IBM System Storage

IBM

Multipath Subsystem Device Driver User's Guide

Multipath Subsystem Device Driver User's Guide

Contents

Figures

Tables

Introduction

This guide provides information that pertains to Chapter 1, Chapter 3, and Chapter 4 of the *IBM® System Storage™ Multipath Subsystem Device Driver User's Guide*, version 1.8.0. The primary changes are to Chapter 1 and Chapter 3, to add support for DS3950 and DS5020.

Changes in these chapters are indicated by a vertical bar (|).

Refer to the *IBM System Storage Multipath Subsystem Device Driver 1.7.x User's Guide* for general information and if you are using the AIX, HP-UX, SUN Solaris, Linux, Windows, or Netware operating systems, and you're using one of the following versions of SDD:
- SDD Version 1 Release 8 Modification 0 Level *x* for HP-UX
- SDD Version 1 Release 7 Modification 2 Level *x* for AIX®
- SDD Version 1 Release 6 Modification 4 Level *x* for Solaris and Windows®
- SDD Version 1 Release 6 Modification 3 Level *x* for Linux®
- SDD Version 1 Release 6 Modification 0 Level *x* for Netware
- SDD Device Specific Module Version 2 Release 4 Modification 0 Level *x*
- SDD Path Control Module Version 2 Release 5 Modification 1 Level *x*

Who should use this book

This errata is intended for those who are using the SDDPCM loadable path control module on an AIX host system, or for those who are using SDD on an HP-UX host system.

Last update

This guide was last updated: December 15, 2009.

Chapter 1. Overview of the SDD

The IBM System Storage Multipath Subsystem Device Driver (SDD) supports a storage-redundant configuration environment for a host system that is attached to the following devices:

- IBM TotalStorage® Enterprise Storage Server® (ESS)
- IBM System Storage DS8000®
- IBM System Storage DS6000™
- IBM System Storage SAN Volume Controller
- IBM System Storage DS5000
- IBM System Storage DS4000®
- IBM BladeCenter S SAS RAID Controller Module
- IBM System Storage DS3950

This guide uses the following terminology:

- The phrase *supported storage devices* refers to the following types of devices:
  - SAS RAID Controller Module for BladeCenter S
  - DS3950, DS4100 (AIX only), DS4200, DS4300, DS4500, DS4700, DS4800, DS5020, DS5100, DS5300, DS6000, and DS8000
  - ESS
  - SAN Volume Controller
- The phrase *disk storage system* refers to ESS, DS8000, or DS6000 devices.
- The phrase *virtualization product* refers to the SAN Volume Controller. Table 1 indicates the products that different SDD platforms support.
- The phrase *DS4000* refers to DS4100 (AIX only), DS4200, DS4300, DS4500, DS4700, and DS4800 devices.
- The phrase *DS5000* refers to DS5020, DS5100 and DS5300 devices.
- The phrase *BladeCenter* refers to IBM BladeCenter S SAS RAID Controller Module devices.

Table 1. SDD platforms that are supported by supported storage devices

	Supported storage device							
Platform	ESS	DS8000	DS6000	DS5000	DS4000	DS3950	BladeCenter	SAN Volume Controller
AIX SDD	✔	✔	✔					✔
AIX SDDPCM	✔	✔	✔	✔	✔	✔		✔
HP	✔	✔	✔	✔	✔	✔		✔
Linux	✔	✔	✔					✔
Novell	✔	✔	✔					
SUN	✔	✔	✔					✔
Windows NT® SDD	✔							✔

Table 1. SDD platforms that are supported by supported storage devices (continued)

	Supported storage device							
Platform	ESS	DS8000	DS6000	DS5000	DS4000	DS3950	BladeCenter	SAN Volume Controller
Windows 2000 and Windows 2003 SDD	✔	✔	✔					✔
Windows 2003 and Windows 2008 SDDDSM	✔	✔	✔				✔	✔

The SDD supports a storage-redundant configuration environment for a host system that is attached to storage devices. It provides enhanced data availability, dynamic input/output (I/O) load-balancing across multiple paths, and automatic path failover protection.

This guide provides step-by-step procedures on how to install, configure, and use SDD features on the following host systems:

- IBM AIX (SDD and SDDPCM)
- HP-UX
- Supported Linux distributions, levels, and architectures. For up to date information about specific kernel levels supported in this release, see the Readme file on the CD-ROM or visit the SDD Web site:

  www.ibm.com/servers/storage/support/software/sdd
- Novell Netware (disk storage systems only)
- Sun Solaris
- Microsoft® Windows NT, Windows 2000, or Windows 2003 SDD
- Microsoft Windows Server 2003 or Windows Server 2008 SDD
- Microsoft Windows Server 2003 or Windows Server 2008 SDDDSM

The SDD architecture

The SDD is a software solution to support the multipath configuration environments in supported storage devices. It resides in a host system with the native disk device driver and provides the following functions:

- Enhanced data availability
- Dynamic input/output (I/O) load-balancing across multiple paths
- Automatic path failover protection
- Concurrent download of licensed machine code

Table 2 on page 3 shows the position of the SDD in the protocol stack. I/O operations that are sent to the SDD proceed to the host disk driver after path selection. When an active path experiences a failure (such as a cable or controller failure), the SDD dynamically switches to another path.

Table 2. SDD in the protocol stack

Raw disk I/O → Logical Volume Manager (LVM) I/O → LVM device driver → Subsystem Device Driver → AIX SCSI/FCP disk driver → SCSI/FCP adapter driver

S008996Q

Disk I/O / File system → Subsystem Device Driver → Linux disk SCSI driver → SCSI adapter driver

S009318

Raw disk I/O → Logical Volume Manager I/0 → LVM device driver → Subsystem Device Driver → HP disk driver → SCSI adapter driver

S008998Q

Raw disk I/O → Logical Volume Manager I/O → LVM device driver → Subsystem Device Driver → Sun Solaris disk driver → SCSI adapter driver

S008999Q

Table 2. SDD in the protocol stack (continued)

System disk I/O	System disk I/O
Subsystem Device Driver → Windows NT disk driver → adapter driver → (disk) S008997Q	Windows 2000 disk driver → Subsystem Device Driver → adapter driver → (disk)
System disk I/O ↓ Windows Server 2003 disk driver → Subsystem Device Driver → adapter driver → (disk)	This space intentionally blank
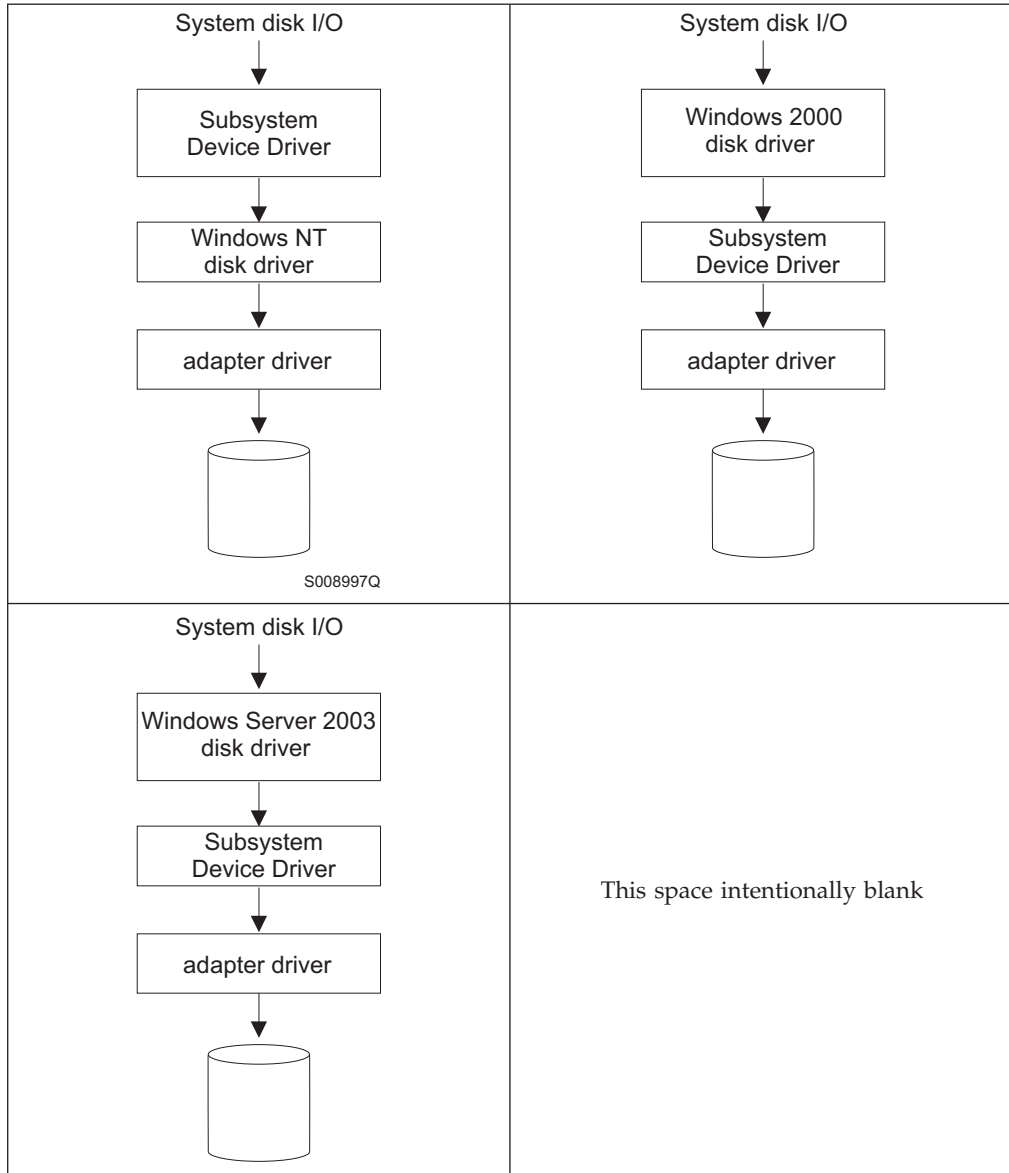
Each SDD vpath device represents a unique physical device on the storage server.
Each physical device is presented to the operating system as an operating system
disk device. There can be up to 32 operating system disk devices that represent up
to 32 different paths to the same physical device.

The SDD vpath devices behave almost like native operating system disk devices.
You can use most disk device operations of operating systems on the SDD vpath
devices, including commands such as **open**, **close**, **dd**, or **fsck**.

Enhanced data availability

Figure 1 on page 5 shows a host system that is attached through small computer
system interface (SCSI) or fibre-channel adapters to a disk storage system that has
internal component redundancy and multipath configuration. The SDD, residing in
the host system, uses this multipath configuration to enhance data availability. That
is, when there is a path failure, the SDD reroutes I/O operations from the failing

path to an alternate operational path. This capability prevents a single failing bus adapter on the host system, SCSI or fibre-channel cable, or host-interface adapter on the disk storage system from disrupting data access.
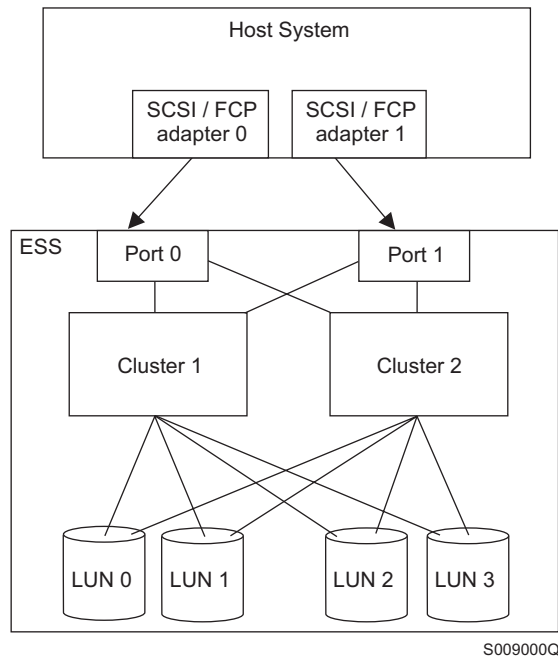
Figure 1. Multipath connections between a host system and the disk storage in a disk storage system

Figure 2 shows a host system that is attached through fibre-channel adapters to a SAN Volume Controller that has internal components for redundancy and multipath configuration. The SDD, residing in the host system, uses this multipath configuration to enhance data availability. That is, when there is a path failure, the SDD reroutes I/O operations from the failing path to an alternate operational path. This capability prevents a single failing bus adapter on the host system, fibre-channel cable, or host-interface adapter on the SAN Volume Controller from disrupting data access.
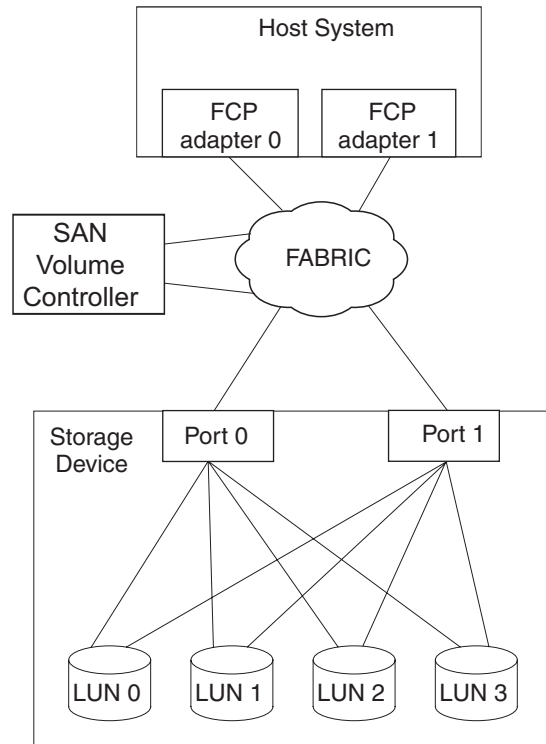
Figure 2. Multipath connections between a host system and the disk storage with the SAN Volume Controller

Note: SAN Volume Controller does not support parallel SCSI attachment.

Dynamic I/O load balancing

By distributing the I/O workload over multiple active paths, the SDD provides dynamic load balancing and eliminates dataflow bottlenecks. In the event of failure in one data path, the SDD automatically switches the affected I/O operations to another active data path, ensuring path-failover protection.

Automatic path-failover protection

The SDD failover protection feature minimizes any disruptions in I/O operations and recovers I/O operations from a failing data path. The SDD provides path-failover protection using the following process:

- Detects a path failure.
- Notifies the host system of the path failure.
- Selects and uses an alternate data path.

The SDD dynamically selects an alternate I/O path when it detects a software or hardware problem.

Some operating system drivers report each detected error in the system error log. With the SDD automatic path-failover feature, some reported errors are actually recovered from an alternative path.

Concurrent download of licensed machine code for disk storage systems

With the SDD multipath mode (configured with at least two paths per multipath device), you can concurrently download and install the licensed machine code (LMC) while applications continue to run. For certain disk storage system LMC, the disk storage system I/O bay or tower are quiesced and resumed. Its adapters might not respond for the duration of the service action, which could be 30 minutes or more.

Note: The SDD does not support single-path mode during the concurrent download of LMC. Also, the SDD does not support single-path mode during any disk storage system concurrent maintenance that impacts the path attachment, such as a disk storage system host-bay-adapter replacement.

Attention: Do not shut down the host or reconfigure the SDD during the concurrent download of licensed machine code or you might lose your initial SDD configuration.

For information about performing the concurrent download of LMC for ESS, see the microcode installation instructions for your specific type and model.

Concurrent download of licensed machine code for virtualization products

With the SDD multipath mode (configured with at least two paths per multipath device), you can concurrently download and install the licensed machine code while your applications continue to run.

At least one path must be configured through each node of a virtualization product group. That is, if only two paths exist, they must go to separate nodes for each I/O group. However, at least two paths to each node are recommended.

During the code upgrade, each node of an I/O group is upgraded sequentially. The node that is being upgraded is temporarily unavailable, and all I/O operations to that node fail. However, failed I/O operations are directed to the other node of the I/O group, and applications do not see any I/O failures.

Attention: Do not shut down the host or reconfigure the SDD during the concurrent download of licensed machine code or you might lose your initial SDD configuration.

For information about performing concurrent download of LMC for virtualization products, see the Configuration Guide for your specific type and model.

Concurrent download of licensed machine code for DS3950, DS4000, and DS5000

If you are using the SDD multipath mode, you can concurrently download and install the licensed machine code while your applications continue to run, as long as you configure redundant paths to each storage controller port in addition to the multiple host adapter ports.

Because switching a device to another controller is a time-consuming recovery action and affects I/O performance, you can use this redundancy to avoid an unnecessary controller failover if a path fails. Therefore, configure a minimum of four paths for each LUN with two host adapter ports and two storage controller ports where each host adapter port has redundancy to each storage controller port and vice versa.

Attention: Do not shut down the host or reconfigure the SDD during the concurrent download of licensed machine code or you might lose your initial SDD configuration.

Active/Passive dual array controller path-selection algorithm for DS3950, DS4000, and DS5000 products

The DS4000 and DS5000 products are dual array controller disk subsystems. Each LUN is assigned to one controller, which is considered the owner, or the active controller, of a particular LUN. The other controller is considered as an alternate, or passive, controller. Thus, the SDD distinguishes the following paths to the DS4000 and DS5000 product LUN:

- Paths on the ownership (active) controller
- Paths on the alternate (passive) controller

With this type of active/passive dual-controller subsystem device, I/O can be sent only to the ownership controller. When the SDD selects paths for I/O, it selects paths that are connected only to the ownership controller. If there is no path on the ownership controller that can be used, SDD changes the LUN controller ownership to an alternate controller, switches the paths that were passive to active, and then selects these active paths for I/O.

Preferred node path-selection algorithm for the virtualization products

Virtualization products are two-controller disk subsystems. The SDD distinguishes the paths to a virtualization product LUN:

- Paths on the preferred controller
- Paths on the alternate controller

When the SDD selects paths for I/O, preference is always given to a path on the preferred controller. Therefore, in the selection algorithm, an initial attempt is made to select a path on the preferred controller. Only if no path can be used on the preferred controller is a path selected on the alternate controller. This means that the SDD automatically fails back to the preferred controller any time a path on the preferred controller becomes available during either manual or automatic recovery. Paths on the alternate controller are selected at random. If an error occurs and a path retry is required, retry paths are first selected on the preferred controller. If all retries fail on the preferred controller's paths, paths on the alternate controller are selected for retry. The following is the path selection algorithm for the SDD:

1. With all paths available, I/O is only routed to paths on the preferred controller.
2. If no path on the preferred controller is available, I/O failover processing to the alternate controller occurs.
3. When failover processing to the alternate controller has occurred, automatic I/O failback processing to the preferred controller occurs if a path on the preferred controller is available.

The following output of the **datapath query device** command shows that the preferred paths are being selected and shows the format of the virtualization product serial number.

```
DEV#:   0  DEVICE NAME: Disk0 Part0    TYPE: 2145        POLICY: OPTIMIZED
SERIAL: 6005676801800005F800000000000004
============================================================================
Path#            Adapter/Hard Disk    State    Mode     Select      Errors
    0   Scsi Port4 Bus0/Disk0 Part0    OPEN     NORMAL   501876           0
    1   Scsi Port4 Bus0/Disk0 Part0    OPEN     NORMAL   501238           0
    2   Scsi Port4 Bus0/Disk0 Part0    OPEN     NORMAL        0           0
    3   Scsi Port4 Bus0/Disk0 Part0    OPEN     NORMAL        0           0
    4   Scsi Port5 Bus0/Disk0 Part0    OPEN     NORMAL   499575           0
    5   Scsi Port5 Bus0/Disk0 Part0    OPEN     NORMAL   500698           0
    6   Scsi Port5 Bus0/Disk0 Part0    OPEN     NORMAL        0           0
    7   Scsi Port5 Bus0/Disk0 Part0    OPEN     NORMAL        0           0
```

Chapter 3. Using SDDPCM on an AIX host system

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

This chapter provides a general view of the SDDPCM path control module, including where it resides on the I/O stack in the operating system and the features and functions that it supports. This chapter also provides procedures for the following management tasks:

- Install SDDPCM
- Configure SDDPCM MPIO-capable devices
- Uninstall SDDPCM
- Migrate disk storage MPIO-capable devices from the AIX default PCM to SDDPCM
- Migrate disk storage MPIO-capable devices from SDDPCM to the AIX default PCM or SDD
- Installing SDDPCM from AIX NIM SPOT to client SAN boot disk or internal boot disk

Figure 3 on page 12 shows the position of SDDPCM in the protocol stack. I/O operations are sent to the AIX disk driver. The SDDPCM path selection routine is invoked to select an appropriate path for each I/O operation.
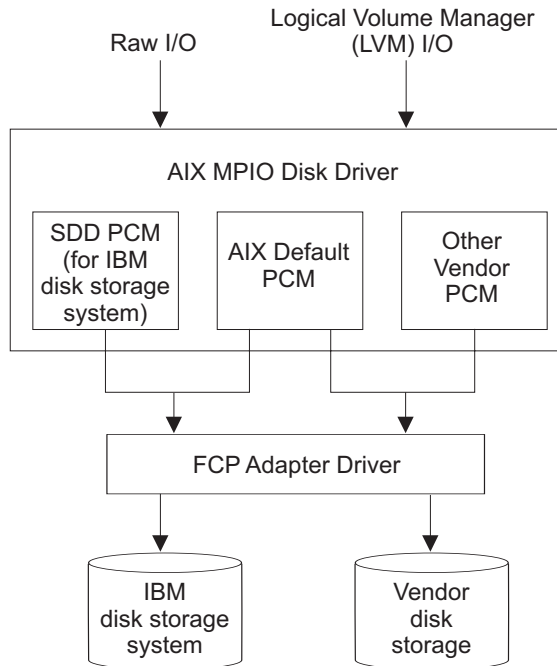
Figure 3. SDDPCM in the protocol stack

For detailed information about MPIO support on AIX 5.2 TL07 (or later), AIX 5.3 TL03 (or later), or AIX 6.1, visit the following Web site:

http://publib16.boulder.ibm.com/pseries/en_US/aixbman/baseadmn/manage_MPIO.htm

AIX MPIO-capable device drivers automatically discover, configure and make available every storage device path. SDDPCM manages the paths to provide:
- High availability and load balancing of storage I/O
- Automatic path-failover protection
- Concurrent download of supported storage devices licensed machine code
- Prevention of a single-point-failure

For updated and additional information that is not included in this chapter, see the Readme file on the CD-ROM or visit the SDD Web site:

www.ibm.com/servers/storage/support/software/sdd

SDD and SDDPCM are exclusive software packages on a server. You cannot install both software packages on a server for supported storage devices. When supported storage devices are configured as non-MPIO-capable devices (that is, multiple logical device instances are created for a physical LUN), you should install SDD to get multipath support.

You must install SDDPCM in order to configure supported storage devices into MPIO-capable-devices (where only one logical device instance is created for a physical LUN). Before you install SDDPCM, make sure that you meet all the required hardware and software requirements. See "Verifying the hardware and software requirements" and "Preparing for SDDPCM installation for supported storage devices."

Note: SDDPCM does not support SCSI storage devices.

With SDD 1.6.0.0 (or later), SDDPCM and SDD cannot coexist on a AIX server. If a server connects to any supported storage devices, all devices must be configured either as non-MPIO-capable devices or as MPIO-capable devices.

Supported SDDPCM features

The following SDDPCM features are supported in this release:

- 32- and 64-bit kernels
- Four types of reserve policies:
  - No reserve policy
  - Exclusive host access single path policy
  - Persistent reserve exclusive host policy
  - Persistent reserve shared host access policy
- Four path-selection algorithms:
  - Failover
  - Round robin
  - Load balancing
  - Load balancing port

    This algorithm is supported only with DS8000, DS6000, ESS, and SAN Volume Controller device models.
- Automatic failed paths reclamation by healthchecker
- Failback error-recovery algorithm
- Fibre-channel dynamic device tracking
- Support for all DS4000, DS5000, DS3950, DS6000, DS8000, ESS FCP, and SAN Volume Controller devices
- Support SAN boot device on MPIO supported storage devices

    Notes:

    1. SDDPCM supports ESS devices as SAN boot devices, beginning with AIX52 TL06 and AIX53 TL02.
    2. SDDPCM supports DS8000, DS6000, and SAN Volume Controller devices as SAN boot devices, beginning with AIX52 TL07 and AIX53 TL03.
    3. SDDPCM supports DS4000, DS5000, and DS3950 devices as SAN boot devices, beginning with AIX53 TL8 and AIX61 TL2
- Support for external supported MPIO storage devices as the primary or secondary dump device
- Support storage devices multipath devices as system paging space
- Support SDDPCM server daemon enhanced path health check function
- Support a maximum of 1200 LUNs
- Dynamically adding paths or adapters
- Dynamically removing paths or adapters
- Dynamically changing the device path selection algorithm
- Dynamically changing the device hc_interval
- Dynamically enabling or disabling the device health check function
- Dynamically changing the device hc_mode
- Web-based System Manager (WebSM) for MPIO supported storage devices (See www-1.ibm.com/servers/aix/wsm/ for more information about WebSM.)
- Reserve last path of a device in OPEN mode

- Support the essutil Product Engineering tool in the SDDPCM **pcmpath** command line program
- Support HACMP™ with Enhanced Concurrent Mode volume group in concurrent resource groups and nonconcurrent resource groups
- Support GPFS™ in AIX 5.2 TL06 (or later), 5.3 TL02 (or later), and AIX 6.1.
- Support Virtual I/O server with AIX 5.3 or later and AIX 6.1.

  Note: This support does not include DS4000, DS5000, and DS3950 storage devices.

Unsupported SDDPCM features

The following SDDPCM features are not currently supported.
- HACMP with persistent reservation policies
- Multiple hosts sharing supported storage MPIO devices without supported clustering software installed
- HACMP with non-Enhanced Concurrent Mode volume group
- Mirror Write Consistency with volume group in a concurrent resource group
- Virtual I/O server is not supported with DS4000, DS5000, and DS3950 storage subsystems

Verifying the hardware and software requirements

You must install the following hardware and software components to ensure that SDDPCM installs and operates successfully.

Hardware

The following hardware components are needed:
- Supported storage devices (FCP devices only)
- One or more switches, if the supported storage devices is not direct-attached
- Host system
- Fibre-channel adapters and cables

Software

The following software components are needed:
- AIX 5.2 TL10 (or later), AIX 5.3 TL08 (or later), or AIX 6.1 TL02 operating system, with all the latest PTFs. Refer to the Readme file of the SDDPCM level that you plan to install for the required AIX TL for that level.
- If your attached storage is SAN Volume Controller version 4.2.1.6 or above and you require SAN Volume Controller APAR IC55826, you must install SDDPCM 2.2.0.3 or above with the required AIX TL and APAR. Refer to the Readme file of the SDDPCM level that you plan to install for the required AIX TL and APARs.
- Fibre-channel device drivers
- One of the following installation packages:
  - devices.sddpcm.52.rte (version 2.5.1.0 or later)
  - devices.sddpcm.53.rte (version 2.5.1.0 or later)
  - devices.sddpcm.61.rte (version 2.5.1.0 or later)
- Supported storage devices: devices.fcp.disk.ibm.mpio.rte (version of 1.0.0.20 or later) host attachment package for SDDPCM

Unsupported environments

SDDPCM does not support:
- ESS SCSI devices
- A host system with both a SCSI and fibre-channel connection to a shared ESS logical unit number (LUN)
- Single-path mode during code distribution and activation of LMC nor during any supported storage devices concurrent maintenance that impacts the path attachment, such as a supported storage device host-bay-adapter replacement.
- AIX5.2 is not supported with DS4000, DS5000, and DS3950 storage subsystems.

Host system requirements

To successfully install SDDPCM for supported storage devices, you must have the required AIX TL installed on your host system along with the AIX required fixes, APARs, and microcode updates that are identified on the following Web site:

www-1.ibm.com/servers/storage/support/

Starting with SDDPCM 2.2.0.0, SDDPCM requires AIX 5.2 TL10 (or later), AIX 5.3 TL06 (or later), or AIX 6.1, with all the latest PTFs. Always refer to the Readme file for the required AIX TL for the SDDPCM level that you plan to install.

Supported storage device requirements

To successfully install SDDPCM, ensure that the devices.fcp.disk.ibm.mpio.rte package is installed on the server. Refer to "Installing and upgrading the AIX SDDPCM host attachment" for more information.

Fibre requirements

You must check for and download the latest fibre-channel device driver APARs, maintenance-level fixes, and microcode updates from the following Web site:

www-1.ibm.com/servers/eserver/support/

If your host has only one fibre-channel adapter, it requires you to connect through a switch to multiple supported storage device ports. You should have at least two fibre-channel adapters to prevent data loss due to adapter hardware failure or software failure.

For information about the fibre-channel adapters that can be used on your AIX host system, go to the following Web site:

www.ibm.com/servers/storage/support

To use the SDDPCM fibre-channel support, ensure that your host system meets the following requirements:
- The AIX host system is an IBM RS/6000® or IBM System p® with AIX 5.2 TL10 (or later), AIX 5.3 TL06 (or later), or AIX 6.1.
- The AIX host system has the fibre-channel device drivers installed along with all latest APARs.
- The host system can be a single processor or a multiprocessor system, such as SMP.
- A fiber-optic cable connects each fibre-channel adapter to a supported storage system port.

- If you need the SDDPCM I/O load-balancing and failover features, ensure that a minimum of two paths to a device are attached.

Preparing for SDDPCM installation

The SDDPCM installation package installs a number of major files on your AIX system. The major files that are part of the SDDPCM installation package are:

File name	Description
sddpcmrtl	A dynamically loaded module added to the device configuration methods to extend the supported storage devices device configuration methods to facilitate the configuration operations of the PCM KE
sddpcmke	A dynamically-loaded module added to the AIX 5L™ kernel that provides path management functions for supported storage devices
sdduserke	A dynamically-loaded module added to the AIX 5L kernel that provides the API to sddpcmke
pcmpath	SDDPCM command line tool
pcmsrv	Daemon for enhanced path healthcheck, and First Time Data Capture
sample_pcmsrv.conf	The sample SDDPCM server daemon configuration file
fcppcmmap	Collects supported storage devices fibre-channel device information through SCSI commands
pcmquerypr	SDDPCM persistent reserve command tool
pcmgenprkey	SDDPCM persistent reserve command tool to generate persistent reserve key
relbootrsv	Release SCSI-2 reserve on boot devices or on active nonboot volume groups
	Note: Beginning with SDDPCM version 2.5.1.0, the relbootrsv program also can be used to release SCSI-2 reserve on an MPIO hdisk.
relSDDPCMbootrsv	Automatically executed upon reboot to release leftover SCSI-2 reserve on SDDPCM boot device, beginning with SDDPCM version 2.2.0.1.
sddpcmgetdata	Script to collect SDDPCM information, trace log files, and system error logs into an sddpcmdata_*host*_date_time.tar file for problem determination
sddpcm_get_config	Display configuration information for DS4000, DS5000, and DS3950 storage and MPIO-based devices.
lspcmcfg	Display SDDPCM MPIO device/path configuration/state information.

Preparing for SDDPCM installation for supported storage devices

Before you install SDDPCM, you must complete the following tasks:

- Connect the supported storage device to your host system and the required fibre-channel adapters that are attached.
- Configure the supported storage device for single-port or multiple-port access for each LUN. To use the load-balancing and failover features, SDDPCM requires a minimum of two independent paths that share the same logical unit. **With a single path, failover protection is not available.**

For more information about how to configure your supported storage device, see the Planning Guide for your product.

Before you install SDDPCM, you must:

- Determine that you have the correct installation package
- Remove the SDD package, if it is installed.
- Remove the ibm2105.rte (version 32.6.100.*x*) and/or devices. fcp.disk.ibm.rte (version 1.0.0.*x*), if they are installed.
- Install the AIX fibre-channel device drivers, if necessary.
- Verify and upgrade the fibre channel adapter firmware level
- Install the SDDPCM Host Attachment: devices.fcp.disk.ibm.mpio.rte (version 1.0.0.15)

Determining the correct installation package

The package name of SDDPCM for different AIX operating system levels is the following:

- devices.sddpcm.52.rte for AIX 5.2
- devices.sddpcm.53.rte for AIX 5.3
- devices.sddpcm.61.rte for AIX 6.1

Determining if the SDD package is installed

To determine if the SDD is installed:

1. Use the **lslpp -l *ibmSdd*** and **lslpp -l devices.sdd*** commands to determine if any SDD package is installed on the system.
2. If SDD is installed for supported storage device configuration, you must unconfigure and remove all SDD vpath devices, and then uninstall the SDD package. See "Removing SDD from an AIX host system."

Determining if the ibm2105.rte package is installed

To determine if the ibm2105.rte package is installed:

1. Use the **lslpp -l *ibm2105*** command to determine if any ibm2105.rte with VRMF 32.6.100.*xx* is installed.
2. If ibm2105.rte is installed, you must:
   a. Unconfigure and remove all supported storage devices.
   b. Use **smitty** to uninstall the ibm2105.rte package. If ibm2105.rte is installed, you must remove all supported storage devices and uninstall the ibm2105.rte package. See "Removing SDD from an AIX host system."

Determining if the devices.fcp.disk.ibm.rte package is installed

To determine if the devices.fcp.disk.ibm.rte package is installed:

1. Use the **lslpp -l devices.fcp.disk.ibm*** command to determine if any devices.fcp.disk.ibm.rte with VRMF 1.0.0.*x* is installed.
2. If devices.fcp.disk.ibm.rte is installed, you must:
   a. Unconfigure and remove all supported storage devices.
   b. Use **smitty** to uninstall the devices.fcp.disk.ibm.rte package. If devices.fcp.disk.ibm.rte is installed, you must remove all supported storage devices and uninstall the devices.fcp.disk.ibm.rte package. See "Removing SDD from an AIX host system."

Installing the AIX fibre-channel device drivers

You must check for the latest information on fibre-channel device driver APARs, maintenance-level fixes, and microcode updates at the following Web site:

www-1.ibm.com/servers/storage/support/

Perform the following steps to install the AIX fibre-channel device drivers from the AIX compact disk:

1. Log in as the root user.
2. Load the compact disc into the CD-ROM drive.
3. From your desktop window, enter **smitty install_update** and press **Enter** to go directly to the installation panels. The **Install and Update Software** menu is displayed.
4. Highlight **Install Software** and press **Enter**.
5. Press **F4** to display the **INPUT Device/Directory for Software** panel.
6. Select the compact disc drive that you are using for the installation; for example, /dev/cd0, and press **Enter**.
7. Press **Enter** again. The **Install Software** panel is displayed.
8. Highlight **Software to Install** and press **F4**. The **Software to Install** panel is displayed.
9. The fibre-channel device drivers include the following installation packages:

   devices.pci.df1080f9
   > The adapter device driver for RS/6000 or IBM System p with feature code 6239.

   devices.pci.df1000f9
   > The adapter device driver for RS/6000 or IBM System p with feature code 6228.

   devices.pci.df1000f7
   > The adapter device driver for RS/6000 or IBM System p with feature code 6227.

   devices.common.IBM.fc
   > The FCP protocol driver.

   devices.fcp.disk
   > The FCP disk driver.

   Select each one by highlighting it and pressing **F7**.
10. Press **Enter**. The **Install and Update from LATEST Available Software** panel is displayed with the name of the software you selected to install.
11. Check the default option settings to ensure that they are what you need.
12. Press **Enter** to install. SMIT responds with the following message:

```
+----------------------------------------------------------------------+
|      ARE YOU SURE??                                                   |
|      Continuing may delete information you may want to keep. 413      |
|      This is your last chance to stop before continuing. 415         |
+----------------------------------------------------------------------+
```

13. Press **Enter** to continue. The installation process can take several minutes to complete.

14. When the installation is complete, press **F10** to exit from SMIT. Remove the compact disc.

15. Check to see if the correct APARs are installed by entering the following command:

    instfix   -iv | grep   IY*nnnnn*

    where *nnnnn* represents the APAR numbers.

    If the APARs are listed, that means that they are installed. If they are installed, go to "Configuring supported storage MPIO-capable devices." Otherwise, go to step 3.

16. Repeat steps 1 through 14 to install the APARs.

Verifying and upgrading the fibre channel adapter firmware level

Use the following procedures to verify and upgrade your current fibre channel adapter firmware level.

