Compaq StorageWorks™

HSJ80 Array Controller
ACS Version 8.5J-2

Configuration Planning Guide

Second Edition (August 2000)
Part Number: EK–HSJCP–PA. B01
Compaq Computer Corporation

Printed in the U.S.A.

Compaq StorageWorks HSJ80 Array Controller ACS Version 8.5J-2 Configuration Planning Guide
Second Edition (August 2000)
Part Number EK–HSJCP–PA. B01

Contents

Chapter 2

Planning Storage

Figures

Tables

About This Guide

This document presents the configuration planning guidelines for the HSJ80 Array Controller and storagesets running Array Controller Software (ACS) Version 8.5J-2.

The first chapter describes the configuration rules within which the HSJ80 Array Controller needs to operate. It also provides a description of the available cabling options to the cluster. The selection of a particular cabling option affects how you will configure the controller to operate within the cluster.

The second chapter contains information to guide you in the selection of the types of storage for the subsystem. Together, these two chapters should enable you to plan the total subsystem configuration needed for your business.

This book does *not* contain information about the operating environments to which the controller may be connected, nor does it contain detailed information about subsystem enclosures or their components. See the documentation that accompanied these hosts and peripherals for their respective information.

Conventions

This book uses the text conventions and special notices described in the following sections.

Text Conventions

Convention	Meaning
UPPERCASE	Command syntax that must be entered exactly as shown, for example: SET FAILOVER COPY=OTHER_CONTROLLER
UPPERCASE	CLI command name discussed within text, for example: “Use the SHOW SPARESET command to show the contents of the spareset.”
`Monospaced`	`Screen display as shown in the text.`
sans serif italic or *serif italic*	Command variable or numeric value that you supply, for example: SHOW *RAIDset-name* or SET THIS_CONTROLLER CHUNKSIZE=*n*
serif italic	Reference to other book titles, for example: See the *HSJ80 Array Controller Troubleshooting Resources Guide* for details.
. . .	Indicates that a portion of an example or figure has been omitted.

Special Notices

This book does not contain detailed descriptions of standard safety procedures. However, it does contain warnings for procedures that could cause personal injury, and cautions for procedures that could damage the controller or its related components. Look for these symbols when you are carrying out the procedures in this book:

WARNING: A *Warning* contains information essential to human safety. It advises users that failure to take or avoid a specific action could result in physical harm to the user. ***Use a warning, not a caution, when such damage is possible.***

CAUTION: A *Caution* contains information that the user needs to know to avoid damaging the software or hardware.

IMPORTANT: An *Important* note is a type of note that provides information essential to the completion of a task. Users can disregard information in an important note and still complete a task, but some error or failure may result.

NOTE: A *Note* emphasizes or supplements important points of the main text. It supplies information that may apply only in special cases—for example, memory limitations, equipment configurations, or details that apply to specific versions of a program.

Related Publications

The following table lists some of the documents related to the use of the controller, cache module, and external cache battery.

Document Title	6-3 Part Number	2-5-2 Part Number
Compaq StorageWorks HSJ80 Array Controller ACS Version 8.5J-2 Installation and Configuration Guide	N/A	AA-RN17A-TE
Compaq StorageWorks HSJ80 Array Controller ACS Version 8.5J-2 Solution Software Kit Overview	N/A	EK-HSJSO-OA
Compaq StorageWorks HSJ80 Array Controller ACS Version 8.5J-2 Release Notes	N/A	EK-HSJAA-RA
Compaq StorageWorks HSJ80 Array Controller ACS Version 8.5J-2 CLI Reference Guide	N/A	EK-HSJCL-RA
Compaq StorageWorks HSJ80 Array Controller ACS Version 8.5J-2 Maintenance and Service Guide	N/A	EK-J80MS-SA
Compaq StorageWorks HSJ80 Array Controller ACS Version 8.5J-2 Troubleshooting Resources Guide	N/A	EK-J80TR-SA
Compaq StorageWorks HSJ80 Array Controller ACS Version 8.5J-2 Command Console 2.3 Application Notes	N/A	EK-HSJAN-AA

Revision History

First Edition: December 1999

Second Edition: August 2000

Chapter 1

Planning a Subsystem

This chapter contains concepts that will help you plan your subsystem:

- "Terminology" on page 1–1
    - "Controller Designations A and B" on page 1–2
    - "Controller Designations: this and other" on page 1–2
- "Failover" on page 1–3
- "Configuration Rules" on page 1–3
- "Selecting a Cache Mode" on page 1–4
- "Assigning Unit Numbers" on page 1–8
- "Cluster Interconnect Options" on page 1–8

When you have conceptually planned the subsystem, you can use Chapter 2 as a guide to plan the subsystem storage devices. You will then need to use the plans developed from Chapters 1 and 2 to configure both the controller and subsystem storage. The configuration procedures to use are located in the *Compaq StorageWorks HSJ80 Array Controller ACS Version 8.5J-2 Installation and Configuration Guide*.

Terminology

When configuring the subsystem, you will encounter the following terms and concepts that you must understand:

- controller A and controller B
- this controller and other controller

Controller Designations A and B

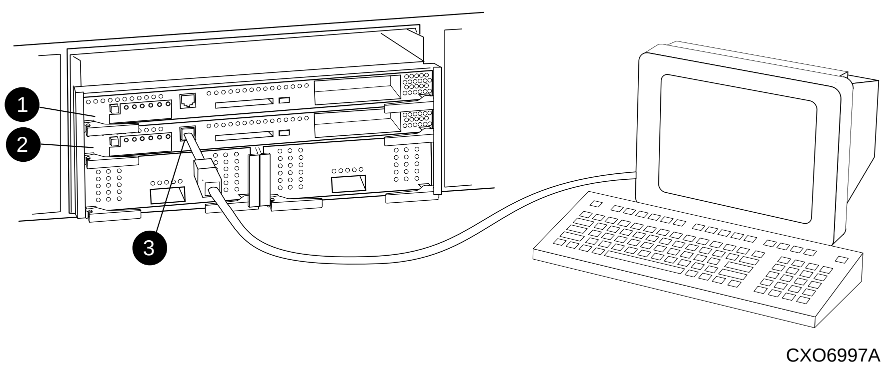
Controllers and cache modules are designated either A or B, depending on their location within the storage enclosure. The relationship is a physical one and is illustrated in Figure 1–1, which shows a horizontal mounting of the controller enclosure. Vertical mountings have the controller pair on the visual left: A = left, B = right.

Controller Designations: this and other

Some CLI commands use the terms *this* and *other* to identify one controller or the other in a dual-redundant pair. These designations are defined as follows:

- this controller—the controller that is the focus of the CLI session. That is, the controller through which the CLI commands are being entered (may be controller A or B). The maintenance (local) terminal is connected to the maintenance port of this controller.
- other controller—the controller that is not the focus of the CLI session and through which CLI commands are not being entered. The maintenance terminal is not connected to the other controller.

The relationship is a relative one defined by the location of the maintenance port cable, as illustrated in Figure 1–1.



❶ Controller A (other controller in this sample maintenance port connection)

❷ Controller B (this controller in this sample maintenance port connection)

❸ Controller Maintenance Port (with serial cable attached)

Figure 1–1. Controller Locations and Terminology

Failover

Dual-redundant controller configurations have a feature called *failover*. Failover keeps the storage array available to the cluster in the event of a single-controller failure by allowing the surviving controller to take control of the entire subsystem.

CI bus-based subsystems have both controllers connected to the same storage units. Because both controllers service the same storage units, either controller can continue to service subsystem storage if its companion controller fails.

If you configure the subsystem storage devices before setting the controllers to a failover condition, make sure you know which controller has the good storage configuration information before entering the CLI command SET FAILOVER COPY=this_controller. This command places the configuration information into the "other" controller.

Configuration Rules

Before you can configure controller storage, you will need to review the following configuration rules to ensure your configuration meets the requirements and conditions:

- Maximum 36 physical storage devices using DS-BA356-S and DS-BA356-MW StorageWorks enclosures (dual-redundant controller configurations)
- Maximum 42 physical storage devices using DS-BA356-S and DS-BA356-MW StorageWorks enclosures (single-controller configurations)
- Maximum 72 storage devices in the Model 4200 enclosure
- Maximum 2 host ports per HSJ80 controller
- Maximum 16 nodes per cluster
- Maximum 32 nodes per cluster when using a Computer Interconnect Star Coupler Extender (CISCE) on the cluster
- Maximum 512 GB unit capacity
- Maximum 8 partitions per storageset or individual disk
- Storageset Maximums (general)
  - 24 members per storageset
  - 30 RAID-1 storagesets
  - 20 RAID-5 storagesets
  - 45 RAID-0/1/5 storagesets
  - 30 RAID-1/5 storagesets

- ❒ 20 RAID-5 and RAID-1 storagesets for single controller configurations
- ❒ Up to 6 members per mirrorset (RAID-1)
- ❒ From 2 to 24 members per stripeset (RAID-0)
- ❒ From 3 to 14 members per RAIDset (RAID-3/5)
- ❒ 48 devices per mirrored stripeset (24 devices x 2)

Selecting a Cache Mode

Before selecting a cache mode, you should understand the caching techniques supported by the cache module. The cache module supports the following caching techniques to increase the performance of the subsystem read and write operations:

- Read caching
- Read-ahead caching
- Write-through caching
- Write-back caching

These cache modes are described in the following paragraphs.

Read Caching

Read caching decreases the subsystem response time to a read request by allowing the controller to satisfy the read request from the cache memory, rather than from the disk drives.

When the controller receives a read request from the host, it reads the data from the disk drives, delivers it to the host, and stores it in cache memory. If the host requests the same data again, the controller can satisfy the read request from the cached data, rather than re-reading it from the disk drives. By default, read caching is enabled for all storage units.

Refer to the commands SET *unit* MAXIMUM_CACHED_TRANSFER and SET *unit* MAX_READ_CACHED_TRANSFER_SIZE in the *Compaq StorageWorks HSJ80 Array Controller ACS Version 8.5J-2 CLI Reference Guide* for more detail.

Read-Ahead Caching

Read-ahead caching begins once the controller has processed a read request and it receives a sequential read request from the host. If the controller does not find the data in the cache memory, it reads the data from the disks and sends it to the cache memory.

The controller then anticipates subsequent read requests and fetches the next block or blocks of data from the disks as it sends the requested read data to the host. This is a parallel action. The controller notifies the host of the read completion; subsequent sequential read requests are satisfied through the cache memory. By default, read-ahead caching is enabled for all disk units. Refer to the command SET *unit* READAHEAD_CACHE in the *Compaq StorageWorks HSJ80 Array Controller ACS Version 8.5J-2 CLI Reference Guide* for more detail.

Write-Through Caching

Write-through caching also decreases the subsystem response time to a read request by allowing the controller to satisfy the request from the cache memory rather than from the disk drives.

When the controller receives a write request from the host, it stores the data in its cache memory, writes the data to the disk drives, then notifies the host when the write operation is complete. This process is called write-through caching because the data actually passes through—and is stored in—the cache memory on its way to the disk drives. If the host requests the recently written data, the controller satisfies the read request from its cache memory, rather than from the disk drives.

If read caching is enabled for a storage unit, write-through caching is also enabled. Also, because both caching techniques enhance the controller's read performance, write-through caching is automatically disabled when you disable read caching.

By default, read caching (and, therefore, write-through caching) is enabled for all storage units.

Write-Back Caching

This caching technique decreases the subsystem response time to write requests by allowing the controller to declare the write operation complete as soon as the data reaches its cache memory. The controller performs the slower operation of writing the data to the disk drives at a later time.

By default, write-back caching is disabled for all storagesets. The controller will not provide write-back caching to a unit unless you ensure that the cache memory is non-volatile as described in the "Non-Volatile Memory" section that follows.

Refer to the commands SET *unit* MAXIMUM_CACHED_TRANSFER and SET *unit* MAX_WRITE_CACHED_TRANSFER_SIZE in the *Compaq StorageWorks HSJ80 Array Controller ACS Version 8.5J-2 CLI Reference Guide* for more detail.

Non-Volatile Memory

The controller cannot provide write-back caching to a unit unless its cache memory is non-volatile. In other words, you must provide a back-up power source to the cache module to preserve the unwritten cache data in the event of a power failure. If the cache memory is volatile—that is, if it has no back-up power—the unwritten cache data is lost during a power failure.

By default, the controller expects to use an external cache battery (ECB) as the cache module back-up power source to enable it to be non-volatile (see the *Compaq StorageWorks HSJ80 Array Controller ACS Version 8.5J-2 Maintenance and Service Guide* for more information about the ECB).

If the subsystem is backed up with a UPS as an alternative to the ECB, a switch must be set in the CLI command SET this_controller UPS=DATACENTER_WIDE or SET this_controller UPS=NODE_ONLY or SET this_controller NOUPS. See the *Compaq StorageWorks HSJ80 Array Controller ACS Version 8.5J-2 CLI Reference Guide* for details about using the CLI Command: SET *this_controller* UPS.

