HITACHI
Inspire the Next

Hitachi NAS Platform™ and Hitachi High-performance NAS Platform™, powered by BlueArc

Storage Subsystem Guide

Hitachi Data Systems

MK-96BA006-06

Trademarks

Hitachi is a registered trademark of Hitachi, Ltd. in the United States and other countries.

Hitachi Data Systems is a registered trademark and service mark of Hitachi, Ltd., and the Hitachi Data Systems design mark is a trademark and service mark of Hitachi, Ltd.

High-Performance NAS Platform is a registered trademark of Hitachi Data Systems

HiCommand is a registered trademark of Hitachi, Ltd.

This product includes software developed by the OpenSSL Project for use in the OpenSSL Toolkit (http://www.openssl.org/). Some parts of ADC use open source code from NetApp, Inc. and Traakan, Inc.

The product described in this guide may be protected by one or more U.S. patents, foreign patents, or pending applications.

BlueArc is a registered trademark of BlueArc Corporation in the United States and other countries. The following are trademarks licensed to BlueArc Corporation, registered in the USA and other countries: BlueArc™, the BlueArc™ logo and the BlueArc Storage System™.

Microsoft, MS-DOS, Windows, Windows NT, Windows 2000/2003/2008, and WIndows XP are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries.

UNIX is a registered trademark in the United States and other countries, licensed exclusively through The Open Group.

All other trademarks appearing in this document are the property of their respective owners.

Notice of Export Controls

Export of technical data contained in this document may require an export license from the United States government and/or the government of Japan. Contact the Hitachi Data Systems Legal Department for any export compliance questions.

Document Revision Level

Revision	Date	Description
00	June 2007	Initial release
01	November 2007	Revision 01, supersedes and replaces MK-97BA006-00
02	May 2008	Revision 02, supersedes and replaces MK-97BA006-01
03	October 2008	Revision 03, supersedes and replaces MK-97BA006-02
04	June 2009	Revision 04, supersedes and replaces MK-97BA006-03
05	August 2009	Revision 05, supersedes and replaces MK-97BA006-04
06	January 2010	Revision 06, supersedes and replaces MK-97BA006-05

Contact

Hitachi Data Systems
750 Central Expressway
Santa Clara, California 95050-2627
https://portal.hds.com

North America

800-446-0744

Outside of North America

1-858-547-4526

About This Guide

The following types of messages are used throughout this guide. We recommend that these messages are read and clearly understood before proceeding.

Tip: A tip contains supplementary information that is useful in completing a task.

Note: A note contains information that helps to install or operate the system effectively.

Caution: A caution indicates the possibility of damage to data or equipment. Do not proceed beyond a caution message until the requirements are fully understood.

Other Related Documents

- Hitachi NAS Platform™ ***and*** Hitachi High-performance NAS Platform™ ***System Administration Guide***: In PDF and HTML format, this guide provides a full specification of the system and instructions on how to administer the High-performance NAS Platform using Web Manager.
- Hitachi NAS Platform™ ***Hardware Reference***: This guide (in PDF format) provides an overview of the hardware, describes how to resolve any problems, and shows how to replace faulty components.
- Hitachi High-performance NAS Platform™ ***Hardware Reference***: This guide (in PDF format) provides an overview of the hardware, describes how to resolve any problems, and shows how to replace faulty components.
- **Hitachi NAS Platform™ *and* Hitachi High-performance NAS Platform™** ***Storage Subsystem Guide*****:** In PDF format, this guide provides information about using the storage subsystems attached to the NAS storage server.
- Hitachi High-performance NAS Platform™ ***Command Line Reference***: This guide (in HTML format) describes how to administer the system by typing commands at a command prompt.
- **Release Notes**: This document gives late-breaking news on the system.

Browser Support

Any of the following browsers can be used to run Web Manager, the High-performance NAS Platform System Management Unit (SMU) web-based graphical user interface.

- Microsoft Internet Explorer: version 6.0 or later.
- Mozilla Firefox: version 1.5 or later.

The following Java Runtime Environment is required to enable some advanced Web Manager functionality.

- Sun Microsystems Java Runtime Environment: version 5.0, update 6, or later.

A copy of all product documentation is included for download or viewing through Web Manager. The following software is required to view this documentation:

- Adobe Acrobat: version 7.0.5 or later.

Table of Contents

1 The Hitachi Data Systems NAS Storage System

2 System Drives and System Drive Groups

1 The Hitachi Data Systems NAS Storage System

System Overview

The Hitachi NAS Platform™ and the High-performance NAS Platform are highly scalable and modular network attached storage (NAS) servers, with multi-gigabit throughput from network to disk. These systems consist of the following elements:

- System Management Unit (SMU)
- Hitachi NAS Platform™s and/or High-performance NAS Platforms
- Virtual Servers (EVSs)
- Private Management Network
- Public Data Network
- Storage Subsystem(s)

System Management Unit

The System Management Unit's (SMU's) Web Manager interface provides front-end server administration and monitoring tools. It supports clustering, data migration, and replication, and acts as the Quorum Device in a cluster. Although integral to the system as a whole, the SMU does not move data between the network client and the storage server.

There are two kinds of SMU; external and internal.

- An external SMU can manage up to eight (8) storage servers/clusters in any combination. Each external SMU can manage both Hitachi NAS Platforms/clusters and High-performance NAS Platforms/clusters.

  An external SMU is a separate device in the storage server system. To eliminate the SMU as a single point of failure, you can configure your system with a second external SMU as a standby SMU.

- An internal SMU can manage a single stand-alone Hitachi NAS Platform (an external SMU is required to manage more than a single Hitachi NAS Platform).

  An internal SMU is a service that runs on the Hitachi NAS Platform and provides the same management and monitoring functionality as an external SMU. When using an internal SMU, there is no way to configure a standby SMU.

Storage Server(s)

The patented architecture of the Hitachi NAS Platform and the High-performance NAS Platform is structured around bi-directional data pipelines and a hardware-based file system. It scales to 4 petabytes, supporting higher sustained access loads without compromising performance. Each storage server can be configured as a single stand alone server or as a node of a cluster. All network clients communicate directly with the storage server.

The server processes file access requests from network clients via Gigabit Ethernet (GE) or 10 Gigabit Ethernet (10GbE) links, reading and writing from/to one or multiple storage devices, connected through Fibre Channel (FC) links. Storage servers can be configured as stand alone servers or as a cluster with multiple nodes, which share the same storage devices, so that network requests can be distributed across cluster nodes.