Verifying the adapter firmware level: You must verify that your current adapter firmware is at the latest level. If your current adapter firmware is not at the latest level, you must upgrade to a new adapter firmware (microcode). To check the current supported firmware level for fibre-channel adapters, go to the following Web site:

http://www14.software.ibm.com/webapp/set2/firmware/gjsn

Perform the following steps to verify the firmware level that is currently installed:

1. Enter the **lscfg -vl fcsN** command. The vital product data for the adapter is displayed.

2. Look at the **ZB** field. The **ZB** field should look similar to:

   ```
   (ZB)............S2F3.30X1
   ```

   To verify the firmware level, ignore the second character in the **ZB** field. In the example, the firmware level is sf330X1.

3. If the adapter firmware level is at the latest level, there is no need to upgrade; otherwise, the firmware level must be upgraded. To upgrade the firmware level, go to "Upgrading the adapter firmware level."

Upgrading the adapter firmware level: Upgrading the firmware level consists of downloading the firmware (microcode) from your AIX host system to the adapter. Before you upgrade the firmware, ensure that you have configured any fibre-channel-attached devices (see "Configuring supported storage MPIO-capable devices"). After the devices are configured, download the firmware from the AIX host system to the FCP adapter by performing the following steps:

1. Verify that the correct level of firmware is installed on your AIX host system. Go to the /etc/microcode directory and locate the file called df1000f7.*XXXXXX* for feature code 6227 and df1000f9.*XXXXXX* for feature code 6228, where

XXXXXX is the level of the microcode. This file was copied into the /etc/microcode directory during the installation of the fibre-channel device drivers.

2. From the AIX command prompt, enter **diag** and press **Enter**.
3. Highlight the **Task Selection** option.
4. Highlight the **Download Microcode** option.
5. Press **Enter** to select all the fibre-channel adapters to which you want to download firmware. Press **F7**. The Download panel is displayed with one of the selected adapters highlighted. Press **Enter** to continue.
6. Highlight /etc/microcode and press **Enter**.
7. Follow the instructions that are displayed to download the firmware, one adapter at a time.

Fibre-channel HBA attributes for DS4000, DS5000, and DS3950 storage devices: You must set the FC HBA **dyntrk** attribute to **yes** instead of the default settting, which is **no**.

Installing and upgrading the AIX SDDPCM host attachment

The host attachment for SDDPCM adds supported storage devices information to allow AIX to properly configure these devices as MPIO-capable hdisks. The supported storage devices information allows AIX to perform the following actions:

- Identify the hdisks as 2105, 2145, 1750, 2107, DS4100, DS4200, DS4300, DS4500, DS4700, DS4800, DS5000, DS5020, or DS3950 hdisks.
- Set default hdisk attributes, such as queue_depth and timeout values.
- Indicate to the AIX device driver configure method to set 2105, 2145, 1750, 2107, DS3950, DS4100, DS4200, DS4300, DS4500, DS4700, DS4800, DS5000, or DS5020 hdisks as MPIO devices.

The AIX SDDPCM Host Attachment package name is devices.fcp.disk.ibm.mpio.rte. See the SDDPCM Readme file for the latest Host Attachment level for the SDDPCM version that you are installing.

For AIX SDDPCM host attachment installation and upgrade instruction from the CD, see the *Host System Attachment Guide* for your storage device. Do not reboot the system before the SDDPCM package is installed on the system.

For AIX SDDPCM host attachment installation and upgrade from the SDD download Web site, use following procedure:

1. Download code from the Web site that is appropriate for your system.
2. Move the downloaded file to the /usr/sys/inst.images directory.
3. Change the directory to the /usr/sys/inst.images directory.
4. Untar the downloaded file. For example:

    ```
    tar -xvf devices.fcp.disk.ibm.mpio.rte.tar
    ```

5. Update the .toc file. For example:

    ```
    pwd
    rm -i .toc
    inutoc .
    grep -i devices.fcp.disk.ibm .toc
    ```

    This command should reflect the newer SDDPCM host attachment version that will be uploaded.

6. From your desktop window, type **smitty install_update** and press **Enter** to go directly to the installation panels. The Install and Update Software menu is displayed.

7. Highlight **Install Software** and press **Enter**.

8. Type **.** to indicate the current directory and press **Enter**.

9. Highlight **Software to Install** and press **F4**. The Software to Install panel is displayed.

10. Select the devices.fcp.disk.ibm.mpio.rte package.

11. Press **Enter**. The Install and Update from the LATEST Available Software panel is displayed with the name of the software that you selected to install.

12. Check the default option settings to ensure that they are what you need.

13. Press **Enter** to install. SMIT responds with the following message:

```
+------------------------------------------------------------------+
|ARE YOU SURE??                                                    |
|Continuing may delete information you may want to keep.           |
|This is your last chance to stop before continuing.              |
+------------------------------------------------------------------+
```

14. Press **Enter** to continue. The installation process may take a few minutes to complete.

15. When the installation or upgrade is complete, press **F10** to exit from SMIT.

16. If this is a host attachment installation, do not reboot the system until you have the SDDPCM package installed. If this is a host attachment upgrade, and the SDDPCM package is already installed on the system, then reboot the system to complete the host attachment upgrade.

    Attention: Do not reboot the system if you only have the SDDPCM host attachment package installed.

Installing and upgrading SDDPCM

Use these topics to install and upgrade SDDPCM.

Installing SDDPCM from CD-ROM

SDDPCM is released as an AIX installation image. The SDDPCM install image resides in the /usr/sys/inst.images/SDDPCM directory on CD-ROM directory. Because the package does not reside in the /usr/sys/inst.images directory, which is the default directory for the AIX install program, you must mount the CD-ROM file system before you can use SMIT to install SDDPCM from the CD-ROM directory.

Notes:

1. To mount the CD-ROM and install SDDPCM, you must have root access and AIX system administrator knowledge.

2. The AIX SDDPCM host attachment package (devices.fcp.disk.ibm.mpio.rte) must be installed before you install the SDDPCM package (devices.sddpcm.52.rte, devices.sddpcm.53.rte, or devices.sddpcm.61.rte package).

Creating and mounting the CD-ROM filesystem

To install SDDPCM from the CD-ROM, you must first create and mount the CD-ROM filesystem. Use SMIT to perform the following steps to create and mount the CD-ROM to CD-ROM file system.

Note: Throughout this procedure, /dev/cd0 is used for the compact disc driver address. The driver address can be different in your environment.

1. Log in as the root user
2. Insert the compact disc into the CD-ROM drive.
3. From your desktop window, enter **smitty fs** and press **Enter**.
4. Select **Add / Change / Show / Delete File Systems** and press **Enter**.
5. Select **CDROM File System** and press **Enter**.
6. Select **Add a CDROM File System** and press **Enter**. The Add a CDROM File System panel is displayed.
7. Select **DEVICE name** and select **F4**. The DEVICE name panel is displayed.
8. Select the compact disc drive that you are using for the installation, (for example, cd0), and press **Enter**.
9. Select **MOUNT POINT** and enter a directory where you want the CDROM File System to be mounted, (for example, /cdmnt).
10. Click the default option settings for the other fields to ensure that they are what you need.

```
+------------------------------------------------------------+
+ Add a CDROM File System                                    +
+                                                            +
+ Type or select values in entry fields.                     +
+ Press Enter AFTER making all desired changes.              +
+                                                            +
+                                         [Entry Fields]     +
+ * DEVICE name                           cd0                +
+ * MOUNT POINT                           [/cdmnt]           +
+ Mount AUTOMATICALLY at system restart?     no              +
+                                                            +
+------------------------------------------------------------+
```

11. Press **Enter** to create the CDROM File System.
12. When the CDROM File System has been created, press **F10** to exit from smit.
13. From your desktop window, enter **smitty mount** and press **Enter**.
14. Select **Mount a File System** and press **Enter**. The Mount a File System panel is displayed.
15. Select **FILE SYSTEM name** and press **F4**
16. Select the CDROM File System that you created and press **Enter**.
17. Select **DIRECTORY** on which to mount and press **F4**.
18. Select the CDROM File System that you created and press **Enter**.
19. Select **TYPE of file system** and press **Enter**.
20. Select **cdrfs** as the type of file system and press **Enter**.
21. Select **Mount as a REMOVABLE file system?** and press **TAB** to change the entry to *yes*.
22. Select **Mount as a READ-ONLY system?** and press **TAB** to change entry to *yes*.
23. Click to check the default option settings for the other fields to ensure that they are what you need.

```
+------------------------------------------------------------------+
+                Mount a File System                               +
+ Type or select values in entry fields.                          +
+ Press Enter AFTER making all desired changes.                   +
+                                          [Entry Fields]      +
+ FILE SYSTEM name                         [/dev/cd0]          +
+ DIRECTORY over which to mount            [/cdmnt]            +
+ TYPE of file system                      cdrfs               +
```

```
+ FORCE the mount?                              no               +
+ REMOTE NODE containing the file system    []                   +
+    to mount                                                    +
+ Mount as a REMOVABLE file system?             yes              +
+ Mount as a READ-ONLY system?                  yes              +
+ Disallow DEVICE access via this mount?        no               +
+ Disallow execution of SUID and sgid programs no               +
+ in this file system?                                           +
+                                                                +
+----------------------------------------------------------------+
```

24. Press **Enter** to mount the file system.

25. When the file system has been mounted successfully, press **F10** to exit from smit.

    Attention: Do not reboot the system if you only have the SDDPCM host attachment package installed.

Using the System Management Interface Tool facility to install SDDPCM

Use the System Management Interface Tool (SMIT) facility to install SDDPCM. The SMIT facility has two interfaces: nongraphical (enter **smitty** to invoke the nongraphical user interface) and graphical (enter **smit** to invoke the graphical user interface).

Throughout this SMIT procedure, /dev/cd0 is used for the compact disc drive address. The drive address can be different in your environment. Perform the following SMIT steps to install the SDDPCM package on your system.

1. From your desktop window, cd to the directory where the CD-ROM file system is mounted, for example /cdmnt.

2. Go to the directory usr/sys/inst.images/SDDPCM.

3. From your desktop window, enter **smitty install_update** and press **Enter** to go directly to the installation panels. The Install and Update Software menu is displayed.

4. Highlight **Install Software** and press **Enter**.

5. Enter **.** to indicate the current directory and press **Enter**.

6. Highlight **Software to Install** and press **F4**. The Software to Install panel is displayed.

7. Select the devices.sddpcm.52.rte, devices.sddpcm.53.rte, or devices.sddpcm.61.rte package, based on the OS level.

8. Press **Enter**. The Install and Update from LATEST Available Software panel is displayed with the name of the software that you selected to install.

9. Check the default option settings to ensure that they are what you need.

10. Press **Enter** to install. SMIT responds with the following message:

    ```
    ARE YOU SURE??
    Continuing may delete information you may want to keep.
    This is your last chance to stop before continuing.
    ```

11. Press **Enter** to continue. The installation process can take several minutes to complete.

12. When the installation is complete, press **F10** to exit from SMIT.

Unmounting the CD-ROM File System

After successfully installing SDDPCM, use the following procedure to unmount CD-ROM file system in order to remove the CD-ROM:

1. Go to the root (/) directory.

2. Enter **umount /cdmnt** and press **Enter** to unmount the CD-ROM file system from the /cdmnt directory.

3. Enter **rmfs /cdmnt** and press **Enter** to remove the CD-ROM file system.

4. Remove the CD-ROM.

Installing SDDPCM from downloaded code

Use this procedure to install SDDPCM from downloaded code:

1. Download code from the Web site that is appropriate for your system

2. Move the downloaded files to the /usr/sys/inst.images directory

3. Change directory to the /usr/sys/inst.images directory.

4. Untar the downloaded file: for example, tar -xvf devices.sddpcm.53.rte.tar

5. Update the .toc file. For example:

   ```
   pwd
   rm -i .toc
   inutoc .
   grep -i sdd .toc
   ```

   This command should reflect the newer SDDPCM code version that will be updated.

6. Continue the installation by following the instructions beginning in step 3 on page 23.

Installing SDDPCM with the AIX OS from an AIX NIM SPOT server to the client SAN boot disk or the internal boot disk

You can install SDDPCM from an AIX Network Installation Management (NIM) server to the client SAN boot disk or the internal boot disk at the same time that the AIX OS is installed. You must set up the NIM master and create the lpp_source and Shared Product Object Tree (SPOT) resources with the images on a file system, which is either NFS-exported or is obtained from a CD or DVD.

Prepare for the NIM SPOT installation with AIX OS and SDDPCM on the client's SAN boot disk or the internal boot disk. To do this, first set up a NIM master and create the lpp_source and SPOT resource. You can use the System Management Interface Tool (SMIT) facility to implement the following procedures:

1. Install the following filesets to set up the system as an NIM master:

   bos.sysmgt.min.master
   bos.sysmgt.nim.spot

2. Initialize the NIM master system by running the **smitty nim_config_env** command.

3. Create a new lpp_source and SPOT resource by running the **smitty nim_config_env** command.

4. Add the SDDPCM fileset to the newly created lpp_source by running the **smitty nim_task_inst** command.

5. Create a SPOT from the new lpp_source by running the **smitty nim_config_env** command.

6. Define an NIM client by running the **smitty nim** command.

See the NIM task roadmap on the Web for detailed information on how to complete these tasks:

publib16.boulder.ibm.com/pseries/en_US/aixins/insgdrf/
nim_roadmap.htm#nim_roadmap

After you have successfully prepared for the NIM SPOT installation, you are ready
to use the SMIT tool to start the NIM installation on the client system:

1. Run the **smitty nim** command.
   a. Click **Perform NIM Administration Tasks > Manage Network Install
      Resource Allocation > Manage Machines > Allocate Network Install
      Resources**.
   b. Select the hostname of the client that you defined previously.
   c. Select the lpp_source and SPOT resources that you created previously, and
      then press **Enter**.
2. Run the **smitty nim** command again.
   a. Click **Perform NIM Administration Tasks > Manage Machines > Perform
      Operations on Machines**.
   b. Select the hostname of the client that you selected previously.
   c. Click **bos_inst**.
   d. Set the **ACCEPT new license agreements** field to **Yes**, and then press **Enter**.

The system automatically reboots after the **smitty nim** task completes. Use the
following command to check the SAN boot disk and make sure the boot disk is
configured with SDDPCM:

lsattr -El hdiskX (*SAN boot disk device name*)

From the output of this command, check the ODM attribute PCM to ensure that
the value is *PCM/friend/sddpcm*.

Updating SDDPCM

The following sections discuss methods of updating SDDPCM, verifying the
currently installed version of SDDPCM, and the maximum number of devices that
SDDPCM supports.

Updating SDDPCM packages by installing a newer base package or a program temporary fix

SDDPCM allows you to update SDDPCM by installing a newer base package or a
program temporary fix (PTF). A PTF file has a file extension of .bff (for example,
devices.sddpcm.52.rte.2.1.0.1.bff) and can either be applied or committed when it is
installed. If the PTF is committed, the update to SDDPCM is permanent; to remove
the PTF, you must uninstall SDDPCM. If the PTF is applied, you can choose to
commit or to reject the PTF at a later time. If you decide to reject the PTF, you will
not need to uninstall SDDPCM from the host system.

Note: If you are not upgrading the operating system, regardless of whether you
      have SAN boot devices, you can update SDDPCM packages by installing a
      newer base package or a program temporary fix. Otherwise, see "Migrating
      SDDPCM during an AIX OS upgrade with multipath SAN boot devices (on
      supported storage hdisks)."

After applying the base package or the PTF, reboot the system. The SDDPCM
server daemon should automatically start after restarting the system. If it does not
start automatically, start the SDDPCM server daemon manually.

Use the SMIT facility to update SDDPCM. The SMIT facility has two interfaces, nongraphical (enter **smitty** to invoke the nongraphical user interface) and graphical (enter **smit** to invoke the GUI).

If the base package or PTF is on a CD-ROM, you must mount the CD file system. See "Creating and mounting the CD-ROM filesystem" for directions on how to mount the CD file system. In the following procedure, /dev/cd0 is used for the CD drive address. The drive address can be different in your environment.

Perform the following SMIT steps to update the SDDPCM package on your system:

1. Log in as the root user.
2. Type **smitty install_update** and press **Enter** to go directly to the installation panels. The Install and Update Software menu is displayed.
3. Select **Install Software** and press **Enter**.
4. Press **F4** to display the INPUT Device/Directory for Software panel.
5. Select either a CD drive that you are using for the installation or a local directory where the packages reside; for example, /dev/cd0, and press **Enter**.
6. Press **Enter** again. The Install Software panel is displayed.
7. Select **Software to Install** and press **F4**. The Software to Install panel is displayed.
8. Select the base package or the PTF package that you want to install.
9. Press **Enter**. The Install and Update from LATEST Available Software panel is displayed with the name of the software that you selected to install.
10. If you only want to apply the PTF, select **Commit software Updates?** and tab to change the entry to *no*. The default setting is to commit the PTF. If you specify *no* to Commit Software Updates?, ensure that you specify *yes* to Save Replaced Files?.
11. Check the other default option settings to ensure that they are what you need.
12. Press **Enter** to install. SMIT responds with the following message:

```
+----------------------------------------------------------------+
|ARE YOU SURE??                                                  |
|Continuing may delete information you may want to keep.         |
|This is your last chance to stop before continuing.            |
+----------------------------------------------------------------+
```

13. Press **Enter** to continue. The installation process can take several minutes to complete.
14. When the installation is complete, press **F10** to exit from SMIT.
15. Unmount the CD-ROM file system and remove the compact disc.

Committing or rejecting a program temporary fix update

Before you reject a PTF update, you need to unconfigure and remove all supported storage devices from your host system. Committing a PTF does not require this extra step. Perform the following steps to commit or reject a PTF update with the SMIT facility. The SMIT facility has two interfaces: nongraphical (enter **smitty** to invoke the nongraphical user interface) and graphical (enter **smit** to invoke the GUI).

1. Log in as the root user.
2. From your desktop window, enter **smitty install** and press **Enter** to go directly to the installation panels. The Software Installation and Maintenance menu is displayed.
3. Select **Software Maintenance and Utilities** and press **Enter**.

4. Select **Commit Applied Software Updates** to commit the PTF or select **Reject Applied Software Updates** to reject the PTF.

5. Press **Enter**. The Commit Applied Software Updates panel is displayed or the Reject Applied Software Updates panel is displayed.

6. Select **Software name** and press **F4**. The software name panel is displayed.

7. Select the Software package that you want to commit or reject.

8. Check the default option settings to ensure that they are what you need.

9. Press **Enter**. SMIT responds with the following message:

```
+----------------------------------------------------------------------+
|ARE YOU SURE??                                                        |
|Continuing may delete information you may want to keep.               |
|This is your last chance to stop before continuing.                  |
+----------------------------------------------------------------------+
```

10. Press **Enter** to continue. The commit or reject process can take several minutes to complete.

11. When the installation is complete, press **F10** to exit from SMIT.

   Note: You do not need to restart the system even though the bosboot message may indicate that a restart is necessary.

Verifying the currently installed version of SDDPCM

You can verify your currently-installed version of SDDPCM by issuing one of the following commands:

lslpp -l *sddpcm*

or

pcmpath query version

Note: Using the **pcmpath** command requires that at least one SDDPCM-supported MPIO device is configured on the system.

Maximum number of devices that SDDPCM supports

SDDPCM supports a maximum of 1200 configured devices and a maximum of 16 paths per device. That means the maximum number of host adapter ports that are supported is 16. However, with the round robin or load balance path selection algorithms, configuring more than four paths per device might impact the I/O performance. Use the minimum number of paths necessary to achieve sufficient redundancy in the SAN environment. The recommended number of paths per device is four.

In order to support 1200 supported storage device LUNs, system administrators should first determine whether the system has sufficient resources to support a large number of devices. See "Preparing your system to configure more than 600 supported storage devices or to handle a large amount of I/O after queue depth is disabled "for more information.

Migrating SDDPCM

The following sections discuss the methods of migrating SDDPCM with and without SAN boot devices

- "Migrating the supported storage SAN boot device or nonboot volume group from AIX default PCM to SDDPCM" on page 28
- "Migrating from SDDPCM to the AIX default PCM or to SDD" on page 29

- "Migrating from SDD with SAN boot devices (on supported storage hdisks) to SDDPCM with multipath SAN boot devices" on page 29

Migrating the supported storage SAN boot device or nonboot volume group from AIX default PCM to SDDPCM

The default reserve policy of the AIX base PCM is a single-path policy, which is SCSI-2 reserve. The path selection algorithm is fail_over, which means only one path is opened at a time and that path made SCSI-2 reserve to the disk. All I/O is routed to this path. This reserve policy and path selection algorithm can cause problems if, after the SDDPCM packages are installed, you open the default PCM device before you restart the system. The default PCM device can be opened if:

- You build a volume group and file system with AIX default PCM devices, and leave the volume groups active and file system mounted
- You configure default PCM devices as the backing devices of a virtual target device when you are in a VIOS environment

After the system starts, you might see some paths in the INVALID state. The INVALID state means that the path failed to open. This is because the SCSI-2 reserve is not released during the system restart; thus, only the paths previously opened with SCSI-2 reserve can be opened and used for I/O after system restart. You might not see paths in the INVALID state if your system is at AIX 5.2 TL10 or later or at AIX 5.3 TL07 or later, or if you have IY83717 or IY83847 installed on your system. Instead, you might see a heavier I/O select count on one path. This is because the SCSI-2 reserve is not released during the system restart. Even if all the paths are allowed to be opened, only opened paths that previously made SCSI-2 reserve can actually be used for I/O.

If you have supported storage SAN boot devices that are configured with AIX default PCM, and the reserve policy is single_path (SCSI-2 reserve), switching the boot devices from AIX default PCM to SDDPCM might result in this reservation conflict situation. If you install an SDDPCM version earlier than 2.2.0.1, you must always run the **relbootrsv** command to release the SCSI-2 reserve on SAN boot devices after you install the SDDPCM host attachment package and the SDDPCM package. Run the **relbootrsv** command before you restart the system and then run the following command against the hdisks that are part of the rootvg to verify that they are no longer reserved.

```
# pcmquerypr -Vh /dev/hdisk6
connection type: fscsi0
open dev: /dev/hdisk6

Attempt to read reservation key...

Attempt to read registration keys...
Read Keys parameter
        Generation :  0
        Additional Length:  0
resrvpolicy= no_reserve
Reserve Key provided by current host = none (hex)02bbb003
Not reserved.
```

If you install SDDPCM version 2.2.0.1 or later, the SCSI-2 reserve on SAN boot devices is released automatically during system boot.

In a VIOS environment, reservation conflict problems can occur on a virtual target device that used to have a default PCM device as its backing device. To prevent this problem, perform one of the following actions:

- Switch from the AIX default PCM to SDDPCM before you use the AIX default PCM device as the backing device of a virtual target device.
- Before switching from the AIX default PCM to SDDPCM, put the virtual target devices in the **Define** state. This properly closes the AIX default PCM and releases the SCSI-2 reserve before migrating to SDDPCM.

Reservation conflict problems can also occur on nonboot volume groups. To prevent this problem, perform one of the following actions:

- Switch from the AIX default PCM to SDDPCM before you make any volume groups and file systems.
- To switch from the AIX default PCM to SDDPCM, you must unmount file systems and vary off the volume group of the AIX default PCM to release the SCSI-2 reserve on the volume group before system restart.
- Issue **relbootrsv** *VGname* to release the SCSI-2 reserve on the active, nonboot volume group devices before you restart the system.

Note: If you specify a *VGname* (volume group name), **relbootrsv** releases the SCSI-2 reserve of the specified non-SAN boot volume group; otherwise, it releases the SCSI-2 reserve of a SAN boot volume group (`rootvg`).

Migrating from SDDPCM to the AIX default PCM or to SDD

Note: If you have supported storage devices configured with SDDPCM as MPIO boot devices, you need to contact IBM Customer Support for migration from SDDPCM to the AIX default PCM. The following procedure is applicable only for a system without SDDPCM MPIO devices configured as the SAN boot device.

To migrate from SDDPCM to the AIX default PCM or to SDD, you must first unconfigure the devices, stop the SDDPCM server daemon, and then uninstall the SDDPCM package and the SDDPCM host attachment package. See "Removing SDDPCM from an AIX host system" for directions on uninstalling SDDPCM. After you uninstall SDDPCM, you can then restart the system to migrate supported storage MPIO devices to the AIX default PCM. If you want to migrate supported storage devices to SDD devices, you must then install the supported storage device host attachment for SDD and the appropriate SDD package for your system. Then restart the system to configure the supported storage devices to SDD vpath devices.

Migrating from SDD with SAN boot devices (on supported storage hdisks) to SDDPCM with multipath SAN boot devices

If you have supported storage devices configured with SDD and there are SAN boot devices with supported storage hdisk devices, you need to contact IBM Customer Support for migration from SDD to SDDPCM.

Migrating SDDPCM during an AIX OS upgrade with multipath SAN boot devices (on supported storage hdisks)

SDDPCM provides different packages to match the AIX OS level. If an AIX system is going to be upgraded to a different OS level; for example, from AIX 5.3 to AIX 6.1, you need to install the corresponding SDDPCM package for that OS level.

If you want to upgrade AIX OS and there are SAN boot devices with SDDPCM supported storage hdisk devices, you need to contact IBM Customer Support for migration from SDDPCM during the OS upgrade.

If you are not in SAN boot environment, or you are only upgrading the AIX OS Technology Level or Server Pack; for example, from AIX 5.3 TL04 to AIX 5.3 TL06, you can follow the procedures in "Updating SDDPCM" on page 25.

Configuring and unconfiguring supported storage MPIO-capable devices

After installing the MPIO-supported storage device host attachment and the SDDPCM package, you need to restart the system in order to configure supported storage devices as MPIO-capable devices. After the first system restart, you can then use the normal AIX command line configure programs to configure and unconfigure supported storage MPIO-capable devices.

After the system restarts, the SDDPCM server daemon (**pcmsrv**) should automatically start.

Configuring supported storage MPIO-capable devices

The newly installed supported storage devices must be configured as MPIO-capable devices before you can use them. Use one of the following commands to configure these devices:

* **cfgmgr** command

    Note: If operating in a switched environment, issue the **cfgmgr** command once for each host adapter each time a device is added.

    If you use the **cfgmgr** command to configure supported storage MPIO devices, you might need to start the SDDPCM server daemon manually, if it has not already started. See "SDDPCM server daemon" for information describing how to check the daemon status and how to manually start the daemon.

* **shutdown -rF** command to restart the system.

    After the system restarts, the SDDPCM server daemon (**pcmsrv**) should automatically start.

Verifying the SDDPCM configuration

To verify the SDDPCM configuration, you can use one of the following:

* SMIT MPIO management submenu
* SDDPCM **pcmpath query device** command

Perform the following steps using SMIT to verify the SDDPCM configuration on an AIX host system:

1. Enter **smitty MPIO** from your desktop window. The **MPIO management** menu is displayed.
2. Select MPIO Device Management and press **Enter**. The MPIO Device Management panel is displayed.
3. Select List ALL MPIO Devices and press **Enter**. All MPIO devices on the host are listed.
4. Search for all IBM MPIO FC *XXXX* devices, where *XXXX* can be 2105, 2107, 1750, or SAN Volume Controller MPIO devices. Also search for MPIO FC 2145 if you have a host attachment package for SDDPCM with version 1.0.0.8 or later. Ensure that these devices are in the *Available* state.

Search for all IBM MPIO DS*XXXX* array disks, where *XXXX* can be 4100, 4200, 4300, 4400, 4700, 4800, 5000, 5020, or 3950 devices. Ensure that these devices are in the **Available** state

You can also use the SDDPCM **pcmpath query device** command to query the configuration status of supported storage devices.

Notes:

1. If none of the supported storage devices are configured successfully as MPIO devices, the **pcmpath query device** command will fail.

2. Because the AIX disk-configure method supports parallel device configuration, the same MPIO logical device name can configured with different supported storage devices after each device configuration. If you have a program or script that has the MPIO logical device name hardcoded, ensure that you check the MPIO logical device name binding with the storage device LUN. If there is a change to the binding between an MPIO logical device name and the storage device LUN, you must revise your program or script.

Dynamically adding and removing paths or adapters

SDDPCM supports dynamically adding or removing extra paths or adapters while I/O is running. To add extra paths that are attached to an adapter to existing available devices, enter:

mkpath -l hdiskX -p fscsiY

When the command returns successfully, the paths are added to the devices.

To check the device configuration status, enter:

lspath -l hdiskX -H -F "*name path_id parent connection status***"**

or

pcmpath query device X

To add a new adapter to existing available supported storage MPIO devices, enter: **cfgmgr -vl fscsiX**

To check the adapter configuration status, enter:

pcmpath query adapter

or

pcmpath query device

To dynamically remove all paths under a parent adapter from a supported storage MPIO device, enter: **rmpath -dl hdiskX -p fscsiY**

To dynamically remove an adapter and all child devices from supported storage MPIO devices, use **smit mpio**, or enter the following on the command line:

rmdev -l fscsiX -R

To dynamically remove a particular path, run **smit mpio**, or enter one of the following commands on the command line:

rmpath -l hdiskX -p fscsiY -w *connection location code*

or

rmpath -dl hdiskX -p fscsiY -w *connection location code*

Issue the following command to get a particular path connection location code:

lspath -l hdiskX -H -F "*name path_id parent connection status***"**

Note: You cannot remove the last path from a supported storage MPIO device. The command will fail if you try to remove the last path from a supported storage MPIO device.

Unconfiguring supported storage MPIO-capable devices

Note: Contact IBM support if your system has SAN boot MPIO devices configured with SDDPCM on supported storage system disks. The SAN boot MPIO devices cannot be unconfigured with following procedures.

To remove all non-SAN supported storage MPIO-capable devices:
1. Unmount the file systems of all supported storage devices.
2. Vary off all supported storage device volume groups.
3. Enter the **stopsrc -s pcmsrv** command to stop **pcmsrv**.
4. Enter the following command for each adapter:

   rmdev -dl fcsX -R

Switching DS4000, DS5000, or DS3950 storage device configurations

For SDDPCM version 2.5.1.0 and subsequent releases, with systems having AIX 5.3 TL08 and AIX 6/1 TL02 and required AIX APAPs installed, users can use the AIX **manage_disk_drivers** command to display the configured drivers for DS3950, DS4100, DS4200, DS4300, DS4500, DS4700, DS4800, DS5100, and DS5020 storage devices, and a list of supported drivers for these device models. Supported drivers are SDDPCM (displayed as AIX_SDDAPPCM), AIX native PCM (displayed as AIX_APPCM), and AIX_FCP array driver (displayed as AIX_fcparray; only supports DS4K storage devices). For example:

```
[root@telefone] /home/root # manage_disk_drivers -l

Device Present Driver Driver Options
2810XIV AIX_AAPCM AIX_AAPCM,AIX_non_MPIO
DS4100 AIX_SDDAPPCM AIX_APPCM,AIX_fcparray
DS4200 AIX_SDDAPPCM AIX_APPCM,AIX_fcparray
DS4300 AIX_SDDAPPCM AIX_APPCM,AIX_fcparray
DS4500 AIX_SDDAPPCM AIX_APPCM,AIX_fcparray
DS4700 AIX_SDDAPPCM AIX_APPCM,AIX_fcparray
DS4800 AIX_SDDAPPCM AIX_APPCM,AIX_fcparray
DS3950 AIX_SDDAPPCM AIX_APPCM
DS5020 AIX_SDDAPPCM AIX_APPCM
DS5100 AIX_SDDAPPCM AIX_APPCM
```

This AIX command also allows users to switch DS4000, DS5000, and DS3950 storage device configuration drivers among these driver options. However, switching device configuration drivers requires a system reboot to reconfigure the devices after running this command. The syntax of this command function is:

`manage_disk_drivers -d device -o driver_option`

Removing SDDPCM from an AIX host system

Before you remove the SDDPCM package from your AIX host system, all supported storage devices must be unconfigured and removed from your host system, or migrated to the AIX default PCM. The SDDPCM server daemon must be stopped.

Note: SDDPCM supports MPIO-supported storage devices as SAN boot devices. If your server has a supported storage device configured with SDDPCM as the SAN boot device, contact IBM support for your SAN boot configuration environment.

To remove SDDPCM, you must remove both the host attachment for SDDPCM and the SDDPCM software packages before you reconfigure supported storage devices or restart the system. Otherwise, the devices can be in the *Defined* state and cannot be configured as either MPIO or non-MPIO devices.

After all the supported storage devices are removed or migrated to the AIX default PCM and the SDDPCM server daemon (**pcmsrv**) is stopped, perform the following steps to remove the SDDPCM software package:

1. Enter **smitty deinstall** from your desktop window to go directly to the Remove Installed Software panel.
2. Press **F4** in the SOFTWARE name field to bring up a list of packages and press the **F7** key to select the package to uninstall.
3. Press **Tab** in the **PREVIEW Only?** field to toggle between Yes and No. Select **No** to remove the software package from your AIX host system.

   Note: If you select **Yes**, the process stops at this point and previews what you are removing. The results of your precheck are displayed without removing the software. If the state for any supported storage MPIO device is either *Available* or *Defined*, the process fails.
4. Select **No** for the remaining fields on this panel.
5. Press **Enter**. SMIT responds with the following message:

   ```
   ARE YOU SURE??
   Continuing may delete information you may want to keep.
   This is your last chance to stop before continuing.
   ```

6. Press **Enter** to begin the removal process. This might take a few minutes.
7. When the process is complete, the SDDPCM software package and the supported storage device host attachment for SDDPCM are removed from your system.

SDDPCM support for HACMP with Enhanced Concurrent Mode volume groups

Beginning with SDDPCM 2.1.2.0, SDDPCM supports HACMP V5.1, V5.3, and V5.4 on an AIX 5.2 TL07 (or later) and AIX 5.3 TL03 (or later) system with both concurrent and nonconcurrent resource groups. SDDPCM 2.4.0.0 supports HACMP V5.3, and V5.4 on an AIX 6.1 system with both concurrent and nonconcurrent resource groups.

This support requires definition of the shared volume groups as Enhanced Concurrent Mode volume groups, in either concurrent or nonconcurrent resource groups. This means that no reserve needs to be broken in response to a node failure, and hence any requirement on breaking reserves is removed. A special interaction between HACMP and LVM ensures that if the volume group is used in a nonconcurrent resource group, applications are allowed to access it on one node at a time. Only no_reserve policy is supported in both concurrent and nonconcurrent resoure groups.

The Enhanced Concurrent Mode volume groups are sufficient to ensure high availability. However, If system ECM volume groups are in nonconcurrent resource groups, you should configure your SAN using the following guidelines:

* The interaction between HACMP and LVM to ensure that only one node has nonconcurrent access at a time is *advisory locking*. This is in contrast to the *mandatory locking* provided by SCSI reserves. To ensure that production data is not inadvertently modified by nodes that are not in the HACMP cluster, the following should be done:

  1. Use either physical cabling or zoning to ensure that only HACMP nodes have access to the shared LUNs. That is, non-HACMP nodes should be prevented by hardware from accessing the shared LUNs.
  2. Start HACMP on the cluster nodes at boot time. This ensures that HACMP will activate the appropriate access controls on the shared disk before applications have a chance to modify the access controls.

* Configure disk heartbeating to reduce the likelihood of one node considering the other *dead* and attempting to take over the shared disks. (This is known as a *partitioned cluster*, or *split brain syndrome*). If the shared disks consist of multiple enclosures, use one disk in each enclosure as a heartbeat path.

Different storage systems or models might support different versions of HACMP and SDDPCM. For information, see the interoperability matrix for your storage:

http://www.ibm.com/systems/storage/disk/ess/
http://www.ibm.com/systems/storage/disk/ds6000/
http://www.ibm.com/systems/storage/disk/ds8000/
http://www.ibm.com/systems/storage/software/virtualization/svc/

For HACMP v5.1, v5.2, v5.3, and v5.4 for AIX5L support information, go to the following Web site:

http://publib.boulder.ibm.com/infocenter/clresctr/vxrx/index.jsp?topic=/
com.ibm.cluster.hacmp.doc/hacmpbooks.html

For HACMP up-to-date APAR information, go to the following Web site:

http://www14.software.ibm.com/webapp/set2/sas/f/hacmp/download/
aix53.html

MPIO reserve policies

SDDPCM supports four types of MPIO reserve policies. You can select one of the four reserve policies based on their configuration environment or application needs. The supported reserve policies are:

- No Reserve reservation policy
- Exclusive Host Access Single Path reservation policy
- Persistent Reserve Exclusive Host Access reservation policy
- Persistent Reserve Shared Host Access reservation policy

No Reserve reservation policy

If you set MPIO devices with this reserve policy, there is no reserve being made on MPIO devices. A device without reservation can be accessed by any initiators at any time. I/O can be sent from all the paths of the MPIO device. This is the default reserve policy of SDDPCM.

