IBM

IBM System Storage SAN Volume Controller

Software Installation and Configuration Guide

Version 6.1.0

GC27-2286-00

IBM

IBM System Storage SAN Volume Controller

Software Installation and Configuration Guide

Version 6.1.0

GC27-2286-00

IBM

> **Note**
>
> Before using this information and the product it supports, read the information in "Notices" on page 269.

Contents

Chapter 8. IBM System Storage support for Microsoft Volume Shadow Copy Service and Virtual Disk Service for Windows 257

Appendix. Accessibility 267

Notices 269

Index 273

Figures

Tables

About this guide

This publication provides information that helps you configure and use the IBM® System Storage® SAN Volume Controller.

This publication also describes the configuration tools, both command-line and Web-based, that you can use to define, expand, and maintain the storage of the SAN Volume Controller.

Who should use this guide

This guide is intended for system administrators or others who install, configure, and use the IBM System Storage SAN Volume Controller.

Before using the SAN Volume Controller, you should have an understanding of storage area networks (SANs), the storage requirements of your enterprise, and the capabilities of your storage units.

Summary of changes

This summary of changes describes new functions that have been added to this release. Technical changes or additions to the text and illustrations are indicated by a vertical line to the left of the change. This document also contains terminology, maintenance, and editorial changes.

Summary of changes for GC27-2286-00, SAN Volume Controller Software Installation and Configuration Guide

This summary of changes provides a list of new, modified, and changed information since the last version of the guide. This topic describes the changes to this guide since the previous edition, SC23-6628-05.

New information

The following sections summarize the changes that have since been implemented from the previous version.

This version includes the following new information:
- Support statements for RAID

Changed information

The following updates were made in this document:
- Terminology changes:
1 To coincide with new and existing IBM products and functions, several common terms have changed
1 and are incorporated in the SAN Volume Controller information. Certain SAN Volume Controller
1 information, particularly command-line interface (CLI) documentation, remains primarily unchanged.
1 The following table shows the current and previous usage of the changed common terms.

1 *Table 1. Terminology mapping table*

6.1.0 SAN Volume Controller term	Previous SAN Volume Controller term	Description
event	error	An occurrence of significance to a task or system. Events can include completion or failure of an operation, a user action, or the change in state of a process.
host mapping	VDisk-to-host mapping	The process of controlling which hosts have access to specific volumes within a cluster.
storage pool	managed disk (MDisk) group	A collection of storage capacity that provides the capacity requirements for a volume.
thin provisioning (or thin-provisioned)	space-efficient	The ability to define a storage unit (full system, storage pool, volume) with a logical capacity size that is larger than the physical capacity assigned to that storage unit.
volume	virtual disk (VDisk)	A discrete unit of storage on disk, tape, or other data recording medium that supports some form of identifier and parameter list, such as a volume label or input/output control.

Removed information

The following information was removed from this book:
- SAN fabric and LAN configuration terms
- iSCSI overview information
- Limiting queue depth in large SANs topics
- Performance of fibre-channel extenders
- Web interface and command-line interface (CLI) tasks
- Upgrading the web interface information

Summary of changes for SC23-6628-04 and SC23-6628-05, SAN Volume Controller Software Installation and Configuration Guide

This summary of changes provides a list of new, modified, and changed information since the last version of the guide. This topic describes the changes to this guide since the previous edition, SC23-6628-03.

New information

The following sections summarize the changes that have since been implemented from the previous version.

This version includes the following new information:
- Support statements for the SAN Volume Controller 2145-CF8 node and the optional solid-state drive (SSD)
- Support statements for syslog servers to receive error, warning, and informational notifications from the cluster
- Support statements for reverse IBM FlashCopy® mapping with multiple targets
- Support statements for iSCSI 1Gb host attachment using on-board Ethernet ports
- Support statements for remote authentication for users of SAN Volume Controller clusters
- Configuring information for Texas Memory Systems RamSan Solid State Storage systems
- Configuring information for Xiotech Emprise systems

Changed information

The following updates were made in this document:
- Support statements for new managed disk (MDisk) states
- Support statements for viewing quorum disks and setting active quorum disks
- Support statements for SNMP servers to receive error, warning, and informational notifications from the cluster
- Enhancements to e-mail notification and inventory reporting
- Enhancements to statistics collection
- Support statements for second cluster IP and service IP addresses
- Support statements for a second Ethernet port on SAN Volume Controller nodes
- Support statements for extending the capability of creating partnerships among SAN Volume Controller clusters from one to three partners
- Support statements for embedded CIMOM on a cluster

Removed information

The following information was removed from this book:
- The SAN Volume Controller 2145-4F2 node is not supported in this version of SAN Volume Controller

Emphasis

Different typefaces are used in this guide to show emphasis.

The following typefaces are used to show emphasis:

Boldface	Text in **boldface** represents menu items and command names.
Italics	Text in *italics* is used to emphasize a word. In command syntax, it is used for variables for which you supply actual values, such as a default directory or the name of a cluster.
`Monospace`	Text in `monospace` identifies the data or commands that you type, samples of command output, examples of program code or messages from the system, or names of command flags, parameters, arguments, and name-value pairs.

SAN Volume Controller library and related publications

Product manuals, other publications, and websites contain information that relates to SAN Volume Controller.

SAN Volume Controller Information Center

The IBM System Storage SAN Volume Controller Information Center contains all of the information that is required to install, configure, and manage the SAN Volume Controller. The information center is updated between SAN Volume Controller product releases to provide the most current documentation. The information center is available at the following website:

publib.boulder.ibm.com/infocenter/svc/ic/index.jsp

SAN Volume Controller library

Unless otherwise noted, the publications in the SAN Volume Controller library are available in Adobe portable document format (PDF) from the following website:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Other IBM publications

Table 2 lists IBM publications that contain information related to the SAN Volume Controller.

Table 2. Other IBM publications

Title	Description	Order number
IBM System Storage Productivity Center Introduction and Planning Guide	This guide introduces the IBM System Storage Productivity Center hardware and software.	SC23-8824
Read This First: Installing the IBM System Storage Productivity Center	This guide describes how to install the IBM System Storage Productivity Center hardware.	GI11-8938
IBM System Storage Productivity Center User's Guide	This guide describes how to configure the IBM System Storage Productivity Center software.	SC27-2336
IBM System Storage Multipath Subsystem Device Driver User's Guide	This guide describes the IBM System Storage Multipath Subsystem Device Driver for IBM System Storage products and how to use it with the SAN Volume Controller.	GC52-1309

IBM documentation and related websites

Table 3 lists websites that provide publications and other information about the SAN Volume Controller or related products or technologies.

Table 3. IBM documentation and related websites

Website	Address
Support for SAN Volume Controller (2145)	Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145
Support for IBM System Storage and IBM TotalStorage® products	www.ibm.com/storage/support/
IBM Publications Center	www.ibm.com/e-business/linkweb/publications/servlet/pbi.wss
IBM Redbooks® publications	www.redbooks.ibm.com/

Related accessibility information

To view a PDF file, you need Adobe Acrobat Reader, which can be downloaded from the Adobe website:

www.adobe.com/support/downloads/main.html

How to order IBM publications

The IBM Publications Center is a worldwide central repository for IBM product publications and marketing material.

The IBM Publications Center offers customized search functions to help you find the publications that you need. Some publications are available for you to view or download at no charge. You can also order publications. The publications center displays prices in your local currency. You can access the IBM Publications Center through the following Web site:

www.ibm.com/e-business/linkweb/publications/servlet/pbi.wss

Sending your comments

Your feedback is important in helping to provide the most accurate and highest quality information.

To submit any comments about this book or any other SAN Volume Controller documentation:

- Go to the feedback page on the website for the SAN Volume Controller Information Center at publib.boulder.ibm.com/infocenter/svc/ic/index.jsp?topic=/com.ibm.storage.svc.console.doc/feedback.htm. There you can use the feedback page to enter and submit comments or browse to the topic and use the feedback link in the running footer of that page to identify the topic for which you have a comment.
- Send your comments by email to starpubs@us.ibm.com. Include the following information for this publication or use suitable replacements for the publication title and form number for the publication on which you are commenting:
  - Publication title: *IBM System Storage SAN Volume Controller Software Installation and Configuration Guide*
  - Publication form number: GC27-2286-00
  - Page, table, or illustration numbers that you are commenting on
  - A detailed description of any information that should be changed

Chapter 1. SAN Volume Controller overview

The SAN Volume Controller combines software and hardware into a comprehensive, modular appliance that uses symmetric virtualization.

