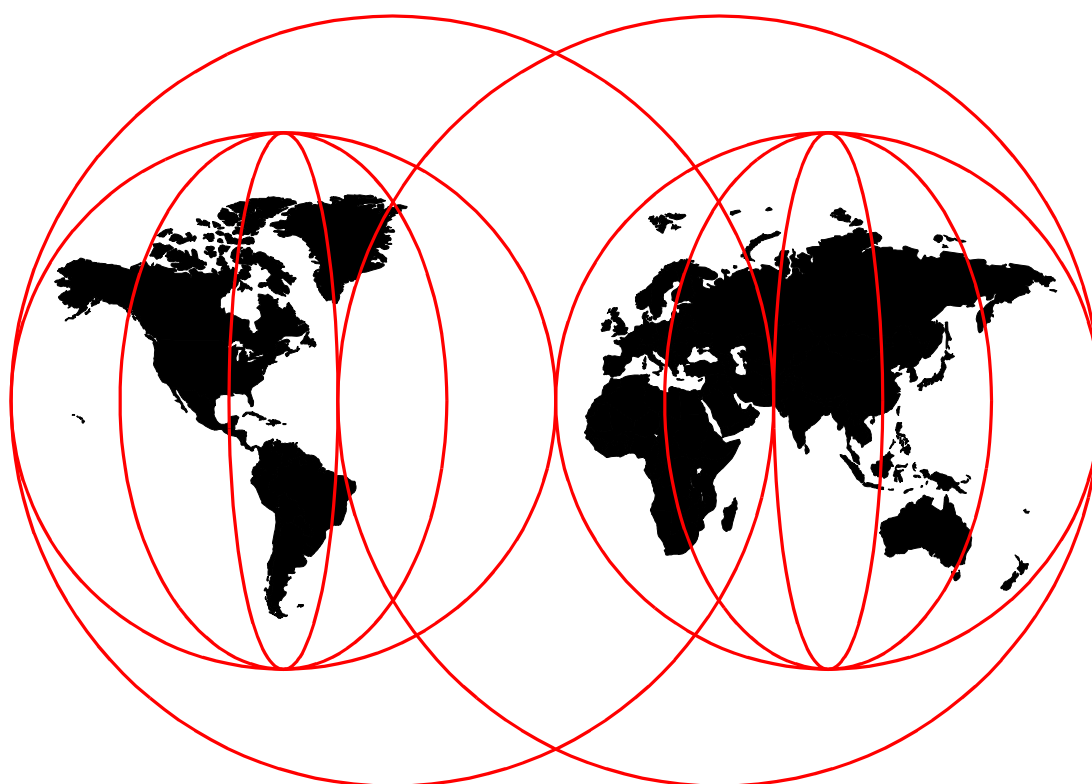




# Configuring the IBM VSS for Performance and Availability

*Mark Blunden, Alberto Chechi, SoonKwon Ha, Bill Nevill, Cesar Maciel*



**International Technical Support Organization**

<http://www.redbooks.ibm.com>





International Technical Support Organization

SG24-5279-00

**Configuring the IBM VSS for  
Performance and Availability**

December 1998

**Take Note!**

Before using this information and the product it supports, be sure to read the general information in Appendix A, "Special Notices" on page 111.

**First Edition (December 1998)**

Comments may be addressed to:  
IBM Corporation, International Technical Support Organization  
Dept. 471F 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.

# Contents

<b>Figures</b> .....	vii
<b>Tables</b> .....	ix
<b>Preface</b> .....	xi
The Team That Wrote This Redbook .....	xi
Comments Welcome .....	xii
<b>Chapter 1. Versatile Storage Server Overview</b> .....	1
1.1 Structure and Components .....	2
1.1.1 Racks .....	4
1.1.1.1 2105-B09 .....	4
1.1.1.2 2105-100 .....	5
1.1.2 Clusters .....	5
1.1.2.1 332 MHz PowerPC CPUs .....	6
1.1.2.2 Read/Write Cache .....	7
1.1.2.3 CPU Level 1 Cache .....	7
1.1.2.4 CPU Level 2 Cache .....	8
1.1.2.5 Internal SCSI adapter .....	8
1.1.2.6 SCSI Disk Drive .....	8
1.1.2.7 CD-ROM .....	8
1.1.2.8 Diskette Drive .....	8
1.1.2.9 Ethernet .....	8
1.1.2.10 Serial Ports .....	8
1.1.2.11 LCD Display .....	9
1.1.2.12 Service Processor .....	9
1.1.3 Bays .....	9
1.1.3.1 PCI Buses .....	9
1.1.3.2 Interface Adapters .....	10
1.1.3.3 Host Interface Adapter .....	11
1.1.3.4 Disk Interface Adapter .....	11
1.1.3.5 SSA Adapter Memory .....	11
1.1.4 SSA Disk Drawer .....	13
1.1.5 SSA Disk Drive Module .....	14
1.1.6 Arrays .....	15
1.1.6.1 6+P+S .....	16
1.1.6.2 7+P .....	16
1.1.6.3 524-Byte Sector Format .....	16
1.1.7 User Interfaces .....	17
1.1.7.1 Versatile Storage Specialist .....	17
1.1.7.2 Service Terminal .....	18
1.2 VSS Installation Procedures: An Overview .....	19
<b>Chapter 2. Networking with Versatile Storage Server</b> .....	21
2.1 Configuring TCP/IP .....	21
2.2 Configuring Versatile Storage Specialist .....	22
2.3 Configuring e-Mail .....	23
2.4 Configuring the Simple Network Management Protocol .....	24
2.5 Serial Connections .....	25
2.5.1 Call-Home Support .....	25
2.5.2 Remote Diagnostics and Support .....	25

<b>Chapter 3. Versatile Storage Specialist (VS Specialist)</b> . . . . .	27
3.1 Prerequisites . . . . .	27
3.2 Components . . . . .	28
3.3 Security . . . . .	29
3.4 Connecting to Your VS Specialist . . . . .	30
3.4.1 VSS Connections and Configuration Status . . . . .	30
3.5 Panels . . . . .	32
3.5.1 Introduction Panel . . . . .	32
3.5.2 Status Panel . . . . .	34
3.5.3 Problem Notification Panel . . . . .	36
3.5.4 Communications Panel . . . . .	37
3.5.5 Storage Allocation Panel . . . . .	38
3.5.5.1 Storage Allocation -- Graphical View Panel . . . . .	38
3.5.5.2 Storage Allocation -- Graphical View Panel: View Utilization . . . . .	39
3.5.5.3 Storage Allocation -- Tabular View . . . . .	41
3.5.6 Storage Allocation Tasks . . . . .	43
3.5.6.1 Modify Host Systems . . . . .	43
3.5.6.2 Configure Host Ports . . . . .	44
3.5.6.3 Add Volumes . . . . .	46
3.5.6.4 Modify Volume Assignments . . . . .	50
3.5.7 Users Panel . . . . .	52
3.5.8 Licensed Internal Code Panel . . . . .	54
<b>Chapter 4. Host Connections</b> . . . . .	57
4.1 Defining a Host . . . . .	57
4.1.1 Examples . . . . .	58
4.2 Connecting a Host to a Port . . . . .	60
4.2.1 Examples . . . . .	62
4.3 Managing Logical Volumes on the VSS . . . . .	64
4.3.1 Examples . . . . .	65
4.4 Modifying Volume Assignments . . . . .	67
4.4.1 Examples . . . . .	68
4.5 Understanding Port and Cluster Affinity . . . . .	69
4.5.1 What Is Cluster Affinity? . . . . .	69
4.5.2 Physical Structure . . . . .	70
4.5.3 Setting the Affinity . . . . .	70
4.5.4 Changing the Affinity . . . . .	71
4.5.5 Example . . . . .	73
4.6 Server Storage Attachment Examples . . . . .	75
4.6.1 IBM AIX . . . . .	75
4.6.2 Microsoft Windows NT . . . . .	75
4.6.3 Sun Example . . . . .	76
4.6.4 HP Example . . . . .	76
<b>Chapter 5. VSS Availability</b> . . . . .	77
5.1 Types of Failure . . . . .	77
5.1.1 Host Connection Failures . . . . .	77
5.1.1.1 IBM I/O Path Optimizer . . . . .	77
5.1.2 Disk Interface Failures . . . . .	78
5.1.3 Storage Processor Cluster Failures . . . . .	80
5.1.4 Power Failures . . . . .	81
5.2 Data Integrity . . . . .	82
5.3 System Availability . . . . .	82
5.3.1 Planning for Maximum Availability . . . . .	82

5.3.1.1 Minimal Disruption in Case of an Adapter Failure . . . . .	82
5.3.1.2 Minimal Disruption Caused by Maintenance Actions . . . . .	82
5.3.2 High Availability Configuration . . . . .	83
5.4 VSS Availability Test Scenarios . . . . .	85
5.4.1 Storage Relocation . . . . .	85
5.4.2 SSA Cable Failures . . . . .	86
5.4.3 Storage Cluster Failures . . . . .	87
<b>Chapter 6. Performance and Tuning . . . . .</b>	<b>89</b>
6.1 Performance Fundamentals . . . . .	89
6.1.1 4 Way SMP Clusters . . . . .	89
6.1.2 Large Cache Memory . . . . .	89
6.1.3 PCI Bus Structure . . . . .	89
6.1.4 SSA Adapters . . . . .	90
6.2 Performance Configuration Options . . . . .	90
6.2.1 Cache Size . . . . .	90
6.2.2 Cluster Load Balancing . . . . .	91
6.2.3 SCSI Interface Configuration . . . . .	91
6.2.4 SSA Adapter Configuration . . . . .	92
6.2.5 Logical Volume Configuration . . . . .	92
6.3 Performance Expectations . . . . .	95
6.3.1 Array Throughput Rules of Thumb . . . . .	97
6.3.2 Subsystem Throughput Expectations . . . . .	97
6.4 Performance Monitoring and Reporting . . . . .	98
6.4.1 IOSTAT Reports . . . . .	98
<b>Chapter 7. Sizing Logical Volumes . . . . .</b>	<b>101</b>
7.1 Logical Volume Sizes . . . . .	101
7.2 Choosing Logical Volume Sizes . . . . .	102
<b>Chapter 8. Error Notification and Recovery Procedures . . . . .</b>	<b>103</b>
8.1 Error Log and Reporting Tools . . . . .	103
8.1.1 VSS Error Logging Architecture . . . . .	103
8.1.2 Alerts Generation . . . . .	103
8.1.2.1 e-Mail . . . . .	104
8.1.2.2 SNMP . . . . .	104
8.1.2.3 Call Home Support . . . . .	104
8.2 Recovery Procedures . . . . .	104
8.2.1 Recovery Menus in the Service Terminal Interface . . . . .	104
8.2.2 Recovering from Adapter Failures . . . . .	106
8.2.2.1 Storage Relocation . . . . .	106
8.2.2.2 Adapter Replacement . . . . .	107
8.2.3 Recovering from SSA Disk Failures . . . . .	107
8.2.4 Recovering from Cluster Failures . . . . .	108
8.2.5 Recovering from Internal SMP Disk Failures . . . . .	108
8.2.6 Recovering from Network Failures . . . . .	109
8.2.7 Performing an Emergency Shutdown . . . . .	109
<b>Appendix A. Special Notices . . . . .</b>	<b>111</b>
<b>Appendix B. Related Publications . . . . .</b>	<b>113</b>
B.1 International Technical Support Organization Publications . . . . .	113
B.2 Redbooks on CD-ROMs . . . . .	113
B.3 Other Publications . . . . .	113

<b>How to Get ITSO Redbooks</b> .....	115
How IBM Employees Can Get ITSO Redbooks .....	115
How Customers Can Get ITSO Redbooks .....	116
IBM Redbook Order Form .....	117
<b>List of Abbreviations</b> .....	119
<b>Index</b> .....	121
<b>ITSO Redbook Evaluation</b> .....	125



## Figures

1. The Team That Wrote This Redbook: (l to r) Cesar, Bill, Key, Mark, and SK	xii
2. The VSS Structure and Components	3
3. Front View of a VSS 2105-B09 Rack and Two 2105-100 Racks	4
4. Front View of the Clusters	6
5. Rear View of the Clusters	7
6. PCI Buses in the VSS	10
7. SSA Adapter Memory	12
8. 7133-020 SSA Disk Drawer	14
9. 7133 SSA Disk Drawer and Two Arrays	16
10. 524-Byte Sector Format in the VSS	17
11. VS Specialist Main Window	18
12. The Main Service Menu	19
13. The VSS Network Environment	21
14. VS Specialist Main Panel	29
15. Entering the URL of Your VSS Machine	30
16. VS Specialist: Server Starting Message Panel	31
17. Logical View of the VS Specialist Panels	32
18. VS Specialist: Introduction Panel	33
19. VS Specialist: Username and Password Required Panel	34
20. VS Specialist: Status -- Graphical View Panel	35
21. VS Specialist: Status -- Problem Log Panel	36
22. VS Specialist: Problem Notification Panel	37
23. VS Specialist: Communications Panel	38
24. VS Specialist: Storage Allocation Panel; Graphical View	39
25. VS Specialist: Storage Allocation -- Graphical View: View Utilization	40
26. Disposition of Elements in the Storage Allocation -- Graphical View Panel	41
27. VS Specialist: Storage Allocation -- Tabular View Panel	43
28. Modify Host Systems Panel	44
29. Storage Allocation Tasks: Configure Host Ports Panel	45
30. Storage Allocation Tasks: Add Volumes (1 of 2) Panel	47
31. Vs Specialist: Add Volumes (1 of 2) Panel - Object Selection	48
32. Vs Specialist: Add Volumes (2 of 2) Panel	50
33. VS Specialist: Modify Volume Assignments Panel	51
34. VS Specialist: The User Administration Panel	53
35. VS Specialist: Sorting Columns in the User Administrator Panel	54
36. VS Specialist: Licensed Internal Code Panel	55
37. VS Specialist: README File for an LIC Source	56
38. Host Types List	57
39. Adding a Host	59
40. Removing a Host	60
41. Selecting a New Host Type for a Port	62
42. Adding a Host to a Port: Part 1	63
43. Adding a Host to a Port: Part 2	63
44. Removing a Host from a Port	64
45. Adding Volumes: Part 1	66
46. Adding Volumes: Part 2	66
47. Adding Volumes: Part 3	67
48. Adding a Volume to a Host Port	68
49. Removing a Volume from a Host Port	69
50. Changing Cluster Affinity: Warning Message	72

51. Changing Volume Affinity . . . . .	72
52. Changing the Cluster Affinity of a Host Port . . . . .	74
53. Implementation of Multiple Paths to a VSS Logical Volume . . . . .	78
54. Software-based Disk Mirroring . . . . .	79
55. Simplified Diagram of the Failover Process . . . . .	81
56. High Availability Configuration: Two Hosts and Mirrored Logical Volumes. . . . .	84
57. Four Hosts in a High Availability Environment . . . . .	84
58. Error Notification via e-Mail . . . . .	86
59. Status Panel after Disconnecting an SSA Drawer . . . . .	87
60. Sample Configuration for Availability . . . . .	93
61. Sample Basic Performance Configuration . . . . .	94
62. Sample Advanced Performance Configuration . . . . .	95
63. I/O Rate and Response Time for 100% Random Read Test . . . . .	96
64. I/O Rate and Response Time for 70% Random Read Test . . . . .	96
65. I/O Rate and Response Time for 50% Random Read Test . . . . .	97
66. IOSTAT Report from AIX . . . . .	99
67. IOSTAT Report from Sun OS . . . . .	99
68. IOSTAT Report from HP-UX . . . . .	100
69. Repair Menu . . . . .	105
70. Configuration Options Menu . . . . .	105
71. Utility Menu . . . . .	106
72. Storage Relocation Used to Recover from SCSI Failures . . . . .	106
73. Quiesce a Resource Option . . . . .	107
74. Simplified Diagram of the Failback Process . . . . .	108

## **Tables**

1. DDM Capacity of 4.5 GB and 9.1 GB SSA Disk Drives in an AIX Environment	15
2. VSS Installation Procedures: An Overview . . . . .	19
3. SCSI Ports: Local Affinity . . . . .	70
4. SCSI Ports: Recommended Affinity (Remote) . . . . .	70
5. VSS Cache Sizing for Less Effective Host Caching Environments . . . . .	91
6. VSS Cache Sizing for Effective Host Caching Environments . . . . .	91
7. I/O Rate per Array Rule of Thumb for Various Read to Write Ratios. . . . .	97
8. VSS Throughput Expectations for Various I/O Characteristics . . . . .	98
9. LUN Capacities Related to Their Physical and Logical Sizes . . . . .	101



---

## Preface

This redbook will help you install, tailor, and configure the new IBM Versatile Storage Server (VSS). It is written for storage administrators responsible for configuring the VSS for performance and availability.

In this book we provide an overview of the VSS and detailed information about configuring it to meet your data storage requirements for performance and availability. We present step-by-step examples of the configuration procedures to use to support those requirements. We also provide performance and tuning tips that will enable you to get the most out of your VSS subsystem.

---

## 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 coauthored five previous redbooks and managed two other redbooks. Mark also teaches IBM classes worldwide on all areas of storage. He has worked for IBM for 18 years in many areas of the IT business. Before joining the ITSO in 1998, Mark worked in Sydney, Australia, as an Advisory Storage Specialist

**Alberto Chechi** is a Advisory I/TAP System Software Specialist in Italy. He joined IBM in 1989, working on mainframe MVS operating systems and moved on to DFHSM and DFSMS. Alberto has been assigned to world wide software support several times. In 1992 he started working on ADSM V1 and has focused in this area. He was assigned to V2 and V3 ADSM's ESPs. Alberto is involved in developing customer storage solutions for DFSMS/HSM and ADSM and is currently the IBM Global Services South Region support for storage products.

**SoonKwon Ha (SK)** is an Advisory I/TAP Specialist in Korea. SK has 10 years of experience in UNIX systems and has worked at IBM for 3 years. His areas of expertise include Open Systems Storage, AIX, PSSP for RS/6000, and RS/6000 SP. He holds a degree in Computer Science. This is SK's first ITSO redbook.

**Bill Nevill** is a consulting IT Specialist in the Storage Systems Division in Canada. He has specialized in I/O subsystems and storage management for the last 11 years and has focused on open storage solutions for the last 2 years. Bill has extensive experience in I/O benchmarking and I/O architectures. He has contributed and presented to such organizations as CMG and COUG. This is Bill's first ITSO redbook.

**Cesar Maciel** is an RS/6000 Product Specialist in Brazil. He joined IBM in 1996, and has four years of experience in AIX and RS/6000. Cesar holds a degree in Electrical Engineering. His areas of expertise include UNIX systems and open systems storage. Cesar has authored papers in process control technologies, and has written extensively on availability and recovery procedures of the VSS.

Figure 1 on page xii shows the ITSO, San Jose team at the Almaden Research Center.

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

Maggie Cutler, International Technical Support Organization, San Jose Center

John Aschoff, System Performance and Design, IBM San Jose

Ruth Azevedo, System Performance and Design, IBM San Jose

Brian Smith, System Performance and Design, IBM San Jose

Dean Underwood, Open Systems Strategy and Plans, IBM Tucson

Jeff Ferrier, Device RAS, IBM San Jose

Mike Benhase, Microcode Development, IBM Tucson

David Short, System Performance and Design, IBM San Jose

Kevin Smith, System Performance and Design, IBM San Jose

Jerry Coale, Array RAS, IBM San Jose



Figure 1. The Team That Wrote This Redbook: (l to r) Cesar, Bill, Keki, Mark, and SK

---

## 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 "ITSO Redbook Evaluation" on page 125 to the fax number shown on the form.
- Use the electronic evaluation form found on the Redbooks Web sites:

For Internet users <http://www.redbooks.ibm.com>

For IBM Intranet users <http://w3.itso.ibm.com>

- Send us a note at the following address:

[redbook@us.ibm.com](mailto:redbook@us.ibm.com)

---

## Chapter 1. Versatile Storage Server Overview

It is a new world out there--a connected world at that. Information flows among departments, across divisions, throughout the enterprise and beyond as never before. With e-business, information must be available 24 hours a day, 7 days a week. Business requirements frequently change with new applications spurring an astonishingly rapid rate of data growth, and demand for information access.

Fostering access to data in an open systems world represents a sizable challenge, given the patchwork of servers and storage systems that exist within most companies today. The ability to manage and share information is a pressing concern, as administrative costs continue to climb and the management of multiple servers and data storage grows more complex every day. These concerns and the spiraling costs associated with them have already led many companies to a strategy of server consolidation--the relocation of distributed servers and related storage into data centers with centralized management provided by corporate IT departments. These companies need versatile storage solutions designed to simplify storage management, provide sharing of storage resources, enhance data sharing, and support their plans for server consolidation.

To provide server consolidation solutions IBM developed the Versatile Storage Server (VSS), a centrally managed disk storage solution for multiple UNIX, AS/400, and Windows NT based servers that is uniquely designed to allow companies to ease the work load involved in server consolidation. The VSS provides centralized storage for most servers using UNIX based operating systems, as well as most servers using Windows NT, and IBM's AS/400 server. For specific information about VSS support for UNIX based servers and Intel based PC servers see: <http://www.ibm.com/storage/vss>.

The VSS provides true data sharing for like UNIX and Windows NT based servers. With the true data sharing capabilities of the VSS and the appropriate host software to manage access, storage consolidation is also possible. For example, two IBM RS/6000 servers can share a single copy of data within the VSS.

In this book, we describe the characteristics, installation activities, configuration procedures, and performance expectations of the VSS. Essentially, there are two areas of activities, those performed by IBM Customer Service, and those performed by the customer. The sequence of activities is:

1. Provide power requirements, telephone line, TCP/IP connections  
These tasks are the responsibility of the customer.
2. Physically install the VSS  
This is job of the IBM CE.
3. Choose array configuration  
This is a choice of 6+P+S and 6+P+S or 6+P+S and 7+P. This information is provided by the customer for input by the IBM CE.
4. Define TCP/IP configuration, SNMP configuration, user IDs, e-mail definitions  
These values are defined by the IBM CE from input obtained from the customer.

5. Set up Call Home configuration  
This is done by the IBM CE.
6. Enable VS Specialist  
This is done by the IBM CE.
7. Define new Host Types  
This is done by the CE from input provided by the customer.
8. Connect host SCSI interface  
This is done by the CE from definitions provided by the customer.
9. Set up logical volumes, SCSI and host configurations  
This is done by the customer through the VS Specialist interface.

In this book we describe each of these activities in detail. In addition, we provide some information for performance configuration and monitoring.

The VSS includes the IBM StorWatch Versatile Storage Specialist (VS Specialist), which allows storage administrators to manage and configure storage from the desktop or remote locations.

In summary, VSS provides the following benefits:

- Simplification of server consolidation tasks
- Up to 2 TB of shared SSA storage for UNIX and Windows NT based servers, and IBM's AS/400.
- Data sharing for UNIX and Windows NT based servers
- High performance through the use of serial storage architecture (SSA) disk technology, intelligent algorithms, and high-speed read/write cache
- High availability using RAID-5 technology, dual active processing clusters with integrated failover, and redundant power/cooling components
- Flexible management through the VS Specialist
- Support for existing IBM 7133 SSA disk subsystems (Models 010 and 020), providing investment protection and enhanced value

The VSS is the first disk product to use the IBM Seascape architecture.

---

## 1.1 Structure and Components

In this section we describe the building blocks of the VSS (see Figure 2 on page 3).

The VSS includes the following components:

- Racks - 2105-B09 and 2105-100
- Clusters with PowerPC 604e four-way CPUs and read/write cache
- Bays
- SSA disk drawers
- Disk modules
- Arrays



- User interfaces

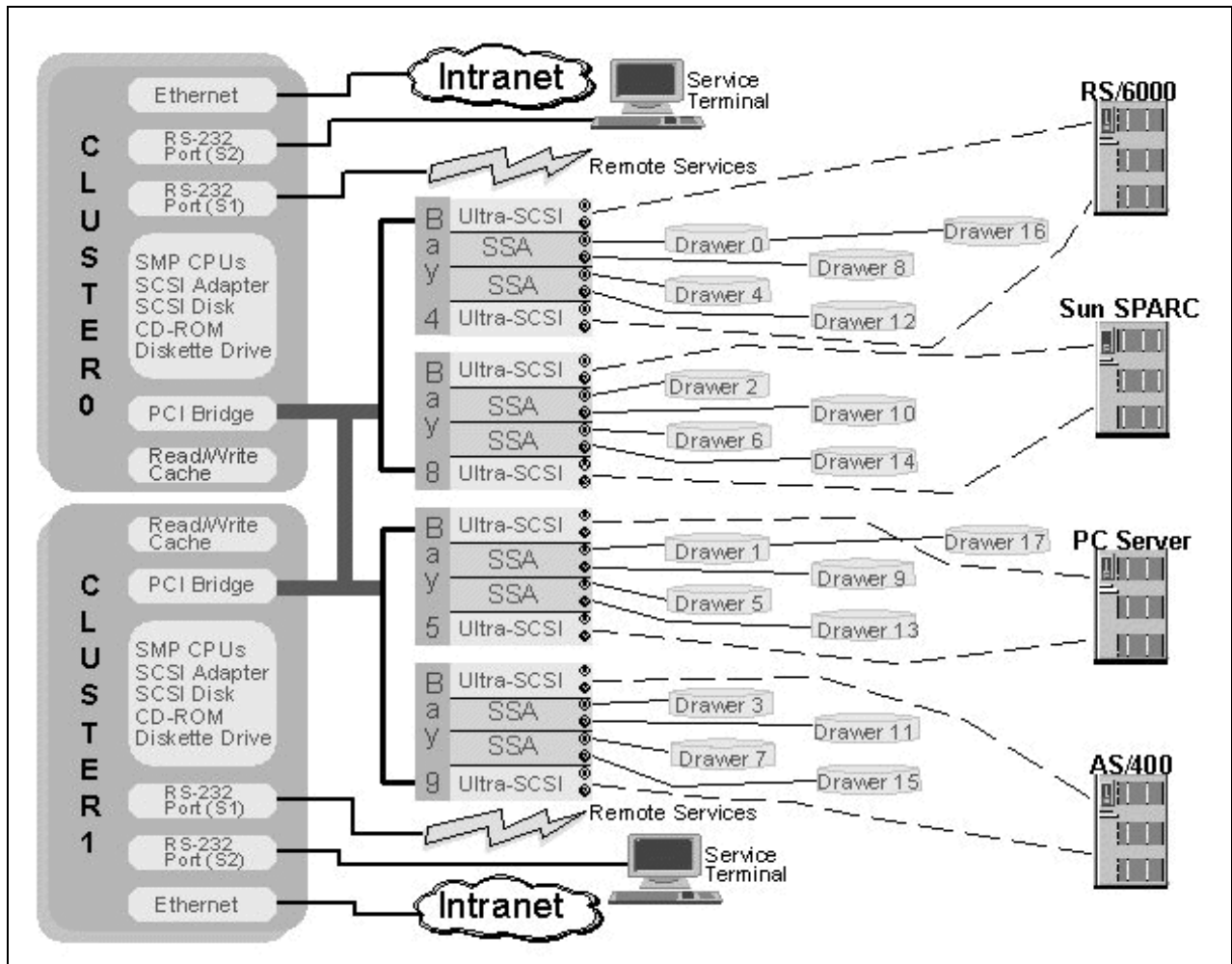


Figure 2. The VSS Structure and Components

## 1.1.1 Racks

Two types of racks are used in the VSS subsystem: 2105-B09 and 2105-100. In Figure 3 on page 4, the 2105-B09 rack is in the middle. Two 2105-100 racks are on either side of the 2105-B09 rack.



Figure 3. Front View of a VSS 2105-B09 Rack and Two 2105-100 Racks

### 1.1.1.1 2105-B09

The 2105-B09 rack is 33 in. wide but supports 19 in. drawers. It houses up to four SSA disk drawers, a fully redundant power control subsystem, and a storage server containing two clusters with read/write cache and four adapter bays. The four SSA disk drawers provide usable storage of 456 GB (if all drives are 9.1 GB in size).

Power is supplied to the 2105-B09 rack through two power cords carrying either 50 amp single-phase or 50/60 amp three-phase alternating current (AC) power, with each cord providing 100% of the required power.

There are two 350 V direct current (DC) bulk power supplies. One of these is enough to supply power to the subsystem, which provides for uninterruptable service if one of the DC power supplies fails. A 48 V DC transformer on each 350 V rail drops the voltage to the correct level for the individual components such as the clusters and the adapter bays. The SSA disk drawers are supplied with 350 V DC provided by the bulk power supplies. Because the drawers must have at least two of their three power supplies available to power up, each drawer has a single connection to each 350 V rail, and a Y connection that connects to both rails. In the event that one of the 350 V bulk power supplies fails, the disk drawers can still power up.

In case of emergency, an emergency power off (EPO) switch is located on the rack, as well as *power* and *service required* indicators.

To guard against AC voltage sags and short outages, an optional battery can be fitted to the cabinet, providing 5 to 15 minutes of power, depending on the configuration, in the event of AC power loss. The battery is not meant to provide uninterruptable service indefinitely. The storage servers will not initiate data destage or shutdown procedures in the event of an AC power loss for longer than the battery can provide backup. The loss of power does not produce a data integrity issue, as any unchanged data not destaged to disk is saved in the nonvolatile fast write cache on the SSA adapter. Typically, the host initiates shutdown procedures, which in turn ensure destaging of any modified data.

Power cables for the four disk drawers are included in the base rack, regardless of how many drawers are initially configured in the rack.

#### **1.1.1.2 2105-100**

The 2105-100 rack is similar in configuration to the 2105-B09, except that it does not have the storage server that houses the clusters and adapter bays. Instead, three more SSA disk drawers can be installed in the space not occupied by the storage server. Seven SSA disk drawers in total can be installed in the 2105-100 rack, providing a maximum usable storage of 800 GB.

A subsystem always consists of one 2105-B09, and optionally up to two 2105-100 racks. Existing 7014-S00, 7015-R00, and 7202-900 racks can also be used as expansion racks to a 2105-B09 rack, but they do not support the power sequencing features and redundancy function of the 2105-B09.

The 2105-100 racks are placed next to the 2105-B09 rack, to allow connection of the power sequencing cables (the frames are bolted together). If existing 7015-R00 or 7202-900 racks are required in the subsystem, they should be placed next to the 2105-100 or 2105-B09 racks.

Where there are only 2105-100 expansion racks in the subsystem, they may be placed up to 20 m (65 ft) away from the 2105-B09 if the appropriate RPQ is specified (the RPQ supplies the cable that enables the 20 m distance). The standard configuration has the 2105 racks bolted together.

The 2105-100 rack contains the same fully redundant power subsystem as the 2105-B09 rack. Power is supplied to the 2105-100 rack through two power cords carrying either 50 amp single-phase or 50/60 amp three-phase AC power, with each cord providing 100% of the needed power.

The 2105-100 rack has no operator power on-power off switch. Power is sequenced from the 2105-B09 rack. For emergencies, an EPO switch is located on the rack, as well as *power* and *service required* indicators.

The 2105-100 rack also has an optional battery that provides the same function it provides for the 2105-B09 rack--5 to 15 minutes of backup to facilitate host shutdown.

### **1.1.2 Clusters**

The VSS has two independent symmetric multiprocessor (SMP) clusters (cluster 0 and cluster 1). Under normal conditions, each cluster manages approximately

half of the attached storage. For most cluster failures, the VSS routes all transactions to all the drives through the remaining cluster.

The clusters provide management functions for the VSS and memory for cache. The clusters also manage the interfaces to host system devices, external local area networks (LANs), and remote services (call home) modems.

The cluster cache and the nonvolatile storage provide enhanced performance. When the host system sends data, the VSS initially stores it in the cluster cache. Then the VSS transfers the data to nonvolatile storage on the SSA adapter. The SSA adapter destages data to the disk drive module (DDM).

Each cluster includes an internal SCSI disk drive. The disk drive provides storage for Licensed Internal Code (LIC) and error logs.

Figure 4 on page 6 shows the front view of the clusters, and Figure 5 on page 7 shows the rear view of the clusters.

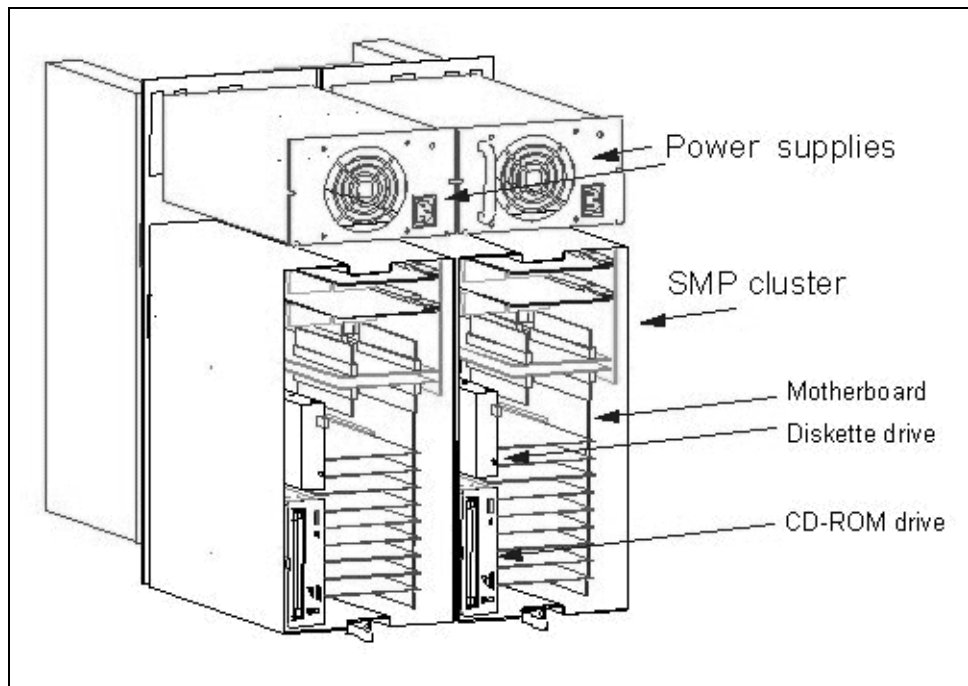


Figure 4. Front View of the Clusters

#### 1.1.2.1 332 MHz PowerPC CPUs

Each cluster uses four PowerPC 604e CPU processors, running at a clock speed of 332 MHz, to provide high-performance computing. These processors are produced using 0.5 micron lithography technology. The 0.5 micron lithography process allows for efficient, high-density packaging. The 604e chip contains approximately 5.1 million transistors. It also uses split voltages of 2.5 V DC for

core functions and 3.3 V DC for I/O. Split voltages greatly reduce power consumption.

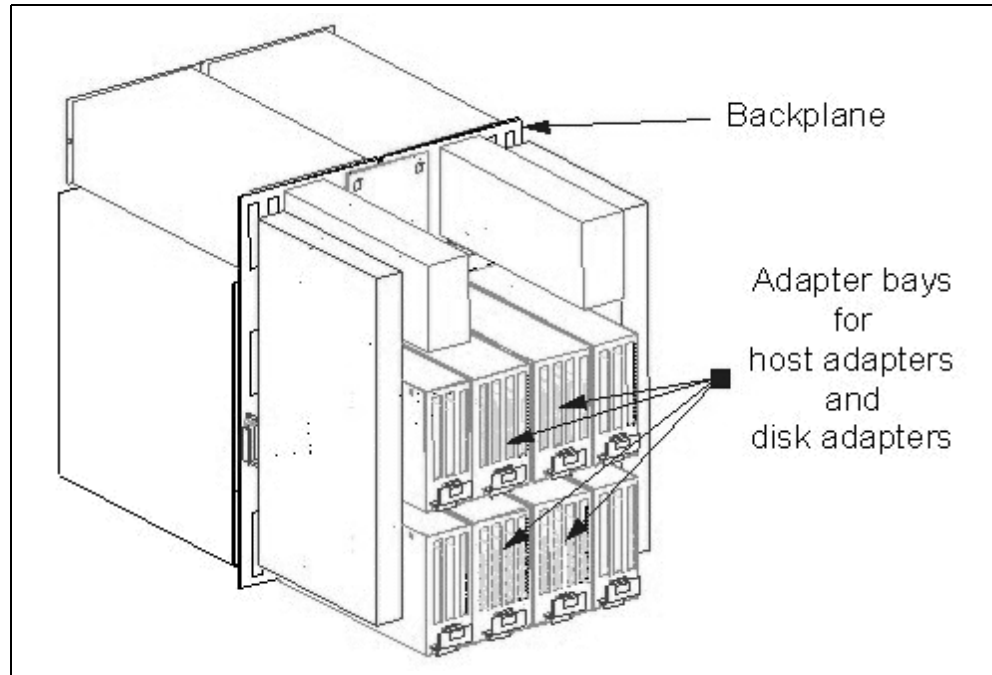


Figure 5. Rear View of the Clusters

#### 1.1.2.2 Read/Write Cache

Each cluster can be configured with up to 3 GB of ECC synchronous dynamic random access memory (SDRAM) as cache for temporary storage of read and write data from the host and the VSS disk drives. A combined total of up to 6 GB of cache memory can be installed per VSS subsystem.

The basic VSS subsystem is configured with 256 MB of cache per cluster, or 512 MB in total. Total cache options are 512 MB, 1 GB, 2 GB, 4 GB, or 6 GB.

The cache is used as read and write cache but is considered volatile; that is, it has no long-term battery backup.

If an application or the host operating system requires the read after write function to verify data integrity, the read request is satisfied from the cluster cache, thus increasing the performance of this typically performance degrading procedure.

The ECC SDRAM provides for single-bit error correction and double-bit error detection. When data is written to memory, the controller calculates check bits and stores those along with the data. On a read operation, the controller calculates the check bits again and compares them with the originals. Therefore if a single bit of the memory location goes bad, the memory controller reconstructs the missing data. If two bits of the memory location go bad, the memory controller detects this and generates a check condition.

#### 1.1.2.3 CPU Level 1 Cache

Each CPU has 64 KB of level 1 (L1) cache internally. This four-way cache is set-associative split into 32 KB of instruction cache and 32 KB of data cache.

The instruction cache can provide up to four instruction executions in a single clock cycle. The RISC architecture defines special instructions for maintaining instruction cache coherency. The 604e processor implements these instructions.

The data cache is also four-way set-associative, with hardware supported write-back for reloading on cache misses. To ensure cache coherency in a multiprocessor configuration, the modified/exclusive/shared/invalid (MESI) protocol is implemented on a per cache-block basis.

#### **1.1.2.4 CPU Level 2 Cache**

Each CPU has 256 KB of level 2 (L2) cache externally. This cache consists of error correcting code (ECC) memory and is eight-way set-associative. L2 cache provides an extension of the L1 cache and improves overall performance by caching more data and instructions.

#### **1.1.2.5 Internal SCSI adapter**

Each cluster has an internal 16-bit, fast SCSI adapter used for controlling the internal SCSI disk drive and CD-ROM drive.