Exclusive Host Access single-path reservation policy

This is the SCSI-2 reservation policy. If you set this reserve policy for MPIO devices, only the fail_over path selection algorithm can be selected for the devices. With this reservation policy, an MPIO device has all paths being opened; however, only one path made a SCSI-2 reservation on the device. I/O can only be sent through this path. When this path is broken, reserve is released, another path is selected, and SCSI-2 reserve is reissued by this new path. All input and output is now routed to this new path.

Persistent Reserve Exclusive Host Access reservation policy

If you set an MPIO device with this persistent reserve policy, a persistent reservation is made on this device with a persistent reserve (PR) key. Any initiators who register with the same PR key can access this device. Normally, you should pick a unique PR key for a server. Different servers should have different unique PR key. I/O is routed to all paths of the MPIO device, because all paths of an MPIO device are registered with the same PR key. In a nonconcurrent clustering environment, such as HACMP, this is the reserve policy that you should select.

Current® HACMP clustering software supports no_reserve policy with Enhanced Concurrent Mode volume group. HACMP support for persistent reserve policies for supported storage MPIO devices is not available.

Persistent Reserve Shared Host Access reservation policy

If you set an MPIO device with this persistent reserve policy, a persistent reservation is made on this device with a persistent reserve (PR) key. However, any initiators that implemented persistent registration can access this MPIO device, even if the initiators are registered with different PR keys. In a concurrent clustering environment, such as HACMP, this is the reserve policy that you should select for sharing resources among multiple servers.

Current HACMP clustering software supports no_reserve policy with Enhanced Concurrent Mode volume group. HACMP support for persistent reserve policies for supported storage MPIO devices is not available.

SDDPCM ODM attribute settings

The following sections discuss the SDDPCM ODM attribute default settings, and how to change the attributes of the supported storage MPIO devices:

- "SDDPCM ODM attribute default settings"
- "Changing device reserve policies"
- "Changing the path selection algorithm" on page 37
- "Changing SDDPCM path healthcheck mode" on page 38

SDDPCM ODM attribute default settings

SDDPCM has following default attributes settings:

Attribute	Default value
device reserve policy	no_reserve
path selection algorithm	load balance
healthcheck mode	nonactive
healthcheck time interval	60 seconds

Changing device reserve policies

Use the **chdev** command to change the reserve policy for a device. Because **chdev** requires the device to be unconfigured and reconfigured, this is a disruptive operation.

The following reserve policies can be used with any of the supported path selection algorithms (see "Supported SDDPCM features"):

- no_reserve
- persistent reserve exclusive host access
- persistent reserve shared host access

When the reserve policy of a device is exclusive host access single path (SCSI-2), the only supported path selection algorithm is the fail_over algorithm. The fail_over algorithm selects one path at a time for all I/Os. When the active path fails, then an alternative path is selected. The SCSI-2 reserve is reissued by this alternative path

To change the device reserve policy to no_reserve, enter:

chdev -l hdiskX -a reserve_policy=no_reserve

If you want to change the reserve policy to one of the persistent reserve policies, you must provide a persistent reserve key at the same time that you change the device policy to one of the persistent reserve types. For example, to change the reserve policy to PR_shared:

chdev -l hdiskX -a PR_key_value=0x1234 -a reserve_policy=PR_shared

Note: Beginning with SDDPCM 2.1.0.0, 2 persistent reserve tools are provided to manage supported storage system MPIO devices. See "Persistent reserve command tools" on page 47 for more information.

Changing the path selection algorithm

Starting with SDDPCM 2.1.0.0, you can use the **pcmpath set device algorithm** command to dynamically change the path selection algorithm. See "pcmpath set device algorithm" on page 78 for information about this command.

You can also use the **chdev** command to change the path selection algorithm of a device. Because **chdev** requires that the device be unconfigured and then reconfigured, this is a disruptive operation.

Use the following command to change the device path selection algorithm to round robin:

chdev -l hdiskX -a algorithm=round_robin

You can change the reserve_policy and algorithm for a device with one command. For example, to change the reserve policy to no_reserve and the path selection algorithm to round robin:

chdev -l hdiskX -a reserve_policy=no_reserve -a algorithm=round_robin

Using the load balancing port algorithm

Beginning with SDDPCM 2.4.0.0, you can use the load balancing port algorithm to achieve better balanced I/O performance in certain specific SAN configuration environments. This path selection algorithm supports only 2105, 2107, 1750, and 2145 storage devices. This algorithm is based on the existing load balancing algorithm, but it also uses the incorporated I/O statistics on each target port to which the host is connected. You can use this algorithm for SAN configurations where I/O throughput is unbalanced on the storage targets.

The following examples demonstrate the possible configurations in which this algorithm can help to improve performance.

In Figure 4, link 2 between the switch and the storage is receiving twice the load of the other links. This condition creates a throughput imbalance on link 2 that prevents optimal load balancing on both hosts.
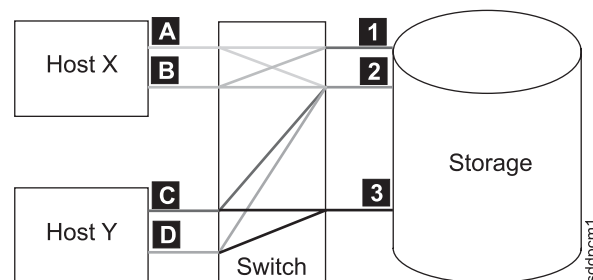
Figure 4. Workload imbalance when one link receives twice the load of the other links

Host X is attempting to balance the load across four paths (A1, A2, B1, B2). Host Y is attempting to balance the load across four paths (C2, C3, D2, D3).

In Figure 5 on page 38, link 2 between the switch and the storage is more heavily loaded than link 1. This condition creates a throughput imbalance on link 2 that prevents optimal load balancing on Host X.
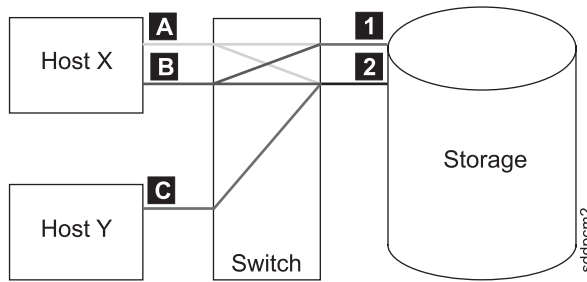
Figure 5. Workload imbalance when one link is more heavily loaded than another link

Host X is attempting to balance the load across four paths (A1, A2, B1, B2). Host Y has only one active path to the storage (C2).

In Figure 6, one path is lost from a failed component, which results in the same workload imbalance shown in Figure 5.
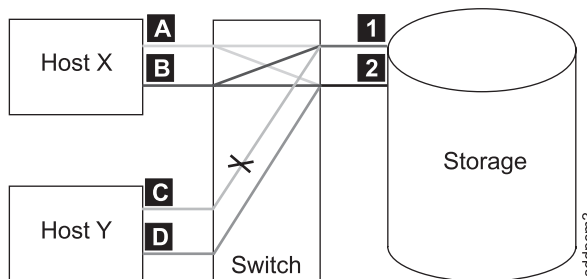
Figure 6. Workload imbalance when one host sharing workload across two paths loses one path

Host X is attempting to balance the load across four paths (A1, A2, B1, B2). Host Y has two paths to the storage, but only one path is active (C1, D2).

You can either use **pcmpath set device algorithm lbp** or **chdev -l hdiskX -a algorithm=load_balance_port** to configure the device with the load balancing port algorithm. To ensure that the load balancing port algorithm produces the desired result, all the hdisks on a host system that is attached to a given storage device must be configured with the same algorithm.

Note: This algorithm improves performance by balancing drastically unbalanced throughput on multiple target ports of the same storage. It cannot completely balance the throughput on the target ports because SDDPCM manages I/O load balancing only on the host where it is installed.

```
pcmpath query port target port number
pcmpath query portstats target port number
```

To display performance information about the target ports, you can use the **pcmpath query port** *target port number* command and the **pcmpath query portstats** *target port number* command. For more information, see "Using SDDPCM pcmpath commands" on page 51.

Changing SDDPCM path healthcheck mode

SDDPCM supports the path healthcheck function. If this function is enabled, SDDPCM tests opened paths and reclaims failed paths based on the value that is set in the following device healthcheck attribute:

hc_mode

Healthchecking supports the following modes of operations:

- Enabled - When this value is selected, the healthcheck command will be sent to paths that are opened with a normal path mode.
- Failed - When this value is selected, the healthcheck command is sent to paths that are in failed state.
- Nonactive - When this value is selected, the healthcheck command will be sent to paths that have no active I/O. This includes paths that are opened or in failed state.

If the algorithm selected is round robin or load balance, the healthcheck command will only be sent to failed paths, because the round robin and load-balancing algorithms route I/O to all opened paths that are functional. The default value setting of SDDPCM is *nonactive*.

Starting with SDDPCM 2.1.0.0, the **pcmpath set device hc_mode** command allows you to dynamically change the path healthcheck mode. See "pcmpath set device hc_mode" on page 80 for information about this command.

You can also use the **chdev** command to change the device path healthcheck mode. Because **chdev** requires that the device be unconfigured and then reconfigured, this is a disruptive operation. To change the path healthcheck mode to *failed*, issue following command:

chdev -l hdiskX -a hc_mode=failed

Changing SDDPCM path healthcheck time interval

The hc_interval attribute will determine how often the paths of a device should be health-checked. The hc_interval attribute has a range of values from 0 - 3600 seconds. When a value of 0 is selected, the healthcheck function is disabled. Beginning with SDDPCM v2.1.2.3, the default value setting is changed from 20 (seconds) to 60 (seconds).

Starting with SDDPCM 2.1.0.0, the **pcmpath set device hc_interval** command allows you to dynamically change the path healthcheck time interval. See "pcmpath set device hc_interval" on page 79 for information about this command.

You can also use the **chdev** command to change the device path healthcheck time interval. Because **chdev** requires that the device be unconfigured and then reconfigured, this is a disruptive operation. To disable the path healthcheck interval function, issue the following command:

chdev -l hdiskX -a hc_interval=0

Note: Currently, the SDDPCM healthcheck function only checks the paths being opened. It does not healthcheck any path that is in the close state. The SDDPCM server daemon healthchecks close_failed paths. If a device healthcheck function is disabled, the SDDPCM server daemon healthchecks the device's failed path that is opened. See "Using the SDD server and the SDDPCM server" for more information.

From SDDPCM v2.1.2.3 and later, a new feature of SDDPCM server daemon
healthcheck function is introduced. The SDDPCM server daemon automatically
starts or stops the healthcheck function on a device if you issue one of the
following:

- **pcmpath set device m hc_interval 0** on the fly to disable a device internal
  healthcheck function
- **pcmpath set device m hc_interval n** to enable a device internal healthcheck
  function on the fly

Note: The SDDPCM server daemon only healthcheck FAILED paths. It does not
implement healthcheck on opened or idle paths as the SDDPCM internal
healthceck does. This is the difference between the SDDPCM internal
healthcheck and the SDDPCM server daemon healthcheck.

Supported AIX fibre-channel device driver features

SDDPCM supports the following fibre channel adapter driver features:

- Fast I/O failure of fibre channel devices
- Fibre-channel dynamic device tracking

Fast I/O failure of fibre-channel devices

The AIX fibre channel adapter driver supports the new fc_err_recov attribute in
AIX. If enabled, this attribute can reduce the amount of time needed to fail I/O in
certain conditions. By default, this attribute has a value of *delayed_fail*, which
means that the fast failure is disabled.

To enable fast failure, you must perform the following steps:

1. Change all the child devices of fscsiX on your system to the defined state by
   running the **rmdev -l fscsiX -R** command for all adapters on your system.
2. Enter the **chdev -l fscsiX -a fc_err_recov=fast_fail** command for all adapters on
   your system.
3. Enter the **cfgmgr** command to reconfigure all devices back to the available
   state.

You need to have the following Fibre Channel device driver APARs applied to
your system in order to have all dynamic tracking and fast fail fixes:

APAR	Description
IY37183	Dynamic Tracking and Fast I/O Failure of Fibre Channel Devices
IY44139	Dynamic Tracking: PERM I/O errors after Host HBA Cable Pull
IY44142	Dynamic Tracking: I/O Hang after Host HBA Cable Pull
IY44342	Dynamic Tracking: Ioctl call may fail after N_Port ID Change
IY44344	Dynamic Tracking: Back-to-Back Cable Move May Delay Error Recovery
IY45368	Fast Fail/Dynamic Tracking: FC Device Inaccessible after Move
IY46701	Dynamic Tracking & MPIO: Multiple Cable Swap Cause Path Failure

Note: You must set the FC HBA **fc_err_recov** attribute to **fast_fail** for DS4000,
DS5000, and DS3950 storage devices.

Fibre channel dynamic device tracking

Dynamic Tracking support allows for the following scenarios without I/O errors or loss of access:

- Combine two switches in two different SANS into one SAN by connecting the switches with a cable and cascading the switches.
- Change the supported storage device FC cable from one switch port to a different switch port; Disconnected cable must be reconnected within 15 seconds.
- Swap two supported storage device FC cables on the switch ports. The disconnected cables must be reconnected within 15 seconds.
- Swap two supported storage device FC cables on the supported storage device ports. The disconnected cables must be reconnected within 15 seconds.

Notes:

1. This 15 seconds window actually includes the time required to bring up the FC link after the cable is reconnected. For example, if it takes 4 seconds to bring up the FC link, the time allowed to leave the cable disconnected is 11 seconds. I/O will fail if the cable is disconnected longer than 11 seconds.

2. Fibre-channel device dynamic tracking does not support the case when a cable is moved from one adapter on the supported storage device to another free, previously unseen adapter on the supported storage device. The reason is that the world wide port name (WWPN) will be different for that adapter, and the supported storage device cannot track that adapter. The WWPN must be a static identifier of a remote port.

AIX has added a new attribute that must be changed to enable dynamic tracking. By default, dynamic tracking is disabled. To enable dynamic tracking, you must do the following:

1. Change all the child devices of fscsiX on your system to the defined state by issuing the **rmdev -l fscsiX -R** command for all adapters on your system.

2. Issue the **chdev -l fscsiX -a dyntrk=yes** command for all adapters on your system.

3. Issue **cfgmgr** to reconfigure all devices back to the available state.

   Note: You must set the FC HBA **dyntrk** attribute to **yes** for DS4000, DS5000, and DS3950 storage devices.

For more information about Dynamic Tracking and Fast I/O Failure, see the AIX release notes at:

http://publib.boulder.ibm.com/pseries/aixgen/relnotes/52RELNOTES/10073902.htm

Installing APAR IY37183 also installs the file:

/usr/lpp/bos/README.FIBRE-CHANNEL

This file has more information about the Dynamic Tracking and Fast I/O Failure features.

Multipath SAN boot support

This section describes how to configure supported storage system MPIO devices as the SAN boot device.

Attention: The sddpcm default reserve policy is no_reserve and the default path selection algorithm is load_balance. If the SAN boot devices are configured with the default reserve policy and path selection algorithm, you must not configure these SAN boot devices on other AIX servers. This is because no scsi-2 reserve or scsi-3 persistent reserve is implemented on these SAN boot devices to prevent access from another AIX server.

Configuring supported storage system MPIO devices as the SAN boot device

A supported storage MPIO device can be used as the system boot device. To configure the supported storage device boot device with the SDDPCM module:

1. Select one or more supported storage system devices as the boot device.
2. Install one of the following AIX operating systems on the selected supported storage devices:
   - If the selected supported storage device is ESS, the required operating system is AIX 5.2 TL06 (or later), AIX 5.3 TL02 (or later), or AIX 6.1 TL0 (or later).
   - If the selected supported storage device is DS6000, the required operating system is AIX 5.2 TL07 (or later), AIX 5.3 TL03 (or later), or AIX 6.1 TL0 (or later).
   - If the selected supported storage device is DS8000, the required operating system is AIX 5.2 TL07 (or later), AIX 5.3 TL03 (or later), or AIX 6.1 TL0 (or later).
   - If the selected supported storage device is SAN Volume Controller, the required operating sytsem is AIX 5.2 TL07 (or later), AIX 5.3 TL03 (or later), or AIX 6.1 TL0 (or later).
   - If the selected supported storage device is DS4000, DS5000, or DS3950, the required operating system is AIX53 TL8 (or later) or AIX61 TL2 (or later).
3. Restart the system. The supported storage boot device is configured as an MPIO device with AIX default PCM.

   Note: For IBM DS4000 storage devices, if the OS is AIX53, or if the system was migrated from AIX53 to AIX61 (or later), the DS4000 devices are configured by the AIX FCPArray(RDAC) device driver instead of the AIX native MPIO PCM.

4. Install the supported storage device host attachment for SDDPCM and the SDDPCM packages.
5. To release scsi-2 reserve on boot devices with SDDPCM v2.2.0.0 or earlier, you must issue the **relbootrsv** command. For boot devices with SDDPCM v2.2.0.1 or later, scsi-2 reserve is automatically released during the system reboot. If you want to release non-rootvg scsi-2 reserves, provide the volume group name as a parameter. For example: **relbootrsv vgname**.
6. Restart the system.

All supported storage MPIO devices, including supported storage MPIO SAN boot devices, are now configured with SDDPCM.

When you convert a boot device with SDDPCM v2.2.0.0 and earlier from the AIX default PCM to SDDPCM, you must issue the **relbootrsv** command, as shown in

step 5. If you fail to do so, you might encounter a problem where either all paths of the boot device cannot be opened successfully or they can be opened but cannot be used for I/O. This problem occurs because the AIX default PCM has a default reserve policy of single-path (SCSI-2). See "Migrating the supported storage SAN boot device or nonboot volume group from AIX default PCM to SDDPCM" on page 28 for information about solving this problem.

There is a known problem during the SAN boot configuration. After the system is restarted followed by the operating system installation on the supported storage MPIO devices, you might see that some paths of the rootvg are in Failed path state. This can happen even if the system is successfully restarted.

This problem is corrected in AIX 5.2 TL08 or later and AIX 5.3 TL04 or later. Apply the following APARs on these OS levels after the first reboot followed by the operating system installation:
- AIX 5.2 TL08 or later: apply APAR IY83717
- AIX 5.3 TL04 or later: apply APAR IY83847

No APAR is available to correct this problem on AIX52 TL07 and AIX53 TL03. If you configure a SAN boot device with supported storage MPIO devices on one of these operating system levels and experience this problem, you can manually recover these failed paths by issuing one of the following commands:
- **chpath -s E -l hdiskX -p fscsiX**
- **pcmpath set device M path N online**

Support system dump device with the supported storage system MPIO device

You can choose a supported storage MPIO device to configure with the system primary and secondary dump devices. You can configure the system dump device with the supported SAN boot device, or with the nonboot device. The path selection algorithm for the system dump device automatically defaults to *failover_only* when the system dump starts.

During the system dump, only one path is selected for dump requests. If the first path fails, I/O is routed to the next path being selected.

You must apply all the APARs for AIX 5.2 TL08 and later, or AIX 5.3 TL04 and later.

Dynamically enabling and disabling paths or adapters

This section describes how to dynamically enable and disable paths and adapters.

Dynamically enabling or disabling a path

There are three ways to dynamically enable (place online) or disable (place offline) a path:
1. Use the following **pcmpath** commands to change the path state:

   pcmpath set device M path N online

   or

   pcmpath set device M path N offline
2. Use the path control commands provided by AIX.

AIX 5.2 TL04 (or later) or AIX 5.3 TL01 (or later) provides several new path control commands. These commands can be used to configure or remove paths, change path state (enable or disable), and display the path current state.

Use the following AIX path command to change the path state:

chpath -l hdiskX -s E|D -p fscsiX -w ″**5005076300c99b0a,5200000000000000**″

Note: To find the parent adapter and connect location of a path, use the following AIX path command:

> **lspath -l hdiskX -H -F "***name path_id parent connection***"**

3. Use the smitty MPIO management submenu.

   a. Enter **smitty MPIO** and press **Enter**. This displays the MPIO Management panel.

   b. Select MPIO Path Management and press **Enter**. This displays the MPIO Path Management panel.

   c. Select Enable Paths or Disable paths to enable or disable paths.

Note: If you issue the **pcmpath set device M path N offline** command to a DS4000 or a DS5000 device and the path to be set *offline* is the last active path, it causes the device to fail over to the alternate controller. This is a time-consuming action, and might take a few seconds to complete; thus, after the **pcmpath set device M path N offline** command returns, the device might not have completed the controller failover yet. You must issue the **pcmpath query device M** command to check the result and to ensure that this set offline command completed successfully. If the command completed successfully, you can see that the active paths and passive paths are swapped. This is an asynchronous event.

Dynamically enabling or disabling an adapter

The SDDPCM pcmpath command can be used to enable (place online) or disable (place offline) an adapter.

To disable an adapter, use the following command:

pcmpath set adapter N offline

Notes:

1. Beginning with SDDPCM version 2.4.0.0, storage devices DS4000 and DS5000 are supported. These are active/passive storage devices, and behave differently than ESS, DS6000, DS8000 and SAN Volume Controller storage devices. To support the new devices, the syntax of the **pcmpath set adapter N online|offline** command changed by adding the extra parameter **aa|ap**. To set the adapter that connects to ESS, DS6000, DS8000, or SAN Volume Controller storage devices to online or offline , enter the **aa** parameter. To set the adapter that connects to DS4000, DS5000, or DS3950 storage devices to online or offline, enter the **ap** parameter.

2. SDDPCM reserves the last path of a device. This command will fail if there is any device using the last path attached to this adapter.

3. For DS4000 and DS5000 storage devices, this set adapter offline command might cause multiple devices to fail over to an alternate controller if the device path that is attached to this adapter is the last active path. Furthermore, if some devices are owned by controller A and some devices are owned by controller B, the set adapter offline command might cause some devices to fail over from controller A to controller B and some devices to fail over from controller B to

controller A. It is possible that the **pcmpath set adapter M offline** command returns a failure message, but some number of devices are failed over successfully and some devices are not. You must issue the **pcmpath query device** command to find out if the command is fully completed or partially completed.

Attention: If the AIX server has more than one DS4000, DS5000, or DS3950 device model configured, the **pcmpath set adapter N offline** command might cause multiple device models to fail over to the alternate controllers. If this is the case, the **pcmapth set adapter N offline** command is not supported.

Using the SDDPCM trace function

SDDPCM supports AIX trace function. The SDDPCM trace ID is 5A7. Trace ID 5A7 traces routine entry, exit, and error paths of the algorithm. To use it, manually turn on the trace function before the program starts to run, then turn off the trace function either after the program stops, or any time you need to read the trace report.

To start the trace function, enter:

trace -a -j 5A7

To stop the trace function, enter:

trcstop

To read the report, enter:

trcrpt | pg

To save the trace data to a file, enter:

trcrpt > filename

Note: To perform the SDDPCM trace function, you must have the bos.sysmgt.trace installation package installed on your system.

SDDPCM server daemon

The SDDPCM server (also referred to as *pcmsrv*) is an integrated component of SDDPCM 2.1.0.0 (or later). This component consists of a UNIX® application daemon that is installed in addition to the SDDPCM path control module. The SDDPCM server daemon provides path-recovery function for SDDPCM devices and the First Time Data Capture function. For detailed information about the SDDPCM server daemon, see "Using the SDD server and the SDDPCM server."

Verifying if the SDDPCM server has started

After you have installed SDDPCM and restarted the system, verify if the SDDPCM server (**pcmsrv**) has automatically started by entering **lssrc -s pcmsrv**. If **pcmsrv** has automatically started, the output from the **lssrc –s pcmsrv** command looks like this:

```
+-----------------------------------------------------------------+
| Subsystem   GROUP   PID  Status                                 |
| pcmsrv              NNN  Active                                  |
+-----------------------------------------------------------------+
```

where *NNN* is the process ID number.

The status of **pcmsrv** should be Active if the SDDPCM server has automatically started. If the SDDPCM server has not started, the status will be *Inoperative*. Go to "Starting the SDDPCM server manually" to proceed.

Because **pcmsrv** is bound to the SDDPCM kernel extension module, **pcmsrv** can fail to start if the SDDPCM is installed and the supported storage MPIO devices have not been configured yet. In this case, you can either restart the system or you can start **pcmsrv** manually after supported storage MPIO devices are configured.

Because **pcmsrv** is bound to the SDDPCM kernel extension module, in order to uninstall or upgrade SDDPCM, you must stop **pcmsrv** so that the SDDPCM kernel extension can be unloaded from the system. During an upgrade, the new SDDPCM kernel extension can be loaded into the system when supported storage MPIO devices are configured.

Starting the SDDPCM server manually

If **pcmsrv** did not start automatically after you performed the SDDPCM Installation and configured supported storage devices, you can start **pcmsrv** by entering:

startsrv -s pcmsrv -e XPG_SUS_ENV=ON

Go to "Verifying if the SDDPCM server has started" on page 45 to see if you successfully started the SDDPCM server.

Stopping the SDDPCM server

You can temporarily disable **pcmsrv** by entering:

stopsrc -s pcmsrv.

This will stop the current version of **pcmsrv**, but **pcmsrv** will start again if the system is restarted.

The SDDPCM server (**pcmsrv**) recovers failed paths for devices that have disabled the SDDPCM healthcheck function. Furthermore, SDDPCM server provides the First Time Data Capture function, which saves SDDPCM internal trace into a log file. You should not stop **pcmsrv** permanently. However, in certain cases when you want to permanently (not to start even after a system restart) disable **pcmsrv**, you must comment out the following line in the system init table (/etc/inittab):

`srv:2:wait:/usr/bin/startsrc -s pcmsrv -e XPG_SUS_ENV=ON > /dev/null 2>&1`

See "Using the SDD server and the SDDPCM server" for information about the following tasks:
- Enabling or disabling a TCP/IP port for SDDPCM server ("Enabling or disabling the sddsrv or pcmsrv TCP/IP port")
- Changing to a different port number for the SDDPCM server ("Changing the sddsrv or pcmsrv TCP/IP port number")

SDDPCM utility programs

This section describes the persistent reserve command tools, the SDDPCM pcmpath commands, and a tool to query DS4000 and DS5000 storage subsystem and MPIO device configuration information.

Persistent reserve command tools

Starting with SDDPCM 2.1.0.0, SDDPCM supports two persistent reserve command tools. The following sections describe the persistent reserve command tools.

pcmquerypr

The **pcmquerypr** command provides a set of persistent reserve functions. This command supports the following persistent reserve service actions:

- Read persistent reservation key
- Release persistent reserve
- Preempt-abort persistent reserve
- Clear persistent reserve and registration keys

This command can be issued to all system MPIO devices, including MPIO devices not supported by SDDPCM.

The **pcmquerypr** command can be used in the following situation: the reserve policy of the SDDPCM MPIO devices is set to either persistent reserve exclusive host access (PR_exclusive) or persistent reserve shared host access (PR_shared), and persistent reserve has been left on the device by a node, blocking access by another node. The **pcmquerypr** command can be used in this situation to query, preempt, or clear the persistent reserve left by a node or server on the devices.

There are more cases when you might need this tool to solve persistent reserve related problems, such as unexpected persistent reserve left on devices because of failing to release the persistent reserve. Caution must be taken with the command, especially when implementing preempt-abort or clear persistent reserve service action. With preempt-abort service action not only the current persistent reserve key is preempted; it also aborts tasks on the LUN that originated from the initiators that are registered with the preempted key. With clear service action, both persistent reservation and reservation key registrations are cleared from the device.

The following information describes in detail the syntax and examples of the **pcmquerypr** command.

Description

> The **pcmquerypr** command implements certain SCSI-3 persistent reservation commands on a device. The device can be a supported storage MPIO device. This command supports persistent reserve IN and OUT service actions, such as read, reservation key, release persistent reservation, preempt-abort persistent reservation, or clear persistent reservation and reservation key registrations.

Syntax

```
►►──pcmquerypr──┬──────┬──┬──────┬──┬──────┬──┬──────────────────┬──►◄
                ├─ -p ─┤  └─ -v ─┘  └─ -V ─┘  └─ -h /dev/PVname ──┘
                ├─ -c ─┤
                └─ -r ─┘
```

Flags:

-p   If the persistent reservation key on the device is different from the current host reservation key, the existing persistent reservation key on the device is preempted. This option can be issued only when the device is not already open.

-c   If there is a persistent reservation on the device, the persistent reservation is removed and all reservation key registrations on the device are cleared. This option can be issued only when the device is not already open.

-r   Removes the persistent reservation key on the device made by this host. This option can be issued only when the device is not already open.

-v   Displays the persistent reservation key if it exists on the device.

-V   Verbose mode. Prints detailed message.

Return code

If the command issued without options of -p, -r or -c, the command will return:

0   There is no persistent reservation key on the device, or the device is reserved by the current host

1   The persistent reservation key is different from the host reservation key

2   The command failed.

If the command issued with one of the options of -p, -r or -c, it returns:

0   The command was successful.

2   The command failed.

Examples

1. To query the persistent reservation on a device, enter **pcmquerypr -h /dev/hdisk30**.

   This command queries the persistent reservation on the device without displaying. If there is a persistent reserve on a disk, it returns 0 if the device is reserved by the current host. It returns 1 if the device is reserved by another host.

2. To query and display the persistent reservation on a device, enter **pcmquerypr -vh /dev/hdisk30**.

   Same as Example 1. In addition, it displays the persistent reservation key.

3. To query and display which type of persistent reservation is on a device, enter **pcmquerypr -Vh /dev/hdisk#**.

   The following output indicates there is SCSI-2 reserve on the device:

   ```
   # pcmquerypr -Vh /dev/hdisk27
   connection type: fscsi3
   open dev: /dev/hdisk27

   Attempt to read reservation key...
   *> ioctl(PR_READ) error; errno = 5 (I/O error)
   *> status_validity=0x1, scsi_bus_status=0x2

   Attempt to read reservation key...
   *> ioctl(PR_READ) error; errno = 5 (I/O error)
   *> status_validity=0x1, scsi_bus_status=0x18
   ```

```
                    Attempt to read reservation key...
                    *> ioctl(PR_READ) error; errno = 5 (I/O error)
                    *> status_validity=0x1, scsi_bus_status=0x18

                    Attempt to read reservation key...
                    *> ioctl(PR_READ) error; errno = 5 (I/O error)
                    *> status_validity=0x1, scsi_bus_status=0x18
```

The following output indicates that there is SCSI-3 reserve on the device:

```
# pcmquerypr -Vh /dev/hdisk43
connection type: fscsi0
open dev: /dev/hdisk43

Attempt to read reservation key...
*> ioctl(PR_READ) error; errno = 5 (I/O error)
*> status_validity=0x1, scsi_bus_status=0x2
Attempt to read reservation key...

Attempt to read registration keys...
Read Keys parameter
Generation : 12
Additional Length: 32
Key0 : 0x3236303232344446
Key1 : 0x3236303232344446
Key2 : 0x3236303232344446
Key3 : 0x3236303232344446
resrvpolicy= no_reserve
Reserve Key provided by current host = none (hex)0924ffff
Reserve Key on the device: 0x3236303232344446
Reservation key type: 0x6
Device is reserved by SDD device.
```

4. To release the persistent reservation if the device is reserved by the current host, enter **pcmquerypr -rh /dev/hdisk30**.

   This command releases the persistent reserve if the device is reserved by the current host. It returns 0 if the command succeeds or the device is not reserved. It returns 2 if the command fails.

5. To reset any persistent reserve and clear all reservation key registrations, enter **pcmquerypr -ch /dev/hdisk30**.

   This command resets any persistent reserve and clears all reservation key registrations on a device. It returns 0 if the command succeeds, or 2 if the command fails.

6. To remove the persistent reservation if the device is reserved by another host, enter **pcmquerypr -ph /dev/hdisk30**.

   This command removes an existing registration and persistent reserve from another host. It returns 0 if the command succeeds or if the device is not persistent reserved. It returns 2 if the command fails.

pcmgenprkey

Description

The **pcmgenprkey** command can be used to set or clear the PR_key_value ODM attribute for all SDDPCM MPIO devices. It also can be used to query and display the reservation policy of all SDDPCM MPIO devices and the persistent reserve key, if those devices have a PR key.

Syntax

```
►►──pcmgenprkey────────────────────────────────────────────────►◄
              │  ┌─ -v ──────────────────────┐ │
              └──┤                            ├─┘
                 └─ -u ─┘   └─ -k prkeyvalue ─┘
```
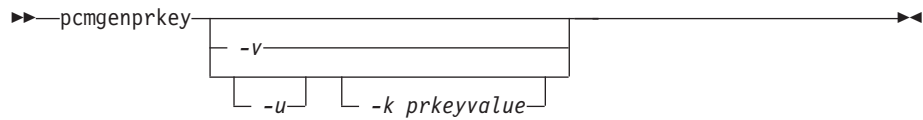
Examples

1. To set the persistent reserve key to all SDDPCM MPIO devices with a
   provided key value, issue **pcmgenprkey -u -k 0x1234567890abcedf**.
   This creates a customized PR_key_value attribute with the provided
   key value for all SDDPCM MPIO devices, except the devices that
   already have the same customized PR key attribute. The provided key
   must contain either a decimal integer or a hexadecimal integer.

2. To clear the PR_key_value attribute from all SDDPCM MPIO devices,
   issue **pcmgenprkey -u -k** *none* .

3. To update the customized PR_key_value attribute with the
   HACMP-provided Preserve key or the output string from the **uname**
   command for all the SDDPCM MPIO devices, issue **pcmgenprkey -u**.
   When the **-u** option is used without the **-k** option, this command
   searches for the HACMP-provided Preservekey attribute and uses that
   value as the PR key if that attribute is available; otherwise, it uses the
   output string from the **uname** command as the PR key.

4. To display the reserve_policy, the PR_key_value attribute, and the
   persistent reserve key attribute of all the SDDPCM devices, issue
   pcmgenprkey -v. If the MPIO device does not have a persistent reserve
   key, a value of *none* is displayed.

sddpcm_get_config

Description

The **sddpcm_get_config** command can be used to display information
about MPIO-based DS4000 or DS5000 subsystems and the hdisks
associated with them. Specifically, it displays information about the frame
(subsystem), including the frame's assigned name and worldwide name,
and a list of hdisks (only those currently in the *Available* state) that are
associated with that subsystem, including the hdisk name, LUN number,
current ownership, preferred path, and the user-assigned label for that
volume.

Attention: To protect the configuration database, the **sddpcm_get_config**
command is not interruptible, because stopping this command before it
completes could result in a corrupted database.

Syntax

```
►►──sddpcm_get_config──────────────────────────────────────────►◄
                        │  ┌─ -v ──────────────┐ │
                        └──┤                    ├─┘
                           └─ -l hdiskN ─┘ └─ -A ─┘
```
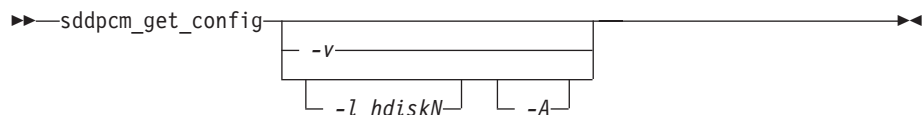
Flags

-l hdiskN

List information for the subsystem which includes hdiskN.

-A   List information for all attached subsystems.

-V List additional information, largely of limited value, including the MPIO SDDPCM internal frame number, number of controllers, partition number, and partition count.

