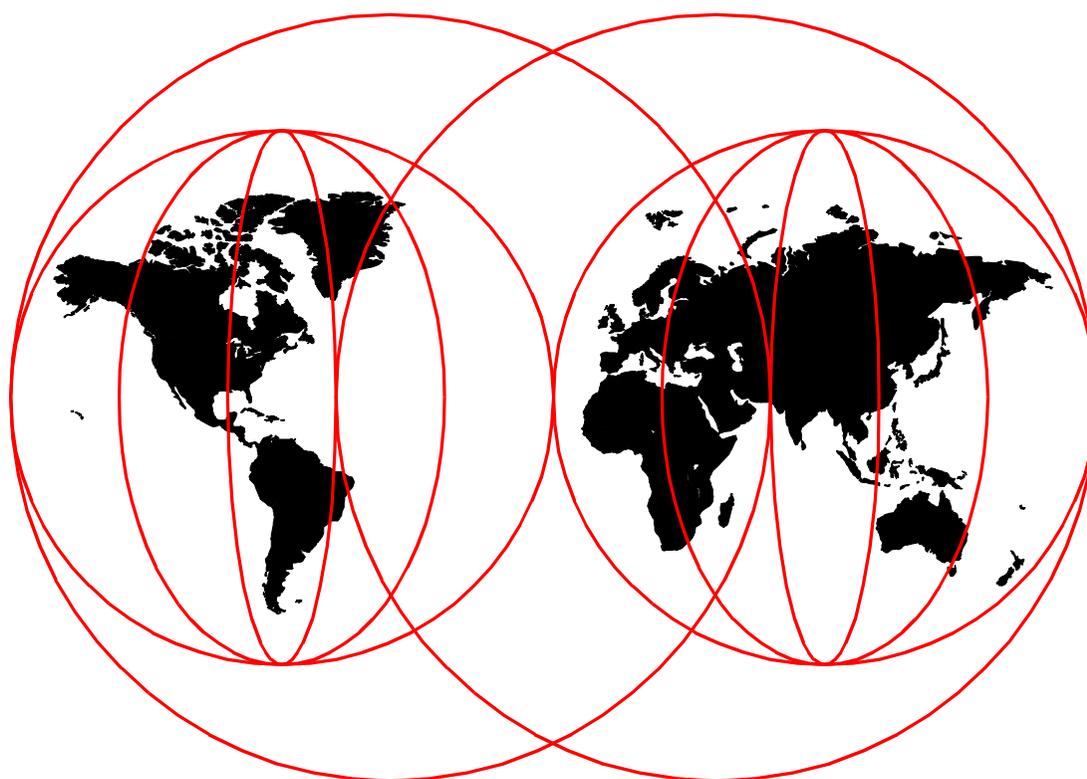


Understanding SSA Subsystems in Your Environment

Mark Blunden, Paul Staples



International Technical Support Organization

www.redbooks.ibm.com



International Technical Support Organization

SG24-5750-00

**Understanding SSA Subsystems
in Your Environment**

April 2000

Take Note!

Before using this information and the product it supports, be sure to read the general information in Appendix C, "Special notices" on page 141.

First Edition (April 2000)

This edition applies to the Serial Storage Architecture disk as available at the publishing date. For more information about product documentation, see the Publications section of the IBM Programming announcements for SSA.

Comments may be addressed to:
IBM Corporation, International Technical Support Organization
Dept. QXXE Building 80-E2
650 Harry Road
San Jose, California 95120-6099

When you send information to IBM, you grant IBM a non-exclusive right to use or distribute the information in any way it believes appropriate without incurring any obligation to you.

© Copyright International Business Machines Corporation 2000. All rights reserved.

Note to U.S Government Users - Documentation related to restricted rights - Use, duplication or disclosure is subject to restrictions set forth in GSA ADP Schedule Contract with IBM Corp.

Contents

Figures	vii
Tables	ix
Preface	xi
The team that wrote this redbook	xi
Comments welcome	xii
Chapter 1. Serial Storage Architecture (SSA) overview	1
1.1 What is SSA?	1
1.2 SSA subsystem	1
1.2.1 SSA operation	2
1.2.2 Loop availability	3
1.2.3 Spatial reuse	3
1.2.4 Compatibility	3
1.2.5 Summary	3
1.3 Storage area network (SAN)	4
Chapter 2. SSA hardware components	7
2.1 SSA adapters	7
2.2 SSA disk subsystems	12
2.2.1 3527 SSA storage subsystem for PC servers	12
2.2.2 7131 SSA Multi-storage Tower Model 405	13
2.2.3 7133 Serial Storage Architecture (SSA) disk subsystem	13
2.3 SSA disks	21
2.4 Other SSA components	21
2.4.1 IBM 7190-100 SCSI host to SSA loop attachment	22
2.4.2 IBM 7190-200 ultra SCSI host to SSA loop attachment	24
2.4.3 IBM 2108 SAN Data Gateway Model S20	24
2.4.4 VICOM SLIC Router FC-SL	26
2.4.5 Copper cables, fibre optic extenders and fibre cables	28
Chapter 3. Configuring SSA loops	33
3.1 Some rules for cabling SSA loops	33
3.2 Bypass cards	34
3.3 Loops with multiple hosts	40
3.3.1 Cabling example 1	40
3.3.2 Cabling example 2	41
3.3.3 Cabling example 3 (recommended for highest availability)	42
3.4 Data domains and disk placement	43
3.5 Mirroring configurations	44
3.5.1 Mirroring example 1	45
3.5.2 Mirroring example 2	45
3.5.3 Mirroring example 3	46
3.6 SSA adapter cards	48
3.7 Fibre optic extender	50
3.8 Using RAID	51
3.9 RISC 6000 systems with internal SSA disks	53
3.10 How to develop an SSA configuration	57
Chapter 4. Tools for managing SSA	61
4.1 SSA device drivers and the management process	61

4.2	Microcode maintenance	63
4.3	SSA commands and services	64
4.3.1	SMIT commands in AIX	64
4.3.2	AIX command line entries	68
4.3.3	Utilities	70
4.3.4	Remote Systems Management	71
4.3.5	SSA cable	72
4.4	SSA spare tool	78
4.5	Managing SSA disks in a Sun Solaris environment	79
4.5.1	Other tools	80
4.6	Managing SSA disks in an HP-UX environment	80
4.7	Managing SSA disks in a Windows NT environment	81
4.8	Managing SSA disks in other environments	85
4.9	IBM StorWatch Serial Storage Expert (StorX)	87
Chapter 5. Monitoring and managing SSA environment with StorX		89
5.1	Customizing a StorX live view	89
5.2	Event monitor	91
5.3	Monitoring and managing a shared configuration	92
5.4	Monitoring and managing a complex configuration	94
5.5	Using StorX for subsystem troubleshooting	97
5.6	Monitoring and managing an hdisk failure	98
5.6.1	Recovery procedure (hdisk failure)	100
5.7	Monitoring and managing an SSA open loop problem	102
5.8	Monitoring and managing an SSA adapter failure	103
5.8.1	Two SSA loops connected to one SSA adapter card	103
5.8.2	One SSA loop shared by two systems	104
5.9	Monitoring and managing an SSA RAID array creation	105
5.10	Monitoring and managing disk problems in a RAID 5 configuration	109
5.10.1	Crash of pdisk 1	109
5.10.2	Recovery procedure (RAID 5 configuration)	110
5.10.3	New hot spare disk definition	111
5.10.4	Completed recovery procedure	112
5.11	Monitoring, managing, and diagnosing with StorX	114
5.11.1	Disk properties window	114
5.11.2	Adapter Properties window	119
Appendix A. Disk replacement methodology		121
A.1	Disk replacement procedure summary	121
A.2	Methods summary	121
A.2.1	Method 1	122
A.2.2	Method 2	126
A.2.3	Method 3	130
A.2.4	Method 4	134
Appendix B. SSA RAID disk change flowchart		139
Appendix C. Special notices		141
Appendix D. Related publications		143
D.1	IBM Redbooks publications	143
D.2	IBM Redbooks collections	143
D.3	Other resources	143

How to get IBM Redbooks	145
IBM Redbook Fax Order Form	146
List of abbreviations	147
Index	149
IBM Redbooks evaluation	151

Figures

1. The authors of this book—Paul and Mark	xi
2. SSA spatial reuse	2
3. SSA adapter connection ports and LEDs	7
4. 7133 Stand-alone Disk Tower and 7133 Disk Drawers in S00 frame	14
5. 7133 model 10/500 disk and connection layout	16
6. 7133 model 20/600 disk and connector layout	18
7. 7133 D/T40 operator panel	20
8. The 2108 Model S20 Drawer and stand-alone box versions	25
9. VICOM SLIC Router FC-SL and drawer	27
10. 7133 schematic layout	35
11. Simple loop with two adapters	36
12. Simple loop with three adapters	36
13. Simple loop, poor availability configuration	37
14. Advised disk unit layout in partially populated enclosures	38
15. Poor availability configuration	38
16. Result of loss of adapter	39
17. Correct cabling for this configuration	39
18. Logical cabling diagram: example 1 (not recommended)	40
19. Physical cabling diagram: example 1	41
20. Logical cabling diagram: example 2 (not recommended)	41
21. Physical cabling diagram: example 2	42
22. Logical cabling diagram: example 3 (recommended)	42
23. Physical cabling diagram: example 3	43
24. Optimal disk allocation within multi host loop	43
25. Poor disk allocation in multi host loop	44
26. Disaster waiting to happen! NOT recommended	44
27. Example 1, simple mirroring, disk redundancy only	45
28. Example 2, simple mirroring, disk and 7133 enclosure redundancy	46
29. Example 3, simple mirroring, best redundancy for single host	46
30. Mirrored configurations — total redundancy for all components	47
31. Mirrored configurations — total redundancy and maximum access	48
32. Fibre-Optic Extender (OE) layout	50
33. Fibre-Optic extender usage	50
34. Layout using RAID and non-RAID arrays for 6217 and 6218	52
35. Layout using RAID and non-RAID arrays for 6215 and 6219 (single adapter)	53
36. Internal SSA layout RS6000 7025 model F30	54
37. Internal SSA layout RS6000 7025 model F40	55
38. Internal SSA layout RS6000 702X models F50 and H50	56
39. Internal SSA layout RS6000 7026 model H10	57
40. Suggested conceptual diagram, example 1	58
41. Actual cabling schematic, example 1	59
42. Suggested conceptual diagram, example 2	59
43. Actual cabling schematic, example 2	60
44. SMIT Devices Menu	65
45. SSA Adapter Menu	65
46. SSA Disks Menu	66
47. Logical Disk Menu	66
48. SSA Physical Disk Menu	67
49. SSA RAID Arrays Menu	67
50. SSA RSM initial screen view	82

51. Detailed view of Adapter product information	82
52. Physical view from Adapter.	83
53. VPD of Individual disk.	83
54. Logical view from Adapter: part 1	84
55. Logical view from Adapter: part 2	84
56. View of Event Logger	85
57. Uncustomized StorX live view	89
58. Management Set Properties panel	90
59. Customized StorX viewer with annotations.	91
60. Event Monitor display	91
61. Event Monitor showing the associated device	92
62. StorX live view from RS/6000-1: logical disk option	93
63. StorX live view from RS/6000-2: logical disk option	93
64. StorX live view from RS/6000-1, pdisk name option	94
65. StorX live view from RS/6000-2, pdisk name option	94
66. SSA planning without using StorX	95
67. Live view of uncustomized StorX	96
68. Live view of customized StorX from System-1, pdisk name option.	97
69. Initial configuration, logical disk name option	98
70. Initial configuration, physical disk name option	98
71. StorX live view of an hdisk problem	99
72. hdisk replacement StorX live view	100
73. Configuration of the new SSA disk, pdisk name option	101
74. Recovery procedure completed	102
75. Initial configuration of two SSA loops connected to one SSA adapter	102
76. StorX live view of an SSA open loop.	103
77. Initial configuration with two loops — StorX live view	104
78. SSA adapter problem affecting two loops — StorX live view	104
79. Initial configuration with one loop and two adapters — StorX live view	105
80. SSA adapter problem on System1 — StorX live view.	105
81. Initial configuration before SSA RAID array creation — StorX live view.	106
82. Hot spare definition — StorX live view	106
83. Hot spare and array candidate disks — StorX live view	107
84. SSA RAID array — StorX live view	108
85. SSA RAID array, pdisk option — StorX live view	108
86. Volume group RAID5VG — StorX live view	109
87. Crash of pdisk 1 in an SSA RAID array	110
88. Replacement of pdisk 1 — StorX live view	111
89. Definition of pdisk 8 — StorX live view	112
90. New SSA RAID 5 environment, pdisk option	113
91. New SSA RAID5 environment, logical disk option	113
92. Two loops connected to an SSA adapter with RAID5 array as one loop	114
93. General page of Disk Properties window	115
94. Description page of Disk Properties windows.	116
95. System page of Disk Properties window.	116
96. Advanced page of Disk Properties window.	117
97. General page of Disk Properties window	118
98. General page of Disk Properties windows	118
99. Adapter Properties window	119
100.SSA RAID disk change flowchart : part 1	139
101.SSA RAID disk change flowchart: part 2	140

Tables

1. RS/6000 non-RAID SSA adapters	8
2. RS/6000 RAID SSA adapters	8
3. RS/6000 enhanced RAID SSA adapters	8
4. RS/6000 advanced serial RAID adapter	9
5. RS/6000 advanced serial RAID plus adapter	10
6. Sun Systems SBus SSA adapters	11
7. Open systems PCI RAID SSA adapters	11
8. Advanced serial RAID/X adapter	11
9. 3527 SSA storage subsystem for PC server	12
10. 7131 SSA Multi-storage Tower Model 405 specifications	13
11. 7133 Models 10, 20, 500, 600, D40 and T40 specifications	15
12. Disk to cable connector assignments models 10/500	16
13. Link LEDs	18
14. Mode LEDs	18
15. LED state and meanings models D40/T40	20
16. Supported servers and operating systems for 7190-100	23
17. Supported servers and operating systems for the 2108-S20	25
18. Adapter cards summarized	49
19. VICOM and IBM product comparison	86

Preface

This redbook describes how to design, configure, monitor, and control SSA disk subsystems that are connected to IBM and non-IBM operating system servers.

The material demonstrates how to manage SSA loop configurations using various storage management facilities such as AIX Logical Volume Manager, StorX, and the latest Web user-interface technology.

Platform-specific storage management offerings are discussed, and customers are provided with a complete and comprehensive guide showing how to perform flexible, low-risk, remote-site, open storage subsystem design, configuration, and management in a variety of SSA disk subsystem environments.

The team that wrote this redbook

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

Mark Blunden is the project leader for Open Systems Storage at the International Technical Support Organization, San Jose Center. He has co-authored 7 previous redbooks and co-managed 4 other redbooks. Mark also teaches IBM classes worldwide on all areas of Storage. He has worked for IBM for 19 years in many areas of the IT business. Before joining the ITSO in 1998, Mark worked in Sydney, Australia, as an Advisory Storage Specialist.

Paul Staples is a Senior Support Specialist (H/W) in the UK. He has 16 years of experience in mid-range systems, the last 9 years working in the AIX System Support Centre supporting Business Partners and IBM Field Engineers in the UK and Ireland. Paul's areas of expertise include high end RS/6000 systems (SP2, Rxx and S70) and Storage subsystems. He also teaches IBM classes for Field CEs. He has gained his Certified AIX Support Professional accreditation for AIX 4.3.



Figure 1. The authors of this book—Paul and Mark

A special thanks to Derrick Daines for his invaluable support in providing input to this book. Derrick was working remotely from the UK to provide information and guidance. Most of the time that Derrick spent was his own.

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

Barry Mellish
International Technical Support Organization, San Jose Center

Joe Emery
VICOM Systems, Mountain View, CA

Comments welcome

Your comments are important to us!

We want our redbooks to be as helpful as possible. Please send us your comments about this or other redbooks in one of the following ways:

- Fax the evaluation form found in “IBM Redbooks evaluation” on page 151 to the fax number shown on the form.
- Use the online evaluation form found at <http://www.redbooks.ibm.com/>
- Send your comments in an Internet note to redbook@us.ibm.com

Chapter 1. Serial Storage Architecture (SSA) overview

A brief description of SSA is given along with its history and components. How to design and implement SSA subsystems are covered in later chapters. For a full description of SSA and its functionality, read *A Practical Guide to Serial Storage Architecture for AIX*, SG24-4599.

1.1 What is SSA?

SSA is a high performance, serial interconnect technology used to connect disk devices and host adapters. SSA is an open standard, and SSA specifications have been approved by the SSA Industry Association and are approved as an ANSI standard through the ANSI X3T10.1 subcommittee.

Before SSA, the method of transmitting data to storage devices used a parallel architecture, such as SCSI (Small Computer System Interface). It was recognized that although this is a fast and low-cost method of transmission, there are a number of limitations that prevent speed increases and connectivity. These are:

- Address limitation — only up to 16 addresses available.
- Data Skew — as the bits are transmitted in parallel, the faster you drive them, the more likely you are to receive packets with a bit missing.
- Total bus length — this is due to the above. To prevent skew the bus length is reduced. Typically, even with Differential SCSI, it can be no longer than 30 meters, which must include internal SCSI cabling in the devices.
- Crosstalk and electromagnetic interference — to keep cable cost low, the data lines are rarely shielded, which means the more the bandwidth, the greater the possibility of unwanted signals being induced on the lines.

SSA eliminates this by serializing the SCSI data set and using loop architecture that only requires two wires: transmit and receive (SCSI, with data and control lines requires up to 68). The SSA interface also supports full-duplex. This means that it can transmit and receive data simultaneously at full speed. To see how this is applied in practice, we discuss the SSA subsystem.

1.2 SSA subsystem

SSA subsystems are built up of loops of adapters and disks. Contained in these devices are Interface chips called nodes. The adapter contains two SSA nodes and can support two SSA loops. Each disk drive also contains a single SSA node. A node can be either an initiator or a target. An *initiator* issues commands, while a *target* responds with data and status. The SSA nodes in the adapter are, therefore, initiators, while the SSA nodes in the disk drives are targets. Each SSA node is given a unique address at manufacturing time, known as the *UID*. This allows the initiators in the loop to determine what other SSA nodes have been attached to the loop and to understand the SSA loop topology.

SSA is a full-duplex loop based architecture, with two physical read paths and two physical write paths to every disk attached to the loop. This means that when two adjacent SSA nodes are communicating, the rest of the SSA network is free to allow other node to node communications to take place. This concept is known as

spatial reuse; see Figure 2. The SSA architecture allows more than one initiator to be present in a loop. In that case, commands and data from multiple initiators can be directed to the same or different targets and intermixed freely. The SSA network is managed by one particular initiator, known as the *master initiator*. This is the initiator with the highest UID. If a new initiator is added to the network with a higher UID than those already present, it will take over the master responsibilities for that network.

Similarly, if a master initiator is removed from the SSA network, then the initiator with the next highest UID takes over master responsibility. This master takeover and usurpation occurs automatically without any user intervention. Unlike SCSI and fibre channel arbitrated loop, SSA does not use bus phases. Also, SSA does not use arbitration to contend for the bus. Instead, it makes use of frame multiplexing. All data is transported as frames, which are the basic units of data transferred between SSA nodes. A frame can contain up to 128 data bytes. The loop network has several advantageous characteristics:

- Full-duplex support is provided on each link, so that traffic can be present in both directions on the loop simultaneously.
- Since a loop topology is used, should a break occur in the loop (for example a cable is disconnected), each adapter on the loop adjusts its routing algorithms, under direction from the master initiator, so that frames are automatically rerouted to avoid the break. This allows devices to be removed or added to the loop while the subsystem continues to operate without interruption.

SSA operation

- 4 links per loop
 - 2 read and 2 write operations simultaneously in each direction
 - 20/40 MB/sec on each link

Loop Availability

- Loop reconfigures itself dynamically

Spatial Reuse

- Up to 8 simultaneous operations to local group of disks (domains)

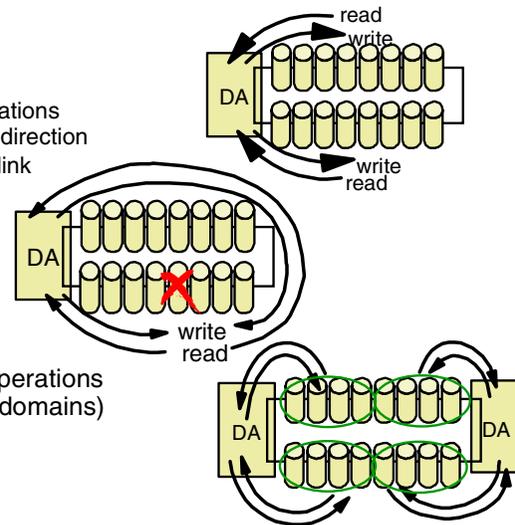


Figure 2. SSA spatial reuse

1.2.1 SSA operation

Data is sent from the adapter card to the first disk on the loop and then passed around the loop by the disks until it arrives at the target disk. Unlike bus based designs, which reserve the whole bus for data transfer, SSA only uses the part of the loop between adjacent disks for data transfer. This means that many simultaneous data transfers can take place on an SSA loop, and it is one of the main reasons that SSA performs so much better than SCSI. This simultaneous transfer capability is known as spatial reuse.

Each read or write path on the loop operates at 40MB/s, providing a total loop bandwidth of 160MB/s.

1.2.2 Loop availability

The loop is a self-configuring, self-repairing design which allows genuine hot-plugging. If the loop breaks for any reason, then the adapter card will automatically reconfigure the loop into two single loops. The most likely scenarios for a broken loop are SSA cables becoming detached or if the actual disk drive interface electronics should fail. If this should happen, the adapter card will dynamically reconfigure the loop into two single loops, effectively isolating the failed disk. If the disk were part of a RAID array, the adapter card would automatically regenerate the missing disk using the remaining data and parity disks to the spare disk. Once the failed disk has been replaced, the loop will automatically be reconfigured into full-duplex operation.

1.2.3 Spatial reuse

Spatial reuse allows domains to be set up on the loop. A domain means that one or more groups of disks “belong” to one of the two adapter cards, as is the case during normal operation. The benefit of this is that each adapter card can talk to its domains (or disk groups) using only part of the loop. The use of domains allows each adapter card to operate at maximum capability, because it is not limited by I/O operations from the other adapter. Theoretically, each adapter card could drive its domains at 160MB/s, giving 320MB/s throughput on a single loop!

1.2.4 Compatibility

Because SSA allows SCSI-2 mapping, all functions associated with initiators, targets, and logical units are translatable. Therefore, SSA can use the same command descriptor blocks, status codes, command queuing, and all other aspects of current SCSI systems. The effect of this is to make the type of disk subsystem transparent to the application. No porting of applications is required to move from traditional SCSI I/O subsystems to high performance SSA. SSA and SCSI I/O systems can coexist on the same host running the same applications.

1.2.5 Summary

The advantages of SSA are summarized as follows:

- Dual paths to devices.
- Simplified cabling — cheaper, smaller cables and connectors, no separate terminators.
- Faster interconnect technology.
- Not an arbitrated system.
- Full-duplex, frame multiplexed serial links.
- 40MB/s total per port, resulting in 80MB/s total per node, and 160MB/s total per adapter.
- Spatial reuse allows concurrent activity in different portions of the same loop.
- Hot-pluggable cables and disks.

- Very high capacity per adapter — up to 127 devices per loop, although most adapter implementations limit this. For example, current IBM SSA adapters provide 96 disks per Micro Channel or PCI slot.
- Distance between devices of up to 25m (82ft) with copper cables, 2.4km (1.5 miles) with optical links.
- Auto-configuring — no manual address allocation.
- SSA is an open standard.
- SSA switches can be introduced to produce even greater fan-out and more complex topologies.

1.3 Storage area network (SAN)

SSA is an important component of the SAN architecture. For a comprehensive overview of this subject, read *Introduction to Storage Area Network, SAN*, SG24-5470.

Storage area networks are an emerging architecture which are only now starting to be implemented in the IT industry. SANs are fibre-based networks which will be used primarily for data transfer between servers and storage devices of similar, or non-similar, manufacturers.

SSA has been a part of the initial SAN implementations throughout the world, but the industry is heading towards a fibre-based storage network. SSA will still play an important part in a SAN solution, because SSA can be connected to a fibre SAN through the use of SSA to FC converters. The installation of these converters allows you to take advantage of the tremendous I/O throughput capabilities, as well as the availability benefits of SSA in your SAN environment.

The SSA to FC protocol converters are available from either IBM or VICOM.

The IBM product is the IBM 7139 Model 111. Options include a rack mount enclosure for up to four units. Either one short-wave GBIC (Gigabit Interface Converter) for up to 500 meters or one long-wave GBIC for up to 10 kilometers between the server or IBM SAN fibre channel switch is required.

The Fibre Channel SLIC Router provides fibre channel attachment for Serial Disk storage to a wide range of UNIX servers including Sun and HP and to Intel-based servers with Windows NT. Attachment to IBM SAN Fibre Channel Switches is also provided. For current supported servers and switched fabric, visit:

www.vicom.com

Like the IBM SAN Data Gateway for Serial Disk, the Fibre Channel SLIC Router provides host-independent disk mirroring for high availability, disk concatenation for large volume groups, and Instant Copy for creating a separate copy of data for tape backup.

The Fibre Channel SLIC Router maximizes fibre channel host adapter bandwidth potential by minimizing FC-AL arbitration. Only one fibre channel target (SLIC Router) exists on the FC-AL loop, while up to 64 IBM Serial Disks may be attached to the serial disk loop. The router provides SSA-160 disk performance with IBM Advanced 7133 Models D40 and T40. Up to 90MB/s sustained throughput, or up to 25,000 I/Os per second per logical channel in full-duplex, simultaneous read/write mode has

been measured by VICOM. Scalable performance is provided by adding additional routers, up to eight host attachments are supported for each serial loop.

The VICOM product is called a SLIC Router FC-SL. For more information on how to implement these converters, refer to 2.4.4, "VICOM SLIC Router FC-SL" on page 26, or visit the VICOM Web page at:

www.vicom.com

Chapter 2. SSA hardware components

This chapter discusses the various components required to make up a functioning SSA subsystem.

2.1 SSA adapters

As mentioned before, SSA is an open standard and has adapters and devices available to allow connections to different hardware platforms and operating systems, and it offers RAID and non-RAID options.

All adapters require device drivers and microcode to make them function properly. When you have multiple SSA adapters in the loop, it is essential that they are all at the same level, or, when you have a supported mixed adapter configuration, they are at a compatible level. This code along with installation instructions can be found at the SSA Web site:

<http://www.hursley.ibm.com/ssa/index.html>

From time to time, revisions are made to this code, and we highly recommend that you keep your code as current as possible. To keep informed of the revisions, you can subscribe to the SSA Web site and receive email when changes have been added.

Figure 3 shows the back view of the SSA adapter as seen from the outside of the server machine.

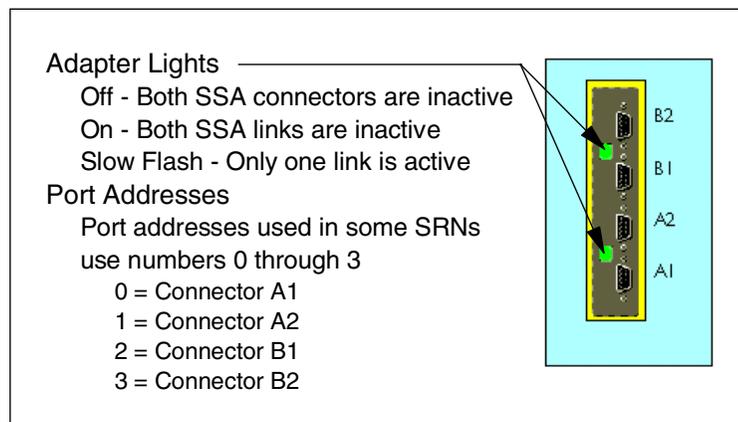


Figure 3. SSA adapter connection ports and LEDs

The tables (Table 1 on page 8 through Table 8 on page 11) list the various adapters and their characteristics. All adapters except for feature code 4003 support two loops. All adapters, except for feature code 4012, can support 48 disks per loop. The total loop bandwidth is either 80MB/s or 160MB/s, depending on adapter type. For Micro Channel (MCA) systems you can have up to four adapters per MCA bus in the system. Also, for PCI-type systems, you can have four adapters per system unless otherwise indicated, either in the table, or in the *PCI Adapter Placement Reference*, SA38-0538. For the latest copy, visit:

http://www.rs6000.ibm.com/resource/hardware_docs/#index4

Table 1 shows specifications for the non-RAID SSA adapters on the RS/6000.

Table 1. RS/6000 non-RAID SSA adapters

Features	Adapter label 4-D	Adapter label 4-G
Feature code	6214	6216
Code Name	Mayflower	Mayflower
AIX Software level	3.2.5, 4.1.4, 4.1.5, 4.2.1, 4.3.X	
System interface	MCA	
Link speed	20MB/s	20MB/s
Max adapters per loop	2	8
Adapter co-existence	4-D and 4-G	4-G ¹ and 4-D ²
RAID type supported	NONE	
Boot support	YES	
Dump tool	NO	

Table 2 shows the specifications for the RAID SSA adapters on the RS/6000 platform.

Table 2. RS/6000 RAID SSA adapters

Features	Adapter label 4-I	Adapter label 4-J
Feature code	6217	6218
Code Name	Sarratoga	Beaulieu
AIX Software level	4.1.5, 4.2.1, 4.3.X	
System interface	MCA	PCI
Link speed	20MB/s	20MB/s
Max adapters per loop	1	1
Adapter co-existence	None	
RAID type supported	RAID 5	
Hot Spare supported	YES	
Fast Write Cache (FWC) option	NO	
Hot Spares span loops	YES	
Boot support	Non RAID	NO
Dump tool	NO	

Table 3 shows the RS/6000 enhanced RAID SSA adapters specifications.

Table 3. RS/6000 enhanced RAID SSA adapters

Features	Adapter label 4-M	Adapter label 4-N
Feature code	6219	6215

Features	Adapter label 4-M		Adapter label 4-N	
	Fremont	Fremont 2	Campbell	Campbell 2
Code Name	Fremont	Fremont 2	Campbell	Campbell 2
AIX Software level	4.1.5 4.2.1 4.3.0 4.3.X	4.1.5 4.2.1 4.3.1 4.3.X	4.1.5 4.2.1 4.3.0 4.3.X	4.1.5 4.2.1 4.3.1 4.3.X
System interface	MCA		PCI	
Link speed	20MB/s		20MB/s	
Max adapters per loop	2 ³	8 ⁵	2 ³	8 ⁵
Adapter co-existence	4-M ³ 4-N ³	4-M ⁶ 4-N ⁶ 4-P ⁶	4-N ³ 4-M ³	4-M ⁶ 4-N ⁶ 4-P ⁶
RAID type supported	RAID 5			
Hot Spare supported	YES			
Fast Write Cache (FWC) option	4MB			
Hot Spares span loops	NO			
Boot support	Non RAID		NO	
Dump tool	Requires Hot Spare configured			

Table 4 shows the RS/6000 Advanced RAID SSA adapter specifications.

Table 4. RS/6000 advanced serial RAID adapter

Features	Adapter label 4-P
Feature code	6225
Code Name	Santa Cruz
AIX Software level	4.2.1, 4.3.2
System interface	PCI
Link speed	40MB/s or 20MB/s
Max adapters per loop	8 (5)
Adapter co-existence	4-N ⁶ , 4-M ⁶ , 4-P ⁶
RAID type supported	RAID 0, RAID 5
Hot Spare supported	YES
Fast Write Cache (FWC) option	32MB
Hot Spares span loops	NO ⁴
Boot support	NO
Dump tool	Requires Hot Spare configured

Table 5 shows the RS/6000 advanced serial RAID plus adapter specifications.

Table 5. RS/6000 advanced serial RAID plus adapter

Features	Adapter label 4-P
Note: This adapter is identical to FC6225, but with higher microcode and larger cache. Updating microcode on FC6225 will give the extra functionality but not the performance uplift.	
Feature code	6230
Code Name	Santa Cruz 2
AIX Software level	4.2.1, 4.3.2, 4.3.3
System interface	PCI
Link speed	40MB/s or 20MB/s
Max adapters per loop	8 (5)
Adapter co-existence	4-N ⁶ , 4-M ⁶ , 4-P ⁶
RAID type supported	RAID 0, RAID 1 ⁷ , RAID 5, RAID 0+1 ⁸
Hot Spare supported	YES
Fast Write Cache (FWC) option	32MB
Optional DRAM (FC 6231)	128MB DRAM (standard is 64MB)
Hot Spares span loops	NO (4)
Boot support	NO
Dump tool	Requires Hot Spare configured

Notes for RS/6000 adapters

1. Up to eight 4-G adapters may be connected in a loop.
2. This is the maximum configuration. A 4-D adapter can only have two adapters in a loop.
3. If enhanced RAID adapters are connected together, then RAID and fast-write cache cannot be used.
4. Members of a RAID array must be on the same loop. If hot spare is enabled for the array, the spare disk must be on the same loop as the array.
5. If more than two adapters are connected, RAID and fast-write cache cannot be used.
6. Two adapters per loop with RAID but no fast-write cache, up to eight adapters per loop then RAID and fast-write cache cannot be used.
7. 1- or 2-way RAID-1 with 2 disks per array.
8. 1- or 2-way RAID-0+1 with 4-16 disks (up to 8 mirrored pairs of disks per array).

Table 6 shows the specifications for the SBus SSA adapter for the Sun systems servers.

Table 6. Sun Systems SBus SSA adapters

Features	Adapter label
Feature code	4003
Code Name	Sunlight
Software level	Solaris 2.4, 2.5.1
System interface	SBus
Link speed	20MB/s
Max adapters per loop	4 (only 1 per system)
Adapter co-existence	NONE
This adapter only supports one loop and only two adapters per system which cannot be joined in the same loop.	

Table 7 shows the open systems PCI RAID SSA adapters specifications.

Table 7. Open systems PCI RAID SSA adapters

Features	Adapter label		
Feature code	4011	4012	4013
Code Name	Los Gatos 1	Los Gatos 2	Los Gatos 3
Software levels	Windows NT OS/2 Novel NetWare DOS	Windows NT	Windows NT OS/2 Novel NetWare
System interface	PCI	PCI	PCI
Max adapters per loop	1	2	2
Adapter co-existence	NONE	NONE	NONE
RAID type supported	RAID 0, RAID 1 RAID 5 ³	RAID 1 ^{1,4}	RAID 5 ^{2,4}

Notes on open systems adapters

1. RAID only with single adapter
2. Two adapters per loop with RAID and non Fast Write Cache
3. Two adapters per system bus
4. Three adapters per bus

Table 8 shows the advanced serial RAID/X SSA adapter specifications.

Table 8. Advanced serial RAID/X adapter

Features	Adapter label
7133 feature code Option number	4040 p/n 09L2123
Code Name	Santa Cruz 2

Features	Adapter label
Operating system	Windows NT 4.0, Novel NetWare 4.2, 5.0
System interface	PCI
Link speed	40MB/s or 20MB/s
RAID type supported	RAID 0, RAID 1, RAID 5, RAID 0+1
Adapters per loop	RAID 1, RAID 5, RAID 0+1 -- two RAID 0 -- one
Fast Write Cache (FWC) option	32MB

2.2 SSA disk subsystems

Three SSA subsystems are discussed in this section:

- “3527 SSA storage subsystem for PC servers” on page 12
- “7131 SSA Multi-storage Tower Model 405” on page 13
- “7133 Serial Storage Architecture (SSA) disk subsystem” on page 13

2.2.1 3527 SSA storage subsystem for PC servers

The IBM 3527 SSA storage subsystem for PC servers uses SSA technology to deliver outstanding performance, highly scalable capacity, and enhanced availability.

The 3527 SSA storage subsystem uses the SSA RAID adapter for PC servers to run under a variety of operating systems, such as:

- IBM Windows NT 3.5.1 or 4.0, Novel Netware 4.1 or 4.1 SMP, OS/2 Warp Server or Warp Server SMP, and OS/2 2.11 SMP
- Compaq Windows NT 4.0
- HP Windows NT 4.0

The 3527 SSA storage subsystem uses the SSA RAID adapter for PC servers and supports various PC servers, such as:

- IBM PC Server families PC320, PC325, PC330, PC520, PC704 and PC720
- Compaq ProLiant 1500, 2500, 5000 and 6000
- HP NetServer LH Pro and LX Pro

Table 9 lists the 3527 specifications.

Table 9. 3527 SSA storage subsystem for PC server

Item	Specification
Transfer Rate SSA Interfaces	80 MB/s drive interface bandwidth
Configuration	2 to 5 disks (4.5 GB or 9.1 GB) per subsystem
Configuration range	9.0 GB-22.5 GB (with 4.5 GB disks) 18 GB-45.5 GB (with 9.1 GB disks)
Supported RAID level	0, 1, and 5
Hot swappable disks	Optional

2.2.2 7131 SSA Multi-storage Tower Model 405

The 7131 SSA Multi-storage Tower Model 405 provides a solution for those who require the high performance of SSA and for applications that exceed the capacity of direct SCSI attached storage. The 7131 Model 405 delivers outstanding performance, high capacity, and enhanced availability.

Note: The model 405 may no longer be available for purchase, but as many of you still have them in your configuration, information is provided here to help you.

The 7131 SSA Multi-storage Tower Model 405 runs under various operating systems, such as:

- IBM — AIX 4.1.4 (and above) and AIX 4.2, Windows NT 3.51 or 4.0, Novel NetWare 4.1 or 4.1 SMP, OS/2 Warp Server or Warp Server SMP, and OS/2 2.11 SMP
- HP — HP-UX 10.01, 10.10, or 10.20
- Sun — Solaris 2.4, 2.5, or 2.51

The 7131 SSA Multi-storage Tower Model 405 is supported on various systems, such as:

- IBM — Supported on RS/6000 and RS/6000 SP systems with adapters feature 6214, 6216, 6217, 6218, and PC Server systems with the SSA RAID adapter for PC Server.
- HP — HP 9000 Series 800 with the IBM 7190 SCSI Host to SSA Loop Attachment
- Sun — Supported on selected SPARCstation, SPARCserver, SPARCcenter, and Ultra Enterprise models with the IBM 7190 SCSI Host to SSA Loop Attachment

Table 10 lists the 7131 SSA Multi-storage Tower Model 405 specifications.

Table 10. 7131 SSA Multi-storage Tower Model 405 specifications

Item	Specification
Transfer rate SSA interface	80 MB
Configuration	2 to 5 disk drives (4.5 GB or 9.1 GB) per subsystem
Configuration range	Up to 4.4 GB to 11 GB (with 2.2 GB disk drives) Up to 9.0 GB to 22.5 GB (With 4.5 GB disk drives) Up 18.2 GB to 45.5 GB (With 9.1 GB disk drives)
Supported RAIDS	0, 1, and 5
Hot swap disks	Yes

2.2.3 7133 Serial Storage Architecture (SSA) disk subsystem

The IBM 7133 SSA disk subsystem uses SSA technology to deliver outstanding performance, highly scalable capacity, and enhanced availability. The 7133 contains four strings of four disk drives. Each string of drives has two SSA connections, one at each end. Strings can be connected together to form larger configurations. Figure 4 shows a photograph of a 7133 tower and drawer unit.



Figure 4. 7133 Stand-alone Disk Tower and 7133 Disk Drawers in S00 frame

The 7133 models 10 and 500 were the first IBM products announced in 1995 with the revolutionary new serial storage architecture. Some IBM customers still use the Models 10 and 500, but these have been replaced, respectively, by the 7133 Model 20 and 7133 Model 600, and now the 7133 model D40 and 7133 model T40.

The 7133 Models 10, 20, 500, 600, D40 and T40 have redundant, hot swappable, power and cooling.

The 7133 runs under various operating systems, such as:

- IBM — AIX 3.2.5 (with PTFs), AIX 4.2, Windows NT 3.51 or 4.0, Novel NetWare 4.1 or 4.1 SMP, OS/2 Warp Server or Warp Server SMP, and OS/2 2.11 SMP
- HP — HP-UX 10.01, 10.10, or 10.20
- Sun — Solaris 2.4, 2.5, or 2.5.1

The 7133 supports various systems, such as:

- IBM — RS/6000 and RS/6000 SP systems with SSA adapters FC 6214, FC 6216, FC 6217, and FC 6218 and on PC Server systems with the SSA RAID adapter for PC Server.
- HP — HP 9000 series 800 with the IBM 7190 SCSI Host to SSA Loop Attachment
- Sun — Supported on selected SPARCstation, SPARCcenter, and Ultra Enterprise models with the IBM 7190 SCSI Host to SSA Loop Attachment

Table 11 lists the 7133 Models specifications.

Table 11. 7133 Models 10, 20, 500, 600, D40 and T40 specifications

Item	Specification
Transfer rate SSA interface	80 MB/s for Models 10, 500, 20, 600 80/160MB/s for Models D40 and T40
Configuration	4 to 16 disks (1.1 GB, 2.2 GB, 4.5 GB, for Models 10, 20, 500, and 600, and 9.1 GB for Models 20 and 600 only. Models D40 and T40 use 4.5GB, 9.1GB, 18.2GB and 36.4GB new style disks only. With 1.1 GB disk drives you must have 8 to 16 disks.
Configuration range	8.8 to 17.6 (with 1.1 GB disks) 8.8 to 35.2 GB (with 2.2 GB disks) 18 to 72 GB (with 4.5 GB disks) 36.4 to 145.6 GB (with 9.1 GB disks) 72.8 to 291.2 GB (with 18.2 GB disks) 145.6 to 582.4 GB (with 36.4 GB disks)
Supported RAID level	Dependent on adapter, 0, 1, 5 and 10
Hot swappable disk	Yes (and hot swappable redundant power and cooling)

These rules must be followed when connecting a 7133:

- Each SSA loop must be connected to a valid pair of connectors on the SSA adapter card:
 - A1 and A2 to form one loop, or B1 and B2 to form one loop.
 - Only one pair of connectors of a SSA adapter can be connected in a particular SSA loop.
 - A1 or A2 with B1 or B2 of the same adapter cannot be in the same SSA loop.
- A maximum of 48 disks can be connected in a SSA loop.
- A maximum of three dummy disk drive modules can be connected next to each other.
- The maximum number of adapters per SSA loop. See adapter tables (Table 1 on page 8 through Table 8 on page 11).
- SSA adapter loop compatibility for connecting different SSA adapters in the same loop. See adapter tables (Table 1 on page 8 through Table 8 on page 11).
- Maximum of two adapters in the same host per SSA loop.
- There is no cable length restriction.
- There is no addressing setup for any SSA device.
- There is no termination.

2.2.3.1 7133 models 10/500

The 7133 Model 010/500 subsystem has the following components:

- Front panel with:
 - On/Off switch
 - Power and check indicator lights

- Four strings of four drives for a total of 16 drive locations. Each string has two SSA connections, one at each end of the string.
- Two fan and power supplies and dummy fan and power supply assemblies, or three fan and power supply assemblies.

Any slot not having a disk drive assembly must have a dummy disk drive assembly installed. These components are:

- External Indicators
 - Power light, green when 7133 is turned on
 - Check light, amber if a fault occurs in the 7133 subsystem
- Power Supply/Fan Assembly Indicators
 - Power light, green when power is present to the assembly
 - Check light, amber if the power supply is defective, or flashes if the fan has failed
- Power Card Indicators
 - Power light, green when power is present to the assembly

Figure 5 shows the internal cabling within the Model 010/500 drawer.

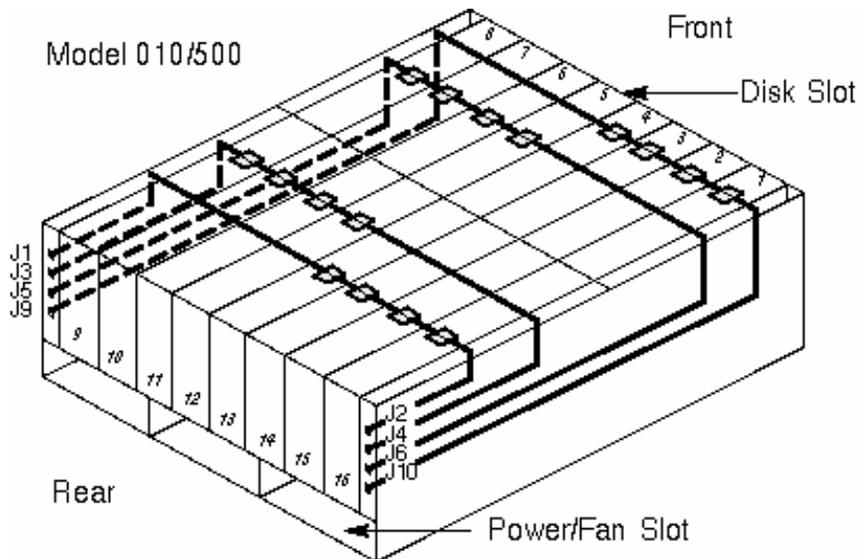


Figure 5. 7133 model 10/500 disk and connection layout

Table 12 shows the disk to cable connector assignments.

Table 12. Disk to cable connector assignments models 10/500

Drives	1st connector	2nd connector
1 - 4	J10	J9
5 - 8	J6	J5
9 - 12	J3	J4
13 - 16	J1	J2

Jumper cables have to be used to connect the strings, if not attached to an adapter.

2.2.3.2 7133 models 20/600

The 7133 models 020/600 are similar to 7133 Models 010/500. The 7133 Model 020/600 has the high availability hardware known as the bypass card which allows the drawer to function in either bypass mode or forced inline mode allowing continuous access to the SSA links.

Bypass circuit

There are four bypass circuit cards which can operate in two modes independently selected by jumpers on each card.

- **Bypass mode**

In normal operation, the SSA links will be through the external connectors at the rear of the 7133-020/600. This is called the Inline State. On loss of power to both SSA connectors forming a pair, the SSA connection will be made internally between the two connectors. This is called the bypass state.

Therefore, if an SSA adapter is connected to a valid pair of connectors, in the event of the system that the adapter is installed is powered down, the link that would have been through the adapter will now be through the bypass circuit, thus maintaining a continuous SSA loop.

- **Forced inline mode**

In this mode the bypass circuit acts as a 7133-010/500. The link will never be made through the bypass circuit.

On the 7133 model 020/600 the SSA connectors were rationalized and are labeled 1, 3, 5, 8, 9, 12, 13 and 16. Each label corresponds to the disk slot position in the 7133 to where the connector is wired.

Figure 6 shows the internal cabling within the Model 020/600 drawer.