#### **1.1.2.6 SCSI Disk Drive**

Each cluster contains an internal SCSI disk drive that is used to store its operating code and configuration data. It also contains the boot loader, diagnostics, and the reliability, availability and serviceability (RAS) code. The cluster boots from the internal disk drive during its initialization sequence and uses it during operation for error logging and paging.

#### **1.1.2.7 CD-ROM**

Each cluster contains an 8X speed CD-ROM that is used to initially load and install the operating code. The CR-ROM is also used to load diagnostics and microcode or operating code updates.

#### **1.1.2.8 Diskette Drive**

A 1.44 MB 3.5 in. diskette drive is installed in each cluster. The diskette drive is used for loading microcode and updates, when a diskette is the preferred distribution medium. It is also used for saving and loading configuration data. The diskette allows for transfer of configurations between clusters and for offsite storage of configuration data.

#### **1.1.2.9 Ethernet**

A 10 Mbit/s Ethernet adapter is installed in each cluster. The Ethernet adapter is used to connect to the customer's intranet (through 10BaseT) and is the primary interface for configuring the VSS subsystem.

#### **1.1.2.10 Serial Ports**

Two RS-232 serial ports per cluster are used for access by service representatives.

One port is connected to a modem to enable remote diagnostic access and call home capability. Call home enables the VSS to contact the IBM service center in the event of a failure or if the subsystem requires attention by a Customer Engineer (CE).

The second port is used by the CE for maintenance and repair actions to the subsystem.

#### **1.1.2.11 LCD Display**

A multidigit liquid crystal display (LCD) is used to display status codes while the cluster boots and while the storage server is operational. Typically, the status codes provide information to the CE when performing maintenance and repair actions on the VSS.

#### **1.1.2.12 Service Processor**

The service processor provides a means of low-level diagnosis of the storage server, even when the main CPUs are inoperable. The service processor allows access to cluster, nonvolatile random access memory (NVRAM), vital product data (VPD), and error data. The service processor can also be used to call home when the main CPUs are not running.

### **1.1.3 Bays**

Each cluster connects two bays through PCI bridges. Bay 4 and bay 8 belong to cluster 0, and bay 5 and bay 9 belong to cluster 1. Each bay has a PCI 4-slot planar and supports up to two Ultra-SCSI adapters for host interface and two SSA adapters for SSA disk interface.

The cross-cluster interconnection is built between the clusters and the bays. It supports the cluster failover if a problem is detected.

#### **1.1.3.1 PCI Buses**

The PCI local bus was developed by IBM and other companies to create an industry standard local bus that could be used across various processor architectures as well as support the architectures as they evolved. PCI provides significantly improved performance over other common bus architectures, for example, by being able to complete a 4-byte write in as few as two bus clock cycles.

PCI is a clock synchronous bus that operates at 32 MHz. It provides a combination address and data bus that is multiplexed. The basic transfer operation for the PCI bus is a burst operation that enables a contiguous block of data to be transferred on the bus immediately after the address.

Each cluster has two 4-byte PCI buses running at 133 MB/sec.

Figure 6 on page 10 shows PCI buses in the VSS.

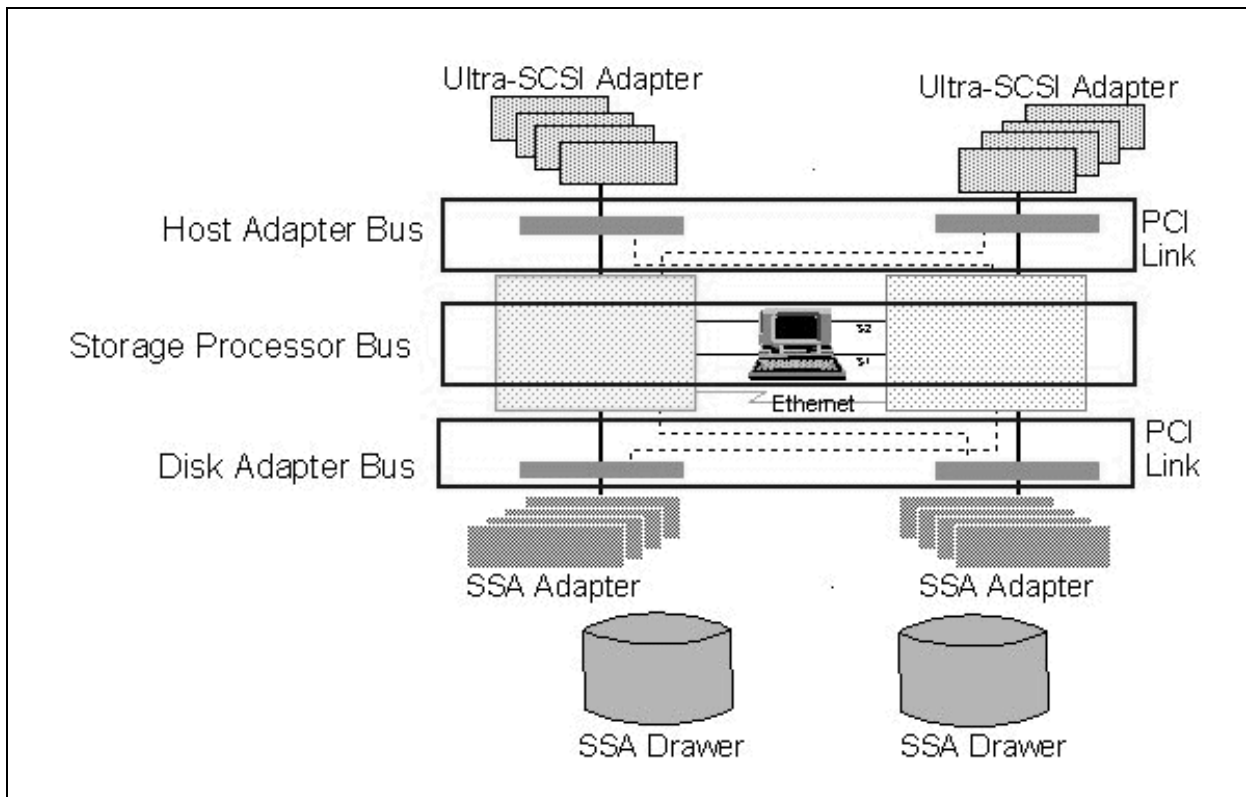


Figure 6. PCI Buses in the VSS

In essence, there are three interconnected buses:

- Host adapter bus

The host adapter bus is a PCI bridge, into which host interface adapters can be plugged. As new and faster technologies emerge, the new host interface adapters can be plugged.

- Storage processor bus

The storage processor bus is a backplane into which the cluster adapters are plugged. In the VSS the storage server is based on a RS/6000 system planar with four-way PowerPC 604e SMP processors. As RISC technology changes, the planars can either be upgraded or changed and new adapters inserted into the backplane. The storage server contains the system read/write cache, (minimum size 512 MB, maximum size 6 GB). As well as controlling the disk I/Os, the storage server carries out availability and configuration management.

- Disk adapter bus

The adapter cards used are similar to the PCI SSA RAID adapters used in RS/6000 products. This allows for easy adaptation to SSA adapter technology enhancement.

### 1.1.3.2 Interface Adapters

The interface adapters installed in the VSS connect to both hosts and SSA disks. They are installed in the hot insertion and removal adapter bays, which are



located in the rear of the 2105-B09 rack. Up to 16 adapters can be installed: 8 host interface adapters and 8 disk interface adapters.

### 1.1.3.3 Host Interface Adapter

The host interface adapter is a 32-bit PCI Ultra-SCSI adapter. It supports the SCSI-3 protocol and command set and provides two 16-bit-wide, 40 MB/s differential channels, capable of reading and writing concurrently. Each channel of the host interface adapter can support up to 16 target IDs like any 16-bit SCSI bus, and also supports up to 64 logical unit numbers (LUNs) per target ID (in conformance with the SCSI-3 standard).

Each of the channels can be attached to a single host, two homogeneous hosts, or two heterogeneous hosts. The hosts must attach either through an Ultra-SCSI differential interface or a SCSI-2 fast and wide differential interface.

Depending on the host system to which it is attached, the VSS can emulate one or more existing device types for data storage:

Host System	Device Emulated
RS/6000	2105 Model B09
HP-9000	Generic SCSI
Sun SPARC	Generic SCSI
AS/400	9337 Model 580 and 590
Data General	Generic SCSI
Windows NT	Generic SCSI

Host system types and volume sizes that the VSS supports are:

Host System Type	Volume Size (GB)
AS/400 V3R1, V3R6	4.19
AS/400 V3R2, V3R7 or later	4.19, 8.59
All other host systems	0.5, 1, 2, 4, 8, 12, 16, 20, 24, 28, 32, 40, 48, 56

### 1.1.3.4 Disk Interface Adapter

The disk interface adapter is a 32-bit PCI SSA RAID-5 adapter. This adapter is similar to the 6215 Enhanced PCI SSA-EL adapter used in RS/6000 products but has a special firmware load for use in the VSS.

The disk interface adapter supports two separate loops of SSA disks. As a general VSS configuration rule, 16 SSA disks are supported per loop, because 16 SSA disks are supported in one SSA disk drawer. In a maximum configuration of 18 SSA disk drawers, however, the first two adapters support two loops each with support for 32 SSA disk drives in two SSA disk drawers.

### 1.1.3.5 SSA Adapter Memory

The SSA adapter contains both volatile and nonvolatile memory to assist in caching, staging, destaging, and RAID parity calculations.

Figure 7 on page 12 shows a simple diagram of SSA adapter memory.

### **Volatile Storage**

The adapter contains 32 MB of DRAM, which is used for transferring data between the SSA disk arrays and the cluster read/write cache. The SSA adapter DRAM is the only place where RAID parity data is stored once it is staged from disk. Parity data is never transferred to the VSS cluster read/write cache. It is cached in adapter DRAM to assist in calculation of the new parity data if the data is modified. The DRAM is also used for storing the SSA adapter microcode and its code control blocks and a mirror of the SSA adapter's fast write cache. The DRAM is not partitioned in a physical sense--specific amounts are not set aside for read cache, fast write cache, adapter microcode, or control blocks. The size of the read cache depends on the amount of unused DRAM at any given time.

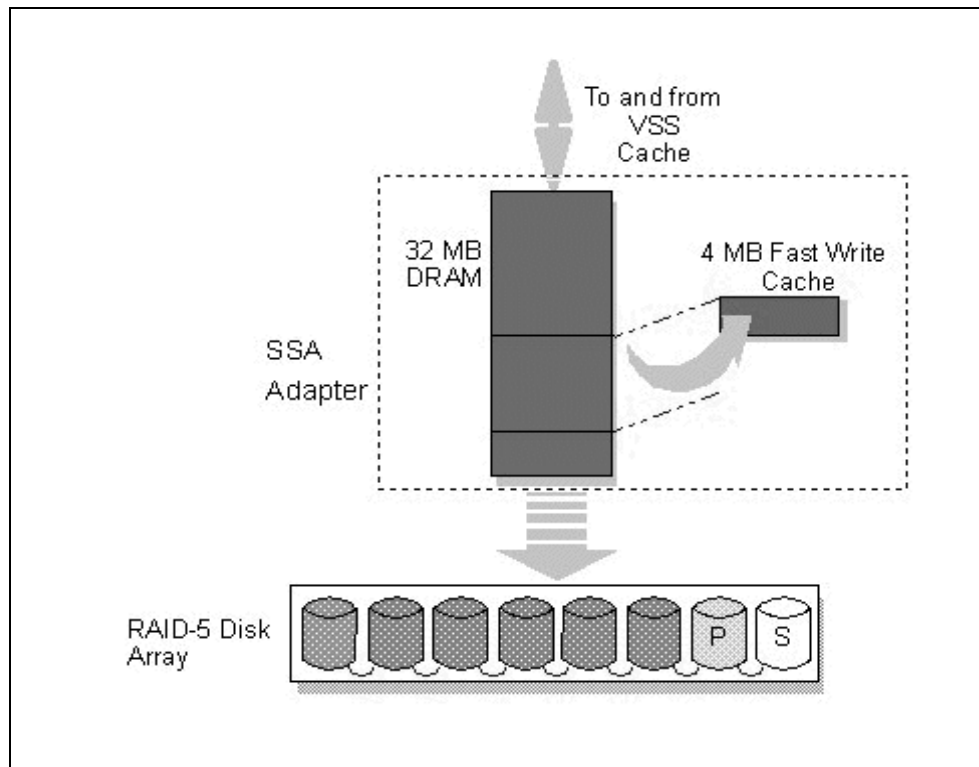


Figure 7. SSA Adapter Memory

### **Nonvolatile Storage (Fast Write Cache)**

To prevent data loss in the event of power outages or cluster failure, the SSA adapter contains 4 MB of SRAM. The synchronous random access memory (SRAM) has a lithium battery backup and is designed to maintain the data for 10 years. SRAM uses far less power than DRAM to maintain data integrity. When data is written from the cluster cache to the adapter DRAM, it is copied into the fast write cache to guard against data loss before the data is actually written to SSA disk. As soon as the fast write cache has a copy of the data, the SSA adapter signals the cluster that the write operation is complete, and in turn the cluster signals the host that the write is complete. Fast write cache provides a significant increase in performance over the usual practice of the host having to wait until the data is committed to disk. Without fast write cache it would not be possible to perform this fast write function and maintain data integrity.

In some circumstances, however, it may be undesirable to copy data into the fast write cache. The following conditions cause data to bypass the fast write cache:

- Full stripe write

If a full stripe of data is written from the VSS read/write cache, it is beneficial to simply calculate parity and write straight to disk before signaling write complete without suffering from the RAID-5 write penalty.

- Transfer size larger than full stripe or larger than fast write cache size

If a write from the VSS read/write cache is larger than a full stripe, or larger than the fast write cache size (4 MB), parity is calculated, and the data is written straight to disk to avoid flooding the fast write cache.

#### 1.1.4 SSA Disk Drawer

The 7133 SSA disk drawer can contain up to 16 SSA disk drives. The slots support hot insertion and removal of disk drives. If SSA disk drives are not installed in drive slots, a dummy drive must be installed to propagate the signal from adjoining slots. A maximum of three consecutive dummy slots is supported.

In the VSS subsystem, only one loop of 8 or 16 SSA disk drives per SSA disk drawer is supported, so a drawer is installed in the VSS as one loop.

Two types of SSA disk drawers are supported in the VSS subsystem: the 7133-010 and the 7133-020. Although the 7133-010 is no longer available, customers who already have 7133-010 drawers can use them in the VSS subsystem. See announcement of 7133-D40 for additional support. Figure 8 on page 14 shows 7133-020 SSA disk drawer.

There are some considerations when using existing SSA disk drawers--all disks within the drawers must be reformatted to the VSS specification. Thus the data cannot be migrated directly from existing RAID drawers to the VSS subsystem.

An SSA disk drawer has three separate slots for power supplies and cooling fans. In the VSS subsystem, for high availability all three power supply and cooling fan slots must be populated.

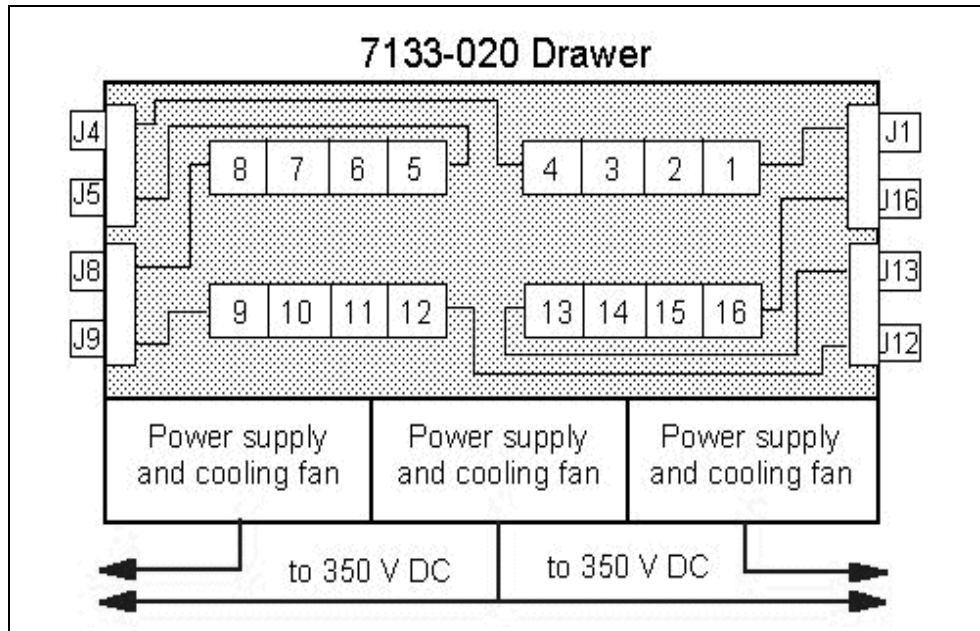


Figure 8. 7133-020 SSA Disk Drawer

The 7133-010 only supports 4.5 GB SSA disk drives in the VSS.

The 7133-020 is similar in design to the 7133-010, but it has some upgraded features and supports 9.1 GB SSA disk drives, due to improved cooling air flow within the drawer, in addition to the 4.5 GB SSA disk drives.

The 7133-010 and 7133-020 differ slightly in design. The 7133-020 has the following enhancements:

- Automatic host bypass circuits
- Third power supply module is standard
- Redesigned front cover and on-off switch
- Different cable connector numbering

### 1.1.5 SSA Disk Drive Module

Two types of SSA disk drives are supported in the VSS subsystem: 4.5 GB and 9.1 GB SSA disk drives.

Table 1 on page 15 shows the usable capacity of 4.5 GB and 9.1 GB SSA disk drive arrays in an AIX environment. Capacity is based on 1 MB = 1,000,000 bytes.

DDM Size (GB)	DDM Capacity (GB)	Spare DDM per Drawer	Usable DDM Capacity Group 1 (GB)	Usable DDM Capacity Group 2 (GB)
4.5	4.4	1	25.71	29.99
4.5	4.4	2	25.71	25.71
9.1	8.89	1	52.62	61.39
9.1	8.89	2	52.62	52.61

Table 1. DDM Capacity of 4.5 GB and 9.1 GB SSA Disk Drives in an AIX Environment

Both SSA disk drives come in an industry-standard 3.5 in. form factor package, which is compact and robust.

The SSA disk drive module contains all of the hardware to physically connect the disk drive to the backplane of the 7133 SSA disk drawer. The module provides power and signals to the disk drive. The module is designed to be installed and removed while the 7133 is powered on and running, also known as *hot pluggable*.

### 1.1.6 Arrays

An array is defined as a number of disks connected in a single loop that are part of a RAID configuration. In the VSS, two types of arrays are supported. Each array contains eight disks in a RAID-5 configuration, but the function of the disks varies slightly in the two configurations. The two types are the six-drives-plus-parity-plus-spare (6+P+S) array, and the seven-drives-plus-parity (7+P) array. Figure 9 on page 16 shows a typical configuration of a 7133 SSA disk drawer and two arrays.

As a single loop of SSA disks in a VSS can contain up to 16 drives, one or two arrays are supported per loop. For each drawer, the first array must be a 6+P+S array. The other can be either a 6+P+S array, or a 7+P array.

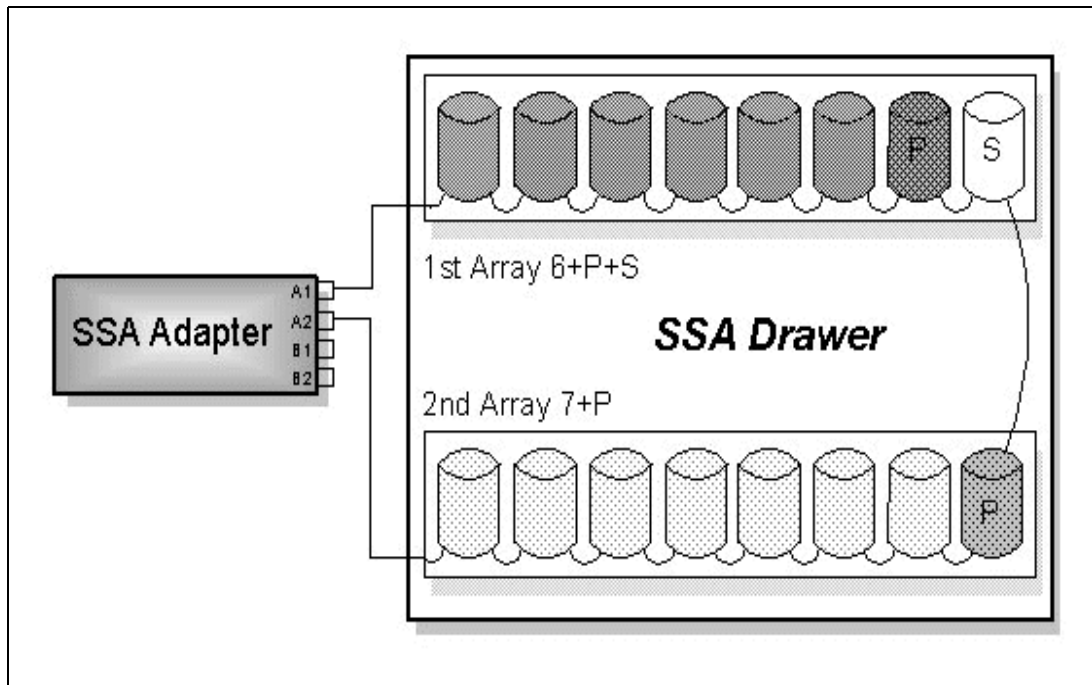


Figure 9. 7133 SSA Disk Drawer and Two Arrays

#### 1.1.6.1 6+P+S

The 6+P+S array contains six disks used to store customer data, one parity drive, and one spare drive. In the SSA RAID-5 implementation, parity is not stored on one single drive but is written on a track basis rotating across all drives in the array. The spare drive is required to support a single loop in case of drive loss during normal operations. The adapter can detect the loss of a drive and perform an automatic reconstruction of the data for the lost disk. The reconstructed data is then written to the spare, allowing the defective disk to be replaced as soon as possible. The spare drive must be the same size as the data and parity disks.

#### 1.1.6.2 7+P

The 7+P array is used for storing customer and parity data and can only exist in a loop that already has a 6+P+S array.

#### 1.1.6.3 524-Byte Sector Format

Most fixed-block disk architectures use a fixed-byte sector of 512 bytes. Most UNIX systems, including AIX, use a fixed-byte sector of 512 bytes. When used in a VSS subsystem, an SSA disk drive has a fixed-byte sector of 524 bytes (see Figure 10 on page 17). The 524-byte sector format enables the same array to contain volumes attached to UNIX, Windows NT, and AS/400 hosts without the requirement to reformat.

The data portion of the sector remains at 512 bytes, for all systems except the AS/400, which uses 520 bytes. A 2-byte sequence number and a 2-byte longitudinal redundancy check (LRC) increase the size of the sector to 524 bytes. The sequence number is a modulo-64,000 value of the logical block address

(LBA) of this sector and is used as an additional method of ensuring that the correct block is being accessed. The LRC, generated on the SCSI host interface adapter, is calculated on the 520 data and header bytes and is used as an error checking mechanism as the data progresses from the host, through the VSS storage server, into the SSA adapter, and on to the RAID array. The LRC is also used as an error checking mechanism as the data is read from the array and sent to the host adapter. The sequence number and LRC are never transferred to the host system.

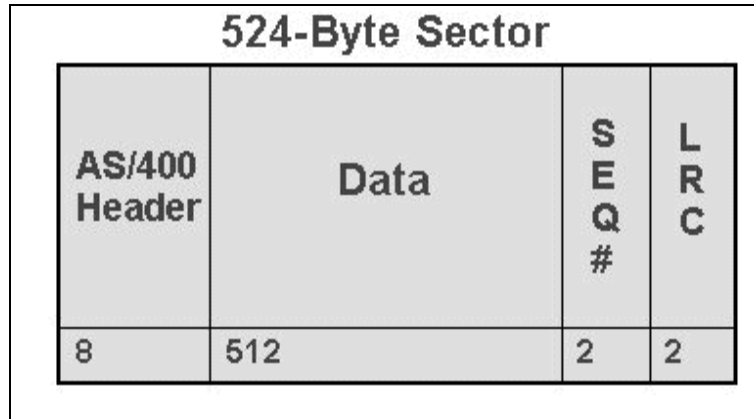


Figure 10. 524-Byte Sector Format in the VSS

## 1.1.7 User Interfaces

In this section we describe the two available user interfaces: the VS Specialist and the Service Terminal.

### 1.1.7.1 Versatile Storage Specialist

The VS Specialist is a Web-based, integrated storage management tool that enables you to monitor and manage the VSS subsystem through an Internet browser, such as Netscape Navigator, Microsoft Internet Explorer, or Sun Hot Java (the browser must support Java 1.1.2). You can access the VSS from the desktop or other locations supported by the network.

With the VS Specialist you can view and update the configuration and monitor error logs.

Figure 11 on page 18 shows the main window of the VS Specialist.



Figure 11. VS Specialist Main Window

### 1.1.7.2 Service Terminal

The VSS has service interface ports for external connection of the service terminal.

The service representative uses the service terminal to perform the following tasks:

- Install or remove hardware features
- Format the SSA disk arrays
- Define user IDs
- Define TCP/IP configuration
- Define telephone numbers and e-mail addresses
- Define Simple Network Management Protocol (SNMP) routing
- Make repairs
- Update LIC

This interface can also be used to perform system management functions normally performed through the VS Specialist interface. Some of these functions are:

- Configure host system connections and disk arrays
- Add VSS logical volumes



Figure 12 on page 19 shows the Main Service Menu of the service terminal.

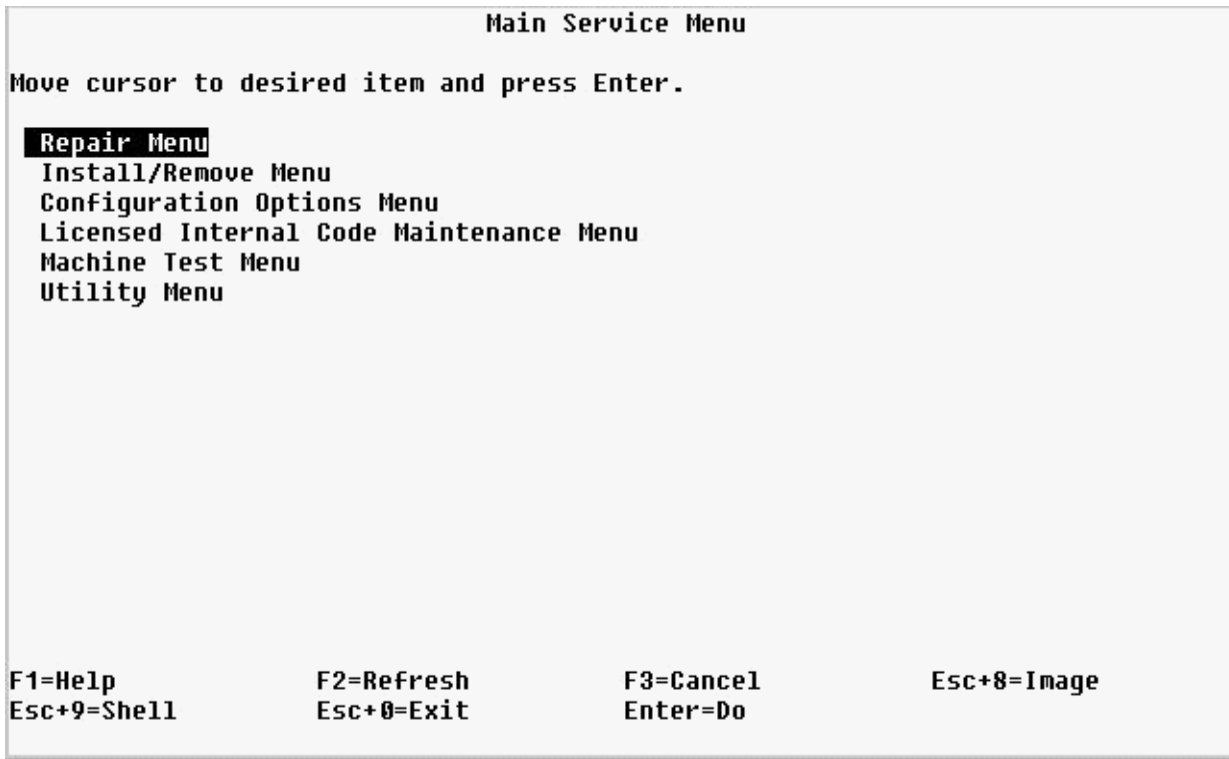


Figure 12. The Main Service Menu

## 1.2 VSS Installation Procedures: An Overview

Table 2 on page 19 provides an overview of the installation procedures. In subsequent chapters we describe in detail how to configure the VSS for optimum performance.

Table 2. VSS Installation Procedures: An Overview

Description	Things to Do	By whom
Preinstallation	Fill in the configuration worksheets in the <i>IBM Versatile Storage Server Introduction and Planning Guide</i> : <ul style="list-style-type: none"> <li>- Communication resource worksheet</li> <li>- Storage resource worksheets</li> <li>- System attachment resource worksheets</li> </ul>	IBM CE and customer
VSS hardware installation	Install the VSS hardware and verify the VSS installation	IBM CE
Service terminal connection	Connect the service terminal to the VSS	IBM CE and customer
VSS network configuration	Configure the following: <ul style="list-style-type: none"> <li>- TCP/IP</li> <li>- VS Specialist</li> <li>- e-mail</li> <li>- SNMP</li> <li>- Call home and remote services</li> </ul>	IBM CE and customer

Description	Things to Do	By whom
Storage allocation	Do the following to allocate storage to the host system that attaches to the VSS: <ul style="list-style-type: none"> <li>- Format SSA disk drawer</li> <li>- Add arrays</li> <li>- Set the cluster affinity for each SCSI port</li> </ul>	IBM CE and customer
Host system attachment	Do the following to attach storage in the VSS to the host system: <ul style="list-style-type: none"> <li>- Define host system</li> <li>- Configure SCSI port</li> <li>- Define logical volume and logical volume size</li> <li>- Specify number of logical volumes</li> <li>- Connect logical volume to SCSI LUN</li> </ul>	Customer

## Chapter 2. Networking with Versatile Storage Server

In this chapter we describe how to configure TCP/IP, the VS Specialist, e-mail, and the SNMP on the VSS and explain the purpose of serial connections for the VSS. Figure 13 on page 21 shows the network environment of the VSS.

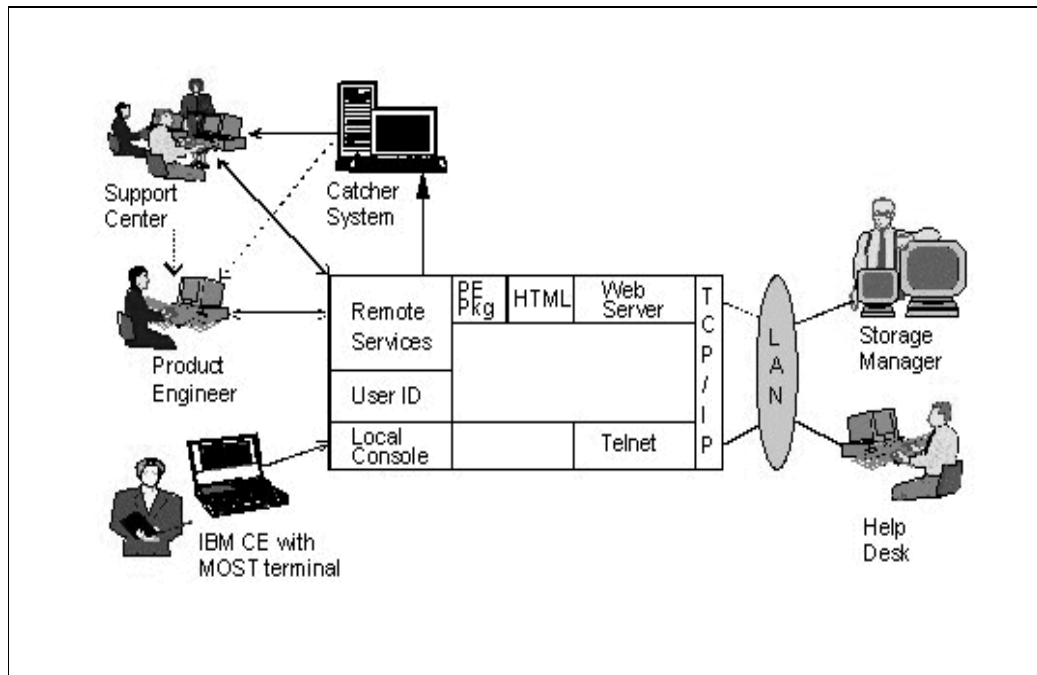


Figure 13. The VSS Network Environment

**Note:** Before configuring the VSS network environment, fill in the communication resource worksheet in the *IBM Versatile Storage Server Introduction and Planning Guide*.

### 2.1 Configuring TCP/IP

The VSS has a 10 Mbit/s Ethernet adapter in each SMP cluster. The Ethernet adapter is primarily used for service communication between clusters. It also connects to the intranet (through 10BaseT) and is the interface for configuring the VSS subsystem through the VS Specialist.

The IBM CE configures the TCP/IP environment at the service terminal, using the required information supplied by the customer, by following these steps:

1. Select network interface

Select either the en0 Standard Ethernet Network Interface or the et0 IEEE 802.3 Ethernet Network Interface. The difference between these two interfaces is the header format. On the same local network, all stations must use the same header format. Consult with your networking group for the correct header format.

2. Specify cluster 0 hostname

Enter the hostname that you are assigning to cluster 0 for access over the network. For example: tarp98c0.sanjose.ibm.com.

**Note:** The hostname you define here is also the URL name for access to the VS Specialist in cluster 0 or cluster 1.

3. Specify cluster 0 Internet address

Enter the dotted decimal address you are assigning to cluster 0 of this machine. For example: 9.113.24.232.

4. Specify cluster 1 hostname

Enter the hostname that you are assigning to cluster 1 for access over the network. For example: tarp98c1.sanjose.ibm.com.

5. Specify cluster 1 Internet address

Enter the dotted decimal address you are assigning to cluster 1 of this machine. For example: 9.113.24.231.

6. Specify network mask

Enter the dotted decimal network mask that you are applying to TCP/IP on this machine. For example: 255.255.255.0.

7. Specify nameserver Internet address

Enter the dotted decimal address of the nameserver for this machine to access. For example: 9.113.42.250.

8. Specify nameserver domain name

Enter the domain name of the nameserver. For example: sanjose.ibm.com.

9. Specify default gateway address

Enter a dotted decimal or symbolic name as the address of the default gateway for this machine. For example: 9.113.24.11 or sanjosegate.sanjose.ibm.com.

10. Select Ethernet adapter mode

Select either half-duplex or full-duplex mode for the Ethernet adapter. The mode must match the mode of the external switch or router to which you are connecting the adapter.

---

## 2.2 Configuring Versatile Storage Specialist

The VS Specialist enables you to manage and access storage from the desktop or remote locations on your LAN with a Web browser, such as Netscape Navigator, Microsoft Internet Explorer, or Sun Hot Java (the browser must support Java 1.1.2).

Using the VS Specialist you can perform the following tasks:

- Monitor error logs
- View and update the configuration:
  - Add host systems or delete host systems
  - Configure host ports
  - Add volumes
  - Assign volumes to be accessible to more than one host system

- Modify volume assignments
- View communication resource settings, such as TCP/IP configuration and users
- View cluster LIC levels

When the CE configures the VSS at the service terminal, he or she has these options:

- Enable VS Specialist connection in the VSS

The VSS contains a Web server that you can access over the network. To enable or disable the Web server, the CE either checks "enable" to permit access or "disable" to deny access.

- Add VS Specialist users

Identify each user you are adding with a unique alphanumeric character name. Select an authorization level for each user:

- Viewer

A viewer can view the current configuration and status information.

- Configurator

A configurator can view the current configuration and status information and can make changes to the configuration.

- Administrator

An administrator can define new user IDs, delete old IDs, and assign, change, and revoke passwords and levels of authorization.

Enter a VS Specialist password of one to eight alphanumeric characters for each user.

- Remove a VS Specialist user. To remove a VS Specialist user, identify each user you are removing by their one to eight alphanumeric characters name.

For detailed information about using the VS Specialist, refer to Chapter 3, "Versatile Storage Specialist (VS Specialist)" on page 27.

---

## 2.3 Configuring e-Mail

The VSS generates two types of e-mail messages: error notifications and information messages.

The VSS sends error notifications when it detects a situation that requires correction. Information messages provide information such as:

- Notification when the VSS receives a new level of LIC
- Notification when the service representative runs diagnostic tests to verify that your addresses are receiving e-mail messages
- Notification of configuration changes

The CE can set up an e-mail environment at the service terminal by following these steps:

1. Specify maximum error notification count per problem

The VSS sends error notifications (0 to 9) to e-mail addresses. Enter the total number of notifications you want sent to each recipient for each problem. A number greater than 1 increases opportunities for delivery.

## 2. Add e-mail destinations

The VSS sends error notifications and information messages to the destinations you enter here:

- Enter the full e-mail address for each destination, for example, skha@kr.ibm.com.
- For each recipient, check one box--Errors, Information, All, or None--to specify which error notification he or she will receive.

**Note:** We strongly recommend that you list at least one e-mail address. This address will receive information when product engineers or service support personnel connect to your machine.

## 3. Change e-mail settings

To change the notification for a destination:

- Enter the full e-mail address for the destination, for example, achechi@it.ibm.com.
- Check one box: Errors, Information, All, or None.

## 4. Remove e-mail destinations

To stop sending notifications to a recipient, enter the full e-mail address of the destination you are removing.

---

## 2.4 Configuring the Simple Network Management Protocol

The VSS generates SNMP traps and supports a read-only management information base (MIB). The VSS generates both generic and product-specific SNMP traps that are sent to the addresses. Product-specific traps provide information about problems that the VSS detects and that require corrective action.

The CE can configure an SNMP environment at the service terminal by following these steps:

### 1. Allow MIB access

The VSS provides access to an MIB. To allow read access to this information, check Read-only. To disallow or deny read access to this information, check None. These records are available to the service terminal.

### 2. Allow generic traps

The VSS can generate generic SNMP traps. To allow the VSS to send these traps, check Yes. To disallow the VSS to send these traps, check No.

### 3. Allow product-specific traps

The VSS can generate product-specific SNMP traps. To allow the VSS to send these traps, check Yes. To disallow the VSS to send these traps, check No.

### 4. Specify maximum trap notifications per problem

Enter the total number of trap notifications you want sent for each trap condition. To ensure that traps are not lost, you may want to send each trap notification more than once.

5. Specify SNMP community name

Enter the SNMP community name (up to 31 characters) to use for authenticating requests. If you do not enter a name, this entry defaults to public access.