Symmetric virtualization is achieved by creating a pool of managed disks (MDisks) from the attached storage systems. Those storage systems are then mapped to a set of volumes for use by attached host systems. System administrators can view and access a common pool of storage on the storage area network (SAN). This functionality helps administrators to use storage resources more efficiently and provides a common base for advanced functions.

A *SAN* is a high-speed fibre-channel network that connects host systems and storage devices. In a SAN, a host system can be connected to a storage device across the network. The connections are made through units such as routers and switches. The area of the network that contains these units is known as the *fabric* of the network.

SAN Volume Controller software

The SAN Volume Controller software performs the following functions for the host systems that attach to SAN Volume Controller:

- Creates a single pool of storage
- Provides logical unit virtualization
- Manages logical volumes
- Mirrors logical volumes

The SAN Volume Controller also provides the following functions:

- Large scalable cache
- Copy Services
1 – IBM FlashCopy (point-in-time copy), including thin-provisioned FlashCopy to make multiple targets
1   affordable
  – Metro Mirror (synchronous copy)
  – Global Mirror (asynchronous copy)
  – Data migration
- Space management
2 – IBM System Storage Easy Tier to migrate the most frequently used data to higher performing
2   storage
3 – Metering of service quality when combined with IBM Tivoli® Storage Productivity Center
t – Thin-provisioned logical volumes

1 Figure 1 on page 2 shows hosts, SAN Volume Controller nodes, and RAID storage systems connected to a
1 SAN fabric. The redundant SAN fabric comprises a fault-tolerant arrangement of two or more
1 counterpart SANs that provide alternate paths for each SAN-attached device.
1
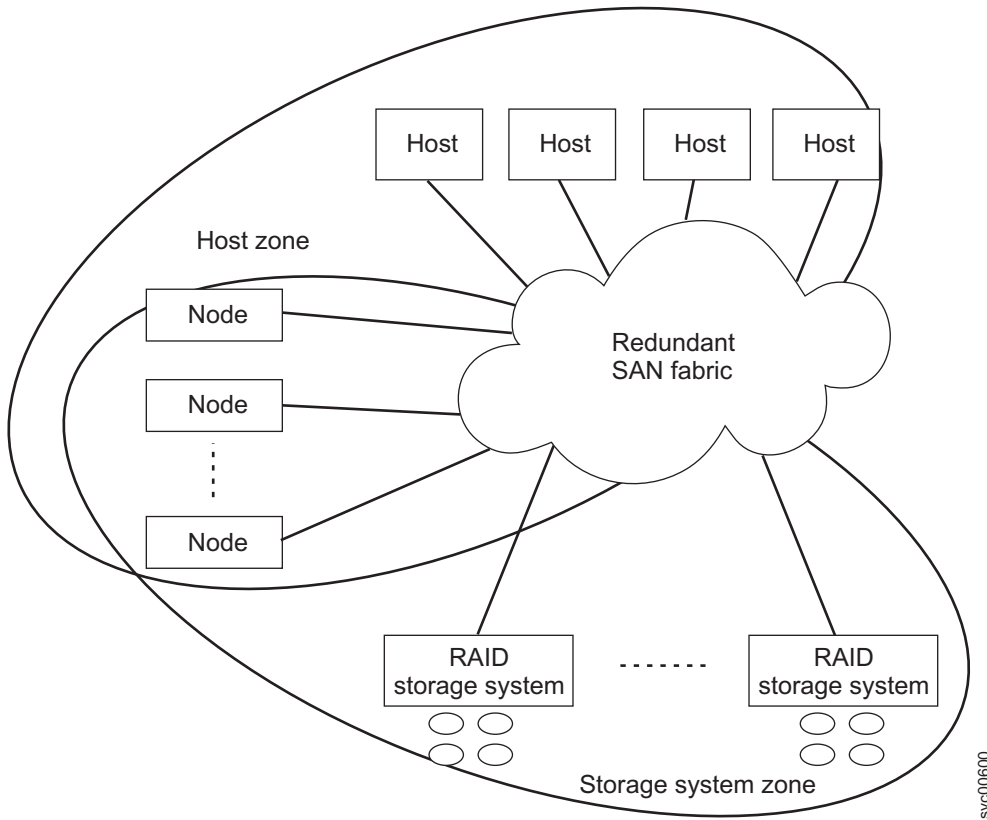
1
1 *Figure 1. SAN Volume Controller system in a fabric*
1
1 **Volumes**

1 A cluster of SAN Volume Controller nodes present volumes to the hosts. Most of the advanced functions
1 that SAN Volume Controller provides are defined on volumes. These volumes are created from managed
1 disks (MDisks) that are presented by the RAID storage systems. All data transfer occurs through the SAN
1 Volume Controller nodes, which is described as symmetric virtualization.

1 Figure 2 shows the data flow across the fabric.
1

Host   Host   Host   Host

Hosts send I/O
to volumes.

Node

Redundant
SAN fabric

Node

I/O is sent to
managed disks.

Data transfer

RAID
storage system   ....   RAID
storage system
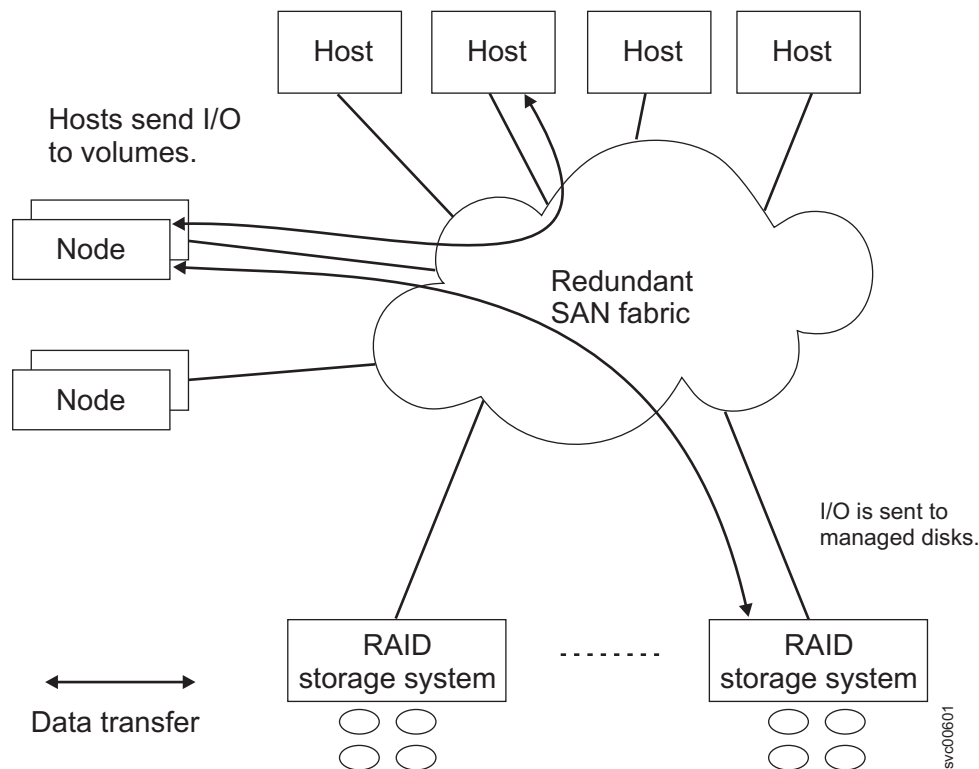
svc00601

Figure 2. Data flow in a SAN Volume Controller system

The nodes in a cluster are arranged into pairs known as *I/O groups*. A single pair is responsible for serving I/O on a given volume. Because a volume is served by two nodes, there is no loss of availability if one node fails or is taken offline.

System management

The SAN Volume Controller nodes in a cluster operate as a single system and present a single point of control for system management and service. System management and error reporting are provided through an Ethernet interface to one of the nodes in the cluster, which is called the *configuration node*. The configuration node runs a web server and provides a command-line interface (CLI). The configuration node is a role that any node can take. If the current configuration node fails, a new configuration node is selected from the remaining nodes. Each node also provides a command-line interface and web interface for performing hardware service actions.

