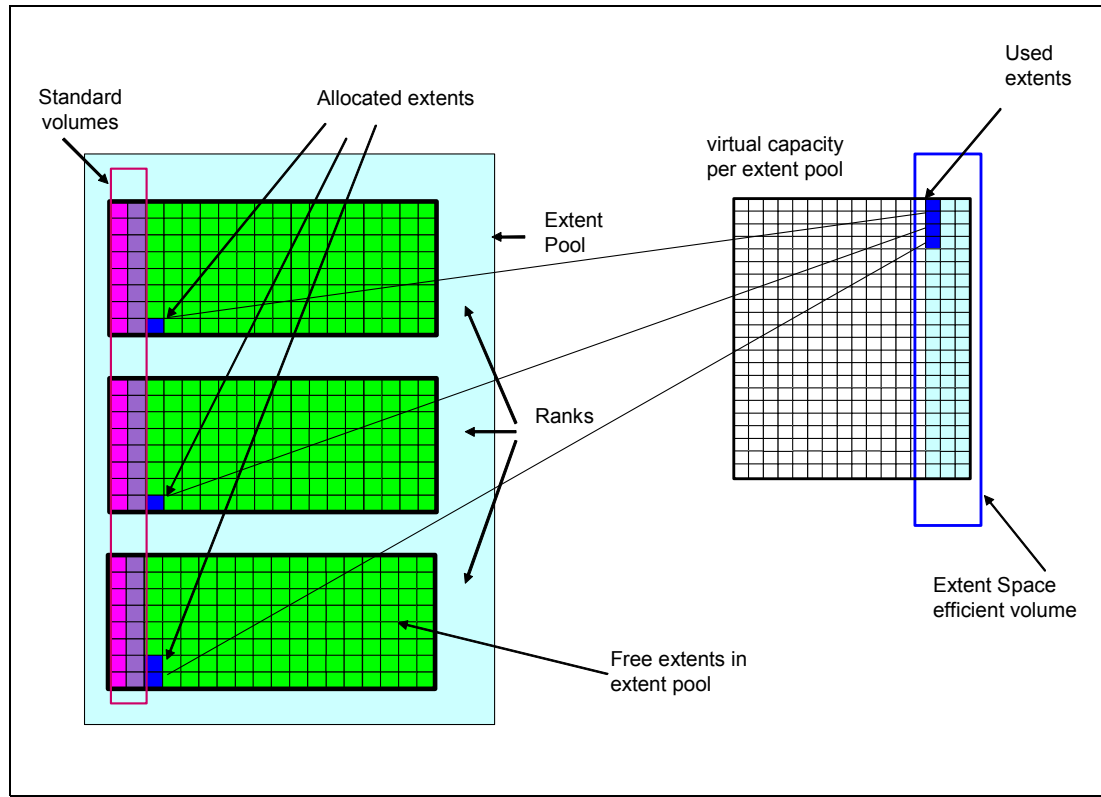


Figure 5-10 Concept of ESE logical volumes

Use of Extent Space Efficient volumes

Like standard volumes (which are fully provisioned), ESE volumes can be mapped to hosts. However, they are not supported in combination with Copy Services functions at this time.

Use of Track Space Efficient volumes

Track Space Efficient volumes are supported as FlashCopy target volumes only.

For detailed information about ESE and TSE volumes concepts, refer to *DS8000 Thin Provisioning*, REDP-4554.

Important: Space Efficient volumes (ESE or TSE) are supported but not managed by the IBM System Storage Easy Tier function.

5.2.7 Allocation, deletion, and modification of LUNs/CKD volumes

All extents of the ranks assigned to an Extent Pool are independently available for allocation to logical volumes. The extents for a LUN/volume are logically ordered, but they do not have to come from one rank and the extents do not have to be contiguous on a rank.

This construction method of using fixed extents to form a logical volume in the DS8000 series allows flexibility in the management of the logical volumes. We can delete LUNs/CKD volumes, resize LUNs/volumes, and reuse the extents of those LUNs to create other

LUNs/volumes, maybe of different sizes. One logical volume can be removed without affecting the other logical volumes defined on the same Extent Pool.

Because the extents are *cleaned* after you have deleted a LUN or CKD volume, it can take some time until these extents are available for reallocation. The reformatting of the extents is a background process.

There are two extent allocation algorithms for the DS8000: Rotate volumes and Storage Pool Striping (Rotate extents).

Tip: The default for extent allocation method is Storage Pool Striping (Rotate extents) for Licensed Machine Code 6.6.0.xx and later. In prior releases of Licensed Machine Code the default allocation method was Rotate volumes.

Storage Pool Striping: Extent rotation

The preferred storage allocation method is Storage Pool Striping. Storage Pool Striping is an option when a LUN/volume is created. The extents of a volume can be striped across several ranks. An Extent Pool with more than one rank is needed to use this storage allocation method.

The DS8000 maintains a sequence of ranks. The first rank in the list is randomly picked at each power on of the storage subsystem. The DS8000 keeps track of the rank in which the last allocation started. The allocation of the first extent for the next volume starts from the next rank in that sequence. The next extent for that volume is taken from the next rank in sequence and so on. Thus, the system rotates the extents across the ranks.

Rotate volumes allocation method

Extents can be allocated sequentially. In this case all extents are taken from the same rank until we have enough extents for the requested volume size or the rank is full, in which case the allocation continues with the next rank in the Extent Pool.

If more than one volume is created in one operation, the allocation for each volume starts in another rank. When allocating several volumes, we *rotate* through the ranks.

You might want to consider this allocation method when you prefer to manage performance manually. The workload of one volume is going to one rank. This makes the identification of performance bottlenecks easier; however, by putting all the volumes data onto just one rank, you might introduce a bottleneck, depending on your actual workload.

Tip: Rotate extents and rotate volume EAMs provide distribution of volumes over ranks. Rotate extents does this at a granular (1 GB extent) level, which is the preferred method to minimize hot spots and improve overall performance.

In a mixed disk characteristics (or *hybrid*) Extent Pool containing exactly two *classes* (or *tiers*) of ranks, the Storage Pool striping EAM is used independently of the requested EAM, and EAM is set to *managed*.

For extent pools that contain SSD disks, extent allocation is done on HDD ranks (either Enterprise or Near Line) while space remains available, and before allocating on SSD ranks. For extent pools that contain a mix of Enterprise and Nearline ranks, extent allocation is done on Enterprise ranks first.

Figure 5-11 shows an example of how volumes are allocated within the Extent Pool.

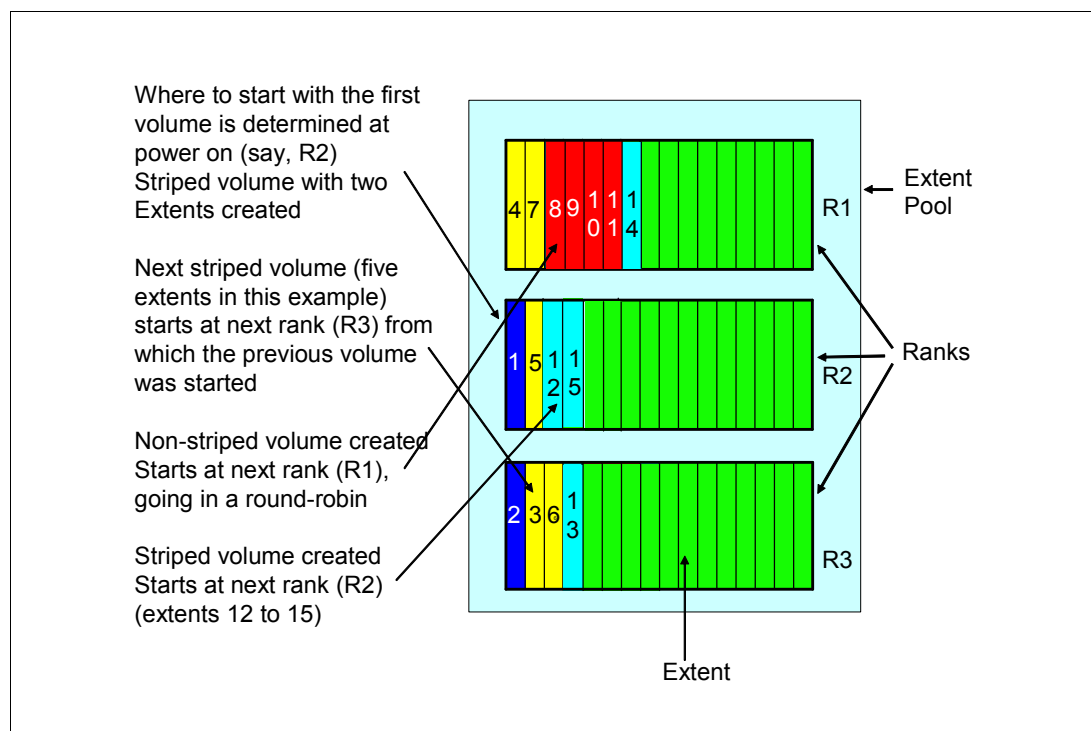


Figure 5-11 Extent allocation methods

When you create striped volumes and non-striped volumes in an Extent Pool, a rank could be filled before the others. A full rank is skipped when you create new striped volumes.

Tip: If you have to add capacity to an Extent Pool because it is nearly full, it is better to add several ranks at the same time, not just one. This allows new volumes to be striped across the newly added ranks.

With the Easy Tier manual mode facility, if the extent pool is a non-hybrid one, the user can request an Extent Pool merge followed by a volume relocation with striping to perform the same function. In case of hybrid managed extent pool, extents will be automatically relocated over time, according to performance needs. For more details, refer to the *IBM System Storage DS8000: Easy Tier*, REDP-4667.

Figure 5-12 shows both standard volumes and Extent Space Efficient volumes in an Extent Pool, both using the rotate volumes allocation method (red standard volume 8-12 and pink ESE volume 17-19, 26-27) and Storage Pool Striping extent allocation methods (all other standard and ESE volumes).

Tip: The rotate volume EAM is not allowed if one Extent Pool is composed of SSD disks and has a Space Efficient repository or virtual capacity configured.

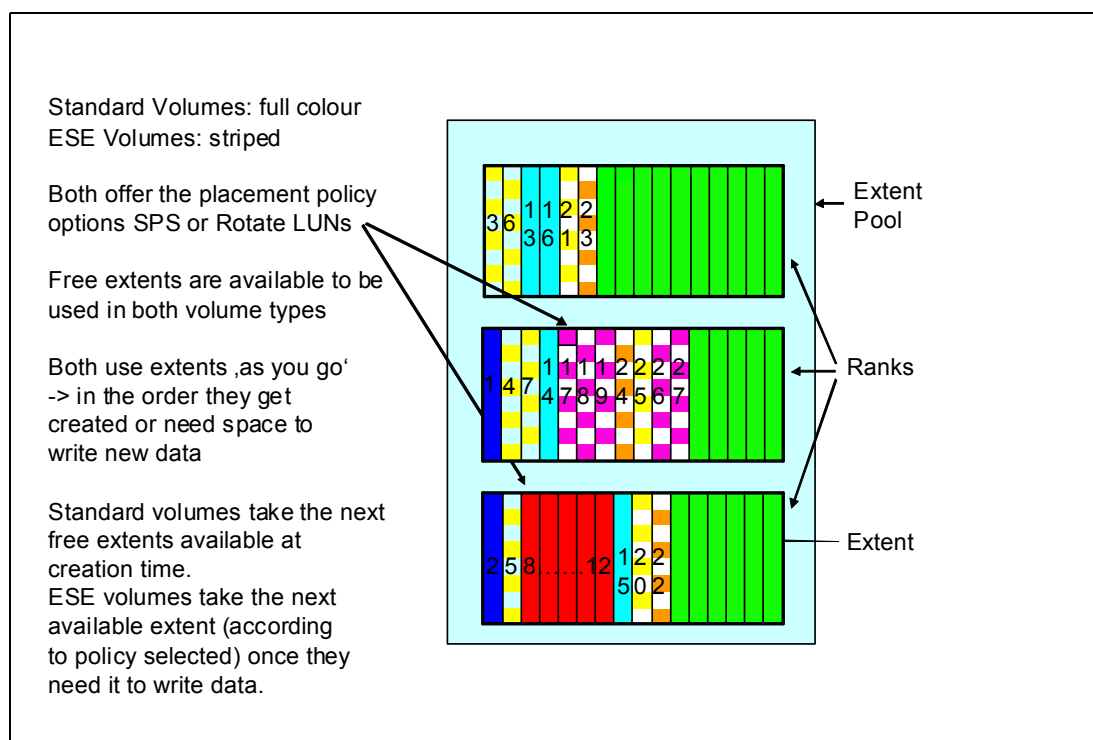


Figure 5-12 Free extents could be used for Standard Volumes or Extent Space Efficient Volumes

By using striped volumes, you distribute the I/O load of a LUN/CKD volume to more than just one set of eight disk drives. The ability to distribute a workload to many physical drives can greatly enhance performance for a logical volume. In particular, operating systems that do not have a volume manager that can do striping will benefit most from this allocation method.

However, if you have Extent Pools with many ranks and all volumes are striped across the ranks and you lose just one rank, for example, because there are two disk drives in the same rank that fail at the same time and it is not a RAID 6 rank, you will lose much of your data.

On the other hand, if you do, for example, Physical Partition striping in AIX already, double striping probably will not improve performance any further. The same can be expected when the DS8000 LUNs are used by an SVC striping data across LUNs.

If you decide to use Storage Pool Striping it is probably better to use this allocation method for all volumes in the Extent Pool to keep the ranks equally filled and utilized.

Tip: When configuring a new DS8000, do not mix volumes using the storage pool striping method and volumes using the rotate volumes method in the same Extent Pool.

For more information about how to configure Extent Pools and volumes for optimal performance see Chapter 7, “Performance” on page 157.

Logical volume configuration states

Each logical volume has a configuration state attribute. The configuration state reflects the condition of the logical volume relative to user requested configuration operations, as shown in Figure 5-13.

When a logical volume creation request is received, a logical volume object is created and the logical volume's configuration state attribute is placed in the *configuring* configuration state. After the logical volume is created and available for host access, it is placed in the *normal* configuration state. If a volume deletion request is received, the logical volume is placed in the *deconfiguring* configuration state until all capacity associated with the logical volume is deallocated and the logical volume object is deleted.

The *reconfiguring* configuration state is associated with a volume expansion request. See “Dynamic Volume Expansion” on page 121 for more information. The *transposing* configuration state is associated with an Extent Pool merge, as described in “Dynamic Extent Pool merge” on page 110. The *migrating*, *migration paused*, *migration error*, and *migration cancelled* configuration states are associated with a volume relocation request, as described in “Dynamic volume migration” on page 122.

As shown, the configuration state serializes user requests with the exception that a volume deletion request can be initiated from any configuration state.

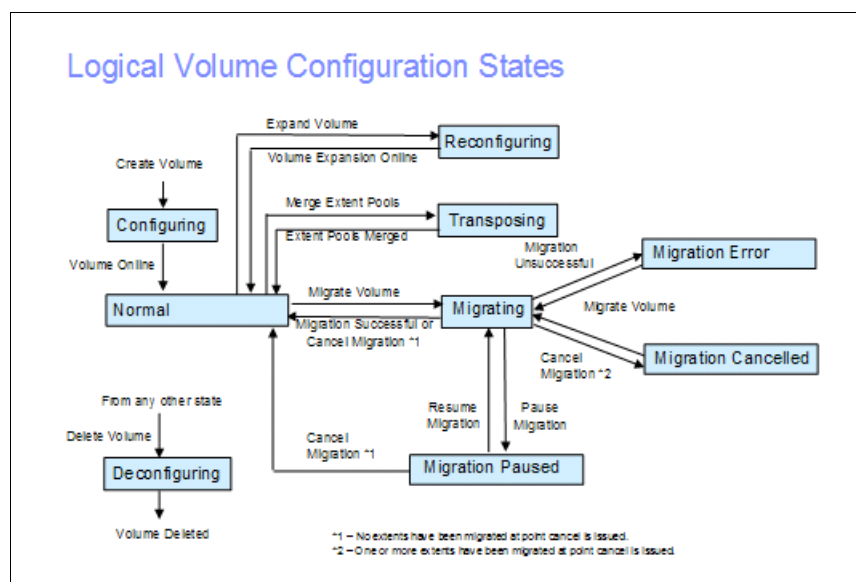


Figure 5-13 Logical volume configuration states

Dynamic Volume Expansion

The size of a LUN or CKD volume can be expanded without destroying the data. On the DS8000, you simply add extents to the volume. The operating system will have to support this re-sizing.

A logical volume has the attribute of being striped across the ranks or not. If the volume was created as striped across the ranks of the Extent Pool, then the extents that are used to increase the size of the volume are striped. If a volume was created without striping, the

system tries to allocate the additional extents within the same rank that the volume was created from originally.

Because most operating systems have no means of moving data from the end of the *physical* disk off to unused space at the beginning of the disk, and because of the risk of data corruption, IBM does not support shrinking a volume. The DS8000 configuration interfaces DS CLI and DS GUI will *not* allow you to change a volume to a smaller size.

Consideration: Before you can expand a volume, you have to delete any copy services relationship involving that volume.

Dynamic volume migration

Dynamic volume migration or Dynamic Volume Relocation (DVR) is a capability provided as part of the Easy Tier manual mode facility.

Dynamic Volume Relocation allows data stored on a logical volume to be migrated from its currently allocated storage to newly allocated storage while the logical volume remains accessible to attached hosts. The user can request Dynamic Volume Relocation using the Migrate Volume function that is available through the DS8000 Storage Manager GUI or DS CLI. Dynamic Volume Relocation allows the user to specify a target Extent Pool and an extent allocation method (EAM). The target Extent Pool can be a separate Extent Pool than the Extent Pool where the volume is currently located, or the same extent pool, only if it is a non-hybrid (or single-tier) pool.

Important: DVR in the same extent pool is not allowed in the case of a managed hybrid pool. In managed hybrid extent pools, Easy Tier Automatic Mode automatically relocates extents within the ranks to allow performance rebalancing.

Dynamic volume migration provides:

- ▶ The ability to change the Extent Pool in which a logical volume is provisioned, which provides a mechanism to change the underlying storage characteristics of the logical volume to include the disk class (Solid State Drive, enterprise disk, or SATA disk), disk rpm, and RAID array type. Volume migration can also be used to migrate a logical volume into or out of an Extent Pool.
- ▶ The ability to specify the extent allocation method for a volume migration allowing the extent allocation method to be changed between the available extent allocation method any time after volume creation. Volume migration specifying the rotate extents EAM can also be used, in non-hybrid extent pools, to re-distribute a logical volume's extent allocations across the currently existing ranks in the Extent Pool if additional ranks are added to an Extent Pool.

Each logical volume has a configuration state, as described in “Logical volume configuration states” on page 121. To initiate a volume migration, the logical volume must initially be in the normal configuration state. The volume migration will follow each of the states discussed.

There are additional functions that are associated with volume migration that allow the user to pause, resume, or cancel a volume migration. Any and all logical volumes can be requested to be migrated at any given time as long as there is sufficient capacity available to support the pre-allocation of the migrating logical volumes in their specified target Extent Pool.

For additional information about this topic, refer to *IBM System Storage DS8000 Easy Tier*, REDP-4667.

5.2.8 Logical subsystem

A *logical subsystem* (LSS) is another logical construct. It groups logical volumes and LUNs, in groups of up to 256 logical volumes.

On the DS8000 series, there is no fixed binding between any rank and any logical subsystem. The capacity of one or more ranks can be aggregated into an Extent Pool and logical volumes configured in that Extent Pool are not bound to any specific rank. Different logical volumes on the same logical subsystem can be configured in separate Extent Pools. As such, the available capacity of the storage facility can be flexibly allocated across the set of defined logical subsystems and logical volumes. You can now define up to 255 LSSs for the DS8000 series.

For each LUN or CKD volume, you can now choose an LSS. You can have up to 256 volumes in one LSS. There is, however, one restriction. We already have seen that volumes are formed from a number of extents from an Extent Pool. Extent Pools, however, belong to one server (CEC), server 0 or server 1, respectively. LSSs also have an affinity to the servers. All even-numbered LSSs (X'00', X'02', X'04', up to X'FE') belong to server 0 and all odd-numbered LSSs (X'01', X'03', X'05', up to X'FD') belong to server 1. LSS X'FF' is reserved.

System z users are familiar with a *logical control unit* (LCU). System z operating systems configure LCUs to create device addresses. There is a one to one relationship between an LCU and a CKD LSS (LSS X'ab' maps to LCU X'ab'). Logical volumes have a logical volume number X'abcd' where X'ab' identifies the LSS and X'cd' is one of the 256 logical volumes on the LSS. This logical volume number is assigned to a logical volume when a logical volume is created and determines the LSS that it is associated with. The 256 possible logical volumes associated with an LSS are mapped to the 256 possible device addresses on an LCU (logical volume X'abcd' maps to device address X'cd' on LCU X'ab'). When creating CKD logical volumes and assigning their logical volume numbers, consider whether Parallel Access Volumes (PAVs) are required on the LCU and reserve addresses on the LCU for alias addresses.

For open systems, LSSs do not play an important role except in determining which server manages the LUN (and in which Extent Pool it must be allocated) and in certain aspects related to Metro Mirror, Global Mirror, or any of the other remote copy implementations.

Certain management actions in Metro Mirror, Global Mirror, or Global Copy, operate at the LSS level. For example, the freezing of pairs to preserve data consistency across all pairs, in case you have a problem with one of the pairs, is done at the LSS level. With the option to put all or most of the volumes of a certain application in just one LSS, makes the management of remote copy operations easier (Figure 5-14 on page 124).

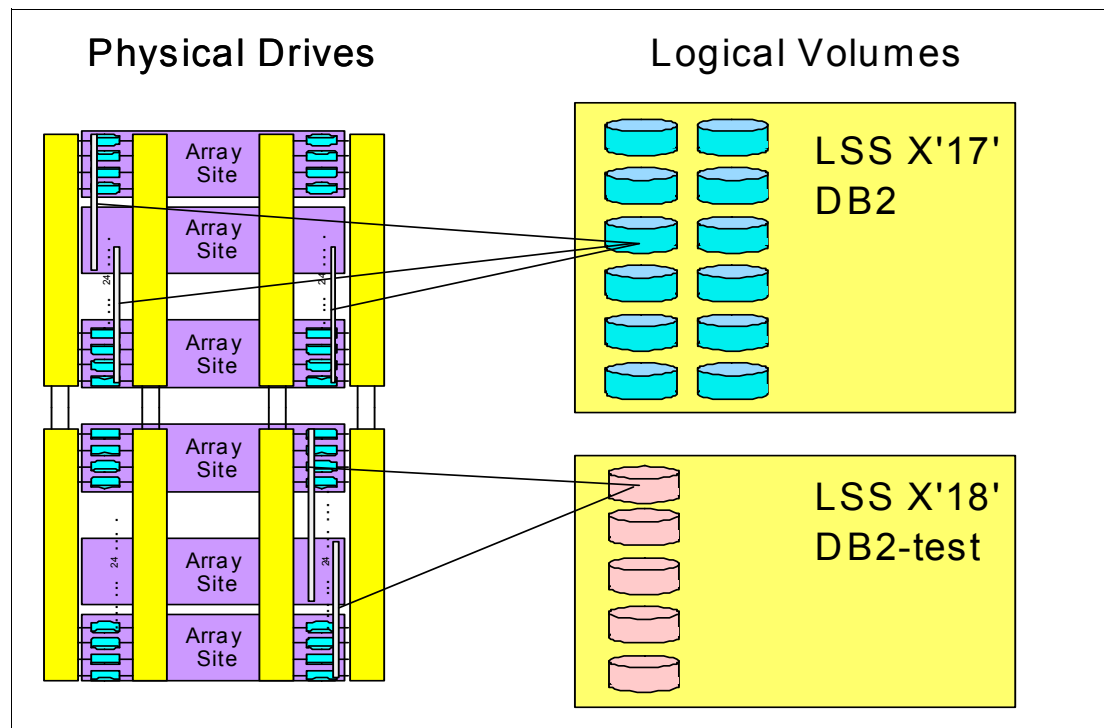


Figure 5-14 Grouping of volumes in LSSs

Fixed block LSSs are created automatically when the first fixed block logical volume on the LSS is created, and deleted automatically when the last fixed block logical volume on the LSS is deleted. CKD LSSs require user parameters to be specified and must be created before the first CKD logical volume can be created on the LSS; they must be deleted manually after the last CKD logical volume on the LSS is deleted.

Address groups

Address groups are created automatically when the first LSS associated with the address group is created, and deleted automatically when the last LSS in the address group is deleted.

All devices in an LSS must be either CKD *or* FB. This restriction goes even further. LSSs are grouped into address groups of 16 LSSs. LSSs are numbered X'ab', where *a* is the address group and *b* denotes an LSS within the address group. So, for example, X'10' to X'1F' are LSSs in address group 1.

All LSSs within one address group have to be of the same type, CKD or FB. The first LSS defined in an address group sets the type of that address group.

Important: System z users who still want to use ESCON® to attach hosts to the DS8000 series should be aware that ESCON supports only the 16 LSSs of address group 0 (LSS X'00' to X'0F'). Therefore, this address group should be reserved for ESCON-attached CKD devices in this case and not used as FB LSSs. The DS8800 does not support ESCON channels. ESCON devices can only be attached by using FICON/ESCON converters.

Figure 5-15 on page 125 shows the concept of LSSs and address groups.

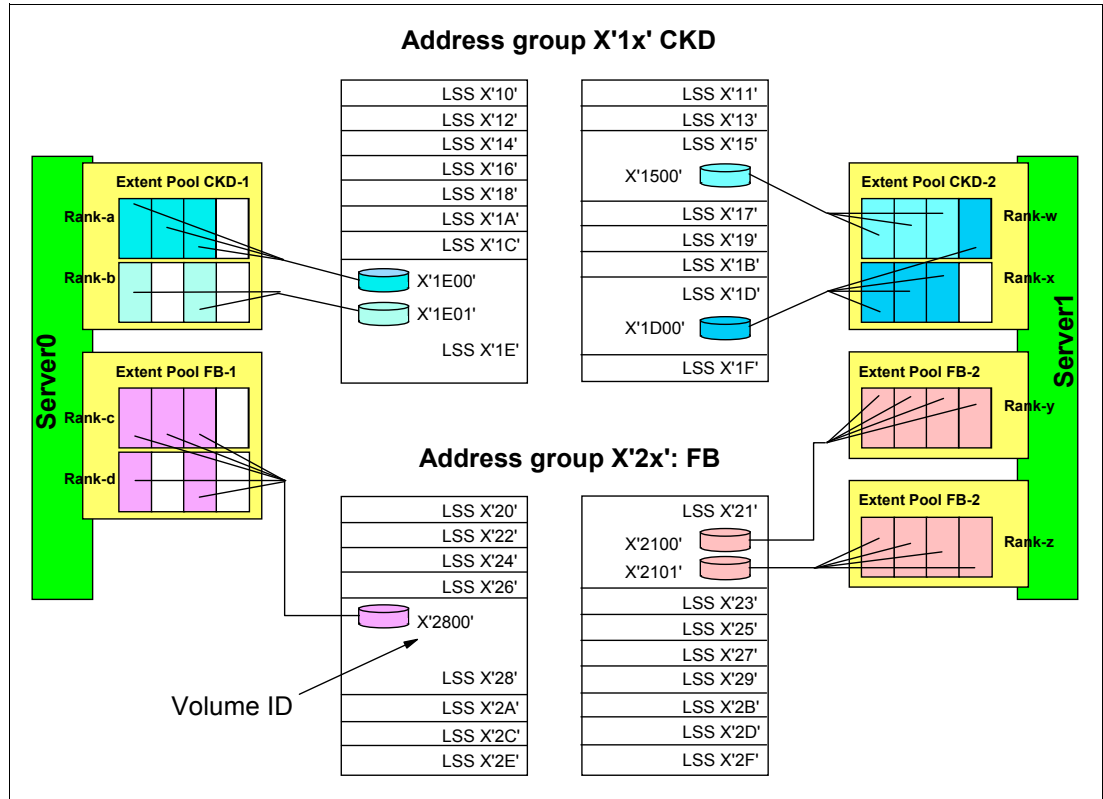


Figure 5-15 Logical storage subsystems

The LUN identifications $X'gabb'$ are composed of the address group $X'g'$, and the LSS number within the address group $X'a'$, and the position of the LUN within the LSS $X'bb'$. For example, FB LUN $X'2101'$ denotes the second ($X'01'$) LUN in LSS $X'21'$ of address group 2.

5.2.9 Volume access

A DS8000 provides mechanisms to control host access to LUNs. In most cases, a server has two or more HBAs and the server needs access to a group of LUNs. For easy management of server access to logical volumes, the DS8000 introduced the concept of host attachments and volume groups.

Host attachment

Host bus adapters (HBAs) are identified to the DS8000 in a host attachment construct that specifies the HBAs' World Wide Port Names (WWPNs). A set of host ports can be associated through a port group attribute that allows a set of HBAs to be managed collectively. This port group is referred to as a host attachment within the GUI.

Each host attachment can be associated with a volume group to define which LUNs that HBA is allowed to access. Multiple host attachments can share the same volume group. The host attachment can also specify a port mask that controls which DS8000 I/O ports the HBA is allowed to log in to. Whichever ports the HBA logs in on, it sees the same volume group that is defined on the host attachment associated with this HBA.

The maximum number of host attachments on a DS8000 is 8192.

Volume group

A *volume group* is a named construct that defines a set of logical volumes. When used in conjunction with CKD hosts, there is a default volume group that contains all CKD volumes and any CKD host that logs in to a FICON I/O port has access to the volumes in this volume group. CKD logical volumes are automatically added to this volume group when they are created and automatically removed from this volume group when they are deleted.

When used in conjunction with open systems hosts, a host attachment object that identifies the HBA is linked to a specific volume group. You must define the volume group by indicating which fixed block logical volumes are to be placed in the volume group. Logical volumes can be added to or removed from any volume group dynamically.

There are two types of volume groups used with open systems hosts and the type determines how the logical volume number is converted to a host addressable LUN_ID on the Fibre Channel SCSI interface. A *map volume group* type is used in conjunction with FC SCSI host types that poll for LUNs by walking the address range on the SCSI interface. This type of volume group can map any FB logical volume numbers to 256 LUN_IDs that have zeroes in the last six Bytes and the first two Bytes in the range of X'0000' to X'00FF'.

A *mask volume group* type is used in conjunction with FC SCSI host types that use the Report LUNs command to determine the LUN_IDs that are accessible. This type of volume group can allow any and all FB logical volume numbers to be accessed by the host where the mask is a bitmap that specifies which LUNs are accessible. For this volume group type, the logical volume number X'*abcd*' is mapped to LUN_ID X'40ab40cd00000000'. The volume group type also controls whether 512 Byte block LUNs or 520 Byte block LUNs can be configured in the volume group.

When associating a host attachment with a volume group, the host attachment contains attributes that define the logical block size and the Address Discovery Method (LUN Polling or Report LUNs) that are used by the host HBA. These attributes must be consistent with the volume group type of the volume group that is assigned to the host attachment so that HBAs that share a volume group have a consistent interpretation of the volume group definition and have access to a consistent set of logical volume types. The GUI typically sets these values appropriately for the HBA based on your specification of a host type. You must consider what volume group type to create when setting up a volume group for a particular HBA.

FB logical volumes can be defined in one or more volume groups. This allows a LUN to be shared by host HBAs configured to separate volume groups. An FB logical volume is automatically removed from all volume groups when it is deleted.

The maximum number of volume groups is 8320 for the DS8000.

Figure 5-16 on page 127 shows the relationships between host attachments and volume groups. Host AIXprod1 has two HBAs, which are grouped together in one host attachment and both are granted access to volume group DB2-1. Most of the volumes in volume group DB2-1 are also in volume group DB2-2, accessed by server AIXprod2. In our example, there is, however, one volume in each group that is not shared. The server in the lower left part has four HBAs and they are divided into two distinct host attachments. One can access volumes shared with AIXprod1 and AIXprod2. The other HBAs have access to a volume group called "docs".

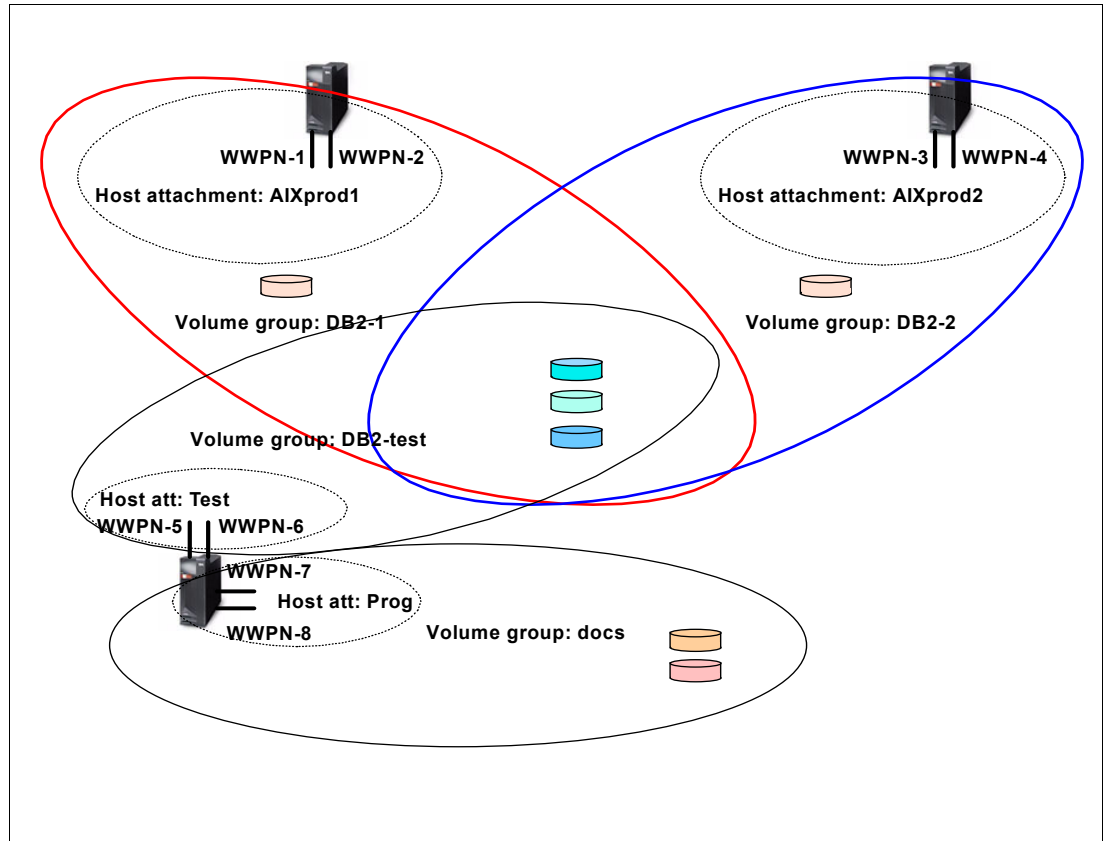


Figure 5-16 Host attachments and volume groups

5.2.10 Virtualization hierarchy summary

Going through the virtualization hierarchy, we start with just a bunch of disks that are grouped in array sites. An array site is transformed into an array, with spare disks. The array is further transformed into a rank with extents formatted for FB data or CKD.

- ▶ The extents from selected ranks are added to an Extent Pool. The combined extents from those ranks in the Extent Pool are used for subsequent allocation to one or more logical volumes. Within the Extent Pool, we can reserve space for Track Space Efficient (TSE) volumes by means of creating a repository. Both ESE and TSE volumes require virtual capacity to be available in the Extent Pool.
- ▶ Next, we create logical volumes within the Extent Pools (by default, striping the volumes), assigning them a logical volume number that determines which logical subsystem they would be associated with and which server would manage them. This is the same for both Standard volumes (fully allocated) and Extent Space Efficient volumes. Track Space Efficient volumes for use with FlashCopy SE can only be created within the repository of the Extent Pool.
- ▶ The LUNs are then assigned to one or more volume groups.
- ▶ Finally, the host HBAs are configured into a host attachment that is associated with a volume group.

This virtualization concept provides much more flexibility than in previous products. Logical volumes can dynamically be created, deleted, and resized. They can be grouped logically to simplify storage management. Large LUNs and CKD volumes reduce the total number of volumes, which contributes to the reduction of management effort.

Figure 5-17 summarizes the virtualization hierarchy.

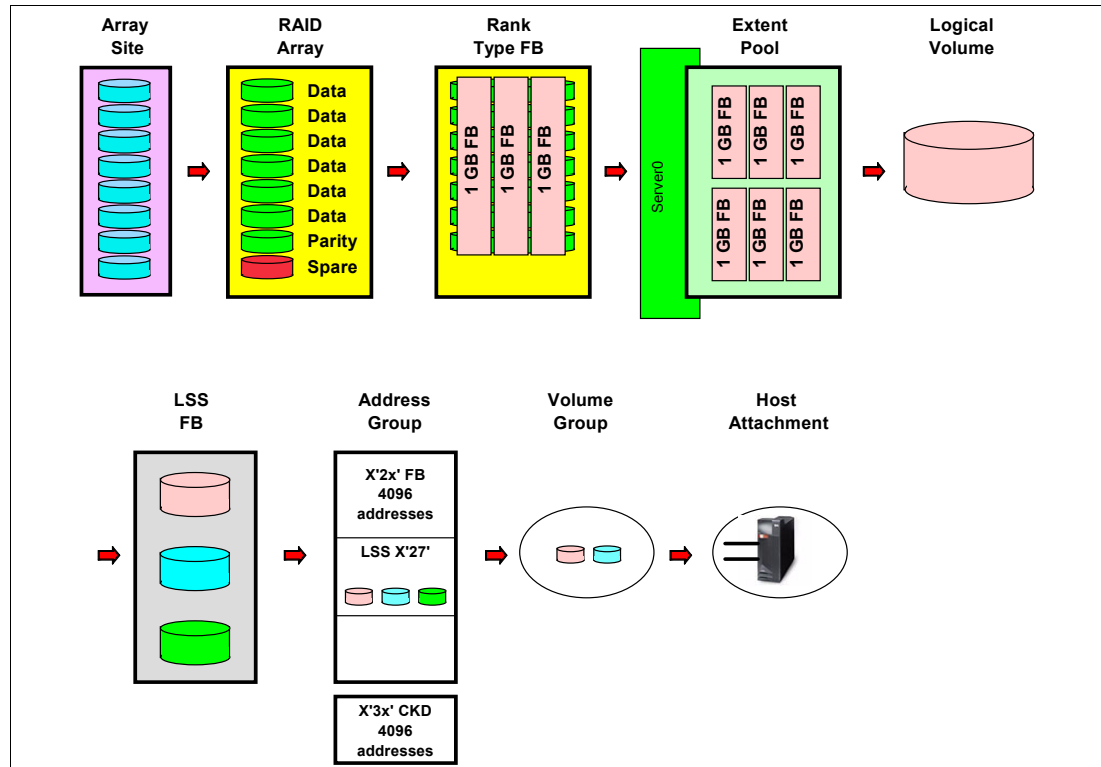


Figure 5-17 Virtualization hierarchy

5.3 Benefits of virtualization

The DS8000 physical and logical architecture defines new standards for enterprise storage virtualization. The main benefits of the virtualization layers are:

- ▶ Flexible LSS definition allows maximization and optimization of the number of devices per LSS.
- ▶ No strict relationship between RAID ranks and LSSs.
- ▶ No connection of LSS performance to underlying storage.
- ▶ Number of LSSs can be defined based upon device number requirements:
 - With larger devices, significantly fewer LSSs might be used.
 - Volumes for a particular application can be kept in a single LSS.
 - Smaller LSSs can be defined if required (for applications requiring less storage).
 - Test systems can have their own LSSs with fewer volumes than production systems.
- ▶ Increased number of logical volumes:
 - Up to 65280 (CKD)
 - Up to 65280 (FB)
 - 65280 total for CKD + FB
- ▶ Any mixture of CKD or FB addresses in 4096 address groups.

- ▶ Increased logical volume size:
 - CKD: 223 GB (262,668 cylinders), architected for 219 TB
 - FB: 16 TB, architected for 1 PB
- ▶ Flexible logical volume configuration:
 - Multiple RAID types (RAID 5, RAID 6, and RAID 10)
 - Storage types (CKD and FB) aggregated into Extent Pools
 - Volumes allocated from extents of Extent Pool
 - Storage pool striping
 - Dynamically add and remove volumes
 - Logical Volume Configuration States
 - Dynamic Volume Expansion
 - Extent Space Efficient volumes for Thin Provisioning
 - Track Space Efficient volumes for FlashCopy SE
 - Extended Address Volumes (CKD)
 - Dynamic Extent Pool merging for Easy Tier
 - Dynamic Volume Relocation for Easy Tier
- ▶ Virtualization reduces storage management requirements.

5.4 zDAC - z/OS FICON discovery and Auto-Configuration

Both, the DS8700 and DS8800 supports the z/OS FICON Discovery and Auto-Configuration Feature (zDAC) which is deployed by the new z/Enterprise z196.

This function has been developed to reduce the complexity and skills needed in a complex FICON production environment for changing the I/O configuration.

With zDAC you can add storage subsystems to an existing I/O configuration in less time, depending on the policy you have defined. zDAC proposes new configurations that incorporate the current contents of your I/O Definition File (IODF) with additions for new and changed subsystems and their devices based on the policy you have defined in the Hardware Configuration Definition (HCD).

The following requirements must be met for using zDAC:

- ▶ Your System z must be a z/Enterprise (z196) running z/OS V1 R12
- ▶ LPAR must be authorized to make dynamic I/O Configuration (zDCM) changes on each processor hosting a discovery system.
- ▶ Hardware Configuration Definition (HCD) and Hardware Configuration Management (HCM) users need to have authority for making dynamic I/O configuration changes

As its name implies, zDAC provides two capabilities:

- ▶ Discovery:
 - Provides capability to discover attached disk connected to FICON fabrics
 - Detects new and older storage subsystems
 - Detects new control units on existing storage subsystems
 - Proposes control units and device numbering

- Proposes paths for all discovery systems to newly discovered control units including the Sysplex scope
- ▶ Auto-Configuration:
 - For high availability reasons, when zDAC proposes channel paths, it looks at single point of failure only. It does not consider any channel or port speed, or any current performance information.
 - After a storage subsystem has been explored, the discovered information is compared against the target IODF, paths are proposed to new control units, and devices are displayed to the user. With that scope of discovery and autoconfiguration the target work IODF is being updated

When using zDAC, keep in mind the following considerations:

- ▶ Physical planning is still our responsibility
- ▶ Logical configurations of the Storage Subsystem are still done by you
- ▶ What z/OS image should be allowed to use the new devices
- ▶ How should the new devices be numbered
- ▶ How many paths to new control units should be configured

Figure 5-18 gives a schematic overview of the zDAC concept

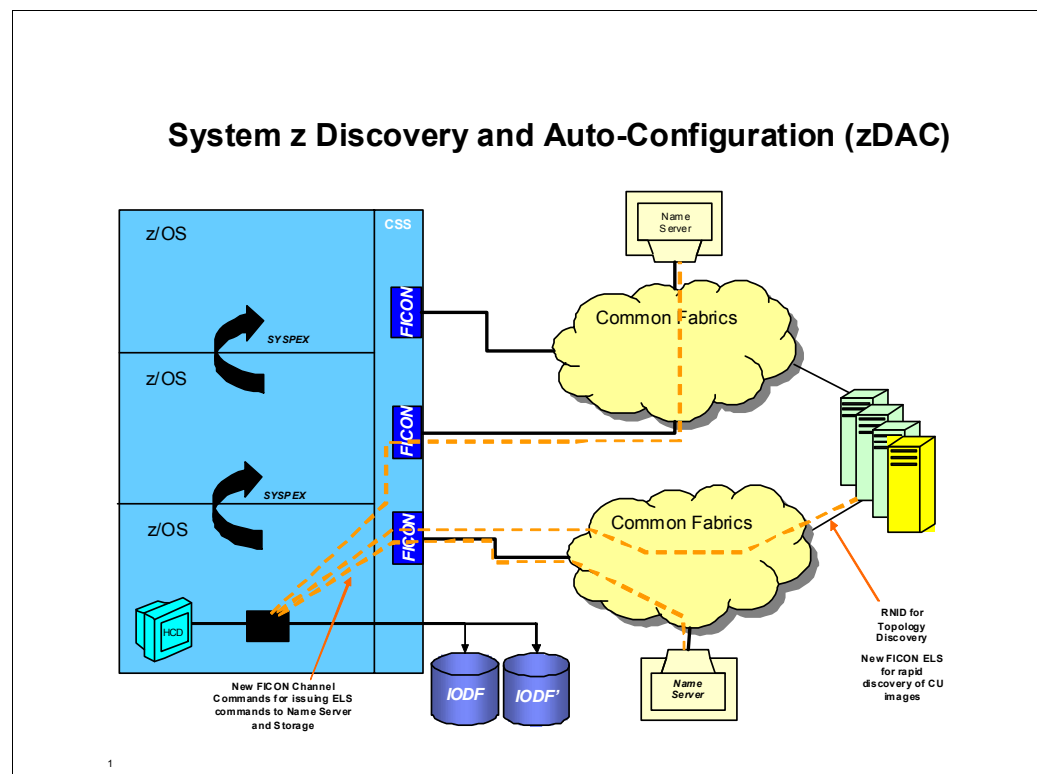


Figure 5-18 zDAC concept

Tip: The zDAC support is included in the DS8000 Licensed Machine Code R5.1 and later. For more detailed Information, refer to *z/OS V1R12 HCD User's Guide SC33-7988*.

5.5 EAV V2 - Extended Adresse Volumes (CKD)

Today's large storage facilities tend to expand to larger CKD Volume capacities, and some installations are running out of z/OS addressable UCB's 64 K limitation disk storage. Because of the 4-digit device addressing limitation, it is necessary to define larger CKD volumes by increasing the number of cylinders per volume.

With the LVS (Large Volume Support) IBM has expanded the CKD volumes to 65,520 cylinders using the existing 16-bit cylinder addressing. These volumes are often referred to as 3390-9 and 64 K cylinder volumes. Products and components like DADSM/CVAF, DSSMSdss, ICKDSF and DFSORT also support 64K cylinder large volume. With LVS support it is possible to address theoretically a capacity of up to 3,964 PB.

Starting with DS8000 LMC 4.1, IBM has developed an even greater volume, an Extended Address Volume (EAV) called 3390 Model A. For that volume the support has been enhanced to expand the volume to 262,668 cylinders, using 28-bit cylinder addressing, which currently has a theoretical limitation of 256 M tracks.

Note: Starting with EAV, we will partially change from track to cylinder addressing.

On the following page we remind you of general addressing method for an Extended Address Volume (EAV). The addressing method is identical for both EAV capacities.

The partial change from track to cylinder addressing creates two address areas on EAV volumes:

- ▶ **Track Managed Space:** The area on an EAV located within the first 65,520 cylinders. Using the 16-bit cylinder addressing allows a theoretical maximum address of 65,535 cylinders. In order to allocate more cylinders, we need to have a *new* format to address the area above 65,520 cylinders.
 - 16-bit cylinder numbers: Existing track address format: CCCCHHHH
 - HHHH: 16-bit track number
 - CCCC: 16-bit track cylinder
- ▶ **Cylinder Managed Space:** The area on an EAV located above the first 65,520 cylinders. This space is allocated in so-called Multicylinder Units (MCU), which currently have a size of 21 cylinders.
 - New cylinder-track address format to address the extended capacity on an EAV:
 - 28-bit cylinder numbers: CCCCcccH
 - H: A 4-bit track number (0-14)
 - ccc: The high order 12 bits of a 28-bit cylinder number
 - CCCC: The low order 16 bits of a 28-bit cylinder number

Components and products, such as DADSM/CVAF, DFSMSdss, ICKDSF, and DFSORT, will also now support 262,668 cylinders.

- ▶ DS8700 on LMC 5.1 and current z/OS limit CKD EAV volume size:
 - 3390 Model A: 1 - 262,668 cylinders (about 223 GB addressable Storage)
 - 3390 Model A: Up to 236 x 3390 Model 1 (Four times the size we have with LVS)
- ▶ Configuration Granularity
 - 1 Cylinder boundary sizes: 1 to 56,520 cylinders

- 1113 Cylinders boundary sizes: 56,763 (51 x 1113) to 262,668 (236 x 1113) cylinders

The size of an existing Mod 3/9/A volume can be increased to its maximum supported size using Dynamic Volume Expansion (DVE). That can be done with the DS CLI command, as shown in Example 5-1.

Example 5-1 Dynamically expand CKD volume

```
dsccli> chckdvol -cap 262268 -captype cyl 9ab0
```

```
Date/Time: 10. Mai 2010 07:52:55 CEST IBM DSCLI Version: 6.5.1.193 DS:
IBM.2107-75KAB25
CMUC00022I chckdvol: CKD Volume 9AB0 successfully modified.
```

Keep in mind that Dynamic Volume Expansion can be done while the volume remains online (to the host system). A VTOC refresh through ICKDSF is a best practice, as it shows the newly added free space. When the relevant volume is in a Copy Services relationship, then that Copy Services relationship must be terminated until both the source and target volumes are at their new capacity, and then the Copy Service pair must be re-established.

The VTOC allocation method for an EAV volume has been changed compared to the VTOC used for LVS volumes. The size of an EAV VTOC index has been increased four-fold, and now has 8,192 blocks instead of 2,048 blocks. Because there is no space left inside the format 1 DSCB, new DSCB formats, Format 8 and Format 9, have been created to protect existing programs from seeing unexpected track addresses. These DSCBs are called extended attribute DSCBs. Format 8 and 9 DSCBs are new for EAV. The existing Format 4 DSCB has been changed also to point to the new format 8 DSCB.

Comparing the volume size we support on LVS with have a 4 times improvement on EAV Release 1.

EAV Release 2

Starting with DS8700 and DS8800 on LMC 6.2 and z/OS V1 R12, we improved the size of an EAV R2 CKD volume to 1062 cylinders using the same addressing method as in the previous EAV R1. This is also referenced to IOCP as an 3390A.

- ▶ DS8000 on LMC 6.2 and current z/OS V1.R12 limit CKD EAV volume size:
 - 3390 Model A: 1 - 1,182,006 cylinders (about 1,004 TB addressable Storage)
 - 3390 Model A: Up to 1062 x 3390 Model 1 (Four times the size we have with EAV R1)
- ▶ Configuration granularity
 - 1 Cylinder boundary sizes: 1 to 56,520 cylinders
 - 1113 Cylinders boundary sizes: 56,763 (51 x 1113) to 1,182,006(1062 x 1113) cylinders

Data set type dependencies on an EAV R2

EAV R2 has the following data set type dependencies:

- ▶ All VSAM, sequential data set types: extended and large format, BDAM, PDS, PDSE, VVDS and BCS are eligible to be placed on the extended address space (EAS) (cylinder managed space) of an EAV R2 volume running on z/OS V1.12 and above:
 - This includes all VSAM data types, such as KSDS, RRDS, ESDS, Linear DS, and also covers DB2, IMS™, CICS®, and zFS data sets.
 - The VSAM data sets placed on an EAV volume can be either SMS or non-SMS managed.

- For EAV Release 2 volume, the following data sets may exist, but are not eligible to have extents in the extended address space (cylinder managed space) in z/OS V1.12:
 - VSAM data sets with incompatible CA sizes
 - VTOC (it is still restricted to the first 64K-1 tracks)
 - VTOC index
 - Page data sets
 - A VSAM data set with imbedded or keyrange attributes, are currently not supported
 - HFS file system
 - SYS1.NUCLEUS

All other data sets are eligible to be placed on an EAV R2 Extended Address Space(EAS).

In the actual releases, you can expand all Mod 3/9/A to an large EAV 2 using DVE and in case of a sequential data set VTOC reformat is been performed automatically, if REFVTOC=ENABLE is been enabled in the DEVSUPxx parmlib member.

Figure 5-19 illustrates the Data set placement on EAV as supported on z/OS V1 R12

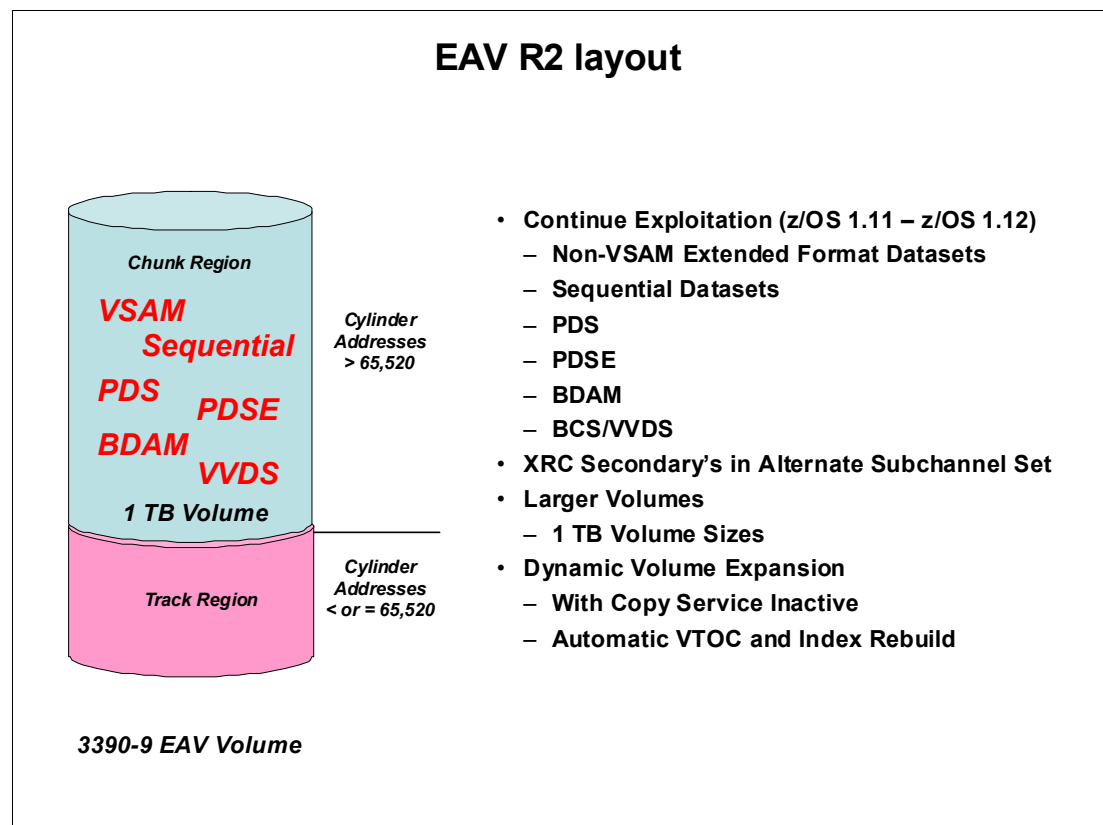


Figure 5-19 Data set placement on EAV supported on z/OS V1 R12.

z/OS prerequisites for EAV volumes

EAV volumes have the following prerequisites:

- ▶ EAV volumes are only supported on z/OS V1.10 and above. If you try to bring an EAV volume online for a system with a pre-z/OS V1.10 release, the EAV Volume will not come online.
- ▶ If you want to use the improvements of EAV R2, its only supported on z/OS V1.12 and above. Keep in mind, a non-VSAM data set allocated with EADSCB on z/OS V1.12 will not be able to be opened on pre z/OS V1.12.
- ▶ After upgrading an Large Volume to a Mod1062 (EAV2 with 1,182,006Cyls) and the system has permission an automatic VTOC refresh and Index rebuild is performed. The permission is granted by REFVTOC=ENABLE in Parmlib member DEVSUPxx. The trigger to the system is a state change interrupt after the volume expansion that is presented by the storage subsystem to the z/OS. Now this function allows you the VTOC is automatically refreshed and the Index is rebuild by the system.
- ▶ There are no additional HCD considerations for the 3390 Model A definitions.
- ▶ On parmlib member IGDSMSxx, the parameter USEEAV(YES) must be set to allow data set allocations on EAV volumes. The default value is NO and prevents allocating data sets to an EAV volume. Example 5-2 shows a sample message that you receive when trying to allocate a data set on EAV volume and USEEAV(NO) is set.

Example 5-2 Message IEF021I with USEEVA set to NO

```
IEF021I TEAM142 STEP1 DD1 EXTENDED ADDRESS VOLUME USE PREVENTED DUE TO SMS USEEAV  
(NO) SPECIFICATION.
```

- ▶ There is a new parameter called Break Point Value (BPV). It determines which size the data set must have to be allocated on a cylinder-managed area. The default for that parameter is 10 cylinders and it can be set on parmlib member IGDSMSxx and in the Storage Group definition (Storage Group BPV overrides system-level BPV). The BPV value can be 0-65520: 0 means that cylinder-managed area is always preferred and 65520 means that a track-managed area is always preferred.

How to identify an EAV 2

Any EAV has more than 65,520 cylinder, in order to address this volumes, the format 4 DSCB was updated to x'FFFE' and DSCB 8+9 is used for cylinder managed address space. Most of the EAV eligible data sets has been modified by SW with EADSCB=YES.

An easy way to identify any EAV being used is to list VTOC Summary in TSO/ISPF option 3.4. Example 5-3 shows the VTOC summary of a 3390 Model A with a size of 1TB CKD usage

Note: Before implementing EAV volumes, have the latest maintenance and z/OS V1.10 and V1.11 coexisting maintenance levels applied. For EAV 2, have the latest maintenance for z/OS V1.12 installed. Refer also to latest PSP information.

Example 5-3 shows a typical TOC summary of an EAV 2 volume.

Example 5-3 TSO/ISPF 3.4 screen for an 1TB EAV volume: VTOC Summary

```
Menu   Reflist  Refmode   Utilities   Help

Volume . : SL9F05
Command ==>

Unit . . : 3390                               Free Space

VTOC Data                                     Total          Tracks          Cyls
Tracks . :      2,069                          Size . . :      14,732,210      982,147
%Used . . :          1                        Largest . :      13,751,640      916,776
Free DSCBS:    103,386                        Free
Extents . . :                               3

Volume Data                                Track Managed      Tracks          Cyls
Tracks . :    17,730,090                      Size . . :      980,570      65,371
%Used . . :          16                      Largest . :      979,070      65,271
Trks/Cyls:      15                          Free
F1=Help   F3=Exit   F12=Cancel
```

When the data set list is displayed, enter either:

- "/" on the data set list command field for the command prompt pop-up,
 - an ISPF line command, the name of a TSO command, CLIST, or REXX exec, or
 - "=" to execute the previous command.
-

EAV R2 Migration considerations

Consider the following items:

► Assistance:

Migration assistance will be provided through use of the Application Migration Assistance Tracker. For more details about Assistance Tracker, see APAR II13752, which can be found at the following address:

<http://www.ibm.com/support/docview.wss?uid=isg1II13752>

► Recommended actions:

- Review your programs and take a look at the calls for the macros OBTAIN, REALLOC, CVAFDIR, CVAFSEQ, CVAFDSM, and CVAFFILT. Those macros were changed and you need to update your program to reflect those changes.
- Look for programs that calculate volume or data set size by any means, including reading a VTOC or VTOC index directly with a BSAM or EXCP DCB. This task is important because now we have new values returning for the volume size.
- Review your programs and look for EXCP and STARTIO macros for DASD channel programs and other programs that examine DASD channel programs or track addresses. Now that we have new addressing mode, programs must be updated.
- Look for programs that examine any of the many operator messages that contain a DASD track, block address, data set, or volume size. The messages now show new values.

► Migrating data:

- Define new EAVs by creating them on the DS8700 or expanding existing volumes using Dynamic Volume Expansion.

- Add new EAV volumes to storage groups and storage pools, and update ACS routines.
- Copy data at the volume level: TDMF®, DFSMSdss, PPRC, DFSMS, Copy Services Global Mirror, Metro Mirror, Global Copy, and FlashCopy.
- Copy data at the data set level: SMS attrition, LDMF, DFSMSdss, and DFSMSHsm.
- With z/OS V1.12 all data set types are currently good volume candidates for EAV R2 except:
a view, that are already mentioned on Page 124.

DFSMSdss, DFSMSHsm, DFSORT a.s.o are now eligible for EAV R2 on z/OS V1.12



IBM System Storage DS8000

Copy Services overview

This chapter provides an overview of the Copy Services functions that are available with the DS8000 series models, including Remote Mirror and Copy functions and Point-in-Time Copy functions.

These functions make the DS8000 series a key component for disaster recovery solutions, data migration activities, and for data duplication and backup solutions.

This chapter covers the following topics

- ▶ Introduction to Copy Services
- ▶ FlashCopy and FlashCopy SE
- ▶ Remote Pair FlashCopy (Preserve Mirror)
- ▶ Remote Mirror and Copy:
 - Metro Mirror
 - Global Copy
 - Global Mirror
 - Metro/Global Mirror
 - z/OS Global Mirror
 - z/OS Metro/Global Mirror

The information provided in this chapter is only an overview. It is covered to a greater extent and in more detail in the following IBM Redbooks and IBM Redpapers™ publications:

- ▶ *IBM System Storage DS8000: Copy Services for Open Systems*, SG24-6788
- ▶ *IBM System Storage IBM System Storage DS8000: Copy Services for IBM System z*, SG24-6787
- ▶ *IBM System Storage DS8000 Series: IBM FlashCopy SE*, REDP-4368

6.1 Copy Services

Copy Services is a collection of functions that provide disaster recovery, data migration, and data duplication functions. With the Copy Services functions, for example, you can create backup data with little or no disruption to your application, and you can back up your application data to the remote site for disaster recovery.

The Copy Services functions run on the DS8800 storage unit and support open systems and System z environments. They are also supported on other DS8000 family models.

DS8000 Copy Services functions

Copy Services in the DS8000 include the following optional licensed functions:

- ▶ IBM System Storage FlashCopy and IBM FlashCopy SE, which are point-in-time copy functions
- ▶ Remote mirror and copy functions, which include:
 - IBM System Storage Metro Mirror, previously known as synchronous PPRC
 - IBM System Storage Global Copy, previously known as PPRC eXtended Distance
 - IBM System Storage Global Mirror, previously known as asynchronous PPRC
 - IBM System Storage Metro/Global Mirror, a three-site solution to meet the most rigorous business resiliency needs
 - For migration purposes on a RPQ base, consider IBM System Storage Metro/Global Copy. Understand that this combination of Metro Mirror and Global Copy is not suited for disaster recovery solutions; it is only intended for migration purposes.
- ▶ Additionally for IBM System z users, the following options are available:
 - z/OS Global Mirror, previously known as eXtended Remote Copy (XRC)
 - z/OS Metro/Global Mirror, a three-site solution that combines z/OS Global Mirror and Metro Mirror

Many design characteristics of the DS8000, its data copy and mirror capabilities, and features contribute to the protection of your data, 24 hours a day and seven days a week.

Copy Services management interfaces

You control and manage the DS8000 Copy Services functions using the following interfaces:

- ▶ DS Storage Manager, the graphical user interface of the DS8000 (DS GUI)
- ▶ DS Command-Line Interface (DS CLI), which provides a set commands that cover all Copy Service functions and options
- ▶ Tivoli Storage Productivity Center for Replication (TPC-R), which allows you to manage large Copy Services implementations easily and provides data consistency across multiple systems
- ▶ DS Open Application Programming Interface (DS Open API)

System z users can also use the following interfaces:

- ▶ TSO commands
- ▶ ICKDSF utility commands
- ▶ ANTRQST application programming interface (API)
- ▶ DFSMSdss utility

6.2 FlashCopy and FlashCopy SE

FlashCopy and FlashCopy SE provide the capability to create copies of logical volumes with the ability to access both the source and target copies immediately. Such kind of copies are called point-in-time copies.

FlashCopy is an optional licensed feature of the DS8000. Two variations of FlashCopy are available:

- ▶ Standard FlashCopy, also referred to as the Point-in-Time Copy (PTC) *licensed function*
- ▶ FlashCopy SE *licensed function*

To use FlashCopy, you must have the corresponding licensed function indicator feature in the DS8800, and you must acquire the corresponding DS8000 function authorization with the adequate feature number license in terms of physical capacity. For details about feature and function requirements, see 10.1, “IBM System Storage DS8000 licensed functions” on page 252.

In this section, we discuss the FlashCopy and FlashCopy SE basic characteristics and options.

6.2.1 Basic concepts

FlashCopy creates a point-in-time copy of the data. When a FlashCopy operation is invoked, it takes only a few seconds to establish the FlashCopy relationship, consisting of the source and target volume pairing and the necessary control bitmaps. Thereafter, a copy of the source volume is available as though all the data had been copied. As soon as the pair has been established, you can read and write to both the source and target volumes.

Two variations of FlashCopy are available:

- ▶ Standard FlashCopy uses a normal volume as target volume. This target volume has to have at least the same size as the source volume and that space is fully allocated in the storage system.
- ▶ FlashCopy Space Efficient (SE) uses Space Efficient volumes (see 5.2.6, “Space Efficient volumes” on page 114) as FlashCopy targets. A Space Efficient target volume has a virtual size that is at least that of the source volume. However, space is not allocated for this volume when the volume is created and the FlashCopy initiated. Space is allocated just for updated tracks only when the source or target volume are written.

Be aware that both FlashCopy and FlashCopy SE can coexist on a DS8000.

Tip: In this chapter, *track* means a piece of data in the DS8800. The DS8000 uses the concept of logical tracks to manage Copy Services functions.

Figure 6-1 on page 140 and the subsequent section explain the basic concepts of a standard FlashCopy.

If you access the source or the target volumes while the FlashCopy relation exists, I/O requests are handled as follows:

- ▶ Read from the source volume
When a read request goes to the source, data is directly read from there.
- ▶ Read from the target volume
When a read request goes to the target volume, FlashCopy checks the bitmap and:

- If the requested data has already been copied to the target, it is read from there.
 - If the requested data has not been copied yet, it is read from the source.
- Write to the source volume
- When a write request goes to the source, the data is first written to the cache and persistent memory (write cache). Later, when the data is destaged to the physical extents of the source volume, FlashCopy checks the bitmap for the location that is to be overwritten and:
- If the point-in-time data was already copied to the target, the update is written to the source directly.
 - If the point-in-time data has not been copied to the target yet, it is now copied immediately and only *then* is the update written to the source.

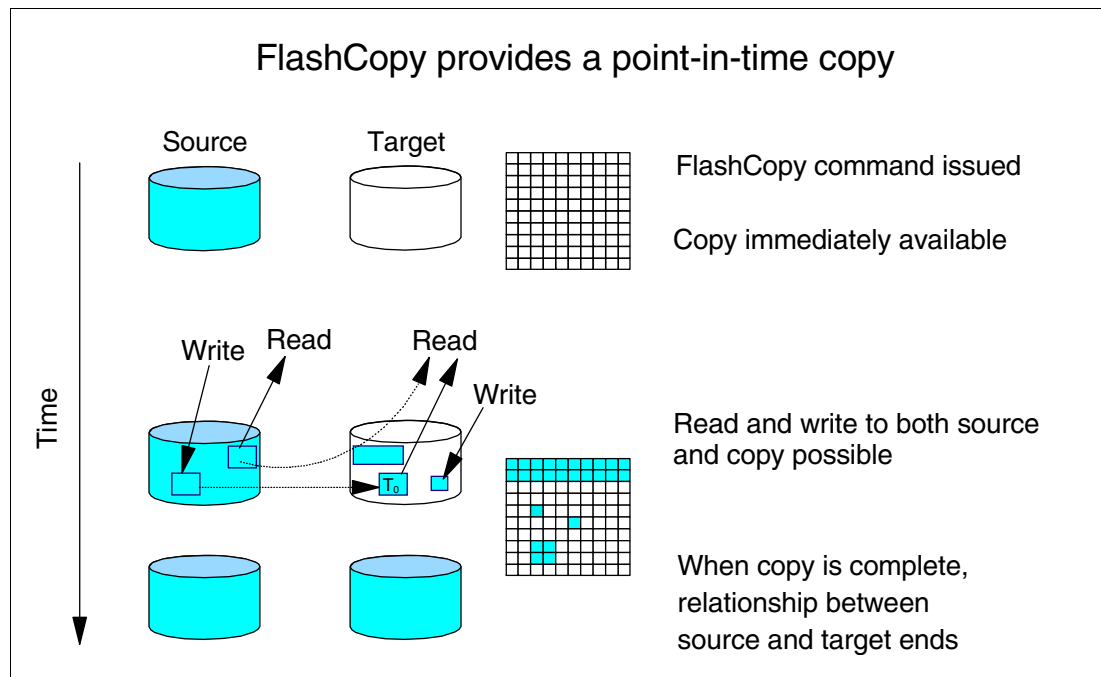


Figure 6-1 FlashCopy concepts

- Write to the target volume
- Whenever data is written to the target volume while the FlashCopy relationship exists, the storage system checks the bitmap and updates it if necessary. This way, FlashCopy does not overwrite data that was written to the target with point-in-time data.

The FlashCopy background copy

By default, standard FlashCopy invokes a background copy process that copies all point-in-time data to the target volume. After the completion of this process, the FlashCopy relation ends and the target volume becomes independent of the source.

The background copy can slightly impact application performance because the physical copy needs storage resources. The impact is minimal because host I/O always has higher priority than the background copy.

No background copy option

A standard FlashCopy relationship can also be established with the NOCOPY option. With this option FlashCopy does not initiate a background copy. Point-in-time data is copied only when required due to an update to either source or target. This eliminates the impact of the background copy.

This option is useful in the following situations:

- ▶ When the target will not be needed as an independent volume
- ▶ When repeated FlashCopy operations to the same target are expected

FlashCopy SE is automatically invoked with the NOCOPY option, because the target space is not allocated and the available physical space is smaller than the size of the volume. A full background copy would contradict the concept of space efficiency.

6.2.2 Benefits and use

The point-in-time copy created by FlashCopy is typically used where you need a copy of the production data produced with little or no application downtime. Use cases for the point-in-time copy created by FlashCopy include online backup, testing new applications, or creating a copy of transactional data for data mining purposes. To the host or application, the target looks exactly like the original source. It is an instantly available, binary copy.

IBM FlashCopy SE is designed for temporary copies. FlashCopy SE is optimized for use cases where only about 5% of the source volume data is updated during the life of the relationship. If more than 20% of the source data is expected to change, standard FlashCopy would likely be the better choice.

Standard FlashCopy will generally have superior performance to FlashCopy SE. If performance on the source or target volumes is important, using standard FlashCopy is a more desirable choice.

Scenarios where using IBM FlashCopy SE is a good choice include:

- ▶ Creating a temporary copy and backing it up to tape.
- ▶ Creating temporary point-in-time copies for application development or DR testing.
- ▶ Performing regular online backup for different points in time.
- ▶ FlashCopy target volumes in a Global Mirror (GM) environment. Global Mirror is explained in 6.3.3, “Global Mirror” on page 146.

In all scenarios, the write activity to both source and target is the crucial factor that decides whether FlashCopy SE can be used.

6.2.3 FlashCopy options

FlashCopy provides many additional options and functions. We explain the following options and capabilities in this section:

- ▶ Incremental FlashCopy (refresh target volume)
- ▶ Persistent FlashCopy
- ▶ Data Set FlashCopy
- ▶ Multiple Relationship
- ▶ Consistency Group FlashCopy
- ▶ FlashCopy on existing Metro Mirror or Global Copy primary
- ▶ Inband commands over remote mirror link

Incremental FlashCopy (refresh target volume)

Refresh target volume provides the ability to refresh a FlashCopy relation without copying all data from source to target again. When a subsequent FlashCopy operation is initiated, only the changed tracks on both the source and target need to be copied from the source to the target. The direction of the refresh can also be reversed, from (former) target to source.

In many cases only a small percentage of the entire data is changed in a day. In this situation, you can use this function for daily backups and save the time for the physical copy of FlashCopy.

Incremental FlashCopy requires the background copy and the Persistent FlashCopy option to be enabled.

Persistent FlashCopy

Persistent FlashCopy allows the FlashCopy relationship to remain even after the copy operation completes. You must explicitly delete the relationship to terminate it.

Data Set FlashCopy

Data Set FlashCopy allows you to create a point-in-time copy of individual data sets instead of complete volumes in an IBM System z environment.

Multiple Relationship FlashCopy

FlashCopy allows a source to have relationships with up to 12 targets simultaneously. A usage case for this feature is create regular point-in-time copies as online backups or time stamps. Only one of the multiple relations can be incremental.

Consistency Group FlashCopy

Consistency Group FlashCopy allows you to freeze and temporarily queue I/O activity to a volume. Consistency Group FlashCopy helps you to create a consistent point-in-time copy without quiescing the application across multiple volumes, and even across multiple storage units.

Consistency Group FlashCopy ensures that the order of dependent writes is always maintained and thus creates host-consistent copies, not application-consistent copies. The copies have power-fail or crash level consistency. To recover an application from Consistency Group FlashCopy target volumes, you need to perform the same kind of recovery as after a system crash.

FlashCopy on existing Metro Mirror or Global Copy primary

This option allows you to establish a FlashCopy relationship where the target is a Metro Mirror or Global Copy primary volume. This enables you to create full or incremental point-in-time copies at a local site and then use remote mirroring to copy the data to the remote site.

Tip: You cannot FlashCopy from a source to a target if the target is also a Global Mirror primary volume.

Metro Mirror and Global Copy are explained in 6.3.1, “Metro Mirror” on page 145 and in 6.3.2, “Global Copy” on page 146.

Inband commands over remote mirror link

In a remote mirror environment, commands to manage FlashCopy at the remote site can be issued from the local or intermediate site and transmitted over the remote mirror Fibre

Channel links. This eliminates the need for a network connection to the remote site solely for the management of FlashCopy.

Tip: This function is available by using the DS CLI, TSO, and ICKDSF commands, but not by using the DS Storage Manager GUI.

6.2.4 FlashCopy SE-specific options

Most options for standard FlashCopy (see 6.2.3, “FlashCopy options” on page 141) work identically for FlashCopy SE. The options that differ are discussed in this section.

Incremental FlashCopy

Because Incremental FlashCopy implies an initial full volume copy and a full volume copy is not possible in an IBM FlashCopy SE relationship, Incremental FlashCopy is not possible with IBM FlashCopy SE.

Data Set FlashCopy

FlashCopy SE relationships are limited to full volume relationships. As a result, data set level FlashCopy is not supported within FlashCopy SE.

Multiple Relationship FlashCopy SE

Standard FlashCopy supports up to 12 relationships per source volume and one of these relationships can be incremental. A FlashCopy onto a Space Efficient volume has a certain overhead because additional tables and pointers have to be maintained. Therefore it might be advisable to avoid utilizing all 12 possible relations.

6.2.5 Remote Pair FlashCopy

Remote Pair FlashCopy or Preserve Mirror is available for the DS8700 and DS8800 with Licensed Machine Code (LMC) level 7.6.1.xx.xx. Remote Pair FlashCopy is also available with the DS8100 and DS8300, but only on specific firmware (release 4.25).

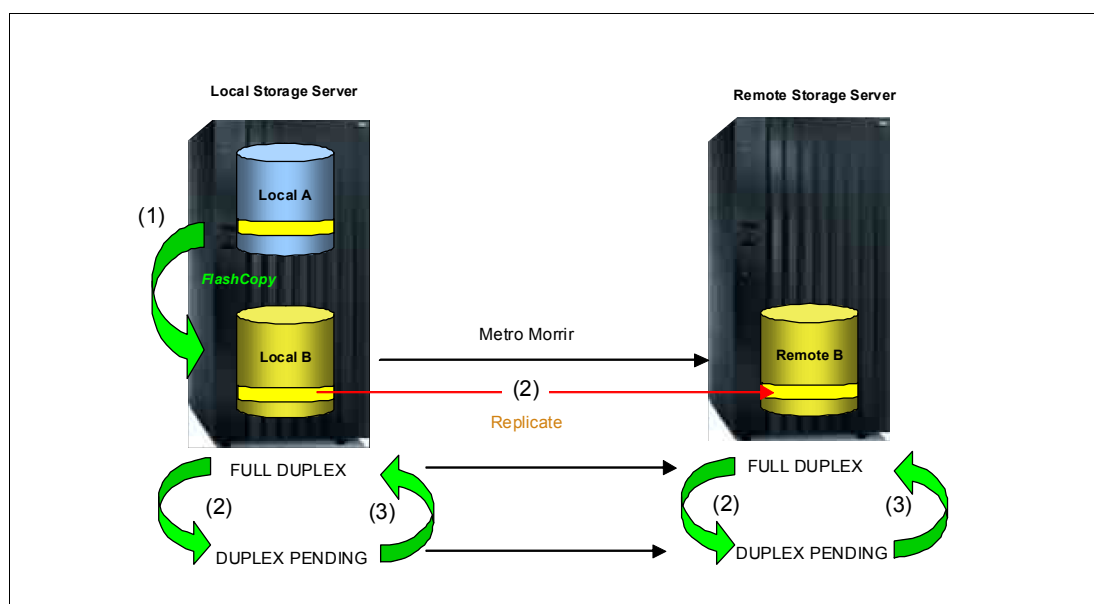


Figure 6-2 The FlashCopy target is also an MM or GC source

Remote Pair FlashCopy or Preserve Mirror overcomes the shortcomings of the previous solution to FlashCopy onto a Metro Mirror source volume. Figure 6-2 on page 143 illustrates the following behavior:

1. FlashCopy is issued at Local A volume, which starts a FlashCopy relationship between the Local A and the Local B volumes.
2. As soon as the FlashCopy operation starts and replicates the data from Local A to Local B volume, the Metro Mirror volume pair status changes from FULL DUPLEX to DUPLEX PENDING. During the DUPLEX PENDING window, the Remote Volume B does not provide a defined state regarding its data status and is unusable from a recovery viewpoint.
3. After FlashCopy finishes replicating the data from Local A volume to Local B volume, the Metro Mirror volume pair changes its status from DUPLEX PENDING back to FULL DUPLEX. The remote Volume B provides a recoverable state and can be used in case of an planned or unplanned outage at the local site.

As the name implies, Preserve Mirror does preserve the existing Metro Mirror status of FULL DUPLEX. Figure 6-3 shows this approach, which guarantees that there is no discontinuity of the disaster recovery readiness.

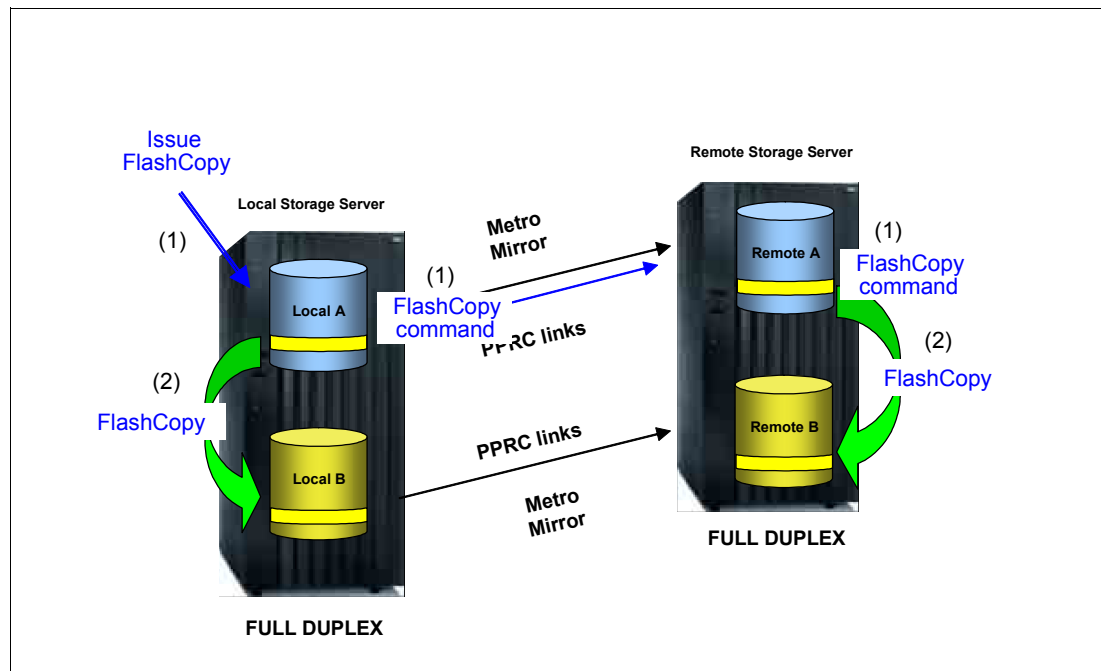


Figure 6-3 Remote Pair FlashCopy preserves the Metro Mirror FULL DUPLEX state

The behavior consists of the following:

1. The FlashCopy command is issued by an application or by you to the Local A volume with Local B volume as the FlashCopy target. The DS8000 firmware propagates the FlashCopy command through the PPRC links from the Local Storage Server to the Remote Storage Server. This inband propagation of a Copy Services command is only possible for FlashCopy commands.
2. Independently of each other, the Local Storage Server and the Remote Storage Server then execute the FlashCopy operation. The Local Storage Server coordinates the activities at the end and takes action if the FlashCopies do not succeed at both Storage Servers. Figure 6-3 shows an example where Remote Pair FlashCopy might have the most relevance: A data set level FlashCopy in a Metro Mirror CKD volumes environment

where all participating volumes are replicated. Usually the user has no influence where the newly allocated FlashCopy target data set is going to be placed. The key item of this configuration is that disaster recovery protection is not exposed at any time and FlashCopy operations can be freely taken within the disk storage configuration. If using Remote Pair FlashCopy, the Metro Mirror volume pair status keeps FULL DUPLEX, the DR viewpoint and GDPS® recovery standpoint is fully assured.

For a more detailed description about Remote Pair FlashCopy, refer to *IBM System Storage DS8000: Remote Pair FlashCopy (Preserve Mirror)*, REDP-4504.

6.3 Remote Mirror and Copy

The Remote Mirror and Copy functions of the DS8000 are a set of flexible data mirroring solutions that allow replication between volumes on two or more disk storage systems. These functions are used to implement remote data backup and disaster recovery solutions.

The Remote Mirror and Copy functions are optional licensed functions of the DS8000 that include:

- ▶ Metro Mirror
- ▶ Global Copy
- ▶ Global Mirror
- ▶ Metro/Global Mirror

In addition, System z users can use the DS8000 for:

- ▶ z/OS Global Mirror
- ▶ z/OS Metro/Global Mirror

In the following sections, we discuss these Remote Mirror and Copy functions.

For a more detailed and extensive discussion about these topics, refer to the IBM Redbooks publications listed in “Related publications” on page 469.

Licensing requirements

To use any of these Remote Mirror and Copy optional licensed functions, you must have the corresponding licensed function indicator feature in the DS8000, and you must acquire the corresponding DS8800 function authorization with the adequate feature number license in terms of physical capacity. For details about feature and function requirements, see 10.1, “IBM System Storage DS8000 licensed functions” on page 252.

Also, consider that certain of the remote mirror solutions, such as Global Mirror, Metro/Global Mirror, or z/OS Metro/Global Mirror, integrate more than one licensed function. In this case, you need to have all of the required licensed functions.

6.3.1 Metro Mirror

Metro Mirror, previously known as Synchronous Peer-to-Peer Remote Copy (PPRC), provides real-time mirroring of logical volumes between two DS8000s, or any other combination of DS8100, DS8300, DS6800, and ESS800, that can be located up to 300 km from each other. It is a synchronous copy solution where a write operation must be carried out on both copies, at the local and remote sites, before it is considered complete.

Figure 6-4 on page 146 illustrates the basic operational characteristics of Metro Mirror.

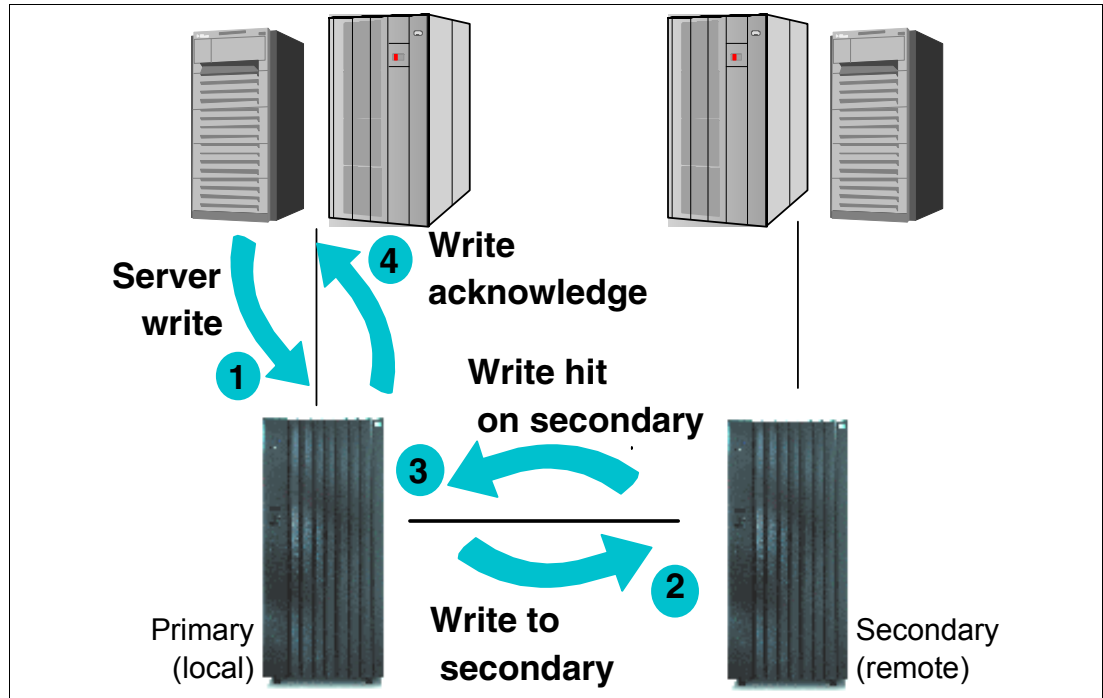


Figure 6-4 Metro Mirror basic operation

6.3.2 Global Copy

Global Copy, previously known as Peer-to-Peer Remote Copy eXtended Distance (PPRC-XD), copies data non-synchronously and over longer distances than is possible with Metro Mirror. When operating in Global Copy mode, the source does not wait for copy completion on the target before acknowledging a host write operation. Therefore, the host is not impacted by the Global Copy operation. Write data is sent to the target as the connecting network allows and independent of the order of the host writes. This makes the target data lag behind and be inconsistent during normal operation.

You have to take extra steps to make Global Copy target data usable at specific points in time. These steps depend on the purpose of the copy.

Here are two examples:

- **Data migration**

You can use Global Copy to migrate data over long distances. When you want to switch from old to new data, you have to stop the applications on the old site, tell Global Copy to synchronize the data, and wait until it is finished.

- **Asynchronous mirroring**

Global Mirror, the IBM solution for asynchronous data replication, uses Global Copy to transport the data over the long distance network. Periodic FlashCopies are used to provide consistent data points. Global Mirror is described in the next section.

6.3.3 Global Mirror

Global Mirror, previously known as Asynchronous PPRC, is a two-site, long distance, asynchronous, remote copy technology for both System z and Open Systems data. This solution integrates the Global Copy and FlashCopy technologies. With Global Mirror, the data

that the host writes at the local site is asynchronously mirrored to the storage unit at the remote site. With special management steps, under control of the local master storage unit, a consistent copy of the data is automatically maintained and periodically updated on the storage unit at the remote site.

Global Mirror benefits

- ▶ Support for virtually unlimited distances between the local and remote sites, with the distance typically limited only by the capabilities of the network and the channel extension technology. This unlimited distance enables you to choose your remote site location based on business needs and enables site separation to add protection from localized disasters.
- ▶ A consistent and restartable copy of the data at the remote site, created with minimal impact to applications at the local site.
- ▶ Data currency where, for many environments, the remote site lags behind the local site typically 3 to 5 seconds, minimizing the amount of data exposure in the event of an unplanned outage. The actual lag in data currency that you experience will depend upon a number of factors, including specific workload characteristics and bandwidth between the local and remote sites.
- ▶ Dynamic selection of the desired recovery point objective (RPO), based upon business requirements and optimization of available bandwidth.
- ▶ Session support: data consistency at the remote site is internally managed across up to eight storage units located at both the local site and the remote site.
- ▶ Efficient synchronization of the local and remote sites with support for failover and failback operations, helping to reduce the time that is required to switch back to the local site after a planned or unplanned outage.

How Global Mirror works

Figure 6-5 illustrates the basic operational characteristics of Global Mirror.

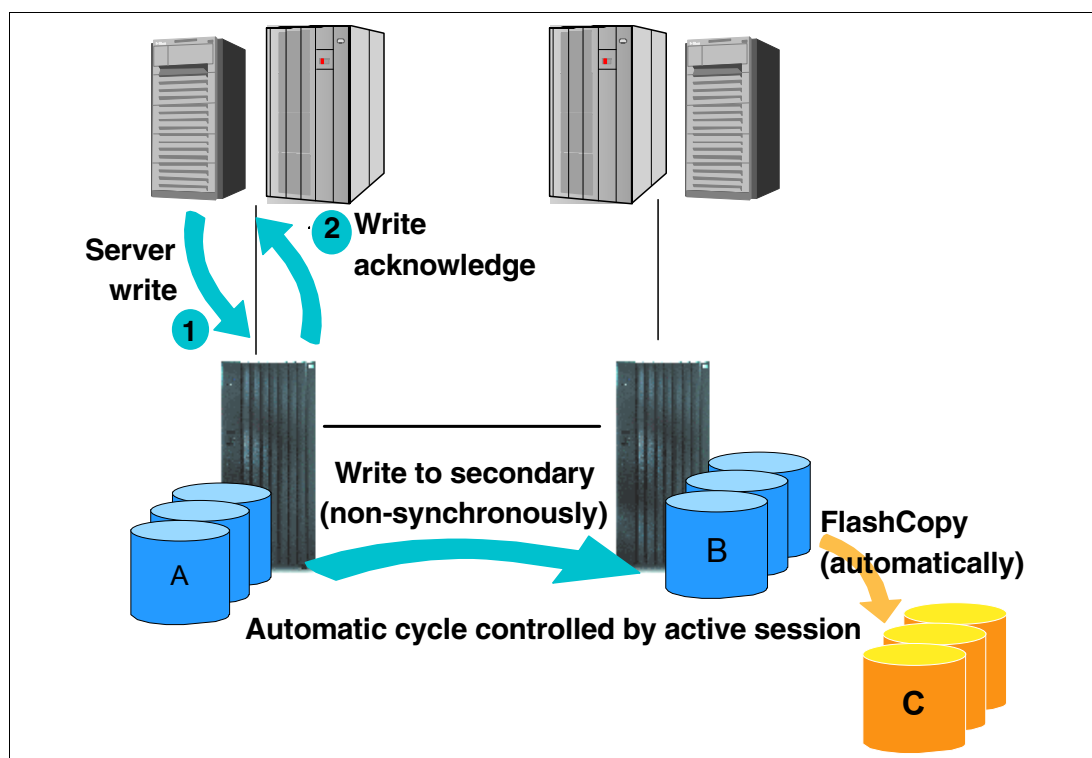


Figure 6-5 Global Mirror basic operation

The A volumes at the local site are the production volumes and are used as Global Copy primaries. The data from the A volumes is replicated to the B volumes using Global Copy. At a certain point in time, a Consistency Group is created from all the A volumes, even if they are located in separate storage units. This has little application impact, because the creation of the Consistency Group is quick (on the order of a few milliseconds).

After the Consistency Group is created, the application writes can continue updating the A volumes. The missing increment of the consistent data is sent to the B volumes using the existing Global Copy relations. After all data has reached the B volumes, Global Copy is halted for brief period while Global Mirror creates a FlashCopy from the B to the C volumes. These now contain a consistent set of data at the secondary site.

The data at the remote site is current within 3 to 5 seconds, but this recovery point depends on the workload and bandwidth available to the remote site.

With its efficient and autonomic implementation, Global Mirror is a solution for disaster recovery implementations where a consistent copy of the data needs to be available at all times at a remote location that can be separated by a long distance from the production site.

6.3.4 Metro/Global Mirror

Metro/Global Mirror is a three-site, multi-purpose, replication solution for both System z and Open Systems data. Local site (site A) to intermediate site (site B) provides high availability replication using Metro Mirror, and intermediate site (site B) to remote site (site C) supports long distance disaster recovery replication with Global Mirror. See Figure 6-6.

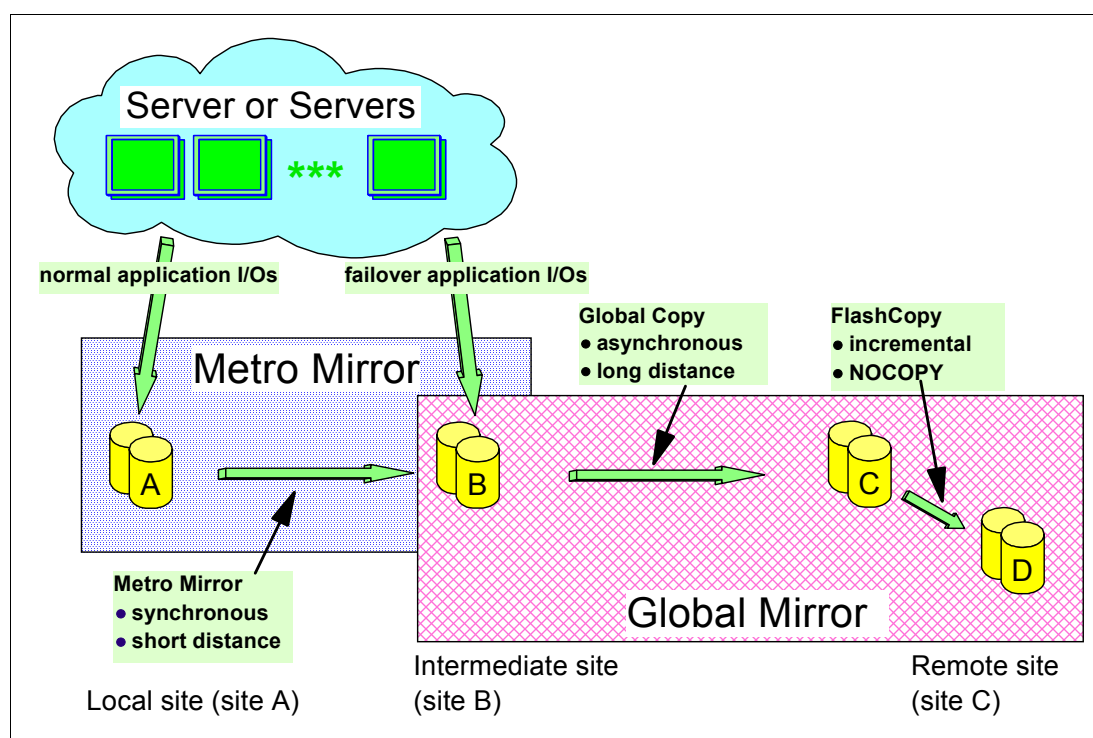


Figure 6-6 Metro/Global Mirror elements

Both Metro Mirror and Global Mirror are well established replication solutions. Metro/Global Mirror combines Metro Mirror and Global Mirror to incorporate the best features of the two solutions:

► Metro Mirror

- Synchronous operation supports zero data loss.
- The opportunity to locate the intermediate site disk systems close to the local site allows use of intermediate site disk systems in a high availability configuration.

Tip: Metro Mirror can be used for distances of up to 300 km. However, when used in a Metro/Global Mirror implementation, a shorter distance for the Metro Mirror connection is more appropriate to effectively guarantee high availability of the configuration.

► Global Mirror

- Asynchronous operation supports long distance replication for disaster recovery.
- The Global Mirror methodology has no impact to applications at the local site.
- This solution provides a recoverable, restartable, and consistent image at the remote site with an RPO, typically within 3 to 5 seconds.

6.3.5 Multiple Global Mirror sessions

With the DS8000 and Licensed Machine Code (LMC) level 7.6.1.xx. or later, you are no longer limited to one Global Mirror session within a storage system (SFI). Up to 32 Global Mirror hardware sessions can be supported within the same DS8000 (Figure 6-7).

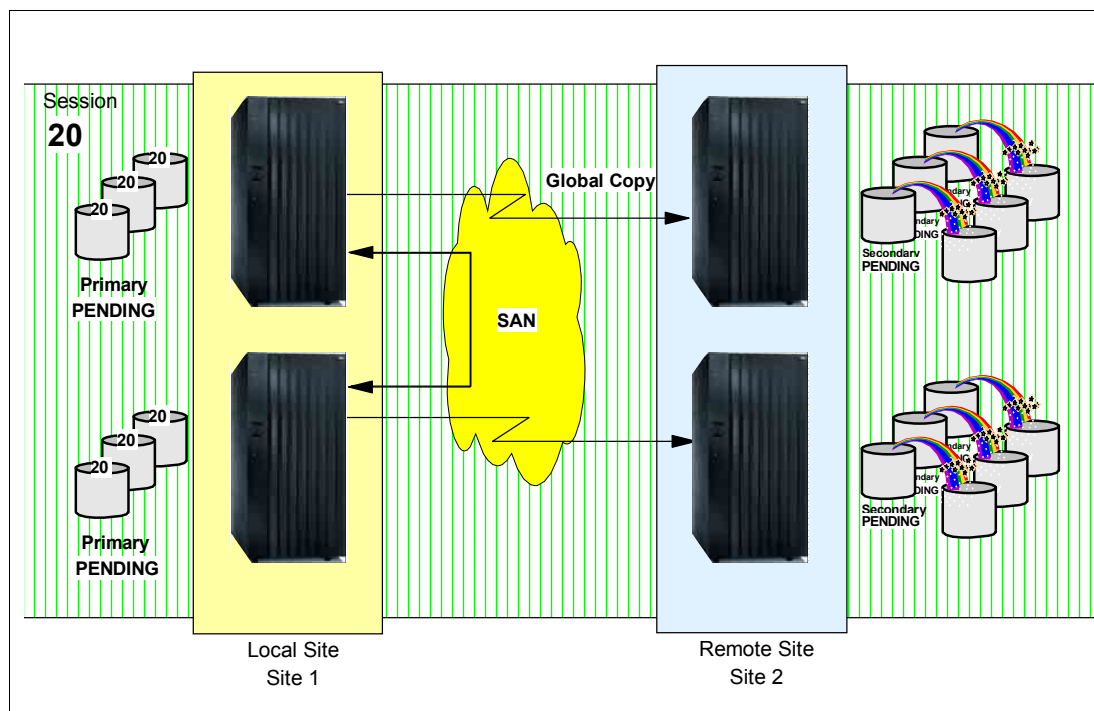


Figure 6-7 Single GM hardware session support

The session in Figure 6-7 is meant to be a GM master session that controls a GM session. A GM session is identified by a GM session id (in this example, number 20). The session ID applies to any LSS at Site 1, containing Global Copy primary volumes that belong to session

20. The two storage systems configuration consists of a GM master in the DS8000 at the bottom and a subordinate DS8000 that contains also Global Copy primary volumes that belong to session 20. The GM master controls the subordinate through PPRC FCP-based paths between both DS8000 storage systems. Consistency is provided across all primary subsystems.

With the DS8100 and DS8300, it is not possible to create more that one GM session per GM master. Potential impacts with such a single GM session are shown in Figure 6-8.

Assume a disk storage consolidated environment which is commonly used by various application servers. To provide good performance, all volumes are spread across the primary DS8300s. For disaster recovery purposes, a remote site exists with corresponding DS8300s and the data volumes are replicated through a Global Mirror session with the Global Mirror master function in a DS8100 or a DS8300.

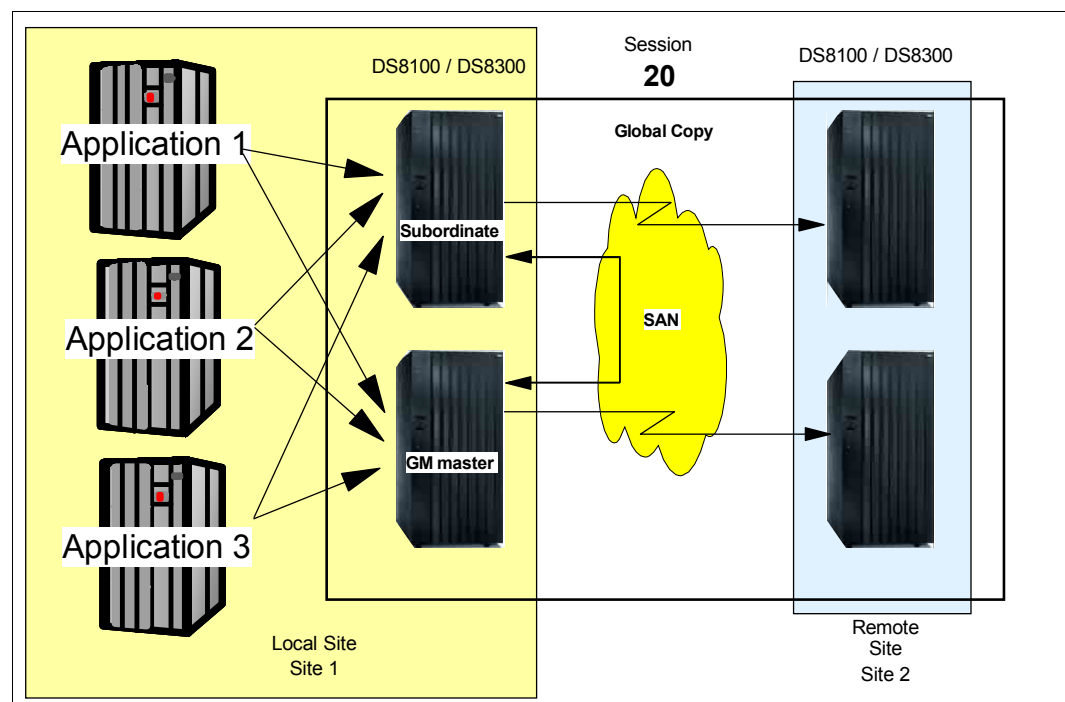


Figure 6-8 Multiple applications - single GM session

When server 2 with Application 2 fails and the participating volumes that are connected to Application 2 are not accessible from the servers in the remote site or Site 2, the entire GM session 20 must fail over to the remote site

Figure 6-9 on page 151 shows the impact on the other two applications, Application 1 and Application 3. Because there is only one GM session possible with a DS8100 or DS8300 on one SFI, the entire session has to be failed over to the remote site to restart Application 2 on the backup server at the remote site. The other two servers with Application 1 and Application 3 are affected and must also be swapped over to the remote site.

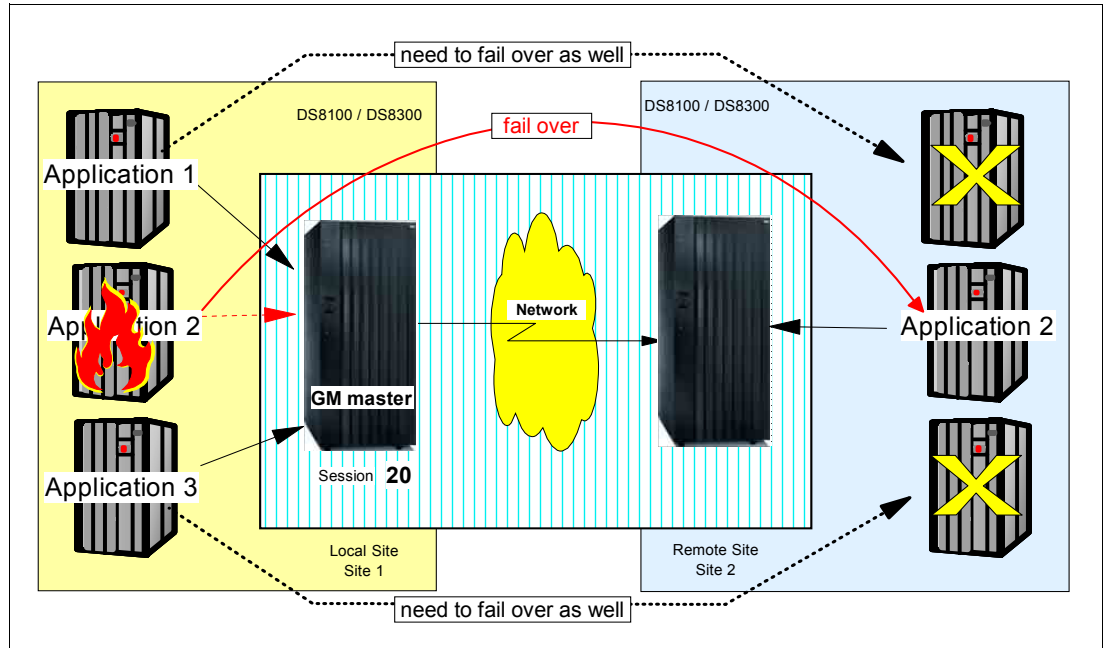


Figure 6-9 Multiple applications - single GM sessions - fail over requirements

This implies services interruption not only to the failed server with Application 2, but also service impacts to Application 1 and Application 3 which need to shut down in Site 1 as well and restart in Site 2 after the GM session fail over process is completed.

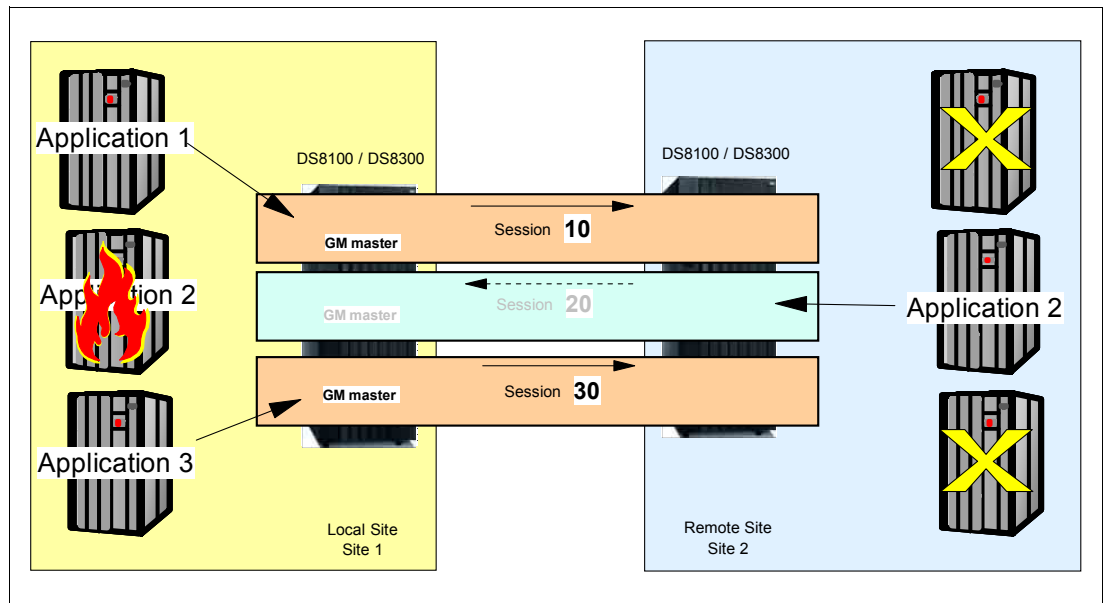


Figure 6-10 DS8000 provides multiple GM master sessions support within R6.1

Figure 6-10 shows the same server configuration. However, the storage subsystems DS8100 or DS8300 are exchanged by DS8700 or DS8800 with release 6.1 on LMC: 7.6.1.xx.xx. This allows you to use up to 32 dedicated GM master sessions. In the example in Figure 6-10 Application 1 is connected to volumes that reside in LSS number 00 to LSS number 3F. Application 2 connects to volumes in LSS 40-7F and the server with Application 3 connects to volumes in LSS 80-BF.

Each set of volumes on an Application server resides in its own GM session, which is controlled by the concerned GM master session within the same DS8700. Only a GM session can reside in a certain LSS, which you need to consider when planning on how to divide up volumes into separate GM sessions.

Now when the Application 2 server fails, only GM session 20 is failed over to the remote site and the concerned server in Site 2 restarts with Application 2 after the fail over process completes.

DS8000 and release 6.1 and later allows for a finer granularity and dedicated recovery actions. This is not uncommon because different applications might have different RPO requirements. The ability to fail over only the configuration of a failing server or applications does improve the availability of other applications compared to the situation before DS8000.

An installation can now have one or more test sessions in parallel with one or more productive GM sessions within the same SFI to test and gain experience on possible management tools and improvements.

Notice that the basic management of a GM session does not change. The GM session builds on the existing Global Mirror technology and microcode of the DS8000.

6.3.6 z/OS Global Mirror

z/OS Global Mirror, previously known as eXtended Remote Copy (XRC), is a copy function available for the z/OS operating systems. It involves a System Data Mover (SDM) that is found only in z/OS. z/OS Global Mirror maintains a consistent copy of the data asynchronously at a remote location, and can be implemented over unlimited distances. It is a combined hardware and software solution that offers data integrity and data availability and can be used as part of business continuance solutions, for workload movement, and for data migration. z/OS Global Mirror function is an optional licensed function of the DS8000.

Figure 6-11 illustrates the basic operational characteristics of z/OS Global Mirror.

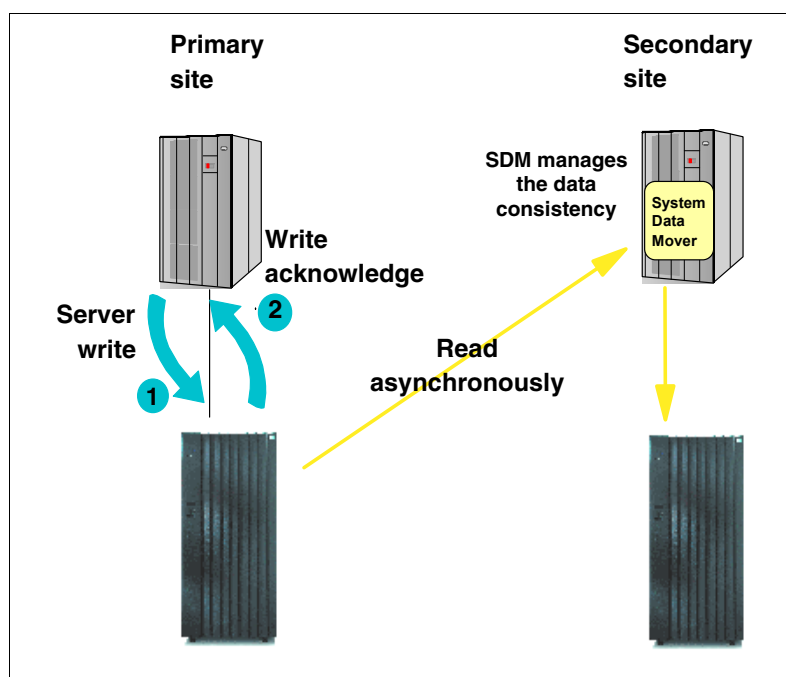


Figure 6-11 z/OS Global Mirror basic operations

z/OS Global Mirror on zIIP

The IBM z9® Integrated Information Processor (zIIP) is a special engine available for System z since the z9 generation. z/OS now provides the ability to utilize these processors to handle eligible workloads from the System Data Mover (SDM) in an z/OS Global Mirror (zGM) environment.

Given the appropriate hardware and software, a range of zGM workload can be offloaded to zIIP processors. The z/OS software must be at V1.8 and later with APAR OA23174, specifying zGM PARMLIB parameter zIIPEnable(YES).

6.3.7 z/OS Metro/Global Mirror

This mirroring capability implements z/OS Global Mirror to mirror primary site data to a location that is a long distance away and also uses Metro Mirror to mirror primary site data to a location within the metropolitan area. This enables a z/OS three-site high availability and disaster recovery solution for even greater protection against unplanned outages.

Figure 6-12 illustrates the basic operational characteristics of a z/OS Metro/Global Mirror implementation.

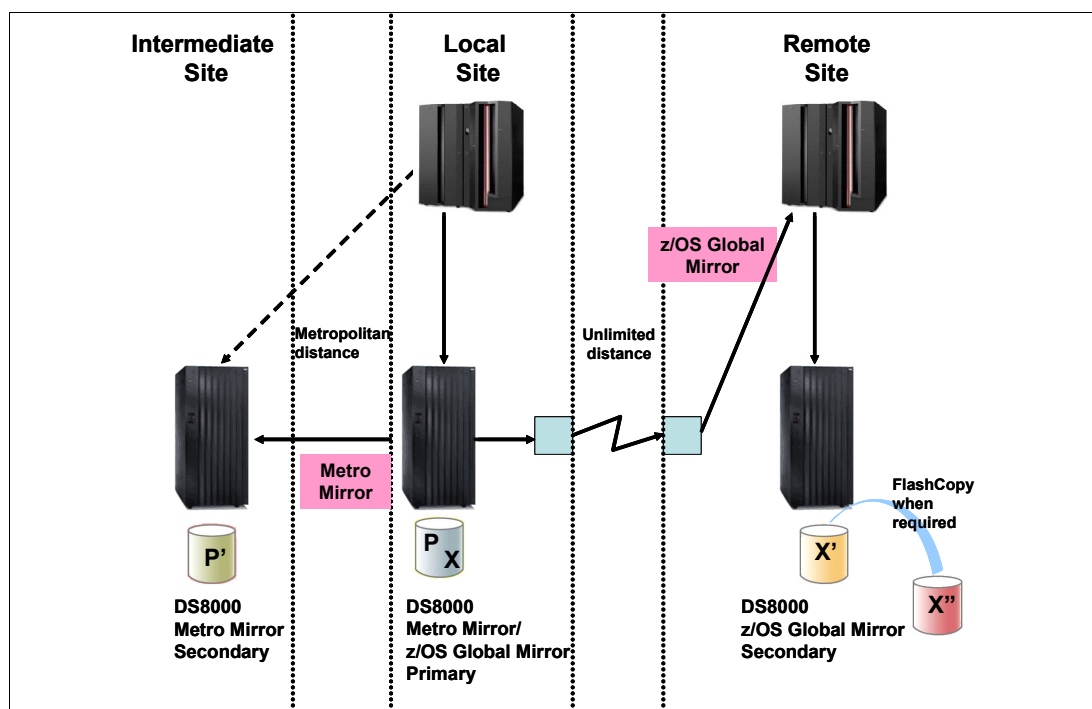


Figure 6-12 z/OS Metro/Global Mirror

6.3.8 Summary of Remote Mirror and Copy function characteristics

In this section, we summarize the use of and considerations for the set of Remote Mirror and Copy functions available with the DS8000 series.

Metro Mirror

Metro Mirror is a function for synchronous data copy at a limited distance. The following considerations apply:

- ▶ There is no data loss, and it allows for rapid recovery for distances up to 300 km.
- ▶ There will be a slight performance impact for write operations.

Global Copy

Global Copy is a function for non-synchronous data copy at long distances, which is only limited by the network implementation. The following considerations apply:

- ▶ It can copy your data at nearly an unlimited distance, making it suitable for data migration and daily backup to a remote distant site.
- ▶ The copy is normally *fuzzy* but can be made consistent through a synchronization procedure.
- ▶ Global Copy is typically used for data migration to new DS8000s using the existing PPRC FC infrastructure.

Global Mirror

Global Mirror is an asynchronous copy technique; you can create a consistent copy in the secondary site with an adaptable Recovery Point Objective (RPO). RPO specifies how much data you can afford to recreate if the system needs to be recovered. The following considerations apply:

- ▶ Global Mirror can copy to nearly an unlimited distance.
- ▶ It is scalable across multiple storage units.
- ▶ It can realize a low RPO if there is enough link bandwidth; when the link bandwidth capability is exceeded with a heavy workload, the RPO will grow.
- ▶ Global Mirror causes only a slight impact to your application system.

z/OS Global Mirror

z/OS Global Mirror is an asynchronous copy technique controlled by z/OS host software called System Data Mover. The following considerations apply:

- ▶ It can copy to nearly unlimited distances.
- ▶ It is highly scalable.
- ▶ It has low RPO; the RPO might grow if the bandwidth capability is exceeded, or host performance might be impacted.
- ▶ Additional host server hardware and software is required.

6.4 Resource Groups for copy services

Resource Groups are implemented in such a way that each copy service volume is separated and protected from other volumes in a copy service relationship. Therefore, in a multi customer environment we protect the customer data logically from each other. During Resource Groups definition, we define an aggregation of resources and define certain policies depending how the resources are configured or managed. This gives you the ability of multi-tenancy by assigning specific resources to specific tenants, limiting copy services relationship so that they exist only between resources within each tenant's scope of resources.

Resource Groups provides additional policy-based limitations to DS8000 users, to secure partitioning of copy services resources between user-defined partitions. This process of specifying the appropriate rules is performed by an administrator using resource group functions. A resource scope (RS) specifies a selection criteria for a set of resource groups.

Using resource group on DS8000 introduces the following concepts:

- ▶ Resource Group Label (RGL): The RGL is a text string from 1 to 32 characters.
- ▶ Resource Scope (RS): The RS is a text string from 1 to 32 characters that selects one or more resource group labels by matching the RS to RGL string.
- ▶ Resource Group (RG): A RG consists of new configuration objects. It has a unique RGL within an storage facility image (SFI). An RG contains specific policies volumes and LSS/LCUs are that associated with a single RG.
- ▶ User Resource Scope (URS): Each user has an ID assigned to URS that contains an RS. The URS cannot equal zero.

The Resource Groups concept has been implemented into IBM Storage System DS8700 and DS8800 with microcode Release 6.1. RG is supported on DS8100 and DS8800 using LMC: 4.3.xx.xx. The RG environments can also be managed by TPC-R starting on level 4.1.1.6 and later.

Figure 6-13 on page 156 shows a sample how the multi tenancy is used in a mixed DS8000 environment and how the OS environment is separated.

For a more detailed description about implementation, planning, and using Resource Groups, refer to *IBM System Storage DS8000 Series: Resource Groups*, REDP-4758.

Tip: Resource Groups is implemented in the Code by default and is available at no extra fee.