6. Specify SNMP SysContact

This is an optional entry. If you want, enter the location of the machine.

7. Add SNMP trap destinations

Enter the hostnames or dotted decimal addresses of the destinations that are to receive SNMP traps. For example: host.com, or 9.113.152.254.

8. Remove SNMP trap destinations

To remove previously defined trap destinations, enter the hostname or dotted decimal address of each destination that you are removing. The entry for removal must match the entry that you used for setting the destination.

---

## 2.5 Serial Connections

Each SMP cluster has two RS-232 serial ports. One port (S1) is connected to a modem to enable remote diagnostic access and call home. The other port (S2) is connected to an RS-232 switch, to allow connection of ASCII terminals or CE laptops for maintenance purposes. Port S2 is the service terminal.

### 2.5.1 Call-Home Support

Call-home support is a continuous self-monitoring system that initiates a call to the service representative if a failure or potential failure is detected. With this feature, a service representative can automatically be sent out to replace the failed or failing component and can bring along the required replacement parts. Repair time is thus minimized.

### 2.5.2 Remote Diagnostics and Support

Using the RS-232 connection, the service representative can initiate a call to the VSS storage server and remotely analyze potential failures and determine the type of failure and the proper fix. This is particularly important when microcode problems are suspected. Using the remote support facility, IBM service support staff can initiate a call to the VSS storage server and correct many types of microcode problems. When a microcode problem is reported, a fix can be created and then applied from a remote location.

It is important to stress that this remote access does not compromise security of the data stored on disk.

Two levels of support can access the system through the RS-232 port. CE support does not have access to any customer data. Product engineer (PE) support, the highest level of support, may require root access to the system, which gives the PE access to the data. This level of access must be specifically unlocked by the customer on site; it relocks automatically after one week. For all support levels, the customer has to authorize the use of remote support. If remote

support is activated, the customer is notified by e-mail. Call-back modems and other security procedures (such as passwords) should be used to prevent unauthorized access to the system.



---

## Chapter 3. Versatile Storage Specialist (VS Specialist)

In this chapter we describe the VS Specialist, the Web interface for the VSS. It is the configuration and administration interface of the VSS. It is the most important interface between the hardware, the software, and the storage administrators.

---

### 3.1 Prerequisites

The VS Specialist consists of a set of Java applets. Therefore you need a Java-capable browser that supports Java 1.1.2. We used Netscape Communicator 4.06 installed on Windows NT 4.0. We suggest that you use the same or a newer version, which you can download from the Netscape Web site.

The VS Specialist is not enabled by default. At installation time the IBM CE should have defined the hardware and part of your VSS configuration, enabled the VS Specialist, and defined the userid to connect to the VS Specialist. The userids can be defined only after the interface has been enabled. Section 2.2, "Configuring Versatile Storage Specialist" on page 22 details the VS Specialist setup.

The interface the IBM CE uses is a direct connection to the VSS. It is necessary to connect and perform the initial steps before the VS Specialist is enabled. Once the VS Specialist is enabled, the user can connect and finish or modify the VSS configuration by using the VS Specialist through the Web interface. When the service terminal is in use, you cannot use the VS Specialist to perform configuration changes.

**Note:** The service terminal is not intended for use by customers, and you should not use it after the installation of the VSS has been completed. Perform any further changes from the VS Specialist.

---

## 3.2 Components

The VS Specialist is a software package residing inside the VSS and includes these components.

- The Storage Function Operation Interface (SFOI)
- The information server
- The Web server
- Client code

The SFOI, the information server, and the Web server belong to the VSS itself and run inside the cluster controllers. The client software is a set of Java applets running inside a Web browser on the client workstation.

The SFOI is an internal application programming interface (API) to the VSS server functions required by the VS Specialist. The VS Specialist uses this API to perform queries of the current configuration and service setup for the facility. All microcode, including the VS Specialist, SFOI, and VSS information server functional microcode runs on both cluster controllers of the VSS. However, the VSS information server functions provide control over only the local cluster's configuration. To enable the VS Specialist to present a consolidated view of the storage facility, the SFOI maintains a cross-cluster socket connection. This ensures that the configuration information from both clusters is available to you even though a network connection to only one of the two cluster controllers is established.

The information server is essentially a communication link that communicates on one side with the client applets through a TCP/IP socket connection and on the other side with the SFOI. It provides the most up-to-date configuration information to the client applets, so that it can be displayed to the end user, or to allow the end user to change the configuration.

The information server constantly listens for socket connections being made by the clients. When a socket connection is established, a new thread is spawned to handle the "servicing" for that client. Each client provides a user name and access token to the server, which are then used for authentication purposes. After client authentication, the information server processes any information requests received over the socket connection. The requested information is retrieved from the SFOI and then returned to the client, after which the socket connection is closed and that particular server thread is ended. This ensures that no zombie threads will persist on the server side, in case the browser terminated the applet unexpectedly (or the browser itself terminated).

The information server also keeps an activity log, where it records in a log file all actions and error conditions encountered during the operation of the server. In case of catastrophic failure, the log file can be used to determine the cause of the problem.

The VS Specialist Web server is set up for secure connection through the Secure Sockets Layer (SSL). The data sent between the Web server and the Web browser on the client side is encrypted through the use of public and private key pairs.

The client code of the VS Specialist consists of a set of Hypertext Markup Language (HTML) pages and Java applets, which are executed from inside a Java-enabled Web browser running on the end-user's network-connected workstation. The user has the option of connecting to either of the two cluster controllers, by specifying a universal resource locator (URL) consistent with the controller's IP address or hostname.

The VS Specialist is frame enabled. When running inside a Web browser, it presents two main frames on the main panel: a navigation frame, on the left with the hiperlink buttons for accessing VS Specialist functions, and the working frame on the right (see Figure 14). At the bottom of the panel is the message area where messages are displayed while applets are processing or the client browser is waiting to receive data from the server.



Figure 14. VS Specialist Main Panel

### 3.3 Security

Accessing the VSS through the VS Specialist requires a valid userid and password. As an option, the VS Specialist can be configured to require that a user originate from a specific IP address or a range of IP addresses. Users can be restricted to one of three different authorization levels: viewer, configurator, and administrator.

A viewer can view the current configuration and status information, while a configurator can view the current configuration and status information and can make changes to the configuration. An administrator can define new user IDs, delete old IDs, and assign, change, and revoke passwords and levels of authorization.

The connections are secured through the SSL. SSL implements HTTPS, a variation of the standard HTTP, which encrypts data to and from the VSS and the client through an asymmetric encryption algorithm.

### 3.4 Connecting to Your VS Specialist

To connect to the VS Specialist through the browser, enter the URL of one of the two clusters of your VSS machine (see Figure 15). Usually the two clusters have the same hostname root with a suffix, such as c0 and c1, to identify the cluster. You do not need to be connected to a specific cluster, even if you want to work with a cluster different from the cluster to which you connected.

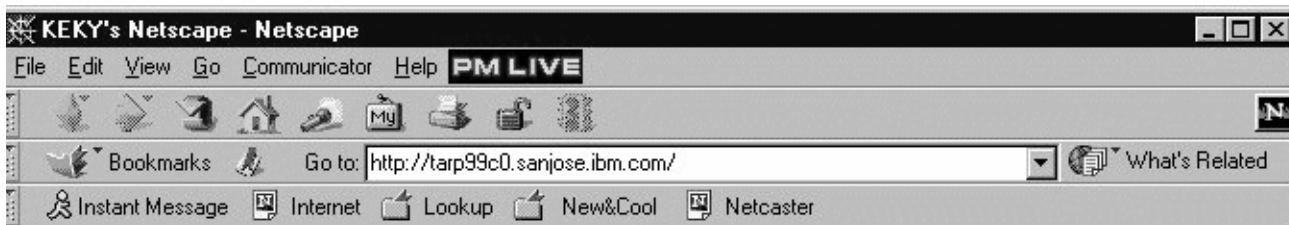


Figure 15. Entering the URL of Your VSS Machine

The information server is a process that starts inside the VSS and enables the connection with the VS Specialist users. If the information server is not ready to allow the connections, you are asked to wait until the startup process completes (see Figure 16).

The first time you enable the VS Specialist, the process of starting the information server may take up to 20 minutes for a machine that contains a full capacity of arrays and a large number of volumes.

#### 3.4.1 VSS Connections and Configuration Status

Although it does not matter which cluster you connect to to perform configuration changes, there is a performance implication. When you update the configuration from one cluster, the other cluster is updated too. However, the VS Specialist Java code running in the second cluster does not automatically update its version of the configuration if you sign on to the other cluster. The Java code notices that the configuration has changed and regenerates its configuration data. This may take some minutes. If a configuration change is made through the CE service terminal, the next time you sign on to either cluster with the VS Specialist, the same update to the VS Specialist configuration will be triggered and you will have to wait. In both cases you will get the screen shown in Figure 16.

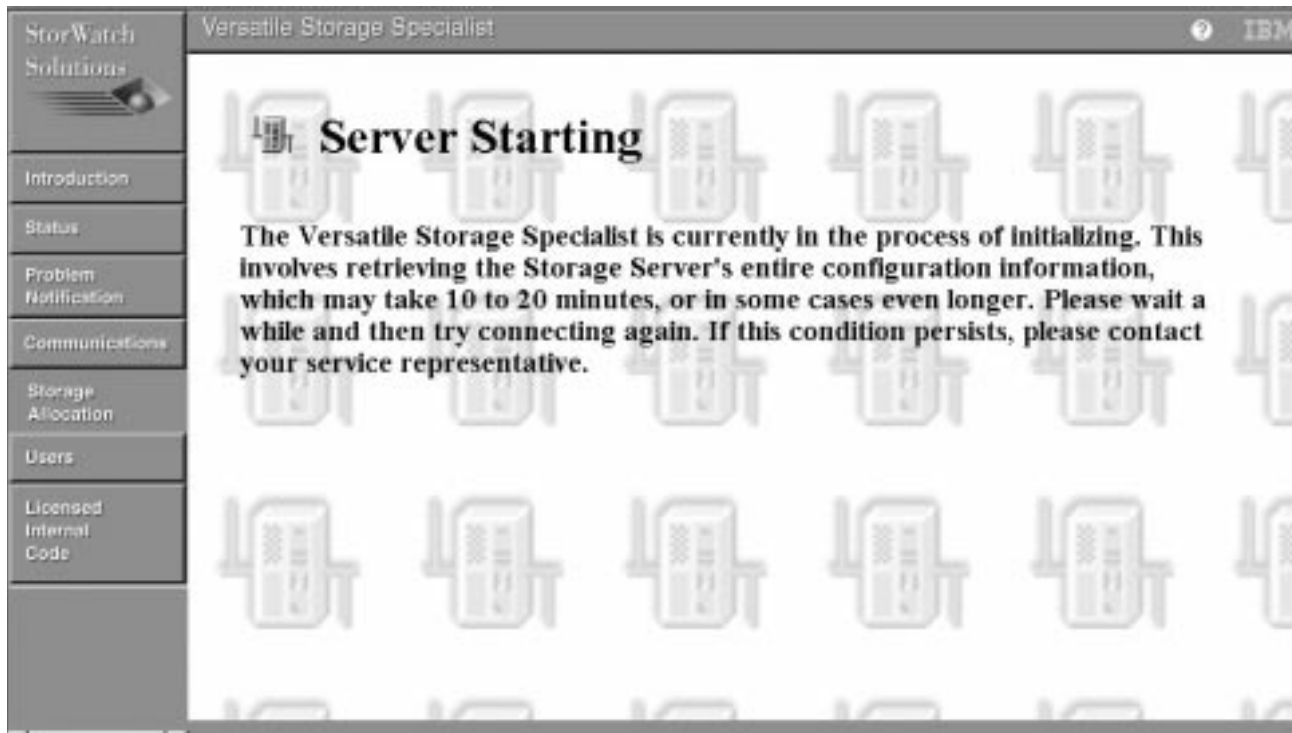


Figure 16. VS Specialist: Server Starting Message Panel

## 3.5 Panels

To access the VS Specialist panels simply click on a button in the navigation frame. Figure 17 presents a logical view of all of the VS Specialist panels, the connections among them, and the main functions.

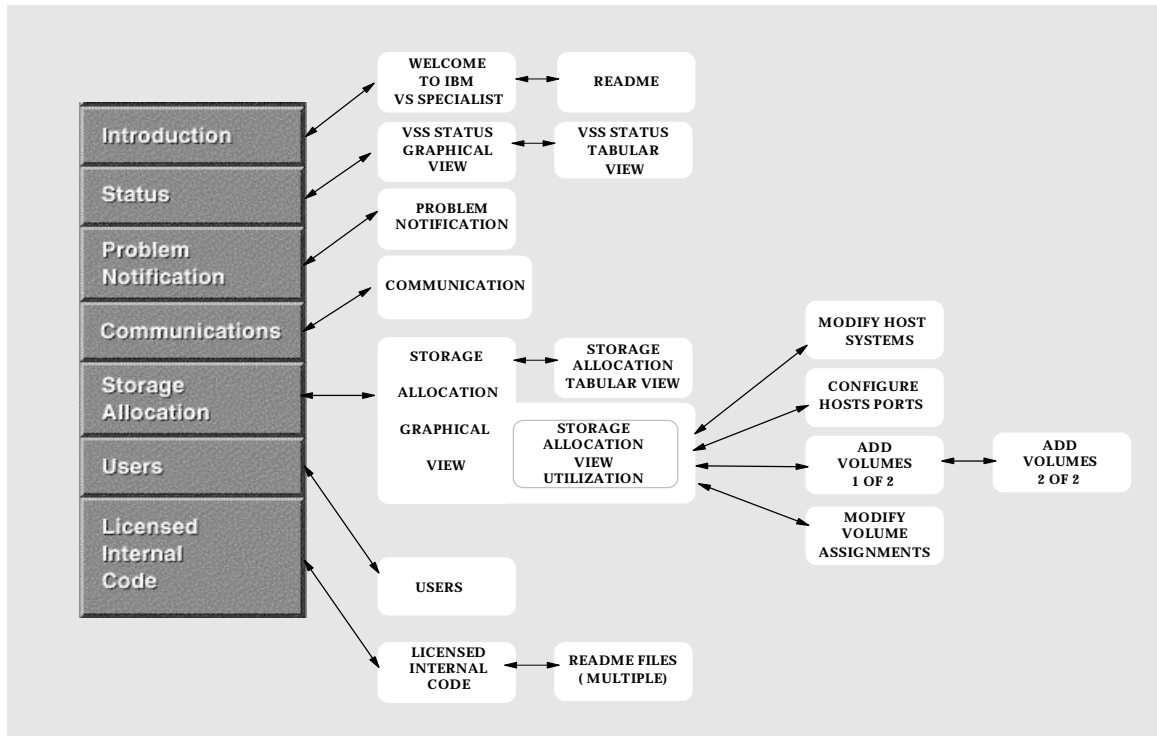


Figure 17. Logical View of the VS Specialist Panels

### 3.5.1 Introduction Panel

The introduction panel, also called the *information panel*, is the default panel. It shows the machine type, model, and serial number set at manufacturing time.

The title bar contains a question mark, which you can use to call the help panel. On the left side of the panel, the navigation frame contains these selectable buttons:



Figure 18. VS Specialist: Introduction Panel

- Introduction
- Status
- Problem Notification
- Communications
- Storage Allocation
- Users
- Licensed internal code

The Introduction panel comes up by default when you connect to the VS Specialist.

The Introduction panel is the only VS Specialist panel that does not require user authentication. When you enter the URL of your VSS machine, the Introduction panel appears, and you are not prompted to authenticate with the server. When you click on any other hiperlink button in the navigation frame, the Username and Password Required panel appears (see Figure 19 on page 34), and you must authenticate before you can work with the panel you requested.



Figure 19. VS Specialist: Username and Password Required Panel

### 3.5.2 Status Panel

When you select the Status button on the VS Specialist navigation frame, the Status -- Graphical View panel appears as shown in Figure 20. In this panel you can see the logical components of the VSS machine: host attachment ports, controller and cache stores, SSA adapters, SSA storage devices. A Hosts icon completes the logical schematic of the VSS environment.

The solid lines represent the current affinity connections. The broken lines represent the alternative affinity connections. In failover mode the broken line of the operational cluster controller changes to solid, and the solid line of the failed cluster changes to a broken line.

A failed component is highlighted in different colors. The severity of the problem determines the color used. The different colors are:

- White - Normal status. No failure has occurred.
- Yellow - Failure status. A component has failed, but access to the data is not affected. Performance may be degraded, however.
- Orange - Failure status. A component has failed, and access to the data is affected. One or more hosts have temporarily lost access to one or more logical volumes defined in the storage facility.



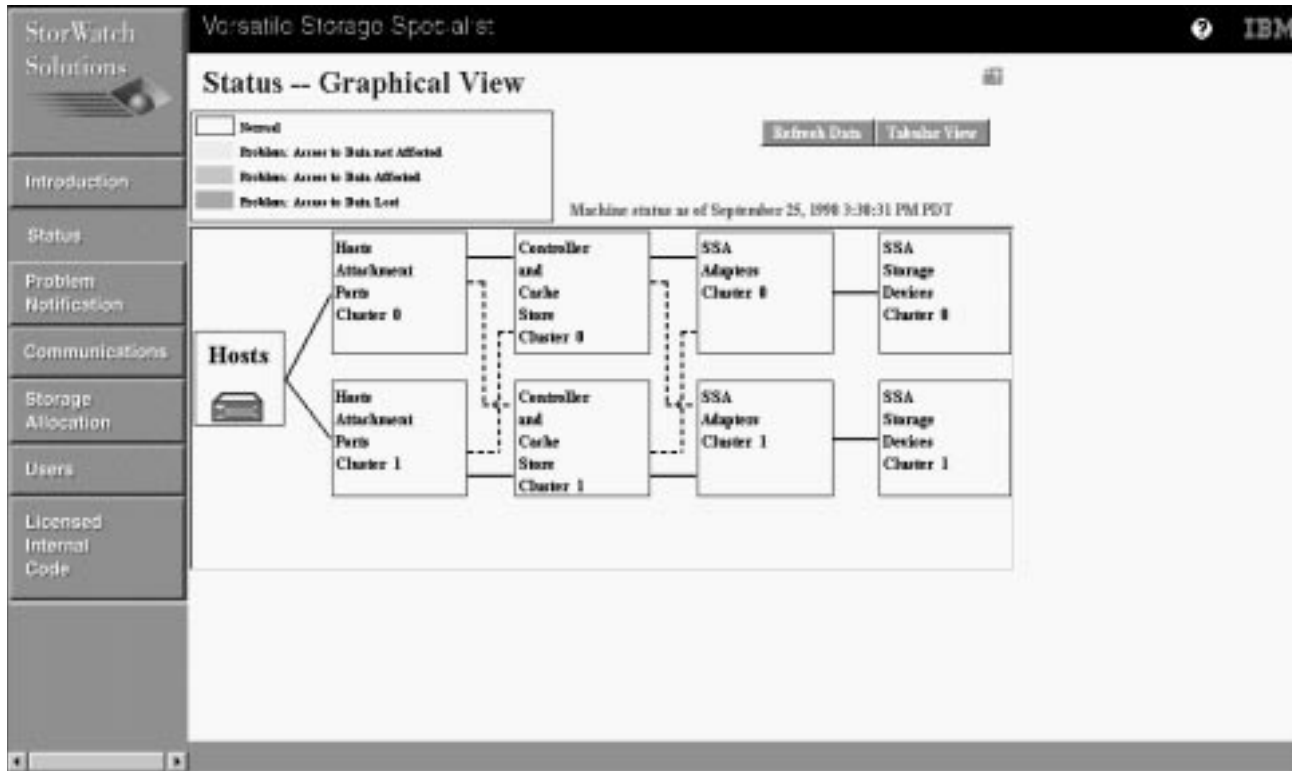


Figure 20. VS Specialist: Status -- Graphical View Panel

On the Status -- Graphical View panel, click on **Tabular View** button to access the Status -- Problem Log panel (Figure 21 on page 36). The panel contains a list of the problems that occurred on the date recorded in the Problem Age field.

**Note:** Switching from a refreshed Status -- Graphical View panel to the Tabular View (Status -- Problem Log) does not mean that the list you will receive is refreshed as well. Use the Refresh Data button every time you switch between the Status panel and the Problem Log, and refresh the panel every time you want to see the current status.

The current status can differ, according to the successful completion or failure of subsequent recovery actions. The fixed or expired problems are not listed on the Status -- Problem Log panel.

Here is a brief description of each field:

- Problem Id: a unique ID for each problem
- Problem Description: Brief description of problem
- Problem Age: Date and time of problem record creation
- Problem Status: Severity of the problem. This field is highlighted in different colors. The possible colors are:
  - Gray - No impact
  - Yellow - Access to data not affected
  - Orange - Access to data affected
  - Red - Data loss

- Service Status: Shows whether problem is "in service" or "in pending service"
- User Action: When displayed, contains suggestions of the type of service required by the IBM CE or the customer
- Probable Cause: When displayed, indicates which component or machine function may have failed
- Additional Information: This field is not always used. When displayed it contains additional information that may be useful for further diagnosing or resolving the problem. It is primarily used to identify a failed data location in terms of LUN and LBA.

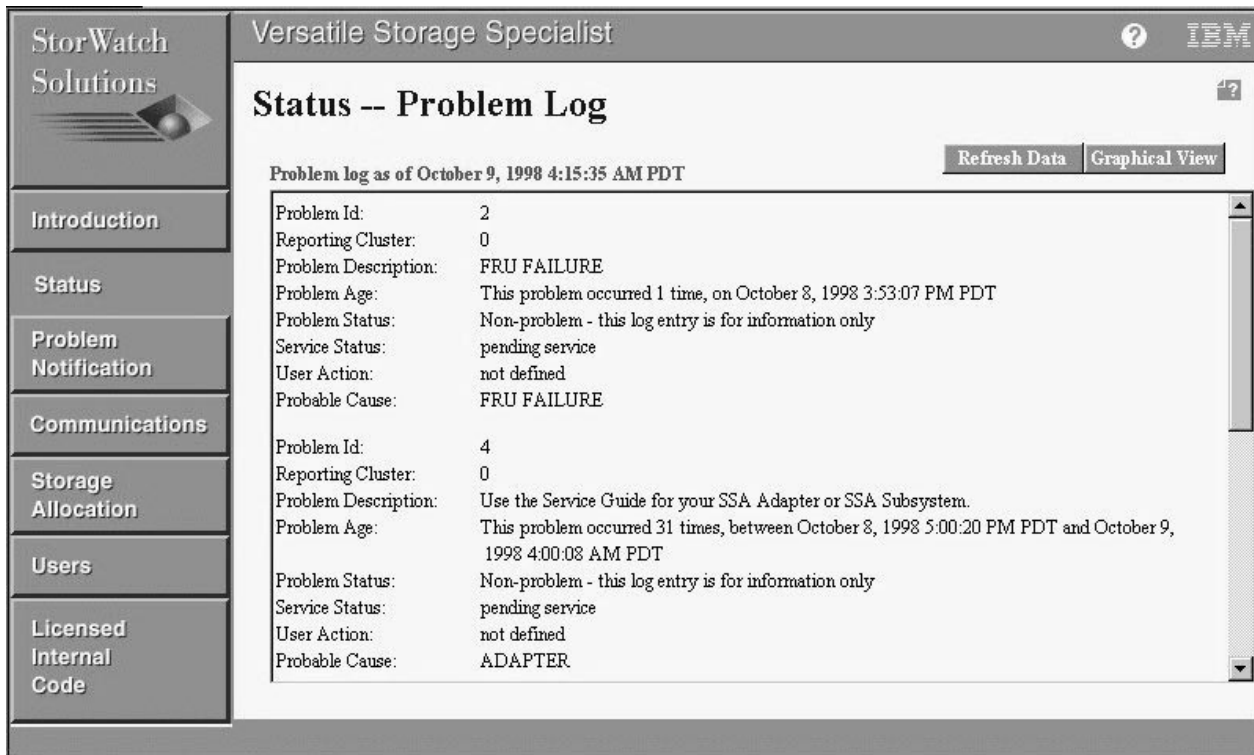


Figure 21. VS SPecialist: Status -- Problem Log Panel

### 3.5.3 Problem Notification Panel

The VSS sends notification of a problem to a user, either through e-mail or by calling the specific phone number of the service provider. The VS Specialist Problem Notification panel (Figure 22 on page 37) shows the e-mail addresses and phone numbers as well as the SNMP trap destinations and the status of remote service access.

The settings on the Problem Notification panel are enabled when the IBM CE installs the VSS machine. The information must be entered from the service terminal and can be updated later.

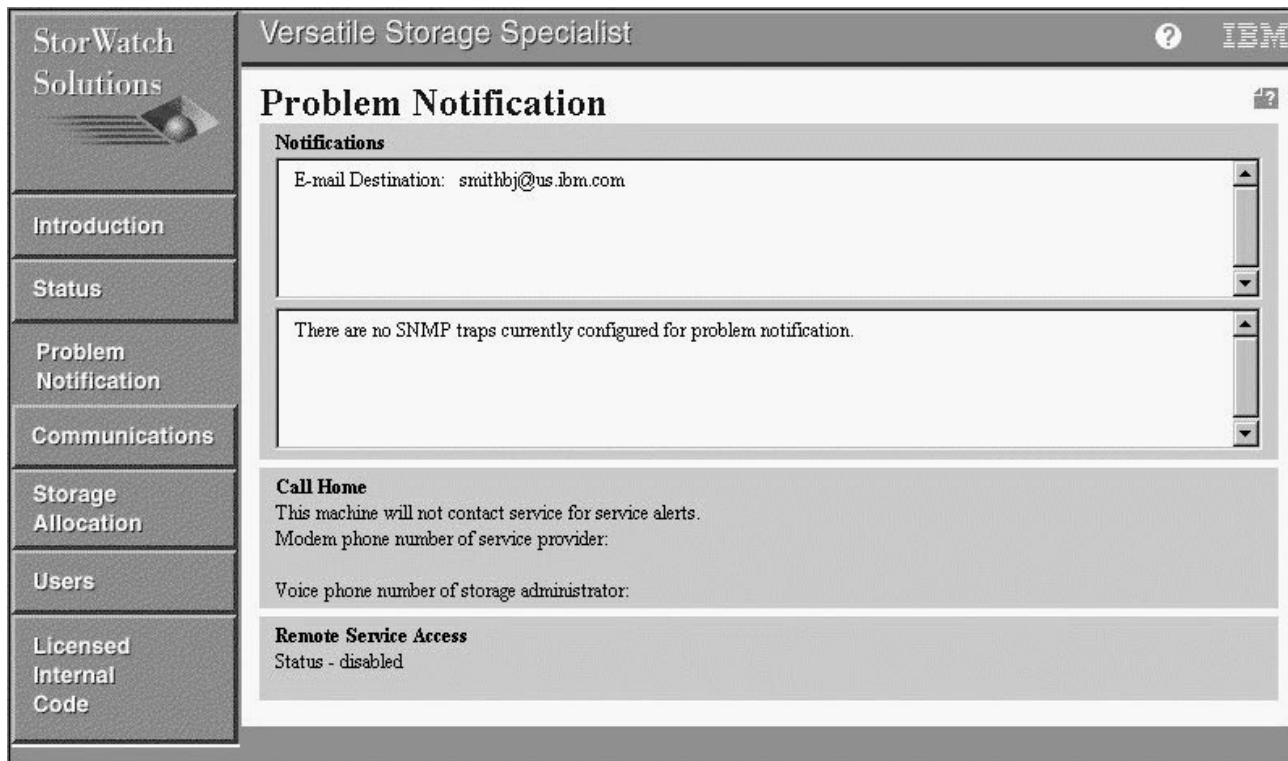


Figure 22. VS Specialist: Problem Notification Panel

### 3.5.4 Communications Panel

The Communications panel (Figure 23 on page 38) shows the following information regarding the network and communications-related information:

- TCP/IP setup for each cluster
  - Hostname
  - IP (dotted decimal) Address
  - Subnet Mask
  - Default Router
  - Domain Name Server address (if any)
- The VSS current modem setup, and local modem phone numbers for dial-in Remote Service Access

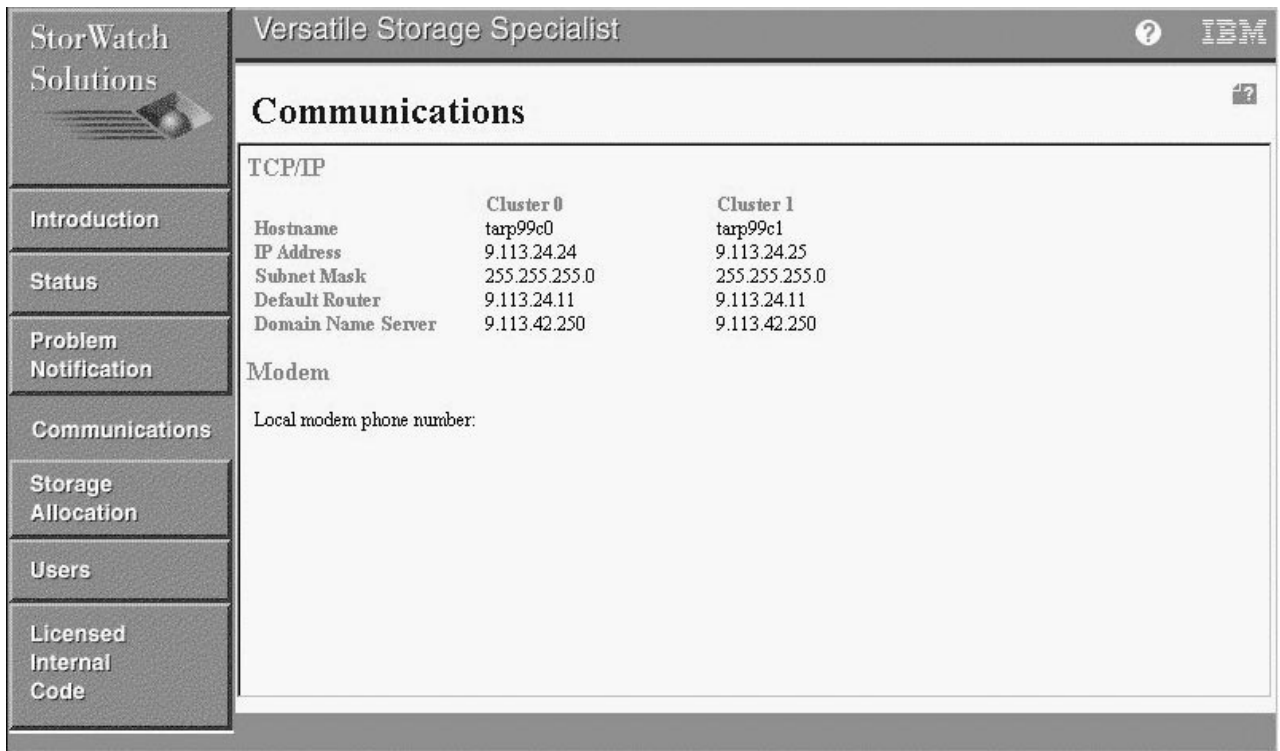


Figure 23. VS Specialist: Communications Panel

### 3.5.5 Storage Allocation Panel

The Storage Allocation panel is the most important and detailed of the VS Specialist panels. When you click the Storage Allocation hiperlink on the VS Specialist main panel, the Storage Allocation -- Graphical View panel shows up (see Figure 24 on page 39).

#### 3.5.5.1 Storage Allocation -- Graphical View Panel

The Storage Allocation -- Graphical View panel provides an authorized user with a graphical representation of the hosts, host adapters (SCSI adapters), storage adapters (SSA Adapters), clusters, drawers, and arrays on the VSS machine.

The first object line represents the hosts. It is a horizontally scrollable bar where the defined hosts are shown. With the 2105-B09 model of VSS, you can define up to 64 hosts. If more hosts are defined than can be displayed in the panel, use the arrows on the left and right side of the host display to scroll to the remaining hidden host object representation.

The second object line shows the host adapters. In the current VSS configuration, 8 dual-ported SCSI adapters are available. The ports are labeled A and B. The SCSI ports are numbered from port 0 to port 15, left to right. Even numbered ports are the adapter A ports, and odd numbered ports are adapter B ports. The adapters are physically connected to the bays. The VSS has a maximum of two SCSI and two SSA adapters per bay. As installed in the bay, the sequence is SCSI, SSA, SSA, and SCSI. See Chapter 1.1, "Structure and Components" on page 2 for further details. That is why SCSI adapters in each bay are numbered as adapter 1 and adapter 4, and the SSA adapters are numbered as adapter 2 and adapter 3. A maximum of four bays can be

connected to a VSS B09 model. Each pair of bays is connected to different clusters. Bays 4 and 8 are in cluster zero, and bays 5 and 9 are in cluster 1.

In the Storage Allocation -- Graphical View panel, each time you select one of these elements, the information window positioned at the bottom right corner of the panel shows the details of the selected object.

The third object line consists of two elements: Cluster 0 and Cluster 1. The clusters are two separate components internally connected to each other. You see these as separate entities. On the rear side of the two clusters are the four device adapters (SSA) for each cluster. The SSA adapters are numbered 2 and 3 and are in the middle of the four adapter slots, between the two SCSI adapters. Next to each adapter is the drawer, shown by the two arrays of RAID-5 SSA disks. Refer to Chapter 1.1.6, "Arrays" on page 15 for more information about setting up the arrays.

The Clear View button does not have a connecting panel. You use it to clear any selection that you have made on the panel. The Storage Allocation -- Graphical View panel is the only interactive, nonstatic panel of the VS Specialist. You can select hosts, host ports, and arrays and if you want to change the selection, you can use the Clear View button to refresh the Storage Allocation -- Graphical view panel.

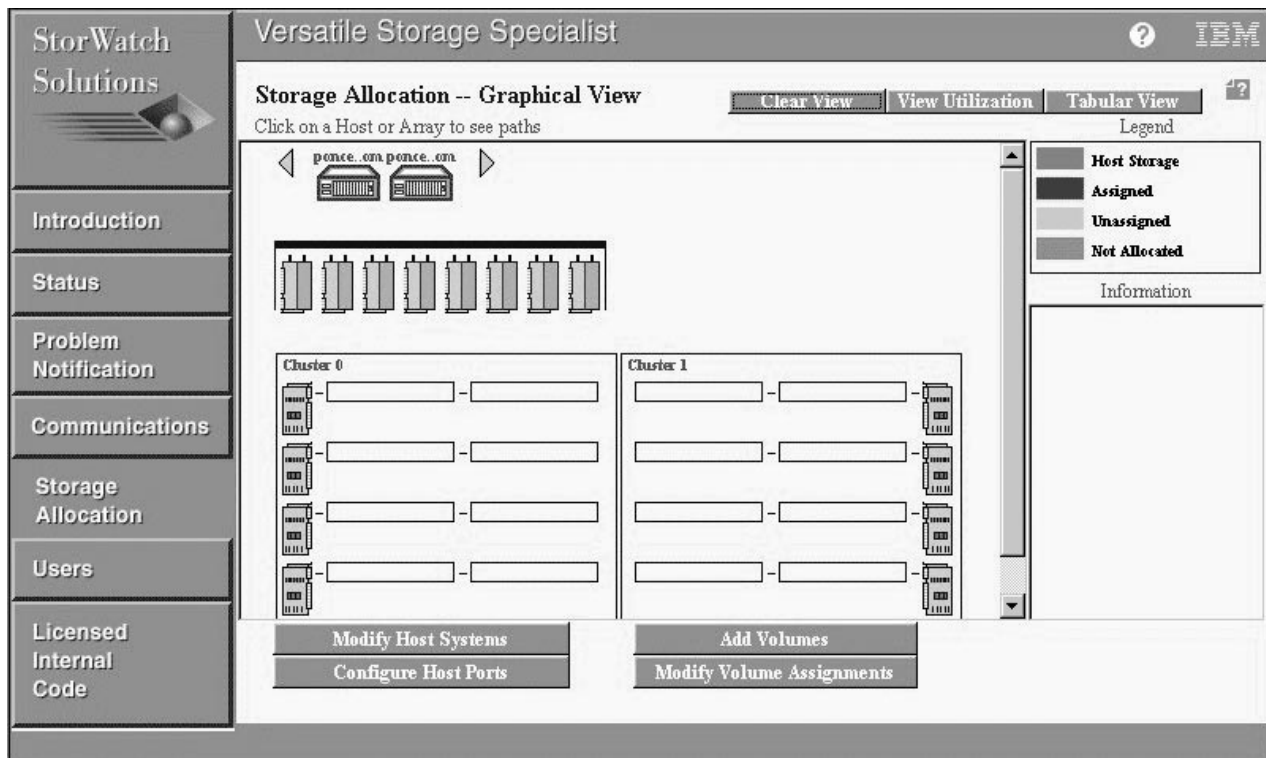


Figure 24. VS Specialist: Storage Allocation Panel; Graphical View

### 3.5.5.2 Storage Allocation -- Graphical View Panel: View Utilization

When you select the View Utilization button on the Storage Allocation -- Graphical View panel you are not switched to another panel. Rather the arrays of the clusters are scaled in color to signify different capacity information (see

Figure 25 on page 40). The legend in the top right corner of the frame explains the colors:

- Red: Space allocated to logical volumes and accessible to one or more hosts
- Orange: Space allocated to logical volumes but not yet assigned to any host
- Green: Space not yet allocated to logical volumes
- Purple: This color has the same meaning as red, but it is used to show how much of the space is assigned to a particular host, when a host is selected in the view. Purple is used to distinguish between space that is allocated to volumes accessible to that particular host, as opposed to the space that is allocated to all the other logical volumes (which is shown in red).