Hitachi NAS Platforms support clusters with up to two nodes, and current generation High-performance NAS Platforms support clusters with up to eight nodes. Should one cluster node fail, its file services and server administration functions are transferred to other nodes.

Note: All nodes of a cluster must be of the same series. A cluster cannot be made up of a different models of storage servers.

Both the Hitachi NAS Platform and the High-performance NAS Platform are rack mountable and contain three hot-swappable fan assemblies, and two hot-swappable redundant power supplies. The front panel of each storage server displays system status with a green power LED and an amber fault LED. The rear panel has additional status LEDs, and includes connectors (power, Ethernet, Fibre Channel, RS-232). See the *Hardware Reference* for your series of storage server for more information about the storage server hardware.

Virtual Servers (EVSs)

All file services are provided by logical server entities referred to as EVSs (virtual servers). A server or cluster supports up to 64 EVSs. Each EVS is assigned unique network settings and storage resources. In clusters, EVSs are automatically migrated between servers when faults occur to ensure maximum availability. When multiple servers or clusters are configured with common storage access, they are referred to as *server farms*. EVSs can be manually migrated between servers in a Server Farm based on performance and availability requirements.

Private Management Network

To minimize the performance impact of auxiliary devices, a private management network connects the SMU and devices such as FC switches, and uninterruptible power supply (UPS) units. The private management network connects the private management interface of the SMU, the Ethernet management interface on the storage server(s), and all of the Ethernet managed devices that make up the storage system.

The private management network is isolated from the public data network by the SMU, which uses network address translation (NAT) technology. Devices on the private network are only accessible from the public data network through the SMU, which provides NAT, NTP, and email relay services.

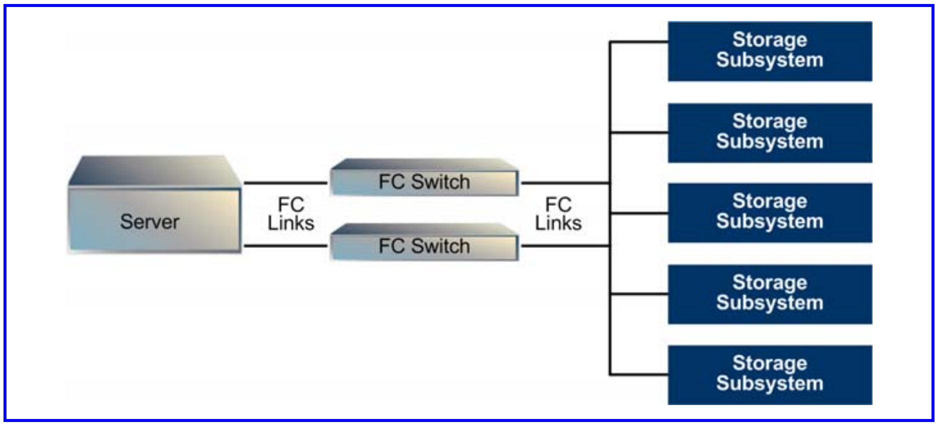
Public Data Network

The public data network, from the storage server perspective, consists of the public Ethernet port on the SMU, and management access can be enabled on individual Gigabit Ethernet (GE) interfaces on the storage server.

Clients connect to the SMU through the public data network, and client connections are made through the public data network. Typically, storage servers/clusters are configured to be managed through the public data network.

Storage Subsystem

Storage subsystems contain the devices that store the data managed by the storage server. The server allows you to simultaneously connect multiple diverse storage subsystems behind a single server (or cluster), which integrates all physical storage resources into one or more logical file systems.



Each storage subsystem is made up of RAID controllers, storage devices, and the Fibre Channel (FC) infrastructure (such as FC switches and cables) used to connect these devices to a single storage server or cluster.

Storage Subsystem Characteristics

The storage subsystems use hardware RAID controllers, which provide complete RAID functionality and enhanced disk failure management. The number of controllers in the system depends on its storage capacity.

Some storage subsystems use only one type of disk (FC, SATA, or SAS). Some storage subsystems can use different disk technologies, and may even be able to mix disk types within a storage subsystem (but mixing drive types within a storage enclosure is not supported).

Hot Spare Disk

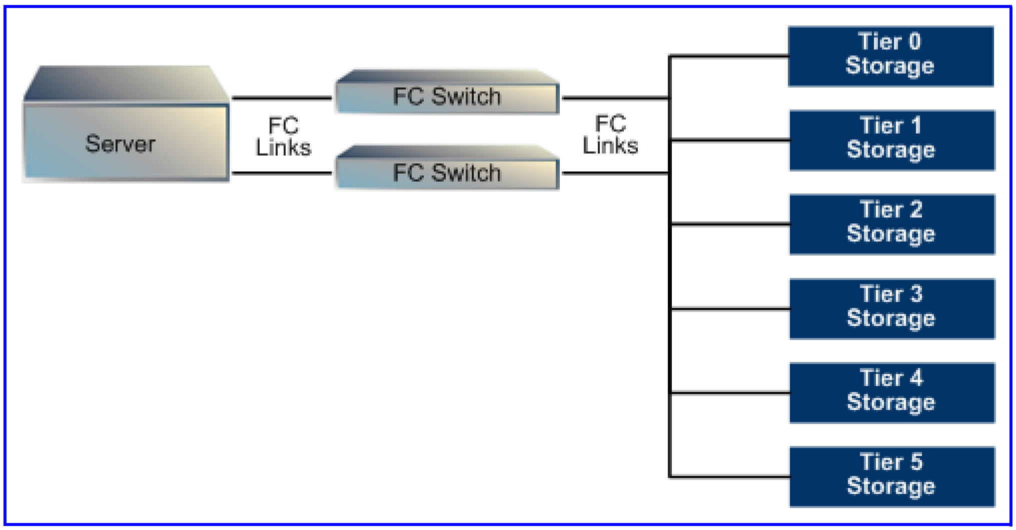
A hot spare disk is a physical disk that is configured for automatic use in the event that another physical disk fails. Should this happen, the system automatically switches to the hot spare disk and rebuilds on it the data that was located on the failed disk. The hot spare disk must have at least as much capacity as the other disks in the system drive. The hot spare disk is a global hot spare, meaning that only one hot spare disk is required per RAID rack, but Hitachi Data Systems recommends having several hot spares per rack to maintain a higher margin of safety against hardware failures.

If it is necessary to remove and replace a failed disk, some storage subsystems support “hot swap” operations. In a hot swap, an offline or failed disk is removed and a replacement disk is inserted while the power is on and the system is operating.