Fabric types

I/O operations between hosts and SAN Volume Controller nodes and between SAN Volume Controller nodes and RAID storage systems are performed by using the SCSI standard. The SAN Volume Controller nodes communicate with each other by using private SCSI commands.

SAN Volume Controller uses the SCSI commands over the fibre-channel SAN and 1 Gbps Ethernet. Table 4 shows the fabric type that can be used for communicating between hosts, nodes, and RAID storage systems. These fabric types can be used at the same time.

Table 4. SAN Volume Controller communications types

Communications type	Host to SAN Volume Controller	SAN Volume Controller to storage system	SAN Volume Controller to SAN Volume Controller
Fibre Channel SAN	Yes	Yes	Yes

Table 4. SAN Volume Controller communications types (continued)

Communications type	Host to SAN Volume Controller	SAN Volume Controller to storage system	SAN Volume Controller to SAN Volume Controller
iSCSI (1 Gbps Ethernet)	Yes	No	No

Solid-state drives (SSDs)

Some SAN Volume Controller nodes contain solid-state drives (SSDs). These internal SSD drives can be used to create RAID-managed disks (MDisks) that in turn can be used to create volumes. SSDs provide host servers a pool of high-performance storage for critical applications.

Figure 3 shows this configuration. Internal SSD MDisks can also be placed in a storage pool with MDisks from regular RAID storage systems, and IBM System Storage Easy Tier will perform automatic data placement within that storage pool by moving high-activity data onto better performing storage.
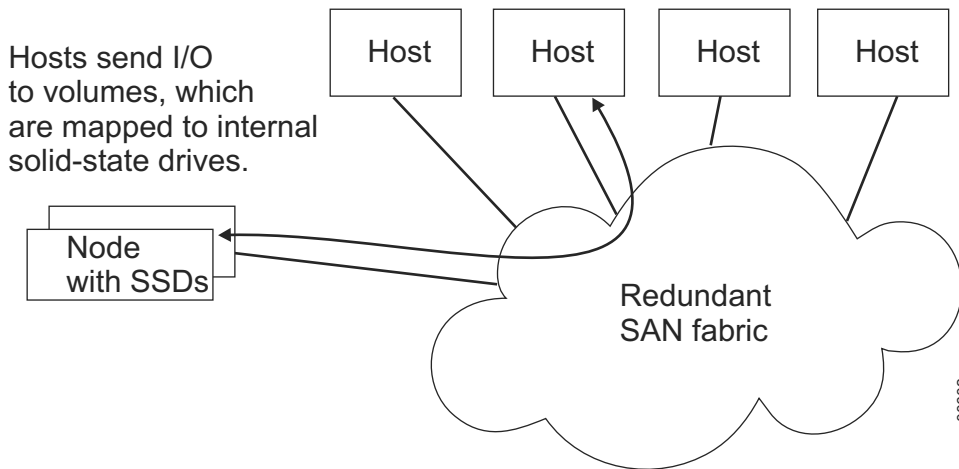
Figure 3. SAN Volume Controller nodes with internal SSDs

SAN Volume Controller hardware

Each SAN Volume Controller node is an individual server in a SAN Volume Controller cluster on which the SAN Volume Controller software runs.

The nodes are always installed in pairs, with a minimum of one and a maximum of four pairs of nodes constituting a *cluster*. Each pair of nodes is known as an *I/O group*. All I/O operations that are managed by the nodes in an I/O group are cached on both nodes.

I/O groups take the storage that is presented to the SAN by the storage systems as MDisks and translates the storage into logical disks (volumes) that are used by applications on the hosts. A node is in only one I/O group and provides access to the volumes in that I/O group.

Virtualization

Virtualization is a concept that applies to many areas of the information technology industry.

For data storage, virtualization includes the creation of a pool of storage that contains several disk systems. These systems can be supplied from various vendors. The pool can be split into volumes that are visible to the host systems that use them. Therefore, volumes can use mixed back-end storage and provide a common way to manage a storage area network (SAN).

Historically, the term *virtual storage* has described the virtual memory techniques that have been used in operating systems. The term *storage virtualization*, however, describes the shift from managing physical volumes of data to logical volumes of data. This shift can be made on several levels of the components of storage networks. Virtualization separates the representation of storage between the operating system and its users from the actual physical storage components. This technique has been used in mainframe computers for many years through methods such as system-managed storage and products like the IBM Data Facility Storage Management Subsystem (DFSMS). Virtualization can be applied at the following four main levels:

At the server level
> Manages volumes on the operating systems servers. An increase in the amount of logical storage over physical storage is suitable for environments that do not have storage networks.

At the storage device level
> 3   Uses RAID to create disk systems. This type of virtualization can range from simple RAID
> 3   controllers to advanced volume management such as that provided by the IBM System Storage
> 3   DS8000®. The Virtual Tape Server (VTS) is another example of virtualization at the device level.

At the fabric level
> Enables storage pools to be independent of the servers and the physical components that make up the storage pools. One management interface can be used to manage different storage systems without affecting the servers. SAN Volume Controller performs virtualization at the fabric level.

At the file system level
> Provides the highest benefit because data is shared, allocated, and protected at the data level rather than the volume level.

Virtualization is a radical departure from traditional storage management. In traditional storage management, storage is attached directly to a host system, which controls storage management. SANs introduced the principle of networks of storage, but storage is still primarily created and maintained at the RAID system level. Multiple RAID controllers of different types require knowledge of, and software that is specific to, the given hardware. Virtualization provides a central point of control for disk creation and maintenance.

One problem area that virtualization addresses is unused capacity. Before virtualization, individual host systems each had their own storage, which wasted unused storage capacity. Using virtualization, storage is pooled so that jobs from any attached system that need large amounts of storage capacity can use it as needed. Virtualization makes it easier to regulate the amount of available storage without having to use host system resources or to turn storage devices off and on to add or remove capacity. Virtualization also provides the capability to move storage between storage systems transparently to host systems.

Types of virtualization

Virtualization can be performed either asymmetrically or symmetrically. Figure 4 on page 6 provides a diagram of the levels of virtualization.

Asymmetric
> 1   A virtualization engine is outside the data path and performs a metadata style service. SAN
> 1   Volume Controller does not use asymmetric virtualization.

Symmetric
> A virtualization engine sits in the data path and presents disks to the hosts, but hides the physical storage from the hosts. Advanced functions, such as cache and Copy Services, can
> 1   therefore be implemented in the engine itself. SAN Volume Controller uses symmetric
> 1   virtualization.

Virtualization at any level provides benefits. When several levels are combined, the benefits of those levels can also be combined. For example, you can combine benefits by attaching a RAID controller to a virtualization engine that provides virtual volumes for a virtual file system.

1 **Note:** The SAN Volume Controller implements fabric-level *virtualization*. Within the context of the SAN
1 Volume Controller, *virtualization* refers to symmetric fabric-level virtualization.
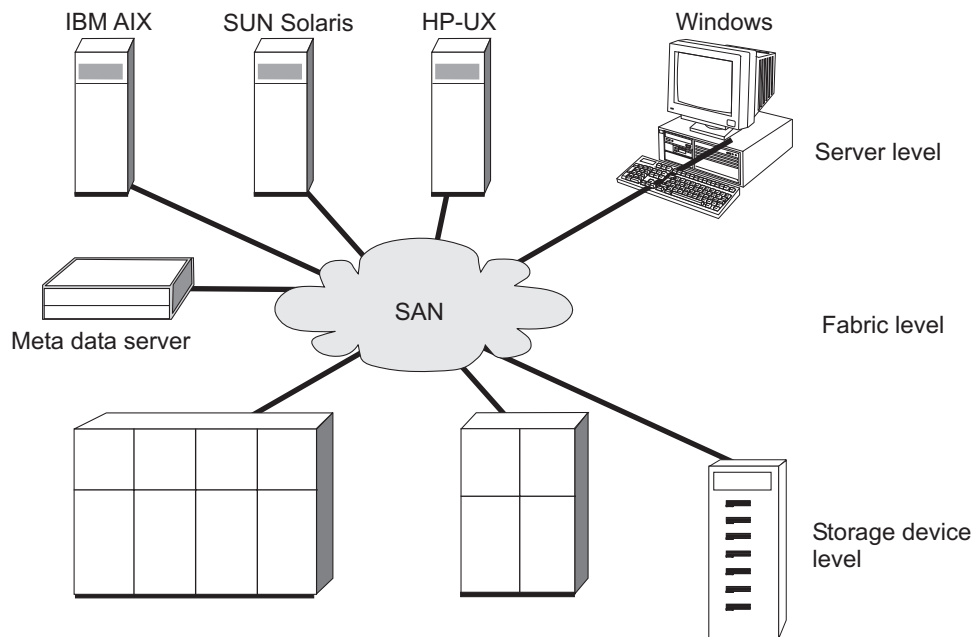
Figure 4. Levels of virtualization

Symmetric virtualization

The SAN Volume Controller provides symmetric virtualization.