> **NOTE:** The controller executes multiple write operations to satisfy a single write request for a RAIDset or mirrorset. For this reason, a RAIDset or mirrorset requires non-volatile cache memory to ensure data redundancy until the write request is satisfied.

Regardless of the back-up power source you choose, the cache-memory power LED flashes about once every three seconds to indicate the cache module memory array is receiving power from its primary power source.

Fault-Tolerance for Write-Back Caching

The cache module supports the following features to protect the availability of its unwritten (write-back) data:

- Non-volatile memory (required for write-back caching)
- Mirrored caching (optional)
- Dynamic caching techniques (automatic)

Mirrored Caching

To further ensure the availability of unwritten cache data, you can use a portion of each cache module memory to mirror the other cache module’s write-back data in a dual-redundant configuration. Refer to the command SET this_controller (or other_controller) MIRRORED_CACHE command in the *Compaq StorageWorks HSJ80 Array Controller ACS Version 8.5J-2 CLI Reference Guide* for more detail.

Figure 1–2 shows the principle of mirrored caching: half of cache A mirrors cache B write-back data and vice versa. This arrangement ensures that the write-back data is preserved if a cache module or any of its components fail.

> **NOTE:** When your controllers are configured to use mirrored write-back cache, the cache capacity is half of the total amount of cache in the configuration. If each cache module has 512 MB of cache for a total of 1024 MB of cache in the configuration, the cache capacity is 512 MB.

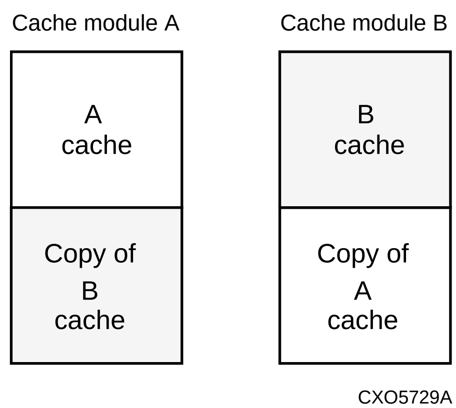


Figure 1–2. Mirrored Caching

Before configuring dual-redundant controllers and enabling mirrored write-back cache, make sure the following conditions are met:

- Both controllers have the same size cache (512 MB).
- Diagnostics indicate that both cache modules are good.
- Both cache modules have a battery present (unless you have enabled the UPS switch). A battery does not have to be present for either cache if you enable the UPS switch.
- No unit errors are outstanding; for example, lost data or data that cannot be written to devices.
- Both controllers are started and configured in failover mode.

For important considerations when replacing or upgrading memory in a mirrored cache configuration, see the *Compaq StorageWorks HSJ80 Array Controller ACS Version 8.5J-2 Maintenance and Service Guide*.

Dynamic Caching Techniques

If the controller detects a full or partial failure of its cache module or ECB, it automatically reacts to preserve the cached write-back data. Then, depending upon the severity of the failure, the controller chooses an interim caching technique—also called the cache policy—which it uses until you repair or replace the cache module or ECB.

Assigning Unit Numbers

Every container (storageset, partition, or JBOD disk) needs a unit number to communicate to the hosts. Each unit number contains the following:

- A letter that indicates the device type in the storage unit.

  D for disk drives (including optical drives)

  T for tape drives

- A number that must be in the range 0 through 4094, which uniquely identifies the storage unit (4095 and 4096 are reserved).

The unit number is placed into the system with the ADD UNIT *unit-number container-name* command. This command assigns a unit number to the container name established with the ADD DISK or ADD TAPE command.

For a detailed description of these commands, refer to the *Compaq StorageWorks HSJ80 Array Controller ACS Version 8.5J-2 CLI Reference Guide*.

Cluster Interconnect Options

The HSJ80 may be cabled to the cluster in several more options than were previously available with HSJ40 and HSJ50 subsystems. The supported cabling options are listed below and described in the paragraphs that follow:

- Single HSJ80 Array Controller, single host port, single Star Coupler, single cluster
- Dual HSJ80 Array Controllers, single host port, single Star Coupler, single cluster
- Dual HSJ80 Array Controllers, dual host port, dual Star Couplers, single cluster
- Dual HSJ80 Array Controllers, dual host port, dual Star Couplers, multiple clusters

Single or Dual Controller, Single Cluster

A single or dual HSJ80 controller subsystem configured to use only one of the two available host ports functions in a manner similar to that of the earlier HSJ40 and HSJ50 subsystems. This mode of operation requires that the second port (the unused port) on the controller be disabled.

Either CI port may be used, but the port not being used must be disabled using the SET this_controller NOPORT_1_PATH_A (or NOPORT_2_PATH_B) CLI command. Figure 1–3 shows the single controller configuration and Figure 1–4 the dual-controller configuration.

This cabling configuration is the basic HSJ80 upgrade option; there are no subsystem configuration changes required.

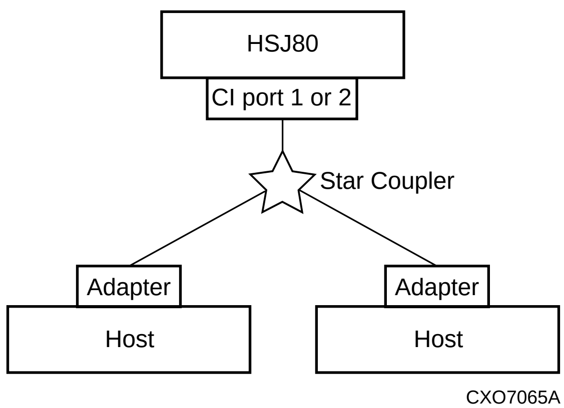


Figure 1–3. Single Controller, Single Host Port, Single Cluster Configuration

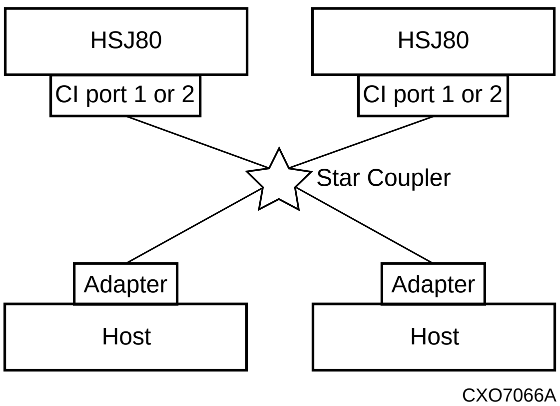


Figure 1–4. Dual Controller, Single Host Port, Single Cluster Configuration

Dual Controller, Dual Host Port, Single Cluster

This cabling configuration may be used to correct the problem in which the host data rate exceeds the capability of a single host adapter. Adding the second host adapter and cabling it to the second HSJ80 host port (by way of a second Star Coupler) allows a parallel path to storage.

The dual controller, dual host port, dual Star Coupler, dual host adapters option allows a single cluster to effectively double the throughput of storage available to it. This option also improves availability by being capable of surviving a failure of either an HSJ80 CI host port or host adapter. Figure 1–5 shows the configuration.

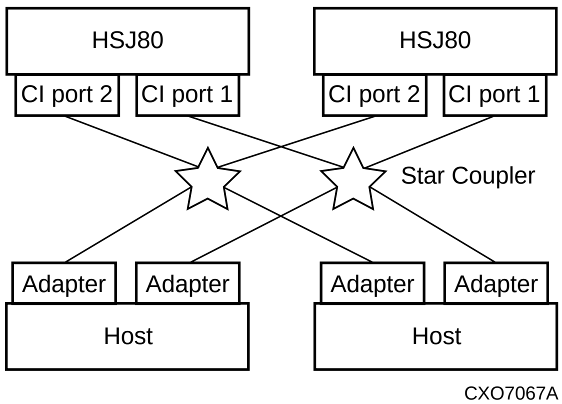


Figure 1–5. Dual Controller, Dual Host Port, Dual Star Coupler Configuration

Dual Controller, Multiple-Cluster Configuration

The approved cabling method for a multiple-cluster subsystem configuration is shown in Figure 1–6. This configuration gives the subsystem better throughput, better space utilization, and reduced maintenance cost. The reduced maintenance cost is derived from the use of a single dual-redundant HSJ80 subsystem being capable of replacing two HSJ50 storage subsystems.

CAUTION: When using this mode of operation, the port access of the controllers and the logical units must be set correctly or data loss may occur. See the following description for additional information.

An HSJ80 Array Controller storage subsystem may be used in a multiple-cluster configuration using dual Star Couplers and both host ports on both controllers cabled in a dual-redundant configuration.

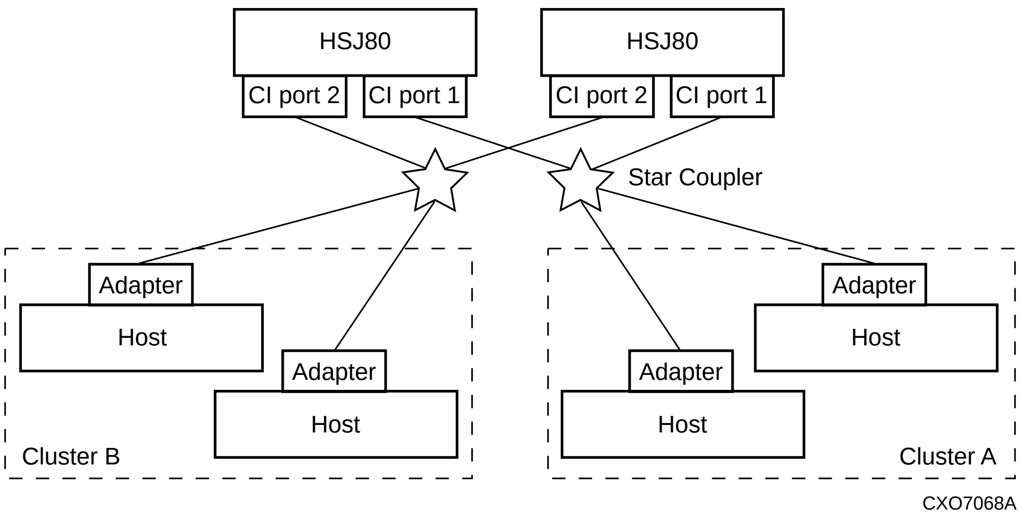


Figure 1–6. Dual Controller, Dual Host Port, Multiple-Cluster Configuration

Multiple-cluster configurations should be configured with caution because data stored on a disk with one cluster may be read and written by the other cluster if not configured properly. To avoid the problems that this could cause, you must:

- Select the containers to be used with a particular cluster.
- Enter a series of CLI commands for each container in the subsystem to make them accessible by only one of the two clusters.

The SET *unit_number* THIS_PORT_1_ACCESS= FULL (or NONE) CLI command assigns specific storage units for access by only one of the two clusters. Each container in the subsystem must be set up in the manner described in the next section.

The following example of a set of commands shows the setting of the access to port 1 to full on this controller (and other) for container D100. This is followed by the syntax to turn on the access to port 2 to full on this controller (and other) for container D200:

```
Set D100 THIS_PORT_1_ACCESS=FULL
Set D100 THIS_PORT_2_ACCESS=NONE
Set D100 OTHER_PORT_1_ACCESS=FULL
Set D100 OTHER_PORT_2_ACCESS=NONE

Set D200 THIS_PORT_2_ACCESS=FULL
Set D200 THIS_PORT_1_ACCESS=NONE
Set D200 OTHER_PORT_2_ACCESS=FULL
Set D200 OTHER_PORT_1_ACCESS=NONE
```

Using Figure 1–6 as the example, cluster A now has full access to D100, while cluster B does not have access to D100. Conversely, cluster B now has full access to D200, while cluster A does not have access to D200.

Chapter 2

Planning Storage

This chapter provides information to help you plan the storage configuration of your subsystem. Use the guidelines found in this section to plan the various types of storage containers needed.

The following information is contained in this chapter:

Where to Start

Containers are defined as individual disk drives (JBODs), storagesets of varying types (mirrorsets, stripesets, and so on), and partitioned drives. The following is a structure you can follow to plan your storage configuration. The references in each step locate details about specific commands and concepts. Appendix A provides blank templates that you may use to keep track of the containers being configured.