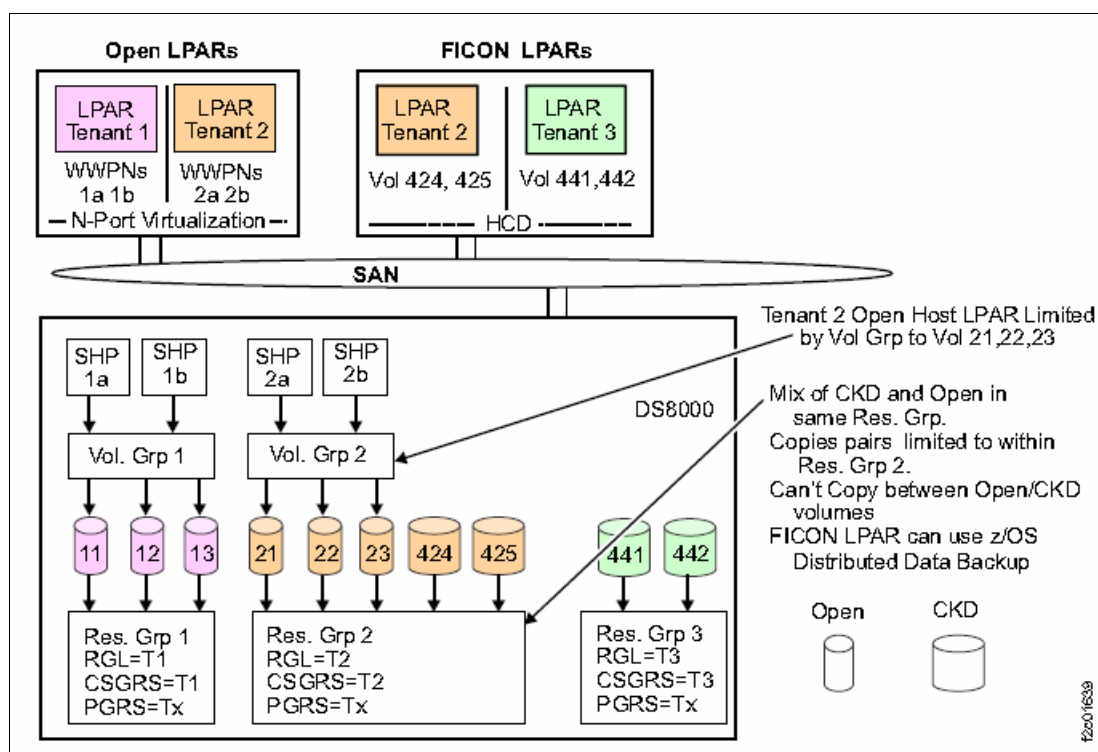


Figure 6-13 Example of a multi-tenancy configuration



Performance

This chapter discusses the performance characteristics of the IBM System Storage DS8800 regarding physical and logical configuration. The considerations presented in this chapter can help you plan the physical and logical setup.

For a detailed discussion about performance, refer to *DS8000 Performance Monitoring and Tuning*, SG24-7146.

For a detailed discussion about I/O Priority Manager, refer to *DS8000 I/O Priority Manager*, REDP-4760.

This chapter covers the following topics:

- ▶ DS8700 hardware: Performance characteristics
- ▶ Software performance: Synergy items
- ▶ Performance and sizing considerations for open systems
- ▶ Performance and sizing considerations for System z

7.1 DS8700 hardware: Performance characteristics

The DS8700 features IBM POWER6 server technology and a PCI Express I/O infrastructure to help support high performance. Compared to the POWER5+ processor in previous models, the POWER6 processor can enable over a 50% performance improvement in I/O operations per second in transaction processing workload environments. Additionally, sequential workloads can receive as much as 150% bandwidth improvement, which is an improvement factor of 2.5 compared to the previous models. The DS8700 offers either a dual 2-way processor complex or a dual 4-way processor complex.

The DS8700 overcomes many of the architectural limits of the predecessor disk subsystems. In this section, we go through the architectural layers of the DS8000 and discuss the performance characteristics that differentiate the DS8000 from other disk subsystems.

7.1.1 Fibre Channel switched disk interconnection at the back end

Fibre Channel-connected disks are used in the DS8000 back end. This technology is commonly used to connect a group of disks in a daisy-chained fashion in a Fibre Channel Arbitrated Loop (FC-AL).

FC-AL has certain shortcomings. The most obvious ones are:

- ▶ Arbitration, that is, disks compete for loop bandwidth.
- ▶ Failures within the FC-AL loop, particularly with intermittently failing components on the loops and disks.
- ▶ The increased time that it takes to complete a loop operation as the number of loop devices increase.

These shortcomings limit the effective bandwidth of an FC-AL implementation, especially for cases of highly parallel operations with concurrent reads and writes of various transfer sizes.

The DS8700 series and FC-AL shortcomings

The DS8000 uses the same Fibre Channel drives that are used in conventional FC-AL-based storage systems. To overcome the arbitration issue within FC-AL, the architecture is enhanced by adding a switch-based approach and creating FC-AL switched loops, as shown in Figure 7-1. This is called a *Fibre Channel switched disk subsystem*.

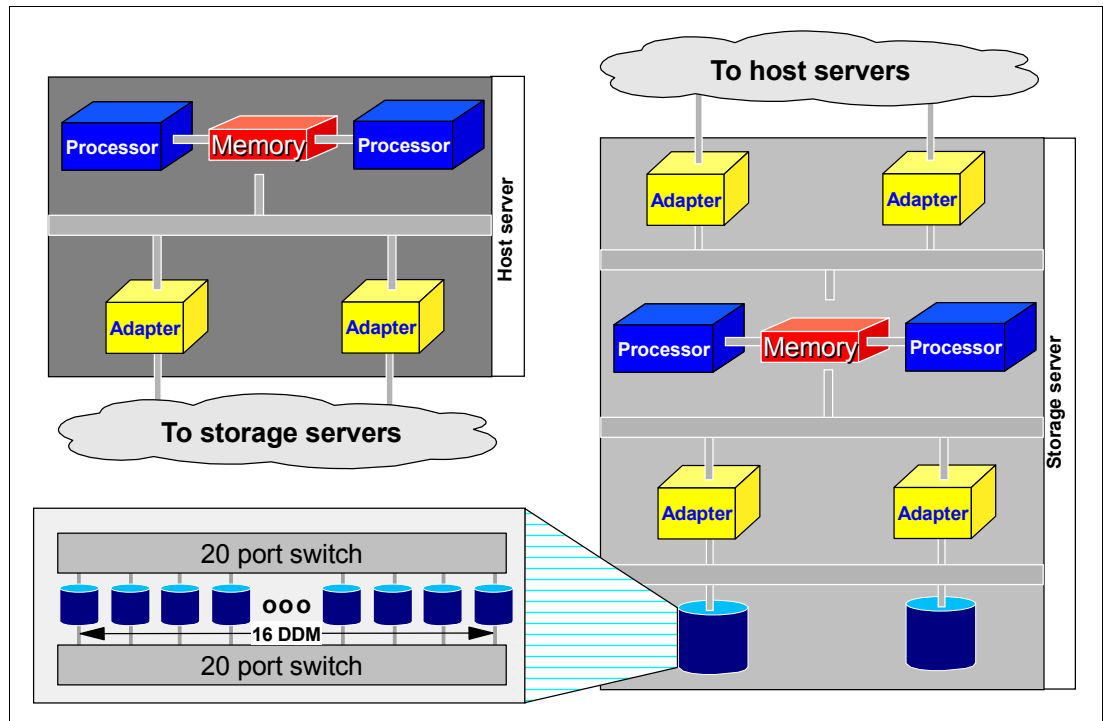


Figure 7-1 Switched FC-AL disk subsystem

These switches use the FC-AL protocol and attach FC-AL drives through a point-to-point connection. The arbitration message of a drive is captured in the switch, processed, and propagated back to the drive, without routing it through all the other drives in the loop.

Performance is enhanced, because both device adapters (DAs) connect to the switched Fibre Channel disk subsystem back end, as shown in Figure 7-2. Note that each DA port can concurrently send and receive data.

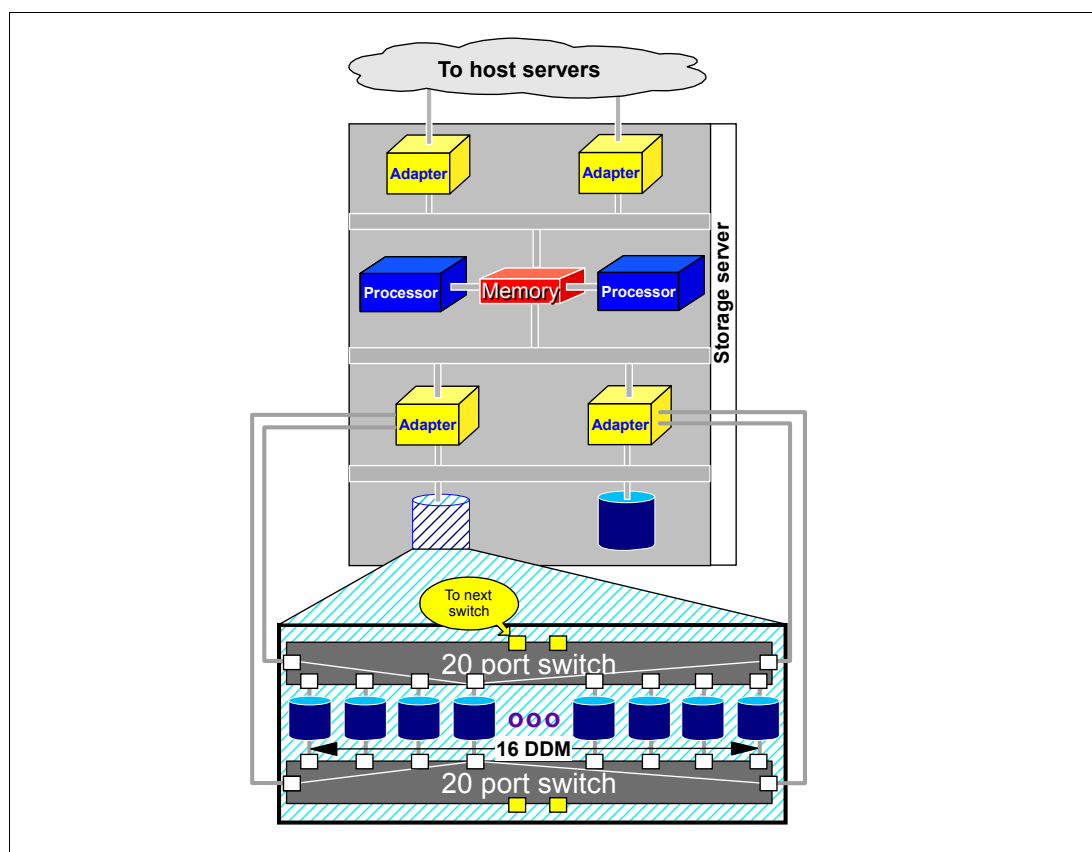


Figure 7-2 High availability and increased bandwidth connect both DAs to two logical loops

These two switched point-to-point connections to each drive, which also connect both DAs to each switch, mean the following:

- ▶ There is no arbitration competition and interference between one drive and all the other drives, because there is no hardware in common for all the drives in the FC-AL loop. This leads to an increased bandwidth, which utilizes the full speed of a Fibre Channel for each individual drive.
- ▶ This architecture doubles the bandwidth over conventional FC-AL implementations due to two simultaneous operations from each DA to allow for two concurrent read operations and two concurrent write operations at the same time.
- ▶ In addition to the superior performance, we must not forget the improved reliability, availability, and serviceability (RAS) that this setup has over conventional FC-AL. The failure of a drive is detected and reported by the switch. The switch ports distinguish between intermittent failures and permanent failures. The ports understand intermittent failures, which are recoverable, and collect data for predictive failure statistics. If one of the switches fails, a disk enclosure service processor detects the failing switch and reports the failure using the other loop. All drives can still connect through the remaining switch.

This discussion has just outlined the physical structure. A virtualization approach built on top of the high performance architectural design contributes even further to enhanced performance, as discussed in Chapter 5, "Virtualization concepts" on page 103.

7.1.2 DS8700 Fibre Channel device adapter

The DS8000 relies on eight disk drive modules (DDMs) to form a RAID 5, RAID 6, or a RAID 10 array. These DDMs are actually spread over two Fibre Channel fabrics. With the virtualization approach and the concept of extents, the DS8000 device adapters (DAs) are mapping the virtualization scheme over the disk subsystem back end, as shown in Figure 7-3. For a detailed discussion about disk subsystem virtualization, see Chapter 5, “Virtualization concepts” on page 103.

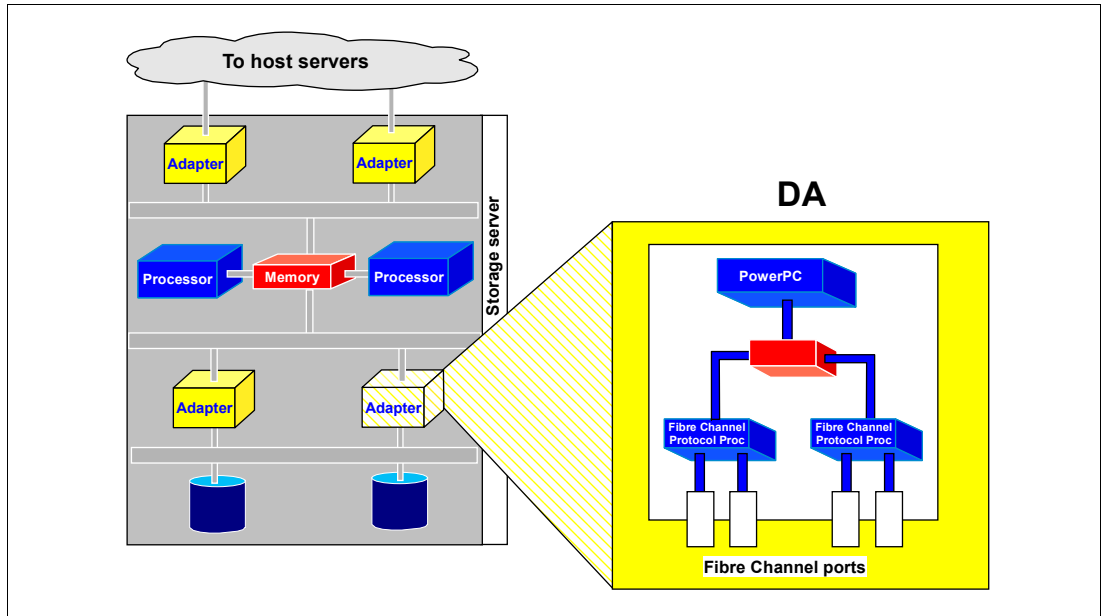


Figure 7-3 Fibre Channel device adapter

The RAID device adapter is built on PowerPC technology with four Fibre Channel ports and high function and high performance ASICs.

Note that each DA performs the RAID logic and frees up the processors from this task. The actual throughput and performance of a DA is not only determined by the port speed and hardware used, but also by the firmware efficiency.

For the DS8700, the device adapters have been upgraded with a processor that is twice as fast on the adapter card compared to DS8100 and DS8300, providing a much higher throughput on the device adapter.

7.2 DS8800 hardware: Performance characteristics

The DS8800 features IBM POWER6+ server technology and a PCI Express I/O infrastructure to help support high performance. Compared to the POWER5+ processor in previous models, the POWER6 and POWER6+ processors can enable a more than 50% performance improvement in I/O operations per second in transaction processing workload environments. Additionally, peak large-block sequential workloads can receive as much as 200% bandwidth improvement, which is an improvement factor of 3 compared to the DS8300 models.

The DS8800 offers either a dual 2-way processor complex or a dual 4-way processor complex. The DS8800 overcomes many of the architectural limits of the predecessor disk subsystems.

In this section, we go through the architectural layers of the DS8000 and discuss the performance characteristics that differentiate the DS8000 from other disk subsystems.

7.2.1 DS8800 Fibre Channel switched interconnection at the back-end

DS8800 works with SAS disks. Shortly before the FC-to-SAS conversion is made, Fibre Channel switching is used in the DS8800 back-end.

The FC technology is commonly used to connect a group of disks in a daisy-chained fashion in a Fibre Channel Arbitrated Loop (FC-AL). To overcome the arbitration issue within FC-AL, the DS8800 architecture is enhanced by adding a switch-based approach and creating FC-AL switched loops, as shown in Figure 4-6 on page 90. It is called a *Fibre Channel switched disk system*.

These switches use the FC-AL protocol and attach to the SAS drives (bridging to SAS protocol) through a point-to-point connection. The arbitration message of a drive is captured in the switch, processed, and propagated back to the drive, without routing it through all the other drives in the loop.

Performance is enhanced because both device adapters (DAs) connect to the switched Fibre Channel subsystem back-end, as shown in Figure 7-4 on page 163. Note that each DA port can concurrently send and receive data.

These two switched point-to-point connections to each drive, which also connect both DAs to each switch, mean the following:

- ▶ There is no arbitration competition and interference between one drive and all the other drives, because there is no hardware in common for all the drives in the FC-AL loop. This leads to an increased bandwidth, which goes with the full 8 Gbps FC speed up to the back-end place where the FC-to-SAS conversion is made, and which utilizes the full SAS 2.0 speed for each individual drive.
- ▶ This architecture doubles the bandwidth over conventional FC-AL implementations due to two simultaneous operations from each DA to allow for two concurrent read operations and two concurrent write operations at the same time.
- ▶ In addition to superior performance, note the improved reliability, availability, and serviceability (RAS) that this setup has over conventional FC-AL. The failure of a drive is detected and reported by the switch. The switch ports distinguish between intermittent failures and permanent failures. The ports understand intermittent failures, which are recoverable, and collect data for predictive failure statistics. If one of the switches fails, a disk enclosure service processor detects the failing switch and reports the failure using the other loop. All drives can still connect through the remaining switch.

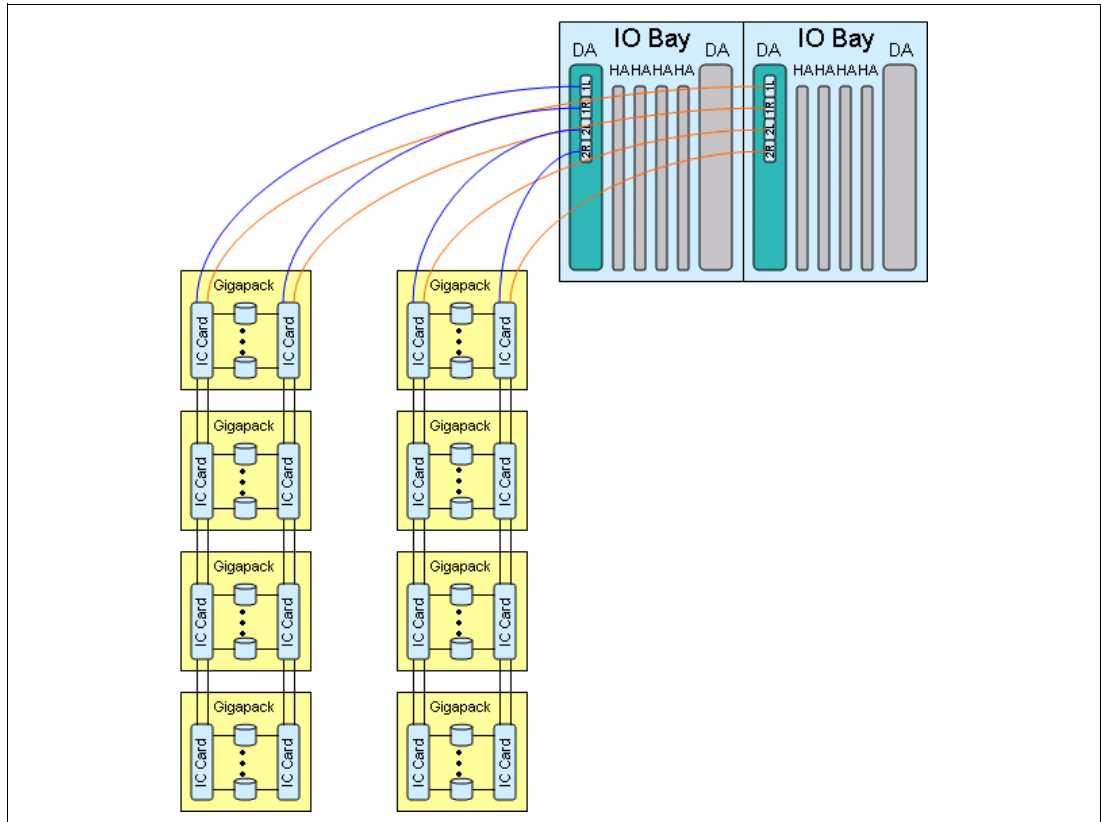


Figure 7-4 High availability and increased bandwidth connect both DAs to two logical loops

This discussion outlines the physical structure. A virtualization approach built on top of the high performance architectural design contributes even further to enhanced performance, as discussed in Chapter 5, “Virtualization concepts” on page 103.

7.2.2 Fibre Channel device adapter

The DS8000 relies on eight disk drive modules (DDMs) to form a RAID 5, RAID 6, or RAID 10 array. These DDMs are spread over two Fibre Channel fabrics. With the virtualization approach and the concept of extents, the DS8000 device adapters (DAs) are mapping the virtualization scheme over the disk subsystem back-end, as shown in Figure 7-5 on page 164. For a detailed discussion about disk subsystem virtualization, refer to Chapter 5, “Virtualization concepts” on page 103.

The RAID device adapter is built on PowerPC technology with four Fibre Channel ports and high function and high performance ASICs. It is PCIe Gen.-based and runs at 8 Gbps.

Note that each DA performs the RAID logic and frees up the processors from this task. The actual throughput and performance of a DA is not only determined by the port speed and hardware used, but also by the firmware efficiency.

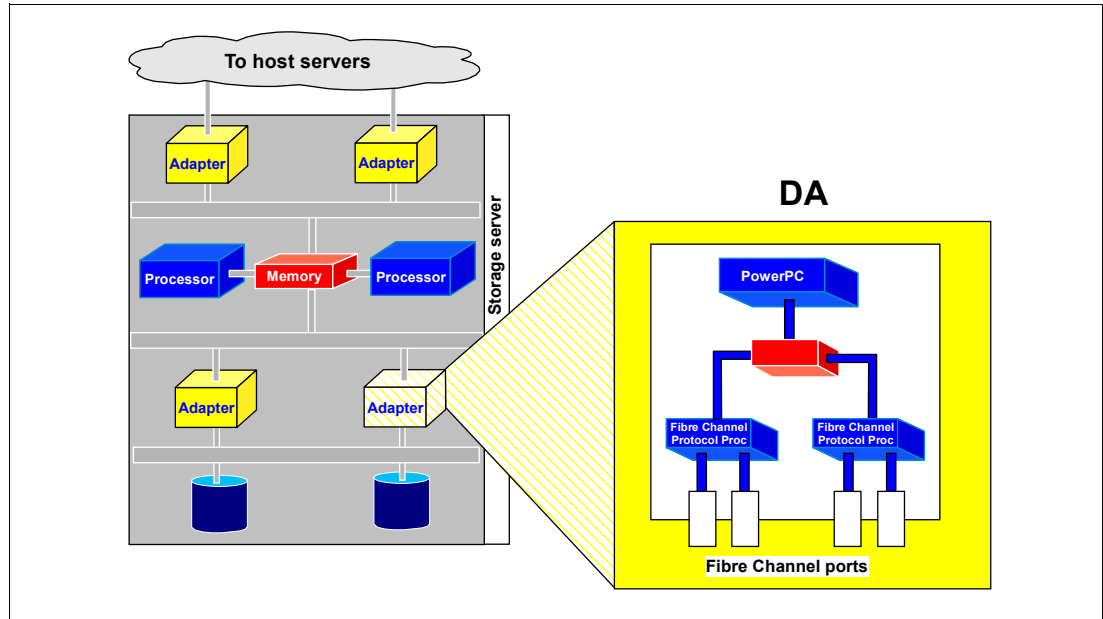


Figure 7-5 Fibre Channel device adapter

Figure 7-6 shows the detailed cabling between the Device Adapters and the 24-drive Gigapacks. The ASICs seen there provide the FC-to-SAS bridging function from the external SFP connectors to each of the ports on the SAS disk drives. The processor is the controlling element in the system.

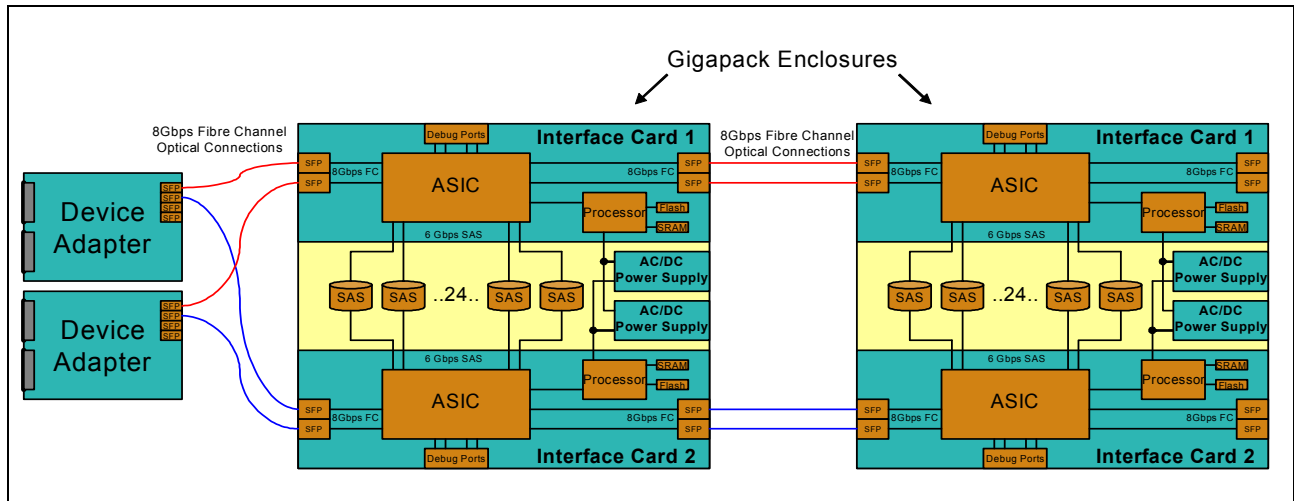


Figure 7-6 Detailed DA-disk back-end diagram

Already for the DS8700, the device adapters had been upgraded with a twice-as-fast processor on the adapter card compared to DS8100 and DS8300, providing much higher throughput on the device adapter. For the DS8800, additional enhancements to the DA bring a major performance improvement compared to DS8700: For DA limited workloads, the maximum IOps throughput (small blocks) per DA has been increased by 40% to 80%, and DA sequential throughput in MB/s (large blocks) has increased by approximately 85% to 210% from DS8700 to DS8800. For instance, a single DA under ideal workload conditions can process a sequential large-block read throughput of up to 1600 MB/s. These improvements are of value in particular when using Solid-State Drives (SSDs), but also give the DS8800

system very high sustained sequential throughput, for instance in High-Performance Computing configurations.

Technically, the improvements (processor, architecture) are similar to those designed for the Host Adapters, and are described in 7.2.3, “Eight-port and four-port host adapters” on page 165.

7.2.3 Eight-port and four-port host adapters

Before looking into the heart of the DS8000 series, we briefly review the host adapters and their enhancements to address performance. Figure 7-7 shows the host adapters. These adapters are designed to hold either eight, or four Fibre Channel (FC) ports, which can be configured to support either FCP or FICON.

Each port provides industry-leading throughput and I/O rates for FICON and FCP.

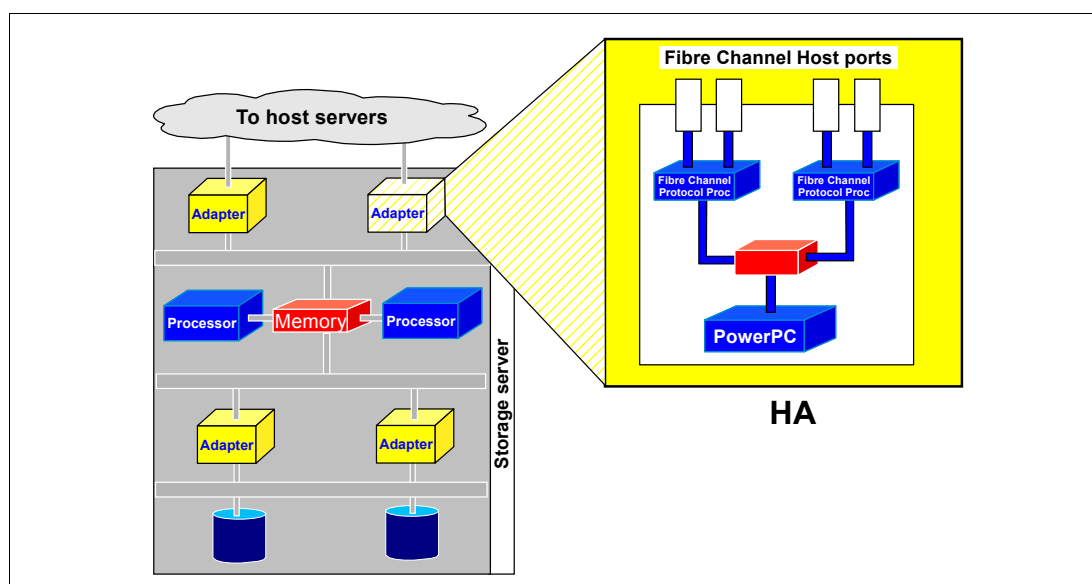


Figure 7-7 Host adapter with four Fibre Channel ports

With FC adapters that are configured for FICON, the DS8000 series provides the following configuration capabilities:

- ▶ Either fabric or point-to-point topologies
- ▶ A maximum of 128 host adapter ports, depending on the DS8800 processor feature
- ▶ A maximum of 509 logins per Fibre Channel port
- ▶ A maximum of 8192 logins per storage unit
- ▶ A maximum of 1280 logical paths on each Fibre Channel port
- ▶ Access to all control-unit images over each FICON port
- ▶ A maximum of 512 logical paths per control unit image

FICON host channels limit the number of devices per channel to 16,384. To fully access 65,280 devices on a storage unit, it is necessary to connect a minimum of four FICON host channels to the storage unit. This way, by using a switched configuration, you can expose 64 control-unit images (16,384 devices) to each host channel.

The front-end with the 8 Gbps ports scales up to 128 ports for a DS8800, using the 8-port HBAs. This results in a theoretical aggregated host I/O bandwidth of 128 times 8 Gbps.

The following improvements have been implemented on the architecture of the Host Adapter, leading to HA throughputs which are more than double compared to DS8700:

- ▶ The architecture is fully on 8 Gbps.
- ▶ x8 Gen2 PCIe interface; no PCI-X-to-PCIe bridge carrier is needed.
- ▶ The single-core 1 GHz PowerPC processor (750 GX) has been replaced by a dual-core 1.5 GHz (Freescale MPC8572).
- ▶ Adapter memory has increased fourfold.

The 8 Gbps adapter ports can negotiate to 8, 4, or 2 Gbps (1 Gbps not possible). For attachments to 1 Gbps hosts, use a switch in between.

7.2.4 IBM System p POWER6: Heart of the DS8000 dual cluster design

The DS8700 model incorporates the System p POWER6 processor technology. The new DS8800 model incorporates POWER6+ processor technology. All the DS8000 models can be equipped with the 2-way processor feature or the 4-way processor feature for highest performance requirements.

Whereas the DS8100 and DS8300 used the RIO-G connection between the clusters as a high bandwidth interconnection to the device adapters, the DS8700 and DS8800 uses dedicated PCI Express connections to the I/O enclosures and the device adapters. This increases the bandwidth to the storage subsystem backend by a factor of up to 16 times to a theoretical bandwidth of 64 GBps.

High performance and high availability interconnect to the disk subsystem

Figure 7-8 shows how the I/O enclosures connect to the processor complex.

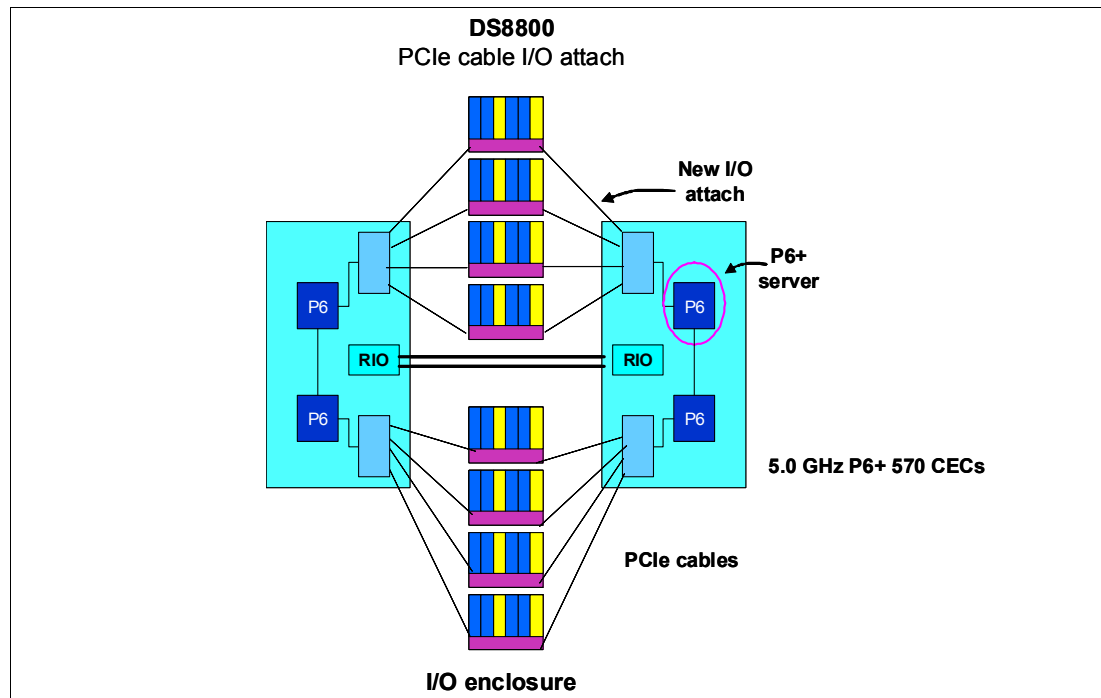


Figure 7-8 PCI Express connections to I/O enclosures

All I/O enclosures are equally served from either processor complex.

Each I/O enclosure contains two DAs. Each DA, with its four ports, connects to four switches to reach out to two sets of 16 drives or disk drive modules (DDMs) each. Note that each switch interface card has two ports to connect to the next card with 24 DDMs when vertically growing within a DS8000. As outlined before, this dual two-logical loop approach allows for multiple concurrent I/O operations to individual DDMs or sets of DDMs and minimizes arbitration through the DDM/switch port mini-loop communication.

7.2.5 Vertical growth and scalability

Figure 7-9 shows a simplified view of the basic DS8800 structure and how it accounts for scalability.

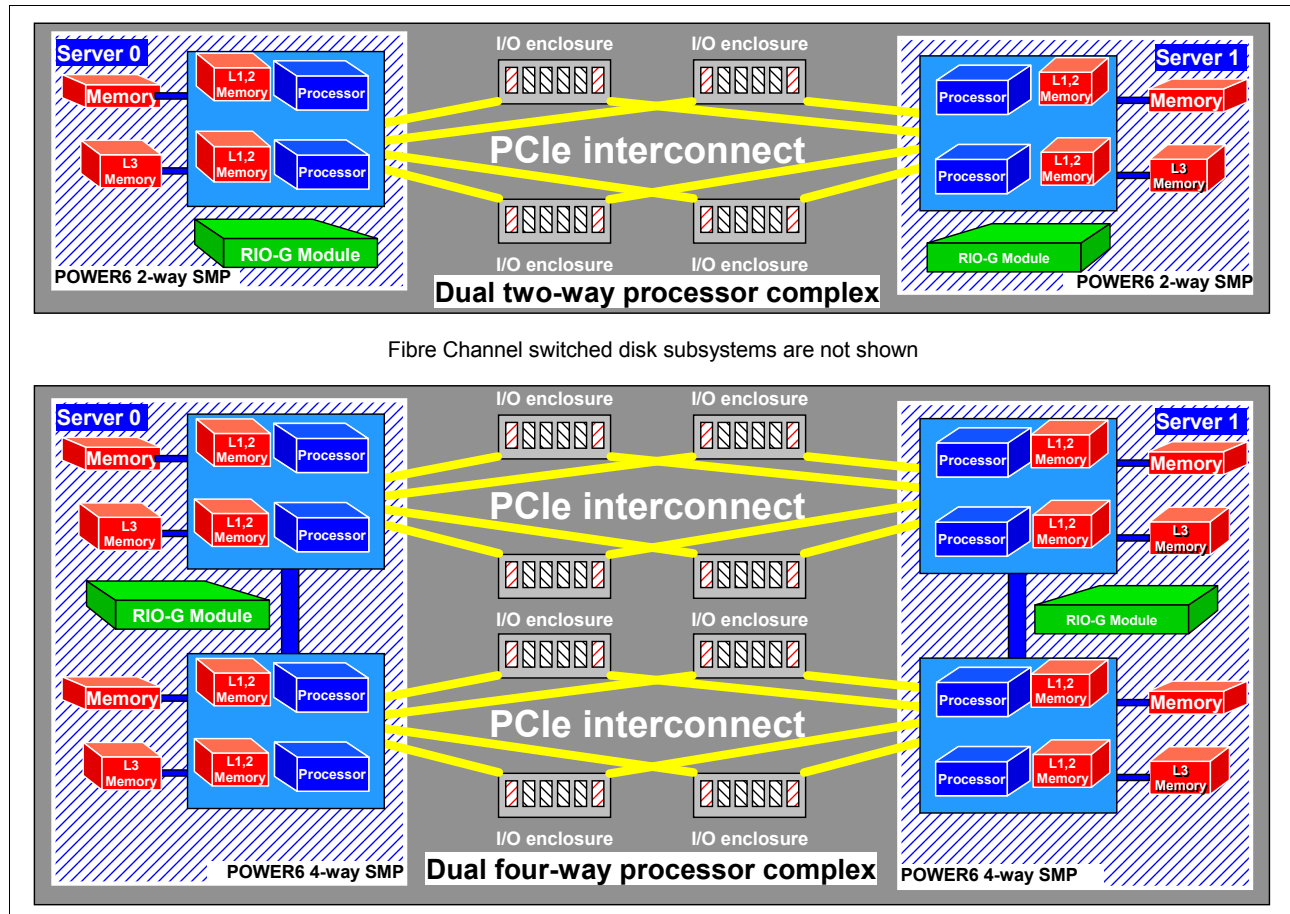


Figure 7-9 DS8800 scale performance linearly: view without disk subsystems

Although Figure 7-9 does not display the back-end part, it can be derived from the number of I/O enclosures, which suggests that the disk subsystem also doubles, as does everything else, when switching from a DS8800 2-way system with four I/O enclosures to an DS8800 4-way system with eight I/O enclosures. Doubling the number of processors and I/O enclosures accounts also for doubling the potential throughput.

Again, note that a virtualization layer on top of this physical layout contributes to additional performance potential.

7.3 Software performance: Synergy items

There are a number of performance features in the DS8000 that work together with the software on the host and are collectively referred to as *synergy items*. These items allow the DS8000 to cooperate with the host systems in manners beneficial to the overall performance of the systems.

7.3.1 End-to-end I/O priority: Synergy with AIX and DB2 on System p

End-to-end I/O priority is a new addition, requested by IBM, to the SCSI T10 standard. This feature allows trusted applications to override the priority given to each I/O by the operating system. This is only applicable to raw volumes (no file system) and with the 64-bit kernel. Currently, AIX supports this feature in conjunction with DB2. The priority is delivered to storage subsystem in the FCP Transport Header.

The priority of an AIX process can be 0 (no assigned priority) or any integer value from 1 (highest priority) to 15 (lowest priority). All I/O requests associated with a given process inherit its priority value, but with end to end I/O priority, DB2 can change this value for critical data transfers. At the DS8000, the host adapter will give preferential treatment to higher priority I/O, improving performance for specific requests deemed important by the application, such as requests that might be prerequisites for others, for example, DB2 logs.

7.3.2 Cooperative caching: Synergy with AIX and DB2 on System p

Another software-related performance item is *cooperative caching*, a feature which provides a way for the host to send cache management hints to the storage facility. Currently, the host can indicate that the information just accessed is unlikely to be accessed again soon. This decreases the retention period of the cached data, allowing the subsystem to conserve its cache for data that is more likely to be reaccessed, improving the cache hit ratio.

With the implementation of cooperative caching, the AIX operating system allows trusted applications, such as DB2, to provide cache hints to the DS8000. This improves the performance of the subsystem by keeping more of the repeatedly accessed data within the cache. Cooperative caching is supported in System p AIX with the Multipath I/O (MPIO) Path Control Module (PCM) that is provided with the Subsystem Device Driver (SDD). It is only applicable to raw volumes (no file system) and with the 64-bit kernel.

7.3.3 Long busy wait host tolerance: Synergy with AIX on System p

Another new addition to the SCSI T10 standard is SCSI long busy wait, which provides a way for the target system to specify that it is busy and how long the initiator should wait before retrying an I/O.

This information, provided in the Fibre Channel Protocol (FCP) status response, prevents the initiator from retrying too soon. This in turn reduces unnecessary requests and potential I/O failures due to exceeding a set threshold for the number of retries. IBM System p AIX supports SCSI long busy wait with MPIO, and it is also supported by the DS8000.

7.3.4 PowerHA Extended distance extensions: Synergy with AIX on System p

The PowerHA™ SystemMirror Enterprise Edition (former HACMP/XD) provides server and LPAR failover capability over extended distances. It can also take advantage of the Metro Mirror or Global Mirror functions of the DS8000 as a data replication mechanism between the primary and remote site. PowerHA System Mirror with Metro Mirror supports distances of up to 300 km. The DS8000 requires no changes to be used in this fashion.

7.4 Performance considerations for disk drives

You can determine the number and type of ranks required based on the needed capacity and on the workload characteristics in terms of access density, read to write ratio, and hit rates.

You can approach this task from the disk side and look at basic disk figures. Current SAS 15K RPM disks, for example, provide an average seek time of approximately 3.1 ms and an average latency of 2 ms. For transferring only a small block, the transfer time can be neglected. This is an average 5.1 ms per random disk I/O operation or 196 IOPS. A combined number of eight disks (as is the case for a DS8000 array) will thus potentially sustain 1568 IOPS when spinning at 15 K RPM. Reduce the number by 12.5% when you assume a spare drive in the eight pack.

Back on the host side, consider an example with 1000 IOPS from the host, a read-to-write ratio of 3 to 1, and 50% read cache hits. This leads to the following IOPS numbers:

- ▶ 750 read IOPS.
- ▶ 375 read I/Os must be read from disk (based on the 50% read cache hit ratio).
- ▶ 250 writes with RAID 5 results in 1,000 disk operations due to the RAID 5 write penalty (read old data and parity, write new data and parity).
- ▶ This totals to 1375 disk I/Os.

With 15K RPM DDMs doing 1000 random IOPS from the server, we actually do 1375 I/O operations on disk compared to a maximum of 1440 operations for 7+P configurations or 1260 operations for 6+P+S configurations. Thus, 1000 random I/Os from a server with a standard read-to-write ratio and a standard cache hit ratio saturate the disk drives. We made the assumption that server I/O is purely random. When there are sequential I/Os, track-to-track seek times are much lower and higher I/O rates are possible. We also assumed that reads have a hit ratio of only 50%. With higher hit ratios, higher workloads are possible. This shows the importance of intelligent caching algorithms as used in the DS8000.

Important: When sizing a storage subsystem, you should consider the capacity and the number of disk drives needed to satisfy the performance requirements.

For a single disk drive, various disk vendors provide the disk specifications on their websites. Because the access times for the disks are the same for same RPM speeds, but they have different capacities, the I/O density is different. 146 GB 15K RPM disk drives can be used for access densities up to, and slightly over, 1 I/O per GB-s. For 450 GB drives, it is approximately 0.5 I/O per GBs. Although this discussion is theoretical in approach, it provides a first estimate.

After the speed of the disk has been decided, the capacity can be calculated based on your storage capacity needs and the effective capacity of the RAID configuration you will use. Refer to Table 8-11 on page 221 for information about calculating these needs.

Solid-State Drives

From a performance point of view, the best choice for your DS8800 disks would be the new Solid-State Drives (SSDs). SSDs have no moving parts (no spinning platters and no actuator arm). The performance advantages are the fast seek time and average access time. They are targeted at applications with heavy IOPS, bad cache hit rates, and random access workload, which necessitates fast response times. Database applications with their random and intensive I/O workloads are prime candidates for deployment on SSDs.

Enterprise drives (SAS and FC disk drives)

Enterprise drives (SAS drives in the DS8800 or Fibre Channel disk drives in the DS8700) provide high performance, reliability, availability, and serviceability. Enterprise drives rotate at 15,000 or 10,000 RPM. If an application requires high performance data throughput and continuous, intensive I/O operations, enterprise drives are the best price performance option.

Near line drives (SATA disk)

When analyzing disk alternatives, keep in mind that the 2 TB near line drives are both the largest and slowest of the drives available for the DS8700. Near line drives are a cost-efficient storage option for lower intensity storage workloads and are available with the DS8700. This, combined with the lower utilization and the potential for drive protection throttling, means that these drives are a poor choice for high performance or I/O intensive applications.

Tip: The near line drives offer a cost-effective option for lower priority data, such as various fixed content, data archival, reference data, and near-line applications that require large amounts of storage capacity for lighter workloads. These new drives are meant to complement, not compete with, existing Fibre Channel drives, because they are not intended for use in applications that require drive utilization duty cycles greater than 20 percent.

SSD Usage

For detailed recommendations about SSD usage and performance, refer to *DS8000: Introducing Solid State Drives*, REDP-4522.

Easy Tier

For detailed recommendations about Easy Tier, refer to *DS8000 Easy Tier*, REDP-4667

I/O Performance Manager

For detailed recommendations about I/O Performance Manager usage and performance, refer to *DS8000 I/O Performance Manager*, REDP-4760

RAID level

The DS8000 series offers RAID 5, RAID 6, and RAID 10.

RAID 5

Normally, RAID 5 is used because it provides good performance for random and sequential workloads and it does not need much additional storage for redundancy (one parity drive). The DS8000 series can detect sequential workload. When a complete stripe is in cache for destage, the DS8000 series switches to a RAID 3-like algorithm. Because a complete stripe has to be destaged, the old data and parity do not need to be read. Instead, the new parity is calculated across the stripe, and the data and parity are destaged to disk. This provides good sequential performance. A random write causes a cache hit, but the I/O is not complete until a copy of the write data is put in NVS. When data is destaged to disk, a write in RAID 5 causes four disk operations, the so-called *write penalty*:

- ▶ Old data and the old parity information must be read.
- ▶ New parity is calculated in the device adapter
- ▶ Data and parity are written to disk.

Most of this activity is hidden to the server or host because the I/O is complete when data has entered cache and NVS.

RAID 6

RAID 6 is an option that increases data fault tolerance. It allows additional failure, compared to RAID 5, by using a second independent distributed parity scheme (dual parity). RAID 6 provides a Read Performance similar to RAID 5, but has more write penalty than RAID 5 because it has to write a second parity stripe.

RAID 6 should be considered in situations where you would consider RAID 5, but need increased reliability. RAID 6 was designed for protection during longer rebuild times on larger capacity drives to cope with the risk of having a second drive failure within a rank while the failed drive is being rebuilt. It has the following characteristics:

Sequential Read	About 99% x RAID 5 Rate
Sequential Write	About 65% x RAID 5 Rate
Random 4K 70%R/30%W IOPs	About 55% x RAID 5 Rate

The performance is significantly degraded with two failing disks.

RAID 10

A workload that is dominated by *random writes* will benefit from RAID 10. Here data is striped across several disks and at the same time mirrored to another set of disks. A write causes only two disk operations compared to RAID 5's four operations. However, you need nearly twice as many disk drives for the same capacity when compared to RAID 5. Thus, for twice the number of drives (and probably cost), we can do four times more random writes, so it is worth considering using RAID 10 for high performance random write workloads.

The decision to configure capacity as RAID 5, RAID 6, or RAID 10, and the amount of capacity to configure for each type, can be made at any time. RAID 5, RAID 6, and RAID 10 arrays can be intermixed within a single system and the physical capacity can be logically reconfigured at a later date (for example, RAID 6 arrays can be reconfigured into RAID 5 arrays).

Disk Magic and Capacity Magic

Apart from the general guidance we provide in this chapter, the best approach is to use your installation workload as input to the Disk Magic modelling tool (see Appendix A, "Tools and service offerings" on page 455). With your workload data and current configuration of disk subsystem units, Disk Magic can establish a base model. From this base model, Disk Magic can project the DS8000 units (and their configuration) that will be needed to absorb the present workload and also any future growth that you anticipate.

To estimate the number and capacity of disk drive sets needed to fulfill your storage capacity requirements, use the Capacity Magic tool. It is an easy-to-use tool that will also help you determine the requirements for any growth in capacity that you can foresee.

7.5 DS8000 superior caching algorithms

Most, if not all, high-end disk systems have an internal cache integrated into the system design, and some amount of system cache is required for operation. Over time, cache sizes have dramatically increased, but the ratio of cache size to system disk capacity has remained nearly the same. The DS8800 can be equipped with up to 384 GB of cache.

7.5.1 Sequential Adaptive Replacement Cache

The DS8000 series uses the Sequential Adaptive Replacement Cache (SARC) algorithm, which was developed by IBM Storage Development in partnership with IBM Research. It is a self-tuning, self-optimizing solution for a wide range of workloads with a varying mix of sequential and random I/O streams. SARC is inspired by the Adaptive Replacement Cache (ARC) algorithm and inherits many features of it. For a detailed description about ARC, see “Outperforming LRU with an adaptive replacement cache algorithm” by N. Megiddo et al. in IEEE Computer, volume 37, number 4, pages 58–65, 2004. For a detailed description about SARC, see “SARC: Sequential Prefetching in Adaptive Replacement Cache” by Binny Gill, et al, in the Proceedings of the USENIX 2005 Annual Technical Conference, pages 293–308.

SARC basically attempts to determine four things:

- ▶ When data is copied into the cache.
- ▶ Which data is copied into the cache.
- ▶ Which data is evicted when the cache becomes full.
- ▶ How the algorithm dynamically adapts to different workloads.

The DS8000 series cache is organized in 4 KB pages called cache pages or slots. This unit of allocation (which is smaller than the values used in other storage systems) ensures that small I/Os do not waste cache memory.

The decision to copy data into the DS8000 cache can be triggered from two policies: demand paging and prefetching.

- ▶ *Demand paging* means that eight disk blocks (a 4K cache page) are brought in only on a cache miss. Demand paging is always active for all volumes and ensures that I/O patterns with some locality discover at least recently used data in the cache.
- ▶ *Prefetching* means that data is copied into the cache speculatively even before it is requested. To prefetch, a prediction of likely future data accesses is needed. Because effective, sophisticated prediction schemes need an extensive history of page accesses (which is not feasible in real systems), SARC uses prefetching for sequential workloads. Sequential access patterns naturally arise in video-on-demand, database scans, copy, backup, and recovery. The goal of sequential prefetching is to detect sequential access and effectively preload the cache with data so as to minimize cache misses. Today prefetching is ubiquitously applied in web servers and clients, databases, file servers, on-disk caches, and multimedia servers.

For prefetching, the cache management uses tracks. A track is a set of 128 disk blocks (16 cache pages). To detect a sequential access pattern, counters are maintained with every track to record whether a track has been accessed together with its predecessor. Sequential prefetching becomes active only when these counters suggest a sequential access pattern. In this manner, the DS8000 monitors application read-I/O patterns and dynamically determines whether it is optimal to stage into cache:

- ▶ Just the page requested
- ▶ That page requested plus the remaining data on the disk track
- ▶ An entire disk track (or a set of disk tracks), which has not yet been requested

The decision of when and what to prefetch is made in accordance with the Adaptive Multi-stream Prefetching (AMP) algorithm, which dynamically adapts the amount and timing of prefetches optimally on a per-application basis (rather than a system-wide basis). AMP is described further in 7.5.2, “Adaptive Multi-stream Prefetching” on page 174.

To decide which pages are evicted when the cache is full, sequential and random (non-sequential) data is separated into separate lists. Figure 7-10 on page 174 illustrates the SARC algorithm for random and sequential data.

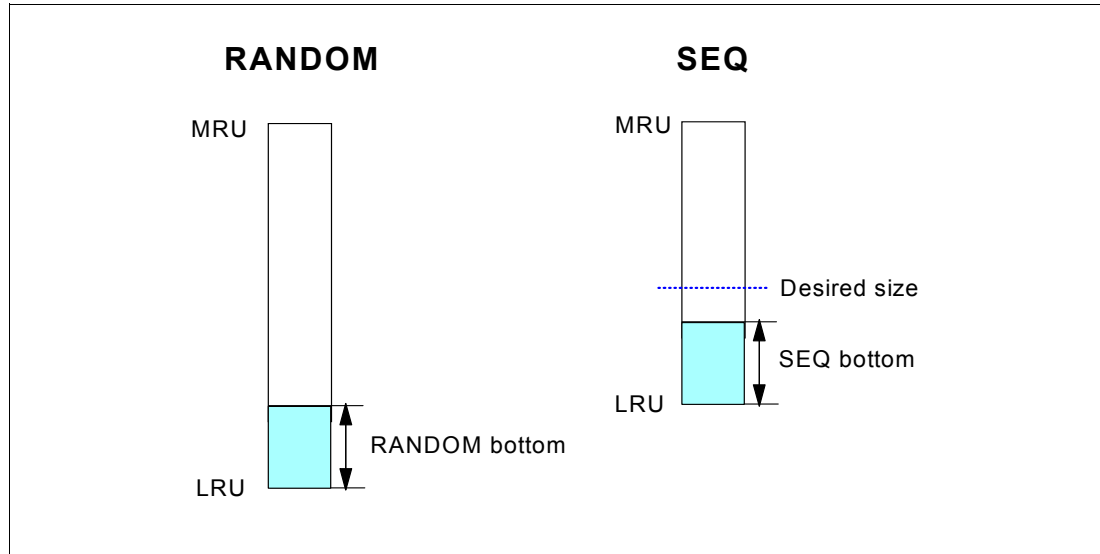


Figure 7-10 Sequential Adaptive Replacement Cache

A page that has been brought into the cache by simple demand paging is added to the head of Most Recently Used (MRU) of the RANDOM list. Without further I/O access, it goes down to the bottom of Least Recently Used (LRU). A page that has been brought into the cache by a sequential access or by sequential prefetching is added to the head of MRU of the SEQ list and then goes in that list. Additional rules control the migration of pages between the lists so as to not keep the same pages in memory twice.

To follow workload changes, the algorithm trades cache space between the RANDOM and SEQ lists dynamically and adaptively. This makes SARC scan-resistant, so that one-time sequential requests do not pollute the whole cache. SARC maintains a desired size parameter for the sequential list. The desired size is continually adapted in response to the workload. Specifically, if the bottom portion of the SEQ list is found to be more valuable than the bottom portion of the RANDOM list, then the desired size is increased; otherwise, the desired size is decreased. The constant adaptation strives to make optimal use of limited cache space and delivers greater throughput and faster response times for a given cache size.

Additionally, the algorithm dynamically modifies the sizes of the two lists and the rate at which the sizes are adapted. In a steady state, pages are evicted from the cache at the rate of cache misses. A larger (respectively, a smaller) rate of misses effects a faster (respectively, a slower) rate of adaptation.

Other implementation details take into account the relationship of read and write (NVS) cache, efficient destaging, and the cooperation with Copy Services. In this manner, the DS8000 cache management goes far beyond the usual variants of the Least Recently Used/Least Frequently Used (LRU/LFU) approaches.

7.5.2 Adaptive Multi-stream Prefetching

As described previously, SARC dynamically divides the cache between the RANDOM and SEQ lists, where the SEQ list maintains pages brought into the cache by sequential access or sequential prefetching.

In DS8800 and DS8700, Adaptive Multi-stream Prefetching (AMP), which is a tool from IBM Research, manages the SEQ. AMP is an autonomic, workload-responsive, self-optimizing

prefetching technology that adapts both the amount of prefetch and the timing of prefetch on a per-application basis to maximize the performance of the system. The AMP algorithm solves two problems that plague most other prefetching algorithms:

- ▶ *Prefetch wastage* occurs when prefetched data is evicted from the cache before it can be used.
- ▶ *Cache pollution* occurs when less useful data is prefetched instead of more useful data.

By wisely choosing the prefetching parameters, AMP provides optimal sequential read performance and maximizes the aggregate sequential read throughput of the system. The amount prefetched for each stream is dynamically adapted according to the application's needs and the space available in the SEQ list. The timing of the prefetches is also continuously adapted for each stream to avoid misses and at the same time avoid any cache pollution.

SARC and AMP play complementary roles. While SARC is carefully dividing the cache between the RANDOM and the SEQ lists so as to maximize the overall hit ratio, AMP is managing the contents of the SEQ list to maximize the throughput obtained for the sequential workloads. Whereas SARC impacts cases that involve both random and sequential workloads, AMP helps any workload that has a sequential read component, including pure sequential read workloads.

AMP dramatically improves performance for common sequential and batch processing workloads. It also provides excellent performance synergy with DB2 by preventing table scans from being I/O bound and improves performance of index scans and DB2 utilities like Copy and Recover. Furthermore, AMP reduces the potential for array hot spots, which result from extreme sequential workload demands.

For a detailed description about AMP and the theoretical analysis for its optimal usage, see "AMP: Adaptive Multi-stream Prefetching in a Shared Cache" by Binny Gill, et al. in USENIX File and Storage Technologies (FAST), February 13 - 16, 2007, San Jose, CA. For a more detailed description, see "Optimal Multistream Sequential Prefetching in a Shared Cache" by Binny Gill, et al, in the ACM Journal of Transactions on Storage, October 2007.

7.5.3 Intelligent Write Caching

Another additional cache algorithm, referred to as Intelligent Write Caching (IWC), has been implemented in the DS8000 series. IWC improves performance through better write cache management and a better destaging order of writes. This new algorithm is a combination of CLOCK, a predominantly read cache algorithm, and CSCAN, an efficient write cache algorithm. Out of this combination, IBM produced a powerful and widely applicable write cache algorithm.

The CLOCK algorithm exploits *temporal* ordering. It keeps a circular list of pages in memory, with the "hand" pointing to the oldest page in the list. When a page needs to be inserted in the cache, then a R (recency) bit is inspected at the "hand's" location. If R is zero, the new page is put in place of the page the "hand" points to and R is set to 1; otherwise, the R bit is cleared and set to zero. Then, the clock hand moves one step clockwise forward and the process is repeated until a page is replaced.

The CSCAN algorithm exploits *spatial* ordering. The CSCAN algorithm is the circular variation of the SCAN algorithm. The SCAN algorithm tries to minimize the disk head movement when servicing read and write requests. It maintains a sorted list of pending requests along with the position on the drive of the request. Requests are processed in the current direction of the disk head, until it reaches the edge of the disk. At that point, the direction changes. In the CSCAN algorithm, the requests are always served in the same direction. After the head has

arrived at the outer edge of the disk, it returns to the beginning of the disk and services the new requests in this one direction only. This results in more equal performance for all head positions.

The basic idea of IWC is to maintain a sorted list of write groups, as in the CSCAN algorithm. The smallest and the highest write groups are joined, forming a circular queue. The additional new idea is to maintain a recency bit for each write group, as in the CLOCK algorithm. A write group is always inserted in its correct sorted position and the recency bit is set to zero at the beginning. When a write hit occurs, the recency bit is set to one. The destage operation proceeds, where a destage pointer is maintained that scans the circular list looking for destage victims. Now this algorithm only allows destaging of write groups whose recency bit is zero. The write groups with a recency bit of one are skipped and the recent bit is then turned off and reset to zero, which gives an “extra life” to those write groups that have been hit since the last time the destage pointer visited them; Figure 7-11 gives an idea of how this mechanism works.

In the DS8000 implementation, an IWC list is maintained for each rank. The dynamically adapted size of each IWC list is based on workload intensity on each rank. The rate of destage is proportional to the portion of NVS occupied by an IWC list (the NVS is shared across all ranks in a cluster). Furthermore, destages are smoothed out so that write bursts are not translated into destage bursts.

Enhancement to IWC in this release level 7.6.1.xx.xx. code is the WOW (Wise Ordered Write). This new update to cache algorithm by increasing residency time of data in NVS. This improvement focuses on maximizing throughput with good average response time.

In summary, IWC has better or comparable peak throughput to the best of CSCAN and CLOCK across a wide gamut of write cache sizes and workload configurations. In addition, even at lower throughputs, IWC has lower average response times than CSCAN and CLOCK.

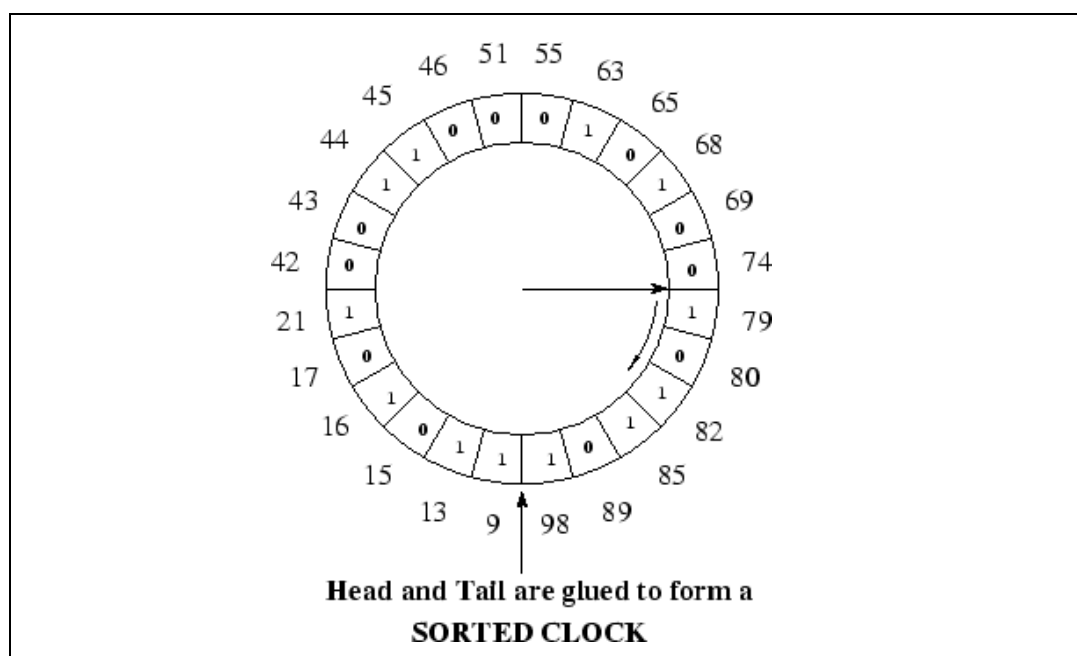


Figure 7-11 Intelligent Write Caching

7.6 Performance considerations for logical configuration

To determine the optimal DS8000 layout, the I/O performance requirements of the servers and applications should be defined up front, because they will play a large part in dictating both the physical and logical configuration of the disk subsystem. Prior to designing the disk subsystem, the disk space requirements of the application should be well understood.

7.6.1 Workload characteristics

The answers to questions such as “How many host connections do I need?” and “How much cache do I need?” always depend on the workload requirements, such as how many I/Os per second per server, I/Os per second per gigabyte of storage, and so on.

The information you need to conduct detailed modeling includes:

- ▶ Number of I/Os per second
- ▶ I/O density
- ▶ Megabytes per second
- ▶ Relative percentage of reads and writes
- ▶ Random or sequential access characteristics
- ▶ Cache hit ratio

7.6.2 Data placement in the DS8000

After you have determined the disk subsystem throughput, the disk space, and the number of disks required by your hosts and applications, you have to make a decision regarding data placement.

As is common for data placement, and to optimize DS8000 resource utilization, follow these guidelines:

- ▶ Equally spread the LUNs and volumes across the DS8000 servers. Spreading the volumes equally on rank group 0 and 1 will balance the load across the DS8000 units.
- ▶ Use as many disks as possible. Avoid idle disks, even if all storage capacity will not be initially utilized.
- ▶ Distribute capacity and workload across DA pairs.
- ▶ Use multirank Extent Pools.
- ▶ Stripe your logical volume across several ranks (the default for large Extent Pools).
- ▶ Consider placing specific database objects (such as logs) on separate ranks.
- ▶ For an application, use volumes from both even and odd numbered Extent Pools (even numbered pools are managed by server 0, and odd numbers are managed by server 1).
- ▶ For large, performance-sensitive applications, consider using two dedicated Extent Pools (one managed by server 0, the other managed by server 1).
- ▶ Consider using separate Extent Pools for 6+P+S arrays and 7+P arrays. If you use the default Storage Pool Striping, this will ensure that your ranks are equally filled.

Important: Balance your ranks and Extent Pools between the two DS8000 servers. Half of the ranks should be managed by each server (see Figure 7-12).

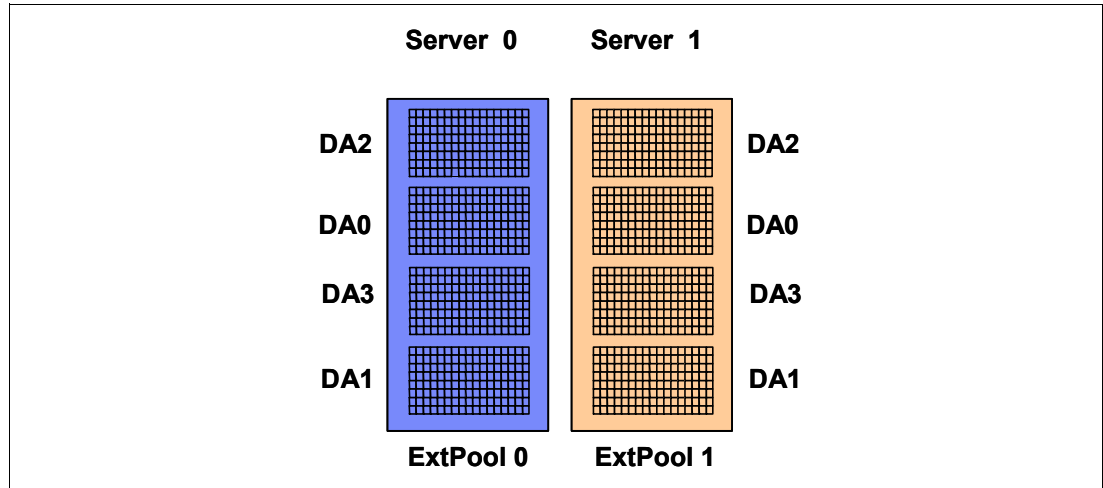


Figure 7-12 Ranks in a multirank Extent Pool configuration balanced across DS8000 servers

Tip: Database logging usually consists of sequences of synchronous sequential writes. Log archiving functions (copying an active log to an archived space) also tend to consist of simple sequential read and write sequences. Consider isolating log files on separate arrays.

All disks in the storage disk subsystem should have roughly equivalent utilization. Any disk that is used more than the other disks will become a bottleneck to performance. A practical method is to use predefined Storage Pool Striping. Alternatively, make extensive use of volume-level striping across disk drives.

7.6.3 Data placement

There are several options for creating logical volumes. You can select an Extent Pool that is owned by one server. There could be just one Extent Pool per server or you could have several. The ranks of Extent Pools can come from arrays on different device adapter pairs.

For optimal performance, your data should be spread across as many hardware resources as possible. RAID 5, RAID 6, or RAID 10 already spreads the data across the drives of an array, but this is not always enough. There are two approaches to spreading your data across even more disk drives:

- ▶ Storage Pool Striping
- ▶ Striping at the host level

Storage Pool Striping

Striping is a technique for spreading the data across several disk drives in such a way that the I/O capacity of the disk drives can be used in parallel to access data on the logical volume.

The easiest way to stripe is to use Extent Pools with more than one rank and use Storage Pool Striping when allocating a new volume (see Figure 7-13 on page 179). This striping method is independent of the operating system.

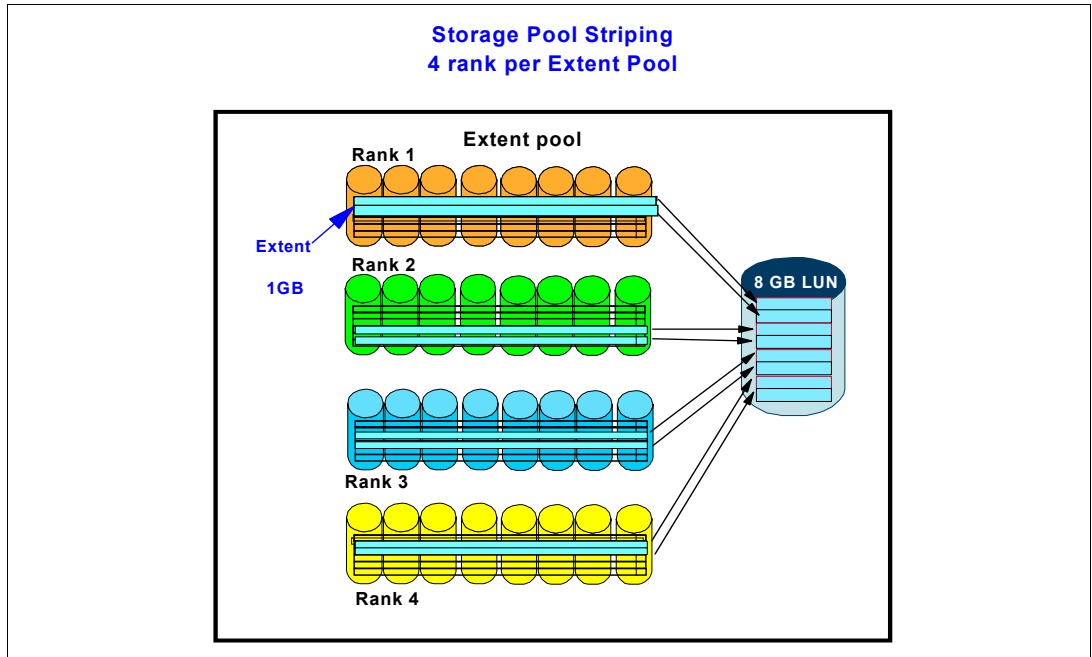


Figure 7-13 Storage Pool Striping

In 7.4, “Performance considerations for disk drives” on page 170, we discuss how many random I/Os can be performed for a standard workload on a rank. If a volume resides on just one rank, this rank’s I/O capability also applies to the volume. However, if this volume is striped across several ranks, the I/O rate to this volume can be much higher.

The total number of I/Os that can be performed on a given set of ranks does not change with Storage Pool Striping.

On the other hand, if you stripe all your data across all ranks and you lose just one rank, for example, because you lose two drives at the same time in a RAID 5 array, *all* your data is gone. Remember that with RAID 6 you can increase reliability and survive two drive failures, but the better choice is to mirror your data to a remote DS8000.

Tip: Use Storage Pool Striping and Extent Pools with a minimum of four to eight ranks of the same characteristics (RAID type and disk RPM) to avoid hot spots on the disk drives.

Figure 7-14 on page 180 shows a good configuration. The ranks are attached to DS8000 server 0 and server 1 in a half and half configuration, ranks on separate device adapters are used in a multi-rank Extent Pool, and there are separate Extent Pools for 6+P+S and 7+P ranks.

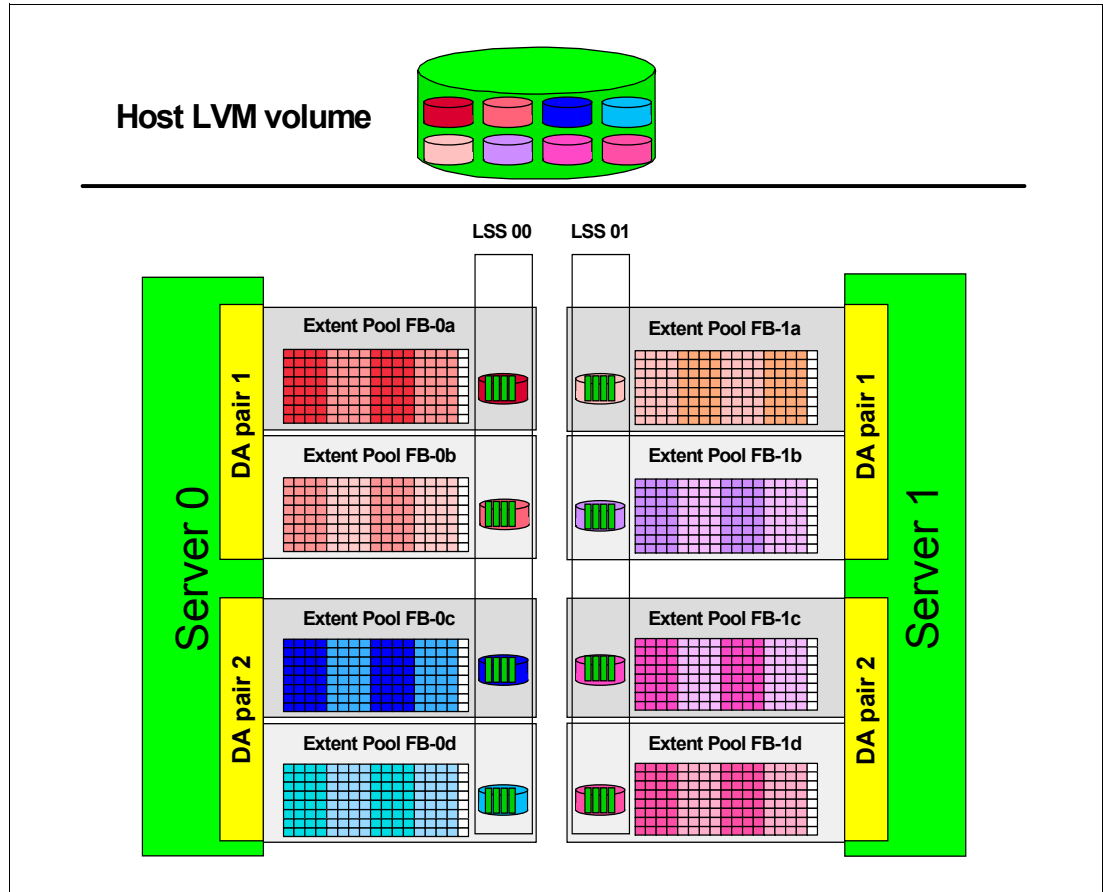


Figure 7-15 Optimal placement of data

Figure 7-15 shows an optimal distribution of eight logical volumes within a DS8000. Of course, you could have more Extent Pools and ranks, but when you want to distribute your data for optimal performance, you should make sure that you spread it across the two servers, across different device adapter pairs, and across several ranks.

To be able to create large logical volumes or to be able to use Extent Pool striping, you must consider having Extent Pools with more than one rank.

If you use multirank Extent Pools and you do not use Storage Pool Striping, you have to be careful where to put your data, or you can easily unbalance your system (see the right side of Figure 7-16 on page 182).

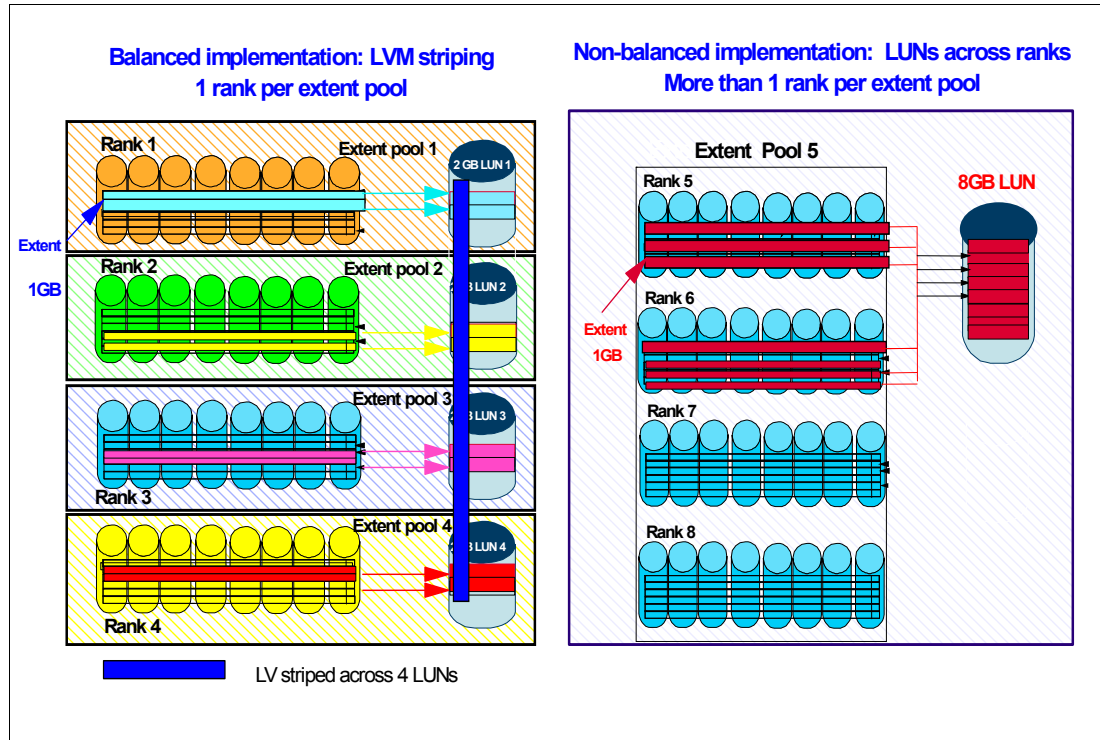


Figure 7-16 Spreading data across ranks

Combining Extent Pools made up of one rank and then LVM striping over LUNs created on each Extent Pool will offer a balanced method to evenly spread data across the DS8000 without using Extent Pool striping, as shown on the left side of Figure 7-16.

The stripe size

Each striped logical volume that is created by the host's logical volume manager has a stripe size that specifies the fixed amount of data stored on each DS8000 logical volume (LUN) at one time.

The stripe size has to be large enough to keep sequential data relatively close together, but not too large so as to keep the data located on a single array.

We recommend that you define stripe sizes using your host's logical volume manager in the range of 4 MB to 64 MB. You should choose a stripe size close to 4 MB if you have a large number of applications sharing the arrays and a larger size when you have few servers or applications sharing the arrays.

Combining Extent Pool striping and logical volume manager striping

Striping by a logical volume manager is done on a stripe size in the MB range (about 64 MB). Extent Pool striping is done at a 1 GiB stripe size. Both methods could be combined. LVM striping can stripe across Extent Pools and use volumes from Extent Pools that are attached to server 0 and server 1 of the DS8000 series. If you already use LVM Physical Partition (PP) striping, you might want to continue to use that striping. Double striping will probably not increase performance.

7.7 Performance and sizing considerations for open systems

In these sections, we discuss topics that are particularly relevant to open systems.

7.7.1 Determining the number of paths to a LUN

When configuring an IBM System Storage DS8000 for an open systems host, a decision must be made regarding the number of paths to a particular LUN, because the multipathing software allows (and manages) multiple paths to a LUN. There are two opposing factors to consider when deciding on the number of paths to a LUN:

- ▶ Increasing the number of paths increases availability of the data, protecting against outages.
- ▶ Increasing the number of paths increases the amount of CPU used because the multipathing software must choose among all available paths each time an I/O is issued.

A good compromise is between two and four paths per LUN.

7.7.2 Dynamic I/O load-balancing: Subsystem Device Driver

The Subsystem Device Driver (SSD) is an IBM-provided pseudo-device driver that is designed to support the multipath configuration environments in the DS8000. It resides in a host system with the native disk device driver.

The dynamic I/O load-balancing option (default) of SSD is recommended to ensure better performance because:

- ▶ SSD automatically adjusts data routing for optimum performance. Multipath load balancing of data flow prevents a single path from becoming overloaded, causing input/output congestion that occurs when many I/O operations are directed to common devices along the same input/output path.
- ▶ The path to use for an I/O operation is chosen by estimating the load on each adapter to which each path is attached. The load is a function of the number of I/O operations currently in process. If multiple paths have the same load, a path is chosen at random from those paths.

For more information about the SSD, refer to *DS8000: Host Attachment and Interoperability*, SG24-8887.

7.7.3 Automatic port queues

When there is I/O between a server and a DS8800 Fibre Channel port, both the server host adapter and the DS8800 host bus adapter support queuing I/Os. How long this queue can be is called the *queue depth*. Because several servers can and usually do communicate with few DS8800 ports, the queue depth of a storage host bus adapter should be larger than the one on the server side. This is also true for the DS8800, which supports 2048 FC commands queued on a port. However, sometimes the port queue in the DS8800 HBA can be flooded.

When the number of commands sent to the DS8000 port has exceeded the maximum number of commands that the port can queue, the port has to discard these additional commands.

This operation is a *normal* error recovery operation in the Fibre Channel protocol to prevent more damage. The normal recovery is a 30-second timeout for the server, after that time the

command is resent. The server has a *command retry* count before it will fail the command. Command Timeout entries will be seen in the server logs.

Automatic Port Queues is a mechanism the DS8800 uses to self-adjust the queue based on the workload. This allows higher port queue oversubscription while maintaining a fair share for the servers and the accessed LUNs.

The port that the queue is filling up goes into SCSI Queue Fill mode, where it accepts no additional commands to slow down the I/Os.

By avoiding error recovery and the 30 second blocking SCSI Queue Full recovery interval, the overall performance is better with Automatic Port Queues.

7.7.4 Determining where to attach the host

When determining where to attach multiple paths from a single host system to I/O ports on a host adapter to the storage facility image, the following considerations apply:

- ▶ Choose the attached I/O ports on separate host adapters.
- ▶ Spread the attached I/O ports evenly between the four I/O enclosure groups.

The DS8000 host adapters have no server affinity, but the device adapters and the rank have server affinity. Figure 7-17 shows a host that is connected through two FC adapters to two DS8000 host adapters located in separate I/O enclosures.

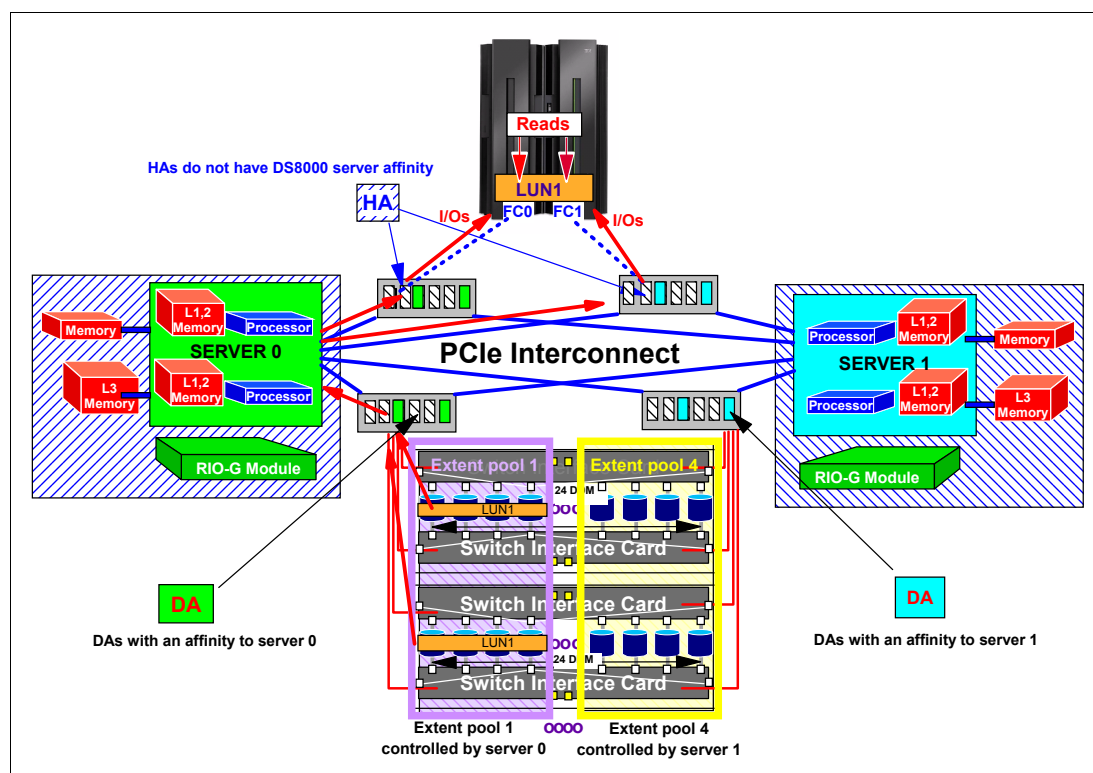


Figure 7-17 Dual-port host attachment

The host has access to LUN1, which is created in the Extent Pool 1 controlled by the DS8000 server 0. The host system sends read commands to the storage server.

When a read command is executed, one or more logical blocks are transferred from the selected logical drive through a host adapter over an I/O interface to a host. In this case, the

logical device is managed by server 0, and the data is handled by server 0. The read data to be transferred to the host must first be present in server 0's cache. When the data is in the cache, it is then transferred through the host adapters to the host.

7.7.5 I/O Priority Manager

I/O Priority Manager enables more effective storage performance management combined with the ability to align quality of service levels to separate workloads in the system. The system will prioritize access to system resources to achieve the volume's desired quality of service based on defined performance goals of high, standard, or low of any volume. I/O Priority Manager will constantly monitor and balance system resources to help applications meet their performance targets automatically, without operator intervention.

For more detailed information, refer to *DS8000 I/O Priority Manager*, REDP-4760.

7.8 Performance and sizing considerations for System z

Here we discuss several System z specific topics regarding the performance potential of the DS8000 series. We also discuss the considerations you must have when you configure and size a DS8000 that replaces older storage hardware in System z environments.

7.8.1 Host connections to System z servers

Figure 7-18 on page 186 partially shows a configuration where a DS8000 connects to FICON hosts. Note that this figure only indicates the connectivity to the Fibre Channel switched disk subsystem through its I/O enclosure, symbolized by the rectangles.

Each I/O enclosure can hold up to four HAs. The example in Figure 7-18 on page 186 shows only eight FICON channels connected to the first two I/O enclosures. Not shown is a second FICON director, which connects in the same fashion to the remaining two I/O enclosures to provide a total of 16 FICON channels in this particular example. The DS8800 disk storage system provides up to 128 FICON channel ports. Again, note the efficient FICON implementation in the DS8000 FICON ports.

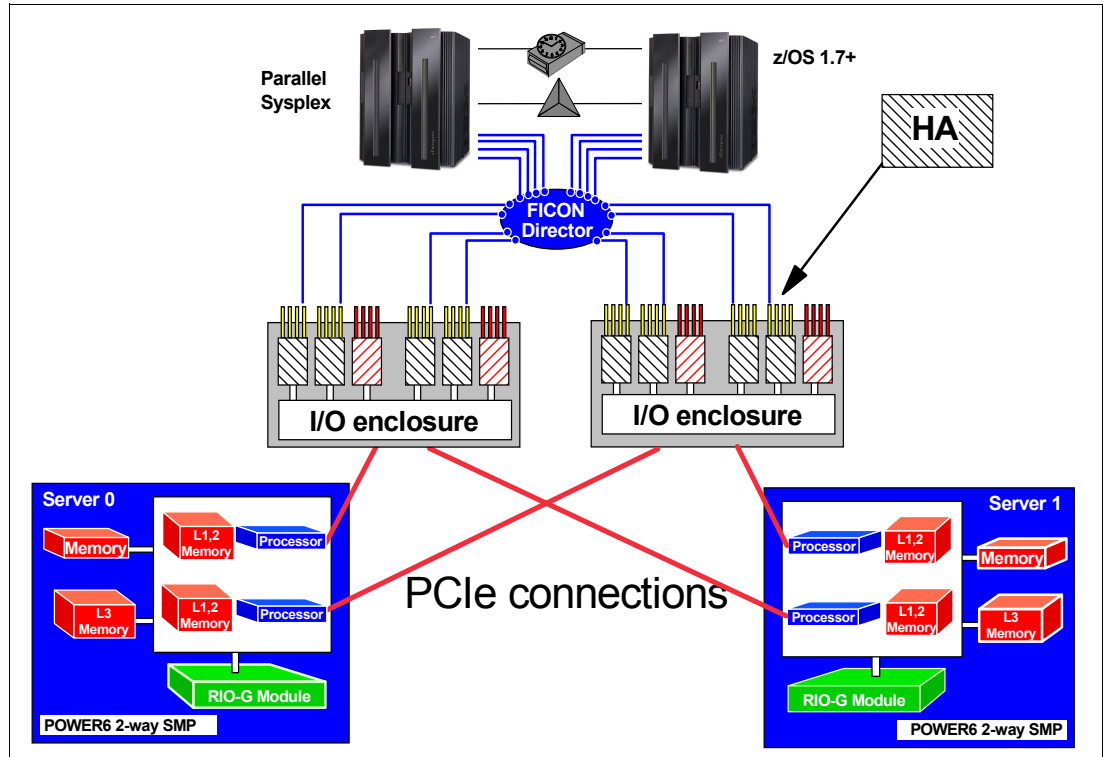


Figure 7-18 DS8000 front-end connectivity example (partial view)

Note the following performance factors:

- ▶ Do not mix ports connected to a FICON channel with a port connected to a PPRC link in the same Host Adapter.
- ▶ For large sequential loads (and with large block sizes), only use two ports per Host Adapter.

7.8.2 Parallel Access Volume

Parallel Access Volume (PAV) is one of the features that was originally introduced with the IBM TotalStorage® Enterprise Storage Server® (ESS) and that the DS8000 series has inherited. PAV is an optional licensed function of the DS8000 for the z/OS and z/VM operating systems, helping the System z servers that are running applications to concurrently share the same logical volumes.

The ability to do multiple I/O requests to the same volume nearly eliminates I/O supervisor queue delay (IOSQ) time, one of the major components in z/OS response time. Traditionally, access to highly active volumes has involved manual tuning, splitting data across multiple volumes, and more. With PAV and the Workload Manager (WLM), you can almost forget about manual performance tuning. WLM manages PAVs across all the members of a Sysplex too. This way, the DS8000 in conjunction with z/OS has the ability to meet the performance requirements by its own.

Traditional z/OS behavior without PAV

Traditional storage disk subsystems have allowed for only one channel program to be active to a volume at a time to ensure that data being accessed by one channel program cannot be altered by the activities of another channel program.

Figure 7-19 illustrates the traditional z/OS behavior without PAV, where subsequent simultaneous I/Os to volume 100 are queued while volume 100 is still busy with a preceding I/O.

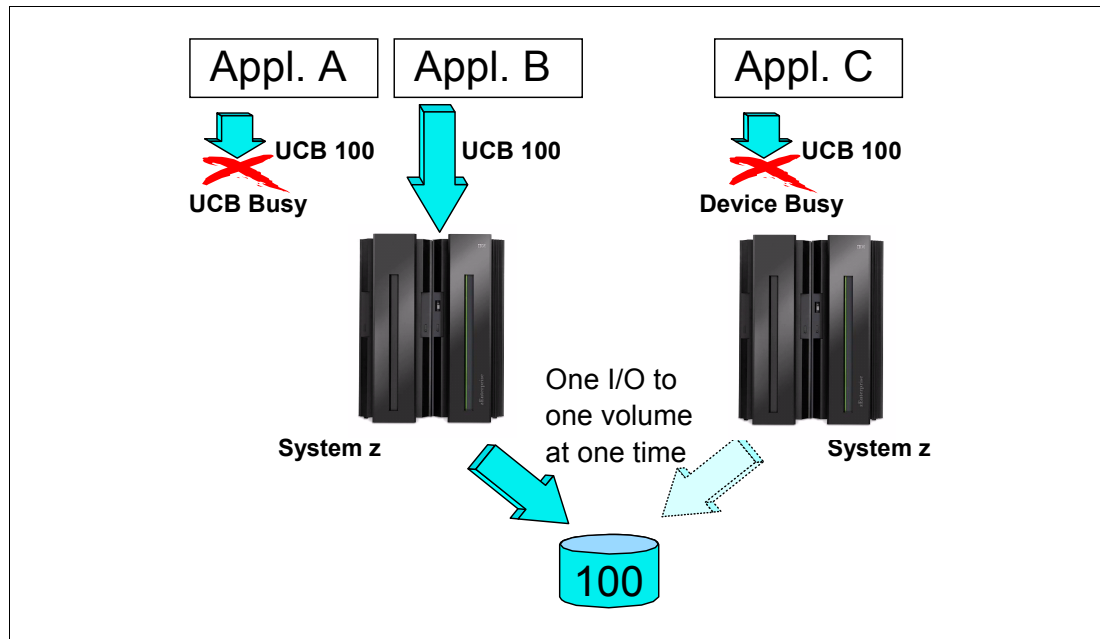


Figure 7-19 Traditional z/OS behavior

From a performance standpoint, it did not make sense to send more than one I/O at a time to the storage disk subsystem, because the hardware could process only one I/O at a time. Knowing this, the z/OS systems did not try to issue another I/O to a volume, which, in z/OS, is represented by a Unit Control Block (UCB), while an I/O was already active for that volume, as indicated by a UCB busy flag (Figure 7-19).

Not only were the z/OS systems limited to processing only one I/O at a time, but also the storage subsystems accepted only one I/O at a time from different system images to a shared volume, for the same reasons previously mentioned (Figure 7-19).

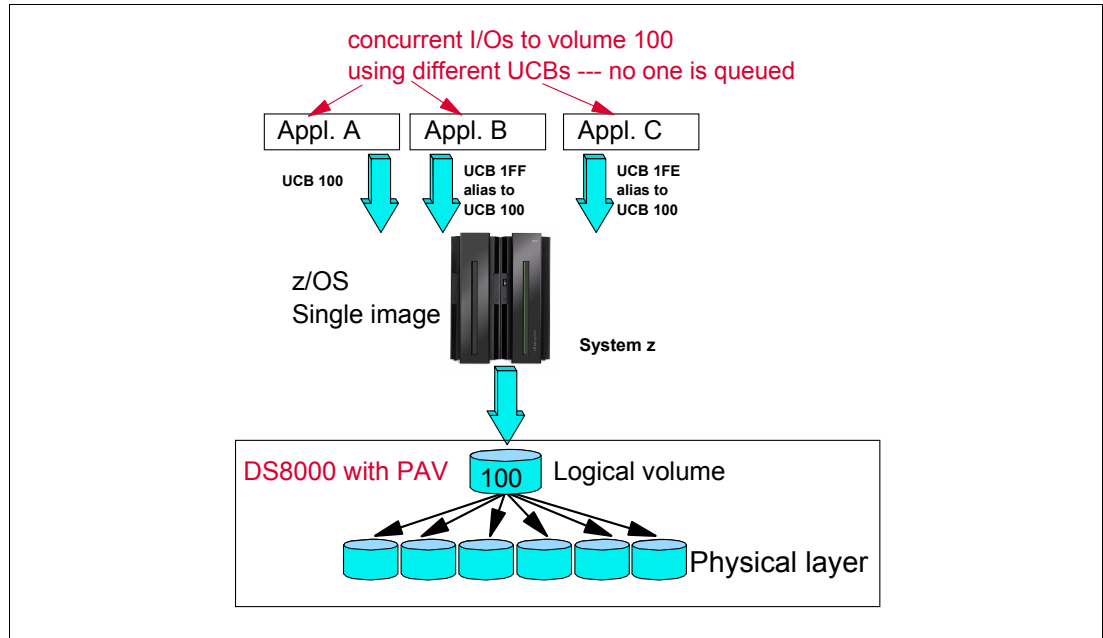


Figure 7-20 z/OS behavior with PAV

Parallel I/O capability z/OS behavior with PAV

The DS8000 has the ability to perform more than one I/O to a CKD volume. Using the alias address in addition to the conventional base address, a z/OS host can use several UCBs for the same logical volume instead of one UCB per logical volume. For example, base address 100 might have alias addresses 1FF and 1FE, which allows for three parallel I/O operations to the same volume (Figure 7-20).

This feature that allows parallel I/Os to a volume from one host is called Parallel Access Volume (PAV).

There are two concepts that are basic in PAV functionality:

- Base address

The base device address is the conventional unit address of a logical volume. There is only one base address associated with any volume.

- Alias address

An alias device address is mapped to a base address. I/O operations to an alias run against the associated base address storage space. There is no physical space associated with an alias address. You can define more than one alias per base.

Alias addresses have to be defined to the DS8000 and to the I/O definition file (IODF). This association is predefined, and you can add new aliases nondisruptively. Still, the association between base and alias is not fixed; the alias address can be assigned to another base address by the z/OS Workload Manager.

For guidelines about PAV definition and support, see *DS8000: Host Attachment and Interoperability*, SG24-8887.

PAV is an optional licensed function on the DS8000 series. PAV also requires the purchase of the FICON Attachment feature.

7.8.3 z/OS Workload Manager: Dynamic PAV tuning

It is not always easy to predict which volumes should have an alias address assigned, and how many. Your software can automatically manage the aliases according to your goals. z/OS can exploit automatic PAV tuning if you are using the z/OS Workload Manager (WLM) in Goal mode. The WLM can dynamically tune the assignment of alias addresses. The Workload Manager monitors the device performance and is able to dynamically reassign alias addresses from one base to another if predefined goals for a workload are not met.

z/OS recognizes the aliases that are initially assigned to a base during the Nucleus Initialization Program (NIP) phase. If dynamic PAVs are enabled, the WLM can reassign an alias to another base by instructing the IOS to do so when necessary (Figure 7-21).

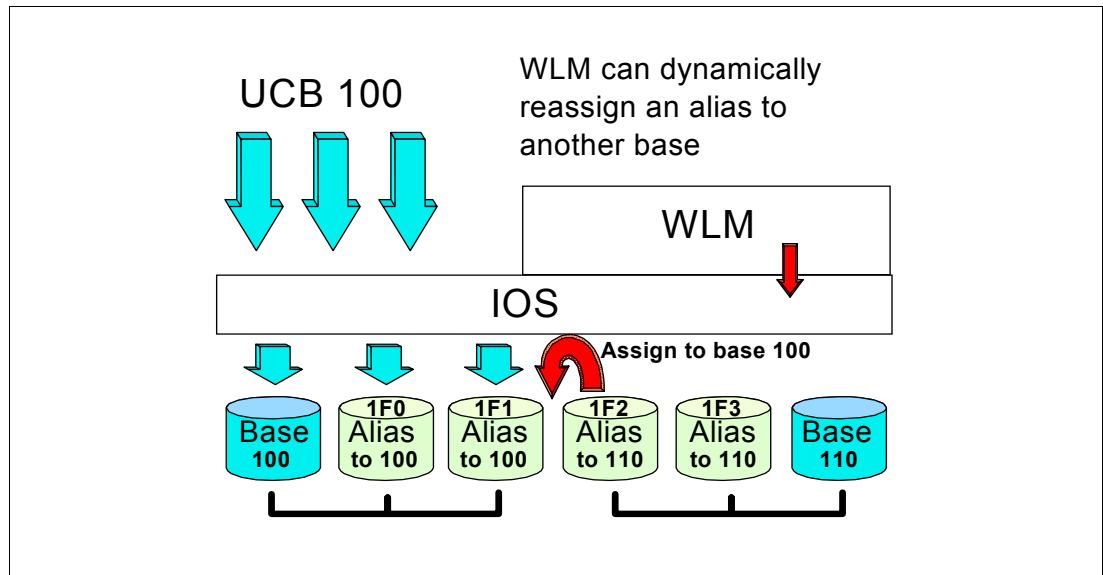


Figure 7-21 WLM assignment of alias addresses

z/OS Workload Manager in Goal mode tracks system workloads and checks whether workloads are meeting their goals as established by the installation.

WLM also keeps track of the devices utilized by the workloads, accumulates this information over time, and broadcasts it to the other systems in the same sysplex. If WLM determines that any workload is not meeting its goal due to IOS queue (IOSQ) time, WLM will attempt to find an alias device that can be reallocated to help this workload achieve its goal (Figure 7-22 on page 190).

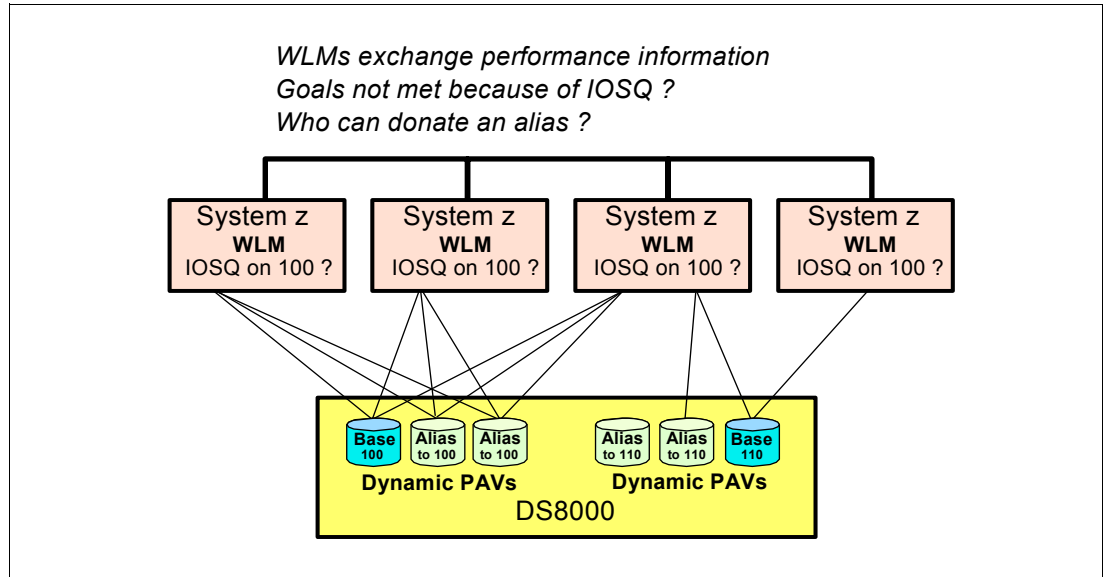


Figure 7-22 Dynamic PAVs in a sysplex

7.8.4 HyperPAV

Dynamic PAV requires the WLM to monitor the workload and goals. It takes time until the WLM detects an I/O bottleneck. Then the WLM must coordinate the reassignment of alias addresses within the sysplex and the DS8000. All of this takes time, and if the workload is fluctuating or has a burst character, the job that caused the overload of one volume could have ended before the WLM had reacted. In these cases, the IOSQ time was not eliminated completely.

With HyperPAV, an on demand proactive assignment of aliases is possible, as shown in Figure 7-23.

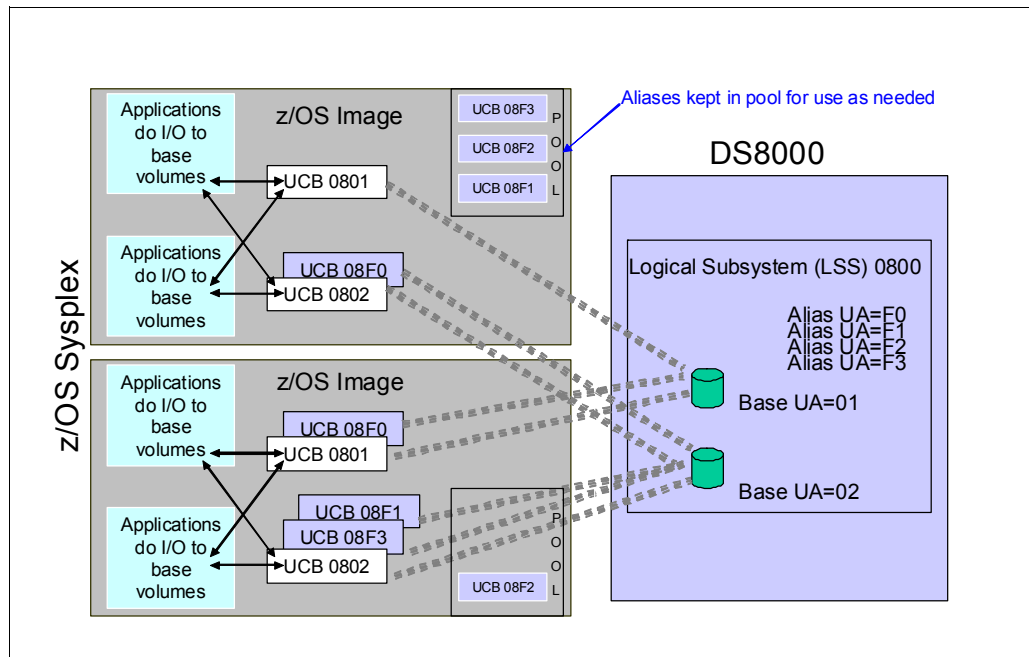


Figure 7-23 HyperPAV: Basic operational characteristics

With HyperPAV, the WLM is no longer involved in managing alias addresses. For each I/O, an alias address can be picked from a pool of alias addresses within the same LCU.

This capability also allows multiple HyperPAV hosts to use one alias to access different bases, which reduces the number of alias addresses required to support a set of bases in an IBM System z environment, with no latency in assigning an alias to a base. This functionality is also designed to enable applications to achieve better performance than is possible with the original PAV feature alone, while also using the same or fewer operating system resources.

Benefits of HyperPAV

HyperPAV has been designed to:

- ▶ Provide an even more efficient Parallel Access Volumes (PAV) function
- ▶ Help clients who implement larger volumes to scale I/O rates without the need for additional PAV alias definitions
- ▶ Exploit the FICON architecture to reduce impact, improve addressing efficiencies, and provide storage capacity and performance improvements:
 - More dynamic assignment of PAV aliases improves efficiency
 - Number of PAV aliases needed might be reduced, taking fewer from the 64 K device limitation and leaving more storage for capacity use
- ▶ Enable a more dynamic response to changing workloads
- ▶ Simplified management of aliases
- ▶ Make it easier for users to make a decision to migrate to larger volume sizes

Optional licensed function

HyperPAV is an optional licensed function of the DS8000 series. It is required in addition to the normal PAV license, which is capacity dependent. The HyperPAV license is independent of the capacity.

HyperPAV alias consideration on EAV

HyperPAV provides a far more agile alias management algorithm, as aliases are dynamically bound to a base for the duration of the I/O for the z/OS image that issued the I/O. When I/O completes, the alias is returned to the pool in the LCU. It then becomes available to subsequent I/Os.

Our rule of thumb is that the number of aliases required can be approximated by the peak of the following multiplication: I/O rate multiplied by the average response time. For example, if the peak of the above calculation happened when the I/O rate is 2000 I/O per second and the average response time is 4 ms (which is 0.004 sec), then the result of above calculation will be:

$$2000 \text{ IO/sec} \times 0.004 \text{ sec/IO} = 8$$

This means that the average number of I/O operations executing at one time for that LCU during the peak period is eight. Therefore, eight aliases should be able to handle the peak I/O rate for that LCU. However, because this calculation is based on the average during the RMF™ period, you should multiply the result by two, to accommodate higher peaks within that RMF interval. So in this case, the recommended number of aliases would be:

$$2 \times 8 = 16$$

Depending on the kind of workload, there is a huge reduction in PAV-alias UCBs with HyperPAV. The combination of HyperPAV and EAV allows you to significantly reduce the constraint on the 64 K device address limit and in turn increase the amount of addressable storage available on z/OS. In conjunction with Multiple Subchannel Sets (MSS) on IBM System z196 (zEnterprise), z10, and z9, you have even more flexibility in device configuration. Keep in mind that the EAV volumes will be supported only on IBM z/OS V1.10 and later. Refer to *IBM System Storage DS8000: Host Attachment and Interoperability*, SG24-8887, for more details about EAV specifications and considerations.

Tip: For more details about MSS, see *Multiple Subchannel Sets: An Implementation View*, REDP-4387, found at:

<http://www.redbooks.ibm.com/abstracts/redp4387.html?Open>

HyperPAV implementation and system requirements

For support and implementation guidance, see *DS8000: Host Attachment and Interoperability*, SG24-8887.

RMF reporting on PAV

RMF reports the number of exposures for each device in its Monitor/DASD Activity report and in its Monitor II and Monitor III Device reports. If the device is a HyperPAV base device, the number is followed by an 'H', for example, 5.4H. This value is the average number of HyperPAV volumes (base and alias) in that interval. RMF reports all I/O activity against the base address, not by the individual base and associated aliases. The performance information for the base includes all base and alias I/O activity.

HyperPAV would help minimize the Input/Output Supervisor Queue (IOSQ) Time. If you still see IOSQ Time, then there are two possible reasons:

- ▶ There are more aliases required to handle the I/O load compared to the number of aliases defined in the LCU.
- ▶ There is Device Reserve issued against the volume. A Device Reserve would make the volume unavailable to the next I/O, causing the next I/O to be queued. This delay will be recorded as IOSQ Time.

7.8.5 PAV in z/VM environments

z/VM provides PAV support in the following ways:

- ▶ As traditionally supported, for VM guests as dedicated guests through the CP ATTACH command or DEDICATE user directory statement.
- ▶ Starting with z/VM 5.2.0, with APAR VM63952, VM supports PAV minidisks.

Figure 7-24 and Figure 7-25 on page 193 illustrate PAV in a z/VM environment.

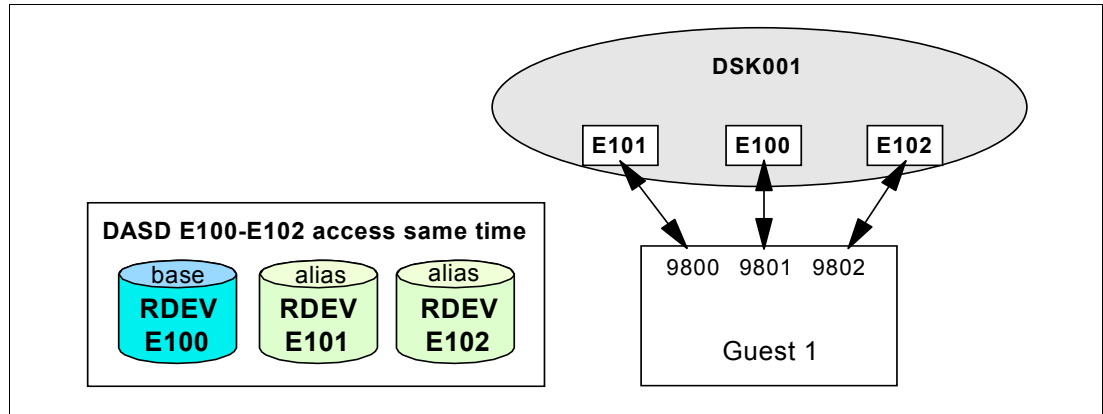


Figure 7-24 z/VM support of PAV volumes dedicated to a single guest virtual machine

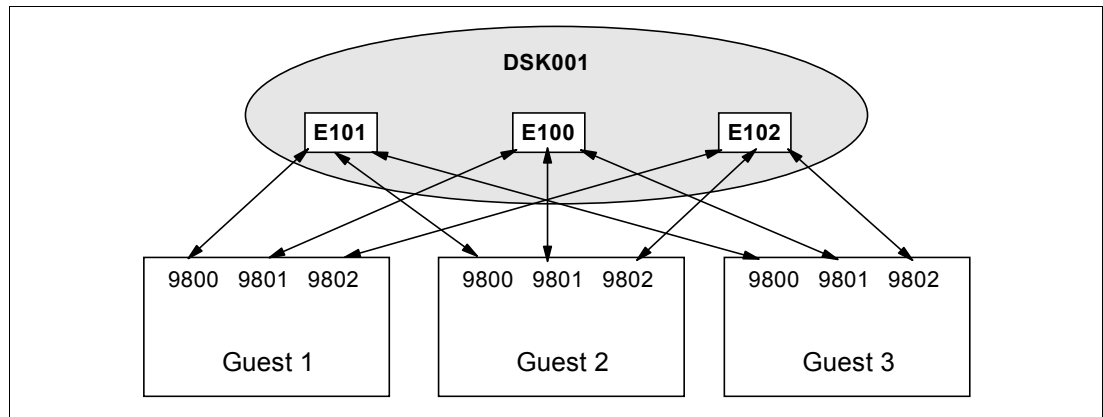


Figure 7-25 Linkable minidisks for guests that exploit PAV

In this way, PAV provides to the z/VM environments the benefits of a greater I/O performance (throughput) by reducing I/O queuing.

With the small programming enhancement (SPE) introduced with z/VM 5.2.0 and APAR VM63952, additional enhancements are available when using PAV with z/VM. For more information, see 10.4, “z/VM considerations” in *DS8000: Host Attachment and Interoperability*, SG24-8887.

7.8.6 Multiple Allegiance

Normally, if any System z host image (server or LPAR) does an I/O request to a device address for which the storage disk subsystem is already processing an I/O that came from another System z host image, then the storage disk subsystem will send back a *device busy* indication, as shown in Figure 7-19 on page 187. This delays the new request and adds to the overall response time of the I/O; this delay is shown in the Device Busy Delay (AVG DB DLY) column in the RMF DASD Activity Report. Device Busy Delay is part of the Pend time.

The DS8000 series accepts multiple I/O requests from different hosts to the same device address, increasing parallelism and reducing channel impact. In older storage disk subsystems, a device had an implicit allegiance, that is, a relationship created in the control unit between the device and a channel path group when an I/O operation is accepted by the device. The allegiance causes the control unit to guarantee access (no busy status

presented) to the device for the remainder of the channel program over the set of paths associated with the allegiance.

With Multiple Allegiance, the requests are accepted by the DS8000 and all requests are processed in parallel, unless there is a conflict when writing to the same data portion of the CKD logical volume, as shown in Figure 7-26.

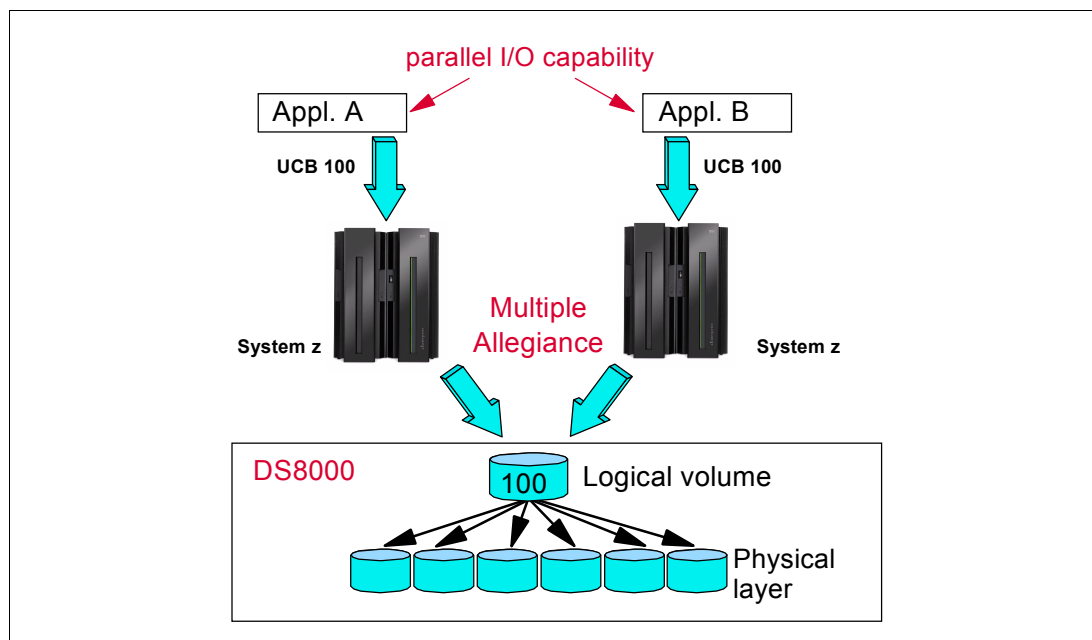


Figure 7-26 Parallel I/O capability with Multiple Allegiance

Nevertheless, good application software access patterns can improve global parallelism by avoiding reserves, limiting the extent scope to a minimum, and setting an appropriate file mask, for example, if no write is intended.

In systems without Multiple Allegiance, all except the first I/O request to a shared volume are rejected, and the I/Os are queued in the System z channel subsystem, showing up in Device Busy Delay and PEND time in the RMF DASD Activity reports. Multiple Allegiance will allow multiple I/Os to a single volume to be serviced concurrently. However, a device busy condition can still happen. This will occur when an active I/O is writing a certain data portion on the volume and another I/O request comes in and tries to either read or write to that same data. To ensure data integrity, those subsequent I/Os will get a busy condition until that previous I/O is finished with the write operation.

Multiple Allegiance provides significant benefits for environments running a sysplex, or System z systems sharing access to data volumes. Multiple Allegiance and PAV can operate together to handle multiple requests from multiple hosts.

7.8.7 I/O priority queuing

The concurrent I/O capability of the DS8000 allows it to execute multiple channel programs concurrently, as long as the data accessed by one channel program is not altered by another channel program.

Queuing of channel programs

When the channel programs conflict with each other and must be serialized to ensure data consistency, the DS8000 will internally queue channel programs. This subsystem I/O queuing capability provides significant benefits:

- ▶ Compared to the traditional approach of responding with a *device busy* status to an attempt to start a second I/O operation to a device, I/O queuing in the storage disk subsystem eliminates the impact associated with posting status indicators and redriving the queued channel programs.
- ▶ Contention in a shared environment is eliminated. Channel programs that cannot execute in parallel are processed in the order that they are queued. A fast system cannot monopolize access to a volume also accessed from a slower system. Each system gets a fair share.

Priority queuing

I/Os from different z/OS system images can be queued in a priority order. It is the z/OS Workload Manager that makes use of this priority to privilege I/Os from one system against the others. You can activate I/O priority queuing in WLM Service Definition settings. WLM has to run in Goal mode.