Example

Display information about the frame of which hdisk4 is a member.

```
sddpcm_get_config -l hdisk4
```

The system displays a message similar to the following:

```
Storage Subsystem Name = 'SDD_DS5300'
         hdisk     LUN #   Ownership          User Label
         hdisk34      0   B (preferred)     ocotillop2_1
         hdisk35      1   A (preferred)      ocotillop2_2
         hdisk36      2   B (preferred)     ocotillop2_3
         hdisk37      3   A (preferred)     ocotillop2_4

--------------------------------------------------------------------
```

Using SDDPCM pcmpath commands

SDDPCM provides commands that you can use to display the status of adapters that are used to access managed devices, to display the status of devices that the device driver manages, or to map supported storage MPIO devices or paths to a supported storage device location. You can also set individual path conditions either to online or offline, set all paths that are connected to an adapter either to online or offline, or set all paths that are connected to a supported storage device port or ports to online or offline. This section includes descriptions of these commands. Table 3 provides an alphabetical list of these commands, a brief description, and where to go in this chapter for more information.

Table 3. Commands

Command	Description	Page
pcmpath clear device count	Dynamically clears the error count or error/select counts to zero.	53
pcmpath disable ports	Places paths connected to certain ports offline.	54
pcmpath enable ports	Places paths connected to certain ports online.	54
pcmpath open device path	Opens an INVALID path.	58
pcmpath query adapter	Displays information about adapters.	60
pcmpath query adaptstats	Displays performance information for all FCS adapters that are attached to SDDPCM devices.	62
pcmpath query device	Displays information about devices.	64
pcmpath query devstats	Displays performance information for a single SDDPCM device or all SDDPCM devices.	67
pcmpath query essmap	Displays each device, path, location, and attributes.	69
pcmpath query port	Displays information about a single target port or all target ports that are attached to SDDPCM-configured MPIO devices.	70
pcmpath query portmap	Displays the status of the logic paths that are managed by SDDPCM between the host and the storage ports.	72

Table 3. Commands (continued)

Command	Description	Page
pcmpath query portstats	Displays performance information about a single target port or all target ports that are attached to SDDPCM-configured MPIO devices.	73
pcmpath query version	Displays the version of the currently installed SDDPCM.	75
pcmpath query wwpn	Displays the world wide port name (WWPN) for all fibre-channel adapters.	76
pcmpath set adapter	Sets all device paths that are attached to an adapter to online or offline.	77
pcmpath set device path	Sets the path of a device to online or offline.	81
pcmpath set device algorithm	Set all or some of supported storage MPIO device path selection algorithm	78
pcmpath set device hc_interval	Set all or some of supported storage MPIO device health check time interval	79
pcmpath set device hc_mode	Set all or some of supported storage MPIO device health check mode	80

pcmpath clear device count

The **pcmpath clear device count** command dynamically sets the error counter or error/select counters to zero.

Syntax:

```
►►──pcmpath clear─device number 1─┬─────────────────┬─count─┬─error─┬───────►◄
                                  └─device number 2─┘       └─all───┘
```

Parameters:

device number 1 <device number 2>
> When two device numbers are entered, this command applies to all the devices whose index numbers fit within the range of these two device index numbers.

error
> Clears the error counter of the specified SDDPCM MPIO device or devices.

all Clears both the select counter and the error counter of the specified SDDPCM MPIO device or devices.

Examples: If you have a non-zero select counter or error counter, entering **pcmpath query device 20** causes the following output to be displayed:

```
DEV#:  20  DEVICE NAME: hdisk20  TYPE: 2145  ALGORITHM:  Load Balance
SERIAL: 60050768018180235800000000000463
============================================================================
Path#         Adapter/Path Name     State     Mode     Select      Errors
   0            fscsi1/path0         CLOSE     NORMAL       14           0
   1*           fscsi1/path1         CLOSE     NORMAL        8           0
   2            fscsi3/path2         CLOSE     NORMAL    10009           0
   3*           fscsi3/path3         CLOSE     NORMAL        8           0
```

If you enter the **pcmpath clear device 20 count all** and then enter **pcmpath query device 20**, the following output is displayed:

```
DEV#:  20  DEVICE NAME: hdisk20  TYPE: 2145  ALGORITHM:  Load Balance
SERIAL: 60050768018180235800000000000463
============================================================================
Path#         Adapter/Path Name     State     Mode     Select      Errors
   0            fscsi1/path0         CLOSE     NORMAL        0           0
   1*           fscsi1/path1         CLOSE     NORMAL        0           0
   2            fscsi3/path2         CLOSE     NORMAL        0           0
   3*           fscsi3/path3         CLOSE     NORMAL        0           0
```

pcmpath disable ports

The **pcmpath disable ports** command sets SDDPCM MPIO device paths offline for the specified supported storage system location code.

Syntax:

```
►►──pcmpath disable ports─connection─ess essid──────────────────────►◄
```

Parameters:

connection

    The connection code must be in one of the following formats:

- Single port = R1-Bx-Hy-Zz
- All ports on card = R1-Bx-Hy
- All ports on bay = R1-Bx

    Use the output of the **pcmpath query essmap** command to determine the connection code.

essid

    The supported storage device serial number, given by the output of **pcmpath query portmap** command.

Examples: If you enter the **pcmpath disable ports R1-B1-H3 ess 12028** command and then enter the **pcmpath query device** command, the following output is displayed:

```
DEV#:   3 DEVICE NAME: hdisk3  TYPE: 2105E20  ALGORITHM:  Load Balance
SERIAL: 20712028
==========================================================================
Path#         Adapter/Path Name     State     Mode      Select      Errors
    0             fscsi0/path0       CLOSE     OFFLINE        6           0
    1             fscsi0/path1       CLOSE     NORMAL         9           0
    2             fscsi1/path2       CLOSE     OFFLINE       11           0
    3             fscsi1/path3       CLOSE     NORMAL         9           0

DEV#:   4 DEVICE NAME: hdisk4  TYPE: 2105E20  ALGORITHM:  Load Balance
SERIAL: 20712028
==========================================================================
Path#         Adapter/Path Name     State     Mode      Select      Errors
    0             fscsi0/path0       CLOSE     OFFLINE     8702           0
    1             fscsi0/path1       CLOSE     NORMAL      8800           0
    2             fscsi1/path2       CLOSE     OFFLINE     8816           0
    3             fscsi1/path3       CLOSE     NORMAL      8644           0

DEV#:   5 DEVICE NAME: hdisk5  TYPE: 2105E20  ALGORITHM:  Load Balance
SERIAL: 20912028
==========================================================================
Path#         Adapter/Path Name     State     Mode      Select      Errors
    0             fscsi0/path0       CLOSE     OFFLINE     8917           0
    1             fscsi0/path1       CLOSE     NORMAL      8919           0
    2             fscsi1/path2       CLOSE     OFFLINE     9008           0
    3             fscsi1/path3       CLOSE     NORMAL      8944           0

DEV#:   6 DEVICE NAME: hdisk6  TYPE: 2105E20  ALGORITHM:  Load Balance
SERIAL: 20B12028
==========================================================================
Path#         Adapter/Path Name     State     Mode      Select      Errors
    0             fscsi0/path0       CLOSE     OFFLINE     9044           0
    1             fscsi0/path1       CLOSE     NORMAL      9084           0
    2             fscsi1/path2       CLOSE     OFFLINE     9048           0
    3             fscsi1/path3       CLOSE     NORMAL      8851           0
```

```
DEV#:   7  DEVICE NAME: hdisk7  TYPE: 2105E20  ALGORITHM:  Load Balance
SERIAL: 20F12028
============================================================================
Path#        Adapter/Path Name     State     Mode      Select      Errors
    0         fscsi0/path0          CLOSE     OFFLINE     9089           0
    1         fscsi0/path1          CLOSE     NORMAL      9238           0
    2         fscsi1/path2          CLOSE     OFFLINE     9132           0
    3         fscsi1/path3          CLOSE     NORMAL      9294           0

DEV#:   8  DEVICE NAME: hdisk8  TYPE: 2105E20  ALGORITHM:  Load Balance
SERIAL: 21012028
============================================================================
Path#        Adapter/Path Name     State     Mode      Select      Errors
    0         fscsi0/path0          CLOSE     OFFLINE     9059           0
    1         fscsi0/path1          CLOSE     NORMAL      9121           0
    2         fscsi1/path2          CLOSE     OFFLINE     9143           0
    3         fscsi1/path3          CLOSE     NORMAL      9073           0
```

pcmpath enable ports

The **pcmpath enable ports** command sets MPIO device paths online for the specified supported storage device location code.

Syntax:

```
►►──pcmpath enable ports─connection─ess essid───────────────────────────────────◄
```

Parameters:

connection

> The connection code must be in one of the following formats:
> - Single port = R1-Bx-Hy-Zz
> - All ports on card = R1-Bx-Hy
> - All ports on bay = R1-Bx
>
> Use the output of the **pcmpath query essmap** command to determine the connection code.

essid

> The supported storage device serial number, given by the output of **pcmpath query portmap** command.

Examples: If you enter the **pcmpath enable ports R1-B1-H3 ess 12028** command and then enter the **pcmpath query device** command, the following output is displayed:

```
DEV#:   3  DEVICE NAME: hdisk3  TYPE: 2105E20  ALGORITHM:  Load Balance
SERIAL: 20112028
==========================================================================
Path#          Adapter/Path Name     State     Mode     Select     Errors
    0            fscsi0/path0         CLOSE    NORMAL        6          0
    1            fscsi0/path1         CLOSE    NORMAL        9          0
    2            fscsi1/path2         CLOSE    NORMAL       11          0
    3            fscsi1/path3         CLOSE    NORMAL        9          0


DEV#:   4  DEVICE NAME: hdisk4  TYPE: 2105E20  ALGORITHM:  Load Balance
SERIAL: 20712028
==========================================================================
Path#          Adapter/Path Name     State     Mode     Select     Errors
    0            fscsi0/path0         CLOSE    NORMAL     8702          0
    1            fscsi0/path1         CLOSE    NORMAL     8800          0
    2            fscsi1/path2         CLOSE    NORMAL     8816          0
    3            fscsi1/path3         CLOSE    NORMAL     8644          0

DEV#:   5  DEVICE NAME: hdisk5  TYPE: 2105E20  ALGORITHM:  Load Balance
SERIAL: 20912028
==========================================================================
Path#          Adapter/Path Name     State     Mode     Select     Errors
    0            fscsi0/path0         CLOSE    NORMAL     8917          0
    1            fscsi0/path1         CLOSE    NORMAL     8919          0
    2            fscsi1/path2         CLOSE    NORMAL     9008          0
    3            fscsi1/path3         CLOSE    NORMAL     8944          0

DEV#:   6  DEVICE NAME: hdisk6  TYPE: 2105E20  ALGORITHM:  Load Balance
SERIAL: 20B12028
==========================================================================
Path#          Adapter/Path Name     State     Mode     Select     Errors
    0            fscsi0/path0         CLOSE    NORMAL     9044          0
    1            fscsi0/path1         CLOSE    NORMAL     9084          0
    2            fscsi1/path2         CLOSE    NORMAL     9048          0
    3            fscsi1/path3         CLOSE    NORMAL     8851          0
```

```
DEV#:   7  DEVICE NAME: hdisk7  TYPE: 2105E20  ALGORITHM:  Load Balance
SERIAL: 20F12028
===========================================================================
Path#          Adapter/Path Name     State    Mode     Select     Errors
    0             fscsi0/path0        CLOSE    NORMAL     9089          0
    1             fscsi0/path1        CLOSE    NORMAL     9238          0
    2             fscsi1/path2        CLOSE    NORMAL     9132          0
    3             fscsi1/path3        CLOSE    NORMAL     9294          0

DEV#:   8  DEVICE NAME: hdisk8  TYPE: 2105E20  ALGORITHM:  Load Balance
SERIAL: 21012028
===========================================================================
Path#          Adapter/Path Name     State    Mode     Select     Errors
    0             fscsi0/path0        CLOSE    NORMAL     9059          0
    1             fscsi0/path1        CLOSE    NORMAL     9121          0
    2             fscsi1/path2        CLOSE    NORMAL     9143          0
    3             fscsi1/path3        CLOSE    NORMAL     9073          0
```

pcmpath open device path

The **pcmpath open device path** command dynamically opens a path that is in the `invalid` state. You can use this command to open an invalid path even when I/O is actively running on the devices.

Syntax:

```
►►──pcmpath open device─device number─path─path number─────────────────────────►◄
```

Parameters:

device number
> The logical device number of this hdisk, as displayed by the **pcmpath query device** command.

path number
> The path ID that you want to change, as displayed under "Path Name" by the **pcmpath query device** command.

Examples: If you enter the **pcmpath query device 23** command, the following output is displayed:

```
DEV#:  23 DEVICE NAME: hdisk23 TYPE: 2105E20 ALGORITHM: Load Balance
SERIAL: 20112028
==============================================================================
Path#         Adapter/Path Name      State     Mode     Select     Errors
    0            fscsi1/path0         OPEN      NORMAL      557          0
    1            fscsi1/path1         OPEN      NORMAL      568          0
    2            fscsi0/path2         INVALID   NORMAL        0          0
    3            fscsi0/path3         INVALID   NORMAL        0          0
```

Note that the current state of path 2 and path 3 is INVALID, which means that open path 2 and path 3 failed.

If the root cause of the path 2 open failure is fixed and you enter the **pcmpath open device 23 path 2** command, the following output is displayed:

```
Success: device 23 path 2 opened

DEV#:  23  DEVICE NAME: hdisk23 TYPE: 2105E20 ALGORITHM: Load Balance
SERIAL: 20112028
==============================================================================
Path#         Adapter/Path Name      State     Mode     Select     Errors
    0            fscsi1/path0         OPEN      NORMAL      557          0
    1            fscsi1/path1         OPEN      NORMAL      568          0
    2            fscsi0/path2         OPEN      NORMAL        0          0
    3            fscsi0/path3         INVALID   NORMAL        0          0
```

After issuing the **pcmpath open device 23 path 2** command, the state of path 2 becomes OPEN.

The terms used in the output are defined as follows:

Dev#   The logical device number of this hdisk.

Device name
> The name of this device.

Type   The device product ID from inquiry data.

Algorithm
>   The current path selection algorithm for the device. The algorithm selected
>   is one of the following types: load balancing, load balancing port, round
>   robin, or failover.

Serial   The LUN for this device.

Path#   The path index displayed by the **pcmpath query device** command.

Adapter
>   The name of the adapter to which the path is attached.

Path Name
>   The name of the path. The number displayed as part of the name is the
>   path ID of this path that is used by the **pcmpath open device path** and
>   **pcmpath set device path** commands.

State   The condition of each path of the named device:
>   **Open**   Path is in use.
>   **Close**   Path is not being used.
>   **Close_Failed**   Path is broken and is not being used.
>   **Failed**   Path is opened, but no longer functional because of error.
>   **Invalid**   The path failed to open.

Mode   The mode of the named path, which is either *Normal* or *Offline*.

Select   The number of times this path was selected for I/O.

Errors   The number of I/O errors that occurred on this path.

pcmpath query adapter

The **pcmpath query adapter** command displays information about a single adapter or all adapters that are attached to SDDPCM-configured MPIO devices.

Syntax:

```
►►──pcmpath query adapter─adapter number─┬──────┬─────────────────────────►◄
                                         ├─aa─┤
                                         └─ap─┘
```

Parameters:

adapter number
> The index number of the adapter for which you want information displayed. If you do not enter an adapter index number, information about all adapters is displayed.

aa   The adapter of active/active storage controller devices.

ap   The adapter of active/passive storage controller devices.

Examples: If you enter the **pcmpath query adapter** command and your system has both Dual Active or Active/Asymmetrc (for example, ESS) and Active/Passive (for example, DS4800) devices, the following output is displayed:

Total Dual Active and Active/Asymmetrc Adapters : 2

Adpt#	Name	State	Mode	Select	Errors	Paths	Active
0	fscsi2	NORMAL	ACTIVE	920506	0	80	38
1	fscsi0	NORMAL	ACTIVE	921100	0	80	38

Total Active/Passive Adapters : 2

Adpt#	Name	State	Mode	Select	Errors	Paths	Active
0	fscsi0	NORMAL	ACTIVE	0	0	30	0
1	fscsi1	NORMAL	ACTIVE	0	0	30	0

The terms used in the output are defined as follows:

Adpt #
> The index number of the adapter.

Name   The name of the adapter.

State   The condition of the named adapter. It can be either:
> **Normal**   Adapter is in use.
> **Degraded**   One or more opened paths are not functioning.
> **Failed**   All opened paths that are attached to this adapter are not functioning.

Mode   The mode of the named adapter, which is either *Active* or *Offline*.

Select   The number of times this adapter was selected for I/O.

Errors   The number of errors that occurred on all paths that are attached to this adapter.

Paths   The number of paths that are attached to this adapter.

Active   The number of functional paths that are attached to this adapter. The

number of functional paths is equal to the number of opened paths attached to this adapter minus any that are identified as failed or disabled (offline).

pcmpath query adaptstats

The **pcmpath query adaptstats** command displays information about a single or all fibre-channel adapters that are attached to SDDPCM-configured MPIO devices. If you do not enter an adapter number, information about all adapters is displayed.

Syntax:

```
►►──pcmpath query adaptstats─adapter number─┬──────────┬──────────────►◄
                                            ├──aa──┤
                                            └──ap──┘
```

Parameters:

adapter number
> The index number of the adapter for which you want information displayed. If you do not enter an adapter index number, information about all adapters is displayed.

aa The adapter of active/active storage controller devices.

ap The adapter of active/passive storage controller devices.

Examples: If you enter the **pcmpath query adaptstats** command and your system only has Active/Passive devices (for example, DS4800), the following output is displayed:

```
Total Active/Passive Adapters : 2
Adapter #:  0
=============
             Total Read  Total Write  Active Read  Active Write   Maximum
I/O:                  0            0            0             0         0
SECTOR:               0            0            0             0         0
Adapter #:  1
=============
             Total Read  Total Write  Active Read  Active Write   Maximum
I/O:                  0            0            0             0         0
SECTOR:               0            0            0             0         0


/*-----------------------------------------------------------------------*/
```

The terms used in the output are defined as follows:

Total Read
- I/O: total number of completed read requests
- SECTOR: total number of sectors that have been read

Total Write
- I/O: total number of completed write requests
- SECTOR: total number of sectors that have been written

Active Read
- I/O: total number of read requests in process
- SECTOR: total number of sectors to read in process

Active Write
- I/O: total number of write requests in process
- SECTOR: total number of sectors to write in process

Maximum

- I/O: the maximum number of queued I/O requests
- SECTOR: the maximum number of queued sectors to Read or Write

pcmpath query device

The **pcmpath query device** command displays information about a single SDDPCM MPIO device, a set of SDDPCM MPIO devices between device M and N, a specific model of SDDPCM MPIO devices, or all SDDPCM MPIO devices. If two device numbers are entered, the command displays device information for a set of devices between these two device numbers. The option to specify a model cannot be used when you query with device numbers. If you do not enter a device number, information about all devices is displayed. If a device number is entered, the command displays the device information about the hdisk that is associated with this number.

Beginning with SDDPCM 2.1.0.7, a new option, -d, is provided to display devices of a particular device model. The valid device models are:

- 1750 - All 1750 models (DS6000)
- 2105 - All 2105 models (ESS)
- 2107 - All 2107 models (DS8000)
- 2145 - All 2145 models (SAN Volume Controller)

Beginning with SDDPCM 2.4.0.0, the following device models are added for option -d:

- 1722 - All 1722 devices (DS4300)
- 1724 - All 1724 devices (DS4100)
- 1742 - All 1742 devices (DS4400 and DS4500)
- 1814 - All 1814 devices (DS4200 and DS4700)
- 1815 - All 1815 devices (DS4800)
- 1818 - All 1818 devices (DS5100 and DS5300)

Beginning with SDDPCM 2.1.0.7, a new feature is available to verify SAN configurations in a controller environment, such as DS6000 or SAN Volume Controller. A nonpreferred path will be marked with "*".

For example, if you configure 4 paths per MPIO hdisk device and you want to configure an equal distribution between preferred controller and nonpreferred controller of DS6000 devices, you can configure your environment with 2 paths connected to a preferred controller and 2 paths connected to a nonpreferred controller. This feature will help you verify your configuration by indicating which path is connected to the nonpreferred controller before the device starts operation.

Note: You must open SDDPCM MPIO devices of DS6000 and SAN Volume Controller once before the '*' mark on nonpreferred paths can be displayed in the following three commands:

- **pcmpath query device**
- **pcmpath query essmap**
- **pcmpath query portmap**

For DS4000 and DS5000 SDDPCM MPIO devices, only opened passive paths are displayed with the '*' mark in the pcmpath query device command.

Beginning with SDDPCM 2.1.3.0, two new options are added to the **device query** command. The first option lets you specify two numbers to query a set of devices; the second option *-i x y* lets you repeat the query command every *x* seconds for *y* times.

The **pcmpath query device** commands display only supported storage MPIO devices that are configured with the SDDPCM module. Any AIX internal disks or non-SDDPCM-configured MPIO devices are not displayed.

Syntax:
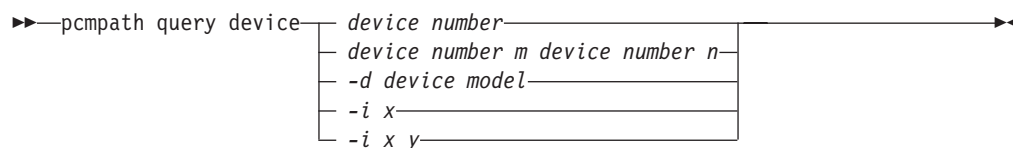
```
►►──pcmpath query device──┬─ device number ─────────────────────┬──────────────►◄
                          ├─ device number m device number n ─┤
                          ├─ -d device model ─────────────────┤
                          ├─ -i x ─────────────────────────────┤
                          └─ -i x y ───────────────────────────┘
```

Parameters:

device number
> The device number refers to the logical device number of the hdisk.

device number_m device_number_n
> Use the *device_number_m device_number_n* option to provide a range of device index numbers.

device model
> Displays devices of a particular device model. The valid device models are:
> - 1722 - All 1722 devices (DS4300)
> - 1724 - All 1724 devices (DS4100)
> - 1742 - All 1742 devices (DS4400 and DS4500)
> - 1750 - All 1750 models (DS6000)
> - 1814 - All 1814 devices (DS4200 and DS4700)
> - 1815 - All 1815 devices (DS4800)
> - 1818 - All 1818 devices (DS5100 and DS5300)
> - 2105 - All 2105 models (ESS)
> - 2107 - All 2107 models (DS8000)
> - 2145 - All 2145 models (SAN Volume Controller)

i   Repeats the command every *x* seconds for *y* times. If you do not specify *y*, the command repeats indefinitely every *x* seconds.

Examples: If you enter the **pcmpath query device 65 66** command, the following output is displayed:

For the supported storage devices:
```
DEV#:  65  DEVICE NAME: hdisk65  TYPE: 1814      ALGORITHM:  Load Balance
SERIAL: 600A0B800011796C0000B68148283C9C
==========================================================================
Path#         Adapter/Path Name      State     Mode      Select      Errors
   0*            fscsi0/path0         OPEN      NORMAL        42           0
   1*            fscsi1/path2         OPEN      NORMAL        44           0
   2             fscsi0/path1         OPEN      NORMAL        54           0
   3             fscsi1/path3         OPEN      NORMAL        52           0

DEV#:  66  DEVICE NAME: hdisk66  TYPE: 1814      ALGORITHM:  Load Balance
SERIAL: 600A0B80001179440000C1FE48283A08
==========================================================================
Path#         Adapter/Path Name      State     Mode      Select      Errors
   0             fscsi0/path0         CLOSE     NORMAL        44           1
```

```
1              fscsi1/path2        CLOSE    NORMAL       57              0
2              fscsi0/path1        CLOSE    NORMAL        0              0
3              fscsi1/path3        CLOSE    NORMAL        1              0
```

The terms used in the output are defined as follows:

Dev#   The logical device number of this hdisk.

Name   The logical name of this device.

Type   The device product ID from inquiry data.

Algorithm
>   The current path selection algorithm selected for the device. The algorithm selected is one of the following: load balancing, load balancing port, round robin, or failover.

Serial   The LUN for this device.

Path#   The path index displayed by device query command.

Adapter
>   The name of the adapter to which the path is attached.

Path Name
>   The name of the path. The number displayed as part of the name is the path ID that is used by **pcmpath open device path** and **pcmpath set device path** commands.

State   The condition of the path attached to the named device:
>   **Open**   Path is in use.
>   **Close**   Path is not being used.
>   **Failed** Path is no longer being used. It has been removed from service due to errors.
>   **Close_Failed**
>   >   Path was detected to be broken and failed to open when the device was opened. The path stays in Close_Failed state when the device is closed.
>   **Invalid**
>   >   The path is failed to open, but the MPIO device is opened.

Mode   The mode of the named path. The mode can be either *Normal* or *Offline*.

Select   The number of times this path was selected for I/O.

Errors   The number of input and output errors that occurred on a path of this device.

pcmpath query devstats

The **pcmpath query devstats** command displays performance information for a single SDDPCM MPIO device or all SDDPCM MPIO devices. If you do not enter a device number, information about all devices is displayed. If a device number is entered, the command displays the device information about the hdisk that is associated with this number.

Note: The **pcmpath query devstats** command displays only SDDPCM MPIO devices. Any AIX internal disks or non-SDDPCM configured MPIO devices will not be displayed.

Beginning with SDDPCM 2.1.3.0, two new options are added to the **query devstats** command. The first option lets you specify two numbers to query a set of devices; the second option -*i x y* lets you repeat the query command every *x* seconds for *y* times.

Syntax:

```
►►──pcmpath query devstats──┬─ device number ─────────────────┬──────────────►◄
                            ├─ device number m device number n ─┤
                            ├─ -d device model ─────────────────┤
                            ├─ -i x ─────────────────────────────┤
                            └─ -i x y ───────────────────────────┘
```

Parameters:

device number
>   The device number refers to the logical device number of the hdisk.

device number_m device_number_n
>   Use the *device_number_m device_number_n* option to provide a range of device index numbers.

device model
>   Displays devices of a particular device model. The valid device models are:
>   - 1722 - All 1722 devices (DS4300)
>   - 1724 - All 1724 devices (DS4100)
>   - 1742 - All 1742 devices (DS4400 and DS4500)
>   - 1750 - All 1750 models (DS6000)
>   - 1814 - All 1814 devices (DS4200 and DS4700)
>   - 1815 - All 1815 devices (DS4800)
>   - 2105 - All 2105 models (ESS)
>   - 2107 - All 2107 models (DS8000)
>   - 2145 - All 2145 models (SAN Volume Controller)

i
>   Repeats the command every *x* seconds for *y* times. If you do not specify *y*, the command repeats indefinitely every *x* seconds.

Examples: If you enter the **pcmpath query devstats 2** command, the following output about hdisk2 is displayed:

```
DEV#:   2  DEVICE NAME: hdisk2
=================================
            Total Read  Total Write  Active Read  Active Write   Maximum
I/O:               60           10            0             0          2
SECTOR:           320            0            0             0         16
```

```
Transfer Size:      <= 512        <= 4k        <= 16K        <= 64K       > 64K
                      30            40            0             0           0
/*---------------------------------------------------------------------------*/
```

The terms used in the output are defined as follows:

Total Read

- I/O: total number of completed read requests
- SECTOR: total number of sectors that have been read

Total Write

- I/O: total number of completed write requests
- SECTOR: total number of sectors that have been written

Active Read

- I/O: total number of read requests in process
- SECTOR: total number of sectors to read in process

Active Write

- I/O: total number of write requests in process
- SECTOR: total number of sectors to write in process

Maximum

- I/O: the maximum number of queued I/O requests
- SECTOR: the maximum number of queued sectors to read or write

Transfer size

- <= 512: the number of I/O requests received, whose transfer size is 512 bytes or less
- <= 4k: the number of I/O requests received, whose transfer size is 4 KB or less (where KB equals 1024 bytes)
- <= 16K: the number of I/O requests received, whose transfer size is 16 KB or less (where KB equals 1024 bytes)
- <= 64K: the number of I/O requests received, whose transfer size is 64 KB or less (where KB equals 1024 bytes)
- > 64K: the number of I/O requests received, whose transfer size is greater than 64 KB (where KB equals 1024 bytes)

pcmpath query essmap

The **pcmpath query essmap** command displays data collected from the fcppcmmap program after the system has been restarted, or by the pcmpath program if any device configuration change occurs after the system has been restarted. The fcppcmmap program issues SCSI commands directly to the disks to collect the information that is displayed. This command is supported for disk storage systems only.

Syntax:

```
►►──pcmpath query essmap─────────────────────────────────────────►◄
```

Examples: If you enter the **pcmpath query essmap** command, the following output is displayed:

```
Disk    Path  P Location    adapter   LUN SN       Type         Size LSS  Vol  Rank  C/A  S ...
------  ----  - ----------  ------    ------ ----- -----------  ---- ---- ---  ----- ---- - ...
hdisk5  path0 * 30-60-01[FC] fscsi1   13AAAKA1200 IBM 1750-500  1.1  18   0    0000  01   Y ...
hdisk5  path1   30-60-01[FC] fscsi0   13AAAKA1200 IBM 1750-500  1.1  18   0    0000  01   Y ...
hdisk5  path2 * 20-60-01[FC] fscsi0   13AAAKA1200 IBM 1750-500  1.1  18   0    0000  01   Y ...
hdisk5  path3   20-60-01[FC] fscsi1   13AAAKA1200 IBM 1750-500  1.1  18   0    0000  01   Y ...
```

The terms used in the output are defined as follows:

Disk	The logical device name assigned by the host.
Path	The logical path name of a MPIO device.
P	Indicates the logical paths and whether the path is preferred and nonpreferred. "*"indicates that the path is a nonpreferred path.
Location	The physical location code of the host adapter through which the LUN is accessed.
Adapter	The logical adapter name assigned by the host LUN.
LUN SN	The unique serial number for each LUN within the supported storage device.
Type	The device and model.
Size	The capacity of the configured LUN.
LSS	The logical subsystem where the LUN resides. (Beginning with 2.1.3.0, the value displayed is changed from decimal to hexadecimal.)
Vol	The volume number within the LSS.
Rank	The unique identifier for each RAID array within the supported storage device.
C/A	The cluster and adapter accessing the array.
S	Indicates that the device is shared by two and more supported storage device ports. Valid values are *yes* or *no*.
Connection	The physical location code of supported storage device adapter through which the LUN is accessed.
Port	The supported storage device port through which the LUN is accessed.
RaidMode	The disk RAID mode.

pcmpath query port

The **pcmpath query port** command displays information about a single target port or all target ports that are attached to SDDPCM-configured MPIO devices.

Syntax:

```
►►──pcmpath query port─target port number───────────────────────────────────►◄
```

Parameters:

target port number
> Use the target port number option to display information about the target port. If you do not enter a target port number, information about all target ports is displayed.

Examples: If you have 12 active ports and enter the **pcmpath query port** command, the following output is displayed:

```
Port#         Wwpn    State    Mode       Select    Errors  Paths  Active
   0  5005022300ca031e  NORMAL  ACTIVE          0         0     60       0
   1  5005022307044750  NORMAL  ACTIVE          0         0     64       0
   2  5005022307155750  NORMAL  ACTIVE          0         0     64       0
   3  500502230e02ac54  NORMAL  ACTIVE          0         0     64       0
   4  500502230e01fe54  NORMAL  ACTIVE          0         0     64       0
   5  500502230eb3d54   NORMAL  ACTIVE          0         0     64       0
   6  500503380110046b  NORMAL  ACTIVE          0         0     60       0
   7  5005033801100d64  NORMAL  ACTIVE          0         0     60       0
   8  50050338011003d1  NORMAL  ACTIVE          0         0     60       0
   9  5005033801100118  NORMAL  ACTIVE          0         0     60       0
  10  500503330323006a  NORMAL  ACTIVE       3167         0     62       4
  11  500503330323406a  NORMAL  ACTIVE       3220         1     62       4
```

The terms used in the output are defined as follows:

Port # The index number of the port.

WWPN The worldwide port name of the target port.

State The condition of the named target port, which can be one of the following types:

> **Normal**
> > The target port that is in use.

> **Degraded**
> > One or more of the opened paths are not functioning.

> **Failed** All opened paths that are attached to this target port are not functioning.

Mode The mode of the named target port, which is either *Active* or *Offline*.

Select The number of times this target port was selected for input or output.

Errors The number of errors that occurred on all paths that are attached to this target port.

Paths The number of paths that are attached to this target port.

Active The number of functional paths that are attached to this target port. The number of functional paths is equal to the number of opened paths that are attached to this target port minus any that are identified as failed or disabled (offline).

Note: This command is supported only with 1750, 2105, 2107, and 2145 device types.

pcmpath query portmap

The **pcmpath query portmap** command displays the status of the logic paths that are managed by SDDPCM between the host and the storage ports. This command is supported for disk storage systems only.

Syntax:

```
►►──pcmpath query portmap──────────────────────────────────────────────►◄
```

Examples: If you enter the **pcmpath query portmap** command, the following output is displayed:

```
                BAY-1(B1)              BAY-2(B2)              BAY-3(B3)              BAY-4(B4)
ESSID   DISK    H1   H2   H3   H4      H1   H2   H3   H4      H1   H2   H3   H4      H1   H2   H3   H4
                ABCD ABCD ABCD ABCD    ABCD ABCD ABCD ABCD    ABCD ABCD ABCD ABCD    ABCD ABCD ABCD ABCD
                BAY-5(B5)              BAY-6(B6)              BAY-7(B7)              BAY-8(B8)
                H1   H2   H3   H4      H1   H2   H3   H4      H1   H2   H3   H4      H1   H2   H3   H4
                ABCD ABCD ABCD ABCD    ABCD ABCD ABCD ABCD    ABCD ABCD ABCD ABCD    ABCD ABCD ABCD ABCD

    13AAAKA hdisk5 O--- ---- ---- ----    o--- ---- ---- ----    ---- ---- ---- ----    ---- ---- ---- ----
    13AAAKA hdisk6 Y--- ---- ---- ----    y--- ---- ---- ----    ---- ---- ---- ----    ---- ---- ---- ----

Y  =  online/open           y = (alternate path) online/open
O  =  online/closed         o = (alternate path) online/closed
N  =  offline               n = (alternate path) offline
-  =  path not configured
PD =  path down
```

The terms used in the output are defined as follows:

Y	The port is online and open, meaning that at least one path attached to this port is functional.
y	Paths connected to this port are nonpreferred paths. The port is online and open, meaning that at least one path attached to this port is functional.
O	The port is online and closed, meaning that at least one path state and mode is closed and online.
o	Paths connected to this port are nonpreferred paths. The port is online and closed, meaning that at least one path state and mode is closed and online.
N	The port is offline, meaning that all paths attached to this port are offline.
n	Paths connected to this port are nonpreferred paths. The port is offline, meaning that all paths attached to this port are offline.
-	The path is not configured.
PD	The path is down. It is either not functional or has been placed offline.

Note: The following fields apply only 1750 devices and can only be shown after the device is opened once:

- y
- o
- n

The serial number of ESS devices is five digits, whereas the serial number of DS6000 and DS8000 devices is seven digits.

pcmpath query portstats

The **pcmpath query portstats** command displays performance information about a single target port or all target ports that are attached to SDDPCM-configured MPIO devices. If you do not enter a device number, information about all devices is displayed.

Syntax:

►►──pcmpath query portstats─*target port number*─────────────────────────────►◄

Parameters:

target port number
> Use the target port number option to display information about the target port. If you do not enter a target port number, information about all target ports is displayed.

Examples: If you have four target ports and enter the **pcmpath query portstats** command, the following output is displayed:

```
Port #:  0
=============
          Total Read  Total Write  Active Read  Active Write   Maximum
I/O:              0            0            0             0           0
SECTOR:           0            0            0             0           0

Port #:  1
=============
          Total Read  Total Write  Active Read  Active Write   Maximum
I/O:              0            0            0             0           0
SECTOR:           0            0            0             0           0

Port #: 2
=============
          Total Read  Total Write  Active Read  Active Write   Maximum
I/O:           3109           58            0             0           2
SECTOR:       28292         4084            0             0        2096

Port #: 3
=============
          Total Read  Total Write  Active Read  Active Write   Maximum
I/O:           3169           51            0             0           2
SECTOR:       24762         6470            0             0        2096
```

The terms used in the output are defined as follows:

Total Read
> - I/O: The total number of completed read requests
> - SECTOR: The total number of sectors that have been read

Total Write
> - I/O: The total number of completed write requests
> - SECTOR: The total number of sectors that have been written

Active Read
> - I/O: The total number of read requests in process
> - SECTOR: The total number of sectors to read in process

Active Write
> - I/O: The total number of write requests in process
> - SECTOR: The total number of sectors to write in process.