1. Familiarize yourself with the current physical layout of the devices and their addressing scheme. See "Device PTL Addressing Convention" on page 2–3.
2. Determine your storage requirements. Use the questions in "Determining Storage Requirements" on page 2–5 to help you.
3. Choose the type of storage containers you need in your subsystem. See "Choosing a Container Type" on page 2–6 for a comparison and description of each type of storageset.
4. Create a storageset profile (described in "Creating a Storageset Profile" on page 2–7). Fill out the storageset profile while you read the sections that pertain to your chosen storage type:
   - "Storageset Planning Considerations" on page 2–9
     - "Stripeset Planning Considerations" on page 2–9
     - "Mirrorset Planning Considerations" on page 2–12
     - "RAIDset Planning Considerations" on page 2–15
     - "Striped Mirrorset Planning Considerations" on page 2–17
   - "Partition Planning Considerations" on page 2–19
5. Decide which CLI command switches should be turned on for your subsystem. CLI command device switches apply to all devices, including those configured as single-disk units (JBOD). General information on switches is detailed in "Choosing Switches for Storagesets and Devices" on page 2–21. Detailed information is also provided in the *Compaq StorageWorks HSJ80 Array Controller ACS Version 8.5J-2 CLI Reference Guide*.
   - Determine which unit switches you want for your units (see "Device Switches" on page 2–23).
   - Determine which initialization switches you want for your planned storage containers (see "Initializing Switches" on page 2–23).
6. Create a storage map (see "Storage Maps" on page 2–29).

7. Configure the storage you have now planned using one of the following methods:
   - StorageWorks Command Console (SWCC) graphical user interface (GUI).
   - Command Line Interpreter (CLI) commands by way of a terminal or PC connected to the maintenance port of the controller. This method allows you greater flexibility in defining and naming storage containers. The *Compaq StorageWorks HSJ80 Array Controller ACS Version 8.5J-2 CLI Reference Guide* provides CLI command details.

Device PTL Addressing Convention

The controller has six I/O ports (SCSI device ports), each of which connects to a SCSI bus. Each SCSI bus connects to a storage enclosure that supports up to seven storage devices (targets) in the BA35x enclosure or up to 14 in the 4200 series enclosure. In dual-controller subsystems, these device buses are shared between the two controllers.

NOTE: The SWCC graphical user interface (GUI) calls the device ports "channels."

The controller identifies storage devices based on the SCSI Port-Target-LUN (PTL) numbering scheme. The physical location of a storage device in its enclosure determines its PTL:

- P—Designates the controller's I/O port number (1 through 6).
- T—Designates the target identification (ID) number of the device. Valid target ID numbers are 0 through 15.
- L—Designates the logical unit number (LUN) of the device. For disk devices, the LUN is always 0. Passthrough devices may have LUNs other than 0.

PTL addressing is used in the CLI commands to add device types (Disk, CD-ROM, and so on).

Figure 2–1 shows the structure of a PTL address.

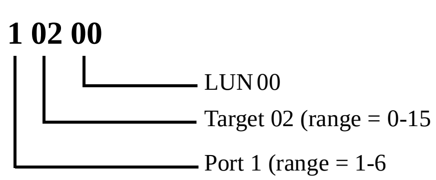


Figure 2–1. PTL Naming Convention

Figure 2–2 shows a sample SA500/SA800 rack arrangement of four device enclosures, each cabled to an I/O cable from the controller enclosure (I/O ports 1 through 4). Note that the connection of the I/O cables are in sequence from top to bottom on the illustration (the suggested manner). The target addresses are fixed within the device enclosures, with 0 on the right and 6 on the left (looking at the front of the SA500/SA800). Detailed information about this example is located in "Creating a Storage Map" on page 2–30.

	Target							PORT
		5	4	3	2	1	0	
Power Supply	Power Supply	D106 Disk10500	D105 RAID2 Disk10400	D104 S1 M3 Disk10300	D103 S1 Disk10200	D101 M1 Disk 10100	D100 RAID1 Disk10000	1
Power Supply	Power Supply	D107 Disk20500	D105 RAID2 Disk20400	D104 S1 M3 Disk20300	D103 S1 Disk20200	D101 M1 Disk20100	D100 RAID1 Disk20000	2
Power Supply	Power Supply	D108 Disk30500	D105 RAID2 Disk30400	D104 S1 M4 Disk30300	D103 S1 Disk30200	D102 M2 Disk30100	D100 RAID1 Disk30000	3
Power Supply	Power Supply	D109 Disk40500	D105 RAID2 Disk40400	D104 S1 M4 Disk40300	D103 S1 Disk40200	D102 M2 Disk40100	Spare Disk40000	4

Figure 2–2. Example PTL Addressing with a Dual-Redundant Controller Configuration

When the controller receives an I/O request, it identifies the storageset unit number for the request, and reconciles the unit number to the storageset name. From the storageset name, the controller locates the appropriate devices for the I/O request. For example, the storage unit named D100 (Figure 2–3) is a RAIDset named “RAID1.” RAID1 in this example contains DISK10000, DISK20000, and DISK30000. The controller generates the read or write request to the appropriate devices using the PTL addressing convention.

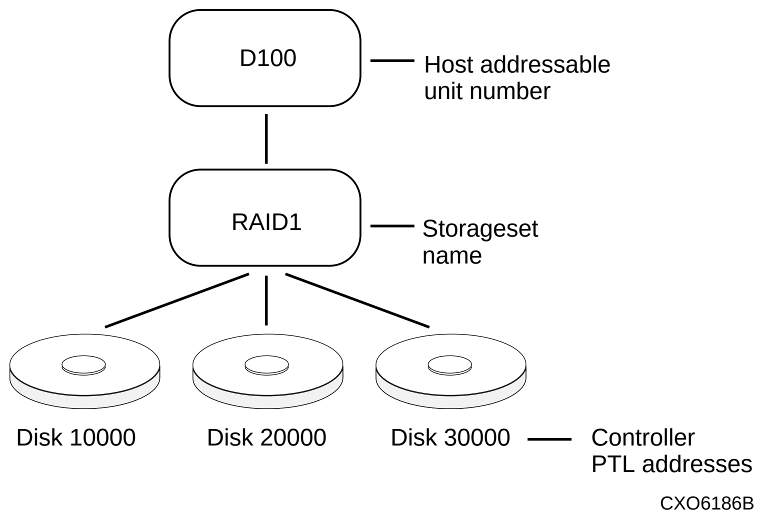


Figure 2–3. Mapping a Unit to Physical Disk Drives

Determining Storage Requirements

You cannot adequately plan your subsystem storage without determining what your storage requirements are. Here are a few questions you should ask about subsystem usage:

- What applications or user groups will access the subsystem? How much capacity do they need?
- What are the I/O requirements? If an application is data-transfer intensive, what is the required transfer rate? If it is I/O-request intensive, what is the required response time? What is the read/write ratio for a typical request?
- Are most I/O requests directed to a small percentage of the disk drives? Do you want to keep it that way or balance the I/O load?

- Do you store mission-critical data? Is availability the highest priority or would standard backup procedures suffice?

Choosing a Container Type

Different applications may have different storage requirements, so you will probably want to configure more than one kind of container within your subsystem.

In choosing a container, you choose between independent disks (JBODs) or one of several storageset types. The independent disks and the selected storageset may also be partitioned (see Figure 2–4).

Storagesets implement RAID (Redundant Array of Independent Disks) technology. Consequently, they all share one important feature: each storageset, whether it contains two disk drives or ten, looks like one large virtual disk drive to the host.

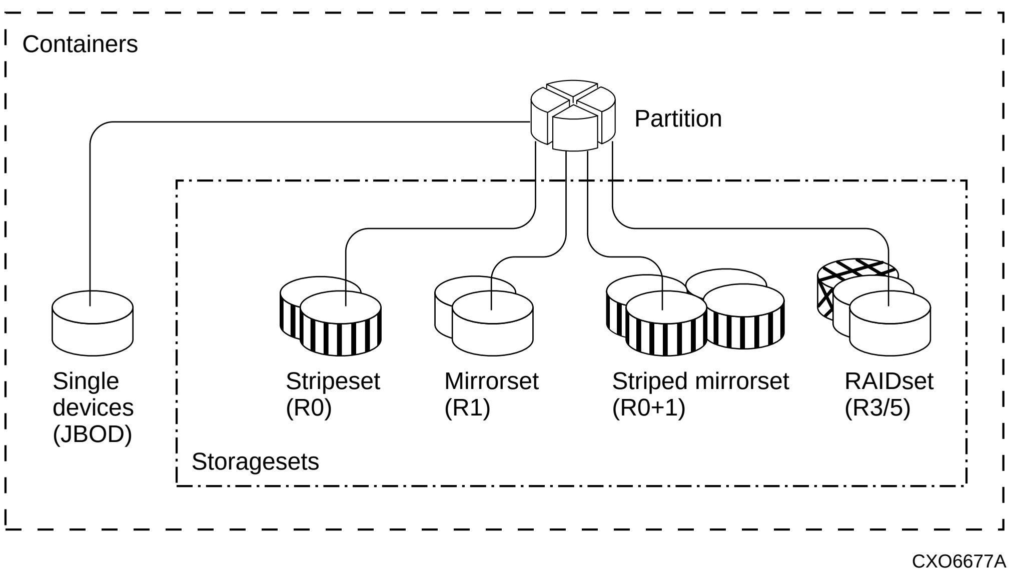


Figure 2–4. Container Types

Table 2–1 compares the different kinds of containers to help you determine which ones satisfy your requirements.

Table 2–1 A Comparison of Container Types

Container Name	Relative Availability	Request Rate (Read/Write) I/O per second	Transfer Rate (Read/Write) MB per second	Applications
Independent disk drives (JBOD)	Equal to number of JBOD disk drives	Identical to single disk drive	Identical to single disk drive	
Stripeset (RAID 0)	Proportionate to number of disk drives; worse than single disk drive	Excellent if used with large chunk size	Excellent if used with small chunk size	High performance for non-critical data
Mirrorset (RAID1)	Excellent	Good/Fair	Good/Fair	System drives; critical files
RAIDset (RAID 3/5)	Excellent	Excellent/Poor	Read: excellent (if used with small chunk sizes) Write: good (if used with small chunk sizes)	High request rates, read-intensive, data lookup
Striped Mirrorset (RAID 0+1)	Excellent	Excellent if used with large chunk size	Excellent if used with small chunk size	Any critical response-time application

For a comprehensive discussion of RAID, refer to *The RAIDBOOK—A Source Book for Disk Array Technology* (published by the RAID Advisory Board; ISBN # 1-879936-90-9).

Creating a Storageset Profile

Creating a profile for your storagesets, partitions, and devices can help simplify the configuration process. Filling out a storageset profile helps you choose the storagesets that best suit your needs and make informed decisions about the switches that can be enabled for each storageset or storage device configured in your subsystem.

See the sample storageset profile shown in Figure 2–5.

Appendix A contains blank profiles that you can copy and use to record the details for your storagesets. Use the information in this chapter to help you make decisions when creating storageset profiles.

Type of Storageset:

_____ Mirrorset __X_ RAIDset _____ Stripeset _____ Striped Mirrorset ____ JBOD

Storageset Name *R1*

Disk Drives *D10300, D20300, D30300, D40300, D50300, D60300*

Unit Number*D101*

Partitions:

Unit #	Unit #	Unit #	Unit #	Unit #	Unit #	Unit #	Unit #
%	%	%	%	%	%	%	%

RAIDset Switches:

Reconstruction Policy	Reduced Membership	Replacement Policy
_X_Normal (default)	_X_No (default)	_X_Best performance (default)
___Fast	___Yes, missing:	___Best fit
		___None

Mirrorset Switches:

Replacement Policy	Copy Policy	Read Source
___Best performance (default)	___Normal (default)	___Least busy (default)
___Best fit	___Fast	___Round robin
___None		___Disk drive:

Initialize Switches:

Chunk size	Save Configuration	Metadata
_X_ Automatic (default)	___No (default)	_X_Destroy (default)
___ 64 blocks	_X_Yes	___Retain
___ 128 blocks		
___ 256 blocks		
___ Other:		

Unit Switches:

Read Cache	Write Cache	Maximum Cache Transfer
_X_ Yes (default)	___Yes (default)	_X_32 blocks (default)
___ No	_X_No	___Other:
Write-Back Cache	**Write Protection**	**Availability**
___Yes (default)	_X_No (default)	_X_Run (default)
_X_No	Yes	___NoRun
Host Access Enabled: ____		

Figure 2–5. Sample Storageset Profile

Storageset Planning Considerations

This section provides guidelines to help you choose storageset types for your subsystem:

- Stripeset Planning Considerations
- Mirrorset Planning Considerations
- RAIDset Planning Considerations
- Striped Mirrorset Planning Considerations

Stripeset Planning Considerations

Stripesets (RAID 0) enhance I/O performance by spreading the data across multiple disk drives. Each I/O request is divided into small segments called *chunks*. These chunks are then simultaneously "striped" across the disk drives in the storageset, thereby allowing several disk drives to participate in one I/O request.

For example, in a three-member stripeset that contains disk drives Disk10000, Disk20000, and Disk30000, the first chunk of an I/O request is written to Disk10000, the second to Disk20000, the third to Disk30000, the fourth to Disk10000, and so forth until all of the data has been written to the drives (see Figure 2–6).