When a channel program with a higher priority comes in and is put in front of the queue of channel programs with lower priority, the priority of the low-priority programs will be increased (Figure 7-27). This prevents high-priority channel programs from dominating lower priority ones and gives each system a fair share.

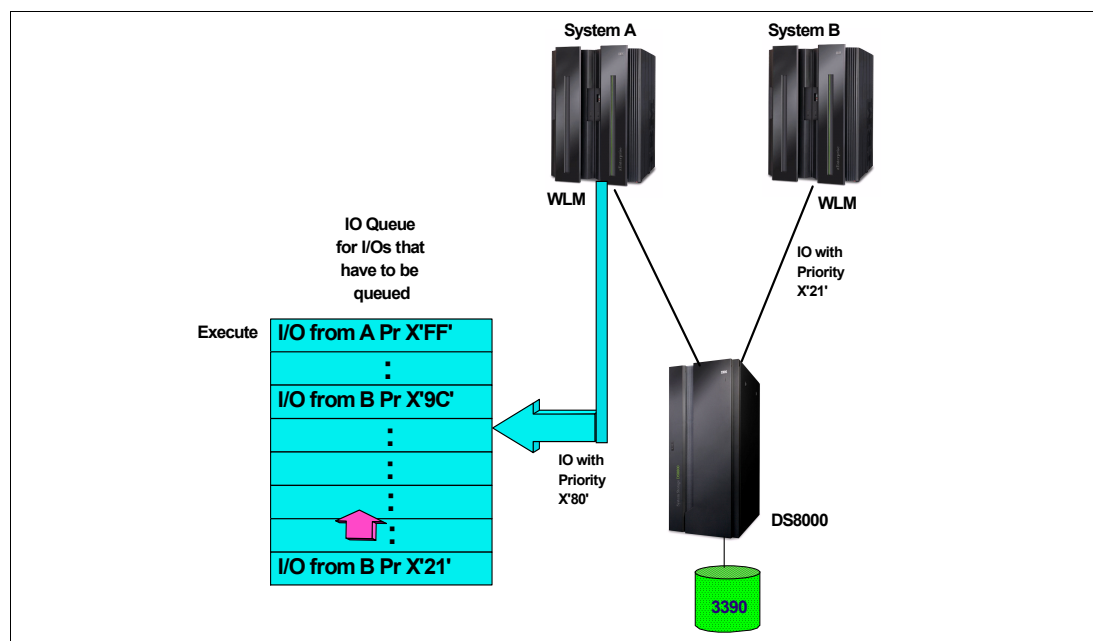


Figure 7-27 I/O priority queuing

7.8.8 Performance considerations on Extended Distance FICON

The function known as Extended Distance FICON produces performance results similar to z/OS Global Mirror (zGM) Emulation/XRC Emulation at long distances. Extended Distance FICON does not really extend the distance supported by FICON, but can provide the same benefits as XRC Emulation. In other words, with Extended Distance FICON, there is no need to have XRC Emulation running on the Channel extender.

For support and implementation discussions, see 10.6, “Extended Distance FICON” in *DS8000: Host Attachment and Interoperability*, SG24-8887.

Figure 7-28 shows Extended Distance FICON (EDF) performance comparisons for a sequential write workload. The workload consists of 64 jobs performing 4 KB sequential writes to 64 data sets with 1113 cylinders each, which all reside on one large disk volume. There is one SDM configured with a single, non-enhanced reader to handle the updates. When turning the XRC Emulation off (Brocade emulation in the diagram), the performance drops significantly, especially at longer distances. However, after the Extended Distance FICON (Persistent IU Pacing) function is installed, the performance returns to where it was with XRC Emulation on.

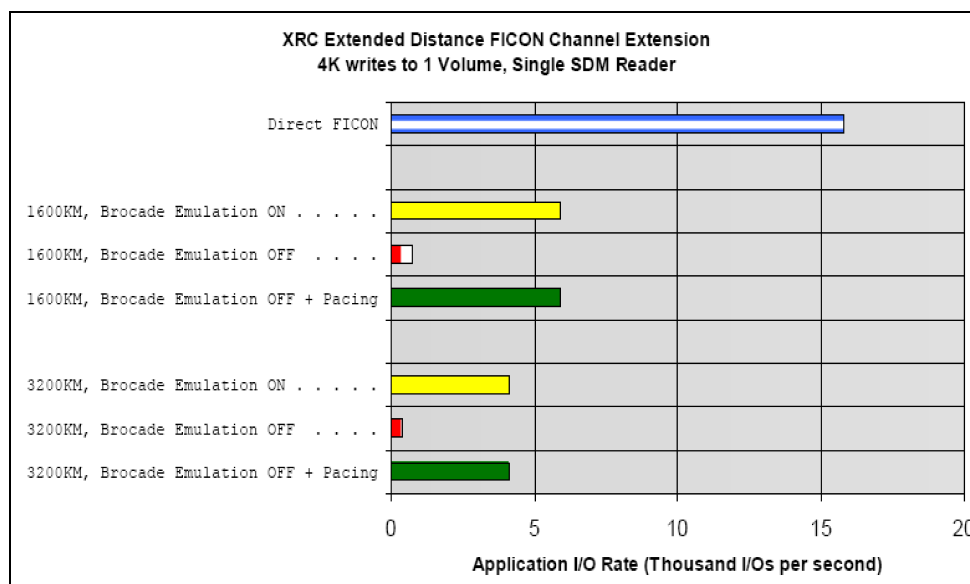


Figure 7-28 Extended Distance FICON with small data blocks sequential writes on one SDM reader

Figure 7-29 shows EDF performance, this time used in conjunction with Multiple Reader support. There is one SDM configured with four enhanced readers.

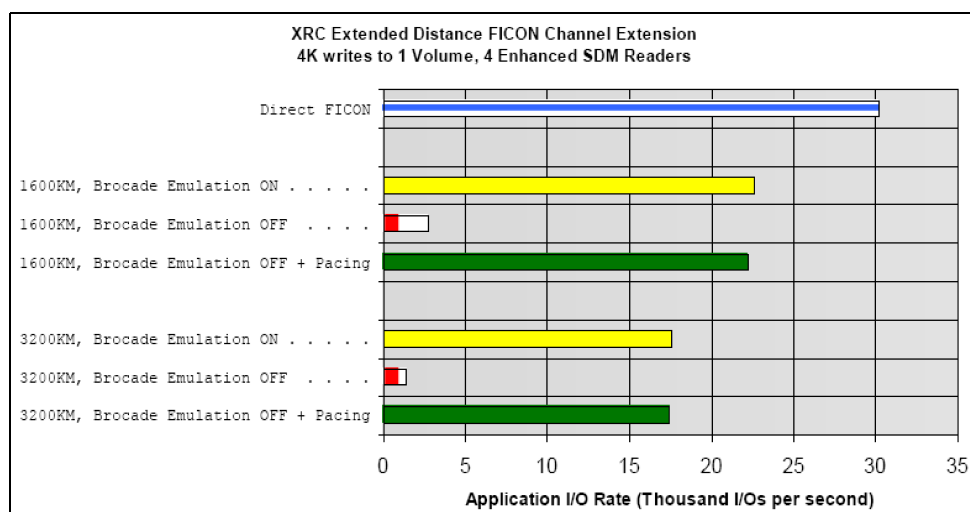


Figure 7-29 Extended Distance FICON with small data blocks sequential writes on four SDM readers

These results again show that when the XRC Emulation is turned off, performance drops significantly at long distances. When the Extended Distance FICON function is installed, the performance again improves significantly.

7.8.9 High Performance FICON for z

The FICON protocol involved several exchanges between the channel and the control unit. This led to unnecessary overhead. With High Performance FICON, the protocol has been streamlined and the number of exchanges has been reduced (Figure 7-30).

High Performance FICON for z (zHPF) is an enhanced FICON protocol and system I/O architecture that results in improvements for small block transfers (a track or less) to disk using the device independent random access method. Instead of Channel Command Word (CCWs), Transport Control Words (TCWs) can be used. I/O that is using the Media Manager, like DB2, PDSE, VSAM, zFS, VTOC Index (CVAF), Catalog BCS/VVDS, or Extended Format SAM, will benefit from zHPF.

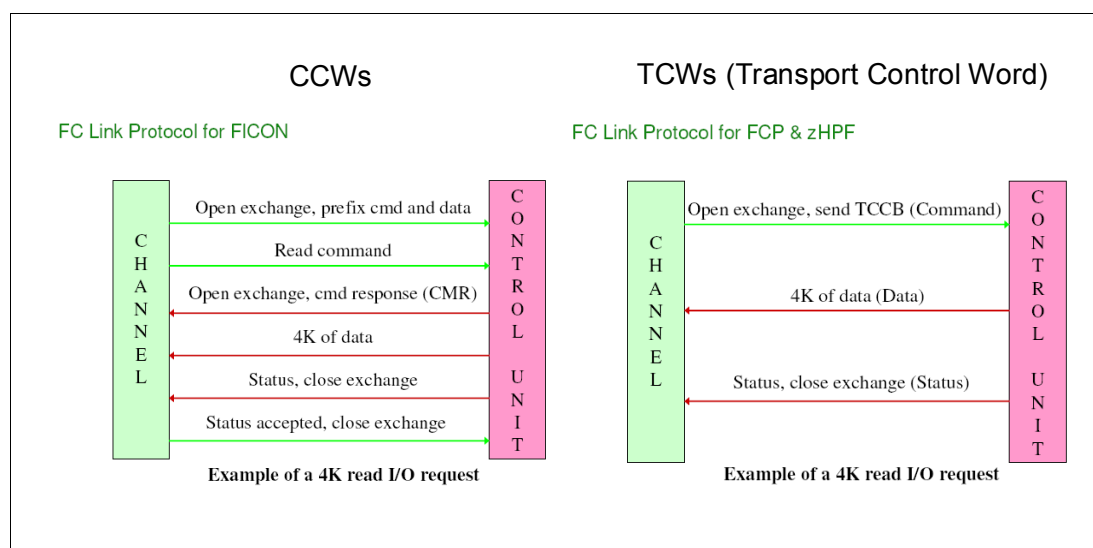


Figure 7-30 zHPF protocol

High Performance FICON for z (zHPF) is an optional licensed feature.

In situations where this is the exclusive access in use, it can improve FICON I/O throughput on a single DS8000 port by 100%. Realistic workloads with a mix of data set transfer sizes can see a 30% to 70% increase in FICON I/Os utilizing zHPF, resulting in up to a 10% to 30% channel utilization savings.

Although clients can see I/Os complete faster as the result of implementing zHPF, the real benefit is expected to be obtained by using fewer channels to support existing disk volumes, or increasing the number of disk volumes supported by existing channels.

Additionally, the changes in architecture offer end-to-end system enhancements to improve reliability, availability, and serviceability (RAS).

Only the System z196 (zEnterprise) or z10 processors support zHPF, and only on the FICON Express8, FICON Express 4, or FICON Express2 adapters. The FICON Express adapters are *not* supported. The required software is z/OS V1.7 with IBM Lifecycle Extension for z/OS V1.7 (5637-A01), z/OS V1.8, z/OS V1.9, or z/OS V1.10 with PTFs, or z/OS 1.11 and higher.

IBM Laboratory testing and measurements are available at the following website:

http://www.ibm.com/systems/z/hardware/connectivity/ficon_performance.html

zHPF is transparent to applications. However, z/OS configuration changes are required: Hardware Configuration Definition (HCD) must have Channel path ID (CHPID) type FC defined for all the CHPIDs that are defined to the 2107 control unit, which also supports zHPF. For the DS8000, installation of the Licensed Feature Key for the zHPF Feature is required. After these items are addressed, existing FICON port definitions in the DS8000 will function in either FICON or zHPF protocols in response to the type of request being performed. These are nondisruptive changes.

For z/OS, after the PTFs are installed in the LPAR, you must then set ZHPF=YES in IECIOSxx in SYS1.PARMLIB or issue the SETIOS ZHPF=YES command. ZHPF=NO is the default setting.

IBM suggests that clients use the ZHPF=YES setting after the required configuration changes and prerequisites are met. For more information about zHPF in general, refer to:

<http://www.ibm.com/systems/z/resources/faq/index.html>

zHPF multitrack support

Although the original zHPF implementation supported the new Transport Control Words only for I/O that did not span more than a track, the DS8800 supports TCW also for I/O operations on multiple tracks.

7.8.10 Extended distance High Performance FICON

This feature allows clients to achieve equivalent FICON write performance at a distance, because certain existing clients running multiple sites at long distances (10 to 100 km) cannot exploit zHPF due to the large impact to the write I/O service time.

Figure 7-31 shows that on the base code, without this feature, going from 0 km to 20 km will increase the service time by 0.4 ms. With the extended distance High Performance FICON, the service time increase will be reduced to 0.2 ms.

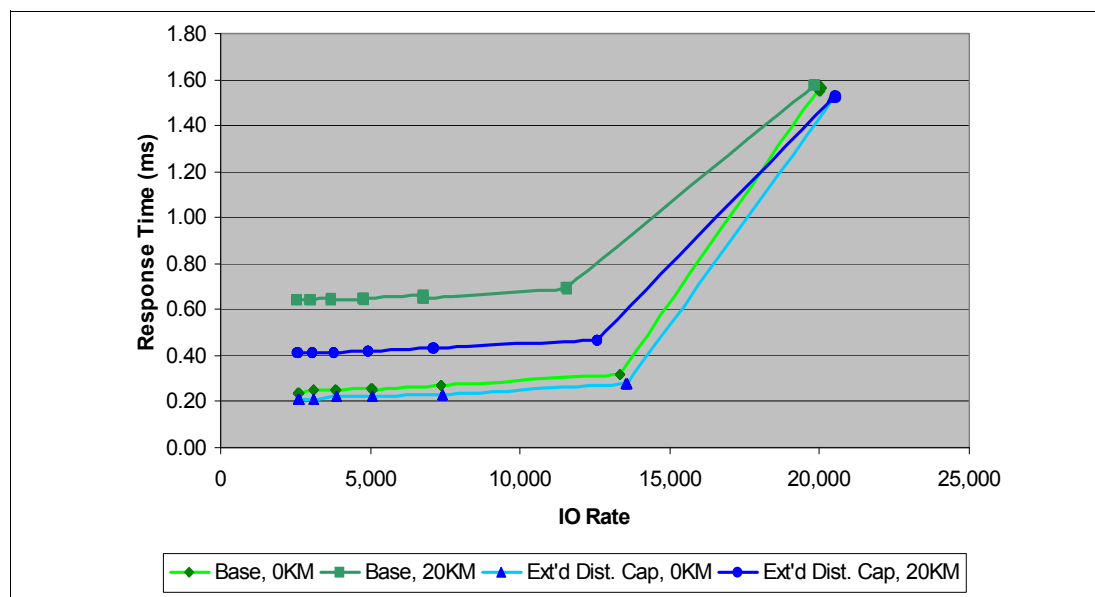


Figure 7-31 Single port 4 K write hit



Part 2

Planning and installation

In this part, we discuss matters related to the installation planning process for the IBM System Storage DS8000 series. We cover the following topics:

- ▶ Physical planning and installation
- ▶ DS8000 HMC planning and setup
- ▶ IBM System Storage DS8000 features and license keys



Physical planning and installation

This chapter discusses the various steps involved in the planning and installation of the IBM System Storage DS8700 and DS8800, including a reference listing of the information required for the setup and where to find detailed technical reference material. The topics covered include:

- ▶ Considerations prior to installation
- ▶ Planning for the physical installation
- ▶ Network connectivity planning
- ▶ Secondary HMC, TPC BE, SSPC, TKLM, LDAP, and Business-to-Business VPN planning
- ▶ Remote mirror and copy connectivity
- ▶ Disk capacity considerations
- ▶ Planning for growth

Refer to *IBM System Storage DS8700 and DS8800 Introduction and Planning Guide*, GC27-2297, for additional information and details that you will need during the configuration and installation process.

8.1 Considerations prior to installation

Start by developing and following a project plan to address the many topics needed for a successful implementation. In general, the following items should be considered for your installation planning checklist:

- ▶ Plan for growth to minimize disruption to operations. Expansion frames can only be placed to the right (from the front) of the DS8000.
- ▶ Location suitability, floor loading, access constraints, elevators, doorways, and so on.
- ▶ Power requirements: Redundancy and use of Uninterrupted Power Supply (UPS).
- ▶ Environmental requirements: Adequate cooling capacity.
- ▶ A place and connection for the secondary HMC.
- ▶ A plan for encryption integration if FDE drives are considered for the configuration, including a place and connection for the TKLM server.
- ▶ Integration of LDAP to allow a single user ID / password management.
- ▶ Business to Business VPN for the DS8000 to allow fast data offload and service connections.
- ▶ A plan for type of disks (SSD, Enterprise, Nearline)
- ▶ A plan detailing the desired logical configuration of the storage.
- ▶ Consider TPC monitoring and for DS8000 storage manager management in your environment.
- ▶ Consider the use of the I/O Priority Manager feature to prioritize specific applications.
- ▶ Oversee the available services from IBM to check for microcode compatibility and configuration checks.
- ▶ Available Copy Services and backup technologies.
- ▶ Consider new Resource Groups function feature for the IBM System Storage DS8000.
- ▶ A plan for a dynamic data relocation, IBM System Storage Easy Tier feature.
- ▶ Staff education and availability to implement the storage plan. Alternatively, you can use IBM or IBM Business Partner services.

Client responsibilities for the installation

The DS8000 series is specified as an IBM or IBM Business Partner installation and setup system. However, the following activities are some of the required planning and installation activities for which the client is responsible at a high level:

- ▶ Physical configuration planning is a client responsibility. Your Storage Marketing Specialist can help you plan and select the DS8000 series physical configuration and features.
- ▶ Installation planning is a client responsibility.
- ▶ Integration of LDAP and Business to Business VPN connectivity are client responsibilities. IBM can provide services to set up and integrate these components.
- ▶ Integration of TPC and SNMP into the client environment for monitoring of performance and configuration is a client responsibility. IBM can provide services to set up and integrate these components.
- ▶ Configuration and integration of TKLM servers and DS8000 Encryption for extended data security is a client responsibility. IBM provides services to set up and integrate these components.

- Logical configuration planning and application is a client responsibility. *Logical configuration* refers to the creation of RAID ranks, volumes, and LUNs, and the assignment of the configured capacity to servers. Application of the initial logical configuration and all subsequent modifications to the logical configuration are client responsibilities. The logical configuration can be created, applied, and modified using the DS Storage Manager, DS CLI, or DS Open API.

IBM Global Services will also apply or modify your logical configuration (these are fee-based services).

In this chapter, you will find information that will assist you with the planning and installation activities. Additional information can be found in *IBM System Storage DS8700 and DS8800 Introduction and Planning Guide*, GC27-2297.

8.1.1 Who should be involved

Have a project manager to coordinate the many tasks necessary for a successful installation. Installation will require close cooperation with the user community, the IT support staff, and the technical resources responsible for floor space, power, and cooling.

A storage administrator should also coordinate requirements from the user applications and systems to build a storage plan for the installation. This will be needed to configure the storage after the initial hardware installation is complete.

The following people should be briefed and engaged in the planning process for the physical installation:

- Systems and storage administrators
- Installation planning engineer
- Building engineer for floor loading, air conditioning, and electrical considerations
- Security engineers for business-to-business VPN, LDAP, TKLM, and encryption
- Administrator and operator for monitoring and handling considerations
- IBM or IBM Business Partner installation engineer

8.1.2 What information is required

A validation list to assist in the installation process should include:

- Drawings detailing the positioning as specified and agreed upon with a building engineer, ensuring the weight is within limits for the route to the final installation position.
- Approval to use elevators if the weight and size are acceptable.
- Connectivity information, servers, and SAN and mandatory LAN connections.
- Agreement on the security structure of the installed DS8000 with all security engineers.
- Ensure that you have a detailed storage plan agreed upon. Ensure that the configuration specialist has all the information to configure all the arrays and set up the environment as required.
- License keys for the Operating Environment License (OEL), which are mandatory, and any optional license keys.

Note that *IBM System Storage DS8700 and DS8800 Introduction and Planning Guide*, GC27-2297, contains additional information about physical planning. You can download it from the following URL:

<http://www.ibm.com/systems/storage/disk/ds8000/index.html>

8.2 Planning for the physical installation

This section discusses the physical installation planning process and gives important tips and considerations.

8.2.1 Delivery and staging area

The shipping carrier is responsible for delivering and unloading the DS8000 as close to its final destination as possible. Inform your carrier of the weight and size of the packages to be delivered and inspect the site and the areas where the packages will be moved (for example, hallways, floor protection, elevator size and loading, and so on).

Table 8-1 lists the final packaged dimensions and maximum packaged weight of the DS8000 storage unit shipgroup.

Table 8-1 Packaged dimensions and weight for DS8000 models

Shipping container	Packaged dimensions (in centimeters and inches)	Maximum packaged weight (in kilograms and pounds)
Model 941 (4-way) pallet or crate	Height 207.5 cm (81.7 in.) Width 101.5 cm (40 in.) Depth 137.5 cm (54.2 in.)	1378 kg (3036 lb)
Model 94E expansion unit pallet or crate	Height 207.5 cm (81.7 in.) Width 101.5 cm (40 in.) Depth 137.5 cm (54.2 in.)	1209 kg (2665 lb)
Model 951 pallet or crate	Height 207.5 cm (81.7 in.) Width 101.5 cm (40 in.) Depth 137.5 cm (54.2 in.)	1336 kg (2940 lb)
Model 95E expansion unit pallet or crate	Height 207.5 cm (81.7 in.) Width 101.5 cm (40 in.) Depth 137.5 cm (54.2 in.)	1277 kg (2810 lb)
Shipgroup (height might be lower and weight might be less)	Height 105.0 cm (41.3 in.) Width 65.0 cm (25.6 in.) Depth 105.0 cm (41.3 in.)	up to 90 kg (199 lb)
(if ordered) System Storage Productivity Center (SSPC)	Height 68.0 cm (26.8 in.) Width 65.0 cm (25.6 in.) Depth 105.0 cm (41.3 in.)	47 kg (104 lb)
(if ordered) System Storage Productivity Center (SSPC), External HMC	Height 68.0 cm (26.8 in.) Width 65.0 cm (25.6 in.) Depth 105.0 cm (41.3 in.)	62 kg (137 lb)
(if ordered as MES) External HMC container	Height 40.0 cm (17.7 in.) Width 65.0 cm (25.6 in.) Depth 105.0 cm (41.3 in.)	32 kg (71 lb)

Attention: A fully configured model in the packaging can weight over 1416 kg (3120 lbs). Use of fewer than three persons to move it can result in injury.

8.2.2 Floor type and loading

The DS8000 can be installed on a raised or nonraised floor. Installing the unit on a raised floor is preferable because this allows you to operate the storage unit with better cooling efficiency and cabling layout protection.

The total weight and space requirements of the storage unit will depend on the configuration features that you ordered. You might need to consider calculating the weight of the unit and the expansion box (if ordered) in their maximum capacity to allow for the addition of new features.

Table 8-2 lists the weights of the various DS8000 models.

Table 8-2 DS8000 weights

Model	Maximum weight
Model 941 (2-way)	1200 kg (2640 lb)
Model 941 (4-way)	1256 kg (2770 lb)
Model 941 (with Model 94E expansion model)	2354 kg (5190 lb)
Model 951 (2-way)	1215 kg (2675 lb)
Model 951 (4-way)	1218 kg (2680lb)
Model 951(with Model 95E expansion model)	2375 kg (5225 lb)

Important: You need to check with the building engineer or other appropriate personnel to make sure the floor loading was properly considered.

Raised floors can better accommodate cabling layout. The power and interface cables enter the storage unit through the rear side.

Figure 8-1 on page 206 for DS8700 and Figure 8-2 on page 206 for DS8800 show the location of the cable cutouts. You can use the following measurements when you cut the floor tile:

- Width: 45.7 cm (18.0 in.)
- Depth: 16 cm (6.3 in.)

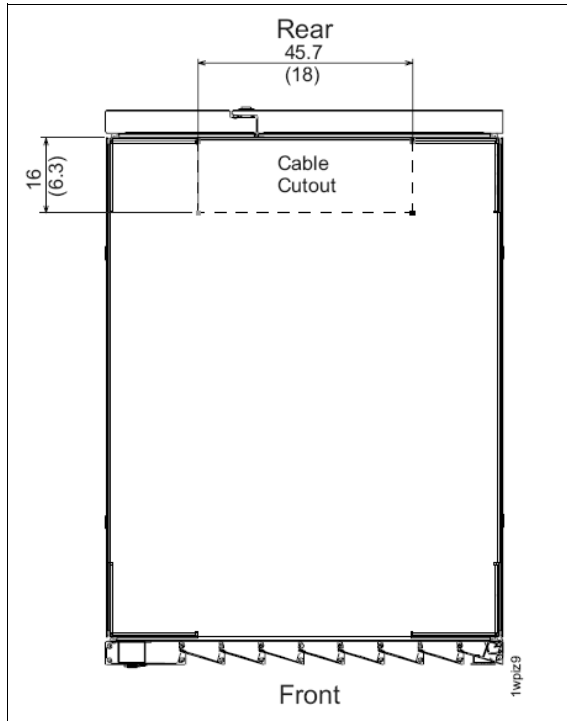


Figure 8-1 Floor tile cable cutout for DS8700

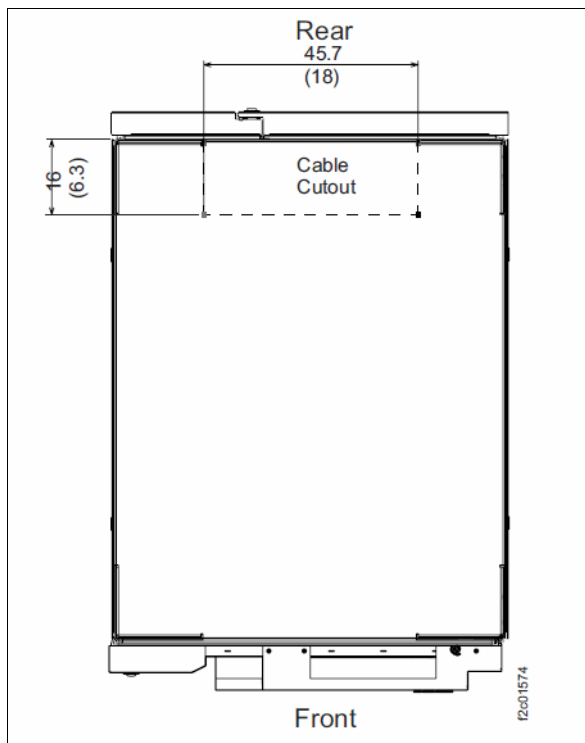


Figure 8-2 Floor tile cable cutout for DS8800

8.2.3 Room space and service clearance

The total amount of space needed by the storage units can be calculated using the dimensions in Table 8-3.

Table 8-3 DS8000 dimensions

Dimension with covers	Model 941	Model 951
Height	193 cm	193.4 cm
Width	84.7 cm	84.8 cm
Depth	118.3 cm	122.8 cm

The storage unit location area should also cover the service clearance needed by IBM service representatives when accessing the front and rear of the storage unit. You can use the following minimum service clearances. The dimensions are also shown as an example in Figure 8-3 for DS8700 and Figure 8-4 on page 208 for DS8800 with two expansion frames:

1. For the front of the unit, allow a minimum of 121.9 cm (48 in.) for the service clearance.
2. For the rear of the unit, allow a minimum of 76.2 cm (30 in.) for the service clearance.
3. For the sides of the unit, allow a minimum of 5.1 cm (2 in.) for the service clearance.

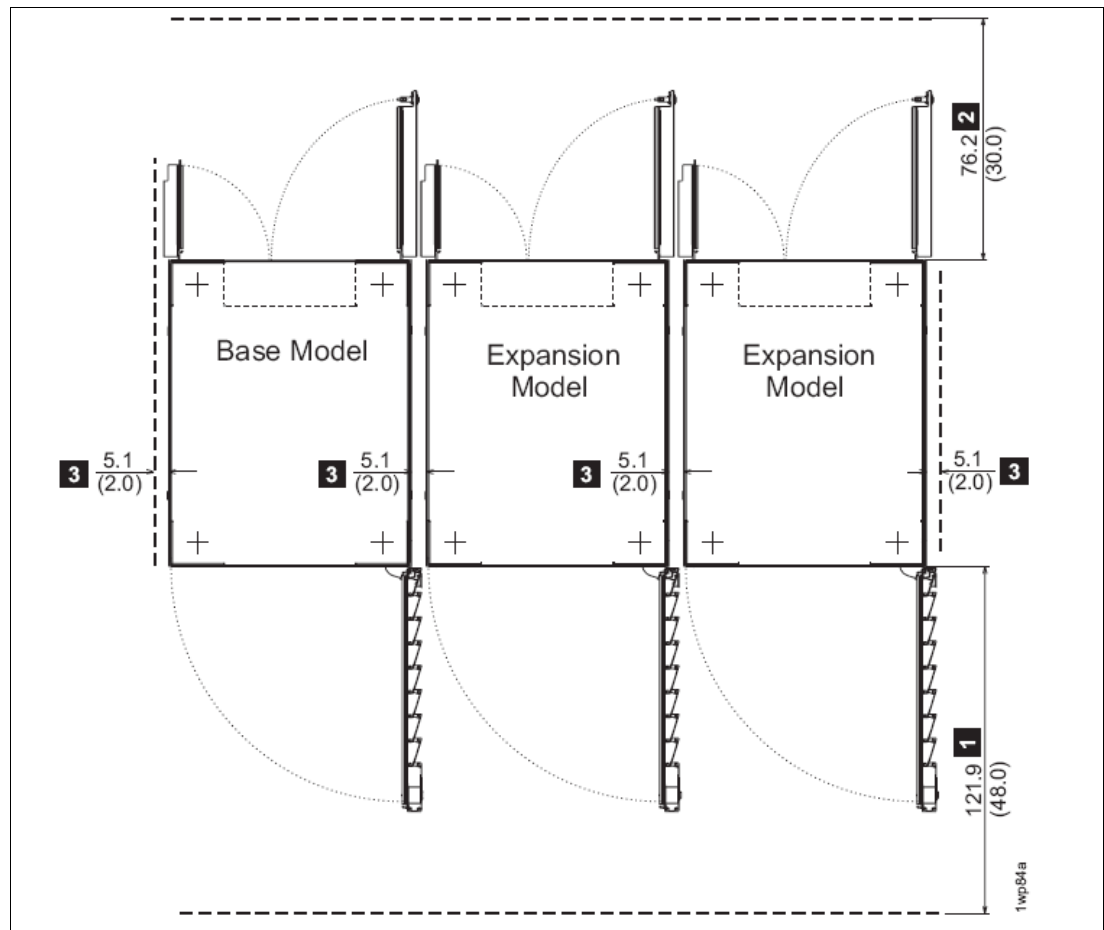


Figure 8-3 DS8700 three frame service clearance requirements

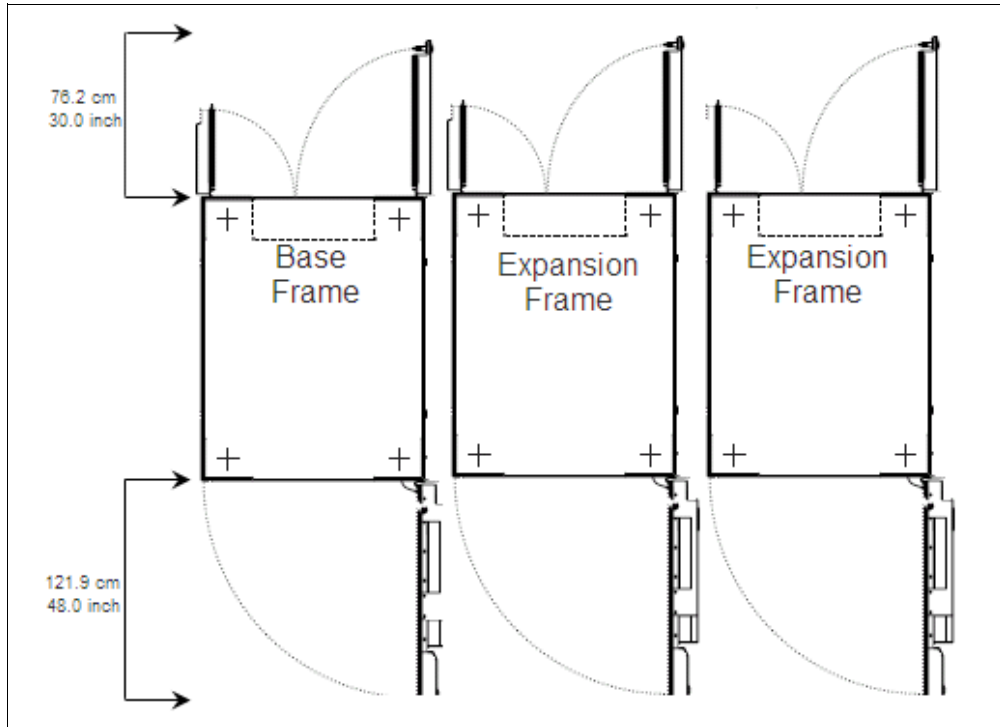


Figure 8-4 DS8000 service clearance requirements

8.2.4 Power requirements and operating environment

Consider the following basic items when planning for the DS8000 power requirements:

- ▶ Power connectors
- ▶ Input voltage
- ▶ Power consumption and environment
- ▶ Power control features
- ▶ Power Line Disturbance (PLD) feature

Power connectors

Each DS8000 base and expansion unit has redundant power supply systems. The two line cords to each frame should be supplied by separate AC power distribution systems. Use a 60 A rating for the low voltage feature and a 25 A rating for the high voltage feature.

For more details regarding power connectors and line cords, refer to *IBM System Storage DS8700 and DS8800 Introduction and Planning Guide*, GC27-2297.

Input voltage

The DS8000 supports a three-phase input voltage source. Table 8-4 lists the power specifications for each feature code.

Table 8-4 DS8000 input voltages and frequencies

Characteristic	Low voltage (Feature 9090)	High voltage (Feature 9091)
Nominal input voltage (3-phase)	200, 208, 220, or 240 RMS Vac	380, 400, 415, or 480 RMS Vac

Characteristic	Low voltage (Feature 9090)	High voltage (Feature 9091)
Minimum input voltage (3-phase)	180 RMS Vac	333 RMS Vac
Maximum input voltage (3-phase)	264 RMS Vac	508 RMS Vac
Steady-state input frequency	50 ± 3 or 60 ± 3.0 Hz	50 ± 3 or 60 ± 3.0 Hz

Power consumption

Table 8-5 lists the power consumption specifications of the DS8000. The power estimates given here are on the conservative side and assume a high transaction rate workload.

Table 8-5 DS8000 power consumption

Measurement	Model 951 (4-way)	Model 95E with I/O enclosure	Model 941 (4-way)	Model 94E with I/O enclosure
Peak electric power	7.3 kVa	7.2 KVa	6.8 KVa	6.5 KVa
Thermal load (BTU/hr)	25,000	24,600	23,000	22,200

The values represent data that was obtained from systems configured as follows:

- ▶ Model 951 base models that contain 15 disk drive sets (240 disk drives) and fibre channel adapters
- ▶ Model 95E first expansion models that contain 21 disk drive sets (336 disk drives) and fibre channel adapters
- ▶ Model 95E second expansion models that contain 30 disk drive sets (480 disk drives)
- ▶ Model 941 base models that contain eight disk drive sets (128 disk drives) and fibre channel adapters
- ▶ Model 94E first expansion models that contain 16 disk drive sets (256 disk drives) and fibre channel adapters
- ▶ Model 94E second to fourth expansion models that contain 16 disk drive sets (256 disk drives) for each expansion.

DS8000: Cooling the storage complex

Important for DS8800 only: Air circulation for the DS8800 is provided by the various fans installed throughout the frame. All of the fans on the DS8800 direct air flow from the front of the frame to the rear of the frame. No air exhausts to the top of the machine. Using a directional air flow in this manner allows for “cool aisles” to the front and “hot aisles” to the rear of the machine.

To optimize the cooling around DS8000, prepare the location of your storage images as follows:

1. Install the storage image on a raised floor. Although you can install the storage image on a non-raised floor, installing the storage image on a raised floor provides increased air circulation for better cooling.
2. Install perforated tiles in the front and back of each base model and expansion model as follows:
 - For a stand-alone base model, install two fully perforated tiles in front of each base model.
 - For a row of machines, install a row of perforated tiles in front of the machines.

- For groupings of machines, where a hot aisle/cold aisle layout is used, use a cold aisle row of perforated tiles in front of all machines.

The recommended operating temperature for the DS8000 is between 20° to 25°C (68° to 78°F) at a relative humidity range of 40% to 50%.

Important: Make sure that air circulation for the DS8000 base unit and expansion units is maintained free from obstruction to keep the unit operating in the specified temperature range.

Power control features

The DS8000 has remote power control features that allow you to control the power of the storage complex through the HMC. Another power control feature is available for the System z environment.

For more details regarding power control features, refer to *IBM System Storage DS8700 and DS8800 Introduction and Planning Guide*, GC27-2297.

Power Line Disturbance feature

The extended Power Line Disturbance (ePLD) feature stretches the available uptime of the DS8000 from 30 milliseconds to 60 seconds during a ePLD event. Generally, install this feature, especially with environments that have no UPS. There is no additional physical connection planning needed for the client with or without the ePLD.

8.2.5 Host interface and cables

The DS8000 can support the following number of Host adapter cards as shown in Table 8-6.

Table 8-6 Maximum host adapter

Base model	Attached expansion model	Maximum host adapter
941 (2-way)	None	2 - 16
941 (4-way)	None	2 - 16
	94E model (1-4)	2 - 32
951 (2-way) business class cabling	None	2 - 4
951 (2-way) standard cabling	None	2 - 4
951 (4-way) standard cabling	None (single rack)	2 - 8
	95E models (1-2)	2 - 16

The DS8000 Model 941/94E (with Release 6.1) supports four ports with the 8 Gb Fibre Channel/FICON PCI Express adapter, which is offered in shortwave and longwave versions.

The result is a maximum of 64 8-Gb ports are supported on a DS8000 Model 941.

Fibre Channel/FICON

The DS8000 Fibre Channel/FICON adapter has four or eight ports (DS8800 only) per card. Each port supports FCP or FICON, but not simultaneously. Fabric components from various vendors, including IBM, CNT, McDATA, Brocade, and Cisco, are supported by both environments.

The Fibre Channel/FICON shortwave Host Adapter, feature 3153, when used with 50 micron multi-mode fibre cable, supports point-to-point distances of up to 300 meters. The Fibre Channel/FICON longwave Host Adapter, when used with 9 micron single-mode fibre cable, extends the point-to-point distance to 10 km for feature 3245 (4 Gb 10 km LW Host Adapter). Feature 3243 (4 Gb LW Host Adapter) supports point-to-point distances up to 4 km. Additional distance can be achieved with the use of appropriate SAN fabric components.

A 31-meter fiber optic cable or a 2-meter jumper cable can be ordered for each Fibre Channel adapter port.

Table 8-7 lists the fiber optic cable features for the FCP/FICON adapters.

Table 8-7 FCP/FICON cable features

Feature	Length	Connector	Characteristic
1410	31 m	LC/LC	50 micron, multimode
1411	31 m	LC/SC	50 micron, multimode
1412	2 m	SC to LC adapter	50 micron, multimode
1420	31 m	LC/LC	9 micron, single mode
1421	31 m	LC/SC	9 micron, single mode
1422	2 m	SC to LC adapter	9 micron, single mode

Tip: The Remote Mirror and Copy functions use FCP as the communication link between the IBM System Storage DS8000 series, DS6000s, and ESS Models 800 and 750.

For more details about IBM-supported attachments, refer to *IBM System Storage DS8000 Host Systems Attachment Guide*, SC26-7917 and *IBM System Storage DS8000: Host Attachment and Interoperability*, SG24-8887.

For the most up-to-date details about host types, models, adapters, and operating systems supported by the DS8000 unit, refer to the DS8000 System Storage Interoperability Center at the following URL:

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

8.3 Network connectivity planning

Implementing the DS8000 requires that you consider the physical network connectivity of the storage adapters and the Hardware Management Console (HMC) within your local area network.

Check your local environment for the following DS8000 unit connections:

- ▶ Hardware Management Console and network access
- ▶ Tivoli Storage Productivity Center Basic Edition (if used) and network access
- ▶ System Storage Productivity Center (if used) and network access
- ▶ DS command-line interface DSCLI
- ▶ Remote support connection
- ▶ Remote power control
- ▶ Storage Area Network connection
- ▶ TKLM connection
- ▶ LDAP connection

For more details about physical network connectivity, see *IBM System Storage DS8000 User's Guide*, SC26-7915, and *IBM System Storage DS8700 and DS8800 Introduction and Planning Guide*, GC27-2297.

8.3.1 Hardware Management Console and network access

Hardware Management Consoles (HMCs) are the focal point for configuration, Copy Services management, and maintenance for a DS8000 unit. The internal HMC included with every base frame rack is mounted in a pull-out tray for convenience and security. The HMC consists of a mobile computer (Lenovo Thinkpad T510) with adapters for modem and 10/100/1000 Mb Ethernet. Ethernet cables connect the HMC to the storage unit.

A second, redundant external HMC is orderable. Having a second HMC is a good idea for environments that use TKLM encryption management and Advanced Copy Services functions. The second HMC is external to the DS8000 rack(s) and consists of a similar mobile workstation as the primary HMC.

Tip: The external HMC rack must be direct connected to the internal DS8000 hub and must also be within 15.2 m (50 ft) of the storage units that are connected to it.

The hardware management console can be connected to your network (eth2 - customer network) for:

- ▶ Remote management of your system using the DS Command-Line Interface (CLI)
- ▶ Remote DS Storage Manager GUI management of your system by using System Storage Productivity Center (SSPC) or from Rel. 6.1 using Tivoli Productivity Center Basic Edition (TPC BE) installed on customer server. Connecting the System Storage Productivity Center (SSPC) or TPC BE server to your LAN allows you to access the DS Storage Manager GUI from any location that has network access.

To connect the hardware management consoles (internal, and external if present) to your network, you need to provide the following settings to your IBM service representative so the management consoles can be configured for attachment to your LAN:

- ▶ Management console network IDs, host names, and domain name
- ▶ Domain Name Server (DNS) settings (if you plan to use DNS to resolve network names)
- ▶ Routing information

For additional information regarding the HMC planning, see Chapter 9, “DS8000 HMC planning and setup” on page 225.

8.3.2 IBM Tivoli Storage Productivity Center

The IBM Tivoli Storage Productivity Center is an integrated software solution that can help you improve and centralize the management of your storage environment through the integration of products. With the Tivoli Storage Productivity Center (TPC), it is possible to manage and fully configure multiple DS8000 storage systems from a single point of control. TPC Basic Edition (BE) is the minimum requirement.

Tip: Tivoli Storage Productivity Center Basic Edition is required at a minimum, to remotely access the DS Storage Manager GUI.

With DS8000 Rel. 6.1, the IBM System Storage DS8000 Storage Manager is accessible using the IBM Tivoli Storage Productivity Center. IBM Tivoli Storage Productivity Center provides a DS8000 management interface. You can use this interface to add and manage multiple DS8000 series storage units from one console.

8.3.3 System Storage Productivity Center and network access

SSPC (2805-MC5), consisting of hardware and software components, is an optional hardware feature for all systems running Release 6.1 level microcode or later, provides a convenient ordering option for TPC-BE

SSPC hardware

The SSPC (IBM model 2805-MC5) server contains the following hardware components:

- ▶ x86 server 1U rack installed
- ▶ Intel Quadcore Xeon processor running at 2.53 GHz
- ▶ 8 GB of RAM
- ▶ Two hard disk drives
- ▶ Dual port Gigabit Ethernet

Optional components are:

- ▶ KVM Unit
- ▶ 8 Gb Fibre Channel Dual Port HBA (this feature enables you to move the Tivoli Storage Productivity Center database from the SSPC server to the IBM System Storage DS8000)
- ▶ Secondary power supply
- ▶ Additional hard disk drives
- ▶ CD media to recover image for 2805-MC5

SSPC software

The IBM System Storage Productivity Center includes the preinstalled (separately purchased) software, running under a licensed Microsoft Windows Server 2008 Enterprise Edition R2 64-bit.

Clients have the option to purchase and install the individual software components to create their own SSPC server.

For additional details, refer to SSPC manuals:

<http://public.dhe.ibm.com/common/ssi/ecm/en/tsd03035usen/TSD03035USEN.PDF>

You can also refer to the SSPC IBM Information Center:

<http://publib.boulder.ibm.com/infocenter/tivihelp/v4r1/index.jsp>

For details, see Chapter 12, “System Storage Productivity Center” on page 279, and *IBM System Storage Productivity Center Deployment Guide*, SG24-7560.

Network connectivity

To connect the System Storage Productivity Center (SSPC) to your network, you need to provide the following settings to your IBM service representative:

- ▶ SSPC network IDs, host names, and domain name
- ▶ Domain Name Server (DNS) settings (if you plan to use DNS to resolve network names)

Routing information

For the GUI management, the networks ports 8451 and 8452 need to be opened between the SSPC console and the DS8000, and the LDAP server if the SSPC is installed behind a firewall.

8.3.4 DS command-line interface DSCLI

The IBM System Storage DSCLI can be used to create, delete, modify, and view Copy Services functions and the logical configuration of a storage unit. These tasks can be performed either interactively, in batch processes (operating system shell scripts), or in DSCLI script files. A DSCLI script file is a text file that contains one or more DSCLI commands and can be issued as a single command. DSCLI can be used to manage logical configuration, Copy Services configuration, and other functions for a storage unit, including managing security settings, querying point-in-time performance information or status of physical resources, and exporting audit logs.

The DSCLI can be installed on and used from a LAN-connected system, such as the storage administrator's workstation or any separate server connected to the storage unit's LAN.

For details about the hardware and software requirements for the DSCLI, refer to *IBM Systems Storage DS8000 Series: Command-Line Interface User's Guide*, SC26-7916.

8.3.5 Remote support connection

Remote support connection is available from the HMC using a modem (dial-up) and the Virtual Private Network (VPN) over the Internet through the Client LAN.

You can take advantage of the DS8000 remote support feature for outbound calls (Call Home function) or inbound calls (remote service access by an IBM technical support representative). You need to provide an analog telephone line for the HMC modem.

Figure 8-5 shows a typical remote support connection.

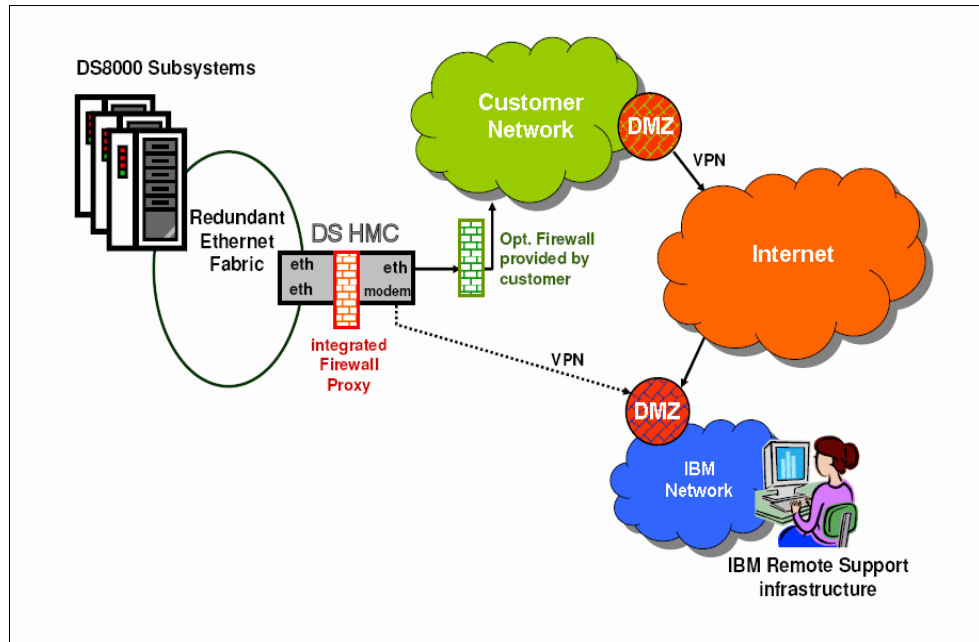


Figure 8-5 DS8000 HMC remote support connection

Note the following guidelines to assist in the preparation for attaching the DS8000 to the Client's LAN:

1. Assign a TCP/IP address and host name to the HMC in the DS8000.
2. If email notification of service alert is allowed, enable the support on the mail server for the TCP/IP addresses assigned to the DS8000.
3. Use the information that was entered on the installation work sheets during your planning.

Generally, use a service connection through the high-speed VPN network utilizing a secure Internet connection. You need to provide the network parameters for your HMC through the installation worksheet prior to actual configuration of the console. See Chapter 9, "DS8000 HMC planning and setup" on page 225 for more details.

Your IBM System Service Representative (SSR) will need the configuration worksheet during the configuration of your HMC. A worksheet is available in *IBM System Storage DS8700 and DS8800 Introduction and Planning Guide*, GC27-2297.

See Chapter 17, "Remote support" on page 425 for further information about remote support connection.

8.3.6 Remote power control

The System z remote power control setting allows you to power on and off the storage unit from a System z interface. If you plan to use the System z power control feature, be sure that you order the System z power control feature. This feature comes with four power control cables.

When you use this feature, you must specify the System z power control setting in the **Power Control Pane** menu, then select the option **Zseries Power Mode** in HMC WUI.

In a System z environment, the host must have the Power Sequence Controller (PSC) feature installed to have the ability to turn on and off specific control units, such as the DS8000. The control unit is controlled by the host through the power control cable. The power control cable comes with a standard length of 31 meters, so be sure to consider the physical distance between the host and DS8000.

8.3.7 Storage Area Network connection

The DS8000 can be attached to a SAN environment through its Fibre Channel ports. SANs provide the capability to interconnect open systems hosts, S/390 and System z hosts, and other storage systems.

A SAN allows your single Fibre Channel host ports to have physical access to multiple Fibre Channel ports on the storage unit. You might need to implement zoning to limit the access (and provide access security) of host ports to your storage ports. Note that shared access to a storage unit Fibre Channel port might come from hosts that support a combination of bus adapter types and operating systems.

8.3.8 Tivoli Key Lifecycle Manager server for encryption

If the DS8000 is configured with FDE drives and enabled for encryption, two isolated Tivoli Key Lifecycle Manager (TKLM) servers are also required.

IBM System Storage DS8000 series offers IBM Tivoli Key Lifecycle Manager Server with hardware feature code #1760. This feature fulfills the encryption environment isolated key server requirement, and supports IBM self-encrypting tape and disk products. A Tivoli Key Lifecycle Manager license must be acquired for use of the Tivoli Key Lifecycle Manager software, which must be ordered separately from the IBM Tivoli Key Lifecycle Manager isolated server hardware.

IBM Tivoli Productivity Center is a suite of products available for DS8000 management. These products are designed to provide centralized, automated, and simplified management of complex and heterogeneous storage environments. IBM Tivoli Productivity Center Basic Edition or Standard Edition is required for remote GUI-based management of DS8000 systems. Other Tivoli products provide additional performance, data, and replication management capabilities.

The isolated TKLM server hardware is equivalent to that of the SSPC 2805 MC5 model.

Tip: No other hardware or software is allowed on this server. An isolated server must use an internal disk for all files necessary to boot and have the TKLM key server become operational.

Table 8-8 lists the general hardware requirements.

Table 8-8 TKLM hardware requirements

System components	Minimum values	Suggested values
System memory (RAM)	4 GB	4 GB
Processor speed	▶ For Linux and Windows systems: 2.66 GHz single processor ▶ For AIX and Sun Solaris systems: 1.5 GHz (2-way)	▶ For Linux and Windows systems: 3.0 GHz dual processors ▶ For AIX and Sun Solaris systems: 1.5 GHz (4-way)
Disk space free for product and prerequisite products, such as DB2 Database and keystore files	15 GB	30 GB

Operating system requirement and software prerequisites

Table 8-9 lists the operating systems requirements for installation.

Table 8-9 TKLM software requirements

Operating system	Patch and maintenance level at time of initial publication
AIX Version 5.3 64-bit, and Version 6.1	For Version 5.3, use Technology Level 5300-04 and Service Pack 5300-04-02
Sun Server Solaris 10 (SPARC 64-bit) Tivoli Key Lifecycle Manager runs in a 32-bit JVM.	None
Windows Server 2003 R2 (32-bit Intel)	None
Red Hat Enterprise Linux AS Version 4.0 on x86 32-bit	None
SUSE Linux Enterprise Server Version 9 on x86 (32-bit) and Version 10 on x86 (32-bit)	None

On Linux platforms, Tivoli Key Lifecycle Manager requires the following package:

compat-libstdc++-33-3.2.3-61 or higher

On Red Hat systems, to determine if you have the package, run the following command:

```
rpm -qa | grep -i "libstdc"
```

For more information regarding the required TKLM server and other requirements and guidelines, refer to *IBM System Storage DS8700 Disk Encryption Implementation and Usage Guidelines*, REDP-4500. For supported products and platform, visit the following URL:

<http://www.ibm.com/support/docview.wss?uid=swg21386446>

For further information about Tivoli Key Lifecycle Manager product and features, visit the following URL:

http://publib.boulder.ibm.com/infocenter/tivihelp/v2r1/index.jsp?topic=/com.ibm.tk1m.doc_2.0/welcome.htm

Encryption planning

Encryption planning is a customer responsibility. There are three major planning components to the implementation of an encryption environment. Review all planning requirements and include them in your installation considerations.

- ▶ Key Server Planning

Introductory information, including required and optional features, can be found in the *IBM System Storage DS8700 and DS8800 Introduction and Planning Guide*, GC27-2297.

- ▶ Tivoli Key Lifecycle Manager Planning

The DS8000 series supports IBM Tivoli Key Lifecycle Manager V1.0 and V2.0.

Isolated key servers ordered with feature code #1760 will have a Linux operating system and TKLM software preinstalled. Customers will need to acquire a TKLM license for use of the TKLM software, which is ordered separately from the stand-alone server hardware.

- ▶ Full Disk Encryption Activation Review Planning

IBM Full Disk Encryption offerings must be activated prior to use. This activation is part of the installation and configuration steps required for use of the technology. This installation and activation review is performed by the IBM Systems and Technology Lab Services group.

TKLM connectivity and routing information

To connect the Tivoli Key Lifecycle Manager to your network, you need to provide the following settings to your IBM service representative:

- ▶ SSPC or TPC server network IDs, host names, and domain name
- ▶ Domain Name Server (DNS) settings (if you plan to use DNS to resolve network names)

There are two network ports that need to be opened on a firewall to allow DS8000 connection and have an administration management interface to the TKLM server. These ports are defined by the TKLM administrator.

8.3.9 Lightweight Directory Access Protocol server for single sign-on

A Lightweight Directory Access Protocol (LDAP) server can be used to provide directory services to the DS8000 through the SSPC TIP (Tivoli Integrated Portal). This can enable a single sign-on interface to all DS8000s in the client environment.

Typically, there is one LDAP server installed in the client environment to provide directory services. For details, refer to *IBM System Storage DS8000: LDAP Authentication*, REDP-4505.

LDAP connectivity and routing information

To connect the LDAP server to the System Storage Productivity Center (SSPC), you need to provide the following settings to your IBM service representative:

- ▶ LDAP network IDs, and host names domain name and port
- ▶ User ID and password of the LDAP server

If the LDAP server is isolated from the SSPC by a firewall, the LDAP port (verified during the TKLM installation) needs to be opened in that firewall. There might also be a firewall between the SSPC and the DS8000 that needs to be opened to allow LDAP traffic between them.

8.4 Remote mirror and copy connectivity

The DS8000 uses the high speed Fibre Channel protocol (FCP) for Remote Mirror and Copy connectivity.

Make sure that you have a sufficient number of FCP paths assigned for your remote mirroring between your source and target sites to address performance and redundancy issues. When you plan to use both Metro Mirror and Global Copy modes between a pair of storage units, use separate logical and physical paths for the Metro Mirror, and another set of logical and physical paths for the Global Copy.

Plan the distance between the primary and secondary storage units to properly acquire the necessary length of fiber optic cables that you need. If needed, your Copy Services solution can use hardware such as channel extenders or dense wavelength division multiplexing (DWDM).

For detailed information, refer to *IBM System Storage DS8000: Copy Services for Open Systems*, SG24-6788 and *IBM System Storage DS8000: Copy Services for IBM System z*, SG24-6787.

8.5 Disk capacity considerations

The effective capacity of the DS8000 is determined by several factors:

- ▶ The spares configuration
- ▶ The size of the installed disk drives
- ▶ The selected RAID configuration: RAID 5, RAID 6, or RAID 10, in two sparing combinations
- ▶ The storage type: Fixed Block (FB) or Count Key Data (CKD)

8.5.1 Disk sparing

On internal storage, RAID arrays automatically attempt to recover from a DDM failure by rebuilding the data for the failed DDM on a spare DDM. In order for sparing to occur, a DDM with a disk capacity equal to or greater than failed disk capacity must be available on the same device adapter pair. After sparing is initiated, the spare and the failing DDM are swapped between their respective array sites such that the spare DDM becomes part of the array site associated with the array at the failed DDM. The failing DDM becomes an failed spare DDM in the array site that the spare came from.

The DS8000 assigns spare disks automatically. The first four array sites (a set of eight disk drives) on a Device Adapter (DA) pair will normally each contribute one spare to the DA pair. A minimum of one spare is created for each array site defined until the following conditions are met:

- ▶ A minimum of four spares per DA pair
- ▶ A minimum of four spares of the largest capacity array site on the DA pair
- ▶ A minimum of two spares of capacity and RPM greater than or equal to the fastest array site of any given capacity on the DA pair

The DDM sparing policies support the over-configuration of spares. This might be useful for certain installations, because it allows the repair of some DDM failures to be deferred until a later repair action is required. Refer to *IBM System Storage DS8700 and DS8800*

Introduction and Planning Guide, GC27-2297 and 4.6.8, “Spare creation” on page 94 for more details about the DS8000 sparing concepts.

8.5.2 Disk capacity

The DS8000 operates in either a RAID 5, RAID 6, or RAID 10 configuration. The following RAID configurations are possible:

- ▶ 6+P RAID 5 configuration: The array consists of six data drives and one parity drive. The remaining drive on the array site is used as a spare.
- ▶ 7+P RAID 5 configuration: The array consists of seven data drives and one parity drive.
- ▶ 5+P+Q RAID 6 configuration: The array consists of five data drives and two parity drives. The remaining drive on the array site is used as a spare.
- ▶ 6+P+Q RAID 6 configuration: The array consists of six data drives and two parity drives.
- ▶ 3+3 RAID 10 configuration: The array consists of three data drives that are mirrored to three copy drives. Two drives on the array site are used as spares.
- ▶ 4+4 RAID 10 configuration: The array consists of four data drives that are mirrored to four copy drives.

Table 8-10 for DS8700 and Table 8-11 on page 221 for DS8800 can help you plan the capacity of your DS8000 system. They show the effective capacity of one rank in the various possible configurations. A disk drive set contains 16 drives, which form two array sites. Hard Disk Drive capacity is added in increments of one disk drive set. Solid State Disk drive capacity can be added in increments of a half disk drive set (eight drives). The capacities in the table are expressed in decimal gigabytes and as the number of extents.

Table 8-10 DS8700 Disk drive set capacity for open systems and System z environments

Disk size/ Rank type	Effective capacity of one rank in decimal GB (Number of extents)					
	Rank of RAID 10 arrays		Rank of RAID 6 arrays		Rank of RAID 5 arrays	
	3 + 3	4 + 4	5 + P + Q	6 + P + Q	6 + P	7 + P
73 GB/ FB	207.23 (193)	277.03 (258)	338.27 (315)	407.94 (380)	414.46 (388)	483.18 (452)
73 GB/ CKD	204.34 (216)	273.39 (289)	333.94 (353)	402.95 (425)	410.57 (434)	479.61 (507)
146 GB/ FB	416.61 (388)	557.27 (519)	680.75 (634)	819.27 (763)	836.44 (779)	976.03 (909)
146 GB/ CKD	411.51 (435)	549.63 (581)	671.66 (710)	808.83 (855)	825.86 (873)	963.03 (1018)
300 GB/ FB	848.26 (790)	1131.72 (1054)	1381.90 (1287)	1663.24 (1549)	1698.66 (1582)	1979.98 (1844)
300 GB/ CKD	836.27 (884)	1116.28 (1180)	1364.13 (1442)	1641.32 (1735)	1675.38 (1771)	1954.45 (2066)
450 GB/ FB	1273.46 (1186)	1699.73 (1583)	2074.47 (1932)	2496.45 (2325)	2549.06 (2374)	2972.12 (2768)
450 GB/ CKD	1256.29 (1328)	1676.31 (1772)	2048.09 (2165)	2464.32 (2605)	2515.41 (2659)	2933.55 (3101)

600 GB/ FB	1726.58 (1608)	2304.25 (2146)	2812.13 (2619)	3384.43 (3152)	3456.38 (3219)	4028.68 (3752)
600 GB/ CKD	1703.76 (1801)	2273.25 (2403)	2777.47 (2936)	3341.29 (3532)	3410.35 (3605)	3977.01 (4204)
600 GB (SSD) / FB	See Note 2	See Note 2	See Note 2	See Note 2	3456.38 (3219)	4028.68 (3752)
600 GB (SSD) / CKD	See Note 2	See Note 2	See Note 2	See Note 2	3410.35 (3605)	3977.01 (4204)
1 TB (SATA)/ FB	2622.08 (2442)	3498.26 (3258)	4269.19 (3976)	5136.80 (4784)	See Note 1	See Note 1
1 TB (SATA)/ CKD	2587.31 (2735)	3450.05 (3647)	4215.38 (4456)	5071.50 (5361)	See Note 1	See Note 1
2 TB (SATA)/ FB	5247.38 (4887)	7000.80 (6520)	8541.62 (7955)	10278.93 (9573)	See Note 1	See Note 1
2 TB (SATA)/ CKD	5177.49 (5473)	6904.89 (7299)	8434.59 (8916)	10146.86 (10726)	See Note 1	See Note 1

Table 8-11 DS8800 Disk drive set capacity for open systems and System z environments

Disk size/ Rank type	Effective capacity of one rank in decimal GB (Number of extents)					
	Rank of RAID 10 arrays		Rank of RAID 6 arrays		Rank of RAID 5 arrays	
	3 + 3	4 + 4	5 + P + Q	6 + P + Q	6 + P	7 + P
146 GB / FB	413.39 (385)	551.90 (514)	674.31 (628)	811.75 (756)	828.93 (772)	966.37 (900)
146 GB / CKD	364.89 (431)	544.89 (576)	665.99 (704)	801.27 (847)	818.30 (865)	954.52 (1009)
450 GB / FB	1275.60 (1188)	1702.95 (1586)	2077.69 (1935)	2500.74 (2329)	2553.35 (2378)	2977.48 (2773)
450 GB / CKD	1259.13 (1331)	1679.15 (1775)	2051.87 (2169)	2468.11 (2609)	2520.14 (2664)	2938.28 (3106)
600 GB / FB	1701.82 (1585)	2270.88 (2115)	2771.22 (2581)	3335.99 (3107)	3406.85 (3173)	3970.54 (3698)
600 GB / CKD	1679.15 (1775)	2240.13 (2368)	2736.78 (2893)	3293.03 (3481)	3361.14 (3553)	3919.28 (4143)
300 GB (SSD)/ FB	See Note 2	See Note 2	See Note 2	See Note 2	1690.07 (1574)	1970.31 (1835)
300 GB (SSD)/ CKD	See Note 2	See Note 2	See Note 2	See Note 2	1667.80 (1763)	1944.98 (2056)

Tips:

- ▶ Effective capacities are in decimal gigabytes (GB). One GB is 1,000,000,000 bytes.
- ▶ Although disk drive sets contain 16 drives, arrays use only eight drives. The effective capacity assumes that you have two arrays for each disk drive set.
- ▶ RAID 5 implementations are not compatible with the use of 1 TB and 2 TB SATA drives.
- ▶ RAID 10 and RAID 6 implementations are not compatible with the use of 300 GB SSD drives for DS8800 and 600 GB SSD for DS8700.

An updated version of Capacity Magic (see “Capacity Magic” on page 456) will aid you in determining the raw and net storage capacities, and the numbers for the required extents for each available type of RAID.

8.5.3 DS8000 Solid State Drive (SSD) considerations

Solid-State Drives (SSDs) are a higher performance option compared to hard disk drives (HDDs).

- ▶ For DS8700, SSD drives are currently available in 600 GB capacity.
- ▶ For DS8800, SSD disks are available in 300 GB capacity.

All disks installed in a storage enclosure pair must be of the same capacity and speed. Feature conversions are available to exchange existing disk drive sets when purchasing new disk drive sets with higher capacity or speed.

SSD drives can be ordered and installed in eight drive install groups (half drive sets) or 16 drive install groups (full drive sets). A half drive set (8) is always upgraded to a full drive set (16) when SSD capacity is added. A frame can contain at most one SSD half drive set.

Tip: A eight drive install increment means that the SSD rank added is assigned to only one DS8000 server (CEC).

To achieve optimal price to performance ratio in DS8000, SSD drives have limitations and considerations that differ from HDDs.

Limitations

- ▶ Drives of different capacity and speed cannot be intermixed in a storage enclosure pair.
- ▶ A DS8700 system is limited to 32 SSDs per DA pair. The maximum number of SSDs in a DS8700 system is 256 drives spread over eight DA pairs.
- ▶ A DS8800 system is limited to 48 SSDs per DA pair. The maximum number of SSDs in a DS8800 system is 384 drives spread over eight DA pairs.
- ▶ RAID 5 is the only supported implementation for SSDs (RAID 6 and RAID 10 implementations are not supported). SSD drives follow normal sparing rules. The array configuration is 6+P+S or 7+P.
- ▶ SSDs are not supported in a DS8800 Business Class machine with feature code 4211 (16 GB memory).

SSD Placement

- ▶ SSD drive sets have a default location when a new machine is ordered and configured.
- ▶ SSD drives are installed in default locations from manufacturing, which is the first storage enclosure pair on each device adapter pair. This is done to spread the SSDs over as many DA pairs as possible to achieve optimal price to performance ratio.
- ▶ The default locations for DS8700 are split among eight DA pairs (if installed) in the first three frames: Two in the first frame, four in the second, and two in the third frame.
- ▶ For DS8700, an SSD feature (drive set or half drive set) is installed in the first disk enclosure pair of an available DA pair. A second SSD feature can be installed on the same DA pair only after the system contains at least eight SSD features, that is, after each of the eight DA pairs contains at least one SSD drive set. This means that the system can have more than 16 SSD drives in a DA pair only if the system has two or more frames. The second SSD feature on the DA pair must be a full drive set.
- ▶ The default locations for DS8800 are split among eight DA pairs (if installed) in the first two frames: Four in the first frame and four in the second frame.
- ▶ Adding SSDs to an existing configuration, to the fourth and fifth frame for DS8700, or the third frame for DS8800 requires a request for price quotation (RPQ). This is to ensure the limitation of 32 SSDs per DA pair for DS8700 and 48 SSDs per DA pair for DS8800 is not exceeded.

Performance consideration

- ▶ Limiting the number of SSD drives to 16 per DA pair provides the optimal price to performance ratio. The DS8800 is the preferable platform for Solid-State Drives due to its better overall system performance.

8.6 Planning for growth

The DS8000 storage unit is a highly scalable storage solution. Features such as total storage capacity, cache size, and host adapters can be easily increased by physically adding the necessary hardware and by changing the needed licensed feature keys (as ordered).

Planning for future growth normally suggests an increase in physical requirements in your installation area, including floor loading, electrical power, and environmental cooling.

A key feature that you can order for your dynamic storage requirement is the Standby Capacity on Demand (CoD). This offering is designed to provide you with the ability to tap into additional storage, and is particularly attractive if you have rapid or unpredictable growth, or if you simply want the knowledge that the extra storage will be there when you need it. Standby CoD allows you to access the extra storage capacity when you need it through a nondisruptive activity. For more information about Capacity on Demand, see 18.2, “Using Capacity on Demand (CoD)” on page 449.



DS8000 HMC planning and setup

This chapter discusses the planning activities needed for the setup of the required DS8000 Hardware Management Console (HMC). This chapter covers the following topics:

- ▶ Hardware Management Console overview
- ▶ Hardware Management Console software
- ▶ HMC activities
- ▶ HMC and IPv6
- ▶ HMC user management
- ▶ External HMC
- ▶ Configuration flow

9.1 Hardware Management Console overview

The HMC is the focal point for DS8000 management with multiple functions, including:

- ▶ DS8000 power control
- ▶ Storage provisioning
- ▶ Advanced Copy Services management
- ▶ Interface for onsite service personnel
- ▶ Call Home and problem management
- ▶ Remote support
- ▶ Connection to TKLM for encryption functions

The HMC is the point where the DS8000 is connected to the client network. It provides the services that the client needs to configure and manage the storage. It also provides the interface where service personnel will perform diagnostics and repair actions. The HMC is the contact point for remote support, both modem and VPN.

9.1.1 Storage Hardware Management Console hardware

The HMC consists of a mobile workstation:

- ▶ Lenovo Thinkpad W500 for DS8700
- ▶ Lenovo Thinkpad T510 for DS8800

The mobile workstation is equipped with adapters for modem and 10/100/1000 Mb Ethernet. The internal HMC included with every Base frame is mounted in a pull-out tray for convenience and security. A second, redundant mobile workstation HMC is orderable and should be used in environments that use TKLM encryption management and Advanced Copy Services functions. The second HMC is external to the DS8000 frame(s). See 9.6, “External HMC” on page 245 for more information regarding adding an external HMC. Figure 9-1 on page 227 for DS8700 and Figure 9-2 on page 227 for DS8800 show the mobile computer HMC and the network connections. This drawing also applies to an external HMC.

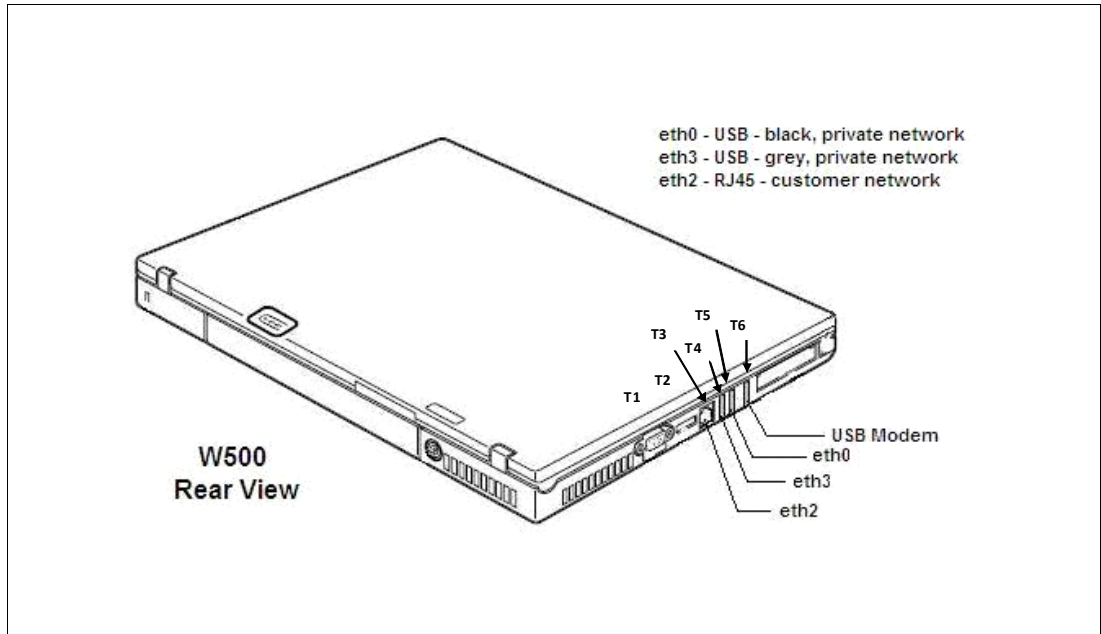


Figure 9-1 DS8700 mobile workstation HMC and connections

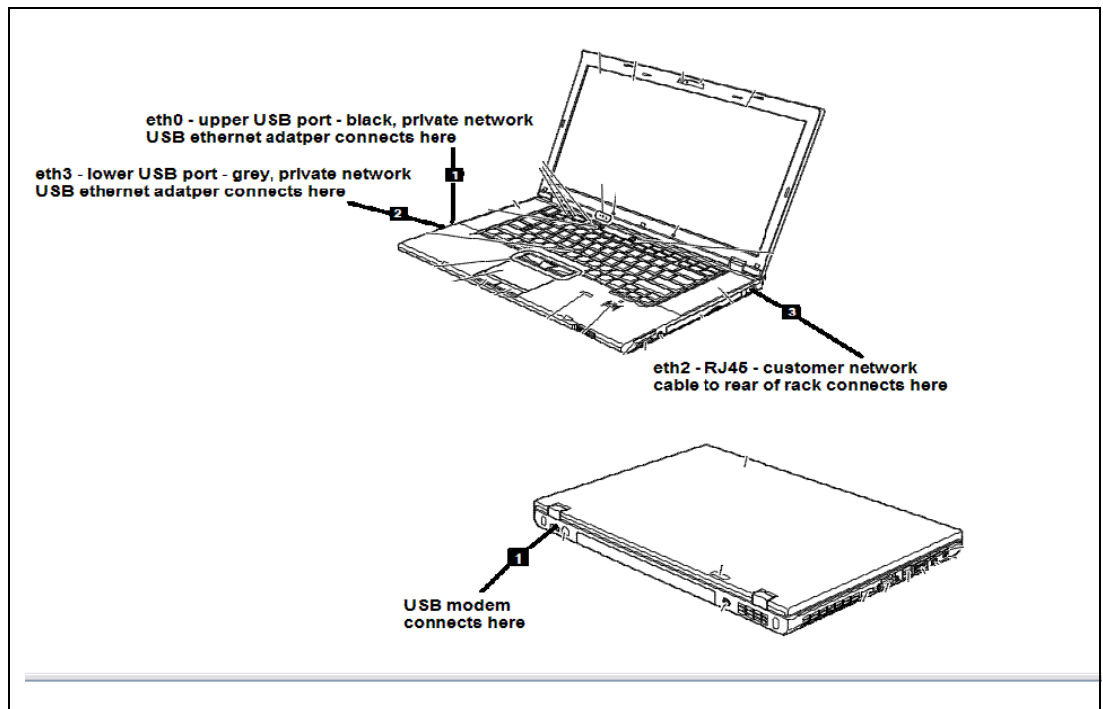


Figure 9-2 DS8800 mobile workstation HMC and connections

9.1.2 Private Ethernet networks

The HMC is connected to the storage facility by way of redundant private Ethernet networks. Figure 9-3 shows the pair of Ethernet switches internal to the DS8000.

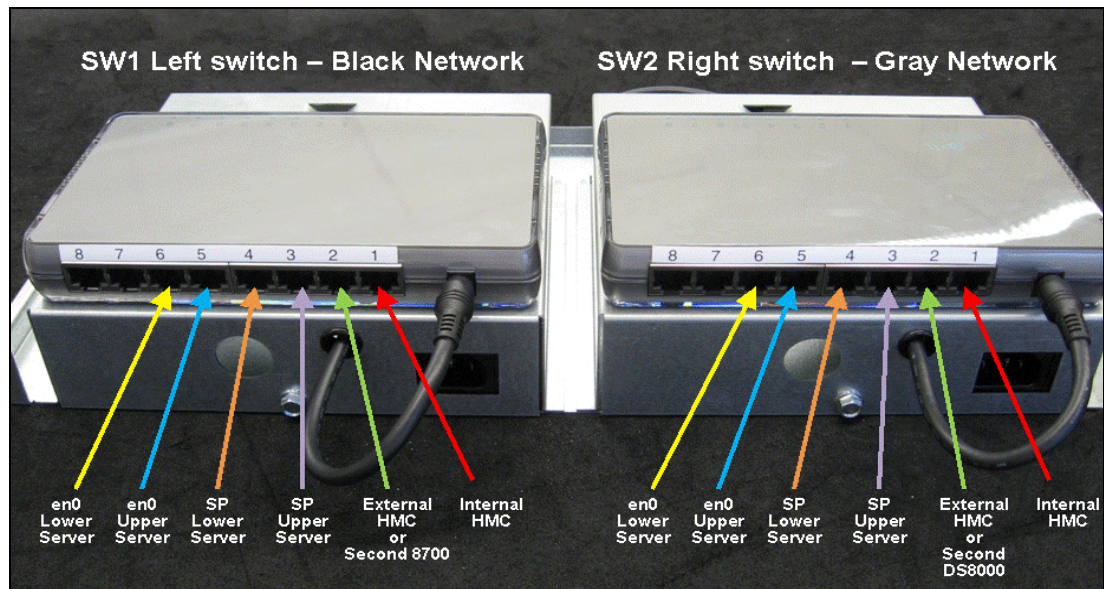


Figure 9-3 Rear view of DS8000 Ethernet switches

The HMC's public Ethernet port, shown as eth2 in Figure 9-2 on page 227, is where the client connects to its network. The HMC's private Ethernet ports, eth0 and eth3, are configured into port 1 of each Ethernet switch to form the private DS8000 network. To interconnect two DS8000s, FC1190 provides a pair of 31 m Ethernet cables to connect each switch in the second system into port 2 of switches in the first frame. Depending on the machine configuration, one or more ports might be unused on each switch.

Important: The internal Ethernet switches pictured in Figure 9-3 are for the DS8000 private network only. No client network connection should ever be made to the internal switches. Client networks are connected to DS8000 internal patch panel

9.2 Hardware Management Console software

The Linux-based HMC includes two application servers that run within a WebSphere® environment: DS Storage Management server and Enterprise Storage Server Network Interface server:

- ▶ DS Storage Management server

The DS Storage Management server is the logical server that communicates with the outside world to perform DS8000-specific tasks.

- ▶ Enterprise Storage Server Network Interface server (ESSNI)

ESSNI is the logical server that communicates with the DS Storage Management server and interacts with the two CECs of the DS8000.

The DS8000 HMC provides several management interfaces. These include:

- ▶ DS Storage Manager graphical user interface (GUI)
- ▶ DS Command-Line Interface (DS CLI)
- ▶ DS Open Application Programming Interface (DS Open API)
- ▶ Web-based user interface (WebUI)

The GUI and the CLI are comprehensive, easy-to-use interfaces for a storage administrator to perform DS8000 management tasks to provision the storage arrays, manage application users, and change HMC options. The two can be used interchangeably, depending on the particular task.

The DS Open API provides an interface for external storage management programs, such as Tivoli Productivity Center (TPC), to communicate with the DS8000. It channels traffic through the IBM System Storage Common Information Model (CIM) agent, a middleware application that provides a CIM-compliant interface.

Older DS8000 family products used a service interface called WebSM. The DS8000 uses a newer, faster interface called WebUI that can be used remotely over a VPN by support personnel to check the health status and perform service tasks.

9.2.1 DS Storage Manager GUI

DS Storage Manager can be accessed using the Tivoli Storage Productivity Center (TPC) Basic Edition (TPC Basic Edition is the minimum software requirement, and can be installed on a customer provided server) or using TPC Element Manager of the SSPC from any network-connected workstation with a supported browser. Login procedures are explained in the following sections.

IBM Tivoli Storage Productivity Center login to DS Storage Manager GUI

The DS Storage Manager graphical user interface (GUI) can be launched using the TPC Element Manager installed on customer server from any supported, network-connected workstation.

IBM Tivoli Storage Productivity Center simplifies storage management by providing the following benefits:

- ▶ Centralizing the management of heterogeneous storage network resources with IBM storage management software
- ▶ Providing greater synergy between storage management software and IBM storage devices
- ▶ Reducing the number of servers that are required to manage your software infrastructure
- ▶ Migrating from basic device management to storage management applications that provide higher-level functions

SSPC login to DS Storage Manager GUI

The DS Storage Manager graphical user interface (GUI) can be launched using the TPC Element Manager of the SSPC from any supported network-connected workstation.

To access the DS Storage Manager GUI through the SSPC, open a new browser window or tab and type the following address:

`http://<SSPC_ipaddress>:9550/ITSRM/app/welcome.html`

A more thorough description of setting up and logging into SSPC can be found in 12.2.2, “Configuring SSPC for DS8000 remote GUI access” on page 284.

9.2.2 Command-line interface

The DS Command-Line Interface (DS CLI), which must be executed in the command environment of an external workstation, is a second option to communicate with the HMC. The DS CLI might be a good choice for configuration tasks when there are many updates to be done. This avoids the web page load time for each window in the DS Storage Manager GUI.

For additional detail regarding the DSCLI use and configuration, see Chapter 14, “Configuration with the DS Command-Line Interface” on page 363 and *IBM Systems Storage DS8000 Series: Command-Line Interface User's Guide*, SC26-7916, for a complete list of DS CLI commands.

9.2.3 DS Open Application Programming Interface

Calling DS Open Application Programming Interfaces (DS Open APIs) from within a program is a third option to implement communication with the HMC. Both DS CLI and DS Open API communicate directly with the ESSNI server software running on the HMC.

The Common Information Model (CIM) Agent for the DS8000 is Storage Management Initiative Specification (SMI-S) 1.1-compliant. This agent is used by storage management applications such as Tivoli Productivity Center (TPC), Tivoli Storage Manager, and VSS/VDS. Also, to comply with more open standards, the agent can be accessed by software from third-party vendors, including VERITAS/Symantec, HP/AppIQ, EMC, and many other applications at the SNIA Interoperability Lab. For more information, visit the following URL:

http://www.snia.org/forums/smi/tech_programs/lab_program/

For the DS8000, the CIM agent is preloaded with the HMC code and is started when the HMC boots. An active CIM agent only allows access to the DS8000s managed by the HMC on which it is running. Configuration of the CIM agent must be performed by an IBM Service representative using the DS CIM Command Line Interface (DSCIMCLI).

9.2.4 Web-based user interface

The Web User Interface (WebUI) is a Internet-browser-based interface used for remote access to system utilities. If a VPN connection has been set up, WebUI can be used by support personnel for DS8000 diagnostic tasks, data offloading, and many service actions. The connection is over port 443 over SSL, providing a secure and full interface to utilities running at the HMC.

Important: Use a secure Virtual Private Network (VPN) or Business-to-Business VPN, that allows service personnel to quickly respond to client needs using the WebUI.

The following procedure can be used to log in to the Hardware Management Console:

1. Log-on at the HMC as shown in Figure 9-4. Click **Log on and launch the Hardware Management Console web application** to open the login window and log in. The default user ID is *customer* and the default password is *cust0mer*.

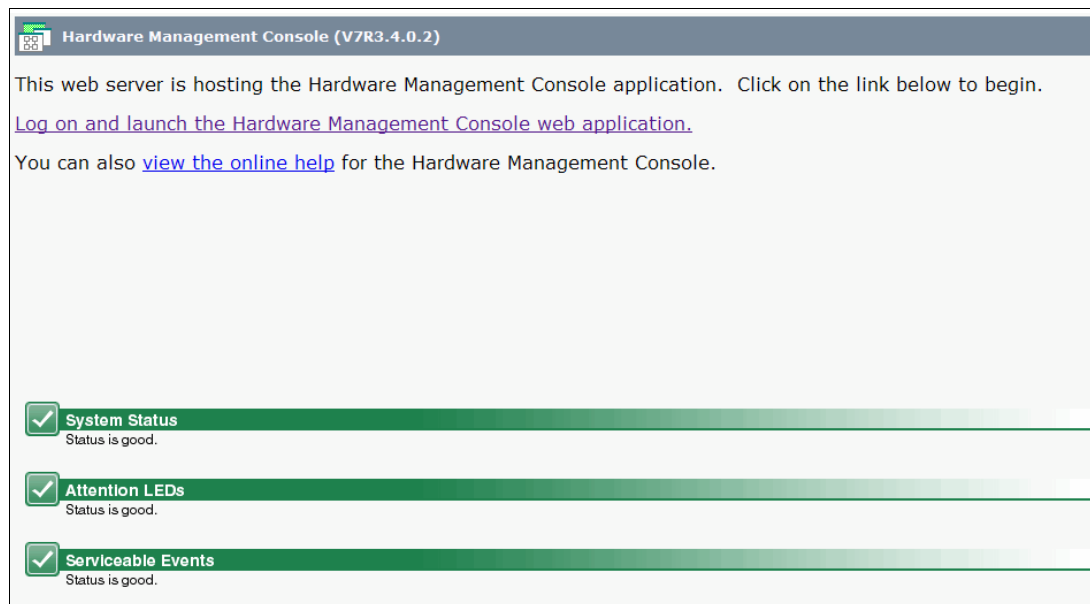


Figure 9-4 Hardware Management Console

2. If you are successfully logged in, you will see the Hardware Management console window, where you can select **Status Overview** to see the status of the DS8000. Other areas of interest are illustrated in Figure 9-5.

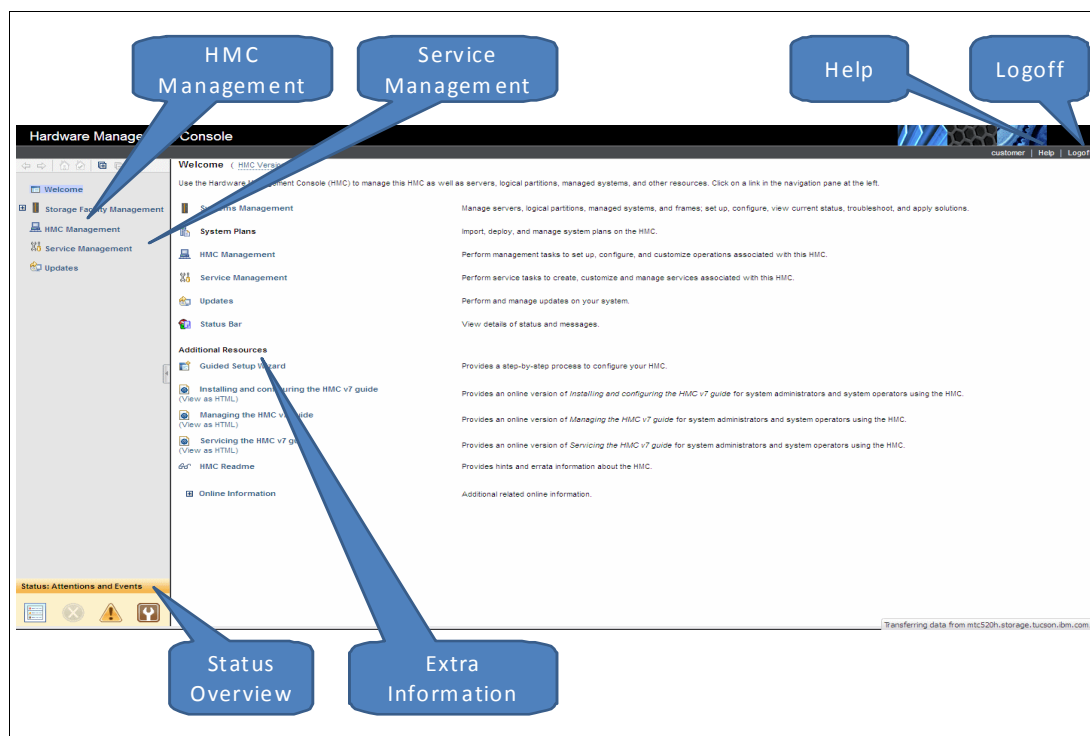


Figure 9-5 WebUI main window

Because the HMC WebUI is mainly a services interface, it will not be covered here. Further information can be obtained through the **Help** menu.

9.3 HMC activities

This section covers planning and maintenance activities for the DS8000 HMC. See Chapter 8, “Physical planning and installation” on page 201 for overall planning information.

9.3.1 HMC planning tasks

The following activities are needed to plan the installation or configuration:

- ▶ The installation activities for the optional external HMC need to be identified as part of the overall project plan and agreed upon with the responsible IBM personnel.
- ▶ A connection to the client network will be needed at the base frame for the internal HMC. Another connection will also be needed at the location of the second, external HMC. The connections should be standard CAT5/6 Ethernet cabling with RJ45 connectors.
- ▶ IP addresses for the internal and external HMCs will be needed. The DS8000 can work with both IPv4 and IPv6 networks. See 9.4, “HMC and IPv6” on page 234 for procedures to configure the DS8000 HMC for IPv6.
- ▶ A phone line will be needed at the base frame for the internal HMC. Another line will also be needed at the location of the second, external HMC. The connections should be standard phone cabling with RJ11 connectors.
- ▶ You can use Tivoli Storage Productivity Center (TPC) Basic Edition in your environment to access the DS GUI. Alternatively you can use the SSPC (machine type 2805-MC5), an integrated hardware and software solution for centralized management of IBM storage products with IBM storage management software. SSPC is described in detail in Chapter 12, “System Storage Productivity Center” on page 279.
- ▶ The web browser to be used on any administration workstation should be a supported one, as mentioned in the *IBM System Storage DS8700 and DS8800 Introduction and Planning Guide*, GC27-2297 or in the Information Center for the DS8000, which can be found at the following URL:

<http://publib.boulder.ibm.com/infocenter/dsichelp/ds8000ic/index.jsp>

A decision should be made as to which web browser should be used. The web browser is the only software that is needed on workstations that will do configuration tasks online using the DS Storage Manager GUI (through the TPC BE or SSPC).

- ▶ The IP addresses of SNMP recipients need to be identified if the client wants the DS8000 HMC to send SNMP traps to a network station.
- ▶ Email accounts need to be identified if the client wants the DS8000 HMC to send email messages for problem conditions.
- ▶ The IP addresses of NTP servers need to be identified if the client wants the DS8000 HMC to utilize Network Time Protocol for time synchronization.
- ▶ When ordering a DS8000, the license and certain optional features need activation as part of the customization of the DS8000. See Chapter 10, “IBM System Storage DS8000 features and license keys” on page 251 for details.

Tip: Applying increased feature activation codes is a concurrent action.

9.3.2 Planning for microcode upgrades

The following activities need to be considered in regard to the microcode upgrades on the DS8000:

- ▶ Microcode changes

IBM might release changes to the DS8000 series Licensed Machine Code. IBM plans to make most DS8000 series Licensed Machine Code changes available for HMC download by FTP from the TESTCASE site. Note that not all Licensed Machine Code changes may be available through the TESTCASE.

- ▶ Microcode install

An IBM service representative can install the changes. Check whether the new microcode requires new levels of DS Storage Manager, DS CLI, and DS Open API, and plan on upgrading them on the relevant workstations if necessary.

- ▶ Host prerequisites

When planning for initial installation or for microcode updates, make sure that all prerequisites for the hosts are identified correctly. Sometimes a new level is required for the SDD as well. DS8000 interoperability information can be found at the IBM System Storage Interoperability Center (SSIC) at the following website:

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

To prepare for the download of drivers, refer to the HBA Support Matrix referenced in the Interoperability Matrix and make sure that drivers are downloaded from the IBM Internet site. This is to make sure that drivers are used with the settings corresponding to the DS8000, not some other IBM storage subsystem.

Important: The Interoperability Center reflects information regarding the latest supported code levels. This does not necessarily mean that former levels of HBA firmware or drivers are no longer supported. If in doubt about any supported levels, contact your IBM representative.

- ▶ Maintenance windows

Normally the microcode update of the DS8000 is a nondisruptive action. However, any prerequisites identified for the hosts (for example, patches, new maintenance levels, or new drivers) could make it necessary to schedule a maintenance window. The host environments can then be upgraded to the level needed in parallel to the microcode update of the DS8000 taking place.

For more information about microcode upgrades, see Chapter 15, “Licensed machine code” on page 401.

9.3.3 Time synchronization

For proper error analysis, it is important to have the date and time information synchronized as much as possible on all components in the DS8000 environment. This includes the DS8000 HMCs, the DS Storage Manager, and DS CLI workstations.

With the DS8000, the HMC has the ability to utilize the Network Time Protocol (NTP) service. Customers can specify NTP servers on their internal network to provide the time to the HMC. It is a client responsibility to ensure that the NTP servers are working, stable, and accurate. An IBM service representative will enable the HMC to use NTP servers, ideally at the time of the initial DS8000 installation.

9.3.4 Monitoring DS8000 with the HMC

A client can receive notifications from the HMC through SNMP traps and email messages. Notifications contain information about your storage complex, such as open serviceable events. You can choose one or both notification methods:

- Simple Network Management Protocol (SNMP) traps

For monitoring purposes, the DS8000 uses SNMP traps. An SNMP trap can be sent to a server in the client's environment, perhaps with System Management Software, which handles the trap based on the MIB delivered with the DS8000 software. A MIB containing all traps can be used for integration purposes into System Management Software. The supported traps are described in more detail in the documentation that comes with the microcode on the CDs provided by the IBM service representative. The IP address to which the traps should be sent needs to be configured during initial installation of the DS8000. For more information about the DS8000 and SNMP, see Chapter 16, "Monitoring with Simple Network Management Protocol" on page 407.

- Email

When you choose to enable email notifications, email messages are sent to all the addresses that are defined on the HMC whenever the storage complex encounters a serviceable event or must alert you to other information.

During the planning process, create a list of who needs to be notified.

SIM notification is only applicable for System z servers. It allows you to receive a notification on the system console in case of a serviceable event. SNMP and email are the only notification options for the DS8000.

9.3.5 Call Home and remote support

The HMC uses both outbound (call home) and inbound (remote service) support.

Call home is the capability of the HMC to contact IBM support center to report a serviceable event. Remote service is the capability of IBM support representatives to connect to the HMC to perform service tasks remotely. If allowed to do so by the setup of the client's environment, an IBM service support representative could connect to the HMC to perform detailed problem analysis. The IBM service support representative can view error logs and problem logs, and initiate trace or dump retrievals.

Remote support can be configured for dial-up connection through a modem or high-speed virtual private network (VPN) Internet connection. Setup of the remote support environment is done by the IBM service representative during initial installation. For more complete information, see Chapter 17, "Remote support" on page 425.

9.4 HMC and IPv6

The DS8000 Hardware Management Console (HMC) can be configured for an IPv6 client network. Note that IPv4 is also still supported.

Configuring the HMC in an IPv6 environment

Usually, the configuration will be done by the IBM service representative during the DS8000 initial installation. See 8.3.3, "System Storage Productivity Center and network access" on page 213 for a thorough discussion about the formatting of IPv6 addresses and subnet masks.

In the remainder of this section, we illustrate the steps required to configure the DS8000 HMC eth2 port for IPv6:

1. Launch and log in to WebUI. Refer to 9.2.4, “Web-based user interface” on page 230 for the procedure.
2. In the HMC welcome window, select **HMC Management** as shown in Figure 9-6.

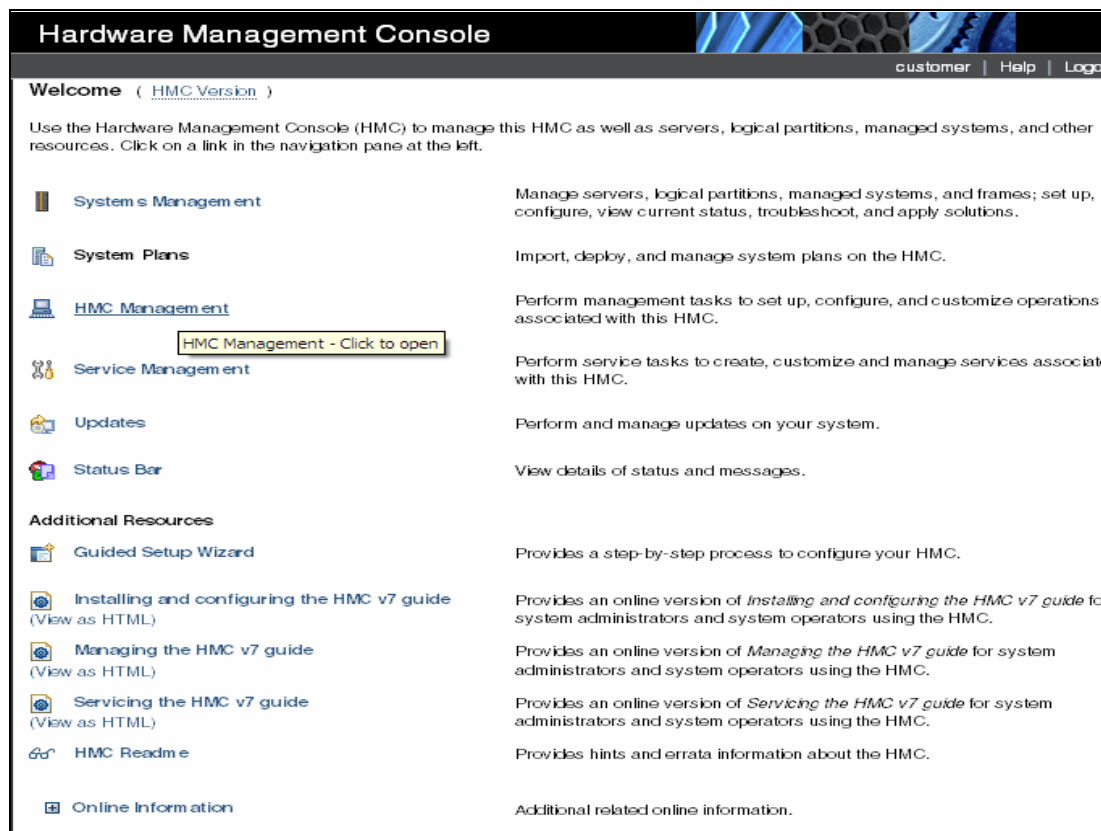


Figure 9-6 WebUI welcome window

3. In the HMC Management window, select **Change Network Settings** as shown in Figure 9-7.

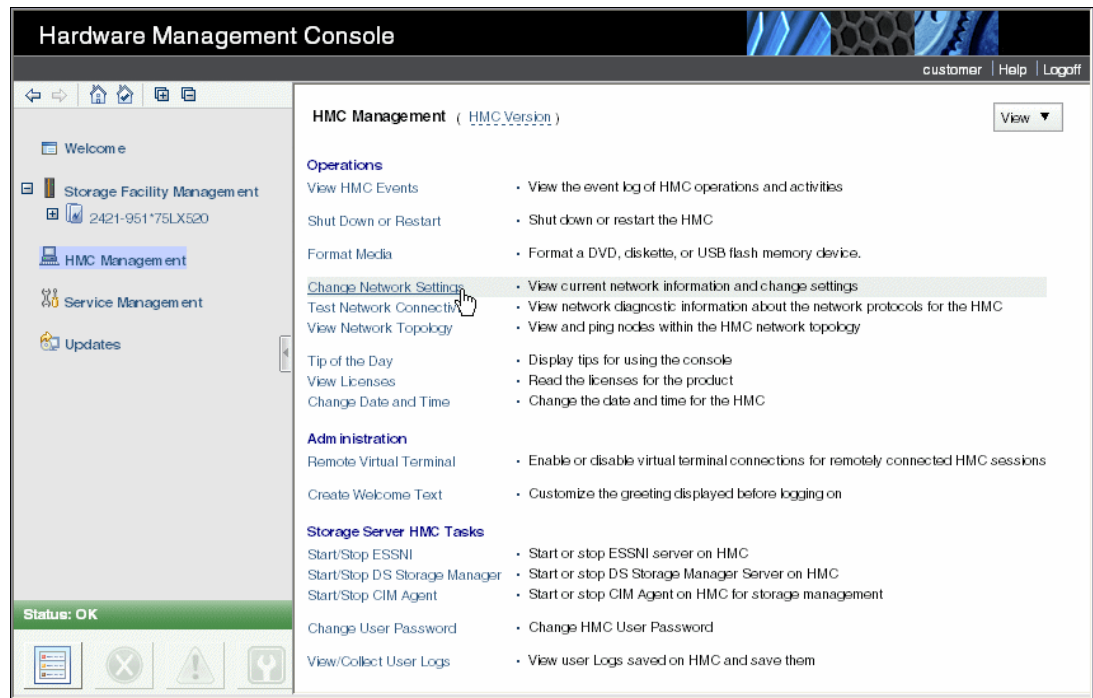


Figure 9-7 WebUI HMC management window

4. Click the LAN Adapters tab.
5. Only eth2 is shown. The private network ports are not editable. Click the **Details** button.
6. Click the IPv6 Settings tab.
7. Click the **Add** button to add a static IP address to this adapter. Figure 9-8 shows the LAN Adapter Details window where you can configure the IPv6 values.

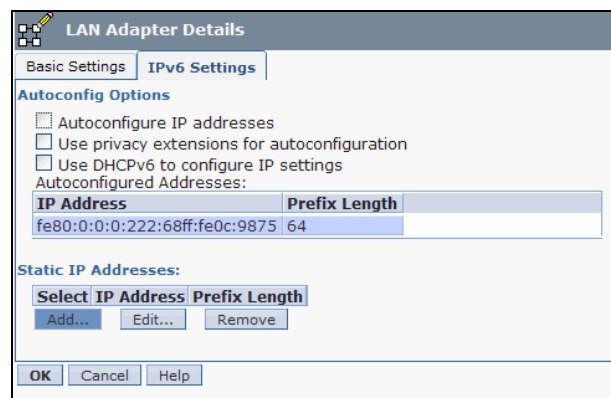


Figure 9-8 WebUI IPv6 settings window

9.5 HMC user management

User management can be performed using the DS CLI or the DS GUI. An administrator user ID is preconfigured during the installation of the DS8000, using the following defaults:

User ID	admin
Password	admin

The password of the admin user ID will need to be changed before it can be used. The GUI will force you to change the password when you first log in. The DS CLI will allow you to log in but will not allow you to issue any other commands until you have changed the password. As an example, to change the admin user's password to `passw0rd`, use the following DS CLI command:

```
chuser-pw passw0rd admin
```

After you have issued that command, you can issue other commands.

Tip: The DS8000 supports the capability to use a Single Point of Authentication function for the GUI and CLI through a proxy to contact the external repository (for example: LDAP Server). Proxy used is a Tivoli Component (Embedded Security Services a.k.a. Authentication Service). This capability requires a minimum TPC Version 4.1 server.

For detailed information about LDAP based authentication, refer to *IBM System Storage DS8000: LDAP Authentication*, REDP-4505.

User roles

During the planning phase of the project, a worksheet or a script file was established with a list of all people who need access to the DS GUI or DS CLI. Note that a user can be assigned to more than one group. At least one person (user_id) should be assigned to each of the following roles:

- ▶ The *Administrator* (admin) has access to all HMC service methods and all storage image resources, except for encryption functionality. This user authorizes the actions of the *Security Administrator* during the encryption deadlock prevention and resolution process.
- ▶ The *Security Administrator* (secadmin) has access to all encryption functions. secadmin requires an Administrator user to confirm the actions taken during the encryption deadlock prevention and resolution process.
- ▶ The *Physical operator* (op_storage) has access to physical configuration service methods and resources, such as managing storage complex, storage image, Rank, array, and Extent Pool objects.
- ▶ The *Logical operator* (op_volume) has access to all service methods and resources that relate to logical volumes, hosts, host ports, logical subsystems, and Volume Groups, excluding security methods.
- ▶ The *Monitor* group has access to all read-only, nonsecurity HMC service methods, such as **list** and **show** commands.
- ▶ The *Service* group has access to all HMC service methods and resources, such as performing code loads and retrieving problem logs, plus the privileges of the Monitor group, excluding security methods.
- ▶ The *Copy Services operator* has access to all Copy Services methods and resources, plus the privileges of the Monitor group, excluding security methods.

Tip: With the DS8700 and DS8800 microcode Release 6.1, Resource Groups offers an enhanced security capability that supports the hosting of multiple customers with Copy Services requirements and the single customer with requirements to isolate the data of multiple operating systems' environments. For additional detail, refer to *IBM Systems Storage DS8000 Series: Resource Groups*, REDP-4758.

- ▶ *No access* prevents access to any service method or storage image resources. This group is used by an administrator to temporarily deactivate a user ID. By default, this user group is assigned to any user account in the security repository that is not associated with any other user group.

Password policies

Whenever a user is added, a password is entered by the administrator. During the first login, this password must be changed. Password settings include the time period in days after which passwords expire and a number that identifies how many failed logins are allowed. The user ID is deactivated if an invalid password is entered more times than the limit. Only a user with administrator rights can then reset the user ID with a new initial password.

General rule: Do not set the values of `chpass` to 0, as this indicates that passwords never expire and unlimited login attempts are allowed.

If access is denied for the administrator due to the number of invalid login attempts, a procedure can be obtained from your IBM support representative to reset the administrator's password. Some password rules changes were introduced with the Release 6.1 microcode.

Tip: Upgrading an existing storage system to Release 6.1 will NOT change old default user acquired rules. Existing default values are retained to prevent disruption. The user might opt to use the new defaults with the `chpass -reset` command. The command resets all default values to new defaults immediately.

The password for each user account is forced to adhere to the following rules:

- ▶ Passwords must contain one character from at least 2 groups and must be between 8 and 16 characters:
 - Groups now include Alphabetic, Numeric, and Punctuation
 - Old rules required at least 5 alphabetic and 1 numeric character
 - Old rules required first and last characters to be alphabetic
- ▶ Passwords may not contain the user's ID.
- ▶ Initial passwords on new user accounts are expired.
- ▶ Passwords "reset" by an administrator are expired.
- ▶ Users must change expired passwords at next logon.

Additional password security implementations on Release 6.1 include the following:

- ▶ Password rules that are checked when changing passwords
- ▶ Valid character set, embedded user ID, age, length, and history
- ▶ Passwords 'invalidated' by change still usable until next password change
- ▶ Users with 'invalidated' passwords not automatically disconnected from DS8000
- ▶ Password rules that are checked when user logs on:

- Password expiration, locked out, and failed attempts
- Users with passwords that expire, or locked out by admin, while logged on are not automatically disconnected from DS8000

Tip: User names and passwords are case sensitive. If you create a user name called *Anthony*, you cannot log in using the user name *anthony*.

9.5.1 User management using the DS CLI

The exact syntax for any DS CLI command can be found in the *IBM Systems Storage DS8000 Series: Command-Line Interface User's Guide*, SC26-7916. You can also use the DS CLI **help** command to get further assistance.

The commands to manage user IDs using the DS CLI are:

► **mkuser**

This command creates a user account that can be used with both DS CLI and the DS GUI. In Example 9-1, we create a user called MaxRos is in the op_storage group. His temporary password is passw0rd. He will have to use the **chpass** command when he logs in for the first time.

Example 9-1 Using the mkuser command to create a new user

```
dscli> mkuser -pw passw0rd -group op_storage MaxRos
Date/Time: April 25, 2011 11:47:54 AM MST IBM DSCLI Version: 7.6.10.464 DS: -
CMUC00133I mkuser: User MaxRos successfully created.
```

► **rmuser**

This command removes an existing user ID. In Example 9-2, we remove a user called JaneSmith.

Example 9-2 Removing a user

```
dscli> rmuser JaneSmith
Date/Time: April 25, 2011 11:52:43 AM MST IBM DSCLI Version: 7.6.10.464 DS: -
CMUC00135W rmuser: Are you sure you want to delete user JaneSmith? [y/n]:y
CMUC00136I rmuser: User JaneSmith successfully deleted.
```

► **chuser**

This command changes the password or group (or both) of an existing user ID. It is also used to unlock a user ID that has been locked by exceeding the allowable login retry count. The administrator could also use this command to lock a user ID. In Example 9-3, we unlock the user, change the password, and change the group membership for a user called JensW. He must use the **chpass** command when he logs in the next time.

Example 9-3 Changing a user with chuser

```
dscli> chuser -unlock -pw time2change -group op_storage JaneSmith
Date/Time: April 25, 2011 11:56:20 AM MST IBM DSCLI Version: 7.6.10.464 DS: -
CMUC00134I chuser: User JaneSmith successfully modified.
```

► **lsuser**

With this command, a list of all user IDs can be generated. In Example 9-4 on page 240, we can see three users and the admin account.

Example 9-4 Using the lsuser command to list users

```
dsccli> lsuser
Date/Time: April 25, 2011 01:26:33 PM MST IBM DSCLI Version: 7.6.10.464 DS: -
Name      Group      State
=====
JaneSmith op_storage active
MaxRos    op_storage active
admin     admin      active
secadmin  secadmin   active
```

► **showuser**

The account details of a user ID can be displayed with this command. In Example 9-5, we list the details of the user MaxRos.

Example 9-5 Using the showuser command to list user information

```
dsccli> showuser MaxRos
Date/Time: April 25, 2011 01:36:23 PM MST IBM DSCLI Version: 7.6.10.464 DS: -
Name      MaxRos
Group      op_storage
State      active
FailedLogin 0
DaysToExpire 365
Scope      PUBLIC
```

► **managepwfile**

This command creates or adds to an encrypted password file that will be placed onto the local machine. This file can be referred to in a DS CLI profile. This allows you to run scripts without specifying a DS CLI user password in clear text. If manually starting DS CLI, you can also refer to a password file with the **-pwfile** parameter. By default, the file is located in the following locations:

Windows	C:\Documents and Settings\<User>\DSCLI\security.dat
Non-Windows	\$HOME/dsccli/security.dat

In Example 9-6, we manage our password file by adding the user ID SJoseph. The password is now saved in an encrypted file called security.dat.

Example 9-6 Using the managepwfile command

```
dsccli> managepwfile -action add -name MaxRos -pw passw0rd
Date/Time: April 25, 2011 01:38:33 PM MST IBM DSCLI Version: 7.6.10.464 DS: -
CMUC00206I managepwfile: Record 10.0.0.1/MaxRos successfully added to password
file C:\Documents and Settings\Administrator\dsccli\security.dat.
```

► **chpass**

This command lets you change two password policies: password expiration (days) and failed logins allowed. In Example 9-7, we change the expiration to 365 days and five failed login attempts.

Example 9-7 Changing rules using the chpass command

```
dsccli> chpass -expire 365 -fail 5
Date/Time: April 25, 2011 01:46:22 PM MST IBM DSCLI Version: 7.6.10.464 DS: -
CMUC00195I chpass: Security properties successfully set.
```

► **showpass**

This command lists the properties for passwords (Password Expiration days and Failed Logins Allowed). In Example 9-8, we can see that passwords have been set to expire in 90 days and that four login attempts are allowed before a user ID is locked.

Example 9-8 Using the showpass command

```
dscli> showpass
Date/Time: April 25, 2011 01:49:45 PM MST IBM DSCLI Version: 7.6.10.464 DS: -
Password Expiration 365 days
Failed Logins Allowed 5
Password Age 0 days
Minimum Length 6
Password History 4
```

9.5.2 User management using the DS GUI

For GUI based user administration, sign on to the DS GUI. See 12.2.2, “Configuring SSPC for DS8000 remote GUI access” on page 284 for procedures about using the TPC BE or SSPC to launch the DS Storage Manager GUI. Then perform the following steps:

1. From the categories in the left sidebar, select **User Administration** under the section Monitor System as shown in Figure 9-9.

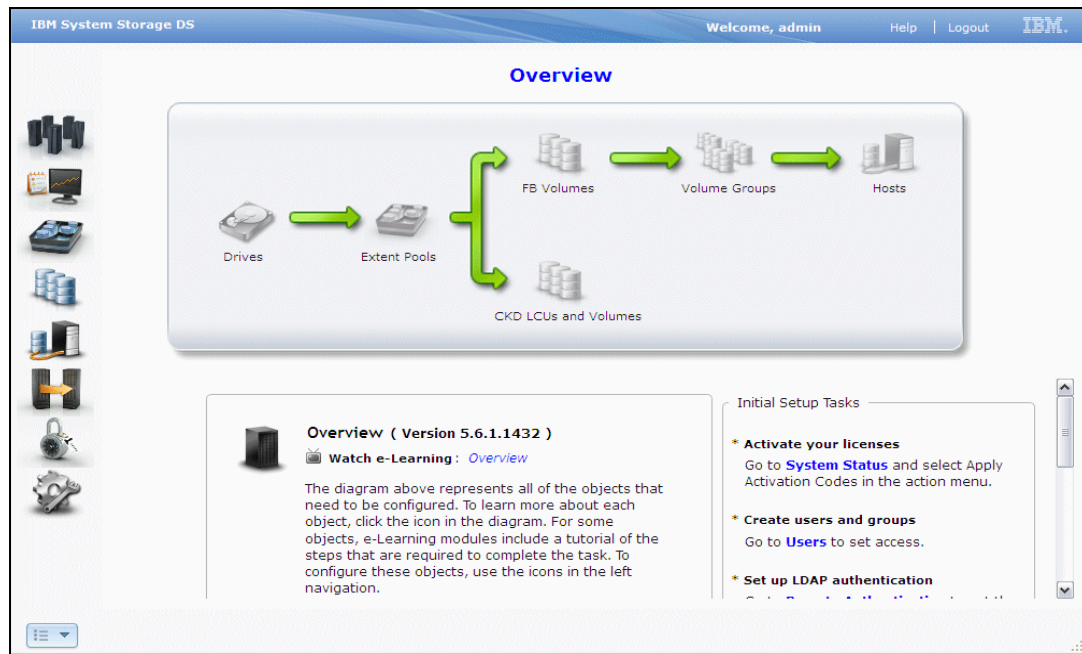


Figure 9-9 DS Storage Manager GUI main window

- From the categories in the left sidebar, select **Remote Authentication** under the section **Configuration** as shown in Figure 9-10.

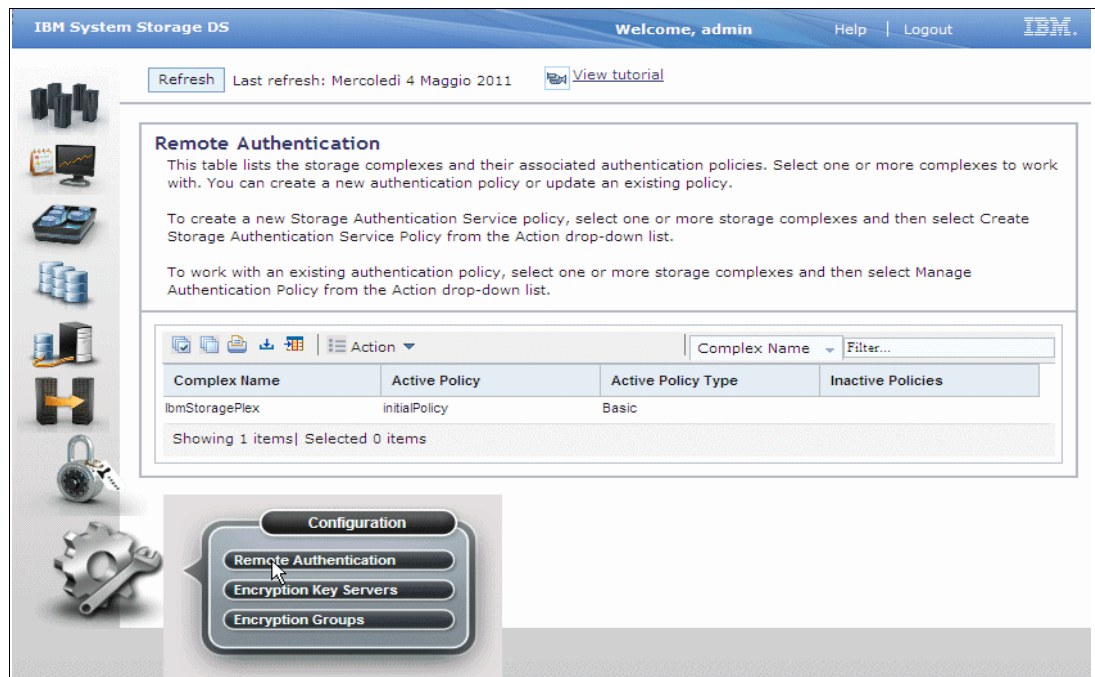


Figure 9-10 Remote Authentication

- You are presented with a list of the storage complexes and their active security policies. Select the complex that you want to modify. You can choose to either create a new security policy or manage one of the existing policies. Do this by selecting **Create Storage Authentication Service Policy** or **Manage Authentication Policy** from the **Select** menu as shown in Figure 9-11.

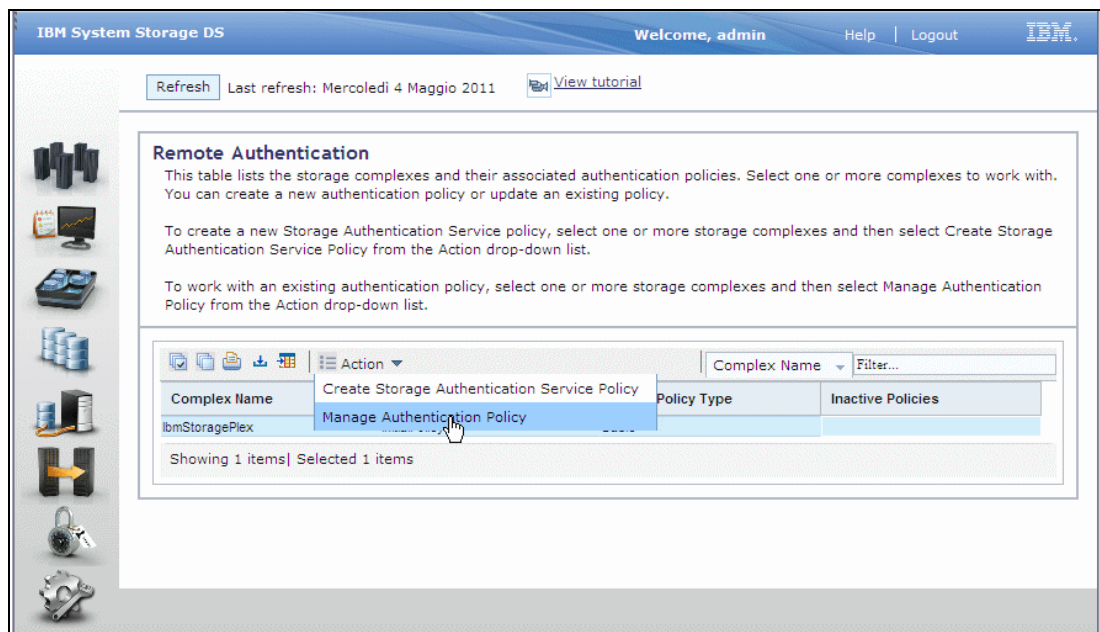


Figure 9-11 Selecting a storage complex

- The next window displays all of the security policies on the HMC for the storage complex you chose. Note that you can create many policies, but only one at a time can be active. Select a policy by highlighting the row. Then select **Properties** from the **Select** menu as shown in Figure 9-12.