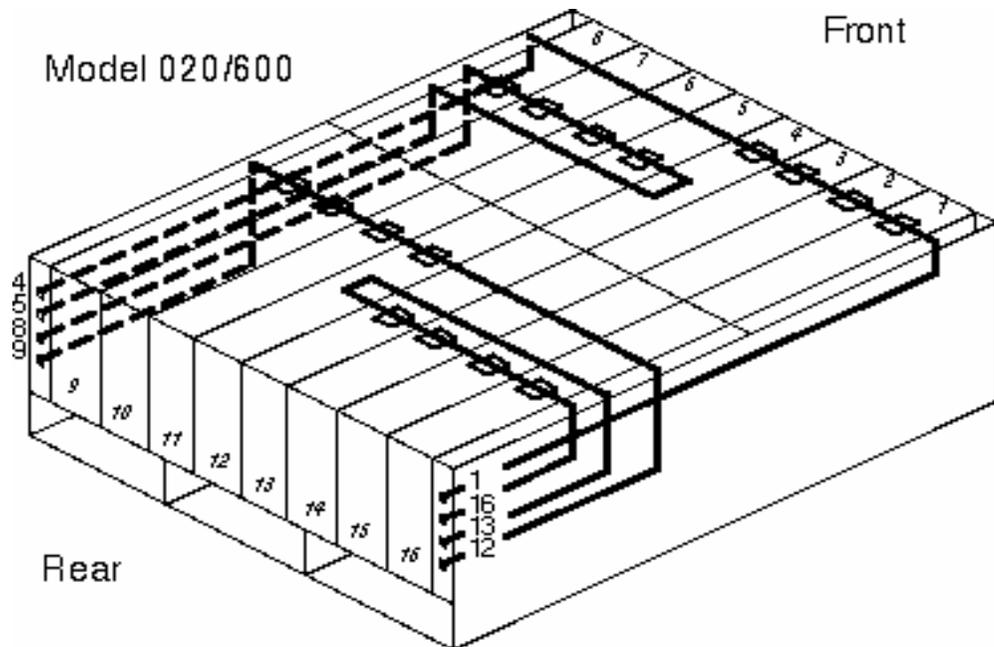


Figure 6. 7133 model 20/600 disk and connector layout

The bypass circuit connector pairings are:

- 1 and 16
- 4 and 5
- 8 and 9
- 12 and 13

For correct operation of the bypass circuit, connect the SSA adapter ports A1 and A2 or B1 and B2 to one of the pairs of connectors listed at the rear of the 7133 model 020/600 and shown in Figure 6.

Each bypass circuit card has three status lights: link status (ready), a green LED for both Port 1 and Port 2, and the mode LED (dual color green/amber). Table 13 and Table 14 show the link and mode LED description for the various status combinations.

Table 13. Link LEDs

Link LED status	Description
Permanently ON	The interface through the bypass card is fully operational.
Slow flash	The interface through the bypass card is not operational.
Permanently OFF	The card is in bypass state or forced inline mode.

Table 14. Mode LEDs

Mode LED status	Description
Permanently ON (amber)	The bypass card is in bypass state.
Permanently ON (green)	The bypass card is jumpered to forced inline state.
Permanently OFF	The bypass card is in inline state.

2.2.3.3 7133 models D40/T40

The 7133 Model D40/T40 storage subsystems are similar to the 7133 models 20 and 600 with the same disk and connector layout. T40 is the stand-alone tower, and D40 is the rack installed drawer. T40/D40s may be intermixed in loops with models 010, 020, 500, and 600.

This drawer is only supported at AIX 4.1.5, 4.2.1, 4.3.1 and higher.

Two power supplies are located at the rear of the subsystem using N+1 technology, so either power supply can operate a fully configured drawer. Three fan assemblies are located at the front of the subsystem. Again these are N+1, and any 2 fans may cool a fully configured drawer. These items are all hot swappable.

The drawer, like the previous models, contains 16 drive positions, supporting 36.4GB, 18.2GB, 9.1GB, and 4.5GB drives. These SSA disk drives are physically different from previous models of the 7133 and are not interchangeable with them.

The drawer also has an Intelligent Controller card to monitor the subsystem environment and provide SES functions which allows the T40/D40 to show up as a device in AIX, for example, enclosure 0 (zero).

With this card it is also capable of deciphering whether a drive or dummy carrier is in each slot. The card is hot-pluggable, therefore, removing the controller card will cause the fan speed to increase.

The card also performs these functions:

- Monitors ambient drawer air temperature with 3 sensors
- Monitors fan speed/status, power supply status, and bypass card settings/status
- Monitors RS-232 interface through power supply 2, if used
- Monitors remote power on feature through power supply 1, if installed
- Controls front panel button and LED information
- Provides interface for SES functions

Each unit has an enhanced control panel displaying error IDs, drive status LEDs, and has buttons to set the ID. A four character alpha-numeric display, shows the enclosure ID. Two rows of eight LEDs (top row for rear drives, bottom row for front drives) indicate the disk drive status. Figure 7 shows the operator panel lights and control options.

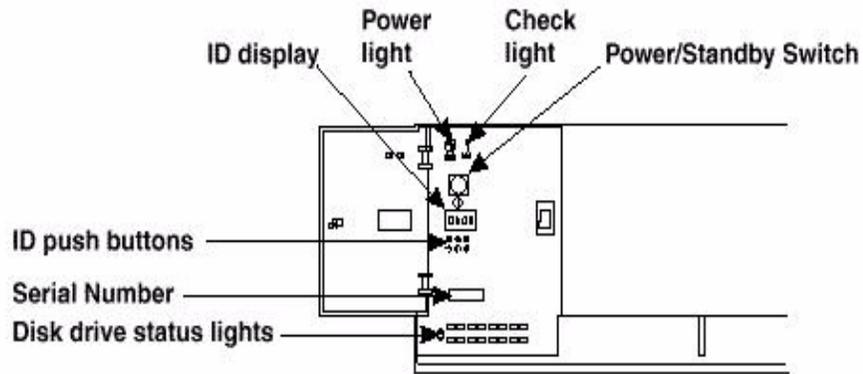


Figure 7. 7133 D/T40 operator panel

If all three operator panel buttons are depressed for greater than two seconds the controller will attempt to set all the bypass card modes to normal. This function is required to allow you or the IBM CE to recover from setting the mode that limited further access to the controller, for example, setting the bypass card that is being used to access the drawer, into forced bypass.

Bypass cards

Four SSA node bypass cards (SNBC) are located in the rear of the unit, with three LEDs each. All four cards are physically and functionally identical in these ways:

- Do not swap between 7133s (EEPROM has bypass configuration data)
- Hot swappable
- Modes (bypass/forced in-line) can be set using RSM (SES commands)
- Forced inline mode can be set manually
- SNBC is fitted with a PAIR of jumpers — one for each half of card
- Forced inline mode allows total isolation of quads
- Three pins — always install jumper to center pin

Table 15 shows the mode LED status indicators.

Table 15. LED state and meanings models D40/T40

Op mode	Status	Status LED	Mode LED	Status LED
Auto	Inline	ON	OFF	ON
Auto	Bypass	OFF	Amber	OFF
Forced Inline	Inline	OFF	Green	OFF
Forced Bypass	Bypass	ON	Amber	ON
Forced Open	Open	OFF	OFF	OFF
Jumper Forced Inline	Inline	OFF	Green	OFF

Power switch-on model D40

This is labeled Power/Standby, and controls electrical power from the power supply units within the 7133 unit to the disk drives and other components. When internal power is removed, the 7133 is in standby mode and the power light is

flashing, which means mainline electrical power is present in the power supply assembly.

Power switch-on model T40

This is the same as above, except when the mainline power is switched off, no mainline power is present in the power supply assembly.

Power light

This is continuously on when the power switch has been pressed to switch on the 7133. The light flashes when the 7133 is in standby mode.

Check light

The amber light comes on if a failure occurs within the 7133 unit. The unit may operate satisfactorily even when the failure of a single part has been detected. The check light can be set to flashing by an SSA service aid to identify a disk drive.

ID display

The ID display is a four-character display that can be set by the three ID push buttons. This display normally shows the ID of the 7133 unit. If a failure occurs within the 7133 unit, an error ID for the failure is shown, then the display alternates between the four-character ID and the failure detected.

ID push button

Used to set the ID display of the 7133 unit that is shown in the ID display.

Enclosure microcode

To enable the controller card to function this model has downloadable microcode. The minimum level for the ROS level and ID must be 0009. The latest level for the ROS level and ID is version 0011. This level, with the 4-P SSA adapter and drivecode and AIX prerequisites, supports 160MB/s bandwidth.

To verify the level code present on the controller, use this *lscfg* command:

```
lscfg -vl enclosureX
```

To download the latest level, refer to the 7133 D40/T40 enclosure microcode at this Web page:

<http://www.hursley.ibm.com/~ssa/rs6k/index.html>

2.3 SSA disks

SSA disks, like most IBM disk storage, have undergone tremendous changes in capacity since they were originally introduced. They have gone from 1.1GB to 36.4GB in the same footprint size in a few years, and they are also now available in 10,000 rpm versions for certain disk sizes.

2.4 Other SSA components

This section lists several other components which will help you take advantage of the SSA architecture.

2.4.1 IBM 7190-100 SCSI host to SSA loop attachment

The 7190-100 can be considered a gateway that enables other systems that support SCSI Fast/Wide attachment to work with most IBM SSA subsystems. So with the 7190-100, you can improve the performance of your Sun, HP, and DEC Fast Wide SCSI adapters and device drivers.

Note: The 7190-100 has now been replaced, but many of you still have them in your configuration, so information is provided here to help you.

2.4.1.1 Powerful SCSI host to SSA loop attachment

As the number of storage devices increases on a SCSI bus, performance decreases dramatically.

The IBM 7190-100 SCSI host to SSA loop attachment is an external device that enables attachment of high-performance SSA subsystems, such as IBM 7131 Model 405 and the 7133 models 010, 500, 020 and 600, to existing SCSI-based systems.

Key features include:

- The 7190-100 overcomes key SCSI limitations by providing increased bandwidth, reliability, number of devices per channel, and cabling flexibility of the SSA.
- The 7190-100 provides maximum SCSI data throughput of as much as 18 MB/s, with up to 1,900 I/O per second.
- The 7190-100 supports up to 48 SSA disks and 437GB on a single loop, but only four 7190-100 units per loop. You can run any combination of the above SSA disks of 2.2GB, 4.5GB, or 9.1GB disks.
- The 7190-100 provides outstanding scalability and performance. It maintains the SCSI features, such as tagged command queuing, scatter/gather, disconnect/reconnect, and synchronous, as well as asynchronous data transfers. It provides the SSA technology and offers up to 80MB/s of bandwidth for faster and more efficient data transfers. It supports the spatial reuse characteristics, and allows the attached adapter and host to transfer data simultaneously to and from the loop SSA disks. In addition, SSA's unique loop architecture and auto-addressing capability enables disks to be added without powering down the systems.

2.4.1.2 High reliability for improved availability

SSA's node architecture provides high reliability for critical storage applications by eliminating single points of path failure on the loop. Should an SSA loop be broken, the 7190 automatically reconfigures the loop into two strings — transparently to the host — with no loss of data. When the defective SSA cable is replaced, the device reconfigures the disk into a loop once again. Additionally, the SSA disk drives are hot swappable, allowing users to add or remove SSA devices while the disk subsystem is powered on. Applications running on the host system continue to have access to existing disks on the SSA loop. Furthermore, the 7190-100 attachment uses parity to ensure that the highest data integrity is maintained. This parity information is designed to protect the data transmitted between the SCSI host adapter and the IBM 7190. Data on the serial SSA loop is protected by cyclical redundancy check bits.

Online service monitoring

Service features built into the 7190 allow Sun and HP systems to efficiently monitor all attached storage subsystems. These online capabilities include:

- Displaying vital product data (VPD)
- Showing the SSA loop topology
- Showing the SSA to SCSI, ID to LUN mapping
- Setting or changing modes for drive identification and diagnostics
- Downloading microcode to the 7190-100 and SSA disk drives as required
- Providing a detailed system wide error log
- Alerting the system to errors in the redundant power and cooling systems and loop topology changes

The IBM 7190-100 SCSI host to SSA loop attachment runs under various operating systems and hardware platforms. Table 16 shows a list of supported servers and operating systems. For the latest supported list, refer to the SSA Web page:

<http://www.hursley.ibm.com/ssa/index.html>

Table 16. Supported servers and operating systems for 7190-100

Servers	Models	Adapters	Operating System
Intel - based			Windows NT 3.5 through 4.0 NT 4.0 requires SP 3 for > 24 drives per loop
Sun	Sun SuperSPARC and UltraSPARC SPARCstation 10 and 20 SPARCserver Models 1000 and 1000E, SPARCcenter Models 2000 and 2000E Ultra Enterprise Models 2, 150, 3000, 4000, 5000, and 6000	SCSI F/W diff X1062A Ultra F/W X1065A (7190-200 only)	Sun solaris 2.4, 2.5 or 2.51 for Sun servers For 7190-200 - Solaris 2.5.1, 2.6
HP-UX	HP series 8000 HP 9000 Enterprise servers - includes D,K,T,E,G,H, and I class servers	HP-PB model 16-bit diff SCSI adapter(28696-600 01) HP-HSC 20 MB FWD SCSI-2 adapter (A2969A) HP K-class native FWD SCSI-2 adapter	7190-100 HP-UX 10.01, 10.10, or 10.20 for HP Servers 7190-200: HP-UX 10.01, 10.10, 10.20, 10.30
DEC	Digital AlphaServer 300, 800, 1000, 1000A, 2000, 2100, 2100A, 4000, 4100, 8200, and 8400 series	SCSIF/W diff KZPSA-BB UltraSCSI model KZPBA-CA (7190-200 only)	Digital UNIX 3.2B through 3.2G and 4.0 through 4.0C for digital servers

2.4.2 IBM 7190-200 ultra SCSI host to SSA loop attachment

The 7190-200 ultra SCSI host loop attachment announced on April 21 1998, provides HP, Sun, and Digital products with the flexibility and performance of serial disk storage.

SSA Management is the same as in 7190-100, but greatly improved. The 7190-200 offers these features:

- Enables SCSI-based servers to benefit from the performance improvements of non-arbitrated serial disks. Each ultra SCSI adapter performs at a sustained speed up to 2600 I/O per second and 29MB/s depending on the server and SCSI adapter used. Up to four 7190-200 interfaces can be connected to the SSA loop. The potential sustained performance is increased by 2600 I/O per second and 29MB/s for each additional 7190-200.
- Supports advanced SCSI features which enhance the performance of HP, Sun, and Digital adapters and device drivers. SCSI tagged command queuing, scatter/gather, and disconnect/reconnect benefit from the 7190-200 interface's ability to perform multiple simultaneous reads and writes to multiple non-arbitrated SSA disks.
- The SCSI adapter sees the 7190-200 as a single SCSI target and the SSA disks as parallel SCSI logical unit numbers (LUNs). The 7190-200 automatically groups up to 48 SSA disks into 15 balanced SCSI LUNs. The LUNs can be configured through administrative utilities and can be used by list-based RAID managers.
- Enables highly scalable configurations for HP, Sun, and Digital servers. Each server may have multiple SCSI adapter connections to a single loop of an SSA disk, providing redundant connections to data. A pair of servers with redundant SCSI adapter connections provides the highest level of availability and facilitates remounting a file system from a failed host to a backup host. Additional software may be required.
- Fibre optic extenders enable the 7133 to be geographically separated by up to 2.4km (1.5 miles) while maintaining the high transaction rates of the 7190-200.
- Supports up to 48 disks and 437 GB on a single SSA loop. Disk drives and servers may be added or removed while the entire configuration remains online and operational. Each SSA component, down to the individual power supply, disk drive, and interface adapter can be remotely configured and diagnosed by using the server's existing remote console facility.
- Includes a drawer option for integrating up to four 7190-200s into a single rack-mounted drawer. The drawer of 7190-200 may be connected to over 1.7TB of 7133 SSA disk drawers delivering up to 10.400 I/O per second and 116MB/s across full ultra SCSI ports.

2.4.3 IBM 2108 SAN Data Gateway Model S20

The IBM 2108 SAN Data Gateway provides a protocol converter from SCSI to SSA for your customers who want to connect the SSA drives to your SCSI adapters in your host servers. This allows you to take advantage of all the SSA benefits if you are not able to use native SSA adapters.

Figure 8 shows a photo of the 2108 S20 SAN Data Gateway.

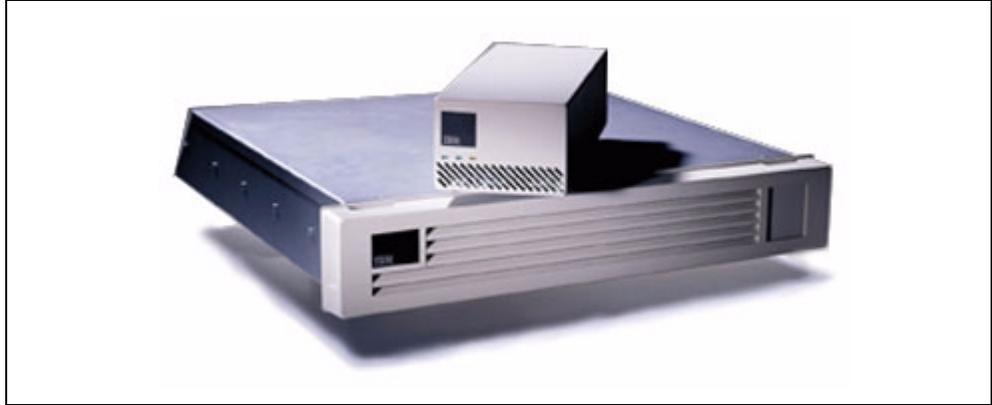


Figure 8. The 2108 Model S20 Drawer and stand-alone box versions

Here is an overview of the 2108 SAN Data Gateway:

- Physically looks the same as the 7190 model 200.
- HVD SCSI to SSA converter box, using industry standard SCSI and SSA protocols.
- Host operating system independent. SSA devices seen as SCSI devices on host by use of host SCSI device driver.
- Sun, HP, DEC, or NT host with the correct SCSI adapter attaches to the 2108-S20, which may then be connected to IBM 7133/7131 SSA Subsystems.
- The same functionality as a 7190-200, but with the additional features of RAID 1 (mirroring), RAID 0 (composite), instant copy drives, and hot spare drives.
- The IBM 2108-S20 Storwatch Specialist Configuration Utility (which only runs on WindowsNT) is used to configure the RAID arrays, and so on.
- The SSA SERV utility (menu driven or command line based, and may be run on all platforms) is used for all other RAS functionality.

Table 17 shows the supported servers and operating systems for the 2108 SAN Data Gateway.

Table 17. Supported servers and operating systems for the 2108-S20

Servers	Models	Adapters	Operating System
Intel - based	IBM Netfinity Servers Compaq Proliant Servers	Adaptec (AHA-2944UW) For Netfinity use part# 59H39	Windows NT 4.0 Server or Enterprise Server Editions with Service Pack 4 or later
Sun	SPARCstaion 10, 20 SPARCserver 1000, 1000E SPARCcenter 2000, 2000E UltraEnterprise 2, 30, 60, 150, 450, 5000, 5500, 6000, 6500, 10000	Sun - SBus (X 1062A or X 1065A Sun - PCI (X6541A)	Solaris 2.5.1, 2.6 or 2.7

Servers	Models	Adapters	Operating System
HP-UX	9000 Series V,T,K,D,G,E,H,I Class systems	HP-HSC (A2969A Or A4107A) HP-PB (28696A) HP-PCI (A4800A)	HP-UX 10.10, 10.20, 11.0
Compaq	Alpha Servers 300, 800, 1000, 1000A, 1200, DS20, 2000, 2100, 2100A, 4000, 4100, 8200, 8400, GSI140 Series	SCSI F/W Diff PCI model KZPSA-BB Ultra SCSI model KZPBA-CB	Tru64 Unix 4.0

Switch Settings and LEDs

Similar to the 7190-200, the 2108-S20 has two sets of switches in the rear of the unit labeled SW1 and SW2.

SW1 - SCSI ID Selection

SCSI devices must be configured so that only one device responds to any given ID. For maximum performance and easy problem determination, we recommend that the 2108-S20 is the only device on the SCSI bus.

SW1 consists of eight dip switches. When a dip switch is toggled, both ID numbers and their LUNs are selected (or deselected). None of the ID numbers (except 6) can be independently enabled.

If the dip switch is down, the 2108-S20 will be able to assign an ID/LUN associated with that switch to a device on the SSA loop.

If the dip switch is up, the 2108-S20 will not assign the ID/LUN associated with the dip switch to a device on the SSA loop.

The 2108-S20 comes with a default that recognizes IDs 0 through 6 and 8 through 15 (all switches down). There is no ID switch for ID 7, because this is normally reserved for the host.

The 2108-S20 must be power cycled for any switch changes to take effect.

SW2 - mode selection

SW2 is another eight dip switch set that determines the mode of operation for the 2108-S20. These modes can be LUN 0 only when in WinNT mode, normal mode, or various service or diagnostic modes.

LEDs

There are three LEDs at the front of the unit (power, status, and fault) and two in the back (status and fault).

For full details of the various switch settings and LED meanings, refer to the *IBM 2108 SAN Data Gateway Model S20, Installation and User's Guide*, SC26-7312.

2.4.4 VICOM SLIC Router FC-SL

The VICOM Serial Loop IntraConnect (SLIC) Router provides connectivity and functionality to allow you to use SSA drives in a fibre channel environment.

The SLIC Router has incorporated drive redundancy, hot-swappable drives and expanded drive storage along with its revolutionary open system which allows

mixed operating systems to access differing loops. This feature provides the SLIC Router architecture with additional benefits such as: storage sharing, host fail-over, and parallel servers.

Each SSA loop may contain up to 64 drives. Multi-host support is provided for up to eight heterogeneous host systems (each requiring a SLIC Router device) per SSA loop. As a stand-alone device, the SLIC Router can create mapped physical drives, mirror drives, and composite drives by utilizing the SLIC Router Mode Switches. The SLIC Router's LED function assures easy maintenance for drive identification and management.

Figure 9 shows a photo of the SLIC Router.

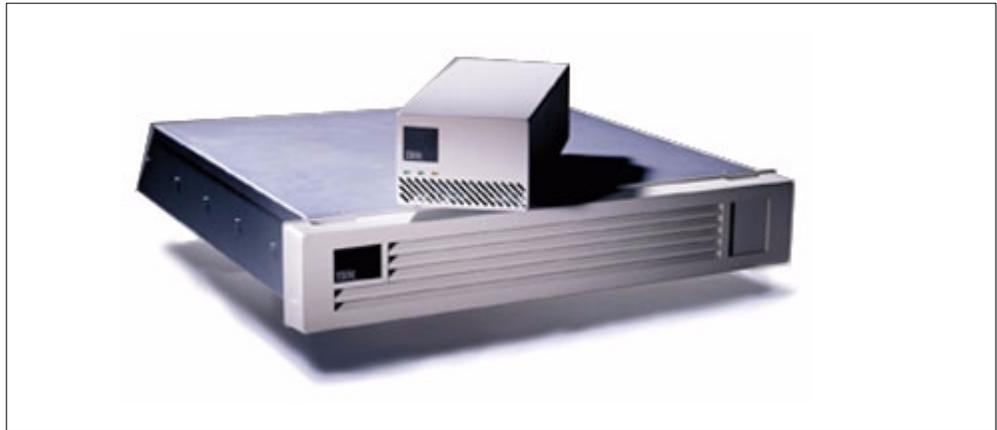


Figure 9. VICOM SLIC Router FC-SL and drawer

When the SLIC Router utilizes its standard SLIC Manager program (compatible with most operating systems), local and/or remote management of the SSA loop is simplified even further. It creates an 'Instant Copy' for any mirror drive from a composite, physical or spare drive. This program enables old and new microcode for both the SLIC Router and the drives to be downloaded. It forces any drive on the SSA loop to identify itself by blinking the drive's LED. The SLIC Manager's ability to download the mapping table enables the user to save their loop design for later use. It also offers diagnostic testing and error logging for every device on the loop. The SLIC daemon (a feature of the SLIC manager) provides customized commands through remote or local access for automatic and continuous monitoring and managing of the subsystem. Refer to the *SLIC Manager Installation and User Guide* which came with the product, or refer to the VICOM Web site for further information.

To the host, the SLIC Router is transparent and the SSA drives on the loop appear as fibre channel disks drives.

2.4.4.1 SLIC Router installation

The SSA bus (storage loop) is designed to be a loop made up of point to point connections. Each device connects to two neighboring devices. When all devices have been attached with standard SSA cables, the bus will become a loop. Ensure that only one SLIC Router is used to configure the loop. Add the other SLIC Routers (if any) to the storage loop after all the SSA drives have been configured.

Before installation begins, confirm with the distributor or reseller from whom the SLIC Router and drives were purchased that both the SLIC Router unit and the drives are loaded with the latest microcode revision. The latest microcode may be downloaded from the CD ROM provided with the unit or from the VICOM Web site:

<http://www.vicom.com>

The user does not need to install the SLIC Manager to setup the SLIC Router. However, if the SSA drives or SLIC Router were not shipped with the latest microcode, the user must install the SLIC Manager to download the latest microcode. (Refer to the *SLIC Manager Installation and User Guide* which came with the product, or refer to the VICOM Web site, for operation on microcode download.)

These steps show the sequence of events to implement the SLIC Router in your environment:

1. Check drive LEDs
2. Clear the node mapping table
3. Run subsystem diagnostic test
4. Prepare the drives
 - a. Create physical drives
 - b. Create any general spares
 - c. Format the drives
5. Create logical drives
 - a. Create any composite drives
 - b. Create any two-way mirrors
 - c. Create any three-way drives
6. Hook up system and set mode
7. Run physical disk scan
8. View drives
9. Upload configuration
10. Add any additional SLIC Routers

For more detailed information on setting up the SLIC Router, please see the *SLIC Router FC-SL Installation and User Guide*, which is provided with the unit.

2.4.5 Copper cables, fibre optic extenders and fibre cables

2.4.5.1 Copper cables

These are the basic building blocks of the SSA loop. There are two types of cable, the "classic" four wire Send/Receive pairs for 20MB loops and six wire cables for use on 40MB loops. They come in several standard lengths up to the 25m maximum allowed between SSA nodes. Caution must be exercised when tightening the extended retaining screws that fasten the cable D shell to the adapter/7133 drawer ports to prevent shearing of the screw. The screw shoulder must be fully seated and tight, but only a touch more than finger tight is sufficient. Always use the special SSA screwdriver provided with every 7133 subsystem.

Classic cable

This cable is grey in color and consists of the micro 'D' type connector at each end and two pairs of transmit/receive wires. It only supports the 20MB (80MB bandwidth) protocol. The shortest lengths are normally only used as jumper cables to interconnect disk quadrants on 7133 models 010 and 500.

160 MB enabling cables

This cable is blue in color and uses the same micro 'D' type connector at each end and two pairs of transmit/receive wires as above. In addition, it has another pair of cables that enable, with supported adapter and disk devices, the loop to run at 160MB/s between those devices. It also enables the extended command set to monitor, change and query the 7133 D40 or T40 disk drawer.

Fibre-Optic Extender 160 (feature code 8851)

The purpose of the Fibre-Optic Extender 160 product is to extend the range of a SSA 40MB/s serial link from 25m, which is the specified limit for a copper link, to a maximum distance of 3km using multi-mode optical fibres (50/125µm and 62.5/125µm) and up to 10km using single-mode optical fibres (9/125µm). Although designed for SSA 40, this product will also operate at SSA 20 speeds on supported products. On the optical connection, a Fibre-Optic Extender 160 (OE) unit can only be connected to another matching device and cannot be used with any other fibre optic extender product.

The Fibre-Optic Extender 160 does not provide any retiming or frame buffering of the SSA data. Consequently, the maximum bandwidth of 40MB/s is only maintained over fibre distances where the propagation delay of the fibre link does not exceed the time taken to fully unload the packet buffers of the SSA devices connected to the OE.

Above this fibre length, the bandwidth will reduce as the fibre length increases up to the maximum distance specified for the fibre type. This product supersedes the SSA fibre optic extender which operates at SSA 20MB/s speeds, providing a maximum link distance of 2.4km, and supports multi-mode optical fibres only.

The maximum fibre length is set by the packet acknowledge (ACK) time-out of the devices driving the OE. On the SSA 20MB/s generation of IBM products, this was set at 25 microseconds which translates to maximum fibre length of 2.4km. In the case of SSA 40MB/s IBM products, the time-out is set at 110 microseconds which will support a maximum fibre distance of 10km. This distance is set by the return path propagation time for a light within the optical fibre and is, therefore, a fixed value. However, in a multi-mode fibre optic system, other factors, such as dispersion, restrict the maximum fibre length achievable before the ACK time-out limit and, therefore, single-mode fibre is required to achieve the 10km distance.

The maximum data rate which is achievable on a fibre optic link also decreases past a certain distance, which is less than the maximum fibre distance, due to the limited frame buffering which exists within current SSA systems. Presently, two frames are buffered on the transmit and receive ends of the link, which means that at 20MB/s the bandwidth will reduce after approximately 320m, because the data in flight along the fibre exceeds this limit. In the case of a SSA 40MB/s link, this distance is reduced to approximately 160m due to the doubling of the data rate.

The Fibre-Optic Extender 160 product consists of two optical extenders that plug into SSA external micro-D sockets. This connector supplies +5V and ground for the OE, the serial SSA data connections, and two option pins, which signal SSA 40 capability and the presence of an optical extender. The two OEs are connected by twin optical fibre cables and emulate a copper SSA serial link without the need for any additional hardware or software.

2.4.5.2 Optical fibre cable definition

The Fibre-Optic Extender 160 incorporates a 1310nm wavelength laser source and is designed to operate with multi-mode fibres up to a maximum distance of 3km and with single-mode fibre up to a maximum distance of 10km.

2.4.5.3 Multimode fibres

The Fibre-Optic Extender 160 will operate with multi-mode optical fibres with a size of 50µm/125µm or 62.5µm/125µm, the full specification of these fibres can be found in the Fibre-Optic Extender 160 Application Guide shipped with the feature or viewed at:

http://xenon.ssd.hursley.ibm.com/extweb/docs/Other_Docs/fibreoptics.html

Different fibre sizes must *not* be mixed along a single path otherwise an unacceptably high attenuation will result at the interface between the fibres. The connector used between the cable and the OE is of the ST 1 standard.

2.4.5.4 Singlemode fibres

The Fibre-Optic Extender 160 is designed for singlemode optical fibres with a size of 9µm/125µm, the full specification of these fibres can be found in the Fibre-Optic Extender 160 Application Guide shipped with the feature or viewed at:

http://xenon.ssd.hursley.ibm.com/extweb/docs/Other_Docs/fibreoptics.html

2.4.5.5 LED states

The *power* LED is coloured green and is illuminated if the power supply to the OE has been above the minimum value for correct operation of the OE for a fixed time interval. The threshold voltage and time interval is set by the internal circuitry of the OE.

The *receive* LED is coloured green and is illuminated if the average optical input power to the receive optical input of the OE is above a threshold value. This threshold is set by the internal circuitry of the OE, and the LED should always be illuminated during normal operation of the OE.

2.4.5.6 Operation

The micro 'D' type connector should be inserted into the external socket of the associated SSA port and the optical cable connected to the ST optical connector on the OE.

On connection to a powered-up system, the *power* LED should illuminate indicating the OE has power. The *receive* LED will only illuminate when the OE is receiving light from the OE at the remote end of the optical link. The receive LED will still be on even if there is no SSA data running across the link since under these conditions the remote OE will be providing a continuous light output.

The order of connection of the extender and optical cable is unimportant. No damage can be caused to either the OE or the connecting system. The OE is also hot-pluggable, so it is not necessary to power down the system to connect the OE.

Chapter 3. Configuring SSA loops

This chapter looks at some considerations and recommendations for configuring SSA loops.

3.1 Some rules for cabling SSA loops

Unlike some other types of disks, SSA disks are configured in "loops". A loop starts at a cable port on an adapter card, goes through a number of disks, and (optionally) other adapter cards, and ends at the adjacent port on the original adapter card.

An adapter card has four ports: A1, A2, B1, and B2. If a loop starts from an A port, it ends at the other A port. If it starts at a B port, it ends at the other B port. The following rules should be observed to create a trouble free loop implementation:

1. Each SSA loop *must* be connected to a *valid pair* of connectors on an SSA adapter card — either connectors A1 and A2, or connectors B1 and B2.

Ports are always used in pairs, either As or Bs. There is no material difference between As and Bs, or between 1s and 2s. The As and Bs may be in the same loop if they are on different adapter cards, but *not* if they are on the same adapter card.
2. Only *one* of the two pairs of connectors on an adapter card can be connected in a particular loop.
3. A maximum of 48 disk drives can be connected in a particular loop. (Note: This allows a maximum of 96 disk drives per adapter card in two loops). In an SSA disk subsystem, all disk slots must be filled with either a disk drive or a "dummy" drive. The function of a dummy drives is to provide a mechanical connection for continuity in the loop. It cannot send or receive signals.
4. A maximum of three consecutive dummy disk drive modules can be connected in a particular SSA loop. Depending on the type of adapter card used, from one to eight adapters can be used in the same loop. The limitations are stated in these rules.
5. If any of your adapter cards are feature code 6214 (type 4-D), a maximum of two may be connected in one loop. For example, one adapter may be a 6214 and one a 6216, but the limit is still two per loop).
6. If your adapter cards are feature code 6216, (type 4-G), a maximum of eight may be connected in one loop.
7. If your adapter cards are feature code 6217 (type 4-I), 6218 (type 4-J), PC part number 32H3811, or 7133 feature code 4011, these adapter cards support RAID and non-RAID configurations, but only one adapter may exist in an SSA loop.
8. If your adapter cards are feature codes 6215 (type 4-N) or 6219 (type 4-M), these adapter cards support both RAID and non-RAID configurations, however, the level of microcode present on the card determines the types of configurations supported by that card.

For adapter microcode levels below 5xxx observe these rules:

- In a RAID configuration, only one 6215 or 6219 adapter card may exist in an SSA loop.
- In a non-RAID configuration, one or two adapter cards may exist in an SSA loop. An exception exists with FC 6222, Fast Write cache. When using this feature, only one adapter may exist in a loop in either RAID or non-RAID mode.

For adapter microcode levels at or above 5xxx, observe these rules:

- In a RAID configuration, one or two 6215 or 6219 adapter cards may exist in an SSA loop.
- In a non-RAID configuration, a maximum of eight adapter cards may exist in an SSA loop. An exception exists with FC 6222, Fast Write cache. When using this feature, only one adapter may exist in a loop in either RAID or non-RAID mode.

When using microcode at or above level 5xxx, you must ensure that you have the proper device driver file sets. These file sets are different from those used with microcode levels below 5xxx.

PC Server Adapter 96H9835, Cluster Adapter, supports up to two adapters on the same loop on certain PC Servers in RAID-1 or non-RAID mode. If two adapters are in the same loop, they cannot be in the same host .

9. When a FC 6215 or FC 6219 SSA Adapter card is connected to two SSA loops, that is, one loop from the A ports and one loop from the B ports, and each loop contains two or more adapters, the additional adapters for each loop must be the same adapters.

If you are using multiple adapter cards (6214, 6216, 6215, or 6219) in a loop, these adapters may be in separate hosts or in the same host.

10. A maximum of 2 adapters in a particular loop may be installed in the same host system. An exception is the PC server cluster adapter; See Rule 8, above.

All adapters and SSA subsystems in a particular loop must be connected to other adapters and/or subsystems. This is done with SSA copper cables which come in lengths of 1, 2.5, 5, 10, and 25 meters or using the Fibre-Optical Extender (OE) and optical cables.

11. The maximum copper cable length between two SSA nodes (disk or adapter) is 25 meters (82 feet).
12. The maximum optical cable length between two SSA nodes (disk or adapter) is 2.4km with FC5500 or 10.5km with FC. The OEs must be a matched pair, different feature codes at either end of the optic cable is not supported.

Following are some cabling examples incorporating single and multiple adapters and hosts. All examples are utilizing bypass cards. Some will demonstrate the preferred method of cabling and others show what *not* to do.

3.2 Bypass cards

The 7133 Models 20 and 600, D40 and T40 contain four bypass cards. The cards provide the connections between the external cable and the disk drives. Figure 10 shows a logical view of the Model 20 and 600 disk arrangement. Each pair of

cable ports (1 and 16, 4 and 5, 8 and 9, and 12 and 13) is mounted on a bypass card.

A SSA bypass card monitors both of its external cable connectors. If it detects the presence of a powered-on SSA adapter or disk on either of the ports making up the pair, it connects the internal SSA links (that is, the connections to the disk drives) to the external cable ports. This is the *inline mode*.

Conversely, if the SSA bypass card fails to detect the presence of a powered-on SSA adapter or disk at the other end of each cable, an electronic switch on the card connects its internal SSA links together and disconnects them from the external cable ports. This is the *bypass mode*. This process takes about 250ms.

For this reason, you should take precautions when powering up or powering down multiple hosts at the same time to prevent breaking the loop in two places at once.

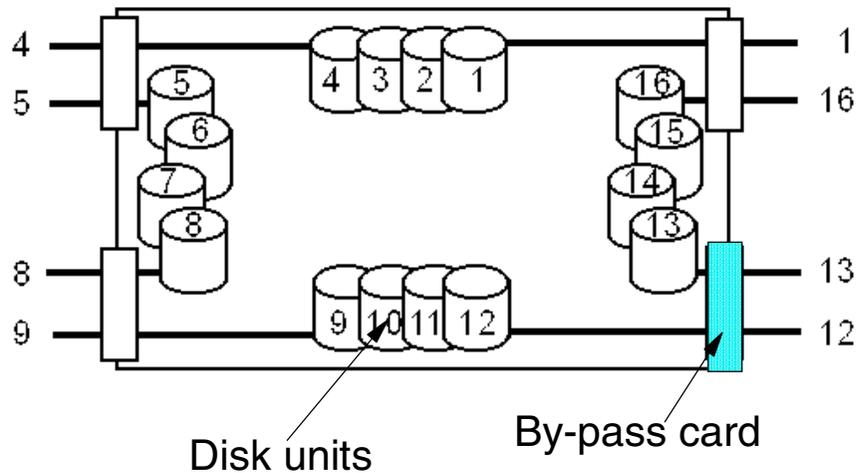


Figure 10. 7133 schematic layout

For example, in Figure 10, if an adapter card is connected to 7133 cable ports 1 and 16, and nothing is connected to any other ports, then ports 1 and 16 would be in inline mode, and all other port pairs would be closed in bypass mode forming one loop of all 16 disks.

In earlier models of the 7133 (Models 10 and 500), jumper cables were used to connect one group of four disk positions to another group of four disk positions. These are not needed on the models with bypass cards. The bypass link performs this function.

In some particular configurations, you may not want the bypass links to close. In that case, the bypass card could be locked in forced inline mode. This means that it will not close even if no active SSA devices exist at either cable port.

The examples in Figure 10 illustrate proper and improper use of the bypass and inline functions.

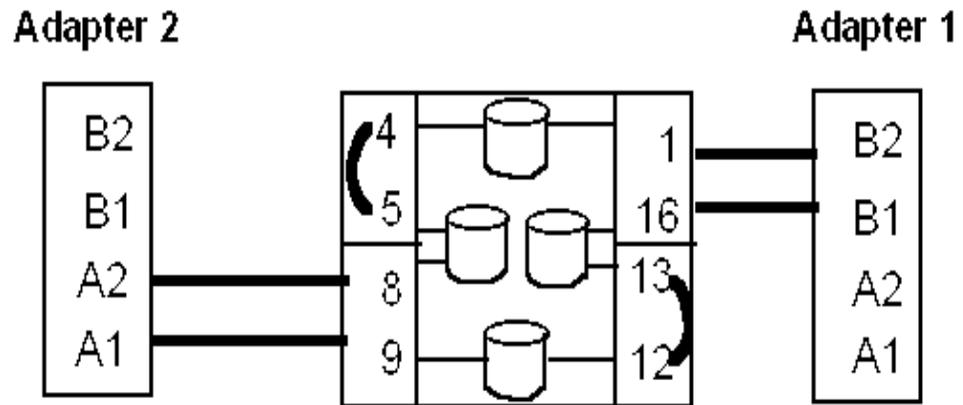


Figure 11. Simple loop with two adapters

The example in Figure 11 above is one 7133, Model 20/600 and two adapters in either the same host or different hosts. If either of the adapters should fail, or are powered off, the bypass link where it is connected will close and re-form a loop giving the remaining adapter two paths to all data. Since nothing is connected to ports 4 and 5, or 12 and 13, the bypass links for these ports are closed.

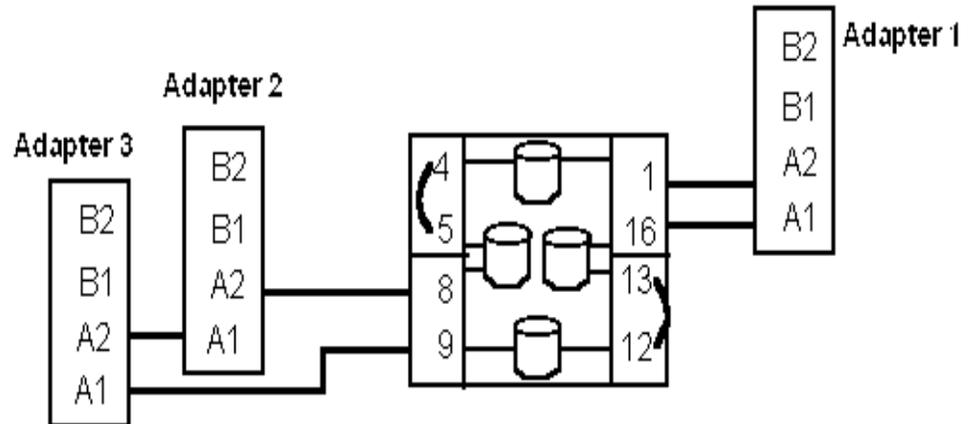


Figure 12. Simple loop with three adapters

In the example in Figure 12, a third adapter has been added to the configuration. If adapter 1 should fail, or is powered off, then bypass link 1 and 16 would close, and adapters 2 and 3 would still have two paths to all disks. If adapter 2 or 3 failed, or was powered off, the 8 and 9 link would not close, because a powered-on device is still attached to one side of the 8 and 9 ports. Adapter 1 and either of the remaining adapters still has access to all disks but only with one path.

If both adapters 2 and 3 failed or were powered off, link 8 and 9 would close and adapter 1 would still have two paths to all disks.

The two examples above illustrate the benefit of the bypass link, that is, to re-form a loop in the event that a device fails or is powered off. However, you should be careful not to create configurations in which the closing of a bypass link creates

an invalid or unwanted configuration. The following examples show how this could occur.

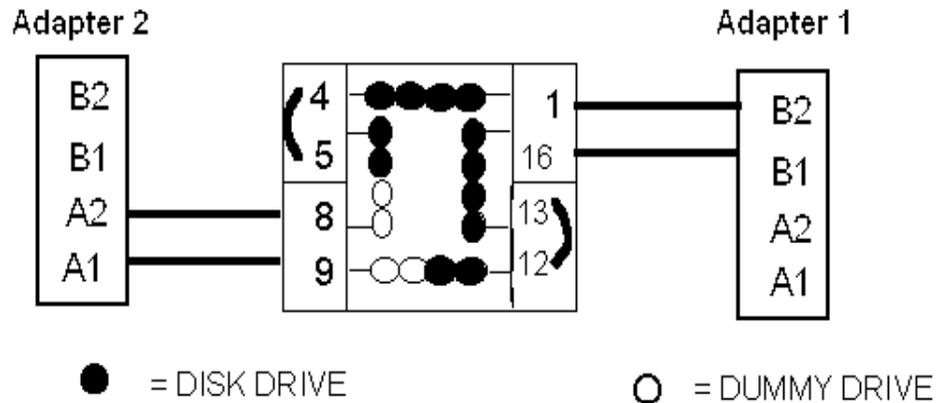


Figure 13. Simple loop, poor availability configuration

The example above in Figure 13 shows a valid configuration with one 7133 and two adapter cards. If adapter 1 fails or is powered off, bypass link 1 and 16 will close and adapter 2 still has two valid paths to all disks. If adapter 2 fails or is powered off, bypass link 8 and 9 closes and puts four dummy disk modules in a row. This violates rule number 4 that allows no more than three consecutive dummy disk modules in a loop. This would result in signal loss across the dummy disks and the loop failing at that point. Adapter 1 will only have a single path to all disks.

Suggestion

To avoid this possibility, when you have fewer than 16 disks in a 7133, they should be spread in the 7133 with a disk in each position that is next to a cable port. If the disks in the above example were spread this way, the 7133 would look like Figure 14.

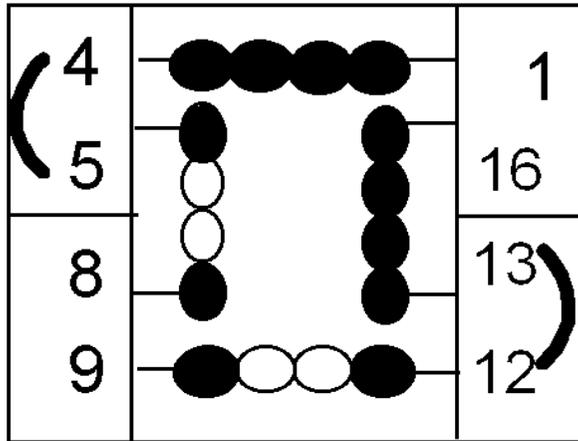


Figure 14. Advised disk unit layout in partially populated enclosures

The next example in Figure 15 shows a valid configuration with one 7133 and two adapter cards. It contains two loops of eight disks each and both adapters have access to all disks.

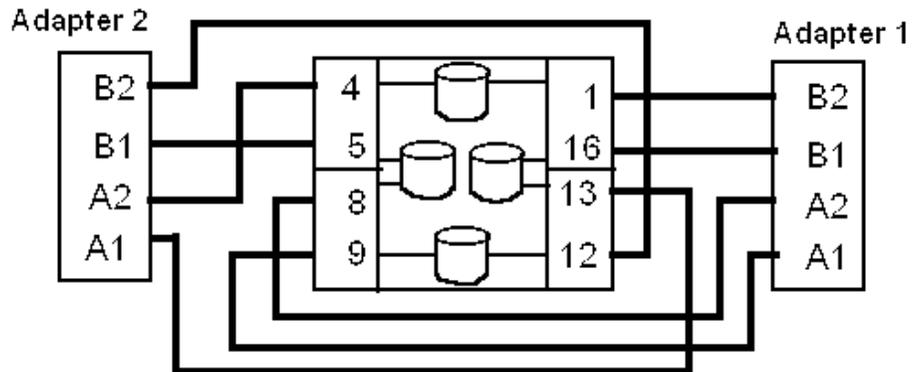


Figure 15. Poor availability configuration

If adapter 2 fails or is powered off, bypass links 4 and 5, and 12 and 13 will close, as the configuration shows in Figure 16.