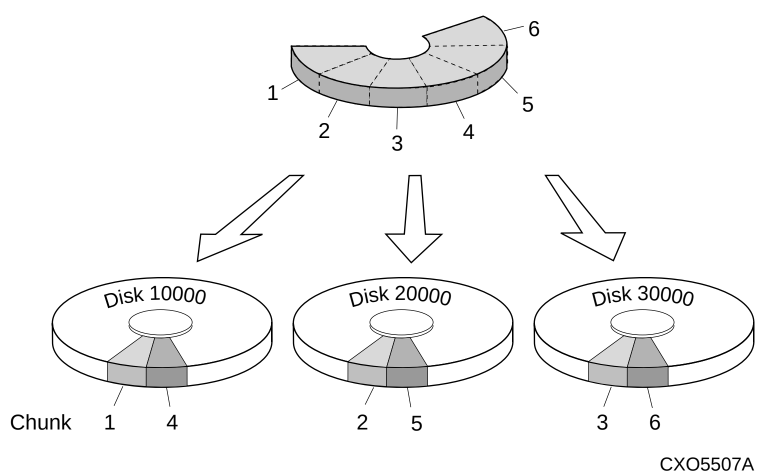


Figure 2–6. A Three-Member RAID 0 Stripeset (Example 1)

The relationship between the chunk size and the average request size determines whether striping maximizes the request rate or the data-transfer rate. You can set the chunk size or use the default setting (see "Chunk Size" on page 2–23 for information about setting the chunk size). Figure 2–7 shows another example of a three-member RAID 0 stripeset.

A major benefit of striping is that it balances the I/O load across all of the disk drives in the storageset. This can increase the subsystem performance by eliminating the hot spots (high localities of reference) that occur when frequently accessed data becomes concentrated on a single disk drive.

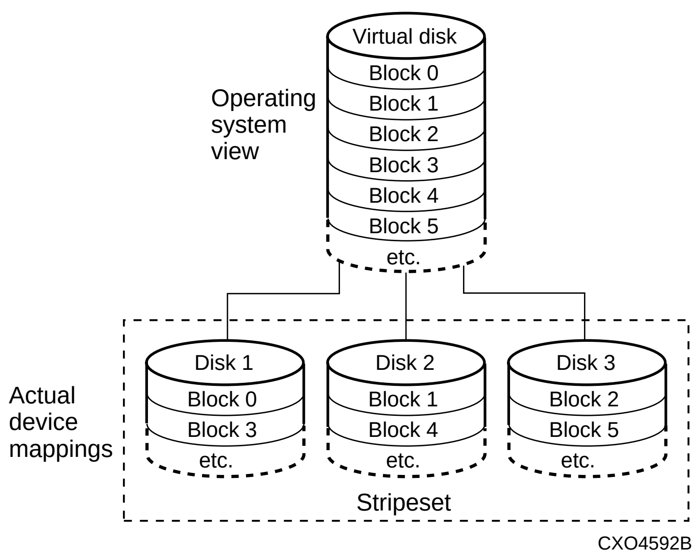


Figure 2–7. A Three-Member RAID 0 Stripeset (Example 2)

Keep the following in mind as you plan your stripesets:

- Reporting methods and size limitations prevent certain operating systems from working with large stripesets. The *Compaq StorageWorks HSJ80 Array Controller ACS Version 8.5J-2 Release Notes* and Chapter 1 in this Guide (“Configuration Rules,” page 1–3) both contain details about these restrictions.

- A storageset must contain only disk drives of the same capacity. The controller limits the capacity of each member to the capacity of the smallest member in the storageset (base member size) when the storageset is initialized. Thus, if you combine 2-GB disk drives with 1-GB disk drives in the same storageset, you will waste 1-GB of capacity on each 2-GB member.

  If you need high performance and high availability, consider using a RAIDset, striped-mirrorset, or a host-based shadow of a stripeset.

- Striping does not protect against data loss. In fact, because the failure of one member is equivalent to the failure of the entire stripeset, the likelihood of losing data is higher for a stripeset than for a single disk drive.

  For example, if the mean time between failures (MTBF) for a single disk is one hour, then the MTBF for a stripeset that comprises N such disks is l/N hours. As another example, if the MTBF of a single disk is 150,000 hours (about 17 years), a stripeset comprising four of these disks would have an MTBF of only slightly more than four years.

  For this reason, you should avoid using a stripeset to store critical data. Stripesets are more suitable for storing data that can be reproduced easily or whose loss does not prevent the system from supporting its critical mission.

- Evenly distribute the members across the device ports to balance load and provide multiple paths as shown in Figure 2–8.
- Stripesets for Array Controllers may contain between two and 24 members.
- Stripesets are well-suited for the following applications:
  - ❒ Storing program image libraries or run-time libraries for rapid loading.
  - ❒ Storing large tables or other structures of read-only data for rapid application access.
  - ❒ Collecting data from external sources at very high data transfer rates.
- Stripesets are not well-suited for the following applications:
  - ❒ A storage solution for data that cannot be easily reproduced or for data that must be available for system operation.
  - ❒ Applications that make requests for small amounts of sequentially located data.
  - ❒ Applications that make synchronous random requests for small amounts of data.

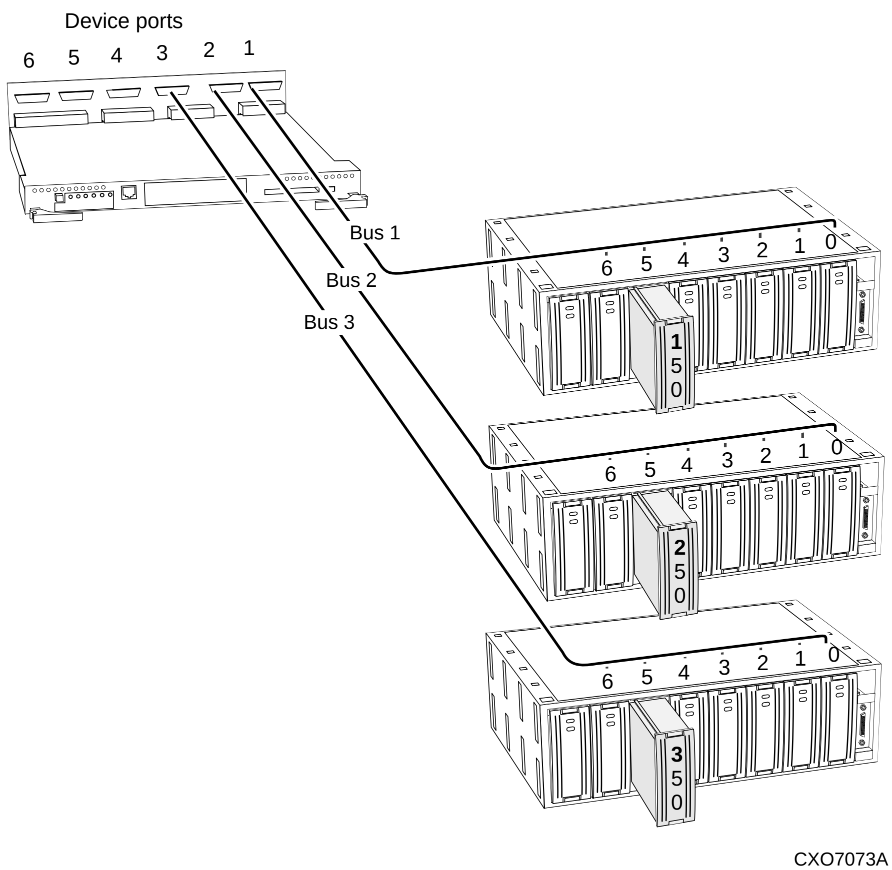


Figure 2–8. Members Distributed Across Controller I/O Device Ports

By spreading the traffic evenly across the SCSI buses (I/O device ports), you ensure that no one bus (port) handles the majority of data to the storageset.

Mirrorset Planning Considerations

Mirrorsets (RAID 1) use redundancy to ensure availability, as illustrated in Figure 2–9. For each primary disk drive, there is at least one mirror disk drive. Thus, if a primary disk drive fails, its mirror drive immediately provides an exact copy of the data. Figure 2–10 shows a second example of a Mirrorset.

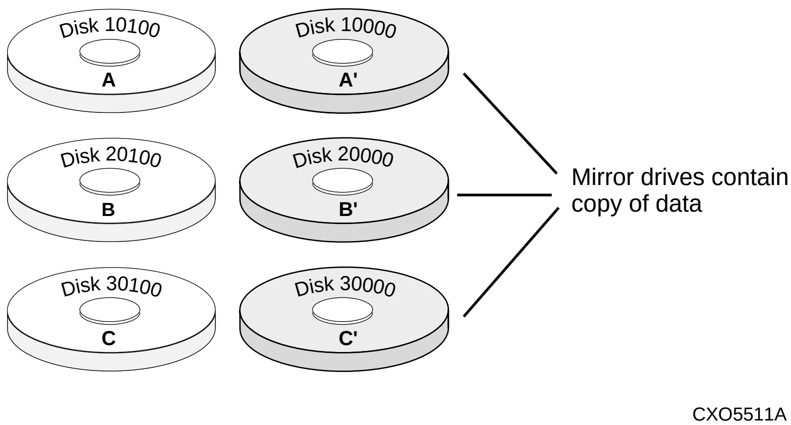


Figure 2–9. Mirrorsets Maintain Two Copies of the Same Data

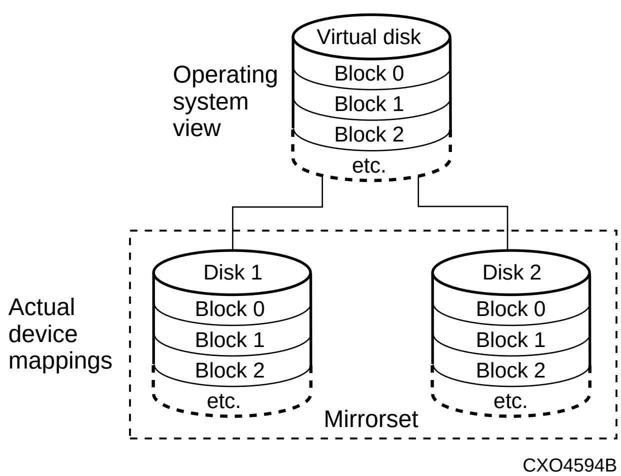
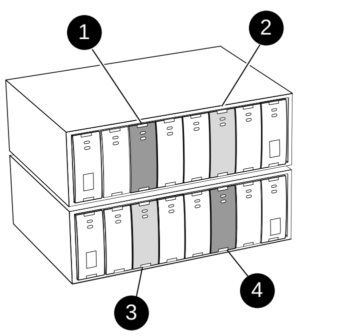


Figure 2–10. Mirrorset Example 2

Keep these points in mind as you plan your mirrorsets:

- Data availability with a mirrorset is excellent but comes with a high cost—you need twice as many disk drives to satisfy a given capacity requirement. If availability is your top priority, consider using dual-redundant controllers and redundant power supplies.
- You can configure up to 12 mirrorsets per controller or pair of dual-redundant controllers. Each mirrorset may contain up to six members.
- Both write-back cache modules must be the same size.
- If you are using more than one mirrorset in your subsystem, you should put the first member of each mirrorset on different buses, as shown in Figure 2–11. The first member of a mirrorset is the first disk drive you add.

  When a controller receives a request to read or write data to a mirrorset, it typically accesses the first member of the mirrorset. If you have several mirrorsets in your subsystem and their first members are on the same bus, that bus will be forced to handle the majority of traffic to your mirrorsets.



❶ First member of first mirrorset
❷ Second member of second mirrorset
❸ First member of second mirrorset
❹ Second member of first mirrorset

Figure 2–11. First Mirrorset Members Placed on Different SCSI Port Addresses

- Place mirrorsets and RAIDsets on different I/O ports to minimize risk in the event of a single port failure.
- Mirrorset units are defaulted to a NOWRITEBACK_CACHE condition. To increase performance, switch to WRITEBACK_CACHE.
- A storageset must contain only disk drives of the same capacity.

- Evenly distribute the members across the I/O ports to balance load and provide multiple paths as shown in Figure 2–8 on page 2–12.
- Mirrorsets are well-suited for the following:
  - Any data for which reliability requirements are extremely high
  - Data to which high-performance access is required
  - Applications for which cost is a secondary issue
- Mirrorsets are not well-suited for:
  - Write-intensive applications (a performance hit of 10% will occur)
  - Applications for which cost is a primary issue

RAIDset Planning Considerations

RAIDsets (RAID 3/5) are enhanced stripesets—they use striping to increase I/O performance and distributed-parity data to ensure data availability. Figure 2–12 illustrates the concept of a three-member RAIDset. Figure 2–13 shows a second example of a RAIDset that uses five members.