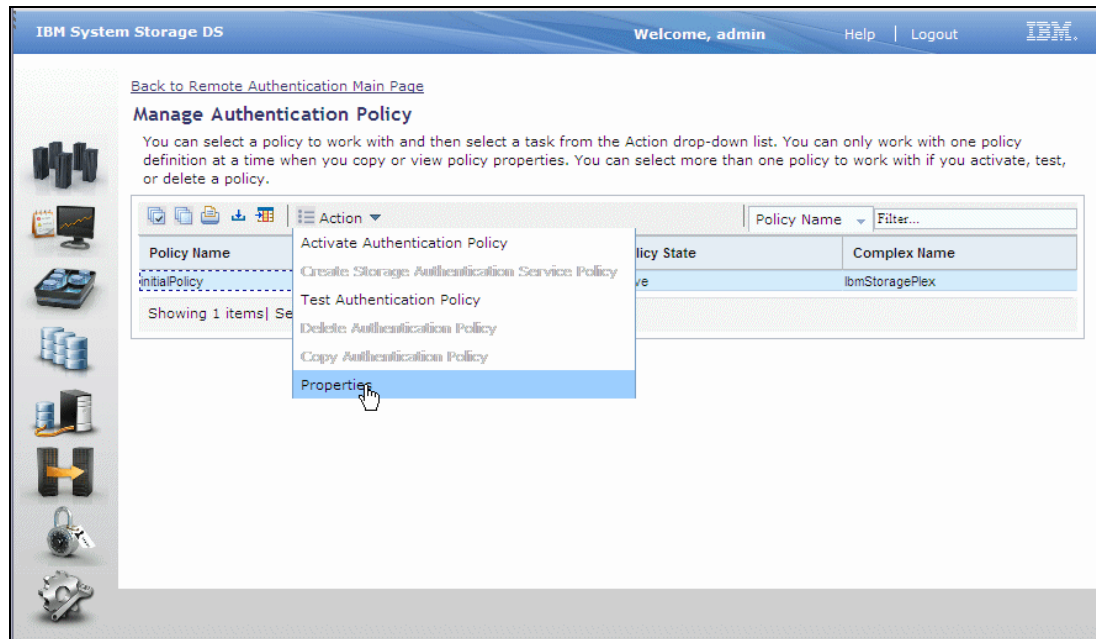


Figure 9-12 Selecting a security policy

- The next window shows you the users defined on the HMC. You can choose to add a new user (click **Select action** → **Add user**) or modify the properties of an existing user as shown in Figure 9-13.

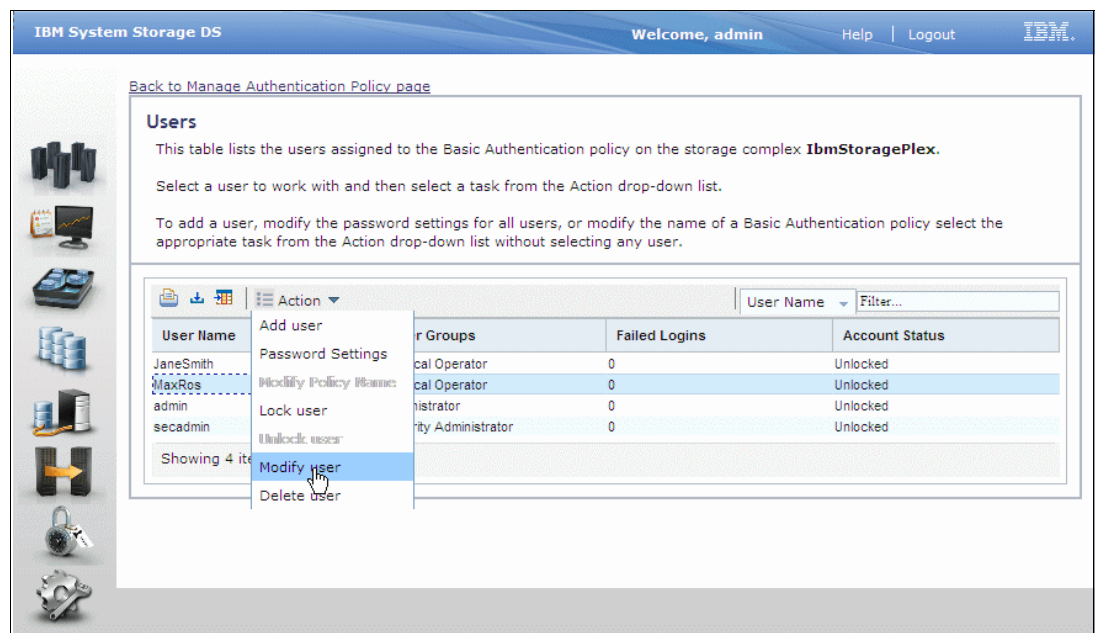


Figure 9-13 Selecting Modify User

The administrator can perform several tasks from this window:

- **Add User** (The DS CLI equivalent is **mkuser**)
- **Modify User** (The DS CLI equivalent is **chuser**)
- **Lock** or **Unlock User**: Choice will toggle (The DS CLI equivalent is **chuser**)
- **Delete User** (The DS CLI equivalent is **rmuser**)
- **Password Settings** (The DS CLI equivalent is **chpass**)

6. The **Password Settings** window is where you can modify the number of days before a password expires, and the number of login retries that a user gets before the account becomes locked, as shown in Figure 9-14.

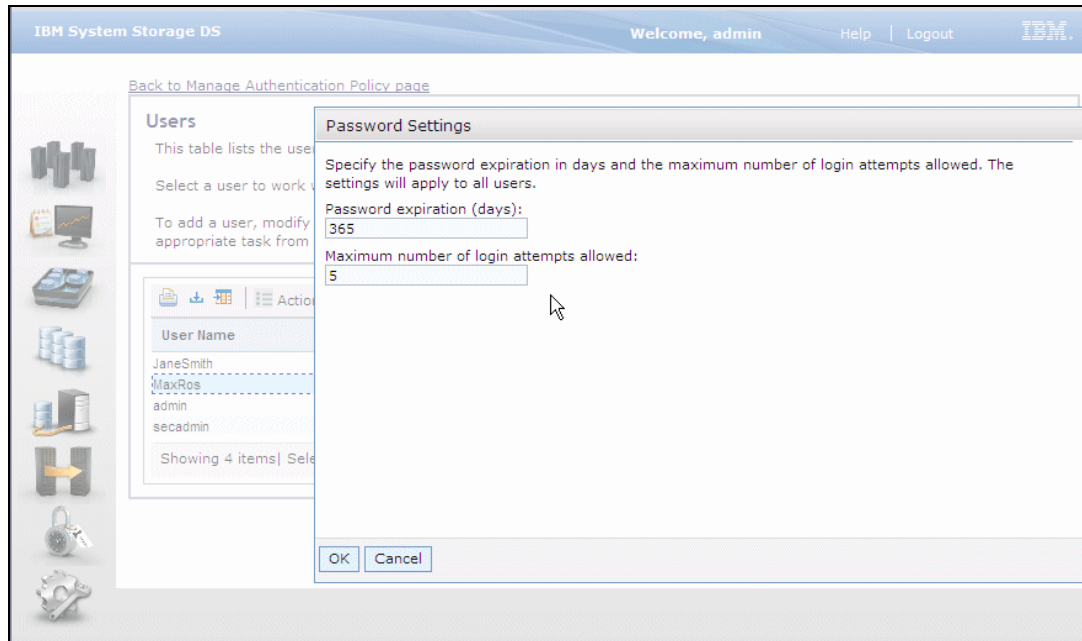


Figure 9-14 Password Settings window

Tip: If a user who is not in the *Administrator* group logs on to the DS GUI and goes to the User Administration window, the user will only be able to see their own user ID in the list. The only action they will be able to perform is to change their password.

7. Selecting **Add user** displays a window in which a user can be added by entering the user ID, the temporary password, and the role. See Figure 9-15 for an example. The role will decide what type of activities can be performed by this user. In this window, the user ID can also be temporarily deactivated by selecting only the **No access** option.

Figure 9-15 Adding a new user to the HMC

Take special note of the new role of the Security Administrator (secadmin). This role was created to separate the duties of managing the storage from managing the encryption for DS8000 units that are shipped with Full Disk Encryption storage drives.

If you are logged in to the GUI as a Storage Administrator, you cannot create, modify, or delete users of the Security Administrator role. Notice how the **Security Administrator** option is disabled in the Add/Modify User window in Figure 9-15. Similarly, Security Administrators cannot create, modify, or delete Storage Administrators. This is a new feature of the microcode for the DS8000.

9.6 External HMC

An external, secondary HMC (for redundancy) can be ordered for the DS8000. The external HMC is an optional purchase, but one that is highly useful. The two HMCs run in a dual-active configuration, so either HMC can be used at any time. For this book, the distinction between the internal and external HMC is only for the purposes of clarity and explanation because they are identical in functionality.

The DS8000 is capable of performing all storage duties while the HMC is down, but configuration, error reporting, and maintenance capabilities become severely restricted. Any organization with extremely high availability requirements should consider deploying an external HMC.

Tip: To help preserve Data Storage functionality, the internal and external HMCs are not available to be used as general purpose computing resources.

9.6.1 External HMC benefits

Having an external HMC provides a number of advantages. Among these are:

- ▶ Enhanced maintenance capability

Because the HMC is the only interface available for service personnel, an external HMC will provide maintenance operational capabilities if the internal HMC fails.

- ▶ Greater availability for power management

Using the HMC is the only way to safely power on or power off the DS8000. An external HMC is necessary to shut down the DS8000 in the event of a failure with the internal HMC.

- ▶ Greater availability for remote support over modem

A second HMC with a phone line on the modem provides IBM with a way to perform remote support if an error occurs that prevents access to the first HMC. If network offload (FTP) is not allowed, one HMC can be used to offload data over the modem line while the other HMC is used for troubleshooting. See Chapter 17, “Remote support” on page 425 for more information regarding HMC modems.

- ▶ Greater availability of encryption deadlock recovery

If the DS8000 is configured for full disk encryption and an encryption deadlock scenario occurs, then using the HMC is the only way to input a Recovery Key to allow the DS8000 to become operational. See 4.8.1, “Deadlock recovery” on page 99 for more information regarding encryption deadlock.

- ▶ Greater availability for Advanced Copy Services

Because all Copy Services functions are driven by the HMC, any environment using Advanced Copy Services should have dual HMCs for operations continuance.

- ▶ Greater availability for configuration operations

All configuration commands must go through the HMC. This is true regardless of whether access is through the TPC BE, SSPC, DS CLI, the DS Storage Manager, or DS Open API with another management program. An external HMC will allow these operations to continue in the event of a failure with the internal HMC.

When a configuration or Copy Services command is issued, the DS CLI or DS Storage Manager will send the command to the first HMC. If the first HMC is not available, it will automatically send the command to the second HMC instead. Typically, you do not have to reissue the command.

Any changes made using one HMC are instantly reflected in the other HMC. There is no caching of host data done within the HMC, so there are no cache coherency issues.

9.6.2 Configuring DS CLI to use a second HMC

The second HMC can either be specified on the command line or in the profile file used by the DS CLI. To specify the second HMC in a command, use the `-hmc2` parameter, as shown in Example 9-9.

Example 9-9 Using the -hmc2 parameter

```
C:\Program Files\IBM\dsccli>dsccli -hmc1 10.0.0.1 -hmc2 10.0.0.5
Enter your username: MaxRos
Enter your password:
Date/Time: April 25, 2011 04:45:33 PM MST IBM DSCLI Version: 7.6.10.464 DS:
IBM.2107-75LX521
```



```
dscli>
```

Alternatively, you can modify the following lines in the `dscli.profile` (or any profile) file:

```
# Management Console/Node IP Address(es)
# hmc1 and hmc2 are equivalent to -hmc1 and -hmc2 command options.
hmc1:10.0.0.1
hmc2:10.0.0.5
```

After you make these changes and save the profile, the DS CLI will be able to automatically communicate through HMC2 in the event that HMC1 becomes unreachable. This change will allow you to perform both configuration and Copy Services commands with full redundancy.

9.7 Configuration worksheets

During the installation of the DS8000, your IBM service representative customizes the setup of your storage complex based on information that you provide in a set of customization worksheets. Each time that you install a new storage unit or management console, you must complete the customization worksheets before the IBM service representatives can perform the installation.

The customization worksheets are important and need to be completed before the installation. It is important that this information is entered into the machine so that preventive maintenance and high availability of the machine are maintained. You can find the customization worksheets in *IBM System Storage DS8000 Introduction and Planning Guide*, GC27-2297.

The customization worksheets allow you to specify the initial setup for the following items:

- ▶ **Company information:** This information allows IBM service representatives to contact you as quickly as possible when they need to access your storage complex.
- ▶ **Management console network settings:** This allows you to specify the IP address and LAN settings for your management console (MC).
- ▶ **Remote support (includes Call Home and remote service settings):** This allows you to specify whether you want outbound (Call Home) or inbound (remote services) remote support.
- ▶ **Notifications (include SNMP trap and email notification settings):** This allows you to specify the types of notifications that you want and that others might want to receive.
- ▶ **Power control:** This allows you to select and control the various power modes for the storage complex.
- ▶ **Control Switch settings:** This allows you to specify certain DS8000 settings that affect host connectivity. You need to enter these choices on the control switch settings worksheet so that the service representative can set them during the installation of the DS8000.

Important: IBM service representatives cannot install a DS8000 system or management console until you provide them with the completed customization worksheets.

9.8 Configuration flow

The following list shows the tasks that need to be done when configuring storage in the DS8000. The order of the tasks does not have to be exactly as shown here, and some of the individual tasks can be done in a different order.

Important: The configuration flow changes when you use the Full Disk Encryption Feature for the DS8000. For details, refer to *IBM System Storage DS8000: IBM System Storage DS8700 Disk Encryption Implementation and Usage Guidelines*, REDP-4500.

1. Install license keys: Activate the license keys for the DS8000.
2. Create arrays: Configure the installed disk drives as either RAID 5, RAID 6, or RAID 10 arrays.
3. Create ranks: Assign each array to either a fixed block (FB) rank or a count key data (CKD) rank.
4. Create Extent Pools: Define Extent Pools, associate each one with either Server 0 or Server 1, and assign at least one rank to each Extent Pool. If you want to take advantage of Storage Pool Striping, you must assign multiple ranks to an Extent Pool. Note that with current versions of the DS GUI, you can start directly with the creation of Extent Pools (arrays and ranks will be automatically and implicitly defined).
5. Create a repository for Space Efficient volumes.
6. Configure I/O ports: Define the type of the Fibre Channel/FICON ports. The port type can be either Switched Fabric, Arbitrated Loop, or FICON.
7. Create host connections for open systems: Define open systems hosts and their Fibre Channel (FC) host bus adapter (HBA) worldwide port names.
8. Create volume groups for open systems: Create volume groups where FB volumes will be assigned and select the host attachments for the volume groups.
9. Create open systems volumes: Create striped open systems FB volumes and assign them to one or more volume groups.
10. Create System z logical control units (LCUs): Define their type and other attributes, such as subsystem identifiers (SSIDs).
11. Create striped System z volumes: Create System z CKD base volumes and Parallel Access Volume (PAV) aliases for them.

The actual configuration can be done using the DS Storage Manager GUI, DS Command-Line Interface, or a mixture of both. A novice user might prefer to use the GUI, whereas a more experienced user might use the CLI, particularly for the more repetitive tasks such as creating large numbers of volumes.

For a more detailed discussion about how to perform the specific tasks, see:

- ▶ Chapter 10, “IBM System Storage DS8000 features and license keys” on page 251
- ▶ Chapter 13, “Configuration using the DS Storage Manager GUI” on page 309
- ▶ Chapter 14, “Configuration with the DS Command-Line Interface” on page 363

General guidelines when configuring storage

Remember the following general guidelines when configuring storage on the DS8000:

- ▶ To achieve a well-balanced load distribution, use at least two Extent Pools, each assigned to one DS8000 internal server (extent Pool 0 and Extent Pool 1). If CKD and FB volumes are required, use at least four Extent Pools.

- ▶ Address groups (16 LCUs/logical subsystems (LSSs)) are all for CKD or all for FB.
- ▶ Volumes of one LCU/LSS can be allocated on multiple Extent Pools.
- ▶ An Extent Pool cannot contain all three RAID 5, RAID 6, and RAID 10 ranks. Each Extent Pool pair should have the same characteristics in terms of RAID type, RPM, and DDM size.
- ▶ Ranks in one Extent Pool should belong to separate Device Adapters.
- ▶ Assign multiple ranks to Extent Pools to take advantage of Storage Pool Striping.
- ▶ CKD guidelines:
 - 3380 and 3390 type volumes can be intermixed in an LCU and an Extent Pool.
- ▶ FB guidelines:
 - Create a volume group for each server unless LUN sharing is required.
 - Place all ports for a single server in one volume group.
 - If LUN sharing is required, there are two options:
 - Use separate volumes for servers and place LUNs in multiple volume groups.
 - Place servers (clusters) and volumes to be shared in a single volume group.
- ▶ I/O ports guidelines:
 - Distribute host connections of each type (FICON and FCP) evenly across the I/O enclosure.
 - Typically, access *any* is used for I/O ports with access to ports controlled by SAN zoning.



IBM System Storage DS8000 features and license keys

This chapter discusses the activation of licensed functions and the following topics:

- ▶ IBM System Storage DS8000 licensed functions
- ▶ Activation of licensed functions
- ▶ Licensed scope considerations

10.1 IBM System Storage DS8000 licensed functions

Many of the functions of the DS8000 that we have discussed so far are optional licensed functions that must be *enabled* for use. The licensed functions are enabled through a 242x licensed function *indicator* feature, plus a 239x licensed function *authorization* feature number, in the following way:

- The licensed functions for DS8000 are enabled through a pair of 242x-9x1 licensed function indicator feature numbers (FC 07xx and FC 7xxx), plus a Licensed Function Authorization (239x-LFA), feature number (FC 7xxx). These functions and numbers are listed in Table 10-1.

Table 10-1 DS8000 licensed functions

Licensed function for DS8000 with Enterprise Choice warranty	IBM 242x indicator feature numbers	IBM 239x function authorization model and feature numbers
Operating Environment License	0700 and 70xx	239x Model LFA, 703x/706x
Thin Provisioning	0707 and 7071	239x Model LFA, 7071
FICON Attachment	0703 and 7091	239x Model LFA, 7091
Database Protection	0708 and 7080	239x Model LFA, 7080
High Performance FICON	0709 and 7092	239x Model LFA, 7092
FlashCopy	0720 and 72xx	239x Model LFA, 725x-726x
Space Efficient FlashCopy	0730 and 73xx	239x Model LFA, 735x-736x
Metro/Global Mirror	0742 and 74xx	239x Model LFA, 748x-749x
Metro Mirror	0744 and 75xx	239x Model LFA, 750x-751x
Global Mirror	0746 and 75xx	239x Model LFA, 752x-753x
z/OS Global Mirror	0760 and 76xx	239x Model LFA, 765x-766x
z/OS Metro/Global Mirror Incremental Resync	0763 and 76xx	239x Model LFA, 768x-769x
Parallel Access Volumes	0780 and 78xx	239x Model LFA, 782x-783x
HyperPAV	0782 and 7899	239x Model LFA, 7899
I/O Priority Manager	0784 and 784x	239x Model LFA, 784x
Easy Tier	0713 and 7083	239x Model LFA, 7083

- The DS8000 provides Enterprise Choice warranty options associated with a specific machine type. The x in 242x designates the machine type according to its warranty period, where x can be either 1, 2, 3, or 4. For example, a 2424-951 machine type designates a DS8800 storage system with a four-year warranty period.
- The x in 239x can be either 6, 7, 8, or 9, according to the associated 242x base unit model. 2396 function authorizations apply to 2421 base units, 2397 to 2422, and so on. For example, a 2399-LFA machine type designates a DS8000 Licensed Function Authorization for a 2424 machine with a four-year warranty period.
- The 242x licensed function indicator feature numbers enable the technical activation of the function, subject to the client applying a feature activation code made available by IBM.

The 239x licensed function authorization feature numbers establish the extent of authorization for that function on the 242x machine for which it was acquired.

IBM offers value-based licensing for the Operating Environment License. It is priced based on the disk drive performance, capacity, speed, and other characteristics that provide more flexible and optimal price/performance configurations. As shown in Table 10-2, each feature indicates a certain number of value units.

Table 10-2 Operating Environment License (OEL): value unit indicators

Feature number	Description
7050	OEL - inactive indicator
7051	OEL - 1 value unit indicator
7052	OEL - 5 value unit indicator
7053	OEL - 10 value unit indicator
7054	OEL - 25 value unit indicator
7055	OEL - 50 value unit indicator
7060	OEL - 100 value unit indicator
7065	OEL - 200 value unit indicator

These features are required in addition to the per TB OEL features (#703x-704x). For each disk drive set, the corresponding number of value units must be configured, as shown in Table 10-3. and Table 10-4

Table 10-3 DS8800 Value unit requirements based on drive size, type, and speed

Drive set feature number	Drive size	Drive type	Drive speed	Encryption drive	Value units required
2208	146 GB	SAS	15K RPM	No	4.8
2608	450 GB	SAS	10K RPM	No	8.6
2708	600 GB	SAS	10K RPM	No	10.9
5608	450 GB	SAS	10K RPM	Yes	8.6
5708	600 GB	SAS	10K RPM	Yes	10.9
6008	300 GB	SSD	N/A	No	6.8

Table 10-4 DS8700 Value unit requirements based on drive size, type, and speed

Drive set feature number	Drive size	Drive type	Drive speed	Encryption drive	Value units required
6016	73 GB	SSD	N/A	No	12
6014	73 GB	SSD half set	N/A	No	6
6116	146 GB	SSD	N/A	No	18
6114	146 GB	SDD half set	N/A	No	9

Drive set feature number	Drive size	Drive type	Drive speed	Encryption drive	Value units required
2216	146 GB	FC	15K RPM	No	4.8
2416	300 GB	FC	15K RPM	No	6.8
2616	450 GB	FC	15K RPM	No	9
2716	600 GB	FC	15K RPM	Yes	11.5
5016	146 GB	FC	15K RPM	Yes	4.8
5116	300 GB	FC	15K RPM	Yes	6.8
5216	450 GB	FC	15K RPM	Yes	9
2816	1 TB	SATA	7.2K RPM	No	11
2916	2 TB	SATA	7.2K RPM	No	20

The HyperPAV license is a flat-fee, add-on license that requires the Parallel Access Volumes (PAV) license to be installed.

The license for Space-Efficient FlashCopy does not require the ordinary FlashCopy (PTC) license. As with the ordinary FlashCopy, the FlashCopy SE is licensed in tiers by gross amount of TB installed. FlashCopy (PTC) and FlashCopy SE can be complementary licenses. FlashCopy SE will serve to do FlashCopies with Track Space-Efficient (TSE) target volumes. When also doing FlashCopies to standard target volumes, use the PTC license in addition.

Metro Mirror (MM license) and Global Mirror (GM) can be complementary features as well.

Tip: For a detailed explanation of the features involved and the considerations you must have when ordering DS8000 licensed functions, refer to these announcement letters:

- ▶ IBM System Storage DS8000 series (IBM 242x)
- ▶ IBM System Storage DS8000 series (M/T 239x) high performance flagship - Function Authorizations.

IBM announcement letters can be found at the following address:

<http://www.ibm.com/common/ssi/index.wss>

Use the *DS8700* or *DS8800* keyword as a search criteria in the Contents field.

10.2 Activation of licensed functions

Activating the license keys of the DS8000 can be done after the IBM service representative has completed the storage complex installation. Based on your 239x licensed function order, you need to obtain the necessary keys from the IBM Disk Storage Feature Activation (DSFA) website at the following address:

<http://www.ibm.com/storage/dsfa>

Important: There is a special procedure to obtain the license key for the Full Disk Encryption feature. It *cannot* be obtained from the DSFA website. Refer to *IBM System Storage DS8700 Disk Encryption*, REDP-4500, for more information.

You can activate all license keys at the same time (for example, on initial activation of the storage unit) or they can be activated individually (for example, additional ordered keys).

Before connecting to the IBM DSFA website to obtain your feature activation codes, ensure that you have the following items:

- ▶ The IBM License Function Authorization documents. If you are activating codes for a new storage unit, these documents are included in the shipment of the storage unit. If you are activating codes for an existing storage unit, IBM will send the documents to you in an envelope.
- ▶ A USB memory device can be used for downloading your activation codes if you cannot access the DS Storage Manager from the system that you are using to access the DSFA website. Instead of downloading the activation codes in softcopy format, you can also print the activation codes and manually enter them using the DS Storage Manager GUI. However, this is slow and error prone, because the activation keys are 32-character long strings.

10.2.1 Obtaining DS8000 machine information

To obtain license activation keys from the DFSA website, you need to know the serial number and machine signature of your DS8000 unit.

To obtain the required information, perform the following steps:

1. Start the DS Storage Manager application. Log in using a user ID with administrator access. If this is the first time you are accessing the machine, contact your IBM service representative for the user ID and password. After a successful login, the DS8000 Storage Manager Overview window opens. Move your cursor to the left top icon that will cause a fish-eye effect and a window pop-up. Select **System Status** (Figure 10-1).

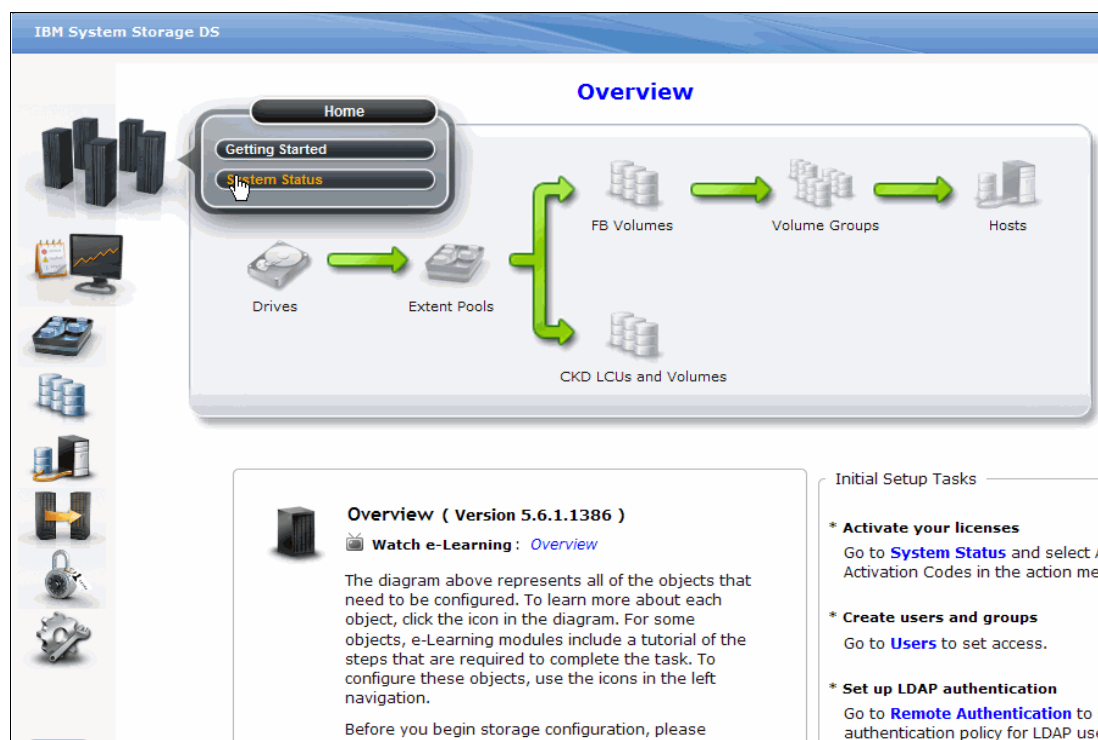


Figure 10-1 DS8000 Storage Manager GUI: Overview window

- Click the Serial number under the **Storage Image** header, then click **Action**. Move your cursor to the **Storage Image** and select **Add Activation Key** (Figure 10-2).

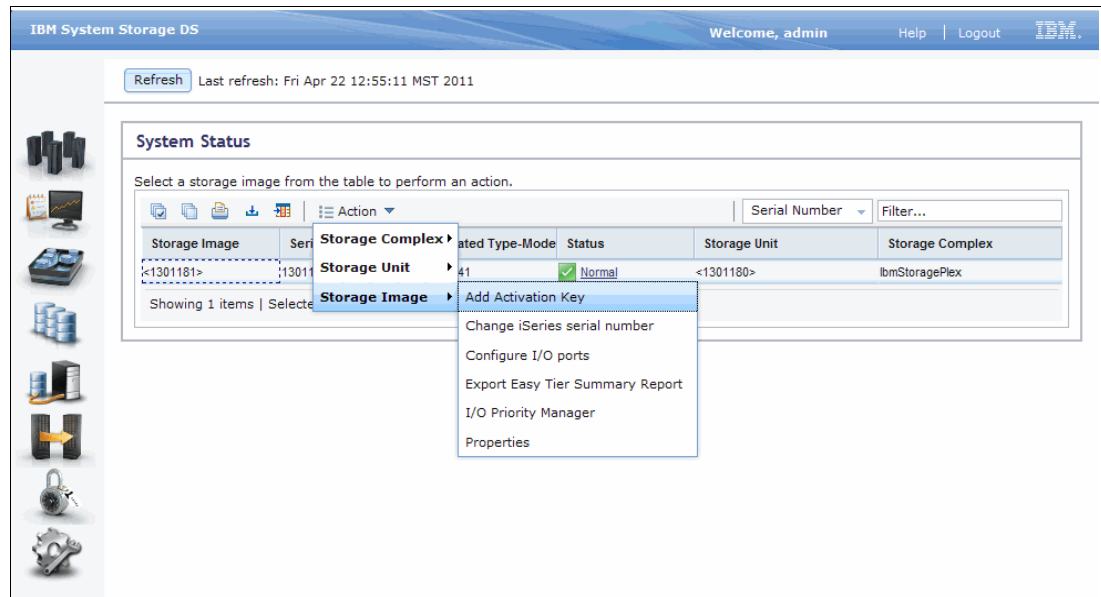


Figure 10-2 DS8000 Storage Manager: Add Activation Key

- This window shows the **Serial Number** and the **Machine Signature** of your DS8000 Storage Image (Figure 10-3).

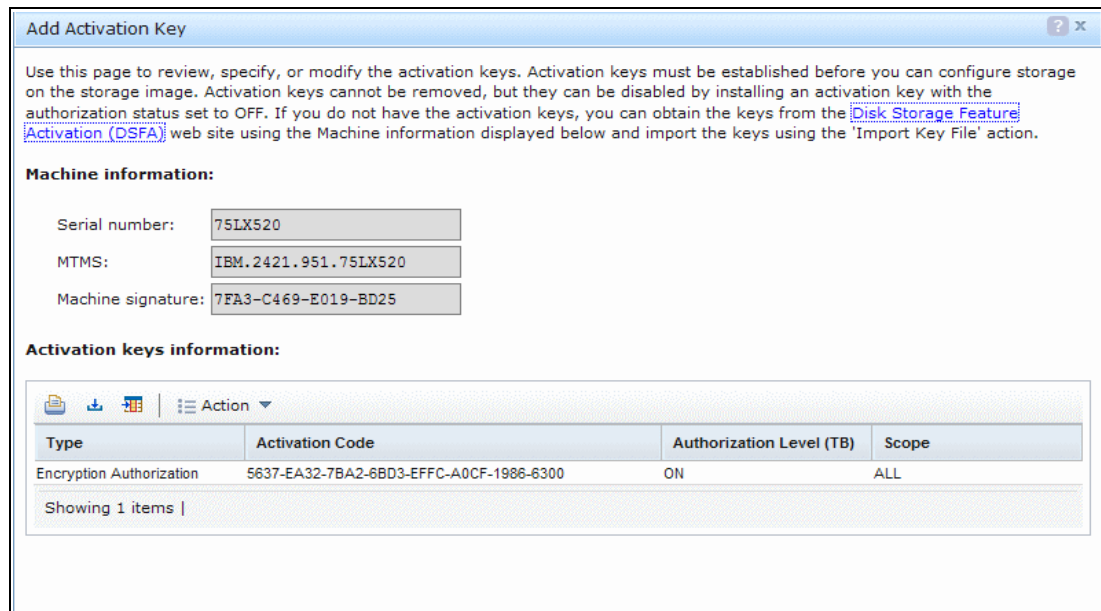


Figure 10-3 DS8000 Machine Signature & Serial Number

Gather the following information about your storage unit:

- The MTMS (Machine Type - Model Number - Serial Number) is a string that contains the machine type, model number, and serial number. The machine type is 242x and the machine mode is 9x1. The last seven characters of the string are the machine's serial number (XYABCDE).

- From the Machine signature field, note the machine signature (ABCD-EF12-3456-7890).

Use Table 10-5 to document this information, which will be entered in the IBM DSFA website to retrieve the activation codes.

Table 10-5 DS8000 machine information

Property	Your storage unit's information
Machine type and model	
Machine's serial number	
Machine signature	

4. Now you can go to 10.2.2, "Obtaining activation codes" on page 257
5. Another way you can link to the DSFA web site is to move your cursor to right-click the blue DSFA area. Select **Open Link in New Window** to launch the DSFA window (Figure 10-4).

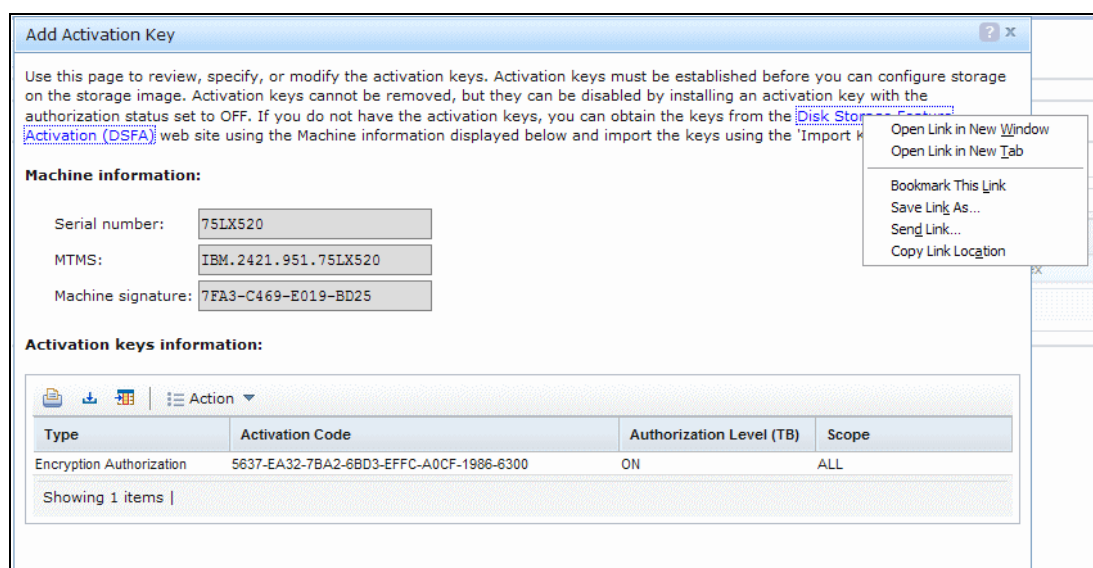


Figure 10-4 Launch DSFA URL link

10.2.2 Obtaining activation codes

Perform the following steps to obtain the activation codes:

1. Connect to the IBM Disk Storage Feature Activation (DSFA) website at the following address:

<http://www.ibm.com/storage/dsfa>

Figure 10-5 on page 258 shows the DSFA website.

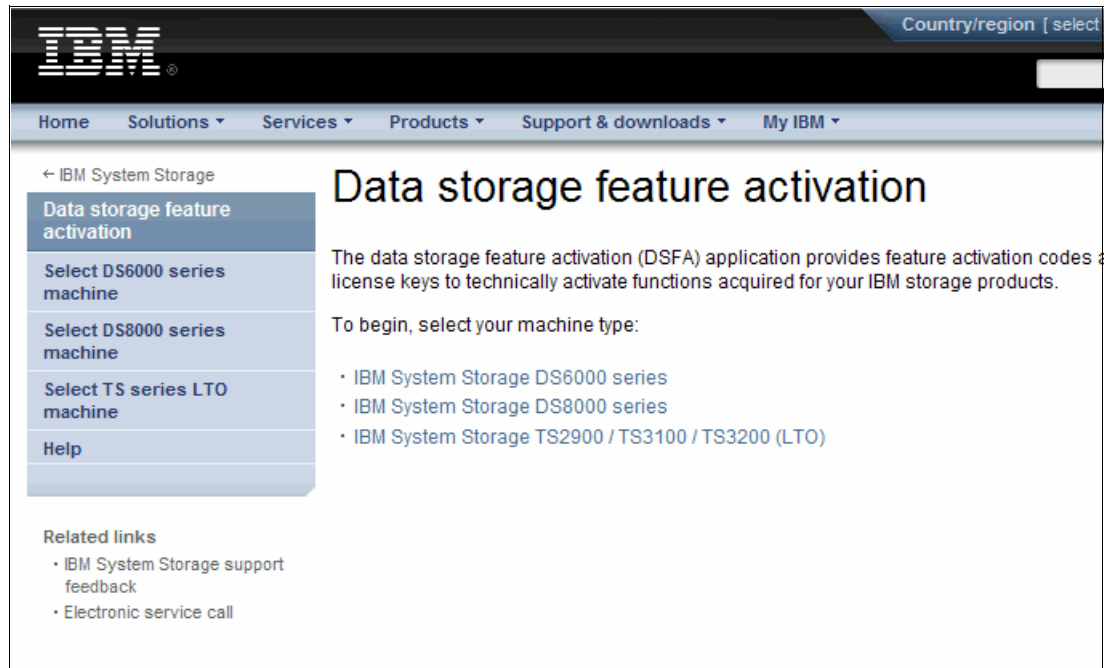


Figure 10-5 IBM DSFA website

2. Click **IBM System Storage DS8000 series**. This brings you to the Select DS8000 series machine window (Figure 10-6). Select the appropriate 242x Machine Type.

Figure 10-6 DS8000 DSFA machine information entry window

3. Enter the machine information collected in Table 10-5 on page 257 and click **Submit**. The View machine summary window opens (Figure 10-7).

[Home](#)
[Products](#)
[Services & solutions](#)
[Support & downloads](#)
[My account](#)

← IBM TotalStorage

Disk storage feature activation

Select DS6000 series machine

Select DS8000 series machine

• View machine summary

• Manage activations

• View activation codes

• Assign function authorization

• View authorization details

Help

Related links

• IBM TotalStorage support feedback

• Electronic service call

View machine summary

IBM TotalStorage DS8300

Use this page to verify the functions on your machine. If applicable, take the following actions before viewing your feature activation codes:

- [Assign a function authorization](#) if you received a DS8000 series licensed function authorization document, and the function is not shown below.
- [Manage an activation](#) if this is the first time you are activating a function or an unassigned value is not 0 TB.

If these conditions do not apply, you can [view your feature activation codes](#).

IBM 2107 Model 9A2 Serial number 75-ABTV0

Feature code	Description
0700	OEL indicator
0720	PTC indicator
0740	RMC indicator
0760	RMZ indicator
0780	PAV indicator

IBM 2244 Model OEL Serial number 75-0DF11

Description	Total license	Assigned	Unassigned
Operating environment	25.0 TB	25.0 TB	0.0 TB

IBM 2244 Model PAV Serial number 75-0DF21

Description	Total license	Assigned	Unassigned
Parallel access volumes	25.0 TB	25.0 TB	0.0 TB

IBM 2244 Model PTC Serial number 75-0DF31

Description	Total license	Assigned	Unassigned
Point in time copy	25.0 TB	25.0 TB	0.0 TB

IBM 2244 Model RMC Serial number 75-0DF41

Description	Total license	Assigned	Unassigned
Remote mirror and copy	25.0 TB	25.0 TB	0.0 TB

IBM 2244 Model RMZ Serial number 75-0DF51

Description	Total license	Assigned	Unassigned
Remote mirror for z/OS	25.0 TB	25.0 TB	0.0 TB

Figure 10-7 DSFA View machine summary window

The View machine summary window shows the total purchased licenses and how many of them are currently assigned. The example in Figure 10-7 shows a storage unit where all licenses have already been assigned. When assigning licenses for the first time, the Assigned field shows 0.0 TB.

4. Click **Manage activations**. The Manage activations window opens. Figure 10-8 on page 260 shows the Manage activations window for your storage images. For each license type and storage image, enter the license scope (fixed block data (FB), count key

data (CKD), or All) and a capacity value (in TB) to assign to the storage image. The capacity values are expressed in decimal terabytes with 0.1 TB increments. The sum of the storage image capacity values for a license cannot exceed the total license value.

← IBM TotalStorage

Disk storage feature activation

Select DS6000 series machine

Select DS8000 series machine

• View machine summary

• Manage activations

• View activation codes

• Assign function authorization

• View authorization details

Help

Related links

• IBM TotalStorage support feedback

• Electronic service call

Manage activations

IBM 2107 Model 9A2

Serial number 75-ABTV0

Use this page to select the license scope and specify the assigned value for your functions. Both of these functions are performed on a storage image basis.

The [license scope](#) defines the type of storage, and therefore the type of servers, the function will be technically enabled for. You must select a license scope for those functions that have a license scope option.

The [assigned value](#) defines the authorization level the function will be technically enabled to. You must initially specify an assigned value for each storage image. You must also update these values in a future session if your total license value has increased or decreased.

Click Submit when you are finished.

Operating environment

Storage image	Scope	Assigned value
Image 1	All	16.0 TB
Image 2	All	9.0 TB
Total license:		25.0 TB
Available for assignment:		0.0 TB

Parallel access volumes

Storage image	Scope	Assigned value
Image 1	CKD	16.0 TB
Image 2	CKD	9.0 TB
Total license:		25.0 TB
Available for assignment:		0.0 TB

Point in time copy

Storage image	Scope	Assigned value
Image 1	All	16.0 TB
Image 2	All	9.0 TB
Total license:		25.0 TB
Available for assignment:		0.0 TB

Remote mirror and copy

Storage image	Scope	Assigned value
Image 1	All	16.0 TB
Image 2	All	9.0 TB
Total license:		25.0 TB
Available for assignment:		0.0 TB

Remote mirror for z/OS

Storage image	Scope	Assigned value
Image 1	CKD	16.0 TB
Image 2	CKD	9.0 TB
Total license:		25.0 TB
Available for assignment:		0.0 TB

Submit

Figure 10-8 DSFA Manage activations window

260 IBM System Storage DS8000: Architecture and Implementation

- When you have entered the values, click **Submit**. The View activation codes window opens, showing the license activation codes for the storage images (Figure 10-9). Print the activation codes or click **Download** to save the activation codes in a file that you can later import in the DS8000.

IBM Disk storage feature activation: View activation codes - DS8000 series - Microsoft Internet Explorer

Address: <https://www-03.ibm.com/storage/dsfa/generateFeatureActivationCode2107.wss?machineType=2107>

Country/region [select] Terms of use

Home Products Services & solutions Support & downloads My account

← IBM TotalStorage

Disk storage feature activation

Select DS6000 series machine

Select DS8000 series machine

- View machine summary
- Manage activations
- View activation codes
- Assign function authorization
- View authorization details

Help

Related links

- IBM TotalStorage support feedback
- Electronic service call

View activation codes

IBM 2107 Model 9A2
Serial number 75-ABTV0

This page displays the feature activation codes for your machine.

Click "Download now" to download these codes to a file that you can import into your machine using the DS Storage Manager application.

Storage image 1	
Description	Feature activation code
Operating environment	9BBE-0EE4-741F-9607-8A77-A856-5E02-7F8D
Parallel access volumes	8654-2A78-5704-0935-8A77-A856-5E02-7F8D
Point in time copy	D0AA-E9CD-4F1D-25AD-8A77-A856-5E02-7F8D
Remote mirror and copy	B5D7-5FF1-121A-445E-8A77-A856-5E02-7F8D
Remote mirror for z/OS	A401-79A1-7722-BCDB-8A77-A856-5E02-7F8D

Storage image 2	
Description	Feature activation code
Operating environment	9CD1-7A33-7A47-C44B-8A77-A856-5E02-7F8D
Parallel access volumes	65B0-C66A-494F-7F63-8A77-A856-5E02-7F8D
Point in time copy	8DD5-284D-CF4A-D562-8A77-A856-5E02-7F8D
Remote mirror and copy	5346-3A0E-490C-B531-8A77-A856-5E02-7F8D
Remote mirror for z/OS	A6CB-5874-6B4F-12FA-8A77-A856-5E02-7F8D

[Download now](#)

About IBM Privacy Contact

Figure 10-9 DSFA View activation codes window

Tip: In most situations, the DSFA application can locate your 239x licensed function authorization record when you enter the DS8000 (242x) serial number and signature. However, if the 239x licensed function authorization record is not attached to the 242x record, you must assign it to the 242x record using the Assign function authorization link on the DSFA application. In this case, you need the 239x serial number (which you can find on the License Function Authorization document).

10.2.3 Applying activation codes using the GUI

Use this process to apply the activation codes on your DS8000 storage images using the DS Storage Manager GUI. After they are applied, the codes enable you to begin configuring storage on a storage image.

Important: The initial enablement of any optional DS8000 licensed function is a concurrent activity (assuming the appropriate level of microcode is installed on the machine for the given function).

The following activation activities are disruptive and require a machine IML or reboot of the affected image:

- ▶ Removal of a DS8000 licensed function to deactivate the function.
- ▶ A lateral change or reduction in the license scope. A *lateral change* is defined as changing the license scope from fixed block (FB) to count key data (CKD) or from CKD to FB. A *reduction* is defined as changing the license scope from all physical capacity (ALL) to only FB or only CKD capacity.

Attention: Before you begin this task, you must resolve any current DS8000 problems. Contact IBM support for assistance in resolving these problems.

The easiest way to apply the feature activation codes is to download the activation codes from the IBM Disk Storage Feature Activation (DSFA) website to your local computer and import the file into the DS Storage Manager. If you can access the DS Storage Manager from the same computer that you use to access the DSFA website, you can copy the activation codes from the DSFA window and paste them into the DS Storage Manager window. The third option is to manually enter the activation codes in the DS Storage Manager from a printed copy of the codes.

Perform the following steps to apply the activation codes:

1. This method is to apply the activation codes using your local computer or a USB drive. Click the **Action** pull-down menu under **Activation Keys Information** and select **Import Key File** as shown in Figure 10-10.

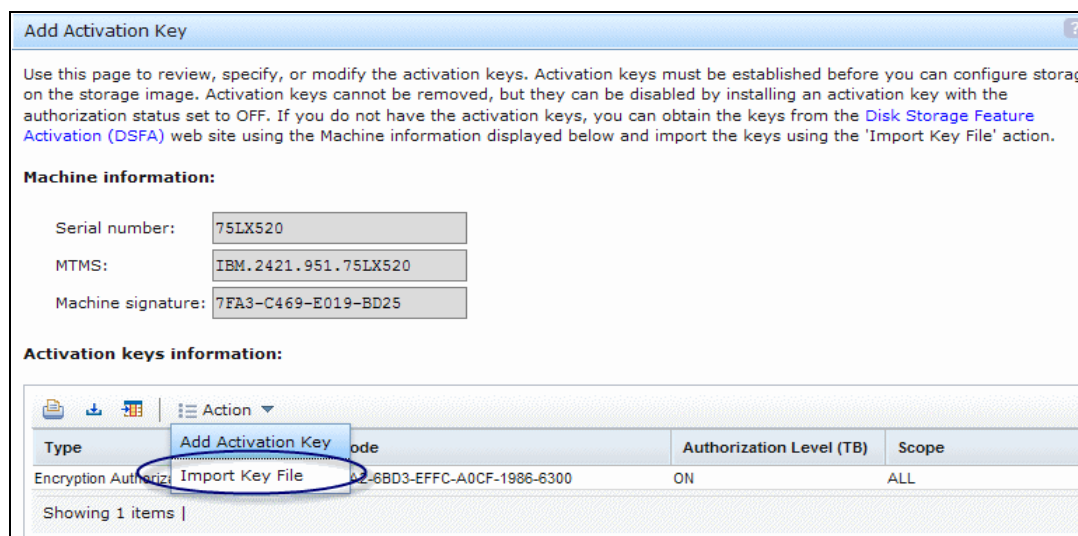


Figure 10-10 DS8000 Storage Manager GUI: select Import Key File

2. Click the **Browse** button and locate the downloaded key file in your computer as shown in Figure 10-11.

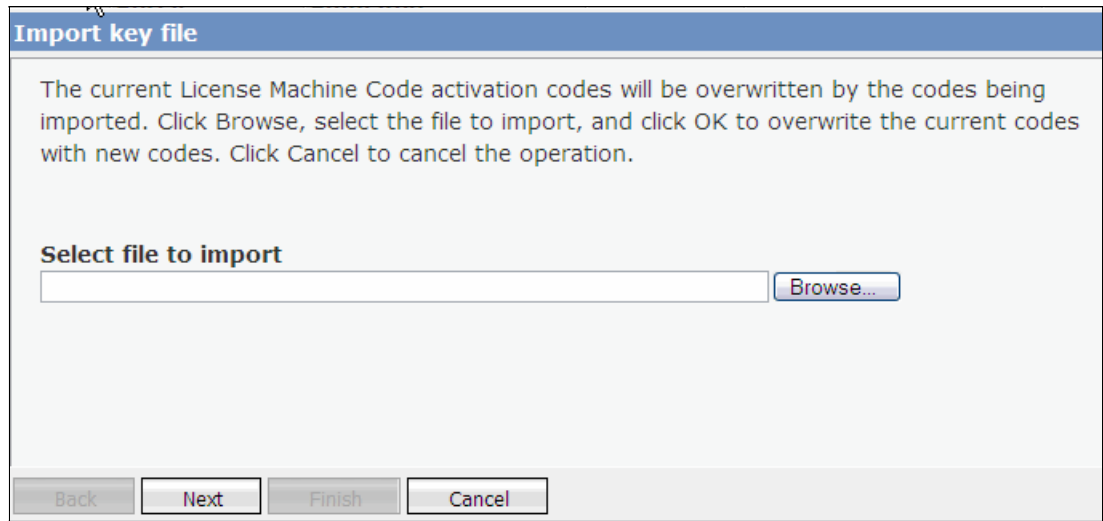


Figure 10-11 Apply Activation Codes by importing the key from the file

3. After the file has been selected, click **Next** to continue. The Confirmation window displays the key name. Click **Finish** to complete the new key activation procedure (Figure 10-12).

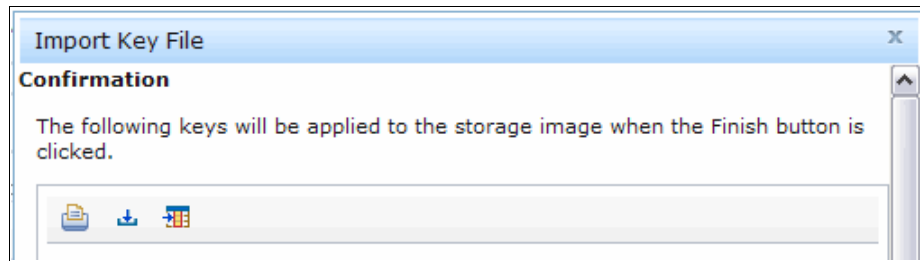
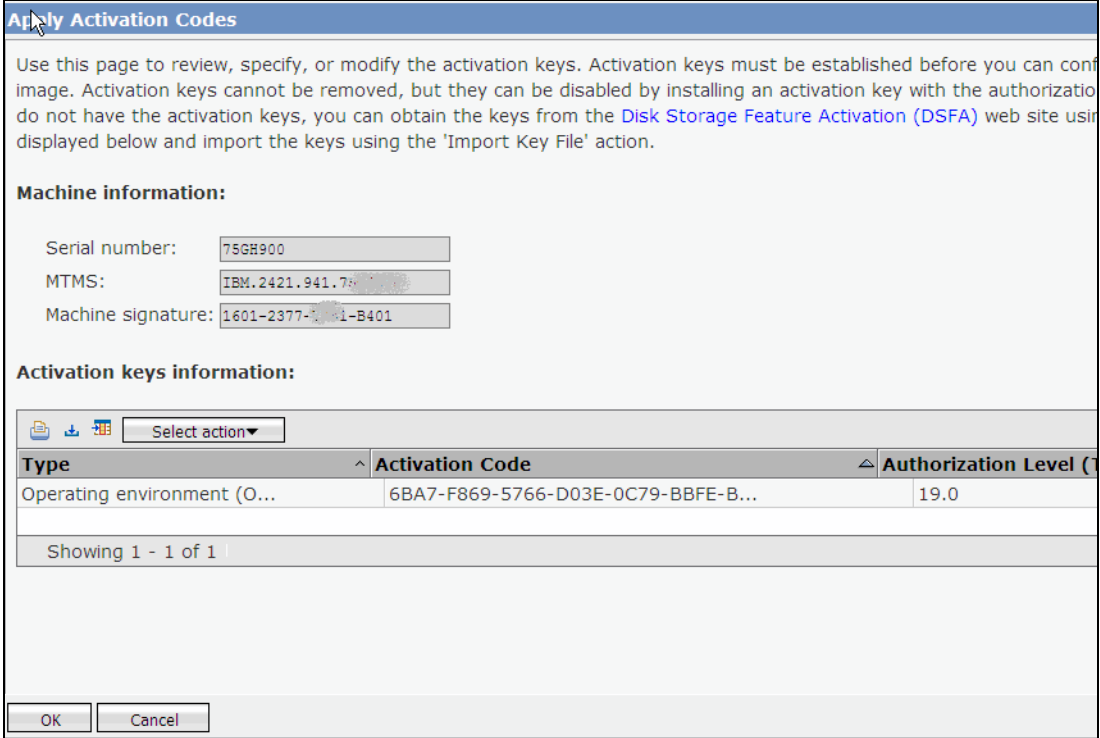


Figure 10-12 Apply Activation Codes: Confirmation window

4. Your license is now listed in the table. In our example, there is one OEL license active, as shown in Figure 10-13 on page 264.



Apply Activation Codes

Use this page to review, specify, or modify the activation keys. Activation keys must be established before you can configure the system. Activation keys cannot be removed, but they can be disabled by installing an activation key with the authorization level of 19.0. If you do not have the activation keys, you can obtain the keys from the [Disk Storage Feature Activation \(DSFA\)](#) web site using the instructions displayed below and import the keys using the 'Import Key File' action.

Machine information:

Serial number: 75GH900
 MTMS: IBM.2421.941.7...
 Machine signature: 1601-2377-...-1-B401

Activation keys information:

Type	Activation Code	Authorization Level (T)
Operating environment (O...	6BA7-F869-5766-D03E-0C79-BBFE-B...	19.0

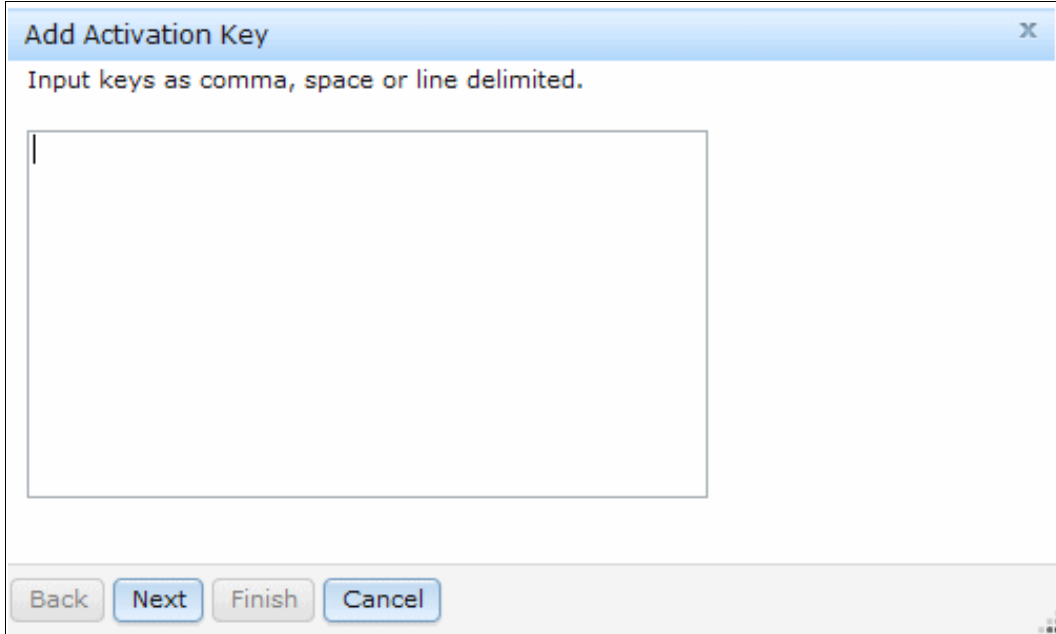
Showing 1 - 1 of 1

OK Cancel

Figure 10-13 Apply Activation Codes window

- Click **OK** to exit Apply Activation Codes wizard.

Another way to enter the Activation Keys is to copy the activation keys from the DSFA window and paste them in the Storage Manager window as shown in Figure 10-14. A third way is to enter the activation keys manually from a printed copy of the codes. Use Enter or Spacebar to separate the keys. Click **Finish** to complete the new key activation procedure.



Add Activation Key

Input keys as comma, space or line delimited.

Back Next Finish Cancel

Figure 10-14 Enter License Key manually

6. The activation codes are displayed as shown in Figure 10-15.

Add Activation Key

Use this page to review, specify, or modify the activation keys. Activation keys must be established before you can configure storage on the storage image. Activation keys cannot be removed, but they can be disabled by installing an activation key with the authorization status set to OFF. If you do not have the activation keys, you can obtain the keys from the [Disk Storage Feature Activation \(DSFA\)](#) web site using the Machine information displayed below and import the keys using the 'Import Key File' action.

Machine information:

Serial number: 1301180

MTMS: IBM.2421.941.1301180

Machine signature: 6DC2-B845-BA42-E494

Activation keys information:

Type	Activation Code	Authorization Level (TB)	Scope
Operating environment (OEL)	3FE3-EEED-CC33-2536-FB5D-4870-C14C-34E9	100.1	ALL
Point in time copy (PTC)	1819-D5E4-567B-3171-FB5D-4870-C14C-34E9	100.1	ALL
FlashCopy SE	86B3-39B3-4E06-9F45-FB5D-4870-C14C-34E9	100.1	ALL
Metro Mirror	BE57-3D56-E0C3-A488-FB5D-4870-C14C-34E9	100.1	ALL
Global Mirror	9402-45F2-FD3F-A276-FB5D-4870-C14C-34E9	100.1	ALL
Metro/Global Mirror (MGM)	203F-20D9-369E-3D5F-FB5D-4870-C14C-34E9	100.1	ALL
Thin Provisioning	DE60-D149-B99D-458A-FB5D-4870-C14C-34E9	ON	ALL
Easy Tier	989B-8B2D-3AA4-5D06-FB5D-4870-C14C-34E9	ON	ALL

Showing 8 items |

OK Cancel

Figure 10-15 Activation codes applied

10.2.4 Applying activation codes using the DS CLI

The license keys can also be activated using the DS CLI. This is available only if the machine Operating Environment License (OEL) has previously been activated and you have a console with a compatible DS CLI program installed.

Perform the following steps:

1. Use the **shows**i command to display the DS8000 machine signature, as shown in Example 10-1.

Example 10-1 DS CLI showsi command

```
dsccli> showsi ibm.2107-75tv181
Date/Time: 10 September 2010 13:22:53 CEST IBM DSCLI Version: 6.6.0.288 DS:
ibm.2107-75tv181
Name          ATS_04
desc          DS8000-R6
ID            IBM.2107-75TV181
Storage Unit   IBM.2107-75TV180
Model         951
WWNN          500507630AFFC29F
```

```

Signature      633b-1234-5678-5baa
State          Online
ESSNet         Enabled
Volume Group   V0
os400Serial    29F
NVS Memory     8.0 GB
Cache Memory   238.3 GB
Processor Memory 253.4 GB
MTS            IBM.2421-75TV180
numegsupported 0

```

2. Obtain your license activation codes from the IBM DSFA website, as discussed in 10.2.2, "Obtaining activation codes" on page 257.
3. Use the **applykey** command to activate the codes and the **lskey** command to verify which type of licensed features are activated for your storage unit.
 - c. Enter an **applykey** command at the dscli command prompt as follows. The -file parameter specifies the key file. The second parameter specifies the storage image.


```
dscli> applykey -file c:\2107_7520780.xml IBM.2107-7520781
```
 - d. Verify that the keys have been activated for your storage unit by issuing the DS CLI **lskey** command, as shown in Example 10-2.

Example 10-2 Using lskey to list installed licenses

```

dscli> lskey ibm.2107-7520781
Date/Time: 05 October 2010 14:19:17 CET IBM DSCLI Version: 6.6.0.220 DS:
ibm.2107-7520781

```

Activation Key	Authorization Level (TB)	Scope
Global mirror (GM)	70	FB
High Performance FICON for System z (zHPF)	on	CKD
IBM FlashCopy SE	100	All
IBM HyperPAV	on	CKD
IBM database protection	on	FB
Metro mirror (MM)	70	FB
Metro/Global mirror (MGM)	70	FB
Operating environment (OEL)	100	All
Parallel access volumes (PAV)	30	CKD
Point in time copy (PTC)	100	All
RMZ Resync	30	CKD
Remote mirror for z/OS (RMZ)	30	CKD

For more details about the DS CLI, refer to *IBM System Storage DS: Command-Line Interface User's Guide*, GC53-1127.

10.3 Licensed scope considerations

For the Point-in-Time Copy (PTC) function and the Remote Mirror and Copy functions, you have the ability to set the scope of these functions to be FB, CKD, or All. You need to decide what scope to set, as shown in Figure 10-8 on page 260. In that example, Image One has 16 TB of RMC, and the user has currently decided to set the scope to All. If the scope was set to FB instead, then you cannot use RMC with any CKD volumes that are later configured. However, it is possible to return to the DSFA website at a later time and change the scope

from CKD or FB to All, or from All to either CKD or FB. In every case, a new activation code is generated, which you can download and apply.

10.3.1 Why you get a choice

Let us imagine a simple scenario where a machine has 20 TB of capacity. Of this capacity, 15 TB is configured as FB and 5 TB is configured as CKD. If we only want to use Point-in-Time Copy for the CKD volumes, then we can purchase just 5 TB of Point-in-Time Copy and set the scope of the Point-in-Time Copy activation code to CKD. There is no need to buy a new PTC license in case you do not need Point-in-Time Copy for CKD anymore, but you would like to use it for FB only. Simply obtain a new activation code from DSFA website by changing the scope to FB.

When deciding which scope to set, there are several scenarios to consider. Use Table 10-6 to guide you in your choice. This table applies to both Point-in-Time Copy and Remote Mirror and Copy functions.

Table 10-6 Deciding which scope to use

Scenario	Point-in-Time Copy or Remote Mirror and Copy function usage consideration	Suggested scope setting
1	This function is only used by open systems hosts.	Select FB.
2	This function is only used by System z hosts.	Select CKD.
3	This function is used by both open systems and System z hosts.	Select All.
4	This function is currently only needed by open systems hosts, but we might use it for System z at some point in the future.	Select FB and change to scope All if and when the System z requirement occurs.
5	This function is currently only needed by System z hosts, but we might use it for open systems hosts at some point in the future.	Select CKD and change to scope All if and when the open systems requirement occurs.
6	This function has already been set to All.	Leave the scope set to All. Changing the scope to CKD or FB at this point requires a disruptive outage.

Any scenario that changes from FB or CKD to All does not require an outage. If you choose to change from All to either CKD or FB, then you must have a disruptive outage. If you are absolutely certain that your machine will only ever be used for one storage type (for example, only CKD or only FB), then you can also quite safely just use the All scope.

10.3.2 Using a feature for which you are not licensed

In Example 10-3 on page 268, we have a machine where the scope of the Point-in-Time Copy license is set to FB. This means we cannot use Point-in-Time Copy to create CKD FlashCopies. When we try, the command fails. We can, however, create CKD volumes, because the Operating Environment License (OEL) key scope is All.

Example 10-3 Trying to use a feature for which you are not licensed

```
dscli> lskey IBM.2107-7520391
Date/Time: 05 October 2010 14:19:17 CET IBM DSCLI Version: 6.6.0.220 DS: IBM.2107-7520391
Activation Key          Authorization Level (TB) Scope
=====
Metro mirror (MM)      5                      All
Operating environment (OEL) 5                      All
Point in time copy (PTC) 5                      FB
The FlashCopy scope is currently set to FB.

dscli> lsckdvol
Date/Time: 05 October 2010 14:19:17 CET IBM DSCLI Version: 6.6.0.220 DS: IBM.2107-7520391
Name  ID  accstate datastate configstate deviceMTM voltype  orgbvols extpool cap (cyl)
=====
-     0000 Online   Normal   Normal   3390-3   CKD Base -      P2      3339
-     0001 Online   Normal   Normal   3390-3   CKD Base -      P2      3339

dscli> mkflash 0000:0001          We are not able to create CKD FlashCopies
Date/Time: 05 October 2009 14:19:17 CET IBM DSCLI Version: 6.5.0.220 DS: IBM.2107-7520391
CMUN03035E mkflash: 0000:0001: Copy Services operation failure: feature not installed
```

10.3.3 Changing the scope to All

As a follow-on to the previous example, in Example 10-4 we have logged onto DSFA and changed the scope for the PTC license to All. We then apply this new activation code. We are now able to perform a CKD FlashCopy.

Example 10-4 Changing the scope from FB to All

```
dscli> lskey IBM.2107-7520391
Date/Time: 05 October 2010 14:19:17 CET IBM DSCLI Version: 6.6.0.220 DS: IBM.2107-7520391
Activation Key          Authorization Level (TB) Scope
=====
Metro mirror (MM)      5                      All
Operating environment (OEL) 5                      All
Point in time copy (PTC) 5                      FB
The FlashCopy scope is currently set to FB

dscli> applykey -key 1234-5678-9FEF-C232-51A7-429C-1234-5678 IBM.2107-7520391
Date/Time: 05 October 2010 14:19:17 CET IBM DSCLI Version: 6.6.0.220 DS: IBM.2107-7520391
CMUC00199I applykey: Licensed Machine Code successfully applied to storage image
IBM.2107-7520391.

dscli> lskey IBM.2107-7520391
Date/Time: 05 October 2010 14:19:17 CET IBM DSCLI Version: 6.6.0.220 DS: IBM.2107-7520391
Activation Key          Authorization Level (TB) Scope
=====
Metro mirror (MM)      5                      All
Operating environment (OEL) 5                      All
Point in time copy (PTC) 5                      All
The FlashCopy scope is now set to All

dscli> lsckdvol
Date/Time: 04 October 2010 15:51:53 CET IBM DSCLI Version: 6.6.0.220 DS: IBM.2107-7520391
Name  ID  accstate datastate configstate deviceMTM voltype  orgbvols extpool cap (cyl)
=====
-     0000 Online   Normal   Normal   3390-3   CKD Base -      P2      3339
-     0001 Online   Normal   Normal   3390-3   CKD Base -      P2      3339
```

```

dscli> mkflash 0000:0001           We are now able to create CKD FlashCopies
Date/Time: 05 October 2010 14:19:17 CET IBM DSCLI Version: 6.6.0.220 DS: IBM.2107-7520391
CMUC00137I mkflash: FlashCopy pair 0000:0001 successfully created.

```

10.3.4 Changing the scope from All to FB

In Example 10-5, we decide to increase storage capacity for the entire machine. However, we do not want to purchase any more PTC licenses, because PTC is only used by open systems hosts and this new capacity is only to be used for CKD storage. We therefore decide to change the scope to FB, so we log on to the DSFA website and create a new activation code. We then apply it, but discover that because this is effectively a downward change (decreasing the scope), it does not apply until we have a disruptive outage on the DS8000.

Example 10-5 Changing the scope from All to FB

```

dscli> lskey IBM.2107-7520391
Date/Time: 05 October 2010 14:19:17 CET IBM DSCLI Version: 6.6.0.220 DS: IBM.2107-7520391
Activation Key          Authorization Level (TB) Scope
=====
Metro mirror (MM)      5                      All
Operating environment (OEL) 5                      All
Point in time copy (PTC) 5                      All
The FlashCopy scope is currently set to All

```

```

dscli> applykey -key ABCD-EFAB-EF9E-6B30-51A7-429C-1234-5678 IBM.2107-7520391
Date/Time: 05 October 2010 14:19:17 CET IBM DSCLI Version: 6.6.0.220 DS: IBM.2107-7520391
CMUC00199I applykey: Licensed Machine Code successfully applied to storage image
IBM.2107-7520391.

```

```

dscli> lskey IBM.2107-7520391
Date/Time: 05 October 2010 14:19:17 CET IBM DSCLI Version: 6.6.0.220 DS: IBM.2107-7520391
Activation Key          Authorization Level (TB) Scope
=====
Metro mirror (MM)      5                      All
Operating environment (OEL) 5                      All
Point in time copy (PTC) 5                      FB
The FlashCopy scope is now set to FB

```

```

dscli> lsckdvol
Date/Time: 05 October 2010 14:19:17 CET IBM DSCLI Version: 6.6.0.220 DS: IBM.2107-7520391
Name  ID  accstate  datastate  configstate  deviceMTM  voltype  orgbvols  extpool  cap (cyl)
=====
-      0000 Online    Normal    Normal      3390-3     CKD Base  -        P2        3339
-      0001 Online    Normal    Normal      3390-3     CKD Base  -        P2        3339

```

```

dscli> mkflash 0000:0001           But we are still able to create CKD FlashCopies
Date/Time: 05 October 2010 14:19:17 CET IBM DSCLI Version: 6.6.0.220 DS: IBM.2107-7520391
CMUC00137I mkflash: FlashCopy pair 0000:0001 successfully created.

```

In this scenario, we have made a downward license feature key change. We must schedule an outage of the storage image. We should in fact only make the downward license key change immediately before taking this outage.

Consideration: Making a downward license change and then not immediately performing a reboot of the storage image is not supported. Do not allow your machine to be in a position where the applied key is different than the reported key.

10.3.5 Applying an insufficient license feature key

In this example, we have a scenario where a DS8000 has a 5 TB Operating Environment License (OEL), FlashCopy (PTC), and Metro Mirror (MM) license. We increased storage capacity and therefore increased the license key for OEL and MM. However, we forgot to increase the license key for FlashCopy (PTC). In Example 10-6, we can see the FlashCopy license is only 5 TB. However, we are still able to create FlashCopies.

Example 10-6 Insufficient FlashCopy license

```
dscli> lskey IBM.2107-7520391
Date/Time: 05 October 2010 14:19:17 CET IBM DSCLI Version: 6.6.0.220 DS:
IBM.2107-7520391
Activation Key          Authorization Level (TB) Scope
=====
Metro mirror (MM)      10                      All
Operating environment (OEL) 10                      All
Point in time copy (PTC)  5                      All

dscli> mkflash 1800:1801
Date/Time: 04 October 2010 17:46:14 CET IBM DSCLI Version: 6.6.0.220 DS:
IBM.2107-7520391
CMUC00137I mkflash: FlashCopy pair 1800:1801 successfully created.
```

At this point, this is still a valid configuration, because the configured ranks on the machine total less than 5 TB of storage. In Example 10-7, we then try to create a new rank that brings the total rank capacity above 5 TB. This command fails.

Example 10-7 Creating a rank when we are exceeding a license key

```
dscli> mkrank -array A1 -stgtype CKD
Date/Time: 05 October 2010 14:19:17 CET IBM DSCLI Version: 6.6.0.220 DS:
IBM.2107-7520391
CMUN02403E mkrank: Unable to create rank: licensed storage amount has been
exceeded
```

To configure the additional ranks, we must first increase the license key capacity of every installed license. In this example, that is the FlashCopy license.

10.3.6 Calculating how much capacity is used for CKD or FB

To calculate how much disk space is currently used for CKD or FB storage, we need to combine the output of two commands. There are some simple rules:

- ▶ License key values are decimal numbers. So, 5 TB of license is 5000 GB.
- ▶ License calculations use the disk size number shown by the **lsarray** command.
- ▶ License calculations include the capacity of all DDMs in each array site.
- ▶ Each array site is eight DDMs.

To make the calculation, we use the **lsrank** command to determine how many arrays the rank contains, and whether those ranks are used for FB or CKD storage. We use the **lsarray** command to obtain the disk size used by each array. Then, we multiply the disk size (146, 300, 450, or 600) by eight (for eight DDMs in each array site).

In Example 10-8 on page 271, **lsrank** tells us that rank R0 uses array A0 for CKD storage. Then, **lsarray** tells us that array A0 uses 300 GB DDMs. So we multiple 300 (the DDM size) by 8, giving us $300 \times 8 = 2400$ GB. This means we are using 2400 GB for CKD storage.

Now, rank R4 in Example 10-8 is based on array A6. Array A6 uses 146 GB DDMs, so we multiply 146 by 8, giving us $146 \times 8 = 1168$ GB. This means we are using 1168 GB for FB storage.

Example 10-8 Displaying array site and rank usage

```
dscli> lsrank
Date/Time: 05 October 2010 14:19:17 CET IBM DSCLI Version: 6.6.0.220 DS:
IBM.2107-75ABTV1
ID Group State  datastate Array RAIDtype extpoolID stgtype
=====
R0      0 Normal Normal    A0          5 P0        ckd
R4      0 Normal Normal    A6          5 P4        fb
```

```
dscli> lsarray
Date/Time: 05 October 2010 14:19:17 CET IBM DSCLI Version: 6.6.0.220 DS:
IBM.2107-75ABTV1
Array State      Data  RAIDtype  arsite Rank DA Pair DDMcap (10^9B)
=====
A0  Assigned  Normal 5 (6+P+S) S1      R0  0          300.0
A1  Unassigned Normal 5 (6+P+S) S2      -  0          300.0
A2  Unassigned Normal 5 (6+P+S) S3      -  0          300.0
A3  Unassigned Normal 5 (6+P+S) S4      -  0          300.0
A4  Unassigned Normal 5 (7+P)  S5      -  0          146.0
A5  Unassigned Normal 5 (7+P)  S6      -  0          146.0
A6  Assigned   Normal 5 (7+P)  S7      R4  0          146.0
A7  Assigned   Normal 5 (7+P)  S8      R5  0          146.0
```

So for CKD scope licenses, we currently use 2,400 GB. For FB scope licenses, we currently use 1168 GB. For licenses with a scope of All, we currently use 3568 GB. Using the limits shown in Example 10-6 on page 270, we are within scope for all licenses.

If we combine Example 10-6 on page 270, Example 10-7 on page 270, and Example 10-8, we can also see why the **mkrank** command in Example 10-7 on page 270 failed. In Example 10-7 on page 270, we tried to create a rank using array A1. Now, array A1 uses 300 GB DDMs. This means that for FB scope and All scope licenses, we use $300 \times 8 = 2400$ GB more license keys.

In Example 10-6 on page 270, we had only 5 TB of FlashCopy license with a scope of All. This means that we cannot have total configured capacity that exceeds 5000 TB. Because we already use 3568 GB, the attempt to use 2400 more GB will fail, because 3568 plus 2400 equals 5968 GB, which is more than 5000 GB. If we increase the size of the FlashCopy license to 10 TB, then we can have 10,000 GB of total configured capacity, so the rank creation will then succeed.



Part 3

Storage configuration

In this part, we discuss the configuration tasks required on your IBM System Storage DS8700 and DS8800. We cover the following topics:

- ▶ System Storage Productivity Center (SSPC)
- ▶ Configuration using the DS Storage Manager GUI
- ▶ Configuration with the DS Command-Line Interface



Configuration flow

This chapter gives a brief overview of the tasks required to configure the storage in a IBM System Storage DS8700 or DS8800.

11.1 Configuration worksheets

During the installation of the DS8700 or DS8800, your IBM service representative customizes the setup of your storage complex based on information that you provide in a set of customization worksheets. Each time that you install a new storage unit or management console, you must complete the customization worksheets before the IBM service representatives can perform the installation.

The customization worksheets are important and need to be completed before the installation. It is important that this information is entered into the machine so that preventive maintenance and high availability of the machine are maintained. You can find the customization worksheets in *IBM System Storage DS8700 and DS8800 Introduction and Planning Guide*, GC27-2297.

The customization worksheets allow you to specify the initial setup for the following items:

- ▶ **Company information:** This information allows IBM service representatives to contact you as quickly as possible when they need to access your storage complex.
- ▶ **Management console network settings:** This allows you to specify the IP address and LAN settings for your management console (MC).
- ▶ **Remote support (includes Call Home and remote service settings):** This allows you to specify whether you want outbound (Call Home) or inbound (remote services) remote support.
- ▶ **Notifications (include SNMP trap and email notification settings):** This allows you to specify the types of notifications that you want and that others might want to receive.
- ▶ **Power control:** This allows you to select and control the various power modes for the storage complex.
- ▶ **Control Switch settings:** This allows you to specify certain DS8800 settings that affect host connectivity. You need to enter these choices on the control switch settings worksheet so that the service representative can set them during the installation of the DS8800.

Important: IBM service representatives cannot install a storage unit or management console until you provide them with the completed customization worksheets.

11.2 Configuration flow

The following list shows the tasks that need to be done when configuring storage in the DS8700. The order of the tasks does not have to be exactly as shown here, and some of the individual tasks can be done in a different order.

Important: The configuration flow changes when you use the Full Disk Encryption Feature for the DS87000 or DS8800. For details, refer to *IBM System Storage DS8000: Disk Encryption Implementation and Usage Guidelines*, REDP-4500.

1. **Install license keys:** Activate the license keys for the storage unit.
2. **Create arrays:** Configure the installed disk drives as either RAID 5, RAID 6, or RAID 10 arrays.
3. **Create ranks:** Assign each array to either a fixed block (FB) rank or a count key data (CKD) rank.

4. Create Extent Pools: Define Extent Pools, associate each one with either Server 0 or Server 1, and assign at least one rank to each Extent Pool. If you want to take advantage of Storage Pool Striping, you must assign multiple ranks to an Extent Pool. With current versions of the DS GUI, you can start directly with the creation of Extent Pools (arrays and ranks will be automatically and implicitly defined).
5. If you plan on using Easy-Tier, in particular in automatic mode, plan for hybrid extent pools.
6. Create a repository for Space Efficient volumes.
7. Configure I/O ports: Define the type of the Fibre Channel/FICON ports. The port type can be either Switched Fabric, Arbitrated Loop, or FICON.
8. Create host connections for open systems: Define open systems hosts and their Fibre Channel (FC) host bus adapter (HBA) worldwide port names.
9. Create volume groups for open systems: Create volume groups where FB volumes will be assigned and select the host attachments for the volume groups.
10. Create open systems volumes: Create striped open systems FB volumes and assign them to one or more volume groups.
11. Create System z logical control units (LCUs): Define their type and other attributes, such as subsystem identifiers (SSIDs).
12. Create striped System z volumes: Create System z CKD base volumes and Parallel Access Volume (PAV) aliases for them.

The actual configuration can be done using either the DS Storage Manager GUI or DS Command-Line Interface, or a mixture of both. A novice user might prefer to use the GUI, whereas a more experienced user might use the CLI, particularly for the more repetitive tasks, such as creating large numbers of volumes.

For a more detailed discussion about how to perform the specific tasks, refer to:

- ▶ Chapter 10, “IBM System Storage DS8000 features and license keys” on page 251
- ▶ Chapter 13, “Configuration using the DS Storage Manager GUI” on page 309
- ▶ Chapter 14, “Configuration with the DS Command-Line Interface” on page 363

General guidelines when configuring storage

Remember the following general guidelines when configuring storage on the DS8800:

- ▶ To achieve a well-balanced load distribution, use at least two Extent Pools, each assigned to one DS8800 internal server (extent Pool 0 and Extent Pool 1). If CKD and FB volumes are required, use at least four Extent Pools.
- ▶ Address groups (16 LCUs/logical subsystems (LSSs)) are all for CKD or all for FB.
- ▶ Volumes of one LCU/LSS can be allocated on multiple Extent Pools.
- ▶ An Extent Pool cannot contain all three RAID 5, RAID 6, and RAID 10 ranks. Each Extent Pool pair should have the same characteristics in terms of RAID type, RPM, and DDM size (Note that there are exceptions for SSDs and SATA drives).
- ▶ Ranks in one Extent Pool should belong to separate Device Adapters.
- ▶ Assign multiple ranks to Extent Pools to take advantage of Storage Pool Striping.
- ▶ CKD:
 - 3380 and 3390 type volumes can be intermixed in an LCU and an Extent Pool.

- ▶ FB:
 - Create a volume group for each server unless LUN sharing is required.
 - Place all ports for a single server in one volume group.
 - If LUN sharing is required, there are two options:
 - Use separate volumes for servers and place LUNs in multiple volume groups.
 - Place servers (clusters) and volumes to be shared in a single volume group.
- ▶ I/O ports:
 - Distribute host connections of each type (FICON and FCP) evenly across the I/O enclosure.
 - Typically, access *any* is used for I/O ports with access to ports controlled by SAN zoning.



System Storage Productivity Center

This chapter discusses how to set up and manage the System Storage Productivity Center (SSPC) to work with the IBM System Storage DS8000 series.

The chapter covers the following topics:

- ▶ System Storage Productivity Center overview
- ▶ System Storage Productivity Center components
- ▶ System Storage Productivity Center setup and configuration
- ▶ Working with a DS8000 system and Tivoli Storage Productivity Center Basic Edition

12.1 System Storage Productivity Center overview

The IBM System Storage Productivity Center (SSPC) is a hardware appliance with pre-installed software that can help you improve and centralize the management of your storage environment through the integration of products. It provides a single point of management integrating the functionality of the IBM Tivoli Storage Productivity Center with storage devices and element managers in an easy-to-use user interface.

Tips:

- ▶ TPC-Basic Edition is required for DS8000 to access the DS8000 GUI.
- ▶ SSPC (2805-MC5) is an optional hardware feature for all systems running the DS8000 R6.0 level microcode or later, and provides a convenient ordering option for IBM Tivoli Productivity Center Basic Edition.

Architecturally, System Storage Productivity Center is a 1U, rack-mounted hardware appliance that consists of an IBM Machine Type 2805 Model MC5 server that is preinstalled with IBM Tivoli Storage Productivity Center Basic Edition software on the Microsoft Windows Server 2008 R2 Standard operating system for 64-bit processors. Also installed on the SSPC appliance is Tivoli Storage Productivity Center for Replication, an application that offers an interface to Copy Services. See Figure 12-1 on page 281 for an overview of the SSPC architecture.

With the System Storage Productivity Center, the IBM DS8000 Storage Manager Interface is accessible from the Tivoli Productivity Center GUI. IBM Tivoli Storage Productivity Center provides a DS8000 element manager window, which allows you to add and manage multiple DS8000 element managers from one console. See 12.3.4, “Configuring the TPC Element Manager to access the DS8000 GUI” on page 301.

IBM System Storage Productivity Center simplifies storage management by:

- ▶ Centralizing the management of storage network resources with IBM storage management software.
- ▶ Providing greater synergy between storage management software and IBM storage devices.
- ▶ Reducing the number of servers that are required to manage the storage infrastructure.
- ▶ Providing a simple migration path from basic device management to using storage management applications that provide higher-level functions.

12.1.1 SSPC components

SSPC is a solution consisting of hardware and software elements.

SSPC hardware

The SSPC (IBM model 2805-MC5) server contains the following hardware components:

- ▶ x86 server 1U rack installed
- ▶ Intel® Quadcore Xeon processor running at 2.53 GHz
- ▶ 8 GB of RAM
- ▶ Two mirrored hard disk drives
- ▶ Dual port Gigabit Ethernet

Optional components are:

- ▶ KVM Unit
- ▶ Secondary power supply
- ▶ Additional hard disk drives
- ▶ CD media to recover image for 2805-MC5
- ▶ 8 Gb Fibre Channel Dual Port HBA, which allows you to move the Tivoli Storage Productivity Center database from the SSPC server to the IBM System Storage DS8000

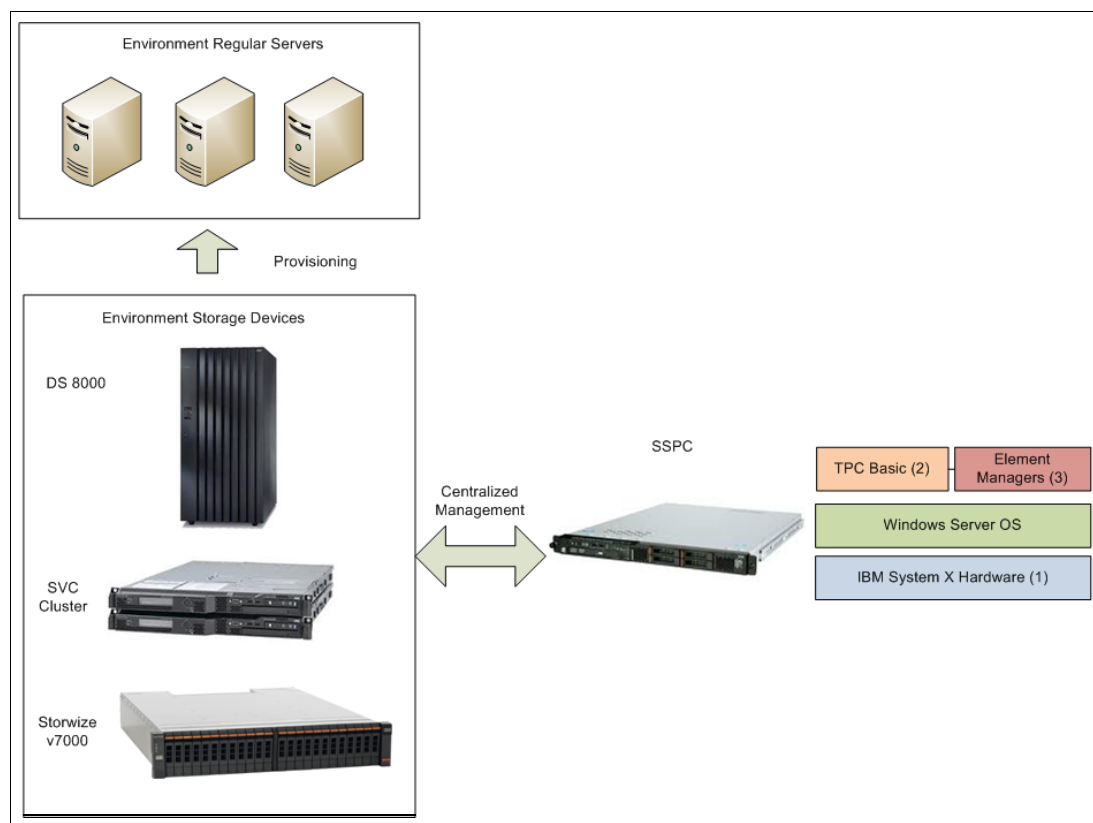


Figure 12-1 SSPC Basic Architecture

SSPC software

The IBM System Storage Productivity Center 1.5 includes the following preinstalled (separately purchased) software, running under a licensed Microsoft Windows Server 2008 Enterprise Edition R2, 64 bit (included):

- ▶ IBM Tivoli Storage Productivity Center V4.2.1 licensed as TPC Basic Edition (includes the Tivoli Integrated Portal). A TPC upgrade requires that you purchase and add additional TPC licenses.
- ▶ DS CIM Agent Command-Line Interface (DSCIMCLI) 5.5.
- ▶ IBM Tivoli Storage Productivity Center for Replication (TPC-R) V4.2.1. To run TPC-R on SSPC, you must purchase and add TPC-R base license for FlashCopy.
- ▶ IBM DB2 Enterprise Server Edition 9.7 64-bit Enterprise.
- ▶ IBM JAVA 1.6 is preinstalled. You do not need to download Java from Sun Microsystems.

Optionally, the following components can be installed on the SSPC:

- ▶ DS8000 Command-Line Interface (DSCLI).
- ▶ Antivirus software.

Customers have the option to purchase and install the individual software components to create their own SSPC server.

12.1.2 SSPC capabilities

The complete SSPC offers the following capabilities:

- ▶ Preinstalled and tested console: IBM has designed and tested SSPC to support interoperability between server, software, and supported storage devices.
- ▶ IBM System Storage DS8000 GUI integration: With Tivoli Storage Productivity Center v4.2.1, the DS Storage Manager GUI for the DS8000 is integrated with Tivoli Storage Productivity Center for remote web access.
- ▶ Automated device discovery: DS8000 and SVC storage devices can be automatically discovered and configured into Tivoli Productivity Center environments. These devices are displayed in the Tivoli Productivity Center through a storage topology viewer.
- ▶ Asset and capacity reporting: Tivoli Storage Productivity Center collects asset and capacity information from storage devices in the SAN, which can be kept for historical reporting, forecasting, and used for other tasks such as analysis and provisioning.
- ▶ Advanced Topology Viewer: Provides a linked graphical and detailed view of the overall SAN, including device relationships and visual notifications.
- ▶ A status dashboard.

12.1.3 SSPC upgrade options

You can upgrade some of the software included with the standard SSPC.

Tivoli Storage Productivity Center Standard Edition

Tivoli Storage Productivity Center Basic Edition (TPC-BE) can be easily upgraded with one or all of the advanced capabilities found in IBM Tivoli Storage Productivity Center Standard Edition (TPC-SE) by purchasing the TPC-SE License. TPC-SE includes IBM Tivoli Storage Productivity Center for Disk, IBM Tivoli Storage Productivity Center for Data, and IBM Tivoli Storage Productivity Center for Fabric.

TPC-SE offers the following capabilities:

- ▶ Device configuration and management of SAN-attached devices from a single console. It also allows users to gather and analyze historical and near real-time performance metrics.
- ▶ Management of file systems and databases, thus enabling enterprise-wide reports, monitoring and alerts, policy-based action, and file system capacity automation in heterogeneous environments.
- ▶ Management, monitoring, and control of SAN fabric to help automate device discovery, topology rendering, error detection fault isolation, SAN error predictor, zone control, real-time monitoring and alerts, and event management for heterogeneous enterprise SAN environments. In addition, it allows collection of performance statistics from IBM Tivoli Storage, Brocade, Cisco, and McDATA fabric switches and directors that implement the SNIA SMI-S specification.

For a detailed comparison of TPC-BE versus TPC-SE refer to *IBM System Storage Productivity Center Deployment Guide*, SG24-7560-01 at:

<http://www.redbooks.ibm.com/cgi-bin/searchsite.cgi?query=SG24-7560-01&SearchOrder=1&SearchFuzzy=FALSE>

Tivoli Storage Productivity Center for Replication

SSPC v1.2 and later come with TPC-R preinstalled. To use the preinstalled TPC-R, the basic license that will allow FlashCopy needs to be applied first to TPC-R on the SSPC. IBM Tivoli Storage Productivity Center for Replication provides management of IBM copy services capabilities. Refer to *IBM System Storage DS8000: Copy Services for Open Environments*, SG24-6788, for more details.

IBM Tivoli Storage Productivity Center for Replication is designed to:

- ▶ Simplify and improve the configuration and management of replication on your network storage devices by performing advanced copy operations in a single action.
- ▶ Manage advanced storage replication services, such as Metro Mirror, Global Mirror, Metro/Global Mirror, FlashCopy, and FlashCopy SE. TPC-R can also monitor copy services.
- ▶ Enable multiple pairing options for source and target volumes.
- ▶ Define session pairs using target and source volume groups, confirm path definitions, and create consistency sets for replication operations.
- ▶ IBM Tivoli Storage Productivity Center for Replication Three Site BC is an addition to the Tivoli Storage Productivity Center for Replication V4 family. Three Site BC provides:
 - Support for three-site IBM DS8000 family Metro Global Mirror configurations.
 - Disaster recovery configurations that can be set up to indicate copy service type (FlashCopy, Metro Mirror, Global Mirror) and the number of separate copies and sites to be set.
- ▶ Manage Open HyperSwap, high availability solution for storage systems based on Metro Mirror. For more information, refer to *IBM System Storage DS8000: Copy Services for Open Systems*, SG24-6788, or the IBM Tivoli Storage Productivity Center V4.2 - Information Center at:

http://publib.boulder.ibm.com/infocenter/tivihelp/v4r1/index.jsp?topic=/com.ibm.tpc_v42.doc/frg_t_manage_hs.html

12.2 SSPC setup and configuration

This section summarizes the tasks and sequence of steps required to set up and configure the DS8000 system, and the SSPC used to manage the DS8000 system.

For detailed information, and additional considerations, see the TPC/SSPC Information Center at the following addresses

<http://publib.boulder.ibm.com/infocenter/tivihelp/v4r1/index.jsp>

You can also find detailed information about SSPC functions and considerations in *IBM System Storage Productivity Center Deployment Guide*, SG24-7560-01 at:

<http://www.redbooks.ibm.com/cgi-bin/searchsite.cgi?query=SG24-7560-01&SearchOrder=1&SearchFuzzy=FALSE>

12.2.1 DS8000 Storage Management

DS8000 series systems purchased after October 23, 2007 will need Tivoli Storage Productivity Center to access the DS8000 GUI from any other system than the HMC.

In IBM Tivoli Productivity Center 4.2, the Enterprise Storage Server Network Interface (ESSNI) replaces the CIM agent as the interface for the Tivoli Storage Productivity Center to communicate with the DS8000.

Tip: You will need to install or upgrade an existing installation of TPC 4.2.1 to TPC 4.2.1 Fixpack 3 to access the latest DS8000 R6.1 DS GUI. Fixpack 3 is found in support/downloads at:

<http://publib.boulder.ibm.com/infocenter/tivihelp/v4r1/index.jsp>

Tivoli Storage Productivity Center Basic Edition V4.2.1 comes pre-installed on the SSPC appliance. SSPC was specifically designed and tested with DS8000 and DS customers in mind. To avoid possible installation delays and potential technical issues, order SSPC if you do not have an instance of Tivoli Storage Productivity Center.

If you already have Tivoli Storage Productivity Center installed and configured in your environment, you can use your existing Tivoli Storage Productivity Center server and Tivoli Storage Productivity Center software license to perform the remote configuration of new DS8000s using the GUI. Of course, your Tivoli Storage Productivity Center installation must have IP connectivity to the new DS8000 that it will manage. To do so, you need to upgrade your Tivoli Storage Productivity Center to V4.2.1.

12.2.2 Configuring SSPC for DS8000 remote GUI access

The steps to configure the SSPC can be found in the *System Storage Productivity Center User's Guide*, SC27-2336, which is shipped with the SSPC. The document is also available at the following address:

<http://publib.boulder.ibm.com/infocenter/tivihelp/v4r1/index.jsp>

After IBM Support physically installs the 2805-MC5 and tests IP connectivity with the DS8000, SSPC is ready for configuration by either the client or IBM Services.

After the initial configuration of SSPC is done, the SSPC user is able to configure the remote GUI access to all DS8000 systems in the TPC Element Manager, as described in 12.3.4, "Configuring the TPC Element Manager to access the DS8000 GUI" on page 301.

12.2.3 Preparing your browser

Perform the following steps to prepare your web browser so that you can configure SSPC for your DS8000s.

Configure your browser to allow pop-up windows

- For Firefox, from the Menu bar:
 - a. Click **Tools** → **Options** → **Content** → **Block Pop-Up Windows**.
 - b. Select the **Block Pop-up Windows** check box.
 - c. Click **OK**.

- For Internet Explorer:

Microsoft Internet Explorer is not available from the SSPC desktop or by clicking **Start → All Programs**. To enable Internet Explorer, follow these steps:

- Click **Start**.
- Right-click Command Prompt and select **Run as administrator**.

Tip: If you are prompted for an administrator password or for confirmation, type the password or click **Continue**.

- At the command prompt, type the following command and press Enter

```
dism /online /Enable-Feature /FeatureName:Internet-Explorer-Optional-amd64
```

The string **Internet-Explorer-Optional-amd64** is case-sensitive.

- When you are prompted to restart the computer, click **Yes**.

To configure Internet Explorer to allow pop-ups, complete the following steps:

- Open Internet Explorer by clicking the Internet Explorer icon on the Quick-Launch toolbar.
- Click **Tools → Pop-up Blocker → Turn Off Pop-up Blocker**.
- If you receive a message that content was blocked because it was not signed by a valid security certificate, click the Information Bar at the top of the window and select **Show blocked content**.

- Additional web browser

If you access SSPC by using a web browser other than Firefox or Internet Explorer, use the configuration instructions for that browser to ensure that new pop-up windows can automatically open when you visit a website. Also, uninstall or turn off any applications that block or suppress pop-up windows.

12.2.4 Accessing the TPC on SSPC

The following methods can be used to access the TPC on the SSPC console:

- Access the complete SSPC:
 - Launch TPC directly at the SSPC Terminal.
 - Launch TPC using Remote Desktop to the SSPC.
- Install the TPC V4.2 GUI by using the TPC installation procedure. The GUI will then connect to the TPC running on the SSPC.

The following steps show the panels invoked when launching the TPC GUI as a Java Webstart session:

1. Launch the TPC GUI front-end as a Java Webstart session.
 - In a browser, enter `http://<SSPC_ipaddress>:9550/ITSRM/app/welcome.html`.
 - Download the correct IBM Java version if it is not installed yet.
 - Select **TPC GUI** and open it with the Java Webstart executable.

For the initial setup, a Java Webstart session will be opened and the `TSRMGUI.jar` file will be offloaded to the workstation on which the browser was started. After the user agrees to unrestricted access to the workstation for the TPC-GUI, the system asks if a shortcut should be created.

Figure 12-2 shows the entry window for installing TPC GUI access through a browser.

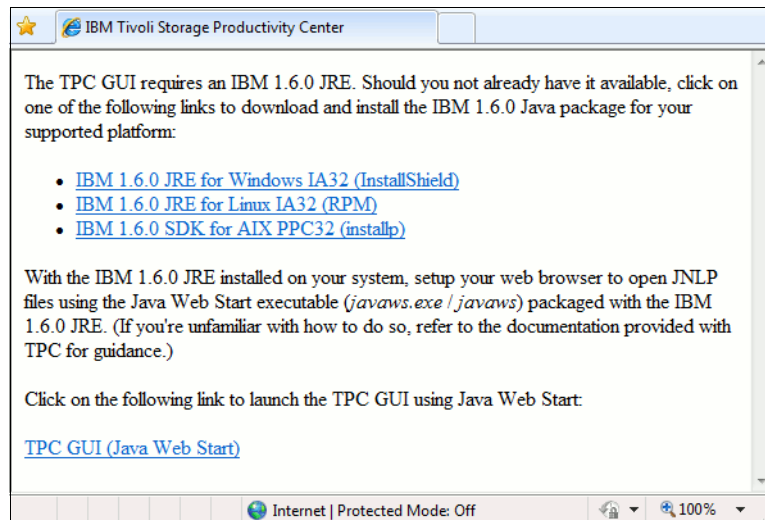


Figure 12-2 Entry window for installing TPC-GUI access through a browser

2. After you click **TPC GUI (Java Web Start)**, a login window appears as shown in Figure 12-3

Tip: You can also launch the TPC GUI from the SSPC (2805-MC5) by double-clicking the Tivoli Productivity Center icon on the desktop or clicking **Start** → **Programs** → **IBM Tivoli Storage Productivity Center** → **Productivity Center**.

Log on to the Tivoli Productivity Center user interface with the Windows Administrator user ID and password. In the Server field of the Tivoli Productivity Center Sign on window, verify that the value that is displayed is the host name of the server to which you want to connect. If not, type the correct host name.

At this point you need to provide the proper credentials for login as show in Figure 12-3

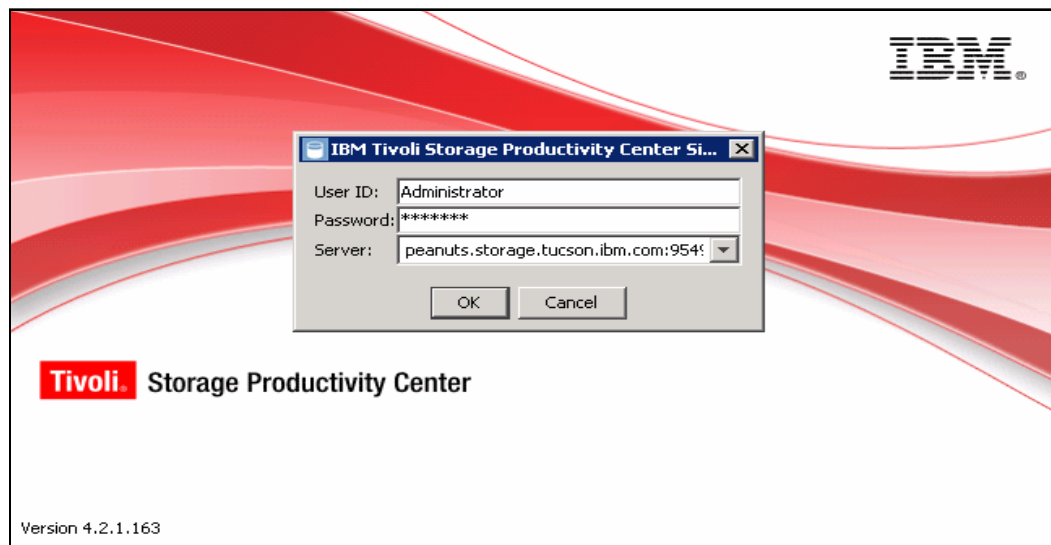


Figure 12-3 TPC GUI login window

If “Change password at first login” was specified by the SSPC administrator for the Windows user account, the user must first log on to the SSPC to change the password. The logon can be done at the SSPC terminal itself or by Windows Remote Desktop authentication to SSPC.

Change the field Server to <SSPC ipaddress>:9549 if there is no nameserver resolution between User Terminal and SSPC.

Tip: Set the SSPC IP address in the TPC-GUI entry window.

- ▶ On the SSPC server, go to:

```
%ProgramFiles%\IBM\TPC\device\apps\was\profiles\deviceServer\installedApps\Default
Node\DeviceServer.ear\DeviceServer.war\app
```

- ▶ Open the file tpcgui.jnlp and change the setting from:

```
<argument>SSPC_Name:9549</argument>
```

to:

```
<argument>SSPC_ipaddress:9549</argument>
```

3. The Welcome window to the IBM Tivoli Productivity Center panel opens if it was not disabled before this procedure, as shown in Figure 12-4.

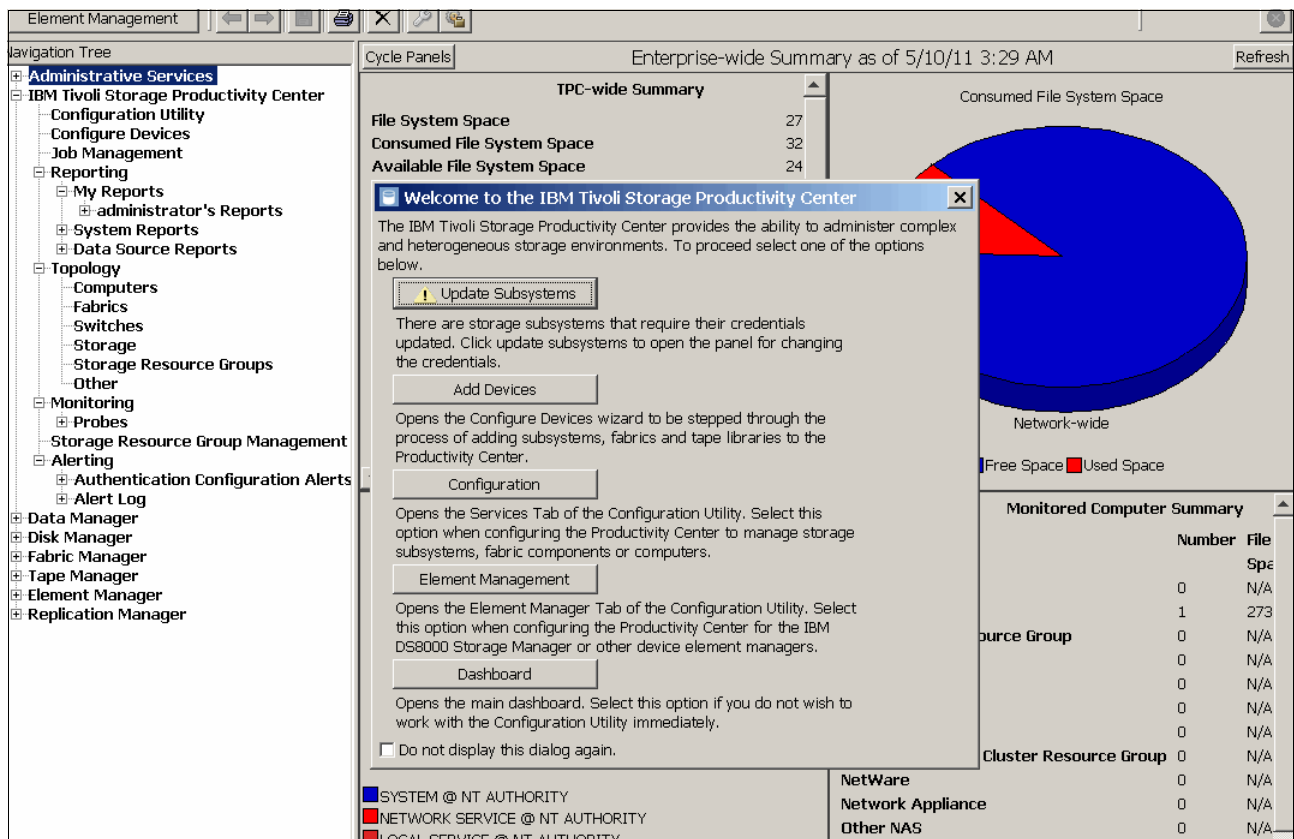


Figure 12-4 Tivoli Storage Productivity Center Welcome window

Figure 12-5 shows the Welcome to the IBM Tivoli Storage Productivity Center panel. This panel provides links to commonly used Tivoli Storage Productivity Center tasks.

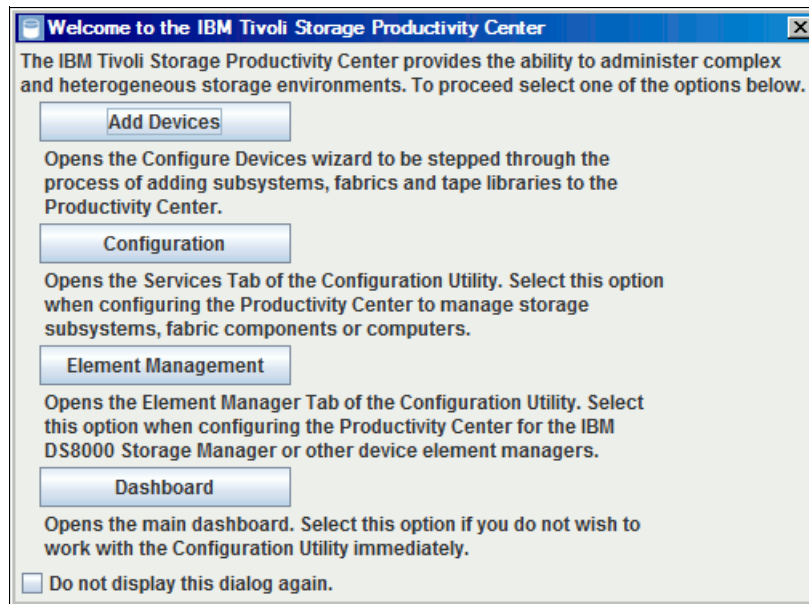


Figure 12-5 Welcome to the IBM Tivoli Storage Productivity Center Tasks panel

12.2.5 Manage embedded CIMOM on DS8000

With DS8000 and LIC Release 6, the embedded CIMOM on DS8000 HMC is enabled by default after the HMC is started. Starting in version 4.2 of IBM Tivoli Storage Productivity Center and SSPC 1.5 respectively, TPC supports ESSNI API and connects to DS8000 HMC without CIMOM.

Tip: With DS8000 and LIC Release 6.1, the DS8000 Hardware Management Console Web application can no longer be launched using SSPC. But you can still access the HMC Console through the DS8000 Web User Interface and the DSCIMCLI as described below.

There is an option to enable or disable the embedded CIMOM manually through the DS8000 HMC Web User Interface (WUI).

To enable or disable the embedded CIMOM, perform these tasks:

1. Log into the DS8000 Web User Interface (WUI) by directing your browser to `https://<DS8000 HMC IP address>`. The HMC WUI window appears.
2. Click **Log on and Launch the Hardware Management Console Web application**. The HMC welcome window opens.
3. Select **HMC Management** and under the Storage Server HMC Tasks section, click **Start/Stop CIM Agent** (Figure 12-6 on page 289).

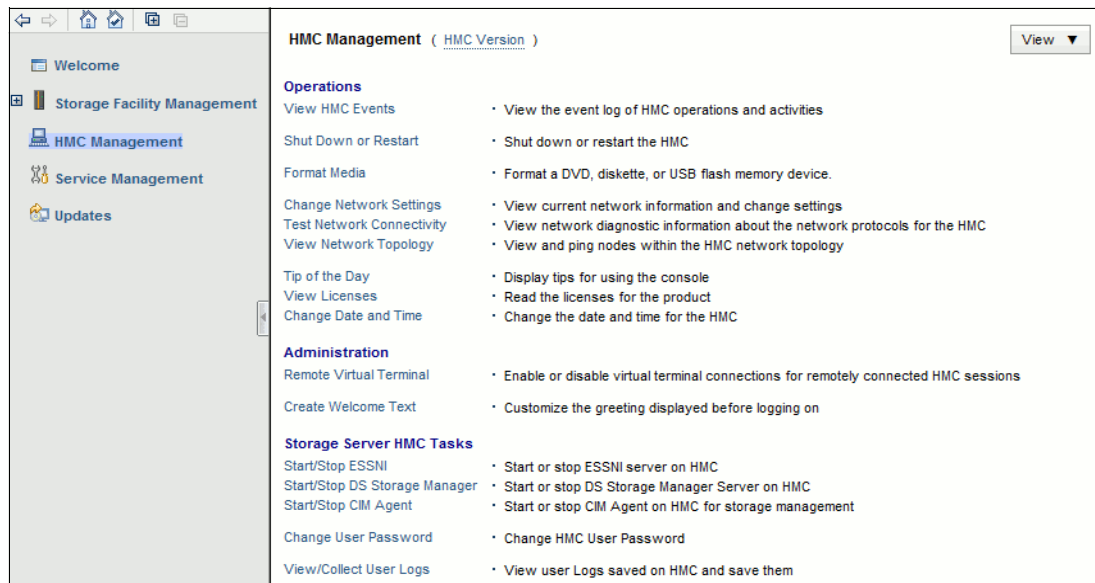


Figure 12-6 HMC WUI: Start/Stop CIM Agent

Test connectivity to DS8000 embedded CIMOM using DSCIMCLI

On the SSPC desktop, double-click the Launch DSCIMCLI icon. A DSCIMCLI command prompt window opens. At the prompt, enter the DS8000 HMC IP address and then the DS8000 Element Manager user name and password. Use the **lsdev** command shown in Example 12-1 to verify connectivity between the CIMOM agent and the primary and secondary DS8000 HMCs. Specify the correct HMC IP address/port and HMC credentials. The status of the **lsdev** command output indicates successful connection.

Example 12-1 DSCIMCLI commands to check CIMOM connectivity to primary and secondary HMC

```
> dscimcli lsdev -l -s https://9.155.70.27:6989 -u <ESSNI user> -p <ESSNI password>
Type IP          IP2          Username Storage Image Status Code Level Min Codelevel
=====
DS    9.155.70.27    -            *          IBM.2107-1305081 successful 5.4.2.540 5.1.0.309

> dscimcli lsdev -l -s https://9.155.70.28:6989 -u <ESSNI user> -p <ESSNI password>
Type IP          IP2          Username Storage Image Status Code Level Min Codelevel
=====
DS    9.155.70.28    -            *          IBM.2107-1305081 successful 5.4.2.540 5.1.0.309
```

Offload embedded CIMOM logs through DSCIMCLI

For problem determination purposes, there is an option to offload the embedded CIMOM logs to the SSPC console using DSCIMCLI commands, as shown in Example 12-2. The file will be offloaded as a compressed file to the SSPC.

Example 12-2 DSCIMCLI commands to offload DSCIMCLI logs from DS8000 HMC onto SSPC

```
C:\Program Files\IBM\DSCIMCLI\Windows> dscimcli collectlog -s
https://<<DS8800_HMC_IP_addr.>:6989 -u <valid ESSNI user> -p <associated ESSNI
password>
Old remote log files were successfully listed.
No one old log file on the DS Agent side.
New remote log file was successfully created.
```

getting log file dscim-logs-2009.3.1-16.57.17.zip from DS Agent: complete 100%
Local log file was successfully created and saved as C:\Program
Files\IBM\DSCIMCLI\WINDOWS\dscim-logs-2009.3.1-16.57.17.zip.
The new created log file dscim-logs-2009.3.1-16.57.17.zip was successfully got
from DS Agent side.
The new created log file dscim-logs-2009.3.1-16.57.17.zip was successfully deleted
on DS Agent
side

Set up users at the OS level

To set up a new SSPC user, the SSPC administrator needs to first grant appropriate user permissions at the operating system level, using the following steps, which are also illustrated in Figure 12-7.

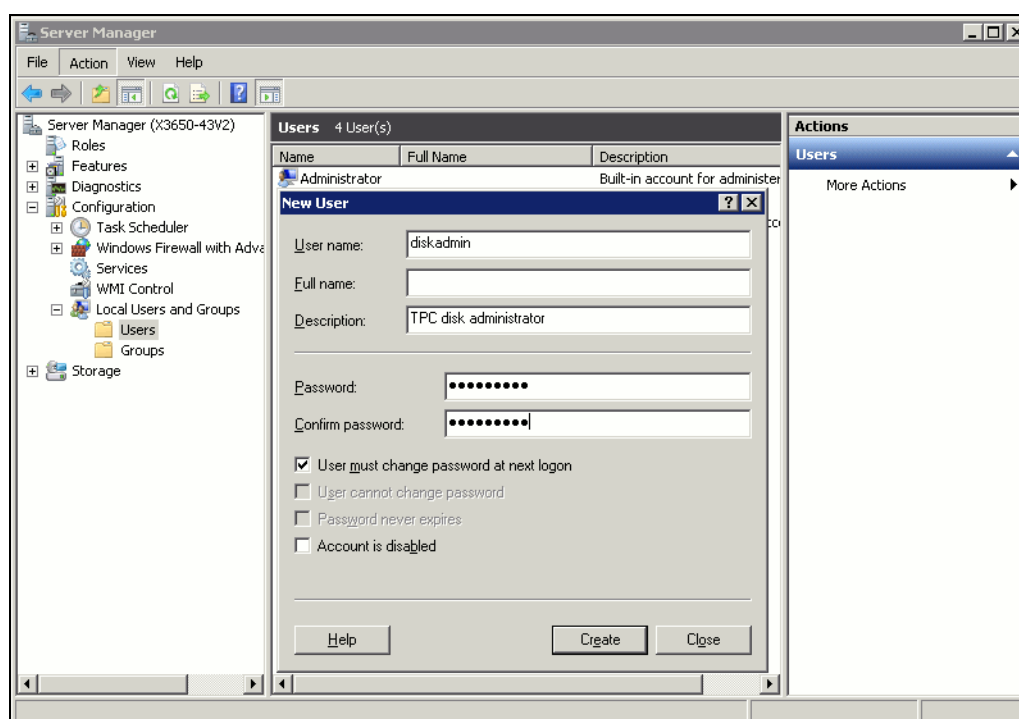


Figure 12-7 Set up a new user on the SSPC

1. From the SSPC Desktop, select **My Computer** → **Manage** → **Configuration** → **Local Users and Groups** → **Users**.
2. Select **Action** → **New User** and:
 - Set the user name and password.
 - If appropriate, check **User must change password at next login**. Note that if this box is checked, further actions are required for the new user to log on.
 - If appropriate, check **Password never expires**.
3. Click **Create** to add the new user.
4. Go back to **Local Users and Groups** → **Groups**.
 - Right-click **Groups** to add a new group or select an existing group.
 - Select **Add** → **Advanced** → **Find Now** and select the user to be added to the group.
5. Click **OK** to add the user to the group, then click **OK** again to exit user management.

Tip: To simplify user administration, use the same name for the Windows user group and the user groups role in TPC. For example, create the Windows user group “Disk Administrator” and assign this group to the TPC role “Disk Administrator”.

Set up user roles in TPC

The group defined at the OS level now needs to be mapped to a role defined in TPC. To do this, perform these steps:

1. Log into TPC with Administrator permissions and select **Administrative Services** → **Configuration** → **Role-to-Group Mapping**.
2. Add the Group you created in Windows to a Role, as shown in Figure 12-8. For DS8000 system management, the recommended roles are Disk Operator, Disk Administrator, and Security Administrator.

After the SSPC administrator has defined the user role, the operator is able to access the TPC GUI.

3. The authorization level in TPC depends on the role assigned to a user group. Table 12-1 shows the association between job roles and authorization levels.