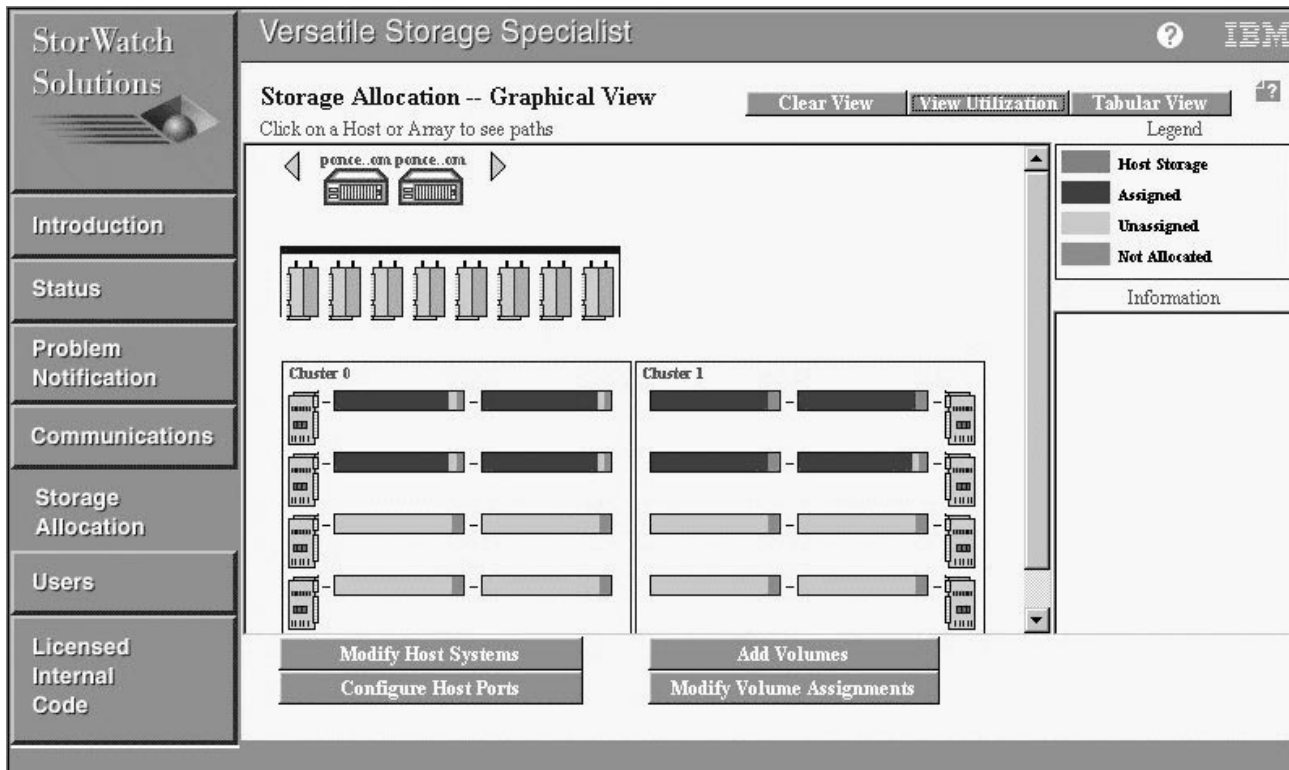


Figure 25. VS Specialist: Storage Allocation -- Graphical View: View Utilization

From the Storage Allocation -- Graphical View panel, you can simply use your mouse and click on the components to see what they are and where they are allocated. See Figure 26 on page 41.

It is important to know some basic rules before configuring a VSS. Each one of the four SCSI adapters of each cluster has an A and B port. The SCSI ports must be assigned to a cluster affinity. We strongly recommend assigning the A port locally (to the cluster where it physically resides) and the B port remotely (to the other cluster). For further details, see Chapter 6.2, "Performance Configuration Options" on page 90. We assume these assignments in our discussion on operations to be performed to configure the VSS.

To understand where a SCSI port resides, select it and view the information that appears in the information window. Do the same for the hosts, the SSA adapters, and the arrays.

The SCSI adapters as represented graphically belong respectively to I/O bay 4, 5, 8, 9, 4, 5, 8, and 9. Bays 4 and 8 are in cluster 0, and bays 5 and 9 are in cluster 1. As per our recommendation, define port A of the SCSI adapter to the local cluster, and port B to the remote cluster. Therefore SCSI ports 0, 3, 4, 7, 8, 11, 12, and 15 will be connected to cluster 0, and SCSI ports 1, 2, 5, 6, 9, 10, 13, and 14 will be connected to cluster 1.

:

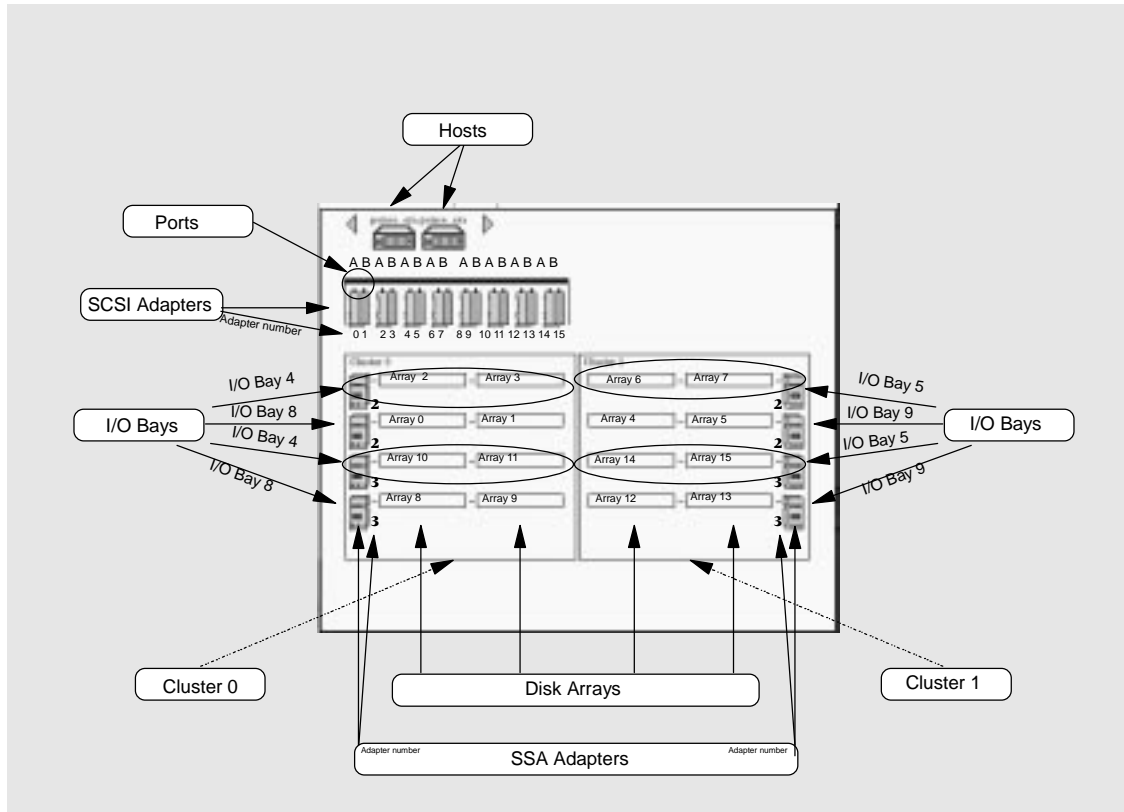


Figure 26. Disposition of Elements in the Storage Allocation -- Graphical View Panel

### 3.5.5.3 Storage Allocation -- Tabular View

Select the Tabular View button to get to the Storage Allocation -- Tabular View List of Assigned Volumes panel (Figure 27 on page 43).

The panel is a table with seven columns, and as many rows as necessary to describe all of the hosts and their assigned volumes in each cluster per port and SCSI LUNs.

The Host column shows the attached host name and the logical volume being described. If the volume is shared with other hosts, there will be an entry for each host in the table, and not multiple hosts shown in a single row.

The Volume column shows the unique number assigned to the volume at the time of its definition. It is also the LUN serial number, which is useful for identifying the host or hosts (when the volume is shared between compatible hosts) whose data may be temporarily inaccessible.

The Size/Type column shows the size in GB of the volume or LUN and the type. Type is displayed only if the host is an AS/400.

The Shared column shows a Yes when the volume is shared with another host. Sorting this column is useful for finding the other host sharing this volume.

Above each column in the table there is a sort facility pull-down list. Each list allows the selection of: No Sort, First, Second, and Third. With the exception of No Sort, which can be selected for multiple columns, the First, Second, and Third values can be selected only once in a column. If one of the values is selected, it will not be available in the other pull-down lists, and the Java applet automatically removes the selection from the other lists. This permits a maximum of three sortable columns. The column with the First value set in the sort list field will be sorted first, then the second and the third, if defined. To run the sort, click on the **Perform Sort** button, and the columns will be ordered in the selected sequence.

The Host Adapter column contains multiple information. Each row shows the bay number (4, 5, 8, or 9); the SCSI adapter ID of the bay (01 or 04); the SCSI port of the adapter being used (A or B, and in parentheses the port ID from 0 to 15); the SCSI ID; and the LUN.

The Cluster column shows the cluster with which the volume is associated. The association is made automatically when a particular array is defined. Consequently, the cluster to which the array belongs is shown here (0 or 1).

The Device Adapter column lists the storage device attachments associated with the volume. It shows the bay (4, 5, 8, or 9); the SSA Adapter (02 or 03); the SSA loop where the volumes are located (SSA adapter 2 in bay 4 cluster 0 and adapter 2 in bay 5 will have two loops if drawer 17 and 18 are installed; otherwise all SSA adapters will have one loop); the array number (00 to 35); and the volume sequence in the array.



**Storage Allocation -- Tabular View**

List of Assigned Volumes

Perform Sort Graphical View

no sort no sort no sort no sort first no sort no sort

Host	Volume	Size/Type	Shared	Host Adapter	Cluster	Device Adapter
ponce.sanjose.ibm.com	3030344641303336	24.0 GB	no	Bay: 4 SCSI Adapter: 01 SCSI Port: A (0) SCSI ID: 00 LUN: 00	0	Bay: 8 SSA Adapter: 02 SSA Loop: A Array: 02 Volume: 000
ponce.sanjose.ibm.com	3030354641303336	24.0 GB	no	Bay: 4 SCSI Adapter: 01 SCSI Port: A (0) SCSI ID: 00 LUN: 01	0	Bay: 8 SSA Adapter: 02 SSA Loop: A Array: 02 Volume: 001
ponce.sanjose.ibm.com	3030454641303336	28.0 GB	no	Bay: 4 SCSI Adapter: 01 SCSI Port: B (1) SCSI ID: 01 LUN: 00	1	Bay: 9 SSA Adapter: 02 SSA Loop: A Array: 07 Volume: 002
ponce.sanjose.ibm.com	3030464641303336	28.0 GB	no	Bay: 4 SCSI Adapter: 01 SCSI Port: B (1) SCSI ID: 01 LUN: 01	1	Bay: 9 SSA Adapter: 02 SSA Loop: A Array: 07 Volume: 002

1533: Downlevel Data ... please wait a few minutes and try again.

Figure 27. VS Specialist: Storage Allocation -- Tabular View Panel

### 3.5.6 Storage Allocation Tasks

The four buttons at the bottom of the Storage Allocation -- Graphical View panel enable you to perform four tasks on the VSS. In this section we describe these tasks. For detailed information about the panels discussed here, refer to Chapter 4, "Host Connections" on page 57.

#### 3.5.6.1 Modify Host Systems

The Modify Host Systems task is logically the first operation you perform when starting to configure a VSS system. The host machine must be defined and connected to the VSS in order to utilize storage space on the VSS. The VS Specialist is used to define the external machine that is going to use the VSS space.

The Modify Host System panel (Figure 28) allows an authorized user to add a new or remove an existing host.

To add a host system:

1. Enter the TCP/IP hostname in the Host Name field.
2. Under Host Types, select one of the types listed. Only the supported device types are shown in the pull-down list.
3. Click on the **Add** button.

The List of Defined Hosts box lists the hosts that have already been defined to VSS. Use the scrollable window if more than the displayed hosts have been defined.

To remove a host system:

1. Select the host in the List of Defined Hosts and click on the **Remove** button.

Neither the **Remove** nor the **Add** button performs the operation. The configuration is updated only when you click on the **Perform Configuration Update** button.

If your entry is incorrect or mistyped or you removed the wrong host, click on the **Cancel Configuration Update** button.

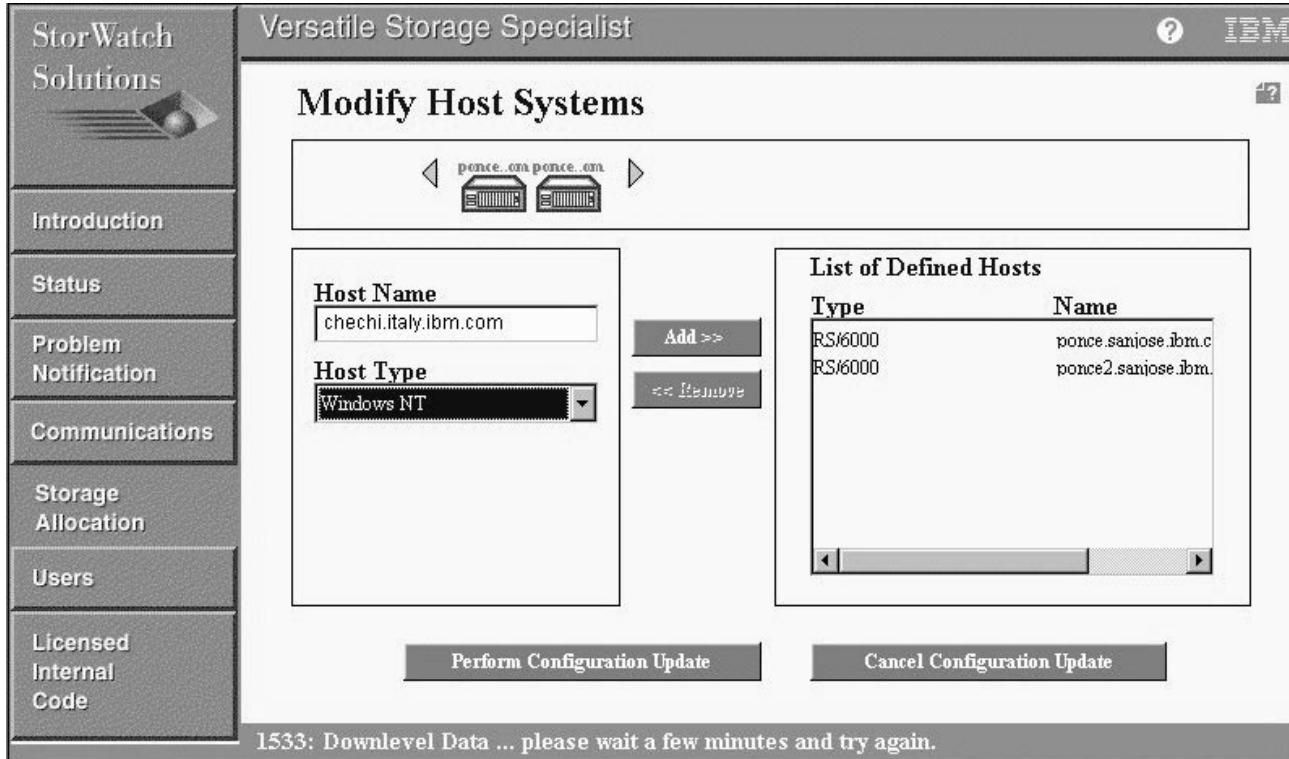


Figure 28. Modify Host Systems Panel

### 3.5.6.2 Configure Host Ports

To configure the host ports, click on the **Configure Host Ports** button. A host port is the SCSI adapter used to physically connect the host machine to the VSS. Once you have plugged the SCSI cable of the host into one of the SCSI adapters in the bay (refer to Chapter 6, “Performance and Tuning” on page 89 and to Chapter 4, “Host Connections” on page 57 for more information about the correct use of the SCSI adapters and ports), you use the Configure Host Ports panel (Figure 29 on page 45) to view the current bus configuration of any port or to configure or modify the configuration for a single selected SCSI port.

The icons on the top of the Configure Host Ports panel represent the SCSI ports. The icons are grouped two by two because each SCSI adapter has two independent ports. The Port IDs are numbered from 0 to 15, left to right.

Next to the First Bus Connection are two fields, each with pull-down lists. The first field is the Port, and the second is the Cluster. Choosing a number in the Port field pull-down list is the same as selecting a port icon. Be aware that changing

Port 0, which belongs to Cluster 0, to Cluster 1, for example, is an allowed selection but causes an error message to appear on your screen, and the operation is not performed.

The entry for the Bus Configuration field must be selected from the pull-down list, which lists the supported BUS configuration host types. Be sure to connect the port to the correct host type. Errors in this field may cause a malfunction and an incorrect LUN configuration.

The Available Hosts box lists the hosts defined to the VSS (through the Modify Host Systems panel). These hosts are not currently assigned to this SCSI port. The box maps the SCSI IDs on the bus. Use the Add and Remove buttons to respectively add or remove a port for a specific host. Use the Perform Configuration Update and the Cancel Configuration Update buttons to apply or cancel the configuration changes.

Each time you select a port icon or a value in a pull-down list, the field is highlighted. When you select a Port icon, the Port and Cluster fields are automatically updated, showing the port ID and cluster to which the port belongs. When you select a port ID in the pull-down list, the corresponding port icon is automatically highlighted.

**Note:** If a CE has previously configured a port through the service terminal, it may be assigned to a dummy host called Unrelated Host or Device. To reassign this port, you must remove this entry if the SCSI ID you want to use is defined to an Unrelated Host or Device. Refer to Chapter 4, “Host Connections” on page 57 for detailed instructions on how to make configuration changes to host ports.

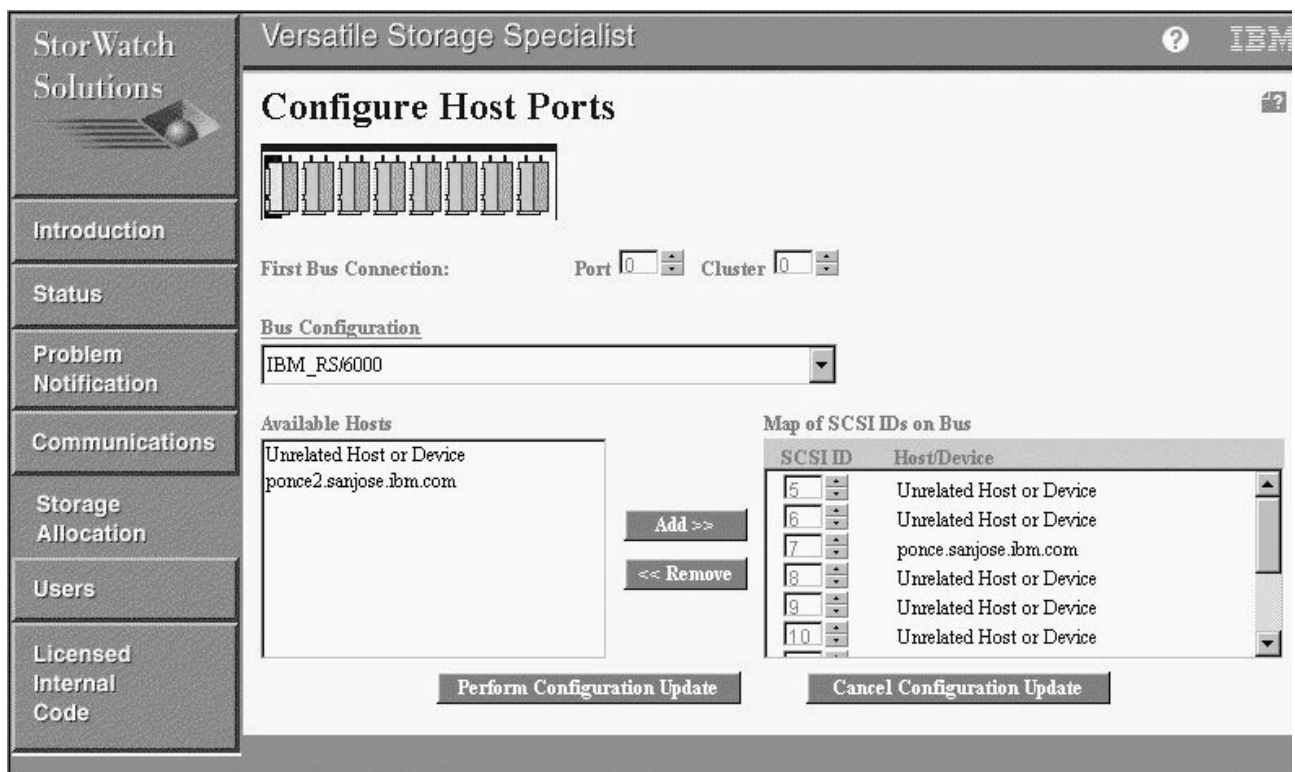


Figure 29. Storage Allocation Tasks: Configure Host Ports Panel

### 3.5.6.3 Add Volumes

The Add Volumes button activates the first of two panels used to define the storage allocation for one or more host systems that are attached through a single host port to VSS (see Figure 30 on page 47). In this panel, changes in selection are not permitted. If you select a host but should have selected another host, you have to use the Clear View button and reselect the correct host. You must select the objects in the following sequence:

1. Host
2. SCSI Port
3. Array

Selecting an array before a host, for example, is not allowed. The selection is not performed, and the object is not highlighted.

Array selection is an optional step. After you select a host system and a SCSI port, all of the arrays within the cluster that owns the port are selected by default. You may optionally select one or more of the arrays in the cluster to control the SSA adapter and arrays used for volume assignment.

The Host icons are included in a horizontally scrollable bar to allow the hosts to be shown. If more hosts are defined to VSS than fit in the row, the arrows at the sides of the row are used to scroll to the other host icons. On top of each icon are the first five and the last two characters of the host name.

**Note:** If you have multiple hosts prefixed with the same five characters in the host name, you can select a wrong host if you just look at the name appearing on the top of the icon. Be sure the host you select is the correct one by double-clicking on the name in the information box of the Add Volumes panel.

When you select a host in the Add Volumes (1 of 2) panel, all arrays that have at least one logical volume belonging to the selected host are shown in color. The arrays that are allocated to hosts other than the host you select and that have no volumes allocated to that host remain uncolored.

Use the Storage Allocation -- Graphical View Utilization panel to determine which arrays have volumes on them and how much space is allocated. Having a clear understanding of the used and free space makes adding a volume easier.

The Information box shows the detailed information related to the last object selected and highlighted in the panel.

To exit the panel without performing any operations, use the **Cancel Adding Volume** or the **Clear View** button.

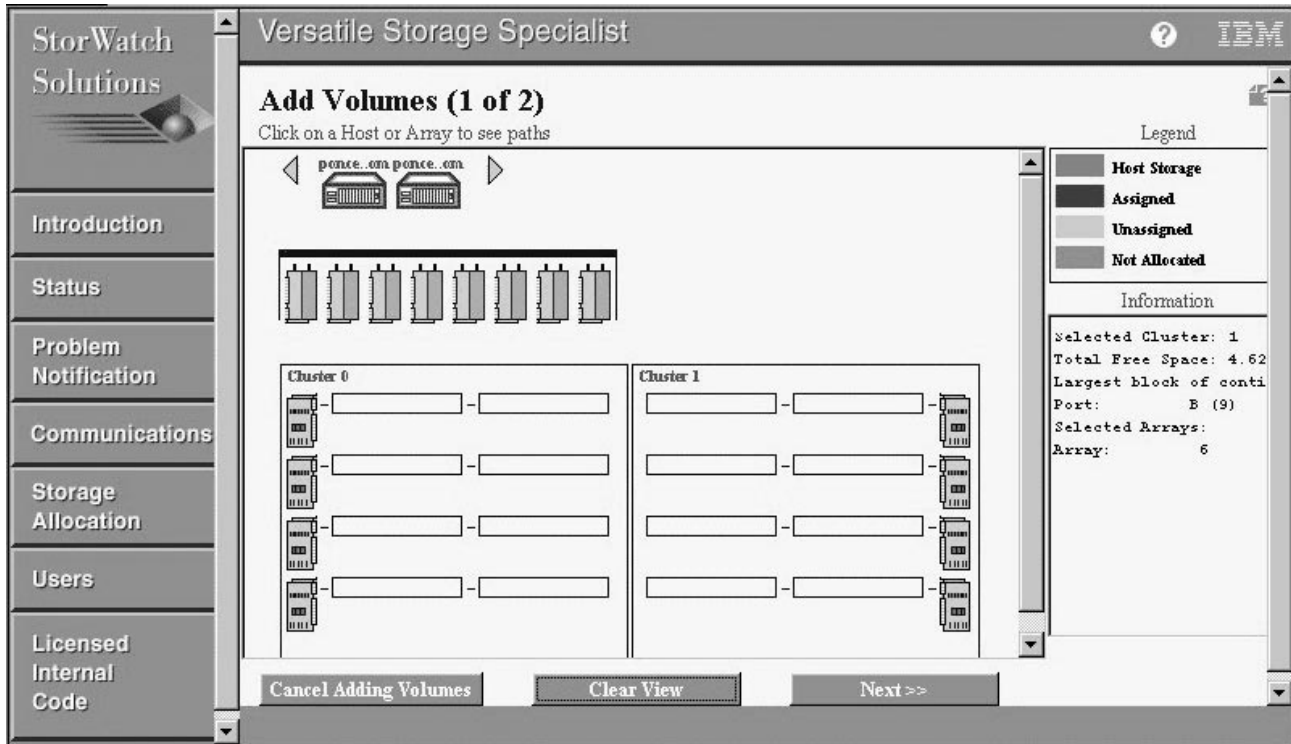


Figure 30. Storage Allocation Tasks: Add Volumes (1 of 2) Panel

When a selection has been made, a host, a port, and an array are highlighted, and a small solid line connects the elements. The host shows its connection with the port, and the port with the device adapter next to the selected disk's array (see Figure 31 on page 48). Click on the **Next** button to go to the Add Volumes (2 of 2) panel and finish the configuration update.

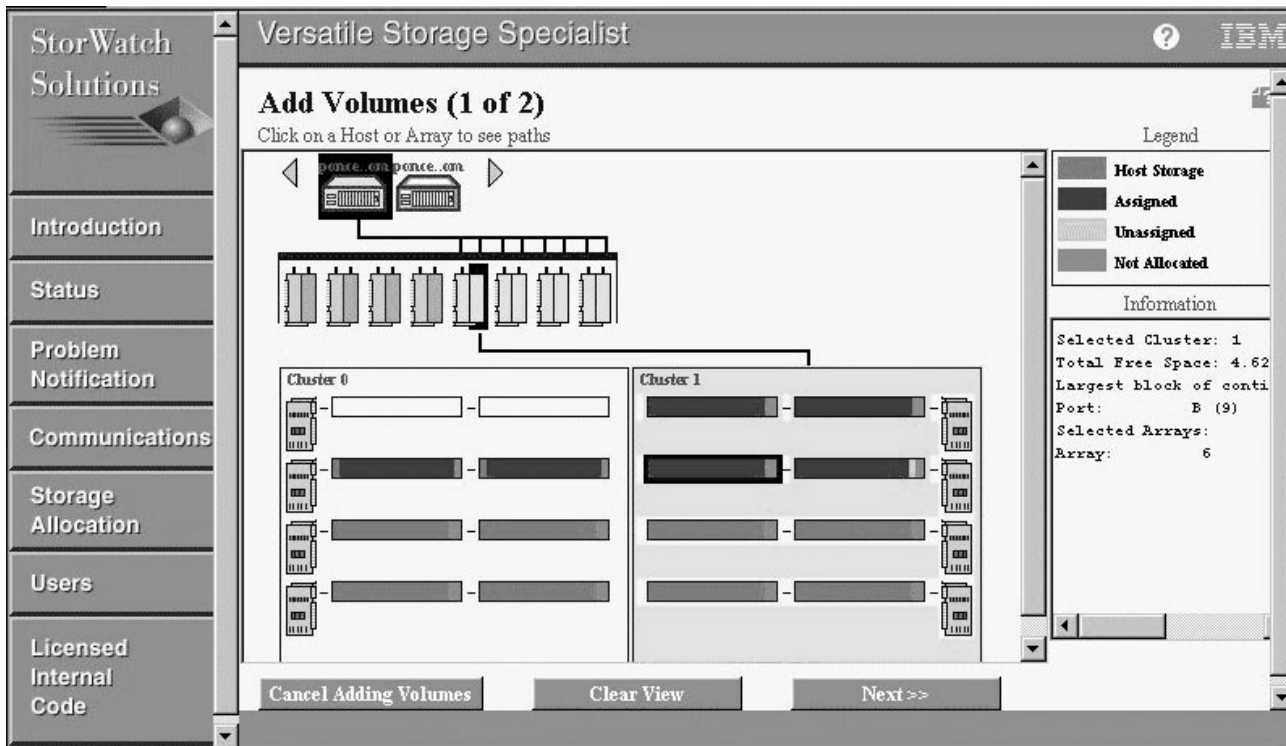


Figure 31. Vs Specialist: Add Volumes (1 of 2) Panel - Object Selection

In the Add Volumes (2 of 2) panel (Figure 32 on page 50) you define the size and the number of volumes you want to create on the selected arrays.

The top section of the panel contains information related to the selections made in the Add Volume (1 of 2) panel: the host to which the volumes are to be assigned, the amount of free space in the cluster to which the SCSI port that has been selected belongs; the total free space in the selected arrays, and the largest possible volume size. The largest volume that can be created is the maximum space that is available in the array. If more arrays are selected, the smallest free space of the arrays is shown in this field.

In the Volume Definition section you create a list of volumes to be defined and assigned to the selected port. The two selector boxes enable you to select as many volumes as required of the desired size (in GB). You cannot type in the values. The Java applet automatically updates the volume size or the number of volumes when the resulting size, multiplied by the number of volumes selected, exceeds the maximum free space available. For example, if just one array is selected, and the maximum volume size is 6.62 GB, the applet allows the creation of a maximum of 13 volumes of 0.5 GB each or 1 volume of 6 GB and one of 0.5 GB, or any other possible combination. The values available will be automatically changed any time a selection is made.

**Note:** When you create a list of volumes to be added, add the largest volume first because the list is processed top to bottom, and the "best fit" program will have a better chance to succeed if the largest volumes are "fit" first.

Once you have decided on number of volumes and the volume size, click on the **Add** button. The Volumes to be added box lists all volumes to be added. At the

bottom of the Volumes to be added box, some information regarding the selection is displayed and is automatically updated each time a volume is added to the list. After you have added all volumes to the Volumes to be added box, you have to decide on where to place the volumes.

In the Volume Placement section you can select: **Place volumes sequentially, starting in first selected array** or **Spread volumes across all selected arrays**. You must select one of the two radio buttons before continuing. By default, the volumes are placed sequentially (refer to Chapter 6, "Performance and Tuning" on page 89 and to Chapter 4, "Host Connections" on page 57).

Use the **Add Volumes** button to add the volumes or the **Cancel Adding Volumes** button to quit without performing the configuration update and to go back to the Storage Allocation -- Graphical View panel.

When you click the **Add Volumes** button, the list of volumes is created and automatically assigned to the selected SCSI port storage map using available SCSI IDs and LUNs. An error message is presented if the storage map does not contain adequate SCSI IDs and/or LUNs. The rules are:

1. The VSS will use a SCSI ID only if it is not assigned to a host initiator or is not assigned to Other Host System or Device. It may be necessary to remove the Other Host System or Device assignment for one or more SCSI IDs to make room in the storage map for new volumes.
2. All of the volumes assigned to a SCSI ID must be members of a common LSS; that is, they must be from arrays that are attached to a common SSA adapter and SSA loop.
3. The maximum number of volumes assigned to a SCSI ID is determined by the maximum\_LUN attribute for the selected Host\_Type.

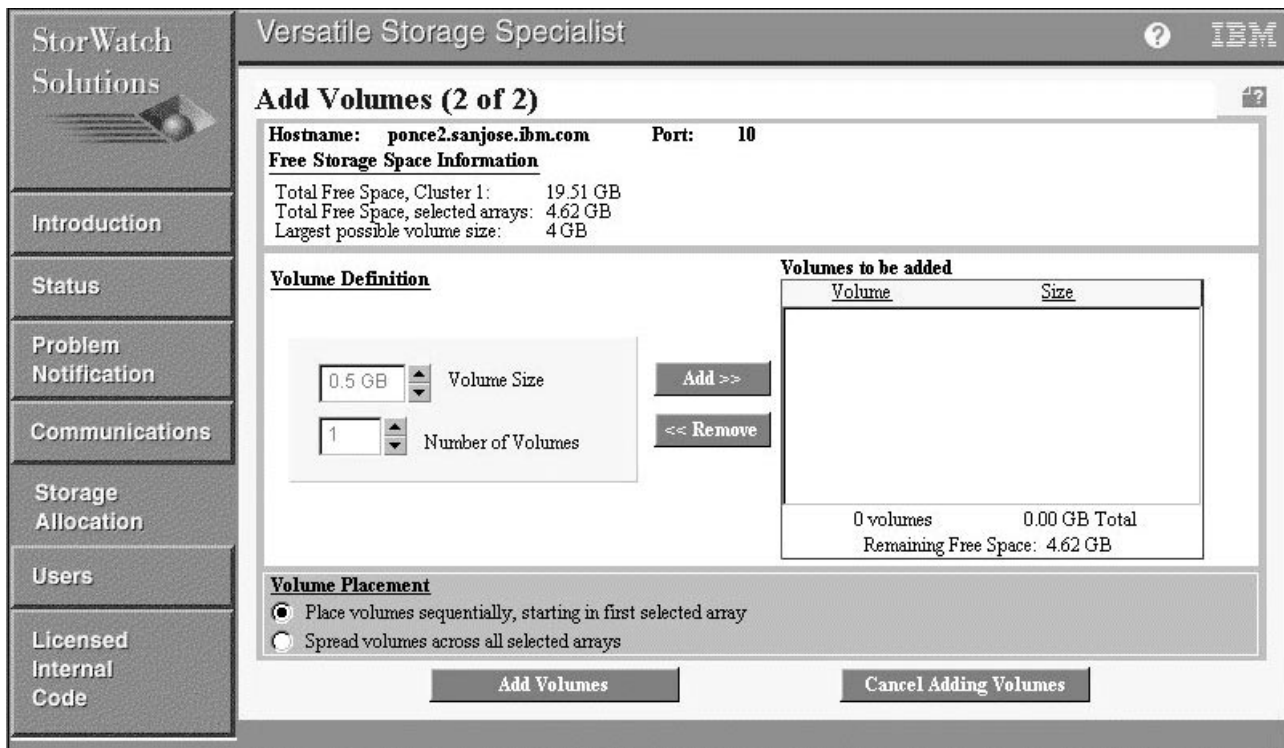


Figure 32. Vs Specialist: Add Volumes (2 of 2) Panel

### 3.5.6.4 Modify Volume Assignments

Click on the **Modify Volume Assignments** button in the Storage Allocation -- Graphical View panel to get to the Modify Volume Assignments panel (Figure 33 on page 51).

This panel is a table with six columns and multiple rows. Each row describes a volume. The reported information for each volume is:

1. Volume: The volume name is the LUN serial number if the volume has already been assigned to a host port. If the volume has not been assigned to a port, the volume shows up as *Unassigned*. The LUN serial number is a VSS generated value created by using the logical volume ID and a portion of the VSS machine serial number. Thus you can create unique names for each volume in each VSS.
2. Location: The array number and the volume number indicate the position of the volume within the array.
3. Size/Type: The size of the volume expressed in GB. If the host to which the volume is attached is an AS/400, the system-attached volume type is shown in this field.
4. Cluster: Contains the cluster number (0 or 1 ) to which the volume belongs, based on where the drawer/array is physically installed.
5. Port: The SCSI port to which the volume is assigned. If a volume is assigned to multiple ports, it is shared among hosts. To identify the volumes defined to more than one host, use the sort function on the Volume column to order the table by volume name and easily identify the volumes with the same name.
6. Host: Displays the host system name associated with the volume



Two actions are allowed in this panel:

- Add selected volume(s) to port
- Remove selected volume assignments

These two actions are available through the two action buttons at the bottom of the panel. The selection of one of the two actions is activated only when at least one volume has been selected in the table. To select a volume click on the desired row. Multiple selections are allowed. Holding down the Shift key on the keyboard, click and release the right mouse button on each row. For a selection of multiple contiguous rows, hold both the Shift key and the right mouse button while dragging the mouse pointer on the rows to be selected.

To add a volume to a host port, highlight the **Add selected volume(s) to port** button. Indicate the port to which the volumes will be assigned by selecting one of the available ports in the scrollable field. To remove a selected volume, highlight the button at the right of the **Remove Selected Volume Assignments** button. To finalize your selections click on **Perform Configuration Update**. To cancel, or exit from the panel, and not perform any action, click on **Cancel Configuration Update**.

To add a selected volume to a port, the list of ports to choose from is limited to those ports that have affinity to the cluster that owns the volumes, is compatible according to Host Type, and has space in the port storage map. If a port is not on the list, it may be because the SCSI IDs are assigned to Other Host System or Device.

**Volume Assignments**

no sort | no sort | no sort | no sort | no sort | no sort

Volume	Location	Size/Type	Cluster	Port	Hosts
3030304641303336	Array 00/ Vol 000	24.0 GB	0	04	ponce.sanjose.ibm.com
3030314641303336	Array 00/ Vol 001	24.0 GB	0	04	ponce.sanjose.ibm.com
Unassigned	Array 00/ Vol 004	04.0 GB	0	n/a	
3030324641303336	Array 01/ Vol 002	28.0 GB	0	07	ponce.sanjose.ibm.com
3030334641303336	Array 01/ Vol 003	28.0 GB	0	07	ponce.sanjose.ibm.com
Unassigned	Array 01/ Vol 005	04.0 GB	0	n/a	
3034444641303336	Array 08/ Vol 000	24.0 GB	0	12	ponce2.sanjose.ibm.com
3034454641303336	Array 08/ Vol 001	24.0 GB	0	12	ponce2.sanjose.ibm.com
3035374641303336	Array 09/ Vol 002	28.0 GB	0	15	ponce2.sanjose.ibm.com
3035384641303336	Array 09/ Vol 003	28.0 GB	0	15	ponce2.sanjose.ibm.com
3030414641303336	Array 05/ Vol 002	28.0 GB	1	05	ponce.sanjose.ibm.com
3030424641303336	Array 05/ Vol 003	28.0 GB	1	05	ponce.sanjose.ibm.com
3030384641303336	Array 04/ Vol 000	24.0 GB	1	06	ponce.sanjose.ibm.com
3030394641303336	Array 04/ Vol 001	24.0 GB	1	06	ponce.sanjose.ibm.com

**Action**

Add selected volume(s) to port

Remove selected volume assignments

Perform Configuration Update | Cancel Configuration Update

Figure 33. VS Specialist: Modify Volume Assignments Panel

### 3.5.7 Users Panel

**Note:** This panel is for information purposes only. No userid information can be changed from this panel. All userid changes must be performed via the CE service function.

The Users hypertext link on the VS Specialist main panel takes you to the User Administration Panel (Figure 34 on page 53). The panel is a table with five columns:

1. Username
2. Password
3. Access level
4. IP Address range (optional)
5. User Comment (optional)

The username is the name given to the user during the initial install from the service terminal.

Users can have three different levels of authorization and can be restricted to accessing the VS Specialist from a specific IP address or a range of IP addresses. If you choose to restrict users to specific IP addresses, you may have to create alternative userids from other IP addresses in case part of the IP network is unavailable. With this security restriction, you may not be able to connect to your VS Specialist if you are connecting through a provider that assigns to you a temporary randomly selected IP address (DHCP - Dynamic Host Configuration Protocol).

User comments are useful when the given usernames are not self-explanatory and do not allow fast recognition of the person accessing the machine. We recommend using the User Comments field to record a contact name for and telephone number of the username.