Understanding Tiered Storage

Tiered storage allows you to connect multiple diverse storage subsystems behind a single server (or cluster). Using tiered storage, you can match application storage requirements (in terms of performance and scaling) to your storage subsystems. This section describes the concept of tiered storage, and explains how to configure the storage server to work with your storage subsystems to create a tiered storage architecture.

Each server supports up to eight FC ports, independently configurable for either 1-, 2- or 4-gigabit operation. Independent configuration allows you to connect to a range of storage subsystems, which allows you to choose the configuration that will best meet application requirements. The server manages all back-end storage as a single system, through an integrated network management interface.



Based on a storage subsystem’s performance characteristics, it is classified as belonging to a certain tier, and each tier is used differently in the enterprise storage architecture. The currently supported storage subsystems are fit into the tiered storage model as follows:

Tier	Performance	Disk Type	Disk RPM
1	Very high	Dual-ported FC	15,000
2	High	Dual- ported FC	10,000
3	Nearline	SATA or SAS	7,200
4	Archival	SATA or SAS	7,200

Tier	Performance	Disk Type	Disk RPM
5	Long-term storage	N/A (Tape)	NA

The storage server supports tiers of storage, where each tier is made up of devices with different performance characteristics or technologies. When used with Hitachi Data Systems storage subsystems, the storage server also supports storage virtualization through Hitachi Universal Storage Platform USP-V and USP-VM technology.

Tiers of storage and storage virtualization are fully supported by Data Migrator, an optional feature which allows you to optimize the usage of tiered storage and remote NFSv3 servers by automatically migrating data among storage subsystems of primary and secondary storage. Based on user-defined policies, Data Migrator monitors file metadata such as size, type, duration of inactivity, access history, and so on. When the criteria of a policy are met, Data Migrator migrates files according to rules specified in the policy as background tasks with minimal impact on server performance. From the perspective of the client workstation, primary versus secondary file location is transparent. Note that Data Migrator does not support migrating data to or from tape library systems. For detailed information about Data Migrator, refer to the *System Administration Guide*.

Fibre Channel Fabric

The server supports fabric FC switches, and when connecting to a FC switch, the server must be configured for N_Port operation. Several FC Switch options are available, contact your Hitachi Data Systems representative for more information on supported switches.

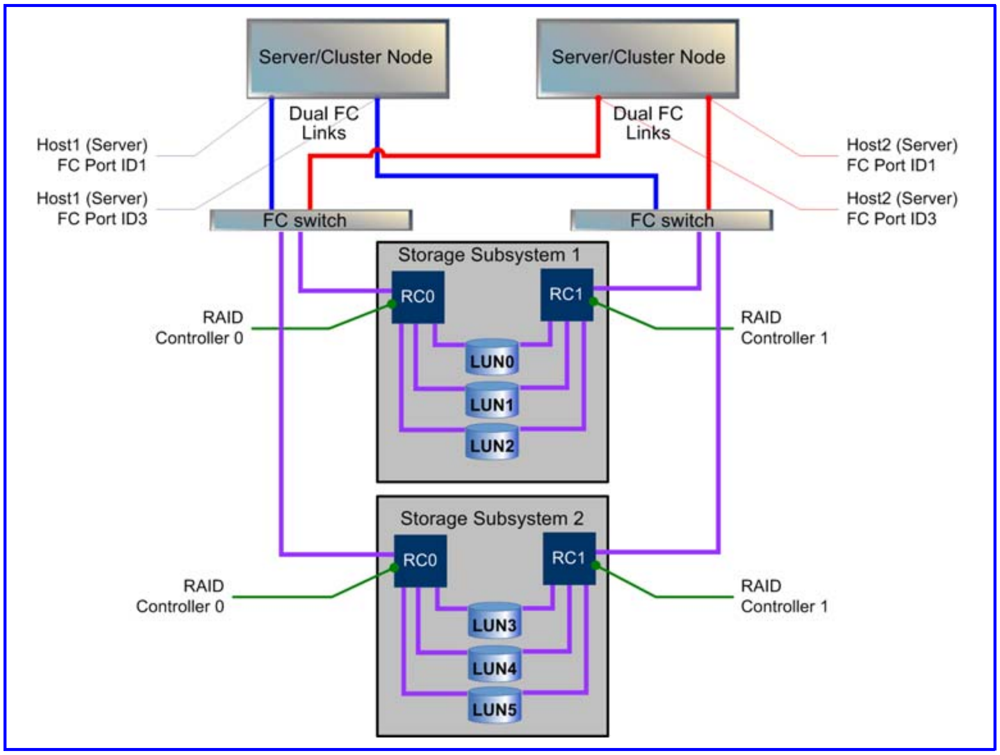
You can manage the FC interface on the server/cluster through the command line interface (CLI), using the following commands:

- `fc-link` to enable/disable the FC link.
- `fc-link-type` to change the FC link type.
- `fc-link-speed` to change the FC interface speed.

For more information about these commands, refer to the *Command Line Reference*.

About FC Paths

Each storage subsystem can be accessed through either of two available FC paths. An FC path is made up of the server’s host port ID, the storage subsystem port WWN (worldwide name), and the SD identifier (ID). The following illustration shows a complete path from the server to each of the SDs on the storage subsystem:



You can view information about the FC paths on the server/cluster through the command line interface (CLI), using the `fc-host-port-load`, `fc-target-port-load`, and the `sdpath` commands.

Load Balancing and Failure Recovery

Load balancing on a storage server is a matter of balancing the loads to the system drives (SDs) on the storage subsystems (RAID arrays) to which the storage server is connected. LUNs are a logical division of a group of the physical disks of the storage subsystem, and LUNs that are visible to the storage server are known as SDs, which are the basic storage unit of the storage subsystem.

The server routes FC traffic to individual SDs over a single FC path, distributing the load across two FC switches and, when possible, across dual active/active or multi-port RAID controllers.

Following the failure of a preferred path, disk I/O is redistributed among other (non-preferred) paths. When the server detects reactivation of the preferred FC path, it once again redistributes disk I/O to use the preferred FC path.

Default load balancing (load balancing automatically performed by the storage server) is performed based on the following criteria:

- "Load" is defined as the number of open SDs, regardless of the level of I/O on each SD. SDs count towards load at the target if they are open to at least one cluster node; the number of nodes (normally all nodes in a cluster, after boot) is not considered.