Table 12-1 TPC roles and TCP administration levels

TCP Role	TCP administration level
Superuser	Has full access to all TPC functions
TPC Administrator	Has full access to all operations in the TPC GUI
Disk Administrator	• Has full access to TPC GUI disk functions, including tape devices • Can launch DS8000 GUI by using stored passwords in TPC Element Manager • Can add and delete volumes by TPC
Disk Operator	• Has access to reports of disk functions and tape devices • Has to enter user name/password to launch the DS8000 GUI • Cannot start CIMOM discoveries or probes • Cannot take actions in TPC, for example, delete and add volumes

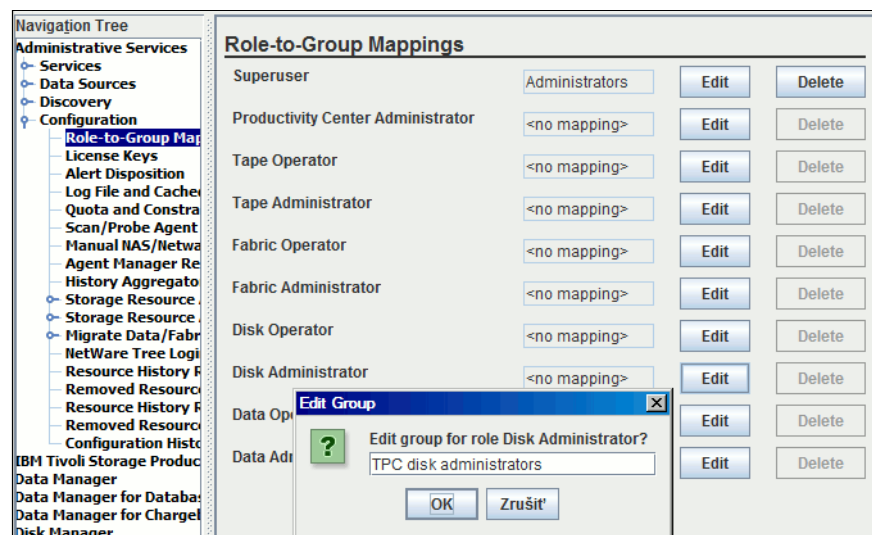


Figure 12-8 Assigning the Windows user group Disk Administrator to the TPC Role Disk Administrator

12.3 Configuring Tivoli Productivity Center For DS8000

This procedure describes how to configure Tivoli Storage Productivity Center that is installed on an SSPC server

12.3.1 Before you start

This procedure assumes the following conditions:

- ▶ You have DS8000 R4.1 or later.

Tip: To access the latest DS8000 R6,1 GUI, you need to install or upgrade an existing installation of TPC 4.2.1 to TPC 4.2.1 Fixpack 3 found in support/downloads at:

<http://publib.boulder.ibm.com/infocenter/tivihelp/v4r1/index.jsp>

- ▶ The DS8000 administrator created user names and passwords for the users of the element managers.
- ▶ Each user name that the DS8000 administrator created is in the administrator group.
- ▶ You completed all steps for setting up the SSPC hardware.
- ▶ You use the Subsystem Standard Group in the Configure Devices wizard when you configure Tivoli Storage Productivity Center to collect data from the DS8000 server.

Tip: Tivoli Storage Productivity Center V4.2 provides a new Configure Devices wizard to configure the devices that Tivoli Storage Productivity Center monitors. The wizard guides you through the steps for adding a device as a data source, running a discovery, including devices in groups, specifying alerts, and setting up schedules for collecting data. By using the wizard, you can configure storage systems, fabrics and switches, computers, and tape libraries.

After you add an IBM System Storage DS8000 server to Tivoli Storage Productivity Center, you can run a probe job to collect DS8000 data. You can also access the DS8000 graphical user interface (GUI).

12.3.2 Adding a DS8000 server by using the Configure Devices wizard

Starting with IBM Tivoli Storage Productivity Center V4.2, the DS8000 can be only be managed using the native device interface. Common Information Model (CIM) agents are no longer required by Tivoli Storage Productivity Center to communicate with the DS8000. Instead, Tivoli Storage Productivity Center uses the Enterprise Storage Server Network Interface (ESSNI) to communicate with DS8000 servers. If you are upgrading Tivoli Storage Productivity Center, a migration tool is provided to help you migrate the existing storage system credentials and to change port numbers.

Tip: With TPC 4.2, if you attempt to define DS8000 CIMOM connection in TPC, a discovery process will fail with an error such as HWN021724W CIMOM https://9.155.54.30:6989 manages Device(s) of type DS8000 which is supported through the native device interface only.

You can add an IBM System Storage DS8000 storage system to Tivoli Storage Productivity Center by using the following steps:

Tip: You can add multiple devices of the same type using a single session of the wizard. For example, you can add a DS8000 server, IBM System Storage SAN Volume Controller, and storage systems that use Common Information Model (CIM) agents by using a single session of the wizard. You cannot configure multiple devices of different types at the same time. For example, you cannot add a storage system and a fabric using a single session of the wizard

1. From the Welcome to the IBM Tivoli Storage Productivity Center window (Figure 12-5 on page 288), click **Add Devices**.
2. From the Select Device Type window, click **Storage Subsystem** and click **Next** (Figure 12-9).

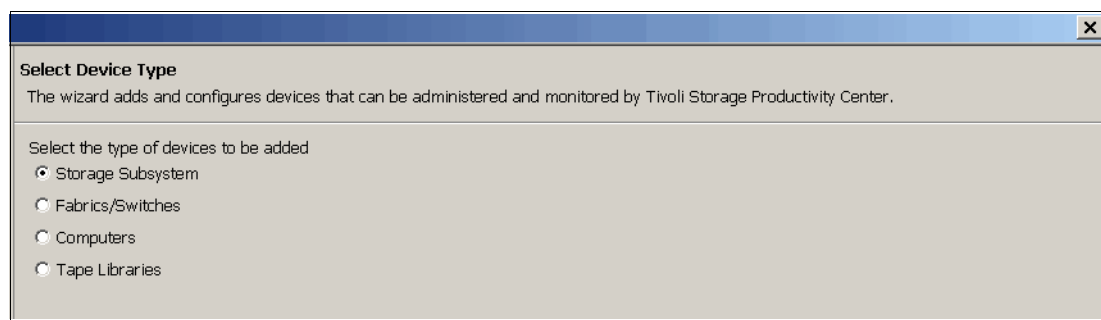


Figure 12-9 Select Device Type window

Tip: You can also access the window shown in Figure 12-9 in the left menu panel navigate to **Administrative services** → **Data sources** → **Storage subsystems** and click the **Add** button (Figure 12-10).

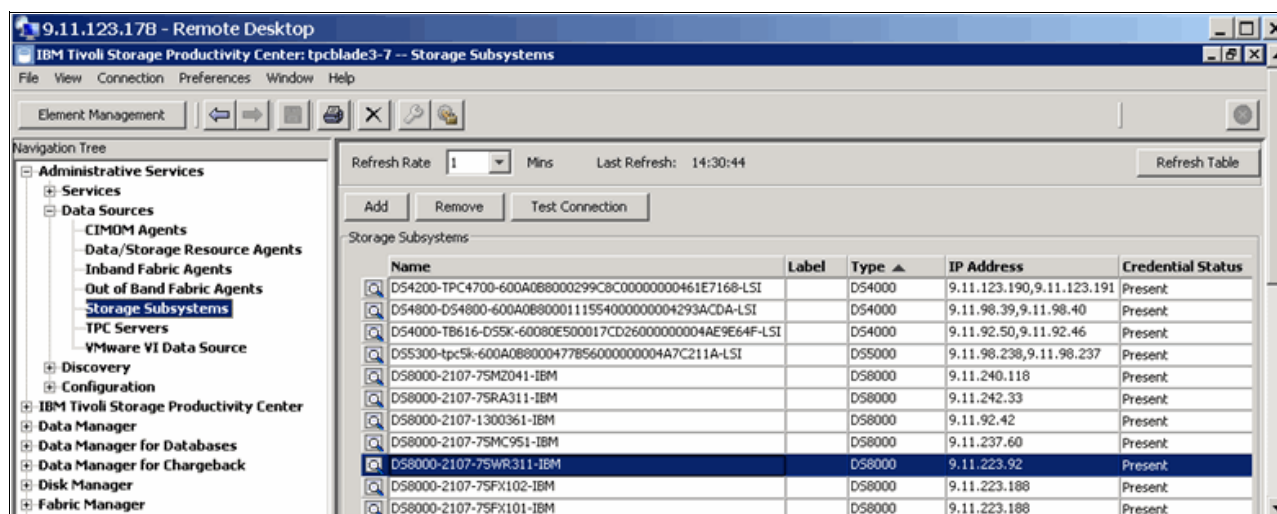


Figure 12-10 Add and configure DS8000 from Administrative Services

- From the Select devices window, click **Add and configure new storage subsystems** and click **Next** to continue, as shown in Figure 12-11.

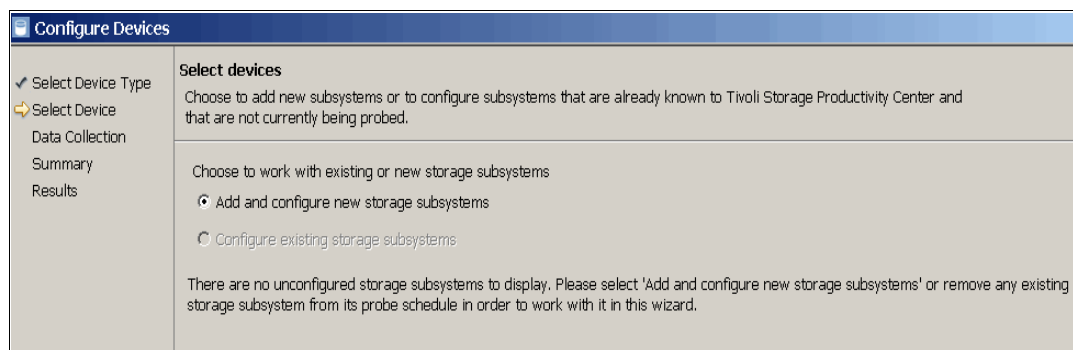


Figure 12-11 Select device type

- From the Configure storage subsystem connections panel (Figure 12-12), locate the *Device Type* field. Click the arrow to view the list of devices, then click **IBM DS8000**. Enter the following connection properties for the DS8000 storage system:
 - HMC Address**
Enter the IP address or host name for the Hardware Management Console (HMC) that manages the DS8000 server.
 - HMC2 Address (Optional)**
Enter the IP address or host name of a second HMC that manages the DS8000 server.
 - Username**
Enter the user name for logging on to the IBM System Storage DS8000 Storage Manager (also known as the DS8000 element manager or GUI) as shown in Figure 12-12. The default user name is admin.

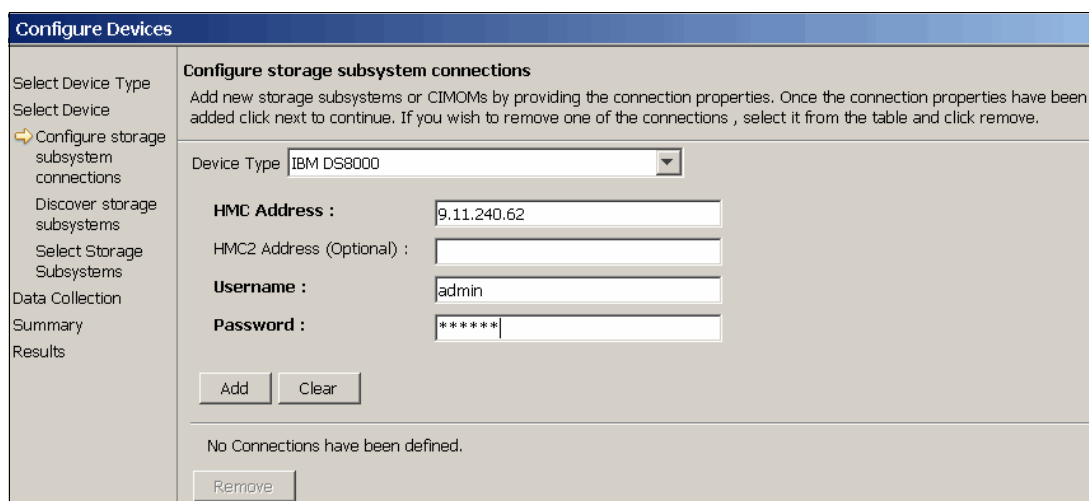


Figure 12-12 Configure DS8000 Storage Subsystem

Tip: This user name is the same as the user name for the Enterprise Storage Server Network Interface (ESSNI).

When you log on to the DS8000 Storage Manager for the first time by using the administrator user ID, you are required to change the password.

You need one user name and password to set up the DS8000 Storage Manager. The user name and the password for it are stored in the Tivoli Storage Productivity Center database. After the user name and password are stored, when you log on to the DS8000 Storage Manager in Tivoli Storage Productivity Center, the stored user name and the privileges associated with it are retrieved.

To set up Lightweight Directory Access Protocol (LDAP) and single sign-on for DS8000 in Tivoli Storage Productivity Center, see “Configuring Tivoli Storage Productivity Center for DS8000 LDAP support and Single sign-on” at the TPC InfoCenter:

<https://publib.boulder.ibm.com/infocenter/tivihelp/v4r1/index.jsp>

– Password

Enter the password for the user name. The default password is admin.

If you want to remove the information that you entered, click **Clear**.

Tip: For information configuring Tivoli Storage Productivity for LDAP refer to *Understanding LDAP Design and Implementation*, SG24-4986 at:

<http://www.redbooks.ibm.com/abstracts/sg244986.html>

For information about DS8000 LDAP Authentication, refer to *IBM System Storage DS8000: LDAP Authentication*, REDP-4505 at:

<http://www.redbooks.ibm.com/abstracts/redp4505.html?open>

5. Click **Add**. The Configure storage subsystem connections window (Figure 12-13) displays the IP (HMC) address and device type connection properties that you entered

IP Address	Device Type
9.11.240.62	IBM DS8000

Figure 12-13 Add DS8000 Storage Subsystem; IP 9.11.240.62

6. To enter connection properties for additional DS8000 servers, repeat steps 4 and 5, then go to step 7.

7. When you finish entering connection properties for the DS8000 servers that you want to add, click **Next**. Tivoli Storage Productivity Center discovers the DS8000 servers and collects initial configuration data from them. When the discovery and collection are completed, the message **Completed successfully** is displayed in the Status column. The devices that are displayed are known to Tivoli Storage Productivity Center as shown in Figure 12-14.

Status	Start Time	End Time	Log FileName
Completed successfully	May 9, 2011 4:33:33 PM MST	May 9, 2011 4:33:37 PM MST	D:\Program Files\IBM\TPC\device\log\msg.NAPIDisco
Completed successfully	May 9, 2011 4:33:33 PM MST	May 9, 2011 4:33:37 PM MST	D:\Program Files\IBM\TPC\device\log\msg.NAPIDisco
Completed successfully	May 9, 2011 4:33:33 PM MST	May 9, 2011 4:33:37 PM MST	D:\Program Files\IBM\TPC\device\log\msg.9.11.123.2

Figure 12-14 Discover Storage Subsystems

8. On the Discover storage subsystems window, click **Next**.
9. On the Select Storage Subsystems window, select the DS8000 server that you want to add as shown in Figure 12-15, and click **Next**.

IP Address	Name
9.11.240.62	DS8000-2107-75LX521-IBM

Figure 12-15 Select Storage Subsystem

10. In the Data Collection window (Figure 12-16 on page 297), indicate how you want Tivoli Storage Productivity Center to collect data from the DS8000 server:
 - a. In the **Use a monitoring group or template** field, select **Monitoring Group**. When you include the DS8000 server in a monitoring group, Tivoli Storage Productivity Center manages the server and a collection of other storage systems in the same manner.
 - b. In the **Select monitoring group** field, click the arrow to select **Subsystem Standard Group**.

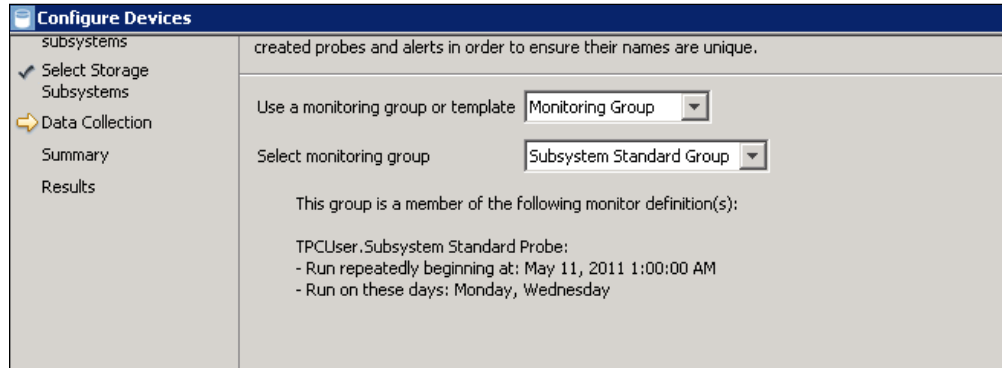


Figure 12-16 Data Collection, Select Monitoring Group: Subsystem Standard Group

Each monitoring group is associated with an existing probe schedule and set of alerts. When you select a monitoring group, its data collection schedule and alerts are automatically applied to the DS8000 server and all storage systems that you are adding. After you complete the wizard, the storage systems remain in the group and you can use that group when working in other parts of the Tivoli Storage Productivity Center user interface, such as in reports.

When you select the **Subsystem Standard Group** option, the following actions occur:

- Tivoli Storage Productivity Center automatically includes the DS8000 in the probe schedule to which the monitoring group is associated. For example, when you select **Subsystem Standard Group**, the DS8000 is included in the Subsystem Standard Probe schedule. The Subsystem Standard Probe schedule runs two times per week on Monday and Wednesday at 1:00 am in our example.
- Tivoli Storage Productivity Center applies a set of alert conditions to the DS8000 and the other storage systems in the monitoring group. When any of these conditions are detected during data collection, an alert is triggered. The group that you select determines the alerts that are associated with the storage systems. For example, if you select Subsystem Standard Group, 18 alert conditions are automatically checked for the storage systems.

In Figure 12-17 on page 298, the **Select monitoring group** field shows information about the group or template that you selected.

To learn more about monitoring groups, search the Information Center at for alerts and schedules associated with monitoring groups, and templates and triggering conditions for alerts. The Information Center is available at the following URL:

<https://publib.boulder.ibm.com/infocenter/tivihelp/v4r1/index.jsp> .

c. Click **Next** to continue.

11. From the Review user selections window (Figure 12-17 on page 298), review the configuration choices that you made for the DS8000 by using the Configure Devices wizard:

- The list of devices that you are adding.
- The name of the monitoring group that you selected. This value is not displayed if you selected a template.
- The name of the probe schedule that is created based on the template you selected. This value is not displayed if you selected a monitoring group.
- Information about the probe schedule created for a template.

- The names of the alerts that are created based on the template you selected. This value is not displayed if you selected a monitoring group.

To change your configuration choices, click **Back** and repeat Step 10.

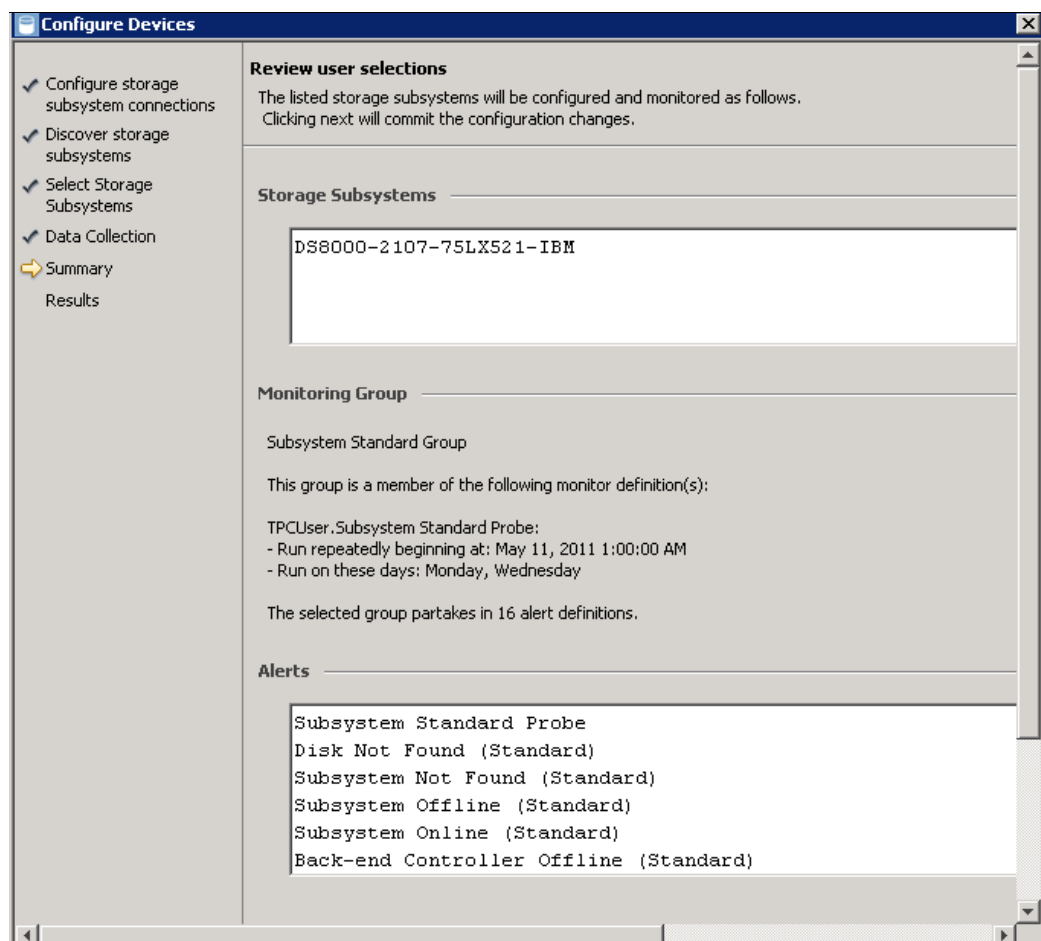


Figure 12-17 Review User Selections from Monitoring Group : Subsystem Standard Group

12. Click **Next** to commit your configuration choices and to complete the configuration. The View results window (Figure 12-18 on page 299) is displayed. It includes the following information:

- A list of the actions performed by the wizard. The window displays a row for the successful actions. If the configuration failed, a row is displayed for each failed action. For example, if the wizard expects to assign a specific alert to five devices, but the operation only succeeds for the three of the devices, this window displays one row for the successful actions and two rows for each failed action.
- Error messages for any failed actions. To resolve the error, search for the message identifier in this information center.

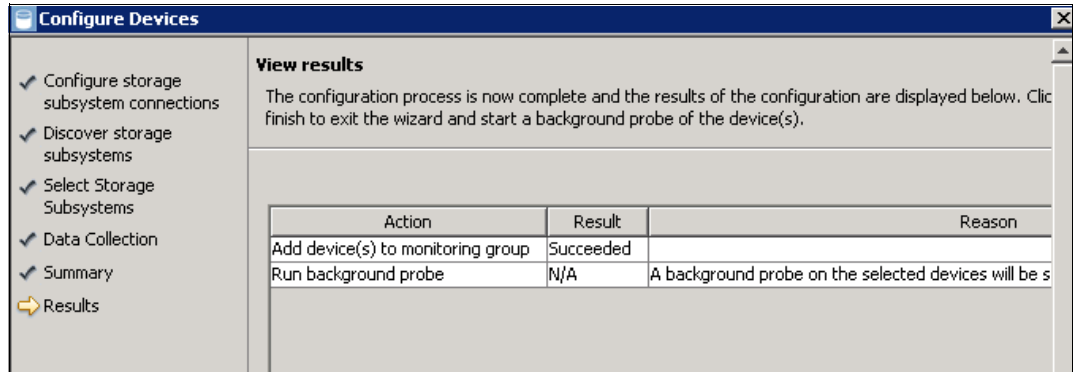


Figure 12-18 View Results

13. Click **Finish** to close the Configure Devices wizard.

12.3.3 Running a probe job to collect DS8000 data

Use these steps to run a Tivoli Storage Productivity Center probe job to collect data about IBM System Storage DS8000:

1. From the View results window (Figure 12-18), click **Finish** to close the wizard and start a probe of the DS8000 server that you configured. To view the status of the probe, from the Job History window click **View Job History**. Depending on your settings, a pop-up window, as shown in Figure 12-19, might appear allowing you to view the discovery logs.

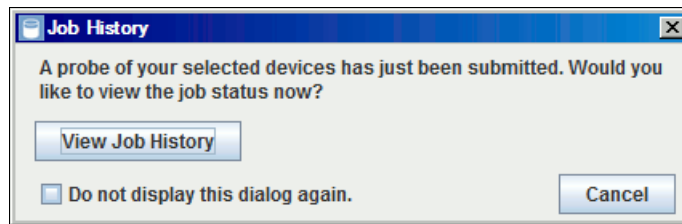


Figure 12-19 Pop-up window will bring you to the jobs list.

With TPC V4.2, there is a significant change in log management. Figure 12-20 on page 300 shows how you can access all TPC logs in one place through the Job Management window.

Use the Job Management window to view and manage all schedules, runs, and jobs that are related to the storage devices that Tivoli Storage Productivity Center monitors. Also, if your devices are not fully monitored (that is, not included in all data collection schedules that are available to the licensed edition of Tivoli Storage Productivity Center that is installed), take advantage of the Job Management - Recommendations window. Messages in the window indicate actions that you can take to more fully utilize Tivoli Storage Productivity Center to monitor your devices.

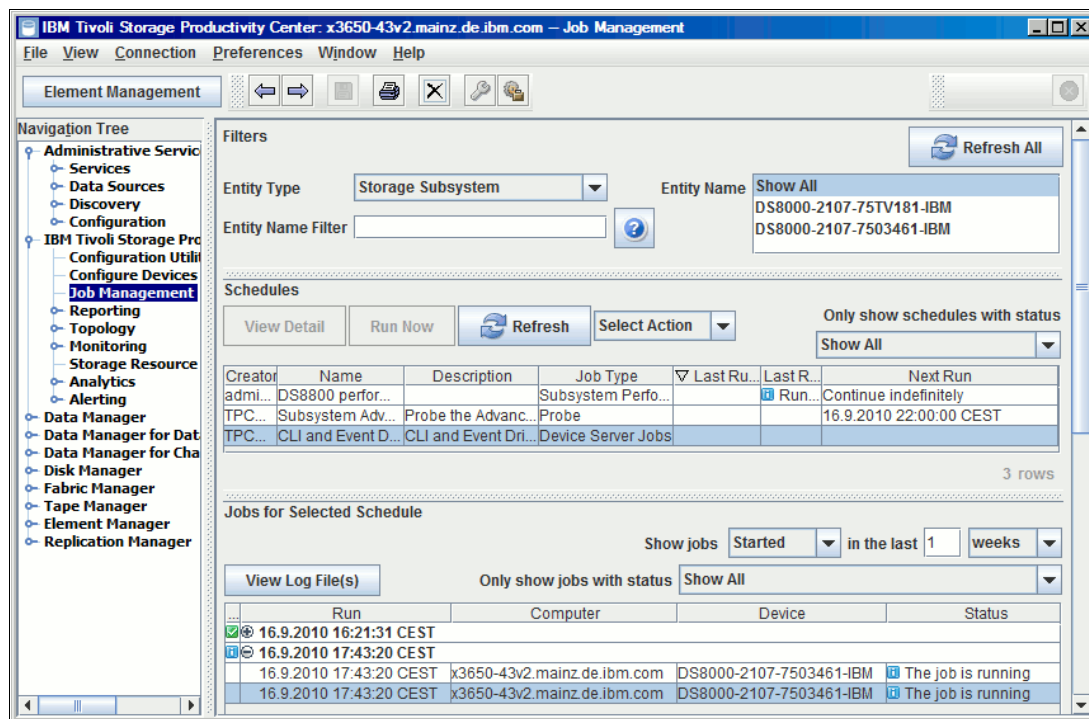


Figure 12-20 TPC 4.2 Job Management Panel

2. Wait for the Probe job to complete. When the Probe is finished, Tivoli Storage Productivity Center adds a row for the Probe in both tables of the Job Management window.

- If the Probe was successful:
 - In the Schedules table, Success is displayed the Last Run Status column. It is accompanied by a green status indicator.
 - In the Jobs for Selected Schedule table, the probe job is displayed in the Run column. It is accompanied by a green status indicator.

The data that results from a probe is stored in the Tivoli Storage Productivity Center database repository. You can use the data to generate reports, including Asset, Data Source, and Storage Subsystem reports. To learn how to generate reports from the data, go to Accessing reports about Storage Resources.

- If the Probe failed:
 - In the Schedules table, Failed is displayed in the Last Run Status column. It is accompanied by a red status indicator.
 - In the Jobs for Selected Schedule table, the probe job is displayed in the Run column. It is accompanied by a red status indicator. Expand the entry and view the information in the Status column to learn more about the failure.

12.3.4 Configuring the TPC Element Manager to access the DS8000 GUI

The TPC Element Manager is a single point of access to the GUI for all the DS8000 systems in your environment. Using the TPC Element Manager for DS8000 remote GUI access allows you to:

- ▶ View a list of Elements (DS8000 GUIs within your environment).
- ▶ Access all DS8000 GUIs by launching an Element with a single action.
- ▶ Add and remove DS8000 Elements. The DS8000 GUI front-end can be accessed by http or https.
- ▶ Save the user and password to access the DS8000 GUI. This option to access the DS8000 GUI without reentering the password allows you to configure SSPC as a single sign-on to all DS8000s in the client environment.

Tip: For storage systems discovered or configured for CIMOM or native device interface, TPC automatically defines Element Manager access. You need to specify the correct username and password in the Element Manager to use it.

Adding and launching Element Managers for the DS8000 in TPC

The Element Manager window displays all storage systems (Element Managers) already defined to TPC. To add a new Element Manager:

1. From the **Select Action** drop-down menu, select **Add Element Manager** as shown in Figure 12-21.

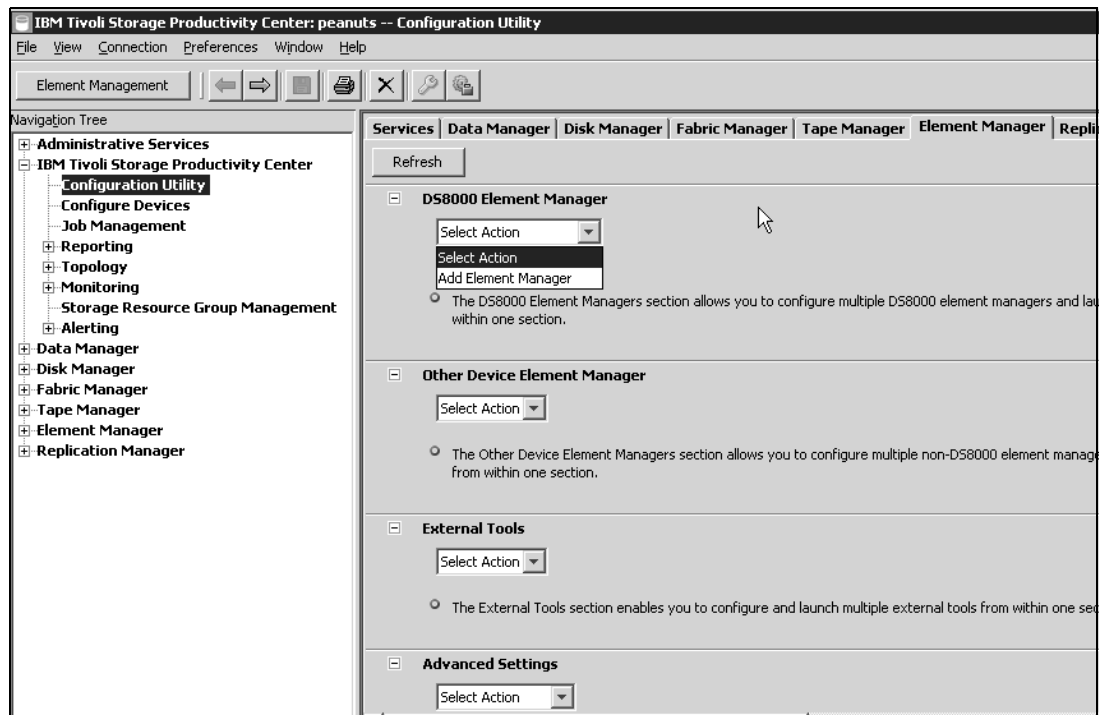
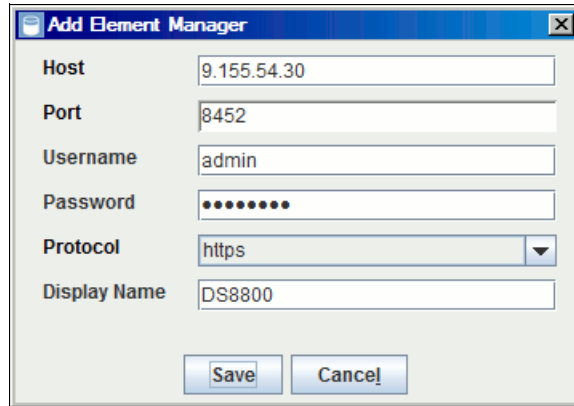


Figure 12-21 TPC Element Manager view: Options to add and launch Elements

2. In the Add Element Manager window (Figure 12-22 on page 302), provide the following information:
 - a. Host: Enter the Domain Name System (DNS) name or IP address of the DS8000 HMC.
 - b. Port: The port number on which the DS8000 HMC listens for requests.

- c. User name and associated Password already defined on DS8000: The default DS8000 user name is admin and password is admin. If this is the first time you try to log on DS8000 with the admin user name, you are prompted to change the password. Be prepared to enter a new password and record it in a safe place.
- d. Protocol: HTTPS or HTTP.
- e. Display Name: Specify a meaningful name of each Element Manager to identify each DS8000 system in the Element Manager table. This is useful, particularly when you have more than one DS8000 system managed by a single SSPC console.

Click **Save** to add the Element Manager.



The 'Add Element Manager' dialog box contains the following fields:

- Host:** 9.155.54.30
- Port:** 8452
- Username:** admin
- Password:** (masked with dots)
- Protocol:** https (dropdown menu)
- Display Name:** DS8800

Buttons: Save, Cancel

Figure 12-22 Configure a new DS8000 Element in the TPC Element Manager view

TPC tests the connection to the DS8000 Element Manager. If the connection was successful, the new DS8000 Element Manager is displayed in the Element Manager table.

3. After the DS8000 GUI has been added to the Element Manager, select the Element Manager you want to work with and, from the **Select Action** drop-down menu, select **Launch Default Element Manager** as shown in Figure 12-23.

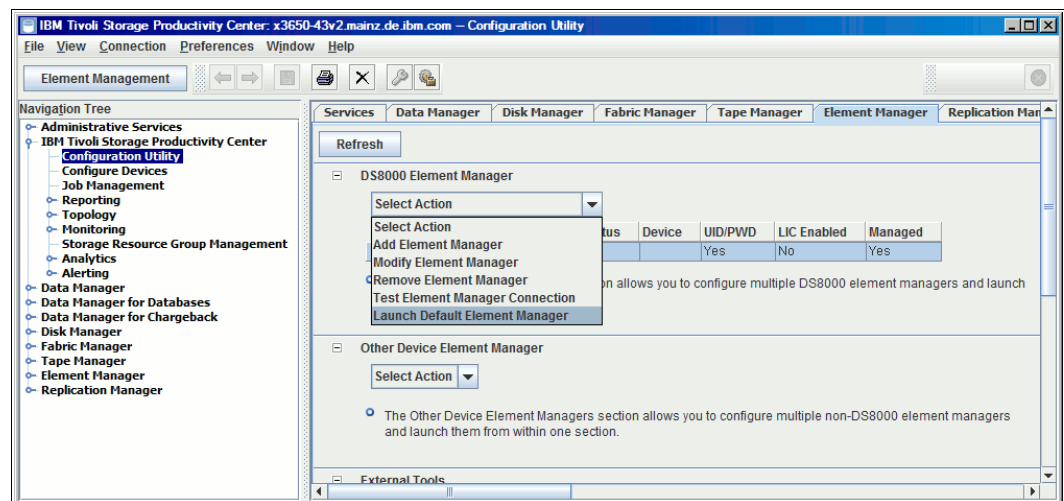


Figure 12-23 Launch the Element Manager

- The DS8000 GUI Overview window for the selected disk system displays. Figure 12-24 shows the latest DS8000 6.1 GUI.

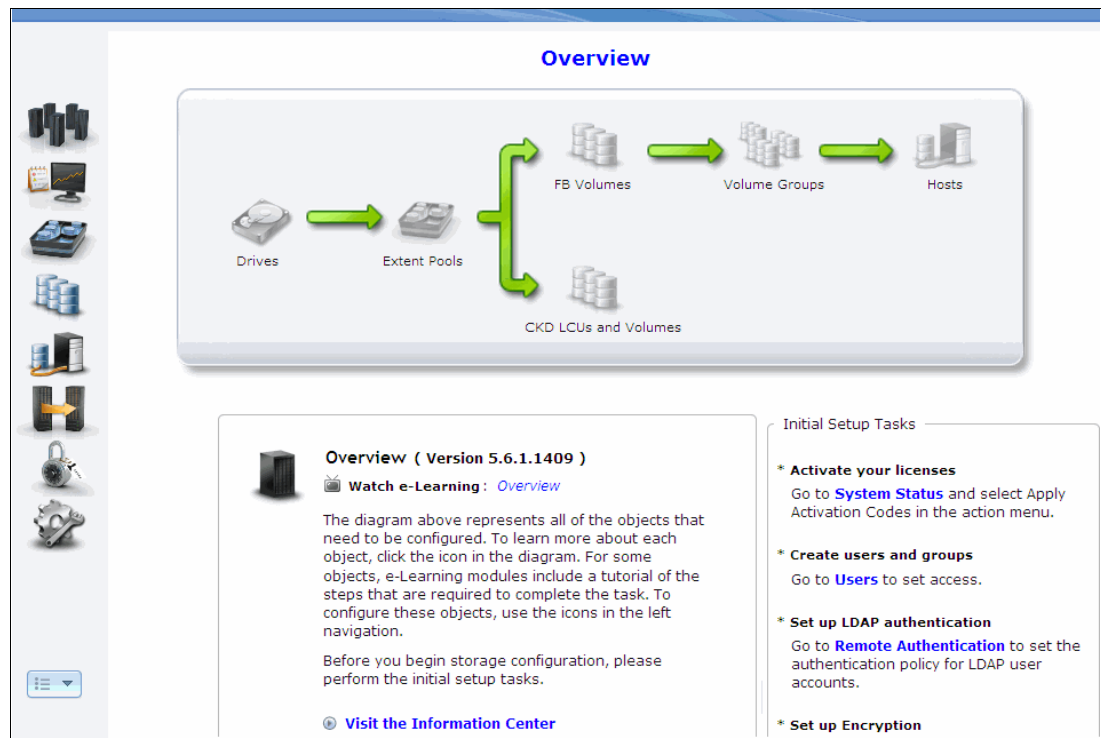


Figure 12-24 SSPC: DS8000 6.1 GUI Overview window

With TPC V4.2, there is a significant change in log management. You can now access all TPC logs in one place in the Job Management window as shown in Figure 12-25.

IBM Tivoli Storage Productivity Center: x3650-43v2.mainz.de.ibm.com – Job Management

File View Connection Preferences Window Help

Element Management

Navigation Tree

- Administrative Services
 - Services
 - Data Sources
 - Discovery
 - Configuration
- IBM Tivoli Storage Productivity Center
 - Configuration Utilities
 - Configuration Devices
 - Job Management**
 - Reporting
 - Topology
 - Monitoring
 - Storage Resource
 - Analytics
 - Alerting
- Data Manager
 - Data Manager for Data
 - Data Manager for Change
 - Disk Manager
 - Fabric Manager
 - Tape Manager
 - Element Manager
 - Replication Manager

Filters

Entity Type: Storage Subsystem Entity Name: Show All

Entity Name Filter: [Search]

Schedules

View Detail Run Now Refresh Select Action Only show schedules with status: Show All

Creator	Name	Description	Job Type	Last Run	Last Run Date	Next Run
admin...	DS8800 perfor...	Subsystem Perfo...	Subsystem Perfo...	Run...	Continue indefinitely	
TPC...	Subsystem Adv...	Probe the Advanc...	Probe		16.9.2010 22:00:00 CEST	
TPC...	CLI and Event D...	CLI and Event Dri...	Device Server Jobs			

3 rows

Jobs for Selected Schedule

Show jobs: Started in the last 1 weeks

View Log File(s) Only show jobs with status: Show All

Run	Computer	Device	Status
16.9.2010 16:21:31 CEST			
16.9.2010 17:43:20 CEST			
16.9.2010 17:43:20 CEST	x3650-43v2.mainz.de.ibm.com	DS8000-2107-7503461-IBM	The job is running
16.9.2010 17:43:20 CEST	x3650-43v2.mainz.de.ibm.com	DS8000-2107-7503461-IBM	The job is running

Figure 12-25 Job management interface

Tip: If not already defined, during this process TPC automatically defines GUI access for the discovered DS8000 in the Element Manager.

12.4 Working with a DS8000 system in TPC-BE

Perform a number of tasks regularly to ensure that operational performance is maintained. In this section, we describe actions to maintain the relationship between TPC and the DS8000 system.

12.4.1 Display disks and volumes of DS8000 Extent Pools

To display the volumes and Disk Drive Modules (DDMs) used by an Extent Pool, double-click that Extent Pool in the Topology viewer. Underneath this topology image, a table view provides further information about the DS8000 devices, as shown in Figure 12-26, Figure 12-27, Figure 12-28 on page 305, and Figure 12-29 on page 305. Details about the displayed health status are discussed in 12.4.3, “Storage health management” on page 308.

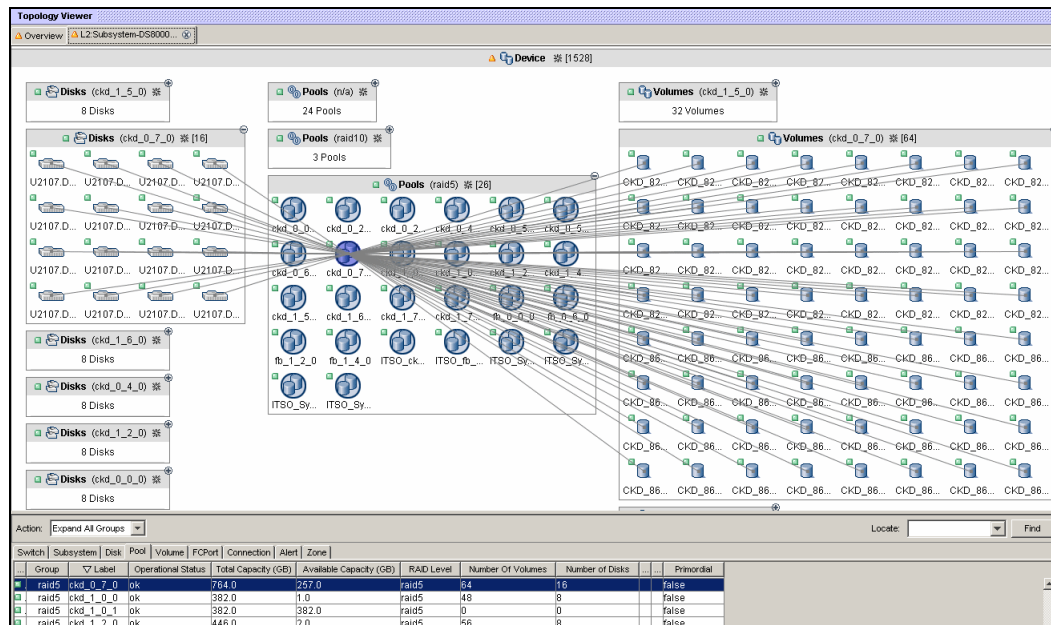


Figure 12-26 Drill-down of the topology viewer for a DS8000 Extent Pool

Figure 12-27 shows a graphical and tabular view with more information

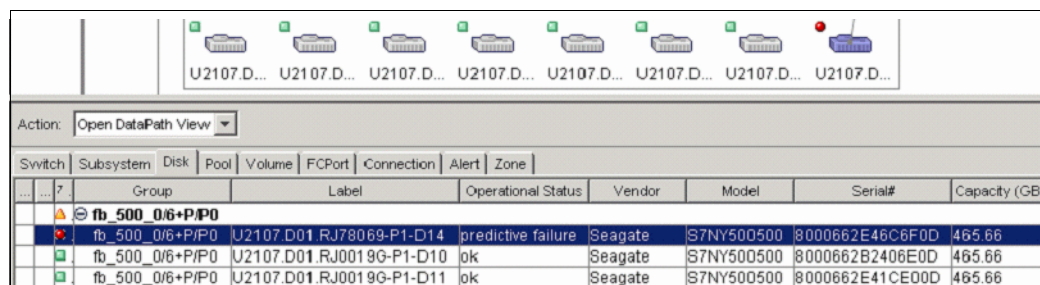


Figure 12-27 Graphical and tabular view of a broken DS8000 DDM set to deferred maintenance

DS8000-2107-75P6151-IBM

Action:

Switch	Subsystem	Disk	Pool	Volume	FCPort	Connection	Alert	Zone			
		Group				Label	Operational	Vendor	Model	Serial#	Capacity (G)
		HP_146GB_15K_P1									
		HP_146GB_15K_P0									
		LP_500GB_7.2K_P5									
		LP_500GB_7.2K_P4									
		MP_300GB_15K_P3									
		MP_300GB_15K_P2									
		n/a									
		n/a				U2107.D01.RJ82234-P1...	starting	Seagate	S7NY50050080006627E...	465.66	
		n/a				U2107.D01.RJ82234-P1...	starting	Seagate	S7NY50050080006627E...	465.66	
		n/a				U2107.D01.RJ82234-P1...	ok	Seagate	S7NY50050080006627E...	465.66	

The DDM displayed as green in Figure 12-28 is a spare DDM, and is not part of the RAID 5 configuration process that is currently in progress.

Chapter 12. System Storage Productivity Center **305**

12.4.2 Display the physical paths between systems

If Out of Band Fabric agents are configured, TPC-BE can display physical paths between SAN components. The view consists of four windows (computer information, switch information, subsystem information, and other systems) that show the physical paths through a fabric or set of fabrics (host-to-subsystem or subsystem-to-subsystem).

To display the path information shown in Figure 12-30 and Figure 12-31, execute these steps:

1. In the topology view, select **Overview** → **Fabrics** → **Fabric**.
2. Expand the Connectivity view of the devices for which you would like to see the physical connectivity.
3. Click the first device.
4. Press Ctrl and click any additional devices to which you would like to display the physical path (Figure 12-30).
5. To obtain more details about the connectivity of dedicated systems, as shown in Figure 12-31, double-click the system of interest and expand the details of the system view.

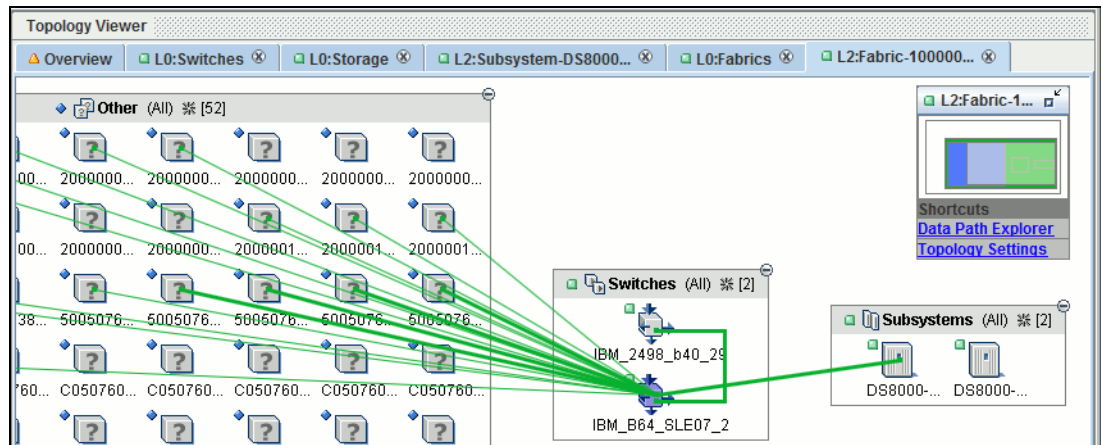


Figure 12-30 Topology view of physical paths between one host and one DS8000 system

In Figure 12-30, the display of the Topology view points out physical paths between the hosts and their volumes located on the DS8000 system. In this view, there are only WWPNs shown in left box labeled Other. To interpret WWPNs of a host in the fabric, data agents must be placed on that host. Upgrading TPC-BE with additional TPC licenses will enable TPC to assess and also warn you about lack of redundancy.

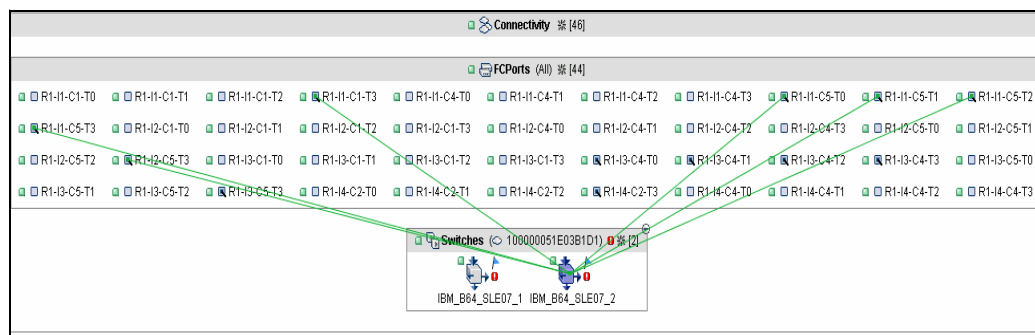


Figure 12-31 Topology view: detailed view of the DS8000 host ports assigned to one of its two switches

In Figure 12-31 on page 306, the display of the Topology viewer points out that the switch connectivity does not match one of the recommendations given by the DS8000 Information Center on host attachment path considerations for a storage image. In this example, we have two I/O enclosures in each I/O enclosure pair (I1/I2 or I3/I4) located on separate RIO loop halves (the DS8000 Information Center¹ mentions that “you can place two host attachments, one in each of the two I/O enclosures of any I/O enclosure pair”). In the example, all switch connections are assigned to one DS8000 RIO loop only (R1-I1 and R1-I2).

As shown in Figure 12-32, the health status function of TPC-BE Topology Viewer allows you to display the individual FC port health inside a DS8000 system.

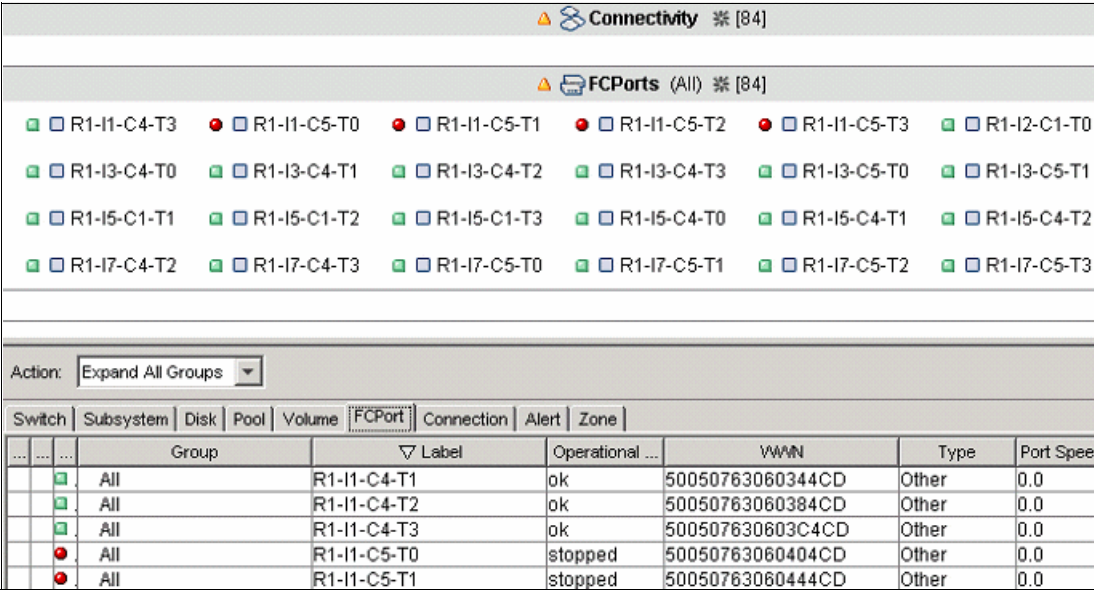


Figure 12-32 TPC graphical view of a broken DS8000 host adapter Card R1-I1-C5 and the associated WWNN as displayed in the tabular view of the topology viewer

As shown in Figure 12-33, the TPC-BE Topology viewer allows you to display the connectivity and path health status of one DS8000 system into the SAN by providing a view that can be broken down to the switch ports and their WWPNs.

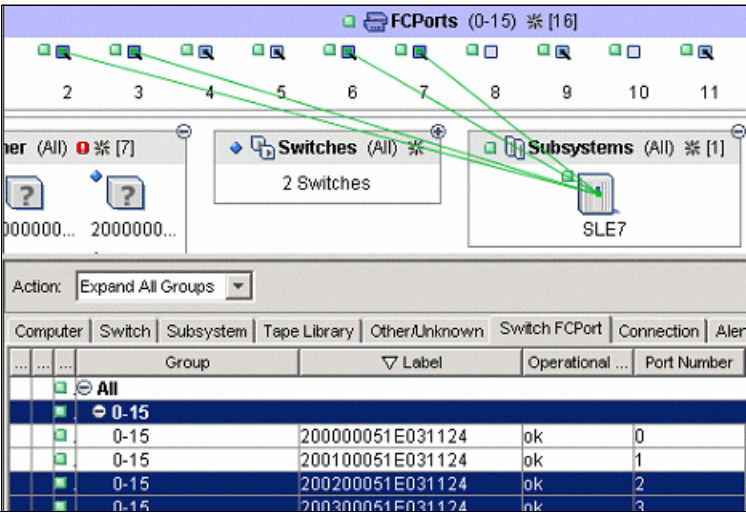


Figure 12-33 Connectivity of a DS8000 system drilled down to the ports of a SAN switch

¹ <http://publib.boulder.ibm.com/infocenter/dsichelp/ds8000ic/index.jsp>

12.4.3 Storage health management

TPC provides a graphical storage health overlay function. This function allows the user to easily spot unhealthy areas through color coding. If the SSPC is monitored on a regular basis, TPC can be configured to show new alerts when the GUI is launched. This can be done by selecting **Preferences → Edit General → On Login Show → All Active Alerts**.

However, if the TPC console is not regularly monitored for health status changes, configure alerts to avoid health issues going unrecognized for a significant amount of time. To configure alerts for the DS8000 system, in the navigation tree, select **Disk Manager → Alerting** and right-click **Storage Subsystem Alerts**. In the window displayed, the predefined alert trigger conditions and the Storage Subsystems can be selected. Regarding the DS8000 system, the predefined alert triggers can be categorized into:

- ▶ Capacity changes applied to cache, volumes, and Extent Pools
- ▶ Status changes to online/offline of storage subsystems, volumes Extent Pools, and disks
- ▶ Device not found for storage subsystems, volumes Extent Pools, and disks
- ▶ Device newly discovered for storage subsystem, volume Extent Pool, and disk
- ▶ Version of storage subsystems changed

12.4.4 Display host volumes through SVC to the assigned DS8000 volume

With SSPC's TPC-BE, you can create a table to display the name of host volumes assigned to an SVC vDisk and the DS8000 volume ID associated to this vDisk. For a fast view, select **SVC vDisks → MDisk → DS8000 Volume ID**. To populate this host, select **Volume name → SVC → DS8000 Volume ID view** (TPC-BE SVC and DS8000 probe setup is required). To display the table, as demonstrated in Figure 12-34, select **TPC → Disk Manager → Reporting → Storage Subsystems → Volume to Backend Volume Assignment → By Volume** and select **Generate Report**.

Storage Subsystem	Jser-Defined Volume Name ▲	Volume Name	Volume Capacity	Storage Pool	Disk	Disk Capacity
SVC_Gondor	Jerry_SG1log	14	1.00 GB	DS8k_group	mdisk_DS8k2	3.00 GB
SVC_Gondor	Jerry_SG1log	14	1.00 GB	DS8k_group	mdisk_DS8K1	3.00 GB
SVC_Gondor	Jerry_SG1log	14	1.00 GB	DS8k_group	mdisk_DS8K4	8.00 GB

Disk Unallocated Space	Backend Storage Subsystem	Backend Storage Subsystem Type	Backend Volume Name
272.00 MB	DS8k_EBC	DS8000	SVC_Gondor_4103 (ID:1007)
272.00 MB	DS8k_EBC	DS8000	SVC_Gondor_4102 (ID:1006)
0	DS8k_EBC	DS8000	SVC_Gondor_4139 (ID:102b)

Figure 12-34 Example of three DS8000 volumes assigned to one vDisk



Configuration using the DS Storage Manager GUI

The DS Storage Manager provides a graphical user interface (GUI) to configure the IBM System Storage DS8000 series and manage DS8000 Copy Services. The DS Storage Manager GUI (DS GUI) is invoked from the System Storage Productivity Center (SSPC) by launching an Element Manager in the TPC GUI. See Chapter 12, “System Storage Productivity Center” on page 279 for details on SSPC.

This chapter explains the possible ways to access the DS GUI, and how to use it to configure the storage on the DS8000.

This chapter includes the following sections:

- ▶ DS Storage Manager GUI overview
- ▶ Logical configuration process
- ▶ Examples of configuring DS8000 storage
- ▶ Examples of exploring DS8000 storage status and hardware

For information about Copy Services configuration in the DS8000 family using the DS GUI, see the following IBM Redbooks publications:

- ▶ *IBM System Storage DS8000: Copy Services for Open Systems*, SG24-6788
- ▶ *IBM System Storage DS8000: Copy Services for IBM System z*, SG24-6787

For information about DS GUI changes related to disk encryption, see *IBM System Storage DS8700: Disk Encryption Implementation and Usage Guidelines*, REDP-4500.

For information about DS GUI changes related to LDAP authentication, see *IBM System Storage DS8000: LDAP Authentication*, REDP-4505.

Tip: Some of the screen captures in this chapter might not reflect the latest version of the DS GUI code.

13.1 DS Storage Manager GUI overview

In this section, we describe the DS Storage Manager GUI (DS GUI) access method design. The DS GUI code resides on the DS8000 Hardware Management Console (HMC) and we discuss access methodologies.

13.1.1 Accessing the DS GUI

The DS GUI code at the DS8000 HMC is invoked at the SSPC from the Tivoli Storage Productivity Center (TPC) GUI, and accessed by launching an Element Manager. The sequence of windows displayed when accessing the DS GUI through SSPC is shown in Figure 13-3 on page 312, Figure 13-4 on page 313, and Figure 13-5 on page 313. The DS8000 HMC that contains the DS Storage Manager communicates with the DS Network Interface Server, which is responsible for communication with the two controllers of the DS8000.

Access to the DS8000 HMC is supported through the IPv4 and the newer IPv6 Internet Protocol.

You can access the DS GUI in any of the following ways:

- ▶ Through the System Storage Productivity Center (SSPC)
- ▶ From TPC on a workstation connected to the HMC
- ▶ From a browser connected to SSPC or TPC on any server
- ▶ Using Microsoft Windows Remote Desktop through the SSPC

These access capabilities, using basic authentication, are shown in Figure 13-1. In our illustration, SSPC connects to two HMCs managing two DS8000 storage complexes. Although you have different options to access DS GUI, SSPC is the preferred access method.

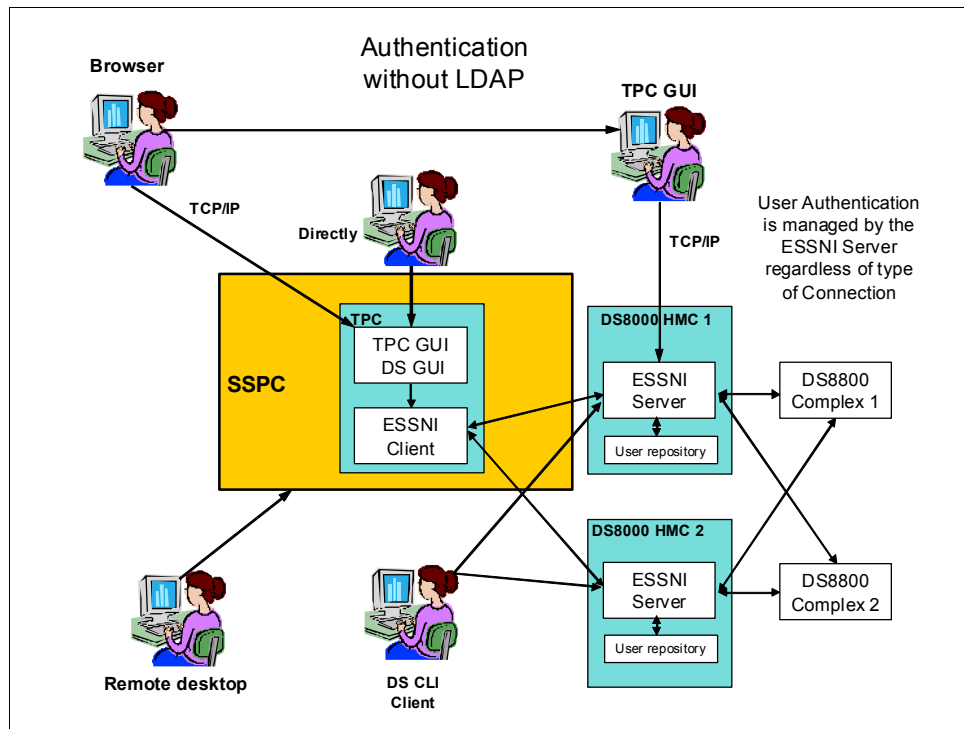


Figure 13-1 Accessing the DS8000 GUI

The DS8000 supports the ability to use a Single Point of Authentication function for the GUI and CLI through an centralized Lightweight Directory Access Protocol (LDAP) server. This capability is supported with SSPC running on 2805-MC5 hardware that has TPC Version 4.2.1 (or later) preloaded. If you have an earlier SSPC hardware version with a earlier TPC version, you have to upgrade TPC to V4.2.1 to take advantage of the Single Point of Authentication function for the GUI and CLI through a centralized LDAP server.

The access capabilities of the LDAP authentication are shown in Figure 13-2. In this illustration, TPC connects to two HMCs managing two DS8000 storage complexes.

Tip: For detailed information about LDAP-based authentication, see *IBM System Storage DS8000: LDAP Authentication*, REDP-4505.

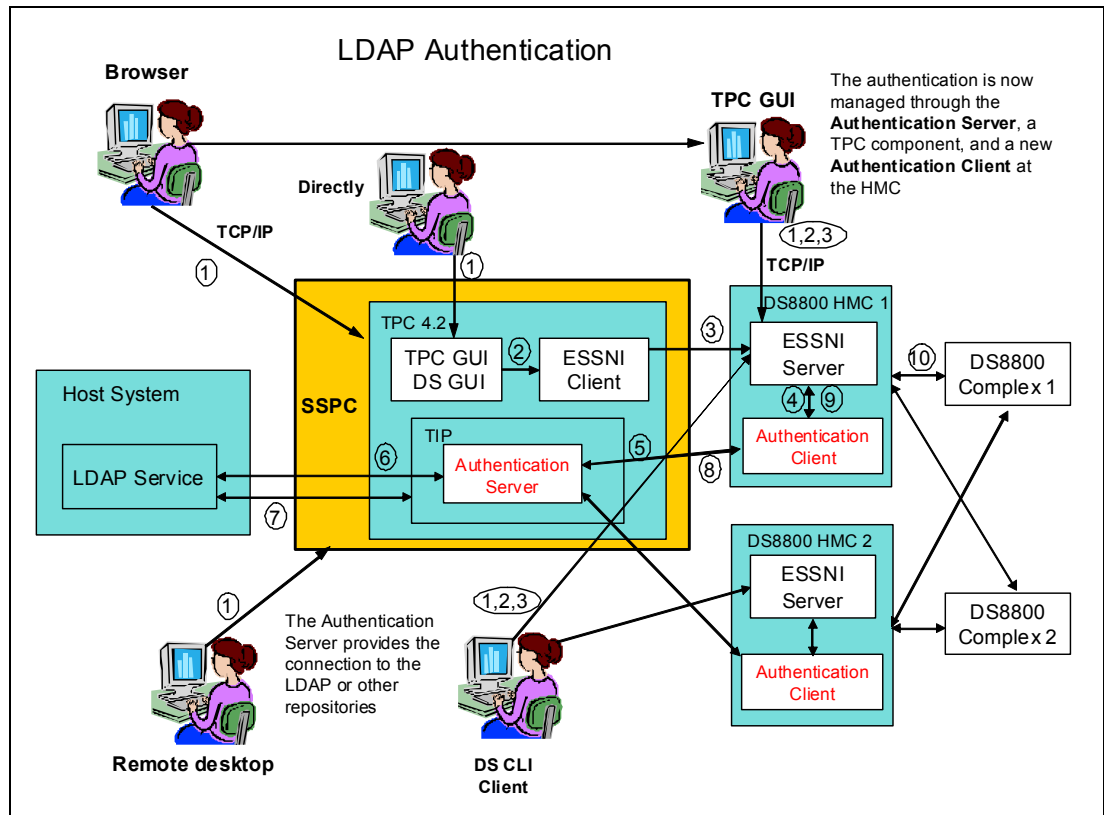


Figure 13-2 LDAP authentication to access the DS8000 GUI and CLI

Accessing the DS GUI through SSPC

As previously stated, the recommended method for accessing the DS GUI is through SSPC.

To access the DS GUI through SSPC, perform the following steps:

1. Log in to your SSPC server and launch the IBM Tivoli Storage Productivity Center.
2. Type in your Tivoli Storage Productivity Center user ID and password.
3. In the Tivoli Storage Productivity Center window shown in Figure 13-3, click **Element Management** (above the Navigation Tree) to launch the Element Manager.

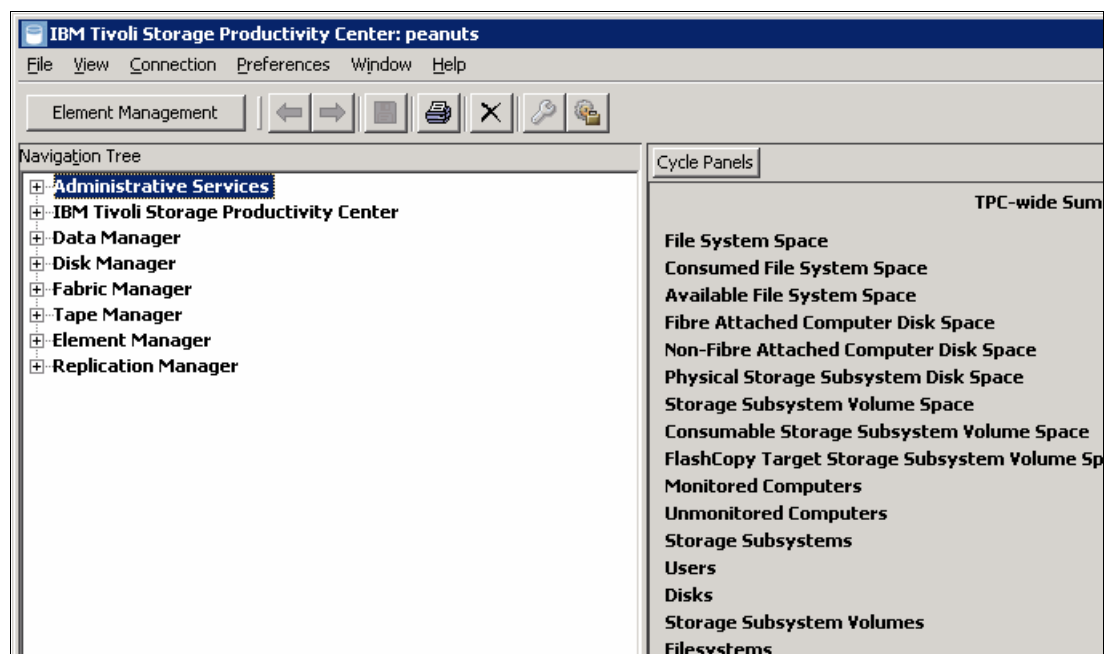


Figure 13-3 SSPC: Launch Element Manager

Tip: Here we assume that the DS8000 storage subsystem (Element Manager) is already configured in TPC, as described in 12.2, “SSPC setup and configuration” on page 283.

- After the Element Manager is launched, click the disk system you want to access, as shown in Figure 13-4.

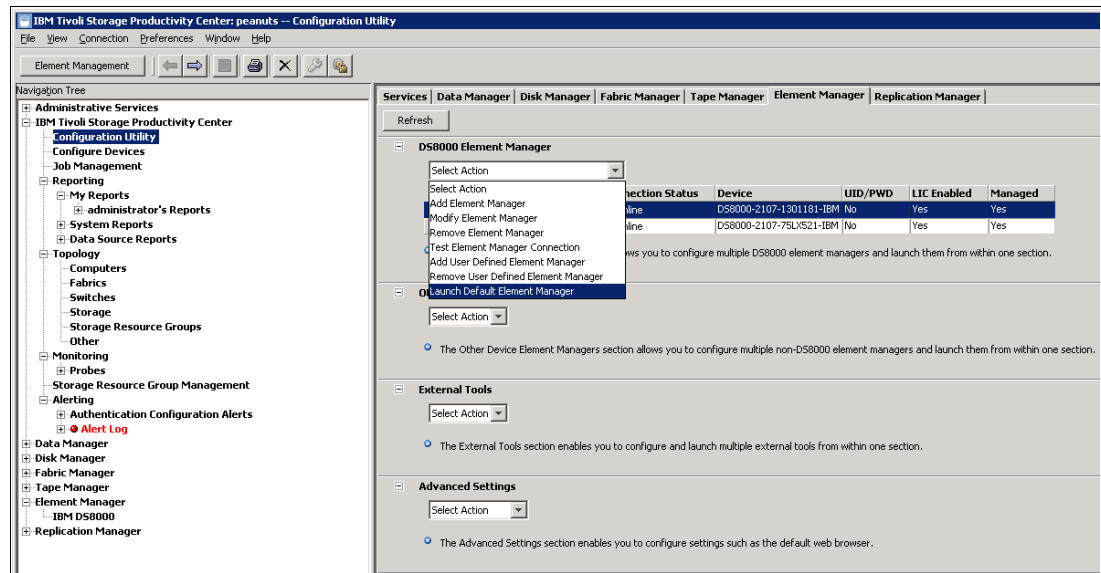


Figure 13-4 SSPC: Select the DS8000

- You are presented with the DS GUI Overview window for the selected disk system as shown in Figure 13-5.

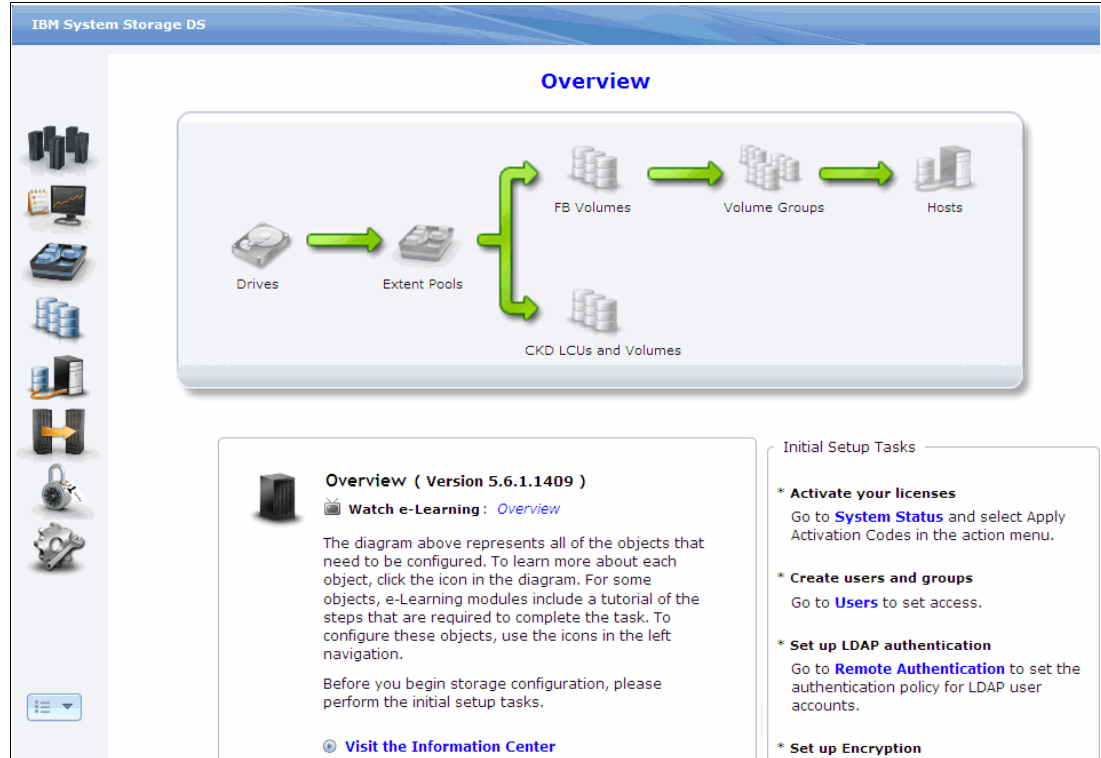


Figure 13-5 SSPC: DS GUI Overview window

Accessing the DS GUI from a browser connected to SSPC

To access the DS GUI, you can connect to SSPC using a web browser, and then use the instructions given in “Accessing the DS GUI through SSPC” on page 312.

Accessing the DS GUI from a browser connected to a TPC workstation

To access the DS GUI, you can connect to a TPC workstation using a web browser, and then use the instructions in “Accessing the DS GUI through SSPC” on page 312. For information about how to access a TPC workstation through a web browser, see 12.2.4, “Accessing the TPC on SSPC” on page 285.

Accessing the DS GUI through a remote desktop connection to SSPC

You can use remote desktop connection to SSPC. After you are connected to SSPC, follow the instructions in “Accessing the DS GUI through SSPC” on page 312 to access the DS GUI. For information how to connect to SSPC using remote desktop, see 12.2.4, “Accessing the TPC on SSPC” on page 285.

13.1.2 DS GUI Overview window

After you log on, the DS Storage Manager Overview window shown in Figure 13-5 on page 313 displays.

In this Overview window, you can see pictures and descriptions of DS8000 configuration components. Pictures in the shaded area can be clicked to view their description in the lower half of the window.

The left side of the window is the *navigation pane*.

DS GUI window options

Figure 13-6 shows an example of the Manage Volumes window. Several important options available on this page are also on many of the other windows of DS Storage Manager. We explain several of these options below.

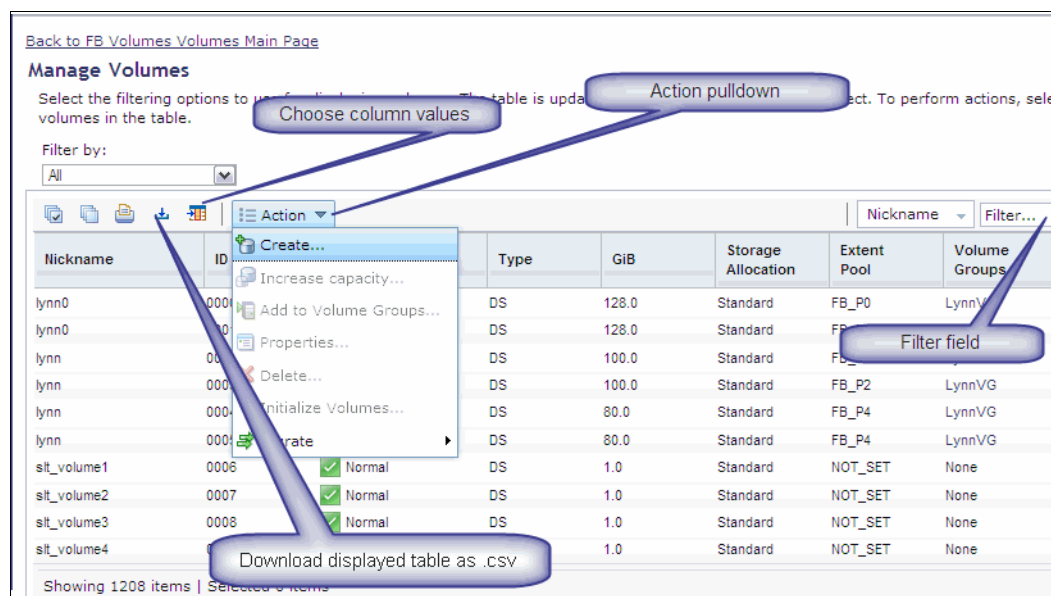


Figure 13-6 Example of the Extentpools window

The DS GUI displays the configuration of your DS8000 in tables. There are several options you can use:

- ▶ To download the information from the table, click **Download**. This can be useful if you want to document your configuration. The file is in comma-separated value (.csv) format and you can open the file with a spreadsheet program. This function is also useful if the table on the DS8000 Manager consists of several pages; the .csv file includes all pages.
- ▶ The **Print** report option opens a new window with the table in HTML format and starts the printer dialog box if you want to print the table.
- ▶ The **Action** drop-down menu provides you with specific actions that you can perform. Select the object you want to access and then the appropriate action (for example, Create or Delete).
- ▶ The **Choose Column Value** button sets and clears filters so that only specific items are displayed in the table (for example, show only FB extentpools in the table). This can be useful if you have tables with a large number of items.
- ▶ To search the table, type the desired criteria in the **Filter** field. The GUI displays entries in the table that match the criteria.

DS GUI navigation pane

The navigation pane, located on the left hand side of the window, allows you to navigate to the various functions of the DS8000 GUI. It has two views to choose from: icon view and original view. The default view is set to icon view, but you can change this by clicking the **Navigation Choice** button in the bottom part of the navigation pane. The two views are pictured in Figure 13-7 on page 316. When you move your mouse over one of the icons in icon view, the icon increases in size and displays the screens that you can navigate to. Figure 13-7 on page 316 shows the options presented when the mouse hovers over the each icon.

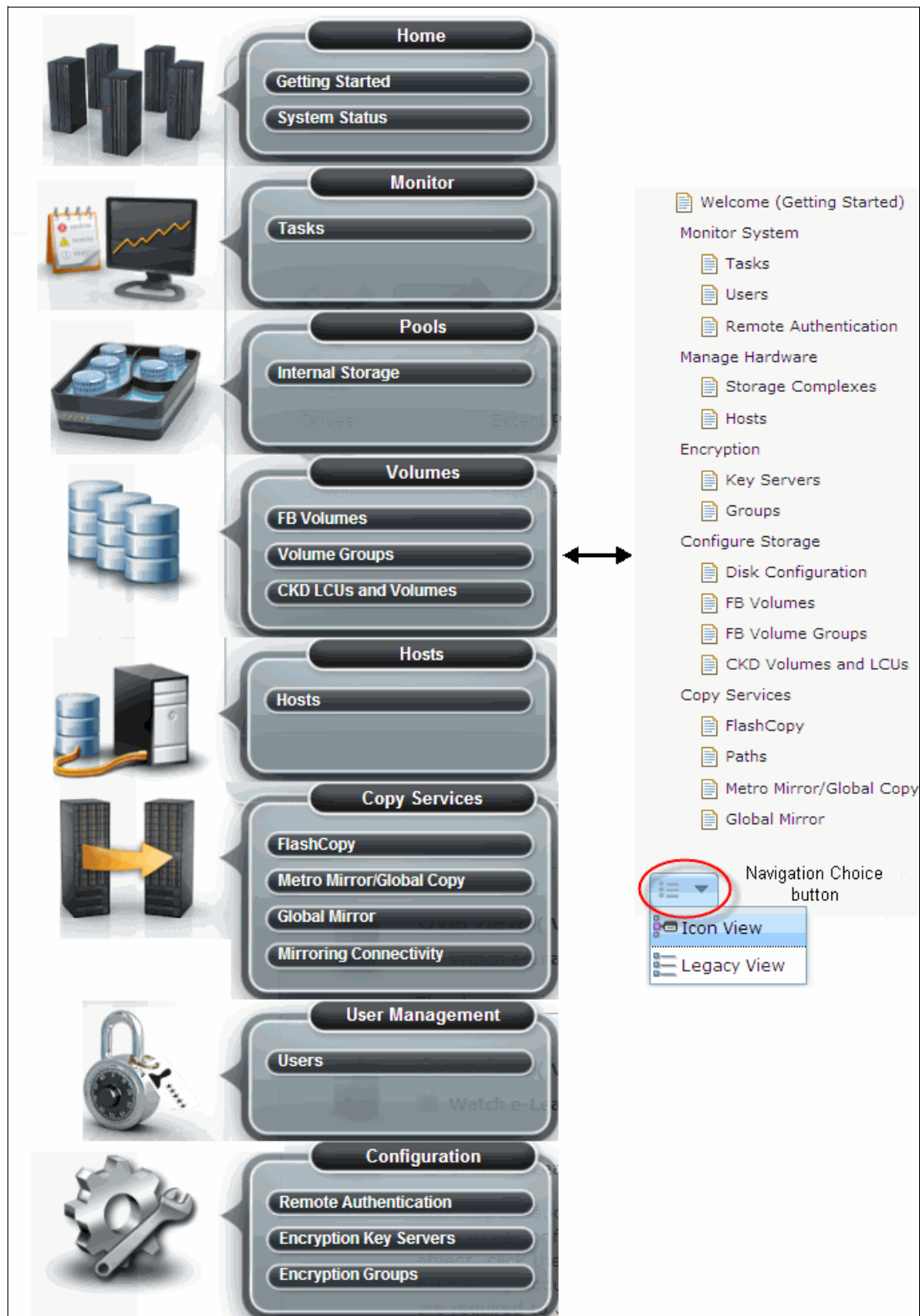


Figure 13-7 Navigation pane. Icon view with options on the left and original view on the right.

13.2 Logical configuration process

When performing the initial logical configuration, the first step is to create the storage complex (processor complex) along with the definition of the hardware of the storage unit.

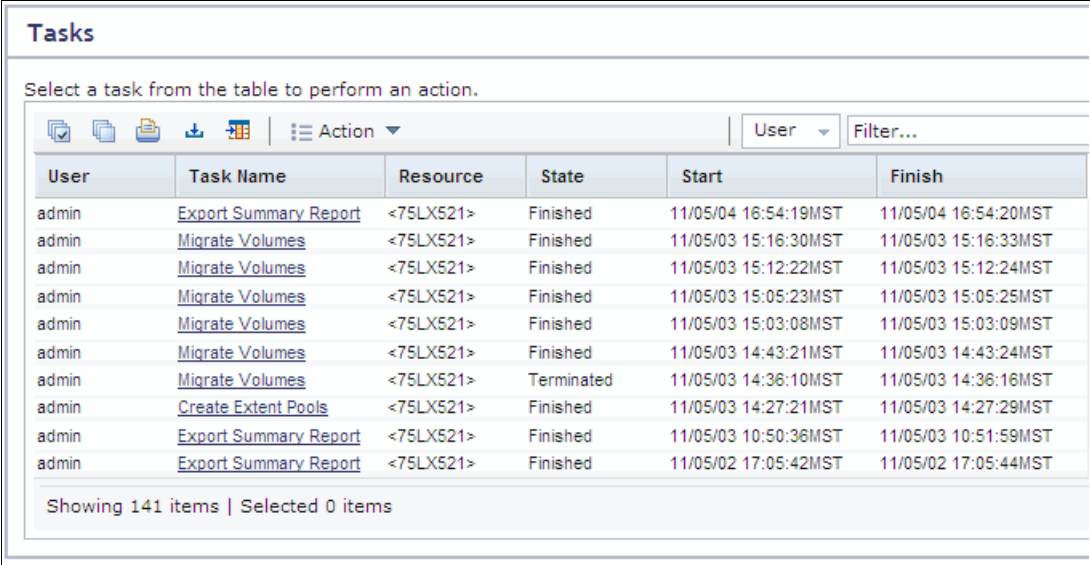
When performing the logical configuration, the following approach is suggested:

1. Start by defining the storage complex.
2. Create Extent Pools.
3. Create open system volumes.
4. Create count key data (CKD) LSSs and volumes.
5. Create host connections and volume groups.

Tasks summary window

Some logical configuration tasks have dependencies on the successful completion of other tasks. For example, you cannot create ranks on arrays until the array creation is complete. The Tasks summary window assists you in this process by reporting the progress and status of these long-running tasks.

Figure 13-8 shows the successful completion of the tasks. Click the specific task link to get more information about the task. The Task Summary window can be seen by hovering over **Monitor** and clicking **Tasks** in the navigation pane.



The screenshot shows a web interface titled "Tasks". Below the title is a prompt: "Select a task from the table to perform an action." There are icons for various actions (copy, paste, print, download, etc.) and a dropdown menu labeled "Action". To the right are "User" and "Filter..." fields. The main part of the window is a table with the following columns: User, Task Name, Resource, State, Start, and Finish. The table contains 10 rows of task data. At the bottom, it says "Showing 141 items | Selected 0 items".

User	Task Name	Resource	State	Start	Finish
admin	Export Summary Report	<75LX521>	Finished	11/05/04 16:54:19MST	11/05/04 16:54:20MST
admin	Migrate Volumes	<75LX521>	Finished	11/05/03 15:16:30MST	11/05/03 15:16:33MST
admin	Migrate Volumes	<75LX521>	Finished	11/05/03 15:12:22MST	11/05/03 15:12:24MST
admin	Migrate Volumes	<75LX521>	Finished	11/05/03 15:05:23MST	11/05/03 15:05:25MST
admin	Migrate Volumes	<75LX521>	Finished	11/05/03 15:03:08MST	11/05/03 15:03:09MST
admin	Migrate Volumes	<75LX521>	Finished	11/05/03 14:43:21MST	11/05/03 14:43:24MST
admin	Migrate Volumes	<75LX521>	Terminated	11/05/03 14:36:10MST	11/05/03 14:36:16MST
admin	Create Extent Pools	<75LX521>	Finished	11/05/03 14:27:21MST	11/05/03 14:27:29MST
admin	Export Summary Report	<75LX521>	Finished	11/05/03 10:50:36MST	11/05/03 10:51:59MST
admin	Export Summary Report	<75LX521>	Finished	11/05/02 17:05:42MST	11/05/02 17:05:44MST

Figure 13-8 Task Summary window

13.3 Examples of configuring DS8000 storage

In the following sections, we show an example of a DS8000 configuration made through the DS GUI.

For each configuration task (for example, creating an array), the process guides you through windows where you enter the necessary information. During this process, you have the ability to go back to make modifications or cancel the process. At the end of each process, you get a

verification window where you can verify the information that you entered before you submit the task.

13.3.1 Define storage complex

During the DS8000 installation, your IBM service representative customizes the setup of your storage complex based on information that you provide in the customization worksheets. After you log into the DS GUI and before you start the logical configuration, check the status of your storage system.

In the navigation pane of the DS GUI, hover over **Home** and click **System Status**. The System Status window opens as shown in Figure 13-9.

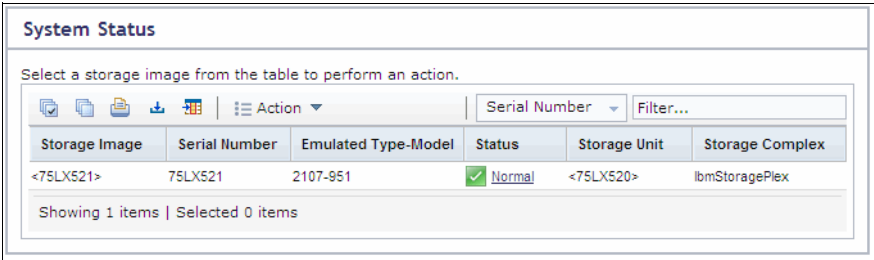


Figure 13-9 System Status window

You should have at least one storage complex listed in the table. If you have more than one DS8000 system in your environment connected to the same network, you can define it here by adding a new storage complex. Select **Storage Complex** → **Add** from the **Action** drop-down menu to add a new storage complex (Figure 13-10).

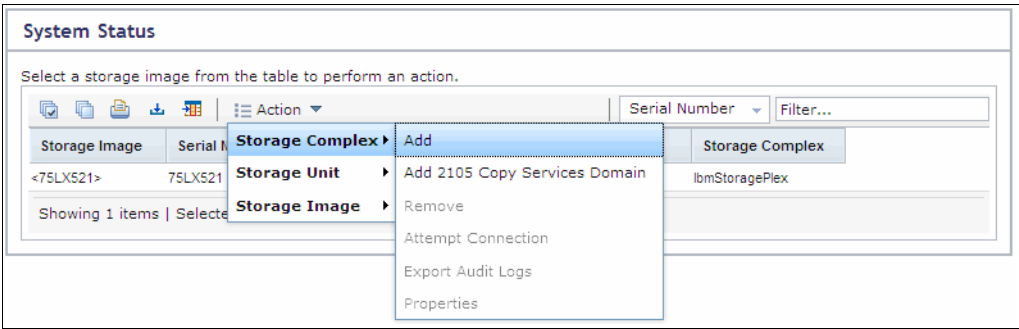


Figure 13-10 Add Storage Complex window

The Add Storage Complex window opens as shown in Figure 13-11.

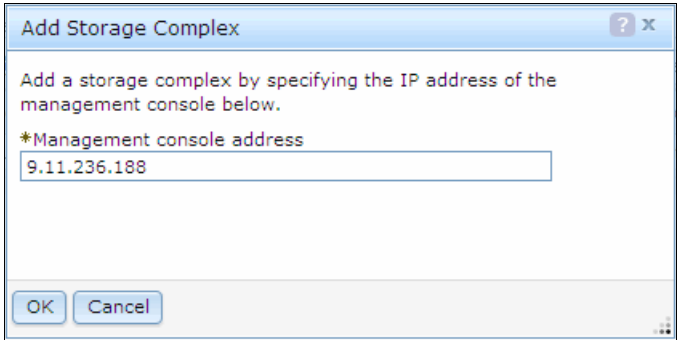


Figure 13-11 Add Storage Complex window

Provide the IP address of the Hardware Management Console (HMC) connected to the new storage complex that you want to add and click **OK** to continue. A new storage complex is added to the table as shown in Figure 13-12.

System Status

Select a storage image from the table to perform an action.

Action

Serial Number

Filter...

Storage Image	Serial Number	Emulated Type-Model	Status	Storage Unit	Storage Complex
<75LX521>	75LX521	2107-054	Normal	<75LX520>	IbmStoragePlex
<1301181>	1301181	2107-941	Normal	<1301180>	IbmStoragePlex

Showing 2 items | Selected 0 items

Figure 13-12 New storage complex is added

Having all the DS8000 storage complexes defined together provides flexible control and management. The status information indicates the healthiness of each storage complex. By clicking the status description link of any storage complex, you can obtain more detailed health check information for various vital DS8000 components (Figure 13-13).

Status

Below is a summary of the respective statuses for the storage complex, storage unit, and storage image.

Connected

Storage complex console 1 (mtc520h/9.11.240.62)

Connected

Storage complex console 2 ()

None

Powered on

Online

Storage Server 0

Online

Storage Server 1

Online

OK

Figure 13-13 Check the status details

Status descriptions can be reported for your storage complexes. These descriptions depend on the availability of the vital storage complexes components. In Figure 13-14, we show an example of various status states.

Storage Complexes Summary

Select a storage image from the table to perform an action.

Select action▼

Select	Storage Image ^	Serial Number ^	Emulated Type-Model ^	Status ^	Storage Unit ^	Storage Complex ^
☐	DS8k_TIC06v1_ATS	75ABTV1	2107-9A2	Attention	ATS_ABTV0	ATS_Mainz_ABTV0
☐	DS8k_TIC06v2_SLE	75ABTV2	2107-9A2	Critical	ATS_ABTV0	ATS_Mainz_ABTV0
☐	ATS_04	75TV181	2107-951	Normal	<75TV180>	DS8K-ITSO
☐	ATS1	7503461	2107-922	Normal	ATS1	IbmStoragePlex
☐	ATS_20780	7520781	2107-922	Normal	<7520780>	IbmStoragePlex

Showing 1 - 5 of 5 Selected 0

Figure 13-14 Different Storage Complex Status states

A Critical status indicates unavailable vital storage complex resources. An Attention status might be triggered by resources being unavailable. Because the DS8000 has redundant components, the storage complex is still operational. One example is when only one storage server inside a storage image is offline as shown in Figure 13-15.

Status

Below is a summary of the respective statuses for the storage complex, storage unit, and storage image.

Storage complex (IbmStoragePlex)

Connected

Storage complex console 1 (m1540h/9.11.236.19)

Connected

Storage complex console 2 ()

None

Storage unit (<75GH900>)

Powered on

Storage image (<75GH901>)

Online

Storage Server 0

Online

Storage Server 1

Offline

OK

Cancel

Figure 13-15 One storage server is offline

Check the status of your storage complex and proceed with logical configuration (create arrays, ranks, Extent Pools, or volumes) only when your HMC consoles are connected to the storage complex, and both storage servers inside the storage image are online and operational.

13.3.2 Create arrays

Tip: You do not necessarily need to create arrays first and then ranks. You can proceed directly with the creation of Extent Pools, as explained in 13.3.4, “Create Extent Pools” on page 329.

To create an array, perform the following steps in the DS GUI:

1. In the GUI, from the navigation pane, hover over **Pools** and click **Internal Storage**. This brings up the Internal Storage window (Figure 13-16).

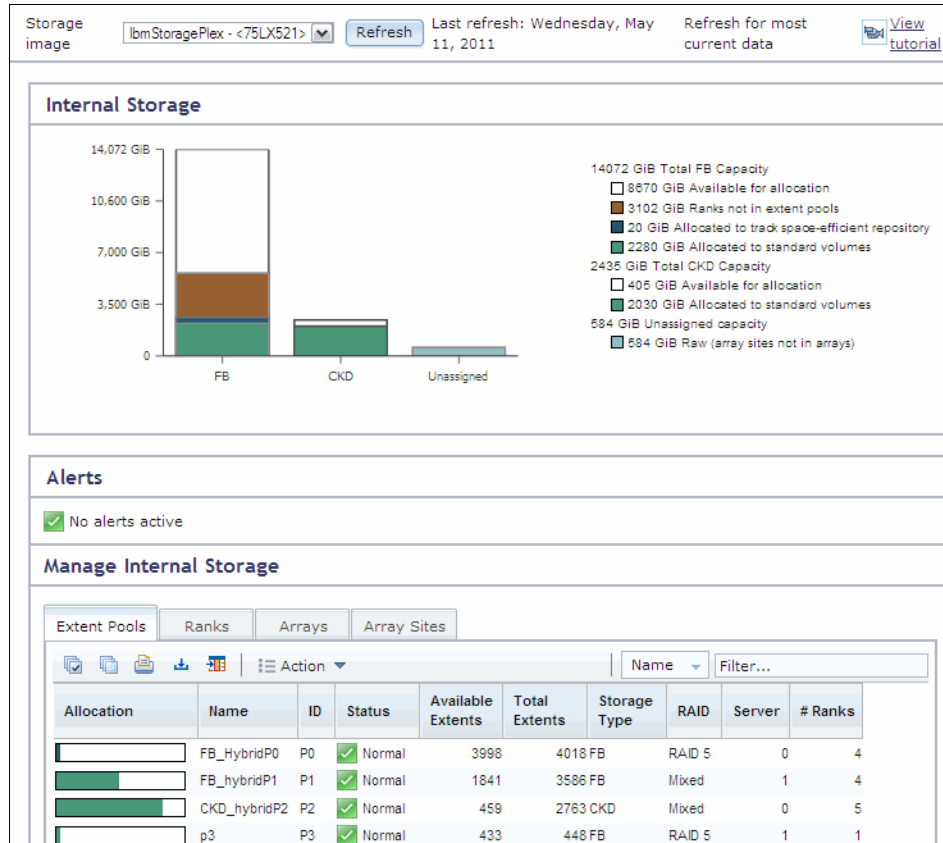
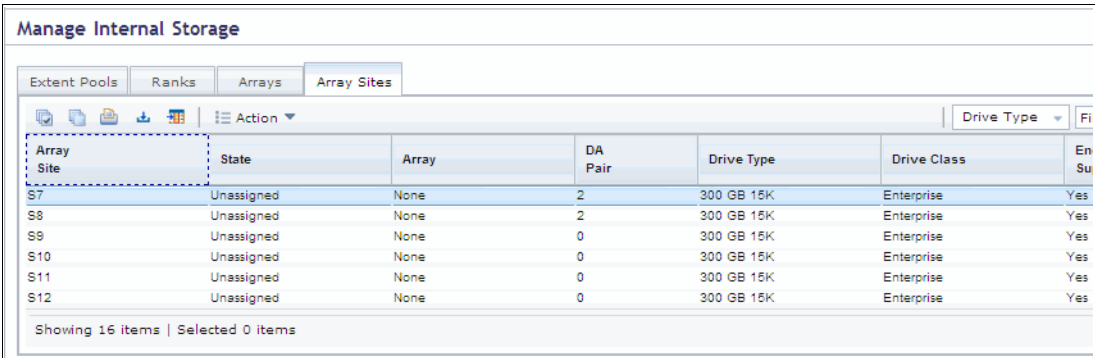


Figure 13-16 Disk Configuration window

Tip: If you have defined more storage complexes or storage images, be sure to select the right storage image before you start creating arrays. From the **Storage image** drop-down menu, select the desired storage image you want to access.

In our example, some of the DS8000 capacity is assigned to open systems, some is assigned to System z, and some is not assigned at all.

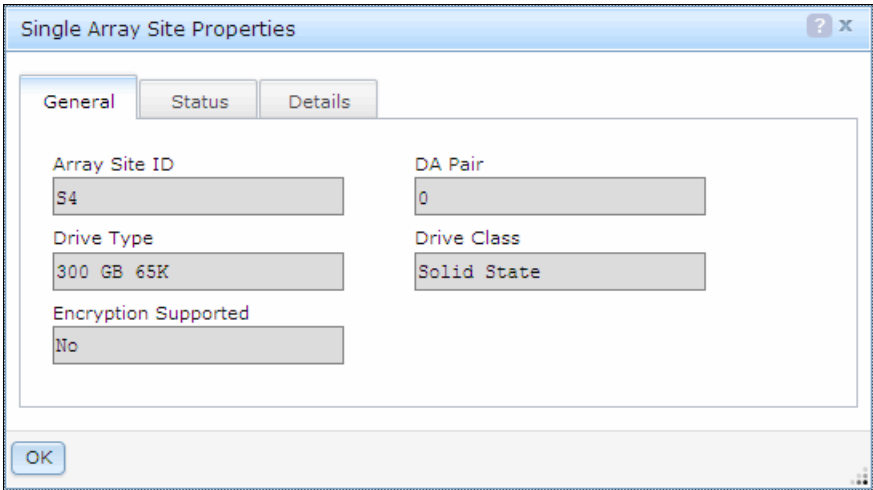
2. Click the **Array Sites** tab to check the available storage that is required to create the array (Figure 13-17).



Array Site	State	Array	DA Pair	Drive Type	Drive Class	En Su
S7	Unassigned	None	2	300 GB 15K	Enterprise	Yes
S8	Unassigned	None	2	300 GB 15K	Enterprise	Yes
S9	Unassigned	None	0	300 GB 15K	Enterprise	Yes
S10	Unassigned	None	0	300 GB 15K	Enterprise	Yes
S11	Unassigned	None	0	300 GB 15K	Enterprise	Yes
S12	Unassigned	None	0	300 GB 15K	Enterprise	Yes

Figure 13-17 Array sites

3. In our example, some array sites are unassigned and therefore eligible to be used for array creation. Each array site has eight physical disk drives. to discover more details about each array site, select the desired array site and click **Properties** under the Action drop-down menu. The Single Array Site Properties window opens. It provides general array site characteristics, as shown in Figure 13-18.



Single Array Site Properties

General Status Details

Array Site ID: S4

DA Pair: 0

Drive Type: 300 GB 65K

Drive Class: Solid State

Encryption Supported: No

OK

Figure 13-18 Select Array Site Properties

- Click the **Status** tab to get more information about the Disk Drive Modules (DDMs) and the state of each DDM, as shown in Figure 13-19.

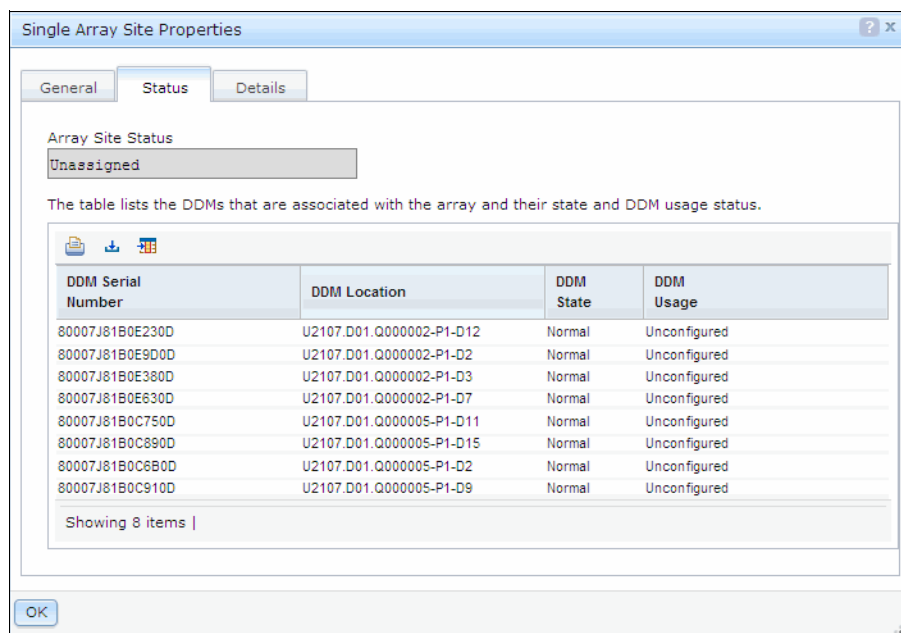


Figure 13-19 Single Array Site Properties: Status

- All DDMs in this array site are in the Normal state. Click **OK** to close the Single Array Site Properties window and go back to the Internal Storage main window.
- After we identify the unassigned and available storage, we can create an array. Click the **Array** tab in the Manage Disk Configuration section and select **Create Arrays** in the **Action** drop-down menu as shown in Figure 13-20.

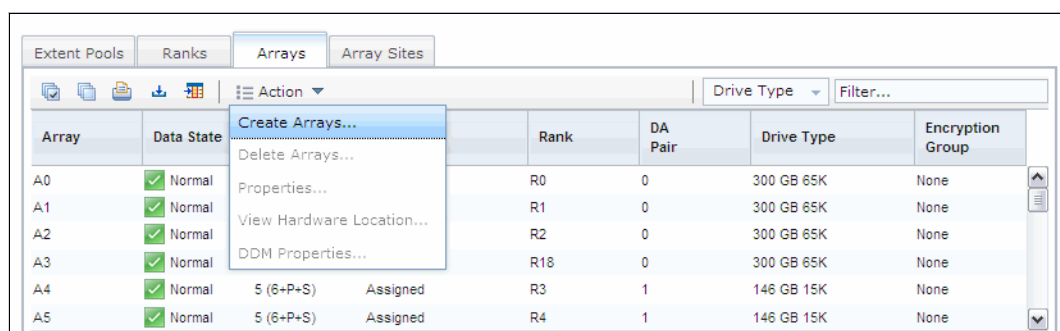


Figure 13-20 Select Create Arrays

The Create New Arrays window opens, as shown in Figure 13-21.

Create New Arrays

Define Storage Characteristics

Choose the RAID type that you want to configure.

RAID Type
RAID 5

Select Available Capacity

The storage resources that you select in this section are used to create the arrays.

Type of Configuration
Automatic

Drive Class
300 GB 65K Solid State drive with 1.5 - 1.8 TiB

Select capacity to configure
1 arrays (1.5 - 1.8 TiB)

DA Pair Usage
Spread among all pairs

Bar graph showing DA Pair Usage (0 to 3) with New Ranks (blue) and Assigned Ranks (green).

DA Pair	New Ranks	Assigned Ranks
0	4	6
1	4	6
2	4	8
3	4	6

Figure 13-21 Create New Arrays window

You need to provide the following information:

- RAID Type: The available RAID types are RAID 5 (default), RAID 6, and RAID 10 for HDDs. Only RAID 5 is available for SSDs.
- Type of configuration: There are two options available:
 - Automatic is the default, and it allows the system to choose the best array site configuration based on your capacity and DDM type.
 - The Manual option can be used if you want to have more control over the resources. When you select this option, a table of available array sites is displayed. You have to manually select array sites from the table.
- If you select the **Automatic** configuration type, you need to provide additional information:
 - From the **DA Pair Usage** drop-down menu, select the appropriate action. The **Spread Among All Pairs** option balances arrays evenly across all available Device Adapter (DA) pairs. The **Spread Among Least Used Pairs** option assigns the array to the least used DA pairs. The **Sequentially Fill All Pairs** option assigns arrays to the first DA pair, then to the second DA pair, and so on. The bar graph displays the effect of your choice.
 - From the **Drive Class** drop-down menu, select the DDM type you want to use for the new array.
 - From the **Select Capacity to Configure** list, select the desired total capacity.

If you want to create many arrays with different characteristics (RAID and DDM type) in one task, select **Add Another Array** as many times as required.

In our example (Figure 13-21), we created one RAID 5 array on 300 GB SSDs.

Click **OK** to continue.

- The Create array verification window is displayed (Figure 13-22). It lists all array sites chosen for the new arrays we want to create. At this stage, you can still change your configuration by deleting the array sites from the lists and adding new array sites if required. Click **Create All** after you decide to continue with the proposed configuration.

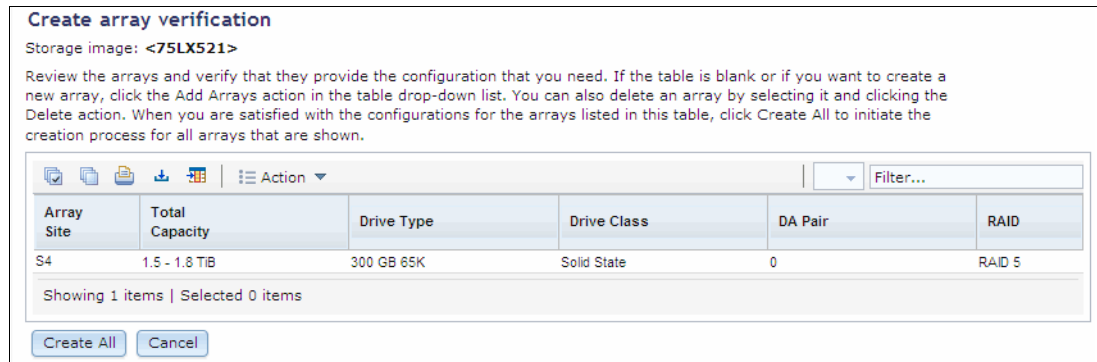


Figure 13-22 Create array verification window

Wait for the message in Figure 13-23 to appear and then click **Close**.

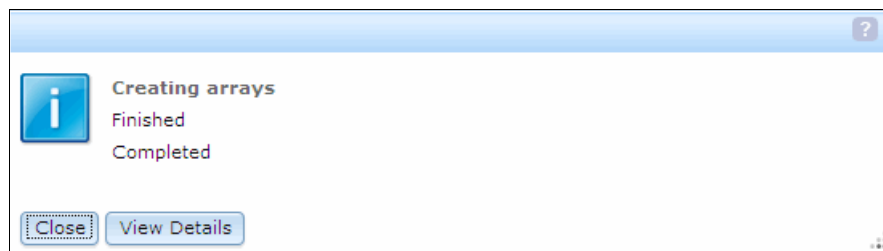


Figure 13-23 Creating arrays completed

- The graph in the Internal Storage summary section has changed to reflect the arrays that were configured.

13.3.3 Create ranks

Tip: You do not necessarily need to create arrays first and then ranks. You can proceed directly with the creation of Extent Pools (see 13.3.4, “Create Extent Pools” on page 329).

To create a rank, perform the following steps in the DS GUI:

1. In the GUI, from the navigation pane, hover over **Pools** and click **Internal Storage**. This brings up the Internal Storage window. Click the Ranks tab to start working with ranks. Select **Create Rank** from the **Action** drop-down menu as shown in Figure 13-24.

Tip: If you have defined more storage complexes/storage images, be sure to select the right storage image before you start creating ranks. From the **Storage image** drop-down menu, select the desired storage image you want to access.

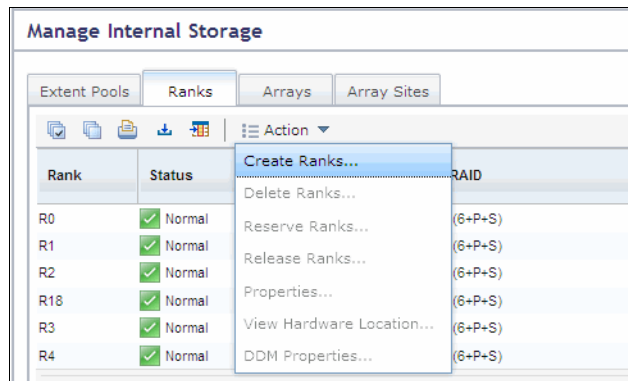


Figure 13-24 Select Create Ranks

2. The Create New Ranks window opens (Figure 13-25).

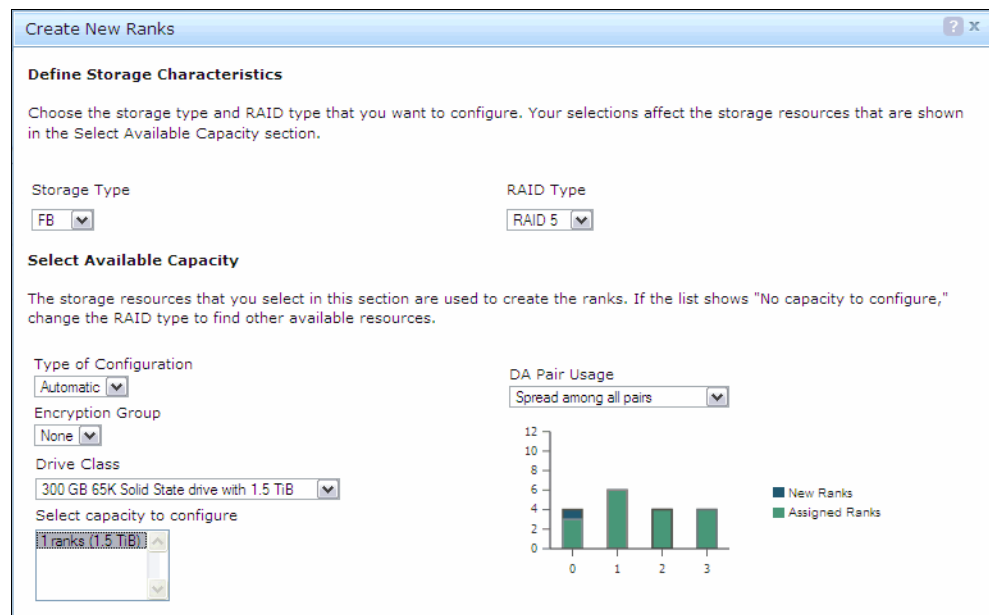


Figure 13-25 Create New Ranks window

To create a rank, you have to provide the following information:

- **Storage Type:** The type of extent for which the rank is to be configured. The storage type can be set to one of the following values:
 - **Fixed block (FB) extents = 1 GB.** In fixed block architecture, the data (the logical volumes) is mapped over fixed-size blocks or sectors.
 - **Count key data (CKD) extents = CKD Mod 1.** In count-key-data architecture, the data field stores the user data.
- **RAID Type:** The available RAID types are RAID 5 (default), RAID 6, and RAID 10 for HDDs. Only RAID 5 is available for SSDs.
- **Type of configuration:** There are two options available:
 - **Automatic** is the default and it allows the system to choose the best configuration of the physical resources based on your capacity and DDM type.
 - The **Manual** option can be used if you want to have more control over the resources. When you select this option, a table of available array sites is displayed. You then manually select resources from the table.
- **Encryption Group** indicates if encryption is enabled or disabled for ranks. Select **1** from the **Encryption Group** drop-down menu if the encryption feature is enabled on this machine. Otherwise, select **None**.
- If you select the Automatic configuration type, you need to provide additional information:
 - From the **DA Pair Usage** drop-down menu, select the appropriate action. The **Spread Among All Pairs** option balances arrays evenly across all available Device Adapter (DA) pairs. The **Spread Among Least Used Pairs** option assigns the array to the least used DA pairs. The **Sequentially Fill All Pairs** option assigns arrays to the first DA pair, then to the second DA pair, and so on. The bar graph displays the effect of your choice.
 - From the **Drive Class** drop-down menu, select the DDM type you want to use for the new array.
 - From the **Select capacity to configure** list, select the desired total capacity.

If you want to create many ranks with different characteristics (Storage, RAID, and DDM type) at one time, select **Add Another Rank** as many times as required.

In our example, we create one FB rank on 300 GB SSDs with RAID 5.

Click **OK** to continue.

3. The Create rank verification window is displayed (Figure 13-26). Each array site listed in the table is assigned to the corresponding array we created in 13.3.2, “Create arrays” on page 321. At this stage, you can still change your configuration by deleting the ranks from the lists and adding new ranks if required. Click **Create All** after you decide to continue with the proposed configuration.

Create rank verification

Storage image: <75LX521>

Review the ranks and verify that they provide the configuration that you need. If the table is blank or if you want to create a new rank, click the Add Ranks action in the table drop-down list. You can also delete a rank by selecting it and clicking the Delete action. When you are satisfied with the configurations for the ranks listed in this table, click Create All to initiate the creation process for all ranks that are shown.

Array Site	Array	Total Capacity	Drive Type	Drive Class	DA Pair	RAID	Storage Type	Encryption Group
S4	A3	1.5 TiB	300 GB 65K	Solid State	0	RAID 5	FB	None

Showing 1 items | Selected 0 items

Create All Cancel

Figure 13-26 Create rank verification window

4. The Creating Ranks window appears. Click the **View Details** button to check the overall progress. It displays the Task Properties window shown in Figure 13-27.

Task Properties

You can view details about the task, view log file details, and save log files.

Task Name Creating ranks	Task Type Real-time
User admin	Resource <75LX521>
State Finished	Status Completed
Start 11/04/26 11:52:15MST	Finish 11/04/26 11:52:37MST

All task details loaded.

Task Details

```
Tue Apr 26 11:52:35 MST 2011 - - Rank R18 is now in assignable state
Tue Apr 26 11:52:35 MST 2011 - Query ranks state until they are in an assignable state (1 of 1 ranks in assignable state) -
Tue Apr 26 11:52:35 MST 2011 - Querying array site data -
Tue Apr 26 11:52:35 MST 2011 - Querying array data -
Tue Apr 26 11:52:36 MST 2011 - Querying rank data -
Tue Apr 26 11:52:36 MST 2011 - Querying extent pool data -
Tue Apr 26 11:52:36 MST 2011 - Querying rank usage data -
Tue Apr 26 11:52:37 MST 2011 - Completed -
```

Close Save to File

Figure 13-27 Creating ranks: Task Properties view

5. After the task is completed, go back to Internal Storage and, under the Rank tab, check the list of newly created ranks.

The bar graph in the Disk Configuration Summary section has changed. There are new ranks, but they are not assigned to Extent Pools yet.

13.3.4 Create Extent Pools

To create an Extent Pool, perform the following steps in the DS GUI:

1. In the GUI, from the navigation pane, hover over **Pools** and click **Internal Storage**. This opens the Internal Storage window. Click the Extent Pool tab.

The bar graph in the summary section provides information about unassigned and assigned capacity.

Select **Create Extent Pools** from the **Action** drop-down menu as shown in Figure 13-28.

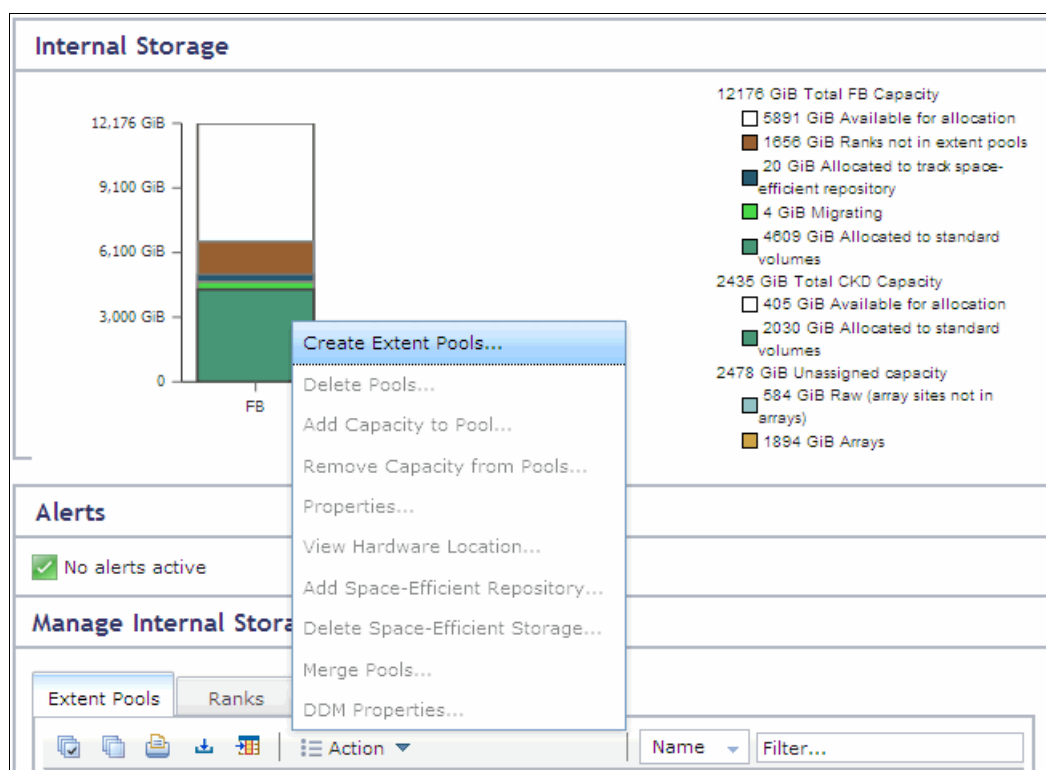


Figure 13-28 Select Create Extent Pools

Tip: If you have defined more storage complexes or storage images, be sure to select the correct storage image before you create Extent Pools. From the **Storage image** drop-down menu, select the desired storage image you want to access.