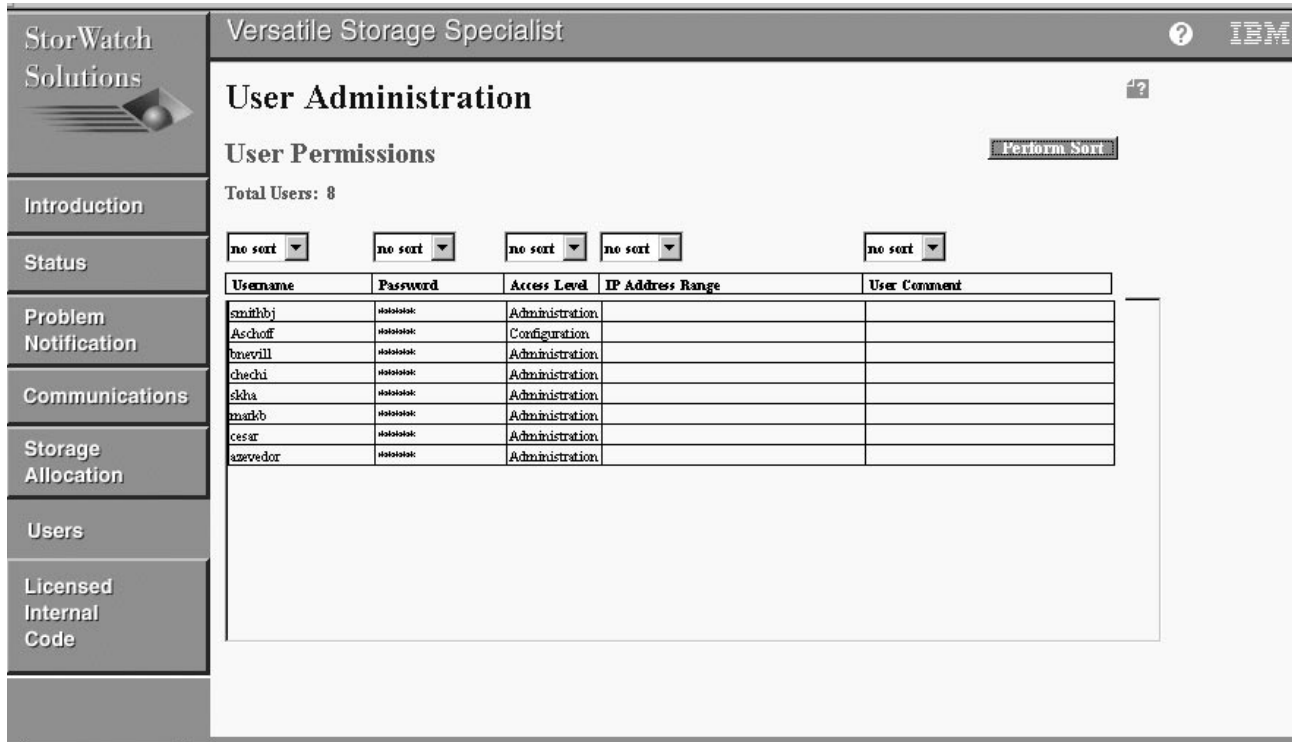


Figure 34. VS Specialist: The User Administration Panel

The User Administration panel allows sorting of all fields, and a maximum of three fields can be sorted concurrently.

The Sort pull-down list for each column has four entries (see Figure 35 on page 54):

1. No Sort
2. First
3. Second
4. Third

No Sort means that column will not be sorted.

First, second, and third represent the sequence of sorting of the column. Say you want to sort both the Username column and the Access Level column. You can select which to sort first, and which second. The column selected first will be sorted in descending order, and the second column will be sorted without altering the sort sequence of the column sorted first.

When you select one of the options in the sort pull-down list, with the exception of "no sort", that option will not be available in the other sort pull-down lists. It is not possible (in fact it is logically incorrect) to have two columns sorted as first, so this condition is avoided by the java applet.

When you have selected the columns to be sorted, click on the **Perform Sort** button to perform the sort function.

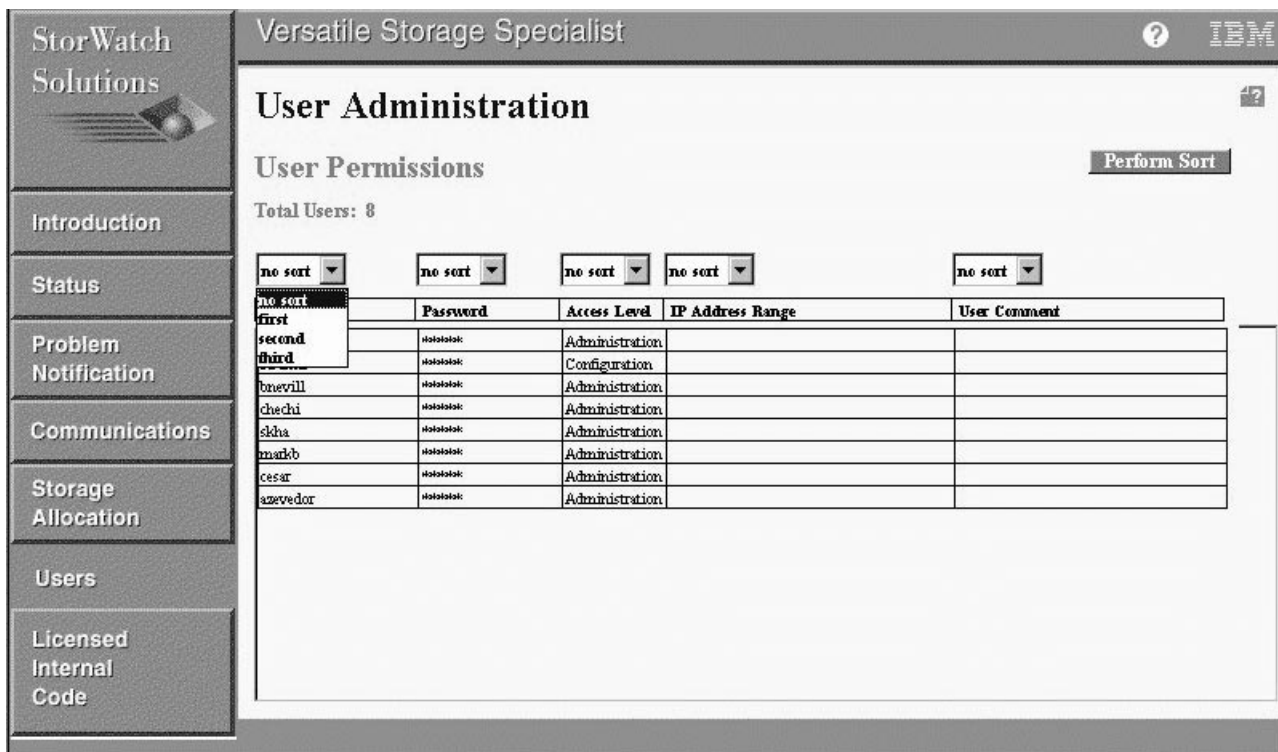


Figure 35. VS Specialist: Sorting Columns in the User Administrator Panel

### 3.5.8 Licensed Internal Code Panel

On the VS Specialist main panel, click on **Licensed Internal Code** to view a list of the levels of LICs in cluster controller 0 and cluster controller1.

**Note:** This is an information panel only.

In each cluster controller section, as shown in Figure 36 on page 55, the LIC sources are:

- Active LIC - The current and active version of LIC that is controlling the operation of the appropriate cluster
- Next LIC - A newer level of LIC, which has been copied to the cluster processor but has not been yet activated
- Previous LIC - An older level of LIC, which is no longer active and has been replaced by the active version. We recommend keeping at least one back-level version of the code. If a newly activated LIC causes problems, it is then possible to revert back to the old code and avoid any outages.

The table on the LIC panel has four columns:

1. LIC Source - One of the five LIC types listed above
2. Version - The version of the LIC in a unique standard format
3. Activation Date - The date of activation (only for LIC source types of Active and Previous)
4. Additional Info - A hypertext link to the README file associated with that LIC (see Figure 37 on page 56).The README file typically contains more detail

about the specific level of LIC, installation details, and other useful information.

**StorWatch Solutions** Versatile Storage Specialist ? IBM

### Licensed Internal Code ?

**Cluster Controller 0**

LIC Source	Version	Activation Date	Additional Info
Active LIC	SA80922	1998/09/24	<a href="#">README</a>
Previous LIC	SA80915	1998/09/23	<a href="#">README</a>
Next LIC	SA80922	1900/01/01	<a href="#">README</a>

**Cluster Controller 1**

LIC Source	Version	Activation Date	Additional Info
Active LIC	SA80922	1998/09/24	<a href="#">README</a>
Previous LIC	SA80915	1998/09/23	<a href="#">README</a>
Next LIC	SA80922	1900/01/01	<a href="#">README</a>

Figure 36. VS Specialist: Licensed Internal Code Panel

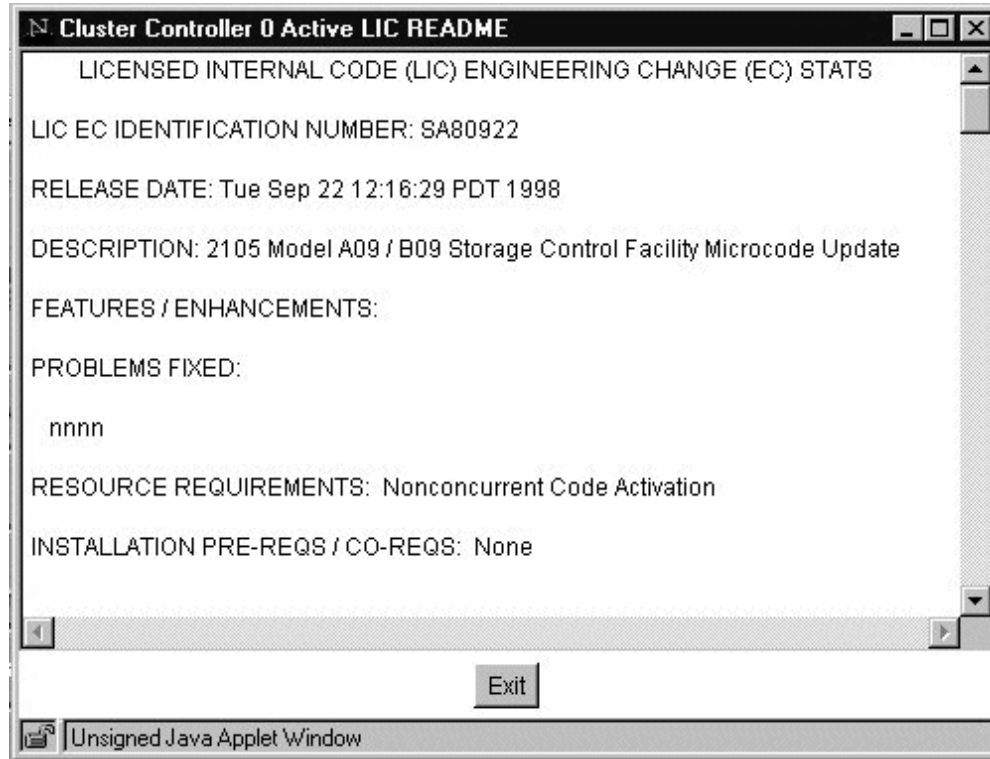


Figure 37. VS Specialist: README File for an LIC Source

---

## Chapter 4. Host Connections

In this chapter we show how to manage the host configuration through the use of the VS Specialist.

---

### 4.1 Defining a Host

A host must be defined to the VSS in order to access VSS disk space. You define a host through the VS Specialist Modify Host Systems subtask panel.

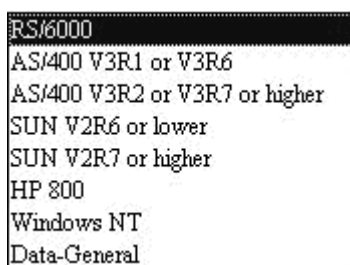
Before defining a host, you must know the kind of host you are going to define. You also need the full hostname and domain name. We recommend using the system's IP host name, but the code does not attempt to communicate with the host to verify the name.

The list of supported host platforms changes frequently as new support is added over time. For a complete list of supported host types, see:

<http://www.storage.ibm.com/hardsoft/products/vss/vssattach.htm>

An IBM CE can modify the Host Types list shown in Figure 38 from the service terminal. The modified host types entered in the list are saved in a custom table inside the VSS, and, when you refresh the VS Specialist Modify Host Systems panel, the custom defined host appears in the available Host Types list.

The Host Type list defines a set of information, such as the number of LUNs that can be created by a specific host type, for all VSS-attached hosts. You can use the Host Type list associated with a host to ensure that the operations carried out through the VS Specialist on that host do not violate any of the platform-specific architectural limits.



```
RS/6000
AS/400 V3R1 or V3R6
AS/400 V3R2 or V3R7 or higher
SUN V2R6 or lower
SUN V2R7 or higher
HP 800
Windows NT
Data-General
```

Figure 38. Host Types List

The VSS provides SCSI III level support. Therefore, from an addressing point of view, this means that 16 SCSI target addresses are supported (0-F) and 64 LUN addresses are supported per target address. SCSI II supports 16 targets but only 8 LUNs per target.

An LUN is an addressing extension to the SCSI target address to allow definition of more devices on the SCSI bus. SCSI II provides 128 possible addresses (16\*8), whereas SCSI III provides 1024 (16\*64) possible addresses.

The VSS microcode supports LUNs 0-63, but the number of LUNs actually used is controlled by a Max\_Lun variable in the Host Type table. Max\_Lun has a range of

0-63, but currently there are no Host Types that support access to more than 32 LUNs.

This is the main reason why it is critical to choose the correct host bus type for the connection. Currently, AIX versions supported by the VSS support 32 LUNs per target whereas current versions of HP-UX, Sun OS, and Windows NT support only 8, the SCSI II standard. Therefore, if you attach 16 volumes to an RS/6000, one target address with 16 LUNs will be generated. For all other host types, two target addresses with 8 LUNs will be generated. If volumes are moved between these two host types, the VS Specialist automatically re-creates the address range used on the basis of the new host type.

SCSI ports can be shared among hosts. Up to four Windows NT servers or four AIX servers can share the same port. Only two HP-UX servers can share the same port at the same time. Sun provides the capability to provide only one host per port. The VSS has 16 SCSI ports and allows up to 64 host connections.

To define a host in VSS, select the **Modify Host Systems** button in the Storage Allocation -- Graphical View panel of the VS Specialist (see Figure 24 on page 39).

The list of available host systems that the user creates using the Modify Host Systems option is a preliminary step for SCSI port configuration. A host system in the list does not become part of the stored machine configuration until it is assigned to a SCSI port and SCSI ID as an initiator, at which time the host system icon appears on the Storage Allocation -- Graphical View panel of both clusters.

To add a host, on the Modify Host Systems panel (Figure 28 on page 44) type in the hostname in the Host Name field and select the **Host Type** from the pull-down list. Click on the **Add** button. The hostname appears in the List of Defined Hosts but has not yet been defined. You can add as many hosts as you want (up to 64), and you can remove any of the already defined hosts that are no longer needed by selecting the hostname from the List of Defined Hosts box and clicking on the **Remove** button. To activate the changes, click on the **Perform Configuration Update** button. If you want to return to the Storage Allocation -- Graphical View panel without activating any of the changes, click on the **Cancel Configuration Update** button.

The defined host is now available to access data inside the VSS.

**Note:** When you define a host through the VS Specialist, it is visible only in the cluster to which the VS Specialist Web interface is connected until a configuration change is made to connect the host to a SCSI port. The information server, by design, does not propagate the simple host definition to the other cluster's object data manager (ODM) database. If you are using cluster 0, for example, you have to define the host port connecting the VS Specialist to the same cluster where the host was defined. When the host is connected to a port through the Configure Host Ports task panel, the new host information is propagated to the other cluster when the configuration is updated.

### 4.1.1 Examples

Figure 39 on page 59 shows how to add a host, and Figure 40 on page 60 shows how to remove a host. Both figures show the sequence of operations to be



performed. They do not represent a real screen; they simply show that you have to select or fill in the requested information.

**Modify Host Systems : Add a Host Example**

1) Type the hostname

---

2) Choose the host type

Host Type

- Windows NT
- RS/6000
- AS/400 V3R1 or V3R6
- AS/400 V3R2 or V3R7 or higher
- SUN V2R6 or lower
- SUN V2R7 or higher
- HP 800
- Windows NT
- Data-General

---

3) Click on Add

---

4) Click on Perform Configuration Update or Cancel Configuration Update

---

5) Expect this message

**! 1502: The Host System list was successfully modified.**

To complete the configuration of newly added hosts, they must be assigned to one or more ports using the Configure Host Ports screen.

Unsigned Java Applet Window

Figure 39. Adding a Host

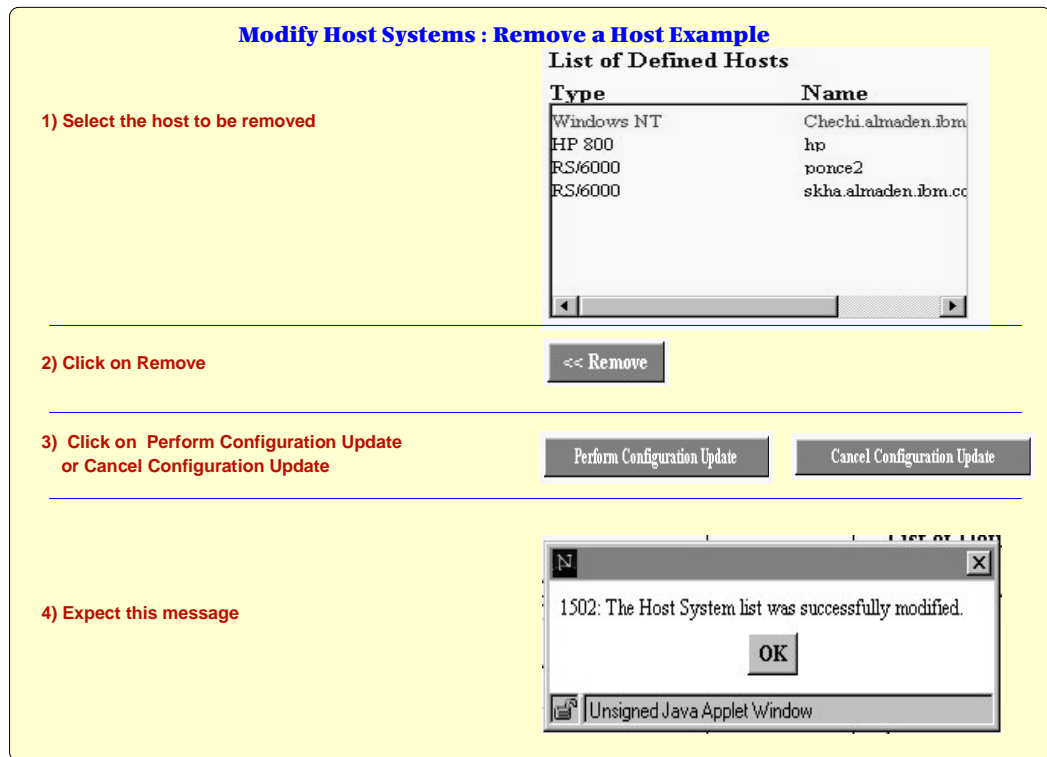


Figure 40. Removing a Host

## 4.2 Connecting a Host to a Port

Once you have defined the host, you must connect it to the VSS. The connection is made by connecting the host to a SCSI adapter. Each SCSI adapter has two ports: A and B. The physical connection is to be made before defining the connection to VSS through the VS Specialist Configure Host Ports task panel. When defining a host port, you have to know the port to which the physical connection has been made. You also have to know the host name, the host type to be used for the bus configuration, and the SCSI ID of the host adapter (normally 7).

The task of connecting a host to port involves notifying VSS that a specific host, or hosts, is physically connected to one of the available SCSI ports. Also, the bus configuration is requested for the host type that is going to be defined to this port. This information is required again because, if you add a different host to the same port, you are given the option to choose the bus configuration again. We recommend that you not connect different host types to the same port. If, for some reason, you have to share a port and assign to it two or more different host types, choose a bus configuration that is compatible with both hosts defined to the port.

The ports have a default bus configuration set to RS/6000. If the port has a different bus configuration from that of the host type you are going to define, you have to change the bus configuration in order to see the list of the host types in the Available Hosts box on the panel.

We recommend updating the Configure Host Ports panel with this information:

1. Select the SCSI port.
2. Select the Bus Configuration. If this is different from the configuration the bus has, and no other hosts have been assigned to this port, select the bus configuration of your new host. If the actual bus configuration is different and is the bus configuration of a host that is already assigned to the port, be careful and select a bus configuration that is compatible with both hosts. This may be a Customized Host Type, to fit the architectural specification of one or more different platforms. When you change the bus configuration, a warning message appears in a window (see Figure 41 on page 62).
3. Select a host. When you select the bus configuration, all Host Types of that category are displayed in the Available Hosts list box. Select the host you want to add.

**Note:** Customized Host Type definitions can be made *only* by the IBM CE through the service terminal.

4. Assign a SCSI ID to the host on the port.

You must execute steps 1 and 2 in sequence. If a different port from the port you meant to assign to your host is selected in the panel, when you change the bus configuration, the port bus configuration changes.

The SCSI ID you are prompted for in the Configure Host Ports panel is the SCSI ID of the adapter residing in the host. The adapter in the host comes with a default address which is usually set to a specific value. Typically for single host SCSI adapters, the SCSI ID is 7, but we recommend consulting the documentation supplied by the vendor of your adapter to check the address and for instructions on how to change it.

When a host port is first assigned to a SCSI port on the VSS, the possible SCSI target IDs for this VSS port may show a status of Unrelated Host or Device (see Figure 41 on page 62). This status must be changed before you can connect a host and use the SCSI IDs for volume assignments. Select one or more of the target IDs with a status of Unrelated Host or Device and click the **Remove** button. Typically you should remove all target IDs in this way and then add the host target ID, normally SCSI ID 7, as in Figure 43 on page 63. This will reserve this SCSI ID for the host adapter and make all other SCSI target and LUN addresses available for volume assignments on this adapter.

You can control volume assignment to SCSI IDs by using the Unrelated Host or Device assignment to prevent certain SCSI IDs from being used by the VSS. For example, you may have assigned the host system initiator to SCSI ID 7 and may want the LUNs to be assigned to SCSI ID 6. The VSS will normally assign the first volume to SCSI ID 0 if it is not reserved, so in order to force volume assignment to ID 6 the SCSI IDs 0-5 should be assigned to Unrelated Host or Device. The Unrelated Host or Device assignment may be removed at any time to free up additional selected SCSI IDs for volume assignment.

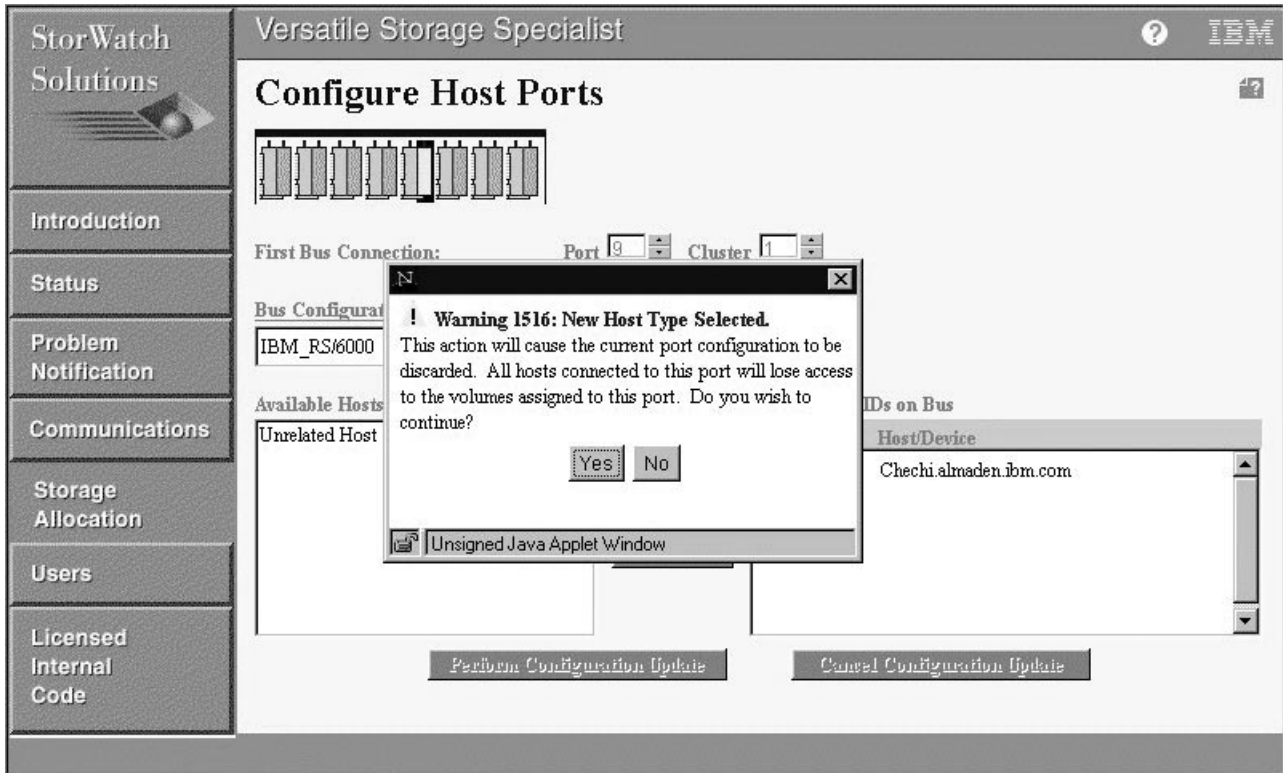


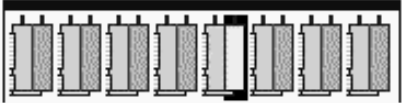
Figure 41. Selecting a New Host Type for a Port

#### 4.2.1 Examples

Figure 42 on page 63 and Figure 43 on page 63 show how to add a host to a port and Figure 44 on page 64 shows how to remove a host from a port. None of the figures shows the VS Specialist panel in its entirety. Instead, they are a simplification of the steps necessary to configure host ports.

### Configure Host Ports : Add a Host to a Port Example


1) Select the port by clicking on the icon  
or  
Select the number of the port, using the spin buttons



First Bus Connection: Part  Charter

---

2) Select the bus configuration (host type)




---

3) Expect this message

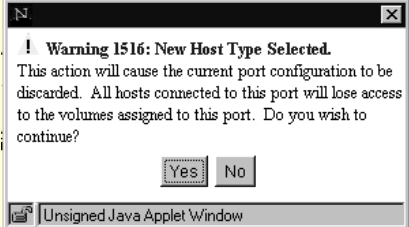



Figure 42. Adding a Host to a Port: Part 1

4) Select the host





---

5) Click on Add

---

6) Choose the SCSI ID for the host bus




---

7) Click on Perform Configuration Update or Cancel Configuration Update

---

8) Expect this message




Figure 43. Adding a Host to a Port: Part 2

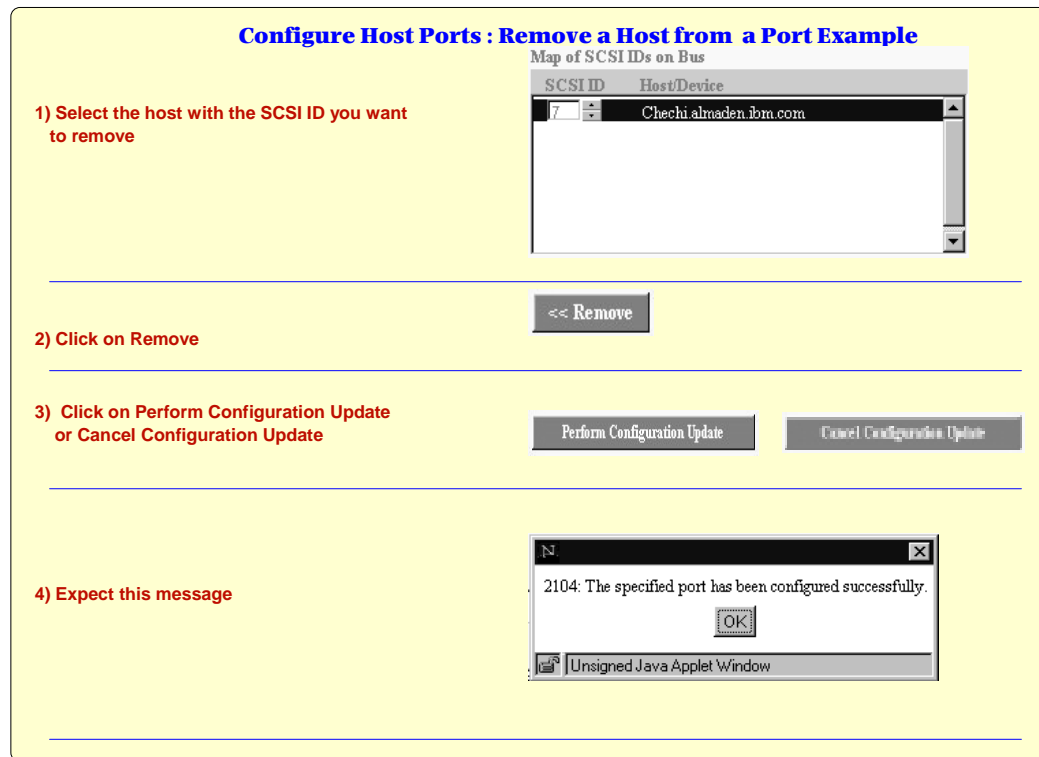


Figure 44. Removing a Host from a Port

### 4.3 Managing Logical Volumes on the VSS

In order for the hosts to access data inside the VSS, the VSS storage area must be partitioned into *logical volumes*. Each logical volume is seen by the host as a SCSI disk, even though the logical volume is physically allocated on all disks of one of the RAID 5 arrays of the drawer. A logical volume cannot span arrays, but you can create several logical volumes in different arrays, connecting all of them to the same host through the same host port.

The logical volumes can be of different sizes, from 0.5 GB up to 56 GB. For the AS/400, only 4.19 GB and 8.59 GB sizes are allowed (see "Host Interface Adapter" on page 11). The logical volume size is restricted to one of the following values: 0.5, 1, 2, 4, 8, 12, 16, 20, 24, 28, 32, 40, 48, 56. Each logical volume is mapped as a SCSI disk to a host and has a target SCSI ID address and an associated LUN. The maximum number of LUNs per SCSI ID that the VSS allows depends on the host type and is 32 for RS/6000 and 8 for all other host types.

The SCSI adapter inside the VSS to which the host is connected does not report itself as a target SCSI ID. All of the logical volumes that this adapter sees are addressed using target ID and LUN. Theoretically, 480 logical volumes can be created for each AIX or Windows NT host (16 IDs minus 1 used by the host adapter multiplied by 32 LUNs) and 120 (16 IDs minus 1 multiplied by 8 LUNs) for all other hosts.

Once created, a logical volume must be assigned to a defined host. It can later be reassigned to other hosts or unassigned for future use.

To create a new logical volume, select the **Add Volumes** button on the Storage Allocation -- Graphical View panel of the VS Specialist. A new panel shows the hosts, ports and arrays defined in VSS. Select the host, then select one port already assigned to that host (once you select the host, all of its assigned ports are highlighted), and then select one or more arrays where the volumes will be created. When a port is selected, the arrays available for selection are based on the cluster affinity for that port (that is, if the port belongs to cluster 0, the arrays from cluster 0 are highlighted). The VS Specialist requires that the selection be done in the order indicated above. An error sound is generated if you select an object out of sequence, and the selection is ignored.

If you select a host or a port and then want to change your selection, before clicking on the **Next** button, you can click on the **Clear View** button and start again. If you select an incorrect array, you can deselect it simply by clicking on it.

When the selection is complete and the host, port(s), and array(s) are highlighted as expected, click on the **Next** button to complete the Add Volume function. The Add Volume (2 of 2) panel is presented. Here you select the desired volume size for the logical volume and the number of logical volumes. You can add or remove several volumes at once, and you can add volumes of different sizes.

If you have selected more than one array, you can choose where the volume(s) should be placed. If you have created more than one volume, you can decide to create the volumes sequentially on the arrays, spreading the allocation to all of the selected arrays, that is, the first volume on one array, the second volume on another array, and so on. Otherwise all volumes are placed on the first available array until it is full. Then another array is selected, filled up, and so on for all volumes and arrays until all of the volumes are allocated. To physically allocate the volumes, click the **Add Volumes** button. See Chapter 3.5.6.3, "Add Volumes" on page 46 for full details on creating volumes.

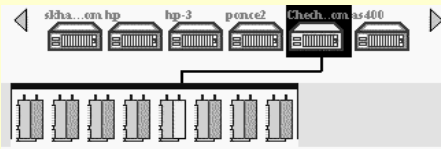
**Note:** The Add Volume function only allows you to create new volumes; it does not allow you to remove volumes created in a previous process. Volumes that are in the Volumes to be added list but not yet created can be removed, but once the volumes are physically allocated, they cannot be removed. You can however, unassign them through the Modify Volume Assignments task panel (see Figure 33 on page 51).

### 4.3.1 Examples

Figure 45 on page 66, Figure 46 on page 66, and Figure 47 on page 67 show the sequence of adding a volume to a host. None of the figures shows the VS Specialist panel in its entirety. Instead, they are a simplification of the steps necessary to add volumes to a host.

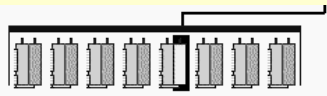
### Add Volumes : Define volumes for a Host

**1) Select the host to which the volume is to be added.**  
A solid line is displayed to show the path to the host port(s).




---

**2) Select one of the host ports the host will use to access the volumes.**




---

**3) The cluster to which the port has affinity is highlighted.**  
Select any of the arrays with available space (scaled in green color) to which to allocate the volumes.

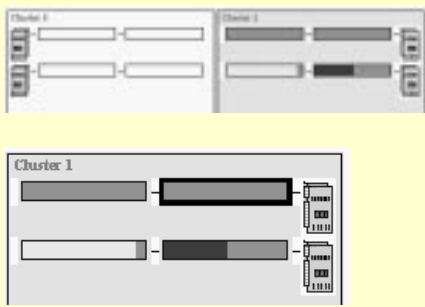
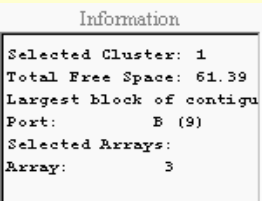


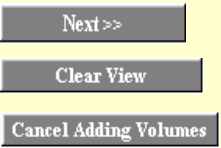
Figure 45. Adding Volumes: Part 1

**4) Check that the Information box reports the correct selection information.**



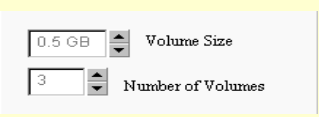

---

**5) Click on Next to proceed with adding volumes in the Add Volumes (2 of 2) panel, or click on Clear View to deselect all the objects and start over again, or click on Cancel Adding Volumes to return to the Storage Allocation – Graphical View panel.**




---

**6) Select the volume size, and the number of volumes to be created.**  
The remaining free space is shown at the bottom of the Volumes to be added box.




---

**7) Click on the Add button to add the volumes to the volumes to be added box.**




Figure 46. Adding Volumes: Part 2



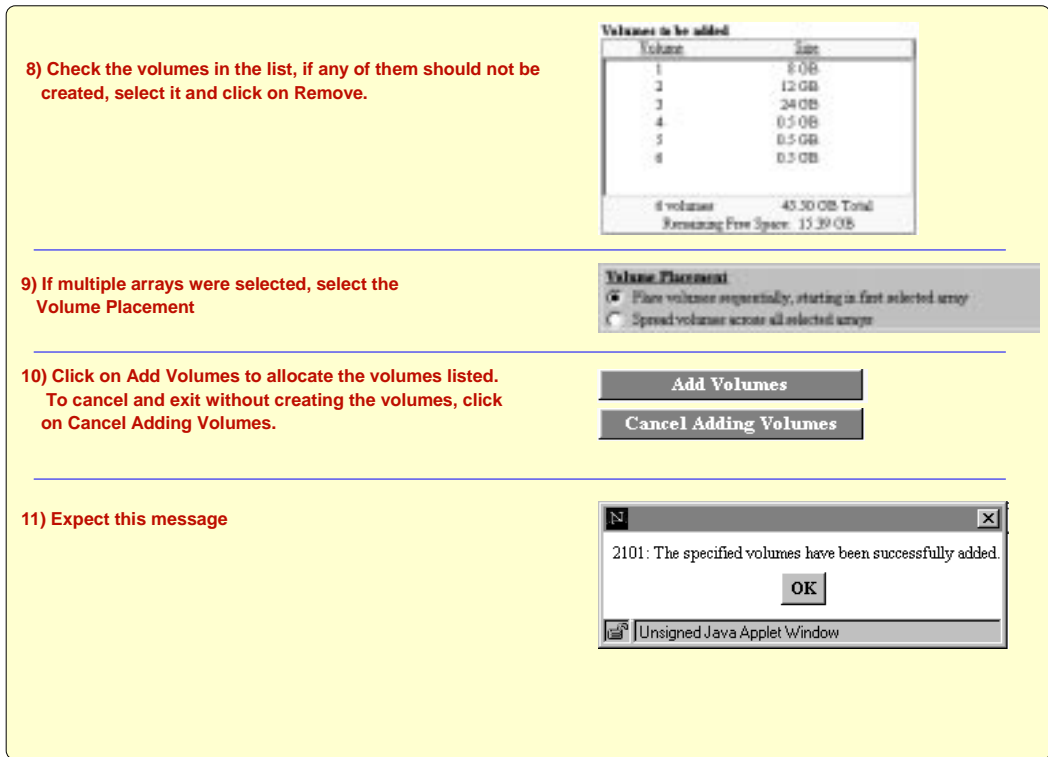


Figure 47. Adding Volumes: Part 3

## 4.4 Modifying Volume Assignments

Through the VS Specialist Modify Volume Assignments panel, you can add a logical volume to a host or remove the host from accessing it. A volume can be accessed by more than one host. When you define a volume in the array, you have to specify the host that will access it. When volumes are initially defined, they have a host accessing them. With the Modify Volume Assignments operation you can define a host to access the same volume that another host is using. You can also remove a host from using that volume.

The information related to the status of the volume is kept in the VSS even if the volume is not accessed by any host. The VSS creates only a small portion of space in its storage for this purpose. The volume, once created, cannot be physically deleted. The access to it can be removed, but the data remains on the volume until someone who has authority to it deletes the data.

To remove logical volumes from an array, you have to reformat the array. Once you reformat the array, all data and all logical volumes on that array will be lost. You can then create new logical volumes.