- Balancing load on RAID controller target ports takes precedence over balancing load on server FC host ports.
- Balancing load among an subsystem's RAID controllers takes precedence over balancing among ports on those controllers.
- In a cluster, choice of RAID controller target port is coordinated between cluster nodes, so that I/O requests for a given SD do not simultaneously go to multiple target ports on the same RAID controller.

You can manually configure load distribution from the CLI (overriding the default load balancing performed by the server), using the `sdpath` command. When manually configuring load balancing using the using the `sdpath` command:

- You can configure a preferred server host port and/or a RAID controller target port for an SD. If both are set, the RAID controller target port preference takes precedence over the server host port preference. When a specified port preference cannot be satisfied, port selection falls back to automatic selection.
- For the SDs visible on the same target port of a RAID controller, you should either set a preferred RAID controller target port for **all SDs** or for **none of the SDs**. Setting the preferred RAID controller target port for only some of the SDs visible on any given RAID controller target port may create a situation where load distribution is suboptimal.

The `sdpath` command can also be used to query the current FC path being used to communicate with each SD. For more information on the `sdpath` command, run the `man sdpath` command.

Fibre Channel Statistics

The server provides per-port and overall statistics, in real time, at 10-second intervals. Historical statistics cover the period since previous server start or statistics reset. The **Fibre Channel Statistics** page of the Web Manager displays a histogram showing the number of bytes/second received and transmitted during the past few minutes.

Supported Storage Subsystems

The storage server supports the following storage subsystems:

- Current offerings:
  - WMS 100
  - AMS 200, AMS 500, AMS 1000, AMS 2000
  - USP 100, USP 600, and USP 1100
  - USP V and USP VM

Due to the specific capacity and performance characteristics of each storage subsystem, it will typically be used in the tiered storage model as follows:

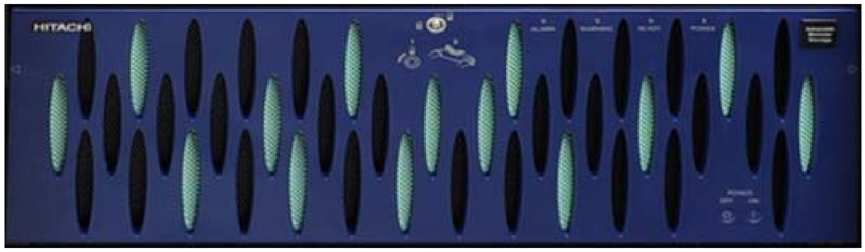
Enclosure	Typically used in Tier(s)
AMS 2000 and AMS 1000	Tier 1, Tier 2, and Tier 3 This subsystem has several configurations, and is suitable for use in several tiers, based on configuration of the individual storage array.
AMS 200 and AMS 500	Tier 2 and Tier 3 These subsystems have several configurations, and are suitable for use in several tiers, based on configuration of the individual storage array.
USP 100, USP 600, and USP 1100	Tier 1 and Tier 2 These subsystems have several configurations, and are suitable for use in several tiers, based on configuration of the individual storage array.
USP V	Tier 1, Tier 2, and Tier 3 This subsystem has several configurations, and is suitable for use in several tiers, based on configuration of the individual storage array.
USP VM	Tier 1 and Tier 3 This subsystem has several configurations, and is suitable for use in several tiers, based on configuration of the individual storage array.
WMS 100	Tier 3

The following table describes the RAID levels supported by each of the currently supported storage subsystems:

Enclosure	RAID Level(s) Supported
USP 100, USP 600, USP 1100, USP V, and USP VM	1/5/6
AMS 2000, AMS 1000, AMS 500, AMS 200, and WMS 100	1/5/6/10

The Adaptable Modular Storage (AMS) 2000 Family

The Adaptable Modular Storage (AMS) 2000 family includes the AMS 2100, AMS 2300 and the AMS 2500 storage subsystems. These storage subsystems can contain up to 480 SAS (Serial Attached SCSI) or SATA (Serial ATA) disk drives in racked or cabinet-mounted storage enclosures. In each subsystem, one storage enclosure contains symmetrical active/active controllers, with up to 16 fibre channel host ports, 32 backend SAS links, and up to 8 active links per drive tray. The AMS 2100, AMS 2300 and the AMS 2500 storage controllers, and each drive tray can contain up to 15 disk drives, and drive trays may contain SAS and/or SATA disk drives. SAS and SATA disks may be installed in the same drive tray, and they may be mixed within a tray. The AMS 2000 controller is symmetric active/active, which allows the host ports to access any LUN without a performance penalty. All host connections are equal in performance, and there are no preferred paths. The dual controller module self-monitors the workload, balancing it between controllers, so that administrators do not have to plan or manage workload distribution.



The AMS 1000 Storage Subsystem

A storage subsystem containing up to 450 Fibre Channel (FC) disk drives or up to 420 SATA (Serial ATA) disk drives in racked or cabinet-mounted storage enclosures. In the subsystem, one storage enclosure contains dual RAID controllers, with eight RAID host ports, 8 backend disk loops, and up to 4 active loops per drive tray. Up to 29 expansion drive trays are supported in addition to the storage controller enclosure. The storage controller and each drive tray can contain up to 15 disk drives, and drive trays may contain FC or SATA disk drives. The storage subsystem may contain a mix of drive trays containing FC disks and drive trays containing SATA disks, but disk types cannot be mixed within a drive tray.

The AMS 500 Storage Subsystem

A storage subsystem containing up to 225 Fibre Channel (FC) disk drives or up to 210 SATA (Serial ATA) disk drives in racked or cabinet-mounted storage enclosures. In the subsystem, one storage enclosure contains one or two RAID controllers, with four RAID host ports and up to 4 active loops per drive tray. Up to 14 expansion drive trays are supported in addition to the storage controller enclosure. The storage controller and each drive tray can contain up to 15 disk drives, and drive trays may contain FC or SATA disk drives. The storage subsystem may contain a mix of drive trays containing FC disks and drive trays containing SATA disks, but disk types cannot be mixed within a drive tray.

The AMS 200 Storage Subsystem

A storage subsystem containing up to 105 Fibre Channel (FC) disk drives or up to 90 SATA (Serial ATA) disk drives in racked or cabinet-mounted storage enclosures. In the subsystem, one storage enclosure contains one or two RAID controllers, with four RAID host ports and up to 2 active loops per drive tray. Up to 6 expansion drive trays are supported in addition to the storage controller enclosure. The storage controller and each drive tray can contain up to 15 disk drives, and drive trays may contain FC or SATA disk drives. The storage subsystem may contain a mix of drive trays containing FC disks and drive trays containing SATA disks, but disk types cannot be mixed within a drive tray.