2. The Create New Extent Pools window displays as shown in Figure 13-29. Scroll down to see the rest of the window and provide input for all the fields.

Create New Extent Pools

Define Storage Characteristics

Choose the storage type and RAID type that you want to configure. Your selections affect the storage resources that are shown in the Select Available Capacity section.

Storage Type: **FB** RAID Type: **RAID 5**

Select Available Capacity

The storage resources that you select in this section are used to create the extent pools. If the list shows "No capacity to configure," change the RAID type or storage type to find other available resources.

Type of Configuration: **Automatic** Encryption Group: **None** Drive Class: **73 GB 15K Enterprise drive with 3.2 - 3.4 TiB**

Select capacity to configure:

- 1 ranks (384 GB)
- 2 ranks (832 GB)
- 3 ranks (1.3 TiB)
- 4 ranks (1.7 TiB)

DA Pair Usage

Spread among all pairs

Define Extent Pool Characteristics

Choose the number of extent pools to create using the resources that you selected above. The ideal configuration creates 2 extent pools, dividing the resources equally among each pool.

Number of extent pools: **Two extent pools (ease of management)**

Nickname prefix: **FB_** Suffix: **ID** Start: **0**

Nickname example: FB_P0, FB_P1

*Storage Threshold: **100** %

*Storage Reserved: **0** %

Server assignment

Automatic

Server 0

Current: 4 ranks
Added: 0 ranks
Total: 4 ranks

Server 1

Current: 0 ranks
Added: 0 ranks
Total: 0 ranks

OK Add Another Pool Cancel

Figure 13-29 Create New Extent Pools window

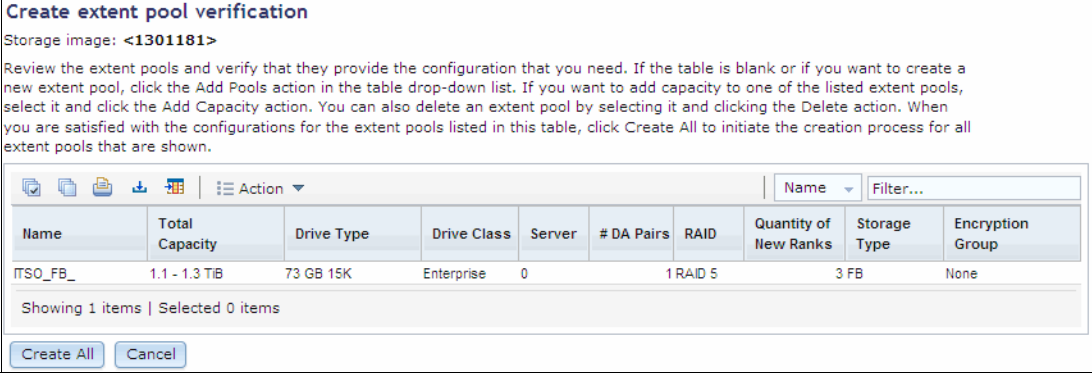
To create an Extent Pool, you must provide the following information:

- Storage Type: The type of extent for which the rank is to be configured. The storage type can be set to one of the following values:
 - Fixed block (FB) extents = 1 GB. In the fixed block architecture, the data (the logical volumes) is mapped over fixed-size blocks or sectors.
 - Count key data (CKD) extents = CKD Mod 1. In the count-key-data architecture, the data field stores the user data.
- RAID Type: The available RAID types are RAID 5 (default), RAID 6, and RAID 10 for HDDs. Only RAID 5 is available for SSDs.

- Type of configuration: There are two options available:
 - **Automatic** is the default and it allows the system to choose the best configuration of physical resources based on your capacity and DDM type.
 - The **Manual** option can be used if you want to have more control over the resources. When you select this option, a table of available array sites is displayed. You have to manually select resources from this table.
 - Encryption Group indicates if encryption is enabled or disabled for ranks. Select **1** from the **Encryption Group** drop-down menu, if the encryption feature is enabled on the machine. Otherwise, select **None**.
 - If you select the **Automatic** configuration type, you need to provide additional information:
 - From the **DA Pair Usage** drop-down menu, select the appropriate action. The **Spread Among All Pairs** option balances ranks evenly across all available Device Adapter (DA) pairs. For example, no more than half of the ranks attached to a DA pair are assigned to each server, so that each server's DA within the DA pair has the same number of ranks. The **Sequentially Fill All Pairs** option assigns arrays to the first DA pair then to the second DA pair etc. The bar graph displays the effect of your choice. The bar graph displays the effects of your choices.
 - From the **Drive Class** drop-down menu, select the DDM type you want to use for the new array.
 - From the Select capacity to configure list, select the desired total capacity.
 - Number of Extent Pools: Here you choose the number of Extent Pools to create. There are three available options, **Two Extent Pools** (ease of management), **Single Extent Pool**, and **Extent Pool** for each rank (physical isolation). The default configuration creates two Extent Pools per storage type, dividing all ranks equally among each pool.
 - Nickname Prefix and Suffix: Provides a unique name for each Extent Pool. This setup is useful if you have multiple Extent Pools, each assigned to separate hosts and platforms.
 - Server assignment: The **Automatic** option allows the system to determine the best server for each Extent Pool. It is the only choice when you select the **Two Extent Pool** option as the number of Extent Pools.
 - Storage Threshold: Specifies the percentage when the DS8000 will generate a storage threshold alert. This allows you to make adjustments before a full storage condition occurs.
 - Storage reserved: Specifies the percentage of the total Extent Pool capacity that is reserved. This percentage is prevented from being allocated to volumes or space-efficient storage.
3. If you have both the FB and CKD storage type, or have different types of DDMs installed, you need to create more Extent Pools accordingly. To create all the required Extent Pools in one task, select **Add Another Pool** as many times as required.

Click **OK** to continue.

- The Create Extent Pool verification window opens (Figure 13-30). Here you can check the names of the Extent Pools that are going to be created, their capacity, server assignments, RAID protection, and other information. If you want to add capacity to the Extent Pools or add another Extent Pool, select the appropriate action from the **Action** drop-down list. After you are satisfied with the specified values, click **Create all** to create the Extent Pools.



Create extent pool verification

Storage image: <1301181>

Review the extent pools and verify that they provide the configuration that you need. If the table is blank or if you want to create a new extent pool, click the Add Pools action in the table drop-down list. If you want to add capacity to one of the listed extent pools, select it and click the Add Capacity action. You can also delete an extent pool by selecting it and clicking the Delete action. When you are satisfied with the configurations for the extent pools listed in this table, click Create All to initiate the creation process for all extent pools that are shown.

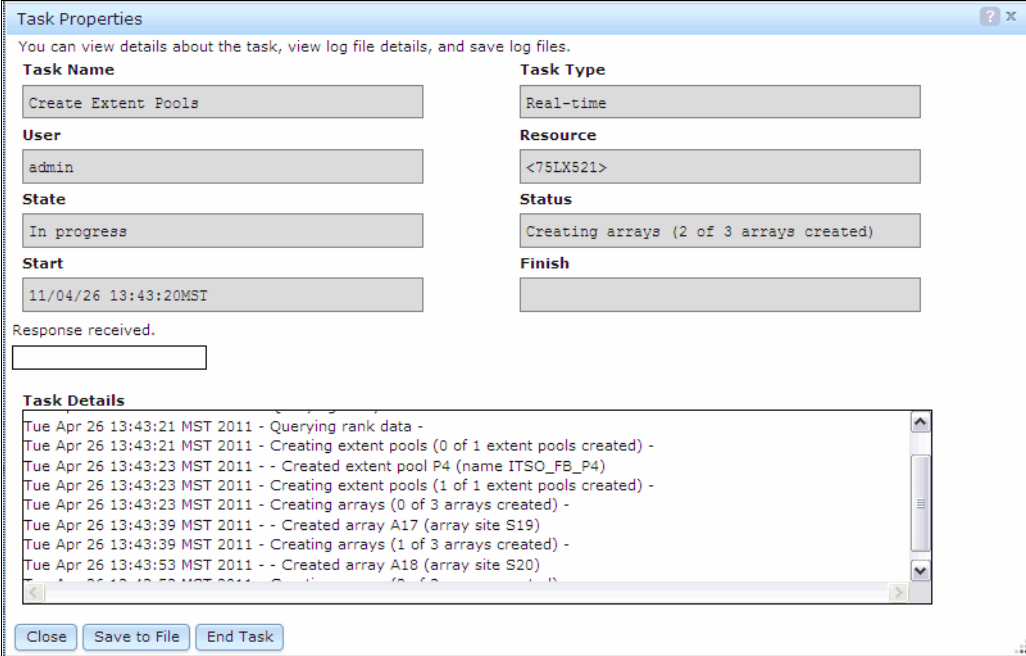
Name	Total Capacity	Drive Type	Drive Class	Server	# DA Pairs	RAID	Quantity of New Ranks	Storage Type	Encryption Group
ITSO_FB_	1.1 - 1.3 TiB	73 GB 15K	Enterprise	0	1	RAID 5	3	FB	None

Showing 1 items | Selected 0 items

Create All **Cancel**

Figure 13-30 Create Extent Pool verification window

- The Creating Extent Pools window appears. Click the **View Details** button to check the overall progress. It displays the Task Properties window shown in Figure 13-31.



Task Properties

You can view details about the task, view log file details, and save log files.

Task Name Create Extent Pools	Task Type Real-time
User admin	Resource <75LX521>
State In progress	Status Creating arrays (2 of 3 arrays created)
Start 11/04/26 13:43:20MST	Finish

Response received.
[]

Task Details

```
Tue Apr 26 13:43:21 MST 2011 - Querying rank data -
Tue Apr 26 13:43:21 MST 2011 - Creating extent pools (0 of 1 extent pools created) -
Tue Apr 26 13:43:23 MST 2011 - - Created extent pool P4 (name ITSO_FB_P4)
Tue Apr 26 13:43:23 MST 2011 - Creating extent pools (1 of 1 extent pools created) -
Tue Apr 26 13:43:23 MST 2011 - Creating arrays (0 of 3 arrays created) -
Tue Apr 26 13:43:39 MST 2011 - - Created array A17 (array site S19)
Tue Apr 26 13:43:39 MST 2011 - Creating arrays (1 of 3 arrays created) -
Tue Apr 26 13:43:53 MST 2011 - - Created array A18 (array site S20)
```

Close **Save to File** **End Task**

Figure 13-31 Creating extent pools: Task Properties window

- After the task is completed, return to the Internal Storage screen, under the Extent Pools tab, and check the list of newly created ranks.

The bar graph in the summary section has changed. There are ranks assigned to Extent Pools and you can create new volumes from each Extent Pool.

- The options available from the **Action** drop-down menu are shown in Figure 13-32. To check the Extent Pool properties, select the desired Extent Pool and, from the **Action** drop-down menu, click **Properties**.

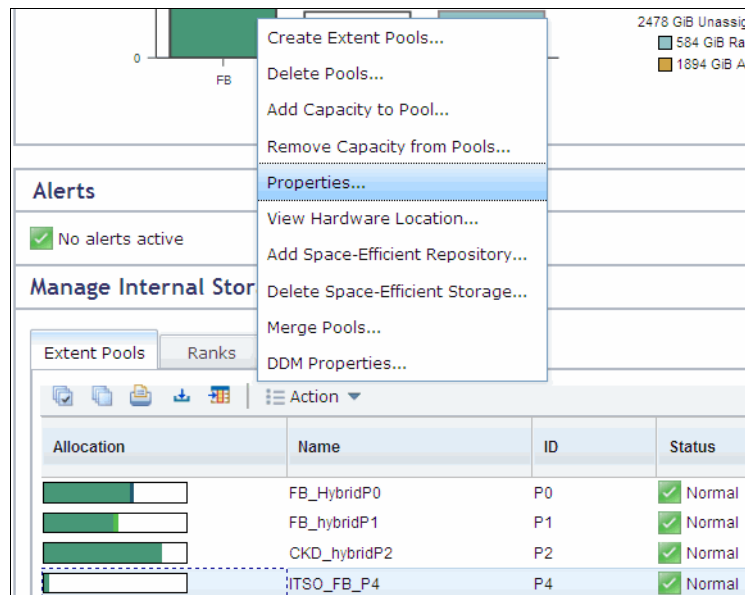


Figure 13-32 Extent pool action Properties

- The Single Pool properties window opens (Figure 13-33). Basic Extent Pool information is provided here and volume relocation related information. You can, if necessary, change the Extent Pool Name, Storage Threshold, and Storage Reserved values and select **Apply** to commit all the changes.

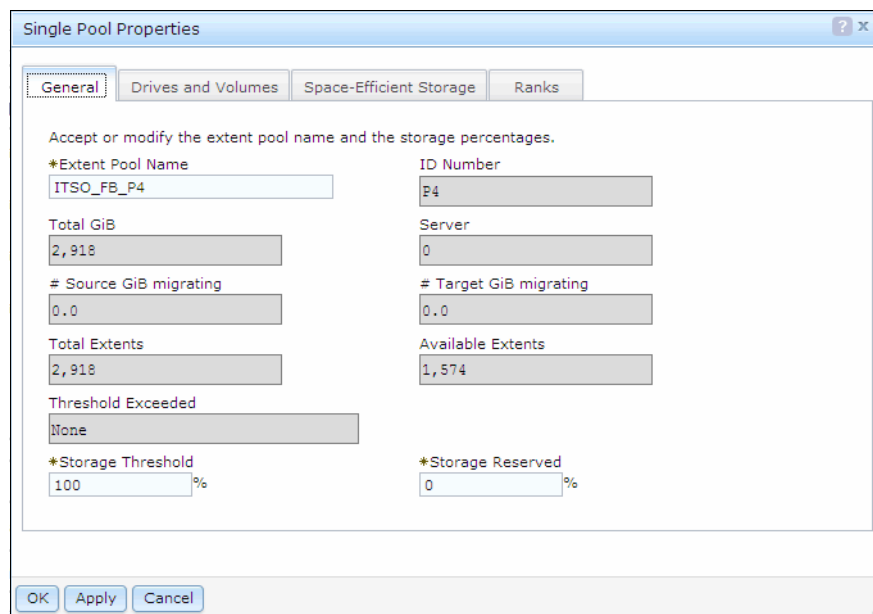
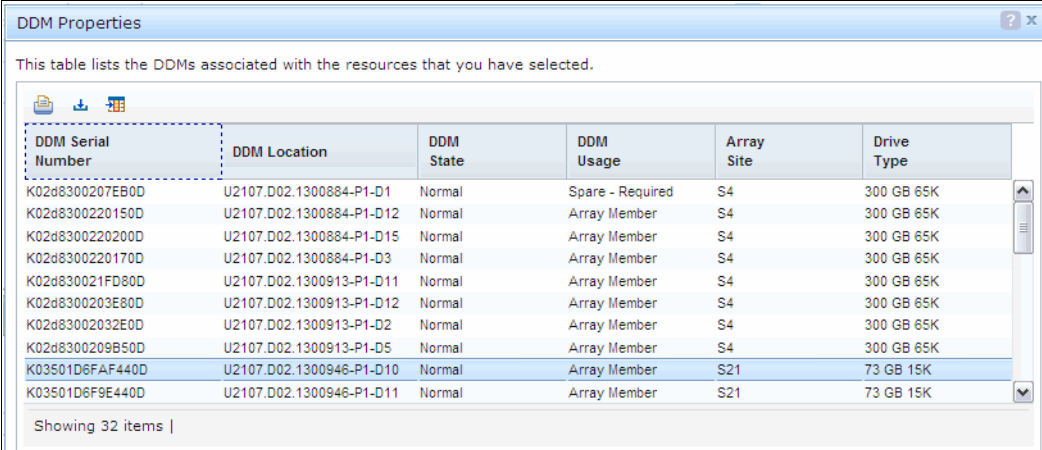


Figure 13-33 Single Pool Properties: General tab

- For more information about drive types or ranks included in the Extent Pool, click the appropriate tab. Click **OK** to return to the Internal Storage window.
- To discover more details about the DDMs, select the desired Extent Pool from the Manage Internal Storage table and, from the **Action** drop-down menu, click **DDM Properties**. The DDM Property window appears as shown in Figure 13-34.



This table lists the DDMs associated with the resources that you have selected.

DDM Serial Number	DDM Location	DDM State	DDM Usage	Array Site	Drive Type
K02d8300207EB0D	U2107.D02.1300884-P1-D1	Normal	Spare - Required	S4	300 GB 65K
K02d8300220150D	U2107.D02.1300884-P1-D12	Normal	Array Member	S4	300 GB 65K
K02d8300220200D	U2107.D02.1300884-P1-D15	Normal	Array Member	S4	300 GB 65K
K02d8300220170D	U2107.D02.1300884-P1-D3	Normal	Array Member	S4	300 GB 65K
K02d830021FD80D	U2107.D02.1300913-P1-D11	Normal	Array Member	S4	300 GB 65K
K02d8300203E80D	U2107.D02.1300913-P1-D12	Normal	Array Member	S4	300 GB 65K
K02d83002032E0D	U2107.D02.1300913-P1-D2	Normal	Array Member	S4	300 GB 65K
K02d8300209B50D	U2107.D02.1300913-P1-D5	Normal	Array Member	S4	300 GB 65K
K03501D6FAF440D	U2107.D02.1300946-P1-D10	Normal	Array Member	S21	73 GB 15K
K03501D6F9E440D	U2107.D02.1300946-P1-D11	Normal	Array Member	S21	73 GB 15K

Showing 32 items |

Figure 13-34 Extent Pool: DDM Properties

Use the DDM Properties window to view all the DDMs that are associated with the selected Extent Pool and to determine the DDMs' state. You can print the table, download it in .csv format, and modify the table view by selecting the appropriate icon at the top of the table.

Click **OK** to return to the Internal Storage window.

13.3.5 Configure I/O ports

Before you can assign host attachments to I/O ports, you must define the format of the I/O ports. There are four or eight FCP/FICON ports on each card depending on the model, and each port is independently configurable using the following steps:

- Hover over the **Home** icon and select **System Status**. The System Status window opens.
- Select the storage image for which you want to configure the ports and, from the **Action** drop-down menu, select **Storage Image** → **Configure I/O Ports** (Figure 13-35).

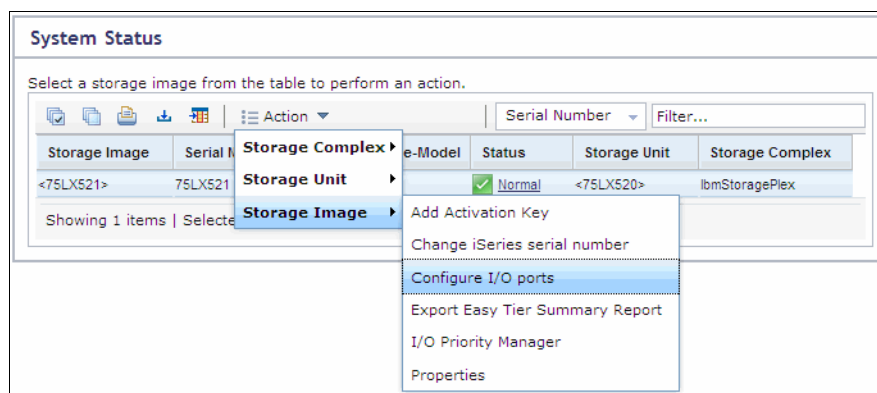


Figure 13-35 System Status window: Configure I/O ports

3. The Configure I/O Port window opens, as shown in Figure 13-36.

Here, you select the ports that you want to format and then click the desired port format (FcSf, FC-AL, or FICON) from the **Action** drop-down menu.

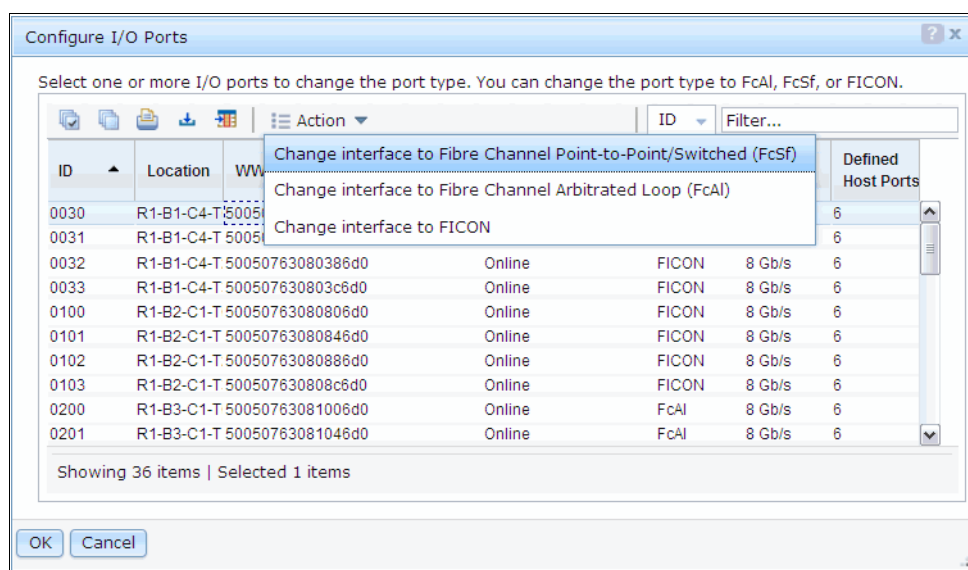


Figure 13-36 Select I/O port format

You get a warning message that the ports might become unusable by the hosts that are currently connected to them.

4. You can repeat this step to format all ports to their required function. Multiple port selection is supported.

13.3.6 Configure logical host systems

In this section, we show you how to configure host systems. This applies only for open systems hosts. A default FICON host definition is automatically created after you define an I/O port to be a FICON port.

To create a new host system, do the following:

1. Hover over **Hosts** and click **Hosts**. The Host connections summary displays as shown in Figure 13-37 on page 336.

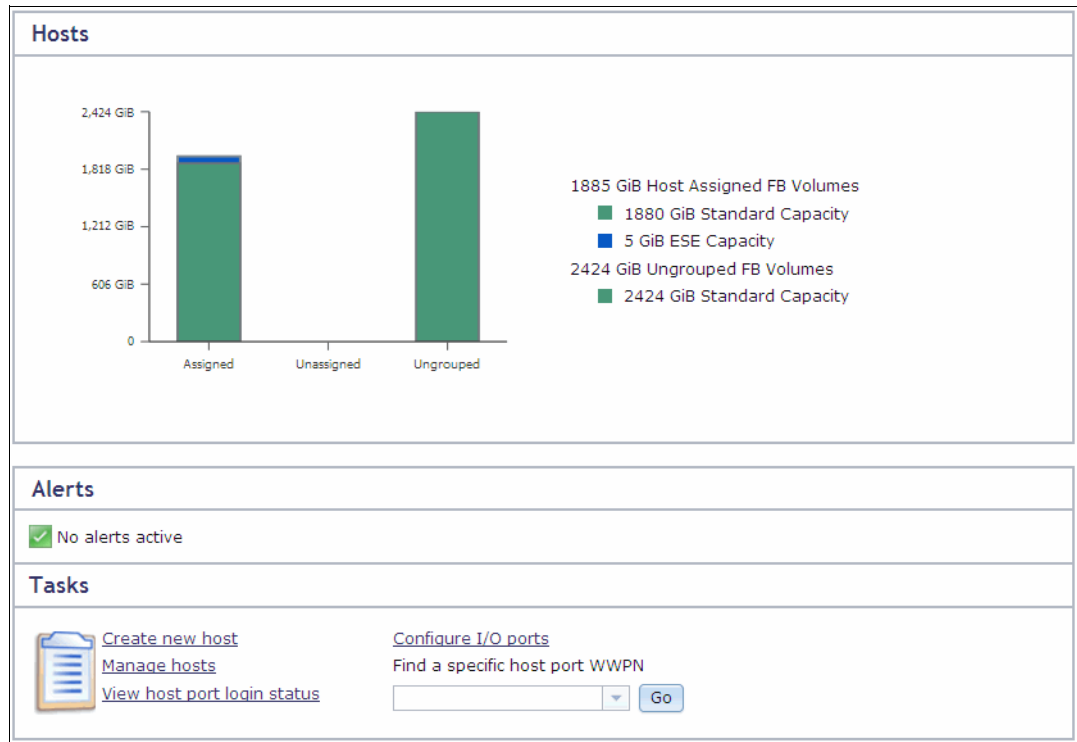


Figure 13-37 Host connections summary

Under the Tasks section, there are shortcut links for various actions. If you want to modify the I/O port configuration previously defined, click the **Configure I/O ports** link.

Tip: You can use the **View host port login status** link to query the host that is logged into the system or use this window to debug host access and switch configuration issues.

If you have more than one storage image, you have to select the right one and then, to create a new host, select the **Create new host connection** link in the Tasks section.

Tip: In the View Host Port Login status window, the list of logged in host posts includes all of the host ports that the storage unit detects, and it will not take into account changes that the storage unit could not detect. For example, the storage unit will not be able to detect that a cable has been disconnected from the host device's port or that a fabric zoning change has occurred. In these cases, the host might not be able to communicate with the storage device anymore; however, the storage device might not detect this and still views the host as logged in.

2. The resulting windows guide you through the host configuration, beginning with the window in Figure 13-38.

Create New Host

Define Host Ports
Define one or more host ports that you will use to map hosts to a volume group in the next step. After you add or remove a host port, the table is updated. All of the host ports in the table will be mapped to the same volume group when you create the connection.

*Host Nickname: Hammerhead16_ *Port Type: Fibre Channel Point-to-Point/Switched (FcSf)

*Host Type: Intel-based Servers (Linux SUSE)(LinuxSuse)

Enter the 16-digit WWPN manually, or select it from the list, and then add it to the table.

*Host WWPN: 21000024ff2fce00 [Add]

WWPN	Nickname	Description
21000024ff2fce01	Hammerhead16__0	
21000024ff2fce00	Hammerhead16__1	
21000024ff2fcdc9	Hammerhead16__2	
21000024ff2fcdc8	Hammerhead16__3	

[Back] [Next] [Finish] [Cancel]

Figure 13-38 Define Host Ports window

In the General host information window, enter the following information:

- a. Host Nickname: Name of the host.
- b. Port Type: You must specify whether the host is attached over an FC Switch fabric (P-P) or direct FC arbitrated loop to the DS8000.
- c. Host Type: The drop-down menu gives you a list of host types from which to select. In our example, we create a Linux host.
- d. Enter the Host WWPN numbers or select the WWPN from the drop-down menu and click the **Add** button.

After the host entry is added into the table, you can manually add a description of each host. When you have entered the necessary information, click **Next**.

3. The Map Host Ports to a Volume Group window displays as shown in Figure 13-39 on page 338. In this window, you can choose the following options:
 - Select the option **Map at a later time** to create a host connection without mapping host ports to a volume group.
 - Select the option **Map to a new volume group** to create a new volume group to use in this host connection.
 - Select the option **Map to an existing volume group** to map to a volume group that is already defined. Choose an existing volume group from the menu. Only volume groups that are compatible with the host type that you selected from the previous window are displayed.

Click **Next** after you select the appropriate option.

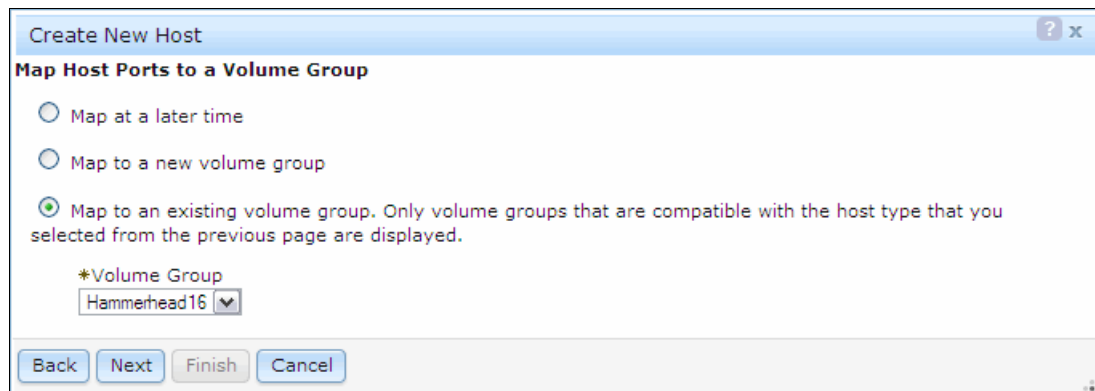


Figure 13-39 Map Host Ports to a Volume Group window

The Define I/O Ports window opens as shown in Figure 13-40.

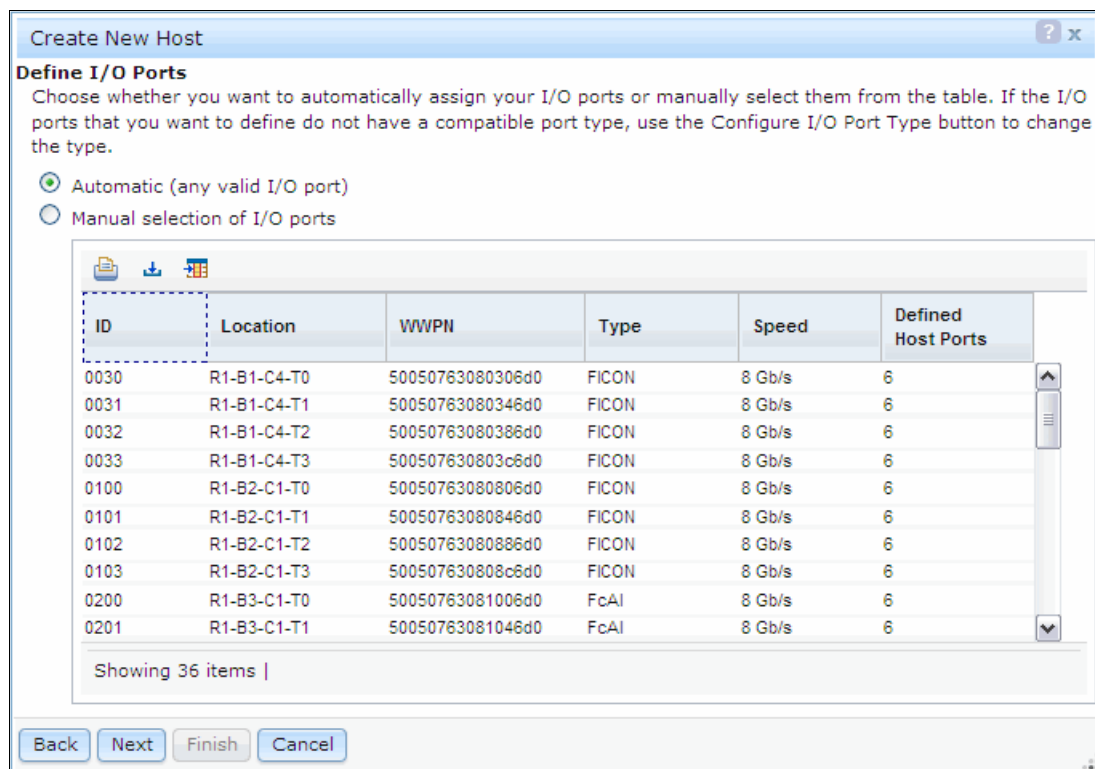


Figure 13-40 Define I/O Ports window

- From the Define I/O ports window, you can choose to automatically assign your I/O ports or manually select them from the table. Defining I/O ports determines which I/O ports can be used by the host ports in this host connection. If specific I/O ports are chosen, the host ports are only able to access the volume group on those specific I/O ports. After defining I/O ports, selecting **Next** directs you to the verification window where you can approve your choices before you commit them.

The Verification window opens as shown in Figure 13-41.

Create New Host

Verification
Review the attributes and verify that they are correct. If they are not correct, click Back to return to the previous page and to specify the correct values. Otherwise, click Finish to complete the process.

Host Ports to be defined

Nickname	WWPN
Hammerhead16__0	21000024ff2fce01
Hammerhead16__1	21000024ff2fcdc9
Hammerhead16__2	21000024ff2fcdc8
Hammerhead16__3	21000024ff2fce00

I/O Ports being used
<Any>

Volume Groups being mapped to

Nickname	# Volumes	Capacity
Hammerhead16	3	120

Back Next Finish Cancel

Figure 13-41 Verification window

- In the Verification window, check the information that you entered during the process. If you want to make modifications, select **Back**, or cancel the process. After you have verified the information, click **Finish** to create the host system. This action takes you to the Manage Host table where you can see the list of all created host connections.

If you need to make changes to a host system definition, select your host in the Manage Host table and choose the appropriate action from the drop-down menu as shown in Figure 13-42.

Manage Hosts
Click on a host row in the table to view details about the connection. Select a host to perform an action.

Nickname	Type	# Host Ports	Volume Groups	# I/O Ports	Storage (GiB)
Hammerhead16	use	4	Hammerhead16	<Any>	400
essfm9	s	2	essfm9_(VG)	<Any>	1765

Showing 2 items | Select

Host Ports defined Volume Groups I/O Ports used

The following volume groups are accessed by all the host ports in the selected host. This host has 4 host ports that are

Figure 13-42 Modify host connections

13.3.7 Create fixed block volumes

This section explains the creation of fixed block (FB) volumes:

1. Hover over **Volumes** and select **FB Volumes**. The FB Volumes Summary window shown in Figure 13-43 displays.

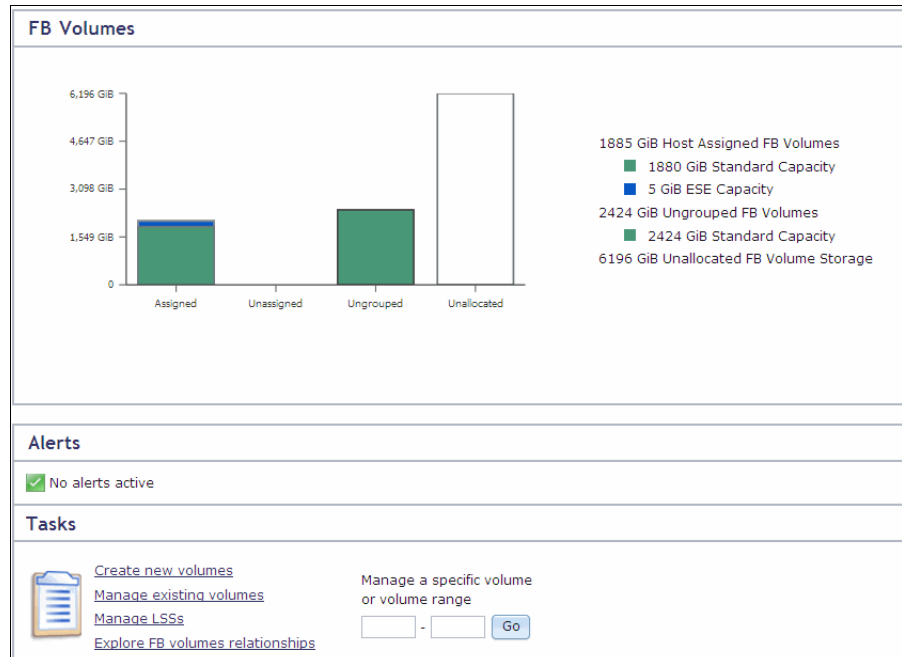


Figure 13-43 FB Volumes summary window

2. If you have more than one storage image, you have to select the appropriate one.
In the Tasks window at the bottom of the window, click **Create new volumes**. The Create Volumes window shown in Figure 13-44 appears.

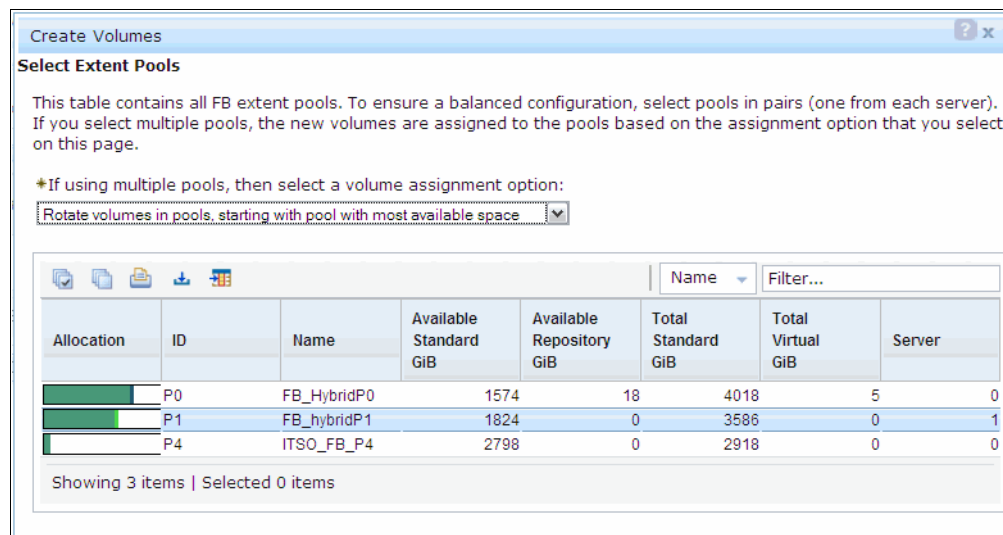


Figure 13-44 Create Volumes: Select Extent Pools

- The table in Create Volumes window contains all the Extent Pools that were previously created for the FB storage type. To ensure a balanced configuration, select Extent Pools in pairs (one from each server). If you select multiple pools, the new volumes are assigned to the pools based on the assignment option that you select on this window.

Click **Next** to continue. The Define Volume Characteristics window appears, as shown in Figure 13-45.

Add Volumes

Define Volume Characteristics

Available capacity: 1,480 GiB Projected remaining capacity: 1,455 GiB

*Volume type: **FB Volumes - DS** *Size (GiB=2[^]30): **25**

*Volume quantity: **1**

*Storage allocation method: **Standard** *Extent allocation method: **Rotate Extents** *Performance Group: **1**

Optionally Choose Nickname

Nickname prefix: **ITSO_** Nickname suffix: **Custom** Start: **00000** ☐ Hexadecimal sequence

Nickname example: ITSO_00000, ITSO_00001

Optionally Assign Volume Groups

To assign the new volume to a volume group, select one or more volume groups and click OK. To add another volume group to the table, select the Create Volume Group action from the drop down.

Nickname	ID	# Volumes	# Host Connections
essfm9_(VG)	0	65	1
Hammerhead16	1	6	1

Showing 2 items | Selected 0 items

OK Add Another Cancel

Figure 13-45 Add Volumes: Define Volume Characteristics

To create a fixed block volume provide the following information:

- Volume type: Specifies the units for the size parameter.
- Size: The size of the volume in the units you specified.
- Volume quantity: The number of volumes to create
- Storage allocation method: This gives you the option to create a standard volume or a space efficient volume. For more information about space efficient volumes see 5.2.6, “Space Efficient volumes” on page 114.
- Extent allocation method: Defines how volume extents are allocated on the ranks in the Extent Pool. This field is not applicable for TSE volumes. The options are:
 - Rotate extents: The extents of a volume are allocated on all ranks in the Extent Pool in a round-robin fashion. This function is called Storage Pool Striping. This allocation method can improve performance because the volume is allocated on multiple ranks. It also helps to avoid hotspots by spreading the workload more evenly on the ranks. This is the default allocation method.

- Rotate volumes: All extents of a volume are allocated on the rank that contains the most free extents. If the volume does not fit on any one rank, it can span multiple ranks in the Extent Pool.
- Performance group: The Performance Group allows you to set the priority level of your volume's I/O operations. For more information, refer to *REDP 4760: DS8000 Performance I/O Manager*.

Optionally, you can provide a Nickname prefix, a Nickname suffix, and one or more volume groups (if you want to add this new volume to a previously created volume group). When your selections are complete, click **Add Another** if you want to create more volumes with different characteristics. Otherwise click **OK** to continue. The Create Volumes window opens as shown in Figure 13-46.

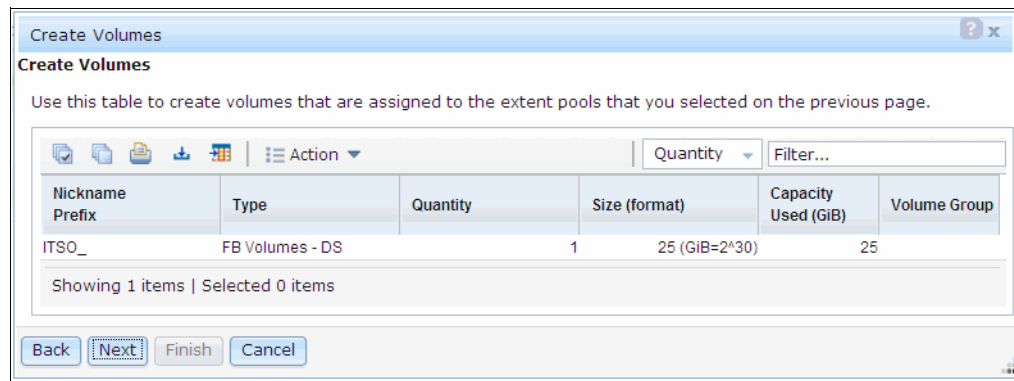


Figure 13-46 Create Volumes window

4. If you need to make any further modifications to the volumes in the table, select the volumes you are about to modify and select the appropriate Action from the **Action** drop-down menu. Otherwise, click **Next** to continue.
5. You need to select an LSS for all created volumes. You can choose **Automatic**, **Manual (Group)**, or **Manual (Fill)**. If you choose **Automatic**, the system assigns the volume addresses for you. If you choose one of the one of the manual assignment methods, select one or more LSSs to assign volume addresses. Scroll down to view the information for additional servers. In our example, we select the **Automatic** assignment method as shown in Figure 13-47 on page 343.

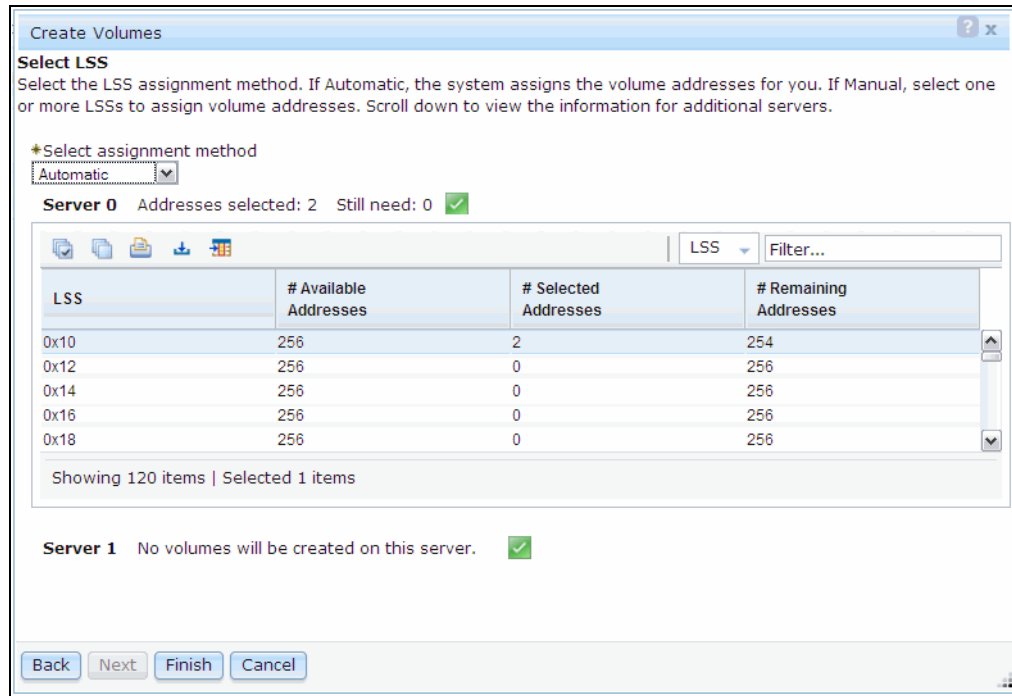


Figure 13-47 Select LSS

6. Click **Finish** to continue.
7. The Create Volumes Verification window shown in Figure 13-48 opens, listing all the volumes that are going to be created. If you want to add more volumes or modify the existing volumes, you can do so by selecting the appropriate action from the **Action** drop-down list. After you are satisfied with the specified values, click **Create all** to create the volumes.

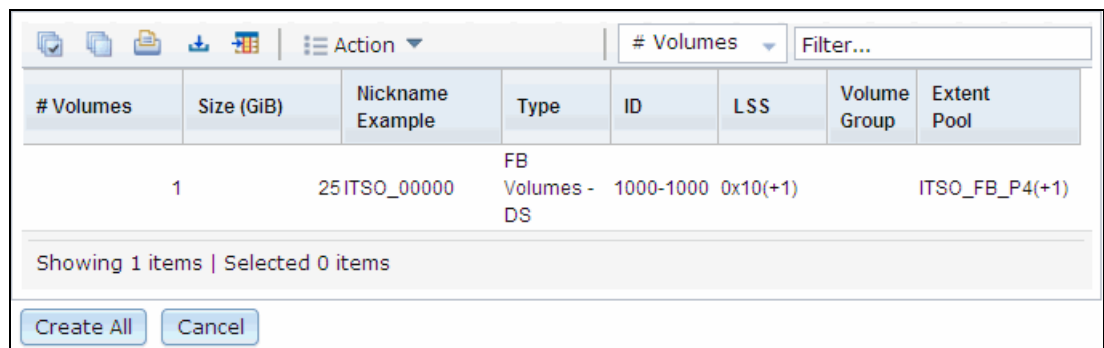


Figure 13-48 Create Volumes Verification window

8. The Creating Volumes information window opens. Depending on the number of volumes, the process can take a while to complete. Optionally, click the **View Details** button to check the overall progress.
9. After the creation is complete, a final window opens. You can select **View Details** or **Close**. If you click **Close**, you return to the main FB Volumes window.

10. The bar graph in the Open Systems - Storage Summary section has changed. From there, you can now select other actions, such as **Manage existing volumes**. The Manage Volumes window is shown in Figure 13-49.

Manage Volumes

Select the filtering options to use for displaying volumes. The table is updated based on the filters that you select. To perform actions, select one or more volumes in the table.

Filter by:

Nickname	ID	Status	Type	GiB	Storage Allocation	Extent Pool	Volume Groups	Host Connections
ITSO0000	1000	✓ Normal	DS	25.0	Standard	ITSO_FB_P4	None	None
	2100	✓ Normal	DS	5.0	Standard	FB_hybridP1	essm9_(VG)essm9	
	2101	✓ Normal	DS	5.0	Standard	FB_hybridP1	essm9_(VG)essm9	
	2102	✓ Normal	DS	5.0	Standard	FB_hybridP1	essm9_(VG)essm9	
	2103	✓ Normal	DS	5.0	Standard	FB_hybridP1	essm9_(VG)essm9	
	2104	✓ Normal	DS	5.0	Standard	FB_hybridP1	essm9_(VG)essm9	
	2105	✓ Normal	DS	5.0	Standard	FB_hybridP1	essm9_(VG)essm9	
	2106	✓ Normal	DS	5.0	Standard	FB_hybridP1	essm9_(VG)essm9	
	2107	✓ Normal	DS	5.0	Standard	FB_hybridP1	essm9_(VG)essm9	
	2108	✓ Normal	DS	5.0	Standard	FB_hybridP1	essm9_(VG)essm9	

Showing 73 items | Selected 0 items

Figure 13-49 FB Volumes: Manage Volumes

If you need to make changes to a volume, select a volume and click the appropriate action from the Action drop down menu.

13.3.8 Create volume groups

To create a volume group, perform this procedure:

1. Hover over **Volumes** and select **Volume Groups**. The Volume Groups window displays.
2. To create a new volume group, select **Create** from the **Action** drop-down menu as shown in Figure 13-50.

Volume Groups

Nickname	ID	Pressing hod	Block size	# Volumes	Capacity	# Host Connections
essm9_(VG)	0		512	65	1765	1
Hammerhead16	1		512	6	400	1

Showing 2 items | Selected 0 items

Figure 13-50 Volume Groups window: Select Create

The Define Volume Group Properties window shown in Figure 13-51 opens.

Create New Volume Group

Define Volume Group Properties
Define the volume group properties. A volume group is a set of logical volumes that can be accessed by a host.

*Volume Group Nickname: ITSO VolumeGroup *Host Type: IBM pSeries, RS/6000 and RS/6000 SP Servers (AIX)(pSeries)

Volumes
Select volumes to be included in the group.
Filter by LSS: 10

Nickname	ID	Storage Allocation	GiB	GB	Extent Pool
ITSO0000	1000	Standard	25.0	26.8	ITSO_FB_P4

Showing 1 items | Selected 0 items

Figure 13-51 Define Volume Group Properties window

3. In the Define Volume Group Properties window, enter the nickname for the volume group and select the host type from which you want to access the volume group. If you select one host (for example, IBM pSeries®), all other host types with the same addressing method are automatically selected. This does not affect the functionality of the volume group; it supports the host type selected.
4. Select the volumes to include in the volume group. If you have to select a large number of volumes, you can specify the LSS so that only these volumes display in the list, and then you can select all.
5. Click **Next** to open the Verification window shown in Figure 13-52.

Create New Volume Group

Verification
Review the attributes and verify that they are correct. If they are not correct, click Back to return to the previous page and to specify the correct values. Otherwise, click Finish to complete the process.

Attribute	Value
Nickname	ITSO VolumeGroup
Accessed by host types	IBM pSeries, RS/6000 and RS/6000 SP Servers (AIX)
Volume quantity	1
Selected capacity (GiB/GB)	25.0 GiB/ 26.8 GB

Showing 4 items |

Figure 13-52 Create New Volume Group Verification window

6. In the Verification window, check the information you entered during the process. If you want to make modifications, select **Back**, or you can cancel the process altogether. After you verify the information, click **Finish** to create the host system attachment. After the creation completes a Create Volume Group completion window appears where you can select **View Details** or **Close**.
7. After you select **Close**, you will see the new volume group in the Volume Group window.

Create Volume Group of scsimap256

In Linux 2.4 kernels, SCSI devices are discovered by scanning the SCSI bus when the host adapter driver is loaded. If there is a gap in the LUN ID sequence, the LUNs after the gap will not be discovered (Figure 13-53). A list of devices that have been discovered and are recognized by the SCSI subsystem are listed in `/proc/scsi/scsi`. Use the `cat` command to display the output of `/proc/scsi/scsi` to verify that the correct number of LUNs has been recognized by the kernel.

Volume Group Properties

Volume Group Properties
You can change the nickname of the volume group, or you can select volumes to add or remove from the volume group. This page displays all of the volumes that are assigned to this volume group.

Volume Group Nickname

ITSO

ID

2

Addressing Method

Map

Block Size

512

Volumes Assigned
Filter by LSS

All

Action

Nickname

Filter...

Nickname	ID	LUN ID	Storage Allocation	GiB	GB
ITSO-0010	0010	0	Standard	5.0	5.4
ITSO-0012	0012	1	Standard	5.0	5.4
Empty Slot	N/A	2	N/A	N/A	N/A
Empty Slot	N/A	3	N/A	N/A	N/A
ITSO-0013	0013	4	Standard	5.0	5.4

Showing 5 items | Selected 0 items

Figure 13-53 Gaps in the LUN ID

If you want to modify the LUN ID of a FB volume that is already in the Volume Group, use **Remove Volumes**. Then use **Add Volumes** to add the volumes back to the Volume Group, and then modify the LUN ID to the new LUN ID.

You can change the LUN ID field when you create the new volume group, or add a new volume to an existing volume group. You can edit the column under the LUN ID (Figure 13-54 on page 347).

Create New Volume Group

LUN ID assignments

Volume Group Nickname: ITSO

Review the LUN ID assignments and verify that they are correct. You may change the assignments on the value you wish to change and typing in the new value. Only integers between 0 and 255 will be accepted. Click Next when done with the assignments.

Nickname	ID	LUN ID	Storage Allocation	GiB	GB
ITSO-0010	0010	0	Standard	5	5.4
ITSO-0012	0012	1	Standard	5	5.4
ITSO-0013	0013	2	Standard	5	5.4

Showing 3 items |

Figure 13-54 Update the LUN ID field

If you do not see the LUN ID column displayed in the window, you can enable it by right-clicking the menu bar. Select the box to the right of LUN ID. The LUN ID field is displayed as shown in Figure 13-55. You can edit this column.

Volume Group Properties

Volume Group Properties

You can change the nickname of the volume group, or you can select volumes to add or remove from the volume group. The table displays all of the volumes that are assigned to this volume group.

Volume Group Nickname: ITSO

ID: 2

Addressing Method: Mask

Block Size: 512

Volumes Assigned

Filter by LSS: All

Action

Nickname

Filter...

Nickname	ID	Storage Allocation	GiB	GB	RAID Type
ITSO-0010	0010	Standard		5.4	RAID 5
ITSO-0012	0012	Standard		5.4	RAID 5

Showing 2 items | Selected 0 items

Figure 13-55 Enable display of LUN ID column

If you enter a LUN ID that is already been used in this volume group, an error message will display as shown in Figure 13-56.

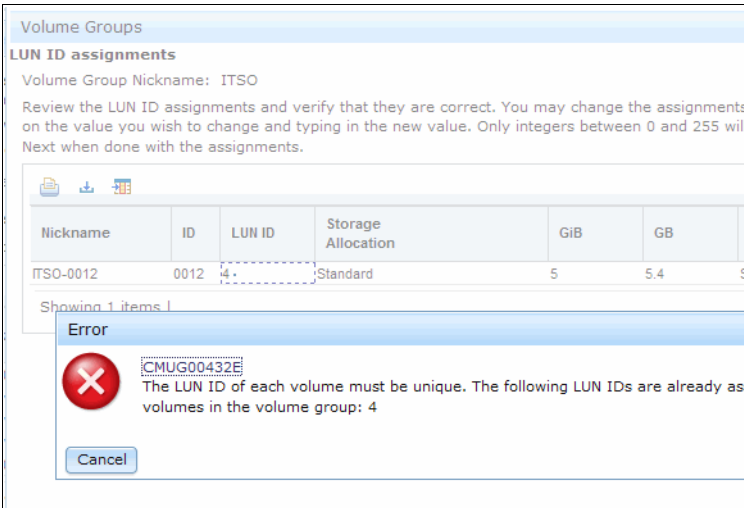


Figure 13-56 Error message for duplicated LUN ID

There are only 256 LUNs in a scsimap256 volume group 0-255. if you enter a number that is larger than 255, you will get the error message shown in Figure 13-57.

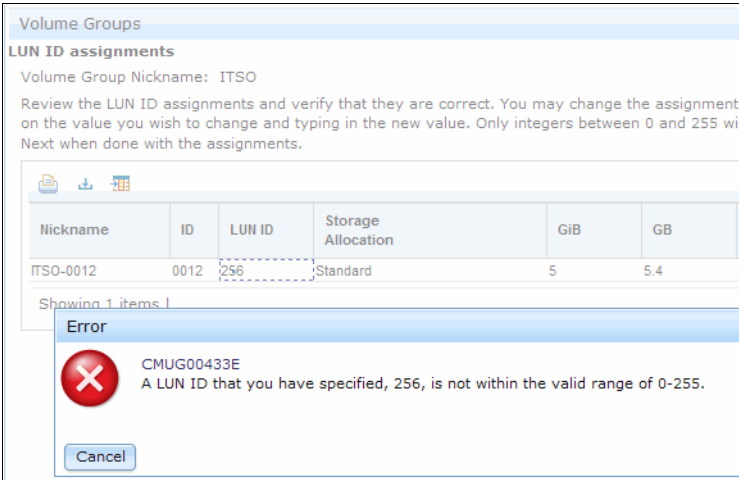


Figure 13-57 Error message for number larger than 255

13.3.9 Create LCUs and CKD volumes

In this section, we show how to create logical control units (LCUs) and CKD volumes. This is only necessary for IBM System z.

Important: The LCUs you create must match the logical control unit definitions on the host I/O configuration. More precisely, each LCU ID number you select during the create process must correspond to a CNTLUNIT definition in the HCD/IOCP with the same CUADD number. It is vital that the two configurations match each other.

Perform the following steps:

1. Hover over **Volumes** and select **CKD LCUs and Volumes**. The CKD LCUs and Volumes window shown in Figure 13-58 opens.

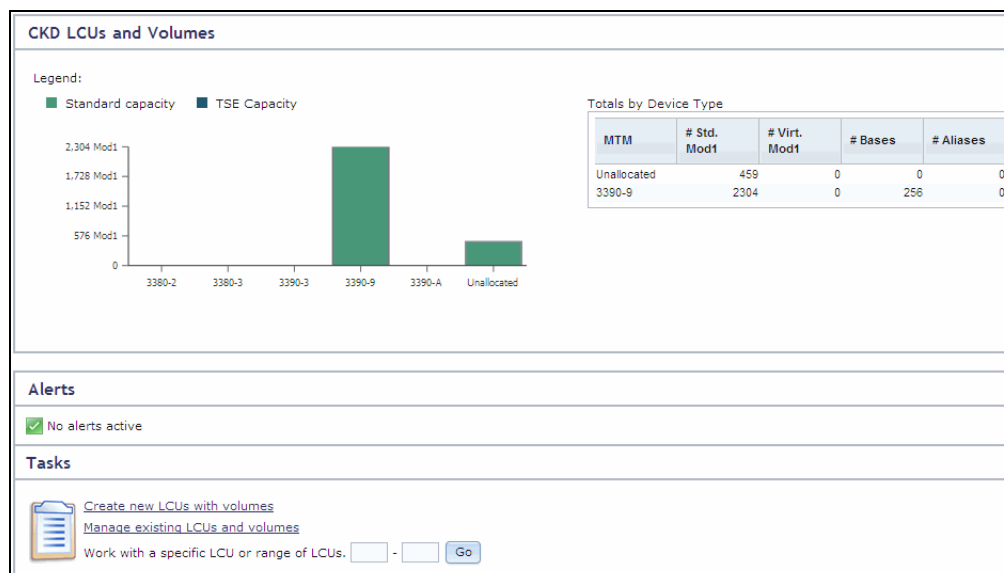


Figure 13-58 CKD LCUs and Volumes window

2. Select a storage image from the Select storage image drop-down menu if you have more than one. The window is refreshed to show the LCUs in the storage image.
3. To create new LCUs, select **Create new LCUs with volumes** from the tasks list. The Create LCUs (Figure 13-59 on page 350) window opens.

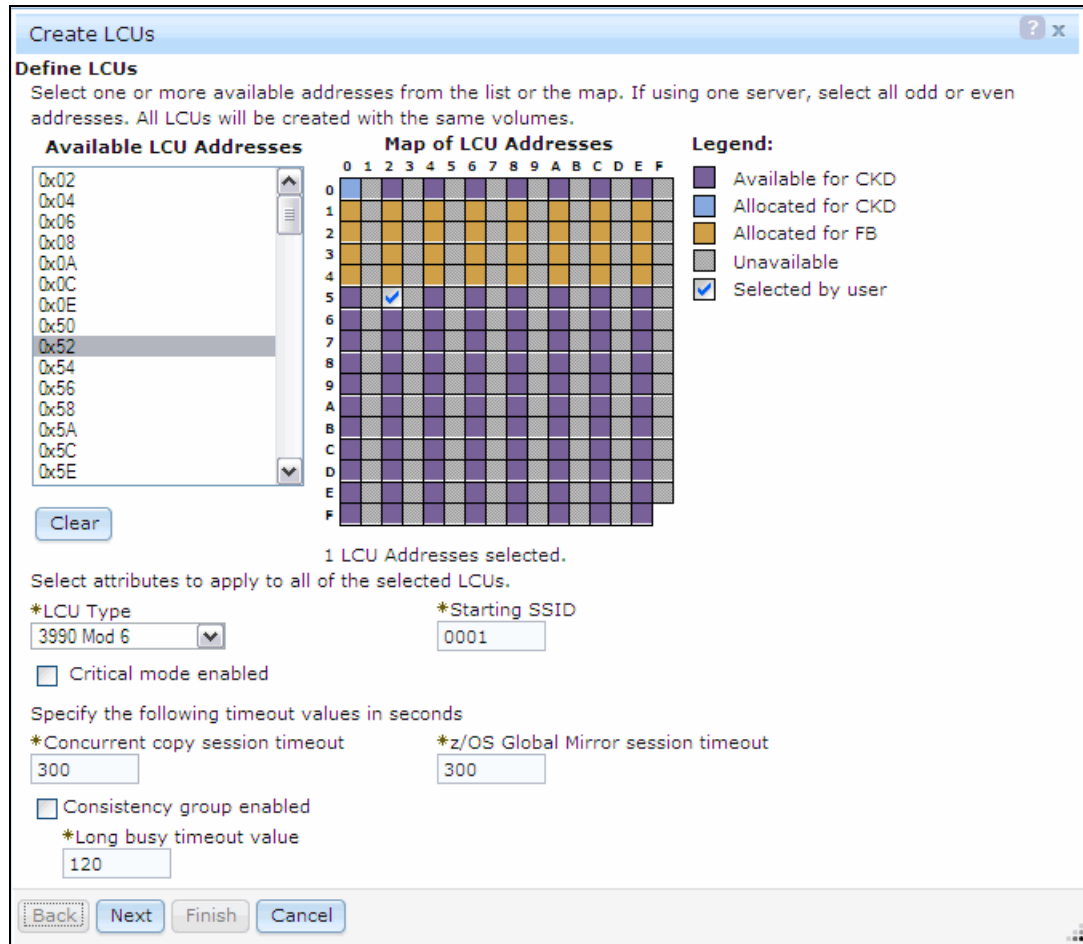


Figure 13-59 Create LCUs window

4. Select the LCUs you want to create. You can select them from the list displayed on the left by clicking the number, or you can use the map. When using the map, click the available LCU square. You must enter all the other necessary parameters for the selected LCUs.
 - Starting SSID: Enter a Subsystem ID (SSID) for the LCU. The SSID is a four character hexadecimal number. If you create multiple LCUs at one time, the SSID number is incremented by one for each LCU. The LCUs attached to the same SYSPLEX must have different SSIDs. Use unique SSID numbers across your whole environment.
 - LCU type: Select the LCU type you want to create. Select 3990 Mod 6 unless your operating system does not support Mod 6. The options are:
 - 3990 Mod 3
 - 3990 Mod 3 for TPF
 - 3990 Mod 6

The following parameters affect the operation of certain Copy Services functions:

- Concurrent copy session timeout: The time in seconds that any logical device on this LCU in a concurrent copy session stays in a long busy state before suspending a concurrent copy session.
- z/OS Global Mirror Session timeout: The time in seconds that any logical device in a z/OS Global Mirror session (XRC session) stays in long busy before suspending the XRC session. The long busy occurs because the data mover has not offloaded data when the logical device (or XRC session) is no longer able to accept additional data.

With recent enhancements to z/OS Global Mirror, there is now an option to suspend the z/OS Global Mirror session instead of presenting the long busy status to the applications.

- Consistency group timeout: The time in seconds that remote mirror and copy consistency group volumes on this LCU stay extended long busy after an error that causes a consistency group volume to suspend. While in the extended long busy state, I/O is prevented from updating the volume.
- Consistency group timeout enabled: Check the box to enable remote mirror and copy consistency group timeout option on the LCU.
- Critical mode enabled: Check the box to enable critical heavy mode. Critical heavy mode controls the behavior of the remote copy and mirror pairs that have a primary logical volume on this LCU.

When all necessary selections have been made, click **Next** to proceed to the next window.

5. In the next window (Figure 13-60), you must configure your base volumes and, optionally, assign alias volumes. The Parallel Access Volume (PAV) license function needs to be activated to use alias volumes.

Figure 13-60 Create Volumes window

Define the base volume characteristics in the first third of this window with the following information:

- Base type:
 - 3380 Mod 2
 - 3380 Mod 3
 - 3390 Standard Mod 3
 - 3390 Standard Mod 9
 - 3390 Mod A (used for Extended Address Volumes - EAV)
 - 3390 Custom

- Volume size: This field must be changed if you use the volume type 3390 Custom or 3390 Mode A.
- Size format: This format only has to be changed if you want to enter a special number of cylinders. This can also only be used by 3390 Custom or 3390 Mod A volume types.
- Volume quantity: Here you must enter the number of volumes you want to create.
- Base start address: The starting address of volumes you are about to create. Specify a decimal number in the range of 0 - 255. This defaults to the value specified in the Address Allocation Policy definition.
- Order: Select the address allocation order for the base volumes. The volume addresses are allocated sequentially, starting from the base start address in the selected order. If an address is already allocated, the next free address is used.
- Storage allocation method: This field only appear on boxes that have the FlashCopy SE function activated. The options are:
 - Standard: Allocate standard volumes.
 - Track Space Efficient (TSE): Allocate Space Efficient volumes to be used as FlashCopy SE target volumes.
- Extent allocation method: Defines how volume extents are allocated on the ranks in the Extent Pool. This field is not applicable for TSE volumes. The options are:
 - Rotate extents: The extents of a volume are allocated on all ranks in the Extent Pool in a round-robin fashion. This function is called Storage Pool Striping. This allocation method can improve performance because the volume is allocated on multiple ranks. It also helps to avoid hotspots by spreading the workload more evenly on the ranks. This is the default allocation method.
 - Rotate volumes: All extents of a volume are allocated on the rank that contains most free extents. If the volume does not fit on any one rank, it can span multiple ranks in the Extent Pool.

Select **Assign the alias volume to these base volumes** if you use PAV or Hyper PAV and provide the following information:

- Alias start address: Enter the first alias address as a decimal number between 0 - 255.
- Order: Select the address allocation order for the alias volumes. The volume addresses are allocated sequentially starting from the alias start address in the selected order.
- Evenly assign alias volumes among bases: When you select this option, you have to enter the number of alias you want to assign to each base volume.
- Assign aliases using a ratio of aliases to base volume: This option gives you the ability to assign alias volumes using a ratio of alias volumes to base volumes. The first value gives the number you assign to each alias volume and the second value selects to which alias volume you want to assign an alias. If you select 1, each base volume will get a alias volume. If you select 2, every second base volume gets an alias volume. If you select 3, every third base volume gets an alias volume. The selection starts always with the first volume.

Tip: You can assign all aliases in the LCU to just one base volume if you have implemented HyperPAV or Dynamic alias management. With HyperPAV, the alias devices are not permanently assigned to any base volume even though you initially assign each to a certain base volume. Rather, they reside in a common pool and are assigned to base volumes as needed on a per I/O basis. With Dynamic alias management, WLM will eventually move the aliases from the initial base volume to other volumes as needed.

If your host system is using Static alias management, you need to assign aliases to all base volumes on this window, because the alias assignments made here are permanent in nature. To change the assignments later, you have to delete and re-create aliases.

In the last section of this window, you can optionally assign the alias nicknames for your volumes:

- Nickname prefix: If you select a nickname suffix of **None**, you must enter a nickname prefix in this field. Blanks are not allowed. If you select a nickname suffix of **Volume ID** or **Custom**, you can leave this field blank.
- Nickname suffix: You can select **None** as described above. If you select **Volume ID**, you have to enter a four character volume ID for the suffix, and if you select **Custom**, you have to enter four digit hexadecimal number or a five digit decimal number for the suffix.
- Start: If you select Hexadecimal sequence, you have to enter a number in this field.

Tip: The nickname is not the System z VOLSER of the volume. The VOLSER is created later when the volume is initialized by the ICKDSF INIT command.

Click **OK** to proceed. The Create Volumes window shown in Figure 13-61 appears.

Create LCU

Create Volumes

Each row of this table displays a set of volumes with the same type and address range. To create volumes that are identical in each LCU, select the Create Volumes action. To modify volumes, select a row and then select the Modify Volumes action. To delete volumes, select one or more rows and then select the Delete Volumes action. Required storage values are updated each time you add or delete volumes.

MTM	# Bases	Start	End	# Aliases	Start	End	Prefix	# Volumes	# Std. Cyls	# Virt. Cyls
3390-9	10	0x00	0x09	0x00	0x00	ITSO	10	100170	0	

Showing 1 items | Selected 0 items

Required storage per LCU: **90 Mod1**, Required server 0 storage: **90 Mod1**, Required server 1 storage: **0 Mod1**

Back Next Finish Cancel

Figure 13-61 Create Volumes window

6. In the Create Volumes window (Figure 13-61), you can select the just created volumes to modify or delete them. You also can create more volumes if this is necessary at the time. Select **Next** if you do not need to create more volumes at this time.

- In the next window (Figure 13-62), you can change the Extent Pool assignment to your LCU. Select **Finish** if you do not want to make any changes here.

Create LCUs

LCU to Extent Pool Assignment

This table contains all pools that have standard or virtual capacity available for allocation. When you select the assignment method, the table is updated automatically. To change the default pool list, select the Change Pool List action. If you are assigning pools manually, you can add or remove LCUs by selecting the pool and then selecting the Change LCU Assignment action.

*Assignment Method
 Rotate LCUs in pools, starting with pool with most available space ☒ LCU may span multiple pools

Allocation	ID	Nickname	# New Std. Mod1	# New Virt. Mod1	# Available Std. Mod1	# Available Virt. Mod1	# Total Std. Mod1	# Total Virt. Mod1
	2	CKD_hybridP2	90	0	369	0	2763	0

Showing 1 items |

Back Next Finish Cancel

Figure 13-62 LCU to Extent Pool Assignment window

- The Create LCUs Verification window appears, as shown in Figure 13-63, where you can see list of all the volumes that are going to be created. If you want to add more volumes or modify the existing ones, you can do so by selecting the appropriate action from the **Action** drop-down list. After you are satisfied with the specified values, click **Create all** to create the volumes.

Create LCUs Verification

Review the LCUs that will be created. To add additional LCUs to the table, select the Add LCUs action. To modify an LCU, select the LCU and then select the Modify LCU action. To delete LCUs, select one or more LCUs and then select the Delete LCUs action.

To create the new LCUs in the table, click Create All. LCUs are created in increasing LCU ID order, one LCU and its volumes at a time.

☒ ☐ ☐ ☐ ☐
 Action Long busy timeout value Filter...

LCU	SSID	EP Name	EP ID	# New Mod1	# New Bases	# New Aliases
0x52	0x0002	CKD_hybridP2	P2	90	10	0

Showing 1 items | Selected 0 items

Create All Cancel

Figure 13-63 Create LCUs Verification window

- The Creating Volumes information window opens. Depending on the number of volumes, the process can take a while to complete. Optionally, click the **View Details** button to check the overall progress.
- After the creation is complete, a final window is displayed. You can select **View details** or **Close**. If you click **Close**, you return to the main CKD LCUs and Volumes window, where you see that the bar graph has changed.

13.3.10 Additional actions on LCUs and CKD volumes

When you select **Manage existing LCUs and Volumes** (Figure 13-64), you can perform additional actions at the LCU or volume level.

As shown in Figure 13-64, you have the following options:

- ▶ **Create:** See 13.3.9, “Create LCUs and CKD volumes” on page 348 for information about this option.
- ▶ **Clone LCU:** See 13.3.9, “Create LCUs and CKD volumes” on page 348 for more information about this option. Here all properties from the selected LCU will be cloned.
- ▶ **Add Volumes:** Here you can add base volumes to the selected LCU. See 13.3.9, “Create LCUs and CKD volumes” on page 348 for more information about this option.
- ▶ **Add Aliases:** Here you can add alias volumes without creating additional base volumes.
- ▶ **Properties:** Here you display the additional properties. You can also change certain of them, such as the timeout value.
- ▶ **Delete:** Here you can delete the selected LCU. This must be confirmed because you will also delete all volumes that will contain data.
- ▶ **Migrate:** This option allows you to migrate volumes from one extent pool to another. For more information about migrating volumes, refer to *REDP 4667 - IBM System Storage DS8000 Easy Tier V2*.

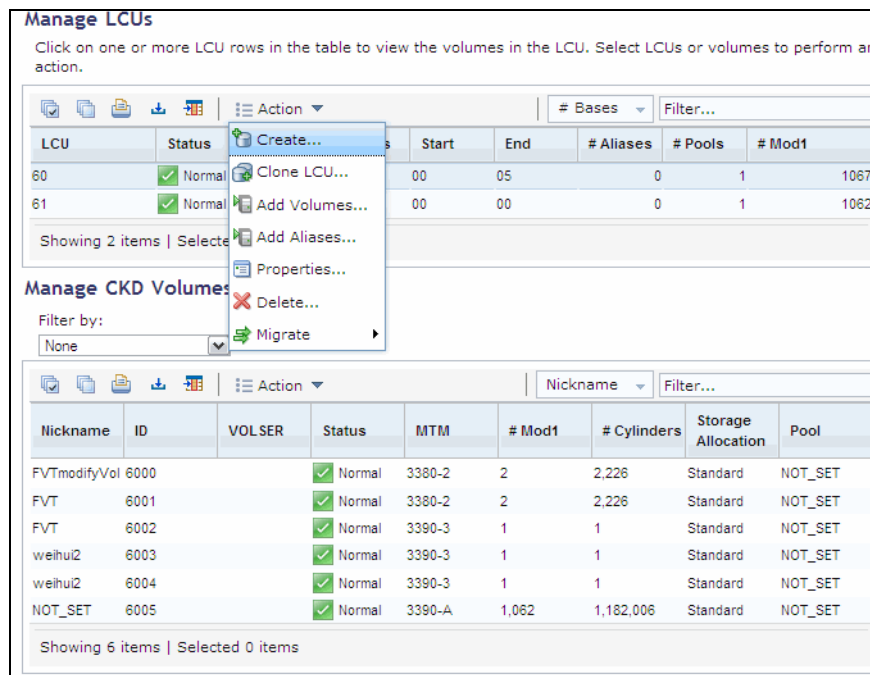


Figure 13-64 Manage LCUs and Volumes window

The next window (Figure 13-65 on page 356) shows that you can take actions at the volume level after you have selected an LCU:

- ▶ **Increase capacity:** Use this action to increase the size of a volume. The capacity of a 3380 volume cannot be increased. After the operation completes, you need to use the ICKDSF REFORMAT REFVTOC command to adjust the volume VTOC to reflect the additional cylinders. Note that the capacity of a volume cannot be decreased.

- ▶ **Add Aliases:** Use this action when you want to define additional aliases without creating new base volumes.
- ▶ **Properties:** Here you can view the volumes properties. The only value you change is the nickname. You can also see if the volume is online from the DS8000 side.
- ▶ **Delete:** Here you can delete the selected volume. This must be confirmed because you will also delete all alias volumes and data on this volume.
- ▶ **Migrate:** This options allows you to migrate volumes from one extent pool to another. For more information about migrating volumes, refer to *REDP 4667 - IBM System Storage DS8000 Easy Tier V2*.

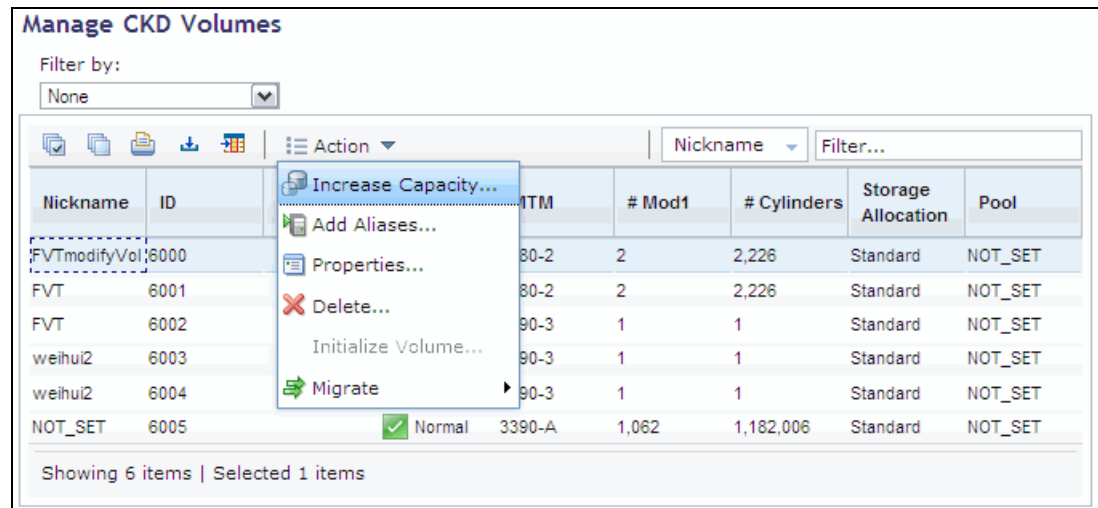


Figure 13-65 Manage CKD Volumes

Tip: After initializing the volumes using the ICKDSF INIT command, you also will see the VOLSERS in this window. This is not done in this example.

The Increase capacity action can be used to dynamically expand volume capacity without needing to bring the volume offline in z/OS. It is good practice to start using 3390 Mod A after you can expand the capacity and change the device type of your existing 3390 Mod 3, 3390 Mod 9, and 3390 Custom volumes. Keep in mind that 3390 Mod A volumes can only be used on z/OS V1.10 or later. After the capacity has been increased on DS8000, you need to run a ICKDSF to rebuild the VTOC Index, allowing it to recognize the new volume size.

13.4 Other DS GUI functions

In this section, we discuss additional DS GUI functions.

13.4.1 Check the status of the DS8000

Perform these steps to display and explore the overall status of your DS8000 system:

1. In the navigation pane in the DS GUI, hover over **Home** and select **System Status**. The System Status window opens.

2. Select your storage complex and, from the **Action** drop-down menu, select **Storage Unit** → **System Summary** as shown in Figure 13-66.

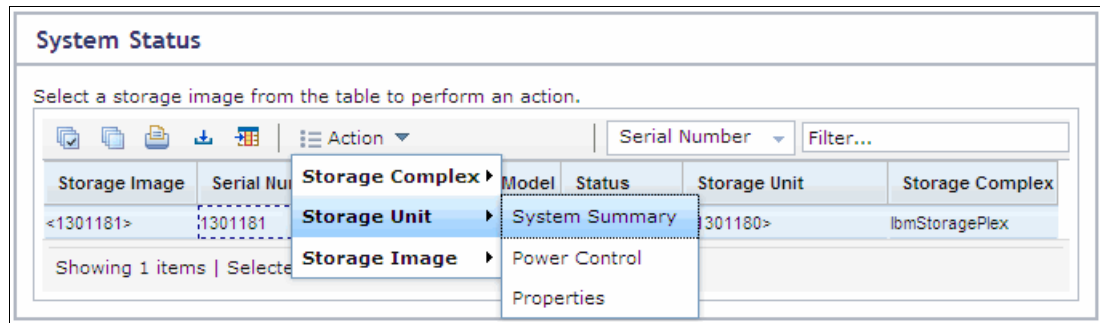


Figure 13-66 Select Storage Unit System Summary

3. The new Storage Complex window provides general DS8000 system information. It is divided into four sections (Figure 13-67):
 - a. **System Summary:** You can quickly identify the percentage of capacity that is currently used, and the available and used capacity for open systems and System z. In addition, you can check the system state and obtain more information by clicking the state link.
 - b. **Management Console information.**
 - c. **Performance:** Provides performance graphs for host MBps, host KIOps, rank MBps, and rank KIOps. This information is periodically updated every 60 seconds.
 - d. **Racks:** Represents the physical configuration.

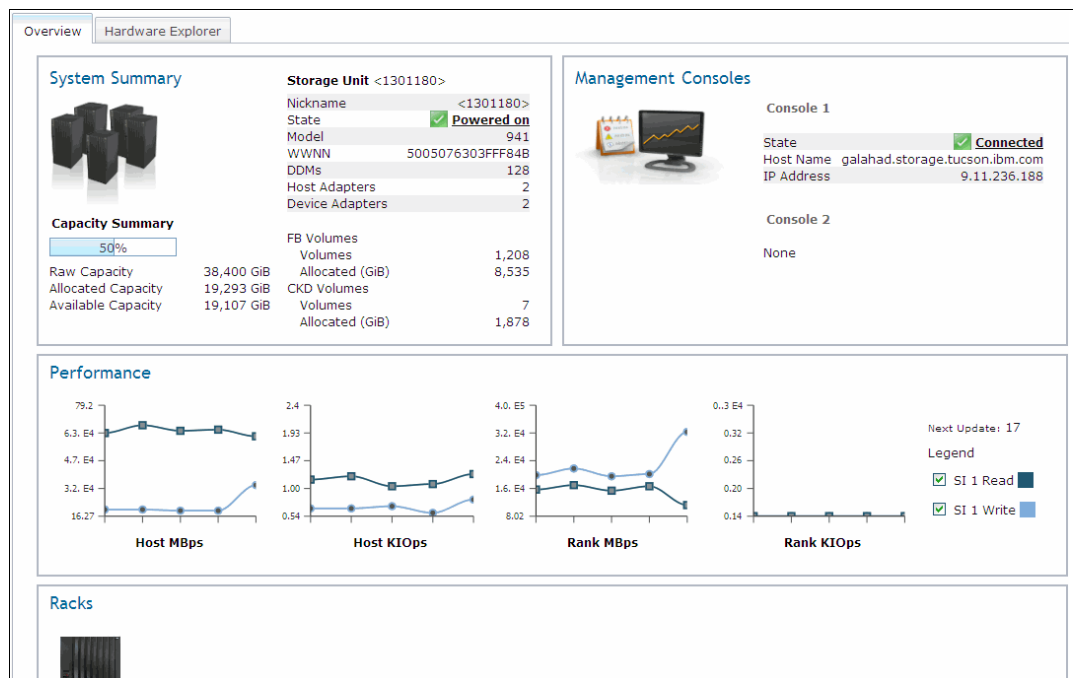


Figure 13-67 System Summary overview

4. In the Rack section, the number of racks shown matches the racks physically installed in the storage unit. If you position the mouse pointer over the rack, additional rack information is displayed, such as the rack number, the number of DDMs, and the number of host adapters (Figure 13-68).



Figure 13-68 System Summary: rack information

13.4.2 Explore the DS8000 hardware

DS8000 GUI allows you to explore hardware installed in your DS8000 system by locating specific physical and logical resources (arrays, ranks, extent pools, and others). The Hardware Explorer shows system hardware and a mapping between logical configuration objects and DDMs.

You can explore the DS8000 hardware components and discover the correlation between logical and physical configuration by performing the following steps:

1. In the navigation pane in the DS GUI, hover over **Home** and select **System Status**.
2. The Storage Complexes Summary window opens. Select your storage complex and, from the **Action** drop-down menu, select **Storage Unit** → **System Summary**.

3. Select the Hardware Explorer tab to switch to the Hardware Explorer window (Figure 13-69).

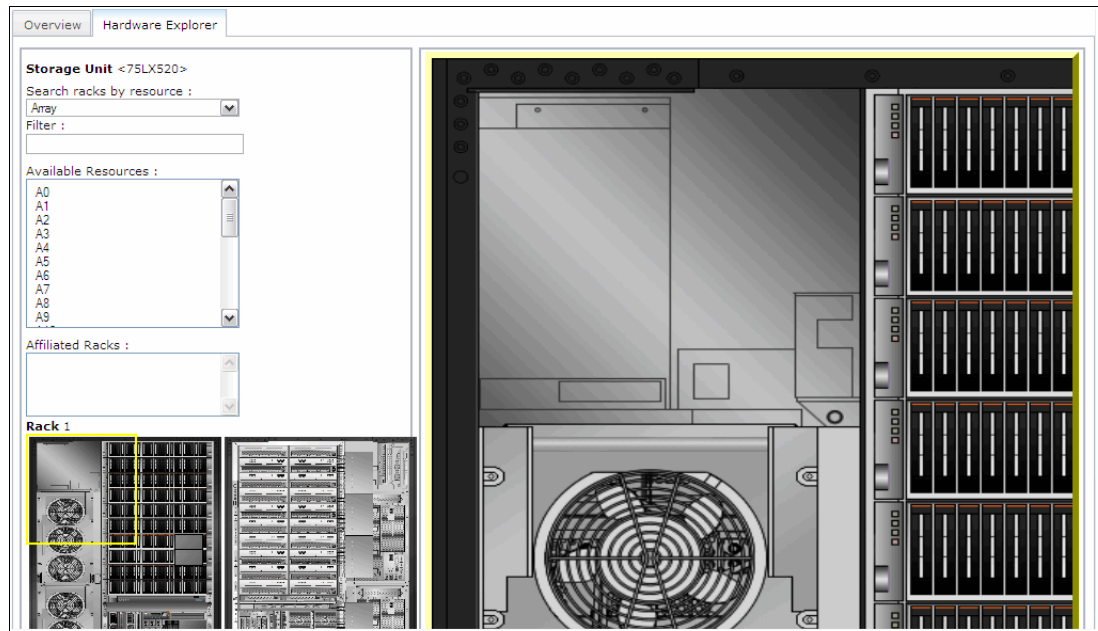


Figure 13-69 Hardware Explorer window

4. In this window, you can explore the specific hardware resources installed by selecting the appropriate component under the **Search racks by resources** drop-down menu. In the Rack section of the window, there is a front and rear view of the DS8000 rack. You can interact with the rack image to locate resources. To view a larger image of a specific location (displayed in the right pane of the window), use your mouse to move the yellow box to the desired location across the DS8000 front and rear view.

- To check where the physical disks of arrays are located, change the search criteria to **Array** and, from the **Available Resources** section, click one or more array IDs that you want to explore. After you click the array ID, the location of each DDM is highlighted in the rack image. Each disk has an appropriate array ID label. Use your mouse to move the yellow box in the rack image on the left to the desired location across the DS8000 front and rear view to view the magnified view of this section as shown in Figure 13-70.

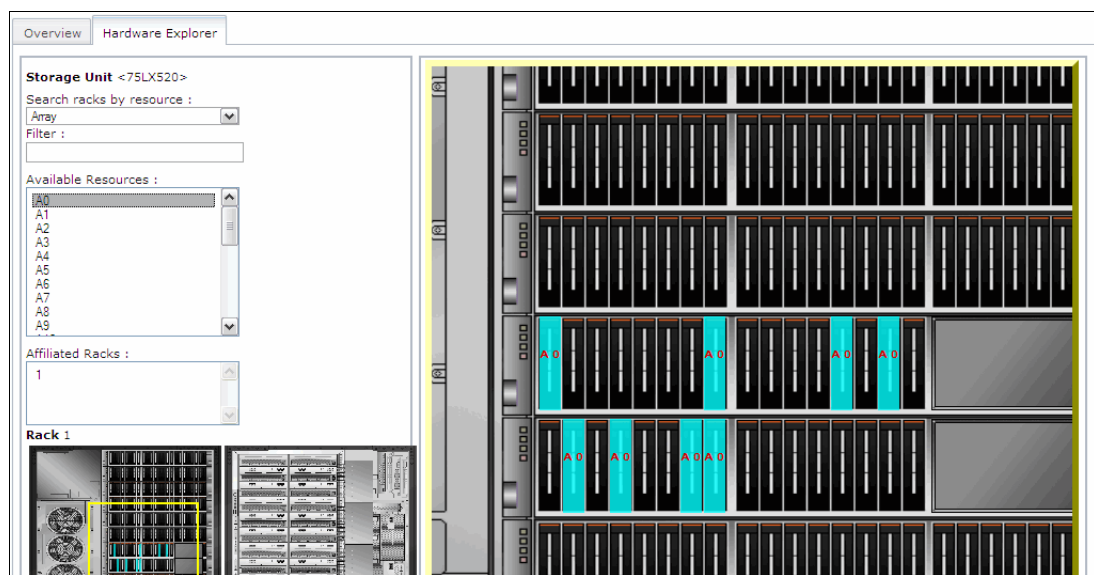


Figure 13-70 View arrays

- After you have identified the location of array DDMs, you can position the mouse pointer over the specific DDM to display more information, as shown in Figure 13-71.

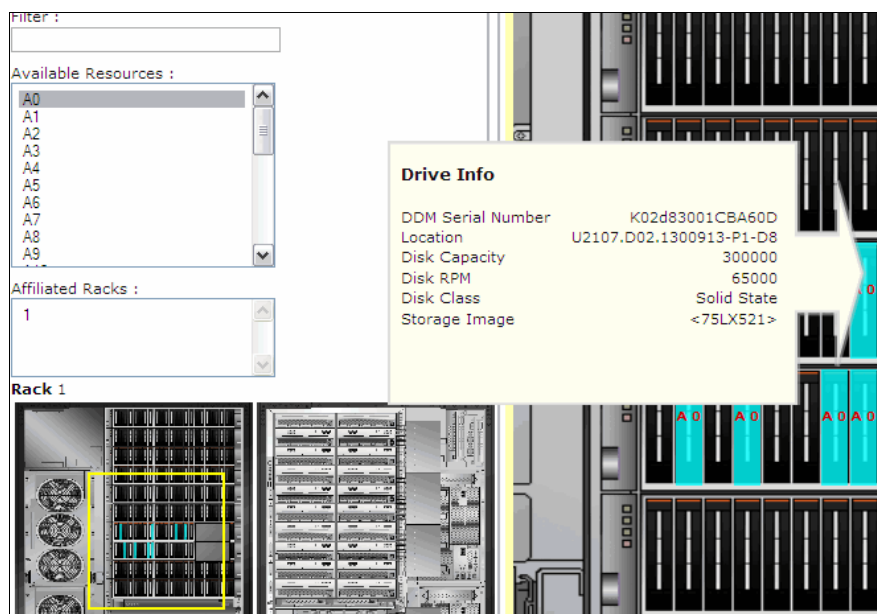


Figure 13-71 DDM information

- Change the search criteria to Extent Pool to discover more about each extent pool location. Select as many extent pools as you need in the Available Resources section and find the physical location of each one as shown in Figure 13-72.

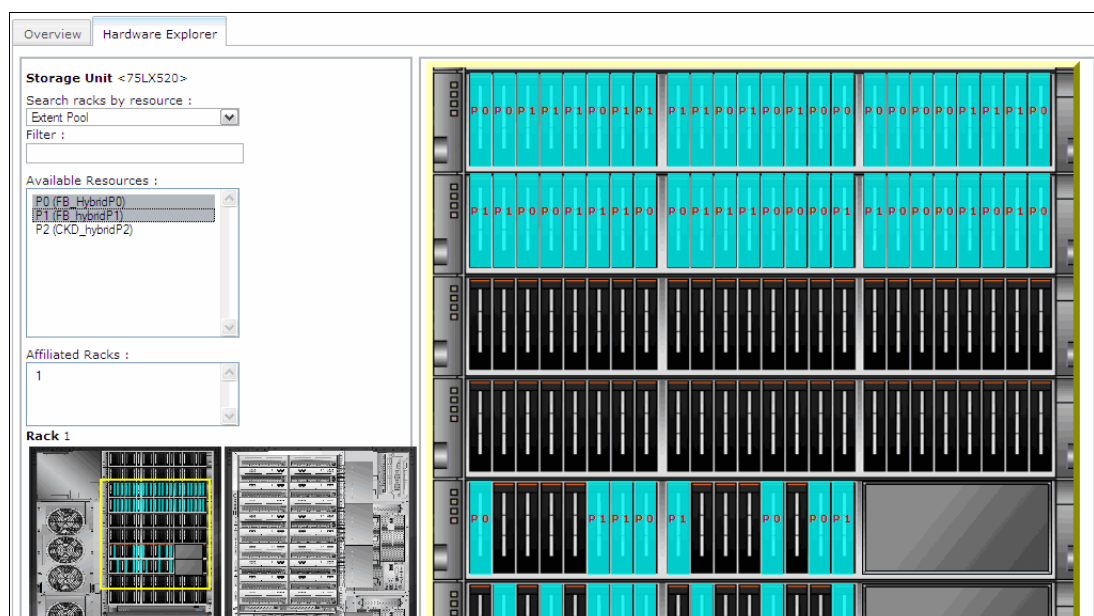


Figure 13-72 View Extent Pools

- Another useful function in the Hardware Explorer GUI section is to identify the physical location of each FCP or FICON port. Change the search criteria to **I/O Ports** and select one or more ports in the **Available Resources** section. Use your mouse to move the yellow box in the rack image to the rear DS8000 view (bottom pane), where the I/O ports are located (Figure 13-73).

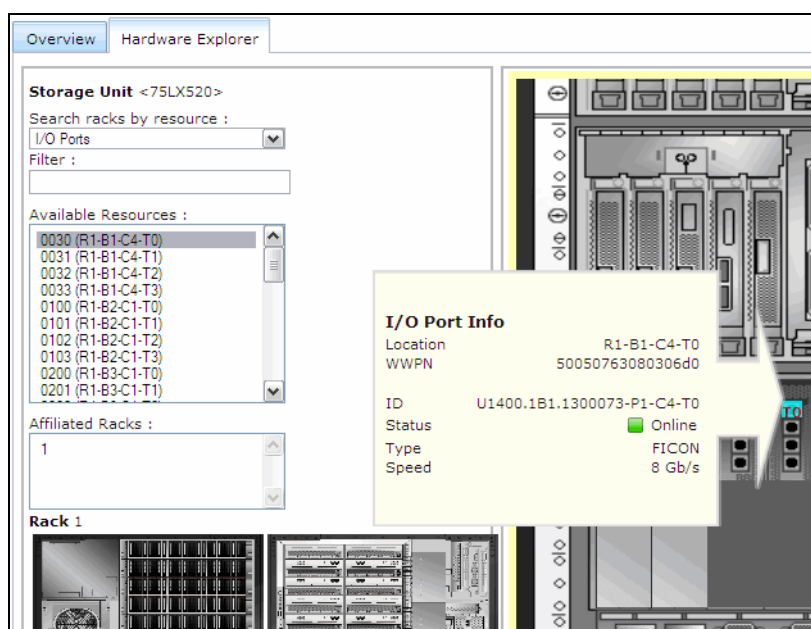


Figure 13-73 View I/O ports

Click the highlighted port to discover its basic properties and status.



Configuration with the DS Command-Line Interface

In this chapter, we explain how to configure storage on the IBM System Storage DS8000 storage subsystem by using the DS Command-Line Interface (DS CLI). We include the following sections:

- ▶ DS Command-Line Interface overview
- ▶ Configuring the I/O ports
- ▶ Configuring the DS8000 storage for FB volumes
- ▶ Configuring DS8000 Storage for Count Key Data Volumes

For information about using the DS CLI for Copy Services configuration, encryption handling, or LDAP usage, refer to the documents listed here.

For Copy Services configuration in the DS8000 using the DS CLI, see the following books:

- ▶ *IBM System Storage DS: Command-Line Interface User's Guide*, GC53-1127
- ▶ *IBM System Storage DS8000: Copy Services for Open Systems*, SG24-6788
- ▶ *IBM System Storage DS8000: Copy Services for IBM System z*, SG24-6787

For DS CLI commands related to disk encryption, refer to *IBM System Storage DS8700 Disk Encryption Implementation and Usage Guidelines*, REDP-4500.

For DS CLI commands related to LDAP authentication, refer to *IBM System Storage DS8000: LDAP Authentication*, REDP-4505.

For DS CLI commands related to Resource Groups, refer to *IBM System Storage DS8000 Resource Groups*, REDP-4758.

For DS CLI commands related to Performance I/O Priority Manager, refer to *IBM System Storage DS8000 Performance I/O Priority Manager*, REDP-4760.

For DS CLI commands related to Easy Tier, refer to *IBM System Storage DS8000 Easy Tier*, REDP-4667.

14.1 DS Command-Line Interface overview

The command-line interface provides a full-function command set that allows you to check your Storage Unit configuration and perform specific application functions when necessary. For detailed information about DS CLI use and setup, refer to *IBM System Storage DS: Command-Line Interface User's Guide*, GC53-1127.

The following list highlights a few of the functions that you can perform with the DS CLI:

- ▶ Create user IDs that can be used with the GUI and the DS CLI.
- ▶ Manage user ID passwords.
- ▶ Install activation keys for licensed features.
- ▶ Manage storage complexes and units.
- ▶ Configure and manage Storage Facility Images.
- ▶ Create and delete RAID arrays, ranks, and Extent Pools.
- ▶ Create and delete logical volumes.
- ▶ Manage host access to volumes.
- ▶ Check the current Copy Services configuration that is used by the Storage Unit.
- ▶ Create, modify, or delete Copy Services configuration settings.
- ▶ Integrate LDAP policy usage and configuration.
- ▶ Implement encryption functionality.

Tip: The DSCLI version must correspond to the LMC level installed on your system. You can have more versions of DSCLI installed on your system, each in its own directory.

14.1.1 Supported operating systems for the DS CLI

The DS Command-Line Interface can be installed on many operating system platforms, including AIX, HP-UX, Red Hat Linux, SUSE Linux, Novell NetWare, IBM i i5/OS, Oracle Solaris, HP OpenVMS, VMware ESX, and Microsoft Windows.

Important: For the most recent information about currently supported operating systems, refer to the IBM System Storage DS8000 Information Center website at:

<http://publib.boulder.ibm.com/infocenter/ds8000ic/index.jsp>

The DS CLI is supplied and installed using a CD that ships with the machine. The installation does not require a reboot of the open systems host. The DS CLI requires Java 1.4.1 or later. Java 1.4.2 is the preferred JRE on Windows, AIX, and Linux, and is supplied on the CD. Many hosts might already have a suitable level of Java installed. The installation program checks for this requirement during the installation process and does not install the DS CLI if you do not have the correct version of Java.

The installation process can be performed through a shell, such as the bash or korn shell, or the Windows command prompt, or through a GUI interface. If performed using a shell, it can be performed silently using a profile file. The installation process also installs software that allows the DS CLI to be completely uninstalled should it no longer be required.

14.1.2 User accounts

DS CLI communicates with the DS8000 system through the HMC console. Either the primary or secondary HMC console can be used. DS CLI access is authenticated using HMC user accounts. The same user IDs can be used for both DS CLI and DS GUI access. See 9.5, "HMC user management" on page 237 for further detail about user accounts.

14.1.3 DS CLI profile

To access a DS8000 system with the DS CLI, you need to provide certain information with the **dsccli** command. At a minimum, the IP address or host name of the DS8000 HMC, a user name, and a password are required. You can also provide information such as the output format for list commands, the number of rows per page in the command-line output, and whether a banner is included with the command-line output.

If you create one or more profiles to contain your preferred settings, you do not have to specify this information each time you use DS CLI. When you launch DS CLI, all you need to do is to specify a profile name with the **dsccli** command. You can override the profile's values by specifying a different parameter value with the **dsccli** command.

When you install the command-line interface software, a default profile is installed in the profile directory with the software. The file name is `dsccli.profile`, for example, `c:\Program Files\IBM\DSCLI\profile\dsccli.profile` for the Windows platform and `/opt/ibm/dsccli/profile/dsccli.profile` for UNIX and Linux platforms.

You have several options for using profile files:

- ▶ You can modify the system default profile `dsccli.profile`.
- ▶ You can make a personal default profile by making a copy of the system default profile as `<user_home>/dsccli/profile/dsccli.profile`. The default home directory `<user_home>` is designated as follows:
 - Windows system: `C:\Documents and Settings\<user_name>`
 - UNIX/Linux system: `/home/<user_name>`
- ▶ You can create specific profiles for different Storage Units and operations. Save the profile in the user profile directory. For example:
 - `c:\Program Files\IBM\DSCLI\profile\operation_name1`
 - `c:\Program Files\IBM\DSCLI\profile\operation_name2`

Attention: The default profile file created when you install the DS CLI will potentially be replaced every time you install a new version of the DS CLI. It is a good practice to open the default profile and then save it as a new file. You can then create multiple profiles and reference the relevant profile file using the `-cfg` parameter.

These profile files can be specified using the DS CLI command parameter `-cfg <profile_name>`. If the `-cfg` file is not specified, the user's default profile is used. If a user's profile does not exist, the system default profile is used.

Tip: If there are two profiles with the same name, one in default system's directory and one in your personal directory, your personal profile will be taken.

Profile change illustration

A simple way to edit the profile is to do the following:

1. From the Windows desktop, double-click the DS CLI icon.
2. In the command window that opens, enter the command **cd profile**.

3. In the profile directory, enter the command **notepad dscli.profile**, as shown in Example 14-1.

Example 14-1 Command prompt operation

```
C:\Program Files\ibm\dscli>cd profile
C:\Program Files\IBM\dscli\profile>notepad dscli.profile
```

4. The notepad opens with the DS CLI profile in it. There are four lines you can consider adding. Examples of these lines are shown in bold in Example 14-2.

Tip: The default newline delimiter is a UNIX delimiter, which can render text in notepad as one long line. Use a text editor that correctly interprets UNIX line endings.

Example 14-2 DS CLI profile example

```
# DS CLI Profile
#
# Management Console/Node IP Address(es)
# hmc1 and hmc2 are equivalent to -hmc1 and -hmc2 command options.
#hmc1:127.0.0.1
#hmc2:127.0.0.1

# Default target Storage Image ID
# "devid" and "remotedevid" are equivalent to
# "-dev storage_image_ID" and "-remotedev storeage_image_ID" command options,
# respectively.
#devid: IBM.2107-AZ12341
#remotedevid:IBM.2107-AZ12341

devid: IBM.2107-75ABCD1
hmc1: 10.0.0.250
username: admin
password: passw0rd
```

Adding the serial number by using the **devid** parameter, and the HMC IP address by using the **hmc1** parameter, is strongly suggested. Not only does this help you to avoid mistakes when using more profiles, but also you do not need to specify this parameter for certain dscli commands that require it. Additionally, if you specify dscli profile for copy services usage, then using the **remotedevid** parameter is strongly suggested for the same reasons. To determine a storage system's id, use the **lssi** CLI command.

Although adding the user name and password parameters will simplify the DS CLI startup, it is not suggested that you add them because a password is saved in clear text in the profile file. It is better to create an encrypted password file with the **managepwfile** CLI command. A password file generated using the **managepwfile** command is located in the directory `user_home_directory/dscli/profile/security/security.dat`.

Important: Use care if adding multiple **devid** and HMC entries. Only one should be uncommented (or more literally, unhashed) at any one time. If you have multiple **hmc1** or **devid** entries, the DS CLI uses the one closest to the bottom of the profile.

There are other customization parameters that affect dscli output; the most important are:

- **banner** - date and time with dscli version is printed for each command.

- ▶ **header** - column names are printed.
- ▶ **paging** - for interactive mode, it breaks output after a certain number of rows (24 by default).

14.1.4 Command structure

This is a description of the components and structure of a command-line interface command.

A command-line interface command consists of one to four types of components, arranged in the following order:

1. The command name: Specifies the task that the command-line interface is to perform.
2. Flags: Modify the command. They provide additional information that directs the command-line interface to perform the command task in a specific way.
3. Flags parameter: Provides information that is required to implement the command modification that is specified by a flag.
4. Command parameters: Provide basic information that is necessary to perform the command task. When a command parameter is required, it is always the last component of the command, and it is not preceded by a flag.

14.1.5 Using the DS CLI application

You have to log into the DS CLI application to use the command modes. There are three command modes for the DS CLI:

- ▶ Single-shot command mode
- ▶ Interactive command mode
- ▶ Script command mode

Single-shot command mode

Use the DS CLI single-shot command mode if you want to issue an occasional command but do not want to keep a history of the commands that you have issued.

You must supply the login information and the command that you want to process at the same time. Follow these steps to use the single-shot mode:

1. Enter:

```
dscli -hmc1 <hostname or ip address> -user <adm user> -passwd <pwd> <command>
```

or

```
dscli -cfg <dscli profile> <command>
```

2. Wait for the command to process and display the end results.

Example 14-3 shows the use of the single-shot command mode.

Example 14-3 Single-shot command mode

```
C:\Program Files\ibm\dscli>dscli -hmc1 10.10.10.1 -user admin -passwd pwd lsuser
Name      Group State
=====
admin     admin locked
admin     admin active
exit status of dscli = 0
```

Tip: When typing the command, you can use the host name or the IP address of the HMC. It is also important to understand that every time a command is executed in single shut mode, the user must be authenticated. The authentication process can take a considerable amount of time.

Interactive command mode

Use the DS CLI interactive command mode when you have multiple transactions to process that cannot be incorporated into a script. The interactive command mode provides a history function that makes repeating or checking prior command usage easy to do.

Perform the following steps:

1. Log on to the DS CLI application at the directory where it is installed.
2. Provide the information that is requested by the information prompts. The information prompts might not appear if you have provided this information in your profile file. The command prompt switches to a **dsccli** command prompt.
3. Begin using the DS CLI commands and parameters. You are not required to begin each command with **dsccli** because this prefix is provided by the **dsccli** command prompt.
4. Use the **quit** or **exit** command to end interactive mode.

Tip: In interactive mode for long outputs, the message Press Enter To Continue... appears. The number of rows can be specified in the profile file. Optionally, you can turn off the paging feature in the profile file by using the `paging:off` parameter.

Example 14-4 shows the use of interactive command mode.

Example 14-4 Interactive command mode

```
# dsccli -cfg ds8800.profile
dsccli> lsarraysite
arsite DA Pair dkcap (10^9B) State   Array
=====
S1      0              450.0 Assigned A0
S2      0              450.0 Assigned A1
S3      0              450.0 Assigned A2
S4      0              450.0 Assigned A3
S5      0              450.0 Assigned A4
S6      0              450.0 Assigned A5
S7      1              146.0 Assigned A6
S8      1              146.0 Assigned A7
S9      1              146.0 Assigned A8
S10     1              146.0 Assigned A9
S11     1              146.0 Assigned A10
S12     1              146.0 Assigned A11
S13     2              600.0 Assigned A12
S14     2              600.0 Assigned A13
S15     2              600.0 Assigned A14
S16     2              600.0 Assigned A15
S17     2              600.0 Assigned A16
S18     2              600.0 Assigned A17
S19     3              146.0 Assigned A18
S20     3              146.0 Assigned A19
S21     3              146.0 Assigned A20
S22     3              146.0 Assigned A21
Press Enter To Continue...
```

```

S23    3                146.0 Assigned A22
S24    3                146.0 Assigned A23
dscli> lssi
Name    ID                Storage Unit    Model WNNN                State  ESSNet
=====
ATS_04  IBM.2107-75TV181  IBM.2107-75TV180  951    500507630AFFC29F  Online Enabled

```

Tip: When typing the command, you can use the host name or the IP address of the HMC. In this case, only a single authentication need to take place.

Script command mode

Use the DS CLI script command mode if you want to use a sequence of DS CLI commands. If you want to run a script that only contains DS CLI commands, then you can start DS CLI in script mode. The script that DS CLI executes can only contain DS CLI commands.

In Example 14-5, we show the contents of a DS CLI script file. Note that it only contains DS CLI commands, although comments can be placed in the file using a hash symbol (#). Empty lines are also allowed. One advantage of using this method is that scripts written in this format can be used by the DS CLI on any operating system into which you can install DS CLI.

For script command mode, you can turn off the banner and header for easier output parsing. Also, you can specify an output format that might be easier to parse by your script.

Example 14-5 Example of a DS CLI script file

```

# Sample ds cli script file
# Comments can appear if hashed
lsarraysite
lsarray
lsrank

```

In Example 14-6, we start the DS CLI using the `-script` parameter and specifying a profile and the name of the script that contains the commands from Example 14-5.

Example 14-6 Executing DS CLI with a script file

```

C:\Program Files\ibm\dscli>dscli -cfg ds8800.profile -script c:\ds8800.script
arsite DA Pair dkcap (10^9B) State    Array
=====
S1      0                450.0 Assigned A0
S2      0                450.0 Assigned A1
S3      0                450.0 Assigned A2
S4      0                450.0 Assigned A3
CMUC00234I lsarray: No Array found.
CMUC00234I lsrank: No Rank found.

```

Tip: The DS CLI script can contain only DS CLI commands. Using shell commands results in process failure. You can add comments in the scripts prefixed by the hash symbol (#). It must be the first non-blank character on the line.

Only one single authentication process is needed to execute all the script commands.

14.1.6 Return codes

When the DS CLI exits, the exit status code is provided. This is effectively a return code. If DS CLI commands are issued as separate commands (rather than using script mode), then a return code will be presented for every command. If a DS CLI command fails (for example, due to a syntax error or the use of an incorrect password), then a failure reason and a return code will be presented. Standard techniques to collect and analyze return codes can be used.

The return codes used by the DS CLI are listed in Table 14-1.

Table 14-1 Return code table

Return code	Category	Description
0	Success	The command was successful.
2	Syntax error	There is a syntax error in the command.
3	Connection error	There was a connection problem to the server.
4	Server error	The DS CLI server had an error.
5	Authentication error	The password or user ID details are incorrect.
6	Application error	The DS CLI application had an error.

14.1.7 User assistance

The DS CLI is designed to include several forms of user assistance. The main form of user assistance is through the **help** command. Examples of usage include:

- ▶ **help** lists all the available DS CLI commands.
- ▶ **help -s** lists all the DS CLI commands with brief descriptions of each one.
- ▶ **help -l** lists all the DS CLI commands with their syntax information.

To obtain information about a specific DS CLI command, enter the command name as a parameter of the **help** command. Examples of usage include:

- ▶ **help <command name>** gives a detailed description of the specified command.
- ▶ **help -s <command name>** gives a brief description of the specified command.
- ▶ **help -l <command name>** gives syntax information about the specified command.

Example 14-7 shows the output of the **help** command.

Example 14-7 Displaying a list of all commands in DS CLI using the help command

```
# dscli -cfg ds8800.profile help
applydbcheck      lsframe           mkpe              setdbcheck
applykey          lshba            mkpprc           setdialhome
chauthpol         lshostconnect    mkpprcpath       setenv
chckdvol          lshosttype       mkrank           setflashrevertible
chextpool         lshostvol        mkrekey          setiport
chfbvol           lsiport          mkremoteflash    setnetworkport
chhostconnect     lskey            mksession        setoutput
chkeymgr          lskeygrp         mksestg          setplex
chlcu             lskeymgr         mkuser           setremoteflashrevertible
chlss            lslcu            mkvolgrp         setrmpw
chpass            lslss            offloadauditlog  setsim
chrank            lsnetworkport    offloaddbcheck   setsmtp
chsession         lspe             offloadss        setsnmp
chsestg           lspportprof      pausegmir        setvpn
chsi              lspprc           pausepprc        showarray
```

chsp	lspprcpath	quit	showarraysite
chsu	lsproblem	resumegmir	showauthpol
chuser	lsrank	resumepprc	showckdvol
chvolgrp	lsremoteflash	resyncflash	showcontactinfo
clearvol	lsserver	resyncremoteflash	showenv
closeproblem	lsession	reverseflash	showextpool
commitflash	lssestg	revertflash	showfbvol
commitremoteflash	lssi	revertremoteflash	showgmir
cpauthpol	lsss	rmarray	showgmircg
diagsi	lsstgenc1	rmauthpol	showgmiroos
dscli	lssu	rmckdvol	showhostconnect
echo	lsuser	rmextpool	showioport
exit	lsvolgrp	rmfbvol	showkeygrp
failbackpprc	lsvolinit	rmflash	showlcu
failoverpprc	lsvpn	rmgmir	showlss
freezepprc	managedbcheck	rmhostconnect	shownetworkport
help	managehostconnect	rmkeygrp	showpass
helpmsg	managepwfile	rmkeymgr	showplex
initckdvol	managereckey	rm1cu	showrank
initfbvol	mkaliasvol	rmpprc	showsestg
lsaddressgrp	mkarray	rmpprcpath	shows1
lsarray	mkauthpol	rmrank	showsp
lsarraysite	mkckdvol	rmreckey	shows1
lsauthpol	mkesconpprcpath	rmremoteflash	showuser
lsavailpprcport	mkextpool	rmsession	showvolgrp
lsckdvol	mkfbvol	rmsestg	testauthpol
lsda	mkflash	rmuser	testcallhome
lsdbcheck	mkgmir	rmvolgrp	unfreezeflash
lsddm	mkhostconnect	sendpe	unfreezepprc
lsxtpool	mkkeygrp	sendss	ver
lsfbvol	mkkeymgr	setauthpol	whoami
lsflash	mk1cu	setcontactinfo	

Man pages

A *man page* is available for every DS CLI command. Man pages are most commonly seen in UNIX-based operating systems and give information about command capabilities. This information can be displayed by issuing the relevant command followed by the `-h`, `-help`, or `-?` flags.

14.2 Configuring the I/O ports

Set the I/O ports to the desired topology. In Example 14-8, we list the I/O ports by using the `lsioport` command. Note that I0000-I0003 are on one adapter card, whereas I0100-I0103 are on another card.

Example 14-8 Listing the I/O ports

```
dscli> lsioport -dev IBM.2107-7503461
ID      WWPN              State Type          topo      portgrp
=====
I0000  500507630300008F Online Fibre Channel-SW SCSI-FCP 0
I0001  500507630300408F Online Fibre Channel-SW SCSI-FCP 0
I0002  500507630300808F Online Fibre Channel-SW SCSI-FCP 0
I0003  500507630300C08F Online Fibre Channel-SW SCSI-FCP 0
I0100  500507630308008F Online Fibre Channel-LW FICON    0
I0101  500507630308408F Online Fibre Channel-LW SCSI-FCP 0
```

I0102	500507630308808F	Online	Fibre Channel-LW	FICON	0
I0103	500507630308C08F	Online	Fibre Channel-LW	FICON	0

There are three possible topologies for each I/O port:

SCSI-FCP	Fibre Channel switched fabric (also called point to point)
FC-AL	Fibre Channel arbitrated loop
FICON	FICON (for System z hosts only)

In Example 14-9, we set two I/O ports to the FICON topology and then check the results.

Example 14-9 Changing topology using setioport

```
dscli> setioport -topology ficon I0001
CMUC00011I setioport: I/O Port I0001 successfully configured.
dscli> setioport -topology ficon I0101
CMUC00011I setioport: I/O Port I0101 successfully configured.
dscli> lsioport
```

ID	WWPN	State	Type	topo	portgrp
I0000	500507630300008F	Online	Fibre Channel-SW	SCSI-FCP	0
I0001	500507630300408F	Online	Fibre Channel-SW	FICON	0
I0002	500507630300808F	Online	Fibre Channel-SW	SCSI-FCP	0
I0003	500507630300C08F	Online	Fibre Channel-SW	SCSI-FCP	0
I0100	500507630308008F	Online	Fibre Channel-LW	FICON	0
I0101	500507630308408F	Online	Fibre Channel-LW	FICON	0
I0102	500507630308808F	Online	Fibre Channel-LW	FICON	0
I0103	500507630308C08F	Online	Fibre Channel-LW	FICON	0

14.3 Monitoring the I/O ports

Monitoring of the I/O ports is one of the most important tasks of the system administrator. Here is the point where the HBAs, SAN, and DS8700 exchange information. If one of these components has problems due to hardware or configuration issues, all the others will be affected as well.

Example 14-10 on page 373 shows the output of the **showiopo** **-metrics** command, which illustrates the many important metrics returned by the command. It provides the performance counter of the port and the FCLink error counter. The FCLink error counter is used to determine the health of the overall communication.

There are groups of errors that point to specific problem areas:

- ▶ Any non-zero figure in the counters LinkFailErr, LossSyncErr, LossSigErr, and PrimSeqErr indicates that the SAN probably has HBAs attached to it that are unstable. These HBAs log in and log out to the SAN and create name server congestion and performance degradation.
- ▶ If the InvTxWordErr counter increases by more than 100 per day, the port is receiving light from a source that is not an SFP. The cable connected to the port is not covered at the end or the I/O port is not covered by a cap.
- ▶ The CRCErr counter shows the errors that arise between the last sending SFP in the SAN and the receiving port of the DS8700. These errors do not appear in any other place in the data center. You must replace the cable that is connected to the port or the SFP in the SAN.

- ▶ The link reset counters LRSent and LRRec also suggest that there are hardware defects in the SAN; these errors need to be investigated.
- ▶ The counters IllegalFrame, OutOrdData, OutOrdACK, DupFrame, InvRelOffset, SeqTimeout, and BitErrRate point to congestions in the SAN and can only be influenced by configuration changes in the SAN.

Example 14-10 Listing the I/O ports with showioport -metrics

```

dscli> showioport -dev IBM.2107-7503461 -metrics I0041
ID                               I0041
Date                             09/30/2009 16:24:12 MST
<< cut here >>
LinkFailErr (FC)                 0
LossSyncErr (FC)                 0
LossSigErr (FC)                  0
PrimSeqErr (FC)                  0
InvTxWordErr (FC)                0
CRCErr (FC)                      0
LRSent (FC)                      0
LRRec (FC)                       0
IllegalFrame (FC)                0
OutOrdData (FC)                  0
OutOrdACK (FC)                   0
DupFrame (FC)                    0
InvRelOffset (FC)                0
SeqTimeout (FC)                  0
BitErrRate (FC)                  0

```

14.4 Configuring the DS8000 storage for FB volumes

This section goes through examples of a typical DS8000 storage configuration when attaching to open systems hosts. We perform the DS8000 storage configuration by going through the following steps:

1. Create arrays.
2. Create ranks.
3. Create Extent Pools.
4. Optionally, create repositories for track space efficient volumes.
5. Create volumes.
6. Create volume groups.
7. Create host connections.

14.4.1 Create arrays

In this step, we create the arrays. Before creating the arrays, it is a best practice to first list the arrays sites. Use the **lsarraysite** to list the array sites, as shown in Example 14-11.

Important: Remember that an array for a DS8000 can only contain one array site, and a DS8000 array site contains eight disk drive modules (DDMs).

Example 14-11 Listing array sites

```

dscli> lsarraysite
arsite DA Pair dkcap (10^9B) State      Array

```

```
=====
S1      0      146.0 Unassigned -
S2      0      146.0 Unassigned -
S3      0      146.0 Unassigned -
S4      0      146.0 Unassigned -
=====
```

In Example 14-11 on page 373, we can see that there are four array sites and that we can therefore create four arrays.