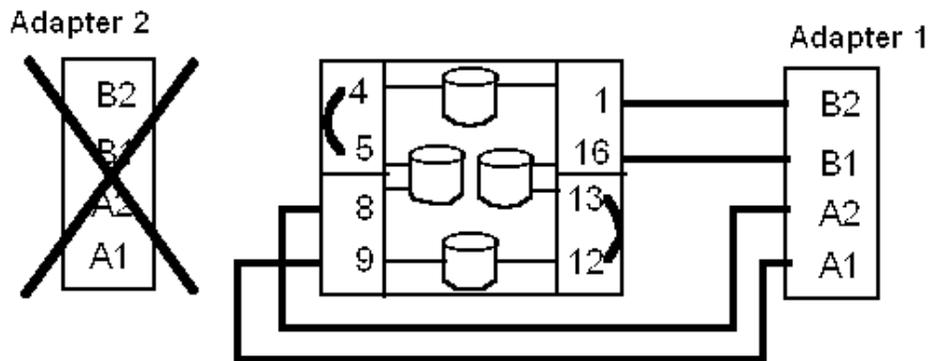


Figure 16. Result of loss of adapter

The result is two, eight disk loops connected to both the A ports and B ports on the same adapter card. This violates the configuration rule that allows only one set of ports (As or Bs) on an adapter card to be in a particular loop.

This invalid configuration can be avoided in two ways. First, ports 4 and 5, and 12 and 13 could be put in forced inline mode, and they would not close if adapter 2 failed or was powered off.

This gives four strings from Adapter 1 instead of loops, but the configuration is still valid.

A better solution is to cable as shown in Figure 17. This provides the same two-loop configuration, but does not result in an invalid configuration.

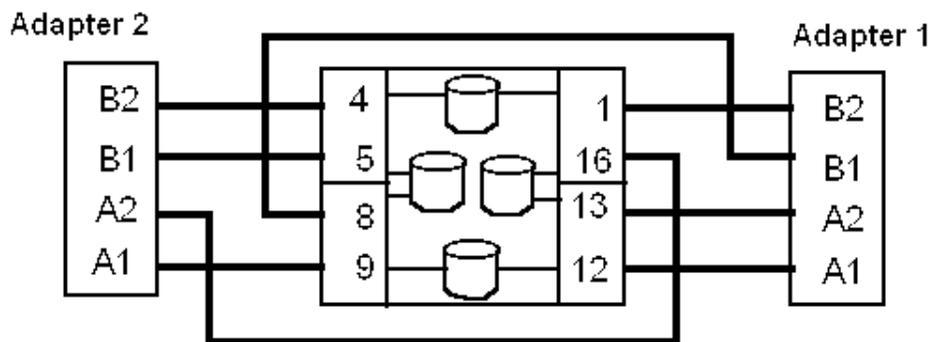


Figure 17. Correct cabling for this configuration

Suggestion

Try to avoid connecting a bypass pair of connectors to both ports (A and B) on the same adapter. If your configuration requires that you do so, then put the ports in *forced inline mode*. This will prevent the invalid configuration as shown in Figure 17.

In general, you should take advantage of the bypass function whenever you can. If your configuration is such that a particular set of bypass ports do not offer a benefit, or you are not certain what the effect might be in case of a failure or when a machine is powered off, it is best to put that set of ports in forced inline mode.

3.3 Loops with multiple hosts

One of the big advantages of SSA comes from being able to create multiple host loop configurations in which hosts can have multiple adapters. In this case, multiple means from two up to eight hosts, depending on adapter type (see rule number 9 in 3.1, “Some rules for cabling SSA loops” on page 33). The two main benefits of this kind of configuration are:

- All hosts can have a view of all disks. This gives great flexibility in data sharing and high-availability backup.
- Dual adapters in each host gives redundancy to eliminate the adapter as a single point of failure for that host.

However, in using this type of configuration, you can create potential problem areas even though your loops obey all of the rules. The following three cabling examples show a two host configuration with two adapters per host. (The same rules would apply for loops with more than two hosts.) The examples are valid configurations and show problems that can occur if one host is powered off. In each example, the logical configuration in the top figure represents the cabling configuration shown in the bottom figure.

3.3.1 Cabling example 1

In the example shown in Figure 18, the hosts are alternated in the loop. If one adapter fails, the second adapter for that host still has access to all disks. However, if one host is powered off, the remaining host could temporarily lose access to half of the disks in the loop.

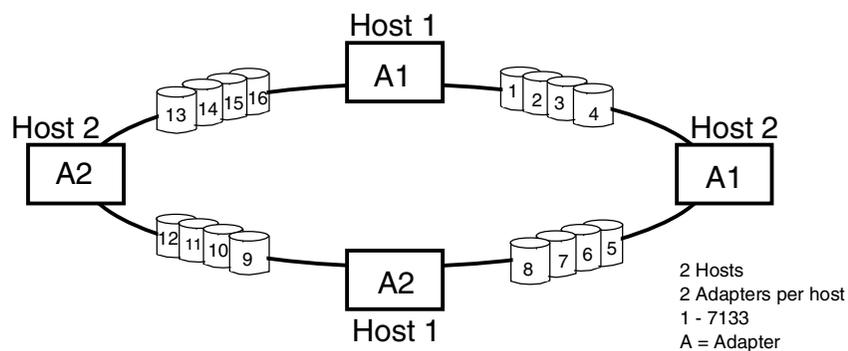


Figure 18. Logical cabling diagram: example 1 (not recommended)

For instance, if host 1 is powered off, two breaks are created in the loop: one for adapter H1A1 and one for adapter H1A2. The breaks are closed by the bypass links at 8 and 9, and 1 and 16 (this takes approximately 250ms). Host 2 must then reconfigure new path lengths for itself, because two nodes (adapters) have been removed. While this is taking place, an error condition could occur if an I/O operation was in process. If host 2, adapter 1 was writing to a disk on the other side of the loop (disks 9 - 16) when host 1 was powered off, both paths from the

adapter to the disk would be broken until the links closed and errors could occur. Depending on the status of the operation in process, host 2 might not be able to recover. Figure 19 shows the physical diagram to match the logical layout above.

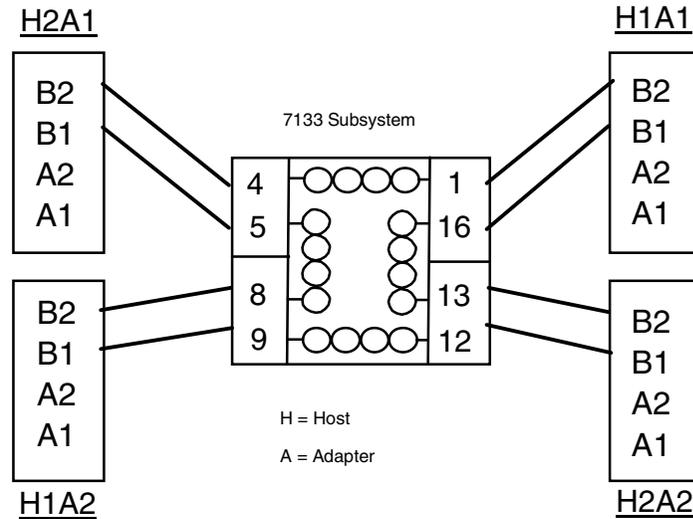


Figure 19. Physical cabling diagram: example 1

3.3.2 Cabling example 2

Note that the logical and cabling configurations in this example look the same as in the previous example, but the hosts are in a different sequence. Figure 20 shows the new logical diagram of these changes.

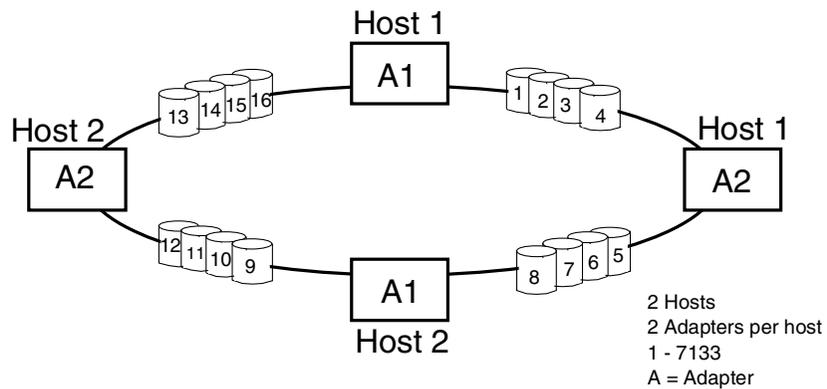


Figure 20. Logical cabling diagram: example 2 (not recommended)

In this example, the hosts are not alternated in the loop. That is, disks are placed between adapters on the same host. If one adapter fails, the second adapter for that host still has access to all disks. However, if host one (for example) is powered off, disks 1 and 4 become temporarily unavailable to the other host. They are unavailable until bypass links 1 and 16, and 4 and 5 close and new path lengths are reconfigured. (If these bypass links happened to be set in forced inline mode, the disks would stay unavailable). This configuration is not

recommended. Figure 21 shows the physical diagram to match the logical layout above.

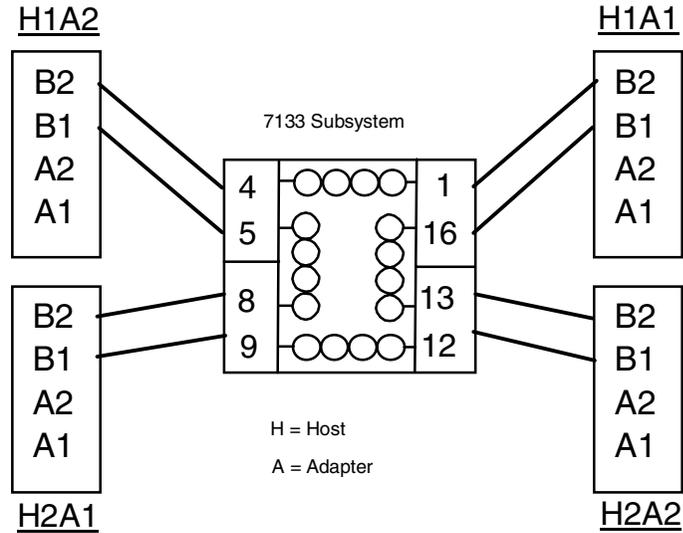


Figure 21. Physical cabling diagram: example 2

3.3.3 Cabling example 3 (recommended for highest availability)

This example shows a configuration that removes the potential problems shown in the previous examples. Figure 22 shows the logical diagram, while Figure 23 shows the physical layout for this configuration.

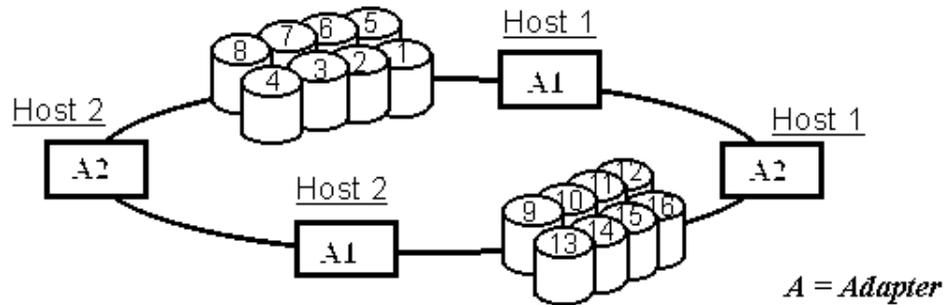


Figure 22. Logical cabling diagram: example 3 (recommended)

In this configuration, the two adapters in the same host are cabled together. If either adapter fails, the remaining adapter in that host still has access to all disks. If a host is powered off, no disks are isolated from the other host.

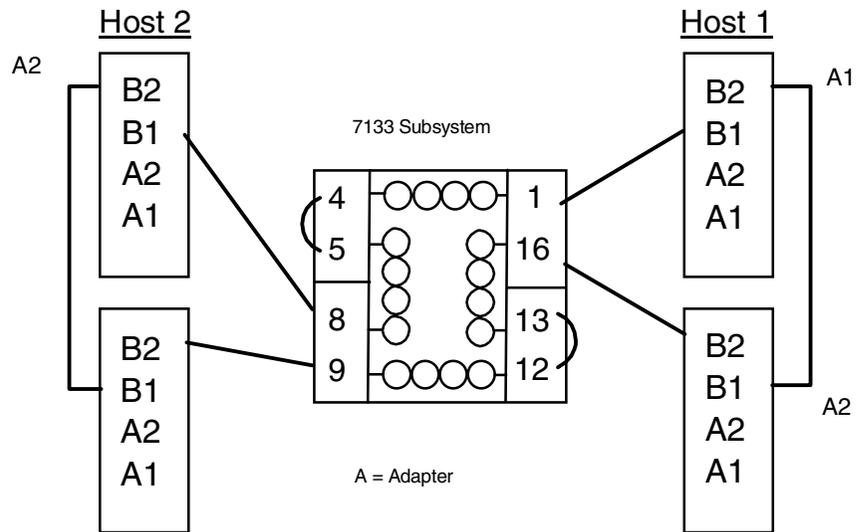


Figure 23. Physical cabling diagram: example 3

Figure 23 shows the recommended configuration for a multi-host, multi-adapter configuration.

3.4 Data domains and disk placement

In a multi-host configuration, the placement of data can have an effect on system performance. To keep your path lengths as short as possible, the general guideline is to place the data used by a host as close as you can to that host. This data is known as the host's domain.

The following examples show one loop with four hosts with one adapter each. They illustrate some things to do and some things not to do. The numbers on the disks show that particular disk's host.

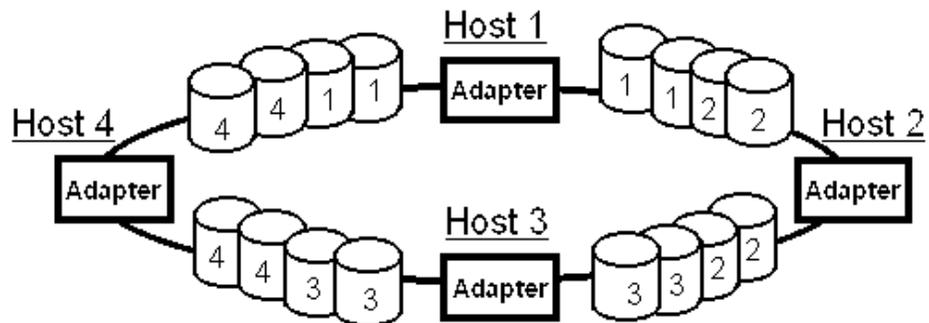


Figure 24. Optimal disk allocation within multi host loop

This configuration (Figure 24) shows four disks allocated to each host. They are placed in the positions closest to the adapter for that host. Obviously, disks are not always evenly allocated as shown here, but this illustrates the objective of placing data close to its using host.

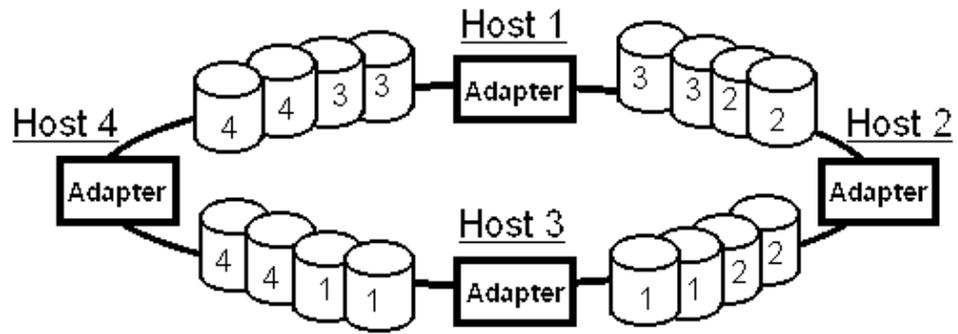


Figure 25. Poor disk allocation in multi host loop

This configuration (Figure 25) has the disks for host 1 and host 3 switched. It is still a valid configuration, and will work fine, but this change has caused the path length for these two hosts to be more than four times as long as it was in the previous example.

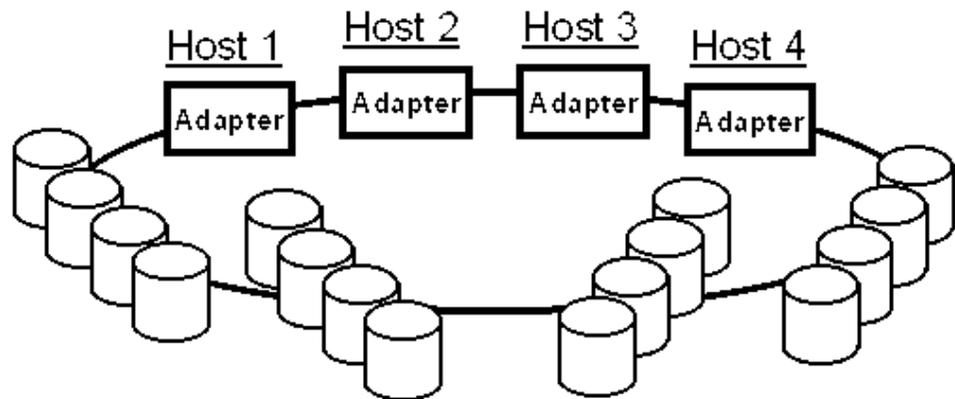


Figure 26. Disaster waiting to happen! NOT recommended

Figure 26 shows a rather extreme example of a very poorly configured, but valid, SSA loop. No matter how disks are allocated to hosts, the path lengths will not be optimum. This example is meant to illustrate that an efficient configuration and proper data placement are equally important for best performance.

This configuration also shows a potential for data isolation. If hosts 1 and 4 are powered off at the same time, hosts 2 and 3 cannot use any of the disks.

3.5 Mirroring configurations

The function of mirroring data, that is, maintaining more than one active and updated copy of data on the system, is a way of providing high availability of data through redundancy. In general, the more hardware redundancy that you have, the greater your availability insurance will be. As the levels of availability insurance increase, so does the cost. Therefore, various levels of redundancy are used in mirrored environments.

The following examples discuss and show how various levels of availability are achieved.

3.5.1 Mirroring example 1

Mirroring example 1 (Figure 27) shows one host with one adapter connected to one 7133. Eight disks in the 7133 could be mirrored to the other eight disks to provide a mirrored copy of each disk. This gives redundancy of disks only, that is, if a disk fails, a back-up copy is available. No redundancy exists for the 7133, the adapter card, or the host system. This is the lowest level of redundancy in a mirrored configuration.

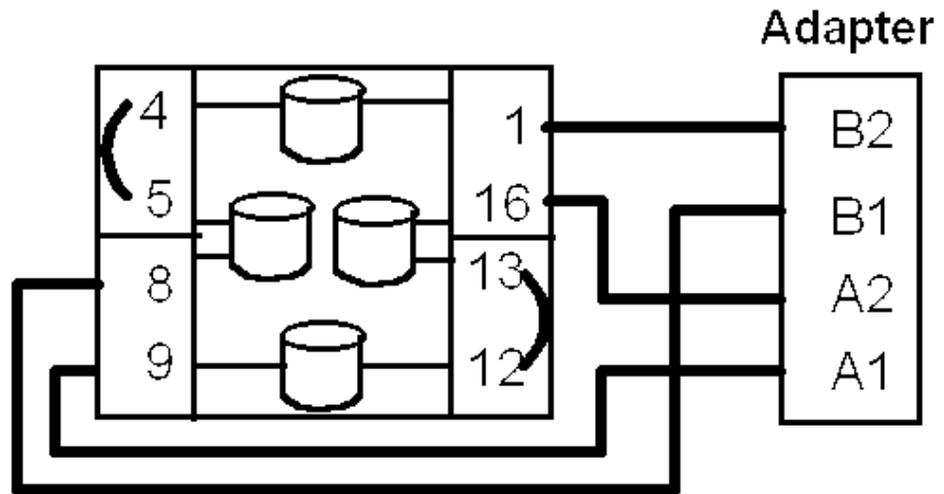


Figure 27. Example 1, simple mirroring, disk redundancy only

3.5.2 Mirroring example 2

Mirroring example 2 (Figure 28) adds a second 7133. In this configuration, you could mirror one 7133 onto the other. This would give redundancy of disks, 7133 circuitry, and cables, but not of the adapter card or the host system.

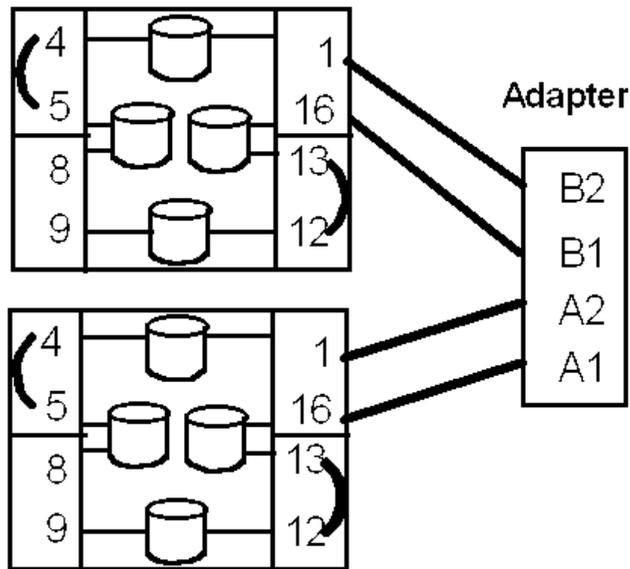


Figure 28. Example 2, simple mirroring, disk and 7133 enclosure redundancy

This configuration provides better protection than in example 1, but it still is not the best. So, we added a second adapter (Figure 29).

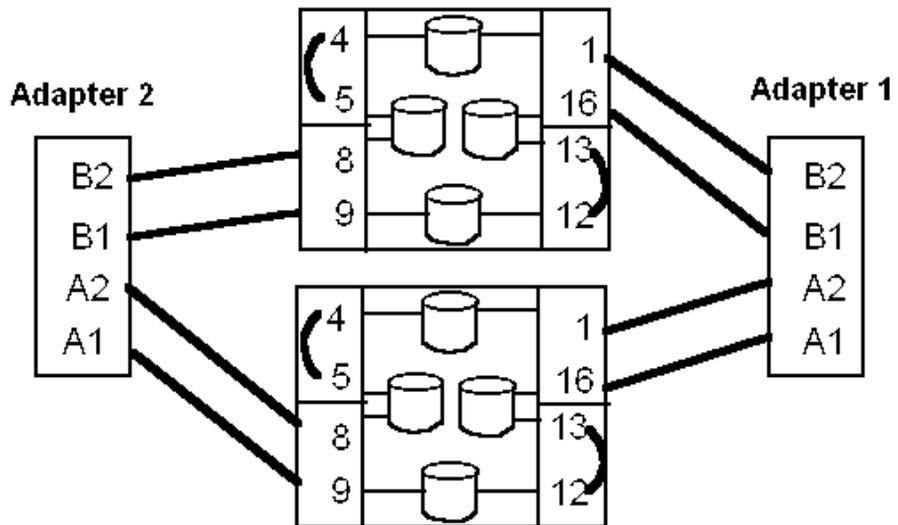


Figure 29. Example 3, simple mirroring, best redundancy for single host

3.5.3 Mirroring example 3

In mirroring example 3, the two adapters could be in the same host or in different hosts. If they are in the same host, we now have disk, adapter card, and 7133 redundancy, and the only exposure is in the host system. If the adapters are in separate systems, we would have redundancy for the host, but not for the adapter card.

The next examples (Figure 30 and Figure 31) show two options which provide total redundancy and the best level of availability. Each shows two host systems with multiple adapter cards attached to multiple 7133s.

Configuration 1

Figure 30 shows two identical configurations. Each host has an adapter card attached to each 7133. One 7133 would be mirrored onto the other.

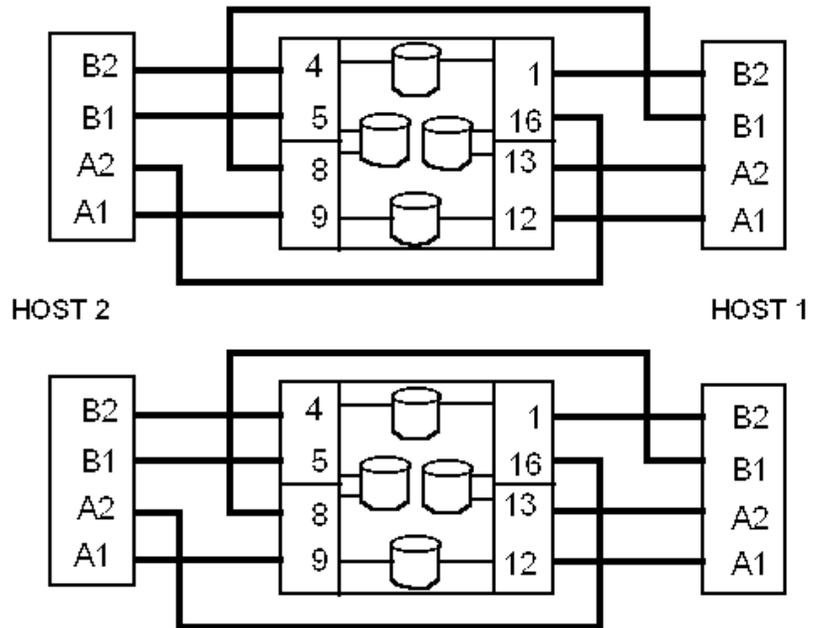


Figure 30. Mirrored configurations — total redundancy for all components

Configuration 2

Figure 31 shows the adapters within a host cabled together. This gives all adapters in a host access to all disks.

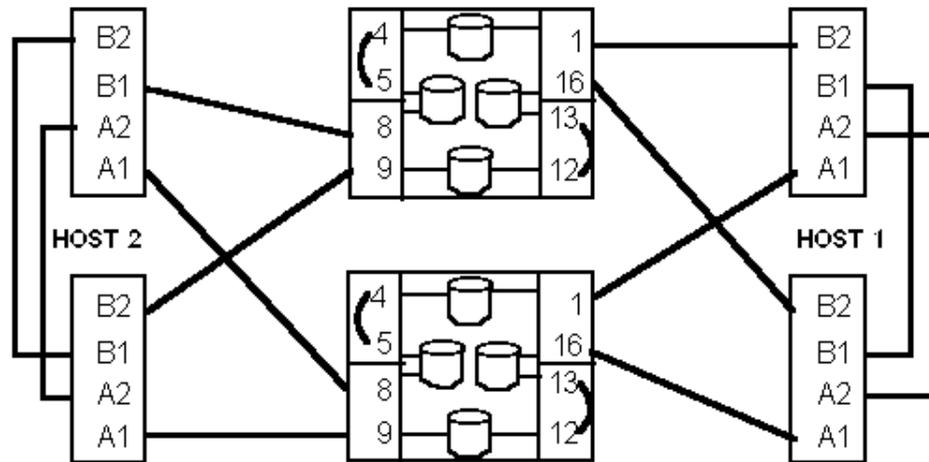


Figure 31. Mirrored configurations — total redundancy and maximum access

The preceding examples show two of the best approaches for a mirrored configuration. Each example shows a good and valid configuration, and each has its own advantages.

Configuration 1 (Figure 30) shows one adapter in each host connected to one 7133. This configuration has a shorter maximum path length than the configuration in Number 2, and would potentially give somewhat better performance. However, if one adapter failed, its copy of the data would become stale and must be updated after the failed adapter is replaced.

Configuration 2 (Figure 31) has both adapters connected to both 7133s. The path lengths are longer than those in configuration one, but if an adapter should fail, the second adapter still has access to all data, and neither copy would become stale. Only, the failed adapter would have to be replaced.

3.6 SSA adapter cards

All SSA subsystems — 7133, 7131, and 3527 — attach to their host systems through an adapter card. So far, nine differently identified adapter cards are available with a variety of functions and attachment hosts. We anticipate the descriptions that follow make it a bit easier to remember the function of each one.

With two exceptions, all adapter cards are "four-port" cards. That is, they have four cable connector ports: two A ports (A1 and A2) and two B ports (B1 and B2). Each card can support two loops, an A loop and a B loop. (The exceptions are the 7190 SCSI Host to SSA Loop attachment, and the FC4003 SBus Interface Controller. Each of these adapters have two cable ports and support one loop.) The configuration rules that you read earlier in this chapter apply to all adapter cards.

The adapter cards are summarized in Table 18, and in the diagrams on the following pages. These are meant only to highlight the functions of the different cards. For more detailed information, you should consult the sales manual or other product specific documentation.

Note: In the charts you will find the term *JBOD*. This may be new to you. It refers to non-RAID disks. It means "Just a Bunch Of Disks".

Table 18. Adapter cards summarized

Feature Code	Characteristics for SSA loop
6214	Micro Channel, non-RAID adapter. You may have up to two in a loop.
6216	Micro Channel, non-RAID adapter. You may have up to eight of these in a loop. A 6214 and a 6216 may be in the same loop but only one of each.
6217	Micro Channel RAID-5 and non-RAID adapter card for certain RISC 6000 systems. It is single initiator, that is, one adapter card in a loop. RAID-5 and non-RAID disks can be intermixed in a loop.
6219	This Micro Channel adapter is a follow-on to the 6217. It supports one or two initiators (adapters) in a loop in RAID mode, and a maximum of eight in non-RAID mode. A fast-write cache option is available. (Note: See cabling rule No. 8 in section 3.1, "Some rules for cabling SSA loops" on page 33 for microcode and file set requirements). Also supports FC 6225.
6218	This adapter has the same function as the 6217, except it is PCI attached rather than Micro Channel. It is supported by certain RISC 6000 systems.
6217	This adapter has the same functions as the 6219, except it is PCI attached rather than Micro Channel.
6215	This adapter has the same function as the 6219, except it is PCI attached rather than Micro Channel. It is supported by certain RISC 6000 systems.
p/n 32H3811	This is the adapter card for certain IBM PC servers. It supports RAID 0, 1, and 5, and non-RAID. It can be the only initiator in the loop.
4011	This is the same card as 32H3811, only it is used on certain Compaq and HP servers. It can be the only initiator in the loop.
p/n 96H9835	This is a follow-on to part number 32H3811, and attaches to selected IBM PC servers. It is called a cluster adapter and it supports RAID 1, and non-RAID with up to two initiators (adapters) per loop.
4012	This is the same card as 96H9835, only it is used in certain Compaq and HP servers.
4003	This two-port card is an SSA to Sun SBus interface controller. It supports up to 48 SSA disks on a single loop in certain Sun systems.
6225	This is PCI and can have two initiators (in RAID mode) or up to eight non-RAIDed.
6230	The same as above but has increased cache.
7190	The two-port 7190 is not an internal host adapter card, but is an external converter which allows SCSI host attachment of SSA subsystems on certain Sun, HP, and DEC systems. It is a single loop and up to four 7190s can be in the loop.
2108	The two-port 2108 is not an internal host adapter card, but is an external converter which allows HVD SCSI host attachment of SSA subsystems on certain Sun, HP, and DEC systems.

3.7 Fibre optic extender

In configurations using SSA copper cables, the maximum distance allowed between nodes is 25 meters. In situations requiring distances of more than 25 meters between nodes, a fibre optic extender can be used. It supports distances of up to 10500 meters (depending on FC) between nodes.

The fibre optical cables connect to the bayonet connectors and the SSA D shell connector connects to an SSA node. Two optical fibre cables and two extenders are required to replace a single copper cable. Four extenders are required to complete an SSA loop (Figure 32).

A fibre optic extender and its connectors are shown in the diagram below. Note that one bayonet connector is longer than the other. Fibres must be connected from the long connector on one end of the fibre to the short connector on the other end of the fibre.

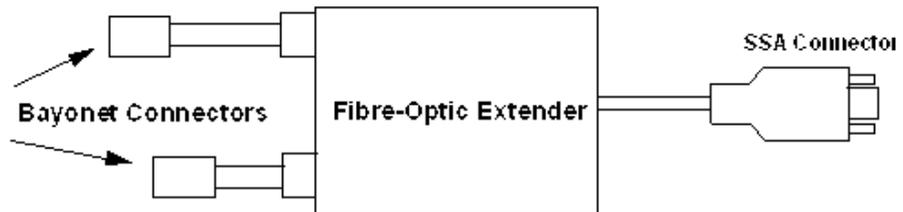
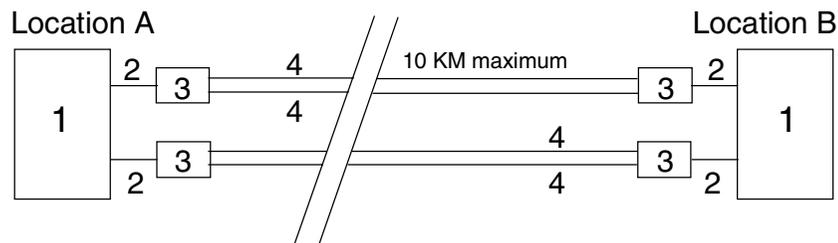


Figure 32. Fibre-Optic Extender (OE) layout

Figure 33 shows a simple SSA configuration with two locations using four fibre optic extenders. Remember, it takes four extenders to complete a loop. (When ordering Extenders, each feature code provides a pair of extenders. You need two pair for each loop).



Legend:

- 1 - SSA disk subsystem or Host Adapter Card
- 2 - SSA Cable connection (comes with Fibre-Optic Extender)
- 3 - Fibre-Optic Extender
- 4 - Optical Fibres (two per path, four per loop)

Figure 33. Fibre-Optic extender usage

Obtaining and installing the fibre cable is a customer responsibility. For the SSA fibre optic extender, a single multi-mode fibre type, either 50/125 µm or 62.5/125

µm, must be used on a single path. Generally, the 50/125 µm fibre has a higher specification than the 62.5/125 µm. If a choice is available, use a 50/125 µm fibre type. This will drive the loop to a maximum of 2.4 kilometers.

Fibre-Optic Extender 160 can also use the above specification cabling with the 2.4km limit, but a maximum distance single-mode fibre type 9/125µm will allow up to 10km between SSA nodes. We advise using this cable specification to give maximum flexibility for your SSA subsystems. See 2.4.5, "Copper cables, fibre optic extenders and fibre cables" on page 28 for full details on cable types.

The ST connectors used to terminate the fibre must be of a suitable specification to meet a 0.05db connection loss over the life of the system.

3.8 Using RAID

We have already talked about mirroring as a way of providing availability insurance. Another way to do that, at a lower cost, is by using RAID-5. In a RAID-5 "array", data records are written across some number of disks (for SSA subsystems, that number is 2 to 15) plus one "parity" disk. Actually, the parity information is spread among all disks in the array, but it is often easier to think of the array as having one parity disk. The purpose of the parity disk is to hold an "and/or" code which allows the system to recreate the data from any single array member, in the case of a disk failure. Data can be recreated from only one array member at a time. If two array members are unavailable at the same time, the data on the entire array is unavailable to the system.

Only one parity disk is used for each array regardless of the size of the array, so the added cost for this availability insurance varies with the size of the array. However, in selecting array sizes there are some things other than cost that should be considered as well:

- Smaller arrays usually give better availability, because they have fewer things to fail.
- Large arrays have a greater chance of having two disks unavailable at the same time.
- In the event of a disk failure, smaller arrays have shorter rebuild times, and larger arrays have longer rebuild times.
- Smaller arrays are more expensive than larger arrays, because they require more parity disks. (An array of 2 + P requires one parity disk for two data members. An array of 15 + P requires one parity disk for fifteen data members).

A good document to help in RAID planning is *Planning SSA RAID Subsystems*, GA33-3271.

RAID-5 configurations can be structured with or without "hot spares". They are not required, but it is a good idea to use them. If hot spares have been identified to the system, and a disk fails, the system can select a hot spare and immediately start rebuilding the data from the failed disk. Without a hot spare, the rebuild process cannot start until the failed disk is replaced.

All RAID adapters are 4-port cards and support two SSA loops. All RAID adapters support disks in RAID and non-RAID mode. When using FC 6217 or FC 6218 RAID adapter cards, arrays may span the loops on that particular adapter card.

When using FC 6215 or FC 6219 RAID adapter cards, all array members must be contained in the same loop.

The following diagrams show a conceptual representation of how arrays might be arranged, and how non-RAID disks might be intermixed.

Figure 34 shows a 6217 or 6218 RAID adapter card with two loops of 16 disks each.

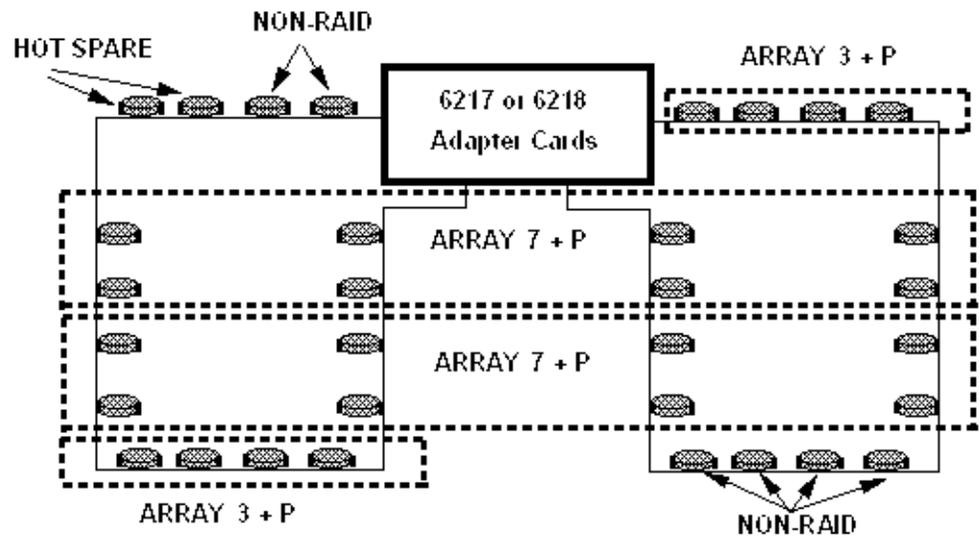


Figure 34. Layout using RAID and non-RAID arrays for 6217 and 6218

Each loop has an array of 3 + P. Two arrays of 7 + P span both loops. Two hot spares can serve all arrays on either loop. The six non-RAID disks could be placed in any position on either loop.

Figure 35 shows one 6215 or 6219 RAID adapter card with two loops of 16 disks each.

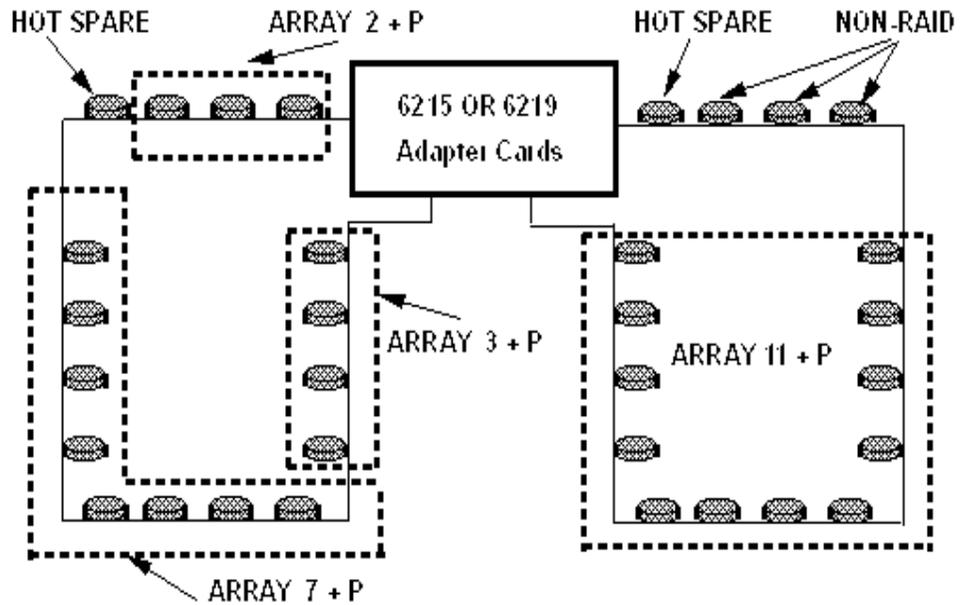


Figure 35. Layout using RAID and non-RAID arrays for 6215 and 6219 (single adapter)

The arrays shown are 2 + P, 3 + P, 7 + P, and 11 + P. Each loop has one hot spare and one loop has three non-RAID disks. Arrays on these adapter cards must be contained on one loop, that is, they cannot span loops as shown in the previous example. A hot spare serves only the arrays on its loop. The non-RAID disks may be placed in any location.

3.9 RISC 6000 systems with internal SSA disks

The RISC systems 7025 F series and the 7026 H series support internal SSA disks in the systems storage bays (with the appropriate SSA backplane feature installed) and 7133 and 7131 SSA subsystems externally. Depending on the system model, a maximum of 12 or 18 internal drives are supported. Both internal and external disks are attached through the 6218 adapter card. In addition to the four external cable ports, the 6218 also has two internal cable ports which are used to attach internal SSA disks. If internal SSA disks are attached then the related external port cannot be used on the 7025 F30.

Configuration examples for each system are shown on the following pages. These configurations should be used for guidance only. You should refer to specific RISC 6000 documentation for additional information. The announcement letter describing these options is 197-068.

7025 Model F30

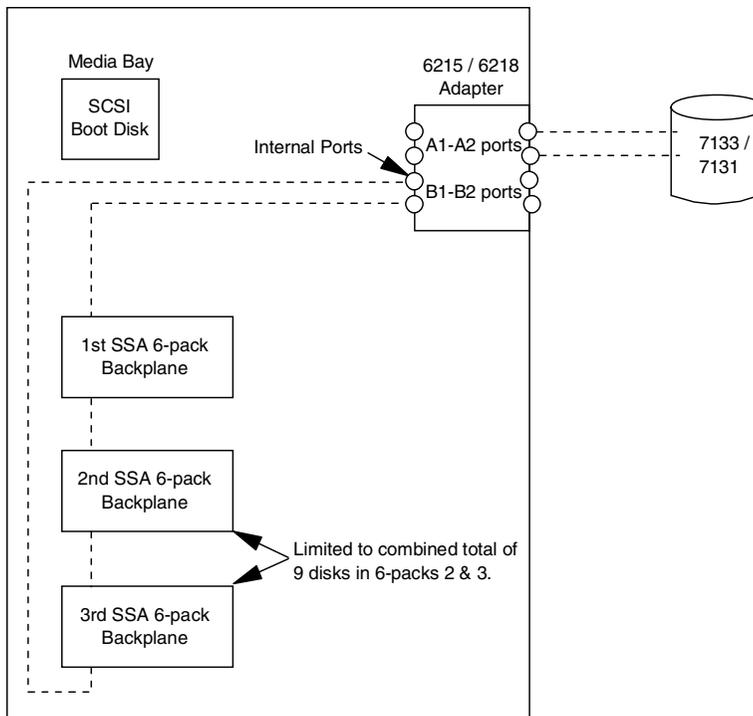


Figure 36. Internal SSA layout RS6000 7025 model F30

In Figure 36 the 7025 model F30, the SSA backplanes are attached through the internal ports on the adapter. If you have additional SSA backplanes installed you are limited to a total of nine disks spread between them. Remember to ensure that there are no more than three disk dummies in a row. The external B1 and B2 ports may *not* be used.

External SSA devices are attached through the A1 and A2 ports only.

7025 Model F40

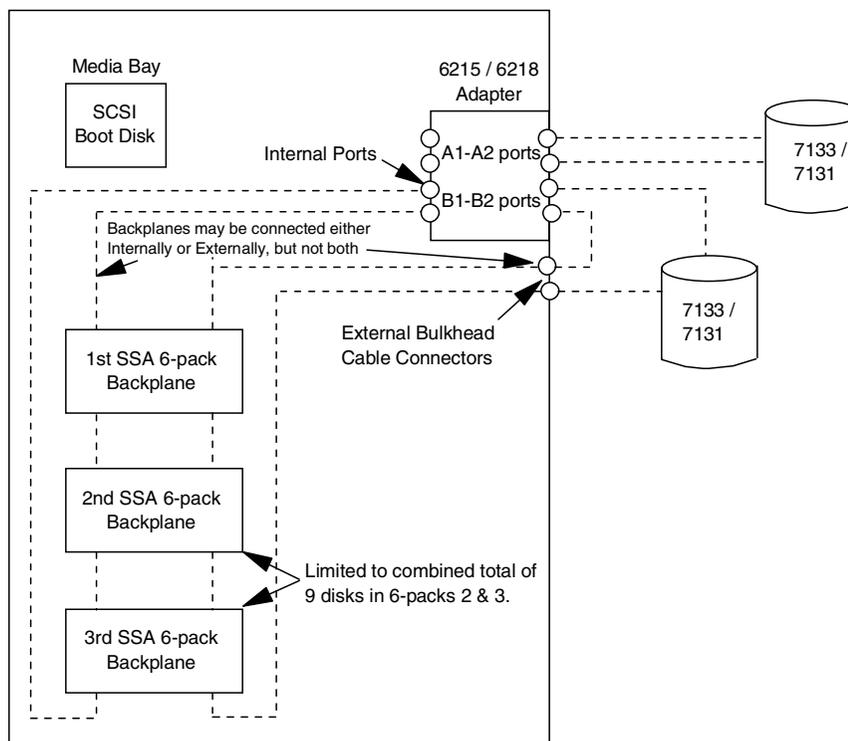


Figure 37. Internal SSA layout RS6000 7025 model F40

In the 7025 model F40 (Figure 37), SSA backplanes may be attached *either* through the internal SSA adapter ports *or* through the external bulkhead connectors with cables to the B1 and B2 ports. If you cable through the bulkhead, then external SSA devices may be included in that loop up to the permitted maximum, taking into account the internal devices. If you have additional SSA backplanes installed, you are limited to a total of nine disks spread between them. Remember to ensure that there are no more than three disk dummies in a row.

The external A1 and A2 ports can be used as normal.