Virtualization splits the storage that is presented by the storage systems into smaller chunks that are
t known as extents. These extents are then concatenated, using various policies, to make volumes. With
symmetric virtualization, host systems can be isolated from the physical storage. Advanced functions,
such as data migration, can run without the need to reconfigure the host. With symmetric virtualization,
the virtualization engine is the central configuration point for the SAN.

Figure 5 on page 7 shows that the storage is pooled under the control of the virtualization engine,
because the separation of the control from the data occurs in the data path. The virtualization engine
performs the logical-to-physical mapping.
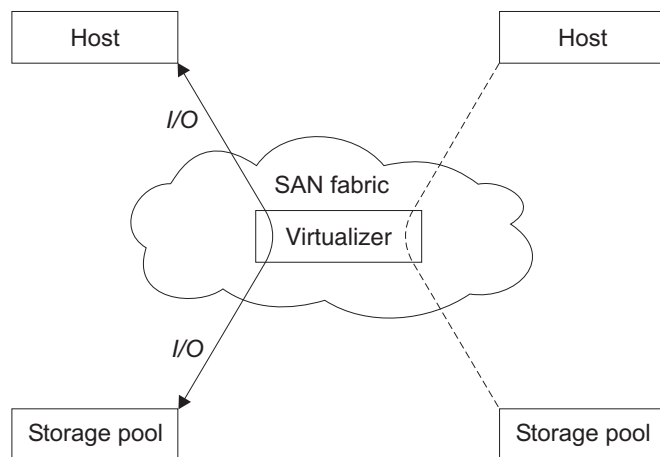
Figure 5. Symmetric virtualization

The virtualization engine directly controls access to the storage and to the data that is written to the storage. As a result, locking functions that provide data integrity and advanced functions, such as cache and Copy Services, can be run in the virtualization engine itself. Therefore, the virtualization engine is a central point of control for device and advanced function management. Symmetric virtualization allows you to build a firewall in the storage network. Only the virtualization engine can grant access through the firewall.

Symmetric virtualization can cause some problems. The main problem that is associated with symmetric virtualization is scalability. Scalability can cause poor performance because all input/output (I/O) must flow through the virtualization engine. To solve this problem, you can use an *n-way* cluster of virtualization engines that has failover capacity. You can scale the additional processor power, cache memory, and adapter bandwidth to achieve the desired level of performance. Additional memory and processing power are needed to run advanced services such as Copy Services and caching.

The SAN Volume Controller uses symmetric virtualization. Single virtualization engines, which are known as *nodes*, are combined to create *clusters*. Each cluster can contain between two and eight nodes.

Object overview

The SAN Volume Controller solution is based on a group of virtualization concepts. Before setting up your SAN Volume Controller environment, you should understand the concepts and the objects in the environment.

Each SAN Volume Controller single processing unit is called a *node*. Nodes are deployed in pairs to make up a cluster. A cluster can consist of one to four pairs of nodes. Each pair of nodes is known as an *I/O group* and each node can be in only one I/O group.

t *Volumes* are logical disks that are presented by the clusters. Each volume is associated with a particular
t I/O group. The nodes in the I/O group provide access to the volumes in the I/O group. When an
t application server performs I/O to a volume, it can access the volume with either of the nodes in the I/O group. Because each I/O group has only two nodes, the distributed cache is only two-way.

Each node does not contain any internal battery backup units and therefore must be connected to an *uninterruptible power supply*, which provides data integrity in the event of a cluster wide power failure. In such situations, the uninterruptible power supply maintains power to the nodes while the contents of the distributed cache are dumped to an internal drive.

1 The nodes in a cluster see the storage that is presented by back-end storage systems as a number of
1 disks, known as *managed disks (MDisks)*.

Each MDisk is divided into a number of *extents* which are numbered, from 0, sequentially from the start
to the end of the MDisk. MDisks are collected into groups, known as storage pools.

Each volume is made up of one or two volume copies. Each volume copy is an independent physical
copy of the data that is stored on the volume. A volume with two copies is known as a *mirrored volume*.
Volume copies are made out of MDisk extents. All the MDisks that contribute to a particular volume
copy must belong to the same storage pool.

A volume can be thin-provisioned. This means that the capacity of the volume as seen by host systems,
called the virtual capacity, can be different from the amount of storage that is allocated to the volume
from MDisks, called the real capacity. Thin-provisioned volumes can be configured to automatically
expand their real capacity by allocating new extents.

At any one time, a single node in the cluster can manage configuration activity. This node is known as
the *configuration node* and manages a cache of the information that describes the cluster configuration and
provides a focal point for configuration.

For a SCSI over fibre-channel connection, the nodes detect the fibre-channel ports that are connected to
the SAN. These correspond to the worldwide port names (WWPNs) of the fibre-channel host bus
adapters (HBAs) that are present in the application servers. You can create logical host objects that group
WWPNs that belong to a single application server or to a set of them.

For a SCSI over Ethernet connection, the iSCSI qualified name (IQN) identifies the iSCSI target
(destination) adapter. Host objects can have both IQNs and WWPNs.

SAN Volume Controller hosts are virtual representations of the physical host systems and application
servers that are authorized to access the cluster volumes. Each SAN Volume Controller host definition
specifies the connection method (SCSI over fibre channel or SCSI over Ethernet), the fibre-channel port or
IQN, and the volumes that the host applications can access.

The cluster provides block-level aggregation and volume management for disk storage within the SAN.
The cluster manages a number of back-end storage systems and maps the physical storage within those
storage systems into logical disk images that can be seen by application servers and workstations in the
SAN. The SAN is configured in such a way that the application servers cannot see the back-end physical
storage. This prevents any possible conflict between the cluster and the application servers both trying to
manage the back-end storage.

Clusters

All your configuration, monitoring, and service tasks are performed at the cluster (system) level.
Therefore, after configuring your cluster, you can take advantage of the virtualization and the advanced
features of the SAN Volume Controller.

A cluster can consist of between two and eight SAN Volume Controller nodes.

All configuration settings are replicated across all nodes in the cluster. Because configuration is performed
at the cluster level, management IP addresses are assigned to the cluster. Each interface accesses the
cluster remotely through the Ethernet cluster-management address.

System management

A system (cluster) is managed by using a command-line session or the management GUI over an
Ethernet connection.

Each SAN Volume Controller node has two Ethernet ports that can be used for management. Ethernet
port 1 must be configured with a management IP address and must be connected on all nodes in the
system. The use of Ethernet port 2 is optional. At any point in time, only one node in the system can
operate as the focal point for configuration and monitoring requests. This node is called the *configuration*

2 *node*. It is the only node that activates the management IP addresses. You can use one or more of these
2 addresses to access the system through the management GUI or the command-line interface (CLI).

3 Each SAN Volume Controller system can have zero to four management IP addresses. You can assign up
3 to two IPv4 addresses and up to two IPv6 addresses.

3 Each SAN Volume Controller node has one or two system IP addresses and up to two Internet Small
3 Computer System Interface over Internet Protocol (iSCSI IP) addresses per node.

3 **Note:** Management IP addresses that are assigned to a system are different from iSCSI IP addresses and
3 are used for different purposes. If iSCSI is used, iSCSI addresses are assigned to node ports. On the
3 configuration node, a port has multiple IP addresses active at the same time.

3 In addition to these IP addresses, you can optionally add two service IP addresses (one per node), to
3 provide access to service assistant.

Cluster IP failover

If the configuration node fails, the cluster IP addresses are transferred to a new node. The cluster services
are used to manage the transfer of the cluster IP addresses from the failed configuration node to the new
configuration node.