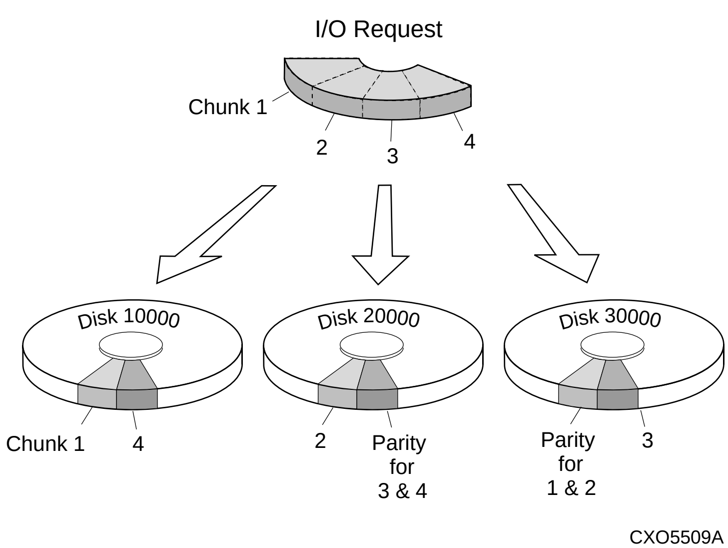


Figure 2–12. Three-Member RAIDset Using Parity (Example 1)

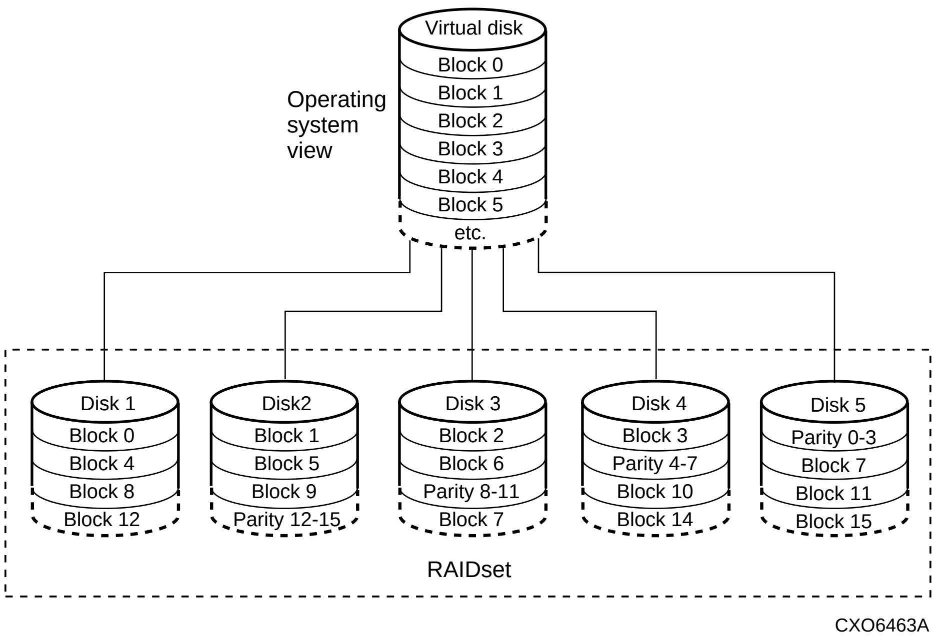


Figure 2–13. Five-Member RAIDset Using Parity (Example 2)

RAIDsets are similar to stripesets in that the I/O requests are broken into smaller “chunks” and striped across the disk drives. RAIDsets also create chunks of parity data and stripe them across all the members of the RAIDset. This parity data is derived mathematically from the I/O data and enables the controller to reconstruct the I/O data if a single disk drive fails. Thus, it becomes possible to lose a disk drive without losing access to the data it contained. Data could be lost, however, if a second disk drive fails before the controller replaces the first failed disk drive.

For example, in a three-member RAIDset that contains disk drives Disk10000, Disk20000, and Disk30000, the first chunk of an I/O request is written to Disk10000 and the second to Disk20000. Parity is then calculated and written to Disk30000, and the third chunk is written to Disk30000, the fourth to Disk10000, and so on until all of the data is saved.

The relationship between the chunk size and the average request size determines whether striping maximizes the request rate or the data-transfer rates. You can set the chunk size or use the default setting. See "Chunk Size" on page 2–23 for information about setting the chunk size.

Keep these points in mind as you plan your RAIDsets:

- Reporting methods and size limitations prevent certain operating systems from working with large RAIDsets. See the *Compaq StorageWorks HSJ80 Array Controller ACS Version 8.5J-2 Release Notes* that came with your platform kit for details about these restrictions.
- A write-back cache module is required for RAIDsets.
- Both cache modules must be the same size.
- A RAIDset must include at least three disk drives, but no more than 14.
- Evenly distribute the members across the device ports to balance the I/O load and provide multiple paths as shown in Figure 2–8 on page 2–12.
- A storageset should contain only disk drives of the same capacity. The controller limits the capacity of each member to the capacity of the smallest member in the storageset. Thus, if you combine 2-GB disk drives with 1-GB disk drives in the same storageset, you will waste 1-GB of capacity on each 2-GB member.
- RAIDset units are set to NOWRITEBACK_CACHE by default. To increase a unit's performance, switch to WRITEBACK_CACHE.
- Place RAIDsets and mirrorsets on different ports to minimize risk in the event of a single port bus failure.
- RAIDsets are particularly well-suited for the following:
  - Small to medium I/O requests
  - Applications requiring high availability
  - High read request rates
  - Inquiry-type transaction processing
- RAIDsets are not particularly well-suited for the following:
  - Write-intensive applications
  - Database applications in which fields are continually updated
  - Transaction processing

Striped Mirrorset Planning Considerations

Striped mirrorsets (RAID 0+1) are a configuration of stripesets whose members are also mirrorsets (Figure 2–14). Consequently, this kind of storageset combines the performance of striping with the reliability of mirroring. The result is a storageset with very high I/O performance and high data availability. Figure 2–15 shows a second example of a striped mirrorset using six members.

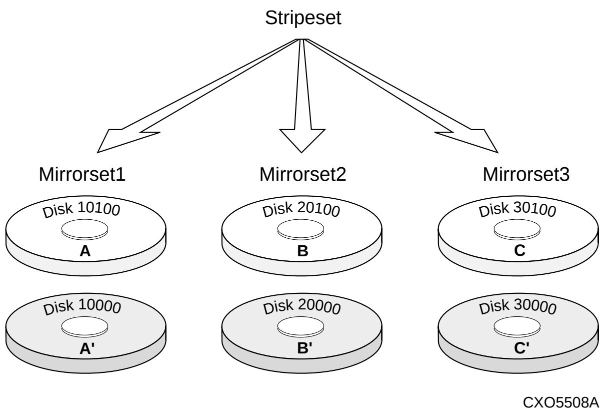


Figure 2–14. Striped Mirrorset (Example 1)

The failure of a single disk drive has no effect on the ability of the storageset to deliver data to the host. Under normal circumstances, it also has very little effect on performance. Because striped mirrorsets do not require any more disk drives than mirrorsets, this storageset is an excellent choice for data that warrants mirroring.

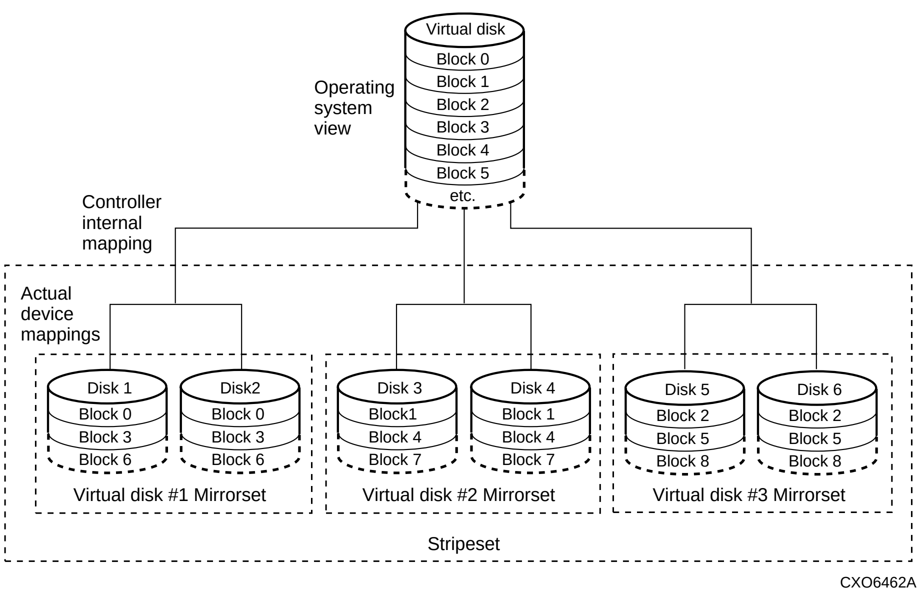
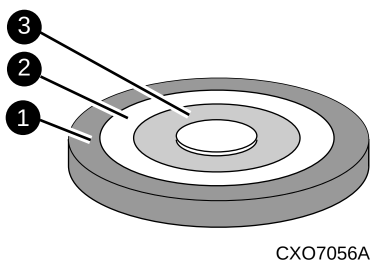


Figure 2–15. Striped Mirrorset (Example 2)

Plan the mirrorset members, then plan the stripeset that will contain them. Review the recommendations in "Storageset Planning Considerations" on page 2–9, and "Mirrorset Planning Considerations" on page 2–12.

Partition Planning Considerations

Use partitions to divide a container (storageset or individual disk drive) into smaller pieces, each of which can be presented to the host as its own storage unit. Figure 2–16 shows the conceptual effects of partitioning a single-disk container.



❶ Partition 1

❷ Partition 2

❸ Partition 3

Figure 2–16. Example of a Partitioned Single Disk Unit

You can create up to eight partitions per storageset (disk drive, RAIDset, mirrorset, stripeset, or striped mirrorset). Each partition has its own unit number so that the host can send I/O requests to the partition just as it would to any unpartitioned storageset or device. Partitions are separately addressable storage units; you can partition a single storageset to service more than one user group or application.

Defining a Partition

Partitions are expressed as a percentage of the storageset or single disk unit that contains them:

- Mirrorsets and single disk units—the controller allocates the largest whole number of blocks that are equal to or less than the percentage you specify.
- RAIDsets and stripesets—the controller allocates the largest whole number of stripes that are less than or equal to the percentage you specify.
  - Stripesets—the stripe size = chunk size × number of members.
  - RAIDsets, the stripe size = chunk size × (number of members minus 1)

An unpartitioned storage unit has more capacity than a partition that uses the whole unit because each partition requires 5 blocks of administrative metadata. Thus, a single disk unit that contains one partition can store *n* minus 5 blocks of user or application data.

Guidelines for Partitioning Storagesets and Disk Drives

Keep these points in mind as you plan your partitions:

- You can create up to eight partitions per storageset or disk drive.
- All of the partitions on the same storageset or disk drive must be addressed through the same controller. Thus, if you set a preferred controller for that unit, all the partitions in that storageset will inherit that preferred controller. This ensures a failover of devices should one of the dual-redundant controllers fail.
- Partitions cannot be combined into storagesets. For example, you cannot divide a disk drive into three partitions, then combine those partitions into a RAIDset.
- Once you partition a container, you cannot unpartition it without reinitializing the container.
- Just as with storagesets, you do not have to assign unit numbers to partitions until you are ready to use them.
- The CLONE utility cannot be used with partitioned mirrorsets or partitioned stripesets.

Loader Planning Considerations

Note the following before configuring a loader:

- If the passthrough device has more than 160 bytes of sense data available for an event, the controller returns only the first 160 bytes from the passthrough device to the host.
- You must begin unit names for passthrough devices with the letter D.
- Refer to the *Compaq StorageWorks HSJ80 Array Controller ACS Version 8.5J-2 Installation and Configuration Guide*, the section titled “Tape Configuration Rules and Recommendations,” on page 8-16, for information regarding the recommended number of tape loaders or libraries per port.

Choosing Switches for Storagesets and Devices

CLI Command switches allow the user another level of command options. The following sections describe how to enable and modify switches, and contain a description of the major CLI command switches.

Enabling Switches

If you use the StorageWorks Command Console (SWCC) to configure the device or storageset, you can set switches from the command console screens during the configuration process. The Command Console automatically applies them to the storageset or device. See the SWCC Help files or Chapter 5 of the *Compaq StorageWorks HSJ80 Array Controller ACS Version 8.5J-2 Installation and Configuration Guide* for information about using the Command Console.

If you use CLI commands to configure the storageset or device manually, the configuration procedure in Chapter 6 of the *Compaq StorageWorks HSJ80 Array Controller ACS Version 8.5J-2 Installation and Configuration Guide* shows when and how to enable each switch. The *Compaq StorageWorks HSJ80 Array Controller ACS Version 8.5J-2 CLI Reference Guide* contains the details of the CLI commands and their switches.