The WMS 100 Storage Subsystem

A storage subsystem containing up to 105 SATA (Serial ATA) disk drives in racked or cabinet-mounted storage enclosures. In the subsystem, one storage enclosure contains one or two RAID controllers, with four RAID host ports and up to 2 active loops per drive tray. Up to 6 expansion drive trays are supported in addition to the storage controller enclosure. The storage controller and each drive tray can contain up to 15 disk drives.

The USP 100, USP 600, and USP 1100 Storage Subsystems

A series of storage subsystems containing up to 1,152 dual-ported FC (Fibre Channel) disk drives (10,000 or 15,000 rpm) in racked or cabinet-mounted storage enclosures. Supports one RAID controller enclosure and up to four disk array enclosures. Supports up to 32 PB of storage and RAID levels 1, 5, and 6. Connectivity through up to 192 Fibre Channel, 96 ESCON, or 96 FICON ports and third-generation Universal Star Network™ crossbar switch architecture.

The USP V and USP VM Storage Subsystems

A series of storage subsystems containing up to 1,152 FC (Fibre Channel) or SATA (Serial ATA) disk drives in racked or cabinet-mounted storage enclosures. One RAID controller enclosure and up to four disk array enclosures. Supports up to 247 PB of storage and RAID levels 1, 5, and 6. Connectivity through up to 224 Fibre Channel, 112 ESCON, or 112 FICON ports and fourth-generation Universal Star Network™ crossbar switch architecture. Also provides virtualization of internally and externally attached storage, intelligent tiering of data between FC and SATA disk drives, and active workload balancing.

2 System Drives and System Drive Groups

System Drives

Logically, system drives (SDs) are the basic storage element used by the storage server. The server assigns each system drive a unique identifying number (ID) and, once assigned, the SD is referenced by that ID number, and the ID number may not be changed.

Physically, each SD is made up of a LUN in a RAID group. The size of the system drive depends on factors such as the RAID level, the number of disks, and their capacity. See Supported Storage Subsystems, on page 7 for information on the RAID level(s) supported by your storage subsystem(s).

On all Hitachi NAS Platform/clusters, and all Series 3000 High-performance NAS Platforms/clusters, system drives can be organized into **system drive groups**, which can improve the performance of a storage server or cluster by optimizing reads and writes.

System Drive Groups

With many storage subsystems, system drives (SDs) are limited to 2TB each. However, with today's large physical disks, RAID arrays must be considerably larger than 2TB in order to make efficient use of space. So it is common for system administrators to build large RAID arrays (often called "RAID groups" or "volume groups") and then divide them into LUNs of 2TB or less. Note that each LUN in a RAID group typically uses some space on each disk in the RAID group. LUNs that are visible to the server/cluster are known as SDs, which are the units of storage that the server sees and manages (the server organizes SDs into Storage Pools, which then contain the file systems).

When performing write operations, if the server were to write simultaneously to multiple SDs in the same RAID group, it would increase head-movement, reducing both performance and the expected life of the disks. So the server has a mechanism to allow it to write to only one SD (LUN) in a RAID group at any one time. By defining SD groups, you tell the server which SDs are in each RAID group and give it the information it needs to optimize write performance.

Note: For High-performance NAS Platform/clusters, system drive groups are supported on Series 2000 and Series 3000 storage servers, but only Series 3000 High-performance NAS Platforms/clusters take advantage of this functionality. All Hitachi NAS Platforms/clusters take advantage of system drive groups.

A system drive that is not in any group is treated as if it were in a group of its own.

The SMU cannot group or ungroup system drives that are used in open Storage Pools. A Storage Pool is open if it has any file system that is mounted or is being checked or fixed anywhere on the cluster.

During EVS migration, the SMU automatically copies the groups from the source storage server or cluster and adds them to the target storage server or cluster.

See Managing System Drive Groups, on page 16 for information on creating and modifying system drive groups.

SD Groups and Read Balancing

When performing read operations, if the server can read file system data that has been distributed across the SD groups, the read operations are more efficient and disk head movement may be decreased. These benefits occur because the data is distributed across different physical disks in the LUNs making up the SD group instead of being on the same physical disks making up the SDs of the SD group.

Typically, file system data is distributed across the SDs in a Storage Pool during normal write operations due to dynamic write balancing. However, after adding SDs to an SD group or extending a Storage Pool by adding SDs that reside on a new/different physical storage subsystem, you may want to manually initiate a data redistribution so that the file system’s data is evenly spread among all SDs in the SD group.

During such a data redistribution, the server’s file serving performance may be impacted, because the server is reading and writing data from the data distribution as well as data from external requests.

The file system data redistribution utility is controlled using the `fs-read-balancer` command, and a separate command, `fs-sdg-utilization`, is used to see a report on how a file system is utilizing each SD group in the underlying Storage Pool. For more information about these commands, refer to the *Command Line Reference.*

Note: After a file system has been expanded to use new storage, you can spread the file system's data into the new storage by running the file system data redistribution utility. Spreading the data into the new storage spreads the read load across all of the file system’s physical disks, which improves read performance.

Read Balancing Utility Considerations

Running the file system data redistribution utility causes data to be re-written to a new location, which will be the least utilized SD groups (the new storage) resulting in more balanced utilization of SD groups.

Note: The file system data redistribution utility can be run only after expanding the file system into the new storage, but it should be run immediately and it may be run only once. If you run the data redistribution utility more than once, or after an application has written a significant amount of data into the recently added storage, the utility will either refuse to run or produce unpredictable results.

The file system data redistribution utility is designed to operate when a file system is expanded into new storage after SDs have been added to a Storage Pool when the file system is nearly full. However, storage may also be added to a Storage Pool for other reasons:

- To increase performance.
- To prevent the file system from becoming 100% full.

To achieve the desired results in either of these situations, you can use the following process after adding the storage (after adding the SDs to the Storage Pool):

1. Create a dummy file system, using all available space.
2. Expand the Storage Pool into the new SDs.
3. Expand the almost full target file system to use some (or all) of the space added to the Storage Pool. Note that the expansion should be at least 50% of the added storage capacity.
4. Run the file system data redistribution utility.
5. Delete the dummy file system.

Snapshots and the File System Data Redistribution Utility

When the file system data redistribution utility is run and snapshots are enabled, the old data is preserved. As a result, snapshots will grow, consuming a lot of disk space. The space used by these snapshots is not freed until all snapshots present when the file system data redistribution utility was started have been deleted.