7025 Model F50 and 7026 Model H50

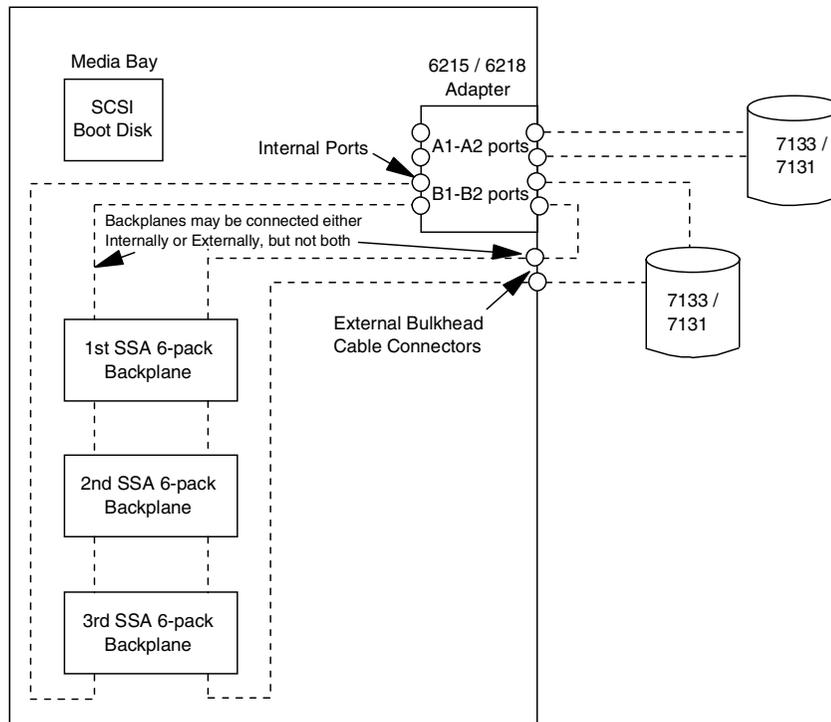


Figure 38. Internal SSA layout RS6000 702X models F50 and H50

In the 7025 model F50 and 7026 model H50 (Figure 38), SSA backplanes may be attached *either* through the internal SSA adapter ports *or* through the external bulkhead connectors with cables to the B1 and B2 ports. If you cable through the bulkhead, then external SSA devices may be included in that loop up to the permitted maximum taking into account the internal devices. If you have additional SSA backplanes installed, these systems can be fully populated with up to 18 devices. The external A1 and A2 ports can be used as normal.

Remember to ensure that there are no more than three disk dummies in a row.

7026 Model H10

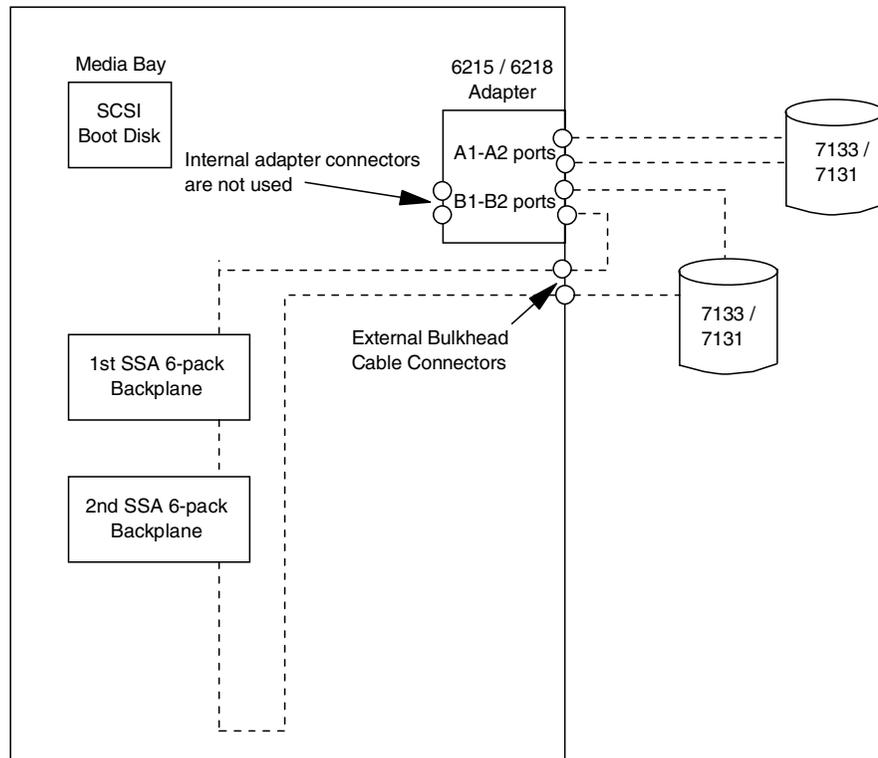


Figure 39. Internal SSA layout RS6000 7026 model H10

In the 7026 model H10 (Figure 39), the internal SSA backplanes are attached through the external bulkhead connectors *only*. It only supports *one* additional SSA backplane, so this way it can be fully populated. External SSA devices can also be added to the B1 and B2 loop up to the permitted maximum, taking into account the internal devices. The A1 and A2 ports can be used as normal.

Remember not to exceed more than three disk dummies in a row.

3.10 How to develop an SSA configuration

Several simple steps should be followed when you are developing an SSA configuration.

The cabling diagram is the last step, and several things should be done prior to this. If you follow these steps, the result is a good configuration that will accomplish your objective.

1. Define what you want to do:

- How many hosts in the configuration?
- How many adapter cards in each host?
- Which adapter cards are they? (6214, 6216, etc.)
- HACMP?
- RAID?
- Mirroring?
- How many 7133/7131 SSA subsystems?
- How many disks in each subsystem?
- If multiple hosts, should all hosts have a view of all disks?

- Any requirement for Fibre-Optic Extender?
2. Create a logical configuration based on your answers to step 1.
In a logical configuration, you can see how all of the pieces fit together. This is where you will define your domains and the sequence of adapters.
 3. Create a cabling configuration based on your logical configuration.
Remember to observe all of the cabling rules which you read earlier. Make use of bypass links whenever possible and provide availability insurance (redundancy) where necessary.

Two simple examples follow: a logical configuration and a cable configuration. Both examples use two 7133s and four adapters.

Logical configuration example 1

Figure 40 and Figure 41 assume:

- Two 7133s.
- Two hosts with two adapters in each host.
- Mirroring from the first 7133 to the second 7133.
- Adapter redundancy — if one adapter fails, the remaining adapter still has access to all disks.
- Both hosts can access all disks.

We have multiple adapters in each host and want adapter redundancy in the configuration, so we need to connect the two adapters together. For a review of this information, see 3.3, “Loops with multiple hosts” on page 40. Figure 40 shows a logical configuration. Figure 41 shows a cabling configuration.

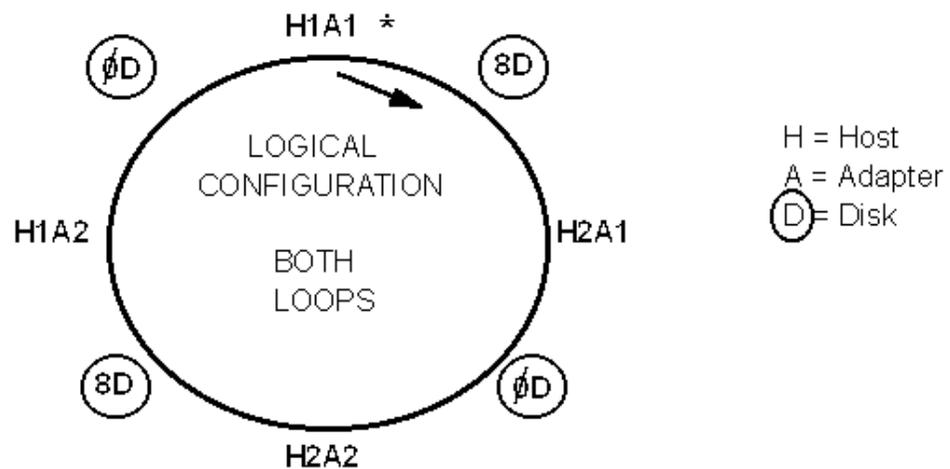


Figure 40. Suggested conceptual diagram, example 1

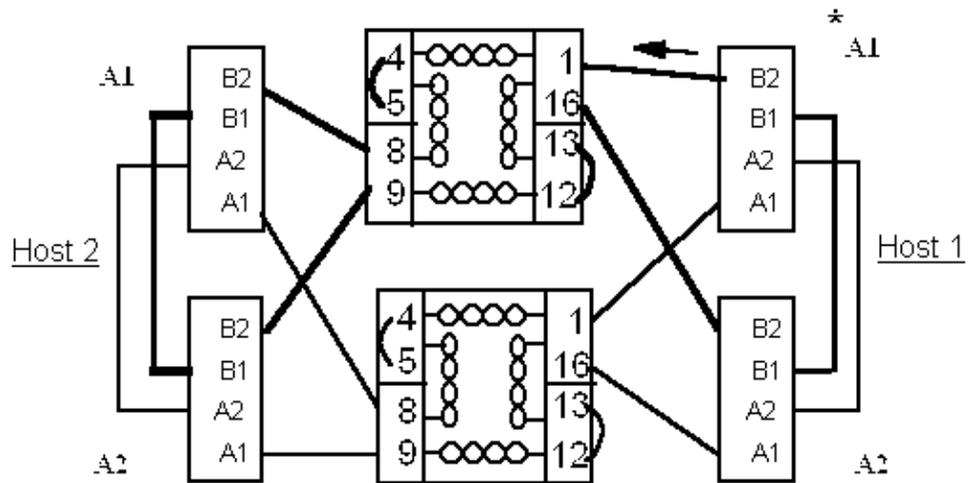


Figure 41. Actual cabling schematic, example 1

Logical configuration, example 2

Figure 42 and Figure 43 on page 60 assumes:

- Two 7133s, four hosts with one adapter in each host
- Mirroring from the first 7133 to the second 7133
- No adapter redundancy
- All hosts have a view of all disks

There are four separate hosts, but adapters should not be connected to each other. Then, disks can be distributed evenly among the hosts. Four bypass links can be used in each loop. Figure 42 and Figure 43 show the conceptual and physical diagrams to match these recommendations.

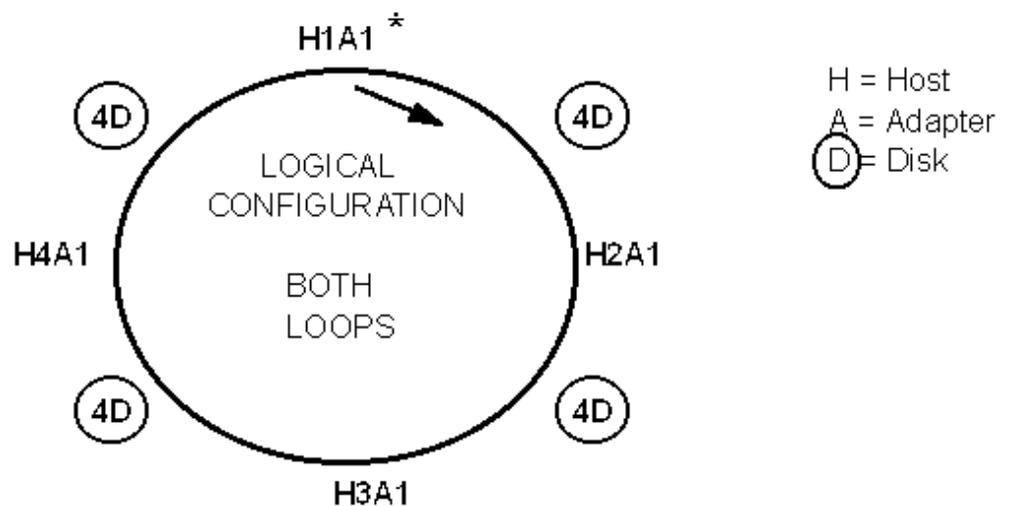


Figure 42. Suggested conceptual diagram, example 2

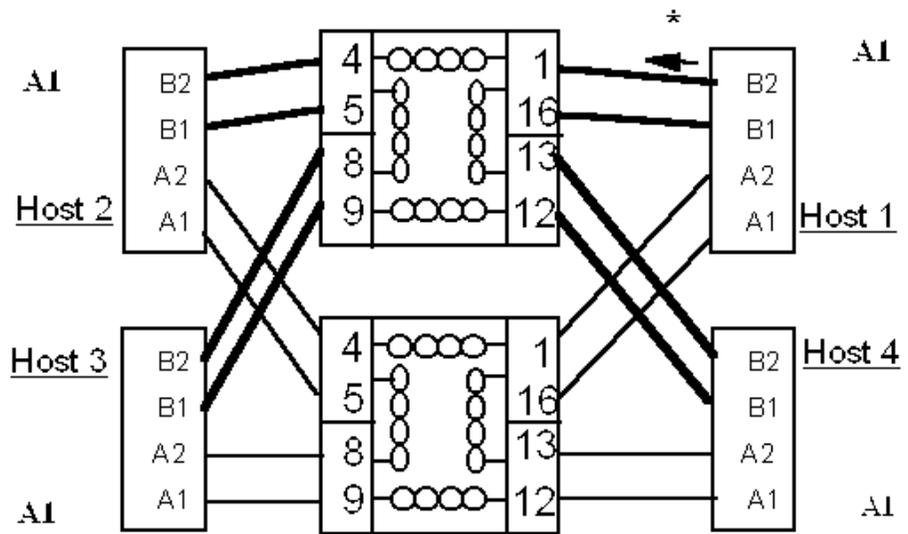


Figure 43. Actual cabling schematic, example 2

Chapter 4. Tools for managing SSA

This chapter describes the various tools and processes for managing and maintaining the SSA adapters and disk subsystems. These range from command line prompts, through diagnostics, to purpose written applications. They enable you to monitor, control, and update loops and subsystems, as well as perform disk recovery and replacement.

4.1 SSA device drivers and the management process

In order to effectively manage SSA disks, it is helpful to have an understanding of how the device driver works and how it helps the management process.

SSA disk drives are represented in AIX as SSA logical disks (hdisk0, hdisk1,...,hdiskN) and SSA physical disks (pdisk0, pdisk1,...,pdiskN). SSA RAID arrays are represented as SSA logical disks (hdisk0, hdisk1,...,hdiskN). SSA logical disks represent the logical properties of the disk drive or array, and can have volume groups and file systems mounted on them. SSA physical disks represent the physical properties of the disk drive. By default one pdisk is always configured for each physical disk drive. One hdisk is configured for each disk drive that is connected to the using system, or for each array. On initial configuration of SSA, pdisk numbers are assigned in ascending order starting with the lowest serial numbered disk (written on disk during manufacture). The pdisks are mapped to hdisks in ascending order, starting with the lowest available number. SCSI disks would have been configured first, and this is why it is very rare for pdisk and hdisk numbers to be the same.

By default, all disk drives are configured as system (AIX) disk drives. The array management software is used to change the disks from hdisks to array candidate disks or hot spares.

SSA logical disks:

- Are configured as hdisk0, hdisk1,...,hdiskN
- Support a character special file (/dev/rhdisk0, /dev/rhdisk1,...,/dev/rhdiskN)
- Support a block special file (/dev/hdisk0, /dev/hdisk1,...,/dev/hdiskN)
- Support the I/O Control (IOCTL) subroutine call for non-service and diagnostic functions only
- Accept the read and write subroutine calls to the special files
- Can be members of volume groups and have file systems mounted on them

SSA physical disks:

- Are configured as pdisk0, pdisk1,...,pdiskN
- Have errors logged against them in the system error log
- Support a character special file (/dev/pdisk0, /dev/pdisk1,...,/dev/p.diskN)
- Support the IOCTLI subroutine for servicing and diagnostic functions
- Do not accept read or write subroutine calls for the character special file

Twin adapters in a loop, single host

Some SSA subsystems allow a disk drive to be controlled by up to two adapters in a particular host system. The disk drive has, therefore, two paths to the host system, and the SSA subsystem can continue to function if one adapter fails. If an adapter fails or the disk drive becomes inaccessible from the original adapter, the SSA disk device driver switches to the alternative adapter without returning an error to any working application. Once a disk drive has been successfully opened, takeover by the alternative adapter does not occur simply because a drive becomes reserved or fenced out. However, during the opening of an SSA logical disk, the device driver attempts to access the disk drive through the alternative adapter, if the path through the original adapter experiences reservation conflict or fenced-out status.

Takeover does not occur because of a medium error on the disk drive. Takeover occurs only after extensive error recovery activity within the adapter and several retries by the device driver. Intermittent errors that last for approximately one second usually do not cause adapter takeover. Once takeover has successfully occurred and the device driver has accessed the disk drive through the alternative adapter, the original adapter becomes the standby adapter. Takeover can, therefore, occur repeatedly from one adapter to another, so long as one takeover event is completed before the next one starts. A takeover event is considered to have completed when the device driver successfully accesses the disk drive through the alternative adapter. Once takeover has occurred, the device driver continues to use the alternative adapter to access the disk drive until either the system is rebooted, or takeover occurs back to the original adapter.

Each time the SSA disks are configured, the SSA disk device driver is told which path or paths are available to each disk drive, and which adapter is to be used as the primary path. By default, primary paths to disk drives are shared equally among the adapters to balance the load. This static load balancing is performed once, when the devices are configured for the first time. You can use the `chdev` command to modify the primary path.

Because of the dynamic nature of the relationship between SSA adapters and disk drives, SSA pdisks and hdisks are not children of an adapter but of an SSA router. This router is called *ssar*. It does not represent any actual hardware, but exists only to be the parent device for the SSA logical disks and SSA physical disks.

Note: When the SSA disk device driver switches from using one adapter to using the other adapter to communicate with a disk, it issues a command that breaks any SSA-SCSI reserve condition that might exist on that disk. The reservation break is performed only if this host had successfully reserved the disk drive through the original adapter. This check is to prevent adapter takeover from breaking reservations held by other using systems.

Disk fencing

If multiple host systems are connected to the SSA disks, SSA-SCSI reserve should not be used as the only method for controlling access to the SSA disks. SSA disks support a fence operation that is used by HACMP to control access to a particular disk within a group of host machines. A disk may either be *fenced in*, in which case only nominated hosts can access that disk, or the disk may be

fenced out, in which case nominated hosts are specifically excluded from accessing it.

Enhanced loop adapters

PCI SSA Multi-Initiator/RAID adapters and Micro Channel Enhanced SSA Multi-Initiator/RAID adapters are capable of reserving to a node number rather than reserving to an adapter. We highly recommend that you make use of this capability by setting the SSA router node number attribute if multiple adapters are to be configured.

Configuring additional drives

Usually, all the disk drives connected to the system are configured automatically by the system boot process and the user need not take any action to configure them. Because SSA devices can be connected to the SSA network while the system is running without taking the system off-line, it may be necessary to configure SSA disks after the boot process has completed. In this case, configure the devices by running the configuration manager with the `cfgmgr` command. When adding multiple disks, ensure that there is only a single break in the loop by removing one disk dummy at a time and replacing it with the new disk before pulling the next disk dummy.

Important note

If you are running a HACMP cluster then you should *never* run the `cfgmgr` command while the cluster is active. Instead use this command:

```
#/usr/lib/methods/cfgssar -l ssar >/dev/null  
#lsdev -Ccssar -S0 -Fname | xargs -n1 mkdev -l
```

Note that the character after the `-S` is a zero, not a capital 'O'.

While it is possible to configure the disks with an active HACMP cluster using the above command, we strongly advise you to add extra drives to take the cluster down if at all possible.

4.2 Microcode maintenance

Periodically, IBM release updates to the disk and adapter microcode and to the SSA device drivers. Generally speaking, these updates will enhance various aspects of your SSA configuration and may provide added functionality. Although updates are not compulsory, we strongly advise that all the code is kept to the latest levels. Code revisions can be found at the SSA Web site:

<http://www.hursley.ibm.com/~ssa/index.html>

Follow the links to your product type, then the adapter type, and then to the disks you have installed. The package comes with a comprehensive readme file, which explains how to upload the code in various configurations (for instance how to load to nodes on a 9076 SP2 system through the control work station). It is possible to have automatic notification of any updates to alert you of changes in code.

In certain cases, for example, updating your entire system with a new operating system and later types of SSA adapters, the adapters and device drivers will be automatically updated, but the existing disk you are re-using is not updated. In

certain cases, the disks must be above a certain level to function properly with the adapter. Use the SSA Web site to check if this needs to be done. Easy to read tables will show you the matrix.

In AIX the Vital Product Data (VPD) contains information about the device including the microcode levels. During the configuration of the complete system, all the VPD files in the system are updated using information resident in the device's EEPROMs before any microcode is downloaded from the using system to the SSA subsystem. This means that if the microcode has been updated the change will not be reflected within VPD until the configuration manager (cfgmgr) is run again, typically, at a subsequent reboot.

4.3 SSA commands and services

This topic supplies you with the commands, menus and management tools used within several different platforms.

4.3.1 SMIT commands in AIX

SSA commands in AIX can be accessed either via SMIT (Systems Management Interface Tool) or from the # prompt on the command line.

The fast path SMIT devices can be used to access the SMIT menu system. Use the last three options on the SMIT devices menu (Figure 44) to control:

- SSA Adapters
- SSA Disks
- SSA RAID Arrays



Figure 44. SMIT Devices Menu

Selecting **SSA Adapters** takes you to the SSA Adapters menu (Figure 45), which provides an additional six options.

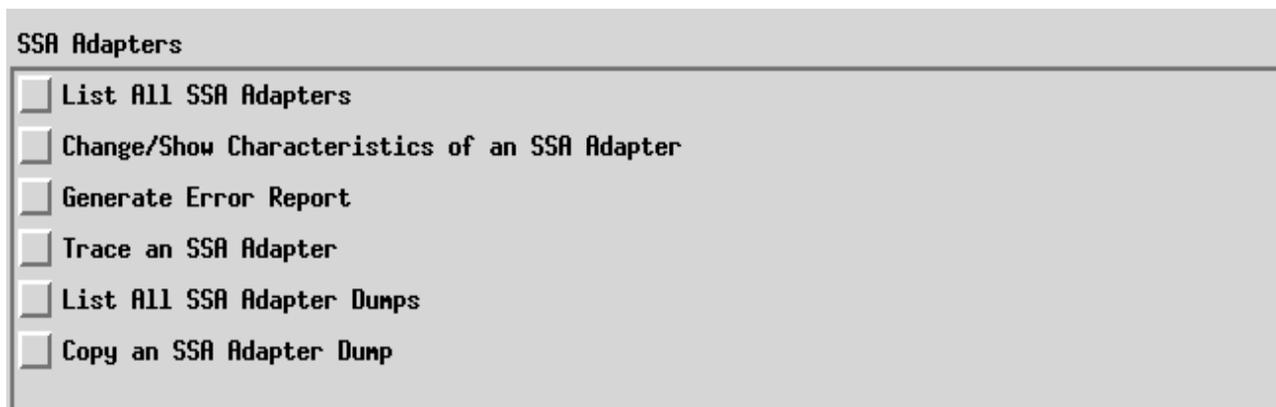


Figure 45. SSA Adapter Menu

Selecting **SSA Disks** on the SMIT Devices menu takes you to the SSA Disks menu (Figure 46), which provides two options.



Figure 46. SSA Disks Menu

Selecting **SSA Logical Disks** on the SSA Disks menu takes you to the SSA Logical Disks menu (Figure 47). This menu gives you 14 further options.

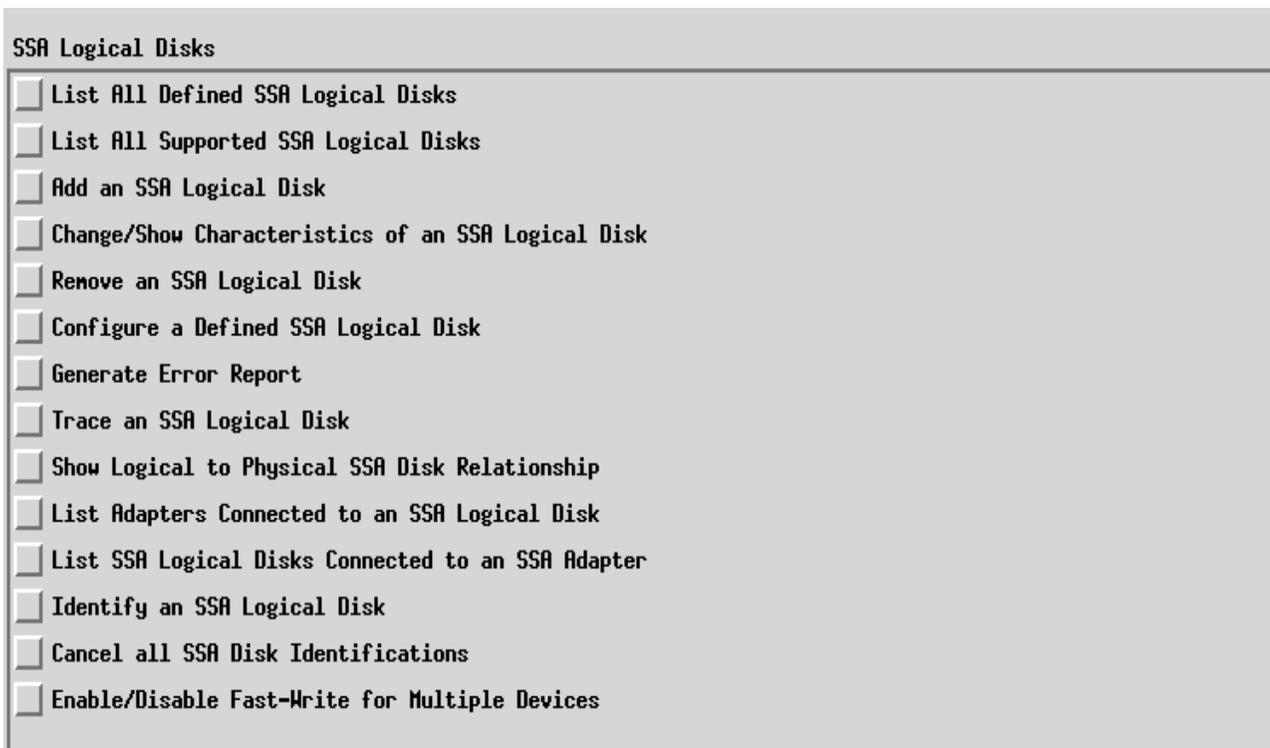


Figure 47. Logical Disk Menu

Selecting **SSA Physical Disks** on the SSA Disks menu takes you to the SSA Physical Disks menu (Figure 48), which gives you 14 further options.

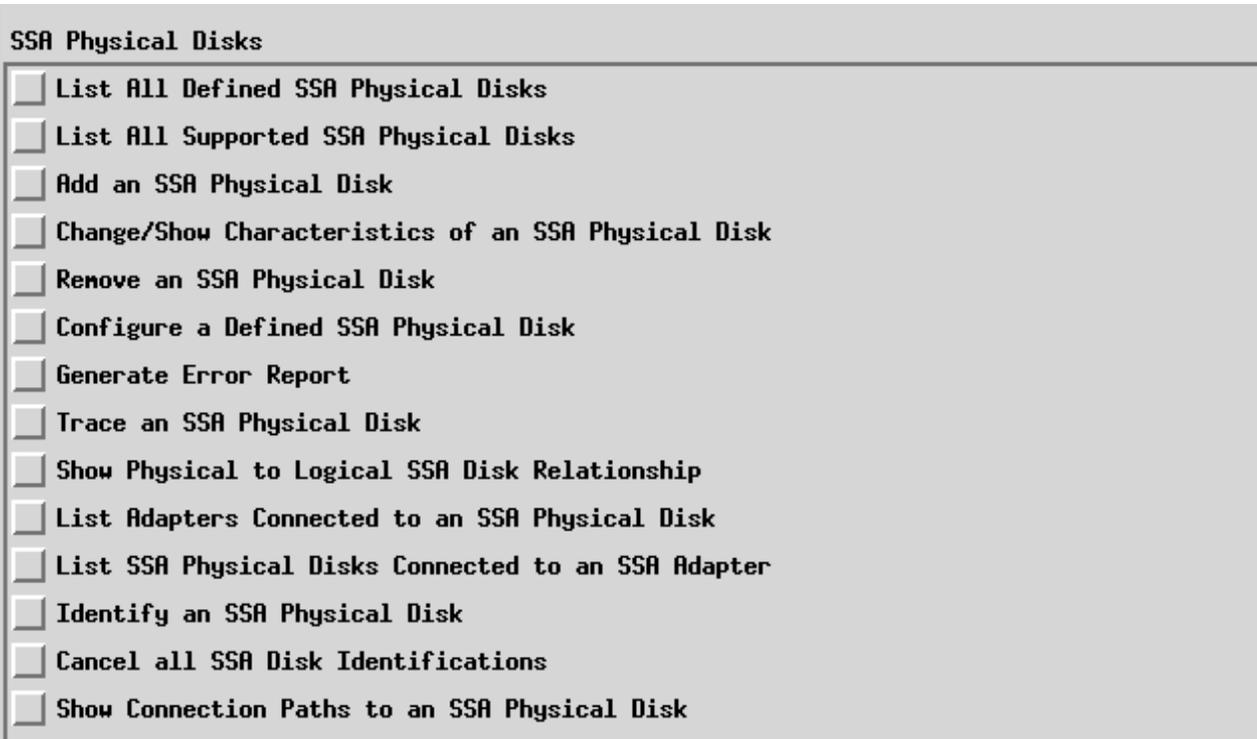


Figure 48. SSA Physical Disk Menu

Selecting **SSA RAID Arrays** on the SMIT Devices menu takes you to the SSA RAID Arrays menu (Figure 49), which gives you 12 further options.

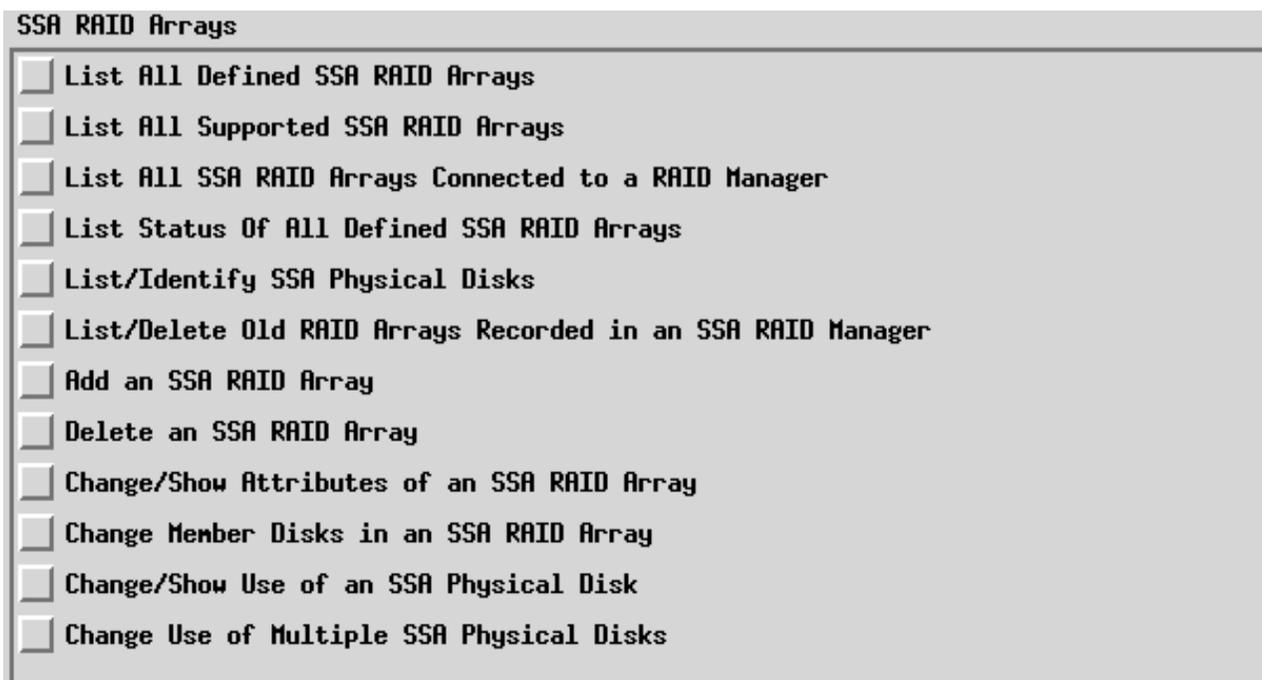


Figure 49. SSA RAID Arrays Menu

4.3.2 AIX command line entries

This topic lists and describes six AIX commands.

ssaxlate command

- Purpose:
Translates between SSA logical disks (hdisks) and SSA physical disks (pdisks)
- Syntax:
ssaxlate -l LogicalDiskName
ssaxlate -l PhysicalDiskName
- Description:
If the parameter is an SSA logical disk, the output will be a list of SSA physical disk names supporting that logical disk. If the parameter is an SSA physical disk, the output will be a list of SSA logical disk names that depend on that physical disk.
- Flags:
-l DiskName — Specifies the SSA logical or physical disk

ssaadap command

- Purpose:
Lists the adapters to which an SSA logical disk or physical disk is connected.
- Syntax:
ssaadap -l LogicalDiskName
ssaadap -l PhysicalDiskName
- Description:
The output is the list of SSA adapters to which the SSA logical or physical disk is connected. If there is more than one adapter, the first adapter listed is the primary adapter.
- Flags:
-l DiskName — Specifies the SSA logical or physical disk

ssaidentify command

- Purpose:
Sets or clears identify mode for an SSA physical disk
- Syntax:
ssaidentify -l PhysicalDiskName -y
ssaidentify -l PhysicalDiskName -n
- Description:
If the -y flag is specified the disk is set to identify mode and flashes an amber light at approximately 1 second intervals. The -n flag turns off identify mode.
- Flags:
-l PhysicalDiskName — Specifies the device to place in identify mode.
-y — Turns on identify mode

-n — Turns off identify mode

ssaconn command

- Purpose:

Displays the SSA connection details for the physical disk

- Syntax:

ssaconn -l PhysicalDiskName -a AdapterName

- Description:

The ssaconn command performs a function similar to the SSA Link Verification service aid.

The output is the PhysicalDiskName followed by the adapter name followed by four hop counts. The four hop counts represent the number of SSA devices between the adapter and the A1, A2, B1, B2 ports of the adapter, respectively. If the disk is not connected to a particular adapter port, the hop count is replaced by a hyphen character.

- Flags:

-l PhysicalDiskName — Specifies the SSA physical disk whose connection details should be listed

-a AdapterName — Specifies the adapter to list the connection details

ssacand command

- Purpose:

Displays the unused connection locations for an SSA adapter

- Syntax:

ssacand -a AdapterName -P -L

- Description:

The ssacand command is used to list out the available connection locations on an SSA adapter. These locations correspond to devices that are connected to the adapter, but for which AIX devices are not configured.

- Flags:

-a AdapterName — Specifies the adapter for which to list the connection locations

-P — Produces a list of possible connection locations for SSA physical disks

-L — Produces a list of possible connection locations for SSA logical disks

ssadisk command

- Purpose:

Displays the names of SSA disks connected to an SSA adapter

- Syntax:

ssadisk -a AdapterName -P -L

- Description:

The ssadisk command is used to list out the names of SSA disks connected to an SSA adapter. These names correspond to devices that are in the

customized device database, and have the SSA adapter as their adapter_a or adapter_b attribute.

- Flags:
 - a AdapterName — Specifies the adapter to which the disks are connected
 - P — Produces a list of SSA physical disks
 - L — Produces a list of SSA logical disks

4.3.3 Utilities

These utilities are a series of script files that primarily assist with problem determination.

vpfe

vpfe is a menu driven application that combines useful commands and SMIT options to assist with problem solving. A set of data collection options is also provided. This utility incorporates Maymap. Here are vpfe's characteristics:

- Virtual PFE v 1.04 README file (English version).

Download vpfe1.04.tar file

- SSA menu featuring options:
 - SSA Diagnostics
 - SSA SERVICE AIDS
 - List Status Of All Defined SSA RAID Arrays
 - List All Defined SSA RAID Arrays
 - List varied on volume groups
 - List SSA Device Drivers
 - List SSA Adapters and Microcode
 - List pdisk Microcode
 - List Rejected Array Disks
 - Error Log Decoder
 - SSA Error Log Analysis
 - Health Check
 - Identify SSA disks.
 - pdisk <--> hdisk translate
- 7135 menu featuring options:
 - List 7135 Devices
 - List 7135 Device Drivers
 - List dac Microcode
 - List varied on volume groups
 - List 7135 Configuration
 - Cache Mirror Status
 - List Array Status
 - smitty raidiant
 - smitty ibmraid
 - List Drive Status
- Data collection options:
 - sweep collection utility
 - ls7133
 - snap -ac
 - maymap

- Simplified version of error log decoder to work on the system error log. Japanese menu options when LANG=Ja_JP.

mergelogs

mergelogs is an AIX utility that will merge multiple errpt -a logs from different systems that are connected to common storage to create a single log in date/time order. Here are its characteristics:

- Merge Logs v 1.0

The AIX mergelog utility takes multiple errpt -a files and creates a single file with all the error entries in date/time sequence. This is useful when multiple systems are connected to a common set of storage. Errors or actions on one system can result in error log entries in another system. By creating a single log file make the process of analyzing the error log easier.

- Syntax

```
mergelogs log1 log2 logn > logs.merged
```

The above example will merge the three error log files log1 log2 and logn and place the output in a file called logs.merged in the local directory

ciao

ciao stands for *check it all out*. It is a utility that will check the SSA device driver, adapter, and pdisk code levels against the current or latest versions available. The utility will be updated as new levels are made available. Here are its characteristics:

- Syntax

```
ciao [-d] [-h]
```

- Flags

-d display down level components only
-h display help

sweep

This is a data collection utility to collect storage related information. You can also send sweep to a customer sweep via email from this site. Here are its flags:

- Flags

-d display down level components only
-h display help

4.3.4 Remote Systems Management

Remote Systems Management (RSM) is a free Web-based configuration tool that allows you to configure and manage SSA disk systems, such as the IBM 7133, using supported HTML Web browsers, thereby, providing true *remote* systems management. It provides an interface to all SSA adapters, disk drives, and enclosures. RSM has been available for Windows NT since 1998, and for Novell NetWare since 1999. Starting March 7, 2000, RSM now supports AIX 4.3.2 and AIX 4.3.3, and it can be downloaded from this Web site:

<http://www.storage.ibm.com/hardsoft/products/ssa/rsm/>

RSM provides an explorer style view of the resources connected to all SSA adapters in the using system. Resources can be listed by type, for example,

JBOD, or RAID5, and by usage, for example, Free, System or other usage values. Selecting a resource from the explorer view displays full resource details in a configuration view. From there, all SSA configuration actions currently available via SMIT panels can be performed. In addition, for RAID-10 customers requiring the split-site function, a confirmation step is provided prior to array creation to verify that the correct components have been assigned to each site.

In addition to the explorer view, resources can be viewed from a logical or physical perspective, that is, by type or by connection. The physical view provides an invaluable tool to aid the diagnosis of link breaks and some cabling errors. This shows resources as they are physically connected to the SSA ports on the adapter(s) including any remote adapters in n-way loops.

RSM also provides an interface to the SCSI Enclosure Services (SES) of the 7133 D40 and T40. It gives detailed information regarding the status of the internal components of these enclosures, including power supplies, cooling fans, bypass cards and disk drives. All the SES controller settings, such as temperature thresholds, four character display names, and bypass card settings can also be modified.

RSM allows customers to use an Internet connected browser to check the level of all their SSA host software and microcode, and, if applicable, to download the latest level directly to the using system's disk drive. The updates residing on the disk drive can then be applied to the disk system.

Figure 50 on page 82 shows a sample RSM screen image.

4.3.5 SSA cable

This is not a tool for controlling SSA networks. Instead, the program is designed to help you cable an SSA subsystem composed of multiple loops and multiple hosts. All you need to specify is which hosts go into which loops, and how many disks you intend to assign to each host in each loop. You may also choose between the defaults 7133-010 and 7133-020.

The program will optimize the order of nodes in the loop. It does this by optimizing to the nearest group of disks, for example to the nearest singleton or pair or quad. It will always try to deliver the minimum number of unassigned disks or gaps in the disk drawers.

The grouping factor may be specified differently for each loop, and is the first numeric field in the "loop:" line. Additionally, if you have a line beginning "*clus:*" it will start a fresh disk tray for every host, and fill these up as best it can, trying to keep drawers associated with hosts.

A node will have the disks that it owns in a loop attached equally to both ports of the connector, as far as this is consistent with the grouping factor. The grouping factor can be any integer greater than zero, and controls how disks assigned to a node in a loop are distributed on each of the adapter ports - for example, B1 and B2. A different grouping factor may be specified for each loop.

For example, a node with six disks in a loop with a grouping factor of 1 will have three disks on port A1 and three on port A2. A grouping factor of 2 will give four disks and two disks, as will a factor of 4. A node with eight disks in a loop and a grouping factor of 1 or 2 or 4 will split its disks four and four, but a grouping factor

of 3 would cause it to split them five and three. The disks are always packaged in quad slots, and the program will display any vacant slots.

Example file ssainxmp4c20

clus:

type: 020

loop: 4 4 8 12 -2

loop: 4 12 8 4 -2

name: Matthew

name: Andrew

name: Nicholas

This input file gives the following output:

```
*****
*
*  SSACABLE V2.1  -  example file ssainxmp4c20
*
*                J S Watts IBM UK Petroleum
*
*****
```

```
Node Matthew   needs  1 adaptors and  16 disks
Node Andrew    needs  1 adaptors and  16 disks
Node Nicholas  needs  1 adaptors and  16 disks
```

This configuration needs:

3 enhanced SSA adaptors, 48 disks in 3 trays and 16 cables

The cabling reflects the use of 7133-020

The disks are grouped within the rack, where possible, by hostname

**** Program limits (changeable on request) ****

```
Number of nodes:  48
Number of loops: 112
Number of trays:  200
Number of disks/loop: 96
```

```
*****
*
*  SSACABLE V2.1  -  example file ssainxmp4c20
*
*  Placement of disks within trays, by host ID
*
*****
```

1 ====Quad=== 4 5 ====Quad=== 8 9 ====Quad===12 13 ====Quad===16

Tray 2: 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2

```

Tray 1:  1  1  1  1      1  1  1  1      1  1  1  1      1  1  1  1
Tray 0:  0  0  0  0      0  0  0  0      0  0  0  0      0  0  0  0

```

1 ====Quad=== 4 5 ====Quad=== 8 9 ====Quad===12 13 ====Quad===16

Legend: Numbers represent hosts, as follows:

```

0 is Matthew
1 is Andrew
2 is Nicholas

```

```

*****
*
* SSACABLE V2.1 - example file ssainxmp4c20
*
* Placement of disks within trays, by loop ownership*
*
*****

```

1 ====Quad=== 4 5 ====Quad=== 8 9 ====Quad===12 13 ====Quad===16

```

Tray 2:  0  0  0  0      0  0  0  0      0  0  0  0      1  1  1  1
Tray 1:  0  0  0  0      0  0  0  0      1  1  1  1      1  1  1  1
Tray 0:  0  0  0  0      1  1  1  1      1  1  1  1      1  1  1  1

```

1 ====Quad=== 4 5 ====Quad=== 8 9 ====Quad===12 13 ====Quad===16

Legend: Numbers represent loops

```

1*****
*
* SSACABLE V2.1 - example file ssainxmp4c20
*
*          HOST/DISK LAYOUT FOR LOOP  0
*
*          (optimized to nearest  4 disks)
*
*****

```

```

*****
*          port A1 *>-----+
Matthew  * adaptor 0      *
(4 disks) *          port A2 *>-----+
*****
*          |
*          *****
*          * 0 *
*          8 disks in 8 slots: *====*
*          * 8 *
*          *****
*          |
*          port A1 *>-----+

```

```

Nicholas      * adaptor 0      *
(12 disks)   *          port A2 *>-----+
             *          *****
                                     |
                                     *****
                                     * 4 *
             8 disks in 8 slots: *====*
                                     * 4 *
                                     *****

             *****
             |
Andrew       *          port A1 *>-----+
(8 disks)   * adaptor 0      *
             *          port A2 *>-----+
             *          *****
                                     |
                                     *****
                                     * 4 *
             8 disks in 8 slots: *====*
                                     * 4 *
                                     *****
                                     |
                                     +-----+

```

```

1*****
*
* SSACABLE V2.1 - example file ssainxmp4c20
*
*          HOST/DISK LAYOUT FOR LOOP 1
*
*          (optimized to nearest 4 disks)
*
*****

```

```

             *****
             |
             *          port B1 *>-----+
Matthew     * adaptor 0      *
(12 disks) *          port B2 *>-----+
             *          *****
                                     |
                                     *****
                                     * 4 *
             8 disks in 8 slots: *====*
                                     * 4 *
                                     *****

             *****
             |
Andrew     *          port B1 *>-----+
(8 disks) * adaptor 0      *
             *          port B2 *>-----+
             *          *****
                                     |
                                     *****
                                     * 4 *
             8 disks in 8 slots: *====*
                                     * 4 *
                                     *****

             *****
             |
Nicholas   *          port B1 *>-----+
(4 disks) * adaptor 0      *
             *          port B2 *>-----+
             *          *****
             |

```

```

*****
* 0 *
8 disks in 8 slots: *===*
* 8 *
*****
|
+-----+

```

```

1*****
*
* SSACABLE V2.1 - example file ssainxmp4c20
*
*          CABLING SCHEDULE FOR LOOP 0
*
*          (optimized to nearest 4 disks)
*
*****

Cable 1: Matthew adapter 0 port A2 to tray 2 port 1
                    tray 2 port 4 bypass to port 5
Cable 2: Nicholas adapter 0 port A1 to tray 2 port 8
Cable 3: Nicholas adapter 0 port A2 to tray 2 port 9
Cable 4:.....tray 2 port 12 to tray 1 port 1
Cable 5: Andrew adapter 0 port A1 to tray 1 port 4
Cable 6: Andrew adapter 0 port A2 to tray 1 port 5
Cable 7:.....tray 1 port 8 to tray 0 port 1
Cable 8: Matthew adapter 0 port A1 to tray 0 port 4

```

```

1*****
*
* SSACABLE V2.1 - example file ssainxmp4c20
*
*          CABLING SCHEDULE FOR LOOP 1
*
*          (optimized to nearest 4 disks)
*
*****

Cable 9: Matthew adapter 0 port B2 to tray 0 port 5
Cable 10:.....tray 0 port 8 to tray 1 port 9
Cable 11: Andrew adapter 0 port B1 to tray 1 port 12
Cable 12: Andrew adapter 0 port B2 to tray 1 port 13
Cable 13:.....tray 1 port 16 to tray 2 port 13
Cable 14: Nicholas adapter 0 port B1 to tray 2 port 16
Cable 15: Nicholas adapter 0 port B2 to tray 0 port 9
                    tray 0 port 12 bypass to port 13
Cable 16: Matthew adapter 0 port B1 to tray 0 port 16

```

```

*****
*
* SSACABLE V2.1 - example file ssainxmp4c20
*
* Cables which interconnect disk ports
*

```

***** Loop 0 *****

Cable 4:.....tray 2 port 12 to tray 1 port 1
Cable 7:.....tray 1 port 8 to tray 0 port 1

***** Loop 1 *****

Cable 10:.....tray 0 port 8 to tray 1 port 9
Cable 13:.....tray 1 port 16 to tray 2 port 13

```
*
* SSACABLE V2.1 - example file ssainxmp4c20
*
* Cables which connect processors to disks
*
*****
```

***** Loop 0 *****

Cable 1: Matthew adapter 0 port A2 to tray 2 port 1
Cable 2: Nicholas adapter 0 port A1 to tray 2 port 8
Cable 3: Nicholas adapter 0 port A2 to tray 2 port 9
Cable 5: Andrew adapter 0 port A1 to tray 1 port 4
Cable 6: Andrew adapter 0 port A2 to tray 1 port 5
Cable 8: Matthew adapter 0 port A1 to tray 0 port 4

***** Loop 1 *****

Cable 9: Matthew adapter 0 port B2 to tray 0 port 5
Cable 11: Andrew adapter 0 port B1 to tray 1 port 12
Cable 12: Andrew adapter 0 port B2 to tray 1 port 13
Cable 14: Nicholas adapter 0 port B1 to tray 2 port 16
Cable 15: Nicholas adapter 0 port B2 to tray 0 port 9
Cable 16: Matthew adapter 0 port B1 to tray 0 port 16

```
*
* SSACABLE V2.1 - example file ssainxmp4c20
*
* Cables which join hosts to hosts
*
*****
```

***** Loop 0 *****

***** Loop 1 *****

```

*****
*
* SSACABLE V2.1 - example file ssainxmp4c20
*
* Cables sorted by hostname
*
*****

Cable  5: Andrew      adapter 0 port A1 to tray 1 port  4
Cable  6: Andrew      adapter 0 port A2 to tray 1 port  5
Cable 11: Andrew      adapter 0 port B1 to tray 1 port 12
Cable 12: Andrew      adapter 0 port B2 to tray 1 port 13
Cable  8: Matthew     adapter 0 port A1 to tray 0 port  4
Cable  1: Matthew     adapter 0 port A2 to tray 2 port  1
Cable 16: Matthew     adapter 0 port B1 to tray 0 port 16
Cable  9: Matthew     adapter 0 port B2 to tray 0 port  5
Cable  2: Nicholas    adapter 0 port A1 to tray 2 port  8
Cable  3: Nicholas    adapter 0 port A2 to tray 2 port  9
Cable 14: Nicholas    adapter 0 port B1 to tray 2 port 16
Cable 15: Nicholas    adapter 0 port B2 to tray 0 port  9

```

4.4 SSA spare tool

The SSA spare tool is a tool for managing Serial Storage Architecture (SSA) storage networks. This tool functions with the family of SSA adapters, SSA disk units, and SSA enclosures that are developed by the IBM Storage System Division.