We can now issue the **mkarray** command to create arrays, as shown in Example 14-12. You will notice that in this case we have used one array site (in the first array, S1) to create a single RAID 5 array. If we wished to create a RAID 10 array, we would have to change the **-raidtype** parameter to 10, and if we wished to create a RAID 6 array, we would have to change the **-raidtype** parameter to 6 (instead of 5).

Example 14-12 Creating arrays with mkarray

```
=====
dscli> mkarray -raidtype 5 -arsite S1
CMUC00004I mkarray: Array A0 successfully created.
dscli> mkarray -raidtype 5 -arsite S2
CMUC00004I mkarray: Array A1 successfully created.
=====
```

We can now see what arrays have been created by using the **lsarray** command, as shown in Example 14-13.

Example 14-13 Listing the arrays with lsarray

```
=====
dscli> lsarray
Array State      Data  RAIDtype  arsite Rank DA Pair DDMcap (10^9B)
=====
A0  Unassigned  Normal 5 (6+P+S) S1      -    0      146.0
A1  Unassigned  Normal 5 (6+P+S) S2      -    0      146.0
=====
```

We can see in this example the type of RAID array and the number of disks that are allocated to the array (in this example 6+P+S, which means the usable space of the array is 6 times the DDM size), the capacity of the DDMs that are used, and which array sites were used to create the arrays.

14.4.2 Create ranks

After we have created all the arrays that are required, we then create the ranks using the **mkrank** command. The format of the command is **mkrank -array Ax -stgtype xxx**, where **xxx** is either fixed block (FB) or count key data (CKD), depending on whether you are configuring for open systems or System z hosts.

After we have created all the ranks, we run the **lsrank** command. This command displays all the ranks that have been created, to which server the rank is attached, the RAID type, and the format of the rank, whether it is Fixed Block (FB) or Count Key Data (CKD).

Example 14-14 shows the **mkrank** commands and the result of a successful **lsrank -1** command.

Example 14-14 Creating and listing ranks with mkrank and lsrank

```
=====
dscli> mkrank -array A0 -stgtype fb
CMUC00007I mkrank: Rank R0 successfully created.
dscli> mkrank -array A1 -stgtype fb
=====
```

```
CMUC00007I mkrank: Rank R1 successfully created.
```

```
dscli> lsrank -l
```

ID	Group	State	datastate	Array	RAIDtype	extpoolID	extpoolnam	stgtype	exts	usedexts
R0	-	Unassigned	Normal	A0	5	-	-	fb	773	-
R1	-	Unassigned	Normal	A1	5	-	-	fb	773	-

14.4.3 Create Extent Pools

The next step is to create Extent Pools. Here are points to remember when creating the Extent Pools:

- ▶ Each Extent Pool has an associated rank group that is specified by the `-rankgrp` parameter, which defines the Extent Pools' server affinity (either 0 or 1, for server0 or server1).
- ▶ The Extent Pool type is either FB or CKD and is specified by the `-stgtype` parameter.
- ▶ The number of Extent Pools can range from one to as many as there are existing ranks. However, to associate ranks with both servers, you need at least two Extent Pools.
- ▶ It is best practice for all ranks in an Extent Pool to have the same characteristics, that is, the same DDM type, size, and RAID type.

For easier management, we create empty Extent Pools related to the type of storage that is in the pool. For example, create an Extent Pool for high capacity disk, create another for high performance, and, if needed, Extent Pools for the CKD environment.

When an Extent Pool is created, the system automatically assigns it an Extent Pool ID, which is a decimal number starting from 0, preceded by the letter P. The ID that was assigned to an Extent Pool is shown in the CMUC00000I message, which is displayed in response to a successful `mkextpool` command. Extent pools associated with rank group 0 get an even ID number. Extent pools associated with rank group 1 get an odd ID number. The Extent Pool ID is used when referring to the Extent Pool in subsequent CLI commands. It is therefore good practice to make note of the ID.

Example 14-15 shows one example of Extent Pools you could define on your machine. This setup would require a system with at least six ranks.

Example 14-15 An Extent Pool layout plan

```
FB Extent Pool high capacity 300gb disks assigned to server 0 (FB_LOW_0)
FB Extent Pool high capacity 300gb disks assigned to server 1 (FB_LOW_1)
FB Extent Pool high performance 146gb disks assigned to server 0 (FB_High_0)
FB Extent Pool high performance 146gb disks assigned to server 1 (FB_High_1)
CKD Extent Pool High performance 146gb disks assigned to server 0 (CKD_High_0)
CKD Extent Pool High performance 146gb disks assigned to server 1 (CKD_High_1)
```

Note that the `mkextpool` command forces you to name the Extent Pools. In Example 14-16, we first create empty Extent Pools using the `mkextpool` command. We then list the Extent Pools to get their IDs. Then we attach a rank to an empty Extent Pool using the `chrank` command. Finally, we list the Extent Pools again using `lsxtpool` and note the change in the capacity of the Extent Pool.

Example 14-16 Extent Pool creation using `mkextpool`, `lsxtpool`, and `chrank`

```
dscli> mkextpool -rankgrp 0 -stgtype fb FB_high_0
CMUC00000I mkextpool: Extent Pool P0 successfully created.
dscli> mkextpool -rankgrp 1 -stgtype fb FB_high_1
```

```
CMUC00000I mkextpool: Extent Pool P1 successfully created.
dscli> lsextpool
Name          ID stgtype rankgrp status availstor (2^30B) %allocated available reserved numvols
=====
FB_high_0 P0 fb          0 below          0          0          0          0          0
FB_high_1 P1 fb          1 below          0          0          0          0          0
dscli> chrnk -extpool P0 R0
CMUC00008I chrnk: Rank R0 successfully modified.
dscli> chrnk -extpool P1 R1
CMUC00008I chrnk: Rank R1 successfully modified.
dscli> lsextpool
Name          ID stgtype rankgrp status availstor (2^30B) %allocated available reserved numvols
=====
FB_high_0 P0 fb          0 below          773          0          773          0          0
FB_high_1 P1 fb          1 below          773          0          773          0          0
```

After having assigned a rank to an Extent Pool, we should be able to see this change when we display the ranks. In Example 14-17, we can see that rank R0 is assigned to extpool P0.

Example 14-17 Displaying the ranks after assigning a rank to an Extent Pool

```
dscli> lsrank -l
ID Group State  datastate Array RAIDtype extpoolID extpoolnam stgtype exts usedexts
=====
R0    0 Normal Normal  A0          5 P0      FB_high_0 fb    773      0
R1    1 Normal Normal  A1          5 P1      FB_high_1 fb    773      0
```

Creating a repository for Track Space Efficient volumes

If the DS8000 has the IBM FlashCopy SE feature, you can create Track Space Efficient (TSE) volumes that can be used as FlashCopy targets. Before you can create TSE volumes, you must create a space efficient repository in the Extent Pool. The repository provides space to store the data associated with TSE logical volumes. Only one repository is allowed per Extent Pool. A repository has a physical capacity that is available for storage allocations by TSE volumes and a virtual capacity that is the sum of LUN/volume sizes of all space efficient volumes. The physical repository capacity is allocated when the repository is created. If there are several ranks in the Extent Pool, the repository's extents are striped across the ranks (Storage Pool Striping).

Example 14-18 shows the creation of a repository. The unit type of the real capacity (-repcap) and virtual capacity (-vircap) sizes can be specified with the -captype parameter. For FB Extent Pools, the unit type can be either GB (default) or blocks.

Example 14-18 Creating a repository for Space Efficient volumes

```
dscli> mksestg -repcap 100 -vircap 200 -extpool p9
CMUC00342I mksestg: The space-efficient storage for the Extent Pool P9 has been
created successfully.
```

You can obtain information about the repository with the **showsestg** command. Example 14-19 shows the output of the **showsestg** command. You might particularly be interested in how much capacity is used within the repository by checking the **repcapalloc** value.

Example 14-19 Getting information about a Space Efficient repository

```
dscli> showsestg p9
extpool          P9
stgtype          fb
```

```

datastate           Normal
configstate         Normal
repcapstatus        below
%repcapthreshold    0
repcap(GiB)         100.0
repcap(Mod1)        -
repcap(blocks)      209715200
repcap(cyl)         -
repcapalloc(GiB/Mod1) 0.0
%repcapalloc        0
viricap(GiB)        200.0
viricap(Mod1)        -
viricap(blocks)     419430400
viricap(cyl)        -
viricapalloc(GiB/Mod1) 0.0
%viricapalloc        0
overhead(GiB/Mod1)  3.0
reqrepcap(GiB/Mod1) 100.0
reqviricap(GiB/Mod1) 200.0

```

Note that more storage is allocated for the repository in addition to repcap size. In Example 14-19 on page 376, the line that starts with overhead indicates that 3 GB had been allocated in addition to the repcap size.

A repository can be deleted with the **rmsestg** command.

Tip: In the current implementation, it is not possible to expand a Space Efficient repository. The physical size or the virtual size of the repository cannot be changed. Therefore, careful planning is required. If you have to expand a repository, you must delete all TSE logical volumes and the repository itself, then recreate a new repository.

14.4.4 Creating FB volumes

We are now able to create volumes and volume groups. When we create them, we should try to distribute them evenly across the two rank groups in the storage unit.

Creating standard volumes

The format of the command that we use to create a volume is:

```
mkfbvol -extpool pX -cap xx -name high_fb_0#h 1000-1003
```

In Example 14-20, we have created eight volumes, each with a capacity of 10 GB. The first four volumes are assigned to rank group 0 and the second four are assigned to rank group 1.

Example 14-20 Creating fixed block volumes using mkfbvol

```

dscli> lsextpool
Name      ID stgtype rankgrp status availstor (2^30B) %allocated available reserved numvols
=====
FB_high_0 P0 fb          0  below          773          0          773          0          0
FB_high_1 P1 fb          1  below          773          0          773          0          0
dscli> mkfbvol -extpool p0 -cap 10 -name high_fb_0#h 1000-1003
CMUC00025I mkfbvol: FB volume 1000 successfully created.
CMUC00025I mkfbvol: FB volume 1001 successfully created.
CMUC00025I mkfbvol: FB volume 1002 successfully created.
CMUC00025I mkfbvol: FB volume 1003 successfully created.

```

```

dscli> mkfbvol -extpool p1 -cap 10 -name high_fb_1_#h 1100-1103
CMUC00025I mkfbvol: FB volume 1100 successfully created.
CMUC00025I mkfbvol: FB volume 1101 successfully created.
CMUC00025I mkfbvol: FB volume 1102 successfully created.
CMUC00025I mkfbvol: FB volume 1103 successfully created.

```

Looking closely at the **mkfbvol** command used in Example 14-20, we see that volumes 1000 - 1003 are in extpool P0. That Extent Pool is attached to rank group 0, which means server 0. Now rank group 0 can only contain even numbered LSSs, so that means volumes in that Extent Pool must belong to an even numbered LSS. The first two digits of the volume serial number are the LSS number, so in this case, volumes 1000 - 1003 are in LSS 10.

For volumes 1100 - 1003 in Example 14-20 on page 377, the first two digits of the volume serial number are 11, which is an odd number, which signifies they belong to rank group 1. Also note that the **-cap** parameter determines size, but because the **-type** parameter was not used, the default size is a binary size. So these volumes are 10 GB binary, which equates to 10,737,418,240 bytes. If we used the parameter **-type ess**, then the volumes would be decimally sized and would be a minimum of 10,000,000,000 bytes in size.

In Example 14-20 on page 377 we named the volumes using naming scheme **high_fb_0_#h**, where **#h** means you are using the hexadecimal volume number as part of the volume name. This can be seen in Example 14-21, where we list the volumes that we have created using the **lsfbvol** command. We then list the Extent Pools to see how much space we have left after the volume creation.

*Example 14-21 Checking the machine after creating volumes by using **lsxtpool** and **lsfbvol***

```

dscli> lsfbvol
Name          ID   accstate  datastate  configstate  deviceMTM  datatype  extpool  cap (2^30B)
=====
high_fb_0_1000 1000 Online    Normal     Normal      2107-922   FB 512    P0        10.0
high_fb_0_1001 1001 Online    Normal     Normal      2107-922   FB 512    P0        10.0
high_fb_0_1002 1002 Online    Normal     Normal      2107-922   FB 512    P0        10.0
high_fb_0_1003 1003 Online    Normal     Normal      2107-922   FB 512    P0        10.0
high_fb_1_1100 1100 Online    Normal     Normal      2107-922   FB 512    P1        10.0
high_fb_1_1101 1101 Online    Normal     Normal      2107-922   FB 512    P1        10.0
high_fb_1_1102 1102 Online    Normal     Normal      2107-922   FB 512    P1        10.0
high_fb_1_1103 1103 Online    Normal     Normal      2107-922   FB 512    P1        10.0
dscli> lsxtpool
Name          ID  stgtype  rankgrp  status  availstor (2^30B)  %allocated  available  reserved  numvols
=====
FB_high_0 P0 fb          0  below          733           5          733         0         4
FB_high_1 P1 fb          1  below          733           5          733         0         4

```

Important: For the DS8000, the LSSs can be ID 00 to ID FE. The LSSs are in address groups. Address group 0 is LSS 00 to 0F, address group 1 is LSS 10 to 1F, and so on. The moment you create an FB volume in an address group, then that entire address group can only be used for FB volumes. Be aware of this fact when planning your volume layout in a mixed FB/CKD DS8000.

Tip: With DS8000 with Licensed Machine Code Release 6.1, you can configure a volume to belong to a certain Resource Group using the **-resgrp <RG_ID>** flag in the **mkfbvol** command. For more details, refer to IBM System Storage DS8000: Copy Services Resource Groups, REDP-4758.

You can configure a volume belong to a certain Performance I/O Priority Manager using **-perfgrp <perf_group_ID>** flag in the **mkfbvol** command. For more details, refer to *IBM System Storage DS8000 I/O Priority Manager*, REDP-4760.

Storage Pool Striping

When creating a volume, you have a choice of how the volume is allocated in an Extent Pool with several ranks. The extents of a volume can be kept together in one rank (as long as there is enough free space on that rank). The next rank is used when the next volume is created. This allocation method is called *rotate volumes*.

You can also specify that you want the extents of the volume you are creating to be evenly distributed across all ranks within the Extent Pool. This allocation method is called *rotate extents*. The Storage Pool Striping spreads the IO of a LUN to multiple ranks, improving performance and also greatly reducing ‘hot spots’.

The extent allocation method is specified with the **-eam rotateexts** or **-eam rotatevols** option of the **mkfbvol** command (see Example 14-22).

Tip: In DS8800 with Licensed Machine Code (LMC) 6.6.xxx, the default allocation policy has changed to *rotate extents*.

Example 14-22 Creating a volume with Storage Pool Striping

```
dsccli> mkfbvol -extpool p53 -cap 15 -name ITS0-XPSTR -eam rotateexts 1720
CMUC00025I mkfbvol: FB volume 1720 successfully created.
```

The **showfbvol** command with the **-rank** option (see Example 14-23) shows that the volume we created is distributed across 12 ranks and how many extents on each rank were allocated for this volume.

Example 14-23 Getting information about a striped volume

```
dsccli> showfbvol -rank 1720
Name           ITS0-XPSTR
ID             1720
accstate       Online
datastate      Normal
configstate    Normal
deviceMTM      2107-900
datatype       FB 512
addrgrp        1
extpool        P53
exts           15
captype        DS
cap (2^30B)    15.0
cap (10^9B)    -
cap (blocks)   31457280
volgrp         -
ranks          12
dbexts         0
sam            Standard
repcapalloc    -
eam            rotateexts
reqcap (blocks) 31457280
=====Rank extents=====
rank extents
=====
```

R24	2
R25	1
R28	1
R29	1
R32	1
R33	1
R34	1
R36	1
R37	1
R38	1
R40	2
R41	2

Track Space Efficient volumes

When your DS8000 has the IBM FlashCopy SE feature, you can create Track Space Efficient (TSE) volumes to be used as FlashCopy target volumes. A repository must exist in the Extent Pool where you plan to allocate TSE volumes (see “Creating a repository for Track Space Efficient volumes” on page 376).

A Track Space Efficient volume is created by specifying the `-sam tse` parameter with the `mkfbvol` command (Example 14-24).

Example 14-24 Creating a Space Efficient volume

```
dscli> mkfbvol -extpool p53 -cap 40 -name ITS0-1721-SE -sam tse 1721
CMUC00025I mkfbvol: FB volume 1721 successfully created.
```

When listing Space Efficient repositories with the `lssestg` command (see Example 14-25), we can see that in Extent Pool P53 we have a virtual allocation of 40 extents (GB), but that the allocated (used) capacity `repcapalloc` is still zero.

Example 14-25 Getting information about Space Efficient repositories

```
dscli> lssestg -l
extentpoolID stgtype datastate configstate repcapstatus %repcapthreshold repcap (2^30B) vircap repcapalloc vircapalloc
=====
P4           ckd      Normal    Normal    below          0          64.0   1.0          0.0          0.0
P47          fb       Normal    Normal    below          0          70.0  282.0         0.0         264.0
P53          fb       Normal    Normal    below          0          100.0 200.0         0.0          40.0
```

This allocation comes from the volume just created. To see the allocated space in the repository for just this volume, we can use the `showfbvol` command (see Example 14-26).

Example 14-26 Checking the repository usage for a volume

```
dscli> showfbvol 1721
Name           ITS0-1721-SE
ID             1721
accstate       Online
datastate      Normal
configstate    Normal
deviceMTM      2107-900
datatype       FB 512
addrgrp        1
extpool        P53
exts           40
captype        DS
cap (2^30B)    40.0
```

```
cap (10^9B)      -
cap (blocks)     83886080
volgrp           -
ranks            0
dbexts           0
sam              TSE
repcapalloc      0
eam              -
reqcap (blocks)  83886080
```

Dynamic Volume Expansion

A volume can be expanded without having to remove the data within the volume. You can specify a new capacity by using the **chfbvol** command (Example 14-27).

The largest LUN size is now 16 TB. Copy services are not supported for LUN sizes larger than 2 TB.

Tip: The new capacity must be larger than the previous one. You *cannot* shrink the volume.

Example 14-27 Expanding a striped volume

```
dscli> chfbvol -cap 40 1720
CMUC00332W chfbvol: Some host operating systems do not support changing the volume
size. Are you sure that you want to resize the volume? [y/n]: y
CMUC00026I chfbvol: FB volume 1720 successfully modified.
```

Because the original volume had the rotateexts attribute, the additional extents are also striped (see Example 14-28).

Example 14-28 Checking the status of an expanded volume

```
dscli> showfbvol -rank 1720
Name           ITS0-XPSTR
ID             1720
accstate       Online
datastate      Normal
configstate    Normal
deviceMTM      2107-900
datatype       FB 512
addrgrp        1
extpool        P53
exts           40
captype        DS
cap (2^30B)    20.0
cap (10^9B)    -
cap (blocks)   41943040
volgrp         -
ranks          2
dbexts         0
sam            Standard
repcapalloc    -
eam            rotateexts
reqcap (blocks) 41943040
=====Rank extents=====
rank extents
=====
R24           20
```

Important: Before you can expand a volume, you must delete all Copy Services relationships for that volume.

Deleting volumes

FB volumes can be deleted by using the **rmfbvol** command.

Starting with Licensed Machine Code (LMC) level 6.5.1.xx, the command includes new options to prevent the accidental deletion of volumes that are in use. A FB volume is considered to be “in use”, if it is participating in a Copy Services relationship or if the volume has received any I/O operation in the previous 5 minutes.

Volume deletion is controlled by the **-safe** and **-force** parameters (they cannot be specified at the same time) as follows:

- ▶ If neither of the parameters is specified, the system performs checks to see whether or not the specified volumes are in use. Volumes that are not in use will be deleted and the ones in use will not be deleted.
- ▶ If the **-safe** parameter is specified, and if any of the specified volumes are assigned to a user-defined volume group, the command fails without deleting any volumes.
- ▶ The **-force** parameter deletes the specified volumes without checking to see whether or not they are in use.

In Example 14-29, we create volumes 2100 and 2101. We then assign 2100 to a volume group. We then try to delete both volumes with the **-safe** option, but the attempt fails without deleting either of the volumes. We are able to delete volume 2101 with the **-safe** option because it is not assigned to a volume group. Volume 2100 is not in use, so we can delete it by not specifying either parameter.

Example 14-29 Deleting a FB volume

```
dsccli> mkfbvol -extpool p1 -cap 12 -eam rotateexts 2100-2101
CMUC00025I mkfbvol: FB volume 2100 successfully created.
CMUC00025I mkfbvol: FB volume 2101 successfully created.
dsccli> chvolgrp -action add -volume 2100 v0
CMUC00031I chvolgrp: Volume group V0 successfully modified.
dsccli> rmfbvol -quiet -safe 2100-2101
CMUC00253E rmfbvol: Volume IBM.2107-75NA901/2100 is assigned to a user-defined volume
group. No volumes were deleted.
dsccli> rmfbvol -quiet -safe 2101
CMUC00028I rmfbvol: FB volume 2101 successfully deleted.
dsccli> rmfbvol 2100
CMUC00027W rmfbvol: Are you sure you want to delete FB volume 2100? [y/n]: y
CMUC00028I rmfbvol: FB volume 2100 successfully deleted.
```

14.4.5 Creating volume groups

Fixed block volumes are assigned to open systems hosts using volume groups, which is not to be confused with the term *volume groups* used in AIX. A fixed block volume can be a member of multiple volume groups. Volumes can be added or removed from volume groups as required. Each volume group must be either SCSI MAP256 or SCSI MASK, depending on the SCSI LUN address discovery method used by the operating system to which the volume group will be attached.

Determining if an open systems host is SCSI MAP256 or SCSI MASK

First, we determine what sort of SCSI host with which we are working. Then we use the **lshosttype** command with the **-type** parameter of **scsimask** and then **scsimap256**.

In Example 14-30, we can see the results of each command.

Example 14-30 Listing host types with the lshosttype command

```
dscli> lshosttype -type scsimask
HostType Profile                               AddrDiscovery LBS
=====
Hp        HP - HP/UX                               reportLUN    512
SVC       San Volume Controller                     reportLUN    512
SanFsAIX  IBM pSeries - AIX/SanFS                       reportLUN    512
pSeries   IBM pSeries - AIX                               reportLUN    512
zLinux    IBM zSeries - zLinux                           reportLUN    512
dscli> lshosttype -type scsimap256
HostType Profile                               AddrDiscovery LBS
=====
AMDLinuxRHEL AMD - Linux RHEL                           LUNPolling   512
AMDLinuxSuse AMD - Linux Suse                     LUNPolling   512
AppleOSX     Apple - OSX                                       LUNPolling   512
Fujitsu      Fujitsu - Solaris                               LUNPolling   512
HpTru64      HP - Tru64                                       LUNPolling   512
HpVms        HP - Open VMS                                    LUNPolling   512
LinuxDT      Intel - Linux Desktop                           LUNPolling   512
LinuxRF      Intel - Linux Red Flag                         LUNPolling   512
LinuxRHEL    Intel - Linux RHEL                             LUNPolling   512
LinuxSuse    Intel - Linux Suse                             LUNPolling   512
Novell       Novell                                           LUNPolling   512
SGI          SGI - IRIX                                       LUNPolling   512
SanFsLinux   - Linux/SanFS                                  LUNPolling   512
Sun          SUN - Solaris                                   LUNPolling   512
VMWare       VMWare                                          LUNPolling   512
Win2000      Intel - Windows 2000                           LUNPolling   512
Win2003      Intel - Windows 2003                           LUNPolling   512
Win2008      Intel - Windows 2008                           LUNPolling   512
iLinux       IBM iSeries - iLinux                           LUNPolling   512
nSeries      IBM N series Gateway                           LUNPolling   512
pLinux       IBM pSeries - pLinux                           LUNPolling   512
```

Create Volume Group

Having determined the host type, we can now make a volume group. In Example 14-31, the example host type we chose is AIX, and in Example 14-30, we can see the address discovery method for AIX is **scsimask**.

Example 14-31 Creating a volume group with mkvolgrp and displaying it

```
dscli> mkvolgrp -type scsimask -volume 1000-1002,1100-1102 AIX_VG_01
CMUC00030I mkvolgrp: Volume group V11 successfully created.
dscli> lsvolgrp
Name                ID Type
=====
ALL CKD             V10 FICON/ESCON A11
AIX_VG_01           V11 SCSI Mask
ALL Fixed Block-512 V20 SCSI A11
ALL Fixed Block-520 V30 OS400 A11
dscli> showvolgrp V11
Name AIX_VG_01
ID    V11
Type  SCSI Mask
```

Vol's 1000 1001 1002 1100 1101 1102

Add or Delete volume on a Volume Group

In this example, we added volumes 1000 to 1002 and 1100 to 1102 to the new volume group. We did this task to spread the workload evenly across the two rank groups. We then listed all available volume groups using **lsvolgrp**. Finally, we listed the contents of volume group V11, because this was the volume group we created.

We might also want to add or remove volumes to this volume group at a later time. To achieve this goal, we use **chvolgrp** with the **-action** parameter. In Example 14-32, we add volume 1003 to volume group V11. We display the results, and then remove the volume.

Example 14-32 Changing a volume group with chvolgrp

```
dscli> chvolgrp -action add -volume 1003 V11
CMUC00031I chvolgrp: Volume group V11 successfully modified.
dscli> showvolgrp V11
Name AIX_VG_01
ID V11
Type SCSI Mask
Vol's 1000 1001 1002 1003 1100 1101 1102
dscli> chvolgrp -action remove -volume 1003 V11
CMUC00031I chvolgrp: Volume group V11 successfully modified.
dscli> showvolgrp V11
Name AIX_VG_01
ID V11
Type SCSI Mask
Vol's 1000 1001 1002 1100 1101 1102
```

Important: Not all operating systems can deal with the removal of a volume. Consult your operating system documentation to determine the safest way to remove a volume from a host.

All operations with volumes and volume groups described previously can also be used with Space Efficient volumes as well.

14.4.6 Creating host connections

The final step in the logical configuration process is to create host connections for your attached hosts. You will need to assign volume groups to those connections. Each host HBA can only be defined once, and each host connection (hostconnect) can only have one volume group assigned to it. Remember that a volume can be assigned to multiple volume groups.

In Example 14-33, we create a single host connection that represents one HBA in our example AIX host. We use the **-hosttype** parameter using the hosttype we have in Example 14-30 on page 383. We allocated it to volume group V11. At this point, provided that the SAN zoning is correct, the host should be able to see the logical unit numbers (LUNs) in volume group V11.

Example 14-33 Creating host connections using mkhostconnect and lshostconnect

```
dscli> mkhostconnect -wwname 100000C912345678 -hosttype pSeries -volgrp V11 AIX_Server_01
CMUC00012I mkhostconnect: Host connection 0000 successfully created.
dscli> lshostconnect
Name          ID      WWPN          HostType Profile      portgrp volgrpID ESSIOport
```

```
=====
AIX_Server_01 0000 100000C912345678 pSeries IBM pSeries - AIX      0 V11      all
=====
```

You can also use simply `-profile` instead of `-hosttype`. However, this is not a best practice. Using the `-hosttype` parameter actually invokes both parameters (`-profile` and `-hosttype`). In contrast, simply using `-profile` leaves the `-hosttype` column unpopulated.

There is also the option in the `mkhostconnect` command to restrict access to only certain I/O ports. This is done with the `-ioport` parameter. Restricting access in this way is usually unnecessary. If you want to restrict access for certain hosts to certain I/O ports on the DS8000, do this by way of zoning on your SAN switch.

Managing hosts with multiple HBAs

If you have a host with multiple HBAs, you have two considerations:

- ▶ For the GUI to consider multiple host connects to be used by the same server, the host connects must have the same name. In Example 14-34, host connects 0010 and 0011 appear in the GUI as a single server with two HBAs. However, host connects 000E and 000F appear as two separate hosts even though in reality they are used by the same server. If you do not plan to use the GUI to manage host connections, then this is not a major consideration. Using more verbose hostconnect naming might make management easier.
- ▶ If you want to use a single command to change the assigned volume group of several hostconnects at the same time, then you need to assign these hostconnects to a unique port group and then use the `managehostconnect` command. This command changes the assigned volume group for all hostconnects assigned to a particular port group.

When creating hosts, you can specify the `-portgrp` parameter. By using a unique port group number for each attached server, you can easily detect servers with multiple HBAs.

In Example 14-34, we have six host connections. By using the port group number, we see that there are three separate hosts, each with two HBAs. Port group 0 is used for all hosts that do not have a port group number set.

Example 14-34 Using the portgrp number to separate attached hosts

dscli> lshostconnect								
Name	ID	WWPN	HostType	Profile		portgrp	volgrpID	
=====								
bench_tic17_fc0	0008	210000E08B1234B1	LinuxSuse	Intel - Linux Suse		8 V1		all
bench_tic17_fc1	0009	210000E08B12A3A2	LinuxSuse	Intel - Linux Suse		8 V1		all
p630_fcs0	000E	10000000C9318C7A	pSeries	IBM pSeries - AIX		9 V2		all
p630_fcs1	000F	10000000C9359D36	pSeries	IBM pSeries - AIX		9 V2		all
p615_7	0010	10000000C93E007C	pSeries	IBM pSeries - AIX		10 V3		all
p615_7	0011	10000000C93E0059	pSeries	IBM pSeries - AIX		10 V3		all

Changing host connections

If we want to change a host connection, we can use the `chhostconnect` command. This command can be used to change nearly all parameters of the host connection except for the worldwide port name (WWPN). If you need to change the WWPN, you need to create a whole new host connection. To change the assigned volume group, use either `chhostconnect` to change one hostconnect at a time, or use the `managehostconnect` command to simultaneously reassign all the hostconnects in one port group.

14.4.7 Mapping open systems host disks to storage unit volumes

When you have assigned volumes to an open systems host, and you have then installed the DS CLI on this host, you can run the DS CLI command **lshostvol** on this host. This command maps assigned LUNs to open systems host volume names.

In this section, we give examples for several operating systems. In each example, we assign several logical volumes to an open systems host. We install DS CLI on this host. We log on to this host and start DS CLI. It does not matter which HMC we connect to with the DS CLI. We then issue the **lshostvol** command.

Important: The **lshostvol** command communicates only with the operating system of the host on which the DS CLI is installed. You cannot run this command on one host to see the attached disks of another host.

AIX: Mapping disks when using Multipath I/O

In Example 14-35, we have an AIX server that uses Multipath I/O (MPIO). We have two volumes assigned to this host, 1800 and 1801. Because MPIO is used, we do not see the number of paths.

In fact, from this display, it is not possible to tell if MPIO is even installed. You need to run the **pcmpath query device** command to confirm the path count.

Example 14-35 lshostvol on an AIX host using MPIO

```
dscli> lshostvol
Disk Name Volume Id          Vpath Name
=====
hdisk3    IBM.2107-1300819/1800 ---
hdisk4    IBM.2107-1300819/1801 ---
```

Tip: If you use Open HyperSwap on a host, the **lshostvol** command might fail to show any devices

AIX: Mapping disks when Subsystem Device Driver is used

In Example 14-36, we have an AIX server that uses Subsystem Device Driver (SDD). We have two volumes assigned to this host, 1000 and 1100. Each volume has four paths.

Example 14-36 lshostvol on an AIX host using SDD

```
dscli> lshostvol
Disk Name          Volume Id          Vpath Name
=====
hdisk1,hdisk3,hdisk5,hdisk7 IBM.2107-1300247/1000 vpath0
hdisk2,hdisk4,hdisk6,hdisk8 IBM.2107-1300247/1100 vpath1
```

Hewlett-Packard UNIX (HP-UX): mapping disks when not using SDD

In Example 14-37, we have an HP-UX host that does not have SDD. We have two volumes assigned to this host, 1105 and 1106.

Example 14-37 lshostvol on an HP-UX host that does not use SDD

```
dscli> lshostvol
Disk Name Volume Id          Vpath Name
=====
c38t0d5    IBM.2107-7503461/1105 ---
```

HP-UX or Solaris: Mapping disks when using SDD

In Example 14-38, we have a Solaris host that has SDD installed. Two volumes are assigned to the host, 4205 and 4206 using two paths. The Solaris command **iostat -En** can also produce similar information. The output of **lshostvol** on an HP-UX host looks exactly the same, with each vpath made up of disks with controller, target, and disk (c-t-d) numbers. However, the addresses used in the example for the Solaris host would not work in an HP-UX system.

Attention: Current releases of HP-UX only support addresses up to 3FFF.

Example 14-38 lshostvol on a Solaris host that has SDD

```
dsccli> lshostvol
Disk Name      Volume Id      Vpath Name
=====
c2t1d0s0,c3t1d0s0 IBM.2107-7520781/4205 vpath2
c2t1d1s0,c3t1d1s0 IBM.2107-7520781/4206 vpath1
```

Solaris: Mapping disks when not using SDD

In Example 14-39, we have a Solaris host that does not have SDD installed. It instead uses an alternative multipathing product. We have two volumes assigned to this host, 4200 and 4201. Each volume has two paths. The Solaris command **iostat -En** can also produce similar information.

Example 14-39 lshostvol on a Solaris host that does not have SDD

```
dsccli> lshostvol
Disk Name Volume Id      Vpath Name
=====
c6t1d0    IBM-2107.7520781/4200 ---
c6t1d1    IBM-2107.7520781/4201 ---
c7t2d0    IBM-2107.7520781/4200 ---
c7t2d1    IBM-2107.7520781/4201 ---
```

Windows: Mapping disks when not using SDD or using SDDDSM

In Example 14-40, we run **lshostvol** on a Windows host that does not use SDD or uses SDDDSM. The disks are listed by Windows Disk number. If you want to know which disk is associated with which drive letter, you need to look at the Windows Disk manager.

Example 14-40 lshostvol on a Windows host that does not use SDD or uses SDDDSM

```
dsccli> lshostvol
Disk Name Volume Id      Vpath Name
=====
Disk2     IBM.2107-7520781/4702 ---
Disk3     IBM.2107-75ABTV1/4702 ---
Disk4     IBM.2107-7520781/1710 ---
Disk5     IBM.2107-75ABTV1/1004 ---
Disk6     IBM.2107-75ABTV1/1009 ---
Disk7     IBM.2107-75ABTV1/100A ---
Disk8     IBM.2107-7503461/4702 ---
```

Windows: Mapping disks when using SDD

In Example 14-41, we run `lshostvol` on a Windows host that uses SDD. The disks are listed by Windows Disk number. If you want to know which disk is associated with which drive letter, you need to look at the Windows Disk manager.

Example 14-41 lshostvol on a Windows host that does not use SDD

```
dsccli> lshostvol
Disk Name   Volume Id           Vpath Name
=====
Disk2,Disk2 IBM.2107-7503461/4703 Disk2
Disk3,Disk3 IBM.2107-7520781/4703 Disk3
Disk4,Disk4 IBM.2107-75ABTV1/4703 Disk4
```

14.5 Configuring DS8000 Storage for Count Key Data Volumes

To configure the DS8000 storage for count key data (CKD) volumes, you follow almost exactly the same steps as for fixed block (FB) volumes.

Note that there is one additional step, which is to create Logical Control Units (LCUs), as displayed in the following list.

1. Create arrays.
2. Create CKD ranks.
3. Create CKD Extent Pools.
4. Optionally, create repositories for Track Space Efficient volumes.
5. Create LCUs.
6. Create CKD volumes.

You do not have to create volume groups or host connects for CKD volumes. If there are I/O ports in Fibre Channel connection (FICON) mode, access to CKD volumes by FICON hosts is granted automatically.

14.5.1 Create arrays

Array creation for CKD is exactly the same as for fixed block (FB). See 14.4.1, “Create arrays” on page 373.

14.5.2 Ranks and Extent Pool creation

When creating ranks and Extent Pools, you need to specify `-stgtype ckd`, as shown in Example 14-42.

Example 14-42 Rank and Extent Pool creation for CKD

```
dsccli> mkrank -array A0 -stgtype ckd
CMUC00007I mkrank: Rank R0 successfully created.
dsccli> lsrank
ID Group State      datastate Array RAIDtype extpoolID stgtype
=====
R0      - Unassigned Normal    A0          6 -          ckd
dsccli> mkextpool -rankgrp 0 -stgtype ckd CKD_High_0
CMUC00000I mkextpool: Extent Pool P0 successfully created.
dsccli> chrnk -extpool P2 R0
CMUC00008I chrnk: Rank R0 successfully modified.
dsccli> lsxtpool
```

Name	ID	stgtype	rankgrp	status	availstor (2^30B)	%allocated	available	reserved	numvol
CKD_High_0	2	ckd	0	below	252	0	287	0	0

Create a Space Efficient repository for CKD Extent Pools

If the DS8000 has the IBM FlashCopy SE feature, you can create Track Space Efficient (TSE) volumes that can be used as FlashCopy targets. Before you can create TSE volumes, you must create a Space Efficient repository in the Extent Pool. The repository provides space to store the data associated with TSE logical volumes. Only one repository is allowed per Extent Pool. A repository has a physical capacity that is available for storage allocations by TSE volumes and a virtual capacity that is the sum of LUN/volume sizes of all Space Efficient volumes. The physical repository capacity is allocated when the repository is created. If there are several ranks in the Extent Pool, the repository's extents are striped across the ranks (Storage Pool Striping).

Space Efficient repository creation for CKD Extent Pools is identical to that of FB Extent Pools, with the exception that the size of the repository's real capacity and virtual capacity are expressed either in cylinders or as multiples of 3390 model 1 disks (the default for CKD Extent Pools), instead of in GB or blocks, which apply to FB Extent Pools only.

Example 14-43 shows the creation of a repository.

Example 14-43 Creating a Space Efficient repository for CKD volumes

```
dscli> mksestg -repcap 100 -viricap 200 -extpool p1
CMUC00342I mksestg: The space-efficient storage for the Extent Pool P1 has been
created successfully.
```

You can obtain information about the repository with the **showsestg** command. Example 14-44 shows the output of the **showsestg** command. You might particularly be interested in how much capacity is used in the repository; to obtain this information, check the **repcapalloc** value.

Example 14-44 Getting information about a Space Efficient CKD repository

```
dscli> showsestg p1
extpool          P1
stgtype          ckd
datastate        Normal
configstate       Normal
repcapstatus     below
%repcapthreshold 0
repcap(GiB)      88.1
repcap(Mod1)     100.0
repcap(blocks)   -
repcap(cyl)      111300
repcapalloc(GiB/Mod1) 0.0
%repcapalloc     0
viricap(GiB)     176.2
viricap(Mod1)    200.0
viricap(blocks)  -
viricap(cyl)     222600
viricapalloc(GiB/Mod1) 0.0
%viricapalloc    0
overhead(GiB/Mod1) 4.0
reqrepcap(GiB/Mod1) 100.0
reqviricap(GiB/Mod1) 200.0
```

Note that storage is allocated for the repository in addition to repcap size. In Example 14-44 on page 389, the line that starts with overhead indicates that 4 GB had been allocated in addition to the repcap size.

A repository can be deleted by using the **rmsestg** command.

Important: In the current implementation, it is not possible to expand a repository. The physical size or the virtual size of the repository cannot be changed. Therefore, careful planning is required. If you have to expand a repository, you must delete all TSE volumes and the repository itself and then create a new repository.

14.5.3 Logical control unit creation

When creating volumes for a CKD environment, you must create logical control units (LCUs) before creating the volumes. In Example 14-45, you can see what happens if you try to create a CKD volume without creating an LCU first.

Example 14-45 Trying to create CKD volumes without an LCU

```
dscli> mkckdvol -extpool p2 -cap 262668 -name ITS0_EAV1_#h C200
CMUN02282E mkckdvol: C200: Unable to create CKD logical volume: CKD volumes require a CKD
logical subsystem.
```

We must use the **mk1cu** command first. The format of the command is:

```
mk1cu -qty XX -id XX -ssXX
```

To display the LCUs that we have created, we use the **ls1cu** command.

In Example 14-46, we create two LCUs using the **mk1cu** command, and then list the created LCUs using the **ls1cu** command. Note that by default the LCUs that were created are 3990-6.

Example 14-46 Creating a logical control unit with mk1cu

```
dscli> mk1cu -qty 2 -id BC -ss BC00
CMUC00017I mk1cu: LCU BC successfully created.
CMUC00017I mk1cu: LCU BD successfully created.
dscli> ls1cu
ID Group addrgrp confgvols subsys conbasetype
=====
BC      0 C              0 0xBC00 3990-6
BD      1 C              0 0xBC01 3990-6
```

Also note that because we created two LCUs (using the parameter **-qty 2**), the first LCU, which is ID BC (an even number), is in address group 0, which equates to rank group 0. The second LCU, which is ID BD (an odd number), is in address group 1, which equates to rank group 1. By placing the LCUs into both address groups, we maximize performance by spreading workload across both rank groups of the DS8000.

Tip: For the DS8000, the CKD LCUs can be ID 00 to ID FE. The LCUs fit into one of 16 address groups. Address group 0 is LCUs 00 to 0F, address group 1 is LCUs 10 to 1F, and so on. If you create a CKD LCU in an address group, then that address group cannot be used for FB volumes. Likewise, if there were, for example, FB volumes in LSS 40 to 4F (address group 4), then that address group cannot be used for CKD. Be aware of this limitation when planning the volume layout in a mixed FB/CKD DS8000.

14.5.4 Create CKD volumes

Having created an LCU, we can now create CKD volumes by using the **mkckdvol** command. The format of the **mkckdvol** command is:

```
mkckdvol -extpool P2 -cap 262668 -datatype 3390-A -eam rotatevols -name  
ITSO_EAV1_#h BC06
```

The major difference to note here is that the capacity is expressed in either cylinders or as CKD extents (1,113 cylinders). to not waste space, use volume capacities that are a multiple of 1,113 cylinders. Also new is the support of DS8000 Licensed Machine Code 5.4.xx.xx for Extended Address Volumes (EAV). This support expands the maximum size of a CKD volume to 262,668 cylinders and creates a new device type, 3390 Model A. This new volume can only be used by IBM z/OS systems running V1.10 or later versions.

Tip: For 3390-A volumes, the size can be specified from 1 to 65,520 in increments of 1 and from 65,667 (next multiple of 1113) to 262,668 in increments of 1113.

In Example 14-47, we create a single 3390-A volume using 262,668 cylinders.

Example 14-47 Creating CKD volumes using mkckdvol

```
dscli> mkckdvol -extpool P2 -cap 262668 -datatype 3390-A -eam rotatevols -name ITSO_EAV1_#h BC06  
CMUC00021I mkckdvol: CKD Volume BC06 successfully created.
```

```
dscli> lsckdvol
```

Name	ID	accstate	datastate	configstate	deviceMTM	voltype	orgbvols	extpool	cap (cyl)
ITSO_BC00	BC00	Online	Normal	Normal	3390-9	CKD Base -	P2		10017
ITSO_BC01	BC01	Online	Normal	Normal	3390-9	CKD Base -	P2		10017
ITSO_BC02	BC02	Online	Normal	Normal	3390-9	CKD Base -	P2		10017
ITSO_BC03	BC03	Online	Normal	Normal	3390-9	CKD Base -	P2		10017
ITSO_BC04	BC04	Online	Normal	Normal	3390-9	CKD Base -	P2		10017
ITSO_BC05	BC05	Online	Normal	Normal	3390-9	CKD Base -	P2		10017
ITSO_EAV1_BC06	BC06	Online	Normal	Normal	3390-A	CKD Base -	P2		262668
ITSO_BD00	BD00	Online	Normal	Normal	3390-9	CKD Base -	P3		10017
ITSO_BD01	BD01	Online	Normal	Normal	3390-9	CKD Base -	P3		10017
ITSO_BD02	BD02	Online	Normal	Normal	3390-9	CKD Base -	P3		10017
ITSO_BD03	BD03	Online	Normal	Normal	3390-9	CKD Base -	P3		10017
ITSO_BD04	BD04	Online	Normal	Normal	3390-9	CKD Base -	P3		10017
ITSO_BD05	BD05	Online	Normal	Normal	3390-9	CKD Base -	P3		10017

Remember, we can only create CKD volumes in LCUs that we have already created.

You also need to be aware that volumes in even numbered LCUs must be created from an Extent Pool that belongs to rank group 0. Volumes in odd numbered LCUs must be created from an Extent Pool in rank group 1.

Tip: With the DS8000 Release 6.1 microcode, you can configure a volume to belong to a certain Resource Groups using the **-resgrp <RG_ID>** flag in the **mkckdvol** command. For more details, refer to *IBM System Storage DS8000: Copy Services Resource Groups*, REDP-4758.

Storage pool striping

When creating a volume, you have a choice about how the volume is allocated in an Extent Pool with several ranks. The extents of a volume can be kept together in one rank (as long as there is enough free space on that rank). The next rank is used when the next volume is created. This allocation method is called *rotate volumes*.

You can also specify that you want the extents of the volume to be evenly distributed across all ranks within the Extent Pool. This allocation method is called *rotate extents*.

The extent allocation method is specified with the `-eam rotateexts` or `-eam rotatevols` option of the `mkckdvol` command (see Example 14-48).

Tip: The default allocation policy has changed to *rotate extents*.

Example 14-48 Creating a CKD volume with Extent Pool striping

```
dsccli> mkckdvol -extpool p4 -cap 10017 -name ITS0-CKD-STRP -eam rotateexts 0080
CMUC00021I mkckdvol: CKD Volume 0080 successfully created.
```

The `showckdvol` command with the `-rank` option (see Example 14-49) shows that the volume we created is distributed across two ranks, and it also displays how many extents on each rank were allocated for this volume.

Example 14-49 Getting information about a striped CKD volume

```
dsccli> showckdvol -rank 0080
Name          ITS0-CKD-STRP
ID            0080
accstate      Online
datastate     Normal
configstate   Normal
deviceMTM     3390-9
volser        -
datatype      3390
voltype       CKD Base
orgbvols      -
addrgrp       0
extpool       P4
exts          9
cap (cyl)     10017
cap (10^9B)   8.5
cap (2^30B)   7.9
ranks         2
sam           Standard
repcapalloc   -
eam           rotateexts
reqcap (cyl)  10017
=====Rank extents=====
rank extents
=====
R4            4
R30           5
```

Track Space Efficient volumes

When your DS8000 has the IBM FlashCopy SE feature, you can create Track Space Efficient (TSE) volumes to be used as FlashCopy target volumes. A repository must exist in the Extent

Pool where you plan to allocate TSE volumes (see “Create a Space Efficient repository for CKD Extent Pools” on page 389).

A Track Space Efficient volume is created by specifying the `-sam tse` parameter with the `mkckdvol` command (see Example 14-50).

Example 14-50 Creating a Space Efficient CKD volume

```
dscli> mkckdvol -extpool p4 -cap 10017 -name ITS0-CKD-SE -sam tse 0081
CMUC00021I mkckdvol: CKD Volume 0081 successfully created.
```

When listing Space Efficient repositories with the `lssestg` command (see Example 14-51), we can see that in Extent Pool P4 we have a virtual allocation of 7.9 GB, but that the allocated (used) capacity `repcapalloc` is still zero.

Example 14-51 Obtaining information about Space Efficient CKD repositories

```
dscli> lssestg -l
extentpoolID stgtype datastate configstate repcapstatus %repcapthreshold repcap (2^30B) vircap repcapalloc vircapalloc
=====
P4           ckd      Normal    Normal    below           0           100.0  200.0      0.0      7.9
```

This allocation comes from the volume just created. To see the allocated space in the repository for just this volume, we can use the `showckdvol` command (see Example 14-52).

Example 14-52 Checking the repository usage for a CKD volume

```
dscli> showckdvol 0081
Name          ITS0-CKD-SE
ID            0081
accstate      Online
datastate     Normal
configstate   Normal
deviceMTM     3390-9
volser        -
datatype      3390
voltype       CKD Base
orgbvols      -
addrgrp       0
extpool       P4
exts          9
cap (cyl)     10017
cap (10^9B)   8.5
cap (2^30B)   7.9
ranks         0
sam           TSE
repcapalloc   0
eam           -
reqcap (cyl)  10017
```

Dynamic Volume Expansion

A volume can be expanded without having to remove the data within the volume. You can specify a new capacity by using the `chckdvol` command (Example 14-53 on page 394). The new capacity must be larger than the previous one; you *cannot* shrink the volume.

Example 14-53 Expanding a striped CKD volume

```
dscli> chckdvol -cap 30051 0080
CMUC00332W chckdvol: Some host operating systems do not support changing the
volume size. Are you sure that you want to resize the volume? [y/n]: y
CMUC00022I chckdvol: CKD Volume 0080 successfully modified.
```

Because the original volume had the rotateexts attribute, the additional extents are also striped (see Example 14-54).

Example 14-54 Checking the status of an expanded CKD volume

```
dscli> showckdvol -rank 0080
Name          ITS0-CKD-STRP
ID            0080
accstate      Online
datastate     Normal
configstate   Normal
deviceMTM     3390-9
volser        -
datatype      3390
voltype       CKD Base
orgbvols      -
addrgrp       0
extpool       P4
exts          27
cap (cyl)     30051
cap (10^9B)   25.5
cap (2^30B)   23.8
ranks         2
sam           Standard
repcapalloc   -
eam           rotateexts
reqcap (cyl)  30051
=====Rank extents=====
rank extents
=====
R4            13
R30           14
```

Tip: Before you can expand a volume, you first have to delete all Copy Services relationships for that volume, and you cannot specify both -cap and -datatype for the **chckdvol** command.

It is possible to expand a 3390 Model 9 volume to a 3390 Model A. You can do that just by specifying a new capacity for an existing Model 9 volume. When you increase the size of a 3390-9 volume beyond 65,520 cylinders, its device type automatically changes to 3390-A. However, keep in mind that a 3390 Model A can only be used in z/OS V1.10 and later (Example 14-55).

Example 14-55 Expanding a 3390 to a 3390-A

*** Command to show CKD volume definition before expansion:

```
dscli> showckdvol BC07
Name          ITS0_EAV2_BC07
```

```

ID          BC07
accstate    Online
datastate   Normal
configstate Normal
deviceMTM   3390-9
volser      -
datatype    3390
voltype     CKD Base
orgbvols    -
addrgrp     B
extpool     P2
exts        9
cap (cyl)   10017
cap (10^9B) 8.5
cap (2^30B) 7.9
ranks       1
sam         Standard
repcapalloc -
eam         rotatevols
reqcap (cyl) 10017

```

*** Command to expand CKD volume from 3390-9 to 3390-A:

```

dscli> chckdvol -cap 262668 BC07
CMUC00332W chckdvol: Some host operating systems do not support changing the volume size.
Are you sure that you want to resize the volume? [y/n]: y
CMUC00022I chckdvol: CKD Volume BC07 successfully modified.

```

*** Command to show CKD volume definition after expansion:

```

dscli> showckdvol BC07
Name          ITS0_EAV2_BC07
ID            BC07
accstate      Online
datastate     Normal
configstate   Normal
deviceMTM     3390-A
volser        -
datatype      3390-A
voltype       CKD Base
orgbvols      -
addrgrp       B
extpool       P2
exts          236
cap (cyl)     262668
cap (10^9B)   223.3
cap (2^30B)   207.9
ranks         1
sam           Standard
repcapalloc   -
eam           rotatevols
reqcap (cyl)  262668

```

You cannot reduce the size of a volume. If you try, an error message is displayed, as shown in Example 14-56.

Example 14-56 Reducing a volume size

```

dscli> chckdvol -cap 10017 BC07

```

```
CMUC00332W chckdvol: Some host operating systems do not support changing the volume size.
Are you sure that you want to resize the volume? [y/n]: y
CMUN02541E chckdvol: BC07: The expand logical volume task was not initiated because the
logical volume capacity that you have requested is less than the current logical volume
capacity.
```

Deleting volumes

CKD volumes can be deleted by using the **rmckdvol** command. FB volumes can be deleted by using the **rmfbvol** command.

Starting with Licensed Machine Code (LMC) level 6.5.1.xx, the command includes a new capability to prevent the accidental deletion of volumes that are in use. A CKD volume is considered to be in use if it is participating in a Copy Services relationship, or if the IBM System z path mask indicates that the volume is in a “grouped state” or online to any host system. A CKD volume is considered to be in use if it is participating in a Copy Services relationship, or if the volume has had any I/O in the last five minutes.

If the **-force** parameter is not specified with the command, volumes that are in use are not deleted. If multiple volumes are specified and some are in use and some are not, the ones not in use will be deleted. If the **-force** parameter is specified on the command, the volumes will be deleted without checking to see whether or not they are in use.

In Example 14-57, we try to delete two volumes, 0900 and 0901. Volume 0900 is online to a host, whereas 0901 is not online to any host and not in a Copy Services relationship. The **rmckdvol 0900-0901** command deletes just volume 0901, which is offline. To delete volume 0900, we use the **-force** parameter.

Example 14-57 Deleting CKD volumes

```
dscli> lsckdvol 0900-0901
Name  ID   accstate  datastate  configstate  deviceMTM  voltype  orgbvols  extpool  cap (cyl)
=====
ITS0_J 0900 Online    Normal     Normal      3390-9     CKD Base -      P1        10017
ITS0_J 0901 Online    Normal     Normal      3390-9     CKD Base -      P1        10017

dscli> rmckdvol -quiet 0900-0901
CMUN02948E rmckdvol: 0900: The Delete logical volume task cannot be initiated because the
Allow Host Pre-check Control Switch is set to true and the volume that you have specified
is online to a host.
CMUC00024I rmckdvol: CKD volume 0901 successfully deleted.

dscli> lsckdvol 0900-0901
Name  ID   accstate  datastate  configstate  deviceMTM  voltype  orgbvols  extpool  cap (cyl)
=====
ITS0_J 0900 Online    Normal     Normal      3390-9     CKD Base -      P1        10017

dscli> rmckdvol -force 0900
CMUC00023W rmckdvol: Are you sure you want to delete CKD volume 0900? [y/n]: y
CMUC00024I rmckdvol: CKD volume 0900 successfully deleted.

dscli> lsckdvol 0900-0901
CMUC00234I lsckdvol: No CKD Volume found.
```

14.5.5 Resource Groups

The Resource Group (RG) feature is designed for multi-tenancy environments. The resources are volumes, LCUs, and LSSs. They are used for access control for Copy Services functions only.

For more information about RG, refer to *IBM System Storage DS8000 Resource Groups*, REDP-4758.

14.5.6 Performance I/O Priority Manager

Performance I/O Priority Manager allows you to control Quality of Service. There are 16 policies for open systems, PG0-PG15.

For more information, refer to *IBM System Storage DS8000 Performance I/O Priority Manager*, REDP-4760.

Maintenance and upgrades

The topics covered in this part include:

- ▶ Licensed machine code
- ▶ Monitoring with Simple Network Management Protocol
- ▶ Remote support
- ▶ Capacity upgrades and CoD



Licensed machine code

In this chapter, we discuss considerations related to the planning and installation of new licensed machine code (LMC) bundles on the IBM System Storage DS8000 series. We cover the following topics in this chapter:

- ▶ How new microcode is released
- ▶ Bundle installation
- ▶ Concurrent and non-concurrent updates
- ▶ Code updates
- ▶ Host adapter firmware updates
- ▶ Loading the code bundle
- ▶ Post-installation activities
- ▶ Summary

15.1 How new microcode is released

The various components of the DS8000 system use firmware that can be updated as new releases become available. These components include device adapters (DA), host adapters (HA), power supplies, and Fibre Channel interface cards (FCIC). In addition, the microcode and internal operating system that run on the HMCs and each central electronics complex can be updated. As IBM continues to develop the DS8000, new functional features will also be released through new licensed machine code (LMC) levels.

When IBM releases new microcode for the DS8000, it is released in the form of a bundle. The term *bundle* is used because a new code release can include updates for various DS8000 components. These updates are tested together, and then the various code packages are bundled together into one unified release. In general, when referring to what code level is being used on a DS8000, the term bundle should be used. Components within the bundle will each have their own revision levels.

For a DS8000 cross-reference table of code bundles, visit the following site:

<http://www.ibm.com/>

- ▶ Select **Products** → **Storage** → **Disk systems**
- ▶ Select **Products**.
- ▶ Select **Enterprise** → **DS8000 series**
- ▶ Select **Product Detail** folder → **DS8X00 Support**
- ▶ Select and click **Documentation** under 'Choose your page'
- ▶ Click → **DS8000 Code Bundle Information** under 'Product documentation'

The Cross-Reference Table shows the levels of code for Release 6.1, which is installed on the DS8800. It should be updated as new bundles are released. It is important to always match your DS CLI version to the bundle installed on your DS8000.

For the DS8000, the naming convention of bundles is PR.MM.FFF.E, where:

P	Product (8 = DS8800, 7= DS8700)
R	Release Major (X)
MM	Release Minor (xx)
FFF	Fix Level (xxx)
E	EFIX level (0 is base, and 1.n is the interim fix build above base level.)

The naming convention is shown in Example 15-1.

Example 15-1 BUNDLE level information

For BUNDLE 86.10.113.0 :	
Product	DS8800
Release Major	6
Release Minor	10
Fix Level	113
EFIX level	0

If using DSCLI, you can obtain the CLI and LMC code level information using the **ver** command. The **ver** command with the following parameters displays the versions of the command-line interface, Storage Manager, and licensed machine code:

- **-s** (Optional) The -s parameter displays the version of the command line interface program. You cannot use the -s and -l parameters together

- **-l** (Optional) The **-l** parameter displays the versions of the command line interface, Storage Manager, and licensed machine code. You cannot use the **-l** and **-s** parameters together. See Example 15-2.
- **-cli** (Optional) Displays the version of the command line interface program. Version numbers are in the format version.release.modification.fixlevel.
- **-stgmgr** (Optional) Displays the version of the Storage Manager.
 - This is not the GUI (Storage Manager GUI). This id is related to HMC (Hardware Master Console code bundle information.) This has no value for users.
- **-lmc** (Optional) Displays the version of the licensed machine code (LMC).

Example 15-2 DSCLI version command

```

dscli> ver -l
Date/Time: April 25, 2011 03:29:23 PM MST IBM DSCLI Version: 7.6.10.464 DS: -
DSCLI          7.6.10.464
StorageManager 7.7.2.0.20110304.1
=====Version=====
Storage Image   LMC
=====
IBM.2107-75LX521 7.6.10.434
dscli>

```

15.2 Bundle installation

Important: Licensed Machine Code is always provided and installed by IBM Service Engineers. Installing a new Licensed Machine Code is *not* a client-serviceable task.

The Bundle package contains all the new levels of code that will update:

- HMC Code Levels
 - HMC OS/Managed System Base
 - DS Storage Manager
 - MM Extension
- Managed System Code Levels
- PTF Code Levels
- Storage Facility Image Code Levels
- Host Attachment Code Levels
- Device Adapter Code Level
- IO Enclosure Code Level
- Power Code Levels
- Fibre Channel Interface Card Code Levels
- Storage Enclosure Power Supply Unit Code Levels
- DDM Firmware Code Level

It is likely that a new Bundle will include updates for the following components:

- ▶ Linux OS for the HMC
- ▶ AIX OS for the central electrical complexes
- ▶ Microcode for HMC and central electrical complexes
- ▶ Microcode/Firmware for Host Adapters

Rarely, a new bundle will include updates for the following components:

- ▶ Firmware for Power subsystem (PPS, RPC, and BBU)
- ▶ Firmware for Storage DDMs
- ▶ Firmware for Fibre Channel interface cards
- ▶ Firmware for Device Adapters
- ▶ Firmware for Hypervisor on central electrical complexes

Code Distribution and Activation (CDA) Preload is a new way of performing Concurrent Code Load distribution. CDA Preload allows the user to perform every non-impacting Concurrent Code Load step for a code load just by inserting the DVD in to the primary HMC drive or running a network acquire using FTP or SFTP (Secure File Transfer) of the desired code level. After the CDA preload is started the following steps will be performed automatically:

- Download of the release bundle
- Prepare the HMC with any code update-specific fixes
- Distribute the code updates to the LPAR and installs them to an alternate BOS
- Perform scheduled precheck scans until the distributed code is ready to be activated by the user for up to 11 days.

Any time after completing the preload, when the user logs in to the primary HMC, they will be guided automatically to correct any serviceable events that might be open, update the HMC, and finally activate the previously distributed code on the storage facility.

The installation process involves several stages:

1. Update the HMC code.
2. Perform updates to the central electrical complex operating system (currently AIX V6.1), plus updates to the internal LMC, performed one at a time. The updates cause each central electrical complex to fail over its logical subsystems to the alternate central electrical complex. This process also updates the firmware running in each device adapter owned by that central electrical complex.
3. Perform updates to the host adapters. For DS8000 host adapters, the impact of these updates on each adapter is less than 2.5 seconds and should not affect connectivity. If an update were to take longer than this, the multipathing software on the host, or Control Unit-Initiated Reconfiguration (CUIR), would direct I/O to another host adapter. If a host is attached with only a single path, connectivity would be lost. See 4.4.2, “Host connections” on page 84 for more information about host attachments.
4. Occasionally, new Primary Power Supply (PPS) and Rack Power Control (RPC) firmware is released. New firmware can be loaded into each RPC card and PPS directly from the HMC. Each RPC and PPS is quiesced, updated, and resumed one at a time until all of them have been updated. There are no service interruptions for power updates.
5. Occasionally, new firmware for the Hypervisor, service processor, system planar, and I/O enclosure planars is released. This firmware can be loaded into each device directly from the HMC. Activation of this firmware might require a shutdown and reboot of each central electrical complex, one at a time. This would cause each central electrical complex to fail over its logical subsystems to the alternate central electrical complex. Certain updates do not require this step, or it might occur without processor reboots. See 4.3, “Central electrical complex failover and fallback” on page 79 for more information.
6. Very occasionally, new DDM firmware is released.

Although the installation process described above might seem complex, it does not require a great deal of user intervention. The IBM Service Representative normally simply starts the Code Distribution and Activation (CDA) process and then monitors its progress using the HMC.

Important: An upgrade of the DS8000 microcode might require that you upgrade the DS CLI on workstations. Check with your IBM Service Representative regarding the description and contents of the release bundle.

15.3 Concurrent and non-concurrent updates

The DS8000 allows for concurrent microcode updates. This means that code updates can be installed with all attached hosts up and running with no interruption to your business applications. It is also possible to install microcode update bundles non-concurrently, with all attached hosts shut down. However, this should not be necessary. This method is usually only employed at DS8000 installation time.

15.4 Code updates

The microcode that runs on the HMC normally gets updated as part of a new code bundle. The HMC can hold up to six versions of code. Each central electrical complex can hold three versions of code (the previous version, the active version, and the next version). Most organizations should plan for two code updates per year.

Best practice: Many clients with multiple DS8000 systems follow the updating schedule detailed here, wherein the HMC is updated 1 to 2 days before the rest of the bundle is applied.

Prior to the update of the central electrical complex operating system and microcode, a pre-verification test is run to ensure that no conditions exist that need to be corrected. The HMC code update will install the latest version of the pre-verification test. Then the newest test can be run. If problems are detected, there are one to two days before the scheduled code installation window to correct them. An example of this procedure is illustrated here:

Thursday	Copy or download the new code bundle to the HMCs. Update the HMC(s) to the new code bundle. Run the updated preverification test. Resolve any issues raised by the preverification test.
Saturday	Update the SFI.

Note that the actual time required for the concurrent code load varies based on the bundle that you are currently running and the bundle to which you are updating. Always consult with your IBM Service Representative regarding proposed code load schedules.

Additionally, check multipathing drivers and SAN switch firmware levels for current levels at regular intervals.

15.5 Host adapter firmware updates

One of the final steps in the concurrent code load process is updating the host adapters. Normally, every code Bundle contains new host adapter code. For DS8000 Fibre Channel cards, regardless of whether they are used for open systems attachment or System z (FICON) attachment, the update process is concurrent to the attached hosts. The Fibre Channel cards use a technique known as *adapter fast-load*. This allows them to switch to the new firmware in less than two seconds. This fast update means that single path hosts, hosts

that boot from SAN, and hosts that do not have multipathing software do not need to be shut down during the update. They can keep operating during the host adapter update because the update is so fast. This also means that no SDD path management should be necessary.

Remote Mirror and Copy path considerations

For Remote Mirror and Copy paths that use Fibre Channel ports, there are no special considerations. The ability to perform a fast-load means that no interruption occurs to the Remote Mirror operations.

Control Unit-Initiated Reconfiguration

Control Unit-Initiated Reconfiguration (CUIR) prevents loss of access to volumes in System z environments due to incorrect or wrong path handling. This function automates channel path management in System z environments in support of selected DS8000 service actions. Control Unit-Initiated Reconfiguration is available for the DS8000 when operated in the z/OS and z/VM environments. The CUIR function automates channel path vary on and vary off actions to minimize manual operator intervention during selected DS8000 service actions.

CUIR allows the DS8000 to request that all attached system images set all paths required for a particular service action to the offline state. System images with the appropriate level of software support respond to these requests by varying off the affected paths, and either notifying the DS8000 subsystem that the paths are offline, or that it cannot take the paths offline. CUIR reduces manual operator intervention and the possibility of human error during maintenance actions, at the same time reducing the time required for the maintenance window. This is particularly useful in environments where there are many systems attached to a DS8800.

15.6 Loading the code bundle

The DS8000 code bundle installation is performed by the IBM Service Representative. Contact your IBM Service Representative to discuss and arrange the required services.

15.7 Post-installation activities

After a new code bundle has been installed, you might need to perform the following tasks:

1. Upgrade the DS CLI of external workstations. For the majority of new release code bundles, there is a corresponding new release of DS CLI. Make sure you upgrade to the new version of DS CLI to take advantage of any improvements IBM has made.
2. Verify the connectivity from each DS CLI workstation to the DS8000.
3. Verify the DS Storage Manager connectivity from the TPC BE or SSPC to the DS8000.
4. Verify the connectivity from any stand-alone TPC Element Manager to the DS8000.
5. Verify the connectivity from the DS8000 to all TKLM Key Servers in use.

15.8 Summary

IBM might release changes to the DS8000 series Licensed Machine Code. These changes might include code fixes and feature updates relevant to the DS8000.

These updates and the information regarding them are detailed on the DS8000 Code Cross-Reference website as previously mentioned.



Monitoring with Simple Network Management Protocol

This chapter provides information about the Simple Network Management Protocol (SNMP) notifications and messages for the IBM System Storage DS8000 series. This chapter covers the following topics:

- ▶ Simple Network Management Protocol overview
- ▶ SNMP notifications
- ▶ SNMP configuration with the HMC
- ▶ SNMP configuration with the DSCLI

16.1 Simple Network Management Protocol overview

SNMP (Simple Network Management Protocol) is an application layer network protocol that allows communication between SNMP managers and SNMP agents using TCP/IP for a transport layer. In this application, the SNMP manager is typically an application program (e.g. NetView®) running on a server in the customer installation and the SNMP agents reside on various network attached units in the customer environment.

SNMP has become a standard for monitoring an IT environment. With SNMP, a system can be monitored, and event management, based on SNMP traps, can be automated.

SNMP is an industry-standard set of functions for monitoring and managing TCP/IP-based networks. SNMP includes a protocol, a database specification, and a set of data objects. A set of data objects forms a Management Information Base (MIB).

Objects contained in the MIB are typically related to management of the network attached units. The objects can be product unique, and can be used to sense information about the product or to control operation of the product. Typically, the SNMP manager provides mechanisms to implement automation code that can react to information communicated through the SNMP interface to provide an appropriate response to certain situations described by such communication.

SNMP provides a standard MIB that includes information such as IP addresses and the number of active TCP connections. The actual MIB definitions are encoded into the agents running on a system.

MIB-2 is the Internet standard MIB that defines over 100 TCP/IP specific objects, including configuration and statistical information, such as:

- ▶ Information about interfaces
- ▶ Address translation
- ▶ IP, Internet-control message protocol (ICMP), TCP, and User Datagram Protocol (UDP)

SNMP can be extended through the use of the SNMP Multiplexing protocol (SMUX protocol) to include enterprise-specific MIBs that contain information related to a specific environment or application. A management agent (a SMUX peer daemon) retrieves and maintains information about the objects defined in its MIB and passes this information about to a specialized network monitor or network management station (NMS).

The SNMP protocol defines two terms, agent and manager, instead of the terms client and server, which are used in many other TCP/IP protocols.

16.1.1 SNMP agent

An *SNMP agent* is a daemon process that provides access to the MIB objects on IP hosts on which the agent is running. The agent can receive SNMP get or SNMP set requests from SNMP managers and can send SNMP trap requests to SNMP managers.

Agents send traps to the SNMP manager to indicate that a particular condition exists on the agent system, such as the occurrence of an error. In addition, the SNMP manager generates traps when it detects status changes or other unusual conditions while polling network objects.

16.1.2 SNMP manager

An *SNMP manager* can be implemented in two ways. An SNMP manager can be implemented as a simple command tool that can collect information from SNMP agents. An SNMP manager also can be composed of multiple daemon processes and database applications. This type of complex SNMP manager provides you with monitoring functions using SNMP. It typically has a graphical user interface for operators. The SNMP manager gathers information from SNMP agents and accepts trap requests sent by SNMP agents.

16.1.3 SNMP trap

A *trap* is a message sent from an SNMP agent to an SNMP manager without a specific request from the SNMP manager.

SNMP defines six generic types of traps and allows definition of enterprise-specific traps. The trap structure conveys the following information to the SNMP manager:

- ▶ Agent's object that was affected
- ▶ IP address of the agent that sent the trap
- ▶ Event description (either a generic trap or enterprise-specific trap, including trap number)
- ▶ Time stamp
- ▶ Optional enterprise-specific trap identification
- ▶ List of variables describing the trap

16.1.4 SNMP communication

The SNMP manager sends SNMP get, get-next, or set requests to SNMP agents, which listen on UDP port 161. The agents send back a reply to the manager. The SNMP agent can be implemented on any kind of IP host, such as UNIX workstations, routers, and network appliances.

You can gather various information about the specific IP hosts by sending the SNMP get and get-next requests, and can update the configuration of IP hosts by sending the SNMP set request.

The SNMP agent can send SNMP trap requests to SNMP managers, which listen on UDP port 162. The SNMP trap1 requests sent from SNMP agents can be used to send warning, alert, or error notification messages to SNMP managers.

You can configure an SNMP agent to send SNMP trap requests to multiple SNMP managers. Figure 16-1 on page 410 illustrates the characteristics of SNMP architecture and communication.

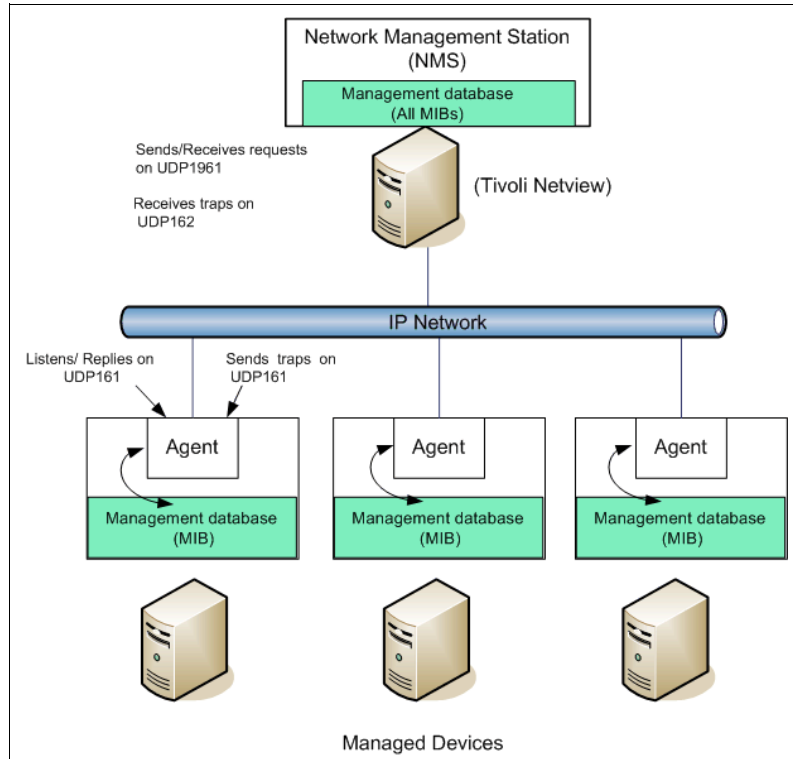


Figure 16-1 SNMP architecture and communication

16.1.5 SNMP Requirements

All SNMP implementations should meet the following functional requirements defined by this section:

1. SNMP trap generation should be operative whenever events that the traps indicate can occur. This should be true independently of the functionality of any other ESSNet GUI or API provided by the product.
2. Any changes to the MIB associated with a given trap should be released concurrently with the supported trap.
3. Certain controls for SNMP traps might be provided. If these controls cannot be altered through the SNMP interface, they should be alterable through the Integrated Configuration Assistant Tool Graphical User Interface or service interface. Their state is reflected in the MIB.
4. Consistency group traps (200 and 201) must be prioritized above all other traps and must be surfaced in less than two seconds from the real time incident.

16.1.6 Generic SNMP security

The SNMP protocol uses the community name for authorization. Most SNMP implementations use the default community name *public* for a read-only community and *private* for a read-write community. In most cases, a community name is sent in a plain-text format between the SNMP agent and the manager. Certain SNMP implementations have additional security features, such as restrictions on the accessible IP addresses.

Therefore, you should be careful about the SNMP security. At the very least, do not allow access to hosts that are running the SNMP agent from networks or IP hosts that do not necessarily require access.

Tip: You might want to physically secure the network to which you send SNMP packets using a firewall because community strings are included as plain text in SNMP packets.

16.1.7 Message Information Base

The objects, which you can get or set by sending SNMP get or set requests, are defined as a set of databases called the *Message Information Base (MIB)*. The structure of MIB is defined as an Internet standard in RFC 1155; the MIB forms a tree structure.

Most hardware and software vendors provide you with extended MIB objects to support their own requirements. The SNMP standards allow this extension by using the private sub-tree, called *enterprise specific* MIB. Because each vendor has a unique MIB sub-tree under the private sub-tree, there is no conflict among vendors' original MIB extensions.

16.1.8 SNMP trap request

An SNMP agent can send SNMP trap requests to SNMP managers to inform them about the change of values or status on the IP host where the agent is running. There are seven predefined types of SNMP trap requests, as shown in Table 16-1.

Table 16-1 SNMP trap request types

Trap type	Value	Description
coldStart	0	Restart after a crash.
warmStart	1	Planned restart.
linkDown	2	Communication link is down.
linkUp	3	Communication link is up.
authenticationFailure	4	Invalid SNMP community string was used.
egpNeighborLoss	5	EGP neighbor is down.
enterpriseSpecific	6	Vendor-specific event happened.

A trap message contains pairs of an OID and a value shown in Table 16-1 to notify the cause of the trap message. You can also use type 6, the *enterpriseSpecific* trap type, when you have to send messages that do not fit other predefined trap types, for example, DISK I/O error and application down. You can also set an integer value field called *Specific Trap* on your trap message.

16.1.9 DS8000 SNMP configuration

SNMP for the DS8000 is designed in such a way that the DS8000 only sends traps in case of a notification. The traps can be sent to a defined IP address.

SNMP alert traps provide information about problems that the storage unit detects. Either you or the service provider must perform corrective action for the trap related problems.

The DS8000 does not have an SNMP agent installed that can respond to SNMP polling. The default Community Name parameter is set to *public*.

The management server that is configured to receive the SNMP traps receives all the generic trap 6 and specific trap 3 messages, which are sent in parallel with the Call Home to IBM.

Before configuring SNMP for the DS8000, you are required to get the destination address for the SNMP trap and also the port information on which the *Trap Daemon* listens.

Tip: The standard port for SNMP traps is port 162.

16.2 SNMP notifications

The HMC of the DS8000 sends an SNMPv1 trap in two cases:

- ▶ A serviceable event was reported to IBM using Call Home.
- ▶ An event occurred in the Copy Services configuration or processing.

A serviceable event is posted as a generic trap 6 specific trap 3 message. The specific trap 3 is the only event that is sent for serviceable events. For reporting Copy Services events, generic trap 6 and specific traps 100, 101, 102, 200, 202, 210, 211, 212, 213, 214, 215, 216, or 217 are sent.

16.2.1 Serviceable event using specific trap 3

In Example 16-1, we see the contents of generic trap 6 specific trap 3. The trap holds the information about the serial number of the DS8000, the event number that is associated with the manageable events from the HMC, the reporting Storage Facility Image (SFI), the system reference code (SRC), and the location code of the part that is logging the event.

The SNMP trap is sent in parallel with a Call Home for service to IBM.

Example 16-1 SNMP special trap 3 of an DS8000

```
Nov 14, 2005 5:10:54 PM CET
Manufacturer=IBM
ReportingMTMS=2107-922*7503460
ProbNm=345
LparName=null
FailingEnclosureMTMS=2107-922*7503460
SRC=10001510
EventText=2107 (DS 8000) Problem
Fru1Loc=U1300.001.1300885
Fru2Loc=U1300.001.1300885U1300.001.1300885-P1
```

For open events in the event log, a trap is sent every eight hours until the event is closed. Use the following link the discover explanations about all System Reference Codes (SRC):

<http://publib.boulder.ibm.com/infocenter/dsichelp/ds8000sv/index.jsp>

In this page, select **Messages and codes** → **List of system reference codes and firmware codes**.

16.2.2 Copy Services event traps

For state changes in a remote Copy Services environment, there are 13 traps implemented. The traps 1xx are sent for a state change of a physical link connection. The 2xx traps are sent for state changes in the logical Copy Services setup. For all of these events, no Call Home is generated and IBM is not notified.

This chapter describes only the messages and the circumstances when traps are sent by the DS8000. For detailed information about these functions and terms, refer to *IBM System Storage DS8000: Copy Services for IBM System z*, SG24-6787 and *IBM System Storage DS8000: Copy Services for Open Systems*, SG24-6788.

Physical connection events

Within the trap 1xx range, a state change of the physical links is reported. The trap is sent if the physical remote copy link is interrupted. The Link trap is sent from the primary system. The *PLink* and *SLink* columns are only used by the 2105 ESS disk unit.

If one or several links (but not all links) are interrupted, a trap 100, as shown in Example 16-2, is posted and indicates that the redundancy is degraded. The RC column in the trap represents the return code for the interruption of the link; return codes are listed in Table 16-2 on page 414.

Example 16-2 Trap 100: Remote Mirror and Copy links degraded

```
PPRC Links Degraded
UNIT: Mnf Type-Mod SerialNm LS
PRI:  IBM 2107-922 75-20781 12
SEC:  IBM 2107-9A2 75-ABTV1 24
Path: Type PP  PLink SP  SLink RC
1:    FIBRE 0143 XXXXXX 0010 XXXXXX 15
2:    FIBRE 0213 XXXXXX 0140 XXXXXX OK
```

If all links are interrupted, a trap 101, as shown in Example 16-3, is posted. This event indicates that no communication between the primary and the secondary system is possible.

Example 16-3 Trap 101: Remote Mirror and Copy links are inoperable

```
PPRC Links Down
UNIT: Mnf Type-Mod SerialNm LS
PRI:  IBM 2107-922 75-20781 10
SEC:  IBM 2107-9A2 75-ABTV1 20
Path: Type PP  PLink SP  SLink RC
1:    FIBRE 0143 XXXXXX 0010 XXXXXX 17
2:    FIBRE 0213 XXXXXX 0140 XXXXXX 17
```

After the DS8000 can communicate again using any of the links, trap 102, as shown in Example 16-4, is sent after one or more of the interrupted links are available again.

Example 16-4 Trap 102: Remote Mirror and Copy links are operational

```
PPRC Links Up
UNIT: Mnf Type-Mod SerialNm LS
PRI:  IBM 2107-9A2 75-ABTV1 21
SEC:  IBM 2107-000 75-20781 11
Path: Type PP  PLink SP  SLink RC
1:    FIBRE 0010 XXXXXX 0143 XXXXXX OK
2:    FIBRE 0140 XXXXXX 0213 XXXXXX OK
```

Table 16-2 on page 414 lists the Remote Mirror and Copy return codes.

Table 16-2 Remote Mirror and Copy return codes

Return code	Description
02	Initialization failed. ESCON link reject threshold exceeded when attempting to send ELP or RID frames.
03	Timeout. No reason available.
04	There are no resources available in the primary storage unit for establishing logical paths because the maximum number of logical paths have already been established.
05	There are no resources available in the secondary storage unit for establishing logical paths because the maximum number of logical paths have already been established.
06	There is a secondary storage unit sequence number, or logical subsystem number, mismatch.
07	There is a secondary LSS subsystem identifier (SSID) mismatch, or failure of the I/O that collects the secondary information for validation.
08	The ESCON link is offline. This is caused by the lack of light detection coming from a host, peer, or switch.
09	The establish failed. It is retried until the command succeeds or a remove paths command is run for the path. Tip: The attempt-to-establish state persists until the establish path operation succeeds or the remove remote mirror and copy paths command is run for the path.
0A	The primary storage unit port or link cannot be converted to channel mode if a logical path is already established on the port or link. The establish paths operation is not retried within the storage unit.
10	Configuration error. The source of the error is one of the following: ► The specification of the SA ID does not match the installed ESCON adapter cards in the primary controller. ► For ESCON paths, the secondary storage unit destination address is zero and an ESCON Director (switch) was found in the path. ► For ESCON paths, the secondary storage unit destination address is not zero and an ESCON director does not exist in the path. The path is a direct connection.
14	The Fibre Channel path link is down.
15	The maximum number of Fibre Channel path retry operations has been exceeded.
16	The Fibre Channel path secondary adapter is not Remote Mirror and Copy capable. This could be caused by one of the following conditions: ► The secondary adapter is not configured properly or does not have the current firmware installed. ► The secondary adapter is already a target of 32 logical subsystems (LSSs).

Trap 210, as shown in Example 16-7, is sent when a Consistency Group in a Global Mirror environment was successfully formed.

Example 16-7 Trap210: Global Mirror initial Consistency Group successfully formed

2005/11/14 15:30:55 CET
Asynchronous PPRC Initial Consistency Group Successfully Formed
UNIT: Mnf Type-Mod SerialNm
 IBM 2107-922 75-20781
Session ID: 4002

Trap 211, as shown in Example 16-8, is sent if the Global Mirror setup got into an severe error state, where no attempts are made to form a Consistency Group.

Example 16-8 Trap 211: Global Mirror Session is in a fatal state

Asynchronous PPRC Session is in a Fatal State
UNIT: Mnf Type-Mod SerialNm
 IBM 2107-922 75-20781
Session ID: 4002

Trap 212, shown in Example 16-9, is sent when a Consistency Group cannot be created in a Global Mirror relationship. The reasons might be:

- ▶ Volumes have been taken out of a copy session.
- ▶ The Remote Copy link bandwidth might not be sufficient.
- ▶ The FC link between the primary and secondary system is not available.

Example 16-9 Trap 212: Global Mirror Consistency Group failure - Retry will be attempted

Asynchronous PPRC Consistency Group Failure - Retry will be attempted
UNIT: Mnf Type-Mod SerialNm
 IBM 2107-922 75-20781
Session ID: 4002

Trap 213, shown in Example 16-10, is sent when a Consistency Group in a Global Mirror environment can be formed after a previous Consistency Group formation failure.

Example 16-10 Trap 213: Global Mirror Consistency Group successful recovery

Asynchronous PPRC Consistency Group Successful Recovery
UNIT: Mnf Type-Mod SerialNm
 IBM 2107-9A2 75-ABTV1
Session ID: 4002

Trap 214, shown in Example 16-11, is sent if a Global Mirror Session is terminated using the DS CLI command **rmgmir** or the corresponding GUI function.

Example 16-11 Trap 214: Global Mirror Master terminated

2005/11/14 15:30:14 CET
Asynchronous PPRC Master Terminated
UNIT: Mnf Type-Mod SerialNm
IBM 2107-922 75-20781
Session ID: 4002

Trap 215, shown in Example 16-12, is sent if, in the Global Mirror Environment, the master detects a failure to complete the FlashCopy commit. The trap is sent after a number of commit retries have failed.

Example 16-12 Trap 215: Global Mirror FlashCopy at Remote Site unsuccessful

Asynchronous PPRC FlashCopy at Remote Site Unsuccessful
A UNIT: Mnf Type-Mod SerialNm
IBM 2107-9A2 75-ABTV1
Session ID: 4002

Trap 216, shown in Example 16-13, is sent if a Global Mirror *Master* cannot terminate the Global Copy relationship at one of his *subordinates*. This might occur if the master is terminated with **rmgmir** but the master cannot terminate the copy relationship on the subordinate. You might need to run a **rmgmir** against the subordinate to prevent any interference with other Global Mirror sessions.

Example 16-13 Trap 216: Global Mirror subordinate termination unsuccessful

Asynchronous PPRC Slave Termination Unsuccessful
UNIT: Mnf Type-Mod SerialNm
Master: IBM 2107-922 75-20781
Slave: IBM 2107-921 75-03641
Session ID: 4002

Trap 217, shown in Example 16-14, is sent if a Global Mirror environment was suspended by the DS CLI command **pausegmir** or the corresponding GUI function.

Example 16-14 Trap 217: Global Mirror paused

Asynchronous PPRC Paused
UNIT: Mnf Type-Mod SerialNm
IBM 2107-9A2 75-ABTV1
Session ID: 4002

Trap 218, shown in Example 16-15, is sent if a Global Mirror has exceeded the allowed threshold for failed consistency group formation attempts.

Example 16-15 Trap 218: Global Mirror number of consistency group failures exceed threshold

Global Mirror number of consistency group failures exceed threshold
UNIT: Mnf Type-Mod SerialNm
IBM 2107-9A2 75-ABTV1
Session ID: 4002

Trap 219, shown in Example 16-16, is sent if a Global Mirror has successfully formed a consistency group after one or more formation attempts had previously failed.

Example 16-16 Trap 219: Global Mirror first successful consistency group after prior failures

```
Global Mirror first successful consistency group after prior failures
UNIT: Mnf Type-Mod SerialNm
      IBM 2107-9A2 75-ABTV1
Session ID: 4002
```

Trap 220, shown in Example 16-17, is sent if a Global Mirror has exceeded the allowed threshold of failed FlashCopy commit attempts.

Example 16-17 Trap 220: Global Mirror number of FlashCopy commit failures exceed threshold

```
Global Mirror number of FlashCopy commit failures exceed threshold
UNIT: Mnf Type-Mod SerialNm
      IBM 2107-9A2 75-ABTV1
Session ID: 4002
```

Trap 221, shown in Example 16-18, is sent when the repository has reached the user-defined warning watermark or when physical space is completely exhausted.

Example 16-18 Trap 221: Space Efficient repository or overprovisioned volume has reached a warning watermark

```
Space Efficient Repository or Over-provisioned Volume has reached a warning
watermark
UNIT: Mnf Type-Mod SerialNm
      IBM 2107-9A2 75-ABTV1
Session ID: 4002
```

Table 16-3 shows the Copy Services suspension reason codes.

Table 16-3 Copy Services suspension reason codes

Suspension reason code (SRC)	Description
03	The host system sent a command to the primary volume of a Remote Mirror and Copy volume pair to suspend copy operations. The host system might have specified either an immediate suspension or a suspension after the copy completed and the volume pair reached a full duplex state.
04	The host system sent a command to suspend the copy operations on the secondary volume. During the suspension, the primary volume of the volume pair can still accept updates but updates are not copied to the secondary volume. The out-of-sync tracks that are created between the volume pair are recorded in the change recording feature of the primary volume.

Suspension reason code (SRC)	Description
05	Copy operations between the Remote Mirror and Copy volume pair were suspended by a primary storage unit secondary device status command. This system resource code can only be returned by the secondary volume.
06	Copy operations between the Remote Mirror and Copy volume pair were suspended because of internal conditions in the storage unit. This system resource code can be returned by the control unit of either the primary volume or the secondary volume.
07	Copy operations between the remote mirror and copy volume pair were suspended when the secondary storage unit notified the primary storage unit of a state change transition to simplex state. The specified volume pair between the storage units is no longer in a copy relationship.
08	Copy operations were suspended because the secondary volume became suspended as a result of internal conditions or errors. This system resource code can only be returned by the primary storage unit.
09	The Remote Mirror and Copy volume pair was suspended when the primary or secondary storage unit was rebooted or when the power was restored. The paths to the secondary storage unit might not be disabled if the primary storage unit was turned off. If the secondary storage unit was turned off, the paths between the storage units are restored automatically, if possible. After the paths have been restored, issue the mkpprc command to resynchronize the specified volume pairs. Depending on the state of the volume pairs, you might have to issue the rmpprc command to delete the volume pairs and reissue a mkpprc command to reestablish the volume pairs.
0A	The Remote Mirror and Copy pair was suspended because the host issued a command to freeze the Remote Mirror and Copy group. This system resource code can only be returned if a primary volume was queried.

16.2.3 I/O Priority Manager SNMP

When the I/O Priority Manager Control switch (with release 6.1) is set to **Monitor** or **Managed**, an SNMP trap alert message can also be enabled. The DS8000 microcode will monitor for *rank saturation*. If a rank is being overdriven to the point of saturation (Very High Usage), an SNMP trap alert message #224 will be posted to the SNMP server.

These SNMPs rules will be followed:

- Up to 8 SNMP traps per SFI server in 24 hour period (max 16 per 24 hours per SFI).

- ▶ Rank enters saturation state if in saturation for five consecutive 1 minute samples.
- ▶ Rank exits saturation state if not in saturation for three of five consecutive 1 minute samples.
- ▶ SNMP message #224 reported when enter saturation or every 8 hours if in saturation. The message identifies the rank and SFI.

16.2.4 Thin Provisioning SNMP

The DS8000 can trigger a specific SNMP trap alert message #223 related to thin provisioning feature. The trap is sent out when certain extent pool capacity thresholds are reached, causing a change in the extent status attribute. The conditions under which a trap is sent are:

- ▶ Extent status is not zero (available space already below threshold) when the first ESE volume is configured
- ▶ Extent status changes state if ESE volumes configured in Extent pool

Example 16-19 shows an illustration of generated event trap 223.

Example 16-19 SNMP trap alert message #223

```
2009/08/01 17:05:29 PDT
Extent Pool Capacity Threshold Reached
UNIT: Mnf Type-Mod SerialNm
IBM 2107-922 75-03460
Extent Pool ID: P1
Limit: 95%
Threshold: 95%
Status: 0
```

For additional information, refer to *DS8000 Thin Provisioning*, REDP-4554-00.

16.3 SNMP configuration

The SNMP for the DS8000 is designed to send traps as notifications. The DS8000 does not have an SNMP agent installed that can respond to SNMP polling. Also, the SNMP community name for Copy Service-related traps is fixed and set to **public**.

16.3.1 SNMP preparation

During the planning for the installation (see 9.3.4, “Monitoring DS8000 with the HMC” on page 234), the IP addresses of the management system are provided for the IBM service personnel. This information must be applied by IBM service personnel during the installation. Also, IBM service personnel can configure the HMC to either send a notification for every serviceable event, or send a notification for only those events that Call Home to IBM.

The network management server that is configured on the HMC receives all the generic trap 6 specific trap 3 messages, which are sent in parallel with any events that Call Home to IBM.

SNMP alerts can contain a combination of a generic and a specific alert trap. The Traps list outlines the explanations for each of the possible combinations of generic and specific alert traps. The format of the SNMP Traps, the list, and the errors reported by SNMP are available at chapter 5 of document available at the following URL:

http://publib.boulder.ibm.com/infocenter/dsichelp/ds8000ic/topic/com.ibm.storage.sic.help.doc/f2c_ictroubleshooting_36mf4d.pdf

SNMP alert traps provide information about problems that the storage unit detects. Either you or the service provider must perform corrective action for the related problems.

16.3.2 SNMP configuration with the HMC using WUI Menu

Customer can configure the SNMP alerting by logging in to the DS8000 HMC Service Management from remote or local through a web browser and launch the web application by using the following customer credentials:

User: customer

Password: cust0mer

1. Log into the Service Management section on the HMC Management console (Figure 16-2).

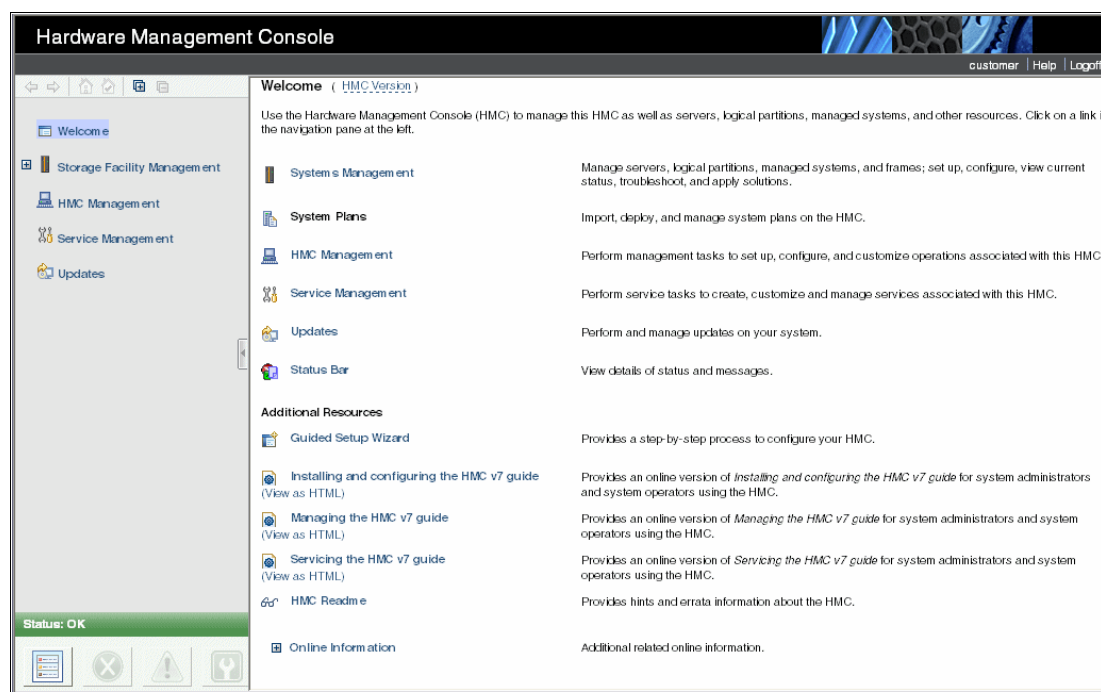


Figure 16-2 HMC Service Management

2. Select **Management Serviceable Event Notification** (Figure 16-3 on page 422) and insert TCP/IP information of SNMP server in the Trap Configuration folder.

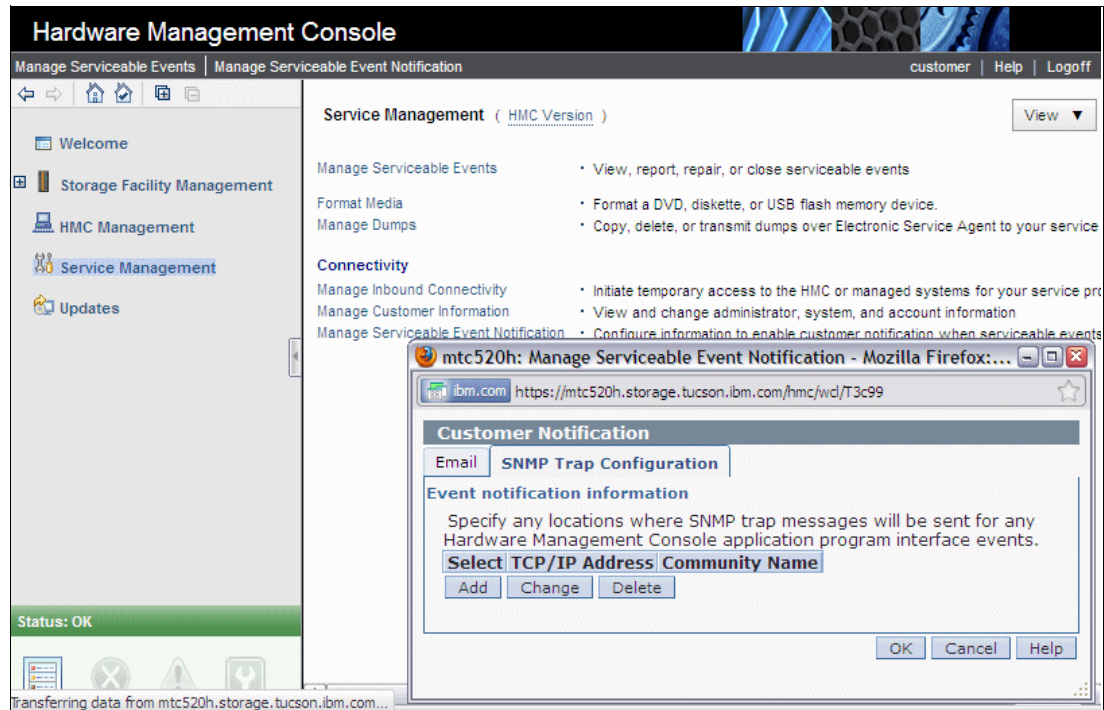


Figure 16-3 HMC Management Serviceable Event Notification

3. To verify the successful setup of your environment, create a Test Event on your DS8000 Hardware Management Console. Select **Storage Facility Management** → **Services Utilities** → **Test Problem Notification** (Figure 16-4).

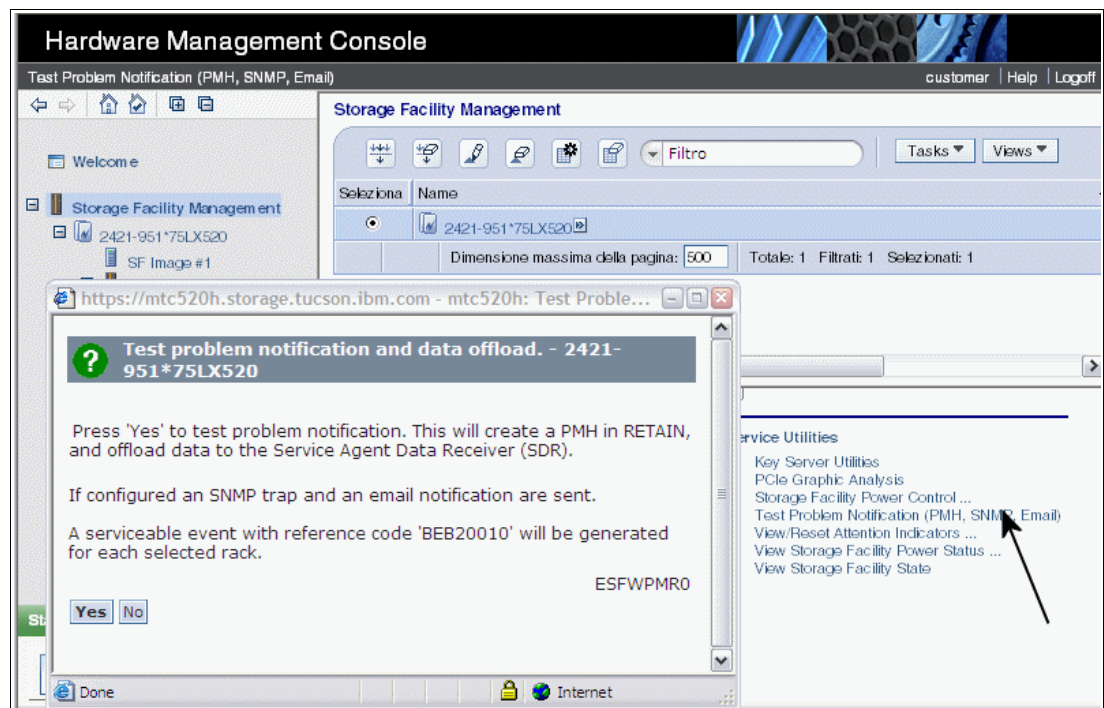


Figure 16-4 HMC Test SNMP trap

The test will generate the Service Reference Code BEB20010 and SNMP server will receive the SNMP trap notification as shown in Figure 16-5.

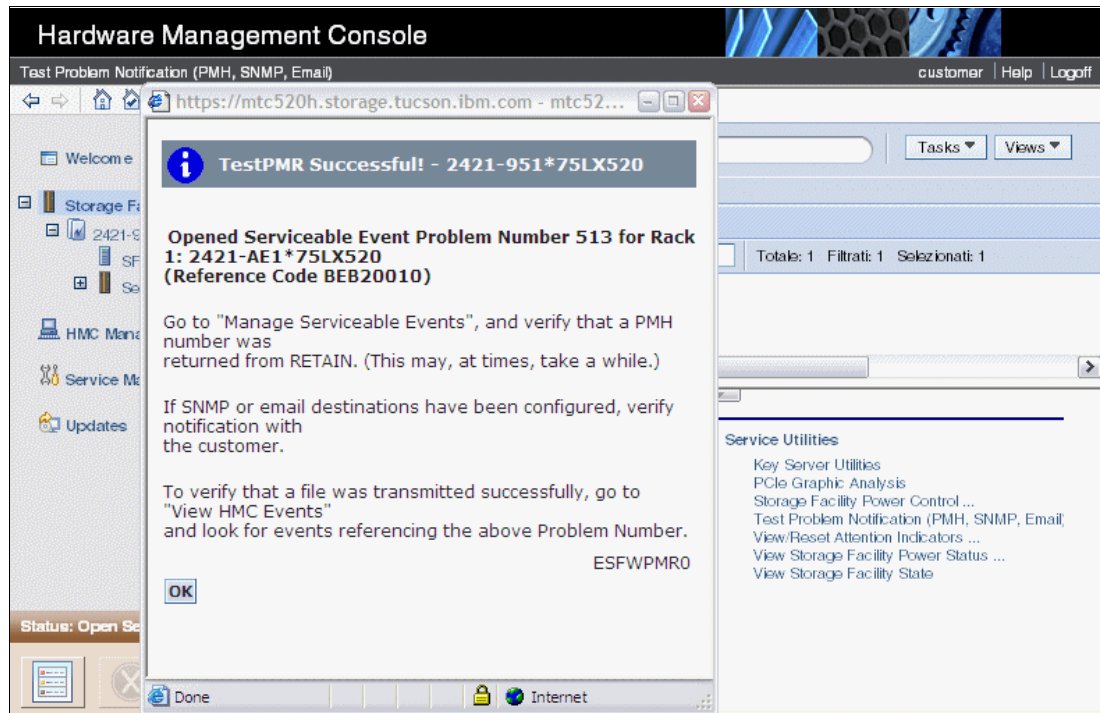


Figure 16-5 HMC SNMP trap test

16.3.3 SNMP configuration with the DS CLI

Perform the configuration for receiving the Copy Services-related traps using the DS CLI. Example 16-20 shows how SNMP is enabled by using the **chsp** command.

Example 16-20 Configuring the SNMP using dscli

```
dscli> chsp -snmp on -snmpaddr 10.10.10.1,10.10.10.2
Date/Time: April 25, 2011 03:55:33 PM MST IBM DSCLI Version: 7.6.10.464 DS: -
CMUC00040I chsp: Storage complex IbmStoragePlex successfully modified.

dscli> showsp
Date/Time: April 25, 2011 56:38:44 PM MST PDT IBM DSCLI Version: 7.6.10.464 DS: -
Name          IbmStoragePlex
desc          -
acct          -
SNMP          Enabled
SNMPadd       10.10.10.1,10.10.10.2
emailnotify   Disabled
emailaddr     -
emailrelay    Disabled
emailrelayaddr -
emailrelayhost -
numkssupported 4
```

SNMP preparation for the management software

To configure an SNMP Console, you need an MIB file. Configuration information for your SNMP manager and MIB can be found in the `SNMP_readme.txt` file located on your DS CLI installation CD-ROM.

For the DS8000, you can use the `ibm2100.mib` file, which is delivered on the DS CLI CD. Alternatively, you can download the latest version of the DS CLI CD image from the following address:

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



Remote support

In this chapter, we discuss the outbound (Call Home and Support Data offload) and inbound (code download and remote support) communications for the IBM System Storage DS8000. This chapter covers the following topics:

- ▶ Introduction to remote support
- ▶ IBM policies for remote support
- ▶ VPN advantages
- ▶ Remote connection types
- ▶ DS8000 support tasks
- ▶ Remote connection scenarios
- ▶ Audit logging

17.1 Introduction to remote support

Remote support is a complex topic that requires close scrutiny and education for all parties involved. IBM is committed to servicing the DS8000, whether it be warranty work, planned code upgrades, or management of a component failure, in a secure and professional manner. Dispatching service personnel to come to your site and perform maintenance on the system is still a part of that commitment. But as much as possible, IBM wants to minimize downtime and maximize efficiency by performing many support tasks remotely.

This plan of providing support remotely must be balanced with the client's expectations for security. Maintaining the highest levels of security in a data connection is a primary goal for IBM. This goal can only be achieved by careful planning with a client and a thorough review of all the options available.

17.1.1 Suggested reading

The following publications might be of assistance in understanding IBM remote support offerings:

- ▶ *IBM System Storage DS8000 Introduction and Planning Guide*, GC35-0515, contains additional information about physical planning. You can download it at the following address:
<http://www.ibm.com/systems/storage/disk/ds8000/index.html>
- ▶ *A Comprehensive Guide to Virtual Private Networks, Volume I: IBM Firewall, Server and Client Solutions*, SG24-5201, can be downloaded at the following address:
<http://www.redbooks.ibm.com/abstracts/sg245201.html?Open>
- ▶ The Security Planning website is available at the following address:
http://publib16.boulder.ibm.com/doc_link/en_US/a_doc_lib/aixbman/security/ipsec_planning.htm
- ▶ *VPNs Illustrated: Tunnels, VPNs, and IPSec*, by Jon C. Snader
- ▶ *VPN Implementation*, S1002693, which can be downloaded at the following address:
<http://www.ibm.com/support/docview.wss?rs=1114&uid=ssg1S1002693>

17.1.2 Organization of this chapter

A list of the relevant terminology for remote support is first presented. The remainder of this chapter is organized as follows:

- ▶ Connections
We review the types of connections that can be made from the HMC to the world outside of the DS8000.
- ▶ Tasks
We review the various support tasks that need to be run on those connections.
- ▶ Scenarios
We illustrate a scenario about how each task is performed over the types of remote connections.

17.1.3 Terminology and definitions

Listed here are brief explanations of some of the terms to be used when discussing remote support. See “Abbreviations and acronyms” on page 465 for a full list of terms and acronyms used in this book. Having an understanding of these terms will contribute to your discussions on remote support and security concerns. A generic definition will be presented here and then more specific information about how IBM implements the idea is given further on in this chapter.

IP network

There are many protocols running on Local Area Networks (LANs) around the world. Most companies use the Transmission Control Protocol/Internet Protocol (TCP/IP) standard for their connectivity between workstations and servers. IP is also the networking protocol of the global Internet. Web browsing and email are two of the most common applications that run on top of an IP network. IP is the protocol used by the DS8000 HMC to communicate with external systems, such as the SSPC or DS CLI workstations. There are two varieties of IP; refer to 8.3.3, “System Storage Productivity Center and network access” on page 213 for a discussion about the IPv4 and IPv6 networks.

SSH

Secure Shell is a protocol that establishes a secure communications channel between two computer systems. The term SSH is also used to describe a secure ASCII terminal session between two computers. SSH can be enabled on a system when regular Telnet and FTP are disabled, making it possible to only communicate with the computer in a secure manner.

FTP

File Transfer Protocol is a method of moving binary and text files from one computer system to another over an IP connection. It is not inherently secure as it has no provisions for encryption and only simple user and password authentication. FTP is considered appropriate for data that is already public, or if the entirety of the connection is within the physical boundaries of a private network.

SFTP

SSH File Transfer Protocol is unrelated to FTP. It is another file transfer method that is implemented inside a SSH connection. SFTP is generally considered to be secure enough for mission critical data and for moving sensitive data across the global Internet. FTP ports (usually ports 20/21) do not have to be open through a firewall for SFTP to work.

SSL

Secure Sockets Layer refers to methods of securing otherwise unsecure protocols such as HTTP (websites), FTP (files), or SMTP (email). Carrying HTTP over SSL is often referred to as HTTPS. An SSL connection over the global Internet is considered reasonably secure.

VPN

A Virtual Private Network is a private “tunnel” through a public network. Most commonly, it refers to using specialized software and hardware to create a secure connection over the Internet. The two systems, although physically separate, behave as though they are on the same private network. A VPN allows a remote worker or an entire remote office to remain part of a company’s internal network. VPNs provide security by encrypting traffic, authenticating sessions and users, and verifying data integrity.

Business-to-Business VPN

Business-to-business is a term for specialized VPN services for secure connections between IBM and its clients. This offering is also known as *Client Controlled VPN* and *Site-to-Site VPN*. This offering is in direct response to client concerns about being in control of VPN sessions with their vendors. It includes the use of a hardware VPN appliance inside the client's network, presumably one that can interact with many vendors' VPN clients.

IPSec

Internet Protocol Security is a suite of protocols used to provide a secure transaction between two systems that use the TCP/IP network protocol. IPSec focuses on authentication and encryption, two of the main ingredients of a secure connection. Most VPNs used on the Internet use IPSec mechanisms to establish the connection.

Firewall

A firewall is a device that controls whether data is allowed to travel onto a network segment. Firewalls are deployed at the boundaries of networks. They are managed by policies which declare what traffic can pass based on the sender's address, the destination address, and the type of traffic. Firewalls are an essential part of network security and their configuration must be taken into consideration when planning remote support activities.

Bandwidth

Bandwidth refers to the characteristics of a connection and how they relate to moving data. Bandwidth is affected by the physical connection, the logical protocols used, physical distance, and the type of data being moved. In general, higher bandwidth means faster movement of larger data sets.

17.2 IBM policies for remote support

The following guidelines are at the core of IBM remote support strategies for the DS8000:

- ▶ When the DS8000 needs to transmit service data to IBM, no host data of any kind is included. Only logs and process dumps are gathered for troubleshooting. The I/O from host adapters and the contents of NVS cache memory are never transmitted.
- ▶ When a VPN session with the DS8000 is needed, the HMC will always initiate such connections and only to predefined IBM servers/ports. There is never any active process that is "listening" for incoming sessions on the HMC.
- ▶ IBM maintains multiple-level internal authorizations for any privileged access to the DS8000 components. Only approved IBM service personnel can gain access to the tools that provide the one-time security codes for HMC command-line access.
- ▶ Although the HMC is based on a Linux operating system, IBM has disabled or removed all unnecessary services, processes, and IDs. This includes standard Internet services such as telnet, ftp, r commands, and rcp programs.

17.3 VPN rationale and advantages

Security is a critical issue for companies worldwide. Having a secure infrastructure requires systems to work together to mitigate the risk of malicious activity from both external and internal sources. Any connection from your network to the public Internet raises the following security concerns:

- ▶ Infection by viruses

- ▶ Intrusion by hackers
- ▶ The accessibility of your data from the remote support site
- ▶ Authorization of the remote users to access your machine when a remote connection is opened
- ▶ Support Center errors that might disrupt your storage network

Fortunately, using the IBM VPN connections along with the security features built-in to DS8000, it is possible to access IBM specialists who can assist you in resolving the most complex problems without the risks that are usually associated with a connection to an external network.

Remote-access support can help to greatly reduce service costs and shorten repair times, which in turn lessens the impact of any failures on your business. Using IBM security access security provides a number of advantages designed help you to save time and money and efficiently solve problems.

Here are just a few of the benefits you can realize:

- ▶ Faster problem solving. You can call on technical experts worldwide to help resolve problems on your DS8000 without having to wait to receive logs, dumps, and traces. As a result, problems can be solved faster.
- ▶ Connection with a worldwide network of experts. IBM Technical support engineers can call on other worldwide subject experts to assist with problem determination. These engineers can then simultaneously view the DS8000 master console.
- ▶ Closer monitoring and enhanced collaboration. You can monitor the actions taken on your master console and join in conference calls with the IBM support engineers as the problem determination process proceeds.
- ▶ Save time and money. Many of your problems can be solved without IBM ever having to send an engineer to your site.

17.4 Remote connection types

The DS8000 HMC has a connection point for the client's network by a standard Ethernet (10/100/1000 Mb) cable. The HMC also has a connection point for a phone line by the modem port. These two physical connections offer four possibilities for sending and receiving data between the DS8000 and IBM. The connection types are:

- ▶ Asynchronous modem connection
- ▶ IP network connection
- ▶ IP network connection with VPN
- ▶ IP network connection with Business-to-Business VPN

In the most secure environments, both of these physical connections (Ethernet and modem) remain unplugged. The DS8000 serves up storage for its connected hosts, but has no other communication with the outside world. This means that all configuration tasks have to be done while standing at the HMC (there is no usage of the SSPC or DS CLI). This level of security, known as an air gap, also means that there is no way for the DS8000 to alert anyone that it has encountered a problem and there is no way to correct such a problem other than to be physically present at the system.

So rather than leaving the modem and Ethernet disconnected, clients will provide these connections and then apply policies on when they are to be used and what type of data they can carry. Those policies are enforced by the settings on the HMC and the configuration of

client network devices, such as routers and firewalls. The next four sections discuss the capabilities of each type of connection.

17.4.1 Asynchronous modem

A modem creates a low-speed asynchronous connection using a telephone line plugged into the HMC modem port. This type of connection favors transferring small amounts of data. It is relatively secure because the data is not traveling across the Internet. However, this type of connection is not terribly useful due to bandwidth limitations. Average connection speed in the US mainland is 28-36 Kbps, and can be less in other parts of the world.

DS8000 HMC modem can be configured to call IBM and send small status messages. Authorized support personnel can call the HMC and get privileged access to the command line of the operating system. Typical PEPackage transmission is not normally performed over a modem line because it can take 15 to 20 hours depending on the quality of the connection. Code downloads over a modem line are not possible.

The client has control over whether or not the modem will answer an incoming call. These options are changed from the WebUI on the HMC by selecting **Service Management** → **Manage Inbound Connectivity** as shown in Figure 17-1.

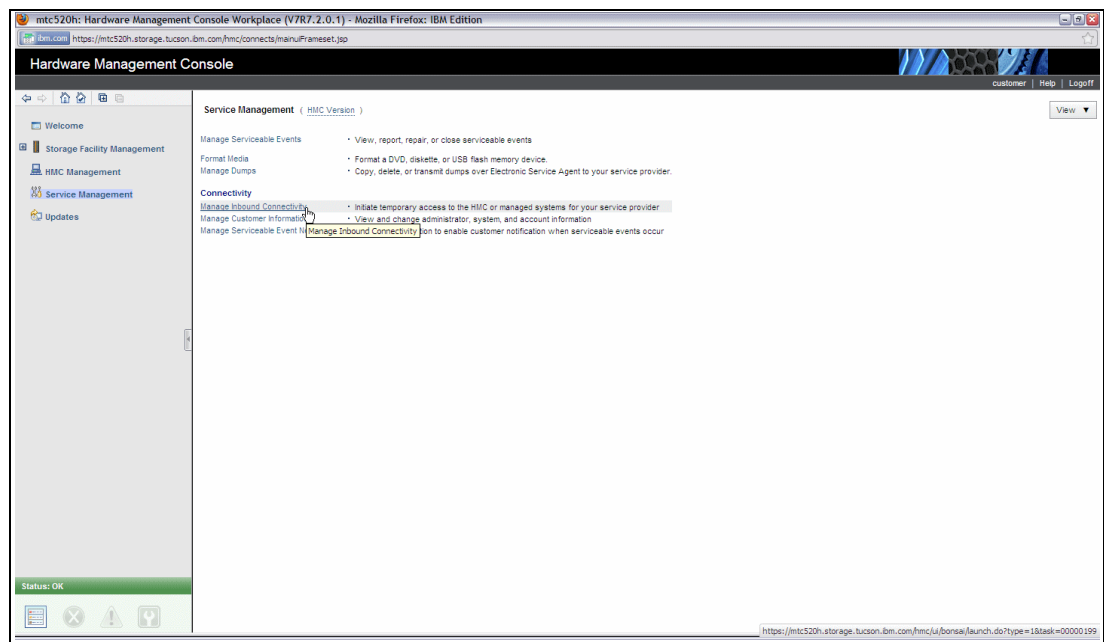


Figure 17-1 Service Management in WebUI

The HMC provides several settings to govern the usage of the modem port:

- ▶ Unattended Session

This check box allows the HMC to answer modem calls without operator intervention. If this is not checked, then someone must go to the HMC and allow for the next expected call. IBM Support must contact the client every time they need to dial in to the HMC.

- ▶ Duration: Continuous

This option indicates that the HMC can answer all calls at all times.

- ▶ Duration: Automatic

This option indicates that the HMC will answer all calls for n days following the creation of any new Serviceable Event (problem).

- **Duration: Temporary**

This option sets a starting and ending date, during which the HMC will answer all calls.

These options are shown in Figure 17-2. See Figure 17-3 on page 438 for an illustration of a modem connection.

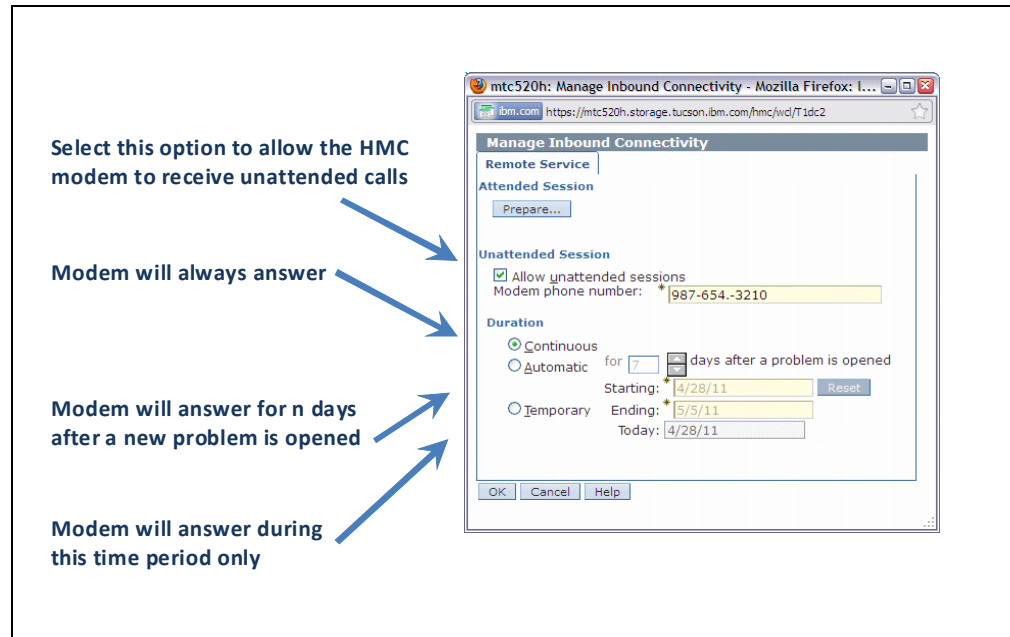


Figure 17-2 Modem settings

17.4.2 IP network

Network connections are considered high speed in comparison to a modem. Enough data can flow through a network connection to make it possible to run a graphical user interface (GUI). Managing a DS8000 from an SSPC would not be possible over a modem line; it requires the bandwidth of a network connection.