There are four options available to recover the space used by snapshots:

1. Allow the snapshots to be deleted according to the snapshot configuration.

   This option recovers space the slowest, but could be used in scenarios where the space won't be required immediately after the file system data redistribution utility completes.

2. Manually delete snapshots after running the file system data redistribution utility.

   This option recovers space more quickly than option 1.

3. Manually kill snapshots after running the file system data redistribution utility.

   This option also recovers space more quickly than options 1or 2, but it requires that the file system is taken offline.

4. Disable snapshots (and therefore backups) and kill/delete existent snapshots before running the file system data redistribution utility.

   This option avoids the snapshot space usage problem altogether.

Creating System Drives on HDS Storage Subsystems

The server's system drives (SDs) are equivalent to the LUNs on the HDS storage subsystems. SDs (LUNs) are created using the HDS Storage Navigator software, which is supplied with the HDS storage subsystem. The LUNs on the HDS storage subsystem are then assigned to a port on the HDS storage subsystem RAID controller, also through the HDS Storage Navigator software.

After the SDs LUNs are created and assigned to a RAID controller port on the HDS storage subsystem, they are detected by the storage server when you run the `scsi-refresh` command. Once detected, they become available for use (as SDs) by the storage server.

Using the **System Drives** page of Web Manager, you then allow or deny access to the SDs on the HDS storage subsystem. The CLI command to allow access is `sd-allow-access`, and the CLI command to deny access is `sd-deny-access`.

Note: When connecting to an HDS storage subsystem, make sure the FC link type and link speed are set identically on the HDS storage subsystem and on the storage server.

For information on grouping SDs and managing SD groups, see Managing System Drive Groups, on page 16.

Managing System Drive Groups

After system drives (SDs) are created, they should be placed into system drive groups to optimize system performance. If your SD is large enough to use all the space in the RAID group, you cannot have multiple SDs in that RAID group, and you must create a system drive group containing just the one SD. You can create a system drive group either automatically or manually. (See Creating SD Groups Automatically, on page 16 or Creating SD Groups Manually, on page 18.)

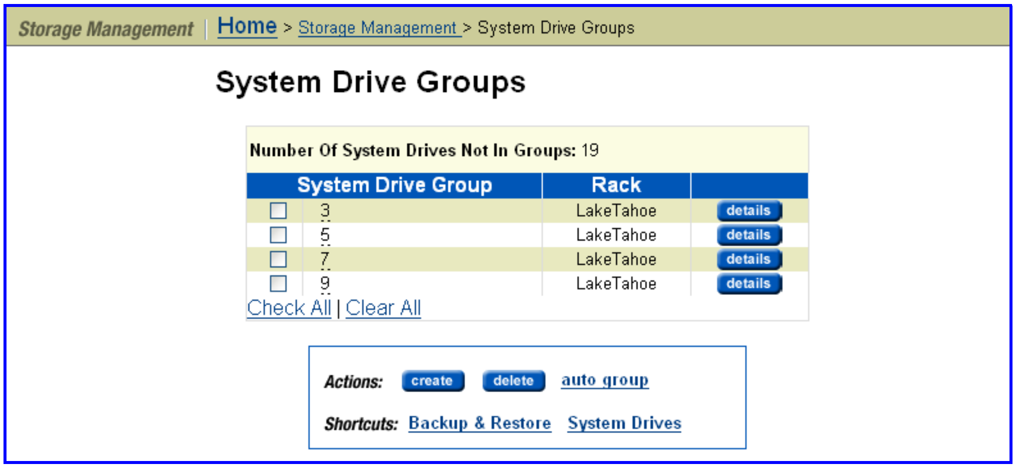
Caution: When creating a system drive group, you must ensure that the system drives being grouped are located on the same physical devices (or RAID group). Otherwise system performance can be adversely affected.

Creating SD Groups Automatically

The following steps describe how to create system drive groups automatically.

1. **Navigate to the System Drives Group page (Home > Storage Management > System Drive Groups).**

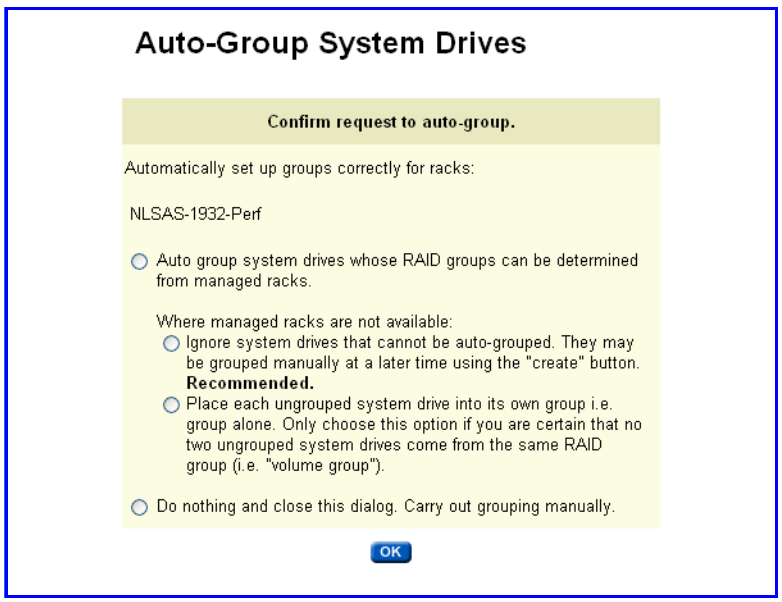
   The **System Drive Groups** page indicates the number of SDs not in groups and lists the SD groups that have already been created.



2. **Start auto-grouping the system drives.**

   In the **System Drives Group** page, click the **auto_group** link (if the "Number of System Drives Not In Groups" is zero, the auto_group link is unavailable because no groups can be created).

   The **Auto-Group System Drives** page opens, allowing you to automatically group system drives.



Note: Grouping system drives automatically can be performed only when the SMU can determine the physical location of the system drives along with the RAID configuration.Without such information, you will need to group system drives manually (see Creating SD Groups Manually, on page 18). The storage server cannot automatically group SDs on HDS storage subsystems, because those SDs (LUNs) are not directly managed by the storage server. Because of this, when you use the auto-group feature of the storage server, each SD on these storage subsystems will be placed into a separate group.