The following changes are performed by the cluster service:

- If software on the failed configuration node is still operational, the software shuts down the cluster IP
  interfaces. If the software cannot shut down the cluster IP interfaces, the hardware service forces the
  node to shut down.
- When the cluster IP interfaces shut down, all remaining nodes choose a new node to host the
  configuration interfaces.
- The new configuration node initializes the configuration daemons, including sshd and httpd, and then
  binds the cluster IP interfaces to its Ethernet ports.
- The router is configured as the default gateway for the new configuration node.
- The routing tables are established on the new configuration node for the cluster IP addresses. The new
  configuration node sends five unsolicited address resolution protocol (ARP) packets for each IP address
  to the local subnet broadcast address. The ARP packets contain the cluster IP and the media access
  control (MAC) address for the new configuration node. All systems that receive ARP packets are forced
  to update their ARP tables. Once the ARP tables are updated, these systems can connect to the new
  configuration node.

  Note: Some Ethernet devices might not forward ARP packets. If the ARP packets are not forwarded,
  connectivity to the new configuration node cannot be established automatically. To avoid this problem,
  configure all Ethernet devices to pass unsolicited ARP packets. You can restore lost connectivity by
  logging into the SAN Volume Controller and starting a secure copy to the affected system. Starting a
  secure copy forces an update to the ARP cache for all systems connected to the same switch as the
  affected system.

Ethernet link failures

If the Ethernet link to the SAN Volume Controller cluster fails because of an event unrelated to the SAN
Volume Controller, such as a cable being disconnected or an Ethernet router failure, the SAN Volume
Controller does not attempt to failover the configuration node to restore cluster IP access. SAN Volume
Controller provides the option for two Ethernet ports, each with its own management IP address, to
protect against this type of failure. If you cannot connect through one IP address, attempt to access the
cluster through the alternate IP address.

3 **Note:** IP addresses that are used by hosts to access the cluster over an Ethernet connection are different
3 from cluster management IP addresses.

Routing considerations for event notification

SAN Volume Controller supports the following protocols that make outbound connections from the cluster:
- Email
- Simple Network Mail Protocol (SNMP)
- Syslog
- Network Time Protocol (NTP)

One or more of these protocols can be configured on the cluster to receive event notifications. When making outbound connections, the SAN Volume Controller uses the following routing decisions:
- If the destination IP address is in the same subnet as one of the cluster IP addresses, then SAN Volume Controller sends the packet immediately.
- If the destination IP address is not in the same subnet as either of the cluster IP addresses, then SAN Volume Controller sends the packet to the default gateway for Ethernet port 1.
- If the destination IP address is not in the same subnet as either of the cluster IP addresses and Ethernet port 1 is not connected to the Ethernet network, then SAN Volume Controller sends the packet to the default gateway for Ethernet port 2.

When configuring any of these protocols for event notifications, use these routing decisions to ensure error notification works correctly in the event of a network failure.

Cluster configuration backup

Cluster configuration backup is the process of extracting configuration data from a cluster and writing it to disk.

Backing up the cluster configuration enables you to restore your cluster configuration in the event that it is lost. Only the data that describes the cluster configuration is backed up. You must back up your application data using the appropriate backup methods.

Configuration restore

Configuration restore is the process of using a backup cluster configuration data file or files to restore a specific cluster configuration.

Restoring the cluster configuration is an important part of a complete backup and disaster recovery solution. You must also regularly back up your application data using appropriate backup methods because you might need to restore your application data after you have restored your cluster configuration. The configuration backup procedure does not back up data.

Powering on and powering off the cluster

Follow these general procedures to power on or to power off the cluster. The procedures must be followed in the order given.

Powering on the cluster

1. Power on fibre-channel switches.
2. Power on external storage systems.
3. Power on SAN Volume Controller nodes.
4. Start hosts.

Powering off the cluster

1. Stop hosts.
2. Shut down the cluster. Using the management GUI, click **Home** > **System Status**. At the bottom of the page, click the cluster name (cluster code level); then click **Manage** > **Shut down Cluster**.

Note: You can use the front panel to shut down the cluster; however, this is not advisable.

3. (Optional) Shut down external storage systems.
4. (Optional) Shut down fibre-channel switches.

Nodes

Each SAN Volume Controller *node* is a single processing unit within a SAN Volume Controller cluster.

For redundancy, nodes are deployed in pairs to make up a cluster. A cluster can have one to four pairs of nodes. Each pair of nodes is known as an I/O group. Each node can be in only one I/O group. A maximum of four I/O groups each containing two nodes is supported.

At any one time, a single node in the cluster manages configuration activity. This configuration node manages a cache of the configuration information that describes the cluster configuration and provides a focal point for configuration commands. If the configuration node fails, another node in the cluster takes over its responsibilities.

Table 5 describes the operational states of a node.

Table 5. Node state

State	Description
Adding	The node was added to the cluster but is not yet synchronized with the cluster state (see Note). The node state changes to Online after synchronization is complete.
Deleting	The node is in the process of being deleted from the cluster.
Online	The node is operational, assigned to a cluster, and has access to the fibre-channel SAN fabric.
Offline	The node is not operational. The node was assigned to a cluster but is not available on the fibre-channel SAN fabric. Run the Directed Maintenance Procedures to determine the problem.
Pending	The node is transitioning between states and, in a few seconds, will move to one of the other states.
Note: A node can stay in the Adding state for a long time. You should wait for at least 30 minutes before taking further action, but if after 30 minutes, the node state is still Adding, you can delete the node and add it again. If the node that has been added is at a lower code level than the rest of the cluster, the node is upgraded to the cluster code level, which can take up to 20 minutes. During this time, the node is shown as adding.	

Configuration node

A *configuration node* is a single node that manages configuration activity of the cluster.

If the configuration node fails, the cluster chooses a new configuration node. This action is called configuration node failover. The new configuration node takes over the cluster IP addresses. Thus you can access the cluster through the same IP addresses although the original configuration node has failed. During the failover, there is a short period when you cannot use the command-line tools or management GUI.

Figure 6 on page 12 shows an example cluster that contains four nodes. Node 1 has been designated the configuration node. User requests (*1*) are handled by node 1.
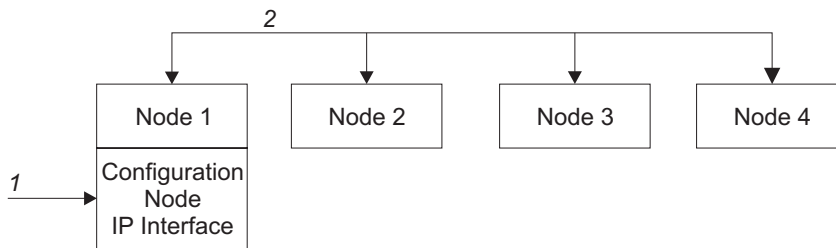
Figure 6. Configuration node

Quorum disks

A quorum disk is a managed disk (MDisk) or a managed drive that contains a reserved area that is used exclusively for cluster (system) management. A cluster automatically assigns quorum disk candidates. When you add new storage to a cluster or remove existing storage, however, it is a good practice to review the quorum disk assignments.

A cluster uses the quorum disk for two purposes:

- To break a tie when a SAN fault occurs, when exactly half of the nodes that were previously a member of the cluster are present.
- To hold a copy of important cluster configuration data. Just over 256 MB is reserved for this purpose on each quorum disk candidate.

A cluster can have only one active quorum disk; however, the cluster uses three managed disks as quorum disk candidates. The cluster automatically selects the actual active quorum disk from the pool of assigned quorum disk candidates. The active quorum disk can be specified by using the **svctask chquorum** command-line interface (CLI) command with the **active** parameter. To view the current quorum disk status, use the **svcinfo lsquorum** command.

The other quorum disk candidates provide redundancy if a quorum disk fails before a cluster is partitioned. To avoid the possibility of losing all the quorum disk candidates with a single failure, assign quorum disk candidates on multiple storage systems.

Note: Mirrored volumes can be taken offline if no quorum disk is available. The synchronization status for mirrored volumes is recorded on the quorum disk.

When you change the managed disks that are assigned as quorum candidate disks, follow these general guidelines:

- When possible, aim to distribute the quorum candidate disks so that each MDisk is provided by a different storage system. For information about which storage systems are supported for quorum disk use, refer to the supported hardware list.
- Before you issue the **svctask chquorum** command, ensure the status of the managed disk that is being assigned as a quorum candidate disk is reported as `online`.
- Allow at least 2 minutes between issuing consecutive **svctask chquorum** commands. This delay provides enough time to ensure that the quorum disk update is finished before another quorum disk change is made.