The SSA Spare Tool works with the Logical Volume Manager (LVM) to automatically identify stale partitions or missing physical volumes in LVM mirrored volume groups. If the SSA Spare Tool finds any stale partitions or missing physical volumes, the tool will:

- Automatically resynchronize the stale partition, if this can be done without requiring a hardware replacement.
- Issue an e-mail message, indicating what has been done, or if a disk must be replaced.
- Logically replace the failed disk in the volume group with a preassigned spare disk.
- Resynchronize the new disk with the remaining disks in the volume group.
- Notify the user when it has logically replaced a failed disk and resynchronized the replacement disks with the other disks in the volume group.

When the failed disk is physically replaced and the data is resynchronized, other scripts are provided to:

- Move the data from the temporary spare disk to the new replacement disk.
- Prepare the spare disk for future use as a spare if another disk failure occurs.

This tool maintains a protected, software mirrored, environment and automatically moves data from a failed disk to a spare disk when required. The SSA Spare Tool supports all SSA disk drive units.

To enable the SSA Spare Tool you must install the filesets associated with this tool on the hot spare server and the related clients. These filesets enable you to setup hot sparing for mirrored SSA disks in a clustered environment. A hot spare is a disk which is made available, dynamically, when a mirrored disk fails. When the mirrored disk fails, these scripts replace the failed disk with the hot spare. The mirrored data is then rebuilt onto the hot spare disk. One machine, designated the server, manages hot sparing, and disk replacement for all of its client machines. This server can also be a client.

Operating requirements

To ensure the proper operation of the SSA Spare Tool, you should ensure the following:

- All the SSA disks in the storage networks that are monitored have the logical volumes 100% mirrored and are on separate physical volumes.
- The volume group is not varied on in concurrent mode.
- The hot spare server must have rsh access to all client nodes for which it is managing hot spares.
- This tool does not support boot or dump logical volume types.
- These scripts handle single failures in a volume group. The scripts do not originate any action if more than one disk fails at the same time in any one volume group. If the sparing process is running when another disk fails, the SSA Spare Tool will complete the current sparing process.
- Appropriate spare SSA disks are available when needed.

Notes

1. The SSA Spare Tool does not support failures of Redundant Array of Independent Disks (RAID) arrays or disks in RAID arrays.
2. Concurrent disk failures in separate volume groups are handled serially.

4.5 Managing SSA disks in a Sun Solaris environment

There are two methods of attaching SSA disks to a SUN host system. For systems that use the Sun SBus, there is a native adapter card, which supports one SSA loop of up to 48 disks. You can have two adapter cards per host, which means that up to 96 SSA disk drives can be attached to a single host. For Sun systems that do not use the SBus, you can use the IBM 7190 SCSI to SSA converter or IBM 2108 SAN Data Gateway. These external attachable devices do not require internal installation or configuration of the host system.

The disk attachment method determines how the disks are managed. With the SSA SBus adapter, you can invoke the SSA Service Aid Utility through the command `ssau`. The utility is automatically installed as part of the adapter installation process. As the Sun operating system does not understand the concept of SSA disks, you must map the SSA IDs to SCSI IDs. Use the SSA Configuration Utility which you invoke through the `ssacf` command.

If you use the IBM 7190 or IBM 2108 to attach the SSA disks, install the SSA Service Functions from the CD-ROM that is supplied with the device. The command `SSA_SERV` invokes the following functions:

- Download microcode
- Display VPD
- Identify a disk drive module
- Set/reset service mode
- Disk drive diagnostics
- 7190 model 100 diagnostics
- Error log analysis
- Clear Check mode
- Activities monitor
- Show SSA topology
- Force SSA master
- Force Web reset
- Write cache option
- Synchronous write
- Quit

The above functions are fully described in the redbook *Exploiting SSA Disk Subsystems in Sun Solaris Platform Environments*, SG24-5083.

4.5.1 Other tools

Other tools that are commonly used in the Sun Solaris environment are DiskSuite and Veritas. These tools are described in the redbook *Exploiting SSA Disk Subsystems in Sun Solaris Platform Environments*, SG24-5083.

4.6 Managing SSA disks in an HP-UX environment

The way of connecting SSA disks to a host running the HP-UX operating system is to use a SCSI to SSA converter, such as the IBM 7190, IBM 2108 SAN Data Gateway, or the VICOM SLIC.

If the SSA disks are attached using the IBM 7190 or IBM 2108, then the SSA service functions should be installed from the CD-ROM that is supplied with the device. To invoke the service functions listed here, use the command `SSA_SERV`:

- Download microcode
- Display VPD
- Identify a disk drive module
- Set/reset service mode
- Disk drive diagnostics
- 7190 Model 100 diagnostics
- Error log analysis
- Clear check mode
- Activities monitor
- Show SSA topology
- Force SSA master
- Force Web reset
- Write cache option
- Synchronous write
- Quit

The above functions are fully described in the redbook *Exploiting SSA Disk Subsystems in Sun Solaris Platform Environments*, SG24-5083.

4.7 Managing SSA disks in a Windows NT environment

The main tool used to manage SSA disks in this environment is SSA Remote Systems Management (SSA RSM). SSA RSM, formerly known as Wabash, is a Web-based SSA device configurator and service aids utility. It is designed as a replacement for the text-based SSA configurator that ships as standard with the SSA PC adapter. Two versions are available, a Netfinity plug-in module for IBM PC servers running TME10 Netfinity Version 5.xx, and a Windows NT stand-alone device. Both versions have the same look, feel, and operability.

With SSA RSM you can easily manage SSA-based storage systems locally and remotely. SSA RSM, is a Web browser based tool, that provides for RAID Array configuration and service activities of the IBM SSA disk storage system. If Netfinity Manager is used, it forms part of the WebGUI panel. The Netfinity Alert Manager already notifies you of critical alerts and threshold warnings. Now all SSA alerts will be forwarded to the Netfinity Alert Manager as well, therefore, simplifying the management of error conditions.

SSA RSM enables you to:

- View the SSA adapter configuration information, including the physical location.
- Show all of the components attached to a particular SSA adapter.
- Create, attach and detach RAID arrays.
- Assign hot spares to RAID arrays.
- Identify individual disks by flashing the disk LEDs, so that you can locate specific disks.
- Check and download the microcode levels of SSA adapters and disk drives.

Figure 50 shows the initial screen view when SSA RSM is started.

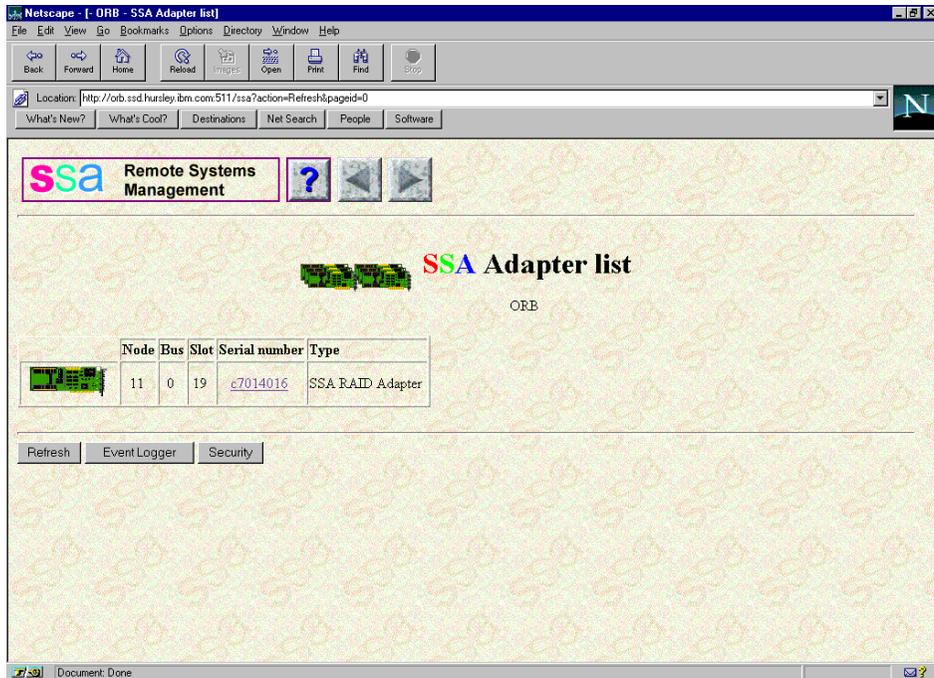


Figure 50. SSA RSM initial screen view

In this example, there is only one SSA adapter in the host PC. By clicking on the serial number of the adapter, you bring up a detailed view of the product information, as shown in Figure 51.

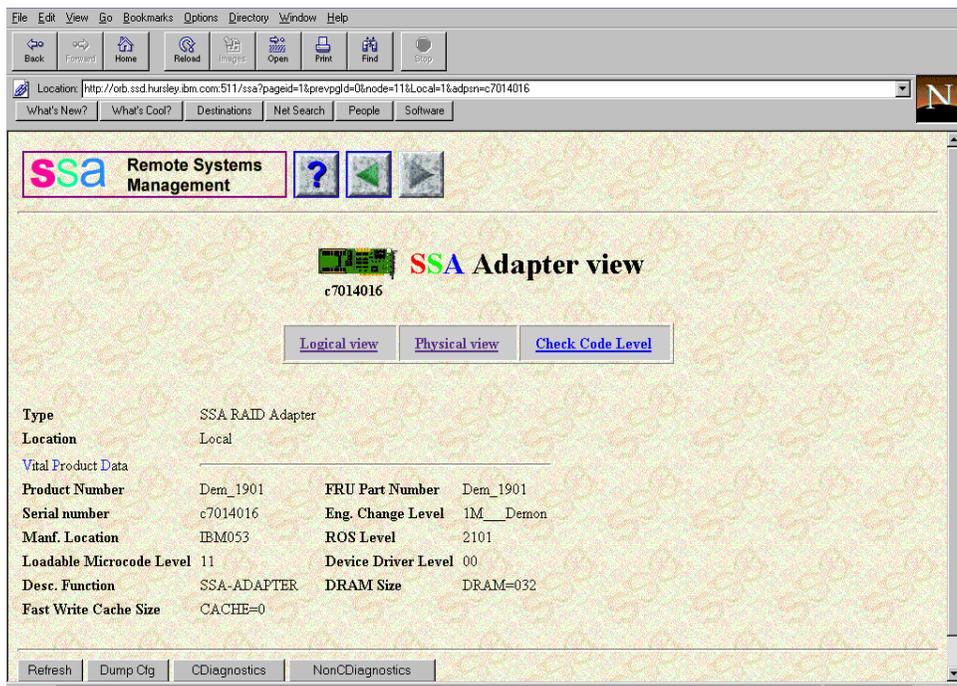


Figure 51. Detailed view of Adapter product information

In the SSA Adapter view, clicking on the **Physical view** of the disks brings up the screen shown in Figure 52. You can see that the four ports on the adapter are listed, with the disks shown in the order as they connect to the adapter.

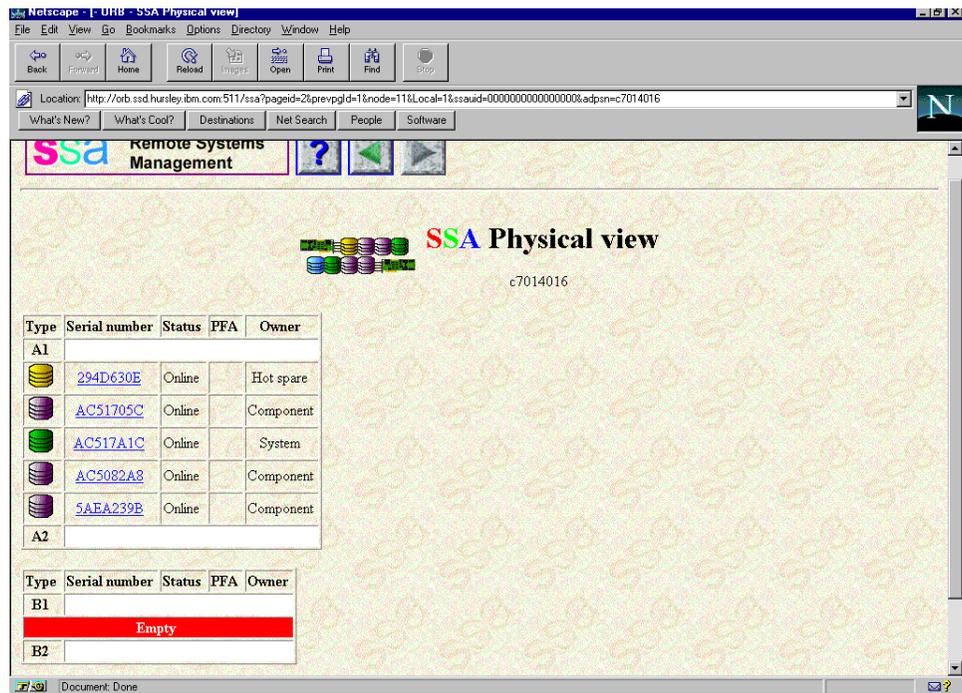


Figure 52. Physical view from Adapter

Click on one of the disks to display the disk VPD, as shown in Figure 53.

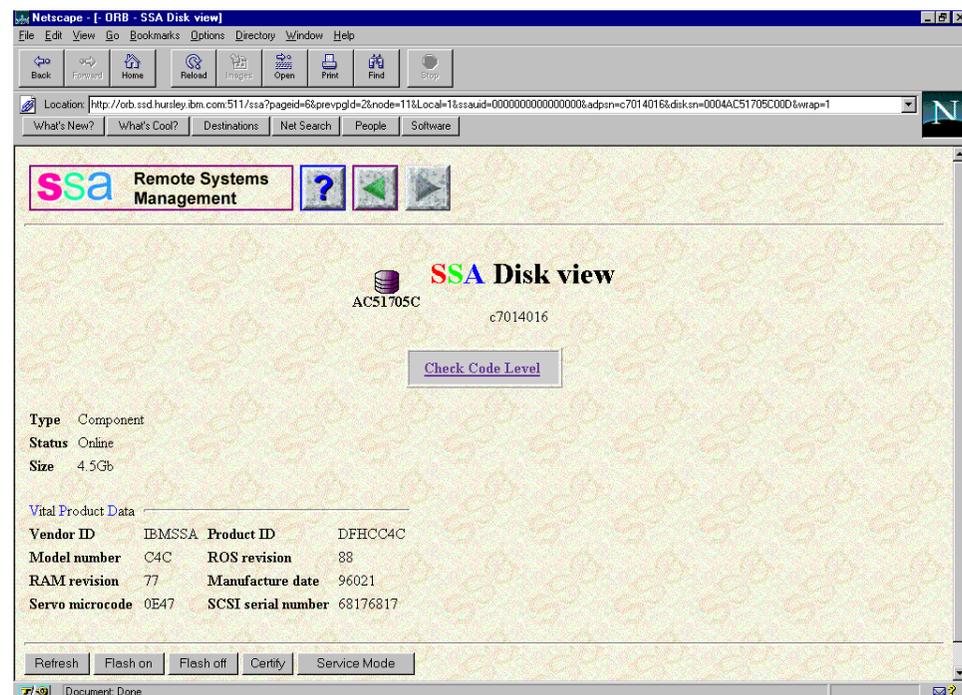


Figure 53. VPD of Individual disk

When you select the **Logical view** in the SSA Adapter view window, you get to the windows shown in Figure 54 and Figure 55.

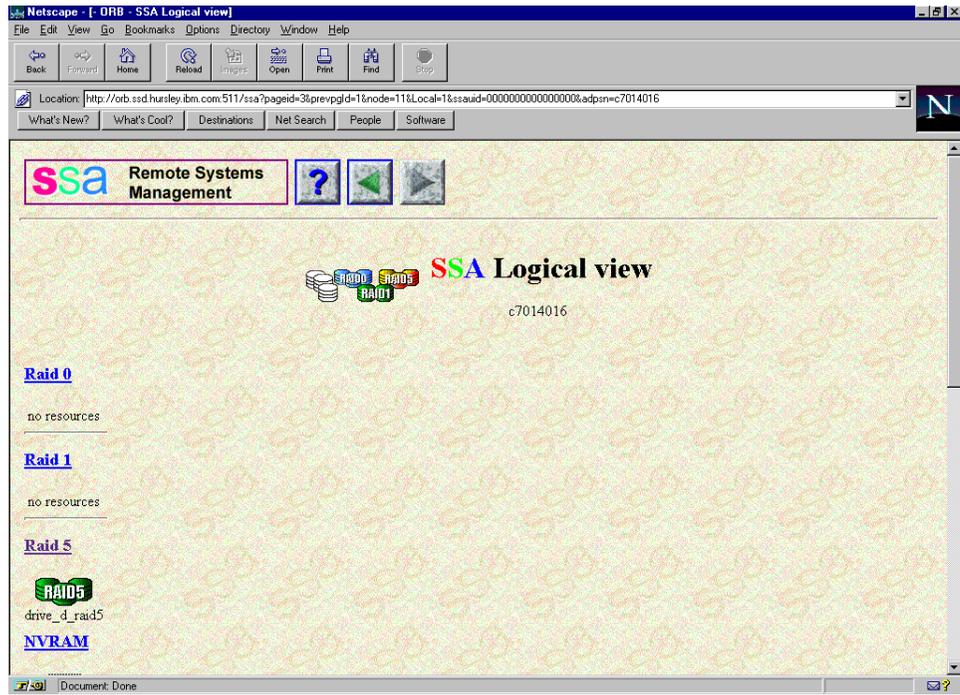


Figure 54. Logical view from Adapter: part 1

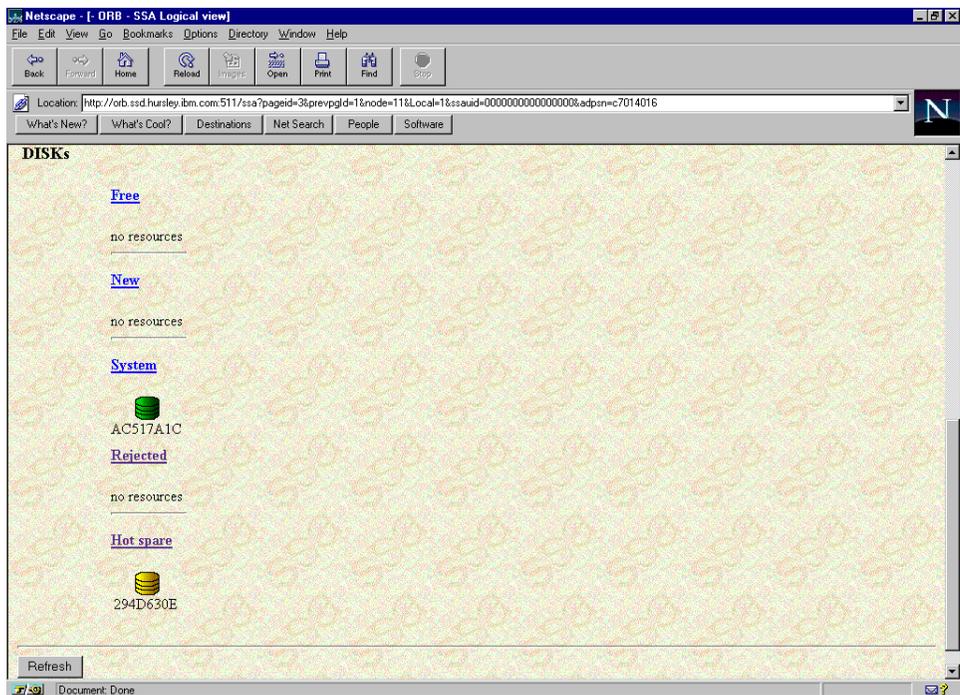


Figure 55. Logical view from Adapter: part 2

While in the SSA Adapter list window, when you select the **Event logger**, you get a detailed view of all recorded SSA events (Figure 56).

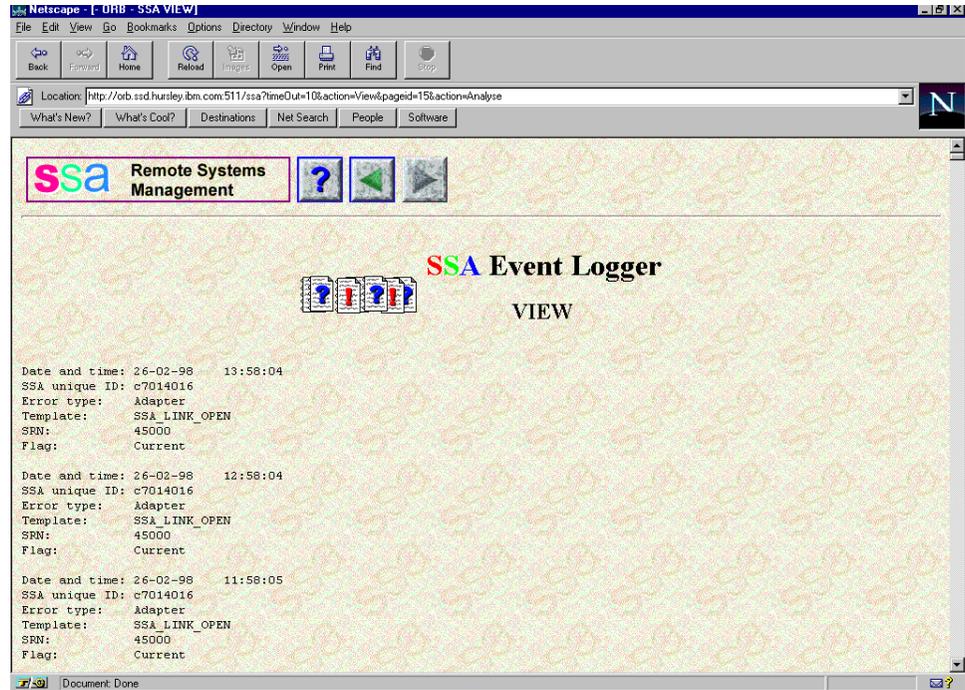


Figure 56. View of Event Logger

You can refresh and analyze this log.

4.8 Managing SSA disks in other environments

In order to attach SSA disks to other operating environments, you have to use non-IBM adapters, such as those made by Pathlight, or SCSI-to-SSA converters, such as the VICOM 2000.

Pathlight adapters

Pathlight makes a range of adapters that can attach SSA disks to Windows NT systems, Apple Macs, and Sun systems. We recommend using an IBM adapter unless the Pathlight adapter has additional functionality that is required by the customer.

The Pathlight Toolbox is a utility for updating the Image80 adapter firmware and SSA disk drive firmware on a Macintosh PowerPC computer. It checks your SSA cables, tells you whether your network topology is functioning properly, and shows you a map of your devices. Apple Guide makes it easy for you to learn how to use the Pathlight Toolbox.

VICOM product range

VICOM produces a range of SCSI-to-SSA converters as well as manufacturing the IBM 7190 and IBM 2108. A product comparison is given in Table 19.

Table 19. VICOM and IBM product comparison

	IBM 7190	IBM 2108 S20	SLIC 1355	UltraLink 1000	UltraLink 2000
Platform tested	SUN, HP, DEC	Windows-NT, SUN, HP, Compaq	Windows-NT, UNIX	Windows-NT UNIX	Windows-NT UNIX
HW RAID-1, RAID-0	No	Yes	No	No	Yes
Multihost support	1 - 4 hosts	1 - 16	1 - 16 hosts	1 - 16 hosts	1 - 16 hosts
Fibre extender support	No	Yes	No	Yes	Yes
19 inch rack mount option	No	Yes	No	Yes	Yes
Burst data rate	20 MB/s	40 MB/s	20 MB/s	40 MB/s	40 MB/s
Sustained data rate	18 MB/s	35 MB/s	18 MB/s	35 MB/s	35 MB/s
I/O throughput	1900 I/O/s	3000 I/O/s	1900 I/O/s	3000 I/O/s	3000 I/O/s
I/O throughput with 2 UltraLinks	N/A	5100 I/O/s	N/A	5100 I/O/s	5100 I/O/s
Capacity per SCSI channel	1 - 48 drives	1 - 64	1 - 64 drives	1 - 64 drives	1 - 64 drives
Capacity per SCSI channel with multiple UltraLinks	N/A	1 - 512 drives	N/A	1 - 120 drives	1 - 512 drives

The tools available to help manage the disk subsystem are those utilities provided with the adapter and any tools provided with the operating system. The utilities provided with the VICOM products enable you to:

- Download microcode
- Display VPD
- Identify a disk drive module
- Set/reset service mode
- Run disk drive diagnostics
- Run VICOM self test diagnostics
- Analyze error log
- Clear check mode
- Monitor activities
- Show SSA topology
- Force SSA master
- Force Web reset
- Write cache option
- Synchronous Write
- Quit

All of the above functions are fully described in the redbook *Exploiting SSA Disk Subsystems in Sun Solaris Platform Environments*, SG24-5083.

4.9 IBM StorWatch Serial Storage Expert (StorX)

The IBM StorWatch Serial Storage Expert (StorX) is a graphical tool for managing Serial Storage Architecture (SSA) storage networks. StorX functions with the family of SSA adapters, SSA disk units, and SSA enclosures that are developed by the IBM Storage System Division. StorX provides:

- A Windows 95, Windows 98, or Windows NT 4.0 graphical user interface. This interface enables you to plan new SSA storage networks, create plans from existing SSA storage networks, or monitor existing SSA storage networks.
- Discovery and display of the SSA topology, capacity, and availability of SSA storage resources.
- Monitoring SSA storage networks for adapter errors and disk errors, topology changes, and redundant array of independent disks (RAID) events.
- A physical view of disk storage in an SSA storage network.
- Support for SSA enclosures (7133, 7131, and 3527).
- Support for SSA internal backplanes.
- Discovery and display of SSA SCSI Enclosure Services (SES) enclosures.
- The ability to simplify the graphical presentation of your SSA storage networks by collapsing storage networks into a single storage network icon.
- Support for SSA adapters and disks that are attached to RS/6000 servers and SP2 systems that are running AIX.
- Support for SSA adapters and disks that are attached to Intel platform servers that are running Windows NT 4.0.
- The ability to extract and plan detailed information about the SSA devices in your storage networks through a report function.
- The ability to export information to a database or a spread sheet.
- The ability to export a graphical view of a management set, in a BMP or GIF format, for use in a graphic program or Web page.
- The ability to automatically locate an SSA device associated with an event that has occurred in your storage network.
- The ability to print cable labels.

You can use the StorX Planner to visually plan your SSA storage network. StorX provides an SSA parts palette for you to add or remove adapters, disks, connections, enclosures, and text and box annotations. Editing tools enable you to select, move, and change attributes for any devices in the storage network configuration you are planning.

You can use the StorX Live Viewer to discover SSA storage networks on AIX or Windows NT 4.0 hosts. StorX automatically discovers topology and attributes and displays them on any network-attached PC that is running Windows 95, Windows 98, or Windows NT 4.0. Monitoring for SSA events allows StorX to report adapter errors, disk errors, topology changes, and RAID events.

For more information on StorX, please refer to Chapter 5, “Monitoring and managing SSA environment with StorX” on page 89.

Chapter 5. Monitoring and managing SSA environment with StorX

This chapter explains how to monitor and manage your various SSA configurations, perform some troubleshooting, and run diagnostics with StorX. We do not cover any other SSA tools. For a description of these other tools, please refer to chapter 3 of *Monitoring and Managing IBM SSA Subsystems*, SG24-5251.

5.1 Customizing a StorX live view

The live viewer in StorX shows most of the system data that you want to see. Each adapter is described and displayed with its host name. The disks can be fully described with serial number, logical name, size and type. A typical StorX viewer screen for SSA disks attached to a Windows NT system is shown in Figure 57. Only part of the canvas has been shown for clarity.

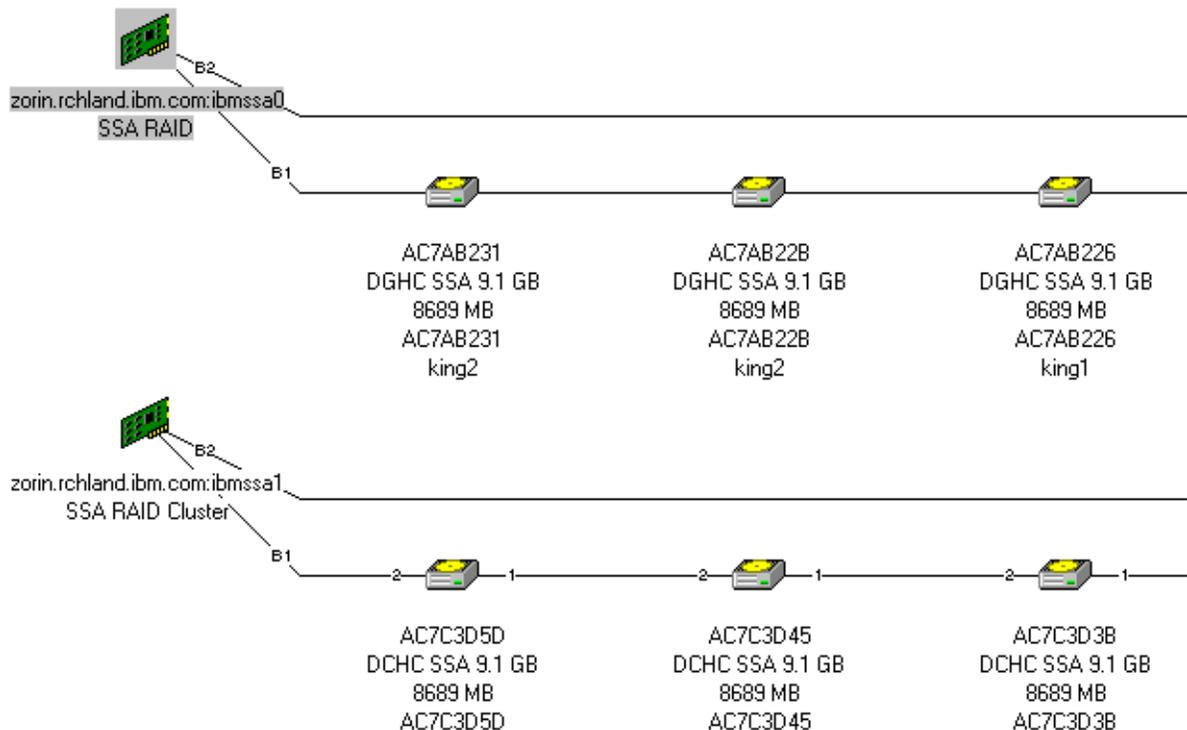


Figure 57. Uncustomized StorX live view

The information displayed for each component in the storage network can be customized in the Management Set Properties panel. The menu items are shown in Figure 58. Any combination of the properties can be selected for display. Similar menus are available for the other components in the system.

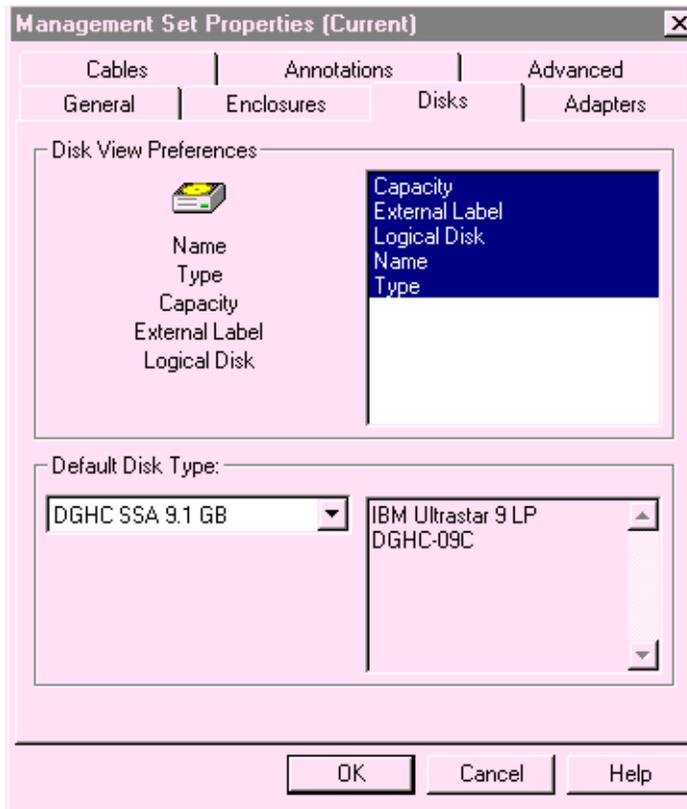


Figure 58. Management Set Properties panel

One of the major problems with the uncustomized canvas is that for large, complex configurations, it can be difficult to see all the devices easily. By adding the 7133 enclosure icon to the canvas, which can be obtained from the parts palette by a simple drag and drop operation, and dragging the disks into the enclosure, the size of the canvas can be reduced. It is possible to further customize the picture by adding descriptive text to the canvas. You may add as much text as required to suit your requirements. It is also possible to show the 7133 enclosure in its four component quadrants. This feature is useful when each quadrant is attached to different adapters. Figure 59 shows a 7133 enclosure broken down into its quadrants.

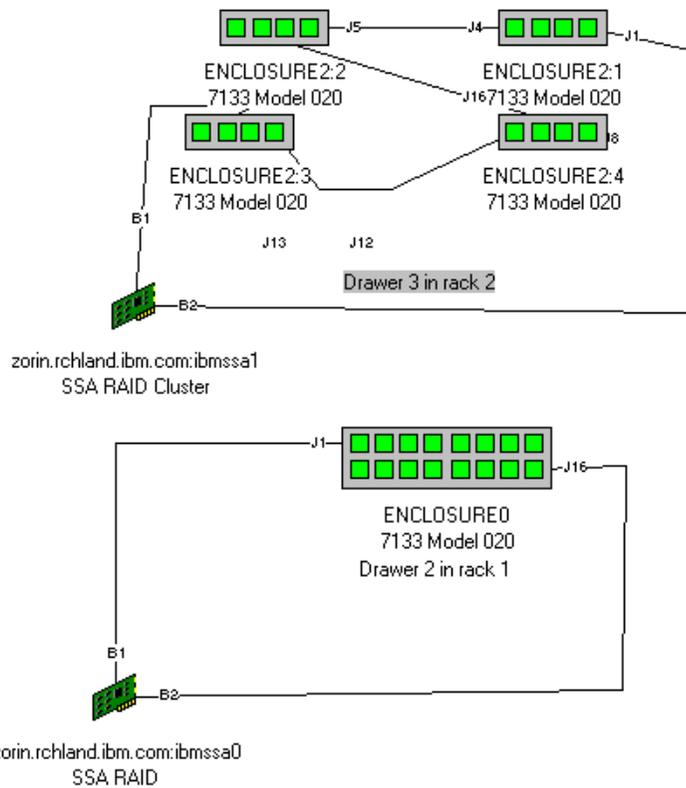


Figure 59. Customized StorX viewer with annotations

5.2 Event monitor

With StorX, there is an event monitor, which displays a list of every event log entry that have occurred while StorX has been running. StorX saves and restores the event log entries as part of the management set.

The event monitor is automatically started and displayed at the bottom of the screen when the live viewer is activated. If the monitor is disabled, it can be re-enabled by selecting **Event Monitor** from the Tools menu on the menu bar, or by clicking on the Event Monitor icon on the toolbar.

An example of the event log monitor is shown in Figure 60.

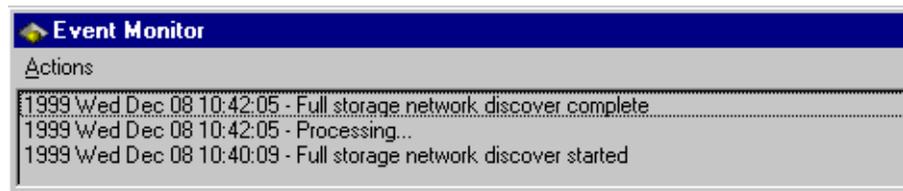


Figure 60. Event Monitor display

It is possible to find a device on the canvas that is associated with a specific event in the event monitor window. StorX will find and highlight the associated device on the canvas and position the canvas to the location of the device. To find a device

that is associated with a specific log entry, click on the event in the Event Monitor window to highlight it, and click on **Actions** on the event monitor menu bar. A drop down menu will appear, click on **Find** and the device associated with the event is highlighted. The result of this can be seen in Figure 61.

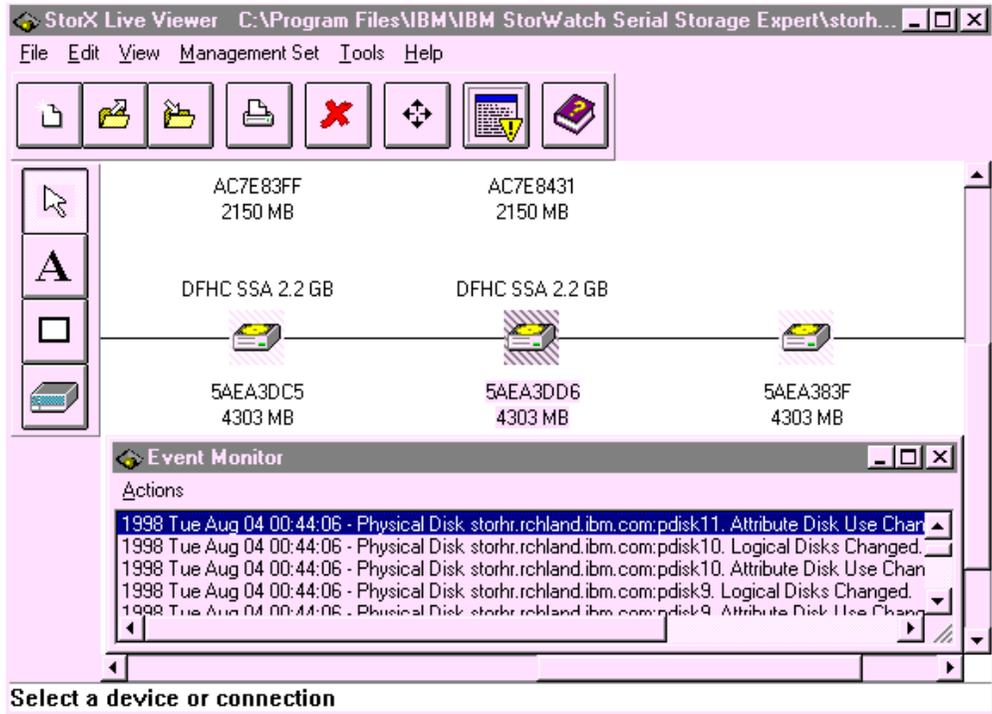


Figure 61. Event Monitor showing the associated device

5.3 Monitoring and managing a shared configuration

In this section, we use a configuration with eight SSA disks in one loop connected to two host systems, RS/6000-1 and RS/6000-2. When you manage your configuration from RS/6000-1, logical disks and pdisk names differ from those used when you manage your configuration from RS/6000-2, Figure 62 and Figure 63, respectively, show the configuration managing from RS/6000-1 and from RS/6000-2 (logical disk option).

Note: For all subsequent examples and diagrams, for ease of use, and explanation, we show all disk drives attached to the adapters. We will not use the 7133 enclosure icon.

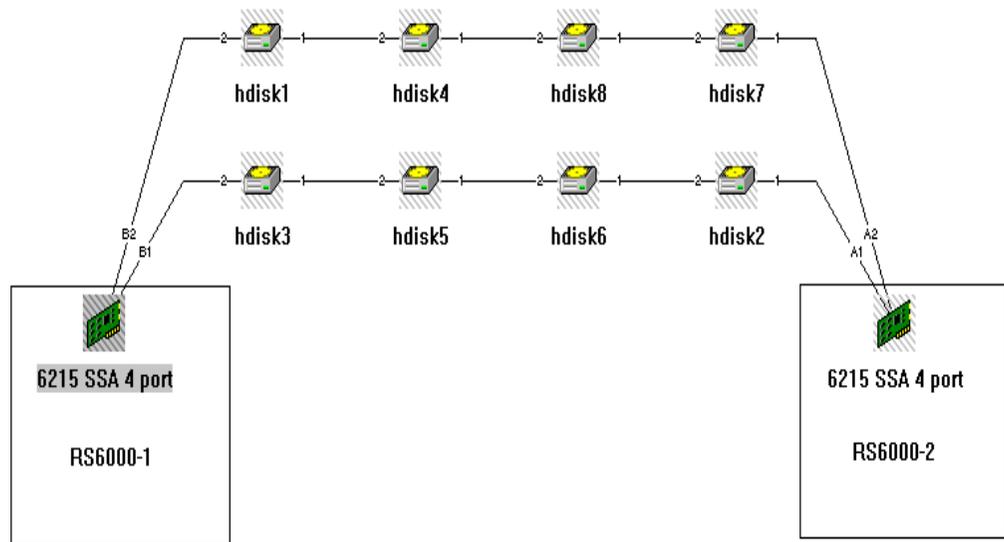


Figure 62. StorX live view from RS/6000-1: logical disk option

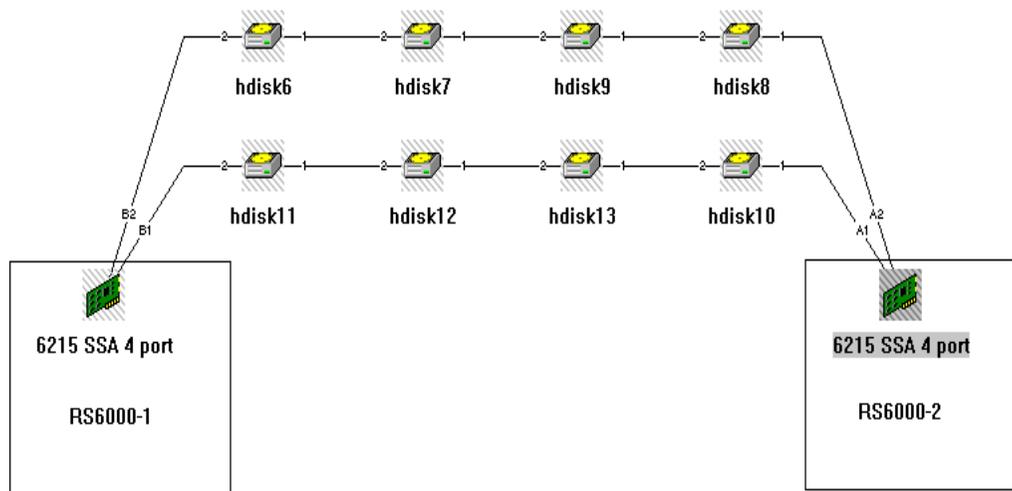


Figure 63. StorX live view from RS/6000-2: logical disk option

If you compare Figure 62 and Figure 63, you see that the names of the logical disks differ according to from which station you manage your configuration.

Figure 64 and Figure 65, respectively, show configuration managing from RS/6000-1 and from RS/6000-2 (pdisk name option).

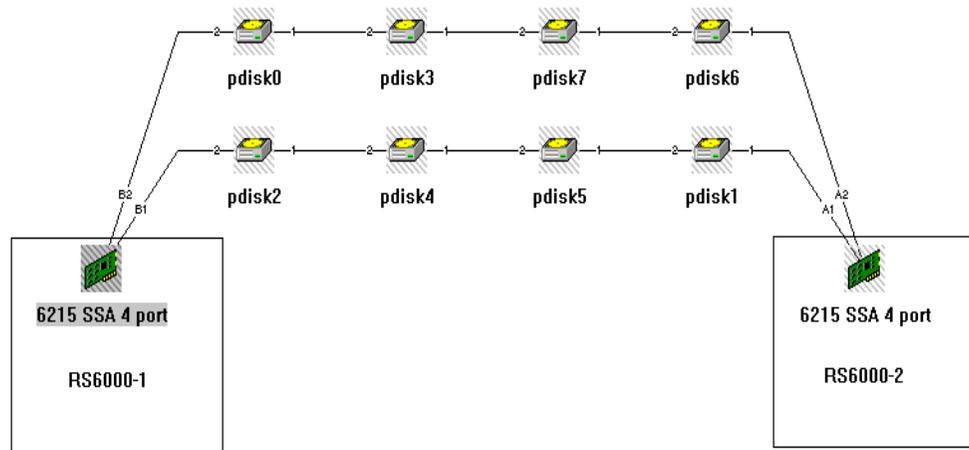


Figure 64. StorX live view from RS/6000-1, pdisk name option

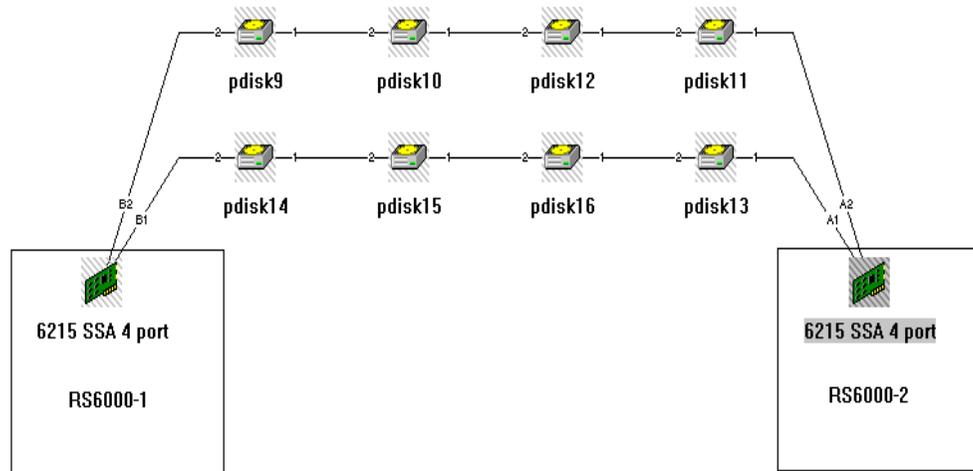


Figure 65. StorX live view from RS/6000-2, pdisk name option

Similarly, in Figure 64 and Figure 65, the pdisk names differ according to whether you manage your configuration from RS/6000-1 or RS/6000-2.