HMCs connected to a client IP network, and eventually to the Internet, can send status updates and offloaded problem data to IBM using SSL sessions. They can also use FTP to retrieve new code bundles from the IBM code repository. It typically take less than an hour to move the information.

Though favorable for speed and bandwidth, network connections introduce security concerns. Care must be taken to:

- Verify the authenticity of data, that is, is it really from the sender it claims to be?
- Verify the integrity of data, that is, has it been altered during transmission?
- Verify the security of data, that is, can it be captured and decoded by unwanted systems?

The Secure Sockets Layer (SSL) protocol is one answer to these questions. It provides transport layer security with authenticity, integrity, and confidentiality, for a secure connection between the client network and IBM. Some of the features that are provided by SSL are:

- Client and server authentication to ensure that the appropriate machines are exchanging data

- ▶ Data signing to prevent unauthorized modification of data while in transit
- ▶ Data encryption to prevent the exposure of sensitive information while data is in transit

See Figure 17-5 on page 440 for an illustration of a basic network connection.

17.4.3 IP network with traditional VPN

Adding a VPN “tunnel” to an IP network greatly increases the security of the connection between the two endpoints. Data can be verified for authenticity and integrity. Data can be encrypted so that even if it is captured enroute, it cannot be “replayed” or deciphered.

Having the safety of running within a VPN, IBM can use its service interface (WebUI) to:

- ▶ Check the status of components and services on the DS8000 in real time
- ▶ Queue up diagnostic data offloads
- ▶ Start, monitor, pause, and restart repair service actions

Performing the following steps result in the HMC creating a VPN tunnel back to the IBM network, which service personnel can then use. There is no VPN service that sits idle, waiting for a connection to be made by IBM. Only the HMC is allowed to initiate the VPN tunnel, and it can only be made to predefined IBM addresses. The steps to create a VPN tunnel from the DS8000 HMC to IBM are listed here:

1. IBM support calls the HMC using the modem. After the first level of authentications, the HMC is asked to launch a VPN session.
2. The HMC hangs up the modem call and initiates a VPN connection back to a predefined address or port within IBM Support.
3. IBM Support verifies that they can see and use the VPN connection from an IBM internal IP address.
4. IBM Support launches the WebUI or other high-bandwidth tools to work on the DS8000.

See Figure 17-6 on page 441 for an illustration of a traditional VPN connection.

17.4.4 IP network with Business-to-Business VPN

The Business-to-Business VPN option does not add any new functionality; IBM Support can perform all of the tasks as with the traditional HMC-based VPN. What a Business-to-Business VPN does provide is a greater measure of control over the VPN sessions by the client. Instead of a VPN tunnel being created between the HMC and the IBM network, a tunnel is created from the client’s VPN appliance to the IBM network. This option has also been referred to as *client controlled VPN*.

Clients who work with many vendors that have their own remote support systems often own and manage a VPN appliance, a server that sits on the edge of their network and creates tunnels with outside entities. This is true for many companies that have remote workers, outside sales forces, or small branch offices. Because the device is already configured to meet the client’s security requirements, they only need to add appropriate policies for IBM support. Most commercially-available VPN servers are interoperable with the IPSec-based VPN that IBM needs to establish. Using a Business-to-Business VPN layout leverages the investment that a client has already made in establishing secure tunnels into their network.

The VPN tunnel that gets created is valid for IBM Remote Support use only and has to be configured both on the IBM and client sides. This design provides several advantages for the client:

- ▶ Allows the client to use Network Address Translation (NAT) so that the HMC is given a non-routable IP address behind the company firewall.
- ▶ Allows the client to inspect the TCP/IP packets that are sent over this VPN.
- ▶ Allows the client to disable the VPN on their device for “lockdown” situations.

Note that the Business-to-Business VPN only provides the tunnel that service personnel can use to actively work with the HMC from within IBM. To offload data or call home, the HMC still needs to have one of the following:

- ▶ Modem access
- ▶ Non-VPN network access (SSL connection)
- ▶ Traditional VPN access

See Figure 17-7 on page 442 for an illustration of a Business-to-Business VPN connection.

17.5 DS8000 support tasks

DS8000 support tasks are tasks that require the HMC to contact the outside world. Some tasks can be performed using either the modem or the network connection, and some can only be done over a network. The combination of tasks and connection types is illustrated in 17.6, “Remote connection scenarios” on page 437. The support tasks that require the DS8000 to connect to outside resources are:

- ▶ Call Home and heartbeat
- ▶ Data offload
- ▶ Code download
- ▶ Remote support

17.5.1 Call Home and heartbeat (outbound)

Here we discuss the Call Home and heartbeat capabilities.

Call Home

Call Home is the capability of the HMC to contact IBM Service to report a service event. This is referred to as *Call Home for service*. The HMC provides machine reported product data (MRPD) information to IBM by way of the Call Home facility. The MRPD information includes installed hardware, configurations, and features. The Call Home also includes information about the nature of a problem so that an active investigation can be launched. Call Home is a one-way communication, with data moving from the DS8000 HMC to the IBM data store.

Heartbeat

The DS8000 also uses the Call Home facility to send proactive *heartbeat* information to IBM. A heartbeat is a small message with basic product information so that IBM knows the unit is operational. By sending heartbeats, both IBM and the client ensure that the HMC is always able to initiate a full Call Home to IBM in the case of an error. If the heartbeat information does not reach IBM, a service call to the client will be made to investigate the status of the DS8000. Heartbeats represent a one-way communication, with data moving from the DS8000 HMC to the IBM data store.

The Call Home facility can be configured to:

- ▶ Use the HMC modem
- ▶ Use the Internet through a SSL connection
- ▶ Use the Internet through a VPN tunnel from the HMC to IBM

Call Home information and heartbeat information are stored in the IBM internal data store so the support representatives have access to the records.

17.5.2 Data offload (outbound)

For many DS8000 problem events, such as a hardware component failure, a large amount of diagnostic data is generated. This data can include text and binary log files, firmware dumps, memory dumps, inventory lists, and timelines. These logs are grouped into collections by the component that generated them or the software service that owns them. The entire bundle is collected together in what is called a *PEPackage*. A DS8000 PEPackage can be quite large, often exceeding 100 MB. In certain cases, more than one might be needed to properly diagnose a problem. In certain cases, the IBM Support center might need an additional dump internally created by DS8000 or manually created through the intervention of an operator.

Tip: From Release 6.1, the On Demand Data (ODD) Dump has been introduced. The On Demand Data (ODD) Dump will provide a mechanism that will allow the collection of debug data for error scenarios traditionally handled using One Shot Panic. With ODD Dump, IBM will be able to collect data after initial error occurrence with no impact to host I/O.

The HMC is a focal point, gathering and storing all the data packages. So the HMC must be accessible if a service action requires the information. The data packages must be offloaded from the HMC and sent in to IBM for analysis. The offload can be done in several ways:

- ▶ Modem offload
- ▶ Standard FTP offload
- ▶ SSL offload
- ▶ VPN offload

Modem offload

The HMC can be configured to support automatic data offload using the internal modem and a regular phone line. Offloading a PEPackage over a modem connection is extremely slow, in many cases taking 15 to 20 hours. It also ties up the modem for this time so that IBM support cannot dial in to the HMC to perform command-line tasks. If this is the only connectivity option available, be aware that the overall process of remote support will be delayed while data is in transit.

Standard FTP offload

The HMC can be configured to support automatic data offload using File Transfer Protocol (FTP) over a network connection. This traffic can be examined at the client's firewall before moving across the Internet. FTP offload allows IBM Service personnel to dial in to the HMC using the modem line while support data is being transmitted to IBM over the network.

Tip: FTP offload of data is supported as an outbound service only. There is no active FTP server running on the HMC that can receive connection requests.

When a direct FTP session across the Internet is not available or desirable, a client can configure the FTP offload to use a client-provided FTP proxy server. The client then becomes responsible for configuring the proxy to forward the data to IBM.

The client is required to manage its firewalls so that FTP traffic from the HMC (or from an FTP proxy) can pass onto the Internet.

SSL offload

For environments that do not permit FTP traffic out to the Internet, the DS8000 also supports offload of data using SSL security. In this configuration, the HMC uses the client-provided network connection to connect to the IBM data store, the same as in a standard FTP offload. But with SSL, all the data is encrypted so that it is rendered unusable if intercepted.

Client firewall settings between the HMC and the Internet for SSL setup require four IP addresses open on port 443 based on geography as detailed here:

► North and South America

129.42.160.48	IBM Authentication Primary
207.25.252.200	IBM Authentication Secondary
129.42.160.49	IBM Data Primary
207.25.252.204	IBM Data Secondary

► All other regions

129.42.160.48	IBM Authentication Primary
207.25.252.200	IBM Authentication Secondary
129.42.160.50	IBM Data Primary
207.25.252.205	IBM Data Secondary

VPN offload

A remote service VPN session can be initiated by HMC for data offload over a modem or an Internet VPN connection. At least one of these methods of connectivity must be configured through the Outbound Connectivity panel. Note that the VPN session is always initiated outbound from the HMC, not inbound.

When there is a firewall in place to shield the customer network from the open internet, the firewall must be configured to allow the SMC to connect to the IBM servers. The HMC establishes connection to the following TCP/IP addresses:

207.25.252.196	IBM Boulder VPN Server
129.42.160.16	IBM Rochester VPN Server

You must also enable the following ports and protocols:

ESP
UDP Port 500
UDP Port 4500

Example 17-1 shows the output of defined permissions that are based on a Cisco PIX model 525 firewall.

Example 17-1 Cisco Firewall configuration

```
access-list DMZ_to_Outside permit esp host 207.25.252.196 host <IP addr for HMC>
access-list DMZ_to_Outside permit esp host 129.42.160.16 host <IP addr for HMC>
access-list DMZ_to_Outside permit udp host 207.25.252.196 host <IP addr for HMC> eq 500
access-list DMZ_to_Outside permit udp host 129.42.160.16 host <IP addr for HMC> eq 500
access-list DMZ_to_Outside permit udp host 207.25.252.196 host <IP addr for HMC> eq 4500
access-list DMZ_to_Outside permit udp host 129.42.160.16 host <IP addr for HMC> eq 4500
```

Only the HMC customer network must be defined for access to the IBM VPN Servers. The IPsec tunneling technology used by the VPN software, in conjunction with the TCP/IP port forwarder on the HMC, provide the ability for IBM Service to access the DS8000 servers themselves through the secure tunnel.

Comparison of DS8000 connectivity options

Table 17-1 shows the benefits and drawbacks of the various types of connection. The terms *remote access* and *remote service* are used interchangeably throughout this document.

Service activities include problem reporting, debug data offload, and remote access. Note that enabling multiple options is allowed, and need to be used for optimal availability.

Table 17-1 Remote support connectivity comparison

Connectivity Option	Pros	Cons	Comments
FTP	- Fast debug data transfer to IBM - Allows proxying	Does not support problem reporting or remote access	To support all service activities, at least one of the following must also be enabled as an adjunct: VPN internet, or modem.
Internet (SSL)	- Fast debug data transfer to IBM - Supports problem reporting - For various reasons, such as proxying, SSL is easier to implement than VPN	Does not support remote access.	To support remote access, at least one of the following must also be enabled as an adjunct: VPN internet, or modem.
VPN Internet	- Fast debug data transfer - Supports all service activities	- Can be difficult to implement in some environments - Does not allow you to inspect packets.	Generally the best option. Might be the only option enabled. However, use the modem as a backup and for initiating remote access sessions.
Modem	- Supports all service activities - Allows IBM service to remotely initiate an outbound VPN session	- Extremely slow debug data transfer to IBM	Might be the only option enabled.

17.5.3 Code download (inbound)

DS8000 microcode updates are published as *bundles* that can be downloaded from IBM. As explained in 15.6, “Loading the code bundle” on page 406, there are three possibilities for acquiring code on the HMC:

- ▶ Load the new code bundle using CDs/DVDs.
- ▶ Download the new code bundle directly from IBM using FTP.
- ▶ Download the new code bundle directly from IBM using SFTP.

Loading code bundles from CDs/DVDs is the only option for DS8000 installations that have no outside connectivity at all. If the HMC is connected to the client network then IBM support will download the bundles from IBM using either FTP or SFTP.

FTP

If allowed, the support representative will open an FTP session from the HMC to the IBM code repository and download the code bundle(s) to the HMC. The client firewall will need to be configured to allow the FTP traffic to pass.

SFTP

If FTP is not allowed, an SFTP session can be used instead. SFTP is a more secure file transfer protocol running within an SSH session, as defined in 17.1.3, “Terminology and definitions” on page 427. If this option is used, the client firewall will need to be configured to allow the SSH traffic to pass.

After the code bundle is acquired from IBM, the FTP or SFTP session will be closed and the code load can take place without needing to communicate outside of the DS8000.

17.5.4 Remote support (inbound and two-way)

Remote support describes the most interactive level of assistance from IBM. After a problem comes to the attention of the IBM Support Center and it is determined that the issue is more complex than a straightforward parts replacement, the problem will likely be escalated to higher levels of responsibility within IBM Support. This could happen at the same time that a support representative is being dispatched to the client site.

IBM might need to trigger a data offload, perhaps more than one, and at the same time be able to interact with the DS8000 to dig deeper into the problem and develop an action plan to restore the system to normal operation. This type of interaction with the HMC is what requires the most bandwidth.

If the only available connectivity is by modem, then IBM Support will have to wait until any data offload is complete and then attempt the diagnostics and repair from a command-line environment on the HMC. This process is slower and more limited in scope than if a network connection can be used.

If a VPN is available, either from the HMC directly to IBM or by using VPN devices (Business-to-Business VPN option), then enough bandwidth is available for data offload and interactive troubleshooting to be done at the same time. IBM Support will be able to use graphical tools (WebUI and others) to diagnose and repair the problem.

17.6 Remote connection scenarios

Now that the four connection options have been reviewed (see 17.4, “Remote connection types” on page 429) and the tasks have been reviewed (see 17.5, “DS8000 support tasks” on page 433), we can examine how each task is performed given the type of access available to the DS8000.

17.6.1 No connections

If both the modem or the Ethernet are not physically connected and configured, then the tasks are performed as follows:

- **Call Home and heartbeat:** The HMC will not send heartbeats to IBM. The HMC will not call home if a problem is detected. IBM Support will need to be notified at the time of installation to add an exception for this DS8000 in the heartbeats database, indicating that it is not expected to contact IBM.

- ▶ **Data offload:** If absolutely required and allowed by the client, diagnostic data can be burned onto a DVD, transported to an IBM facility, and uploaded to the IBM data store.
- ▶ **Code download:** Code must be loaded onto the HMC using CDs carried in by the Service Representative.
- ▶ **Remote support:** IBM cannot provide any remote support for this DS8000. All diagnostic and repair tasks must take place with an operator physically located at the console.

17.6.2 Modem only

If the modem is the only connectivity option, then the tasks are performed as follows:

- ▶ **Call Home and heartbeat:** The HMC will use the modem to call IBM and send the Call Home data and the heartbeat data. These calls are of short duration.
- ▶ **Data offload:** After data offload is triggered, the HMC will use the modem to call IBM and send the data package. Depending on the package size and line quality, this call could take up to 20 hours to complete.
- ▶ **Code download:** Code must be loaded onto the HMC using CDs/DVDs carried in by the Service Representative. There is no method of download if only a modem connection is available.
- ▶ **Remote support:** If the modem line is available (not being used to offload data or send Call Home data), IBM Support can dial in to the HMC and execute commands in a command-line environment. IBM Support cannot utilize a GUI or any high-bandwidth tools.

See Figure 17-3 for an illustration of a modem-only connection.

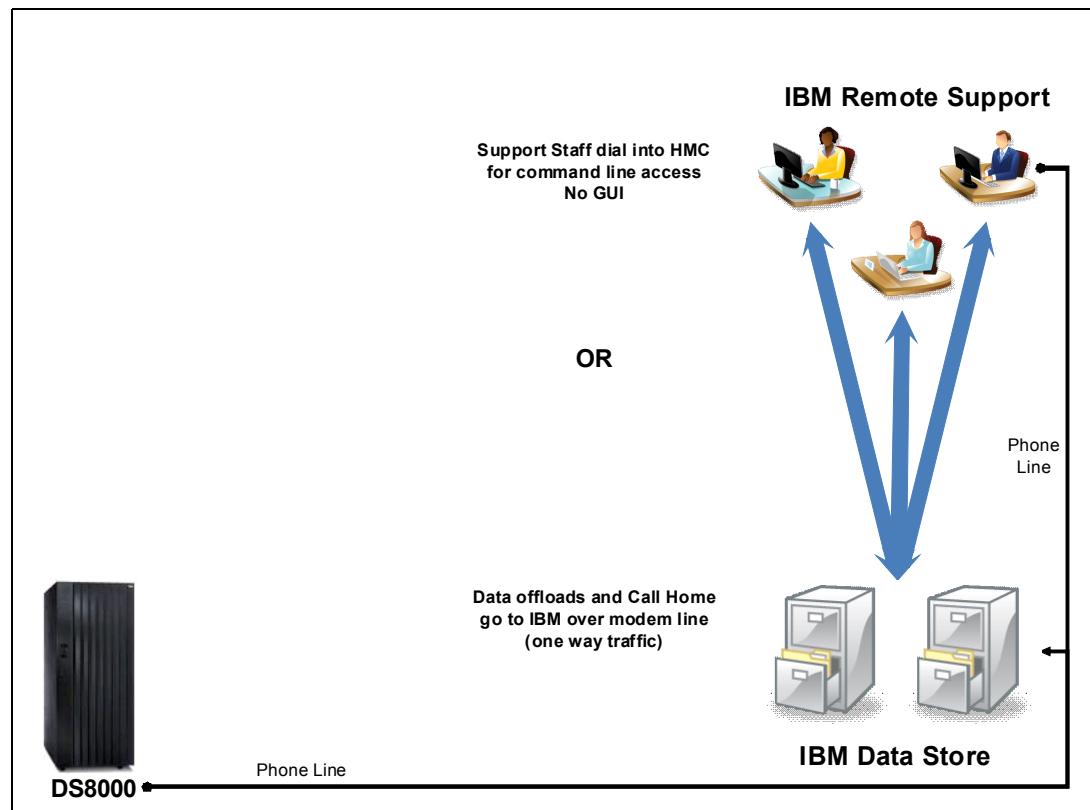


Figure 17-3 Remote support with modem only

17.6.3 VPN only

If the VPN is the only connectivity option, the tasks are performed as follows:

- ▶ Call Home and heartbeat: The HMC will use the VPN network to call IBM and send the Call Home data and the heartbeat data.
- ▶ Data offload: After data offload is triggered, the HMC will use the VPN network to call IBM and send the data package. The package will be sent to IBM server quickly.
- ▶ Code download: Code must be loaded onto the HMC using CDs/DVDs carried in by the Service Representative. There is no method of download if only the VPN connection is available.
- ▶ Remote support: IBM Support center will call you and ask you to open a VPN connection before starting the remote connection. After the VPN has been opened, the IBM Support center can connect to the HMC and execute commands in a command-line environment. IBM Support cannot utilize a GUI.

Figure 17-4 shows an illustration of a VPN-only connection

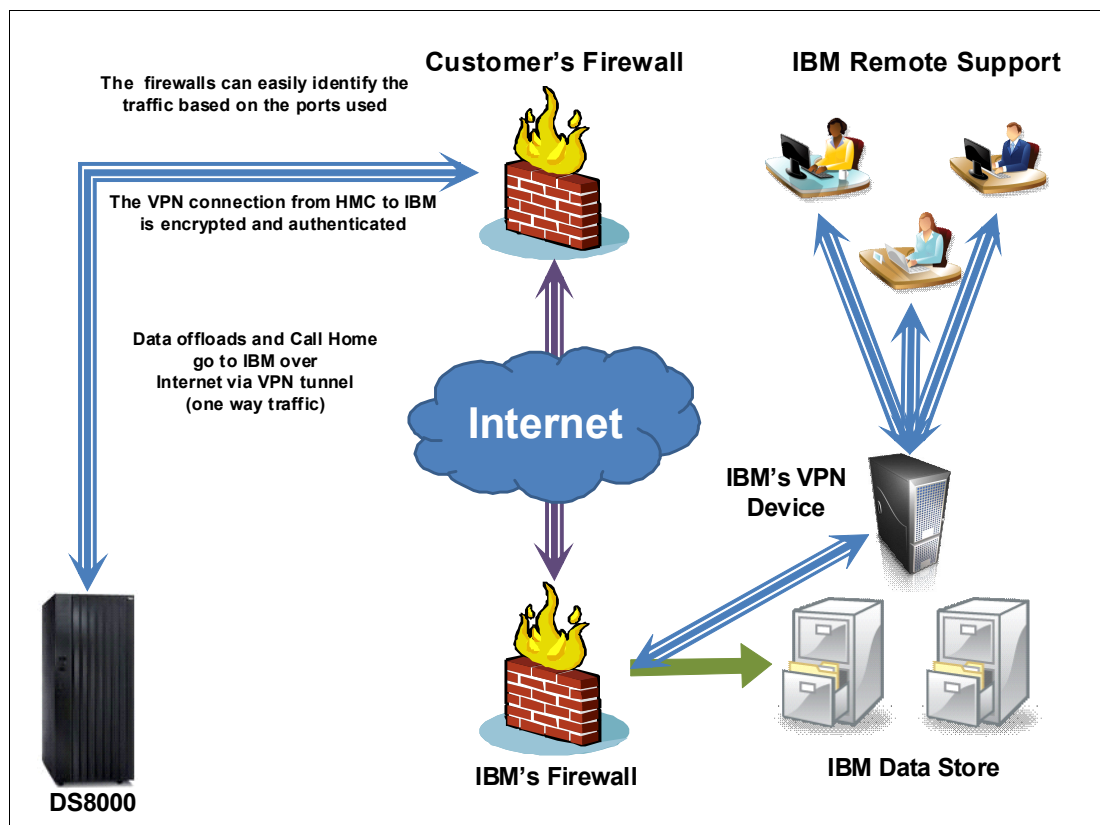


Figure 17-4 Remote support with VPN only

17.6.4 Modem and network with no VPN

If the modem and network access, without VPN, are provided, then the tasks are performed as follows:

- ▶ Call Home and heartbeat: The HMC will use the network connection to send Call Home data and heartbeat data to IBM across the Internet.
- ▶ Data offload: The HMC will use the network connection to send offloaded data to IBM across the Internet. Standard FTP or SSL sockets can be used.

- ▶ **Code download:** Code can be downloaded from IBM using the network connection. The download can be done using FTP or SFTP.
- ▶ **Remote support:** Even though there is a network connection, it is not configured to allow VPN traffic, so remote support must be done using the modem. If the modem line is not busy, IBM Support can dial in to the HMC and execute commands in a command-line environment. IBM Support cannot utilize a GUI or any high-bandwidth tools.

Figure 17-5 shows an illustration of a modem and network connection without using VPN tunnels.

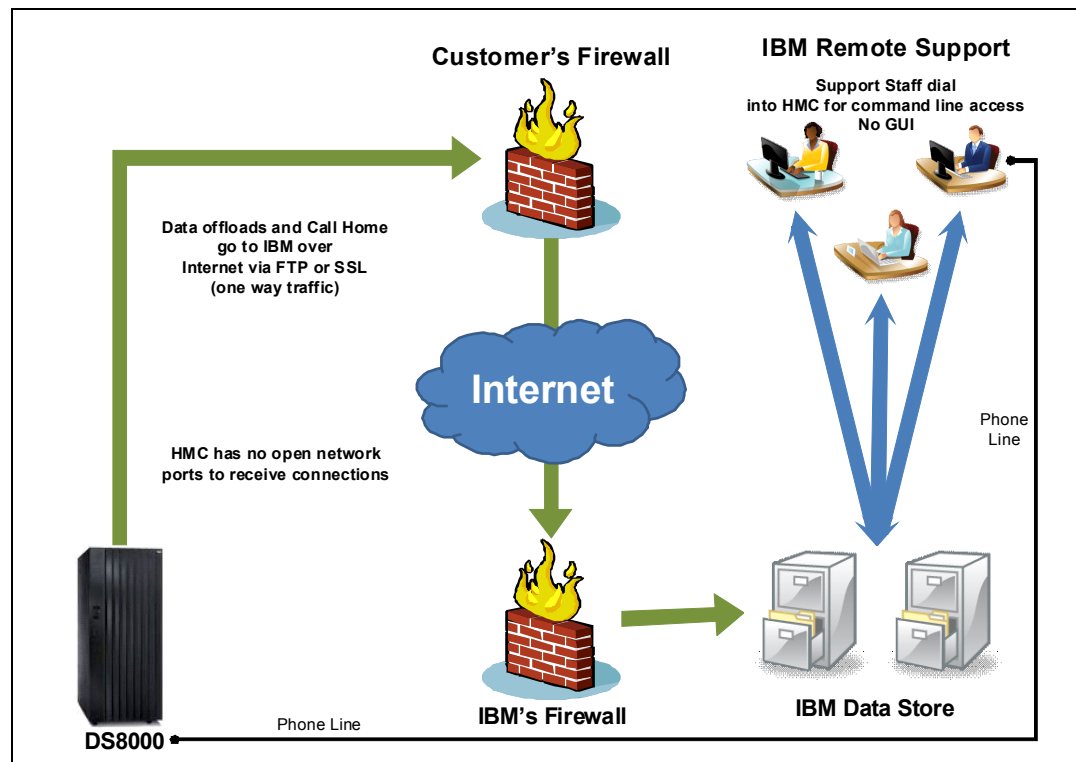


Figure 17-5 Remote support with modem and network (no VPN)

17.6.5 Modem and traditional VPN

If the modem and a VPN-enabled network connection is provided, then the tasks are performed as follows:

- ▶ **Call Home and heartbeat:** The HMC will use the network connection to send Call Home data and heartbeat data to IBM across the Internet, outside of a VPN tunnel.
- ▶ **Data offload:** The HMC will use the network connection to send offloaded data to IBM across the Internet, outside of a VPN tunnel. Standard FTP or SSL sockets can be used.
- ▶ **Code download:** Code can be downloaded from IBM using the network connection. The download can be done using FTP or SFTP outside of a VPN tunnel.
- ▶ **Remote support:** Upon request, the HMC establishes a VPN tunnel across the Internet to IBM. IBM Support can use a GUI and high-bandwidth tools to interact with the HMC at the same time that data is offloading.

Figure 17-6 shows an illustration of a modem and network connection plus traditional VPN.

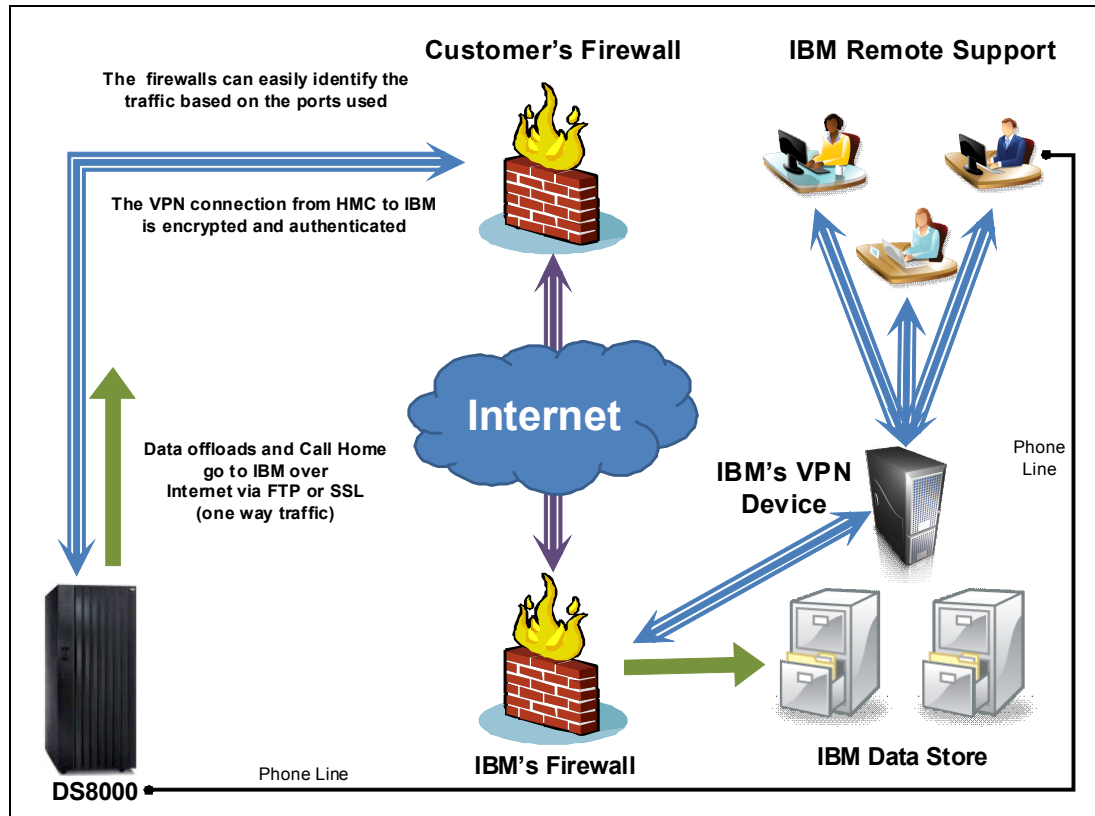


Figure 17-6 Remote support with modem and traditional VPN

17.6.6 Modem and Business-to-Business VPN

If a modem plus a network connection plus a Business-to-Business VPN appliance are installed, then the tasks are performed as follows:

- ▶ **Call Home and heartbeat:** The HMC will use the network connection to send Call Home data and heartbeat data to IBM across the Internet, outside of a VPN tunnel.
- ▶ **Data offload:** The HMC will use the network connection to send offloaded data to IBM across the Internet, outside of a VPN tunnel. Standard FTP or SSL sockets can be used.
- ▶ **Code download:** Code can be downloaded from IBM using the network connection. The download can be done using FTP or SFTP outside of a VPN tunnel.
- ▶ **Remote support:** The VPN tunnel is established between the client's VPN appliance and the IBM VPN appliance. IBM Support can use a GUI and high-bandwidth tools to interact with the HMC at the same time as data offload. The HMC does not have to be involved in establishing the VPN session.

Figure 17-7 shows an illustration of a modem and network connection plus Business-to-Business VPN deployment.

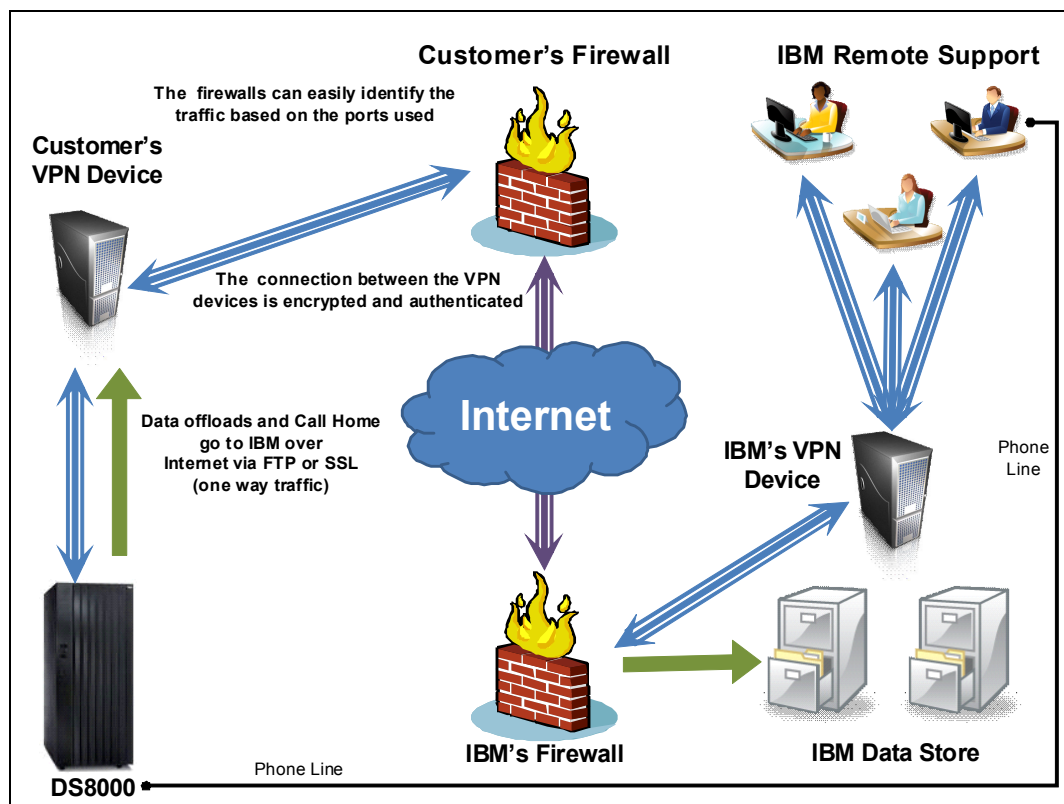


Figure 17-7 Remote support with modem and Business-to-Business VPN

17.7 Audit logging

The DS8000 offers an audit logging security function designed to track and log changes made by administrators using either Storage Manager DS GUI or DS CLI. This function also documents remote support access activity to the DS8000. The audit logs can be downloaded by DS CLI or Storage Manager.

Example 17-2 illustrates the DS CLI **offloadauditlog** command that provides clients with the ability to offload the audit logs to the DS CLI workstation in a directory of their choice.

Example 17-2 DS CLI command to download audit logs

```
dsccli> offloadauditlog -logaddr smc1 c:\75LX520_audit.txt
Date/Time: April 29, 2011 04:33:23 PM MST IBM DSCLI Version: 7.6.10.464 DS: -
CMUC00244W offloadauditlog: The specified file currently exists. Are you sure you
want to replace the file? [y/n]: y
CMUC00243I offloadauditlog: Audit log was successfully offloaded from smc1 to
c:\75LX520_audit.txt.
```

The downloaded audit log is a text file that provides information about when a remote access session started and ended, and what remote authority level was applied. A portion of the downloaded file is shown in Example 17-3.

Example 17-3 Audit log entries related to a remote support event using a modem

```
U,2011/04/27 09:10:57:000
MST,,1,IBM.2107-75LX520,N,8000,Phone_started,Phone_connection_started,,,
U,2011/04/27 09:11:16:000
MST,,1,IBM.2107-75LX520,N,8036,Authority_to_root,Challenge Key = 'Fy31@C37';
Authority_upgrade_to_root,,,
U,2011/04/27 12:09:49:000
MST,customer,1,IBM.2107-75LX520,N,8020,WUI_session_started,,,,
U,2011/04/27 15:35:30:000
MST,customer,1,IBM.2107-75LX520,N,8022,WUI_session_logoff,WUI_session_ended_logged
off,,,
U,2011/04/27 16:49:18:000
MST,,1,IBM.2107-75LX520,N,8002,Phone_ended,Phone_connection_ended,,,
```

The *Challenge Key* shown is not a password on the HMC. It is a token shown to the IBM support representative who is dialing in to the DS8000. The representative must use the Challenge Key in an IBM internal tool to generate a *Response Key* that is given to the HMC. The Response Key acts as a one-time authorization to the features of the HMC. The Challenge and Response Keys change every time a remote connection is made.

The Challenge-Response process must be repeated again if the representative needs to escalate privileges to access the HMC command-line environment. There is no direct user login and no root login through the modem on a DS8000 HMC.

Audit Log Requirements

Entries are added to the audit file only after the operation has completed. All information about the request and its completion status is known. A single entry is used to log both request and response information. It is possible, though unlikely, that an operation does not complete due to an operation timeout. In this case, no entry is made in the log. Below are the main roles for audit log entry:

- ▶ Log users that connect/disconnect to the storage manager.
- ▶ Log user password and user access violations.
- ▶ Log commands that create, remove, or modify logical configuration, including command parameters and user ID.
- ▶ Log commands that modify Storage Facility and Storage Facility settings, including command parameters and user ID.
- ▶ Allow user to add comments to the log.
- ▶ Log Copy Services commands, including command parameters and user ID (TPC-R commands are not supported).
- ▶ Log abnormal system events such as failover, or events that cause a loss of access such as rank down (not in first release).

Audit logs have the following characteristics:

- ▶ Logs should be maintained for a period of 30 days. It is the user's responsibility to periodically extract the log and save it away.
- ▶ Logs are automatically trimmed (FIFO) by the subsystem so they do not consume more than 50 megabytes of disk storage.

For a detailed description about how auditing is used to record “who did what and when” in the audited system, and a guide to log management, visit the following URL:

<http://csrc.nist.gov/publications/nistpubs/800-92/SP800-92.pdf>



Capacity upgrades and CoD

This chapter discusses aspects of implementing capacity upgrades and Capacity on Demand (CoD) with the IBM System Storage DS8800. This chapter covers the following topics:

- ▶ Installing capacity upgrades
- ▶ Using Capacity on Demand (CoD)

18.1 Installing capacity upgrades

Storage capacity can be ordered and added to the DS8800 through disk drive sets. A disk drive set includes 16 disk drive modules (DDM) of the same capacity and spindle speed (RPM). DS8800 disk drive modules are available in the following varieties:

- ▶ Small Form Factor (SFF) Serial Attached SCSI (SAS) DDMs without encryption
 - 146 GB, 15 K RPM
 - 450 GB, 10 K RPM
 - 600 GB, 10 K RPM
- ▶ Small Form Factor (SFF) Serial Attached SCSI (SAS) DDMs with encryption
 - 450 GB, 10 K RPM
 - 600 GB, 10 K RPM
- ▶ Small Form Factor (SFF) Solid State DDMs (SSD)
 - 300 GB

Tip: Full Disk Encryption (FDE) drives can only be added to a DS8800 that was initially ordered with FDE drives installed. See *IBM System Storage DS8700 Disk Encryption Implementation and Usage Guidelines*, REDP-4500, for more information about full disk encryption restrictions.

The disk drives are installed in Storage Enclosures (SEs). A storage enclosure interconnects the DDMs to the controller cards that connect to the device adapters. Each storage enclosure contains a redundant pair of controller cards. Each of the controller cards also has redundant trunking. Figure 18-1 illustrates a Storage Enclosure.

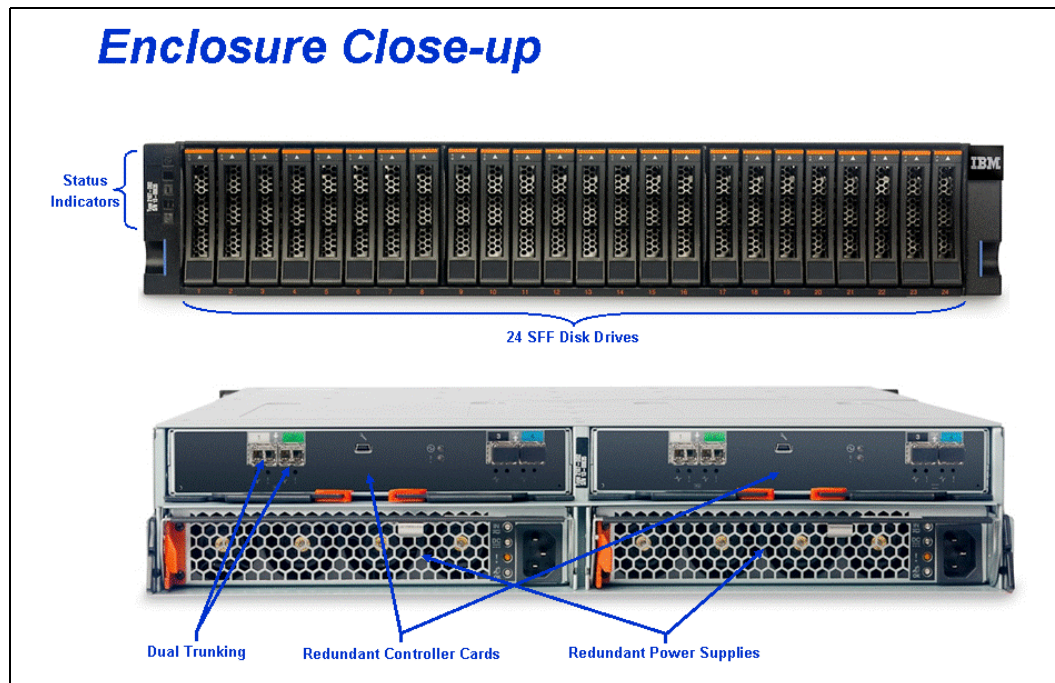


Figure 18-1 DS8800 Storage Enclosure

Storage enclosures are always installed in pairs, with one enclosure in the upper part of the unit and one enclosure in the lower part. A storage enclosure pair can be populated with one, two, or three disk drive sets (16, 32, or 48 DDMs). All DDMs in a disk enclosure pair must be of the same type (capacity and speed). Most commonly, each storage enclosure is shipped full with 24 DDMs, meaning that each pair has 48 DDMs. If a disk enclosure pair is populated with only 16 or 32 DDMs, disk drive filler modules called *baffles* are installed in the vacant DDM slots. This is to maintain the correct cooling airflow throughout the enclosure.

Each storage enclosure attaches to two device adapters (DAs). The DAs are the RAID adapter cards that connect the CECs to the DDMs. The DS8800 DA cards are always installed as a redundant pair, so they are referred to as *DA pairs*.

Physical installation and testing of the device adapters, storage enclosure pairs, and DDMs are performed by your IBM service representative. After the additional capacity is added successfully, the new storage appears as additional unconfigured array sites.

You might need to obtain new license keys and apply them to the storage image before you start configuring the new capacity. See Chapter 10, “IBM System Storage DS8000 features and license keys” on page 251 for more information. You cannot create ranks using the new capacity if this causes your machine to exceed its license key limits. Be aware that applying increased feature activation codes is a concurrent action, but a license reduction or deactivation is often a disruptive action.

Tip: Special restrictions in terms of placement and intermixing apply when adding Solid-State Drives.

18.1.1 Installation order of upgrades

Individual machine configurations vary, so it is not possible to give an exact pattern for the order in which every storage upgrade will be installed. This is because it is possible to order a machine with multiple underpopulated storage enclosures (SEs) across the device adapter (DA) pairs. This is done to allow future upgrades to be performed with the fewest physical changes. Note, however, that all storage upgrades are concurrent, in that adding capacity to a DS8800 does not require any downtime.

As a general rule, when adding capacity to a DS8800, storage hardware is populated in the following order:

1. DDMs are added to underpopulated enclosures. Whenever you add 16 DDMs to a machine, eight DDMs are installed into the upper storage enclosure and eight into the lower storage enclosure. If you add a complete 48 pack, then 24 are installed in the upper storage enclosure and 24 are installed in the lower storage enclosure.
2. After the first storage enclosure pair on a DA pair is fully populated with DDMs (48 DDMs total), the next two storage enclosures to be populated will be connected to a new DA pair. The DA cards are installed into the I/O enclosures that are located at the bottom of the racks. They are not located in the storage enclosures.
3. After each DA pair has two fully populated storage enclosure pairs (96 DDMs total), another storage enclosure pair is added to an existing storage enclosure pair.

18.1.2 Checking how much total capacity is installed

There are four DS CLI commands you can use to check how many DAs, SEs, and DDMs are installed in your DS8800. They are:

- ▶ `lsda`
- ▶ `lsstgenc1`
- ▶ `lsddm`
- ▶ `lsarraysite`

When the `-1` parameter is added to these commands, additional information is shown. In the next section, we show examples of using these commands.

For these examples, the target DS8800 has 2 device adapter pairs (total 4 DAs) and 4 fully-populated storage enclosure pairs (total 8 SEs). This means there are 128 DDMs and 16 array sites because each array site consists of 8 DDMs. In the examples, 10 of the array sites are in use, and 6 are *Unassigned* meaning that no array is created on that array site. The example system also uses full disk encryption-capable DDMs.

In Example 18-1, a listing of the device adapter cards is shown.

Example 18-1 List the device adapters

```
dsccli> lsda -1 IBM.2107-1301511
Date/Time: September 21, 2010 3:21:21 PM CEST IBM DSCLI Version: 6.6.x.xxx DS: IBM.2107-1301511
ID                               State loc                               FC Server DA pair interfcs
=====
IBM.1400-1B3-05065/R0-P1-C3 Online U1400.1B3.RJ05065-P1-C3 - 0      2      0x0230,0x0231,0x0232,0x0233
IBM.1400-1B3-05065/R0-P1-C6 Online U1400.1B3.RJ05065-P1-C6 - 0      3      0x0260,0x0261,0x0262,0x0263
IBM.1400-1B4-05066/R0-P1-C3 Online U1400.1B4.RJ05066-P1-C3 - 1      3      0x0330,0x0331,0x0332,0x0333
IBM.1400-1B4-05066/R0-P1-C6 Online U1400.1B4.RJ05066-P1-C6 - 1      2      0x0360,0x0361,0x0362,0x0363
```

In Example 18-2, a listing of the storage enclosures is shown.

Example 18-2 List the storage enclosures

```
dsccli> lsstgenc1 IBM.2107-1301511
Date/Time: September 21, 2010 3:25:09 PM CEST IBM DSCLI Version: 6.6.x.xxx DS: IBM.2107-1301511
ID                               Interfaces                               interadd stordev cap (GB) RPM
=====
IBM.2107-D02-00086/R3-S15 0x0060,0x0130,0x0061,0x0131      0x1      24 146.0      15000
IBM.2107-D02-00255/R2-S07 0x0460,0x0530,0x0461,0x0531      0x0      24 146.0      15000
IBM.2107-D02-00271/R3-S13 0x0460,0x0530,0x0461,0x0531      0x1      24 146.0      15000
IBM.2107-D02-00327/R2-S05 0x0630,0x0760,0x0631,0x0761      0x1      24 146.0      15000
IBM.2107-D02-00363/R2-S06 0x0632,0x0762,0x0633,0x0763      0x1      24 146.0      15000
```

In Example 18-3, a listing of the storage drives is shown. Because there are 128 DDMs in the example machine, only a partial list is shown here.

Example 18-3 List the DDMs (abbreviated)

```
dsccli> lsddm IBM.2107-75NR571
Date/Time: September 21, 2010 3:27:58 PM CEST IBM DSCLI Version: 6.6.x.xxx DS: IBM.2107-75NR571
ID                               DA Pair dkcap (10^9B) dkuse                               arsite State
=====
IBM.2107-D02-00769/R1-P1-D1 2                               450.0 array member    S5      Normal
IBM.2107-D02-00769/R1-P1-D2 2                               450.0 array member    S5      Normal
IBM.2107-D02-00769/R1-P1-D3 2                               450.0 spare required  S1      Normal
IBM.2107-D02-00769/R1-P1-D4 2                               450.0 array member    S2      Normal
IBM.2107-D02-00769/R1-P1-D5 2                               450.0 array member    S2      Normal
```

IBM.2107-D02-00769/R1-P1-D6	2	450.0	array member	S1	Normal
IBM.2107-D02-00769/R1-P1-D7	2	450.0	array member	S6	Normal
IBM.2107-D02-00769/R1-P1-D8	2	450.0	array member	S6	Normal
IBM.2107-D02-00769/R1-P1-D9	2	450.0	array member	S4	Normal
IBM.2107-D02-00769/R1-P1-D10	2	450.0	array member	S3	Normal
IBM.2107-D02-00769/R1-P1-D11	2	450.0	array member	S1	Normal
IBM.2107-D02-00769/R1-P1-D12	2	450.0	array member	S3	Normal

In Example 18-4, a listing of the array sites is shown.

Example 18-4 List the array sites

```
dscli> lsarraysite -dev IBM.2107-75NR571
Date/Time: September 21, 2010 3:31:08 PM CEST IBM DSCLI Version: 6.6.x.xxx DS: IBM.2107-75NR571
arsite DA Pair dkcap (10^9B) State Array
=====
S1      2      450.0 Assigned A0
S2      2      450.0 Assigned A1
S3      2      450.0 Assigned A2
S4      2      450.0 Assigned A3
S5      2      450.0 Assigned A4
S6      2      450.0 Assigned A5
S7      0      600.0 Assigned A6
S8      0      600.0 Assigned A7
S9      0      600.0 Assigned A8
S10     0      600.0 Assigned A9
S11     0      600.0 Assigned A10
S12     0      600.0 Assigned A11
```

18.2 Using Capacity on Demand (CoD)

IBM offers Capacity on Demand (CoD) solutions that are designed to meet the changing storage needs of rapidly growing e-businesses. This section discusses CoD on the DS8800.

There are various rules about CoD and these are explained in *IBM System Storage DS8000 Introduction and Planning Guide*, GC35-0515. This section explains aspects of implementing a DS8800 that has CoD disk packs.

18.2.1 What is Capacity on Demand

The Standby CoD offering is designed to provide you with the ability to tap into additional storage and is particularly attractive if you have rapid or unpredictable growth, or if you simply want extra storage to be there when you need it.

In many database environments, it is not unusual to have rapid growth in the amount of disk space required for your business. This can create a problem if there is an unexpected and urgent need for disk space and no time to create a purchase order or wait for the disk to be delivered.

With this offering, up to six Standby CoD disk drive sets (96 disk drives) can be factory-installed or field-installed into your system. To activate, you logically configure the disk drives for use. This is a nondisruptive activity that does not require intervention from IBM. Upon activation of any portion of a Standby CoD disk drive set, you must place an order with IBM to initiate billing for the activated set. At that time, you can also order replacement CoD disk drive sets.

This offering allows you to purchase licensed functions based upon your machine's physical capacity, excluding unconfigured Standby CoD capacity. This can help improve your cost of ownership, because your extent of IBM authorization for licensed functions can grow at the same time you need your disk capacity to grow.

Contact your IBM representative to obtain additional information regarding Standby CoD offering terms and conditions.

18.2.2 Determining if a DS8800 has CoD

A common question is how to determine if a DS8800 has CoD disks installed. There are two important indicators that you need to check for:

- ▶ Is the CoD indicator present in the Disk Storage Feature Activation (DSFA) website?
- ▶ What is the Operating Environment License (OEL) limit displayed by the **1skey** DS CLI command?

Verifying CoD on the DSFA website

The data storage feature activation (DSFA) website provides feature activation codes and license keys to technically activate functions acquired for your IBM storage products. To check for the CoD indicator on the DSFA website, you need to perform the following tasks:

1. Get the machine signature using DS CLI. Connect with the DS CLI and execute **showsi -fullid** as shown in Example 18-5. The signature is a unique value that can only be accessed from the machine. You will also need to record the Machine Type displayed and the Machine Serial Number (ending with 0).

Example 18-5 Displaying the machine signature

```
dscli> showsi -fullid IBM.2107-75NR571
Date/Time: September 21, 2010 3:35:13 PM CEST IBM DSCLI Version: 6.6.x.xxx DS:
IBM.2107-75NR571
Name          imaginary1
desc          -
ID            IBM.2107-75NR571
Storage Unit   IBM.2107-75NR570
Model         951
WWNN          5005070009FFC5D5
Signature      b828-2f64-eb24-4f17 <===== Machine Signature
State          Online
ESSNet         Enabled
Volume Group   IBM.2107-75NR571/V0
os400Serial    5D5
NVS Memory     2.0 GB
Cache Memory   50.6 GB
Processor Memory 61.4 GB
MTS            IBM.2421-75NR570 <=====Machine Type (2421) and S/N (75NR570)
numegsupported 1
```

2. Now log on to the DSFA website at:

<http://www.ibm.com/storage/dsfa>

Select **IBM System Storage DS8000 Series** from the DSFA start page. The next window requires you to choose the Machine Type and then enter the serial number and signature, as shown in Figure 18-2.

The screenshot shows a web interface for selecting a DS8000 series machine. The top navigation bar includes links for Home, Solutions, Services, Products, Support & downloads, and My IBM. A left sidebar contains a menu with options: Data storage feature activation, Select DS6000 series machine, Select DS8000 series machine (highlighted), View machine summary, Manage activations, Retrieve activation codes, Assign function authorization, View authorization details, Select TS series LTO machine, and Help. The main content area is titled 'Select DS8000 series machine' and contains instructions: 'Enter your machine information. You can find this information in the DS Sto Manager application under storage unit properties.' and 'You must complete all fields marked with an asterisk (*).'. Below the instructions are three input fields: 'Machine type: *' with a dropdown menu showing options 2107, 2421 (selected), 2422, 2423, and 2424; 'Serial number: *' with a text input field; and 'Machine signature: *' with a text input field. A 'Submit' button is located at the bottom left of the form.

Figure 18-2 DSFA machine specifics

On the **View Authorization Details** window, the feature code 0901 Standby CoD indicator is shown for DS8800 installations with Capacity on Demand. This is illustrated in Figure 18-3 on page 452. If instead you see 0900 Non-Standby CoD, it means that the CoD feature has *not* been ordered for your machine.

The screenshot shows the IBM System Storage website interface. The main navigation bar includes links for Home, Solutions, Services, Products, Support & downloads, and My IBM. The left sidebar contains a menu for IBM System Storage, with options like Data storage feature activation, Select DS6000 series machine, Select DS8000 series machine, View machine summary, Manage activations, Retrieve activation codes, Assign function authorization, View authorization details (selected), Select TS series LTO machine, and Help. Below the sidebar, there are related links for IBM System Storage support feedback and Electronic service call.

View authorization details

IBM 2421 Model 941
Serial number 75-PB930

This page displays information for licensed function indicators and authorizations present on your machine.

Licensed function indicator features

The licensed function indicator feature codes displayed below are used for monthly maintenance billing purposes only and are not used for the creation of your feature activation codes.

An indicator feature code will also be displayed if your machine is participating in the DS8000 series Standby Capacity on Demand program.

Feature	Qty	Description
0901	1	Standby CoD indicator
7035	1	OEL 50TB indicator
7255	1	PTC 50TB indicator

Function authorization feature codes

The function authorization feature codes displayed below define the functions that you can retrieve feature activation codes for.

Operating environment

IBM 2396 Model LFA Serial number 75-12605

Feature	Qty	Description

Figure 18-3 Verifying CoD using DSFA

Verifying CoD on the DS8800

Normally, new features or feature limits are activated using the DS CLI **applykey** command. However, CoD does not have a discrete key. Instead, the CoD feature is installed as part of the Operating Environment License (OEL) key. The interesting thing is that an OEL key that activates CoD will change the feature limit from the limit that you have paid for, to the *largest* possible number.

In Example 18-6 on page 453, you can see how the OEL key is changed. The machine in this example is licensed for 80 TB of OEL, but actually has 82 TB of disk installed because it has 2 TB of CoD disks. However, if you attempt to create ranks using the final 2 TB of storage, the command will fail because it exceeds the OEL limit. After a new OEL key with CoD is installed, the OEL limit will increase to an enormous number (9.9 million TB). This means that rank creation will succeed for the last 2 TB of storage.

Example 18-6 Applying an OEL key that contains CoD

```
dsccli> lskey IBM.2107-75ABCD1
Date/Time: October 21, 2009 2:47:26 PM MST IBM DSCLI Version: 6.5.0.xxx DS: IBM.2107-75ABCD1
Activation Key          Authorization Level (TB) Scope
=====
Operating environment (OEL)    80.3                All

dsccli> applykey -key 1234-5678-9ABC-DEF0-1234-5678-9ABC-DEF0 IBM.2107-75ABCD1
Date/Time: October 21, 2009 2:47:26 PM MST IBM DSCLI Version: 6.5.0.xxx DS: IBM.2107-75ABCD1
CMUC00199I applykey: Licensed Machine Code successfully applied to storage image IBM.2107-75ABCD1

dsccli> lskey IBM.2107-75ABCD1
Date/Time: October 21, 2009 2:47:26 PM MST IBM DSCLI Version: 6.5.0.sss DS: IBM.2107-75ABCD1
Activation Key          Authorization Level (TB) Scope
=====
Operating environment (OEL)    9999999           All
```

18.2.3 Using the CoD storage

In this section, we review the tasks required to start using CoD storage.

CoD array sites

If CoD storage is installed, it will be a maximum of 96 CoD disk drives. Because 16 drives make up a drive set, a better use of terminology is to say a machine can have up to 6 drive sets of CoD disk. Because 8 drives are used to create an array site, this means that a maximum of 12 array sites of CoD can potentially exist in a machine. If a machine has, for example, 384 disk drives installed, of which 96 disk drives are CoD, then there are a total of 48 array sites, of which 12 are CoD. From the machine itself, there is no way to tell how many of the array sites in a machine are CoD array sites as opposed to array sites you can start using right away. During the machine order process, this must be clearly understood and documented.

Which array sites are the CoD array sites

Given a sample DS8800 with 48 array sites, of which 8 represent CoD disks, the client should configure only 40 of the 48 array sites. This assumes that all the disk drives are the same size. It is possible to order CoD drive sets of different sizes. In this case, you would need to understand how many of each size have been ordered and ensure that the correct number of array sites of each size are left unused until they are needed for growth.

How to start using the CoD array sites

Use the standard DS CLI (or DS GUI) commands to configure storage starting with **mkarray**, then **mkrank**, and so on. After the ranks are members of an Extent Pool, then volumes can be created. See Chapter 13, “Configuration using the DS Storage Manager GUI” on page 309, and Chapter 14, “Configuration with the DS Command-Line Interface” on page 363 for more information about this topic.

What if you accidentally configure a CoD array site

Given the sample DS8800 with 48 array sites, of which 8 represent CoD disks, if you accidentally configure 41 array sites but did not intend to start using the CoD disks yet, then use the **rmarray** command immediately to return that array site to an *unassigned* state. If volumes have been created and those volumes are in use, then you have started to use the CoD arrays and should contact IBM to inform IBM that the CoD storage is now in use.

What you do after the CoD array sites are in use

After you have started to use the CoD array sites (and remember that IBM requires that a Standby CoD disk drive set must be activated within a twelve-month period from the date of installation; all such activation is permanent), then contact IBM so that the CoD indicator can be removed from the machine. You must place an order with IBM to initiate billing for the activated set.

At that time, you can also order replacement Standby CoD disk drive sets. If new CoD disks are ordered and installed, then a new OEL key will also be issued and should be applied immediately. If no more CoD disks are desired, or the DS8800 has reached maximum capacity, then an OEL key will be issued to reflect that CoD is no longer enabled on the machine.



A

Tools and service offerings

This appendix provides information about the tools that are available to help you when planning, managing, migrating, and analyzing activities with your DS8800. In this appendix, we also reference the sites where you can find information about the service offerings that are available from IBM to help you in several of the activities related to the DS8800 implementation.

Capacity Magic

Because of the additional flexibility and configuration options storage subsystems provide, it becomes a challenge to calculate the raw and net storage capacity of disk subsystems such as the DS8800. You have to invest considerable time, and you need an in-depth technical understanding of how spare and parity disks are assigned. You also need to consider the simultaneous use of disks with different capacities and configurations that deploy RAID 5, RAID 6, and RAID 10.

Capacity Magic can do the physical (raw) to effective (net) capacity calculations automatically, taking into consideration all applicable rules and the provided hardware configuration (number and type of disk drive sets).

Capacity Magic is designed as an easy-to-use tool with a single, main interface. It offers a graphical interface that allows you to enter the disk drive configuration of a DS8800 and other IBM subsystems, the number and type of disk drive sets, and the RAID type. With this input, Capacity Magic calculates the raw and net storage capacities. The tool also has functionality that lets you display the number of extents that are produced per rank, as shown in Figure A-1.

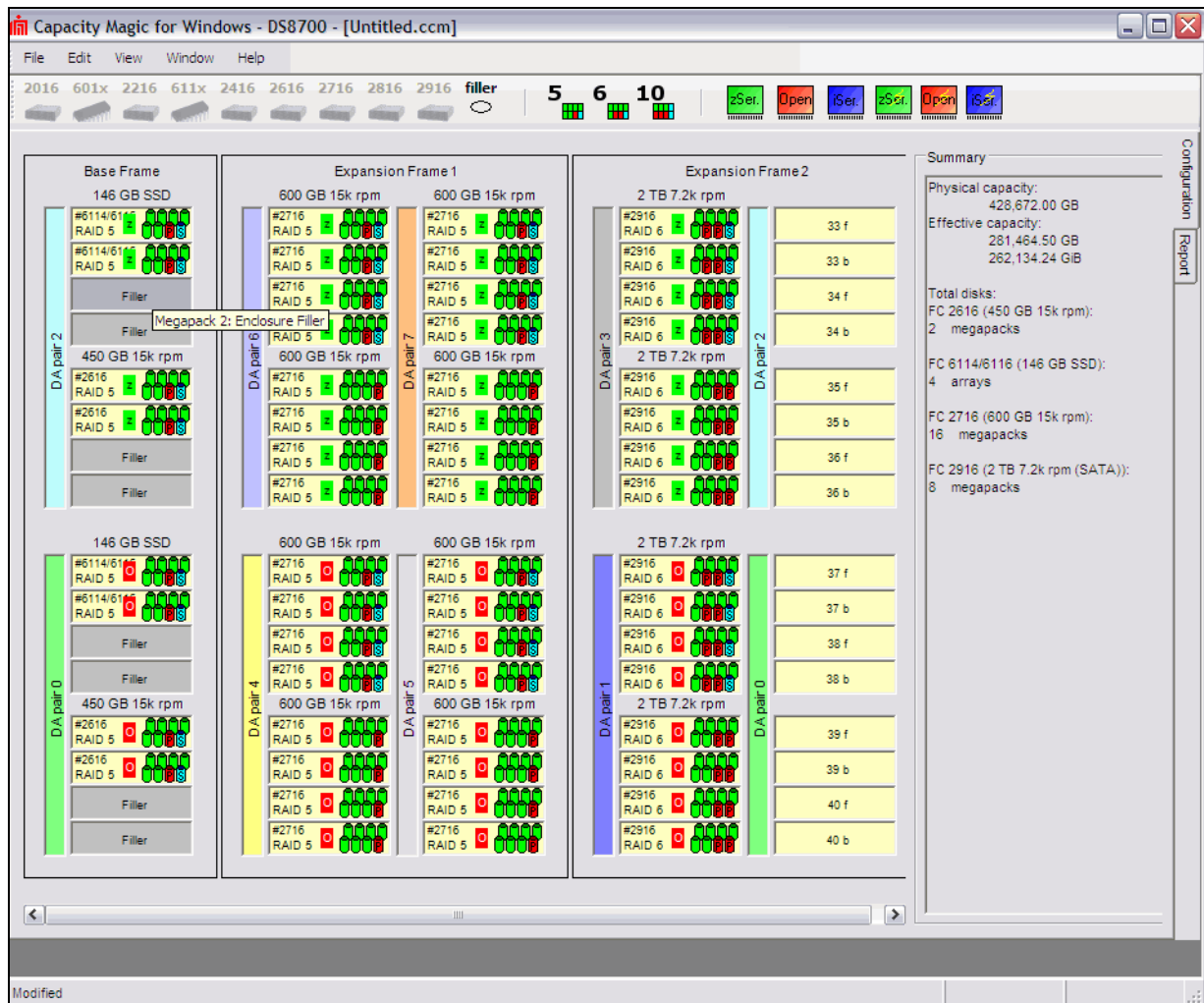


Figure A-1 Configuration window

Figure A-1 on page 456 shows the configuration window that Capacity Magic provides for you to specify the desired number and type of disk drive sets.

Figure A-2 shows the resulting output report that Capacity Magic produces. This report is also helpful in planning and preparing the configuration of the storage in the DS8800, because it also displays extent count information.

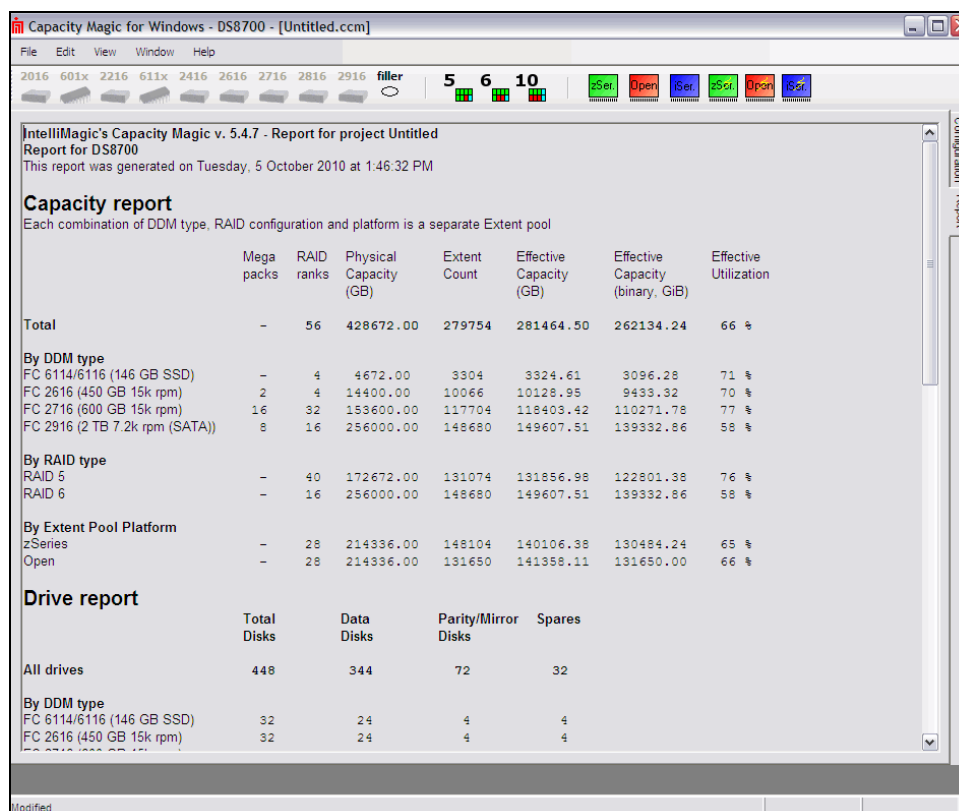


Figure A-2 Capacity Magic output report

Tip: Capacity Magic is a tool used by IBM and IBM Business Partners to model disk storage subsystem effective capacity as a function of physical disk capacity to be installed. Contact your IBM Representative or IBM Business Partner to discuss a Capacity Magic study.

Disk Magic

Disk Magic is a Windows-based disk subsystem performance modeling tool. It supports disk subsystems from multiple vendors, but it offers the most detailed support for IBM subsystems. Currently Disk Magic supports modelling to advanced-function disk subsystems, such as the DS8000 series, DS6000™, ESS, DS4000®, DS5000, N-Series and the SAN Volume Controller.

A critical design objective for Disk Magic is to minimize the amount of input that you must enter, while offering a rich and meaningful modeling capability. The following list provides several examples of what Disk Magic can model, but it is by no means complete:

- ▶ Move the current I/O load to another disk subsystem model.
- ▶ Merge the current I/O load of multiple disk subsystems into a single DS8700.

- ▶ Insert a SAN Volume Controller in an existing disk configuration.
- ▶ Increase the current I/O load.
- ▶ Implement a storage consolidation.
- ▶ Increase the disk subsystem cache size.
- ▶ Change to larger capacity disk drives.
- ▶ Change to higher disk rotational speed.
- ▶ Upgrade from ESCON to FICON host adapters.
- ▶ Upgrade from SCSI to Fibre Channel host adapters.
- ▶ Increase the number of host adapters.
- ▶ Use fewer or more Logical Unit Numbers (LUNs).
- ▶ Activate Metro Mirror.
- ▶ Activate z/OS Global Mirror.
- ▶ Activate Global Mirror.

With the availability of SSD, Disk Magic perform support modelling when SSD ranks are included in the configuration. In a z/OS environment, Disk Magic can provide an estimation of which volumes are good SSD candidates and migrate those volumes to SSD in the model. In an Open System environment, Disk Magic can model the SSD on a server basis.

Tip: Disk Magic is a tool used by IBM and IBM Business Partners to model disk storage subsystem performance. Contact your IBM Representative or IBM Business Partner to discuss a Disk Magic study.

HyperPAV analysis

Traditional aliases allow you to simultaneously process multiple I/O operations to the same logical volume. The question is, how many aliases do you need to assign to the LCUs in your DS8000?

It is difficult to predict the ratio of aliases to base addresses required to minimize IOSQ time. If the ratio is too high, this limits the amount of physical volumes that can be addressed, due to the 64 K addressing limit. If the ratio is too small, then you might see high IOSQ times, which will impact the business service commitments.

HyperPAV can help improve performance by reducing the IOSQ Time and also help in reducing the number of aliases required in an LCU, which would free up more addresses to be used as base-addresses.

To estimate how many aliases are needed, a HyperPAV analysis can be performed using SMF records 70 through 78. The analysis results provide guidance about how many aliases are required. This analysis can be performed against IBM and non-IBM disk subsystems.

Tip: Contact your IBM Representative or IBM Business Partner to discuss a HyperPAV study.

FLASHDA

The FLASHDA is a tool written in SAS that can help with deciding which datasets or volumes are best candidates to be migrated to SSD from HDD.

The prerequisite to use this tool are APARs OA25688 and OA25559, which will report DISC Time separately by Read I/O and Write I/O. The tool uses SMF 42 subtype 6 and SMF 74

subtype 5 records and provides a list by dataset, showing the amount of accumulated DISC Time for the Read I/O operations during the time period selected.

If complete SMF records 70 through 78 are also provided, the report can be tailored to show the report by dataset by each disk subsystem. It can also show the report by volume by disk subsystem. If you are running z/OS v1R10, you can also include the number of cylinders used by volume.

Figure A-3 lists the output of the FLASHDA tool. It shows the Dataset name with the Address and Volser where the dataset resides and the Total DISC Time in milliseconds for all Read I/Os. This list is sorted in descending order to show which datasets would benefit the most when moved to an SSD rank.

Address	Volser	Dataset name	Total Rd DISC
C1D2	I10YY5	IMS10.DXX.WWXPRT11	5,184,281
2198	XAGYAA	DB2PAG.DSNDBD.XX10X97I.CGPL.I0001.YY01	3,530,406
783E	XA2Y58	DB2PA2.DSNDBD.XX40X97E.RESB.I0001.YY02	2,978,921
430A	Y14Y3S	DB214.DSNDBD.ZZXDSSENT.WWXSC.I0001.YY01	2,521,349
21B2	XAGYC6	DB2PAG.DSNDBD.XX40XTKL.MSEG.J0001.YY06	2,446,672
7A10	XA2Y76	DB2PA2.DSNDBD.XX40X97E.RESB.I0001.YY03	2,123,498
7808	XA2Y04	DB2PA2.DSNDBD.XX40X97E.RESB.I0001.YY01	1,971,660
2A13	X39Y60	DB2X39.DSNDBD.YY30X956.MONX.J0001.YY01	1,440,200
2B60	X39Y12	DB2X39.DSNDBD.YY40X975.EQUI.J0001.YY02	1,384,468
C1D2	I10YY5	IMS10.DJX.WWJIFPC1	1,284,444
783B	XA2Y55	DB2PA2.DSNDBD.XX40X97E.RESB.I0001.YY04	1,185,571
2B5A	X39Y06	DB2X39.DSNDBD.YY40X975.EQUI.J0001.YY01	1,016,916

Figure A-3 FLASHDA output

The next report, shown in Figure A-4, is based on the preceding FLASHDA output and from information extracted from the SMF records. This report shows the ranking of the Total Read DISC Time in milliseconds by volume. It also shows the number of cylinders defined for that volume and the serial number of the disk subsystem (DSS) where that volume resides.

Address	Volser	Total Rd DISC	#cyls	DSS
C1D2	I10YY5	6,592,229	32760	IBM-KLZ01
2198	XAGYAA	3,608,052	65520	IBM-MN721
783E	XA2Y58	3,032,377	65520	IBM-OP661
430A	Y14Y3S	2,654,083	10017	IBM-KLZ01
21B2	XAGYC6	2,648,126	65520	IBM-MN721
7A10	XA2Y76	2,389,512	65520	IBM-OP661
7808	XA2Y04	2,102,741	65520	IBM-OP661
22AA	XAGY84	1,458,696	65520	IBM-MN721
2193	XAGYA5	1,455,057	65520	IBM-MN721
2A13	X39Y60	1,444,708	65520	ABC-04749
21B5	XAGYC9	1,429,231	65520	IBM-MN721
2B60	X39Y12	1,387,409	65520	ABC-04749

Figure A-4 Total Read DISC Time report by Volume

Using the report by dataset, you can select the datasets that are used by your critical applications and migrate them to the SSD ranks.

If you use the report by volume, you can decide how many volumes you want to migrate to SSD, and calculate how many SSD ranks are needed to accommodate the volumes that you selected. A Disk Magic study can be performed to see how much performance improvement can be achieved by migrating those volumes to SSD.

Tip: Contact your IBM Representative or IBM Business Partner to discuss a FLASHDA study.

IBM i SSD Analyzer Tool

The SSD Analyzer Tool is designed to help you determine if SSDs could help improve performance on your IBM i system(s). The tool works with the performance data that is collected on your system and works with releases V5R4 and V6R1.

Figure A-5 shows the detailed analysis report by job name, sorted descending by Disk Read Wait Total Seconds. This list can be used to select the data used by the job that would get the highest benefit when migrated to an SSD media.

SSD Data Analysis - Jobs Sorted by Disk Read Time				
Performance member I289130103 in library SB01				
Time period from 2009-10-16-13.01.05.000000 to 2009-10-16-14.00.00.000000				
Job Name	CPU Total Seconds	Disk Read Wait Total Seconds	Disk Read Wait Average Seconds	Disk Read Wait /CPU
FOF00100/STORADNEN/460669	30.426	3,468.207	.006636	114
FOF00000/STORADNEN/516129	33.850	3,461.419	.006237	102
FOF17000/STORADNEN/387280	48.067	3,427.064	.006548	71
FOF17100/STORADNEN/499676	71.951	3,395.609	.007191	47
FOF14000/STORADNEN/487761	33.360	3,295.738	.006799	99
FOF00000/STORADNEN/516028	78.774	2,962.103	.007409	38
FOF00200/STORADNEN/516000	79.025	2,961.518	.007441	37
FOF00400/STORADNEN/516010	78.640	2,957.033	.007412	38

Figure A-5 IBM i SSD Analyzer Tool - DETAIL report

Tip: Contact your IBM Representative or IBM Business Partner to discuss an IBM i SSD analysis.

IBM Tivoli Storage Productivity Center

IBM Tivoli Productivity Center (previously known as the TotalStorage Productivity Center) is a standard software package for managing complex storage environments. One subcomponent of this package is IBM Tivoli Storage Productivity Center for Disk (TPC for Disk), which is designed to help reduce the complexity of managing SAN storage devices by allowing administrators to configure, manage, and performance monitor storage from a single console.

TPC for Disk is designed to:

- ▶ Configure multiple storage devices from a single console
- ▶ Monitor and track the performance of SAN-attached Storage Management Interface Specification (SMI-S) compliant storage devices
- ▶ Enable proactive performance management by setting performance thresholds based on performance metrics and the generation of alerts

IBM Tivoli Productivity Center for Disk centralizes the management of networked storage devices that implement the SNIA SMI-S specification, which includes the IBM System

Storage DS family, XIV®, N series, and SAN Volume Controller (SVC). It is designed to help reduce storage management complexity and costs while improving data availability, centralizing management of storage devices through open standards (SMI-S), enhancing storage administrator productivity, increasing storage resource utilization, and offering proactive management of storage devices. IBM Tivoli Productivity Center for Disk offers the ability to discover storage devices using Service Location Protocol (SLP) and provides the ability to configure devices, in addition to gathering event and errors logs and launching device-specific applications or elements.

For more information, see *Managing Disk Subsystems using IBM TotalStorage Productivity Center*, SG24-7097. Also, refer to the following address:

<http://www.ibm.com/servers/storage/software/center/index.html>

IBM Certified Secure Data Overwrite

STG Lab Services offers IBM Certified Secure Data Overwrite service for the DS8000 series, and ESS models 800 and 750. This offering is meant to overcome the following issues:

- ▶ Deleted data does not mean gone forever. Usually, deleted means that the pointers to the data are invalidated and the space can be reused. Until the space is reused, the data remains on the media and what remains can be read with the right tools.
- ▶ Regulations and business prudence require that the data actually be removed when the media is no longer available.

The service executes a multipass overwrite of the data disks in the storage system:

- ▶ It operates on the entire box.
- ▶ It is three pass overwrite, which is compliant with the DoD 5220.20-M procedure for purging disks.
 - Writes all sectors with zeros.
 - Writes all sectors with ones.
 - Writes all sectors with a pseudo-random pattern.
 - Each pass reads back a random sample of sectors to verify the writes are done.
- ▶ There is a fourth pass of zeros with InitSurf.
- ▶ IBM also purges client data from the server and HMC disks.

Certificate of completion

After the overwrite process has been completed, IBM delivers a complete report containing:

- ▶ A certificate of completion listing:
 - The serial number of the systems overwritten.
 - The dates and location the service was performed.
 - The overwrite level.
 - The names of the engineers delivering the service and compiling the report.
- ▶ A description of the service and the report
- ▶ On a per data drive serial number basis:
 - The G-list prior to overwrite.
 - The pattern run against the drive.

- The success or failure of each pass.
- The G-list after the overwrite.
- Whether the overwrite was successful or not for each drive.

Figure A-6 shows a sample report by drive.

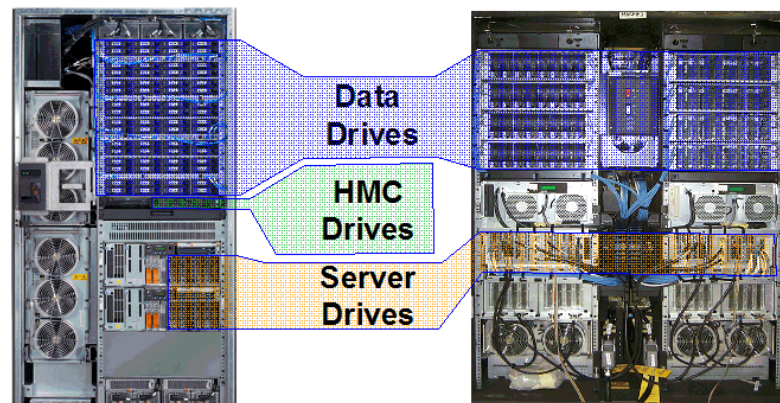
DS8000 PROD1 (Serial # 12345)								
Disk Drive	Disk Drive Serial# (Electronic)	DDM Barcode Serial# (Visible)	Overwrite Status	Sector Defects at Start	Sector Defects After 1st Pass	Sector Defects After 2nd Pass	Sector Defects After 3rd Pass	Sector Defects After 4th Pass
pdisk0	3HY6LFKZ	350A8459	Successful	0	0	0	0	0
pdisk1	3HY6N92M	350B055D	Successful	0	0	0	0	0
pdisk2	3HY6FV79	350B0756	Successful	0	0	0	0	0
pdisk3	3HY6N1NA	350B075C	Successful	0	0	0	0	0
pdisk4	3HY6MAQ0	350B0D07	Successful	0	0	0	0	0
pdisk5	3HY6P1E5	350B0D3D	Successful	0	0	0	0	0
pdisk6	3HY6P2PG	350B0D79	Successful	0	0	0	0	0
pdisk7	3HY6P2QB	350B0D85	Successful	0	0	0	0	0
pdisk8	3HY6P2PR	350B0D86	Successful	0	0	0	0	0
pdisk9	3HY6P16J	350B0D86	Successful	0	0	0	0	0
pdisk10	3HY6P2BX	350B0DA9	Successful	0	0	0	0	0
pdisk11	3HY6P3LF	350B0DCF	Successful	0	0	0	0	0
pdisk12	3HY6P3L2	350B0DDC	Successful	21	21	21	21	21
pdisk13	3HY6M5ZX	350B0DDD	Successful	0	0	0	0	0
pdisk14	3HY6P3K6	350B0DDE	Successful	0	0	0	0	0
pdisk15	3HY6P3JW	350B0DE2	Successful	0	0	0	0	0
pdisk16	3HY6NNZF	350B0DE4	Successful	0	0	0	0	0
pdisk18	3HY6NPQM	350B0E0C	Successful	0	0	0	0	0
pdisk19	3HY6NYZ0	350B0E0D	Failed	0	0	–	–	–
pdisk20	3HY6P484	350B0E1F	Successful	0	0	0	0	0
pdisk21	3HY6P4NZ	350B0E72	Successful	0	0	0	0	0
pdisk22	3HY6P5JB	350B0E73	Successful	0	0	0	0	0
pdisk23	3HY6SVDT	350BF7DC	Successful	0	0	0	0	0
pdisk24	3HY6V1CV	350C91A	Successful	0	0	0	0	0

Figure A-6 Sample report by drive

Drives erased

As a result of the erase service, all disks in the storage system are erased. Figure A-7 on page 463 shows all the drives that are covered by the Secure Data Overwrite Service.

Drives erased and their contents



	Data Drives	HMC Drives	Server Drives
DS8000	Client Data	State Save	State Save NVS
ESS800	Client Data	N/A	State Save
ESS750	Client Data	N/A	State Save

Figure A-7 Drives erased and their contents

IBM Global Technology Services: service offerings

IBM can assist you in deploying IBM System Storage DS8800 subsystems, IBM Tivoli Productivity Center, and SAN Volume Controller solutions. IBM Global Technology Services has the right knowledge and expertise to reduce your system and data migration workload, as well as the time, money, and resources needed to achieve a system-managed environment.

For more information about available services, contact your IBM representative or IBM Business Partner, or visit the following addresses:

<http://www.ibm.com/services/>
<http://www.ibm.com/servers/storage/services/disk.html>

For details about available IBM Business Continuity and Recovery Services, contact your IBM Representative or visit the following address:

<http://www.ibm.com/services/continuity>

For details about educational offerings related to specific products, visit the following address:

<http://www.ibm.com/services/learning/index.html>

Select your country, and then select the product as the category.

IBM STG Lab Services: Service offerings

In addition to the IBM Global Technology Services, the STG Lab Services and Training teams are set up to assist you with one-off, client-tailored solutions and services that will help you in

your daily work with IBM Hardware and Software components. For more information about this topic, go to the following address:

<http://www.ibm.com/systems/services/labservices/>

Abbreviations and acronyms

AAL	Arrays Across Loops	EB	Exabyte
AC	Alternating Current	ECC	Error Checking and Correction
AL-PA	Arbitrated Loop Physical Addressing	EDF	Extended Distance FICON
AMP	Adaptive Multistream Prefetching	EEH	Enhanced Error Handling
API	Application Programming Interface	EPO	Emergency Power Off
ASCII	American Standard Code for Information Interchange	EPOW	Emergency Power Off Warning
ASIC	Application Specific Integrated Circuit	ESCON	Enterprise Systems Connection
B2B	Business-to-Business VPN	ESS	Enterprise Storage Server
BBU	Battery Backup Unit	ESSNI	Enterprise Storage Server Network Interface
CEC	Central Electronics Complex	FATA	Fibre Channel Attached Technology Adapter
CG	Consistency Group	FB	Fixed Block
CHPID	Channel Path ID	FC	Flash Copy
CIM	Common Information Model	FCAL	Fibre Channel Arbitrated Loop
CKD	Count Key Data	FCIC	Fibre Channel Interface Card
CoD	Capacity on Demand	FCP	Fibre Channel Protocol
CPU	Central Processing Unit	FCSE	FlashCopy Space Efficient
CSDO	Certified Secure Data Overwrite	FDE	Full Disk Encryption
CUIR	Control Unit Interface Reconfiguration	FFDC	First Failure Data Capture
DA	Device Adapter	FICON	Fiber Connection
DASD	Direct Access Storage Device	FIR	Fault Isolation Register
DC	Direct Current	FRR	Failure Recovery Routines
DDM	Disk Drive Module	FTP	File Transfer Protocol
DFS	Distributed File System	GB	Gigabyte
DFW	DASD Fast Write	GC	Global Copy
DHCP	Dynamic Host Configuration Protocol	GM	Global Mirror
DMA	Direct Memory Access	GSA	Global Storage Architecture
DMZ	De-Militarized Zone	GUI	Graphical User Interface
DNS	Domain Name System	HA	Host Adapter
DPR	Dynamic Path Reconnect	HACMP	High Availability Cluster Multi-Processing
DPS	Dynamic Path Selection	HBA	Host Bus Adapter
DSCLMCLI	Data Storage Common Information Model Command-Line Interface	HCD	Hardware Configuration Definition
DSCLI	Data Storage Command-Line Interface	HMC	Hardware Management Console
DSFA	Data Storage Feature Activation	HSM	Hardware Security Module
DVE	Dynamic Volume Expansion	HTTP	Hypertext Transfer Protocol
EAV	Extended Address Volume	HTTPS	Hypertext Transfer Protocol over SSL
		IBM	International Business Machines Corporation

IKE	Internet Key Exchange	NAT	Network Address Translation
IKS	Isolated Key Server	NFS	Network File System
IOCDS	Input/Output Configuration Data Set	NIMOL	Network Installation Management on Linux
IOPS	Input Output Operations per Second	NTP	Network Time Protocol
IOSQ	Input/Output Supervisor Queue	NVRAM	Non-Volatile Random Access Memory
IPL	Initial Program Load	NVS	Non-Volatile Storage
IPSec	Internet Protocol Security	OEL	Operating Environment License
IPv4	Internet Protocol version 4	OLTP	Online Transaction Processing
IPv6	Internet Protocol version 6	PATA	Parallel Attached Technology Adapter
ITSO	International Technical Support Organization	PAV	Parallel Access Volumes
IWC	Intelligent Write Caching	PB	Petabyte
JBOD	Just a Bunch of Disks	PCI-X	Peripheral Component Interconnect Extended
JFS	Journaling File System	PCIe	Peripheral Component Interconnect Express
KB	Kilobyte	PCM	Path Control Module
Kb	Kilobit	PFA	Predictive Failure Analysis
Kbps	Kilobits per second	PHYP	POWER Systems Hypervisor
KVM	Keyboard-Video-Mouse	PLD	Power Line Disturbance
L2TP	Layer 2 Tunneling Protocol	PM	Preserve Mirror
LBA	Logical Block Addressing	PMB	Physical Memory Block
LCU	Logical Control Unit	PPRC	Peer-to-Peer Remote Copy
LDAP	Lightweight Directory Access Protocol	PPS	Primary Power Supply
LED	Light Emitting Diode	PSTN	Public Switched Telephone Network
LFU	Least Frequently Used	PTC	Point-in-Time Copy
LIC	Licensed Internal Code	RAM	Random Access Memory
LIP	Loop initialization Protocol	RAS	Reliability, Availability, Serviceability
LMC	Licensed Machine Code	RIO	Remote Input/Output
LPAR	Logical Partition	RPC	Rack Power Control
LRU	Least Recently Used	RPM	Revolutions per Minute
LSS	Logical SubSystem	RPO	Recovery Point Objective
LUN	Logical Unit Number	SAN	Storage Area Network
LVM	Logical Volume Manager	SARC	Sequential Adaptive Replacement Cache
MB	Megabyte	SATA	Serial Attached Technology Adapter
Mb	Megabit	SCSI	Small Computer System Interface
Mbps	Megabits per second	SDD	Subsystem Device Driver
MFU	Most Frequently Used	SDM	System Data Mover
MGM	Metro Global Mirror	SE	Storage Enclosure
MIB	Management Information Block	SFI	Storage Facility Image
MM	Metro Mirror		
MPIO	Multipath Input/Output		
MRPD	Machine Reported Product Data		
MRU	Most Recently Used		

SFTP	SSH File Transfer Protocol	YB	Yottabyte
SMIS	Storage Management Initiative Specification	ZB	Zettabyte
SMP	Symmetric Multiprocessor	zHPF	High Performance FICON for z
SMS	Storage Management Subsystem	zIIP	z9 Integrated Information Processor
SMT	Simultaneous Multithreading		
SMTP	Simple Mail Transfer Protocol		
SNIA	Storage Networking Industry Association		
SNMP	Simple Network Monitoring Protocol		
SOI	Silicon on Insulator		
SP	Service Processor		
SPCN	System Power Control Network		
SPE	Small Programming Enhancement		
SRM	Storage Resource Management		
SSD	Solid State Drive		
SSH	Secure Shell		
SSIC	System Storage Interoperation Center		
SSID	Subsystem Identifier		
SSL	Secure Sockets Layer		
SSPC	System Storage Productivity Center		
SVC	SAN Volume Controller		
TB	Terabyte		
TCE	Translation Control Entry		
TCO	Total Cost of Ownership		
TCP/IP	Transmission Control Protocol / Internet Protocol		
TKLM	Tivoli Key Lifecycle Manager		
TPC	Tivoli Storage Productivity Center		
TPC-BE	Tivoli Storage Productivity Center Basic Edition		
TPC-R	Tivoli Storage Productivity Center for Replication		
TPC-SE	Tivoli Storage Productivity Center Standard Edition		
UCB	Unit Control Block		
UDID	Unit Device Identifier		
UPS	Uninterruptable Power Supply		
VPN	Virtual Private Network		
VTOC	Volume Table of Contents		
WLM	Workload Manager		
WUI	Web User Interface		
WWPN	Worldwide Port Name		
XRC	Extended Remote Copy		

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 publications

For information about ordering these publications, see “How to get IBM Redbooks publications” on page 470. Note that some of the documents referenced here might be available in softcopy only.

- ▶ *IBM System Storage DS8000 Series: IBM FlashCopy SE*, REDP-4368
- ▶ *Multiple Subchannel Sets: An Implementation View*, REDP-4387
- ▶ *IBM System Storage DS8700 Disk Encryption Implementation and Usage Guidelines*, REDP-4500
- ▶ *IBM System Storage DS8000: LDAP Authentication*, REDP-4505
- ▶ *DS8000: Introducing Solid State Drives*, REDP-4522
- ▶ *DS8000 Thin Provisioning*, REDP-4554
- ▶ *A Comprehensive Guide to Virtual Private Networks, Volume I: IBM Firewall, Server and Client Solutions*, SG24-5201
- ▶ *IBM System Storage DS8000: Copy Services for IBM System z*, SG24-6787
- ▶ *IBM System Storage DS8000: Copy Services for Open Systems*, SG24-6788
- ▶ *Managing Disk Subsystems using IBM TotalStorage Productivity Center*, SG24-7097
- ▶ *DS8000 Performance Monitoring and Tuning*, SG24-7146
- ▶ *Migrating to IBM System Storage DS8000*, SG24-7432
- ▶ *IBM System Storage Productivity Center Deployment Guide*, SG24-7560-00
- ▶ *IBM Tivoli Storage Productivity Center V4.2 Release Guide*, SG24-7725-00
- ▶ *IBM System Storage DS8000: Host Attachment and Interoperability*, SG24-8887

Other publications

These publications are also relevant as further information sources. Note that some of the documents referenced here might be available in softcopy only.

- ▶ *IBM System Storage DS8000 Introduction and Planning Guide*, GC27-2297
- ▶ *DS8000 Introduction and Planning Guide*, GC35-0515
- ▶ *IBM System Storage DS: Command-Line Interface User's Guide*, GC53-1127
- ▶ *System Storage Productivity Center Software Installation and User's Guide*, SC23-8823
- ▶ *IBM System Storage Productivity Center Introduction and Planning Guide*, SC23-8824
- ▶ *IBM System Storage DS8000 User's Guide*, SC26-7915
- ▶ *IBM Systems Storage DS8000 Series: Command-Line Interface User's Guide*, SC26-7916

- ▶ *IBM System Storage DS8000 Host Systems Attachment Guide*, SC26-7917
- ▶ *System Storage Productivity Center User's Guide*, SC27-2336
- ▶ *IBM System Storage Multipath Subsystem Device Driver User's Guide*, SC30-4131
- ▶ “Outperforming LRU with an adaptive replacement cache algorithm,” by N. Megiddo and D. S. Modha, in *IEEE Computer*, volume 37, number 4, pages 58–65, 2004
- ▶ “SARC: Sequential Prefetching in Adaptive Replacement Cache” by Binny Gill, et al, *Proceedings of the USENIX 2005 Annual Technical Conference*, pages 293–308
- ▶ “AMP: Adaptive Multi-stream Prefetching in a Shared Cache” by Binny Gill, et al, in *USENIX File and Storage Technologies (FAST)*, February 13 - 16, 2007, San Jose, CA
- ▶ *VPNs Illustrated: Tunnels, VPNs, and IPSec*, by Jon C. Snader, Addison-Wesley Professional (November 5, 2005), ISBN-10: 032124544X

Online resources

These websites and URLs are also relevant as further information sources:

- ▶ IBM Disk Storage Feature Activation (DSFA) website
<http://www.ibm.com/storage/dsfa>
- ▶ Documentation for the DS8000
<http://publib.boulder.ibm.com/infocenter/dsichelp/ds8000ic/index.jsp>
- ▶ System Storage Interoperation Center (SSIC)
<http://www.ibm.com/systems/support/storage/config/ssic>
- ▶ Security Planning website
http://publib16.boulder.ibm.com/doc_link/en_US/a_doc_lib/aixbman/security/ipsec_planning.htm
- ▶ *VPN Implementation*, S1002693:
<http://www.ibm.com/support/docview.wss?rs=1114&uid=ssg1S1002693>

How to get IBM Redbooks publications

You can search for, view, or download IBM Redbooks publications, Redpapers, Hints and Tips, draft publications and Additional materials, as well as order hardcopy IBM Redbooks publications or CD-ROMs, at this website:

ibm.com/redbooks

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

Index

Numerics

2805 69
2805-MC4 280
2805-MC5 213
3390 Model A 394
700 36
951 7–8
95E 7–8, 40

A

AAL
 benefits 64
Accelerated Graphics Port (AGP) 43
activate licenses
 applying activation codes using the DS CLI 265
 applying activation codes using the GUI 262
 functions 252
 obtaining activation code 255
Adaptive Multi-stream Prefetching (AMP) 18, 46, 50, 174
Adaptive Multistream Prefetching (AMP) 7
address groups 124
Advanced Function Licenses 232
 activation 232
affinity 109
air cooling 8
alias devices 353
allocation 117
allocation unit 115
AMP 18, 46, 174
applying activation codes using the DS CLI 265
applying activation codes using the GUI 262
architecture 42
array sites 105
arrays 63, 106
authorization 252

B

baffle 447
base frame 7–8, 38
Base Frame -DS8700 36
battery backup assemblies 67
battery backup unit (BBU) 36, 38
battery booster 40
battery pack 23, 29
BBU 36–38
Break Point Value (BPV) 134
browser 284
bundle 402
Business Class 28
business class 8, 32, 39
Business Class Cabling 28
business class configuration 8
business continuity 14

Business-to-Business 230
Business-to-Business VPN 432

C

cables 210
cache 23, 29, 172
 pollution 175
cache management 49
caching algorithm 172
Call Home 17, 89, 234, 247, 276, 433
Capacity Magic 172, 456
Capacity on Demand 25, 33, 445
capacity upgrade 446
CDA 404
CEC 72
central electronic complex 9, 23, 29
Central Electronic Complex (CEC) 23, 29, 71
Challenge Key 443
chfbvol 381
chpass 244
chuser 244
CIM 229
 agent 230
CIM agent 288
CIMOM 288–289, 301
CKD volumes 112
 allocation and deletion 117
clear key mode 6
CLOCK 175
code
 obtaining activation 255
Code Distribution and Activation (CDA) 404
commands
 structure 367
community name 410
components 35
concurrent copy session timeout 350
configuration
 flow 247, 276
configuration state 121
Configure Devices wizard 292
configuring 121
configuring the DS8000 373, 388
configuring using DS CLI
 configuring the DS8000 373, 388
consistency group
 timeout 351
Consistency Group FlashCopy 142
Control Unit Initiated Reconfiguration see CUIR
Control Unit-Initiated Reconfiguration (CUIR) 406
cooling 8
 disk enclosure 66
cooling fan 37
cooperative caching 169

- Copy Services 112
 - event traps 412
- CSCAN 175
- CUIR 88, 406
- Cylinder Managed Space 131

D

- DA 57
 - Fibre Channel 161, 163
- daemon 409
- data migration
 - Disk Magic 457
- data set FlashCopy 142–143
- DB2 175
- DDM 65, 161, 163
- deadlock 6
- deconfiguring 121
- default profile 365
- demand paging 173
- demand-paged data 50
- destage 42, 115
- device adapter 9
- device adapter (DA) 24, 31, 57
- DFSA 254
- disk drive set 11
- disk drives
 - capacity 220
- disk enclosure 58
 - power and cooling 66
- Disk Magic 172, 457
- Disk Storage Feature Activation (DSFA) 254
- disk subsystem 58
- disk virtualization 104
- DoD 5220.20-M 461
- DS API 138
- DS CIM 281
- DS CIM Command-Line Interface (DSCIMCLI) 230
- DS CLI 138, 214, 230
 - applying activation codes 265
 - command structure 367
 - configuring second HMC 246
 - console
 - default profile 365
 - help 370
 - highlights 364
 - interactive mode 367–368
 - operating systems 364
 - return codes 370
 - script command mode 369
 - script mode 367
 - single-shot mode 367
 - user accounts 364
 - user assistance 370
 - user management 239
- DS Command-Line Interface see DS CLI
- DS GUI 310
- DS HMC
 - external 245
 - planning 225
- DS HMC planning

- activation of Advanced Function Licenses 232
- host prerequisites 233
- latest DS8000 microcode 233
- maintenance windows 233
- time synchronization 233
- DS Open API 230
- DS SM 138
 - user management 241
- DS® family 4
- DS6000
 - business continuity 14
 - Capacity Magic 456
 - dynamic LUN/volume creation and deletion 13
 - large LUN and CKD volume support 14
 - simplified LUN masking 14
 - SNMP configuration 411
 - user management using DS CLI 239
- DS8000 299
 - activate license functions 252
 - activation of Advanced Function Licenses 232
 - applying activation codes using the DS CLI 265
 - applying activation codes using the GUI 262
 - arrays 63
 - base frame 38
 - battery backup assemblies 67
 - Capacity Magic 456
 - components 35
 - configuration
 - flow 247, 276
 - configuring 373, 388
 - considerations prior to installation 202
 - data placement 177
 - DDM 65
 - Disk Magic 457
 - disk subsystem 58
 - DS CLI console
 - DS CLI highlights 364
 - DS HMC
 - planning 225
 - DS8100 Model 921 23
 - ESCON 210
 - expansion frame 40
 - external DS HMC 245
 - Fibre Channel/FICON 210
 - FICON 20
 - floor type and loading 205
 - frames 36, 38
 - HMC Address 294
 - host adapter 10
 - host interface and cables 210
 - host prerequisites 233
 - I/O priority queuing 19
 - IBM Redbooks publications 469
 - input voltage 208
 - maintenance windows 233
 - microcode 233
 - modular expansion 38
 - Multiple Allegiance 19
 - network connectivity planning 212
 - obtaining activation code 255

- online resources 470
- PAV 19
- performance 18, 157
- planning for growth 223
- power and cooling 65
- power connectors 208
- power consumption 209
- power control features 210
- Power Line Disturbance (PLD) 210
- power requirements 208
- POWER5 166
- PPS 65
- processor complex 47
- project planning 279, 309
- Remote Mirror and Copy connectivity 219
- remote power control 216
- remote support 215
- room space and service clearance 207
- SAN connection 216
- scalability 13
- SDD
- server-based 43
- service 17
- service processor 97
- setup 17
- single sign-on 295
- SMP 43
- spare 63
- sparing considerations 219
- SPCN
- stripe size 182
- supported environment 12
- System z performance 19
- technical environment 228
- time synchronization 233
- Username 294
- z/OS Metro/Global Mirror 16
- DS8100
 - Model 921 23
- DS8700
 - disk drive set 11
 - expansion enclosure 62
 - I/O enclosure 50–52
 - RIO-G 50
 - storage capacity 11
 - switched FC-AL 61
- DS8700 frame types 36
- DS8800
 - architecture 42
 - Business Class 28
 - disk enclosure 58
 - EPO 41
 - service processor 50
 - SPCN
- DSCB 132
- DSFA 255, 257
- dual parity 172
- Dynamic alias 353
- Dynamic Extent Pool Merge 110
- dynamic LUN/volume creation and deletion 13

- Dynamic Volume Expansion 5, 121
- Dynamic Volume Expansion (DVE) 121, 129, 132, 381, 393
- dynamic volume migration 122
- Dynamic Volume Relocation 111
- Dynamic Volume Relocation (DVR) 122

E

- EAM 120
- eam rotateexts 379
- eam rotatevols 379
- Easy Tier 5, 12, 110, 119, 122
- EAV 112, 191
- Element Manager 229, 312
- encryption 6
- energy consumption 4
- Enterprise Choice
 - warranty 252
- Enterprise Storage Server Network Interface server (ESSNI) 228
- ePLD 37–38, 210
- EPO 40–41
- EPO switch 42
- ESCON 210
- ESS 800
 - Capacity Magic 456
- ESSNI 228
- Ethernet
 - switches 67
- Ethernet adapter 35
- Expansion frame 37
- expansion frame 22, 24, 28, 31, 37–38, 40
 - 95E 28
- expansion unit 7–8
- Extended Address Volumes (EAV) 7, 14, 112
- extended attribute DSCB 132
- Extended Distance 6
- Extended Distance FICON 195–196
- extended Power Line Disturbance 37
- Extended Remote Copy (XRC) 16, 152
- extent pool merge 119
- extent pools 108, 114
- extent rotation 118
- Extent Space Efficient 14
- Extent Space Efficient Volumes 114
- extent type 107, 109
- external DS HMC 245

F

- failover 81
- fan 37
- fan sense card 36–37
- FATA disk drives
 - capacity 220
 - performance 171
- FC-AL
 - overcoming shortcomings 159
 - switched 10
- FCP 23, 29

- FDE 5, 7, 11, 25, 32
- Fibre Channel
 - distances 57
- Fibre Channel/FICON 210
- FICON 20, 23, 29, 42
- File Transfer Protocol (FTP) 434
- Firefox 284
- fixed block LUNs 112
- FlashCopy 14–15, 138
 - benefits 141
 - Consistency Group 142
 - data set 142
 - establish on existing RMC primary 142
 - inband commands 15, 142
 - incremental 15, 142
 - Multiple Relationship 15, 142
 - no background copy 141
 - options 141
 - persistent 142
 - Refresh Target Volume 142
- FlashCopy SE 15, 139
- flexible service processor (FSP) 50
- floor type and loading 205
- footprint 4
- frames 36, 38
 - base 38
 - expansion 40
- Full Disk Encryption (FDE) 5, 7, 11, 25, 32, 223, 254, 446
- functions
 - activate license 252

G

- Global Copy 14, 16, 138, 146, 154
- Global Mirror 14, 16, 138, 146, 154
 - how it works 147
- GUI
 - applying activation codes 262
- GX+ bus 43

H

- HA 10, 52
- Hardware Configuration Management (HCM) 129
- Hardware Management Console (HMC) 11, 212
- heartbeat 433
- help 371
 - DS CLI 370
- High Performance FICON 6
- High Performance FICON for z (zHPF) 197
- HMC 17, 212
- HMC planning
 - technical environment 228
- host
 - interface 210
 - prerequisite microcode 233
- host adapter (HA) 6
- host adapter see HA
- host adapters 22, 29
 - four port 165
- host attachment 125

- hosttype 384
- HWN021724W 292
- HyperPAV 19, 191–192, 353
- HyperPAV license 254
- Hypervisor (PHYP) 74

I

- I/O Definition File (IODF) 129
- I/O enclosure 50–52, 66
- I/O priority queuing 19, 195
- i5/OS 19, 114
- IBM Certified Secure Data Overwrite 17
- IBM FlashCopy SE 15, 114, 376, 389
- IBM Redbooks publications 469
- IBM System Storage Interoperability Center (SSIC) 233
- IBM System Storage N series 180
- IBM Tivoli Storage Productivity Center Basic Edition (TPC BE) 68
- IBM TotalStorage Multipath Subsystem Device Driver see SDD
- IBM TotalStorage Productivity Center 11
- IBM TotalStorage Productivity Center for Data 282
- IBM TotalStorage Productivity Center for Disk 282
- icon view 315
- IKS 11
- impact 115
- inband commands 15, 142
- increase capacity 356
- Incremental FlashCopy 15, 142
- index scan 175
- indicator 252
- initckdvol 116
- initfbvol 116
- input voltage 208
- installation
 - DS8000 checklist 202
- Intelligent Write Caching (IWC) 6, 50, 175
- interactive mode
 - DS CLI 367–368
- internal fabric 9
- Internet Explorer 285
- IODF 129
- IOPS 170
- IOSQ Time 192
- IPv6 7
- isolated key server (IKS) 11
- IWC 6, 50, 175

L

- lane 43
- large LUN and CKD volume support 14
- lateral change 262
- LCU 348
- LCU type 350
- LDAP 295, 311
- LDAP authentication 7, 363
- LDAP based authentication 237
- Least Recently Used (LRU) 174
- licensed function

- authorization 252
- indicator 252
- Lightweight Directory Access Protocol (LDAP) 311
- logical configuration 203
- logical control unit (LCU) 123
- logical size 114
- logical subsystem see LSS
- logical volumes 112
- long busy state 350
- long busy wait 169
- longwave 57
- LSS 123
- lsuser 239
- LUN ID 346
- LUNs
 - allocation and deletion 117
 - fixed block 112
 - masking 14
 - System i 114
- LVM
 - striping 180

M

- machine reported product data (MPRD) 433
- machine type 22, 29
- maintenance windows 233
- man page 371
- Managed Disk 112
- Management Information Base (MIB) 408
- managepfile 240, 366
- manual mode 122
- Metro Mirror 14, 16, 138, 145, 154
- Metro/Global Mirror 14
- MIB 408, 411
- microcode
 - update 89
- migrate 355
- Migrate Volume 110, 122
- mkckdvol 392
- mkfbvol 378
- mkrank 271
- mkuser 244
- Model 941 22, 28
- Model 94E 22
- Model 951 7–8, 28–29
- Model 95E 28
- modified data 42
- modular expansion 38
- monitoring group 296–297
- Most Recently Used (MRU) 174
- MRPD 433
- Multicylinder Unit (MCU) 131
- Multiple Allegiance 19, 194
- Multiple Reader 16, 20
- Multiple Relationship FlashCopy 15, 142
- Multiple Subchannel Sets (MSS) 192

N

- native device interface 301

- navigation pane 315
- Nearline 110
- network connectivity planning 212
- Network Time Protocol (NTP) 233
- NMS 408
- non-disruptive upgrade 8
- non-volatile storage (NVS) 52, 80
- NTP 233
- Nucleus Initialization Program (NIP) 189
- NVS 52, 80

O

- OEL 253
- offloadauditlog 442
- On Demand Data (ODD) Dump 434
- online resources 470
- Open HyperSwap, 283
- open systems
 - performance 183
 - sizing 183
- Operating Environment License (OEL) 203, 253
- Out of Band Fabric agent 306
- over provisioning 114

P

- Parallel Access Volumes see PAV
- PAV 19, 123
- Payment Card Industry (PCI) 6
- Payment Card Industry Data Security Standard (PCI-DSS) 6
- PCI Express 9, 37, 39, 43, 46
 - adapter 210
- PCI-DSS 6
- PDU 29
- Peer-to-Peer Remote Copy (PPRC) 145
- PEPackage 434
- performance
 - data placement 177
 - FATA disk drives 171
 - open systems 183
 - determining the number of paths to a LUN 183
 - where to attach the host 184
 - workload characteristics 177
 - z/OS 185
 - connect to System z hosts 185
 - disk array sizing 170
- Performance Accelerator feature 27, 34
- Persistent FlashCopy 142
- physical partition (PP) 182
- physical paths 306
- physical planning 201
 - delivery and staging area 204
 - floor type and loading 205
 - host interface and cables 210
 - input voltage 208
 - network connectivity planning 212
 - planning for growth 223
 - power connectors 208
 - power consumption 209

- power control features 210
- Power Line Disturbance (PLD) 210
- power requirements 208
- Remote Mirror and Copy connectivity 219
- remote power control 216
- remote support 215
- room space and service clearance 207
- sparing considerations 219
- storage area network connection 216
- physical size 114
- planning
 - DS Hardware Management Console 199
 - logical 199
 - physical 199
 - project 199
- planning for growth 223
- power 66
 - disk enclosure 66
 - I/O enclosure 66
 - processor enclosure 66
- power and cooling 65
- power connectors 208
- power consumption 209
- power control features 210
- power distribution units (PDU) 29
- Power Line Disturbance 37
- Power Line Disturbance (PLD) 210
- power loss 83
- power requirements 208
- Power Sequence Controller (PSC) 216
- POWER5 166
- POWER6 7
- POWER6+ 8–9
- PPRC-XD 146
- PPS 38
- Predictive Failure Analysis (PFA) 91
- prefetch wastage 175
- prefetched data 50
- prefetching 173
- Preserve Mirror 143
- Primary Power Supply (PPS) 404
- priority queuing 195
- probe job 299
- probe schedule 297
- processor complex 9, 47, 73
- processor complexes (CEC) 42
- processor enclosure
 - power 66
- project plan
 - considerations prior to installation 202
 - physical planning 201
 - roles 203
- project planning 279, 309
 - information required 203
- PTC 141

R

- Rack Power Control (RPC) 404
- rack power control cards (RPC) 36
- RAID 10

- implementation 94
- RAID 6 7, 92, 120
 - implementation 93
 - performance 172
- raidtype 374
- RANDOM 18
- random write 172
- ranks 107, 115
- RAS
 - naming 72
- reconfiguring 121
- recovery key 6
- Recovery Point Objective see RPO
- Redbooks Web site 470
 - Contact us xviii
- reduction 262
- related publications 469
 - help from IBM 470
 - how to get IBM Redbooks publications 470
 - online resources 470
- Remote Mirror and Copy function see RMC
- Remote Mirror and Copy see RMC
- Remote Pair FlashCopy 143
- remote power control 216
- remote support 215
- reorg 180
- repcapalloc 376
- repository 114, 376, 389
- repository size 115
- return codes
 - DS CLI 370
- RIO-G 37, 39, 50
 - interconnect 78
- RIO-G loop 45
- RMC 14, 16, 138, 145, 219
 - Global Copy 146
 - Global Mirror 146
 - Metro Mirror 145
- rmsestg 377
- rmuser 244
- role 291
- room space 207
- rotate extents 5, 118, 341, 352
- rotate volumes 342, 352, 379, 392
- rotated volume 118
- rotateexts 381, 394
- RPC 36, 38
- RPO 154

S

- SAN Volume Controller (SVC) 180
- SARC 18, 50, 173–174
- SAS drive 8
- scalability 13
 - DS8000
 - scalability 168
- script command mode
 - DS CLI 369
- script mode
 - DS CLI 367

- scsimap256 346
- SDD 19, 183
- Secure Data Overwrite 461
- secure key mode 6
- Secure Sockets Layer (SSL) 431
- Security Administrator 237
- self-healing 76
- SEQ 18
- SEQ list 174
- Sequential prefetching in Adaptive Replacement Cache (SARC) 46
- Sequential prefetching in Adaptive Replacement Cache see SARC
- sequential read throughput 8
- sequential write throughput 8
- server affinity 109
- server-based SMP 43
- service clearance 207
- service processor 50, 97
- session timeout 350
- SFI 73
- S-HMC 11
- shortwave 57
- showckdvol 392
- showfbvol 393
- showpass 241
- showsestg 376
- showuser 240
- silicon on insulator (SOI) 9
- Silicon-on Insulator 42
- simplified LUN masking 14
- simultaneous multi-threading (SMT) 43
- Single Point of Authentication 311
- single-shot mode
 - DS CLI 367
- sizing
 - open systems 183
 - z/OS 185
- SMP 43
- SMT 43
- SMUX 408
- SNMP 234, 408
 - agent 408–409
 - configuration 411, 420
 - Copy Services event traps 412
 - manage 409
 - notifications 412
 - preparation for the management software 424
 - preparation on the DS HMC 420
 - preparation with DS CLI 421, 423
 - trap 409, 411
 - trap 101 413
 - trap 202 415
 - trap 210 416
 - trap 211 416
 - trap 212 416
 - trap 213 416
 - trap 214 417
 - trap 215 417
 - trap 216 417
 - trap 217 417
 - trap request 408
- SOI 9
- Solid State Disk 37
- Solid State Drive (SSD) 6, 10, 170, 222
- Space Efficient 373
- Space Efficient repository 120
- Space Efficient volume 129
- spares 63
- sparing 219
- sparing considerations 219
- spatial ordering 175
- SPCN 50, 97
- SSD 6
- SSID 350
- SSL connection 17
- SSPC 11, 68, 73, 232, 279–280
 - install 284
 - user 290
- SSPC ip address 287
- standard cabling 28
- Standby Capacity on Demand see Standby CoD
- Standby CoD 12, 26, 33, 223
- status 320
 - Attention 320
 - Critical 320
- storage area network connection 216
- storage capacity 11
- storage complex 73
- storage facility image 73
- Storage Hardware Management Console see HMC
- Storage Pool Striping 6, 109, 118, 120, 177–178, 180, 341, 352
- Storage Tier Advisor Tool 5
- stripe 115
 - size 182
- striped volume 119
- Subsystem Standard Group 297
- switched FC-AL 10
 - advantages 60
 - DS8700 implementation 61
- System Data Mover (SDM) 153
- System i
 - LUNs 114
- system power control network see SPCN
- System Storage Productivity Center (SSPC) 6, 11, 73, 279
- System z
 - performance 19

T

- temporal ordering 175
- Thin Provisioning 114
- Three Site BC 283
- time synchronization 233
- Tivoli Key Lifecycle Manager (TKLM) 11
- Tivoli Productivity Center 460
- Tivoli Storage Productivity Center for Replication (TPC-R) 138
- TKLM 11

tools

- Capacity Magic 456
- TotalStorage Productivity Center (TPC) 229
- TotalStorage Productivity Center for Fabric 282
- TotalStorage Productivity Center for Replication (TPC-R) 7, 88
- TPC
 - native device interface 292
 - role 291
- TPC for Disk 460
- TPC GUI 285
- TPC-R 138
- track 115, 139
- Track Managed Space 131
- Track Space Efficient (TSE) 14, 114, 352, 376
- Track Space Efficient Volumes 114
- transposing 121
- trap 408–409, 411

U

- user accounts
 - DS CLI 364
- user assistance
 - DS CLI 370
- user management
 - using DS CLI 239
 - using DS SM 241
- user management using DS SM 241
- user role 291

V

- value based licensing 253
- value based pricing 7
- virtual capacity 120
- Virtual Private Network (VPN) 215
- virtual space 114–115
- virtualization
 - abstraction layers for disk 104
 - address groups 124
 - array sites 105
 - arrays 106
 - benefits 128
 - concepts 103
 - definition 104
 - extent pools 108
 - hierarchy 127
 - host attachment 125
 - logical volumes 112
 - ranks 107
 - volume group 126
- Volume Group 346
- volume groups 126, 337
- volume manager 120
- volume relocation 119
- volumes
 - CKD 112
- VPN 17, 428
 - benefits 429

W

- warranty 252
- Web UI 229–230
- web-based user interface (Web UI) 229
- WLM 189
- Workload Manager 189
- write penalty 171

X

- XRC 16, 152
- XRC session 350

Z

- z/OS FICON Discovery and AutoConfiguration Feature (zDAC) 129
- z/OS Global Mirror 14, 16, 20, 23, 29, 138, 152, 154
- z/OS Global Mirror session timeout 350
- z/OS Metro/Global Mirror 16, 148, 153
- z/OS Workload Manager 189
- zDAC 129
- zHPF 6, 20, 197
 - Extended Distance 6
 - extended distance 198
 - multitrack 198
- zIIP 153



IBM System Storage DS8000: Architecture and Implementation

(1.0" spine)

0.875" x 1.498"

460 <-> 788 pages



IBM System Storage DS8000 Architecture and Implementation



Redbooks®

Learn the DS8700 and DS8800 new and common features

Plan, install, and configure the DS8000

Discover the enhanced DS GUI

This IBM Redbooks publication describes the concepts, architecture, and implementation of the IBM System Storage DS8700 and DS8800 storage systems. The book provides reference information to assist readers who need to plan for, install, and configure the DS8700 and DS8800.

The DS8700 includes IBM POWER6-based controllers. The IBM System Storage DS8800 is the most advanced model in the IBM DS8000 lineup and is equipped with IBM POWER6+ based controllers. Both systems feature a dual 2-way or dual 4-way processor complex implementation. They also feature enhanced 8 Gbps device adapters and host adapters. Their extended connectivity, with up to 128 Fibre Channel/FICON ports for host connections, makes them suitable for multiple server environments in both open systems and IBM System z environments. Both systems now support thin provisioning. They also support the Full Disk Encryption (FDE) feature. If desired, they can be integrated in an LDAP infrastructure.

The DS8800 is equipped with high-density storage enclosures populated with 24 small-form-factor SAS-2 drives. The DS8700 and DS8800 storage subsystems can be equipped with Solid-State Drives (SSDs).

The DS8700 and DS8800 can automatically optimize the use of SSD drives through the Easy Tier feature, which is available for no extra fee. For details about Easy Tier, refer to *IBM System Storage DS8000: Easy Tier Concepts and Usage*, REDP-4667.

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

SG24-8886-01

ISBN 0738435864