Changing Switches

You can change the RAIDset, mirrorset, device, and unit switches at any time. You cannot change the initialize switches without destroying the data on the storageset or device. These switches are integral to the formatting and can be changed only by reinitializing the storageset.

CAUTION: Initializing a storageset is similar to formatting a disk drive; all of the data is destroyed during this procedure.

Storage Switch Types

Depending upon the kind of storageset or device being configured, you can enable the following kinds of options or “switches”:

- RAIDset
- Mirrorset
- Partition
- Device

RAIDset Switches

You can enable the following kinds of switches to control how a RAIDset behaves to ensure data availability:

- Replacement policy
- Reconstruction policy
- Membership

For details on the use of these switches, refer to the *Compaq StorageWorks HSJ80 Array Controller ACS Version 8.5J-2 CLI Reference Guide.*

Mirrorset Switches

You can enable the following switches to control how a mirrorset behaves to ensure data availability:

- Replacement policy
- Copy speed
- Read source

For details on the use of these switches, refer to the *Compaq StorageWorks HSJ80 Array Controller ACS Version 8.5J-2 CLI Reference Guide*.

Partition Switches

The geometry parameters of a partition can be specified. The geometry switches are:

- *SIZE*—the percentage of the storage set to use for this partition (1-100)

Device Switches

When you add a disk drive or other storage device to your subsystem, you can enable the following switches:

- Transportability
- Device Transfer rate

For details on the use of these switches, refer to the *Compaq StorageWorks HSJ80 Array Controller ACS Version 8.5J-2 CLI Reference Guide*.

Initializing Switches

You can enable the following kinds of switches to affect the format of a disk drive or storageset:

- Chunk Size (for stripesets and RAIDsets only)
- Save Configuration
- Destroy/Nodestroy

Each of these is described in the following sections.

> **NOTE:** After you initialize the storageset or disk drive, you cannot change these switches without reinitializing the storageset or disk drive.

Chunk Size

Specify the chunk size of the data to be stored to control the stripe size used in RAIDsets and stripesets:

- CHUNKSIZE=DEFAULT lets the controller set the chunk size based on the number of disk drives (d) in a stripeset or RAIDset. If $d \le 9$, then chunk size = 256. If $d > 9$, then chunk size = 128.
- CHUNKSIZE=*n* lets you specify a chunk size in blocks. The relationship between chunk size and request size determines whether striping increases the request rate or the data-transfer rate.

Increasing the Request Rate—A large chunk size (relative to the average request size) increases the request rate by allowing multiple disk drives to respond to multiple requests. If one disk drive contains all of the data for one request, then the other disk drives in the storageset are available to handle other requests. Thus, in principle, separate I/O requests can be handled in parallel, thereby increasing the request rate. This concept is shown in Figure 2–17.

Applications such as interactive transaction processing, office automation, and file services for general timesharing tend to require high I/O request rates.

Large chunk sizes also tend to increase the performance of random reads and writes. It is recommended that you use a chunk size of 10 to 20 times the average request size, rounded to the closest prime number. In general, 113 works well for OpenVMS systems with a transfer size of eight sectors.

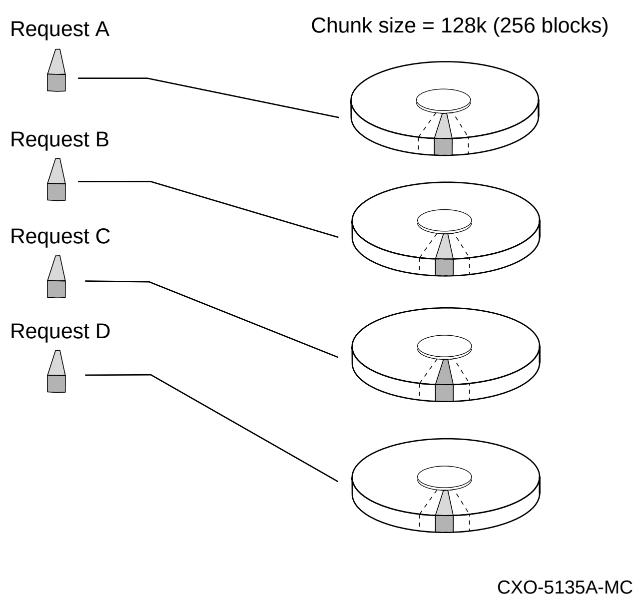


Figure 2–17. Chunk Size Larger Than the Request Size

To calculate the chunk size that should be used for your subsystem, you must first analyze the types of requests that are being made to the subsystem:

- Numerous parallel I/Os that use a small area of disk should use a chunk size of 10 times the average transfer request rate.

- Random I/Os that are scattered over all the areas of the disks should use a chunk size of 20 times the average transfer request rate.
- If you do not know, then you should use a chunk size of 15 times the average transfer request rate.
- If you have mostly sequential reads or writes (like those needed to work with large graphic files), then make the chunk size a small number (that is, 17 sectors).

Table 2–2 shows a few examples of chunk size selection.

Table 2–2 Example Chunk Sizes

Transfer Size (KB)	Small Area of I/O Transfers	Unknown	Random Areas of I/O Transfers
2	41	59	79
4	79	113	163
8	157	239	317

Increasing the Data Transfer Rate—A small chunk size relative to the average request size increases the data transfer rate by allowing multiple disk drives to participate in one I/O request. This concept is shown in Figure 2–18.

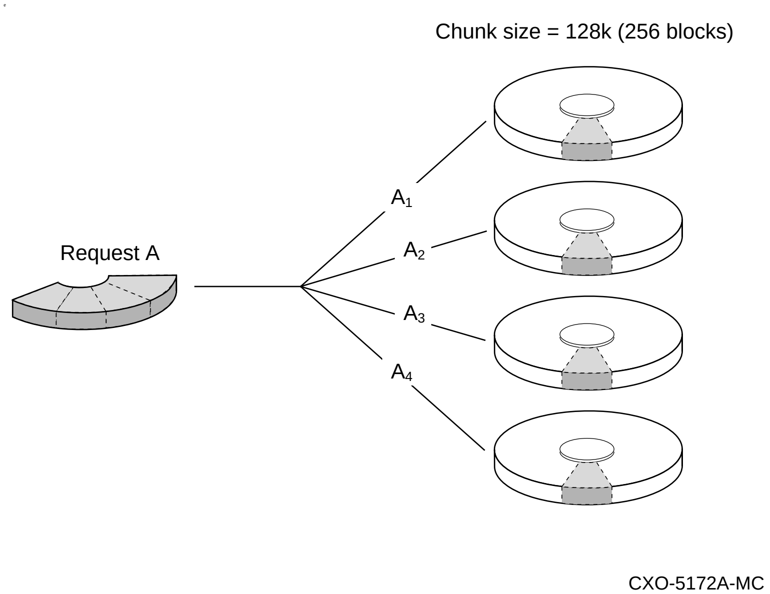


Figure 2–18. Chunk Size Smaller Than the Request Size

Applications such as CAD, image processing, data collection and reduction, and sequential file processing tend to require high data-transfer rates.

Increasing Sequential Write Performance—Sequential write and read requests on stripesets (or striped mirrorsets), should use a small chunk size relative to the I/O size to increase the performance. A chunk size of 17 generally works well.

Save Configuration

Use the CLI command: INITIALIZE *SAVE_CONFIGURATION* to indicate whether to save the subsystem configuration on the storage unit when you initialize it:

- NOSAVE_CONFIGURATION (default) means that the controller stores the subsystem configuration in its non-volatile memory only. Although this is generally secure, the configuration could be jeopardized if the controller fails. For this reason, you should initialize at least one of your disk drives with the SAVE_CONFIGURATION switch enabled.
- SAVE_CONFIGURATION allows the controller to use 256 KB of each device in a storage unit to save the subsystem configuration. The controller saves the configuration every time you change it or add a patch to your controller software. If the controller should fail, you can recover your latest configuration from the storage unit rather than rebuild it from scratch.

  The save configuration option saves:

  - All configuration information normally saved when you restart your controller except the controller serial number, product ID number, vendor ID number, and any manufacturing fault information
  - Patch information

  The save configuration option does not save:

  - Software or hardware upgrades
  - Inter-platform conversions

  Consider the following when saving the configuration:

  - It is not necessary to use the SAVE_CONFIGURATION switch for dual-redundant configurations. Use the SET FAILOVER COPY=*controller* command to restore configuration information in a replacement controller. See the *Compaq StorageWorks HSJ80 Array Controller ACS Version 8.5J-2 CLI Reference Guide* for details.

- ❒ Do not remove and replace disk devices between the time you save and restore your configuration. This is particularly important for devices that you migrate from another system. The controller could recover and use the wrong configuration information on your subsystem.
- ❒ Save your subsystem configuration as soon as possible after removing and replacing any disk devices in your subsystem. This ensures that the devices always contain the latest, valid information for your system.
- ❒ When you incorporate a spare into a storageset that you initialized with the INITIALIZE *container-name* SAVE_CONFIGURATION command, the controller reserves space on the spare for configuration information. The controller updates this information when the configuration changes.
- ❒ You cannot use a storageset that contains user data to save your subsystem configuration unless you back up and restore the user data.
- ❒ If you previously configured storagesets with the SAVE_CONFIGURATION switch, you do not need to initialize them again after you reconfigure your devices with a new controller.
- ❒ When you replace a controller, ensure the replacement controller does not contain any configuration data. If the controller is not new, it must be initialized with the CONFIGURATION RESET command. If you do not take this precaution, you can lose configuration data if non-volatile memory changes.

Destroy or Nodestroy

You must specify whether to destroy or retain the user data and metadata when a disk is initialized that has been previously used in a mirrorset or as a single-disk unit.

NOTE: The *DESTROY* and *NODESTROY* switches are valid only for mirrorsets and striped mirrorsets.

- ■ DESTROY (default) overwrites the user data and forced-error metadata on a disk drive when it is initialized.
- ■ NODESTROY preserves the user data and forced-error metadata when a disk drive is initialized. Use NODESTROY to create a single-disk unit from any disk drive that has been used as a member of a mirrorset. See the REDUCED command in the *Compaq StorageWorks HSJ80 Array Controller ACS Version 8.5J-2 CLI Reference Guide* for information on removing disk drives from a mirrorset.

NOTE: The NODESTROY switch is ignored for members of a RAIDset.

Other Switches

There are several other switches that help control the parameters around which the subsystem operates:

- Partitions
- Maximum Cache Transfer
- Read Cache
- Run
- Write Protection
- Write-Back Cache

Each of these is described in the following sections.

Partitions

Specify the partition number that identifies the partition associated with the host-addressable unit number you are adding.

PARTITION=*partition-number* allows you to identify a partition that is associated with the unit you are adding.

Maximum Cache Transfer

MAXIMUM_CACHED_TRANSFER=*n* lets you decide on the number of data blocks that the controller will not cache to satisfy a request. Any I/O transfers in excess of the specified size will not be cached (*n*= a value from 1 to 1024). The controller will cache reads that are less than or equal to the number placed in the MAXIMUM_CACHED_TRANSFER switch setting. If not specified, the default setting is 32 blocks.

The MAXIMUM_CACHED_TRANSFER switch affects both read and write-back cache when set on a controller that has read and write-back caching.

Read Cache

Enable or disable the caching of read data to the container:

- READ_CACHE (default) enables the caching of read data.
- NOREAD_CACHE disables the caching of read data.
- MAX_READ_CACHED_TRANSFER_SIZE specifies the largest read request to cache.

Run

Specify whether to make the container available to the host. The RUN/NORUN switch is not valid for partitioned units. Do not specify this switch on the SET or ADD commands for a partitioned unit.

- RUN (default) specifies that as soon as you provide a host-addressable unit number the storage unit will be made available to the host.
- NORUN specifies that the container will not be made available to the host until you specify the RUN switch.

Write Protection

Enable or disable write protection for the container:

- NOWRITE_PROTECT (default) enables the controller to write new data to the storage unit.
- WRITE_PROTECT prevents the controller from writing any new data to the storage unit. (The controller can write to a protected unit if it needs to reconstruct data.)

Write-Back Cache

Enable or disable the controller write-back caching for a container:

- WRITEBACK_CACHE enables write-back caching.
- NOWRITEBACK_CACHE (default) disables write-back caching.
- MAX_WRITE_CACHED_TRANSFER_SIZE specifies the largest write request to cache.

NOTE: If you disable write-back caching for a storage unit that previously used it, it may take up to five minutes to flush the unwritten data from the cache to the devices in the storage unit.

Storage Maps

Configuring your subsystem will be easier if you know how the storagesets, partitions, and JBODs correspond to the disk drives in your subsystem. You can more easily see this relationship by creating a hardcopy representation (a storage map). Figure 2–19 is a representative blank storage map showing a simplified physical representation of the enclosure (each cell in the map represents a disk drive in the enclosure). The location of the drive determines the PTL location.