Quorum MDisks or drives in split-cluster configurations

To provide protection against failures that affect an entire location (for example, a power failure), you can use volume mirroring with a configuration that splits a single cluster between two physical locations. For

3 further information, see the split-cluster configuration information. For detailed guidance about
3 split-cluster configuration for high-availability purposes, contact your IBM regional advanced technical
3 specialist.

3 Generally, when the nodes in a cluster have been split among sites, configure the SAN Volume Controller
3 cluster this way:
3 • Site 1: Half of SAN Volume Controller cluster nodes + one quorum disk candidate
3 • Site 2: Half of SAN Volume Controller cluster nodes + one quorum disk candidate
3 • Site 3: Active quorum disk

3 This configuration ensures that a quorum disk is always available, even after a single site failure.

3 The following scenarios describe examples that result in changes to the active quorum disk:
3 • Scenario 1:
3   – Site 3 is either powered off or connectivity to the site is broken.
3   – The cluster selects a quorum disk candidate at site 2 to become the active quorum disk.
3   – Site 3 is either powered on or connectivity to the site is restored.
3   – The SAN Volume Controller administrator must reassign all three quorum disks to ensure that the
3   active quorum disk is now located at site 3 again.
3 • Scenario 2:
3   – The storage system that is hosting the preferred quorum disk at site 3 is removed from the
3   configuration.
3   – If possible, the cluster automatically configures a new quorum disk candidate at site 1 or 2).
3   – The cluster selects a quorum disk candidate at site 1 or 2 to become the active quorum disk.
3   – A new storage system is added to site 3.
3   – The SAN Volume Controller administrator must reassign all three quorum disks to ensure that the
3   active quorum disk is now located at site 3 again.

3 I/O groups and uninterruptible power supply

Nodes are deployed in pairs to make up a cluster. Each pair of nodes is known as an *I/O group*. Each node can only be in one I/O group.

t *Volumes* are logical disks that are presented to the SAN by SAN Volume Controller nodes. Volumes are also associated with an I/O group. The SAN Volume Controller does not contain any internal battery backup units and therefore must be connected to an uninterruptible power supply to provide data integrity in the event of a cluster-wide power failure.

I/O groups

Each pair of nodes is known as an *input/output (I/O) group*. An I/O group is defined during the cluster configuration process.

t *Volumes* are logical disks that are presented to the SAN by SAN Volume Controller nodes. Volumes are also associated with an I/O group.

t When an application server performs I/O to a volume, it can access the volume with either of the nodes
t in the I/O group. When you create a volume, you can specify a preferred node. Many of the multipathing driver implementations that SAN Volume Controller supports use this information to direct I/O to the preferred node. The other node in the I/O group is used only if the preferred node is not accessible.

t If you do not specify a preferred node for a volume, the node in the I/O group that has the fewest
t volumes is selected by SAN Volume Controller to be the preferred node.

3 After the preferred node is chosen, it can be changed only when the volume is moved to a different I/O
3 group.

3 **Attention:** Moving a volume to a different I/O group is disruptive to host I/O.

3 To view the current preferred node for a volume, select **Volumes** → **All Volumes** in the management GUI.
3 Right-click the volume and select **Properties**. To view the current preferred node assignment using the
3 command-line interface, run the **svcinfo lsvdisk** command.

t An I/O group consists of two nodes. When a write operation is performed to a volume, the node that
processes the I/O duplicates the data onto the partner node that is in the I/O group. After the data is
protected on the partner node, the write operation to the host application is completed. The data is
physically written to disk later.

Read I/O is processed by referencing the cache in the node that receives the I/O. If the data is not found,
it is read from the disk into the cache. The read cache can provide better performance if the same node is
t chosen to service I/O for a particular volume.

t I/O traffic for a particular volume is, at any one time, managed exclusively by the nodes in a single I/O
group. Thus, although a cluster can have eight nodes within it, the nodes manage I/O in independent
pairs. This means that the I/O capability of the SAN Volume Controller scales well, because additional
throughput can be obtained by adding additional I/O groups.

Nodes do not contain any internal battery backup units and therefore must be connected to an
uninterruptible power supply to provide data integrity in the event of a cluster-wide power failure. The
uninterruptible power supply only provides power long enough to enable the SAN Volume Controller
cluster to shutdown and save cache data. The uninterruptible power supply is not intended to maintain
power and keep the nodes running during an outage.

t Figure 7 shows a write operation from a host (*1*), that is targeted for volume A. This write is targeted at
the preferred node, Node 1 (*2*). The write is cached and a copy of the data is made in the partner node,
Node 2's cache (*3*). The host views the write as complete. At some later time, the data is written, or
de-staged, to storage (*4*). Figure 7 also shows two uninterruptible power supply units correctly configured
so that each node is in a different power domain.
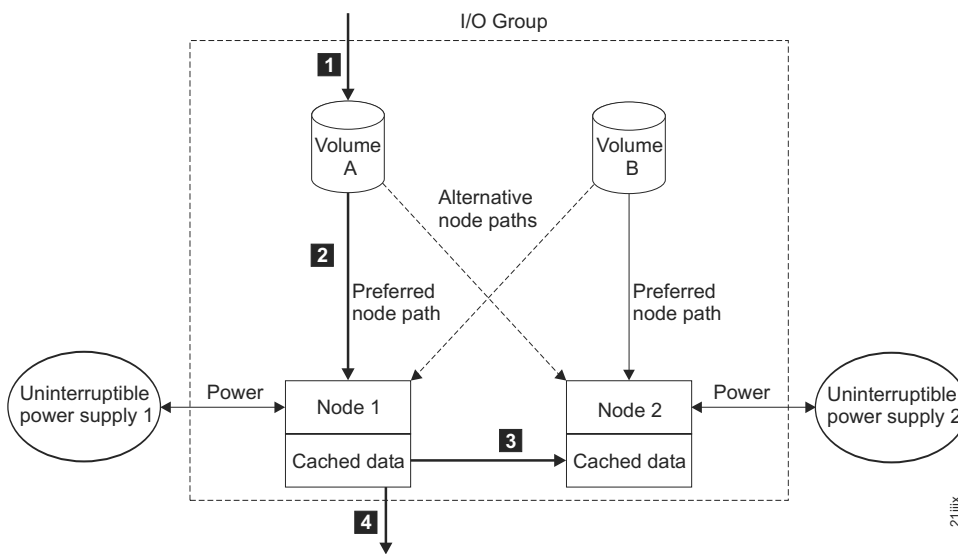
Figure 7. I/O group and uninterruptible power supply

When a node fails within an I/O group, the other node in the I/O group assumes the I/O responsibilities of the failed node. Data loss during a node failure is prevented by mirroring the I/O read and write data cache between the two nodes in an I/O group.

If only one node is assigned to an I/O group or if a node has failed in an I/O group, the cache is flushed to the disk and then goes into write-through mode. Therefore, any writes for the volumes that are assigned to this I/O group are not cached; they are sent directly to the storage device. If both nodes in an I/O group go offline, the volumes that are assigned to the I/O group cannot be accessed.

When a volume is created, the I/O group to provide access to the volume must be specified. However, volumes can be created and added to I/O groups that contain offline nodes. I/O access is not possible until at least one of the nodes in the I/O group is online.

The cluster also provides a virtual recovery I/O group that can be used for certain service actions. You can move the volumes to the recovery I/O group and then into a working I/O group. I/O access is not possible when volumes are assigned to a recovery I/O group.

2145 UPS-1U

A 2145 UPS-1U is used exclusively to maintain data that is held in the SAN Volume Controller dynamic random access memory (DRAM) in the event of an unexpected loss of external power. This use differs from the traditional uninterruptible power supply that enables continued operation of the device that it supplies when power is lost.

With a 2145 UPS-1U, data is saved to the internal disk of the SAN Volume Controller node. The uninterruptible power supply units are required to power the SAN Volume Controller nodes even when the input power source is considered uninterruptible.

Note: The uninterruptible power supply maintains continuous SAN Volume Controller-specific communications with its attached SAN Volume Controller nodes. A SAN Volume Controller node cannot operate without the uninterruptible power supply. The uninterruptible power supply must be used in accordance with documented guidelines and procedures and must not power any equipment other than a SAN Volume Controller node.