As you see, it is very easy to manage your configuration with StorX. You can also select other SSA view preferences as we have done previously, but in this case, the external labels, capacities, and types are the same for both systems.

5.4 Monitoring and managing a complex configuration

In this section, we explain why it may be beneficial to customize your SSA configuration, especially when it is a complex configuration.

A complex configuration with 32 SSA disks from two SSA bays, on one loop and shared by two systems is shown in Figure 66. It shows us a traditionally drawn configuration diagram.

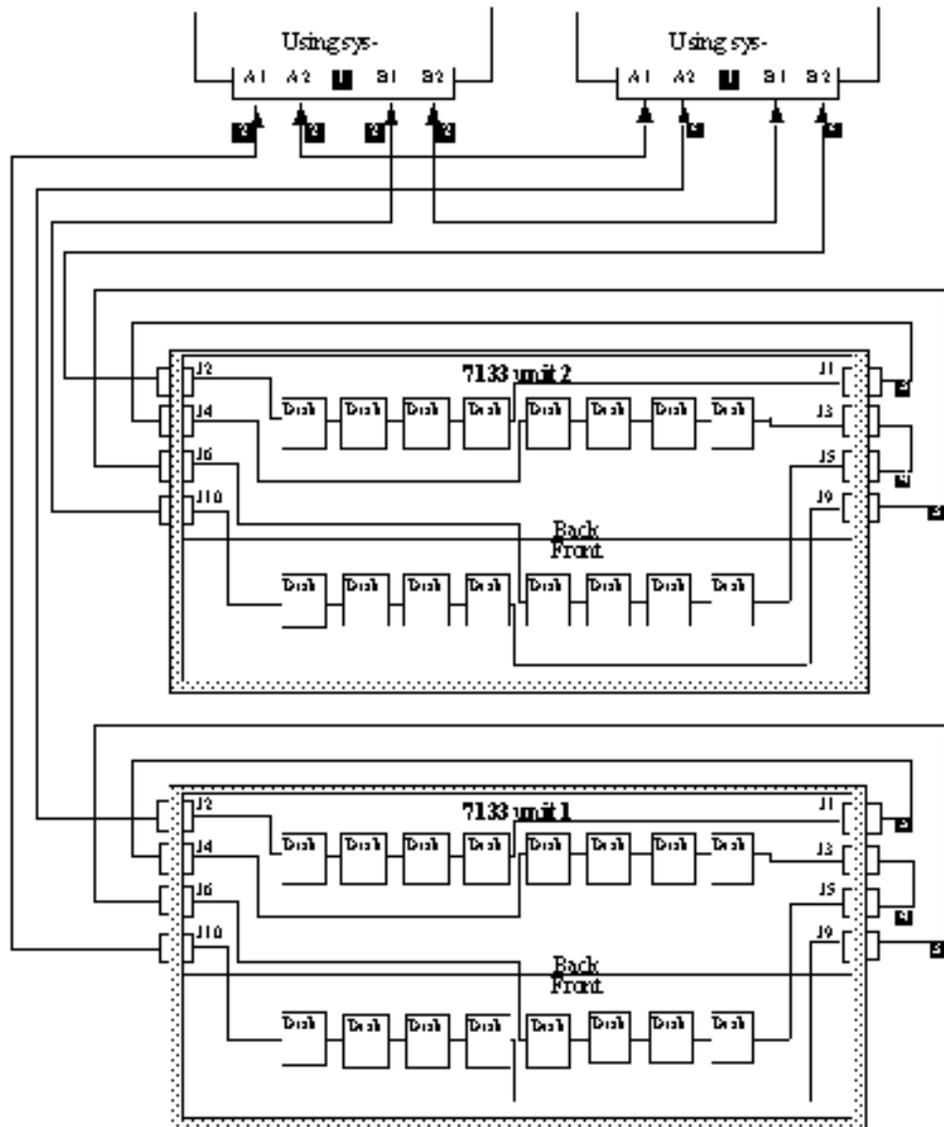


Figure 66. SSA planning without using StorX

Figure 67 shows us the uncustomized StorX live view for the same configuration.

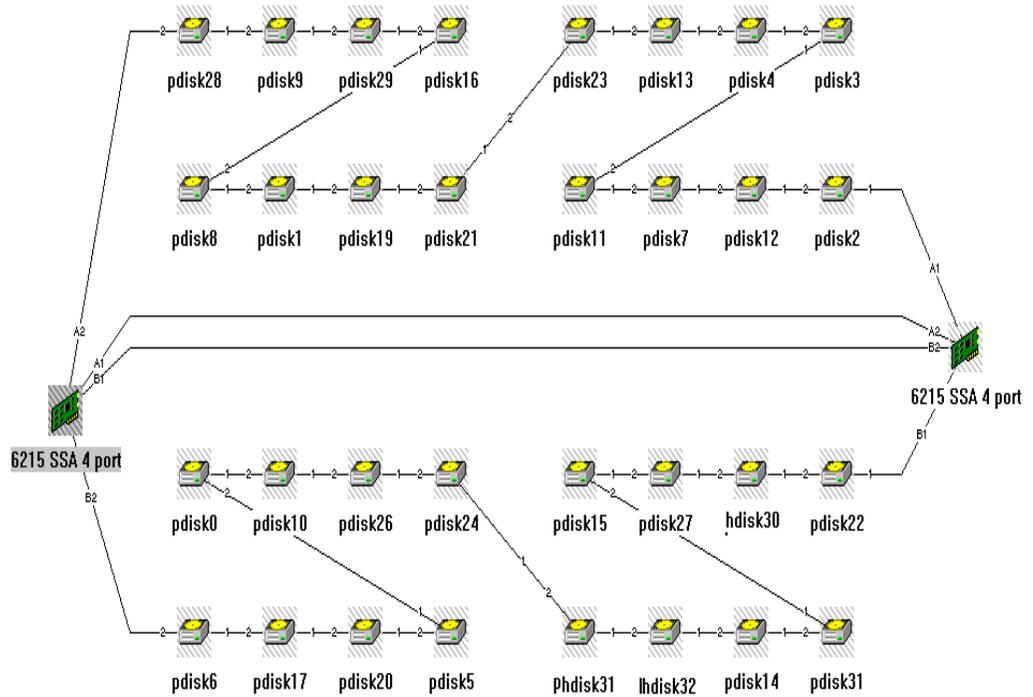


Figure 67. Live view of uncustomized StorX

Figure 68 shows us the same configuration, with customization.

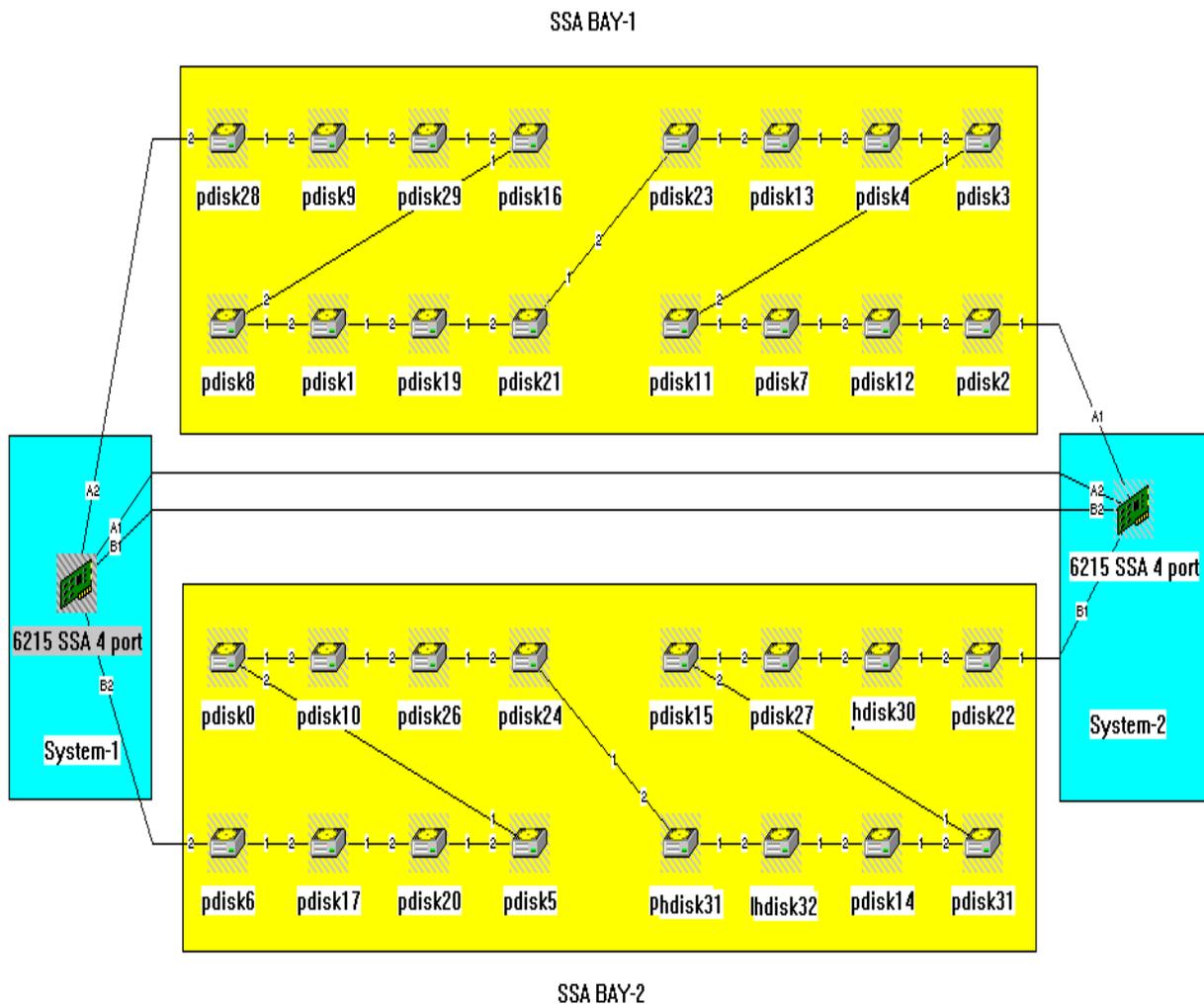


Figure 68. Live view of customized StorX from System-1, pdisk name option

There are many ways to customize a StorX live view, and it is very easy to do. With full drag and drop capabilities, it is easy to show the configuration in a manner that suits you. The important point is that if you manage many SSA configurations, you may want to customize the live view to facilitate your exploitation.

5.5 Using StorX for subsystem troubleshooting

In this section, we discuss SSA subsystem troubleshooting. We show what happens, and what recovery steps are required, when:

- There is a disk failure
- There is an SSA open loop
- There is an adapter failure

Figure 69 and Figure 70, respectively, show us one configuration with two loops connected to one SSA adapter, with a logical disk option and with a pdisk option.

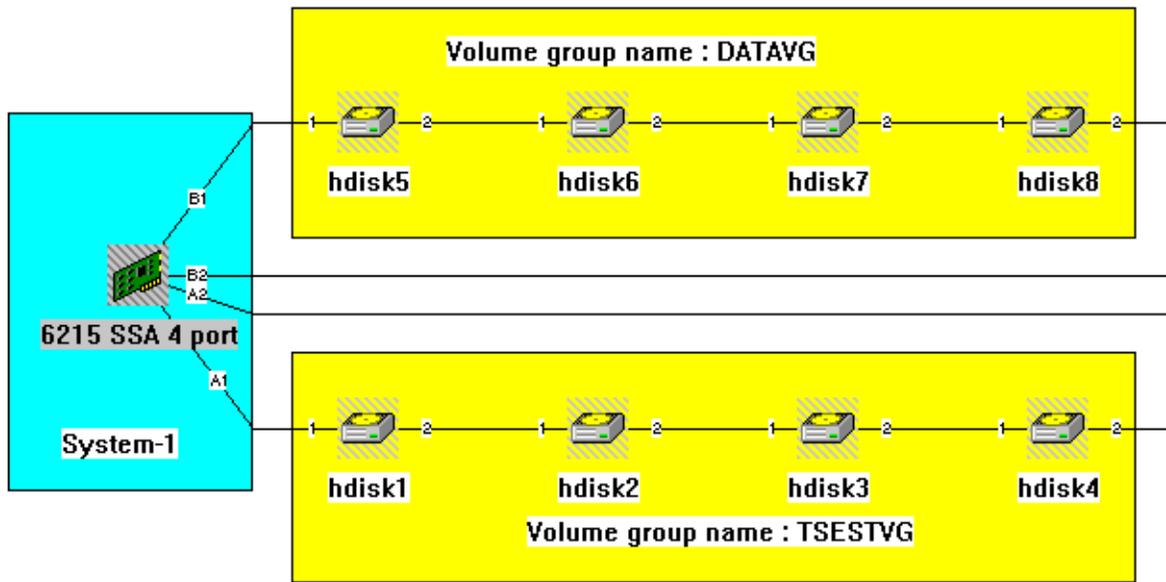


Figure 69. Initial configuration, logical disk name option

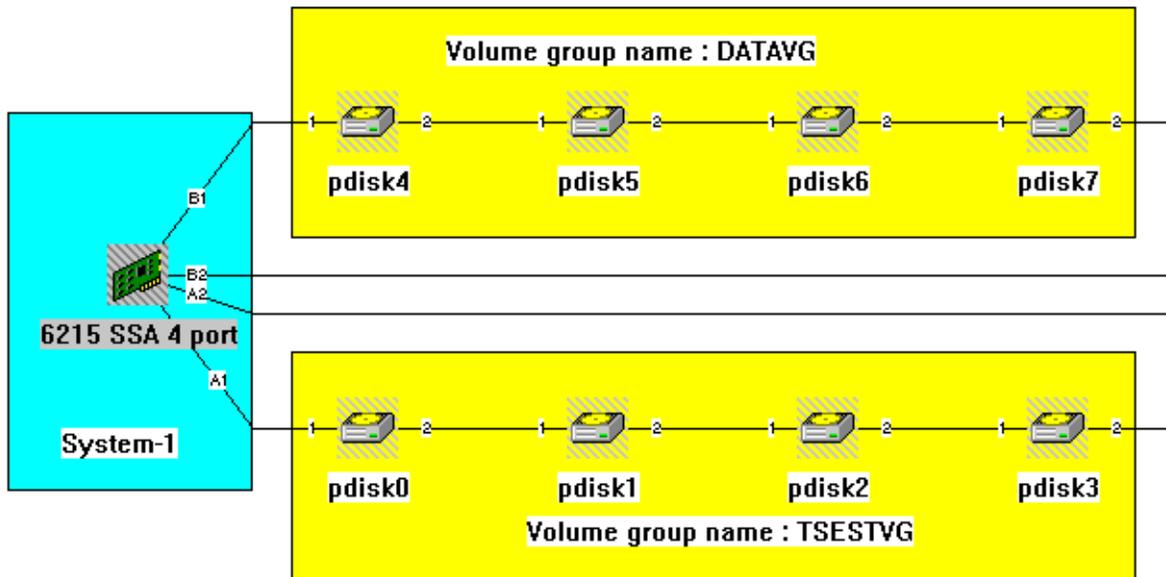


Figure 70. Initial configuration, physical disk name option

5.6 Monitoring and managing an hdisk failure

In our example, a problem occurs on hdisk 7.

The first indication is that the live viewer pops up a window indicating that a change has occurred and a refresh is required to see the event. Figure 71 shows the result after the refresh.

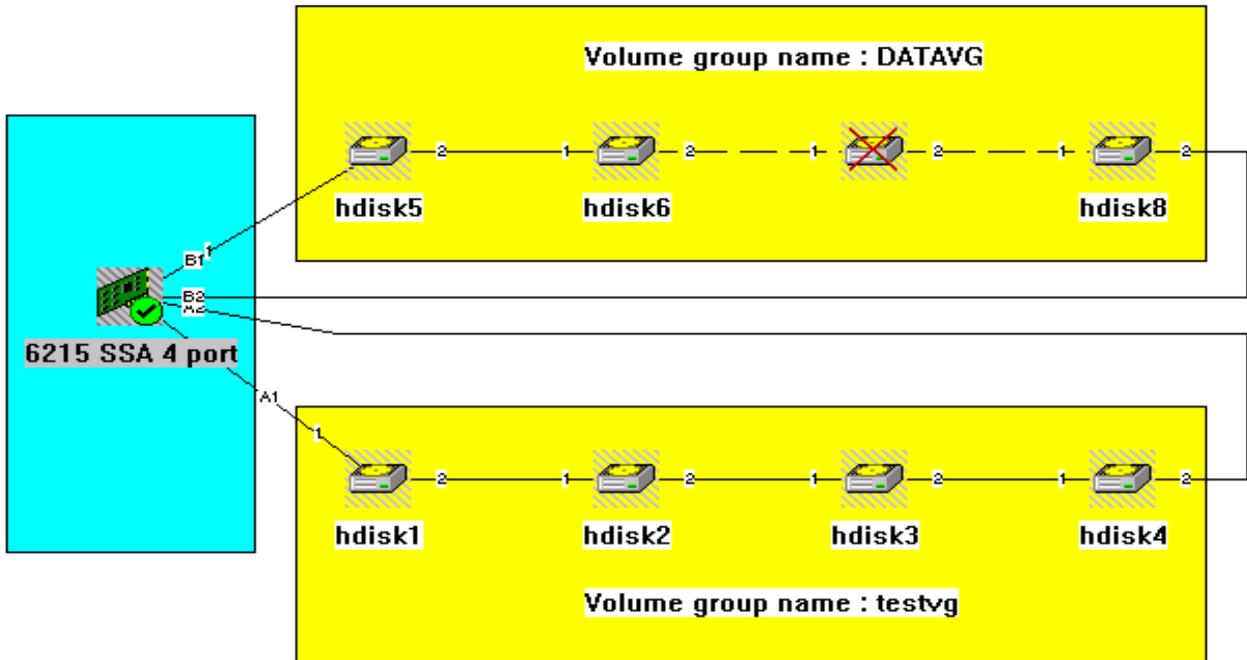


Figure 71. StorX live view of an hdisk problem

You can see that StorX detected the problem immediately, and indicates it by a cross on the defective hdisk (hdisk 7 and its attached pdisk 6), and also by a flag to the SSA adapter. This flag means that a change has occurred in this configuration.

5.6.1 Recovery procedure (hdisk failure)

In this section, we see what actions are required and the StorX view as the problem is resolved. Figure 72 shows us the StorX view after the problem disk drive is replaced with a new one.

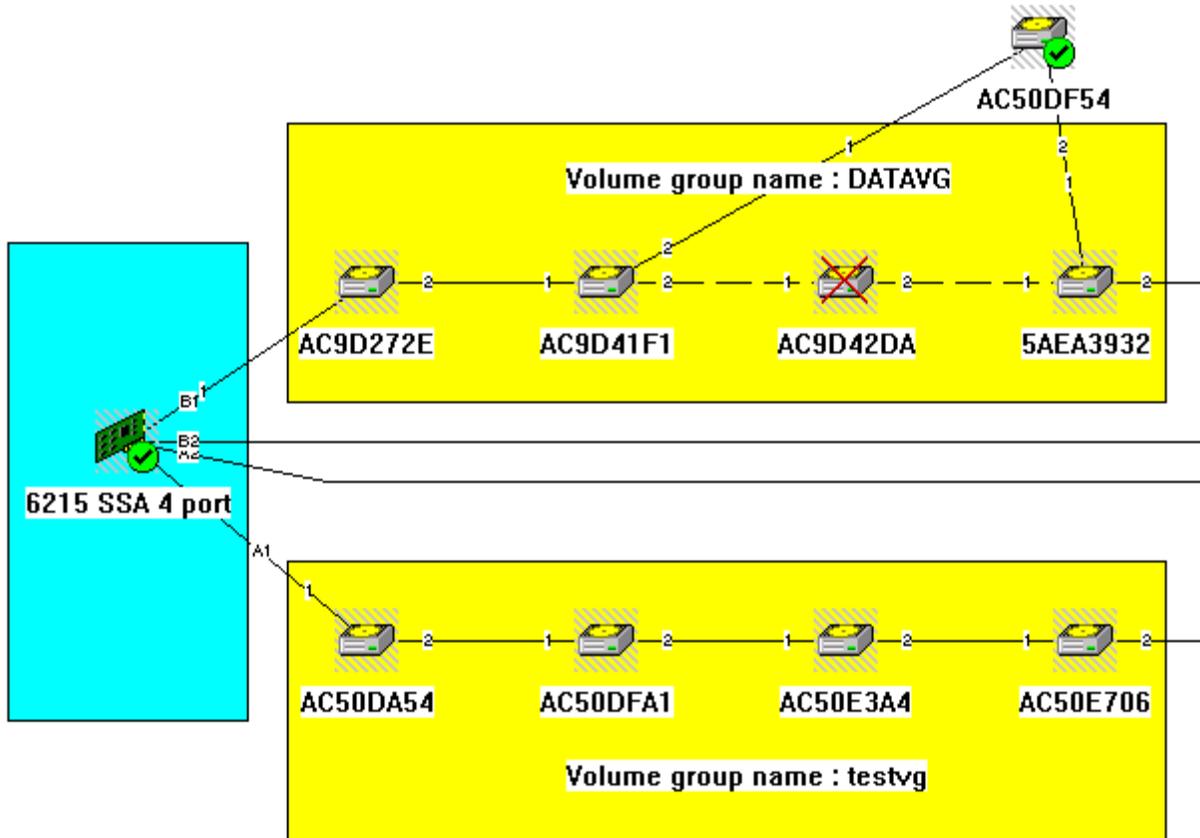


Figure 72. hdisk replacement StorX live view

We see that the bad disk still shows its cross, indicating that it is out of order. There are flags on the SSA adapter and on the new SSA disk to notify us that a change has occurred in our configuration.

To complete the hardware replacement procedure, we must inform the operating system. We, therefore, have to type some AIX commands. The sequence is:

1. `reducevg datavg hdisk7` — to remove hdisk 7 from DATAVG
2. `rmdev -l pdisk6 -d` and `rmdev -l hdisk7 -d` — to remove pdisk 6 and hdisk 7 of the ODM.
3. `cfgmgr` — to configure the new disk
4. Select **Clear change** on the StorX management tool to redisplay the live viewer

Figure 73 shows us the result of the clear change command.

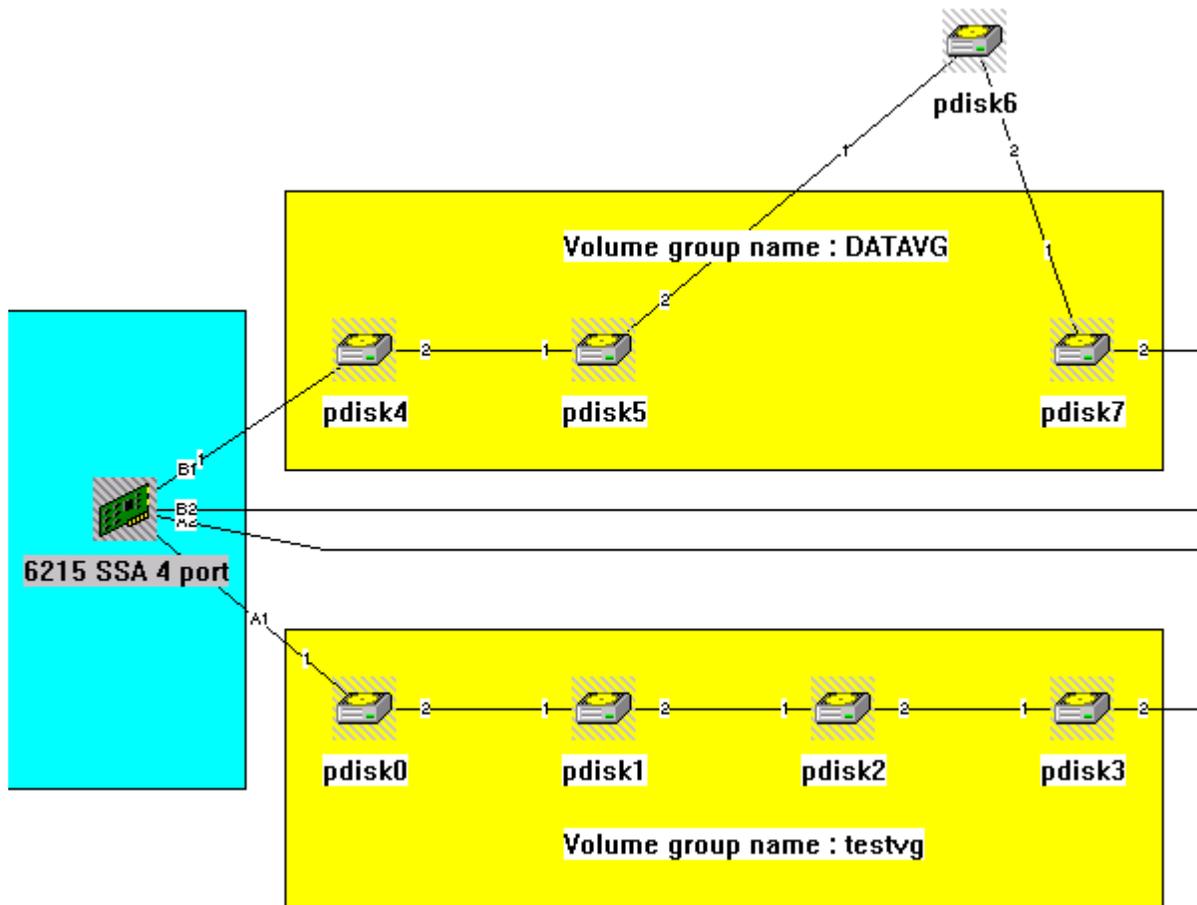


Figure 73. Configuration of the new SSA disk, pdisk name option

The flags have been removed, indicating that the hardware procedure is complete.

hdisk 7 and its pdisk attached (pdisk 6) has been replaced and recognized by the system. We must now complete the software procedure, by integrating hdisk 7 in the DATAVG volume group. This is done by typing the AIX command:

```
extendvg datavg hdisk7
```

The software procedure is now complete. To clean up, we have to drag and drop external hdisk 7 in DATAVG volume group back in line with the other disk drives. Figure 74 shows us the final display.

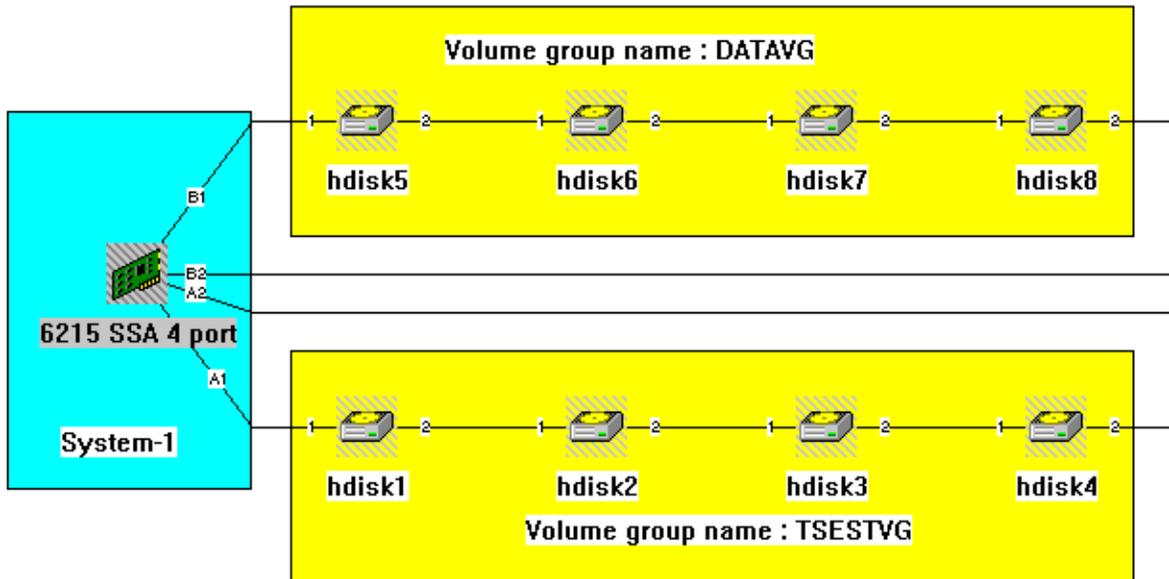


Figure 74. Recovery procedure completed

5.7 Monitoring and managing an SSA open loop problem

In this section, we show what to do in the event of an SSA open loop. Figure 75 shows the initial configuration of two SSA loops connected to one SSA adapter, and we give a recovery procedure.

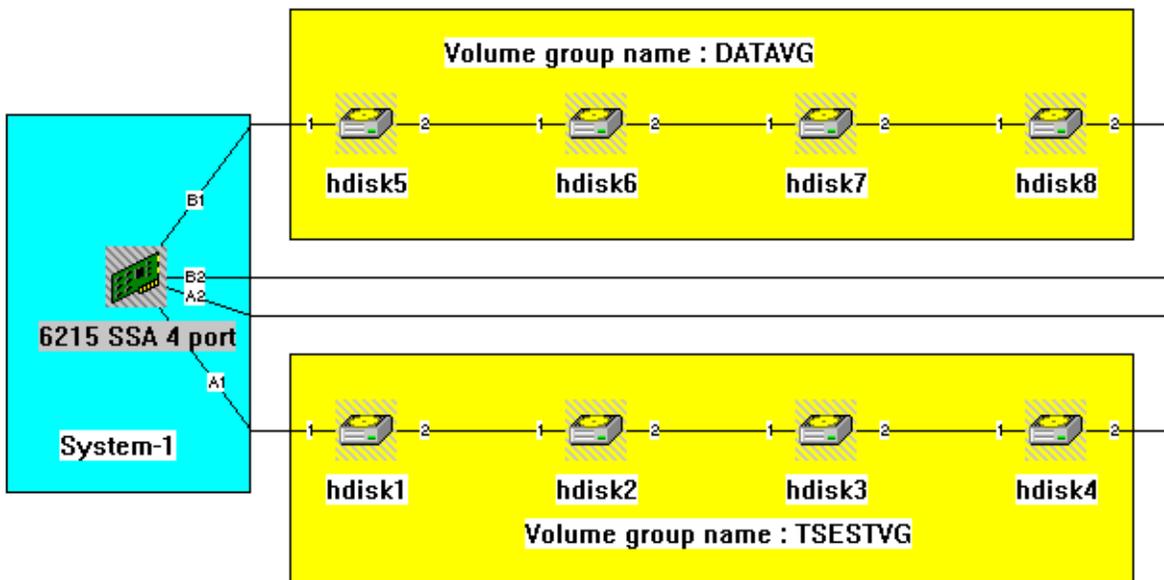


Figure 75. Initial configuration of two SSA loops connected to one SSA adapter

Figure 76 shows the results of an SSA open loop between Ports B1 and B2.

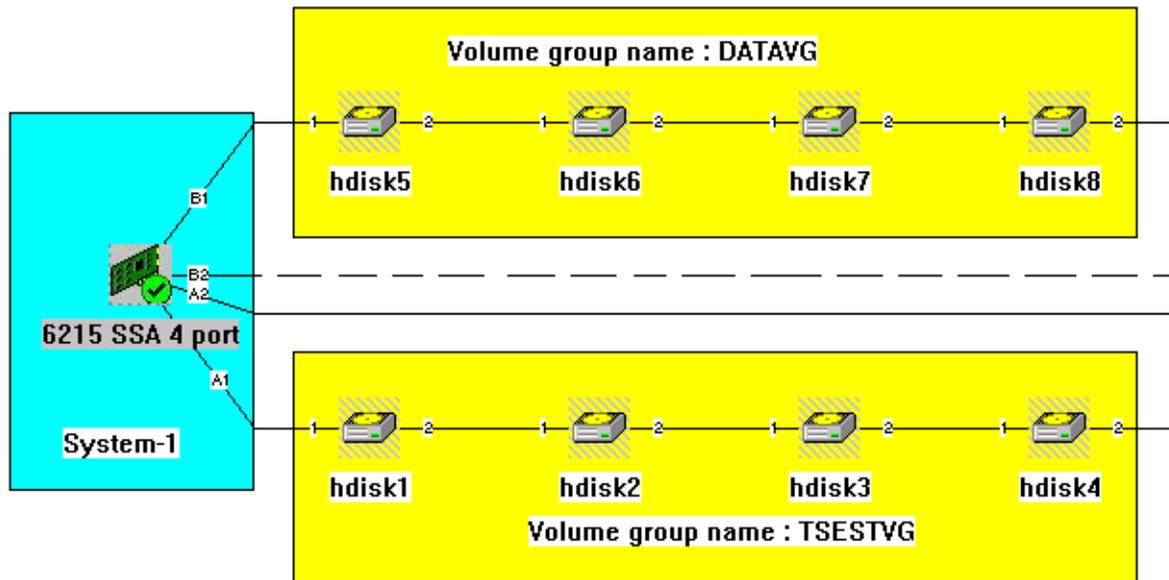


Figure 76. StorX live view of an SSA open loop

We see that StorX detects the problem, and indicates the failing connection port or cable. There is also a flag on the adapter to notify us that a change has occurred in the configuration.

Recovery procedure (SSA open loop problem)

After replacing the defective cable and issuing the clear changes command on StorX, the replacement procedure is complete. We are then back to the initial configuration, which is as shown in Figure 75.

5.8 Monitoring and managing an SSA adapter failure

Because of the flexibility of SSA, and the configuration options available, an SSA adapter failure can be explained through different scenarios. In this section, we show what happens when a failure occurs for a single adapter loop and also what happens for an adapter failure when multiple adapters are on the same loop.

5.8.1 Two SSA loops connected to one SSA adapter card

Figure 77 shows our first configuration of one adapter with two loops.

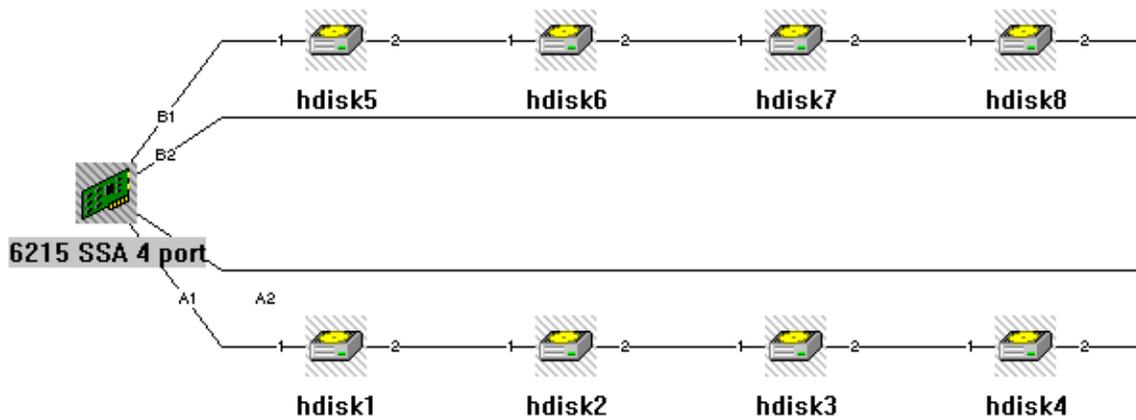


Figure 77. Initial configuration with two loops — StorX live view

An adapter failure was simulated, and Figure 78 shows the resultant StorX display.

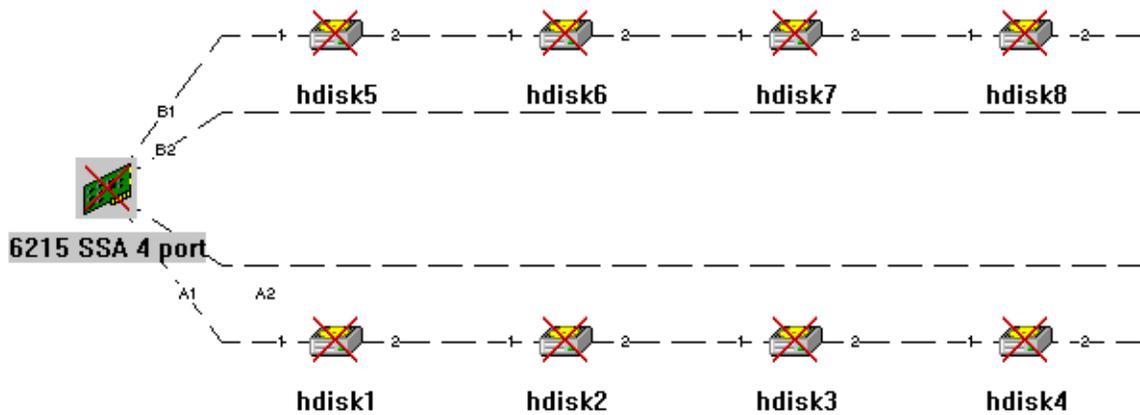


Figure 78. SSA adapter problem affecting two loops — StorX live view

As you see, all the devices have a cross on them to indicate that all this configuration is completely out of order. Access to data has been lost. The resolution of this problem is to replace the broken adapter, or replug the SSA loop into a spare SSA adapter card in the host.

5.8.2 One SSA loop shared by two systems

Figure 79 shows our second configuration, which has one SSA loop shared by SSA adapters from different hosts.

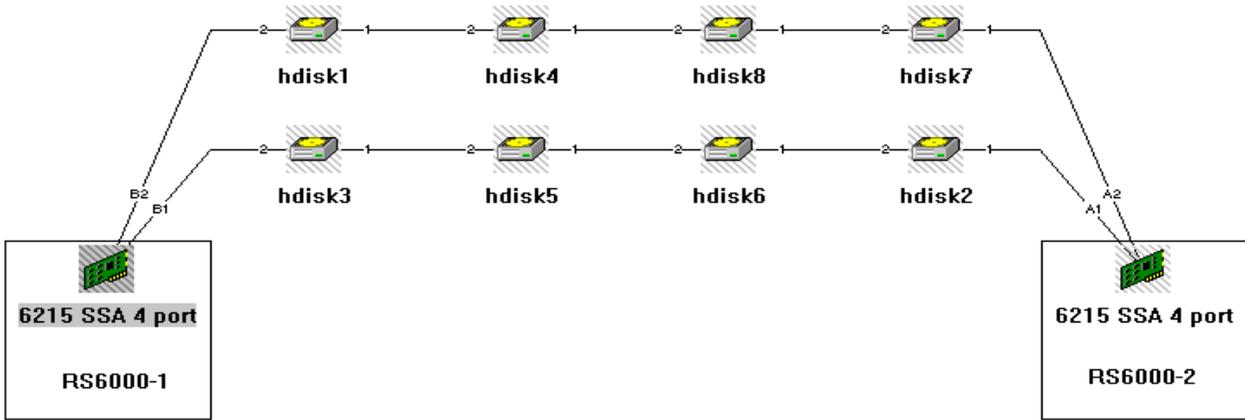


Figure 79. Initial configuration with one loop and two adapters — StorX live view

Again, an adapter failure was simulated, and Figure 80 shows the resultant StorX display.

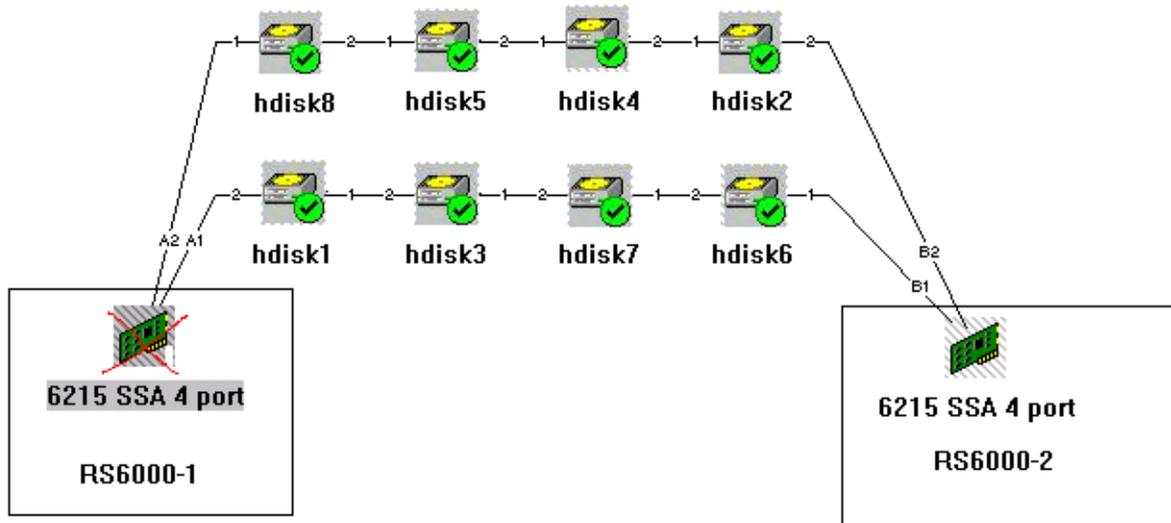


Figure 80. SSA adapter problem on System1 — StorX live view

We see that StorX detects the problem and shows the broken SSA adapter card. It also indicates, by a flag on each hdisk, that changes have occurred in the configuration, but the hdisks are not broken. Data on those drives can still be accessed by RS/6000-2. To resolve this problem, we replace the bad adapter, issue the `cfgmgr AIX` command, and clear changes on StorX live view.

5.9 Monitoring and managing an SSA RAID array creation

In this section, we explain the different steps required to create an SSA RAID array, and how StorX manages and monitors this kind of creation. We keep the two loops connected to one SSA adapter configuration as shown in Figure 81.

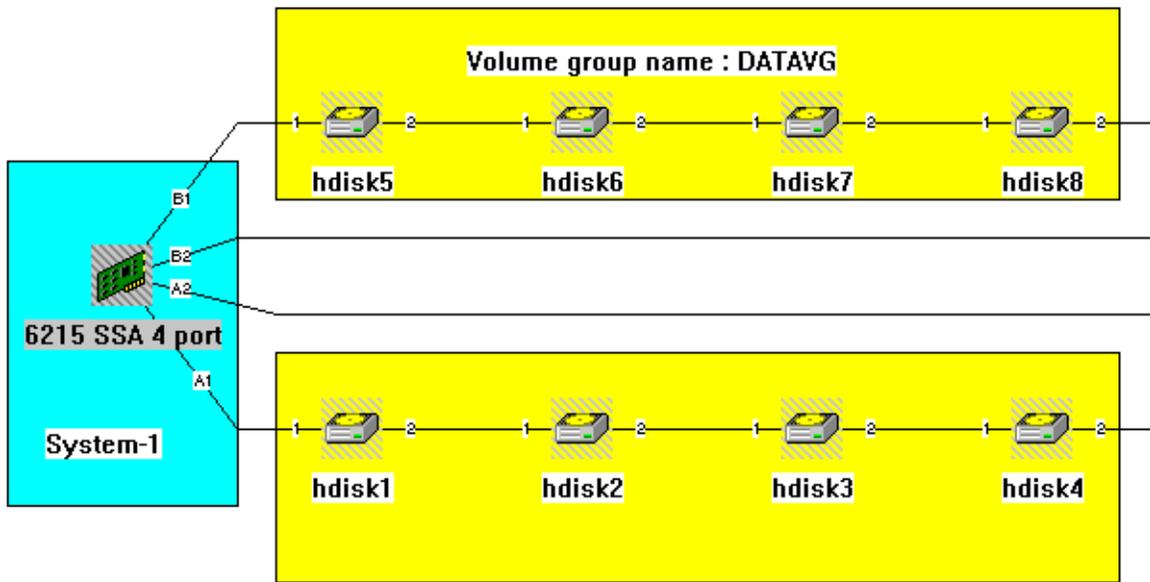


Figure 81. Initial configuration before SSA RAID array creation — StorX live view

In Figure 81 you can see that hdisk 1, hdisk 2, hdisk 3, and hdisk 4 are not included in any volume group.

Step1: hot spare disk definition

In this step, we discuss the process to define a hot spare disk, and we define the SSA RAID array hot spare disk. The AIX command to define a hot spare disk is:

```
chgssadisk_system_command_to_exec -l'ssa2' -n 'pdisk0' '-a use=spare -u'
```

StorX immediately detects that a change has occurred, and shows this change by two device state icons: one on hdisk 1 (hot spare disk), and the second on the SSA adapter. Figure 82 shows these changes.

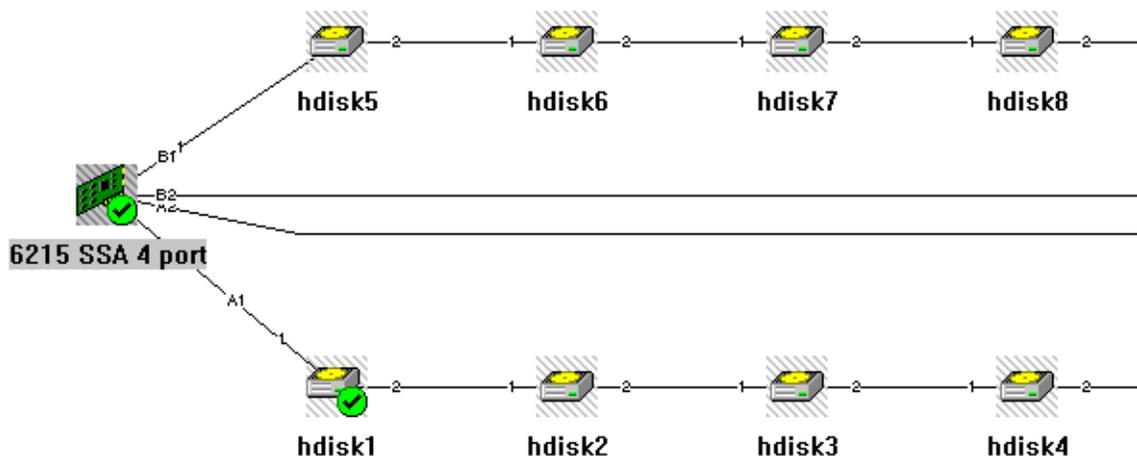


Figure 82. Hot spare definition — StorX live view

Step 2: array candidate disk definitions

The second step is to define the array candidate disks. For this example, hdisk 2 (pdisk 1), hdisk 3 (pdisk 2), and hdisk 4 (pdisk 3) are the array candidate disks. Again, this task is performed under SMIT. The AIX command to define the array candidate disks is:

```
chgssadisk_hr_cmd_to_exec -l 'ssa2' 'pdisk1 pdisk2 pdisk3' '-a use=free -u'
```

Figure 83 shows the StorX changes after the command is executed.