Maximum
> - I/O: The maximum number of queued I/O requests
> - SECTOR: The maximum number of queued sectors to Read or Write

Notes:

1. This command is supported only with 1750, 2105, 2107, and 2145 device types.

2. Data that is displayed by this command is collected only when the device's algorithm is set to lbp. For example, if the algorithm from hdisk10 to hdisk20 is set to lbp, the statistical data for each device is saved in the associated ports and displayed here. If none of the device's algorithm is set to lbp, there is no change in the port statistical output.

pcmpath query version

The **pcmpath query version** command displays the version of SDDPCM that is installed.

Syntax:

►►──pcmpath query version──►◄

Parameters: None

Examples: If you enter the **pcmpath query version** command, the following output is displayed:

```
[root@abc]>    pcmpath query version
IBM SDDPCM Version 2.1.1.0 (devices.sddpcm.52.rte)
```

pcmpath query wwpn

The **pcmpath query wwpn** command displays the World Wide Port Name
(WWPN) of the host fibre-channel adapters.

Syntax:

```
►►──pcmpath query wwpn─────────────────────────────────────────────►◄
```

Parameters: None

Examples: If you enter the **pcmpath query wwpn** command, the following output
is displayed:

```
Adapter Name    PortWWN
fscsi0          10000000C925F5B0
fscsi1          10000000C9266FD1
```

pcmpath set adapter

The **pcmpath set adapter** command sets all device paths attached to an adapter either to online or offline. SDDPCM reserves the last path of a device if the device is opened. This command will fail if there is any device that has the last path attached to this adapter. If all paths of a device are attached to a single fibre-channel adapter that connects to multiple storage ports through a switch, the **pcmpath set adapter 0 offline** command will fail, because the device will lose all its paths.

Attention: If the device is closed, the **pcmpath set adapter offline** command does not have the last path offline protection.

Note: For DS4000 and DS5000 storage devices, this set adapter offline command might cause multiple devices to fail over to an alternate controller if the device path that is attached to this adapter is the last active path. Furthermore, if some devices are owned by controller A and some devices are owned by controller B, the set adapter offline command might cause some devices to fail over from controller A to controller B and some devices to fail over from controller B to controller A. It is possible that the **pcmpath set adapter M offline** command returns a failure message, but some number of devices are failed over successfully and some devices are not. You must issue the pcmpath query device command to find out if the command is fully completed or partially completed.

Attention: If the AIX server has more than one DS4000 or DS5000 device model configured, the **pcmpath set adapter N offline** command might cause multiple device models to fail over to the alternate controller. In this case, the **pcmapth set adapter N offline** command is not supported.

Syntax:

```
►►──pcmpath set adapter─adapter number─┬─ online ──┬──┬─ —aa ─┬─────────────►◄
                                        └─ offline ─┘  └─ —ap ─┘
```

Parameters:

adapter number
    The index number of the adapter that you want to change.

online
    Enables the adapter for service.

offline
    Disables the adapter from service.

aa The adapter of active/active storage controller devices.

ap The adapter of active/passive storage controller devices.

Examples: If you enter the **pcmpath set adapter 0 offline ap** command:
- Adapter 0 of the active/passive controller devices changes to *Offline* mode and, if there are some paths in the opened state, its state might change to *failed*.
- All paths of the active/passive controller devices that are attached to adapter 0 change to *Offline* mode and their states change to *Dead*, if they were in the *Open* state.

pcmpath set device algorithm

The **pcmpath set device algorithm** command dynamically changes the path selection algorithm per MPIO device.

Syntax:

```
►►──pcmpath set device── num1 ──────────────── algorithm─option ──────────────────►◄
                                └─ num2 ─┘
```

Note: If the device reserve policy is set to single_path (SCSI-2 reserve), the device algorithm must be set to fail_over. Any attempt to set the algorithm to round_robin, load_balance, or load_balance_port with single_path reserve policy will fail.

Parameters:

num1 [num2]

- When only *num1* is specified, the command applies to the hdisk specified by *num1*.
- When two device logical numbers are entered, this command applies to all the devices whose logical numbers fit within the range of the two device logical numbers.

option
Specifies one of the following path selection algorithms:

- **rr**, where *rr* indicates round robin
- **lb**, where *lb* indicates load balancing
- **fo**, where *fo* indicates failover policy
- **lbp**, where *lbp* indicates load balancing port

Notes:

1. You can enter the **pcmpath set device** *N* **algorithm** *rr/fo/lb/lbp* command to dynamically change to the path selection algorithm associated with SDDPCM MPIO devices that are in either Close or Open state.
2. Beginning with SDDPCM 2.4.0.0, the algorithm *lbp* incorporates I/O statistics from both host adapters and target ports in the path selection algorithm. This new algorithm is applicable only for device models 1750, 2105, 2107, and 2145.

Examples: If you enter **pcmpath set device 2 10 algorithm rr**, the path-selection algorithm of hdisk 2 to hdisk 10 is immediately changed to the round robin algorithm.

You can also use the **chdev** command to change the path selection algorithm of a device: **chdev -l hdiskX -a algorithm=load_balance_port**

pcmpath set device hc_interval

The **pcmpath set device hc_interval** command dynamically changes the health check time interval for SDDPCM MPIO devices or disables the health check function of a device.

Note: Starting with SDDPCM v2123, when you issue this command to change the SDDPCM internal health check function for one or more devices from *enable* to *disable* or from *disable* to *enable*, the SDDPCM server daemon (pcmsrv) automatically starts or stops the health check function on the devices.

Syntax:

```
►►──pcmpath set device── num1──────────────hc_interval─t──────────────────►◄
                              └─ num2 ─┘
```

Parameters:

num1 [num2]

- When only *num1* is specified, the command applies to the hdisk specified by *num1*.
- When 2 device logical numbers are entered, this command applies to all the devices whose logical numbers fit within the range of the two device logical numbers.

t   The range of supported values for health check interval is 1-3600 seconds. To disable the health check function of a device, set interval time to 0.

Examples: If you enter **pcmpath set device 2 10 hc_interval 30**, the health check time interval of hdisk2 to hdisk10 is immediately changed to 30 seconds.

pcmpath set device hc_mode

The **pcmpath set device hc_mode** command dynamically changes the MPIO device health check mode.

Syntax:

```
►►──pcmpath set device── num1────────────hc_mode─option──────────────────────►◄
                              └─ num2 ─┘
```

Parameters:

num1 [num2]

- When only *num1* is specified, the command applies to the hdisk specified by *num1*.
- When 2 device logical numbers are entered, this command applies to all the devices whose logical numbers fit within the range of the two device logical numbers.

option

Specifies one of the following policies:

- **enabled**, indicates the health check command will be sent to paths that are opened with a normal path mode.
- **failed**, indicates the health check command will be sent to paths that are in failed state.
- **nonactive**, indicates the health check command will be sent to paths that have no active I/O. This includes paths that are opened or in failed state.

Examples: If you enter **pcmpath set device 2 10 hc_mode enabled**, the health check mode of MPIO hdisk2 to hdisk10 is immediately changed to the enabled mode.

pcmpath set device path

The **pcmpath set device path** command sets the path of a device either to *online* or *offline*. You cannot remove the last path to an opened device from service. This prevents a data access failure from occurring.

Attention: If the device is closed, the **pcmpath set adapter offline** command does not have the last path offline protection.

Syntax:

```
►►──pcmpath set device─device number─path─path-ID─┬─ online ──┬─────────────►◄
                                                   └─ offline ─┘
```

Parameters:

device number
    The logical device number of the hdisk.

path ID
    The path ID that you want to change, as displayed under "Path Name" by the **pcmpath query device** command.

online
    Enables the path for service.

offline
    Disables the path from service.

Examples: If you enter the **pcmpath set device 5 path 0 offline** command, path 0 for device 5 changes to Offline mode.

Summary of command syntax

SDDPCM supports the following **pcmpath** commands:

- **pcmpath clear device <n>/(<m> <n>) count error/all**
- **pcmpath query adapter [<n> aa|ap]**
- **pcmpath query adaptstats [<n> aa|ap]**
- **pcmpath query device<n>/[<m> <n>] [-d <device_model>] [-i <x>/(<x> <y>)]**
- **pcmpath query devstats <n>/[<m> <n>] [-d <device_model>] [-i <x>/(<x> <y>)]**
- **pcmpath set adapter n online|offline aa|ap**
- **pcmpath set device M path N online|offline**
- **pcmpath set device <n1> [n2] algorithm <option>**
- **pcmpath set device <n1> [n2] hc_interval <t>**
- **pcmpath set device <n1> [n2] hc_mode <option>**
- **pcmpath disable port <location> ess <essid>**
- **pcmpath enable port <location> ess <essid>**
- **pcmpath open device <m> path <n>**
- **pcmpath query essmap**
- **pcmpath query port <target port number>**
- **pcmpath query portmap**
- **pcmpath query portstats <target port number>**
- **pcmpath query wwpn**
- **pcmpath query version**

Note: If the commands are used for a device, the *n* is the number of the device logical name. For example, **pcmpath query devstats 3** queries the device statistics for hdisk3.

If the commands are used for adapter, the *n* is the index of the adapter. For example, **pcmpath query adapter 2** queries the adapter statistics for the third adapter in adapter list order, which can be fscsi5.

Chapter 4. Using the SDD on an HP-UX host system

This chapter provides step-by-step procedures for you to install, configure, remove, and use SDD on a Hewlett-Packard (HP-UX) host system that is attached to supported storage devices.

For updated and additional information that is not included in this manual, the Readme file on the CD-ROM or go to SDD Web site:

www.ibm.com/servers/storage/support/software/sdd

Verifying the hardware and software requirements

You must install the following hardware and software components to ensure that SDD installs and operates successfully.

Hardware

The following hardware components are needed:

- One or more of the supported storage devices.
- For ESS devices: at least one SCSI host adapter (two are required for load balancing and failover)

  To install SDD and use the input/output (I/O) load balancing and failover features, you need a minimum of two SCSI or fibre-channel adapters.

  A host system with a single fibre-channel adapter that connects through a switch to multiple ESS ports is considered to have multiple fibre-channel SDD vpath devices.

  For information on the fibre-channel adapters that can be used on your HP-UX host system go to:

  www.ibm.com/servers/storage/support

- A SCSI cable to connect each SCSI host adapter to a storage system controller port
- Subsystem LUNs that have been created and confirmed for multiport access
- A fiber-optic cable to connect each fibre-channel adapter to a supported storage device port

Software

SDD supports certain HP-UX kernel levels.

SDD-supported HP-UX kernel levels

SDD supports the following HP-UX kernel levels:

- HP-UX 11.11 (11i) 32-bit and 64-bit
- HP-UX 11.23 (11i v2) PA_RISC 64-bit and IA 64-bit

Ensure that the patches listed in Table 5 on page 86 or newer patches are installed on your HP-UX host system.

Unsupported environments

SDD does not support the following environments:

- HP-UX 11.0 32-bit kernel
- HP-UX 11.0 64-bit kernel
- A system start from an SDD pseudo device
- A system paging file on an SDD pseudo device
- A host system with both a SCSI and fibre-channel connection to a shared LUN
- Single-path mode during concurrent download of licensed machine code nor during any disk storage system concurrent maintenance that impacts the path attachment, such as a disk storage system-host-bay-adapter replacement
- Single-path configuration for fibre channel
- DS8000 and DS6000 with SCSI connectivity

Understanding how SDD works on an HP-UX host system

SDD resides above the HP SCSI disk driver (sdisk) in the protocol stack. For more information about how SDD works, "The SDD architecture."

Preparing for the SDD installation

Before you install the SDD, you must configure the supported storage devices to your host system and attach the required SCSI or fibre-channel adapters. A maximum of 1200 LUNs can be configured by the SDD. The HP-UX SDD supports up to 32 paths per LUN. Be aware that there is no performance gain with a large number of LUNs that are smaller in size. The larger the configuration, the more system resources are required.

Note: 1200 LUNs are supported only by HP-UX 11.23 (PA-RISC 64 and IA64). 600 LUNs are supported by HP-UX 11.11.

Configuring the disk storage system

Before you install SDD, configure your ESS for single-port or multiport access for each LUN. SDD requires a minimum of two independent paths that share the same logical unit to use the load balancing and path-failover-protection features. **With a single path, failover protection is not provided.**

For information about configuring your ESS, the *IBM TotalStorage Enterprise Storage Server Introduction and Planning Guide*.

Configuring the virtualization products

Before you install SDD, configure your SAN Volume Controller devices for single-port or multiport access for each LUN. SDD requires a minimum of two independent paths that share the same logical unit to use the load balancing and path-failover-protection features. **With a single path, failover protection is not provided.**

For information about configuring your SAN Volume Controller, the *IBM System Storage SAN Volume Controller Software Installation and Configuration Guide*.

Configuring DS4000 and DS5000 family disk systems

Before you install the SDD, follow the instructions in the *IBM System Storage DS4x00 Installation, User's and Maintenance Guide* and the *IBM System Storage*

DS4000 Storage Manager Concepts Guide for configuring LUNs that are attached to the HP-UX host system. The SDD requires a minimum of two independent paths that share the same logical unit to use load balance and path failover features. With a single path, failover protection is not provided.

The DS4000 and DS5000 controllers can be set to two different modes: AVT or non-AVT but can operate only in one mode per storage partition. The controller mode is determined by the host type, which is predefined in the DS4000 or DS5000 Storage Manager profile. Only the AVT mode is supported for the HP-UX host type.

To ensure that a single path failure does not result in controller failover during recovery, configure redundant paths to each controller. Because switching to another controller affects performance, configuring the redundant path to a controller can avoid unnecessary controller failover that is caused by a path failure.

Bypassing the boot disk check

The SDD installation process checks the boot disk and then excludes the bootable disk from the SDD configuration. This process attempts to read the boot record from every device instance of the system, so the process can be lengthy if your configuration is large or if the devices are unable to satisfy the I/O request.

If the system boot disk is not an SDD-supported disk, you can bypass the process that checks the boot disk. To do this, enter the following before the SDD installation:

touch /opt/IBMsdd/BYPASS_BOOTDISK_CHECK

You might want to bypass boot disk checking for systems with DS4000 or DS5000 devices, because reading the boot record from every device instance of DS4000 or DS5000 can be very time consuming.

Planning for installation

Before you install SDD on your HP-UX host system, you need to understand what kind of software runs on your host. The way you install SDD depends on the kind of software that you are running. Two types of special device files are supported:
- Block device files
- Character device files

There are three possible scenarios for installing SDD. The scenario you choose depends on the kind of software that you have installed:

Table 4 further describes the various installation scenarios and how you should proceed. Table 5 on page 86 lists the patches necessary for proper operation of SDD on HP-UX .

Table 4. SDD installation scenarios

Installation scenario	Description	How to proceed

Table 4. SDD installation scenarios (continued)

Scenario 1	• SDD is not installed. • The SDD server for Expert is installed. • No software application or DBMS communicates directly to sdisk interface.	Go to: 1. "Determining if the SDD 1.3.1.5 (or later) server for Expert is installed" 2. "Installing SDD" 3. "Standard Unix applications"
Scenario 2	• SDD is not installed. • The SDD server for Expert is installed. • An existing application package or DBMS communicates directly to the sdisk interface.	Go to: 1. "Determining if the SDD 1.3.1.5 (or later) server for Expert is installed" 2. "Installing SDD" 3. "Standard Unix applications"
Scenario 3	• SDD is installed. • The SDD server for Expert is installed.	Go to 1. "Determining if the SDD 1.3.1.5 (or later) server for Expert is installed" 2. "Upgrading the SDD"

Table 5. Patches necessary for proper operation of SDD on HP-UX

HP-UX	Patch bundles
11.23	March 06, standard patch bundles
11.11	September 05, support plus

http://itrc.hp.com for patch details and prerequisites for patches.

Determining if the SDD 1.3.1.5 (or later) server for Expert is installed

If you previously installed the SDD server (the stand-alone version) for IBM TotalStorage Expert V2R1 (ESS Expert) on your HP-UX host system, you must remove this stand-alone version of the SDD server before you proceed with SDD 1.3.1.5 installation. The installation package for SDD 1.3.1.5 includes the SDD server daemon (also referred to as sddsrv), which incorporates the functionality of the stand-alone version of the SDD server (for ESS Expert).

To determine if the stand-alone version of the SDD server is installed on your host system, enter:

swlist SDDsrv

If you previously installed the stand-alone version of the SDD server, the output from the **swlist SDDsrv** command looks similar to this:

```
 SDDsrv  1.0.0.0  SDDsrv bb-bit Version: 1.0.0.0 Nov-14-2001 15:34
```

Notes:

1. The installation package for the stand-alone version of the SDD server (for ESS Expert) is SDDsrvHP*bb*_*yymmdd*.depot (where *bb* represents 32- or 64-bit, and *yymmdd* represents date of installation package). For ESS Expert V2R1, the stand-alone SDD server installation package is SDDsrvHP32_020115.depot for a 32-bit environment, and SDDsrvHP64_020115.depot for a 64-bit environment.)

2. For instructions on how to remove the stand-alone version of the SDD server (for ESS Expert) from your HP-UX host system, the IBM SUBSYSTEM DEVICE DRIVER SERVER 1.0.0.0 (sddsrv) readme for IBM TotalStorage Expert V2R1 at the following Web site:

www.ibm.com/servers/storage/support/software/swexpert/

For more information about the SDD server daemon, go to "SDD Server Daemon."

Installing SDD

Before you install SDD, make sure that you have root access to your HP-UX host system and that all the required hardware and software is ready.

Installing SDD from CD-ROM

Perform the following steps to install SDD on your HP-UX host system:

1. Make sure that the SDD compact disc (CD) is available.
2. Insert the CD-ROM into your CD-ROM drive.

   If CD-ROM directory is not known:

   a. Run `ioscan -funC disk`.
   b. Record the CD-ROM or DVD-ROM directory.
3. Mount the CD-ROM drive using the **mount** command. Here are two examples of the **mount** command:

   mount /dev/dsk/c0t2d0 */cdrom*

   or

   mount /dev/dsk/c0t2d0 */your_installation_directory*

   where */cdrom* or */your_installation_directory* is the name of the directory to which you want to mount the CD-ROM drive.
4. Run the **sam** program.

   > **sam**
5. Select **Software Management**.
6. Select **Install Software to Local Host**.
7. At this point, the **SD Install - Software Selection** panel is displayed. Almost immediately afterwards, a Specify Source menu is displayed:

   a. For **Source Depot Type**, select the local CD-ROM.
   b. For **Source Depot Path**, choose the directory and the IBMsdd.depot file.

      For 32-bit mode applications, use:

      /cdrom/hp32bit/IBMsdd.depot

      or

      /your_installation_directory/hp32bit/IBMsdd.depot

      For 64-bit mode applications, use:

      /cdrom/hp64bit/IBMsdd.depot

      or

      /your_installation_directory/hp64bit/IBMsdd.depot

   c. Click **OK**.

      You will output similar to the following example:

```
IBMsdd_tag 1.7.0.3 IBMsdd Driver 64-bit Version: 1.7.0.3 Sep-24-2007 16:35
```

8. Click the **IBMsdd_tag** product.

9. From the Bar menu, click **Actions** → **Mark for Install**.

10. From the Bar menu, click **Actions** → **Install (analysis)**. An Install Analysis panel is displayed, showing the status of **Ready**.

11. Click **OK** to proceed. A Confirmation window opens and states that the installation will begin.

12. Click **Yes** and press **Enter**. The analysis phase starts.

13. After the analysis phase has finished, another Confirmation window opens informing you that the system will be restarted after installation is complete. Click **Yes** and press **Enter**. The installation of IBMsdd will now proceed.

14. An Install window opens, informing you about the progress of the IBMsdd software installation. The window looks similar to the following:

```
Press 'Product Summary' and/or 'Logfile' for more target information.
Target          : XXXXX
Status          : Executing install setup
Percent Complete    : 17%
Kbytes Installed    :  276 of 1393
Time Left (minutes) : 1
Product Summary      Logfile
Done                              Help
```

The **Done** option is not available when the installation is in progress. It becomes available after the installation process is complete.

15. Click **Done**.

Note: SDD 1.5.0.4 is changed from a static driver to a dynamic loadable kernel module (DLKM) driver. The system does not restart after SDD is installed.

After the installation is finished, the SDD driver is automatically loaded.

You can use the **datapath query device** command to verify the SDD installation. SDD is successfully installed if the command runs successfully.

Installing SDD from downloaded code

The SDD download Web site, www.ibm.com/servers/storage/support/software/sdd, has several SDD packages for different versions of HP-UX. The SDD packages are in tar archives. After you have downloaded the appropriate SDD package into your chosen directory, you can untar the file using the **tar xvf IBMsdd*.tar** command. The perform the following steps to install SDD:

1. Run the sam program.

   > **sam**

2. Select **Software Management**.

3. Select **Install Software to Local Host**.

4. At this point, the SD Install - Software Selection panel is displayed. Then a Specify Source menu is displayed:

   a. Select the Local Directory for **Source Depot Type**.

   b. Select the directory in which you have issued the **tar xvf IBMsdd*.tar** command to untar the file and the IBMsdd.depot file for the **Source Depot Path**. Use the fully-qualified path name for the depot file as shown below.

```
/your_installation_directory/IBMsdd.depot
```

   c. Click **OK**. Output similar to the following example is displayed:

```
Name     Revision        Information                    Size(Kb) Architecture    Category
IBMsdd_tag -> B.11.23.01 IBMsdd Driver 64-bit <version> <date>  nnnn      HPUX_B.11.23_ia64
```

5. Click the **IBMsdd_tag** product and perform the steps beginning with step 9 shown in "Installing SDD from CD-ROM."

Upgrading the SDD

Use these sections to upgrade SDD:

Upgrading from SDD 1.6.0.*x* to SDD 1.6.1.0 or later with concurrent access

The memory management and the installation process have been enhanced to allow installation of the SDD package while the LVM volume groups are active and user applications are running. The concurrent driver upgrade function permits uninterrupted operation when installing SDD.

The installation process:

1. Converts SDD vpath devices to PVLINK devices
2. Unloads and reloads the SDD driver
3. Converts the PVLINK devices back to SDD vpath devices after the new package is installed.

Because the volume groups must be active for the PVLINK conversion process, the following are the limitations:

1. The volume groups must be managed by HP-UX LVM.
2. The MC Service Guard cluster must be halted prior to upgrade. The primary node and the adoptive node or nodes must operate in a single-host environment. The shared volume groups in the adoptive nodes must be exported so that the volumes are not shared; the volume groups can be active in the primary node only. Restore the cluster environment after upgrading SDD.

Performance during upgrade: You should consider the following performance topics while you are upgrading:

- The PVLINK conversion process and the driver reload require additional system resources such as LVM lock, accessing LVM meta data and the kernel memory. With the concurrent I/O, the upgrade process can take longer because the conversion process must wait for the I/O to complete before a link can be removed from PVLINK.
- Reloading the SDD driver can also take longer because of the contention with the kernel memory; the system must wait for a window when the resources become available. The actual time for installation depends on the processor model, physical memory size, I/O intensity, and configuration size. The larger the SDD configuration or the more concurrent I/O activities, the longer it can take to upgrade. The installation time can also take longer if the devices from the ioscan output are not accessible. If there were a lot of inaccessible devices as the result of fabric reconfiguration, you should attempt to clean up the configuration before upgrading.
- After the upgrade, you should check the VPATH_EVENT for allocation failures in syslog.log, /var/adm/IBMsdd/hd2vp.errlog and vp2hd.errlog. These are the indications that the upper limit of the resources have been reached during the conversion process and that you should take a more conservative approach next

time. That is, the concurrent upgrade should be performed during a period when the system load is lighter than the normal operation.

- The installation process also ensures that the current SDD state is not in any degraded state; the recovery process can be lengthy if the upgrade failed due to the hardware errors. Issue the **swjob** command that is indicated at the end of swinstall output to get detailed information about the installation.

- The diagnose message in the package installation and configuration process has been greatly improved to include logs for **cfgvpath**, **vp2hd**, **hd2vp** and the syslog messages. All the SDD related logs have been moved to the /var/adm/IBMsdd directory.

Upgrading from SDD 1.5.0.4 to SDD 1.6.1.0 or later with nonconcurrent access.

Upgrading SDD consists of removing and reinstalling the IBMsdd package. If you are upgrading SDD, go to "Uninstalling SDD" and then go to "Installing SDD."

Upgrading SDD 1.3.0.2 or earlier to SDD 1.5.0.4 or later

SDD 1.3.0.2 (or earlier) could have created a vpath0 device. vpath0 should not be configured. To upgrade your SDD, you need to ensure that you have no vpath0 devices using the following steps:

1. Remove its /etc/vpathsave.cfg (vpath name reservation file) file.
2. Issue **cfgvpath -c** to reconfigure SDD.

To upgrade from SDD 1.3.0.2 (or earlier) to SDD 1.5.0.4 (or later):

1. Uninstall SDD 1.3.0.2 (or earlier). Because SDD 1.3.0.2 (or earlier) does not have the volume group conversion scripts **hd2vp** and **vp2hd**, the uninstallation will not convert the existing volume groups back to cXtXdX devices.
2. Install SDD 1.5.0.4 (or later) on your host.
3. After installation, check to that you have:
   a. The volume group conversion scripts **hd2vp** and **vp2hd** in /opt/IBMdpo/bin
   b. The /etc/vpathsave.cfg file
4. Use the **vp2hd** script to convert volume groups from volume groups containing SDD vpath devices to volume groups containing cXtXdX devices.
5. When the volume groups have been converted:
   a. Remove the /etc/vpathsave.cfg file.
   b. Issue **cfgvpath -c** to reboot.
6. When the host responds, use the **showvpath** command to determine that vpath0 has been removed. If vpath0 no longer exists and the state of all the SDD vpath devices looks good in the **datapath query device** output, use the **hd2vp** script to change volume groups to volume groups containing SDD vpath devices.

   Note: Vpathname vpathN is reserved when it is assigned to a LUN even after the LUN has been removed from the host. The same vpathname, vpathN, will be assigned to the same LUN when it is reconnected to the host.

7. /etc/vpathsave.cfg is the file to reserve vpathnames. Improper removal of the file will invalidate existing volume groups. Do not remove the /etc/vpathsave.cfg file.

Installing an additional package for DS4000 and DS5000 support

The HP SDD support for DS4000 and DS5000 is enabled through the On-Demand-Device function, which recognizes the new devices as a plug-in to the driver. You can install the On-Demand package before or after you install the traditional SDD package.

Note: DS5000 support is limited to HP-UX 11.23 systems, including both IA64 and PA-RISC architectures.

If the On-Demand package is installed later, run **cfgvpath –r** to recognize and reconfigure the DS4000 or DS5000 device. If the On-Demand package was installed earlier, the installation process automatically recognizes and configures the DS4000 and DS5000 devices.

To install the DS On-Demand-Device package, issue the following command:

For DS4000: **swinstall IBMsdd_ds4k_***platform***.depot IBMsdd_ds4k_tag**

For DS5000: **swinstall IBMsdd_ds5k_***platform***.depot IBMsdd_ds5k_tag**

Configuring the SDD

This section provides information necessary to configure the SDD. Use the HP command line interface (CLI) to manage SDD devices.

Maximum number of LUNs

A maximum of 1200 LUNs can be configured by SDD. HP-UX SDD supports up to 32 paths per LUN.

Changing an SDD hardware configuration

When you are adding or removing multiport SCSI devices, you must reconfigure SDD to recognize the new devices. Perform the following steps to reconfigure SDD:

1. Issue the **cfgvpath** command to reconfigure the SDD vpath device by entering:

   /opt/IBMsdd/bin/cfgvpath

2. Restart the system by entering:

   shutdown -r 0

   Note: Restarting the system is no longer required for SDD 1.5.0.4 or later.

The **querysn** command can be used to list all disk storage system devices visible to the host. The **querysn** command reads the unique serial number of a disk storage system device (sdisk). To manually exclude devices from the SDD configuration, their serial number information can be included in the /etc/vpathmanualexcl.cfg text file. For bootable devices, the **get_root_disks** command generates a file called /etc/vpathexcl.cfg to exclude bootable disks from the SDD configuration.

Converting a volume group

SDD provides the following conversion scripts:

hd2vp The **hd2vp** script converts a volume group from supported storage device sdisks into SDD vpath devices.

The syntax for **hd2vp** script is as follows:

```
hd2vp vgname
```

vp2hd The **vp2hd** script converts a volume group from SDD vpath devices into supported storage device sdisks. Use the **vp2hd** program when you want to configure your applications back to original supported storage device sdisks.

The syntax for **vp2hd** script is as follows:

```
vp2hd vgname
```

hd2vp and **vp2hd** convert a volume group from an sdisk pvlink to and from SDD vpaths devices. The **hd2vp** conversion program is invoked at the system boot time to ensure the conversion of sdisks to SDD vpath devices. Beginning with SDD 1.6.0.12, **vp2hd** is not invoked at the system shutdown.

Dynamic reconfiguration

Dynamic reconfiguration provides a way to automatically detect path configuration changes without requiring a reboot.

1. **cfgvpath -r**:

   This operation finds the current hardware configuration and compares it to the SDD vpath device configuration in memory and then identifies a list of differences. It then issues commands to update the SDD vpath device configuration in memory with the current hardware configuration. The commands that **cfgvpath -r** issues to the vpath driver are:

   a. Add an SDD vpath device.
   b. Remove an SDD vpath device; this will fail if device is busy.
   c. Add path to the SDD vpath device.
   d. Remove path from the SDD vpath device; this will fail deletion of the path if the device is busy, but will set path to DEAD and OFFLINE.

2. **rmvpath** command removes one or more SDD vpath devices.

   ```
   rmvpath -all                # Remove all SDD vpath devices

   rmvpath vpath_name          # Remove one SDD vpath device at a time
                               # this will fail if device is busy
   ```

Dynamically changing the SDD path-selection policy algorithm

SDD 1.4.0.0 (or later) supports path-selection policies that increase the performance of multipath-configured supported storage devices and make path failures transparent to applications. The following path-selection policies are supported:

failover only (fo)
All I/O operations for the device are sent to the same (preferred) path until the path fails because of I/O errors. Then an alternate path is chosen for subsequent I/O operations.

load balancing (lb)
The path to use for an I/O operation is chosen by estimating the load on the adapter to which each path is attached. The load is a function of the number of I/O operations currently in process. If multiple paths have the same load, a path is chosen at random from those paths. Load balancing mode also incorporates failover protection.

Note: The load balancing policy is also known as the optimized policy.

round robin (rr)
> The path to use for each I/O operation is chosen at random from those paths that were not used for the last I/O operation. If a device has only two paths, SDD alternates between the two.

The path-selection policy is set at the SDD device level. The default path-selection policy for an SDD device is load balancing. You can change the policy for an SDD device. SDD version 1.4.0.0 (or later) supports dynamic changing of the SDD devices path-selection policy.

Before changing the path-selection policy, determine the active policy for the device. Enter **datapath query device** *N* where *N* is the device number of the SDD vpath device to show the current active policy for that device.

datapath set device policy command

Use the **datapath set device policy** command to change the SDD path-selection policy dynamically:

See "datapath set device policy" for more information about the **datapath set device policy** command.

Preferred node path-selection algorithm for the virtualization products

The virtualization products are two-controller disk subsystems. SDD distinguishes the paths to a virtualization product LUN as follows:

1. Paths on the preferred controller
2. Paths on the alternate controller

When SDD selects paths for I/O, preference is always given to a path on the preferred controller. Therefore, in the selection algorithm, an initial attempt is made to select a path on the preferred controller. Only if no path can be used on the preferred controller will a path be selected on the alternate controller. This means that SDD will automatically fail back to the preferred controller any time a path on the preferred controller becomes available during either manual or automatic recovery. Paths on the alternate controller are selected at random. If an error occurs and a path retry is required, retry paths are first selected on the preferred controller. If all retries fail on the preferred controller's paths, paths on the alternate controller will be selected for retry. The following is the path selection algorithm for SDD:

1. With all paths available, I/O is only routed to paths on the preferred controller.
2. If no path on the preferred controller is available, I/O fails over to the alternate controller.
3. When failover to the alternate controller has occurred, if a path on the preferred controller is made available, I/O automatically fails back to the preferred controller.

SDD datapath query adapter command changes for SDD 1.4.0.0 (or later)

For SDD 1.4.0.0 (or later), the output of some of the datapath commands has changed. See "Using the datapath commands" for details about the datapath commands.

For SDD 1.3.3.11 (or earlier), the output of the **datapath query adapter** command shows all the fibre-channel arrays as different adapters, and you need to determine which hardware paths relate to which adapters. If you need to place an adapter offline, you need to manually issue multiple commands to remove all the associated hardware paths.

For SDD 1.4.0.0 (or later), the output of the **datapath query adapter** command has been simplified.

The following examples show the output resulting from the **datapath query adapter** command for the same configuration for SDD 1.3.3.11 (or earlier) and for SDD 1.4.0.0 (or later).

Example output from **datapath query adapter** command issued in SDD 1.3.3.11 (or earlier):

```
Active Adapters :8
Adapter#          Adapter Name  State   Mode    Select  Error Path Active
0          0/7/0/0.4.18.0.38  NORMAL  ACTIVE       0      0   1     1
1          0/4/0/0.4.18.0.38  NORMAL  ACTIVE       0      0   1     1
2          0/7/0/0.4.18.0.36  NORMAL  ACTIVE       0      0   2     2
3          0/4/0/0.4.18.0.36  NORMAL  ACTIVE       0      0   2     2
4          0/7/0/0.4.18.0.34  NORMAL  ACTIVE       0      0   2     2
5          0/4/0/0.4.18.0.34  NORMAL  ACTIVE       0      0   2     2
6          0/7/0/0.4.18.0.32  NORMAL  ACTIVE       0      0   1     1
7          0/4/0/0.4.18.0.32  NORMAL  ACTIVE       0      0   1     1
```

Adapter #s 0, 2, 4, 6 belong to the same physical adapter. In order to place this adapter offline, you need to issue **datapath set adapter offline** four times. After the four commands are issued, the output of **datapath query adapter** will be:

```
Active Adapters :8
Adapter#          Adapter Name  State   Mode    Select  Error Path Active
0          0/7/0/0.4.18.0.38  NORMAL  OFFLINE      0      0   1     0
1          0/4/0/0.4.18.0.38  NORMAL  ACTIVE       0      0   1     0
2          0/7/0/0.4.18.0.36  NORMAL  OFFLINE      0      0   2     0
3          0/4/0/0.4.18.0.36  NORMAL  ACTIVE       0      0   2     0
4          0/7/0/0.4.18.0.34  NORMAL  OFFLINE      0      0   2     0
5          0/4/0/0.4.18.0.34  NORMAL  ACTIVE       0      0   2     0
6          0/7/0/0.4.18.0.32  NORMAL  OFFLINE      0      0   1     0
7          0/4/0/0.4.18.0.32  NORMAL  ACTIVE       0      0   1     0
```

Example output from **datapath query adapter** command issued in SDD 1.4.0.0 (or later):

```
Active Adapters :2
Adapter#          Adapter Name  State   Mode    Select  Error Path Active
0                    0/7/0/0  NORMAL  ACTIVE       0      0   6     6
1                    0/4/0/0  NORMAL  ACTIVE       0      0   6     6
```

Adapters 0 and 1 represent two physical adapters. To place one of the adapters offline, you need to issue one single command, for example, **datapath set adapter 0 offline**. After the command is issued, the output of **datapath query adapter** will be:

```
Active Adapters :2
Adapter#          Adapter Name  State   Mode    Select  Error Path Active
0                    0/7/0/0  NORMAL OFFLINE      0      0   6     0
1                    0/4/0/0  NORMAL ACTIVE       0      0   6     0
```

SDD datapath query device command changes for SDD 1.4.0.0 (or later)

The following change is made in SDD for the **datapath query device** command to accommodate the serial numbers of supported storage devices. The locations of Serial and Policy are swapped because the SAN Volume Controller serial is too long to fit in the first line.