3. **Click OK to automatically create the SD groups.**

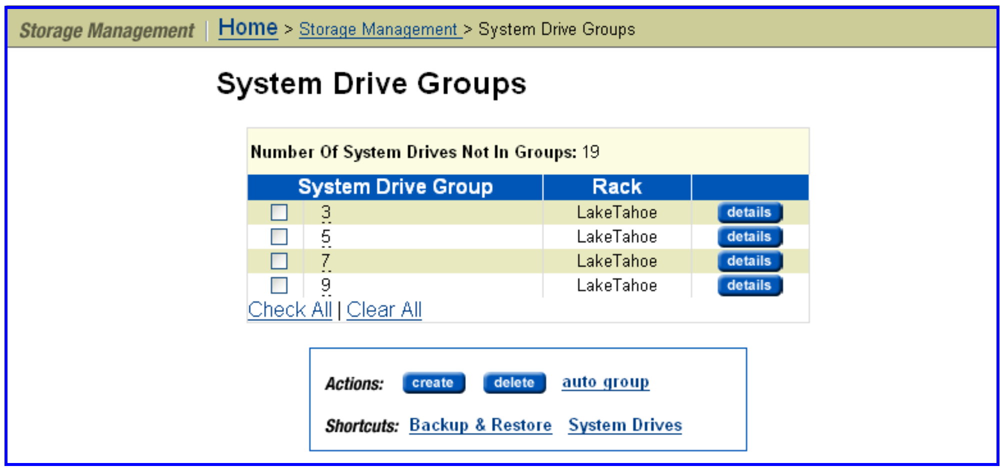
Creating SD Groups Manually

Before creating SD groups manually, you must first determine the physical location of the system drives along with the RAID configuration. This information can be obtained using various third-party applications. With this information, you can determine the SDs that are in the same RAID group, and group them together.

The following steps describe how to create system drive groups manually.

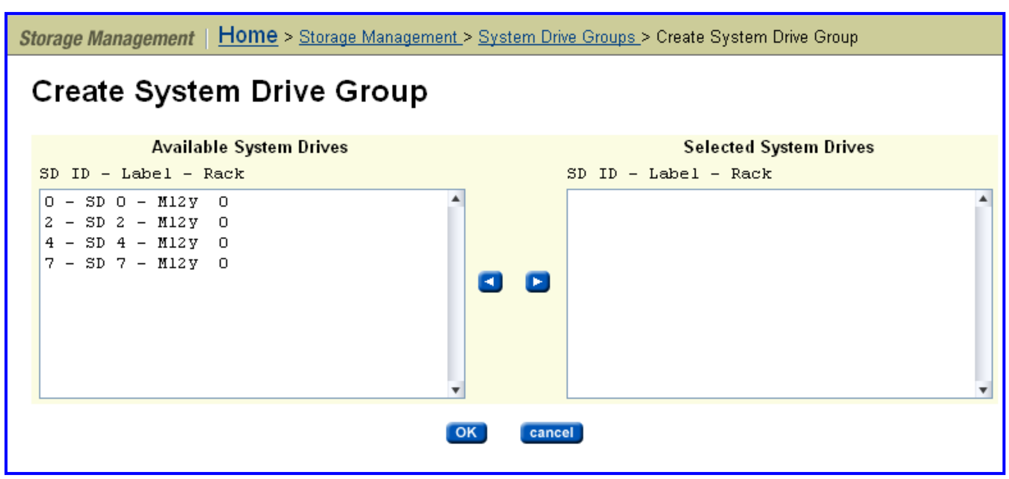
1. **Navigate to the System Drives Group page (Home > Storage Management > System Drive Groups).**

   The **System Drive Groups** page indicates the number of SDs not in groups and lists the SD groups that have already been created.



2. **Create the System Drives Group.**

   In the **System Drives Group** page, click **create** (if the "Number of System Drives Not In Groups" is zero, the **create** button is unavailable because no groups can be created). The **Create System Drive Group** page opens, allowing you to manually group system drives.



The **Available System Drives** list shows all system drives in the storage server or cluster that are not already in a group.

Select the SDs you want to group together from the Available System Drives list and click the right arrow to move them to the Selected System Drives list.

3. **Create the new System Drive Group.**

   In the **Create System Drives Group** page, click **OK** to create a new group containing the SDs in the selected list. Repeat these steps as needed to create the desired SD groups.

Backing up or Restoring SD Groups

Backing up and restoring SD groups is a simple, quick, and error-free method of transferring SD group definitions among clusters or servers that share storage or when moving storage between servers/clusters. SD groups are backed up as a part of the normal server/cluster configuration backup process, so backing up SD groups is not necessary for failure recovery.

Backing up SD groups imposes a very slight management overhead, and may cause other management functions to slow down slightly during the short time the backup is being made. There should be no noticeable effect on file serving throughput.

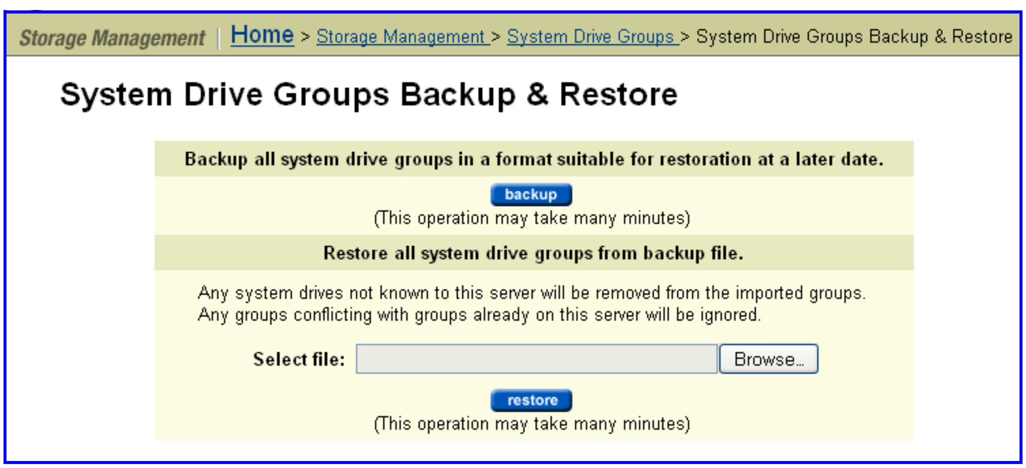
As with backing up SD groups, restoring saved SD groups will not cause a significant overhead, but there are rules about when you can and cannot import groups. For example, you cannot import a group that includes SDs that are used in a storage pool that has currently mounted file systems. For more information on these restrictions, see the `sd-group` command in the *Command Line Reference*.