Creating a Storage Map

If you want to make a storage map, fill out a blank storage map as you add storagesets, partitions, and JBOD disks to your configuration and assign them unit numbers. Appendix B contains templates you may use in the creation of your subsystem storage map. Label each disk drive in the map with the higher levels it is associated with, up to the unit level.

The cables that are routed from the controller I/O module connecting to the device enclosures determine the port address. Normally, the cable connected to the I/O module port 1 is routed to the upper device enclosure, as shown in Figure 2–19 (although that is not mandatory).

Example Storage Map

Figure 2–19 shows an example of a dual-redundant controller subsystem with the following configured storage:

- Unit D100 is a 6-member RAID 3/5 storageset named RAID1, which consists of Disk10000, Disk20000, Disk30000, Disk40000, Disk50000, and Disk60000.
- Unit D101 is a 6-member RAID 3/5 storageset named RAID2, which consists of Disk10100, Disk20100, Disk30100, Disk40100, Disk50100 and Disk60100.
- Unit D102 is a 3-member striped mirrorset named Stripe1, which consists of Mirror1, Mirror2 and Mirror3:
  - ❒ Mirror1 is a 2-member striped mirrorset consisting of Disk10200 and Disk 20200.
  - ❒ Mirror2 is a 2-member mirrorset consisting of Disk30200 and Disk40200.
  - ❒ Mirror3 is a 2-member mirrorset consisting of Disk50200 and Disk60200.
- Unit D103 is a 2-member mirrorset named Mirror4, which consists of Disk10300 and Disk20300.
- Unit D104 is a 4-member stripeset named Stripe2, which consists of Disk30300, Disk40300, Disk 50300, and Disk 60300.
- Unit D105 is a single (JBOD) disk named Disk10400.
- Disk20400 and Disk30400 are autospares.

		Targets						
		5	4	3	2	1	0	PORT
Power Supply	Power Supply		D105 Disk10400	D103 Mirror4 Disk10300	D102 Stripe1 Mirror1 Disk10200	D101 RAID2 Disk10100	D100 RAID1 Disk10000	1
Power Supply	Power Supply		Spare Disk20400	D103 Mirror4 Disk20300	D102 Stripe1 Mirror1 Disk20200	D101 RAID2 Disk20100	D100 RAID1 Disk20000	2
Power Supply	Power Supply		Spare Disk30400	D104 Stripe2 Disk30300	D102 Stripe1 Mirror2 Disk30200	D101 RAID2 Disk30100	D100 RAID1 Disk30000	3
Power Supply	Power Supply			D104 Stripe2 Disk40300	D102 Stripe1 Mirror2 Disk40200	D101 RAID2 Disk40100	D100 RAID1 Disk40000	4
Power Supply	Power Supply			D104 Stripe2 Disk50300	D102 Stripe1 Mirror3 Disk50200	D101 RAID2 Disk50100	D100 RAID1 Disk50000	5
Power Supply	Power Supply			D104 Stripe2 Disk60300	D102 Stripe1 Mirror3 Disk60200	D101 RAID2 Disk60100	D100 RAID1 Disk60000	6

Figure 2–19. Sample Storage Map

Using the LOCATE Command to Find Devices

If you want to complete a storage map at a later time and do not remember where everything is, you can use the CLI command LOCATE. The LOCATE command flashes the amber (fault) LED on the storage devices associated with the specific storageset or unit. To turn off the flashing LEDs, issue the CLI command LOCATE CANCEL.

The following is an example of a command that locates all the drives that make up D104:

```
LOCATE D104
```

The amber LED on the face of each device D104 flashes. After you have noted the position of all the devices contained within D104, issue the following command to turn off the flashing LEDs:

```
LOCATE CANCEL
```

The following is an example of a command that locates all the drives that make up RAIDset RAID1:

```
LOCATE RAID1
```

After you have noted the position of the devices contained within RAID1, issue the following command to turn off the flashing LEDs:

```
LOCATE CANCEL
```

The Next Step

The *Compaq StorageWorks HSJ80 Array Controller ACS Version 8.5J-2 Installation and Configuration Guide* provides process flow and procedures to perform both the controller and storage configurations.

The preferred method for configuring your subsystem is to use the StorageWorks Command Console (SWCC). SWCC is a Graphical User Interface (GUI) used for the set-up and management of StorageWorks RAID subsystems. As an alternative, direct CLI commands may be inserted to configure the subsystem by way of a maintenance terminal setup through the serial port on the front of the controller.

Appendix A

Subsystem Profiles

This appendix contains device and storageset profiles you can use to create your system profiles. It also contains an enclosure template you can use to help keep track of the location of devices and storagesets in your shelves:

- Storageset Profile on page A-2
- Device Profile on page A-3
- Enclosure Templates for Device Mapping on page A-3

Storageset Profile

Type of Storageset:

_____ Mirrorset	_____ RAIDset	_____ Stripeset	_____ Striped Mirrorset

Storageset Name: ____

Disk Drives: ____

Unit Number: ____

Partitions: ____

Unit #	Unit #	Unit #	Unit #	Unit #	Unit #	Unit #	Unit #
%	%	%	%	%	%	%	%

RAIDset Switches:

Reconstruction Policy	Reduced Membership	Replacement Policy
___Normal (default)	___No (default)	___Best performance (default)
___Fast	___Yes, missing:	___Best fit
		___None

Mirrorset Switches:

Replacement Policy	Copy Policy	Read Source
___Best performance (default)	___Normal (default)	___Least busy (default)
___Best fit	___Fast	___Round robin
___None		___Disk drive:

Initialize Switches:

Chunk size	Save Configuration	Metadata
___Automatic (default)	___No (default)	___Destroy (default)
___64 blocks	___Yes	___Retain
___128 blocks		
___256 blocks		
___Other:		

Unit Switches:

Read Cache	Read-Ahead Cache	Maximum Cache Transfer
___Yes (default)	___Yes (default)	___32 blocks (default)
___No	___ No	___Other:
Write Cache	**Write Protection**	**Availability**
___Yes (default)	___No (default)	___Run (default)
___No	___Yes	___NoRun

Device Profile

Type of Storageset:

_____ Platter Disk Drive	_____ Optical Disk Drive	_____ CD-ROM Drive	_____ Tape Drive

Device Name: ____

Unit Number: ____

Device Switches:

Transportability
___ No (default)
___ Yes

Initialize Switches:

Chunk size	Save Configuration	Metadata
___ Automatic (default)	___ No (default)	___ Destroy (default)
___ 64 blocks	___ Yes	___ Retain
___ 128 blocks		
___ 256 blocks		
___ Other:		

Unit Switches:

Read Cache	Write Cache	Maximum Cache Transfer
___ Yes (default)	___ Yes (default)	___ 32 blocks (default)
___ No	___ No	___ Other:
Availability	**Write Protection**	**Read-Ahead Cache**
___ Run (default)	___ No (default)	___ Yes (default)
___ NoRun	___ Yes	___ No

Enclosure Templates for Device Mapping

The templates on the following pages can be used to map the storage devices and their locations within the cabinets. An example is shown in Chapter 2 (Figure 2-2).

Rack Number/Name__________________

NOTE: A single controller can support up to 42 devices.

Dual-redundant controllers can support up to 36 devices.

Each enclosure can support an optional power supply in addition to the ones shown.

Target 7	6	5	4	3	2	1	0	Port
Power Supply								
Power Supply								
Power Supply								
Power Supply								
Power Supply								
Power Supply								

Rack Number/Name________________________

NOTE: A single controller can support up to 42 devices.

Dual-redundant controllers can support up to 36 devices.

Each enclosure can support an optional power supply in addition to the ones shown.

Target 7	6	5	4	3	2	1	0	Port
Power Supply								
Power Supply								
Power Supply								
Power Supply								
Power Supply								
Power Supply								

Rack Number/Name___________________

NOTE: The 4200-series enclosures can support up to 84 devices. Only 72 are currently supported by the software.

Target

0	1	2	3	4	5	8	9	10	11	12	13	14	15

Port

Glossary

This glossary defines terms relevant to the HSJ80 Array Controller.

ACS	*See* array controller software.
agent	Host-based application layer (applet) of the StorageWorks Command Console (SWCC). *See also* client.
array controller	*See* controller.
array controller software (ACS)	Software contained on a removable PCMCIA program card that provides the operating system for the array controller. *See also* PCMCIA.
autospare	A controller feature that automatically replaces a failed disk drive. You can enable the AUTOSPARE switch for the failedset causing physically replaced disk drives to be automatically placed into the spareset. *Also called* autonewspare.
bad block	A data block that contains a physical defect and possibly corrupt data.
bad block replacement (BBR)	A replacement routine that substitutes defect-free disk blocks for those found to have defects. This process takes place in the controller and is transparent to the host.
BBR	*See* bad block replacement.

block A stream of data stored on a disk or tape medium and transferred and error-checked as a unit. In a disk drive, a block is also called a *sector* (the smallest collection of consecutive bytes addressable on a disk drive). In integrated storage elements, a block contains 512 bytes of data, error codes, flags, and the block address header.

cabinet *See* rack**.**

cache memory High-speed memory used as an intermediary between a data user and a larger amount of storage. Cache memory improves performance by placing the most frequently used data in the highest performance memory.

chunk A block of data written to or read by the host. A chunk is usually equal to a disk sector in capacity.

chunk size The number of data blocks, assigned by a system administrator, written to the primary RAIDset or stripeset member before the remaining data blocks are written to the next RAIDset or stripeset member.

CI Bus The Computer Interconnect Bus uses the MSCP Protocol to connect the nodes together in a clustered subsystem. OpenVMS clusters use a star coupler as the connection point. The CI Bus uses two serial paths, each with a data transfer rate of 70 Mb/s (8.75 MB/s). *See also* MSCP, node, star coupler, and cluster.

CLCP Acronym for Code-Load Code-Patch utility. This utility can be used to download patches to the ACS software.

CLI *See* Command Line Interpreter.

cluster A computer system containing a number of hosts and controllers connected together with the star coupler using the CI Bus and MSCP protocol. Each of the hosts have equal access to any of the storage devices on any of the controllers. *See also* nodes, CI Bus, MSCP, and star coupler.

cold swap A method of device replacement that requires the entire system or subsystem to be powered down before the device can be replaced. *See also* hot pluggable and warm swap.

command line interpreter (CLI) The configuration language that communicates with the controller. Also known as *Command Line Interface*.

configuration file A file that contains a representation of the controller setup parameters and storage subsystem configuration. *See also* non-volatile memory.

container 1. Any entity (a single physical device or a group of physical devices) that is capable of storing data. 2. A virtual, internal controller-structure representing either a single disk or a group of disk drives linked as a storageset. Both stripesets and mirrorsets are examples of storageset containers the controller uses to create units.

controller A hardware device that facilitates communication between a host and one or more storage devices. The HS-series StorageWorks family of controllers are all array controllers.

copying A state in which data to be copied to the mirrorset is inconsistent with other members of the mirrorset. *See also* normalizing.

copying member Any member that joins the mirrorset after the mirrorset is created is regarded as a copying member. Once all the data from the normal member (or members) is copied to a normalizing or copying member, the copying member then becomes a normal member. *See also* normalizing member.

data striping The process of segmenting logically sequential data, such as a single file, so that segments can be written to multiple physical devices (usually disk drives) in a round-robin fashion. This technique is useful if the processor is capable of reading or writing data faster than a single disk can supply or accept the data. While data is being transferred from the first disk, the second disk can locate the next segment.

device *See* node and peripheral device.

Diagnostics and Utilities Protocol Acronym for Diagnostic and Utility Protocol. DUP is a host application software utility program that allows a host operator terminal to connect and communicate with the controller. *See also* virtual terminal.

DILX Acronym for Disk Inline Exerciser. DILX is a diagnostic that tests the data transfer capabilities of disk drives in a way that simulates a high level of user activity.

dirty data *See* unwritten cache data.

driver A hardware device or a program that controls or regulates another device. For example, a device driver is a driver developed for a specific device, such as a printer or a disk drive, that allows a computer to operate with the device.

dual-redundant configuration A storage subsystem configuration consisting of two active controllers operating as a single controller. If one controller fails, the other controller assumes control of the failing controller's devices.

DUP *See* Diagnostics and Utilities Protocol.