Example output from **datapath query device** command issued in SDD 1.3.3.11 (or earlier):

```
Dev#:   3    Device Name: vpath5  Type: 2105800    Serial: 14123922
 Policy:    Optimized
 ================================================================================
 Path#         Adapter H/W Path    Hard Disk       State    Mode     Select   Error
 0                 0/7/0/0           c19t8d1        OPEN   NORMAL    3869815     0
 1                 0/7/0/0           c13t8d1        OPEN   NORMAL    3872306     0
 2                 0/3/0/0           c17t8d1        OPEN   NORMAL    3874461     0
 3                 0/3/0/0           c11t8d1        OPEN   NORMAL    3872868     0
Dev#:   3    Device Name: vpath5  Type: 2105800    Policy:    Optimized
 Serial: 14123922
 ================================================================================
 Path#         Adapter H/W Path    Hard Disk       State    Mode     Select   Error
 0                 0/7/0/0           c19t8d1        OPEN   NORMAL    3869815     0
 1                 0/7/0/0           c13t8d1        OPEN   NORMAL    3872306     0
 2                 0/3/0/0           c17t8d1        OPEN   NORMAL    3874461     0
 3                 0/3/0/0           c11t8d1        OPEN   NORMAL    3872868     0
```

Example output from **datapath query device** command issued in SDD 1.4.0.0 (or later): (This example shows a SAN Volume Controller and a disk storage system device.)

```
Dev#:   2    Device Name: vpath4  Type: 2145       Policy:    Optimized
 Serial: 60056768018506870000000000000000
 ================================================================================
 Path#         Adapter H/W Path    Hard Disk       State    Mode     Select   Error
 0                 0/7/0/0           c23t0d0        OPEN   NORMAL    2736767    62
 1                 0/7/0/0            c9t0d0        OPEN   NORMAL          6     6
 2                 0/3/0/0           c22t0d0        OPEN   NORMAL    2876312   103
 3                 0/3/0/0            c8t0d0        OPEN   NORMAL        102   101
Dev#:   3    Device Name: vpath5  Type: 2105800    Policy:    Optimized
 Serial: 14123922
 ================================================================================
 Path#         Adapter H/W Path    Hard Disk       State    Mode     Select   Error
 0                 0/7/0/0           c19t8d1        OPEN   NORMAL    3869815     0
 1                 0/7/0/0           c13t8d1        OPEN   NORMAL    3872306     0
 2                 0/3/0/0           c17t8d1        OPEN   NORMAL    3874461     0
 3                 0/3/0/0           c11t8d1        OPEN   NORMAL    3872868     0
```

Note: vpathname *vpathN* is reserved once it is assigned to a LUN even after the LUN has been removed from the host. The same vpathname, *vpathN*, will be assigned to the same LUN when it is reconnected to the host.

Postinstallation

After SDD is installed, the device driver resides above the HP SCSI disk driver (sdisk) in the protocol stack. In other words, SDD now communicates to the HP-UX device layer. The SDD software installation procedure installs a number of SDD components and updates some system files. Those components and files are listed in Table 6 on page 96 through Table 8 on page 96.

Table 6. SDD components installed for HP-UX host systems

File	Location	Description
mod.o	/opt/IBMsdd/bin	Object file for the SDD driver module
Executables	/opt/IBMsdd/bin	Configuration and status tools
README.sd	/opt/IBMsdd	README file
sddsrv	/sbin/sddsrv	SDD server daemon
sample_sddsrv.conf	/etc/	Sample SDD server configuration file
sddserver	/sbin/init.d	Script to start or stop the SDD daemon at system up/down time
confserver	/sbin/init.d	Script to load SDD driver and run cfgvpath at system boot time
mvserver	/sbin/init.d	Script to move /sbin/rc1.d/S100localmount to /sbin/rc1.d/ S250localmount in order to fix the auto mount problem for SDD vpath device filesystems
datapath.1	/usr/local/man/man1/datapath.1	Manpage for datapath
rmvpath.1	/usr/local/man/man1/rmvpath.1	Manpage for rmvpath
showvpath.1	/usr/local/man/man1/ showvpath.1	Manpage for showvpath
gettrace.1	/usr/local/man/man1/gettrace.1	Manpage for gettrace
sddsrv.1	/usr/local/man/man1/sddsrv.1	Manpage for sddsrv
vp2hd.1	/usr/local/man/man1/vp2hd.1	Manpage for vp2hd
hd2vp.1	/usr/local/man/man1/hd2vp.1	Manpage for hd2vp
cfgvpath.1	/usr/local/man/man1/cfgvpath.1	Manpage for cfgvpath
vpcluster.1	/usr/local/man/man1/vpcluster.1	Manpage for vpcluster
sddgetdata.1	/usr/local/man/man1/ sddgetdata.1	Manpage for sddgetdata

Table 7. System files updated for HP-UX host systems

File	Location	Description
vpath	/usr/conf/master.d	Master configuration file
vpath	/stand/system.d	System configuration file

Table 8. SDD commands and their descriptions for HP-UX host systems

Command	Description
cfgvpath	Configures new SDD vpath devices when there are no existing vpath devices. Do not use the legacy parameter "-c".

Command	Description
cfgvpath [-r] (Dynamic reconfiguration)	The SDD vpath device configuration is updated without system reboot. If initially there is no SDD vpath device configured, **cfgvpath -r** will fail with message "failed to get information from kernel, don't run dynamic configuration, do cfgvpath instead." In this case, issue **cfgvpath** without any option.
showvpath	Lists the configuration mapping between the SDD devices and underlying disks.
datapath	The SDD driver console command tool.
hd2vp	Converts a volume group from sdisks into the SDD vpath devices.
vp2hd	Converts a volume group from the SDD vpath devices into sdisks.
vpcluster	Imports or exports MC Service Guard volume groups.
rmvpath [-all, -vpathname]	Removes the SDD vpath devices from the configuration.
gettrace	Debug tool that gets trace information when problem occurs.
sddgetdata	The SDD data collection tool for problem analysis.
man	Manpage for SDD commands, for example, **man datapath**. Supported SDD commands are **datapath**, **gettrace**, **hd2vp**, **querysn**, **rmvpath**, **sddsrv**, **sddgetdatashowvpath**, **vp2hd**, **vpcluster**, and **sddgetdata**.

If you are not using a DBMS or an application package that communicates directly to the sdisk interface, the installation procedure is nearly complete. However, you still need to customize HP-UX so that standard UNIX® applications can use the SDD. Go to "Standard UNIX applications" for instructions. If you have a DBMS or an application package installed that communicates directly to the sdisk interface, such as Oracle, go to "Using applications with SDD" and read the information specific to the application that you are using.

During the installation process, the following files were copied from the IBMsdd_depot to the system:

 # Kernel-related files

- /opt/IBMsdd/bin/mod.o
- /stand/system.d/vpath
- /usr/conf/master.d/vpath

 # SDD driver-related files

- /opt/IBMsdd
- /opt/IBMsdd/bin
- /opt/IBMsdd/README.sd
- /opt/IBMsdd/bin/cfgvpath

- /opt/IBMsdd/bin/datapath
- /opt/IBMsdd/bin/showvpath
- /opt/IBMsdd/bin/master
- /opt/IBMsdd/bin/system
- /opt/IBMsdd/bin/mod.o
- /opt/IBMsdd/bin/rmvpath
- /opt/IBMsdd/bin/get_root_disks
- /opt/IBMsdd/bin/gettrace
- /opt/IBMsdd/bin/sddgetdata
- /opt/IBMsdd/bin/hd2vp
- /opt/IBMsdd/bin/vp2hd
- /opt/IBMsdd/bin/vpcluster
- /sbin/cfgvpath
- /sbin/datapath
- /sbin/get_root_disks
- /sbin/rmvpath
- /sbin/showvpath
- /sbin/hd2vp
- /sbin/vp2hd
- /sbin/vpcluster
- /sbin/sddgetdata
- /sbin/sddsrv
- /etc/sample_sddsrv.conf

During installation, /opt/IBMsdd/bin/cfgvpath program is initiated to create SDD vpath devices in the /dev/dsk and /dev/rdsk directories for all IBM disks that are available on the system. After installation is done, all SDD vpath devices are configured and the driver is loaded. The system will not reboot.

Note: SDD devices are found in /dev/rdsk and /dev/dsk. The device is named according to the SDD number. A device with a number of 0 would be /dev/rdsk/vpath1.

Uninstalling SDD

The following procedure explains how to remove SDD. You must uninstall the current level of SDD before upgrading to a newer level.

Complete the following procedure to uninstall SDD:
1. Stop applications.
2. If you are using SDD with a database, such as Oracle, edit the appropriate database configuration files (database partition) to remove all the SDD devices.
3. Before running the sam program, run script **vp2hd** to convert volume groups from SDD vpath devices to sdisks.
4. Run the **sam** program.

   > **sam**
5. Click **Software Management**.
6. Click **Remove Software**.

7. Click **Remove Local Host Software**.
8. Click the **IBMsdd_tag** selection.
    a. From the Bar menu, click **Actions** ＞ **Mark for Remove**.
    b. From the Bar menu, click **Actions** ＞ **Remove (analysis)**. A Remove Analysis window opens and shows the status of Ready.
    c. Click **OK** to proceed. A Confirmation window opens and indicates that the uninstallation will begin.
    d. Click **Yes**. The analysis phase starts.
    e. After the analysis phase has finished, another Confirmation window opens indicating that the system will be restarted after the uninstallation is complete. Click **Yes** and press **Enter**. The uninstallation of IBMsdd begins.
    f. An Uninstall window opens showing the progress of the IBMsdd software uninstallation. This is what the panel looks like:

```
Target           : XXXXX
Status           : Executing unconfigure
Percent Complete    : 17%
Kbytes Removed      : 340 of 2000
Time Left (minutes) : 5
Removing Software   : IBMsdd_tag,...........
```

    The **Done** option is not available when the uninstallation process is in progress. It becomes available after the uninstallation process completes.
9. Click **Done**.

When SDD has been successfully uninstalled, the first part of the procedure for upgrading the SDD is complete. To complete an upgrade, you need to reinstall SDD. the installation procedure in "Installing SDD."

Note: The MC Service Guard cluster must be halted prior to uninstall. The primary node and the adoptive node or nodes must operate in a single-host environment. The shared volume groups in the adoptive nodes must be exported so that the volumes are not shared; the volume groups can be active in the primary node only.

SDD server daemon

The SDD server (also referred to as sddsrv) is an integrated component of SDD 1.3.1.5 (or later). This component consists of a UNIX application daemon that is installed in addition to the SDD device driver. See "Using the SDD server and the SDDPCM server" for more information about sddsrv.

Note: The SDD server is started from /etc/inittab. Because the environment variable $TZ is not exported at init time, you should modify /etc/default/tz to reflect the corresponding time zone so that sddsrv.log reflects the correct local time. the man pages for the **ctime** command for more information.

Verifying if the SDD server has started

After you have installed SDD, verify that the SDD server (sddsrv) has automatically started by entering **ps –ef | grep sddsrv**.

If the SDD server (sddsrv) has automatically started, the output will display the process number on which sddsrv has started.

If sddsrv has *not* started, you should uninstall SDD and then reinstall SDD. See
"Installing SDD" for more information.

Starting the SDD server manually

If the SDD server does not start automatically after you perform the SDD
installation, or if you want to start it manually after stopping sddsrv, use the
following process to start sddsrv:

1. Edit /etc/inittab and verify the sddsrv entry.

   For example:

   ```
   srv:23456:respawn:/sbin/sddsrv >/dev/null 2>&1
   ```
2. Save the file /etc/inittab.
3. Issue **init q**.

Go to "Verifying if the SDD server has started" for the steps to if you successfully
started the SDD server.

Changing to a different port number for the SDD server

To change to a different port number for the SDD server, the instructions in
"Changing the sddsrv or pcmsrv TCP/IP port number."

Stopping the SDD server

To stop the SDD server, perform the following steps:

1. Edit /etc/inittab and comment out the SDD server entry:

   ```
   #srv:23456:respawn:/sbin/sddsrv >/dev/null 2>&1 2
   ```
2. Save the file.
3. Issue **init q**.
4. Check if sddsrv is running by issuing **ps -ef |grep sddsrv**. If sddsrv is still
   running, issue **kill -9 pid** of sddsrv.

How to import and export volume groups

Use the HP CLI to manage SDD devices.

You can import volume groups that are created over SDD vpath devices using the
vgimport command. The **vgimport** command is useful in conjunction with the
vgexport command.

Before you can import the specified volume groups, you must perform the
following tasks:

1. Export or move volume groups from one node to another node within a high
   availability cluster by using the **vgexport** command. See "Exporting volume
   groups."
2. FTP the map file to the other node within a high-availability cluster. See
   "Moving the map file."
3. Create the volume group device directory. See "Creating the volume group
   device directory."
4. Create the group special file. See "Creating the group special file."

For more information about the **vgimport** command, Importing volume groups.

Exporting volume groups

The **vgexport** command recognizes the following options and arguments:

–p   The –p option previews the actions to be taken but does not update the /etc/lvmtab file or remove the devices file.

–v   The –v option prints verbose messages including the names of the physical volumes associated with this volume group.

–s   –s is the sharable option (Series 800 only). When you specify the –s option, you must also specify the –p, –v, and –m options. A mapfile is created that can be used to create volume group entries (with the **vgimport** command) on other systems in the high availability cluster.

–m mapfile   By default, a file named mapfile is created in your current directory. The mapfile contains a description of the volume group and its associated logical volumes. Use the –m option to specify a different name for the mapfile. The mapfile serves as input to **vgimport**; When the mapfile is used with the –s option, the volume group specified in the mapfile can be shared with the other systems in the high availability cluster.

vg_name   The vg_name is the path name of the volume group.

vgexport command example:

To export the specified volume group on node 1, enter:

vgexport –p -v –s –m */tmp/vgpath1.map vgvpath1*

where */tmp/vgpath1.map* represents your mapfile, and *vgvpath1* represents the path name of volume group that you want to export.

Moving the map file

You must also FTP the map file to the other node.

For example, to FTP the *vgvpath1.map* map file to node 2, enter:

rcp */tmp/vgvpath1.map node2:/tmp/vgvpath1.map*

Creating the volume group device directory

You must also create the volume group device directory.

For example, to create the volume group device directory */dev/vgvpath1* on node 2, enter:

mkdir */dev/vgvpath1*

Creating the group special file

You must also create the group special file on node 2.

For example, to create the *group c 64* file, enter:

mknod /dev/vgvpath1/group c 64 n

where *n* is the same as that was given when */dev/vgvpath1/group* was created on node 1.

Importing volume groups

The **vgimport** command recognizes the following options and arguments:

–p	The –p option previews the actions to be taken but does not update the /etc/lvmtab file or remove the devices file.
–v	The –v option prints verbose messages including the names of the logical volumes.
-s	–s is the sharable option (disk storage system Series 800 only). When you specify the –s option, you must specify the –p, –v, and –m options. The specified mapfile is the same mapfile specified by using the **vgexport** command also using the –p, –m, and –s options. The mapfile is used to create the volume groups on the importing systems.
–m mapfile	By default, a file named mapfile is created in your current directory. The mapfile contains a description of the volume group and its associated logical volumes. Use the –m option to specify a different name for the mapfile. The mapfile serves as input to **vgimport**; When the mapfile is used with the –s option, the volume group specified in the mapfile can be shared among the exporting system and the importing system.
vg_name	The vg_name is the path name of the volume group.

vgimport command example:

To import the specified volume group on node 2, enter:

vgimport -p -v -s -m */tmp/vgpath1.map vgvpath1*

where */tmp/vgpath1.map* represents your mapfile, and *vgvpath1* represents the path name of the volume group that you want to import.

Note: The **vgimport** command only imports the scsi pvlink devices. Issue the **hd2vp** command after issuing the **vgimport** command.

Exporting or importing MC Service Guard volume groups

The **vpcluster** script makes it easier to export or import MC Service Guard volume groups managed by SDD. The normal Service Guard configuration and SDD device configuration process remain the same. The sharable volume groups used by the HA clusters must be marked

For the primary node operation, **vpcluster** extracts the node and volume group information from the keywords NODE_NAME and VOLUME_GROUP as specified in the cluster configuration file. It ensures that the devices of these volumes groups are SDD-managed vpath devices, performs **vgexport** and creates **vpcluster** control files for the adoptive nodes to import. The input file to **vpcluster** does not have to be the same cluster configuration file for the SG cluster creation. It can be any ASCII file with the line entries that contain NODE_NAME and VOLUME_GROUP keywords without # as a comment. Optionally, the **vpcluster** control file can be copied to each adoptive node with the **rcp** command.

For the adoptive node operation, **vpcluster** uses the control file created by the primary node operation. Prior to volume group import, it validates that the adoptive node is included in the cluster configuration, ensures the importing volume groups are not active volume groups in the adoptive node, creates volume group nodes /dev/vgXXXX using the **mknod** command, and ensures that the same device serial and LUN-id are configured by SDD.

Notes:

1. The device names, either vpath# or C#T#D#, might be different between the primary and adoptive nodes. However, the **vpcluster** process attempts to keep the volume group minor number consistent between the primary and adoptive nodes. In case the same minor number is in use, the next sequential number is assigned. Because the HP **vgimport** process only imports those device names with the standard name C#T#D#, **hd2vp** is invoked to convert sdisk devices to SDD devices after a successful **vgimport**.
2. Use the **cmquerycl** HP command to create the cluster configuration file. This command recognizes only the pvlink scsi devices on both nodes. If you are using the **cmquerycl** command to create the cluster configuration file, you should first issue **vp2hd** to convert vpath devices to pvlink devices.

In addition, the report option for adoptive node validates all volume groups exported by the primary node. A mismatch of volume group minor number or vpath device name is allowed. Other mismatches will be reported.

Syntax:

```
►►──vpcluster──┬─ -primary──┬──┬─────────┬──┬────────┬──┬─────────┬────────────►
               └─ -adoptive─┘  └─ -f file─┘  └─ -dorcp─┘  └─ -report─┘

►──┬─────────┬──┬─────┬─────────────────────────────────────────────────────►◄
   └─ -debug─┘  └─ -h─┘
```

where,

-primary
  Specifies primary node operation. You must specify **-primary** or **-adoptive**.

-adoptive
  Specifies adoptive node operation. You must specify **-primary** or **-adoptive**.

-f file
  For the primary node, specify the cluster configuration file, default is /etc/cmcluster/cmclconf.ascii.

  For the adoptive node, specify the **vpcluster** control file created the primary node, default is /tmp/vpcluster/vpcluster.primary.tar

-dorcp
  Specifies the vpcluster control tar file to be RCPed to the adoptive nodes. The default is no.

-report
  Validates that the exported volume groups from the primary node are imported to the adoptive node and creates a report. This option is valid in the adoptive node.

-debug
  Specifies that a debugging statement is to be printed during **vpcluster** run time.

-h Specifies that detailed help info about the **vpcluster** function is to be displayed.

There are more than one way to configure the SG cluster locking: quorum server, or lock disk. In case the lock disk is chosen, the SDD vpath device should not be used since it is not recognized by the FIRST_CLUSTER_LOCK_PV parameter. Furthermore, it is recommended that SDD vpath devices and sdisk pvlink should not be mixed in the same volume group. The lock device should be excluded from the SDD configuration. See the information about the /etc/vpathmanualexcl.cfg text file.

Using applications with SDD

If your system already has a software application or a DBMS installed that communicates directly with the HP-UX disk device drivers, you need to insert the new SDD device layer between the software application and the HP-UX disk device layer. You also need to customize the software application to have it communicate with the SDD devices instead of the HP-UX devices.

In addition, many software applications and DBMSs need to control certain device attributes such as ownership and permissions. Therefore, you must ensure that the new SDD devices that these software applications or DBMSs access in the future have the same attributes as the HP-UX sdisk devices that they replace. You need to customize the application or DBMS to accomplish this.

This section contains the procedures for customizing the following software applications and DBMSs for use with SDD:
- Standard UNIX applications
- Network File System (NFS) file server

Standard UNIX applications

Note: The SDD path state change depends on the "positive" feedback from an I/O request. The positive feedback is an I/O response, either successful or failed condition, from an I/O request. Due to the standard and designed behavior of the SCSI driver on HP-UX 11i V2 and prior, the following conditions occur when access raw sdisk through SDD:

1. I/O errors are retried indefinitely because pre-HP-UX 11i v3 (B.11.31) has no way to distinguish between a non-volume-manager file system I/O and a raw access.
2. I/O errors returned on certain file system I/Os can result in panic.

SDD can still load balance with the raw device access but failover/failback is not possible, In most cases, an I/O error is likely to hang due to indefinite retries at the SCSI driver layer.

If you have not already done so, install SDD using the procedure in "Installing SDD." When this is done, SDD resides above the HP-UX SCSI disk driver (sdisk) in the protocol stack. In other words, SDD now communicates to the HP-UX device layer. To use standard UNIX applications with SDD, you must make some changes to your logical volumes. You must convert your existing logical volumes or create new ones.

Standard UNIX applications such as newfs, fsck, mkfs, and mount, which normally take a disk device or raw disk device as a parameter, also accept the SDD device as a parameter. Similarly, entries in files such as vfstab and dfstab (in the format of

cntndnsn) can be replaced by entries for the corresponding SDD vpathNs devices. Make sure that the devices that you want to replace are replaced with the corresponding SDD device. Issue the **showvpath** command to list all SDD vpath devices and their underlying disks.

To use the SDD driver for an existing logical volume, you must run the hd2vp conversion script (see "SDD utility programs").

Attention: Do not use the SDD for critical file systems needed at startup, such as /(root), /stand, /usr, /tmp or /var. Doing so may render your system unusable if SDD is ever uninstalled (for example, as part of an upgrade).

Creating new logical volumes

Use the following process to create a new logical volume to use SDD:

Note: You must have superuser privileges to perform these subtasks.

1. Determine the major number of the logical volume device.

   Enter the following command to determine the major number:

   # **lsdev | grep lv**

   A message similar to the following is displayed:

   ```
   64          64          lv              lvm
   ```

   The first number in the message is the major number of the character device, which is the number that you want to use.

2. Create a device node for the logical volume device.

   Note: If you do not have any other logical volume devices, you can use a minor number of 0x010000. In this example, assume that you have no other logical volume devices. A message similar to the following is displayed:

   ```
   # mknod group c 64 0x010000
   ```

   Create a physical volume by performing the procedure in step 3 on page 106.

   a. Create a subdirectory in the /dev directory for the volume group.

      Enter the following command to create a subdirectory in the /dev directory for the volume group:

      # **mkdir /dev/vgIBM**

      In this example, *vgIBM* is the name of the directory.

   b. Change to the /dev directory.

      Enter the following command to change to the /dev directory:

      # **cd /dev/vgIBM**

   c. Create a device node for the logical volume device.

      Enter the following command to re-create the physical volume:

      # **pvcreate /dev/rdsk/vpath1**

A message similar to the following is displayed:

```
Physical volume "/dev/rdsk/vpath1" has been successfully created.
```

In this example, the SDD vpath device associated with the underlying disk is vpath1. Verify the underlying disk by entering the following **showvpath** command:

 # /opt/IBMsdd/bin/showvpath

A message similar to the following is displayed:

```
vpath1:
  /dev/dsk/c3t4d0
```

3. Create a physical volume.

Enter the following command to create a physical volume:

 # pvcreate /dev/rdsk/vpath1

4. Create a volume group.

Enter the following command to create a volume group:

 # vgcreate /dev/vgIBM/dev/dsk/vpath1

5. Create a logical volume.

Enter the following command to create logical volume *lvol1*:

 # lvcreate -L 100 -n lvol1 vgIBM

The -L 100 portion of the command makes a 100-MB volume group; you can make it larger if you want to. Now you are ready to create a file system on the volume group.

6. Create a file system on the volume group.

Use the following process to create a file system on the volume group:

a. If you are using an HFS file system, enter the following command to create a file system on the volume group:

 # newfs -F HFS /dev/vgIBM/rlvol1

b. If you are using a VXFS file system, enter the following command to create a file system on the volume group:

 # newfs -F VXFS /dev/vgIBM/rlvol1

c. Mount the logical volume.

This process assumes that you have a mount point called /mnt.

7. Mount the logical volume.

Enter the following command to mount the logical volume *lvol1*:

 # mount /dev/vgIBM/lvol1 /mnt

Attention: In some cases, it may be necessary to use standard HP-UX recovery procedures to fix a volume group that has become damaged or corrupted. For information about using recovery procedures, such as **vgscan**, **vgextend**, **vpchange**, or **vgreduce**, the following Web site:

http://docs.hp.com/

Click *HP-UX Reference (Manpages)*. Then *HP-UX Reference Volume 2*.

Removing logical volumes

Use the following procedure to remove logical volumes:

1. Remove the existing logical volume.

   Before the logical volume is removed, it must be unmounted. For example, enter the following command to unmount logical volume *lvol1*:

   # umount /dev/vgIBM/lvol1

   Next, remove the logical volume.

   For example, enter the following command to remove logical volume *lvol1*:

   # lvremove /dev/vgIBM/lvol1

   A message similar to the following is displayed:

   ```
   The logical volume "/dev/vgIBM/lvol1" is not empty;
   do you really want to delete the logical volume (y/n)
   ```

   Enter **y** and press **Enter**. A message similar to the following is displayed:

   ```
   Logical volume "/dev/vgIBM/lvol1" has been successfully removed.
   Volume Group configuration for /dev/vgIBM has been saved in
   /etc/lvmconf/vgIBM.conf
   ```

   When prompted to delete the logical volume, enter **y**.

2. Remove the existing volume group.

   Enter the following command to remove the volume group *vgIBM*:

   # vgremove /dev/vgIBM

   A message similar to the following is displayed:

   ```
   Volume group "/dev/vgIBM" has been successfully removed.
   ```

Now, you can re-create the logical volume.

Re-creating the existing logical volumes

Use the following process to convert an existing logical volume to use SDD:

Note: You must have superuser privileges to perform these subtasks.

As an example, suppose you have a logical volume called *lvol1* under a volume group *vgIBM*, which is currently using the disk directly, (for example, through path /dev path /dev/dsk/c3t4d0). You want to convert logical volume *lvol1* to use SDD.

1. Determine the size of the logical volume.

   Enter the following command to determine the size of the logical volume:

   # lvdisplay /dev/vgIBM/lvol1 | grep ″LV Size″

   A message similar to the following is displayed:

```
LV Size (Mbytes) 100
```

In this case, the logical volume size is 100 MB.

2. Re-create the physical volume.

   Enter the following command to re-create the physical volume:

   # **pvcreate /dev/rdsk/vpath1**

   A message similar to the following is displayed:

   ```
   Physical volume "/dev/rdsk/vpath1" has been successfully created.
   ```

   In this example, the SDD vpath device associated with the underlying disk is vpath1. Verify the underlying disk by entering the following command:

   # **/opt/IBMsdd/bin/showvpath**

   A message similar to the following is displayed:

   ```
   vpath1:
     /dev/dsk/c3t4d0
   ```

3. Re-create the volume group.

   Enter the following command to re-create the volume group:

   # **vgcreate /dev/vgibm /dev/dsk/vpath1**

   A message similar to the following is displayed:

   ```
   Increased the number of physical extents per physical volume to 2187.
   Volume group "/dev/vgibm" has been successfully created.
   Volume Group configuration for /dev/vgibm has been saved in
   /etc/lvmconf/vgibm.conf
   ```

4. Re-create the logical volume.

   Re-creating the logical volume consists of a number of smaller steps:

   a. Re-creating the physical volume
   b. Re-creating the volume group
   c. Re-creating the logical volume

      Enter the following command to re-create the logical volume:

      # **lvcreate -L 100 -n lvol1 vgibm**

      A message similar to the following is displayed:

      ```
      Logical volume "/dev/vgibm/lvol1" has been successfully created with
      character device "/dev/vgibm/rlvol1".
      Logical volume "/dev/vgibm/lvol1" has been successfully extended.
      Volume Group configuration for /dev/vgibm has been saved in
      /etc/lvmconf/vgibm.conf
      ```

      The -L 100 parameter comes from the size of the original logical volume, which is determined by using the **lvdisplay** command. In this example, the original logical volume was 100 MB in size.

Attention: The re-created logical volume should be the same size as the original volume; otherwise, the re-created volume cannot store the data that was on the original.

5. Setting the proper timeout value for the logical volume manager.

   The timeout values for the Logical Volume Manager must be set correctly for SDD to operate properly. This is particularly true if the concurrent firmware download has taken place. There are two timeout values: one for logical volume (LV) and one for physical volume (PV).

   The LV timeout value is determined by the application. If the application has no specific timeout requirement, use the HP default value, which is 0 (forever).

   The PV timeout value is recommended by the storage vendor. The HP default PV timeout value is 30 seconds. Generally, this is sufficient during normal operations.

   However, during the concurrent firmware download, you must set the PV timeout value to a minimum of 90 seconds. You can set the timeout value to 90 seconds for normal operations as well. In addition, if you do not use the default LV timeout value, ensure that the LV timeout value is no less than the PV timeout value multiplied by the number of paths. For example, when the default is not used, if a vpath device has four underlying paths and the PV timeout value is 90, the LV timeout value must be at least 360.

   To display the timeout value, use the **lvdisplay** or **pvdisplay** command.

   To change the PV timeout value, use the **pvchange** command after **pvcreate**, and to change the LV timeout value, use the **lvchange** command after **lvcreate**.

   For example:

   - To change the timeout value of all underlying paths of vpathX to 90 seconds, enter **pvchange -t 90 /dev/dsk/vpathX**
   - To change the timeout value for logical volume /dev/vgibm/lvolY to 360 seconds, enter **lvchange -t 360 /dev/vgibm/lvolY**

   In some cases, it might be necessary to use standard HP recovery procedures to fix a volume group that has become damaged or corrupted. For information about using recovery procedures, such as **vgscan**, **vgextend**, **vpchange**, or **vgreduce**, the following Web site:

   http://docs.hp.com/

   Click **HP-UX Reference (Manpages)**. Then, *HP-UX Reference Volume 2*.

Installing SDD on a NFS file server

The procedures in this section show how to install SDD for use with an exported file system (NFS file server).

Setting up NFS for the first time

Perform the following steps if you are installing exported file systems on SDD devices for the first time:

1. If you have not already done so, install SDD using the procedure in "Installing SDD."

2. Determine which SDD (vpathN) volumes that you will use as file system devices.

3. Create file systems on the selected SDD devices using the appropriate utilities for the type of file system that you will use. If you are using the standard HP-UX UFS file system, enter the following command:

   # newfs /dev/rdsk/vpathN

In this example, *N* is the SDD device instance of the selected volume. Create mount points for the new file systems.

4. Install the file systems into the directory /etc/fstab. In the **mount at boot** field, click **yes**.

5. Install the file system mount points into the /etc/exports directory for export.

6. Restart the system.

Installing SDD on a system that already has the NFS file server

Perform the following steps if you have the NFS file server already configured to:

- Export file systems that reside on a multiport subsystem, and
- Use SDD partitions instead of sdisk partitions to access them

1. List the mount points for all currently exported file systems by looking in the /etc/exports directory.

2. Match the mount points found in step 1 with sdisk device link names (files named /dev/(r)dsk/cntndn) by looking in the /etc/fstab directory.

3. Match the sdisk device link names found in step 2 with SDD device link names (files named /dev/(r)dsk/vpathN) by issuing the **showvpath** command.

4. Make a backup copy of the current /etc/fstab file.

5. Edit the /etc/fstab file, replacing each instance of an sdisk device link named /dev/(r)dsk/cntndn with the corresponding SDD device link.

6. Restart the system.

7. Verify that each exported file system:

   a. Passes the start time **fsck pass**

   b. Mounts properly

   c. Is exported and available to NFS clients

If there is a problem with any exported file system after completing step 7, restore the original /etc/fstab file and restart to restore NFS service. Then review your steps and try again.

Notices

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

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

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

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

For license inquiries regarding double-byte (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:

IBM World Trade Asia Corporation
Licensing
2-31 Roppongi 3-chome, Minato-ku
Tokyo 106, Japan

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

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

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

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

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

IBM Corporation
Information Enabling Requests
Dept. DZWA
5600 Cottle Road
San Jose, CA 95193
U.S.A.

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The licensed program described in this document and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement or any equivalent agreement between us.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

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

All statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

This information is for planning purposes only. The information herein is subject to change before the products described become available.

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

If you are viewing this information softcopy, the photographs and color illustrations may not appear.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or ™), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at www.ibm.com/legal/copytrade.shtml.

Adobe is a registered trademark of Adobe Systems Incorporated in the United States, and/or other countries.

Intel is a trademark or registered trademark of Intel Corporation or its subsidiaries in the United States and other countries.

Java is a trademark of Sun Microsystems, Inc. in the United States, other countries, or both.

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

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

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

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

IBM license agreement for machine code

> **Read Before Using**
>
> IMPORTANT
>
> YOU ACCEPT THE TERMS OF THIS IBM LICENSE AGREEMENT FOR MACHINE CODE BY YOUR USE OF THE HARDWARE PRODUCT OR MACHINE CODE. PLEASE READ THE AGREEMENT CONTAINED IN THIS BOOK BEFORE USING THE HARDWARE PRODUCT.

You accept the terms of this Agreement[1] by your initial use of a machine that contains IBM Licensed Internal Code (called "Code"). These terms apply to Code used by certain machines IBM or your reseller specifies (called "Specific Machines"). International Business Machines Corporation or one of its subsidiaries ("IBM") owns copyrights in Code or has the right to license Code. IBM or a third party owns all copies of Code, including all copies made from them.

---

1. Form Z125-4144

If you are the rightful possessor of a Specific Machine, IBM grants you a license to use the Code (or any replacement IBM provides) on, or in conjunction with, only the Specific Machine for which the Code is provided. IBM licenses the Code to only one rightful possessor at a time.

Under each license, IBM authorizes you to do only the following:

1. Run the Code to enable the Specific Machine to function according to its Official Published Specifications (called "Specifications");

2. Make a backup or archival copy of the Code (unless IBM makes one available for your use), provided you reproduce the copyright notice and any other legend of ownership on the copy. You may use the copy only to replace the original, when necessary; and

3. Run and display the Code as necessary to maintain the Specific Machine.

You agree to acquire any replacement for, or additional copy of, Code directly from IBM in accordance with IBM's standard policies and practices. You also agree to use that Code under these terms.

You may transfer possession of the Code to another party only with the transfer of the Specific Machine. If you do so, you must 1) destroy all your copies of the Code that were not provided by IBM, 2) either give the other party all your IBM-provided copies of the Code or destroy them, and 3) notify the other party of these terms. IBM licenses the other party when it accepts these terms. These terms apply to all Code you acquire from any source.

Your license terminates when you no longer rightfully possess the Specific Machine.

Actions you must not take

You agree to use the Code only as authorized above. You must not do, for example, any of the following:

1. Otherwise copy, display, transfer, adapt, modify, or distribute the Code (electronically or otherwise), except as IBM may authorize in the Specific Machine's Specifications or in writing to you;

2. Reverse assemble, reverse compile, or otherwise translate the Code unless expressly permitted by applicable law without the possibility of contractual waiver;

3. Sublicense or assign the license for the Code; or

4. Lease the Code or any copy of it.

Glossary

This glossary includes terms for disk storage system products and Virtualization products.

This glossary includes selected terms and definitions from:

- The *American National Standard Dictionary for Information Systems*, ANSI X3.172–1990, copyright 1990 by the American National Standards Institute (ANSI), 11 West 42nd Street, New York, New York 10036. Definitions derived from this book have the symbol (A) after the definition.
- *IBM Terminology*, which is available online at the following Web site: http://www.ibm.com/ibm/terminology/. Definitions derived from this book have the symbol (GC) after the definition.
- The *Information Technology Vocabulary* developed by Subcommittee 1, Joint Technical Committee 1, of the International Organization for Standardization and the International Electrotechnical Commission (ISO/IEC JTC1/SC1). Definitions derived from this book have the symbol (I) after the definition. Definitions taken from draft international standards, committee drafts, and working papers being developed by ISO/IEC JTC1/SC1 have the symbol (T) after the definition, indicating that final agreement has not been reached among the participating National Bodies of SCI.

This glossary uses the following cross-reference forms:

See This refers the reader to one of two kinds of related information:
- A term that is the expanded form of an abbreviation or acronym. This expanded form of the term contains the full definition.
- A synonym or more preferred term.

See also
This refers the reader to one or more related terms.