2145 UPS-1U operation:

Each SAN Volume Controller node monitors the operational state of the uninterruptible power supply to which it is attached.

If the 2145 UPS-1U reports a loss of input power, the SAN Volume Controller node stops all I/O operations and dumps the contents of its dynamic random access memory (DRAM) to the internal disk drive. When input power to the 2145 UPS-1U is restored, the SAN Volume Controller node restarts and restores the original contents of the DRAM from the data saved on the disk drive.

A SAN Volume Controller node is not fully operational until the 2145 UPS-1U battery state indicates that it has sufficient charge to power the SAN Volume Controller node long enough to save all of its memory to the disk drive. In the event of a power loss, the 2145 UPS-1U has sufficient capacity for the SAN Volume Controller to save all its memory to disk at least twice. For a fully charged 2145 UPS-1U, even after battery charge has been used to power the SAN Volume Controller node while it saves dynamic random access memory (DRAM) data, sufficient battery charge remains so that the SAN Volume Controller node can become fully operational as soon as input power is restored.

Important: Do not shut down a 2145 UPS-1U without first shutting down the SAN Volume Controller node that it supports. Data integrity can be compromised by pushing the 2145 UPS-1U on/off button when the node is still operating. However, in the case of an emergency, you can manually shut down the 2145 UPS-1U by pushing the 2145 UPS-1U on/off button when the node is still operating. Service actions

must then be performed before the node can resume normal operations. If multiple uninterruptible power supply units are shut down before the nodes they support, data can be corrupted.

Internal storage and external storage

A SAN Volume Controller cluster can manage a combination of internal and external storage.

Internal storage

The SAN Volume Controller 2145-CF8 has a number of drives attached to it. These drives are used to create a Redundant Array of Independent Disks (RAID), which are presented as managed disks (MDisks) in the system.

External storage

The SAN Volume Controller can detect logical units (LUs) on an external storage system that is attached via fibre-channel connections. These LUs are detected as managed disks (MDisks) in the system and must be protected from drive failures by using RAID technology on an external storage system.

Storage systems

An external storage system, or *storage controller*, is a device that coordinates and controls the operation of one or more disk drives. A storage system synchronizes the operation of the drives with the operation of the system as a whole.

Storage systems provide the storage that the SAN Volume Controller cluster detects as one or more MDisks.

SAN Volume Controller supports storage systems that implement the use of RAID technology and also those that do not use RAID technology. RAID implementation provides redundancy at the disk level, which prevents a single physical disk failure from causing a RAID-array managed disk (MDisk), storage pool, or associated volume failure. Physical capacity for storage systems can be configured as RAID 1, RAID 0+1, RAID 5, RAID 6, or RAID 10.

Storage systems divide array storage into many Small Computer System Interface (SCSI) logical units (LUs) that are presented on the SAN. Ensure that you assign an entire array to an MDisk as you create the MDisk, to present the array as a single SCSI LU that is recognized by SAN Volume Controller as a single RAID-array MDisk. You can then use the virtualization features of SAN Volume Controller to create volumes from the MDisk.

The exported storage devices are detected by the cluster and reported by the user interfaces. The cluster can also determine which MDisks each storage system is presenting and can provide a view of MDisks that is filtered by the storage system. Therefore, you can associate the MDisks with the RAID that the system exports.

The storage system can have a local name for the RAID or single disks that it is providing. However, it is not possible for the nodes in the cluster to determine this name, because the namespace is local to the storage system. The storage system makes the storage devices visible with a unique ID, called the logical unit number (LUN). This ID, along with the storage system serial number or numbers (there can be more than one controller in a storage system), can be used to associate the MDisks in the cluster with the RAID exported by the system.

The size of an MDisk cannot be changed once it becomes a managed MDisk by adding it to a storage pool. If the size of the LUN that is presented by the storage system is reduced to below the size of the managed MDisk, the MDisk is taken offline by SAN Volume Controller. If the size of the LUN that is presented by the storage system is increased, SAN Volume Controller does not use the additional space. To increase the storage capacity that is managed on a storage system, create a new LU on the storage system and add the MDisk that represents this LU to the storage pool.

3 **Attention:** If you delete a RAID array that is being used by SAN Volume Controller, the storage pool
3 goes offline and the data in that group is lost.

MDisks

A *managed disk (MDisk)* is a logical unit of physical storage. MDisks are either RAID arrays from internal
storage or volumes from external storage systems. MDisks are not visible to host systems.

An MDisk might consist of multiple physical disks that are presented as a single logical disk to the
storage area network (SAN). An MDisk always provides usable blocks of physical storage to the cluster
even if it does not have a one-to-one correspondence with a physical disk.

Each MDisk is divided into a number of extents, which are numbered, from 0, sequentially from the start
t to the end of the MDisk. The extent size is a property of storage pools. When an MDisk is added to a
t storage pool, the size of the extents that the MDisk is divided into depends on the attribute of the storage
t pool to which it has been added.

Access modes

The access mode determines how the cluster uses the MDisk. The following list provides the three types
of possible access modes:

Unmanaged
> The MDisk is not used by the cluster.

Managed
t > The MDisk is assigned to a storage pool and provides extents that volumes can use.

t **Image** The MDisk is assigned directly to a volume with a one-to-one mapping of extents between the
t > MDisk and the volume.

t **Attention:** If you add an MDisk that contains existing data to a storage pool while the MDisk is in
unmanaged or managed mode, you lose the data that it contains. The *image mode* is the only mode that
preserves this data.

Figure 8 shows physical disks and MDisks.
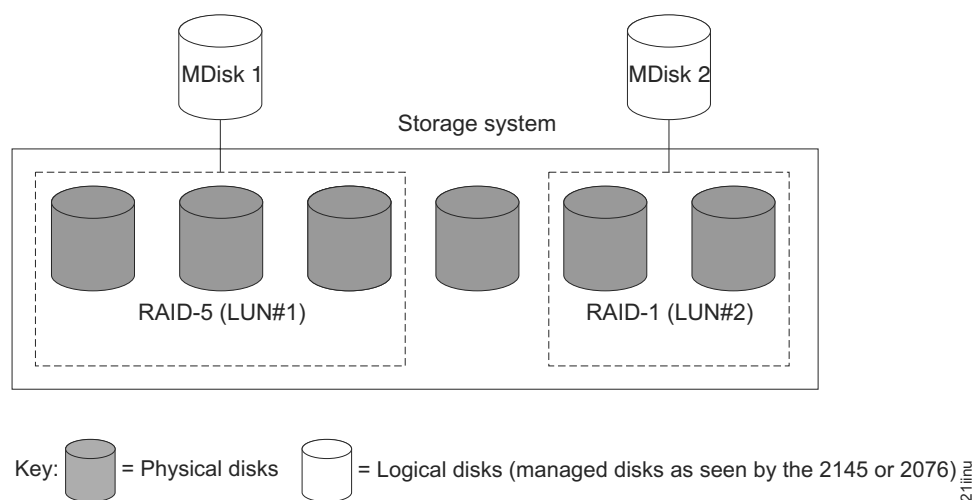
Figure 8. Storage Systems and MDisks

Table 6 on page 18 describes the operational states of an MDisk.

Table 6. MDisk status

Status	Description
Online	The MDisk can be accessed by all online nodes. That is, all the nodes that are currently working members of the cluster can access this MDisk. The MDisk is online when the following conditions are met: • All timeout error recovery procedures complete and report the disk as online. • Logical unit number (LUN) inventory of the target ports correctly reported the MDisk. • Discovery of this LUN completed successfully. • All of the MDisk target ports report this LUN as available with no fault conditions.
Degraded paths	The MDisk is not accessible to one or more nodes in the cluster. Degraded path status is most likely the result of incorrect configuration of either the disk controller or the fibre-channel fabric. However, hardware failures in the disk controller, fibre-channel fabric, or node could also be a contributing factor to this state. Complete the following actions to recover from this state: 1. Verify that the fabric configuration rules for storage systems are correct. 2. Ensure that you have configured the storage system properly. 3. Correct any errors in the event log.
Degraded ports	The MDisk has one or more 1220 errors in the event log. The 1220 error indicates that the remote fibre-channel port has been excluded from the MDisk. This error might cause reduced performance on the storage controller and usually indicates a hardware problem with the storage controller. To fix this problem you must resolve any hardware problems on the storage controller and fix the 1220 errors in the event log. To resolve these errors in the log, select **Troubleshooting** → **Recommended Actions** in the management GUI. This action displays a list of unfixed errors that are currently in the event log. For these unfixed errors, select the error name to begin a guided maintenance procedure to resolve them. Errors are listed in descending order with the highest priority error listed first. Resolve highest priority errors first.
Excluded	The MDisk has been excluded from use by the cluster after repeated access errors. Run the Directed Maintenance Procedures to determine the problem.
Offline	The MDisk cannot be accessed by any of the online nodes. That is, all of the nodes that are currently working members of the cluster cannot access this MDisk. This state can be caused by a failure in the SAN, storage system, or one or more physical disks connected to the storage system. The MDisk is reported as offline if all paths to the disk fail.