Note: When moving storage between stand-alone servers or clusters, be aware that SD device IDs usually change, which may make it seem as if groups have been imported incorrectly.

To back up or restore SD Groups:

1. **Navigate to the System Drive Groups Backup & Restore page.**

   From the **Storage Management** page, select **System Drive Groups**, then click **Backup & Restore** to display the **System Drive Groups Backup & Restore** page.



2. **Backup or restore:**

- **To backup:** Click **backup.** In the browser, specify the name and location of the backup file, then click **OK/Save** (the buttons displayed and the method you use to save the backup file depend on the browser you use).

  A backup file name is suggested, but you can customize it. The suggested file name uses the syntax:

  `SD_GROUPSyyyy-mm-dd_time-UTC-offset.txt`, where the following example illustrates the appropriate syntax: `SD_GROUPS2008-04-30_1729-0700.txt`

- **To restore:** Click **Browse** to display a dialog you can use to choose the backup file, navigate to the directory where the backup file is stored, select the backup text file (`SD_GROUPS2008-04-30_1729-0700.txt`) for the specific export(s) you want to restore, then click **Open**.

  When the **System Drive Groups Backup & Restore** page displays the name and location of the selected file, click **restore.**

Modifying System Drive Groups

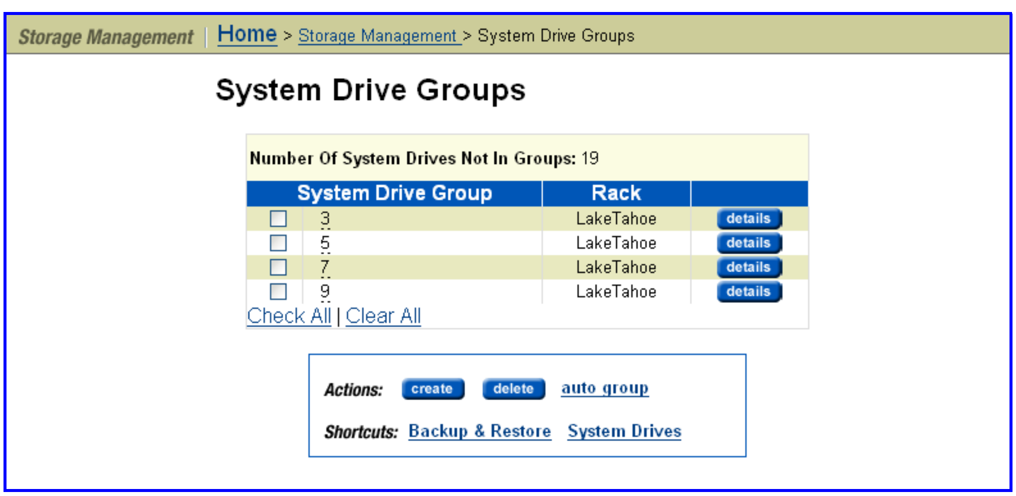
After system drive groups are created, you can modify a group by adding or removing system drives. Before modifying SD groups, you must first determine the physical location of the system drives along with the RAID configuration. This information can be obtained using various third-party applications, such as Storage Navigator.

Caution: When modifying a system drive group, you must ensure that the system drives being added to a group are located on the same physical devices (or RAID group) as the system drives already in the group. Otherwise system performance can be adversely affected.

The following steps describe how to modify system drive groups manually.

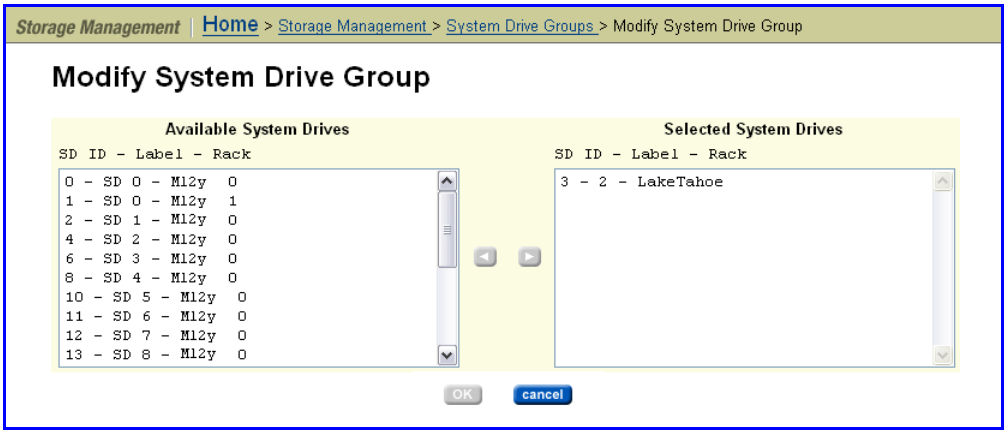
1. **Navigate to the System Drives Group page (Home > Storage Management > System Drive Groups).**

   The **System Drive Groups** page indicates the number of SDs not in groups and lists the SD groups that have already been created.



2. Select the system drive group you want to modify.

In the **System Drives Group** page, click **details** for the system drive group you want to modify. The **Modify System Drive Group** page opens, displaying the system drives that are available to be grouped and the SDs in the selected group.



The **Available System Drives** list shows all SDs in the storage server or cluster that are not already in a group. The **Selected System Drives** list shows the SDs that are already included in the selected group.

3. Modify the SD group.

You can add SDs to the SD group by selecting them in the **Available System Drives** list and clicking the right arrow to move the SDs to the **Selected System Drives** list.

You can remove SDs from the SD group by selecting them in the **Selected System Drives** list and clicking the left arrow to move the SDs to the **Available System Drives** list.

4. Review and save your changes.

In the **Modify System Drives Group** page, ensure that you have the desired SDs in the **Selected System Drives** list, and click **apply** to save the SD group with the modified content. Repeat these steps as needed to modify the desired SD groups.

Hitachi Data Systems

Corporate Headquarters
750 Central Expressway
Santa Clara, California 95050-2627
U.S.A.
Phone: 1 408 970 1000
www.hds.com
info@hds.com

Asia Pacific and Americas
750 Central Expressway
Santa Clara, California 95050-2627
U.S.A.
Phone: 1 408 970 1000
info@hds.com

Europe Headquarters
Sefton Park
Stoke Poges
Buckinghamshire SL2 4HD
United Kingdom
Phone: + 44 (0)1753 618000
info.eu@hds.com

Hitachi Data Systems

MK-96BA006-06