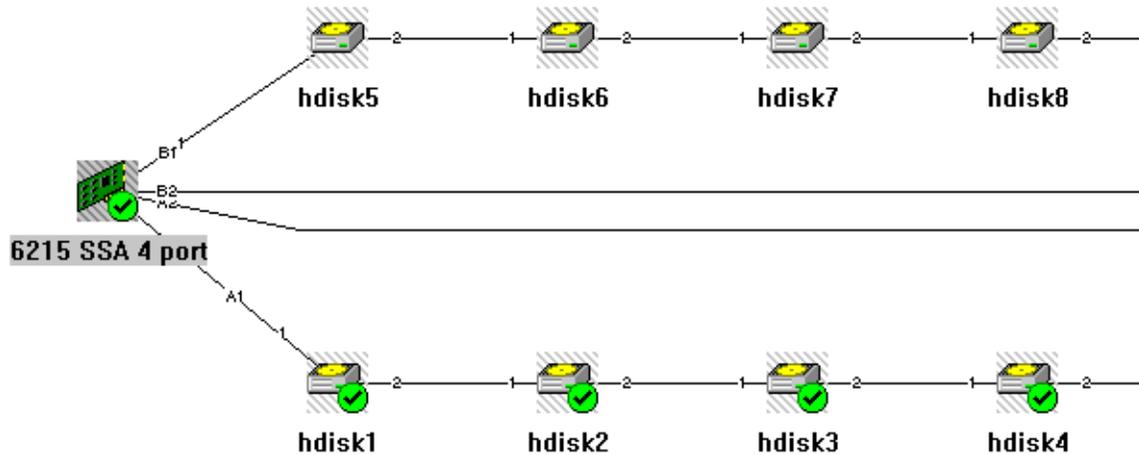


Figure 83. Hot spare and array candidate disks — StorX live view

StorX changes the device state icons on hdisk 2, hdisk 3, and hdisk 4, which are the array candidates.

Step 3: SSA RAID array creation

In the third step we discuss how to create the SSA RAID array. This task also is performed by SMIT. The command is:

```
mkssaraid_raid_5_cmd_to_exec -l 'ssa2' 'raid_5' 'pdisk1 pdisk2 pdisk3' '-a spare=true  
-a spare_exact=false' -a allow_page_splits='true'
```

Figure 84 shows the new StorX live view generated by the SSA RAID Array creation.

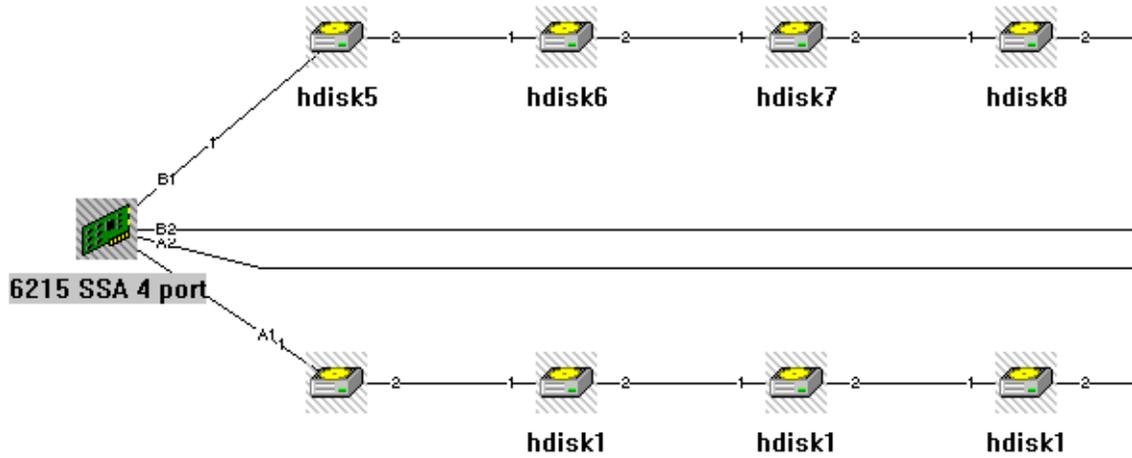


Figure 84. SSA RAID array — StorX live view

You can see that the SSA RAID array contains three disks named hdisk 1, and one hot spare disk (the disk without a name). To facilitate management of the SSA RAID array, we used pdisk names, as shown in Figure 85.

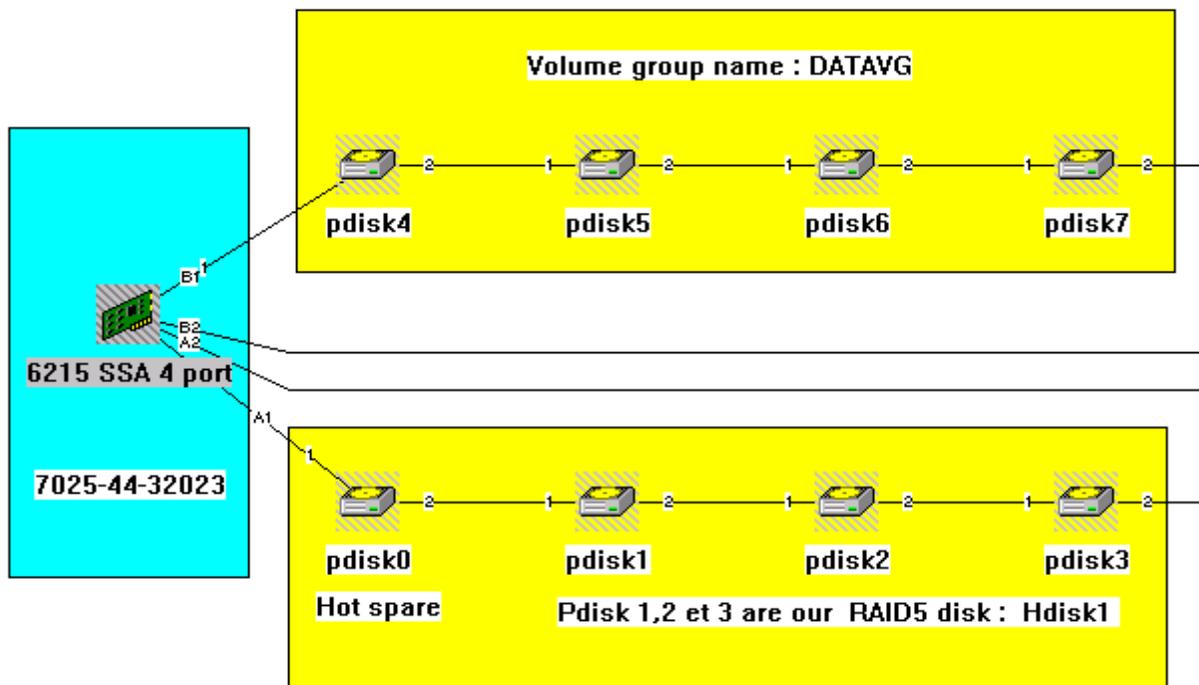


Figure 85. SSA RAID array, pdisk option — StorX live view

Step 4: RAID5VG volume group creation

After creating an SSA array disk, you have to include it in a volume group. We used this SMIT AIX command:

```
VGNAME=mkvg -f 'raid5vg' -s '32' hdisk1
```

We have named this volume group RAID5VG, as shown in Figure 86.

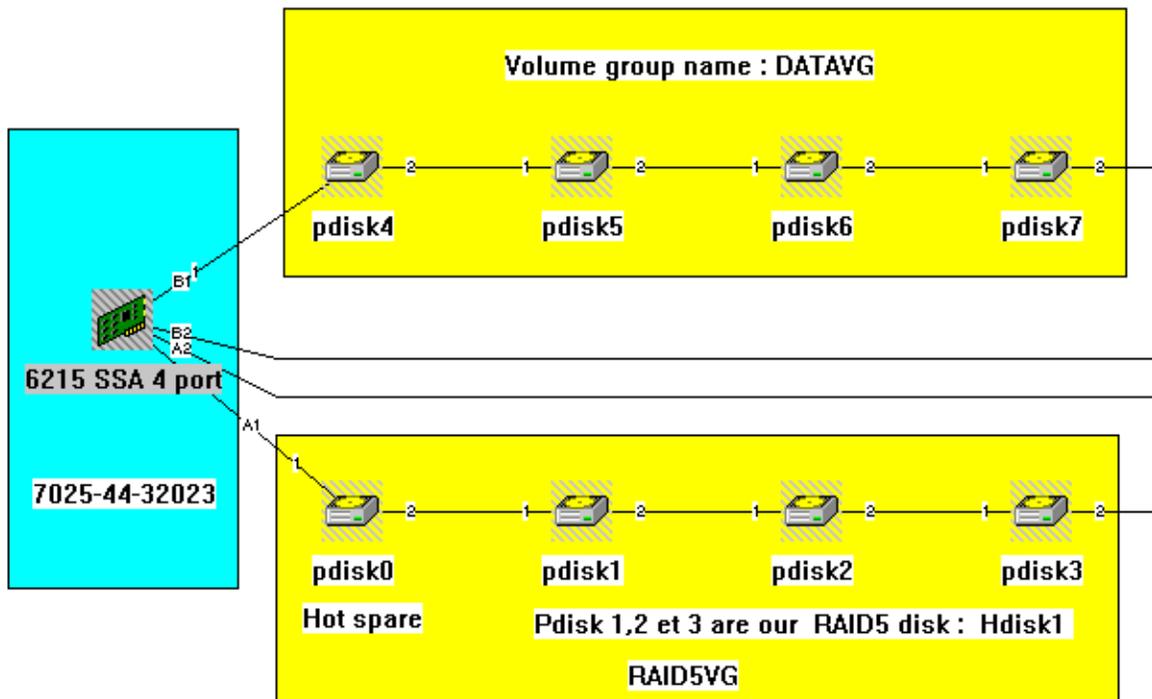


Figure 86. Volume group RAID5VG — StorX live view

5.10 Monitoring and managing disk problems in a RAID 5 configuration

In this section, we discuss how to handle SSA disk problems in a RAID 5 configuration. These are the steps, using StorX, for the configuration just created (Figure 86).

5.10.1 Crash of pdisk 1

In this step, we show what happens when we lose a disk that was part of the SSA RAID array. We selected pdisk 1 as our problem device. Figure 87 shows the StorX live view of the broken disk.

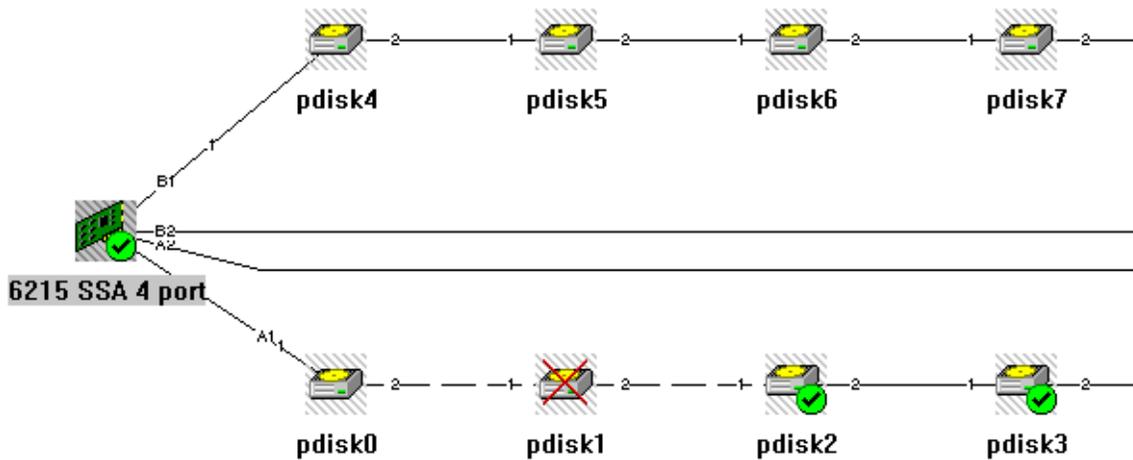


Figure 87. Crash of pdisk 1 in an SSA RAID array

You can see that StorX detects the pdisk crash, and shows us by a red cross on pdisk 1. The device state icons on the SSA adapter, and on pdisks 2 and 3, which are included in SSA RAID 5 hdisk 1 are also changed. These flags indicate that a storage change has occurred on this volume group, but the data is still accessible.

5.10.2 Recovery procedure (RAID 5 configuration)

The first step in the recovery procedure is to replace pdisk 1, and then initialize the new SSA disk by the AIX command: `cfgmgr`. The result of this is shown in Figure 88.

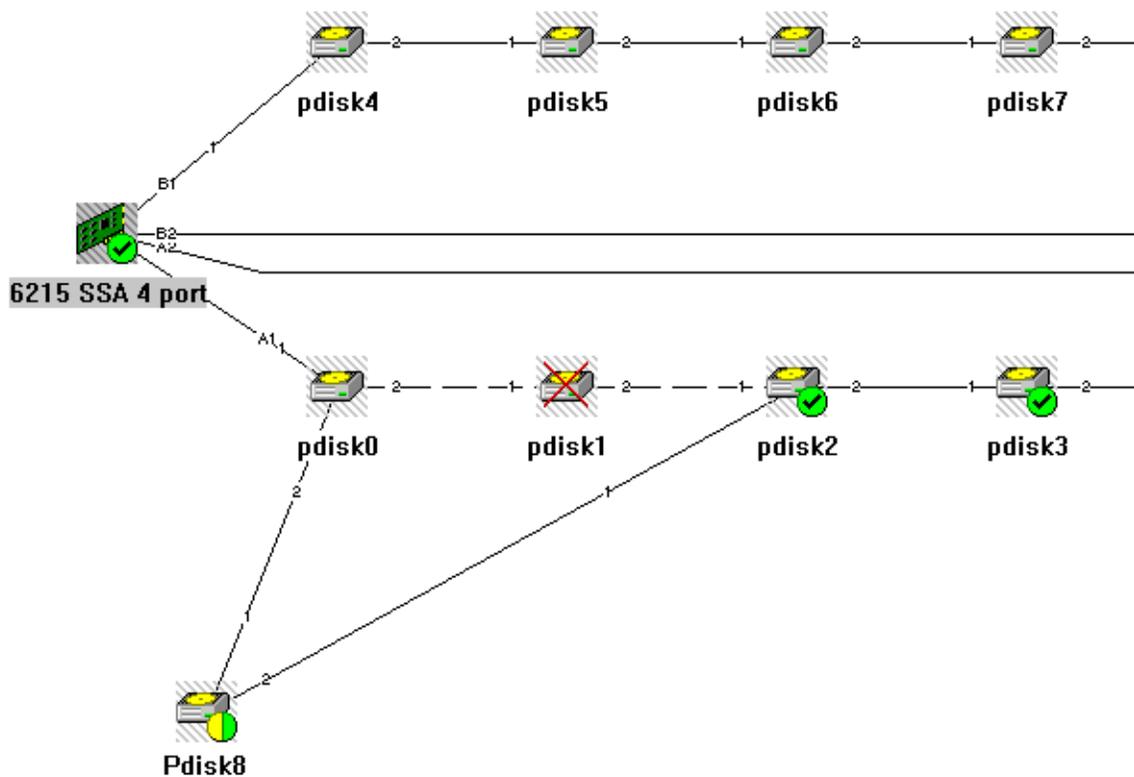


Figure 88. Replacement of pdisk 1 — StorX live view

The new disk, pdisk 8, is recognized by the system but not yet defined in the RAID 5 environment. At this time, pdisk 0, which was the hot spare disk, becomes a member of the SSA RAID 5 array, so currently there is no hot spare disk.

5.10.3 New hot spare disk definition

The second step is to define a new hot spare disk (pdisk 8) for the SSA RAID 5 environment. We could use SMIT or the AIX command:

```
chgssadisk_system_command_to_exec -l 'ssa2' -n 'pdisk8' '-a use=spare -u'
```

On StorX, we used clear change. The result of these two commands is shown in Figure 89.

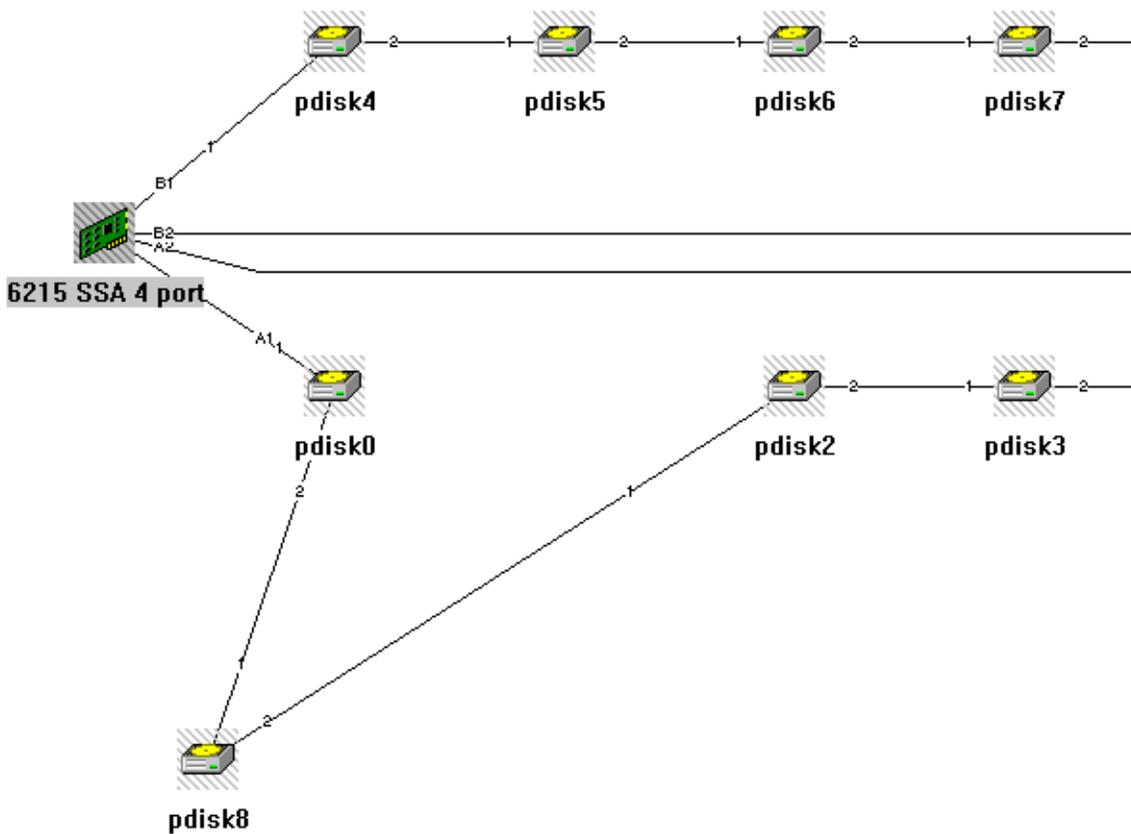


Figure 89. Definition of pdisk 8 — StorX live view

You can see that there are no flags anywhere, and pdisk 8 has been defined as the hot spare disk.

5.10.4 Completed recovery procedure

The recovery procedure is now complete, so all that is left to do is to drag and drop pdisk 8 to its correct place. Figure 90 shows us the new SSA RAID 5 environment from the pdisk perspective.

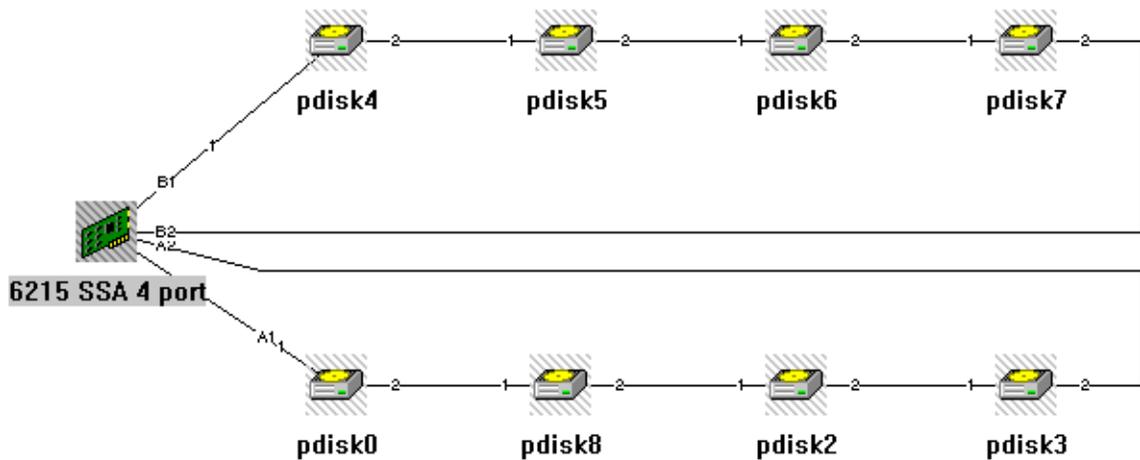


Figure 90. New SSA RAID 5 environment, pdisk option

You can see that pdisk 8 has replaced pdisk 1. Figure 91 shows us the same thing from the hdisk (customized) perspective.

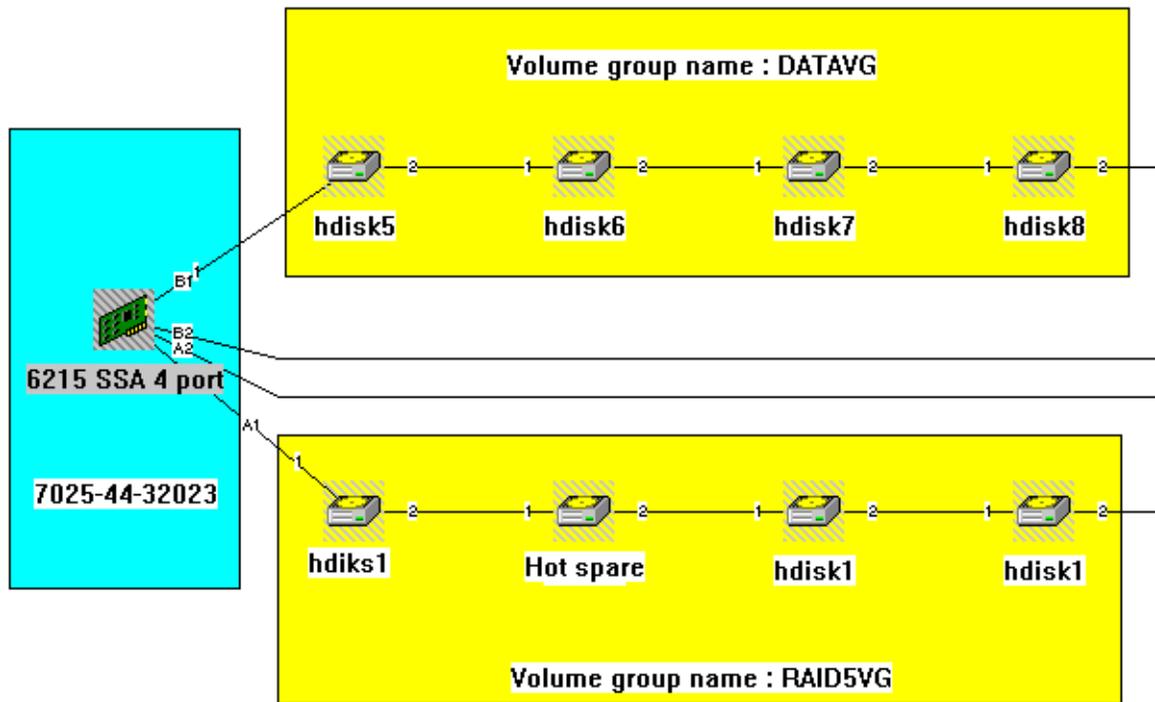


Figure 91. New SSA RAID5 environment, logical disk option

You can see that members of RAID5VG have changed, and that a new hot spare disk has been created. Compare this with Figure 86 on page 109, which represents this configuration before the pdisk 1 crash.

5.11 Monitoring, managing, and diagnosing with StorX

In this section, we show you how to use StorX to manage your SSA environment. With StorX, you can run online diagnostics, view vital product data (VPD), identify whether a disk is an AIX system disk or a hot spare disk, and determine whether disks are members of an SSA RAID array.

We chose the configuration of two loops connected to one SSA adapter, because it is a good example of the difference in the information required when you select the adapter or different disks. Figure 92 shows the configuration.

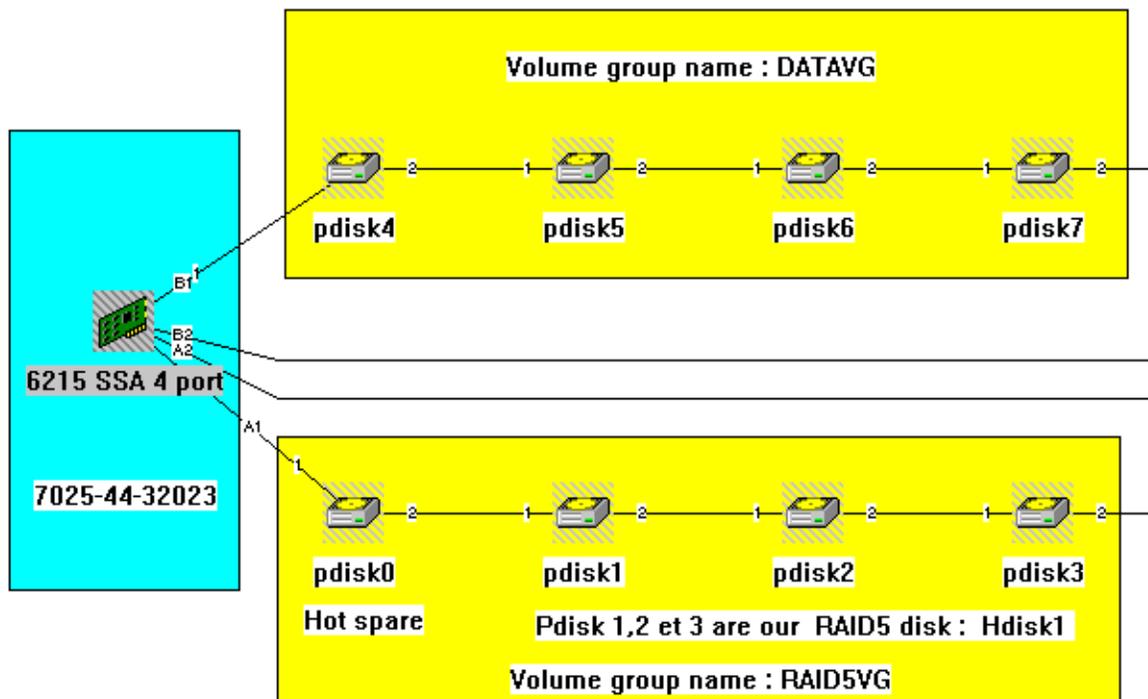


Figure 92. Two loops connected to an SSA adapter with RAID5 array as one loop

5.11.1 Disk properties window

To get to the disk properties window (as shown in Figure 93), double-click on any pdisk or hdisk, or click once with the right mouse button and select **Properties**. For example, we have double-clicked on pdisk 6. Figure 93 through to Figure 97 on page 118 show us the different properties provided by the Disk Properties window.

Figure 93 shows the Disk Properties window.

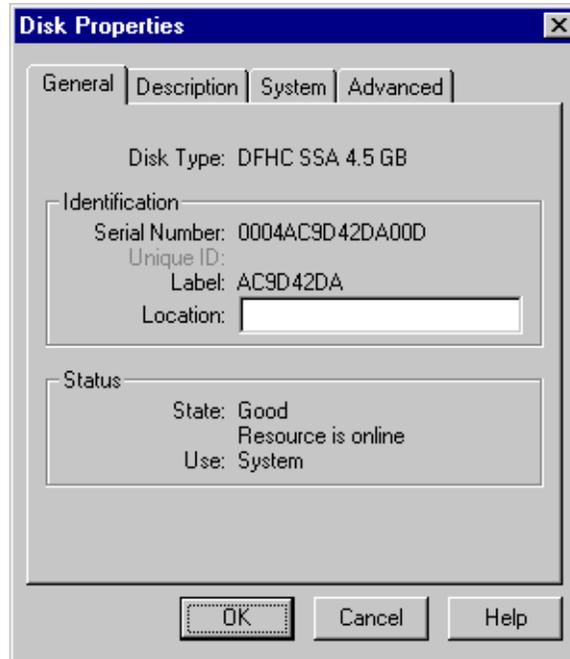


Figure 93. General page of Disk Properties window

There are four different divisions (pages) in the properties window.

- General overview of the selected device
- Description of the selected device
- System information
- Advanced attributes which include identifying and diagnosing the device

We can see that the general page of the disk properties window shows:

- SSA hdisk type
- Serial number
- External label
- Location (you have to describe this location)
- State
- Current use (in our example, pdisk 6 is an AIX system disk)

Figure 94 shows us the Description page of the Disk Properties window.

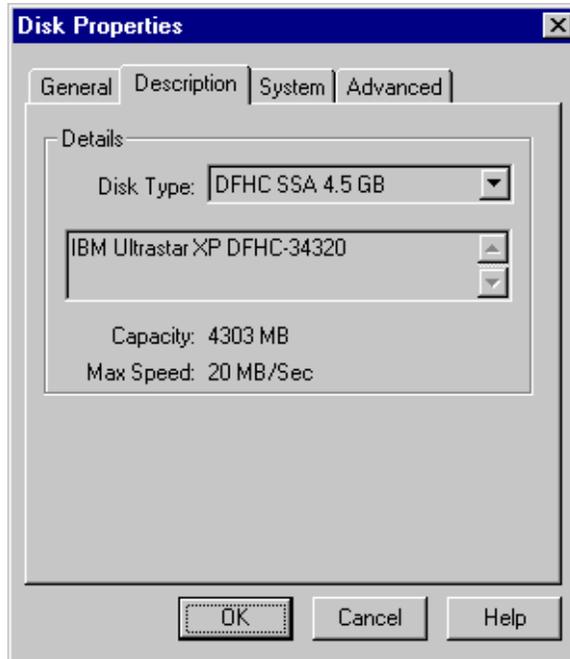


Figure 94. Description page of Disk Properties windows

The Description page of the Disk Properties window indicates:

- Disk type
- IBM product name
- Capacity
- Maximum speed

Figure 95 shows us the System page of the Disk Properties window.

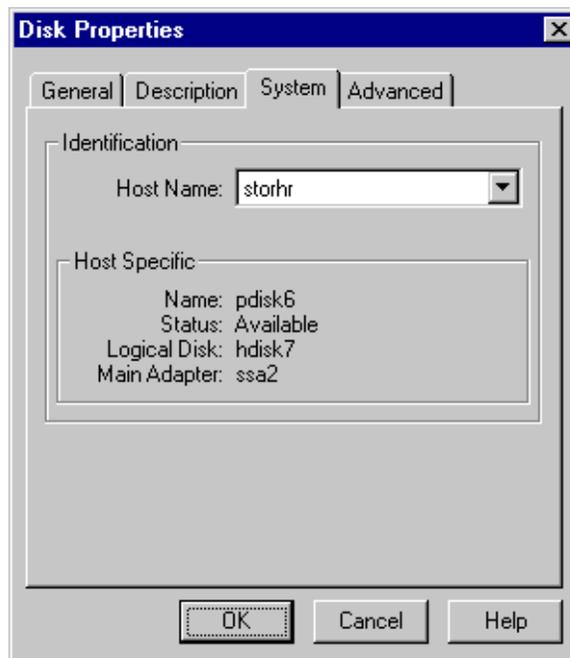


Figure 95. System page of Disk Properties window

The System page of the Disk Properties window indicates:

- System name
- Pdisk name
- Logical disk name
- Main SSA adapter that manage hdisk 7

Figure 96 shows us the Advance page of the Disk Properties window.

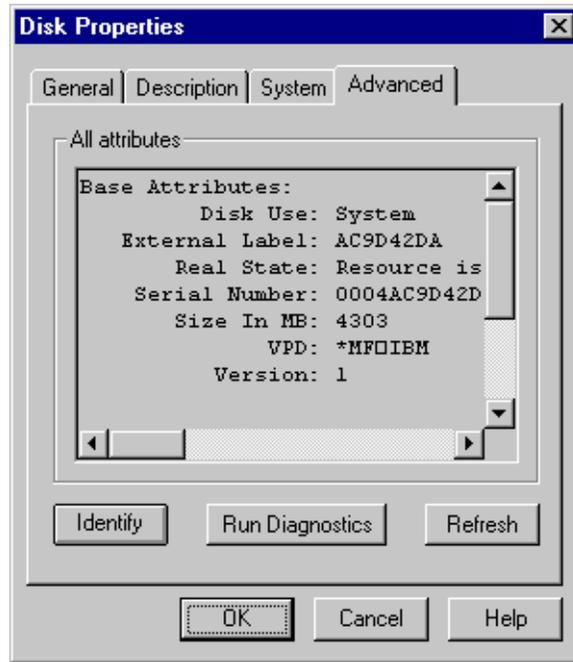


Figure 96. Advanced page of Disk Properties window

The Advanced page of the Disk Properties windows shows:

- Disk use
- External label
- State of the hdisk
- Serial number
- Size (of disk)
- VPD: *MFOIBM (part number)
- StorX version

The Advanced page of the Disk Properties window permits us to make a visual identification of the SSA disk in question by blinking the orange light on the front panel of the drive.

Here you can run concurrent diagnostics on SSA pdisks and the SSA adapter.

You can see which hdisk is used for what purpose. The Disk Properties window can detect the hot spare disk and the SSA RAID array member disks. Figure 97 shows us the General page of the Disk Properties window for pdisk 0 (hot spare disk).

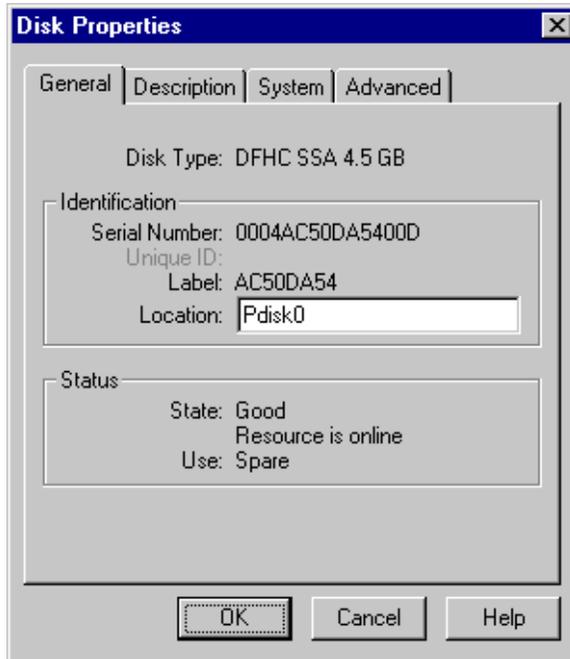


Figure 97. General page of Disk Properties window

You can see that the pdisk 0 usage is set to SPARE.

Figure 98 shows the General page of the Disk Properties window of pdisk 3 (SSA RAID array member).

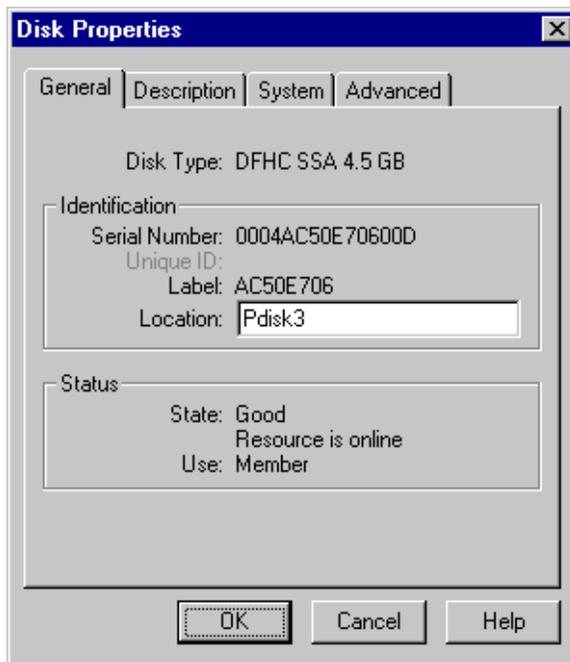


Figure 98. General page of Disk Properties windows

This time pdisk 3 is recognized to be an SSA RAID array member disk, which is correct.

5.11.2 Adapter Properties window

The same type of window is available for the Adapter Properties, as shown in Figure 99.

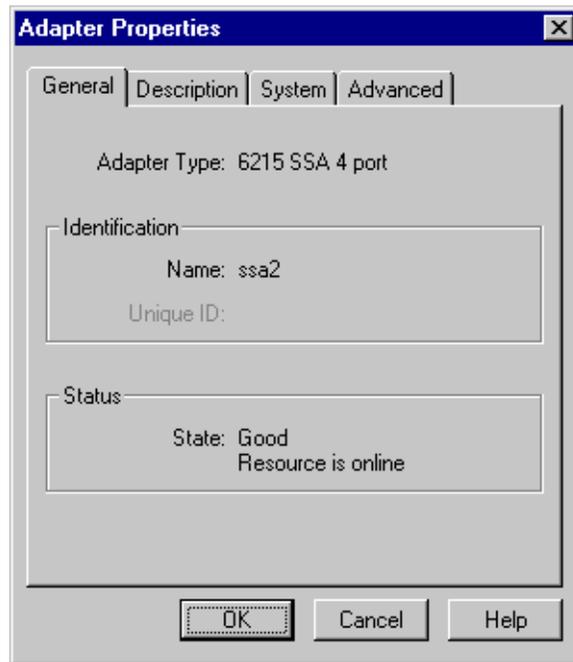


Figure 99. Adapter Properties window

We do not detail the adapter properties in this section, because they offer the same types of information as the Disk Properties windows you have just seen.

Appendix A. Disk replacement methodology

The procedures in the following sections provide a number of methods of using AIX commands to remove a disk and its associated ODM entries cleanly and to enable a replacement disk to be introduced back into the system to the point where data, if any, needs to be restored.

Recently introduced into AIX 4.3.3 is the `copypv` command. This command enables the user to copy a physical volume.

A.1 Disk replacement procedure summary

The disk you are dealing with is probably in an unstable condition and liable to fail at anytime, possibly before the replacement procedure you are attempting has completed. Because of this possibility, you should only attempt to use these methods when you have a full backup of the data. In cases where the disk has failed completely and without warning you may have to rely on a previous backup and so may lose some data.

Appendix B, “SSA RAID disk change flowchart” on page 139 shows a schematic of the decision flowchart for disk replacement.

Remember: Regular backups are the only way to ensure complete safety of your data.

A.2 Methods summary

One of these four methods should be capable of providing a solution to 95 percent of SSA disk replacement situations.

Method 1

This method is the least complicated but requires that you have *two* spare replacement disks and space within the system to accommodate it. The second spare disk simplifies the procedures used to keep the ODM clean.

Use this procedure only if you have only have one replacement disk and space for it be attached to the system. Using this method with only one disk requires only the first half of the procedure and has the disadvantage that the `hdisk` and `pdisk` numbering will either be out of sequence or you will have an additional “ghost” `hdisk` and `pdisk` remaining in the ODM.

Method 2

This method assumes that there is no physical space for any additional disks to be added to the system at all. You do, however, need to have sufficient spare space on other disks within the failing disks Volume Group (VG) to enable all of the data on the failing disk to be moved. It may be that you find yourself just short of enough space to copy all the data across so either by deleting no longer required files or by using a combination of spare space on other disks within the volume group and saving files to tape you can enable the defective disk to be emptied of data prior to removal.

Method 3

This method requires a complete backup of all files completely or partially residing on the disk. The backup will be used during the recovery phase since this method assumes there is no free space within the volume group in which to move the data, or available space to insert an extra disk.

All non-mirrored logical volumes that have data residing on the failing disk will be deleted prior to disk replacement and any mirrored logical volume copies. If the disk being replaced has all of its constituent logical volumes mirrored to different physical volumes, then this method can be used and data recreated without having to resort to the backup, *but still make one anyway!*

Method 4

This method must be used when none of the other methods are suitable. This is normally when the failing disk is no longer usable by AIX and cannot be reactivated. The affected volume group will be removed from the ODM to enable entries for the failed disk to be deleted from the ODM. When the disk has been replaced, it will be necessary to recreate the volume group from scratch and restore from a *previous* backup.

Warning

In *all* of the following methods if the system you are working on is part of a HACMP Cluster *do not* run the `cfgmgr` command. Instead run the following two commands:

```
#/usr/lib/methods/cfgssar -l ssar >/dev/null  
#lsdev -C ssar -S0 -Fname | xargs -n1 mkdev -l
```

(where the character after the -S is a zero, not capital 'O')

A.2.1 Method 1

This procedure allows you to perform a data restore operation by migrating data between disks using a spare SSA disk (if the disk is still accessible). This procedure should *not* be used as a replacement for the standard backup procedure. You should ensure that a valid backup exists prior to using this procedure. The procedure assumes that two spare SSA disks exist (one to be the replacement disk) and that the spare SSA disk is large enough to hold all data currently residing on the failing disk that will be replaced.

It is assumed that both the spare disk and the replacement disk are either new or newly formatted, and if not, that they will be formatted in step 1 or step 13, respectively. This procedure will restore all logical volumes or parts of logical volumes on the replacement disk, but the logical volumes will not necessarily have the same physical partitions allocated to them as on the original disk. Make sure no other users are using the system while this procedure is performed.

Complete these steps:

1. Place the spare SSA disk into the loop and run `cfgmgr` to add it to the system.

```
#cfgmgr
```

Note

The `cfgmgr` command should *not* be executed on any system that is in an HACMP cluster. To do so will seriously damage the configuration of the machine possibly resulting in the crashing of the complete cluster.

Format the disk if necessary using either the `ssa_format` command or by using concurrent diagnostics.

2. Use the `lspv` command to get the name of the spare disk. Also, identify the volume group containing the disk you want to replace. The spare disk will have `none` in the volume group membership column, and if it is a new or newly formatted disk, it will also have `none` in the physical volume ID column. For example:

```
# lspv
hdisk0 00000036960cbdd1 rootvg
hdisk1 00000036960c31de rootvg
hdisk2 00000036960d3007 myvg
hdisk3 none none
```

(`hdisk3` is the name of the spare disk)

3. Make the spare disk a physical volume. For example:

```
# chdev -l hdisk3 -a pv=yes
```

Where `hdisk3` is the name of the spare disk.

4. Add the spare disk to the volume group from which you want to replace another disk. For example:

```
# extendvg -f VGNAME hdisk3
```

(Where `VGNAME` is the name of the VG with the failing disk)

5. Identify which logical volumes reside on the disk you want to replace. **Make a note of their names**, listed in the `LV ID` column, and their types. For example:

```
# lspv -p hdisk0
```

If the disk you are replacing is part of **rootvg** then check step 6 otherwise proceed to step 7.

6. What happens now depends on whether the disk you want to replace contains the specific logical volumes.

Check the output from the command run in step 5 to check if the disk contains the *boot logical volume* (`hd5`) or the *primary dump device* (`hd7`). If either or both of these logical volumes are present on the failing disk perform the following steps for the appropriate LV. If neither of these LVs are present go to step 7.

(It is assumed that the names of the boot logical volume is `hd5` and the primary dump device is `hd7` as originally created during the operating system installation. If they are not, the commands listed in the following steps will have to be modified to reflect the different name used)

For Boot LV:

- d. Migrate the boot logical volume. For example:

```
# migratepv -l hd5 hdisk0 hdisk3
```

- e. Verify that the physical partitions that have been allocated to the migrated boot logical volume are contiguous with the following command:

```
# lslv -m hd5
```

- f. Execute the following commands to recreate the boot image, update the boot record on the spare disk, update the IPL list for normal mode boot in NVRAM, save information about base customized devices in the ODM onto the boot device, and erase the old boot image from the disk you want to replace:

```
# bosboot -d /dev/hdisk3 -a -u -l /dev/hd5
# savebase -d /dev/hdisk3
# mkboot -c -d /dev/hdisk0
```

For Primary Dump device:

- g. Make a system dump logical volume on the spare disk.

```
# mklv -y newdumplv -t sysdump rootvg 2 hdisk3
```

- h. Temporarily change the primary dump device to `newdumplv` using the `sysdumpdev` command.

```
# sysdumpdev -p /dev/newdumplv
```

We are now ready to migrate all remaining LVs to the spare disk.

- i. Move all logical volumes from the disk you want to replace to the spare disk.

```
# migratepv hdisk0 hdisk3
```

This command will take time to complete. It is *very* important to *not* to interrupt the command while it is running.

- j. Change the primary system dump device back to the logical volume identified in step 7.

```
# sysdumpdev -p /dev/hd7
```

- k. Remove the temporary dump logical volume `newdumplv` using the `rmlv` command.

```
# rmlv -f newdumplv
```

Continue with step 7.

- 7. We are now ready to migrate all LVs to the spare disk.

Move all logical volumes from the disk you want to replace to the spare disk.

```
# migratepv hdisk0 hdisk3
```

This command will take time to complete. It is *very* important to *not* to interrupt the command while it is running.

- 8. Check that the failing disk no longer contains any physical partitions that are allocated to any logical volumes.

```
# lsvg -p rootvg
```

The output from the command should confirm that `TOTAL PPs` and `FREE PPs` are identical for the disk you want to replace.

```
PV_NAME. PV STATE. TOTAL PPs. FREE PPs. FREE DISTRIBUTION
hdisk0 __ active _____ 76 _____ 76 _____ 16..15..15..15..15
```

- 9. Reduce the volume group by removing the disk you want to replace from its volume group.

```
# reducevg rootvg hdisk0
```

10. Remove the disk as a device from the system and from the ODM database using the `rmdev` command.

```
# rmdev -l hdisk0 -d
```

11. At this point we will prepare the machine for the failing disk to be removed.

Put the machine into diagnostic mode by using the `diag` command. At the diagnostic entry screen press enter then select Task Selection or Service Aids.

Before removing a SSA disk it is essential to place the loop into string mode either side of the disk we wish to remove. This ensures that the disk being removed passes on any data for other disks contained in its buffer. After the buffer is cleared then the disks either side terminate the loop. Finally the disk is then stopped and the orange LED is switched on to identify the disk to be removed.