If more than one host is accessing a volume, the Modify Volume Assignments panel shows the same volume twice, once for each host port. If more than one host is using the same host port, it is necessary to check the hostname, because the host port number is the same for both hosts.

To remove a volume, select it in the list, checking the hostname, port, array, and volume number, click on the **Remove selected volume assignments** action button, then click on the **Perform Configuration Update** button. It is important to keep track of the array and volume number when removing access to a volume, because, when a volume is unassigned to all hosts, the volume name is removed as well. Only the volume number within the array remains, and that is the information needed to identify the correct volume. It could be difficult to identify a volume in a list with quite a few entries, especially if most of the volumes have the same size.

To add a volume, you must know the host port to which the host is to be connected. Select the volume, either unassigned or already assigned to a host, and then click on the **Add selected volume to port** action button, select the port to which this volume must be connected, and click on the **Perform Configuration Update** button.

If a volume is to be accessed by more than one host, the hosts should be architecturally compatible. The format used to write the data on disk, that is, the file system structure, should be the same. Data corruption is likely to occur if an incorrect configuration is defined.

#### 4.4.1 Examples

Figure 48 shows how to add a volume to a host port, and Figure 49 on page 69 shows how to remove a volume from a host port. Neither figure shows the VS Specialist panel in its entirety. Instead, the figures are a simplification of the steps necessary to modify volume assignments.

**Modify Volume Assignments: Add a Volume to a Host Port**

1) Select any volume(s) (unassigned or already assigned) for reassignment

Volume	Location	Size Type	Cluster	Port	Status
Unassigned	Array 001 Array 001	100 GB	1	001	OK
Array 001A Array 001A	Array 001 Array 001	100 GB	1	001	OK
Array 001B Array 001B	Array 001 Array 001	100 GB	1	001	OK
Array 001C Array 001C	Array 001 Array 001	100 GB	1	001	OK

---

2) Select the action of your host in the Add selected volume(s) to port field

**Action**

Add selected volume(s) to port 001

Remove selected volume assignments

---

3) Select the port to which your host is attached

01

- 01
- 02
- 05
- 06
- 10
- 13
- 14

---

4) Click on Perform Configuration Update or Cancel Configuration Update

Perform Configuration Update

Cancel Configuration Update

---

5) Expect this message

2136: The specified volumes have been successfully assigned to the specified port.

OK

Unsigned Java Applet Window

Figure 48. Adding a Volume to a Host Port



## 4.5.2 Physical Structure

Table 3 on page 70 shows the ports set with their affinity to the local cluster.

**Note: This is not the affinity we suggest. We highly recommend having the affinity shown in Table 4 on page 70.**

Table 3. *SCSI Ports: Local Affinity*

SCSI Adapter	Bay	Port	Port Address	Physical Cluster	Affinity Cluster
1	4	0	A - 600	0	0
1	4	1	B - 601	0	0
4	4	8	A - 630	0	0
4	4	9	B - 631	0	0
1	5	2	A - 700	1	1
1	5	3	B - 701	1	1
4	5	10	A - 730	1	1
4	5	11	B - 731	1	1
1	8	4	A - 400	0	0
1	8	5	B - 401	0	0
4	8	12	A - 430	0	0
4	8	13	B - 431	0	0
1	9	6	A - 500	1	1
1	9	7	B - 501	1	1
4	9	14	A - 530	1	1
4	9	15	B - 531	1	1

## 4.5.3 Setting the Affinity

As shown in Table 4 on page 70, all B ports of each SCSI adapter have their affinity set to the remote cluster. This is the recommended setting.

Table 4. *SCSI Ports: Recommended Affinity (Remote)*

SCSI Adapter	Bay	Port	Port Address	Physical Cluster	Affinity Cluster
1	4	0	A - 600	0	0
1	4	1	B - 601	0	1
4	4	8	A - 630	0	0
4	4	9	B - 631	0	1
1	5	2	A - 700	1	1
1	5	3	B - 701	1	0
4	5	10	A - 730	1	1
4	5	11	B - 731	1	0
1	8	4	A - 400	0	0

SCSI Adapter	Bay	Port	Port Address	Physical Cluster	Affinity Cluster
1	8	5	B - 401	0	1
4	8	12	A - 430	0	0
4	8	13	B - 431	0	1
1	9	6	A - 500	1	1
1	9	7	B - 501	1	0
4	9	14	A - 530	1	1
4	9	15	B - 531	1	0

#### 4.5.4 Changing the Affinity

There are two methods of changing the affinity. At installation time, the IBM CE should have set the affinity as shown in Table 4, through the service terminal.

If during the initial hardware install, the affinity is not set as per the recommendation, it should be changed before defining host ports and volumes.

**Note: If the affinity is changed after volumes have already been defined on the array to which the port is connected, volume access is lost.**

The second method of changing the affinity is to use the VS Specialist Configure Host Ports task panel. This is the method the VSS administrator should use.

From the Configure Host Ports panel, perform the following:

1. Select the icon of the port you want to change or select the port using the Port spin button in the First Bus Connection field.
2. Be sure that no hosts have been defined to this port. If any hosts have been defined, other than Unrelated Host or Device, remove them, keeping in mind that this operation disconnects the host from the volumes. If you proceed with this operation before you remove the hosts, they are removed anyway and access to the data is lost.
3. Select the new Cluster number. Refer to Table 4 on page 70 for the correct affinity selection.
4. A warning message appears (see Figure 50 on page 72).
5. Click on the **Yes** button in the warning message window if you want to continue. Otherwise click on the **No** button to cancel.
6. Select an **Unrelated Host or Device** from the Available Hosts list.
7. Click on **Add**.
8. Select a SCSI ID for this Unrelated Host or Device entry. We recommend not using ID 7. The first ID that appears is the best (probably the number zero). See Figure 51 on page 72 for an example.
9. Click on **Perform Configuration Update** to conclude the operation and perform the update. Otherwise, click on **Cancel Configuration Update** to cancel the entire operation and return to the Storage Allocation -- Graphical View panel.

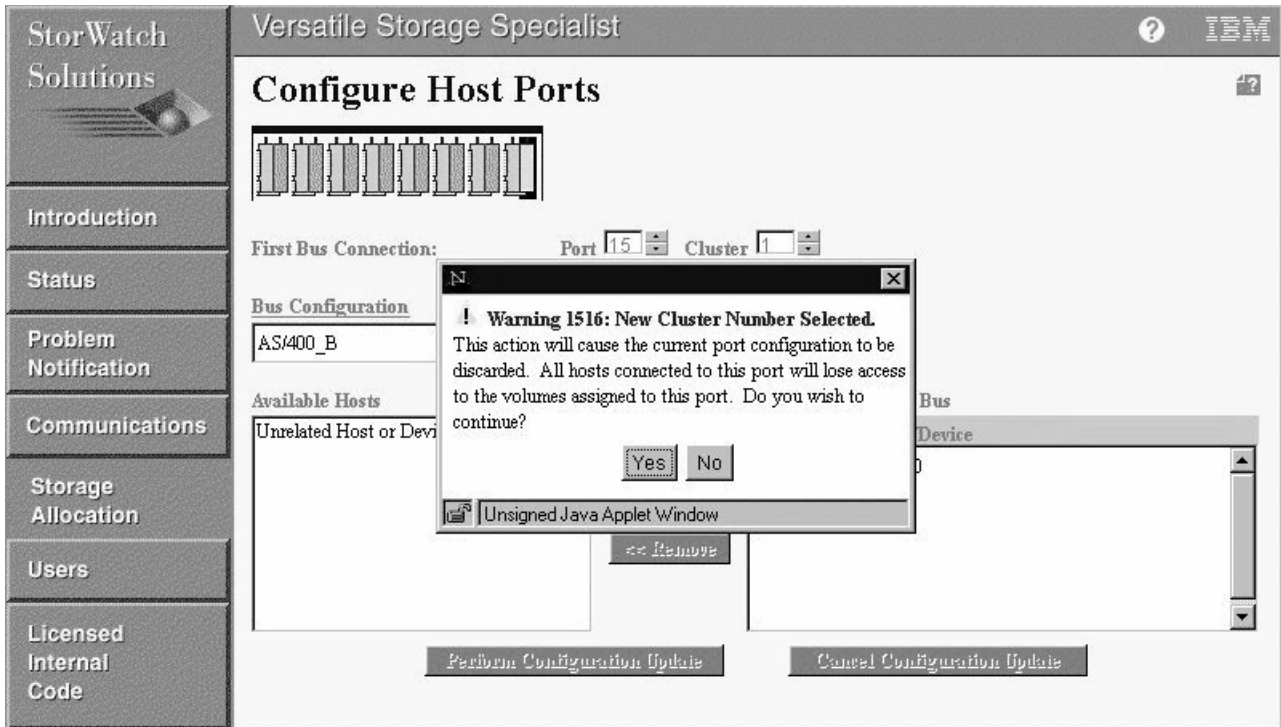


Figure 50. Changing Cluster Affinity: Warning Message

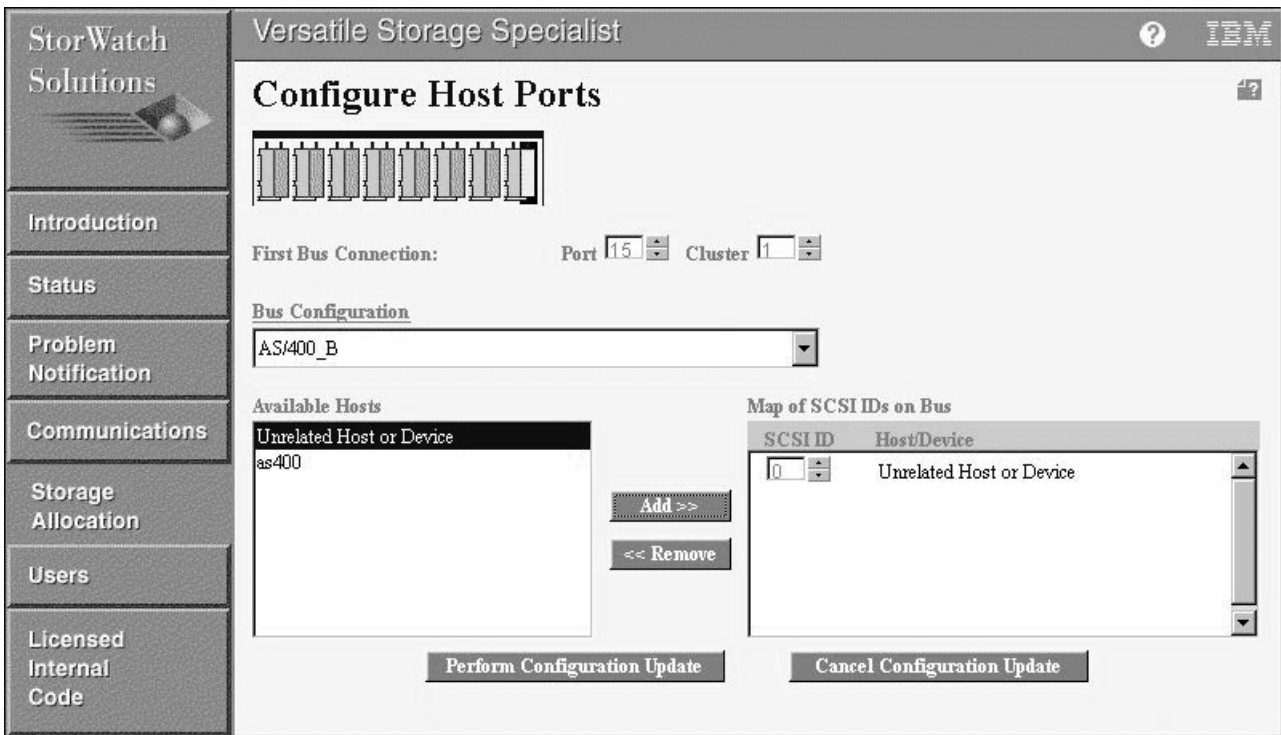


Figure 51. Changing Volume Affinity

Now you can reconnect the volumes to the host to which they were originally connected. You have to configure a port to the host, and then the host to the


cluster where the array containing your volumes resides. Then the volumes can be reassigned to this host.

#### **4.5.5 Example**

Figure 52 on page 74 shows how to change the cluster affinity of a host port.


### Change Cluster Affinity of a Host Port

1) Select the port by clicking on the icon  
or  
Select the number of the port, using the spin buttons



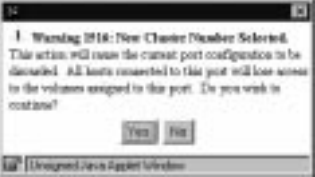

---

2) Change the Cluster number, using the spin button





---

3) Expect this message





---

4) Click on Yes inside the message window



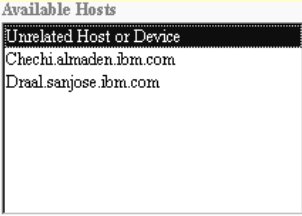

---

5) Expect all host ports to be removed





---

6) Select Unrelated Host or Device from the Available Hosts box





---

7) Click Add





---

9) If you need SCSI ID 0 available, you may change it now.




---

8) Click on Perform Configuration Update




---

9) Expect this message

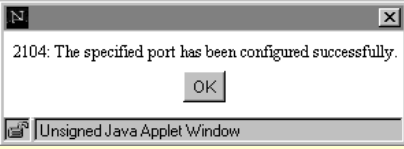


Figure 52. Changing the Cluster Affinity of a Host Port



---

## 4.6 Server Storage Attachment Examples

In this section we give some examples of attaching a volume defined to the VSS to specific host types.

### 4.6.1 IBM AIX

For this test we used an RS/6000 F50, with AIX 4.2.1, connected to VSS through an Ultra SCSI adapter.

The adapter was defined with SCSI ID 7. After we connected the cable between VSS and the adapter and performed the configuration, using the VS Specialist, we configured AIX to recognize the disks.

VSS provides a set of files and a script to be run on AIX to support VSS. These scripts must be run by **root** and execute two main tasks:

1. Create ODM entries for VSS
2. Run **cfgmgr** to make the disks available to AIX

The installation procedures for the script can be found in the *Host Systems Attachment Guide - 2105-B09* manual.

The script to be run is called 2105inst. When executed, it updates ODM to correctly recognize VSS logical volumes (so that AIX can differentiate VSS from a standard SCSI disk) and invokes **cfgmgr** to configure all attached devices to the host system.

Once the script has been installed and executed, the VSS logical volumes are available as **hdisks** on AIX. Standard procedures for creating volume groups, logical volumes, and file systems can then be used.

### 4.6.2 Microsoft Windows NT

For this test we used an IBM PC Server 325 with 64 MB of memory, 3 GB internal hard disk running Microsoft Windows NT Server 4.0 Service Pack 1.

We installed an Adaptec Ultra Wide SCSI adapter (Model AHA2944U/W) in an empty slot of the PC server. We used the default device driver provided by Windows NT for this adapter and restarted the machine. At boot time, the Adaptec device driver showed the SCSI address of our adapter as ID 1.

We modified the VS Specialist to map our host with SCSI ID 1, shut down the Windows NT workstation, powered off, and connected the SCSI cable with address 401 (port 5) to the PC. The host, the port, and the volumes were already defined, using the procedure explained in this chapter. We defined six volumes, three of 0.5 GB size, and three of 8 GB size.

We rebooted Windows NT, started the Disk Administrator Utility, and all of the disks were available. We had to define them and commit the changes (Windows NT does not perform the operation until the changes are committed). We were then able to format the drives and use all of the disks.

### 4.6.3 Sun Example

On a Sun system, attach the VSS SCSI adapter to a valid SCSI adapter in the Sun host. Reboot the Sun host, and the new devices should show up during the boot process with addresses like *Cntndn*, where *Cn* is the controller number, *tn* is the target address provided by the VSS, and *dn* is the LUN used for this volume.

Use the Format command to initialize the disk, and label it. The disk can then be treated as any other SCSI disk, be placed in a volume group, or whatever you choose to do.

### 4.6.4 HP Example

On an HP system, attach the VSS SCSI adapter to a valid SCSI adapter on the HP host. Reboot the HP host or run the INSF command to discover the new disk. The VSS disks will show up as SCSI devices with the physical address associated with the slot the SCSI adapter is in. The volumes can then be partitioned, placed in a volume group, or whatever, just as you would for any other SCSI disk device.

---

## Chapter 5. VSS Availability

In this chapter we describe some availability issues concerning the VSS. We cover component RAS features and guidelines for availability planning.

---

### 5.1 Types of Failure

There are four main categories of possible VSS component failure:

- Host connection failures
- Disk interface failures
- Storage processor cluster failures
- Power failures

VSS is designed to provide data integrity in every failure situation, and data availability in most cases where a single failure occurs. Several components are redundant and provide predictive failure detection, which allows planned maintenance actions. These failures may or may not lead to some problems in data access. Most do not cause an interruption of access, but some may isolate hosts from their disks.

#### 5.1.1 Host Connection Failures

VSS host interface adapters have two SCSI ports that can be connected to either the same host or different hosts. Each port is connected to a storage server cluster, and the SCSI port-to-cluster connection is defined at installation time.

A failure in the host SCSI interface, the VSS host interface adapter, or the cable between them will interrupt access to data from that interface. Because the host interface adapter has two ports, one or more hosts can be affected by such failure. The ongoing I/O operations in the adapter during the failure will not complete. The adapter must be replaced before normal operation is resumed.

You can use disk mirroring across SCSI ports on different host interface adapters to obtain alternative paths to data. Disk mirroring provides two (or more) connections from a given host to copies of the same data inside VSS. Should one interface fail, a mirrored copy of the data can be accessed from another SCSI port.

##### 5.1.1.1 IBM I/O Path Optimizer

It is possible to have two SCSI ports defined to the same host and to the same logical volume inside VSS. However, the host should be able to handle more than one path to the same disks. Software-based failover can be used to provide this functionality. This is the basic function of the I/O Path Optimizer product that IBM has preview announced.

The IBM I/O Path Optimizer takes advantage of redundant paths between a system server and the disk storage inside VSS to optimize availability. The I/O Path Optimizer resides in the server with the disk device driver software for the VSS.

The I/O Path Optimizer provides the intelligence to automatically reroute I/O operations when there is a failure. This eliminates the problem of the data path

between the system server and the VSS disk storage being a potential single point of failure (see Figure 53).

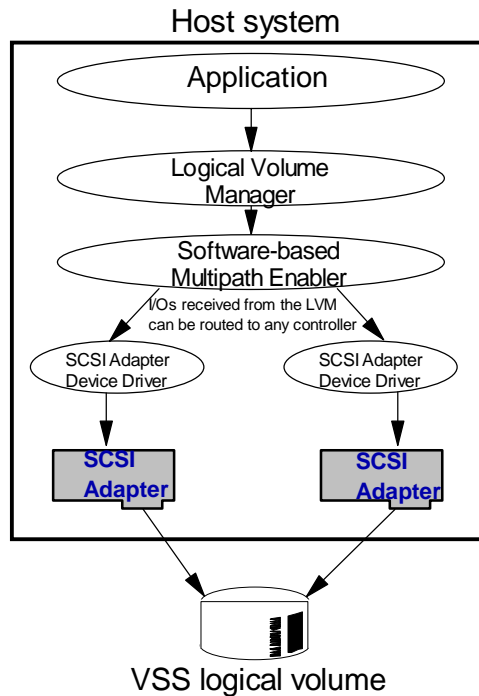


Figure 53. Implementation of Multiple Paths to a VSS Logical Volume

### 5.1.2 Disk Interface Failures

The SSA disk interface used on VSS is a highly reliable adapter, with redundant paths to disks inside one adapter. Each disk can be reached in two ways on the loop, and, if a port fails, the disks can still be accessed by the other port.

If a cable fails, the SSA loop is reconfigured into two strings. After failure, the disks will be accessed by the string to which they are connected. When the problem is resolved, the two strings are automatically reconfigured as a loop. Performance can be reduced when a cable failure occurs, but data integrity and data access are maintained.

The adapter contains a nonvolatile fast write cache (FWC), which is used to reduce penalties in write performance due to RAID-5 parity calculations. Most of the write requests are saved on the nonvolatile FWC (some are written directly to disk, because of some special handling by algorithms inside the adapter). The write requests are first stored in the storage processor cache, then in the adapter DRAM and the nonvolatile FWC. After the VSS has ensured that data is successfully saved in these memories, it signals task (or I/O) completion to the host.

If nonvolatile FWC fails, data is destaged from the adapter DRAM copy, and the adapter can still process I/O requests. There is a performance penalty due to lack of nonvolatile FWC, as the writes are completed only when they are written to the disk, and then VSS sends the signal to the host. An alert and error message that

the nonvolatile FWC in the adapter has failed is generated. Neither data integrity nor data access is compromised when this failure occurs.

If the SSA adapter fails, the disks connected to the failing adapter cannot be accessed. The adapter must be changed in order for the system to resume access to that data. Data previously written to the adapter and yet not destaged to disk will be retained in the fast write cache and will be destaged when the adapter is changed. The fast write cache module is removable from the SSA adapter.

If data access is critical, software-based disk mirroring should be used. Disk mirroring enables two (or more) copies of the data to be stored on different disks, connected to different adapters. Because both disks have the same data, any failures in one of them will not interrupt access to the data. Also, by mirroring across adapters, failures in one of the SSA adapters will not prevent data access.

Figure 54 shows the basic idea of disk mirroring by software. When an applications starts an I/O, it is sent to the underlying piece of software (which can be either an operating system layer or a low-level application) that manages access to the logical volumes. The logical volume manager sends one I/O request to each SCSI controller through its device driver and has both I/Os processed concurrently (assuming the I/O operation is a write. If a read was started, only one logical volume will be read).

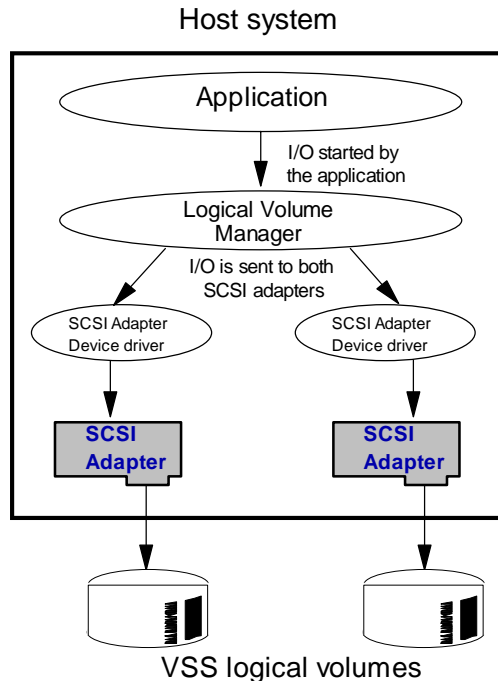


Figure 54. Software-based Disk Mirroring

Because the VSS disk adapters are stored in a hot-swappable bay, they can be changed without bringing the system down. Some disruption may be caused when the bay is removed, because the two SCSI adapters and the other SSA adapter in the bay will be unavailable. After installing the new SSA adapter, disk access will be available again.

Maximum availability, including host failure recovery, is described in “System Availability” on page 82.

### 5.1.3 Storage Processor Cluster Failures

VSS can handle storage processor cluster failures. They do not cause loss of data access or data integrity.

Each cluster has an internal SCSI disk that contains the microcode needed to run the VSS operations, along with diagnostics and configuration information. All adapter, disk and host connection information is stored in the AIX ODM. Because each cluster has its own internal disk, this information is replicated through the network, so that both clusters have the same data inside ODM.

Host adapter interfaces and disk interfaces are connected to both clusters through PCI bridges, so all the adapters can be accessed by one cluster, in a failover scenario. In normal operation, each cluster "owns" the adapters connected to its bays (*Bay 4* and *Bay 8* belongs to *Cluster 0*, *Bay 5* and *Bay 9* belongs to *Cluster 1*). The SCSI ports of the host interface adapters can be logically connected to any cluster, even though they are physically installed on a specific bay.

The two clusters perform a heartbeat operation on each other, to ensure that both are available. They send the heartbeat messages through the PCI bus. Each cluster also heartbeats the adapters it owns.

In the unlikely event of a cluster failure (for example, *Cluster 0*), the heartbeat operation will report the error to *Cluster 1* (if the SMP is still active but it is unable to continue operation), or the heartbeat operation detects the failure (if the entire SMP is inactive and cannot report errors) and starts a failover. After a heartbeat time-out of 10 seconds, *Cluster 1* will begin failover.

Failover consists of several phases (see Figure 55 on page 81). After failure detection, *Cluster 1* checks that the communication with the bays is working properly. Then it puts all the adapters it owns in a "busy" state, so that ongoing I/O operations are stopped. They will finish after the cluster failover completes.

*Cluster 1* then fences *Cluster 0*, configures bays *4* and *8*, and starts configuring the adapters in the bays. First the SCSI adapters and ports are configured, then the SSA adapters.

After finishing this phase, all pending I/Os from *Cluster 1* will be processed and finished.

The I/O operations ongoing in the failed cluster will not have completed. They will be timed out by the host device driver and redriven, at which point the available cluster will process them. Configuration changes are not allowed when just one cluster is operating.

Timeout and failover should complete in about 1 minute. However, the failover times are dependent on the configuration of the VSS (number of racks, I/O adapters) and can be longer than this.

All data flow is routed through *Cluster 1*, and the disks connected to the *Cluster 0* will be accessible after the failover. Normal data access will be resumed.

Once the failed cluster is operational again, a *failback* is performed, to put the cluster back to operation. *Cluster 0* synchronizes the process with *Cluster 1*, which releases the bays and adapters that do not belong to it. It then scans the cache for data related to disks connected to *Cluster 0* (once the failback process has completed, this data is no longer useful). After finishing this process, *Cluster 0* configures its adapters and returns to operation. The VSS returns to its normal state.

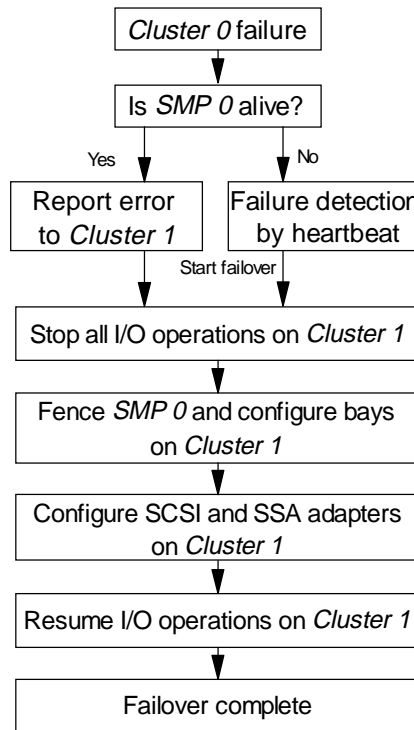


Figure 55. Simplified Diagram of the Failover Process

#### 5.1.4 Power Failures

VSS has redundant power supplies, for both the processing cluster and the SSA drawers, and two separate power cords. In case of a single power supply failure, the remaining power supply can guarantee the normal operation of the system. Cooling fans are also redundant. Both can be replaced without shutting down the system, and without affecting data integrity.

Power can be single-phase or three-phase source.

Power distribution units (PDUs) are redundant and are available on both sides of the VSS rack, providing DC power. In normal operation both units are used. If one PDU fails, the remaining unit provides power to the system.

The 7133-020 SSA drawers inside VSS have three redundant power supplies and cooling fans. They are connected to both PDUs, and the 7133 will operate in case of loss of one of the power supplies. Power supplies and cooling fans are hot swappable and can be changed without powering down the 7133.

An optional battery backup provides power for several minutes, in case of total power loss. The power can then be used to perform a cluster shutdown.

---

## 5.2 Data Integrity

VSS provides full data integrity, regardless of any single failure that may occur. Once data is written to VSS, it is stored either directly to the disk or in the nonvolatile FWC. All data stored in the nonvolatile FWC is also available on the adapter DRAM, so if the nonvolatile FWC fails, the data is still safe in the volatile memory. It is then destaged to disk, and data integrity is maintained.

If a disk adapter fails, the data stored in the nonvolatile FWC is still safe. It can be removed and installed onto a new adapter, and access to the disks will be restored. The data from the nonvolatile FWC is then destaged to disk, without loss.

Finally, a single disk failure does not affect data integrity on the RAID-5 arrays; the failed disk will be replaced by a hot spare in the drawer.

---

## 5.3 System Availability

Adapter failures, or the actions required to repair the system, can disrupt the environment. You can configure the system to have this disruption affect only one host or several hosts. You should consider different options when planning for system availability

### 5.3.1 Planning for Maximum Availability

The VSS can have eight host interface adapters and eight SSA adapters. Each host interface adapter has two SCSI ports, so 16 SCSI connections are available. Using several of these ports and SSA adapters in the same host can increase the availability of the system.

#### 5.3.1.1 Minimal Disruption in Case of an Adapter Failure

When a host interface card fails, data flow through it ceases. The hosts connected to it cannot access the data until the card is replaced, or the logical volumes are reassigned. Logical volumes can be reassigned to another port already connected to the hosts, or to a new port assigned to the hosts (see Chapter 8.2.2.1, “Storage Relocation” on page 106).

It is possible to have two SCSI connections to the same logical volume, in the same host. Without special software, the two connections cannot be used at the same time, but when one fails, the other can be reconfigured to access the data, through operating system commands.

Several vendors have developed software to provide multiple paths to a disk. This software can be used with VSS to provide transparent failover support in case of a SCSI adapter failure. IBM plans to deliver a product called the I/O Path Optimizer that performs this function (see “IBM I/O Path Optimizer” on page 77).

#### 5.3.1.2 Minimal Disruption Caused by Maintenance Actions

If an adapter fails, either SSA or host interface card, the bay in which it resides has to be removed from VSS to replace the failed adapter. The bay can have up to four adapters, so all of them will be unavailable during the maintenance activity. This can cause additional disruption to other systems not affected by the failed card. To create minimal disruption when performing such maintenance, the host interface adapter attached to the host can be in the same bay as the SSA



adapter connected to disks accessed by the host. Also, critical hosts can be connected to VSS through cards on different bays.

### 5.3.2 High Availability Configuration

For maximum availability, use high availability software (such as IBM's High Availability Cluster Multi-Processing - HACMP), along with data mirroring across adapters. The operating system handles single adapter failures, using the mirrors available. The high availability software handles host failures.

With such a high availability solution, data can still be accessed even if a single SSA adapter fails, or a single host interface adapter fails. Adapters should be selected from different bays, to allow nondisruptive maintenance of the system.

Figure 56 shows an example of a high availability configuration. In VSS, ports can be *local* or *remote*. A *local* port is connected to the cluster that owns its bay (for example, *Bay 4* is owned by *Cluster 0*). A *remote* port is connected to the cluster that does not own its bay. Port assignment is done at installation time, and it is better for performance and availability reasons to have *A* ports *local* and *B* ports *remote*. See Chapter 4, "Host Connections" on page 57 and Chapter 6, "Performance and Tuning" on page 89 for further details.

The SCSI ports of a VSS have been configured so that all *A* ports are *local* and all *B* ports are *remote*. Each host has two SCSI adapters, connected to VSS. Two logical volumes are defined inside VSS for both hosts, and disk mirroring is done by software. *Port 1* and *Port 2* are connected to *Disk 1* across *Cluster 0* and *SSA 1*. *Port 3* and *Port 4* are connected to *Disk 2* across *Cluster 1* and *SSA 2*. The dashed lines represent logical connections between SCSI ports and SSA adapters.

In case of a failure of a single SCSI adapter (in either the adapter itself, the host interface adapter to which it connects, or the cable between them), for example, *SCSI 1*, there is no loss of data access. The host can access the mirror across *SCSI 2* and *SSA 2*. If *SSA 1* fails, the host can access the mirror, also across *SCSI 2* and *SSA 2*. Adapters can be replaced without disrupting data access. *Bay 8* can be removed to replace a failed host interface adapter, and access to the data will be through *Bay 9*.

If two SCSI adapters fail on the same host, the high availability software will provides failover support to the other host.

If a VSS cluster fails, it fails over to the other cluster. In this configuration, however, because the connections between the host and the disks are done through both clusters (one mirror in each cluster), the system does not depend on cluster failover for data availability.

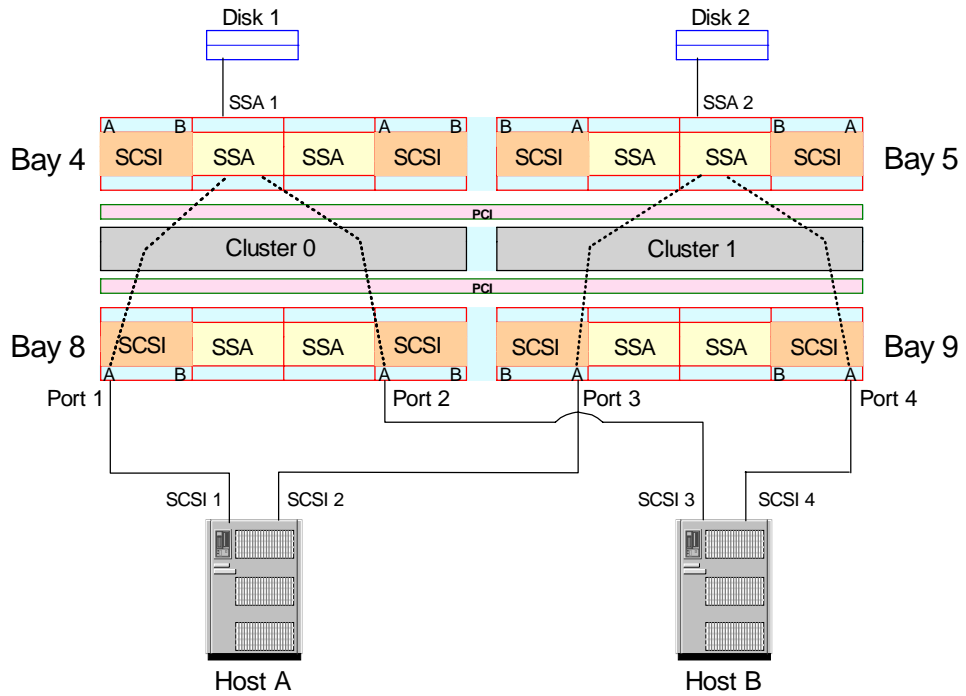


Figure 56. High Availability Configuration: Two Hosts and Mirrored Logical Volumes

If the adapters for a given host are put on different bays, it is possible to have multiple hosts with high availability configurations inside VSS, as shown in Figure 57.

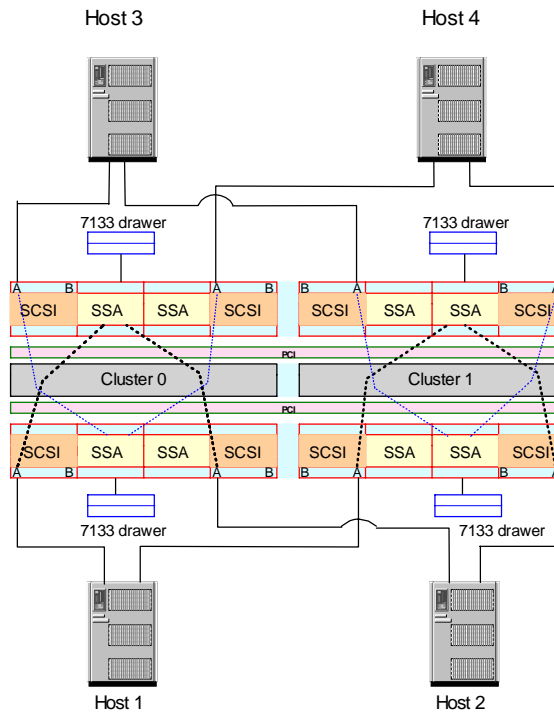


Figure 57. Four Hosts in a High Availability Environment

---

## 5.4 VSS Availability Test Scenarios

We conducted some real-world tests to test availability and failure recovery in the VSS. The test scenarios presented here were developed and run on VSS.

For these tests we used a 2105-B09 VSS machine, with a 2105-100 expansion rack. We configured the VSS with eight SSA adapters and eight host interface adapters. Eight SSA drawers were connected to the eight adapters, and two RS/6000 F50s were used as hosts. The F50s were connected through eight Ultra SCSI adapters each, and AIX level was 4.2.1.

Although the hosts were RS/6000 machines, the tests are related to VSS and would have the same results with any other type of host.

### 5.4.1 Storage Relocation

In this test, a SCSI failure was simulated by disconnecting the SCSI cable, and relocation of the storage was done with the VS Specialist.

We defined two 28 GB logical volumes inside VSS, on the same array. The logical volumes were connected to the same SCSI port on VSS (*Port 0*) to one F50 through *SCSI 2*.

We ran `cfgmgr -l scsi2` on the F50 to have the disks recognized by AIX. The systems recognized them as *hdisk2* and *hdisk3*.

We defined a volume group on both disks, created a file system on it, and started a simple script that produced some I/O (copying a file into another and appending output of the `date` command).

The SCSI cable was then disconnected from VSS port 0, and the system lost access to data. AIX reported the error in the *errlog*, and the application returned an error.

To access the disks, we did a storage relocation inside VSS, using the VS Specialist. We removed the connection between the logical volumes to *Port 0* and connected them to *Port 3*.

On the F50, *SCSI 10* was connected to *Port 3*. Before configuring *SCSI10* to access the disks, we performed some cleanup. We:

- Unmounted the file system
- Varied off the volume group
- Ran `rmdev -l hdisk2 -d` (to remove the *hdisk2*). We also ran this command for *hdisk3*).

Then we ran `cfgmgr -l scsi10`.

The disks were again recognized as *hdisk2* and *hdisk3*. If we had not run the `rmdev` command, the disks would be recognized as two new *hdisk*s. We recommend removing the disk definitions before configuring them on the new adapter.

The disks were again available, and we issued the *varyonvg* command to get access to the volume group. After that we mounted the file system, and the data could be accessed again.

The reconfiguration time (once the failure was detected) took no more than 5 minutes.

The use of software that implements multiple path access to the disk (such as the IBM I/O Path Optimizer) provides much better recovery in failures such as this, allowing continuous data access to applications and load balancing across multiple adapters. Whenever it is possible to have such software, we strongly recommend using it.

### 5.4.2 SSA Cable Failures

The SSA architecture allows redundancy in disk access, providing two independent paths to disks. In the case of failure on a single path, data access can still be performed over the other disk.

We forced a failure by removing one SSA cable from a drawer while accessing data on it. The host machine had not received any error messages, and the application was able to access the data. No impacts were perceived on the host side due to this failure (except a performance impact, because bandwidth was reduced by half. This impact was not noticeable in our tests, but in a production system it may be).

By reconnecting the cable, the loop was reconstructed, and the system returned to its normal performance.