Special characters

1750. The machine type for the IBM System Storage DS6000 series. Models for the DS6000 include the 511 and EX1.

2105. The machine number for the IBM TotalStorage Enterprise Storage Server (ESS). Models of the ESS are expressed as the number 2105 followed by "Model <xxx>", such as 2105 Model 800. The 2105 Model 100 is an ESS expansion enclosure that is typically referred to simply as the Model 100. See also *IBM TotalStorage Enterprise Storage Server* and *Model 100*.

2107. A hardware machine type for the IBM System Storage DS8000 series. Hardware models for the 2107 include base units 921, 922, 931, 932, 9A2, 9B2 and expansion units 92E and 9AE.

2145. A hardware machine type for the IBM System Storage SAN Volume Controller. Models of the SAN Volume Controller are expressed as the number 2145 followed by "-*xxx*", such as 2145-8G4. Hardware models for the 2145 include 2145-4F2, 2145-8F2, 2145-8F4, and 2145-8G4.

3390. The machine number of an IBM disk storage system. The ESS, when interfaced to IBM S/390® or IBM System z® hosts, is set up to appear as one or more 3390 devices, with a choice of 3390-2, 3390-3, or 3390-9 track formats.

3990. The machine number of an IBM control unit.

7133. The machine number of an IBM disk storage system. The Model D40 and 020 drawers of the 7133 can be installed in the 2105-100 expansion enclosure of the ESS.

8-pack. See *disk eight pack*.

/ file system. The root file system; contains files that contain machine-specific configuration data.

/tmp file system. A shared storage location for files.

/usr file system. Contains files and programs necessary for operating the machine.

/var file system. Contains files that are variable on a per-client basis, such as spool and mail files.

A

access. (1) To obtain the use of a computer resource. (2) In computer security, a specific type of interaction

between a subject and an object that results in flow of information from one to the other.

access-any mode. One of the two access modes that can be set for the disk storage system product during initial configuration. It enables all fibre-channel-attached host systems with no defined access profile to access all logical volumes on the disk storage system. With a profile defined in ESS Specialist for a particular host, that host has access only to volumes that are assigned to the WWPN for that host. See also *pseudo-host* and *worldwide port name*.

ACK. See *request for acknowledgement and acknowledgement*.

active Copy Services server. The Copy Services server that manages the Copy Services domain. Either the primary or the backup Copy Services server can be the active Copy Services server. The backup Copy Services server is available to become the active Copy Services server if the primary Copy Services server fails. See also *backup Copy Services server*, *Copy Services client*, and *primary Copy Services server*.

active/active mode. A configuration that enables one controller node of a storage system pair to process I/O requests and provide a standby capability for the other controller node. Generally, an active/active storage system involves a battery-backed mirrored cache, in which the cache content of a controller is mirrored to another for data integrity and availability.

active/passive mode. A configuration that enables one controller node of a storage system pair to process I/O requests, while the other controller node is idle in standby mode ready to take over I/O activity if the active primary controller fails or is taken offline.

alert. A message or log that a storage facility generates as the result of error event collection and analysis. An alert indicates that a service action is required.

allegiance. In Enterprise Systems Architecture/390, a relationship that is created between a device and one or more channel paths during the processing of certain conditions. See also *implicit allegiance*, *contingent allegiance*, and *reserved allegiance*.

allocated storage. In a disk storage system, the space that is allocated to volumes but not yet assigned. See also *assigned storage*.

American National Standards Institute (ANSI). An organization of producers, consumers, and general interest groups that establishes the procedures by which accredited organizations create and maintain voluntary industry standards in the United States. (A)

Anonymous. In ESS Specialist, the label on an icon that represents all connections that are using fibre-channel adapters between the ESS and hosts and

that are not completely defined to the ESS. See also *anonymous host*, *pseudo-host*, and *access-any mode*.

anonymous host. Synonym for "pseudo-host" (in contrast to the Anonymous label that appears on some pseudo-host icons. See also *Anonymous* and *pseudo-host*.

ANSI. See *American National Standards Institute*.

APAR. See *authorized program analysis report*. (GC)

arbitrated loop. For fibre-channel connections, a topology that enables the interconnection of a set of nodes. See also *point-to-point connection* and *switched fabric*.

array. An ordered collection, or group, of physical devices (disk drive modules) that are used to define logical volumes or devices. More specifically, regarding the disk storage system, an array is a group of disks designated by the user to be managed by the RAID-5 technique. See also *redundant array of independent disks*.

ASCII. (American National Standard Code for Information Interchange) The standard code, using a coded character set consisting of 7-bit coded characters (8 bits including parity check), that is used for information interchange among data processing systems, data communication systems, and associated equipment. The ASCII set consists of control characters and graphic characters. (A) Some organizations, including IBM, have used the parity bit to expand the basic code set.

assigned storage. On a disk storage system, the space allocated to a volume and assigned to a port.

authorized program analysis report (APAR). A report of a problem caused by a suspected defect in a current, unaltered release of a program. (GC)

availability. The degree to which a system or resource is capable of performing its normal function. See *data availability*.

B

backup Copy Services server. One of two Copy Services servers in a Copy Services domain. The other Copy Services server is the primary Copy Services server. The backup Copy Services server is available to become the active Copy Services server if the primary Copy Services server fails. A Copy Services server is software that runs in one of the two clusters of an ESS, and manages data-copy operations for that Copy Services server group. See also *active Copy Services server*, *Copy Services client*, and *primary Copy Services server*.

bay. In the disk storage system, the physical space used for installing SCSI, ESCON®, and fibre-channel

host adapter cards. The ESS has four bays, two in each cluster. See also *service boundary*.

bit. (1) Either of the digits 0 or 1 when used in the binary numeration system. (T) (2) The storage medium required to store a single binary digit. See also *byte*.

block. (1) A string of data elements recorded or transmitted as a unit. The elements may be characters, words, or physical records. (T) (2) In the disk storage system, a group of consecutive bytes used as the basic storage unit in fixed-block architecture (FBA). All blocks on the storage device are the same size (fixed size). See also *fixed-block architecture* and *data record*.

byte. (1) A group of eight adjacent binary digits that represent one EBCDIC character. (2) The storage medium required to store eight bits. See also *bit*.

C

cache. A special-purpose buffer storage, smaller and faster than main storage, used to hold a copy of instructions and data obtained from main storage and likely to be needed next by the processor. (T)

cache fast write. In the disk storage system, a form of the fast-write operation in which the storage server writes the data directly to cache, where it is available for later destaging.

cache hit. An event that occurs when a read operation is sent to the cluster, and the requested data is found in cache. The opposite of *cache miss*.

cache memory. Memory, typically volatile memory, that a storage server uses to improve access times to instructions or data. The cache memory is typically smaller and faster than the primary memory or storage medium. In addition to residing in cache memory, the same data also resides on the storage devices in the storage facility.

cache miss. An event that occurs when a read operation is sent to the cluster, but the data is not found in cache. The opposite of *cache hit*.

call home. A communication link established between the disk storage system and a service provider. The disk storage system can use this link to place a call to IBM or to another service provider when it requires service. With access to the machine, service personnel can perform service tasks, such as viewing error logs and problem logs or initiating trace and dump retrievals. See also *heartbeat* and *remote technical assistance information network*.

cascading. (1) Connecting network controllers to each other in a succession of levels, to concentrate many more lines than a single level permits. (2) In high-availability cluster multiprocessing (HACMP), cascading pertains to a cluster configuration in which

the cluster node with the highest priority for a particular resource acquires the resource if the primary node fails. The cluster node relinquishes the resource to the primary node upon reintegration of the primary node into the cluster.

catcher. A server that service personnel use to collect and retain status data that a disk storage system sends to it.

CCR. See *channel command retry*.

CCW. See *channel command word*.

CD. See *compact disc*.

compact disc. An optically read disc, typically storing approximately 660 MB. CD-ROM (compact disc read-only memory) refers to the read-only format used to distribute disk storage system code and documentation.

CEC. See *computer-electronic complex*.

channel. In Enterprise Systems Architecture/390, the part of a channel subsystem that manages a single I/O interface between a channel subsystem and a set of control units.

channel command retry (CCR). In Enterprise Systems Architecture/390, the protocol used between a channel and a control unit that enables the control unit to request that the channel reissue the current command.

channel command word (CCW). In Enterprise Systems Architecture/390, a data structure that specifies an I/O operation to the channel subsystem.

channel path. In Enterprise Systems Architecture/390, the interconnection between a channel and its associated control units.

channel subsystem. In Enterprise Systems Architecture/390, the part of a host computer that manages I/O communication between the program and any attached control units.

channel-subsystem image. In Enterprise Systems Architecture/390, the logical functions that a system requires to perform the function of a channel subsystem. With ESCON multiple image facility (EMIF), one channel subsystem image exists in the channel subsystem for each logical partition (LPAR). Each image appears to be an independent channel subsystem program, but all images share a common set of hardware facilities.

CKD. See *count key data*.

CLI. See *command-line interface*. See also *Copy Services command-line interface*.

cluster. (1) In the disk storage system, a partition capable of performing all disk storage system functions.

With two clusters in the disk storage system, any operational cluster can take over the processing of a failing cluster. (2) In the AIX operating system, a group of nodes within a complex.

cluster processor complex (CPC). In the disk storage system, the unit within a cluster that provides the management function for the disk storage system. It consists of cluster processors, cluster memory, and related logic.

Code Distribution and Activation (CDA). Process of installing licensed machine code on a disk storage system while applications continue to run.

command-line interface (CLI). An interface provided by an operating system that defines a set of commands and enables a user (or a script-like language) to issue these commands by typing text in response to the command prompt (for example, DOS commands, UNIX shell commands). See also *Copy Services command-line interface*.

compression. (1) The process of eliminating gaps, empty fields, redundancies, and unnecessary data to shorten the length of records or blocks. (2) Any encoding that reduces the number of bits used to represent a given message or record. (GC)

computer-electronic complex (CEC). The set of hardware facilities associated with a host computer.

concurrent copy. A facility on a storage server that enables a program to make a backup of a data set while the logical volume remains available for subsequent processing. The data in the backup copy is frozen at the point in time that the server responds to the request.

concurrent download of licensed machine code. Process of installing licensed machine code while applications continue to run.

concurrent maintenance. Service that is performed on a unit while it is operational.

concurrent media maintenance. Service performed on a disk drive module (DDM) without losing access to the data.

configure. In storage, to define the logical and physical configuration of the input/output (I/O) subsystem through the user interface that the storage facility provides for this function.

consistent copy. A copy of a data entity (a logical volume, for example) that contains the contents of the entire data entity at a single instant in time.

console. A user interface to a server, such as can be provided by a personal computer. See also *IBM TotalStorage ESS Master Console*.

contingent allegiance. In Enterprise Systems Architecture/390, a relationship that is created in a control unit between a device and a channel when the channel accepts unit-check status. The allegiance causes the control unit to guarantee access; the control unit does not present the busy status to the device. The allegiance enables the channel to retrieve sense data that is associated with the unit-check status on the channel path associated with the allegiance.

control unit (CU). (1) A device that coordinates and controls the operation of one or more input/output devices, and synchronizes the operation of such devices with the operation of the system as a whole. (2) In Enterprise Systems Architecture/390, a storage server with ESCON, FICON®, or OEMI interfaces. The control unit adapts a native device interface to an I/O interface supported by an ESA/390 host system. (3) In the ESS, the portion of the ESS that supports the attachment of emulated CKD devices over ESCON, FICON, or OEMI interfaces. See also *cluster*.

control-unit image. In Enterprise Systems Architecture/390, a logical subsystem that is accessed through an ESCON or FICON I/O interface. One or more control-unit images exist in each control unit. Each image appears as an independent control unit, but all control-unit images share a common set of hardware facilities. The ESS can emulate 3990-3, TPF, 3990-6, or 2105 control units.

control-unit initiated reconfiguration (CUIR). A software mechanism that the ESS uses to request that an operating system of an IBM System z or S/390 host verify that one or more subsystem resources can be taken offline for service. The ESS can use this process to automatically vary channel paths offline and online to facilitate bay service or concurrent code installation. Depending on the operating system, support for this process might be model-dependent, might depend on the IBM TotalStorage Enterprise Storage Server Subsystem Device Driver, or might not exist.

Coordinated Universal Time (UTC). The international standard of time that is kept by atomic clocks around the world.

Copy Services client. Software that runs on each ESS cluster in the Copy Services server group and that performs the following functions:

- Communicates configuration, status, and connectivity information to the Copy Services server.
- Performs data-copy functions on behalf of the Copy Services server.

See also *active Copy Services server*, *backup Copy Services server*, and *primary Copy Services server*.

Copy Services CLI. See *Copy Services Command-Line Interface*.

Copy Services domain. A collection of user-designated ESS clusters participating in Copy Services functions managed by a designated active Copy Services server. See also *Copy Services server*, *dual-active server*, and *single-active server*.

Copy Services command-line interface (Copy Services CLI). In the ESS, command-line interface software provided with ESS Copy Services and used for invoking Copy Services functions from host systems attached to the ESS. See also *command-line interface*.

Copy Services server. An ESS cluster designated by the copy services administrator to perform the ESS Copy Services functions. See also *active Copy Services server*, *backup Copy Services server*, and *primary Copy Services server*.

Copy Services server group. A collection of user-designated ESS clusters participating in Copy Services functions managed by a designated active Copy Services server. A Copy Services server group is also called a Copy Services domain. See also *active Copy Services server*, *backup Copy Services server*, and *primary Copy Services server*.

count field. The first field of a count key data (CKD) record. This eight-byte field contains a four-byte track address (CCHH). It defines the cylinder and head that are associated with the track, and a one-byte record number (R) that identifies the record on the track. It defines a one-byte key length that specifies the length of the record's key field (0 means no key field). It defines a two-byte data length that specifies the length of the record's data field (0 means no data field). Only the end-of-file record has a data length of zero.

count key data (CKD). In Enterprise Systems Architecture/390, a data-record format employing self-defining record formats in which each record is represented by up to three fields: a *count* field identifying the record and specifying its format, an optional *key* field that can be used to identify the data area contents, and an optional *data* field that typically contains the user data. For CKD records on the ESS, the logical volume size is defined in terms of the device emulation mode (3390 or 3380 track format). The count field is always 8 bytes long and contains the lengths of the key and data fields, the key field has a length of 0 to 255 bytes, and the data field has a length of 0 to 65 535 or the maximum that will fit on the track. See also *data record*.

CPC. See *cluster processor complex*.

CRC. See *cyclic redundancy check*.

CU. See *control unit*.

CUIR. See *control-unit initiated reconfiguration*.

custom volume. In the ESS, a volume in count-key-data (CKD) format that is not a standard volume, which basically means that it does not necessarily present the same number of cylinders and capacity to its assigned logical control unit as provided by one of the following standard S/390 volume types: 3390-2, 3390-3, 3390-9, 3390-2 (3380-track mode), or 3390-3 (3380-track mode). See also *count-key-data*, *interleave*, *standard volume*, and *volume*.

CUT. See *Coordinated Universal Time*.

cyclic redundancy check (CRC). A redundancy check in which the check key is generated by a cyclic algorithm. (T)

cylinder. A unit of storage on a CKD device with a fixed number of tracks.

D

DA. See *device adapter*. See also *SSA adapter*.

daisy chain. See *serial connection*.

DASD. See *direct access storage device*.

DASD fast write (DFW). A function of a storage server in which active write data is stored in nonvolatile cache, thus avoiding exposure to data loss.

data availability. The degree to which data is available when needed, typically measured as a percentage of time that the system would be capable of responding to any data request (for example, 99.999% available).

data compression. A technique or algorithm used to encode data such that the encoded result can be stored in less space than the original data. The original data can be recovered from the encoded result through a reverse technique or reverse algorithm. See also *compression*.

Data Facility Storage Management Subsystem. An operating environment that helps automate and centralize the management of storage. To manage storage, DFSMS™ provides the storage administrator with control over data class, storage class, management class, storage group, and automatic class selection routine definitions.

data field. The optional third field of a count key data (CKD) record. The count field specifies the length of the data field. The data field contains data that the program writes.

data record. The basic unit of S/390 and IBM System z storage on an ESS, also known as a count-key-data (CKD) record. Data records are stored on a track. The records are sequentially numbered starting with 0. The first record, R0, is typically called the track descriptor record and contains data normally used by the operating system to manage the track. See also *count-key-data* and *fixed-block architecture*.

data sharing. The ability of multiple host systems to concurrently utilize data that they store on one or more storage devices. The storage facility enables configured storage to be accessible to any, or all, attached host systems. To use this capability, the host program must be designed to support data that it is sharing.

DDM. See *disk drive module*.

DDM group. See *disk eight pack*.

dedicated storage. Storage within a storage facility that is configured such that a single host system has exclusive access to the storage.

demote. To remove a logical data unit from cache memory. A storage server demotes a data unit to make room for other logical data units in the cache or because the logical data unit is not valid. The ESS must destage logical data units with active write units before they can be demoted.

destaging. Movement of data from an online or higher priority to an offline or lower priority device. The ESS stages incoming data into cache and then destages it to disk.

device. In Enterprise Systems Architecture/390, a disk drive.

device adapter (DA). A physical component of the ESS that provides communication between the clusters and the storage devices. The ESS has eight device adapters that it deploys in pairs, one from each cluster. DA pairing enables the ESS to access any disk drive from either of two paths, providing fault tolerance and enhanced availability.

device address. In Enterprise Systems Architecture/390, the field of an ESCON or FICON device-level frame that selects a specific device on a control-unit image.

device ID. In the ESS, the unique two-digit hexadecimal number that identifies the logical device.

device interface card. A physical subunit of a storage cluster that provides the communication with the attached DDMs.

device number. In Enterprise Systems Architecture/390, a four-hexadecimal-character identifier, for example 13A0, that the systems administrator associates with a device to facilitate communication between the program and the host operator. The device number is associated with a subchannel.

device sparing. A subsystem function that automatically copies data from a failing DDM to a spare DDM. The subsystem maintains data access during the process.

DFS™. See *distributed file service*.

direct access storage device (DASD). (1) A mass storage medium on which a computer stores data. (2) A disk device.

disk cage. A container for disk drives. Each disk cage supports eight disk eight packs (64 disks).

disk drive. Standard term for a disk-based nonvolatile storage medium. The ESS uses hard disk drives as the primary nonvolatile storage media to store host data.

disk drive module (DDM). A field replaceable unit that consists of a single disk drive and its associated packaging.

disk drive module group. See *disk eight pack*.

disk eight pack. In the ESS, a group of eight disk drive modules (DDMs) installed as a unit in a DDM bay.

disk group. In the ESS, a collection of disk drives in the same SSA loop set up by the ESS to be available to be assigned as a RAID array. A disk group can be formatted as CKD or fixed block, and as RAID or non-RAID, or it can be left unformatted. A disk group is a logical assemblage of eight disk drives, in contrast to a disk eight pack. See also *disk eight pack*.

disk storage system. One or more storage devices that are installed with a storage software application to provide a single common pool of storage that is used to store, safeguard, retrieve, and share data. Most disk storage systems also include disaster planning and recovery options. In SDD, a disk storage system refers to an ESS, DS6000, or DS8000 device.

distributed file service (DFS). A service that provides data access over IP networks.

DNS. See *domain name system*.

domain. (1) That part of a computer network in which the data processing resources are under common control. (2) In TCP/IP, the naming system used in hierarchical networks. (3) A Copy Services server group, in other words, the set of clusters designated by the user to be managed by a particular Copy Services server.

domain name system (DNS). In TCP/IP, the server program that supplies name-to-address translation by mapping domain names to internet addresses. The address of a DNS server is the internet address of the server that hosts the DNS software for the network.

drawer. A unit that contains multiple DDMs and provides power, cooling, and related interconnection logic to make the DDMs accessible to attached host systems.

drive. (1) A peripheral device, especially one that has addressed storage media. See also *disk drive module*. (2) The mechanism used to seek, read, and write information on a storage medium.

dual-active mode. (1) With respect to a Copy Services server, the mode of operation of the server when the LIC level of the associated ESS cluster is 2.0 or higher. (2) With respect to a Copy Services domain, the mode of operation of the domain, when the Copy Services servers are dual-active servers. See also *Copy Services server*, *Copy Services domain*, *mixed mode*, and *single-active server*.

duplex. (1) Regarding ESS Copy Services, the state of a volume pair after PPRC has completed the copy operation and the volume pair is synchronized. (2) In general, pertaining to a communication mode in which data can be sent and received at the same time.

dynamic sparing. The ability of a storage server to move data from a failing disk drive module (DDM) to a spare DDM while maintaining storage functions.

E

E10. The predecessor of the F10 model of the ESS. See also *F10*.

E20. The predecessor of the F20 model of the ESS. See also *F20*.

EBCDIC. See *extended binary-coded decimal interchange code*.

EC. See *engineering change*.

ECKD™. See *extended count key data*.

eight pack. See *disk eight pack*.

electrostatic discharge (ESD). An undesirable discharge of static electricity that can damage equipment and degrade electrical circuitry.

emergency power off (EPO). A means of turning off power during an emergency, usually a switch.

EMIF. See *ESCON multiple image facility*.

enclosure. A unit that houses the components of a storage subsystem, such as a control unit, disk drives, and power source.

end of file. A coded character recorded on a data medium to indicate the end of the medium. On a CKD direct access storage device, the subsystem indicates the end of a file by including a record with a data length of zero.

engineering change (EC). An update to a machine, part, or program.

Enterprise Storage Server. See *IBM TotalStorage Enterprise Storage Server*.

Enterprise Systems Architecture/390 (ESA/390). An IBM architecture for mainframe computers and peripherals. Processor systems that follow the ESA/390 architecture include the ES/9000®® family. See also *z/Architecture®*.

Enterprise Systems Connection (ESCON). (1) An Enterprise Systems Architecture/390 and IBM System z computer peripheral interface. The I/O interface uses ESA/390 logical protocols over a serial interface that configures attached units to a communication fabric. (2) A set of IBM products and services that provide a dynamically connected environment within an enterprise.

EPO. See *emergency power off*.

ERDS. See *error-recording data set*.

error-recording data set (ERDS). On S/390 and IBM System z hosts, a data set that records data-storage and data-retrieval errors. A service information message (SIM) provides the error information for the ERDS.

ERP. See *error recovery procedure*.

error recovery procedure (ERP). Procedures designed to help isolate and, where possible, to recover from errors in equipment. The procedures are often used in conjunction with programs that record information on machine malfunctions.

ESA/390. See *Enterprise Systems Architecture/390*.

ESCD. See *ESCON director*.

ESCON. See *Enterprise System Connection*.

ESCON channel. An S/390 or IBM System z channel that supports ESCON protocols.

ESCON director (ESCD). An I/O interface switch that provides for the interconnection of multiple ESCON interfaces in a distributed-star topology.

ESCON host systems. S/390 or IBM System z hosts that attach to the ESS with an ESCON adapter. Such host systems run on operating systems that include MVS™, VSE, TPF, or versions of VM.

ESCON multiple image facility (EMIF). In Enterprise Systems Architecture/390, a function that enables LPARs to share an ESCON channel path by providing each LPAR with its own channel-subsystem image.

EsconNet. In ESS Specialist, the label on a pseudo-host icon that represents a host connection that uses the ESCON protocol and that is not completely defined on the ESS. See also *pseudo-host* and *access-any mode*.

ESD. See *electrostatic discharge*.

eServer™. See *IBM @server*.

ESS. See *IBM TotalStorage Enterprise Storage Server*.

ESS Copy Services. In the ESS, a collection of optional software features, with a Web-browser interface, used for configuring, managing, and monitoring data-copy functions.

ESS Copy Services CLI. See *Copy Services Command-Line Interface*.

ESS Expert. See *IBM TotalStorage Enterprise Storage Server Expert*.

ESS Master Console. See *IBM TotalStorage ESS Master Console*.

ESSNet. See *IBM TotalStorage Enterprise Storage Server Network*.

ESS Specialist. See *IBM TotalStorage Enterprise Storage Server Specialist*.

Expert. See *IBM TotalStorage Enterprise Storage Server Expert*.

extended binary-coded decimal interchange code (EBCDIC). A coding scheme developed by IBM used to represent various alphabetic, numeric, and special symbols with a coded character set of 256 eight-bit codes.

extended count key data (ECKD). An extension of the CKD architecture.

Extended Remote Copy (XRC). A function of a storage server that assists a control program to maintain a consistent copy of a logical volume on another storage facility. All modifications of the primary logical volume by any attached host are presented in order to a single host. The host then makes these modifications on the secondary logical volume.

extent. A continuous space on a disk that is occupied by or reserved for a particular data set, data space, or file. The unit of increment is a track. See also *multiple allegiance* and *parallel access volumes*.

F

F10. A model of the ESS featuring a single-phase power supply. It has fewer expansion capabilities than the Model F20.

F20. A model of the ESS featuring a three-phase power supply. It has more expansion capabilities than the Model F10, including the ability to support a separate expansion enclosure.

fabric. In fibre-channel technology, a routing structure, such as a switch, receives addressed information and routes to the appropriate destination. A fabric can consist of more than one switch. When multiple fibre-channel switches are interconnected, they are said to be *cascaded*.

failback. Cluster recovery from failover following repair. See also *failover*.

failover. (1) In SAN Volume Controller, the function that occurs when one redundant part of the system takes over the workload of another part of the system that has failed. (2) In the ESS, the process of transferring all control of the ESS to a single cluster in the ESS when the other cluster in the ESS fails. See also *cluster*.

fast write. A write operation at cache speed that does not require immediate transfer of data to a disk drive. The subsystem writes the data directly to cache, to nonvolatile storage, or to both. The data is then available for destaging. A fast-write operation reduces the time an application must wait for the I/O operation to complete.

FBA. See *fixed-block architecture*.

FC. See *feature code*. **Note**: *FC* is a common abbreviation for fibre channel in the industry, but the ESS customer documentation library reserves *FC* for feature code.

FC-AL. See *Fibre Channel-Arbitrated Loop*.

FCP. See *fibre-channel protocol*.

FCS. See *fibre-channel standard*.

feature code (FC). A code that identifies a particular orderable option and that is used by service personnel to process hardware and software orders. Individual optional features are each identified by a unique feature code.

fibre channel. A data-transmission architecture based on the ANSI fibre-channel standard, which supports full-duplex communication. The ESS supports data transmission over fiber-optic cable through its fibre-channel adapters. See also *fibre-channel protocol* and *fibre-channel standard*.

Fibre Channel-Arbitrated Loop (FC-AL). An implementation of the fibre-channel standard that uses a ring topology for the communication fabric. See American National Standards Institute (ANSI) X3T11/93-275. In this topology, two or more fibre-channel end points are interconnected through a looped interface. The ESS supports this topology.

fibre-channel connection (FICON). A fibre-channel communications protocol designed for IBM mainframe computers and peripherals.

fibre-channel protocol (FCP). A protocol used in fibre-channel communications with five layers that define how fibre-channel ports interact through their physical links to communicate with other ports.

fibre-channel standard (FCS). An ANSI standard for a computer peripheral interface. The I/O interface defines a protocol for communication over a serial interface that configures attached units to a communication fabric. The protocol has two layers. The IP layer defines basic interconnection protocols. The upper layer supports one or more logical protocols (for example, FCP for SCSI command protocols and SBCON for ESA/390 command protocols). See American National Standards Institute (ANSI) X3.230-199x. See also *fibre-channel protocol*.

FICON. See *fibre-channel connection*.

FiconNet. In ESS Specialist, the label on a pseudo-host icon that represents a host connection that uses the FICON protocol and that is not completely defined on the ESS. See also *pseudo-host* and *access-any mode*.

field replaceable unit (FRU). An assembly that is replaced in its entirety when any one of its components fails. In some cases, a field replaceable unit might contain other field replaceable units. (GC)

FIFO. See *first-in-first-out*.

File Transfer Protocol (FTP). In TCP/IP, an application protocol used to transfer files to and from host computers. See also *Transmission Control Protocol/Internet Protocol*.

firewall. A protection against unauthorized connection to a computer or a data storage system. The protection is usually in the form of software on a gateway server that grants access to users who meet authorization criteria.

first-in-first-out (FIFO). A queuing technique in which the next item to be retrieved is the item that has been in the queue for the longest time. (A)

fixed-block architecture (FBA). An architecture for logical devices that specifies the format of and access mechanisms for the logical data units on the device. The logical data unit is a block. All blocks on the device are the same size (fixed size). The subsystem can access them independently.

fixed-block device. An architecture for logical devices that specifies the format of the logical data units on the device. The logical data unit is a block. All blocks on the device are the same size (fixed size); the subsystem can access them independently. This is the required format of the logical data units for host systems that attach with a SCSI or fibre-channel interface. See also *fibre-channel* and *small computer systems interface*.

FlashCopy®. An optional feature for the ESS that can make an instant copy of data, that is, a point-in-time copy of a volume.

FRU. See *field replaceable unit*.

FTP. See *File Transfer Protocol*.

full duplex. See *duplex*.

G

GB. See *gigabyte*.

GDPS®®. See *Geographically Dispersed Parallel Sysplex®*.

Geographically Dispersed Parallel Sysplex™ (GDPS). An S/390 multisite application-availability solution.

gigabyte (GB). A gigabyte of storage is 10^9 bytes. A gigabyte of memory is 2^{30} bytes.

group. In ESS documentation, a nickname for two different kinds of groups, depending on the context. See *disk eight pack* or *Copy Services server group*.

H

HA. See *host adapter*.

HACMP. See *High-Availability Cluster Multiprocessing*.

hard disk drive (HDD). (1) A storage medium within a storage server used to maintain information that the storage server requires. (2) A mass storage medium for computers that is typically available as a fixed disk (such as the disks used in system units of personal computers or in drives that are external to a personal computer) or a removable cartridge.

hardware service manager (HSM). An option on an AS/400® or IBM System i® host that enables the user to display and work with system hardware resources and to debug input-output processors (IOP), input-output adapters (IOA), and devices.

HBA. See *host bus adapter*.

HDA. See *head and disk assembly*.

HDD. See *hard disk drive*.

hdisk. An AIX term for storage space.

head and disk assembly (HDA). The portion of an HDD associated with the medium and the read/write head.

heartbeat. A status report sent at regular intervals from the ESS. The service provider uses this report to monitor the health of the call home process. See also *call home*, *heartbeat call home record*, and *remote technical assistance information network*.

heartbeat call home record. Machine operating and service information sent to a service machine. These records might include such information as feature code information and product logical configuration information.

hierarchical storage management. (1) A function provided by storage management software such as Tivoli® Storage Management or Data Facility Storage Management Subsystem/MVS (DFSMS/MVS™) to automatically manage free space based on the policy that the storage administrator sets. (2) In AS/400 storage management, an automatic method to manage and distribute data between the different storage layers, such as disk units and tape library devices.

High-Availability Cluster Multiprocessing (HACMP). Software that provides host clustering, so that a failure of one host is recovered by moving jobs to other hosts within the cluster.

high-speed link (HSL). A hardware connectivity architecture that links system processors to system input/output buses and other system units.

home address (HA). A nine-byte field at the beginning of a track that contains information that identifies the physical track and its association with a cylinder.

Note: In the ESS, the acronym HA is shared between home address and host adapter. See also *host adapter*.

hop. Interswitch connection. A hop count is the number of connections that a particular block of data traverses between source and destination. For example, data traveling from one hub over a wire to another hub traverses one hop.

host. See *host system*.

host adapter (HA). A physical subunit of a storage server that provides the ability to attach to one or more host I/O interfaces. The Enterprise Storage Server has four HA bays, two in each cluster. Each bay supports up to four host adapters.

In the ESS, the acronym HA is shared between home address and host adapter. See also *home address*.

host bus adapter. An interface card that connects a host bus, such as a peripheral component interconnect (PCI) bus, to the storage area network.

host name. The Internet address of a machine in the network. In the ESS, the host name can be entered in the host definition as the fully qualified domain name of the attached host system, such as `mycomputer.city.company.com`, or as the subname of the fully qualified domain name, for example, `mycomputer`. See also *host system*.

host processor. A processor that controls all or part of a user application network. In a network, the processing unit in which the data communication access method resides. See also *host system*.

host system. A computer, either of the mainframe (S/390 or IBM system z) or of the open-systems type, that is connected to the ESS. S/390 or IBM System z hosts are connected to the ESS through ESCON or FICON interfaces. Open-systems hosts are connected to the ESS by SCSI or fibre-channel interfaces.

hot plug. Pertaining to the ability to add or remove a hardware facility or resource to a unit while power is on.

HSL. See *high-speed link*.

HSM. See *hierarchical storage management* or *Hardware Service Manager*.

I

IBM @server. The IBM brand name for a series of server products that are optimized for e-commerce. The products include the IBM System i, System p, IBM System x®, and IBM System z.

IBM product engineering (PE). The third-level of IBM service support. Product engineering is composed of IBM engineers who have experience in supporting a product or who are knowledgeable about the product.

IBM System Storage Multipath Subsystem Device Driver (SDD). Software that is designed to provide multipath configuration environment support for a host system that is attached to storage devices. SDD resides in a host system with the native disk device driver.

IBM System Storage Multipath Subsystem Device Driver Path Control Module (SDDPCM). A loadable path control module for disk storage system devices to supply path management functions and error recovery algorithms. When the disk storage system devices are configured as Multipath I/O (MPIO)-devices, SDDPCM becomes part of the AIX MPIO Fibre Channel Protocol device driver during the configuration. The AIX MPIO-capable device driver with the disk storage system SDDPCM module enhances the data availability and I/O load balancing.

IBM System Storage Subsystem Device Driver Device Specific Module (SDDDSM). An IBM storage subsystems multipath I/O solution that is based on Microsoft MPIO technology. It is a device-specific module that is designed to support IBM storage subsystems devices such as SAN Volume Controller, DS8000, and DS6000. SDDDSM resides on a host server with the native disk device driver and provides enhanced data availability, automatic path failover protection, concurrent download of controller firmware code, and path selection policies for the host system.

IBM TotalStorage Enterprise Storage Server (ESS). A member of the Seascape® product family of storage servers and attached storage devices (disk drive modules). The ESS provides for high-performance, fault-tolerant storage and management of enterprise data, providing access through multiple concurrent operating systems and communication protocols. High performance is provided by multiple symmetric multiprocessors, integrated caching, RAID support for the disk drive modules, and disk access through a high-speed serial storage architecture (SSA) interface.

IBM TotalStorage Enterprise Storage Server Expert (ESS Expert). The software that gathers performance data from the ESS and presents it through a Web browser.

IBM TotalStorage Enterprise Storage Server Specialist (ESS Specialist). Software with a Web-browser interface for configuring the ESS.

IBM TotalStorage Enterprise Storage Server Network (ESSNet). A private network providing Web browser access to the ESS. IBM installs the ESSNet software on an IBM workstation called the IBM TotalStorage ESS Master Console, supplied with the first ESS delivery.

IBM TotalStorage ESS Master Console (ESS Master Console). An IBM workstation (formerly named the ESSNet console and hereafter referred to simply as the ESS Master Console) that IBM installs to provide the ESSNet facility when they install your ESS. It includes a Web browser that provides links to the ESS user interface, including ESS Specialist and ESS Copy Services.

ID. See *identifier*.

identifier (ID). A unique name or address that identifies things such as programs, devices, or systems.

IML. See *initial microprogram load*.

implicit allegiance. In Enterprise Systems Architecture/390, a relationship that a control unit creates between a device and a channel path when the device accepts a read or write operation. The control unit guarantees access to the channel program over the set of channel paths that it associates with the allegiance.

initial microcode load (IML). The action of loading microcode for a computer into that computer's storage.

initial program load (IPL). The action of loading software into a computer, typically an operating system that controls the computer.

initiator. A SCSI device that communicates with and controls one or more targets. An initiator is typically an I/O adapter on a host computer. A SCSI initiator is analogous to an S/390 channel. A SCSI logical unit is analogous to an S/390 device. See also *target*.

i-node. The internal structure in an AIX operating system that describes the individual files in the operating system. It contains the code, type, location, and owner of a file.

input/output (I/O). Pertaining to (a) input, output, or both or (b) a device, process, or channel involved in data input, data output, or both.

input/output configuration data set. A configuration definition built by the I/O configuration program (IOCP) and stored on disk files associated with the processor controller.

interleave. In the ESS, to automatically create two striped partitions across the drives in a RAID-5 array, both of which use the count-key-data (CKD) record format.