12. Select the SSA Service Aid, then select the Set Service Mode and Identify. You will then be presented with a list of all the `pdisks`. Select the `pdisk` number associated with the disk you wish to remove and press enter. You will then see a `>` against the `pdisk` entry, this indicates that it is now safe to remove the disk. If you now look at the SSA disks attached to the system you will now see a disk with an orange light on, this is the disk to remove. After removing the disk place the replacement disk in its place (see step 14 if only one replacement disk available).

13. Now return to the Set Service Mode And Identify screen. Select the `pdisk` marked by the `>`, press enter. This will remove the termination mode from the disks either side of the one you have just removed and incorporate it into the loop.

14. The `pdisk` definition of the removed disk should now be removed from the ODM by using the following command.

```
rmdev -l pdiskx -d
```

Note

If you have followed this procedure but only had one disk available then STOP. You will now have a copy of your data but disk numbering will be inconsistent. It is advisable to place the new disk in the same position as the failed disk for performance reasons but by doing this you may isolate other disks temporarily due to having two breaks in the loop. We suggest that the system is shutdown at this point, when shutdown completes then swap the disk position and insert a disk dummy in vacant slot and then reboot.

15. We now need the replacement disk to assume the identity of the removed disk:

```
cfgmgr
```

Note

The `cfgmgr` command should *not* be executed on any system that is in an HACMP cluster. To do so will seriously damage the configuration of the machine possibly resulting in the crashing the complete cluster.

16. If you wish to return the data onto its original disk number and remove the spare disk from the system then do the following.

Repeat steps 2 to 14 (effectively reversing the hdisk numbers in each step). Then remove the spare disk and replace it with the dummy removed at the beginning.

A.2.2 Method 2

This procedure allows you to move data from the failing disk to other disks in the same volume group and save time in restoring the data from a backup tape. In most circumstances, this procedure will shorten the down time of the system and, therefore, increase customer satisfaction. This procedure *should not* be used as a replacement for the standard backup procedure. It is the responsibility of the customer to ensure that a valid backup exists prior to the CE using this procedure. It is assumed that the replacement disk is either new or newly formatted, and if it is not, that it will be formatted in step 12. This procedure will not restore logical volumes to their original places on the replacement disk. You should decide which, if any, logical volumes should be moved back to the replacement disk and perform that operation. Make sure no other users are using the system while this procedure is being followed. Complete the following steps:

1. Identify which disks belong to which volume groups in the system by using the `lspv` command. The output from the command may look something like the following:

```
hdisk0 00000036960cbdd1 rootvg
hdisk1 00000036960c31de rootvg
hdisk2 00000036960d3007 vg01
hdisk3 000003870001328f vg01
hdisk4 00000360ebf34660 vg01
hdisk5 00000360d7c1f19f vg01
hdisk6 00000036628b9724 vg02
```

This indicates that `hdisk0` and `hdisk1` belong to `rootvg`, `hdisk2`, `hdisk3`, `hdisk4`, and `hdisk5` belongs to `vg01`, and `hdisk6` belongs to `vg02`.

2. Identify the number of used physical partitions on the disk you want to replace. For example:

```
# lspv hdisk0 | grep "USED PPs"
```

3. Identify the number of free physical partitions on all other disks belonging to the same volume group as the disk you want to replace. For example:

```
# lspv hdisk1 | grep "FREE PPs"
```

Repeat this command for all disks belonging to the volume group.

4. Check that the number of Used Physical Partitions (UPP) on the disk you want to replace is smaller than the Sum of Free Physical Partitions (SFPP) on the remaining disks in the same volume group. If UPP is larger than SFPP, this procedure can *not* be used to replace the disk. In this case, you should select an alternative procedure.

5. Identify the names of all logical volumes that fully or partially reside on the disk you want to replace. This step involves the use of multiple commands. Use the `lsvg` command to determine which logical volumes belong to the volume group and which of them span multiple disks. For example:

```
# lsvg -l vg01
```

The command may produce output similar to the following:

```
vg01:
LV NAME TYPE LPs PPs PVs LV STATE MOUNT POINT
loglv01 jfslog 1 1 1 open/syncd N/A
lv04 jfs 31 31 1 open/syncd /vol/x11r4
lv27 jfs 53 53 1 open/syncd /vol/X11R5
lv08 jfs 26 26 1 open/syncd /vol/tex
lv10 jfs 63 63 1 open/syncd /vol/abc
lv21 jfs 84 84 2 open/syncd /var/spool/news
lv12 jfs 99 99 1 open/syncd /vol/lpp/3005
lv23 jfs 66 66 2 open/syncd /vol/src
lv07 jfs 92 92 2 open/syncd /vol/misc
```

This means that of all logical volumes in the `vg01` volume group only three of them (`lv21`, `lv23`, and `lv07`) span two disks (PVs), and that they are 84, 66, and 92 physical partitions in size respectively. Use the `lspv` command to tell you which logical volumes fully or partially reside on the disk you want to replace. Make a note of their names and mark those that only partially reside on the disk. For example:

```
# lspv -l hdisk3
```

The command may produce output similar to the following:

```
hdisk3:
LV NAME LPs PPs DISTRIBUTION MOUNT POINT
lv08 26 26 02..10..10..04..00 /vol/tex
lv21 16 16 02..00..12..00..02 /var/spool/news
lv04 31 31 01..08..05..15..02 /vol/x11r4
```

If you compare this output with the output from `lsvg -l vg01`, you can conclude that `lv08` and `lv04` logical volumes reside entirely on the disk `hdisk3`, and that `lv21` only partially resides on the disk `hdisk3`, since only 16 out of 84 physical partitions for `lv21` are placed on `hdisk3`. In the subsequent steps, `lv08`, `lv04`, and 16 physical partitions belonging to `lv21` will be migrated from `hdisk3` to other disks in the same volume group.

6. What happens now depends on whether the disk you want to replace contains the boot logical volume (`hd5`) or not. See the output from the command run in step 5 to check if the disk contains the boot logical volume.

Note

Note. It is assumed that the name of the boot logical volume is `hd5`, as originally created during the operating system installation. If it is not, the commands listed in the following steps will have to be modified to reflect the different name used.

If the disk to be replaced does not belong to `rootvg`, or does not contain the boot logical volume (`hd5`), continue to step 7.

If the disk to be replaced belongs to `rootvg` and contains the boot logical volume (`hd5`), do the following:

- a. Find a disk belonging to the same volume group that has two contiguous free physical partitions. Use the `lspv` command to find out which physical partitions on the disk are free. For example:

```
# lspv -p hdisk3
```

Migrate the boot logical volume. For example:

```
# migratepv -l hd5 hdisk0 hdisk3
```

- b. Verify that the physical partitions that have been allocated to the migrated boot logical volume are contiguous with the following command:

```
# lslv -m hd5
```

- c. Execute the following commands to recreate the boot image, update the boot record on the spare disk, update the IPL list for Normal mode boot in NVRAM, save information about base customized devices in the ODM onto the boot device, and erase the old boot image from the disk you want to replace. For example:

```
# bosboot -d /dev/hdisk3 -a -u -l /dev/hd5
# savebase -d /dev/hdisk3
# mkboot -c -d /dev/hdisk0
```

- d. Continue with step 7.

7. What you do now depends on whether the disk you want to replace contains the primary system dump logical volume or not. Use the `sysdumpdev` command to determine which logical volume is the primary dump logical volume. Check the output from the command run in step 5 to see if the disk contains part of the system dump logical volume.

If the disk to be replaced does not contain the primary system dump logical volume, or any part of it, do the following:

- a. Move all logical volumes from the disk you want to replace to other disks in the same volume group that have enough free physical partitions. You can select some, or all, of the remaining disks provided the criterion described in step 4 is met. For example:

```
# migratepv hdisk0 hdisk1
```

This command will take time to complete. It is *very* important to *not* interrupt the command while it is running.

- b. Continue with step 8.

However, if the disk to be replaced belongs to rootvg and contains the primary system dump logical volume, or any part of it, do the following:

- a. Make a system dump logical volume on one of the remaining disks in the same volume group that has at least two free physical partitions. For example:

```
# mklv -y newdumplv -t sysdump rootvg 2 hdisk1
```

- b. Temporarily change the primary dump device to `newdumplv` using the `sysdumpdev` command. For example:

```
# sysdumpdev -p /dev/newdumplv
```

- c. Move all logical volumes from the disk you want to replace to other disks in the same volume group that have enough free physical partitions. You can select some, or all, of the remaining disks provided the criterion described in step 4 is met. For example:

```
# migratepv hdisk3 hdisk4 hdisk5
```

This command will take time to complete. It is *very* important to *not* interrupt the command while it is running.

- d. Change the primary system dump device back to the primary device identified above. For example:

```
# sysdumpdev -p /dev/hd7
```

- e. Remove the temporary dump logical volume `newdumplv` with the `rmlv` command. For example:

```
# rmlv -f newdumplv
```

- f. Continue with step 8.

8. Check that the disk you want to replace contains no physical partitions that are allocated to any logical volumes. For example:

```
# lsvg -p rootvg
```

The output from the command should confirm that the `TOTAL PPs` and `FREE PPs` for the disk you want to replace are identical. For example:

```
PV_NAME PV STATE TOTAL PPs FREE PPs FREE DISTRIBUTION
hdisk0 active 159 159 32..32..31..32..32
```

9. Reduce the volume group by removing the disk you want to replace from its volume group. For example:

```
# reducevg rootvg hdisk0
```

10. Remove the disk as a device from the system and from the ODM database using the `rmdev` command. For example:

```
# rmdev -l hdisk0 -d
```

11. At this point we will prepare the machine for the disk to be removed. Put the machine into diagnostic mode by using the `diag` command. At the diagnostic entry screen press enter then select Task Selection or Service Aids.

Before removing a SSA disk it is essential to place the loop into string mode either side of the disk we wish to remove. This ensures that the disk being removed passes on any data for other disks contained in its buffer. After the buffer is cleared then the disks either side terminate the loop. Finally the disk is then stopped and the orange light switched on to identify the disk to be removed.

12. Select the SSA Service Aid, then select the Set Service Mode and Identify. You will then be presented with a list of all the `pdisks`. Select the `pdisk` number associated with the disk you wish to remove and press enter. You will then see a `>` against the `pdisk` entry, this indicates that it is now safe to remove the disk. If you look at the SSA disks attached to the system you will now see a disk with an orange light on, this is the disk to remove. After removing the disk place the new disk in its place.

13. After step 12 is completed return to the Set Service Mode And Identify screen. Select the `pdisk` marked by the `>`, press enter. This will now remove the termination mode from the discs either side of the one you have just removed and incorporate into the loop the disc you have just replaced.

14. The `pdisk` entry of the removed disc should now be removed from the ODM at this point by using the following command.

```
rmdev -l pdiskx -d
```

In step 13 you placed another disc in the place of the removed one, then run `cfgmgr` to add this disc to the system.

Note

The `cfgmgr` command should *not* be executed on any system that is in an HACMP cluster. To do so will seriously damage the configuration of the machine possibly resulting in the crashing of the complete cluster.

15. Confirm that the new disk has the same name as the disk that was replaced.
16. Make the new disk a physical volume with the `chdev` command. For example:

```
# chdev -l hdisk0 -a pv=yes
```

Once the procedure is completed, you have a number of options available for using the replacement disk:

- Leave the new disk empty, and use it later on when needed.
- Add the disk to the original volume group, and move some or all of the logical volumes that originally resided on the replaced disk, or any other logical volumes in that volume group, to the new disk.

As a general guideline, you may wish to move only those logical volumes that fully resided on the replaced disk. If you decide to move logical volumes that partially resided on the replaced disk as well, there is a chance that not only the physical partitions that were on the replaced disk, but all physical partitions belonging to these logical volumes, will be moved to the new disk. In some cases, the migration process might fail due to lack of disk space.

A.2.3 Method 3

This procedure describes the proper way of unconfiguring and replacing a hard disk. It should be used when neither Method 1 or Method 2 (as described above) can be applied, due to lack of room in the SSA subsystem or spare disk availability, or because these two procedures are simply not applicable. This method is also applicable when the disc is known to be mirrored. Mirroring is addressed in point 6 of the method below.

This procedure ensures that the system will be left in a sound state after the disk is replaced, that is, no ODM database or VGDA corruptions will occur. The procedure assumes that a valid backup of the data residing on the disk exists. That backup *will be used* to restore the data once the disk is unconfigured and replaced.

It is assumed that the replacement disk is either new or newly formatted. Make sure no other users are using the system while this procedure is being followed. Complete the following steps:

1. Identify to which volume group the disk you want to replace belongs using the `lspv` command. For example, the output from the command may look like the following:

```
hdisk0 00000036960cbdd1 rootvg
hdisk1 00000036960c31de rootvg
hdisk2 00000036960d3007 vg01
hdisk3 000003870001328f vg01
hdisk4 00000360ebf34660 vg01
hdisk5 00000360d7c1f19f vg01
hdisk6 00000036628b9724 vg02
```

This indicates that `hdisk0` and `hdisk1` belong to `rootvg`; `hdisk2`, `hdisk3`, `hdisk4`, and `hdisk5` belong to `vg01`; and `hdisk6` belongs to `vg02`.

2. Identify the names of all logical volumes that fully or partially reside on the disk you want to replace. This step involves the use of multiple commands. Use the `lsvg` command to determine which logical volumes belong to the volume group, which of them span multiple disks, and which of them are mirrored. For example:

```
# lsvg -l vg01
```

The command may produce output similar to the following:

```
vg01:
LV NAME TYPE LPs PPs PVs LV STATE MOUNT POINT
loglv01 jfslog 1 1 1 open/syncd N/A
lv04 jfs 31 31 1 open/syncd /vol/x11r4
lv27 jfs 53 53 1 open/syncd /vol/X11R5
lv08 jfs 26 26 1 open/syncd /vol/tex
lv10 jfs 63 63 1 open/syncd /vol/abc
lv21 jfs 84 84 2 open/syncd /var/spool/news
lv12 jfs 99 99 1 open/syncd /vol/lpp/3005
lv23 jfs 66 66 2 open/syncd /vol/src
lv07 jfs 92 92 2 open/syncd /vol/misc
```

This means that of all logical volumes in the `vg01` volume group only three of them (`lv21`, `lv23`, and `lv07`) span two disks (PVs), and that they are 84, 66, and 92 physical partitions in size, respectively. In this example, no logical volumes are mirrored since the number of LPs and PPs is always the same. When the number of PPs is twice or three times bigger, the number of LPs then a logical volume is singly or doubly mirrored.

If the disk you want to replace is the only disk in its volume group, this command will simply tell you which logical volumes reside on the disk you want to replace. In this case, you need not run any other commands described in this step.

Use the `lspv` command to tell you which logical volumes fully or partially reside on the disk you want to replace. Make a note of their names and mark those that only partially reside on the disk, or have mirror copies on it. For example:

```
# lspv -l hdisk3
```

The command may produce output similar to the following:

```
hdisk3:
LV NAME LPs PPs DISTRIBUTION MOUNT POINT
lv08 26 26 02..10..10..04..00 /vol/tex
lv21 16 16 02..00..12..00..02 /var/spool/news
lv04 31 31 01..08..05..15..02 /vol/x11r4
```

If you compare this output with the output from `lsvg -l vg01`, you can conclude that `lv08` and `lv04` logical volumes reside entirely on the disk `hdisk3`, and that `lv21` only partially resides on the disk `hdisk3`, since only 16 out of 84 physical partitions for `lv21` are placed on `hdisk3`. In the subsequent steps, `lv08`, `lv04`, and 16 physical partitions belonging to `lv21` will be migrated from `hdisk3` to other disks in the same volume group.

3. Check that you have a backup of all file systems (logical volumes of `jfs` type) that either fully or partially reside on the hard disk that will be replaced, or have mirror copies on it. If not, perform such a backup by using your favorite

command, for example `backup`, `tar`, or `cpio`. You should also ensure that you have a backup of all data on raw logical volumes. For example, database packages, such as Oracle, Informix, and Sybase, can be configured to use data on logical volumes that are not of `jfs` type.

4. Unmount all single-copy (non-mirrored) file systems that either fully or partially reside on the hard disk that is to be replaced. For example:

```
# umount /vol/tex
```

5. Delete all single-copy file systems from the disk that is to be replaced. For example:

```
# rdfs /vol/tex
```

6. Remove physical partition copies of mirrored logical volumes from the disk that will be replaced. For example:

```
# rmlvcopy lv100 1 hdisk3
```

7. Close all other logical volumes residing on the disk to be replaced.

If the disk contains a raw logical volume with data on it, it will probably be closed when the application that is using it is stopped. The customer should know how to close such a logical volume.

If the disk contains a paging space, it will have to be inactivated after the next system restart. The following steps will achieve this:

- a. Change the paging space so it will not be used at the next system restart. For example:

```
chps -a n pagingXX
```

Where `pagingXX` is the name of the logical volume containing the paging space.

- b. Shut the system down and reboot the machine using the following command.

```
shutdown -Fr
```

- c. Confirm that `pagingXX` is not active after the reboot with the following command:

```
# lspv -a
```

If the disk contains a non-mirrored journaled file system log (`jfslog`), and it is *not* the only disk in its volume group, you will have to migrate the journaled file system log to another disk in the volume group. This will allow file systems residing on other disks in the volume group to remain mounted and available. The following command will achieve this:

```
# migratepv -l jfslogname pvname1 pvname2
```

Where `jfslogname` is the name of the log logical volume, `pvname1` is the name of the disk to be replaced, and `pvname2` is the name of the target disk that has a free physical partition to accept the `jfslog`.

8. Reduce the volume group by removing the disk you want to replace from its volume group. If the disk you want to replace is the only disk in its volume group, the volume group will also be deleted. For example:

```
# reducevg -df vg01 hdisk3
```

9. Check that the disk you want to replace does not belong to any volume group by using the `lspv` command.

10. Remove the disk as a device from the system and from the ODM database with the `rmdev` command. For example:

```
#rmdev -l hdisk3 -d
```

11. At this point we will prepare the machine for the disk to be removed. Put the machine into diagnostic mode by using the `diag` command. At the diagnostic entry screen press enter then select Task Selection or Service Aids.

Before removing a SSA disk it is essential to place the loop into string mode either side of the disk we wish to remove. This ensures that the disk being removed passes on any data for other disks contained in its buffer. After the buffer is cleared then the disks either side terminate the loop. Finally the disk is then stopped and the orange light switched on to identify the disk to be removed.

12. Select the SSA Service Aid, then select the Set Service Mode and Identify. You will then be presented with a list of all the pdisks. Select the pdisk number associated with the disk you wish to remove and press enter. You will then see a > against the pdisk entry, this indicates that it is now safe to remove the disk. If you look at the SSA disks attached to the system you will now see a disk with an orange light on, this is the disk to remove. After removing the disk place the new disk in its place.

13. After step 12 is completed return to the Set Service Mode And Identify screen. Select the pdisk marked by the >, press enter. This will now remove the termination mode from the discs either side of the one you have just removed and incorporate into the loop the disc you have just replaced.

14. The pdisk entry of the removed disc should now be removed from the ODM at this point by using the following command.

```
rmdev -l pdiskx -d
```

In step 13 you placed another disc in the place of the removed one, then run `cfgmgr` to add this disc to the system.

Note

The `cfgmgr` command should *not* be executed on any system that is in an HACMP cluster. To do so will seriously damage the configuration of the machine possibly resulting in the crashing of the complete cluster.

15. Confirm that the new disk has the same name as the disk that was replaced with the `lspv` command.

16. Make the new disk a physical volume with the `chdev` command. For example:

```
# chdev -l hdisk3 -a pv=yes
```

17. Depending on whether the disk that was replaced was the only disk in its volume group or not, run one of the following:

- If the disk was the only disk in its volume group, create the volume group again, and activate it:

```
# mkvg -y vgname newpvname  
# varyonvg vgname
```

Where `vgname` is the name of the volume group, and `newpvname` is the name of the new disk.

- If the disk was one of several disks in its volume group, add the new disk to the same volume group from which the old one was removed:

```
# extendvg -f vgname newpvname
```

Where `vgname` is the name of the volume group, and `newpvname` is the name of the new disk.

18. Recreate the logical volumes and file systems removed in the previous steps using SMIT `fastpath smitty lvm`, and then restore the data from backup tape(s) or diskettes into the file systems and raw logical volumes if there were any.

19. If a mirror copy was removed in step 6, it can be recreated now by using SMIT, and then synchronizing the mirror copies by running the following command:

```
# syncvg -p newpvname
```

Where `newpvname` is the name of the new hard disk.

A.2.4 Method 4

This procedure describes the proper way of unconfiguring and replacing a hard disk when it belongs to a non-rootvg volume group and the disk or data on it cannot be accessed.

This procedure ensures that the system will be left in a sound state after the disk is replaced, that is, no ODM database or VGDA corruptions will occur, and the disks will have consecutive numbers.

This procedure assumes that a valid backup of the data residing on *all* disks in the volume group exists. This backup *will be used* to restore the data once the disk is unconfigured and replaced, and the volume group is recreated. It is assumed that the replacement disk is either new or newly formatted. Make sure no other users are using the system while this procedure is being followed. Complete the following steps:

1. If the system is still up and running, identify volume group membership for all disks in the system by using the `lspv` command. If the system is not running, ask the customer to provide this information. For example, the output from the command may look like the following:

```
hdisk0 00000036960cbdd1 rootvg
hdisk1 00000036960c31de rootvg
hdisk2 00000036960d3007 vg01
hdisk3 000003870001328f vg01
hdisk5 00000360d7c1f19f vg01
hdisk6 00000036628b9724 vg02
```

This indicates that `hdisk0` and `hdisk1` belong to `rootvg`; `hdisk2`, `hdisk3`, `hdisk4`, and `hdisk5` belong to `vg01`; and `hdisk6` belongs to `vg02`.

2. We now need to remove all of the disks associated with the volume group of the failing disk, therefore, you need to locate some SSA blanks to put into the loop to maintain the integrity of the SSA loop. To translate the corresponding `hdisk`s into `pdisk`s use the following command:

```
ssaxlate -l hdiskx
```

You will then on the return get the `pdisk` number. Make a note of the number and then repeat the command for the remaining disks in the volume group.

3. To Insure we remove the correct disks from the loop we need to find their position in the loop using the SSA Service Aids “Set Service Mode and Identify” facility. Use the commands listed below to perform these actions.

```
diag
```

Select Task Selection or Service Aids, when the list of options appears Ctrl-v down to near the bottom of the list till you see SSA Service Aids. Select this option then select Set Service Mode and Identify. You will now get a list of all the pdisks. Using the information gathered in step 2, move the cursor to cover the first pdisk number on your list. Press enter and take the option Set or Reset Identify. The screen will return to the list you had earlier except that the disk you selected now will have a > by the side of it. Now repeat these actions for the rest of the pdisks on your list. Now view the SSA disks connected to the machine and you should now be able to locate the positions of all the disks you need to remove, the disks you marked will have the orange light on.

4. Having made a note of the locations of the disks to remove in step 3, power off the system using:

```
shutdown -F
```

Remove the disks one at a time making sure you put a blank in its place before removing the next disk. This process is necessary to ensure that no disks are isolated in the loop when the machine is powered back on.

Note

If the SSA loop that you are working on is shared by other systems then you should use the SSA Service Aids “Set Service Mode and Identify” function on one of the other systems in the loop. The procedure you use is the same as described in point 3 except you use the Set Service Mode option and the pdisk numbers used may be different from the numbers used in step 3.

5. Power the system on.
6. Confirm the failed disk and other disks you disconnected in step 3 are in the DEFINED state, rather than AVAILABLE by using the following command:

```
# lsdev -Cc disk
```

7. Confirm the volume group that contains the failed disk that was disconnected in step 4 is still known to the system by using the `lsvg` command.
8. Export the volume group that contains the failed disk using the following command:

```
# exportvg vgroupname
```

Where `vgroupname` is the name of the volume group. This will remove all references to the volume group and all logical volumes and file systems belonging to it from the ODM database, `/dev` directory, and `/etc/filesystems` file.

9. Check that the `/etc/filesystems` file does not have any stanzas left for the file systems that belonged to the exported volume group. If there are any stanzas left, delete them by using your favorite editor.
10. Remove the failed disk as a device from the system and from the ODM database using the `rmdev` command:

```
# rmdev -l oldpvname -d
```

Where `oldpvname` is the name of the failed disk.

11. Confirm the failed disk is not known to the system any more using the `lsdev` command:

```
# lsdev -Cc disk
```

The output should not list the failed disk.

12. At this point the original disks can be refitted preferably in the positions they were removed from. As SSA is hot swappable remove the blank then refit the corresponding disk. Wait for it to power up, signified by both green lights being on. When that disk has started proceed with the other disks in the same manner including the defective disks replacement.

13. Configure the new disk into the system using the configuration manager.

```
cfgmgr.
```

Note

The `cfgmgr` command should *not* be executed on any system that is in an HACMP cluster. To do so will seriously damage the configuration of the machine possibly resulting in the crashing of the complete cluster.

14. Confirm that the new disk has the same name as the disk that was replaced by using the `lspv` command.

15. Make the new disk a physical volume:

```
# chdev -l newpvname -a pv=yes
```

Where `newpvname` is the name of the new disk.

16. Depending on whether the disk that was replaced was the only disk in its volume group or not, run one of the following:

- If the disk was the only disk in its volume group, recreate the volume group on the replacement disk, and activate it:

```
# mkvg -y vgroupname newpvname
# varyonvg vgroupname
```

Where `vgroupname` is the name of the volume group, and `newpvname` is the name of the new disk.

- If the disk was one of several disks in its volume group, recreate the volume group on the replacement disk and all other disks that used to belong to the volume group and activate the volume group:

```
# mkvg -y vgroupname newpvname pvname2 pvname3 ...
# varyonvg vgroupname
```

Where `vgroupname` is the name of the volume group, `newpvname` is the name of the new disk, and `pvname2`, `pvname3`, and so on are the names of other disks that used to belong to the same volume group.

Note

Since the disks that previously belonged to the volume group still contain data, you will be warned that the data on them will be destroyed. You will be asked if you want to continue, to which you should say yes provided you have confirmed that you have a good backup.

17. You should now recreate the logical volumes and file systems in the volume group, and then restore the data from backup tape(s) or diskettes.

Appendix B. SSA RAID disk change flowchart

Figure 100 and Figure 101 show the sequence of events required to replace an SSA RAID disk drive.

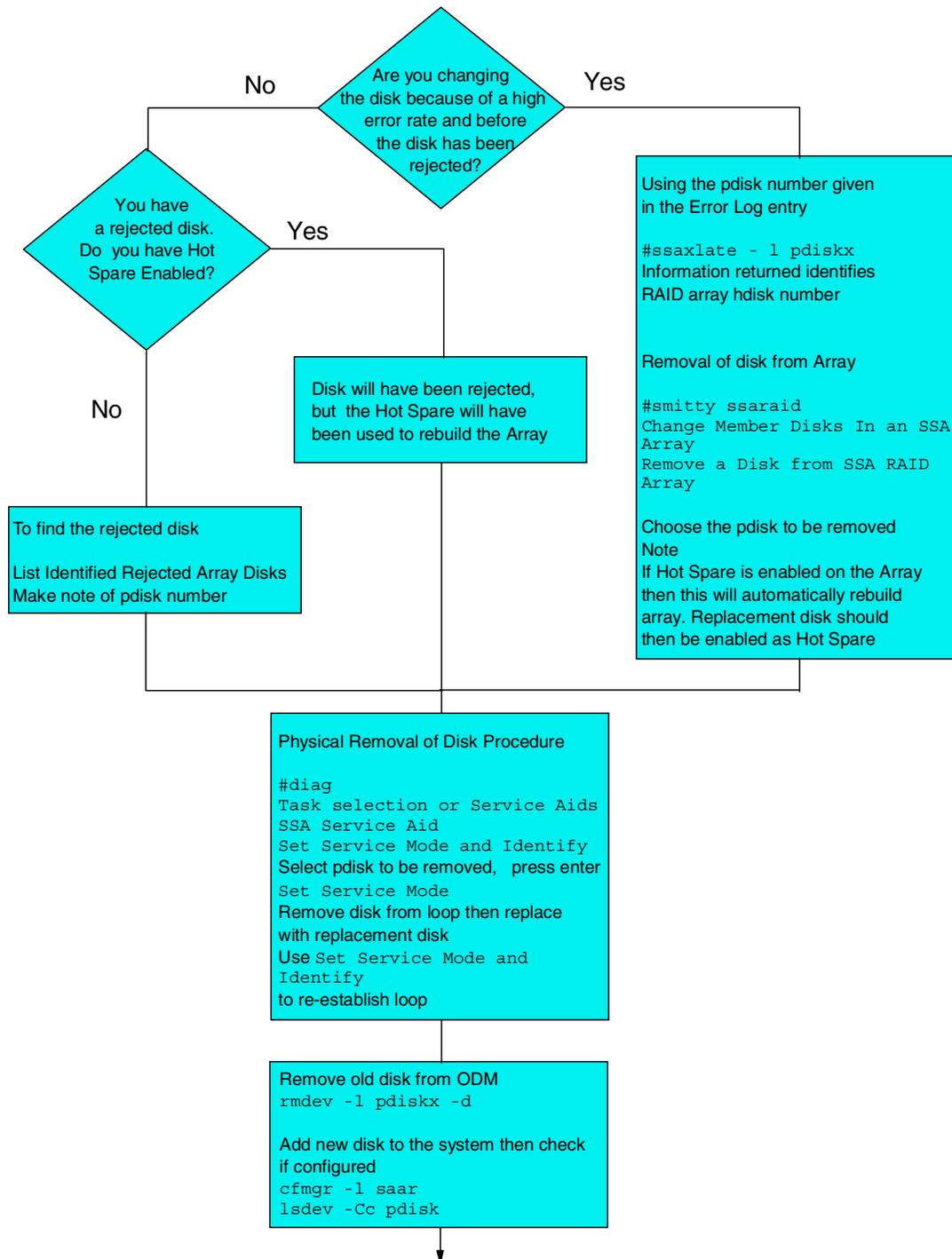


Figure 100. SSA RAID disk change flowchart : part 1

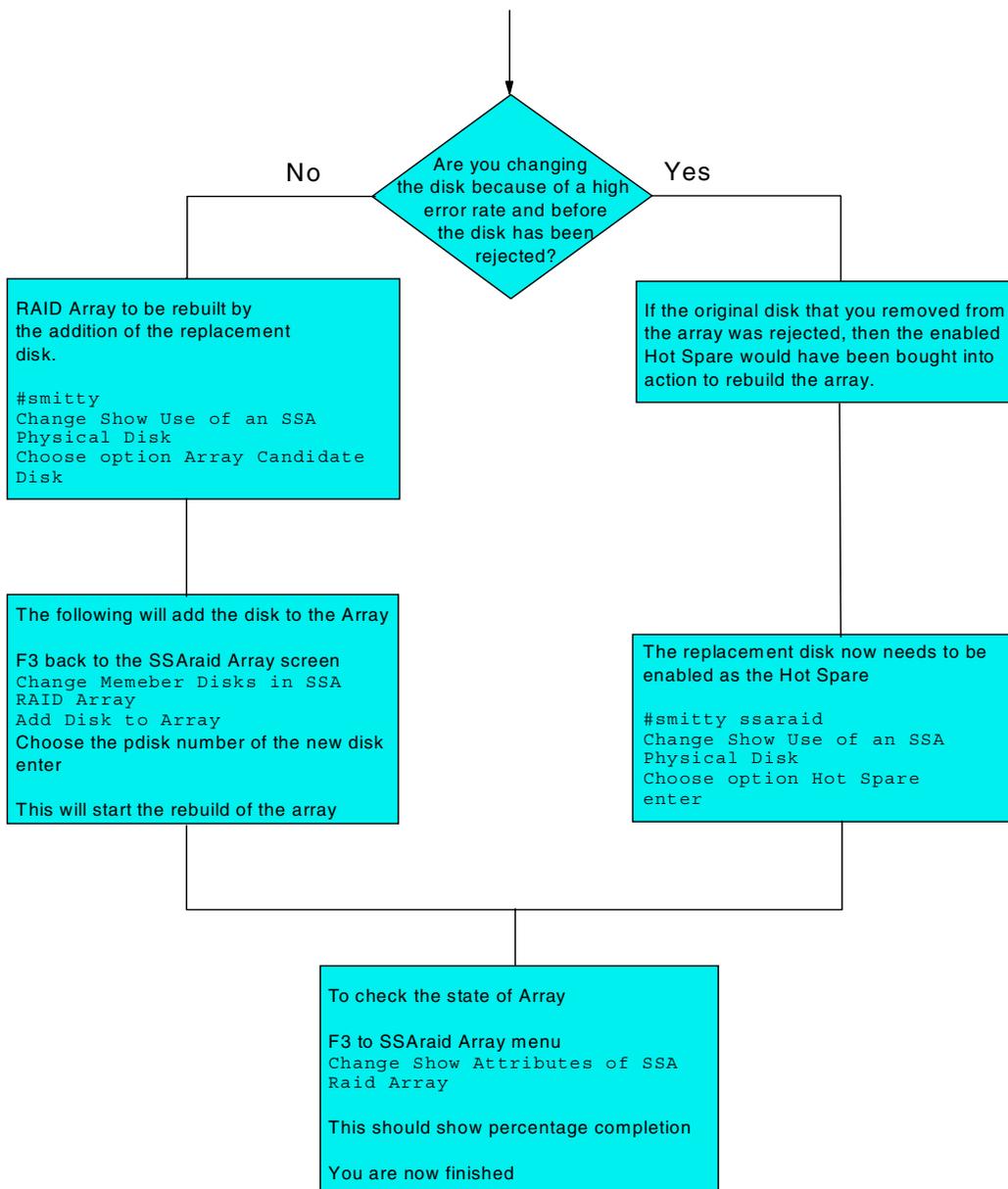


Figure 101. SSA RAID disk change flowchart: part 2

Appendix C. Special notices

This publication is intended to help storage administrators to understand, manage and monitor their SSA storage systems. The information in this publication is not intended as the specification of any programming interfaces that are provided by SSA. See the PUBLICATIONS section of the IBM Programming Announcement for SSA for more information about what publications are considered to be product documentation.

References in this publication to IBM products, programs or services do not imply that IBM intends to make these available in all countries in which IBM operates. Any reference to an IBM product, program, or service is not intended to state or imply that only IBM's product, program, or service may be used. Any functionally equivalent program that does not infringe any of IBM's intellectual property rights may be used instead of the IBM product, program or service.

Information in this book was developed in conjunction with use of the equipment specified, and is limited in application to those specific hardware and software products and levels.

IBM may have patents or pending patent applications covering subject matter 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 the IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact IBM Corporation, Dept. 600A, Mail Drop 1329, Somers, NY 10589 USA.

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The information contained in this document has not been submitted to any formal IBM test and is distributed AS IS. The information about non-IBM ("vendor") products in this manual has been supplied by the vendor and IBM assumes no responsibility for its accuracy or completeness. The use of this information or the implementation of any of these techniques is a customer responsibility and depends on the customer's ability to evaluate and integrate them into the customer's operational environment. While each item may have been reviewed by IBM for accuracy in a specific situation, there is no guarantee that the same or similar results will be obtained elsewhere. Customers attempting to adapt these techniques to their own environments do so at their own risk.

Any pointers in this publication to external Web sites are provided for convenience only and do not in any manner serve as an endorsement of these Web sites.

Any performance data contained in this document was determined in a controlled environment, and therefore, the results that may be obtained in other operating environments may vary significantly. Users of this document should verify the applicable data for their specific environment.

Reference to PTF numbers that have not been released through the normal distribution process does not imply general availability. The purpose of including these reference numbers is to alert IBM customers to specific information relative to the implementation of the PTF when it becomes available to each customer according to the normal IBM PTF distribution process.

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

AIX	AS/400
AT	CT
IBM	Micro Channel
Netfinity	Netfinity Manager
OS/2	RS/6000
ServeRAID	SP
SP2	StorWatch
System/390	400
TME	

The following terms are trademarks of other companies:

C-bus is a trademark of Corollary, Inc. in the United States and/or other countries.

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

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

SLIC, Instant Copy, and SLIC Manager are trademarks of Vicom Systems in the United States and/or other countries.

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

PC Direct is a trademark of Ziff Communications Company in the United States and/or other countries and is used by IBM Corporation under license.

ActionMedia, LANDesk, MMX, Pentium and ProShare are trademarks of Intel Corporation in the United States and/or other countries.

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

SET and the SET logo are trademarks owned by SET Secure Electronic Transaction LLC.

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

Appendix D. Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this redbook.

D.1 IBM Redbooks publications

For information on ordering these publications see “How to get IBM Redbooks” on page 145.

- *Monitoring and Managing IBM SSA Disk Subsystems*, SG24-5251
- *Exploiting SSA Disk Subsystems in Sun Solaris Platform Environments*, SG24-5083
- *Implementing Netfinity Disk Subsystems: ServeRAID SCSI, Fibre Channel and SSA*, SG24-2098
- *A Practical Guide to Serial Storage Architecture for AIX*, SG24-4599
- *Introduction to Storage Area Network, SAN*, SG24-5470

D.2 IBM Redbooks collections

Redbooks are also available on the following CD-ROMs. Click the CD-ROMs button at <http://www.redbooks.ibm.com/> for information about all the CD-ROMs offered, updates and formats.

CD-ROM Title	Collection Kit Number
System/390 Redbooks Collection	SK2T-2177
Networking and Systems Management Redbooks Collection	SK2T-6022
Transaction Processing and Data Management Redbooks Collection	SK2T-8038
Lotus Redbooks Collection	SK2T-8039
Tivoli Redbooks Collection	SK2T-8044
AS/400 Redbooks Collection	SK2T-2849
Netfinity Hardware and Software Redbooks Collection	SK2T-8046
RS/6000 Redbooks Collection (BkMgr Format)	SK2T-8040
RS/6000 Redbooks Collection (PDF Format)	SK2T-8043
Application Development Redbooks Collection	SK2T-8037
IBM Enterprise Storage and Systems Management Solutions	SK3T-3694

D.3 Other resources

These publications are also relevant as further information sources:

- *7133 SSA Subsystem: Operator Guide*, GA33-3259
- *7133 SSA Subsystem: Hardware Technical Information*, SA33-3261
- *7133 Models 010 and 020 SSA Disk Subsystems: Installation Guide*, GA33-3260
- *7133 Models 500 and 600 SSA Disk Subsystems: Installation Guide*, GA33-3263
- *7133 SSA Disk Subsystem: Service Guide*, SY33-0185
- *SSA Adapters: Technical Reference*, S31H-8612
- *Planning SSA RAID Subsystems*, GA33-3271

- *SSA Adapters: User's Guide and Maintenance Information*, SA33-3272
- *Micro Channel SSA RAID Adapters Technical Reference*, SA33-3270
- *PCI SSA RAID Adapters: Technical Reference*, SA33-3225
- *AIX Version 3.2 and V4 Performance Monitoring and Tuning Guide*, SC23-2365
- *PCI Adapter Placement Reference*, SA38-0538
- *IBM 2108 SAN Data Gateway Model S20, Installation and User's Guide*, SC26-7312

How to get IBM Redbooks

This section explains how both customers and IBM employees can find out about IBM Redbooks, redpieces, and CD-ROMs. A form for ordering books and CD-ROMs by fax or e-mail is also provided.

- **Redbooks Web Site** <http://www.redbooks.ibm.com/>

Search for, view, download, or order hardcopy/CD-ROM redbooks from the redbooks Web site. Also read redpieces and download additional materials (code samples or diskette/CD-ROM images) from this redbooks site.

Redpieces are redbooks in progress; not all redbooks become redpieces and sometimes just a few chapters will be published this way. The intent is to get the information out much quicker than the formal publishing process allows.

- **E-mail Orders**

Send orders by e-mail including information from the redbooks fax order form to:

	e-mail address
In United States	usib6fpl@ibmmail.com
Outside North America	Contact information is in the "How to Order" section at this site: http://www.elink.ibm.ibm.com/pbl/pbl

- **Telephone Orders**

United States (toll free)	1-800-879-2755
Canada (toll free)	1-800-IBM-4YOU
Outside North America	Country coordinator phone number is in the "How to Order" section at this site: http://www.elink.ibm.ibm.com/pbl/pbl

- **Fax Orders**

United States (toll free)	1-800-445-9269
Canada	1-403-267-4455
Outside North America	Fax phone number is in the "How to Order" section at this site: http://www.elink.ibm.ibm.com/pbl/pbl

This information was current at the time of publication, but is continually subject to change. The latest information may be found at the redbooks Web site.

IBM Intranet for Employees

IBM employees may register for information on workshops, residencies, and redbooks by accessing the IBM Intranet Web site at <http://w3.itso.ibm.com/> and clicking the ITSO Mailing List button. Look in the Materials repository for workshops, presentations, papers, and Web pages developed and written by the ITSO technical professionals; click the Additional Materials button. Employees may access MyNews at <http://w3.ibm.com/> for redbook, residency, and workshop announcements.

List of abbreviations

ADSM	ADSTAR Distributed Storage Manager	UID	unique identifier
CPU	central processing unit	VG	volume group
DASD	direct access storage device	VPD	vital product data
FW	fast write		
HACMP	high availability for cluster multiprocessing		
HAGEO	high availability geographic cluster		
HSM	hierarchical file system		
HW	hardware		
ID	identification, identity		
I/O	input or output		
ioctl	I/O control		
ITSO	International Technical Support Organization		
JBOD	just a bunch of disks		
JFS	journaled file system		
LED	light-emitting diode		
LUN	logical unit, logical unit number		
LV	logical volume		
LVM	logical volume manager		
MSCS	Microsoft cluster server configuration		
MWC	mirror write consistency		
OE	fibre optical extender		
OS	operating system		
PCI	personal computer interconnect		
PP	physical partition		
PTF	Program Temporary Fix		
PV	physical volume		
RAID	redundant array of independent disks		
RPC	remote procedure call		
RSM	remote systems management		
SCSI	small computer system interface		
SLIC	serial link interface controller		
SMP	symmetrical multiprocessor		
SMIT	system management interface tool		
SSA	serial storage architecture		
StorX	IBM Serial Storage Expert		
Sysback	AIX System Backup and Recovery 6000		
TCP/IP	transmission control protocol/ Internet protocol		

Index

Numerics

160MB/s 3
2108 85
2108 SAN Data Gateway Model S20 24
7133 Model 010/500 15
7133 models 20/600 17
7133 models D40/T40 19
7190 85

A

abbreviations 147
acronyms 147
advantages of SSA 3
ANSI 1
auto-configuring 4

B

bandwidth 7
bypass circuit cards 17
Bypass mode 17

C

cfgmgr 63
Check Light 21
ciao 71
commands 64
Configuring Drives 63
Customized Storx Viewer 90
cyclical redundancy check 22

D

Disk Fencing 62
Disk Properties 114
domain 3

E

Enclosure Microcode 21
enhanced control panel 19
Enhanced Loop Adapters 63
Event Monitor 91

F

Forced Inline mode 17

H

Hot Spare Disk Definition 106

I

ID Display 21
ID Push Button 21

L

LED 18
Link Status 18
Logical Disk Menu 66
loop
 availability 3
LVM 78

M

maintaining SSA 61
maintenance 63
Management Set Properties 90
managing SSA 61
managing SSA disks
 HP-UX environment 80
 in other environments 85
 Sun Solaris environment 79
 Windows NT environment 81
master initiator 2
Menu
 SSA Adapter 65
 SSA Disks 66
 SSA Physical Disk 67
 SSA RAID Arrays 67
mergelogs 71
Micro Channel 7
Monitoring and Managing
 Complex Configuration 94
 Diagnosing with StorX 114
 Disk Problems in a RAID 5 Configuration 109
 hdisk Failure 98
 Shared Configuration 92
 SSA Adapter Failure 103
 SSA Open Loop Problem 102
 SSA RAID Array Creation 105
monitoring SSA storage network 87

O

Online Service Monitoring 23

P

Pathlight Adapters 85
Power Light 21
Power Switch-On
 Model D40 20
 Model T40 21

R

RAID Volume Group Creation 108
Remote Systems Management 71
RSM 71

S

SAN 4
SCSI 1

- SCSI Enclosure Services 72, 87
- SCSI-2 mapping 3
- SES 72
- SES functions 19
- SLIC Router FC-SL 5
- SMIT 64
- SNBC 20
- Spatial Reuse 3
- spatial reuse 2
- specifications
 - 7133 15
 - enhanced RAID SSA adapters 8
 - PCI RAID adapters 11
 - RS/6000 advanced RAID adapters 9
 - RS/6000 advanced RAID plus adapters 10
 - RS/6000 non_RAID adapters 8
 - RS/6000 RAID adapters 8
 - SUN SBus adapter 10
- SSA 2
 - advantages 3
 - loop 3
 - nodes 1
 - operation 2
 - subsystems 1
 - what is it? 1
- SSA Cable 72
- SSA commands 64
- SSA logical disks 61
- SSA Node Bypass cards 20
- SSA physical disks 61
- SSA RAID Array Creation 107
- SSA Spare Tool 78
- ssaadap command 68
- ssacand command 69
- ssaconn command 69
- ssadisk command 69
- ssaidentify command 68
- ssaxlate command 68
- Storage Area Network 4
- StorWatch Serial Storage Expert 87
- StorX 87, 89
 - Customizing Live View 89
 - event monitor 91
- Subsystems
 - 2108 SAN Data Gateway Model S2 24
 - 3527 12
 - 7131 13
 - 7133 13
 - 7190-100 22
 - 7190-200 24
- sweep 71

T

- Tools 61
 - Disksuite 80
 - Remote Systems Management 81
 - SSA Cable 72
 - Veritas 80
 - WebSSA 81
- twin adapters in a Loop 62

U

- UID 1
- utilities 70

V

- VICOM 5
- Vicom and IBM Product Comparison 86
- Vital Product Data 64
- VPD 23
- vpfe 70

IBM Redbooks evaluation

Understanding SSA Subsystems in Your Environment
SG24-5750-00

Your feedback is very important to help us maintain the quality of IBM Redbooks. **Please complete this questionnaire and return it using one of the following methods:**

- Use the online evaluation form found at <http://www.redbooks.ibm.com/>
- Fax this form to: USA International Access Code + 1 914 432 8264
- Send your comments in an Internet note to redbook@us.ibm.com

Which of the following best describes you?

Customer **Business Partner** **Solution Developer** **IBM employee**
 None of the above

Please rate your overall satisfaction with this book using the scale:
(1 = very good, 2 = good, 3 = average, 4 = poor, 5 = very poor)

Overall Satisfaction _____

Please answer the following questions:

Was this redbook published in time for your needs? Yes___ No___

If no, please explain:

What other redbooks would you like to see published?

Comments/Suggestions: (THANK YOU FOR YOUR FEEDBACK!)

SG24-5750-00

Printed in the U.S.A.

Understanding SSA Subsystems in Your Environment

SG24-5750-00