ECB *See* external cache battery.

external cache battery (ECB) Acronym for External Cache Battery, which supplies power to the cache module memory to process the data flow in the event of a power failure to the cache module. *See also* external cache battery.

failback The process of restoring data access to the newly-restored controller in a dual-redundant controller configuration. *See also* failover.

failedset A group of disk drives that have been removed from RAIDsets due to a failure or a manual removal. Disk drives in the failedset should be considered defective and should be tested and repaired before being placed back into the spareset. *See also* spareset.

failover The process that takes place when one controller in a dual-redundant configuration assumes the workload of a failed companion controller. Failover continues until the failed controller is repaired or replaced. *See also* failback.

fault management utility A utility that provides a limited interface to the controller's fault-management software. The utility may be used to:

- Display the last-failure and memory-system-failure entries.
- Translate many of the code values contained in event messages.
- Control the display characteristics of significant events and failures.

flush The act of writing dirty data from cache to a storage medium. *See also* unwritten cache data.

host 1. The primary computer to which a storage subsystem is attached. 2. One of the computers in the same cluster that is connected to the storage subsystem.

host adapter A device that connects the host system I/O bus to the storage subsystem. A host adapter performs the lower layers of the CI bus protocol.

hot disks A disk containing multiple hot spots. Hot disks occur when the workload is poorly distributed across storage devices, which prevents optimum subsystem performance. *See also* hot spots.

hot pluggable A subsystem component that can be replaced without removing power from the enclosure.

hot spots A portion of a disk drive frequently accessed by the host. Because the data being accessed is concentrated in one area, rather than spread across an array of disks providing parallel access, I/O performance is significantly reduced. *See als*o hot disks.

INIT Abbreviation of initialize. The sequence a device goes through upon powering up or the sequence a device goes through when commanded to by a switch or software.

JBOD Acronym for "just a bunch of disks." JBOD describes a single or group of single-device logical units that are not configured into any other container type. *See als*o storageset.

LBN Acronym for logical block number.

local connection A connection to the subsystem using the controller's serial maintenance port or a connection from the host terminal. A local connection enables you to connect to one subsystem controller to perform maintenance tasks. See also DUP, CLI.

local terminal A terminal plugged into the EIA-423 maintenance port located on the front bezel of the controller. *See also* maintenance terminal, local connection.

logical unit A physical or virtual device addressable through a target ID number. LUNs use their target's bus connection to communicate on the SCSI bus.

logical unit number (LUN) A value that identifies a specific logical unit belonging to a SCSI target ID number. A number associated with a physical device unit during a task's I/O operations. Each task in the system must establish its own correspondence between logical unit numbers and physical devices.

mass storage control protocol (MSCP) The protocol by which blocks of information are transferred between the host and the HSJ80 controller on the CI Bus. *See also* CI Bus, node.

maintenance terminal An EIA-423-compatible terminal used with the controller. This terminal is used to identify the controller, enable host paths, enter configuration information, and check the controller's status. The maintenance terminal is not required for normal operations. *See also* local terminal.

member A container that is a storage element in a RAID array.

metadata The data written to a disk for the purposes of controller administration. Metadata improves error detection and media defect management for the disk drive. It is also used to support storageset configuration and partitioning. Non transportable disks also contain metadata to indicate they are uniquely configured for StorageWorks environments. Metadata can be thought of as "data about data."

mirrored write-back caching A method of caching data that maintains two copies of the cached data. The copy is available if either cache module fails.

mirroring The act of creating an exact copy or image of data.

mirrorset *See* RAID level 1.

MSCP *See* Mass Storage Control Protocol.

network A data communication, a configuration in which two or more terminals or devices are connected to enable information transfer.

node In data communications, the point at which one or more functional units connect transmission lines. In a VMS Cluster, a node is a connection point that is common to the cluster. A 16-node cluster could be made up of 2 dual-redundant, dual-ported controllers (4 nodes), and 6 host computers each using dual host port adapters (12 nodes). *See also* MSCP and CI Bus.

nominal membership The desired number of mirrorset members when the mirrorset is fully populated with active devices. If a member is removed from a mirrorset, (said to be *reduced*) the actual number of members may fall below the nominal membership.

non-redundant controller configuration 1. A dual-controller configuration in which the pair is not operating in a failover mode. 2. A subsystem configuration that does not include a second controller.

non-transportable A term representing the class of storage devices whose metadata precludes them from being used in anything other than StorageWorks subsystems. *See also* transportable, metadata.

normal member A mirrorset member that, block-for-block, contains the same data as other normal members within the mirrorset. Read requests from the host are always satisfied by normal members.

normalizing A state in which, block-for-block, data written by the host to a mirrorset member is consistent with the data on other normal and normalizing members. The normalizing state exists only after a mirrorset is initialized. Therefore, no customer data is on the mirrorset.

normalizing member A mirrorset member whose contents is the same as all other normal and normalizing members for data that has been written since the mirrorset was created or lost cache data was cleared. A normalizing member is created by a normal member when either all of the normal members fail or all of the normal members are removed from the mirrorset. *See also* copying member.

NVM Acronym for non-volatile memory. A type of memory where the contents survive power loss. Also known as NVMEM. NVM devices are used in the HSJ80 to contain the subsystem configuration file. *See also* configuration file.

OCP *See* operator control panel.

operator control panel (OCP) The control or indicator panel associated with a device. The OCP is usually mounted on the device and is accessible to the operator. The OCP on the HSJ80 contains the reset button, the I/O port quiesce buttons, and the PCMCIA card.

other controller The controller in a dual-redundant pair that is not connected to the controller serving your current CLI session with a local terminal. *See also* this controller.

parity RAID *See* RAIDset.

partition A logical division of a container, represented to the host as a logical unit.

PCMCIA A term representing Personal Computer Memory Card Industry Association. An international association formed to promote a common standard for PC card-based peripherals to be plugged into notebook computers. A PCMCIA card is about the size of a credit card. It is used in the HSJ80 to load the controller software. *Also called* program card.

port In general terms, the port is: 1. A logical channel in a communications system. 2. The hardware and software used to connect a host controller to a communications bus, such as a SCSI bus or serial bus. With respect to the controller, the controller, the port is: 3. The logical route for data in and out of a controller that can contain one or more channels, all of which contain the same type of data. 4. The hardware and software that connects a controller to a SCSI device.

program card The PCMCIA card containing the controller’s operating software.

PTL Acronym for Port-Target-LUN. The controller’s method of locating a device on the controller device bus.

Rack The StorageWorks enclosure with exterior panels and doors into which cable distribution units, enclosures, fans, and cables are mounted to form a storage subsystem. May also be known as *cabinet*.

RAID level 0 A RAID storageset that stripes data across an array of disk drives. A single logical disk spans multiple physical disks, allowing parallel data processing for increased I/O performance. While the performance characteristics of RAID level 0 are excellent, this RAID level is the only one that does not provide redundancy. Raid level 0 storagesets are sometimes referred to as stripesets.

RAID level 0+1 A RAID storageset that stripes data across an array of disks (RAID level 0) and mirrors the striped data (RAID level 1) to provide high I/O performance and high availability. Raid level 0+1 storagesets are also called *striped mirrorsets*.

RAID level 1 A RAID storageset of two or more physical disks that maintains a complete and independent copy of the entire virtual disk's data. This type of storageset is highly reliable and extremely tolerant of device failure. Raid level 1 storagesets are also called *mirrorsets*.

RAID level 3 A RAID storageset that transfers data parallel across the array's disk drives one byte at a time, causing individual blocks of data to be spread over several disks serving as one enormous virtual disk. A separate redundant check disk for the entire array stores parity on a dedicated disk drive within the storageset. *See also* RAID level 5.

RAID level 3/5 A specially-developed RAID storageset that stripes data and parity across three or more members in a disk array. Also known as *RAIDset*. A RAIDset combines the best characteristics of RAID level 3 and RAID level 5. RAIDsets are the best choice for most applications with small to medium I/O requests, unless the application is write intensive. A RAIDset is also called parity RAID.

RAID Level 5 A RAID storageset that, unlike RAID level 3, stores the parity information across all of the disk drives within the storageset. *See also* RAID level 3.

RAIDset *See* RAID level 3/5.

read caching A cache management method used to decrease the subsystem response time to a read request by allowing the controller to satisfy the request from the cache memory rather than from the disk drives.

read-ahead caching A caching technique for improving performance of synchronous sequential reads by prefetching data from disk.

reconstruction The process of regenerating the contents of a failed member's data. The reconstruction process writes the data to a spareset disk and then incorporates the spareset disk into the mirrorset, striped mirrorset, or RAIDset from which the failed member came. *See also* regeneration.

reduced A term that indicates that a mirrorset or RAIDset is missing one member because the member has failed or has been physically removed. *See also* nominal membership.

redundancy The provision of multiple, interchangeable components to perform a single function in order to cope with failures and errors. A RAIDset is considered to be redundant when user data is recorded directly to one member and all of the other members include associated parity information.

regeneration 1. The process of calculating missing data from redundant data. 2. The process of re-creating a portion of the data from a failing or failed drive using the data and parity information from the other members within the storageset. The regeneration of an entire RAIDset member is called *reconstruction*. *See also* reconstruction.

replacement policy The policy specified by a switch with the SET FAILEDSET command indicating whether a failed disk from a mirrorset or RAIDset is to be automatically replaced with a disk from the spareset. The two switch choices are *AUTOSPARE* and *NOAUTOSPARE*.

request rate The rate at which requests are arriving at a servicing entity.

SBB *See* storage building block.

SCSI *See* small computer system interface.

serial data transmission A method in which each bit of information is sent sequentially on a single channel rather than simultaneously as in parallel transmission.

small computer system interface (SCSI) 1. An American National Standards Institute (ANSI) interface standard defining the physical and electrical parameters of a parallel I/O bus used to connect initiators to devices. 2. A processor-independent standard protocol for system-level interfacing between a computer and intelligent devices, including hard drives, floppy disks, CD-ROMs, printers, scanners, and others.

star coupler The physical hub of the CI cluster subsystem cabling. The star coupler is a set of connection panels contained within a cabinet that contains cable connectors and transformers through which the nodes of a cluster connect to one another through the CI bus. *See also* node, cluster, CI Bus, and MSCP.

storage building block (SSB) 1. A modular carrier plus the interface required to mount the carrier into a standard StorageWorks shelf. 2. Any device conforming to shelf mechanical and electrical standards installed in a 3.5-inch or 5.25-inch carrier, whether it is a storage device or power supply.

storage unit The general term that refers to storagesets, single-disk units, and all other storage devices that are installed in a subsystem and accessed by the host. A storage unit can be any entity that is capable of storing data, whether it is a physical device or a group of physical devices. *See also* container.

storageset 1. A group of devices configured with RAID techniques to operate as a single container. 2. Any collection of containers, such as stripesets, mirrorsets, striped mirrorsets, JBODs, and RAIDsets.

stripe The data divided into blocks and written across two or more member disks in an array. *See also* RAID level 0.

stripe size The stripe capacity as determined by *n*–1 times the chunksize, where *n* is the number of RAIDset members.

striped mirrorset *See* RAID level 0+1.

stripeset *See* RAID level 0.

striping The technique used to divide data into segments (chunks). The segments are striped, or distributed, across members of the stripeset. This technique helps to distribute hot spots across the array of physical devices to prevent hot spots and hot disks.

surviving controller The controller in a dual-redundant configuration pair that serves its companion's devices when the companion controller fails.

target 1. A SCSI device that performs an operation requested by an initiator. 2. Designates the target identification (ID) number of the device.

this controller The controller that is serving your current CLI session through a local or remote terminal. *See also* other controller.

transportable A class of storage devices whose metadata allows them to be used in types of storage subsystems other than StorageWorks subsystems. *See also* non-transportable, metadata.

unit A container made accessible to a host. A unit may be created from a single disk drive or tape drive. A unit may also be created from a more complex container such as a RAIDset. The controller supports a maximum of eight units on each target. *See also* target.

unwritten cached data The write-back cached data that has not been written to storage media, even though the host operation processing the data has completed. Sometimes called *unflushed* data.

virtual terminal A software path from an operator terminal on the host to the controller's CLI interface, sometimes called a *host console*. The path can be established via the host port on the controller (using DUP) or via the maintenance port through an intermediary host.

VTDPY Acronym for the Video Terminal Display utility.

warm swap Denotes the replacement of a system or subsystem component with the power still applied. This is usually accomplished through the use of a utility program. *See also* hot pluggable and cold swap.

write hole The period of time in a RAID level 1 or RAID level 5 write operation when an opportunity emerges for undetectable RAIDset data corruption. Write holes occur under conditions, such as power outages, where the writing of multiple members can be abruptly interrupted. A battery backed-up cache design eliminates the write hole, because data is preserved in cache and unsuccessful write operations can be retried.

write-back caching A cache management method that decreases the subsystem response time to write requests by allowing the controller to declare the write operation "complete" as soon as the data reaches its cache memory. The controller performs the slower operation of writing the data to the disk drives at a later time.

write-through cache A cache management method that decreases the subsystem response time to a read. This method allows the controller to satisfy the request from the cache memory rather than from the disk drives.

When the host requests a write operation, the controller writes data directly to the storage device. This technique allows the controller to complete some read requests from the cache, greatly improving the time to retrieve data. The operation is complete only after the data to be written is received by the target storage device. This cache management method may update, invalidate, or delete data from the cache memory accordingly, to ensure that the cache contains the most current data.

Index

G

H

I

J

L

M

N

O

P

R

S

T

U

W