We then disconnected both cables, interrupting access to the data. The failure was reported by an e-mail sent to the administrator (Figure 58) and in the VS Specialist status panel (Figure 59).

```
# Begin of call home record.....= BEGIN OF FILE (2676)
# Product Manufacture ID and Date...= 0131831998
# Business/Company Name.....=
# Product Machine type and model number= 2305B09
# Product serial number.....= 13FA036
# Remote service call back number...=
# Call back sponsor post ID.....=
# Call back user ID.....= support
# Call back password.....=
# Customer voice phone number.....=
# LJC level of local complex.....= SA81008
# LJC level of remote complex.....= SA83008
# Local Charter Number.....= 1
# Record type.....= 1
# Report time/date stamp.....= 10/12/98 14:40:28 PDT
# Host system type(s).....=
# Level_CPI33 information.....= 1apcs/R3/CPI33test/raab/Act1008
# Possible FRUs to replace:
#
# Engineering FRU Likely FRU Location
# FRU Name Name to fix Description
#-----
#DCM0013 9 308 SSA DDM 40% Drawer5-D13
aa723 SSA Device Adapter 40% R1-D5-E3
ssbck5 Bypass Card, Lower R 10% Drawer5-C3
sssa63_Koth5 SSA Cable 2% R1-R3-E3-A1 to Drawer5-J13
sssa30bbph5 Ser. Backplane Assem. 2% Drawer5-P2
#
# Additional information
#
# Primary Failing Unit...= 7133-20 30-03567
# Reporting Unit.....= 2105-B09 13-FA036
#
# Unit Translation Table
# Drawer5.....= 7133-20 30-03567
# R1.....= 2105-B09 13-FA036
# Problem ID.....= 1
# EDC.....= 1284
# BSN.....= 45800
# Problem Status.....= PENDING
# Description.....= FRU FAILURE
# Additional Message.....= NONE
# Failing Cluster.....= 1
# Reporting Charter.....= 1
# Failing Resource.....= aa720
# First Occurrence.....= Mon Oct 12 14:40:26 1998
# Last Occurrence.....= Mon Oct 12 14:40:26 1998
# Failure Count.....= 1
# Representation Interval = 28800
# Representation Count...= 4
# Isolation Procedure...= 3128
# Failure Actions.....= D521 "SERVICE ACTION REQUIRED"
# Probable Cause.....= D12C "FRU FAILURE"
# Failure Cause.....= D12C "FRU FAILURE"
# User Actions.....= FFFF "NONE"
# End of record flag...= END OF FILE
```

Figure 58. Error Notification via e-Mail

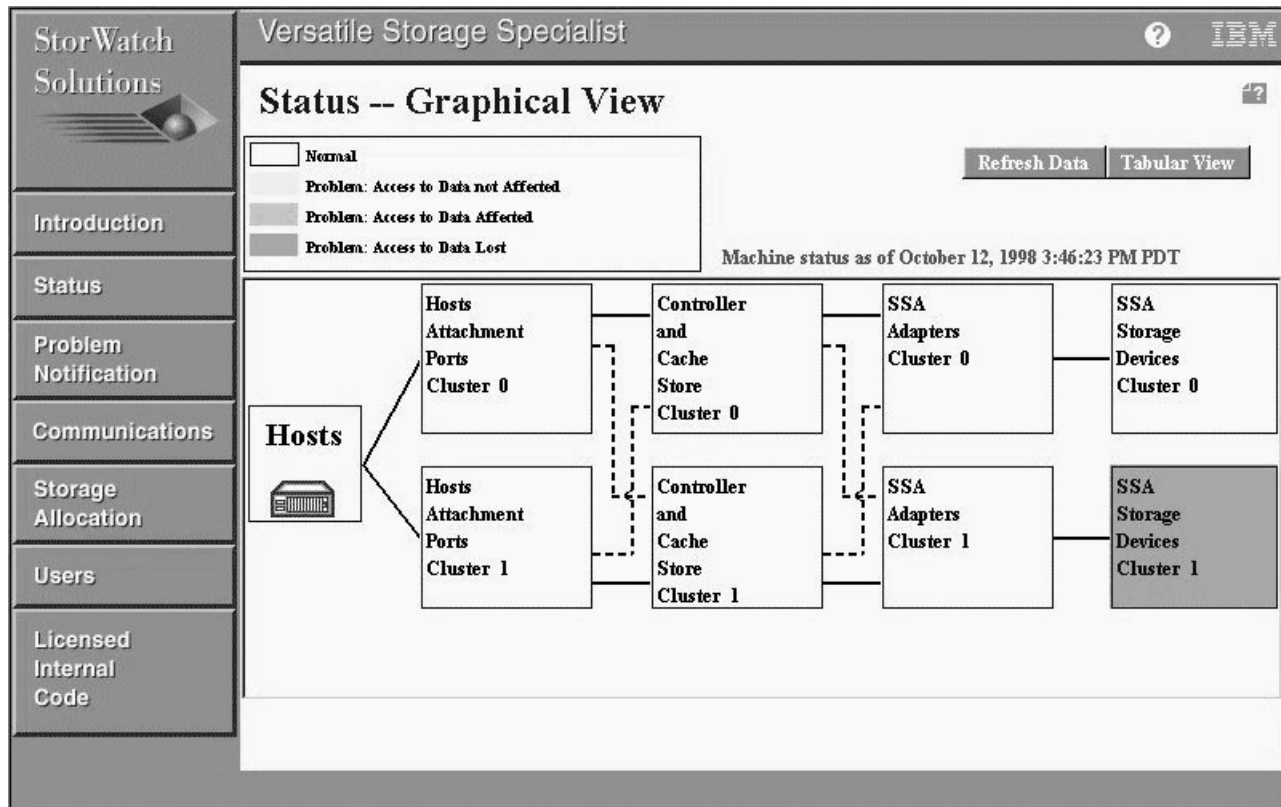


Figure 59. Status Panel after Disconnecting an SSA Drawer

Once the cables were reconnected, access to the disks was possible again. No other actions have to be performed on VSS to recover from these failures. Host actions, such as restarting the application, may be necessary, however.

### 5.4.3 Storage Cluster Failures

We tested how VSS recovers from a single cluster failure and what happens with data access during the failover and failback processes.

We created two volume groups on the host machine, each of them created on arrays in different clusters inside VSS (for example, *vg0* was created in a *Cluster0* array, and *vg1* was created in a *Cluster1* array). In these volume groups, we created a raw logical volume (LV). We also created a raw LV on the internal disk of the host.

Using the **dd** command, we generated four different I/O scenarios:

- Data read from local raw LV and written on *Cluster0* raw LV (I/O writes to *Cluster0*)
- Data read from local raw LV and written on *Cluster1* raw LV (I/O writes to *Cluster1*)
- Data read from *Cluster0* raw LV and written on local raw LV (I/O reads from *Cluster0*)
- Data read from *Cluster1* raw LV and written on local raw LV (I/O reads from *Cluster1*)

The basic syntax of the command is:

```
dd if=/dev/r<LV to be read> of=/dev/r<LV to be read>
```

As an example, we generated I/O writes to *Cluster1* by entering this command:

```
dd if=/dev/rlocallv of=/dev/rlv0
```

where *rlocallv* is the raw LV on the local disk, and *rlv0* is the raw LV inside VSS connected to *Cluster0*.

We generated a cluster failover by powering off one of the clusters while the four I/O tests were being run. In our example, we powered off *Cluster0*.

After powering off, the failover process started immediately. The I/Os for both clusters stopped and remained in this state until the end of the failover. The failover process detected the failure and started its recovery procedures, by reconfiguring the adapters to *Cluster1*. All the adapters previously owned by *Cluster0* are now owned by *Cluster1*. After almost a minute and a half the ongoing I/O operations resume, and all the data can be accessed again.

No I/O was lost during the test failover process. In most cases, an application will not see an I/O error, as the device driver will retry the operation after the failover is complete. In some cases, an application may return an error and end the execution. The **dd** command is very simple and provides few or no recovery procedures, but even so the I/Os were not lost. Applications (like databases) usually have recovery procedures that handle this type of situation and will probably be able to continue running without data loss or I/O errors.

The failed cluster goes to a *fenced* state after the failure and remains powered off until the IBM CE performs maintenance. The failback process is initiated when the failed cluster is put back to operation (by executing the *unfence* action on the CE service terminal). When performed, this action powers up the failed cluster and starts the initial microcode load (IML) process. When IML is finished, the heartbeat process, interrupted during the failure, resumes, and *Cluster1* starts to release the resources of *Cluster0*. After approximately two minutes, the system is back in normal operation, with both clusters working properly. The I/O operations in the failback process are stopped for *Cluster0* devices only and resume after the end of the process.

---

## Chapter 6. Performance and Tuning

In this chapter we describe the configuration performance options and characteristics of the VSS.

---

### 6.1 Performance Fundamentals

The VSS has several inherent performance features: 4 way SMP clusters, large server cache, multiple PCI bus structure, SSA adapters with FWC, and the SSA disks themselves. These features combined into the VSS architecture provide a superior performance configuration for most workloads. In addition to these inherent features, the VSS can be configured to meet various capacity, availability and performance requirements. Below we discuss each performance feature to provide a basis for considering the various options for configuring a VSS.

#### 6.1.1 4 Way SMP Clusters

Each SMP cluster contains a high-performance, 4 way RISC-based processor complex that runs a UNIX kernel and provides the main intelligence for subsystem control. This enables new capabilities for a disk storage controller such as the ability to run a Lotus Domino Go Webserver (LDGW), which allows use of a Web-browser-based interface utilizing Java function for configuration and control. The SMP complex also provides the intelligent cache management function to control the large cache memory.

#### 6.1.2 Large Cache Memory

The cache is managed by an adaptive caching algorithm designed to provide both efficient use of the cache resource and the highest possible hit rates. This is done by varying the amount of data staged (the amount of data moved from the disk backstore into cache) on the basis of the I/O requirement. There are three types of stage operations: stage the required record only, stage the required record plus the remainder of the 32 KB logical track, or stage the entire logical track (sequential prestaging). The method used is dependent on the type of I/O being performed and varies according to the I/O stream. When random I/O is detected, partial track staging is used. This operation is then migrated to record staging if the access patterns detected show access to single records per track. When the VSS detects a sequential pattern, it prestages several logical tracks in an attempt to keep ahead of the applications requirements.

This is a very simplified view of the process but serves to provide a basic understanding of how the cache is managed. Using multiple staging techniques also minimizes the use of other important system resources, such as internal busses, device adapters, processors, and paths.

#### 6.1.3 PCI Bus Structure

The VSS uses a multiple PCI bus structure to provide fast and efficient communication and data movement among the various subsystem components. The use of a PCI bus structure provides important performance and flexibility capabilities. PCI is a widely used industry standard bus architecture and thus enables flexibility in the components used within the VSS architecture. The use of

multiple buses allows multiple components to communicate and or move data concurrently.

#### 6.1.4 SSA Adapters

The SSA adapters used in the VSS are PCI bus RAID 5 FWC adapters. The RAID 5 capability provides dynamic redundancy for physical volume failures as well as a hot spare for additional redundancy. The SSA dual loop architecture allows for large amounts of storage to be attached to a single adapter, up to 48 disks in a VSS configuration, while providing 160 MB/sec bandwidth and 8 concurrent data paths. A fully configured VSS subsystem has 8 SSA adapters, which provide 64 concurrent back-end data paths. The FWC provides a mechanism for partially negating the write penalty usually associated with RAID 5 configurations. The adapter microcode has a full stripe (a group of logical tracks with parity that span an array group) write feature, known as an *express destage*, that provides significantly better performance for large sequential writes than traditional RAID 5 implementations.

---

## 6.2 Performance Configuration Options

When you configure a VSS subsystem, there are several areas where the configuration affects performance and availability. These are cache size, cluster load balancing, SCSI adapter configuration, SSA adapter load balancing, and logical volume configuration. A few basic rules should be applied for these areas when you configure a VSS.

Disclaimer: Various charts are included in this chapter to provide examples of performance expectations under specific conditions such that the reader can extrapolate to his or her own situation. The reported results were obtained in a lab environment with controlled conditions and therefore may not be exactly reproducible in a customer environment. There are many variables in any particular situation and configuration to provide real examples for every case; however, the examples shown should serve as a useful guide.

### 6.2.1 Cache Size

Sufficient cache size is important to the overall throughput of the subsystem. What is "sufficient" depends on two main factors: how much host memory and storage are installed in the VSS. In the UNIX and Windows NT environments, host memory is very important to I/O performance because most I/O is performed through a file system that uses host memory as a cache for I/O. The exception to this are database systems, which tend to use raw logical volumes and control the data placement and format themselves. Databases, however, usually use host memory to provide a buffering capability for I/O requests, thus providing a capability similar to file systems. In addition to these two factors, the basic data structure for both UNIX and Windows NT is a fixed block allocation where an individual block is a discrete entity and may have no logical relationship to adjacent physical blocks. The effect of file system caching, database buffering, and fixed block architecture is to condition the I/O patterns seen at the storage subsystem with a tendency to be very random for random I/O and for sequential I/O to be relatively small sequential streams. There are of course exceptions to these rules of thumb. The more host memory installed on the server and the better its caching effectiveness, normally the less requirement there is for VSS subsystem cache. The demands placed on the cache by different I/O streams are



also different. A random I/O stream can require fairly large amounts of cache to provide good hit rates. Thus a particular file may require some amount of cache to obtain a good cache residency time for its workload. Sequential I/O can normally get good cache hits rates from relatively small amounts of cache because each record is typically read once, and thus, once passed to the host, the cache space for the record can be immediately reused. This is typically done at the track level; that is, once all records from a track have been passed to the host, the portion of cache that held the track can be reused. In this way a relatively small number of tracks of records can be held in cache, and the cache space can be reused as the tracks are passed to the host.

Table 5 on page 91 provides a guide to cache sizing for less effective host caching environments. In an environment where the host memory is constrained increasing the cache size of the VSS relative to the amount of storage it provides can have a significant effect on hit ratios.

*Table 5. VSS Cache Sizing for Less Effective Host Caching Environments*

Cache Size	Drawers 9 GB Disk	Drawers 4.5 GB Disk
512 MB	1	2
1 GB	2	4
2 GB	4	7
4 GB	7	14
6 GB	11	18

Table 6 on page 91 provides a guide to cache sizing for a host environment where the host caching is effective. A large cache size in the VSS is not required as the effect of the host caching is to make the I/Os that goes to the VSS very random thus driving down the hit rates even for a large cache. In this environment there is a low return on investment from increased cache size because doubling the cache size would typically only increase the cache hit rate by a few percentage points.

*Table 6. VSS Cache Sizing for Effective Host Caching Environments*

Cache Size	Drawers 9 GB Disk	Drawers 4.5 GB Disk
512 MB	2	4
1 GB	4	8
2 GB	8	16
4 GB	16	18
6 GB	18	18

## 6.2.2 Cluster Load Balancing

With cluster load balancing the I/O is spread as evenly as possible over the two clusters. Thus if a host has two SCSI connections to a VSS, one should be to each cluster. SSA adapters and storage should also be allocated in a similar fashion whenever possible. This is a simple way of providing load balancing across the two clusters.

## 6.2.3 SCSI Interface Configuration

The SCSI adapters are installed in a logical sequence alternating between clusters; that is, if you have four SCSI adapters (eight ports), they are installed

one per bay or two per cluster. The SCSI interface configuration refers to the internal assignments of the SCSI ports to clusters. Each SCSI port must be associated with a cluster. Because of the internal bus architecture of the VSS, it is best for performance to allocate one port from a physical SCSI adapter board, which contains two ports, to the local cluster, that is, the cluster that owns the bay in which the adapter resides, and the other port to the remote cluster. This configuration provides the best overall throughput for the subsystem. We recommend that you install all available SCSI adapters in the VSS initially. In this way the best possible availability and performance will be provided without later disruption during upgrades.

#### 6.2.4 SSA Adapter Configuration

The SSA adapters must be installed in a logical sequence alternating between clusters just like the SCSI adapters. Each SSA adapter has two independent loops for connecting disks (see Section 1.1.3.4, “Disk Interface Adapter” on page 11 for further discussion of the operation of the SSA adapters). The drawers of SSA disk are attached to the SSA adapters in the same sequence using all A loops first and then rotating through the B loops. Each loop consists of 8 or 16 disks (except in a full configuration where adapters 1 and 2 will have 32 disks) and is configured into one or two Raid-5 arrays accordingly. Each array can be split into a varying number of logical volumes between 0.5 GB and 48 GB or 56 GB depending on the array size (6+P arrays provide 51 GB of usable space and a maximum logical volume size of 48 GB, while 7+P arrays provide 61 GB of usable space and a maximum logical volume size of 56 GB). For best performance, storage use should be spread over as many adapters as possible and across clusters whenever possible. We recommend that you install all available SSA adapters in the VSS initially. In this way the best possible availability and performance will be provided without later disruption for upgrades.

#### 6.2.5 Logical Volume Configuration

There are two items of interest to consider when allocating logical volumes within a VSS: logical volume size and logical volume placement. There is no performance penalty associated with allocating the same amount of storage in varying volume sizes. Therefore if you allocate four 4 GB logical volumes from one array to a host or allocate one 16 GB logical volume from an array and split it into four 4 GB logical volumes at the host through the host volume manager, performance will not be affected by the VSS configuration. Several options are available when configuring the logical volume placement for a particular host attached to a VSS. Volume layout for a host can be configured for availability or for performance or for both.

Figure 60 on page 93 shows a sample configuration for availability. This is a very simplified configuration to enhance clarity. The diagram shows a VSS with eight SCSI adapters, eight SSA adapters, and eight drawers of SSA disk. In this diagram, logical volumes in drawer 0 are connected to port 0 of the SCSI adapter in the same bay, and logical volumes in drawer 4 are connected to port 8 of the other SCSI adapter in this bay. This is shown by the dotted lines connecting the drawer to the SCSI adapter. This type of configuration contains a servers storage attachment to a single bay. This may be desirable for a test system where it is required to isolate the server and its storage from other servers. If all four SCSI ports in a bay are configured within the same bay, the throughput of that bay will be reduced, in contrast to a configuration where the A ports (even-numbered

SCSI ports) are configured to storage in their own cluster and the B ports (odd-numbered SCSI ports) are configured to the alternate cluster. If all four SCSI ports in one bay are to be configured to the same cluster, it would be best if one of the ports on each of the two SCSI adapters are 'standby' ports, that is, they are there to take over in case of failure.

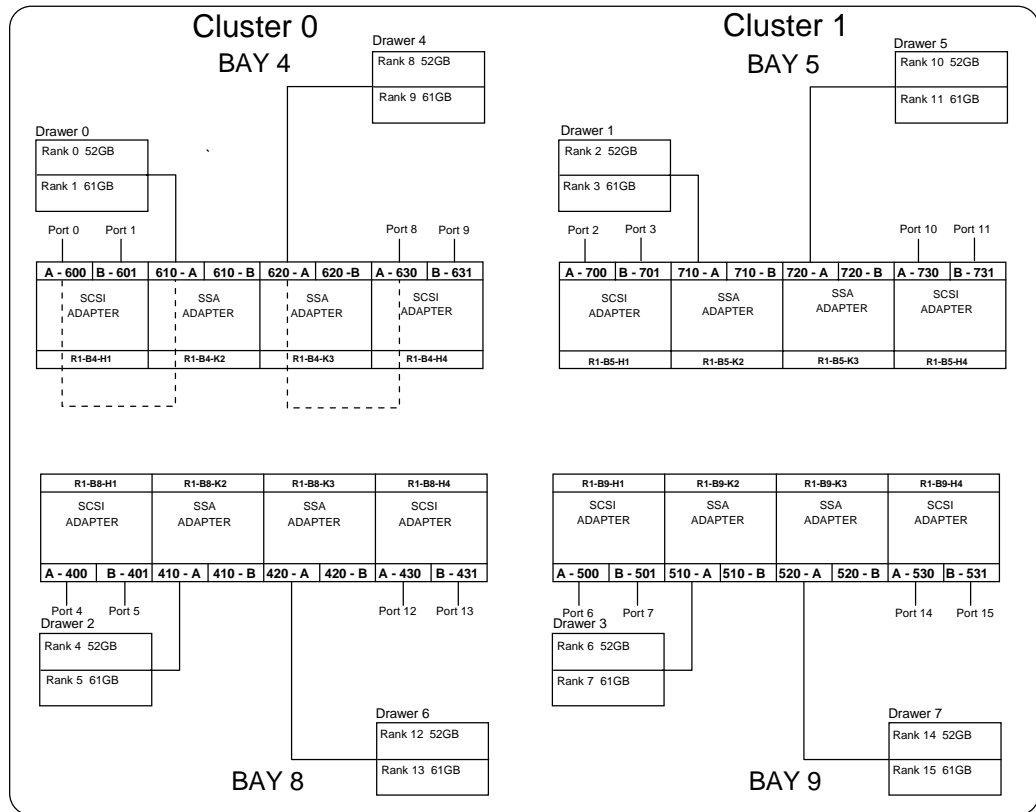


Figure 60. Sample Configuration for Availability

Figure 61 on page 94 shows an example of a basic performance configuration. The diagram shows a VSS with eight SCSI adapters, eight SSA adapters, and eight drawers of SSA disk. Here, port 0 in the first SCSI adapter in bay 4 cluster 0 and port 3 in the first SCSI adapter in bay 5 cluster 1 are attached to storage in bay 8 of cluster 0. Port 1 in the first SCSI adapter in bay 4 cluster 0 and port 3 of the first SCSI adapter in bay 5 cluster 1 are attached to storage in bay 9 of cluster 1. This cross-configuration type of allocation spreads the storage utilization across the available bus structure, SCSI and SSA adapters. However, it does not isolate a particular server unless all of the SCSI ports allocated are in use by one server.

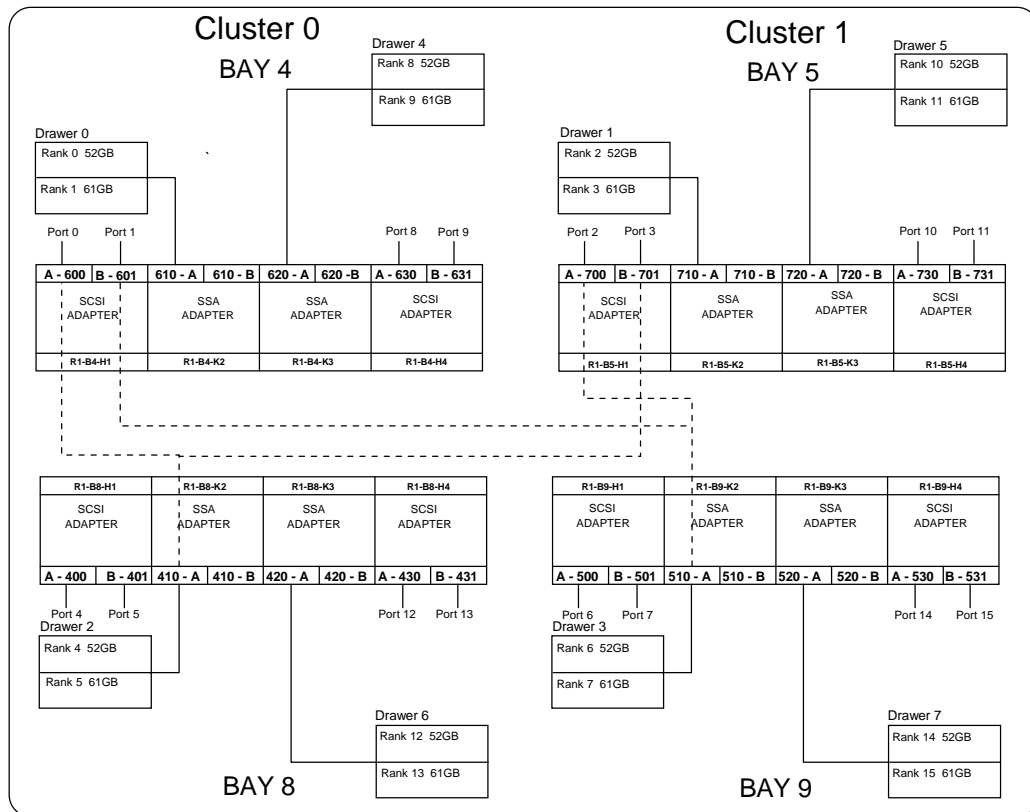


Figure 61. Sample Basic Performance Configuration

Figure 62 on page 95 shows a more advanced performance configuration. The diagram shows a VSS with eight SCSI adapters, eight SSA adapters, and eight drawers of SSA disk. Here, port 0 of Cluster 0 is connected to storage in four arrays in four drawers, two in each Bay of Cluster 0. Port 2 in Cluster 1 is also connected to storage in four arrays in four drawers, two in each bay of Cluster 1. In this way the subject server that owns the two ports will have its storage spread over a large complement of PCI buses, SSA adapters, and arrays. Configuring multiple servers in this way tends to spread the I/O load over the entire machine more evenly. This reduces the likelihood that a busy server will affect a particular bay, bus, SSA adapter, or array's throughput or that a busy bay, bus, SSA adapter, or array will affect a server's throughput. This configuration requires more effort from a planning, allocation, and management point of view and provides the lowest level of isolation between servers and VSS components.

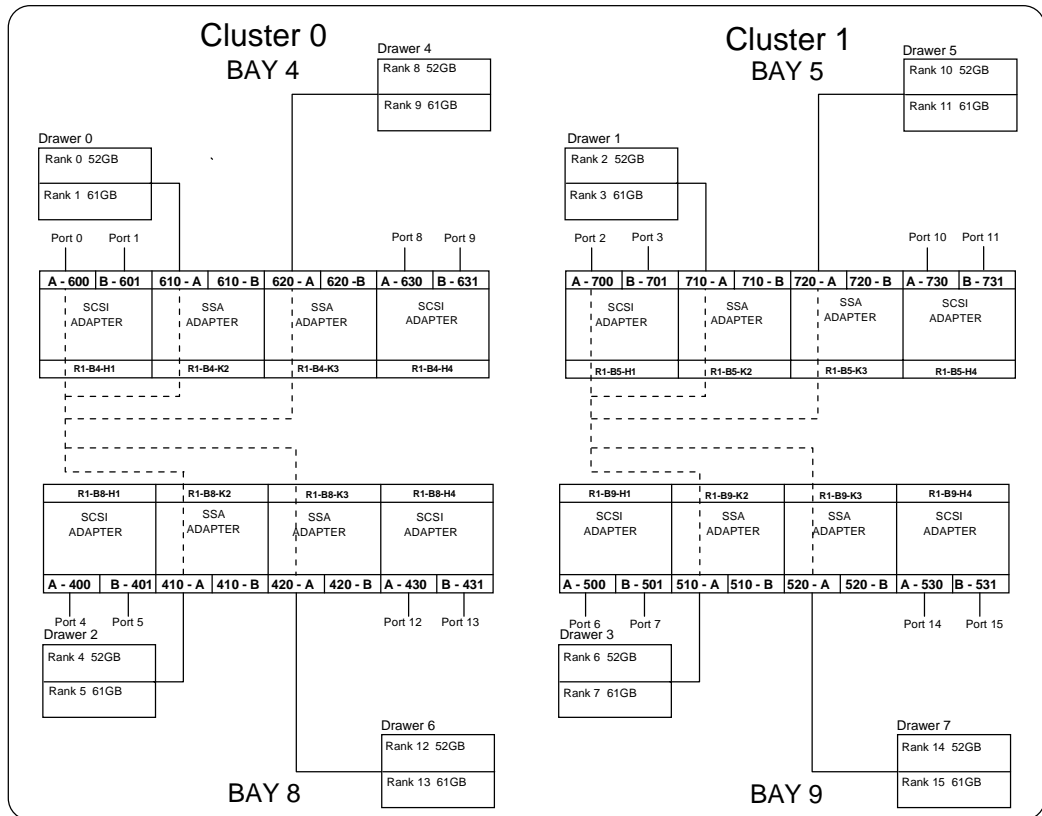


Figure 62. Sample Advanced Performance Configuration

### 6.3 Performance Expectations

In any dynamic disk controller like the VSS attached to a UNIX system with varying applications, the ability to predict performance expectations is somewhat of a black art. The variety of configurations, application mixes, and I/O loads is too great to provide set performance expectations. Therefore we provide a set of benchmark results designed to show the expectations for specific workload characteristics based on block size, read to write ratios, and I/O rates. The figures in this section should give you an idea of what to expect if you know your workload characteristics.

Figure 63 on page 96, Figure 64 on page 96, and Figure 65 on page 97 are based on a VSS configuration with a 4 GB cache size and four arrays on four SSA adapters for a total of approximately 250 GB of usable capacity. The I/O load was driven through four Ultra SCSI adapters from an RS/6000 F50 processor. The I/O was very random over the entire data area, resulting in very low read hit rates of 1-2%. Given the nature of this test, the I/O rates were constrained by the number of arrays and devices available for I/O. Providing more devices and arrays would increase the overall throughput of the subsystem even with the very low read hit rates. This can be seen in the results reported in Table 8 on page 98.

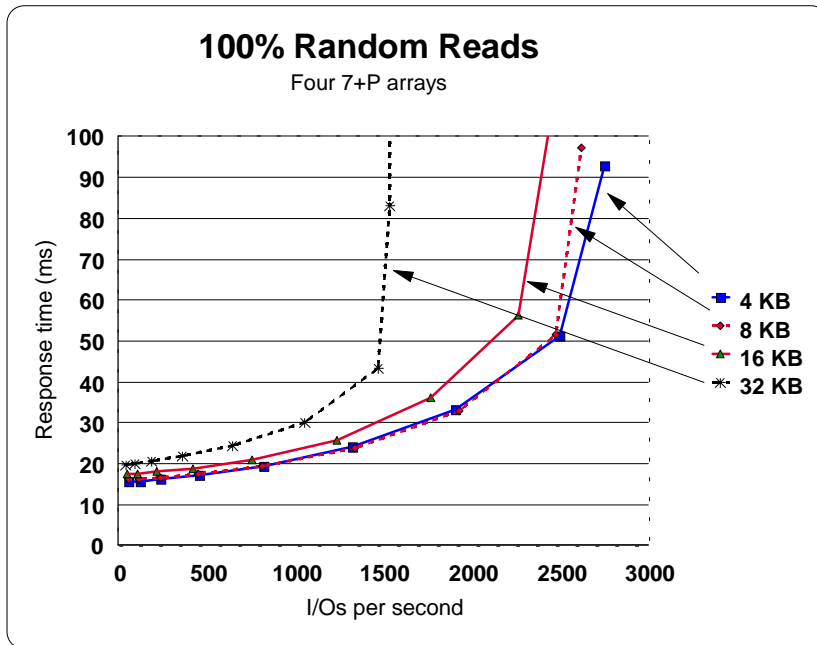


Figure 63. I/O Rate and Response Time for 100% Random Read Test

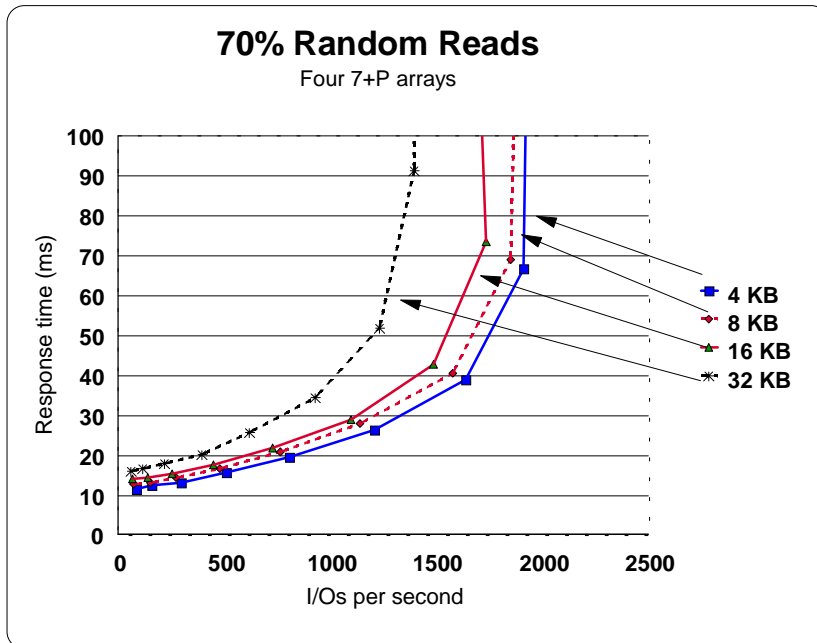


Figure 64. I/O Rate and Response Time for 70% Random Read Test

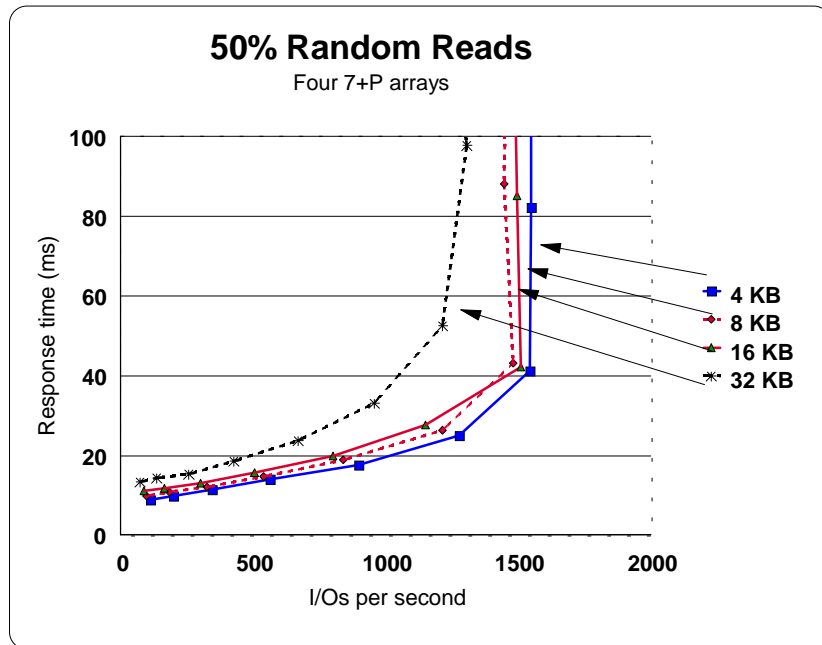


Figure 65. I/O Rate and Response Time for 50% Random Read Test

### 6.3.1 Array Throughput Rules of Thumb

Table 7 on page 97 shows some rules of thumb for I/O intensity rates expected from an array, given common read to write ratios. These numbers are provided for planning purposes to prevent excessive load on a device or array and to ensure good application performance. They are valid for random workloads with record sizes of 16 KB or less; they do not represent the maximum achievable throughput per array. Maximum throughput depends on a number of variables, such as read to write ratio, cache hit rate, and block size.

Table 7. I/O Rate per Array Rule of Thumb for Various Read to Write Ratios

Read Percentage	I/Os per Second per Array
100%	400-500
70%	275-325
50%	200-300

### 6.3.2 Subsystem Throughput Expectations

Table 8 on page 98 shows various throughput capabilities under varying conditions for the subsystem as a whole. These numbers represent the maximum throughput possible under optimal conditions and as such should not be expected to be achievable from normal applications. These throughput numbers are

generally far higher than any normal application would require but serve to show the high end of the achievable envelope.

Table 8. VSS Throughput Expectations for Various I/O Characteristics

Workload	Throughput
100% cache read hits - 4 KB records	26,000 I/Os/sec
Random reads - 4 KB records - no cache hits	8,000 I/Os/sec
Random 70/30 read/write ratio- 50% read cache hits - 4 KB records	9,500 I/Os/sec
Random writes to disk- 4 KB records	2,000 I/Os/sec
Writes - 100% cache hits - 4 KB	9,800 I/Os/sec
Sequential reads from disk - 64 KB	80 MB/sec
Sequential writes to disk - 64 KB	75 MB/sec

## 6.4 Performance Monitoring and Reporting

Performance monitoring and reporting is a very extensive area. The intent of this section is not to try and cover this subject in general but rather to point out a couple of areas where understanding what you may see from performance data requires an understanding of the VSS subsystem. The fundamental difference between the VSS and traditional disk is the use of a RAID 5 array to present a single disk image. Therefore, although the operating system sees one disk, the I/O is actually being performed on several disks, seven or eight in the case of the VSS.

### 6.4.1 IOSTAT Reports

Traditionally open system environments have not provided a wealth of performance monitoring tools compared to the system 390 area. The base tool used for UNIX is IOSTAT, which, although available on most UNIX platforms, varies in its implementation from system to system. I/O activity monitors such as IOSTAT have no way of knowing that the logical disk they see is a physical disk or multiple physical disks in an array. Therefore some numbers reported for a device, for example, % busy, will appear high. It would not be unusual to see a device reported by IOSTAT as 90% - 100% busy because a VSS volume that is spread across an array of eight disks can sustain a much higher I/O rate than for a single physical disk. Having a device 100% busy would generally be a problem for a single device but probably not for a RAID 5 device. To calculate the average device busy you would need to sum the numbers for a group of devices that constitute an array and divide by the number of devices in the array.

Figure 66, Figure 67, and Figure 68 show sample IOSTAT reports from an AIX, Sun OS, and HP-UX system, each attached to a VSS subsystem.

Figure 66 shows the AIX version of IOSTAT and a *%tm\_act* column which indicates the percentage of time the device is busy. The *tps* column shows the transactions per second for the device. This field may also appear higher than would normally be acceptable for a real physical disk device.



Disks:	% tm_act	Kbps	tps	Kb_read	Kb_wrtn
hdisk0	0.2	1.2	0.2	233479	2345859
hdisk1	0.0	0.0	0.0	0	0
hdisk2	5.9	156.4	14.4	222132	527145
hdisk3	5.2	148.7	10.9	298403	708060
hdisk4	0.0	0.0	0.0	11602	0
hdisk5	0.0	0.0	0.0	2115	0

Figure 66. IOSTAT Report from AIX

Figure 67 shows an IOSTAT report from Sun OS. Here you can see an example of a device that appears to be very busy, device sd34. The *r/s* column shows 124.3 reads per second, the *%b* column shows 90% busy for the device, but the *svc\_t* column shows a service time of 15.7 ms, quite reasonable for 124 I/Os per sec. However, the calculations for service time are based on a physical volume, and, as previously mentioned, the physical volume view that the VSS presents to the host is in reality a VSS logical volume, and is comprised of multiple, physical volumes. (Note that the *svc\_t* column indicating service time would normally be called response time as it is the sum of all components of the I/O time.)