Internet Protocol (IP). In the Internet suite of protocols, a protocol without connections that routes data through a network or interconnecting networks and acts as an intermediary between the higher protocol layers and the physical network. The upper layer supports one or more logical protocols (for example, a SCSI-command protocol and an ESA/390 command protocol). See ANSI X3.230-199x. The IP acronym is the IP in TCP/IP. See also *Transmission Control Protocol/Internet Protocol*.

invalidate. To remove a logical data unit from cache memory because it cannot support continued access to the logical data unit on the device. This removal might be the result of a failure within the storage server or a storage device that is associated with the device.

I/O. See *input/output*.

I/O adapter (IOA). In the ESS, an input-output adapter on the PCI bus.

IOCDS. See *input/output configuration data set*.

I/O device. An addressable read and write unit, such as a disk drive device, magnetic tape device, or printer.

I/O interface. An interface that enables a host to perform read and write operations with its associated peripheral devices.

I/O Priority Queueing. Facility provided by the Workload Manager of OS/390® and supported by the ESS that enables the system administrator to set priorities for queueing I/Os from different system images. See also *multiple allegiance* and *parallel access volume*.

I/O processor (IOP). Controls input-output adapters and other devices.

I/O sequential response time. The time an I/O request is queued in processor memory waiting for previous I/Os to the same volume to complete.

IOSQ. See *I/O sequential response time*.

IP. See *Internet Protocol*.

IPL. See *initial program load*.

IBM System i. An IBM @server product that emphasizes integration. It is the successor to the AS/400 family of servers.

J

Java™ Virtual Machine (JVM). A software implementation of a central processing unit (CPU) that runs compiled Java code (applets and applications). (GC)

JVM. See *Java Virtual Machine*.

K

KB. See *kilobyte*.

key field. The second (optional) field of a CKD record. The key length is specified in the count field. The key length determines the field length. The program writes the data in the key field and use the key field to identify or locate a given record. The subsystem does not use the key field.

kilobyte (KB). (1) For processor storage, real, and virtual storage, and channel volume, 2^{10} or 1024 bytes. (2) For disk storage capacity and communications volume, 1000 bytes.

Korn shell. Interactive command interpreter and a command programming language.

KPOH. See *thousands of power-on hours*.

L

LAN. See *local area network*.

last-in first-out (LIFO). A queuing technique in which the next item to be retrieved is the item most recently placed in the queue. (A)

LBA. See *logical block address*.

LCU. See *logical control unit*.

least recently used (LRU). (1) The algorithm used to identify and make available the cache space that contains the least-recently used data. (2) A policy for a caching algorithm that chooses to remove from cache the item that has the longest elapsed time since its last access.

LED. See *light-emitting diode*.

LIC. See *Licensed Internal Code*.

Licensed Internal Code (LIC). Microcode that IBM does not sell as part of a machine, but licenses to the customer. LIC is implemented in a part of storage that is not addressable by user programs. Some IBM products use it to implement functions as an alternate to hard-wired circuitry. See also *licensed machine code (LMC)*.

LIFO. See *last-in first-out*.

light-emitting diode (LED). A semiconductor chip that gives off visible or infrared light when activated.

LMC. See *licensed machine code*.

licensed machine code (LMC). Microcode, basic input/output system code (BIOS), utility programs, device drivers, and diagnostics delivered with an IBM machine.

link address. On an ESCON or FICON interface, the portion of a source or destination address in a frame that ESCON or FICON uses to route a frame through an ESCON or FICON director. ESCON or FICON associates the link address with a specific switch port that is on the ESCON or FICON director. Equivalently, it associates the link address with the channel subsystem or control unit link-level functions that are attached to the switch port.

link-level facility. The ESCON or FICON hardware and logical functions of a control unit or channel subsystem that allow communication over an ESCON or FICON write interface and an ESCON or FICON read interface.

local area network (LAN). A computer network located on a user's premises within a limited geographic area.

local e-mail. An e-mail configuration option for storage servers that are connected to a host-system network that does not have a domain name system (DNS) server.

logical address. On an ESCON or FICON interface, the portion of a source or destination address in a frame used to select a specific channel-subsystem or control-unit image.

logical block address (LBA). The address assigned by the ESS to a sector of a disk.

logical control unit (LCU). See *control-unit image*.

logical data unit. A unit of storage that is accessible on a given device.

logical device. The facilities of a storage server (such as the ESS) associated with the processing of I/O operations directed to a single host-accessible emulated I/O device. The associated storage is referred to as a logical volume. The logical device is mapped to one or

more host-addressable units, such as a device on an S/390 I/O interface or a logical unit on a SCSI I/O interface, such that the host initiating I/O operations to the I/O-addressable unit interacts with the storage on the associated logical device.

logical partition (LPAR). In Enterprise Systems Architecture/390, a set of functions that create the programming environment in which more than one logical partition (LPAR) is established on a processor. An LPAR is conceptually similar to a virtual machine environment except that the LPAR is a function of the processor. Also, the LPAR does not depend on an operating system to create the virtual machine environment.

logical path. In the ESS for Copy Services, a relationship between a source logical subsystem and target logical subsystem that is created over a physical path through the interconnection fabric used for Copy Services functions.

logical subsystem (LSS). In the ESS, a topological construct that consists of a group of up to 256 logical devices. An ESS can have up to 16 CKD-formatted logical subsystems (4096 CKD logical devices) and also up to 16 fixed-block (FB) logical subsystems (4096 FB logical devices). The logical subsystem facilitates configuration of the ESS and might have other implications relative to the operation of certain functions. There is a one-to-one mapping between a CKD logical subsystem and an S/390 control-unit image.

For S/390 or IBM System z hosts, a logical subsystem represents a logical control unit (LCU). Each control-unit image is associated with only one logical subsystem. See also *control-unit image*.

logical unit. In open systems, a logical disk drive.

logical unit number (LUN). In the SCSI protocol, a unique number used on a SCSI bus to enable it to differentiate between up to eight separate devices, each of which is a logical unit.

logical volume. The storage medium associated with a logical disk drive. A logical volume typically resides on one or more storage devices. The ESS administrator defines this unit of storage. The logical volume, when residing on a RAID array, is spread over the drives in the array.

logical volume manager (LVM). A set of system commands, library routines, and other tools that allow the user to establish and control logical volume storage. The LVM maps data between the logical view of storage space and the physical disk drive module (DDM).

longitudinal redundancy check (LRC). (1) A method of error-checking during data transfer that involves checking parity on a row of binary digits that are

members of a set that forms a matrix. Longitudinal redundancy check is also called a longitudinal parity check. (2) In the ESS, a mechanism that the ESS uses for locating errors. The LRC checks the data as it progresses from the host, through the ESS controller, into the device adapter, and to the array.

longwave laser adapter. A connector used between a host and the ESS to support longwave fibre-channel communication.

loop. The physical connection between a pair of device adapters in the ESS. See also *device adapter*.

LPAR. See *logical partition*.

LRC. See *longitudinal redundancy check*.

LRU. See *least recently used*.

LSS. See *logical subsystem*.

LUN. See *logical unit number*.

LVM. See *logical volume manager*.

M

machine level control (MLC. A database that contains the EC level and configuration of products in the field.

machine reported product data (MRPD). Product data gathered by a machine and sent to a destination such as an IBM support server or RETAIN®. These records might include such information as feature code information and product logical configuration information.

mainframe. A computer, usually in a computer center, with extensive capabilities and resources to which other computers may be connected so that they can share facilities.

maintenance analysis procedure (MAP). A hardware maintenance document that gives an IBM service representative a step-by-step procedure for tracing a symptom to the cause of a failure.

Management Information Base (MIB). (1) A collection of objects that can be accessed by means of a network management protocol. (GC) (2) In the ESS, the MIB record conforms to the Open Systems Interconnection (OSI) standard defined by the International Organization for Standardization (ISO) for the exchange of information. See also *simple network management protocol*.

MAP. See *maintenance analysis procedure*.

mass storage. The various techniques and devices for storing large amounts of data in a persisting and machine-readable fashion. Mass storage devices include all types of disk drives and tape drives.

master boot record. The boot program that the BIOS loads. This boot program is located in the first sector of the hard disk and is used to start the boot process

Master Console. See *IBM TotalStorage ESS Master Console*.

MB. See *megabyte*.

MBR. See *master boot record*.

MCA. See *Micro Channel® architecture*.

mean time between failures (MTBF). (1) A projection of the time that an individual unit remains functional. The time is based on averaging the performance, or projected performance, of a population of statistically independent units. The units operate under a set of conditions or assumptions. (2) For a stated period in the life of a functional unit, the mean value of the lengths of time between consecutive failures under stated conditions. (I) (A)

medium. For a storage facility, the disk surface on which data is stored.

megabyte (MB). (1) For processor storage, real and virtual storage, and channel volume, 2^{20} or 1 048 576 bytes. (2) For disk storage capacity and communications volume, 1 000 000 bytes.

MES. See *miscellaneous equipment specification*.

MIB. See *management information base*.

Micro Channel architecture (MCA). The rules that define how subsystems and adapters use the Micro Channel bus in a computer. The architecture defines the services that each subsystem can or must provide.

Microsoft Internet Explorer (MSIE). Web browser software manufactured by Microsoft.

migration. In the ESS, the replacement of a system or subsystem with a different type of system or subsystem, such as replacing a SCSI host adapter with a fibre-channel host adapter. When used in the context of data migration regarding the ESS, the transfer of data from one storage facility to another, such as from a 3390 to the ESS.

MIH. See *missing-interrupt handler*.

mirrored pair. Two units that contain the same data. The system refers to them as one entity.

mirroring. In host systems, the process of writing the same data to two disk units within the same auxiliary storage pool at the same time.

miscellaneous equipment specification (MES). IBM field-installed change to a machine.

missing-interrupt handler (MIH). An MVS and MVS/XA™ facility that tracks I/O interrupts. MIH informs the operator and creates a record whenever an expected interrupt fails to occur before a specified elapsed time is exceeded.

mixed mode. With respect to a Copy Services domain, the mode of operation of the domain when one Copy Services server is a dual-active server and the other Copy Services server is a single-active server. See also *Copy Services server*, *dual-active server*, and *single-active server*.

MLC. See *machine level control*.

mobile solutions terminal (MoST). The mobile terminal used by service personnel.

mode conditioning patch. This cable is used to convert a single mode signal generated by a longwave adapter into a light signal that is appropriate for multimode fibre. Another mode conditioning patch cable is required at the terminating end of the multimode fibre to convert the signal back to single mode light sent into a longwave adapter.

Model 100. A 2105 Model 100, often simply referred to as a Mod 100, is an expansion enclosure for the ESS. See also *2105*.

MoST. See *mobile solutions terminal*.

MRPD. See *machine reported product data*.

MSA. See *multiport serial adapter*.

MSIE. See *Microsoft Internet Explorer*.

MTBF. See *mean time between failures*.

multiple allegiance. An ESS hardware function that is independent of software support. This function enables multiple system images to concurrently access the same logical volume on the ESS as long as the system images are accessing different extents. See also *extent* and *parallel access volumes*.

multiple virtual storage (MVS). Implies MVS/390, MVS/XA, MVS/ESA™, and the MVS element of the OS/390 operating system.

multiplex. The action of transmitting simultaneously.

multiport serial adapter (MSA). An adapter on the ESS Master Console that has multiple ports to which ESSs can be attached.

MVS. See *multiple virtual storage*.

N

name server. A server that stores names of the participating ESS clusters.

Netfinity®. IBM Intel-processor-based server; predecessor to the IBM xSeries® server.

Netscape Navigator. Web browser software manufactured by Netscape.

Network Installation Management (NIM). An environment that provides installation and configuration of software within a network interface.

NIM. See *Network Installation Management*.

node. (1) In a network, a point at which one or more functional units connect channels or data circuits. An ESS is a node in a fibre-channel network. (2) One SAN Volume Controller. Each node provides virtualization, cache, and Copy Services to the storage area network.

node fallover. See *failover*.

non-RAID. A disk drive set up independently of other disk drives and not set up as part of a disk eight pack to store data using the redundant array of disks (RAID) data-striping methodology.

nonremovable medium. A recording medium that cannot be added to or removed from a storage device.

nonvolatile storage (NVS). In the ESS, memory that stores active write data to avoid data loss in the event of a power loss.

NVS. See *nonvolatile storage*.

O

octet. In Internet Protocol (IP) addressing, one of the four parts of a 32-bit integer presented in dotted decimal notation. dotted decimal notation consists of four 8-bit numbers written in base 10. For example, `9.113.76.250` is an IP address containing the octets 9, 113, 76, and 250.

OEMI. See *original equipment manufacturer's information*.

open system. A system whose characteristics comply with standards made available throughout the industry and that therefore can be connected to other systems complying with the same standards. Applied to the ESS, such systems are those hosts that connect to the ESS through SCSI or FCP protocols. See also *small computer system interface* and *fibre-channel protocol*.

operating system (OS). A set of programs that control how the system works. Controls the running of programs and provides such services as resource allocation, scheduling, input and output control, and data management.

organizationally unique identifier (OUI). An IEEE-standards number that identifies an organization with a 24-bit globally unique assigned number

referenced by various standards. OUI is used in the family of 802 LAN standards, such as Ethernet and Token Ring.

original equipment manufacturer's information (OEMI). A reference to an IBM guideline for a computer peripheral interface. The interface uses ESA/390 logical protocols over an I/O interface that configures attached units in a multidrop bus topology.

OS. See *operating system*.

OS/390. The IBM operating system that includes and integrates functions previously provided by many IBM software products (including the MVS operating system) for the IBM S/390 family of enterprise servers.

OS/400®. The IBM operating system that runs the IBM AS/400 and IBM System i @serverfamilies of servers.

OUI. See *organizationally unique identifier*.

P

panel. The formatted display of information that appears on a display screen.

parallel access volume (PAV). An advanced function of the ESS that enables OS/390 and z/OS® systems to issue concurrent I/O requests against a CKD logical volume by associating multiple devices of a single control-unit image with a single logical device. Up to eight device addresses can be assigned to a PAV. The PAV function enables two or more concurrent write operations to the same logical volume, as long as the write operations are not to the same extents. See also *extent*, *I/O Priority Queueing*, and *multiple allegiance*.

parity. A data checking scheme used in a computer system to ensure the integrity of the data. The RAID implementation uses parity to re-create data if a disk drive fails.

path group. In ESA/390 architecture, a set of channel paths that are defined to a control unit as being associated with a single logical partition (LPAR). The channel paths are in a group state and are online to the host. See also *logical partition*.

path group identifier. In ESA/390 architecture, the identifier that uniquely identifies a given logical partition (LPAR). The path group identifier is used in communication between the LPAR program and a device. The identifier associates the path group with one or more channel paths, thereby defining these paths to the control unit as being associated with the same LPAR. See also *logical partition*.

PAV. See *parallel access volume*.

PCI. See *peripheral component interconnect*.

PE. See *IBM product engineering*.

Peer-to-Peer Remote Copy (PPRC). A function of a storage server that constantly updates a secondary copy of a logical volume to match changes made to a primary logical volume. The primary and secondary volumes can be on the same storage server or on separate storage servers. See also *synchronous PPRC* and *PPRC Extended Distance*.

peripheral component interconnect (PCI). An architecture for a system bus and associated protocols that supports attachments of adapter cards to a system backplane.

persistent binding. A feature where a device has the same identification to the operating system after it restarts and after other devices are added to the operating system.

physical path. A single path through the I/O interconnection fabric that attaches two units. For Copy Services, this is the path from a host adapter on one ESS (through cabling and switches) to a host adapter on another ESS.

point-to-point connection. For fibre-channel connections, a topology that enables the direct interconnection of ports. See *arbitrated loop* and *switched fabric*.

port. In the ESS, a physical connection on a host adapter to the cable that connects the ESS to hosts, switches, or another ESS. The ESS uses SCSI and ESCON host adapters that have two ports per adapter, and fibre-channel host adapters that have one port. See also *ESCON*, *fibre channel*, *host adapter*, and *small computer system interface*.

POST. See *power-on self test*.

power-on self test (POST). A diagnostic test that servers or computers run when they are turned on.

PPRC. See *Peer-to-Peer Remote Copy*.

PPRC Extended Distance. An optional feature for the ESS that maintains a fuzzy copy of a logical volume on the same ESS or on another ESS. In other words, all modifications that any attached host performs on the primary logical volume are also performed on the secondary logical volume at a later point in time. The original order of update is not strictly maintained. See also *Peer-to-Peer Remote Copy (PPRC)* and *synchronous PPRC*.

PPRC-XD. See *PPRC Extended Distance*.

predictable write. A write operation that can cache without knowledge of the existing format on the medium. All write operations on FBA DASD devices are predictable. On CKD DASD devices, a write operation is predictable if it does a format write operation for the first data record on the track.

primary Copy Services server. One of two Copy Services servers in a Copy Services server group. The primary Copy Services server is the active Copy Services server until it fails; it is then replaced by the backup Copy Services server. A Copy Services server is software that runs in one of the two clusters of an ESS and performs data-copy operations within that group. See *active Copy Services server* and *backup Copy Services server*.

product engineering. See *IBM product engineering*.

program. On a computer, a generic term for software that controls the operation of the computer. Typically, the program is a logical assemblage of software modules that perform multiple related tasks.

program-controlled interruption. An interruption that occurs when an I/O channel fetches a channel command word with the program-controlled interruption flag on.

program temporary fix (PTF). A temporary solution or bypass of a problem diagnosed by IBM in a current unaltered release of a program. (GC)

promote. To add a logical data unit to cache memory.

protected volume. In the IBM AS/400® platform, a disk storage device that is protected from data loss by RAID techniques. An AS/400 host does not mirror a volume configured as a protected volume, while it does mirror all volumes configured as unprotected volumes. The ESS, however, can be configured to indicate that an AS/400 volume is protected or unprotected and give it RAID protection in either case.

System p. The product name of an IBM @server product that emphasizes performance. It is the successor to the IBM RS/6000® family of servers.

pseudo-host. A host connection that is not explicitly defined to the ESS and that has access to at least one volume that is configured on the ESS. The FiconNet pseudo-host icon represents the FICON protocol. The EsconNet pseudo-host icon represents the ESCON protocol. The pseudo-host icon labelled "Anonymous" represents hosts connected through the FCP protocol. *Anonymous host* is a commonly used synonym for *pseudo-host*. The ESS adds a pseudo-host icon only when the ESS is set to access-any mode. See also *access-any mode*.

PTF. See *program temporary fix*.

PV Links. Short for Physical Volume Links, an alternate pathing solution from Hewlett-Packard providing for multiple paths to a volume, as well as static load balancing.

R

R0.   See *track-descriptor record*.

rack.   See *enclosure*.

RAID.   See *redundant array of independent disks*. RAID is also commonly expanded to redundant array of *independent* disks. See also *array*.

RAID 5.   A type of RAID that optimizes cost-effective performance while emphasizing use of available capacity through data striping. RAID 5 provides fault tolerance for up to two failed disk drives by distributing parity across all the drives in the array plus one parity disk drive. The ESS automatically reserves spare disk drives when it assigns arrays to a device adapter pair (DA pair). See also *device adapter*, *RAID 10*, and *redundant array of independent disks*.

RAID 10.   A type of RAID that optimizes high performance while maintaining fault tolerance for up to two failed disk drives by striping volume data across several disk drives and mirroring the first set of disk drives on an identical set. The ESS automatically reserves spare disk drives when it assigns arrays to a device adapter pair (DA pair). See also *device adapter*, *RAID 5*, and *redundant array of independent disks*.

random access.   A mode of accessing data on a medium in a manner that requires the storage device to access nonconsecutive storage locations on the medium.

rank.   See *array*.

redundant array of independent disks (RAID).   A methodology of grouping disk drives for managing disk storage to insulate data from a failing disk drive.

remote technical assistance information network (RETAIN).   The initial service tracking system for IBM service support, which captures heartbeat and call-home records. See also *support catcher* and *support catcher telephone number*.

REQ/ACK.   See *request for acknowledgement and acknowledgement*.

request for acknowledgement and acknowledgement (REQ/ACK).   A cycle of communication between two data transport devices for the purpose of verifying the connection, which starts with a request for acknowledgement from one of the devices and ends with an acknowledgement from the second device. The REQ and ACK signals help to provide uniform timing to support synchronous data transfer between an initiator and a target. The objective of a synchronous data transfer method is to minimize the effect of device and cable delays.

reserved allegiance.   In Enterprise Systems Architecture/390, a relationship that is created in a control unit between a device and a channel path when the device completes a Sense Reserve command. The allegiance causes the control unit to guarantee access (busy status is not presented) to the device. Access is over the set of channel paths that are associated with the allegiance; access is for one or more channel programs until the allegiance ends.

RETAIN.   See *remote technical assistance information network*.

S

S/390.   IBM enterprise servers based on Enterprise Systems Architecture/390 (ESA/390). *S/390* is the currently accepted shortened form of the original name *System/390*®.

S/390 storage.   (1) Storage arrays and logical volumes that are defined in the ESS as connected to S/390 servers. This term is synonymous with count-key-data (CKD) storage. (2) In ESS documentation, when noted, the term can refer to both S/390 and IBM System z storage. See also *IBM System z storage*.

SAID.   See *system adapter identification number*.

SAM.   See *sequential access method*.

SAN.   See *storage area network*.

SAS.   See *Serial Attached SCSI*.

SBCON.   See *Single-Byte Command Code Sets Connection*.

screen.   The physical surface of a display device upon which information is shown to users.

SCSI.   See *small computer system interface*.

SCSI device.   A disk drive connected to a host through an I/O interface using the SCSI protocol. A SCSI device is either an initiator or a target. See also *initiator* and *small computer system interface*.

SCSI host systems.   Host systems that are attached to the ESS with a SCSI interface. Such host systems run on UNIX, OS/400, Windows NT, Windows 2000, or Novell NetWare operating systems.

SCSI ID.   A unique identifier assigned to a SCSI device that is used in protocols on the SCSI interface to identify or select the device. The number of data bits on the SCSI bus determines the number of available SCSI IDs. A wide interface has 16 bits, with 16 possible IDs.

SCSI-FCP.   Synonym for fibre-channel protocol, a protocol used to transport data between an open-systems host and a fibre-channel adapter on an ESS. See also *fibre-channel protocol* and *small computer system interface*.

SDD. See *IBM System Storage Enterprise Storage Server Subsystem Device Driver*.

SDDDSM. See *IBM System Storage Subsystem Device Driver Device Specific Module*.

SDDPCM. See *IBM System Storage Multipath Subsystem Device Driver Path Control Module*.

Seascape architecture. A storage system architecture developed by IBM for open-systems servers and S/390 and IBM System z host systems. It provides storage solutions that integrate software, storage management, and technology for disk, tape, and optical storage.

self-timed interface (STI). An interface that has one or more conductors that transmit information serially between two interconnected units without requiring any clock signals to recover the data. The interface performs clock recovery independently on each serial data stream and uses information in the data stream to determine character boundaries and inter-conductor synchronization.

sequential access. A mode of accessing data on a medium in a manner that requires the storage device to access consecutive storage locations on the medium.

sequential access method (SAM). An access method for storing, deleting, or retrieving data in a continuous sequence based on the logical order of the records in the file.

Serial Attached SCSI (SAS). A data transfer technology that uses a host bus adapter with four or more channels that operate simultaneously. Each full-duplex channel, known as a SAS port, transfers data in each direction.

serial connection. A method of device interconnection for determining interrupt priority by connecting the interrupt sources serially.

serial storage architecture (SSA). An IBM standard for a computer peripheral interface. The interface uses a SCSI logical protocol over a serial interface that configures attached targets and initiators in a ring topology. See also *SSA adapter*.

server. (1) A host that provides certain services to other hosts that are referred to as clients. (2) A functional unit that provides services to one or more clients over a network. (GC)

service boundary. A category that identifies a group of components that are unavailable for use when one of the components of the group is being serviced. Service boundaries are provided on the ESS, for example, in each host bay and in each cluster.

service information message (SIM). A message sent by a storage server to service personnel through an S/390 operating system.

service personnel. A generalization referring to individuals or companies authorized to service the ESS. The terms "service provider", "service representative", and "IBM service support representative (SSR)" refer to types of service personnel. See also *service support representative*.

service processor. A dedicated processing unit used to service a storage facility.

service support representative (SSR). Individuals or a company authorized to service the ESS. This term also refers to a service provider, a service representative, or an IBM service support representative (SSR). An IBM SSR installs the ESS.

shared storage. In an ESS, storage that is configured so that multiple hosts can concurrently access the storage. The storage has a uniform appearance to all hosts. The host programs that access the storage must have a common model for the information on a storage device. The programs must be designed to handle the effects of concurrent access.

shortwave laser adapter. A connector used between host and ESS to support shortwave fibre-channel communication.

SIM. See *service information message*.

Simple Network Management Protocol (SNMP). In the Internet suite of protocols, a network management protocol that is used to monitor routers and attached networks. SNMP is an application layer protocol. Information on devices managed is defined and stored in the application's Management Information Base (MIB). (GC) See also *management information base*.

simplex volume. A volume that is not part of a FlashCopy, XRC, or PPRC volume pair.

single-active mode. (1) With respect to a Copy Services server, the mode of operation of the server when the LIC level of the associated ESS cluster is below 2.0. (2) With respect to a Copy Services domain, the mode of operation of the domain when the Copy Services servers are single-active servers. See also *Copy Services server*, *Copy Services domain*, *dual-active server*, and *mixed mode*.

Single-Byte Command Code Sets Connection (SBCON). The ANSI standard for the ESCON or FICON I/O interface.

small computer system interface (SCSI). A standard hardware interface that enables a variety of peripheral devices to communicate with one another. (GC)

smart relay host. A mail relay or mail gateway that has the capability to correct e-mail addressing problems.

SMIT. See *System Management Interface Tool*.

SMP. See *symmetric multiprocessor*.

SMS. See *Systems Management Server*.

SNMP. See *simple network management protocol*.

Systems Management Server (SMS). Change and configuration management software from Microsoft that runs on the Microsoft platform and distributes relevant software and updates to users.

software transparency. Criteria applied to a processing environment that states that changes do not require modifications to the host software in order to continue to provide an existing function.

spare. A disk drive on the ESS that can replace a failed disk drive. A spare can be predesignated to allow automatic dynamic sparing. Any data preexisting on a disk drive that is invoked as a spare is destroyed by the dynamic sparing copy process.

spatial reuse. A feature of serial storage architecture that enables a device adapter loop to support many simultaneous read/write operations. See also *serial storage architecture*.

Specialist. See *IBM TotalStorage Enterprise Storage Server Specialist*.

Shared Product Object Tree (SPOT) . (1) A version of the /usr file system that diskless clients mount as their own /usr directory. (2) For NIM, a /usr file system or an equivalent file system that is exported by servers in the NIM environment for remote client use.

SPOT. See *Shared Product Object Tree*.

SSA. See *serial storage architecture*.

SSA adapter. A physical adapter based on serial storage architecture. SSA adapters connect disk drive modules to ESS clusters. See also *serial storage architecture*.

SSID. See *subsystem identifier*.

SSR. See *service support representative*.

stacked status. In Enterprise Systems Architecture/390, the condition when the control unit is in a holding status for the channel, and the last time the control unit attempted to present the status, the channel responded with the stack-status control.

stage operation. The operation of reading data from the physical disk drive into the cache.

staging. To move data from an offline or low-priority device back to an online or higher priority device, usually on demand of the system or on request of the user.

standard volume. In the ESS, a volume that emulates one of several S/390 volume types, including 3390-2, 3390-3, 3390-9, 3390-2 (3380-track mode), or 3390-3 (3380-track mode), by presenting the same number of cylinders and capacity to the host as provided by the native S/390 volume type of the same name.

STI. See *self-timed interface*.

storage area network. A network that connects a company's heterogeneous storage resources.

storage complex. Multiple storage facilities.

storage device. A physical unit that provides a mechanism to store data on a given medium such that it can be subsequently retrieved. See also *disk drive module*.

storage facility. (1) A physical unit that consists of a storage server integrated with one or more storage devices to provide storage capability to a host computer. (2) A storage server and its attached storage devices.

storage server. A physical unit that manages attached storage devices and provides an interface between them and a host computer by providing the function of one or more logical subsystems. The storage server can provide functions that are not provided by the storage device. The storage server has one or more clusters.

striping. A technique that distributes data in bit, byte, multibyte, record, or block increments across multiple disk drives.

subchannel. A logical function of a channel subsystem associated with the management of a single device.

Subsystem Device Driver. See *IBM System Storage Multipath Subsystem Device Driver*.

Subsystem Device Driver Device Specific Module (SDDDSM). An IBM storage subsystems multipath I/O solution that is based on Microsoft MPIO technology. It is a device-specific module that is designed to support IBM storage subsystems devices such as SAN Volume Controller, DS8000, and DS6000. SDDDSM resides on a host server with the native disk device driver and provides enhanced data availability, automatic path failover protection, concurrent download of controller firmware code, and path selection policies for the host system.

subsystem identifier (SSID). A number that uniquely identifies a logical subsystem within a computer installation.

support catcher. See *catcher*.

support catcher telephone number. The telephone number that connects the support catcher server to the

ESS to receive a trace or dump package. See also *support catcher* and *remote technical assistance information network*.

switched fabric. In the ESS, one of three fibre-channel connection topologies that the ESS supports. See also *arbitrated loop* and *point-to-point*.

symmetric multiprocessor (SMP). An implementation of a multiprocessor computer consisting of several identical processors configured in a way that any subset of the set of processors is capable of continuing the operation of the computer. The ESS contains four processors set up in SMP mode.

synchronous PPRC. A function of a storage server that maintains a consistent copy of a logical volume on the same storage server or on another storage server. All modifications that any attached host performs on the primary logical volume are also performed on the secondary logical volume. See also *Peer-to-Peer Remote Copy* and *PPRC Extended Distance*.

synchronous write. A write operation whose completion is indicated after the data has been stored on a storage device.

System/390. See *S/390*.

system adapter identification number (SAID). In the ESS, the unique identification number automatically assigned to each ESS host adapter for use by ESS Copy Services.

System Management Interface Tool (SMIT). An interface tool of the AIX operating system for installing, maintaining, configuring, and diagnosing tasks.

System Modification Program. A program used to install software and software changes on MVS systems.

T

TAP. See *Telocator Alphanumeric Protocol*.

target. A SCSI device that acts as a slave to an initiator and consists of a set of one or more logical units, each with an assigned logical unit number (LUN). The logical units on the target are typically I/O devices. A SCSI target is analogous to an S/390 control unit. A SCSI initiator is analogous to an S/390 channel. A SCSI logical unit is analogous to an S/390 device. See also *small computer system interface*.

TB. See *terabyte*.

TCP/IP. See *Transmission Control Protocol/Internet Protocol*.

Telocator Alphanumeric Protocol (TAP). An industry standard protocol for the input of paging requests.

terabyte (TB). (1) Nominally, 1 000 000 000 000 bytes, which is accurate when speaking of bandwidth and disk storage capacity. (2) For ESS cache memory, processor storage, real and virtual storage, a terabyte refers to 2^{40} or 1 099 511 627 776 bytes.

terminal emulator. In the ESS, a function of the ESS Master Console that allows it to emulate a terminal.

thousands of power-on hours (KPOH). A unit of time used to measure the mean time between failures (MTBF).

time sharing option (TSO). An operating system option that provides interactive time sharing from remote terminals.

TPF. See *transaction processing facility*.

track. A unit of storage on a CKD device that can be formatted to contain a number of data records. See also *home address*, *track-descriptor record*, and *data record*.

track-descriptor record (R0). A special record on a track that follows the home address. The control program uses it to maintain certain information about the track. The record has a count field with a key length of zero, a data length of 8, and a record number of 0. This record is sometimes referred to as R0.

transaction processing facility (TPF). A high-availability, high-performance IBM operating system, designed to support real-time, transaction-driven applications. The specialized architecture of TPF is intended to optimize system efficiency, reliability, and responsiveness for data communication and database processing. TPF provides real-time inquiry and updates to a large, centralized database, where message length is relatively short in both directions, and response time is generally less than three seconds. Formerly known as the Airline Control Program/Transaction Processing Facility (ACP/TPF).

Transmission Control Protocol (TCP). A communications protocol used in the Internet and in any network that follows the Internet Engineering Task Force (IETF) standards for internetwork protocol. TCP provides a reliable host-to-host protocol between hosts in packet-switched communications networks and in interconnected systems of such networks. It uses the Internet Protocol (IP) as the underlying protocol.

Transmission Control Protocol/Internet Protocol (TCP/IP). (1) A combination of data-transmission protocols that provide end-to-end connections between applications over interconnected networks of different types. (2) A suite of transport and application protocols that run over the Internet Protocol. (GC) See also *Internet Protocol* and *Transmission Control Protocol*.

transparency. See *software transparency*.

TSO. See *time sharing option*.

U

UFS. UNIX filing system.

Ultra-SCSI. An enhanced small computer system interface.

unconfigure. To delete the configuration.

unit address. In Enterprise Systems Architecture/390, the address associated with a device on a given control unit. On ESCON or FICON interfaces, the unit address is the same as the device address. On OEMI interfaces, the unit address specifies a control unit and device pair on the interface.

UNIX File System (UFS) . A section of the UNIX file tree that is physically contained on a single device or disk partition and that can be separately mounted, dismounted, and administered.

unprotected volume. An AS/400 term that indicates that the AS/400 host recognizes the volume as an unprotected device, even though the storage resides on a RAID array and is therefore fault tolerant by definition. The data in an unprotected volume can be mirrored. Also referred to as an *unprotected device*.

upper-layer protocol. The layer of the Internet Protocol (IP) that supports one or more logical protocols (for example, a SCSI-command protocol and an ESA/390 command protocol). See ANSI X3.230-199x.

UTC. See *Coordinated Universal Time*.

utility device. The ESA/390 term for the device used with the Extended Remote Copy facility to access information that describes the modifications performed on the primary copy.

V

virtual machine facility. A virtual data processing machine that appears to the user to be for the exclusive use of that user, but whose functions are accomplished by sharing the resources of a shared data processing system. An alternate name for the VM/370 IBM operating system.

virtualization. In the storage industry, a concept in which a pool of storage is created that contains several disk subsystems. The subsystems can be from various vendors. The pool can be split into virtual disks that are visible to the host systems that use them. In SDD, virtualization product refers to SAN Volume Controller.

vital product data (VPD). Information that uniquely defines the system, hardware, software, and microcode elements of a processing system.

VM. The root name of several IBM operating systems, such as VM/370, VM/ESA®, VM/CMS, and VM/SP. See also *virtual machine (VM) facility*.

volume. In Enterprise Systems Architecture/390, the information recorded on a single unit of recording medium. Indirectly, it can refer to the unit of recording medium itself. On a nonremovable-medium storage device, the term can also indirectly refer to the storage device associated with the volume. When multiple volumes are stored on a single storage medium transparently to the program, the volumes can be referred to as logical volumes.

VPD. See *vital product data*.

VSE/ESA™. IBM operating system, the letters of which represent virtual storage extended/enterprise systems architecture.

W

Web Copy Services. See *ESS Copy Services*.

worldwide node name (WWNN). A unique 64-bit identifier for a host containing a fibre-channel port. See also *worldwide port name*.

worldwide port name (WWPN). A unique 64-bit identifier associated with a fibre-channel adapter port. It is assigned in an implementation- and protocol-independent manner.

write hit. A write operation in which the requested data is in the cache.

write penalty. The performance impact of a classical RAID 5 write operation.

WWPN. See *worldwide port name*.

X

XD. See *PPRC Extended Distance*.

XRC. See *Extended Remote Copy*.

xSeries. The product name of an IBM @server product that emphasizes industry-standard server scalability and self-managing server technologies. It is the successor to the Netfinity family of servers.

Z

z/Architecture. An IBM architecture for mainframe computers and peripherals. The IBM @server IBM System z family of servers uses the z/Architecture architecture. It is the successor to the S/390 and 9672 family of servers. See also *Enterprise Systems Architecture/390*.

z/OS. An operating system for the IBM eServer product line that supports 64-bit real storage.

IBM System z. (1) An IBM @server family of servers that emphasizes near-zero downtime. (2) IBM enterprise servers based on z/Architecture.

IBM System z storage. Storage arrays and logical volumes that are defined in the ESS as connected to IBM System z servers. See also *S/390 storage*.

Index

IBM ®

Printed in USA