Device	r/s	w/s	kr/s	kw/s	wait	actv	svc_t	%w	%b
sd15	64.3	0.2	2545.4	0.9	0.0	1.6	24.3	0	90
sd30	40.3	0.0	703.1	0.0	0.0	0.3	7.8	0	21
sd31	117.2	0.0	1849.1	0.0	0.0	0.6	5.1	0	55
sd32	79.6	0.0	3200.6	0.0	0.0	0.5	5.8	0	46
sd33	44.0	0.0	664.8	0.0	0.0	0.2	4.1	0	18
sd34	124.3	0.1	3385.0	0.1	0.0	2.0	15.7	0	90

Figure 67. IOSTAT Report from Sun OS

Figure 68 shows an example of an IOSTAT report from HP-UX. This is a fairly simple format with three columns of statistics; *bps* indicates the kilobytes transferred per second, *sps* indicates the seeks per second, and *msps* indicates milliseconds per average seek. The first two numbers, *bps* and *sps*, are subject to the effects of the VSS RAID-5 architecture. The MAN page for the IOSTAT command on HP-UX states that the *msps* field is set to 1.0 to maintain compatibility because of items such as current technology, striping, and RAID 5, it is no longer possible to calculate an accurate number.

Device	bps	sps	msps
c7t0d4	0	0.1	1.0
c7t0d5	0	0.1	1.0
c7t0d6	143	9.1	1.0
c7t0d7	143	9.2	1.0
c7t1d0	143	9.2	1.0
c7t1d1	143	9.0	1.0

*Figure 68. IOSTAT Report from HP-UX*

## Chapter 7. Sizing Logical Volumes

In this chapter we detail the various logical volume sizes available and how they are reported and appear to RS6000, Sun, HP, and Windows NT hosts. We do not examine the AS/400 volume sizes as they are restricted to the two common volume sizes representing the 4.5 and 9.1 GB disk drives typically available to AS/400s.

### 7.1 Logical Volume Sizes

Because the VSS is based on a RAID 5 array configuration, it is possible to provide varying logical volume sizes as SCSI LUNs to attached hosts. The available sizes are: 0.5, 1, 2, 4, 8, 12, 16, 20, 24, 28, 32, 40, 48, and 56 GB. These sizes are the pure physical space capacities based on the industry standard for disk drives of 1 MB being equal to 1,000,000 bytes and therefore are not the actual sizes of LUNs seen by a particular host. Each host type, SUN, HP, AIX, and Windows NT formats the LUN to its own specifications and therefore sees slightly different available volume sizes. In addition, the way the hosts calculate the volume size for an LUN differs among hosts. Two types of information are provided to the host for an LUN to indicate the LUN size: the read capacity of the device presented in logical blocks in response to the SCSI read capacity command, and the physical size of the device in terms of cylinders, heads, and sectors per track provided in response to the SCSI mode sense command. AIX, HP-UX and Windows NT calculate the LUN size based on the read capacity provided \* 512 (the sector size in bytes). Sun OS calculates the LUN size based on the physical specifications provided (Sun OS reserves two cylinders for alternates):

$$(\text{cylinders}-2) * \text{heads} * \text{sectors} * 512$$

Table 9 shows the various LUN sizes as allocated by the VSS and the characteristics for each LUN size reported in terms of cylinders, heads, sectors, and blocks. The size of each LUN is also shown as calculated by Sun OS and as seen by AIX, HP-UX, and Windows NT, based on the number of blocks reported. This is based on an industry standard memory megabyte of  $1024*1024 = 1048576$  bytes.

Table 9. LUN Capacities Related to Their Physical and Logical Sizes

VSS LUN Size GB 10 <sup>9</sup>	Cylinders	Heads	Sectors	Bytes per Sector	LUN Size Sun OS	512 Byte Blocks	LUN Size GB 2 <sup>30</sup>
0.5	508	30	64	512	0.46	976,576	0.47
1	1017	30	64	512	0.93	1,953,152	0.93
2	2034	30	64	512	1.86	3,906,304	1.86
4	4069	30	64	512	3.72	7,812,544	3.73
8	8138	30	64	512	7.45	15,625,024	7.45
12	12207	30	64	512	11.17	23,437,504	11.18
16	16276	30	64	512	14.90	31,250,048	14.90
20	20345	30	64	512	18.62	39,062,528	18.63

VSS LUN Size GB 10 <sup>^9</sup>	Cylinders	Heads	Sectors	Bytes per Sector	LUN Size Sun OS	512 Byte Blocks	LUN Size GB 2 <sup>^30</sup>
24	24414	30	64	512	22.35	46,875,008	22.35
28	28483	30	64	512	26.08	54,687,552	26.08
32	32552	30	64	512	29.80	62,500,032	29.80
40	4768	64	256	512	37.23	78,125,056	37.25
48	5722	64	256	512	44.69	93,750,016	44.70
56	6675	64	256	512	52.13	109,375,040	52.15

## 7.2 Choosing Logical Volume Sizes

The VSS provides a wide variety of volume sizes or LUNs that can be configured. Which size is best for a particular use is somewhat difficult to say, however, the following observations should be considered: The VSS supports command queuing at the LUN level and allocates logical volumes striped at the track level across seven or eight physical disks in a RAID 5 array, so there is generally no performance problem associated with large LUN sizes. From a performance point of view, configuring two 24 GB volumes or twelve 4 GB volumes in the same array provides the same performance assuming the resultant host volumes are configured the same way and the host supports command queuing. Most current UNIX systems, AIX, HP-UX, Sun OS, and Windows NT, support large LUN sizes as a physical disk. These systems provide a volume manager that allows you to split an LUN (seen as a physical volume) into various logical volume sizes. Thus, a VSS logical volume of 16 GB could be attached to a host, and the resultant LUN of approximately 14.9 GB (see Table 9) could be split into three 4 GB volumes and one 2.9 GB volume or some other combination. Historically on UNIX systems there were problems with file system size limits of 2 GB. Most UNIX operating systems no longer have this limitation, and large file systems can be created. However, it may be desirable from an administration point of view to create discrete file systems for specific uses. Usually it would not be desirable for performance to place some file systems together on the same volume. With the Raid 5 architecture of the VSS subsystem, this issue can be reexamined and more flexibility in file system placement is possible.

Windows NT users usually like to use large LUN sizes to provide file system space because of the limited number of drive letters that LUNs can be associated with, and sometimes because of applications that require a file system to reside on only one drive (Windows NT drive letter). The VSS can currently provide a maximum LUN size of 56 GB, which provides approximately 52 GB of usable space to Windows NT. Allocating a volume in this way provides a single, large volume image and significantly better performance than would be provided by a single large physical disk. Windows NT Disk Administrator also provides the ability to split an LUN into up to four disks and associate a drive letter with each.

---

## Chapter 8. Error Notification and Recovery Procedures

In this chapter we detail some procedures for error recovery and describe some error identification and reporting tools, along with actions to be performed.

---

### 8.1 Error Log and Reporting Tools

VSS has several ways of detecting and reporting errors in operation, hardware failures, or configuration changes. These events can be collected by different methods to provide enough information for system administration and maintenance.

#### 8.1.1 VSS Error Logging Architecture

Because the clusters inside the VSS are RS/6000-based machines at present, the RISC-based hardware error detection and reporting applies. Each cluster has some nonvolatile random access memory (NVRAM) that stores information about hardware microcode levels, low-level errors, and basic hardware details. This memory can be read when running the cluster in service mode, and data can be collected for error analysis.

This information is also shared with the microcode that runs on each cluster, to generate error reports and warnings in case of critical hardware errors or upon the reaching thresholds for recoverable errors. Because the microcode is based on AIX, standard AIX error procedures are used. The basic component of the error logging architecture is the *errdemon*, which collects and stores information in an error log file (called *errlog*).

To create an entry in the error log, the *errdemon* daemon retrieves the appropriate template from the repository, and the resource name of the unit that caused the error, and detailed data. Also, if the error indicates a hardware-related problem, and hardware VPD exists, the daemon retrieves the VPD from the ODM. The error is stored in a special binary file (the *errlog*) that can be retrieved later to generate alerts.

Most entries in the error log are related to hardware and microcode problems, but informational messages can also be logged.

#### 8.1.2 Alerts Generation

Once an error (or a notification) has been detected and stored in the *errlog*, it must be reported to system administrators through:

- e-Mail
- SNMP
- Call home

Customers can select any or all of these options, which can be used simultaneously. They must be enabled through the service terminal. Upon receipt of an alert, system administrators can check whether or not actions must be taken and use the appropriate procedure.

### **8.1.2.1 e-Mail**

VSS uses standard TCP/IP SMTP to send e-mail error reports or notifications. Notifications are sent when configuration changes are performed, or internal microcode is updated, for example.

The users defined to receive e-mail can be selected to receive either error reports or notifications, or both. The configuration steps for e-mail support are described in Section 2.3, “Configuring e-Mail” on page 23.

### **8.1.2.2 SNMP**

When a network monitor is available (such as Tivoli Netview), VSS can be configured to send SNMP traps, allowing the network monitor to perform the VSS monitoring. The VSS also supports a read-only MIB, which can provide additional information about errors in VSS.

The MIB file is available with the VSS microcode and can be used by the SNMP manager to monitor both error reports and notifications.

VSS can be configured to send SNMP traps to more than one host (see Section 2.4, “Configuring the Simple Network Management Protocol” on page 24).

### **8.1.2.3 Call Home Support**

Call home support allows the VSS to call an IBM support center and send data when an error occurs. Data sent from the VSS is used to identify hardware or microcode problems and perform corrective actions.

Call home support can be activated through the service terminal, as described on Section 2.5.1, “Call-Home Support” on page 25.

---

## **8.2 Recovery Procedures**

Whenever a failure occurs in VSS, recovery procedures must be used to correct the problem. Some procedures must also be used as predictive maintenance, to avoid unplanned downtime.

Usually, hardware problems require a physical action such as changing adapters, cables, or connectors. Some workarounds can be performed to allow operation before maintenance is done.

Disk failures, adapter failures, and cluster failures play a major role in system availability and thus are subject to recovery procedures.

Nearly all of the recovery procedures should be done by an IBM CE and are not intended to be performed by the customer. Some of them, however, can be done through the VS Specialist by the customer.

### **8.2.1 Recovery Menus in the Service Terminal Interface**

Most of the recovery actions performed on the VSS are based on the use of the service terminal. The IBM CE uses the service terminal for initial configuration of VSS and for maintenance.

The service terminal can be accessed by connecting a terminal directly to the cluster. Once connected, a main panel is shown, with several menus. These menus are for maintenance purposes, as we show below.

Three basic menus are used for recovery actions.

The Repair Menu (Figure 69) is used for alternate cluster power on and off and cluster quiescing and resuming. In case of emergency power off, the customer should use this menu to perform VSS shutdown.

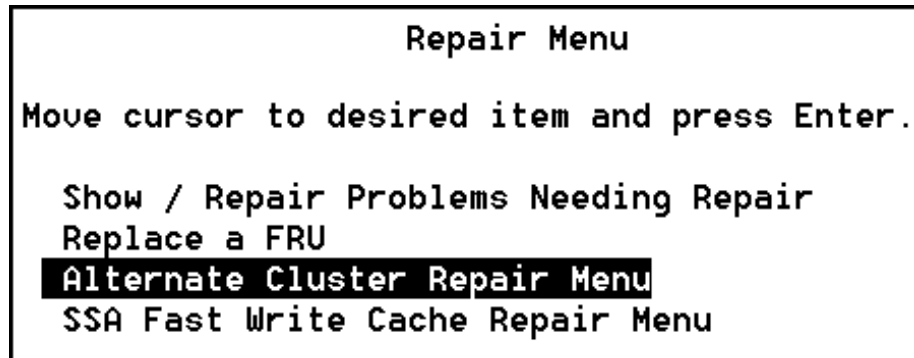


Figure 69. Repair Menu

The Configuration Option Menu (Figure 70) allows you to perform some initial configurations and also save and restore the VSS configuration data on a diskette.

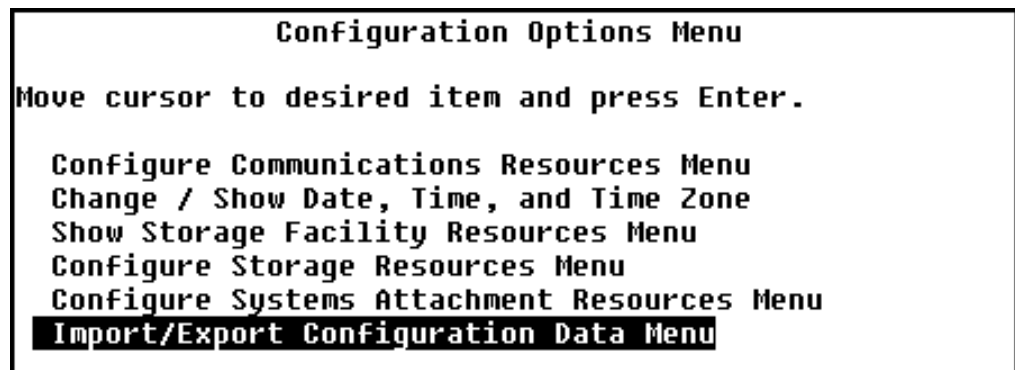


Figure 70. Configuration Options Menu

The Utility Menu (Figure 71) presents some trace facilities and a resource panel and enables you to save the customization data (for example, information about communications and users) on a diskette.

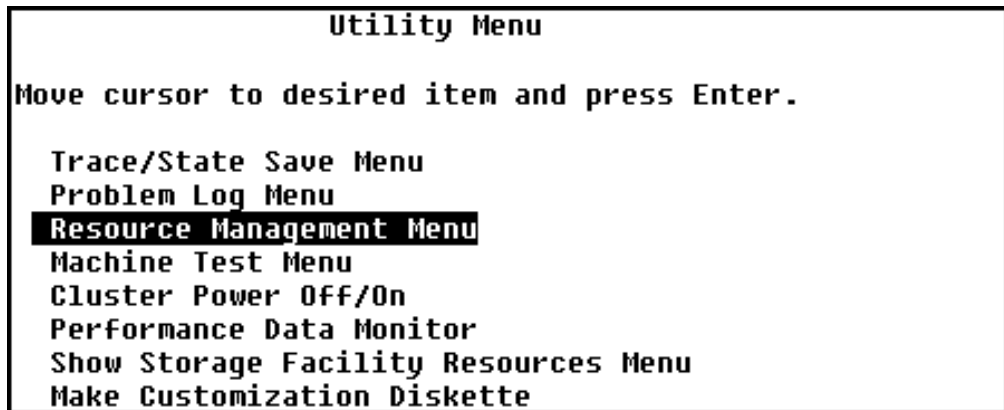


Figure 71. Utility Menu

## 8.2.2 Recovering from Adapter Failures

Usually, when a SCSI adapter fails, it has to be replaced in order to get the functionality back. Within VSS, however, configurations can be defined or changed to provide recovery, in certain situations, even before changing the adapter.

### 8.2.2.1 Storage Relocation

Logical volumes inside the VSS can be relocated between SCSI ports. If the relocation is between ports connected to the same host, you can recover from SCSI failures by relocating the volumes to the other SCSI port. After the relocation, the disk will be accessible again, and, using operating system commands, you can restart the applications that access the data.

In Figure 72 on page 106, SCSI ports 12 and 15 are connected to the same host, and both ports have been configured to connect to *Cluster 0*. The defined logical volume (shown in the fourth rank of *Cluster 0*, from top to bottom) is connected to the host across SCSI port 12. In case of a SCSI failure, the storage can be relocated to SCSI port 15, and data access will be restored.

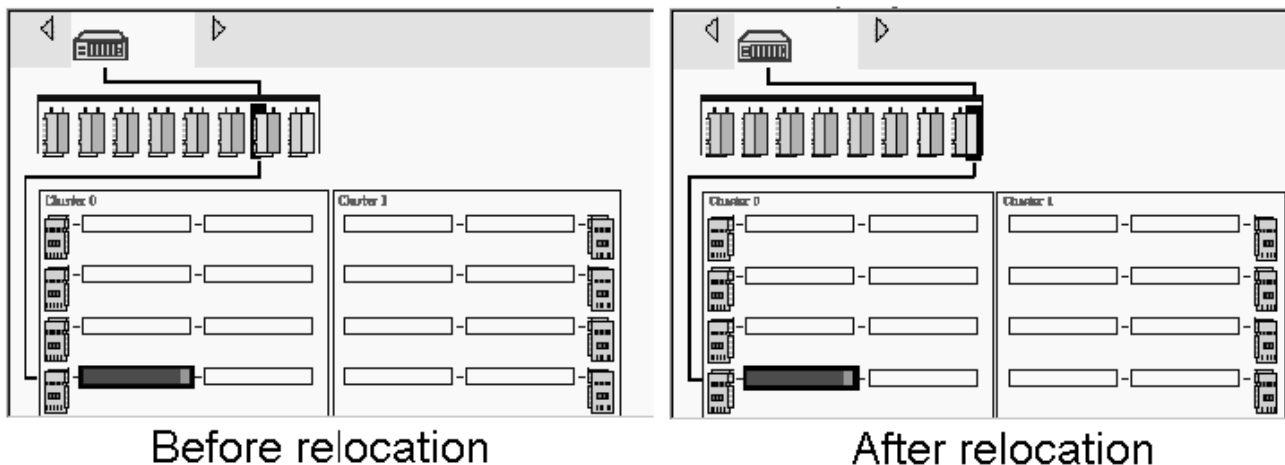


Figure 72. Storage Relocation Used to Recover from SCSI Failures



Note that, if the second SCSI adapter on the host is not connected to any logical volume (as in the example), some performance degradation may occur, due to relocation of the logical volume. The newly relocated logical volume shares the same SCSI adapter that the original logical volume uses. If both logical volumes are used by I/O-intensive applications, some contention may occur. As soon as the failed adapter has been replaced, the original allocation should be returned, to prevent further performance hits.

For the VS Specialist panels used to perform this action, see Section 3.5.6.4, “Modify Volume Assignments” on page 50.

### 8.2.2.2 Adapter Replacement

Before changing the adapter, you must take some action to prevent error generation within the VSS. Because the bay that contains the failed adapter must be removed, it must be quiesced first.

Note: These are CE functions and should not be attempted by customer personnel.

Resources (such as the adapter bays) should be quiesced from the service terminal, under the Resource Management Menu (see Figure 73).

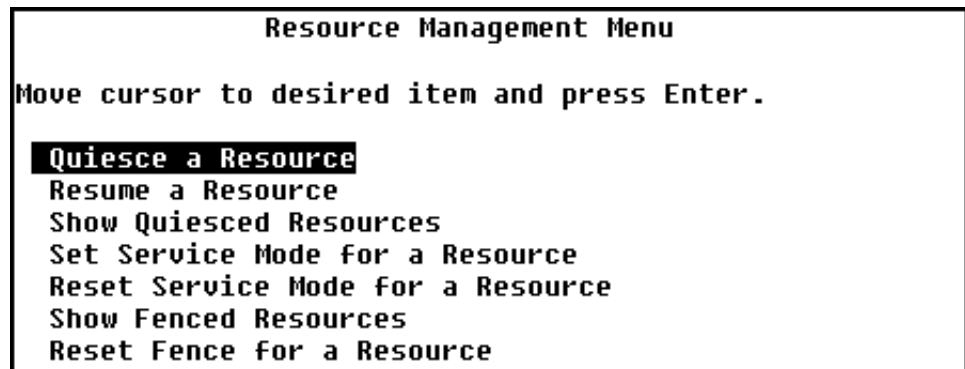


Figure 73. Quiesce a Resource Option

When the **Quiesce a Resource** option is selected, a new panel opens, and the bay in which the adapter resides should be selected. Once quiesced, the adapter bay can be removed, and the failed adapter can be replaced.

In the case of an SSA adapter failure, the nonvolatile FWC card should be removed from the failed adapter and installed on the new adapter. This will ensure data integrity, because data saved in the nonvolatile FWC but not destaged to disk will be destaged when the new adapter starts operation.

After changing the adapter and replacing the bay, the resource should be unquiesced, to be brought back to operation.

### 8.2.3 Recovering from SSA Disk Failures

SSA single disk failures are nondisruptive, because VSS uses RAID-5 protection. Also, whenever an SSA disk fails, the rebuild process starts automatically on the hot spare disk available on the loop.

When the IBM CE replaces the defective drive, it is automatically defined as a hot spare drive, without need of further configuration.

**Note:** When a customer has received notification of a disk failure, replacement of the defective disk should be scheduled at the earliest convenience to reduce the exposure, though unlikely, to a second disk failure on the loop.

#### 8.2.4 Recovering from Cluster Failures

The VSS has internal recovery procedures in case of a cluster failure. No special actions should be required to recover from such a failure. For a description of the failover process, see Section 5.1.3, "Storage Processor Cluster Failures" on page 80.

Once a failover has occurred, all I/Os are done by one cluster. The failed cluster can be serviced and later brought back into operation. This process is called *failback* and should be performed as soon as both clusters are operational (see Figure 74 on page 108). Note, however, that in the failback process, some disruption may occur because access to data that originally belonged to the failed cluster will be interrupted. When the failback process completes, normal data access is restored.

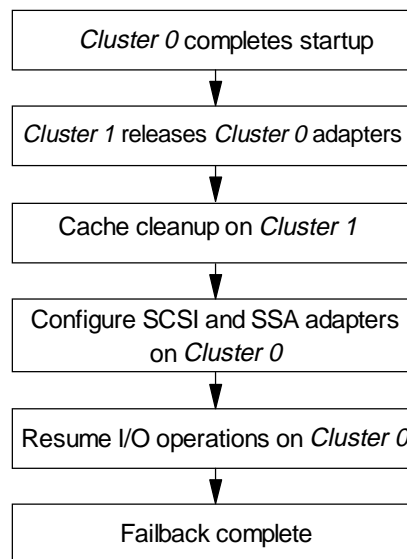


Figure 74. Simplified Diagram of the Failback Process

Cluster failures can be caused by hardware electronics, cable connections between SMP and the bay, and internal disk failures. All these are detectable and repair should be done by the IBM CE.

#### 8.2.5 Recovering from Internal SMP Disk Failures

In case of internal disk failures, the CE must take some actions after replacing the failed disk. Because configuration data is stored on the internal disk, it has to be restored after a disk change. This is not a customer recovery procedure. It is intended to be done by the IBM CE.

The data inside the internal disk consists of three parts: VSS microcode, VSS customization data (such as communications parameters, list of users), and configuration information (hosts, ports, and logical volumes defined). All of the data is stored in both machines. In case of internal disk failure in one cluster, the other cluster can be used to restore the information onto the replaced disk.

Both configuration and customization data can be saved on diskettes and used to reconstruct the information inside the cluster. Because this information is stored on both clusters, there is no need to save it prior to a failure. If one cluster goes down, the information can be stored on diskettes using the remaining cluster.

After replacing the internal disk, the VSS microcode is installed, and the customization and configuration data is restored from the diskettes. After a shutdown, the cluster is ready to be brought to service again.

### 8.2.6 Recovering from Network Failures

A network failure can isolate both clusters from the network or one cluster from another. In both cases, data access is not affected, only the connection between VSS and the VS Specialist.

In the first case, if the network segment that connects the clusters to the outside network is not working, nothing happens within the VSS. The clusters will not notice the failure, and once the network is operational again, users can access the VS Specialist.

In case of a network failure that isolates one cluster from the other, some actions must be performed to return the system to the normal state. If *Cluster0* is isolated from the network (either from adapter or cable failures), it will not be able to send alerts about the failure. At regular intervals, both clusters run a "healthcheck" procedure. When *Cluster1* runs this procedure, it detects that *Cluster0* is unreachable and logs the network failure. It sends an e-mail with the failure information.

If the failure is in the internal network adapter, a cluster failover is required to fix the problem. Hardware maintenance must be done before restoring normal operation.

Once the network failure has been resolved, normal operation resumes.

### 8.2.7 Performing an Emergency Shutdown

In case of an emergency, the VSS can be powered off using the EPO switch. However, this is an extreme procedure and may cause some problems, because the internal microcode was not shut down properly. Whenever a power off is needed, the VSS should be shut down properly.

The VSS has a power switch that, when set to the **0** position, initiates shutdown procedures on both clusters. After three minutes, both clusters and all 7133 drawers are powered down.

When the power switch is set to the **1** position, the 7133 drawers are powered up, and both clusters power on and start loading the microcode. After some minutes the system is ready for operation.

**Note:** Power sequencing does not apply for 7133s in non-VSS racks.



---

## Appendix A. Special Notices

This publication is intended to help storage administrators configure the VSS for performance and availability. The information in this publication is not intended as the specification of any programming interfaces that are provided by the VSS or Versatile Storage Specialist. See the PUBLICATIONS section of the IBM Programming Announcement for Versatile Storage Server 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, 500 Columbus Avenue, Thornwood, NY 10594 USA.

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:

IBM ®  
AIX  
RS/6000  
AS/400

The following terms are trademarks of other companies:

C-bus is a trademark of Corollary, Inc.

Java and HotJava are trademarks of Sun Microsystems, Incorporated.

Microsoft, Windows, Windows NT, and the Windows 95 logo are trademarks or registered trademarks of Microsoft Corporation.

PC Direct is a trademark of Ziff Communications Company and is used by IBM Corporation under license.

Pentium, MMX, ProShare, LANDesk, and ActionMedia are trademarks or registered trademarks of Intel Corporation in the U.S. and other countries.

UNIX is a registered trademark in the United States and other countries licensed exclusively through X/Open Company Limited.

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

---

## Appendix B. Related Publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this redbook.

---

### B.1 International Technical Support Organization Publications

For information on ordering these ITSO publications see “How to Get ITSO Redbooks” on page 115.

- *IBM Versatile Storage Server*, SG24-2221
- *Managing and Monitoring IBM SSA Subsystems*, SG24-5251
- *Exploiting SSA in Sun Solaris Platform Environments*, SG24-5083
- *Implementing PC Serveraid SCSI and SSA RAID Disk Subsystems*, SG24-2098
- *A Practical Guide to Serial Storage Architecture for AIX*, SG24-4599

---

### B.2 Redbooks on CD-ROMs

Redbooks are also available on CD-ROMs. **Order a subscription** and receive updates 2-4 times a year at significant savings.

CD-ROM Title	Subscription Number	Collection Kit Number
System/390 Redbooks Collection	SBOF-7201	SK2T-2177
Networking and Systems Management Redbooks Collection	SBOF-7370	SK2T-6022
Transaction Processing and Data Management Redbook	SBOF-7240	SK2T-8038
Lotus Redbooks Collection	SBOF-6899	SK2T-8039
Tivoli Redbooks Collection	SBOF-6898	SK2T-8044
AS/400 Redbooks Collection	SBOF-7270	SK2T-2849
RS/6000 Redbooks Collection (HTML, BkMgr)	SBOF-7230	SK2T-8040
RS/6000 Redbooks Collection (PostScript)	SBOF-7205	SK2T-8041
RS/6000 Redbooks Collection (PDF Format)	SBOF-8700	SK2T-8043
Application Development Redbooks Collection	SBOF-7290	SK2T-8037

---

### B.3 Other Publications

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 Subsystems: Installation Guide*, GA33-3260
- *7133 Models 500 and 600 SSA Subsystems: Installation Guide*, GA33-3263
- *7133 SSA Subsystem: Service Guide*, SY33-0185
- *SSA 4 Port Adapter and Enhanced SSA 4 Port Adapter: Technical Reference*, S31H-8612
- *SSA RAID Subsystems: Planning SSA RAID Subsystems*, GA33-3271
- *SSA Adapters: User's Guide and Maintenance Information*, SA33-3272

- *MicroChannel SSA RAID Adapters Technical Reference, SA33-3270*
- *PCI SSA RAID Adapters Technical Reference, SA33-3225*
- *AIX Versions 3.2 and 4 Performance Tuning Guide, SC23-2365*



---

## How to Get ITSO Redbooks

This section explains how both customers and IBM employees can find out about ITSO redbooks, CD-ROMs, workshops, and residencies. A form for ordering books and CD-ROMs is also provided.

This information was current at the time of publication, but is continually subject to change. The latest information may be found at <http://www.redbooks.ibm.com/>.

---

## How IBM Employees Can Get ITSO Redbooks

Employees may request ITSO deliverables (redbooks, BookManager BOOKs, and CD-ROMs) and information about redbooks, workshops, and residencies in the following ways:

- **Redbooks Web Site on the World Wide Web**

<http://w3.itso.ibm.com/>

- **PUBORDER** – to order hardcopies in the United States

- **Tools Disks**

To get LIST3820s of redbooks, type one of the following commands:

```
TOOLCAT REDPRINT
TOOLS SENDTO EHONE4 TOOLS2 REDPRINT GET SG24xxxx PACKAGE
TOOLS SENDTO CANVM2 TOOLS REDPRINT GET SG24xxxx PACKAGE (Canadian users only)
```

To get BookManager BOOKs of redbooks, type the following command:

```
TOOLCAT REDBOOKS
```

To get lists of redbooks, type the following command:

```
TOOLS SENDTO USDIST MKTTOOLS MKTTOOLS GET ITSOCAT TXT
```

To register for information on workshops, residencies, and redbooks, type the following command:

```
TOOLS SENDTO WTSCPOK TOOLS ZDISK GET ITSOREGI 1998
```

- **REDBOOKS Category on INEWS**

- **Online** – send orders to: USIB6FPL at IBMMAIL or DKIBMBSH at IBMMAIL

### Redpieces

For information so current it is still in the process of being written, look at "Redpieces" on the Redbooks Web Site (<http://www.redbooks.ibm.com/redpieces.html>). 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.

---

## How Customers Can Get ITSO Redbooks

Customers may request ITSO deliverables (redbooks, BookManager BOOKs, and CD-ROMs) and information about redbooks, workshops, and residencies in the following ways:

- **Online Orders** – send orders to:

In United States  
In Canada  
Outside North America

**IBMMAIL**  
usib6fpl at ibmmail  
caibmbkz at ibmmail  
dkibmbsh at ibmmail

**Internet**  
usib6fpl@ibmmail.com  
lmannix@vnet.ibm.com  
bookshop@dk.ibm.com

- **Telephone Orders**

United States (toll free)  
Canada (toll free)

1-800-879-2755  
1-800-IBM-4YOU

Outside North America  
(+45) 4810-1320 - Danish  
(+45) 4810-1420 - Dutch  
(+45) 4810-1540 - English  
(+45) 4810-1670 - Finnish  
(+45) 4810-1220 - French

(long distance charges apply)  
(+45) 4810-1020 - German  
(+45) 4810-1620 - Italian  
(+45) 4810-1270 - Norwegian  
(+45) 4810-1120 - Spanish  
(+45) 4810-1170 - Swedish

- **Mail Orders** – send orders to:

IBM Publications  
Publications Customer Support  
P.O. Box 29570  
Raleigh, NC 27626-0570  
USA

IBM Publications  
144-4th Avenue, S.W.  
Calgary, Alberta T2P 3N5  
Canada

IBM Direct Services  
Sortemosevej 21  
DK-3450 Allerød  
Denmark

- **Fax** – send orders to:

United States (toll free)  
Canada  
Outside North America

1-800-445-9269  
1-800-267-4455  
(+45) 48 14 2207 (long distance charge)

- **1-800-IBM-4FAX (United States) or (+1) 408 256 5422 (Outside USA)** – ask for:

Index # 4421 Abstracts of new redbooks  
Index # 4422 IBM redbooks  
Index # 4420 Redbooks for last six months

- **On the World Wide Web**

Redbooks Web Site <http://www.redbooks.ibm.com>  
IBM Direct Publications Catalog <http://www.elink.ibm.com/pbl/pbl>

### Redpieces

For information so current it is still in the process of being written, look at "Redpieces" on the Redbooks Web Site (<http://www.redbooks.ibm.com/redpieces.html>). 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.





---

## List of Abbreviations

<b>AC</b>	alternating current	<b>SRAM</b>	synchronous dynamic random access memory
<b>APA</b>	all points addressable	<b>SSA</b>	serial storage architecture
<b>API</b>	application programming interface	<b>SSL</b>	Secure Sockets Layer
<b>CE</b>	customer engineer	<b>VPD</b>	vital product data
<b>DC</b>	direct current	<b>VS Specialist</b>	Versatile Storage Specialist
<b>DDM</b>	disk drive module	<b>VSS</b>	Versatile Storage Server
<b>ECC</b>	error correcting code		
<b>EPO</b>	emergency power off		
<b>FWC</b>	fast write cache		
<b>HTML</b>	Hypertext Markup Language		
<b>HTTP</b>	Hypertext Transfer Protocol		
<b>IBM</b>	International Business Machines Corporation		
<b>ITSO</b>	International Technical Support Organization		
<b>LBA</b>	logical block address		
<b>LCD</b>	liquid crystal display		
<b>LIC</b>	Licensed Internal Code		
<b>LRC</b>	longitudinal redundancy check		
<b>LUN</b>	logical unit number		
<b>MESI</b>	Modified / exclusive / shared /invalid		
<b>MIB</b>	management information base		
<b>NVRAM</b>	nonvolatile random access memory		
<b>NVS</b>	nonvolatile storage		
<b>ODM</b>	object data manager		
<b>PCI</b>	peripheral component interconnect		
<b>PDU</b>	power distribution unit		
<b>PE</b>	product engineer		
<b>RAS</b>	reliability accessibility serviceability		
<b>SDRAM</b>	synchronous dynamic random access memory		
<b>SFOI</b>	Storage Function Operation Interface		
<b>SMP</b>	symmetric multiprocessor		
<b>SNMP</b>	Simple Network Management Protocol		



---

# Index

## Numerics

4 way SMP clusters 89  
524-byte sector format 16  
6+P+S 16  
604e CPU chips 6  
7+P 16  
7133-010 13  
7133-020 13

## A

adapter installation sequence 38  
adapter replacement 107  
add volumes 46  
advanced performance configuration sample 94  
affinity  
    changing 71  
    setting 70  
alerts generation 103  
array throughput rules of thumb 97  
arrays 15  
availability 77  
availability test scenarios 85  
Available hosts list box 61

## B

basic performance configuration sample 93  
bays 9  
best fit 48

## C

cache size 90  
cache sizing for effective host caching environments 91  
cache sizing for less effective host caching environments 91  
call home support 25, 104  
CD-ROM 8  
choosing logical volume (LUN) sizes 102  
Clear View button 39  
client code 28  
cluster affinity 69  
cluster load balancing 91  
clusters 5  
Communications panel 37  
Configuration Options menu 105  
configure  
    e-Mail 23  
    logical volumes 92  
    SCSI interface - performance considerations 91  
    Simple Network Management Protocol 24  
    SSA adapters 92  
    TCP/IP 21  
    VS Specialist 22  
Configure Host Ports panel 44  
connect a host to a port 60  
connections and configuration status 30

CPU Level 1 cache 7  
CPU Level 2 cache 8

## D

data integrity 82  
defining a host 57  
device emulation 11  
disk adapter bus 10  
disk interface adapter 11  
disk interface failures 78  
diskette drive 8

## E

e-mail alerts 104  
emergency shutdown 109  
errdemon 103  
errlog 103  
error log and reporting tools 103  
error logging architecture 103  
error notification and recovery procedures 103  
Ethernet 8

### Examples

    add a host 58  
    add a volume to a host port 68  
    adding a host to a port 62  
    adding a volume to a host 65  
    AIX server storage attachment 75  
    changing the cluster affinity 73  
    HP server storage attachment 76  
    removing a host 60  
    removing a host from a port 64  
    removing a volume from a host port 69  
    Sun server storage attachment 76  
    Windows NT server storage attachment 75

## F

failback process 108  
failures  
    disk interface 78  
    host connection 77  
    power 81  
    storage processor cluster 80  
    types of 77  
fast write cache 12, 78  
    removal 107

## H

HACMP 83  
high availability configuration 83  
host adapter bus 10  
host connection failures 77  
host connections 57  
host interface adapter 11  
host name 46

## I

- IBM I/O Path Optimizer 77
- information server 28
- installation activities 1
- installation procedures
  - overview 19
- interface adapters 10
- internal SCSI adapter 8
- Introduction panel 32
- IOSTAT reports 98

## L

- large cache memory 89
- LCD 9
- Licensed Internal Code panel 54
- logical volume configuration 92
- logical volume sizes 64
- logical volume sizing 101
- longitudinal redundancy check 16
- LUN capacities 101
- LUN support 57

## M

- manage logical volumes on VSS 64
- minimal disruption 82
- modem setup 37
- Modify Host Systems panel 43
- modify volume assignments 50, 67

## N

- network with the VSS 21
- nonvolatile storage 12

## O

- optional battery 5

## P

- PCI bus structure 89
- PCI buses 9
- performance and tuning 89
- performance configuration options 90
- performance expectations 95
- performance fundamentals 89
- performance monitoring and reporting 98
- performing an emergency shutdown 109
- plan for maximum availability 82
- port and cluster affinity 69
- power failures 81
- PowerPC CPUs 6
- Problem Notification panel 36
- problem severity colors 34

## R

- RAID-5 11
- read/write cache 7
- recovery from adapter failures 106

- recovery from cluster failures 108
- recovery from internal SMP disk failures 108
- recovery from network failures 109
- recovery from SSA disk failures 107
- recovery menus 104
- recovery procedures 104
- redundant power supplies 81
- remote diagnostics and support 25
- Repair Menu 105
- rules of thumb 97

## S

- SCSI disk drive 8
- SCSI interface configuration 91
- Secure Sockets Layer 28
- sequence of activities 1
- serial ports 8
- server starting message 31
- server storage attachment examples 75
- service processor 9
- SNMP alerts 104
- software-based failover 77
- sort 42
- SSA adapter
  - configuration 92
  - memory 11
- SSA adapters 90
- SSA cable failure testing 86
- SSA disk drawer 13
- SSA disk drive module 14
- Status panel 34
- Storage Allocation -- Graphical View
  - View Utilization 39
- Storage Allocation -- Graphical View panel 38
- Storage Allocation -- Tabular View panel 41
- storage allocation display colors 40
- storage allocation tasks 43
- storage cluster failures 87
- storage function operation interface 28
- storage processor bus 10
- storage processor cluster failures 80
- storage relocation 106
- StorWatch Versatile Storage Specialist 2
- subsystem throughput expectations 97
- system availability 82

## U

- unrelated host or device 45
- user interfaces
  - service terminal 18
  - Versatile Storage Specialist 17
- Users panel 52
- Utility Menu 106

## V

- volatile storage 12
- VS Specialist
  - components 28



connecting to 30

#### Panels

- Add Volumes 46
- Communications 37
- Configure Host Ports 44
- Introduction 32
- LIC 54
- Modify Host Systems 43
- Modify Volume Assignments 50
- Problem Notification 36
- Status 34
- Storage Allocation 38
- Users 52

prerequisites 27

README file 56

security 29

VSS availability 77

VSS installation procedures

overview 19

VSS overview 1

## W

#### Web

[www.ibm.com/storage/vss](http://www.ibm.com/storage/vss) 1

[www.storage.ibm.com/hardsoft/products/vss/vssattach](http://www.storage.ibm.com/hardsoft/products/vss/vssattach) 57

Web server 28

Web-based tool 17



---

# ITSO Redbook Evaluation

Configuring the IBM VSS for Performance and Availability  
SG24-5279-00

Your feedback is very important to help us maintain the quality of ITSO 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](mailto: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!)**

---

---

---

---

---

**SG24-5279-00**  
**Printed in the U.S.A.**