Attention: If you have observed intermittent breaks in links or if you have been replacing cables or connections in the SAN fabric or LAN configuration, you might have one or more MDisks in degraded status. If an I/O operation is attempted when a link is broken and the I/O operation fails several times, the system partially excludes the MDisk and it changes the status of the MDisk to degraded. You must include the MDisk to resolve the problem.

You can include the MDisk by either selecting **Physical Storage** → **MDisks: Action** → **Include Excluded MDisk** in the management GUI, or by issuing the following command in the command-line interface (CLI):

```
svctask includemdisk mdiskname/id
```

Where *mdiskname/id* is the name or ID of your MDisk.

Extents

Each MDisk is divided into chunks of equal size called *extents*. Extents are a unit of mapping that provide the logical connection between MDisks and volume copies.

MDisk path

Each MDisk from external storage has an online path count, which is the number of nodes that have access to that MDisk; this represents a summary of the I/O path status between the cluster nodes and the storage device. The maximum path count is the maximum number of paths that have been detected by the cluster at any point in the past. If the current path count is not equal to the maximum path count, the MDisk might be degraded. That is, one or more nodes might not see the MDisk on the fabric.

RAID implementation

A Redundant Array of Independent Disks (RAID) is a method of configuring multiple drives either in an external storage system or across internal solid-state drives (SSDs) for high availability and high performance.

RAID is an ordered collection, or group, of physical devices (disk drive modules) that are used to define logical volumes or devices. An array is a group of drives that are designated to be managed with RAID technology. The collection of two or more drives presents the image of a single drive to the system. If a single device fails, data can be read or regenerated from the other drives in the array.

Figure 9 on page 20 shows the relationships of the RAID components on the system.

3

```
                              Cluster
┌──────────────────────────────────────────────────────────────┐
│                                                                │
├────────────────────────────────────────────────────────────── ┤
│                                                                │
└──────────────────────────────────────────────────────────────┘
```

1..4   0..240
I/O Group   Drive

0..128   0..1
Storage pool   Use = Member

1   0..1

0..128   Drive 0..1
1   * Array MDisk   Member

Raid level   1   1..16 Member ID   0..1
Raid status   Spare status
Spare status   Goals   New drive

svc00667
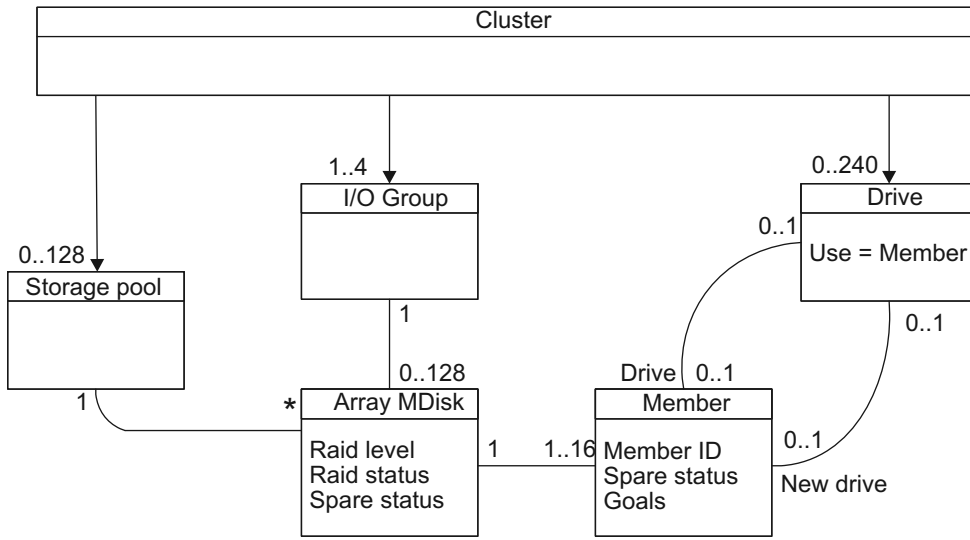
3
3 *Figure 9. RAID objects*
3

2 Physical capacity for SAN Volume Controller can be configured as RAID 0, RAID 1, RAID 5, RAID 6, or
2 RAID 10.

3 **RAID 0**
3   The system supports RAID 0 arrays, which are like the traditional just a bunch of disks (JBOD).
3   JBODs are disks that have not been configured according to the RAID system to increase fault
3   tolerance and improve data access performance. RAID 0 arrays stripe data across one or more
3   drives. RAID 0 arrays have no redundancy and do not support hot-spare takeover or immediate
3   exchange. RAID 0 arrays support concurrent exchange.

2 **RAID 1**
2   RAID 1 provides disk mirroring, which duplicates data between two drives. A RAID 1 array is
2   internally identical to a two-member RAID 10 array.

2 **RAID 5**
2   RAID 5 arrays stripe data over the member drives with one parity strip on every stripe. RAID 5
2   arrays have single redundancy with higher space efficiency than RAID 10 arrays, but with some
2   performance penalty. RAID 5 arrays can tolerate no more than one member drive failure.

2 **RAID 6**
2   RAID 6 arrays stripe data over the member drives with two parity strips on every stripe. The
2   parity strips are known as the P-parity and the Q-parity. A RAID 6 array can tolerate any two
2   concurrent member drive failures.

3 **RAID 10**
3   RAID 10 arrays stripe data over mirrored pairs of drives. RAID 10 arrays have single redundancy.
3   Although RAID 10 arrays can tolerate one failure from every mirrored pair, they cannot tolerate a
3   two-drive failure. The mirrored pairs rebuild independently. One member out of every pair can
3   be rebuilding or missing at the same time. RAID 10 combines the features of RAID 0 and RAID 1
3   to provide a second optimization for fault tolerance.

3 **Spare drive ratio**

3 For every 24 drives in RAID, 23 are members of an array and 1 must be a spare drive.

RAID levels and solid-state drive guidelines

Table 7 lists the various RAID levels available for solid-state drives and guidelines for obtaining the best performance in each level.

Table 7. RAID levels for solid-state drives (SSDs)

Name	Description	RAID level	Width goal	Spare goal	Pool limit	Chain balance	Hybrid pool
SSD RAID 5		RAID 5	8	1	0	Yes	No
SSD RAID 6		RAID 6	12	1	0	Yes	No
SSD RAID 10		RAID 10	8	1	0	Yes	No
SSD RAID 0		RAID 0	8	0	0	Yes	No
SSD Easy Tier	Use this for IBM System Storage Easy Tier hybrid storage pools	RAID 10	2	1	0	Yes	Yes

RAID levels and guidelines for other SAN Volume Controller types

Table 8 lists the RAID levels available for other SAN Volume Controller types guidelines for best performance.

Table 8. RAID levels for other SAN Volume Controller types

Name	Description	RAID level	Width goal	Spare goal	Pool limit	Chain balance	Hybrid pool
Mirrored volumes		RAID 10	8	0	0	No	No
Easy Tier		RAID 1	2	0	0	No	Yes

Storage pools and volumes

Managed disks (MDisks) are collected into groups known as *storage pools*. Volumes are logical disks that are presented to the SAN by SAN Volume Controller nodes. Volumes, like nodes, are associated with an I/O group.

Volume copies are created from the extents of MDisks.

Storage pools

A storage pool is a collection of MDisks that jointly contain all of the data for a specified set of volumes.

Figure 10 on page 22 shows a storage pool that contains four MDisks.
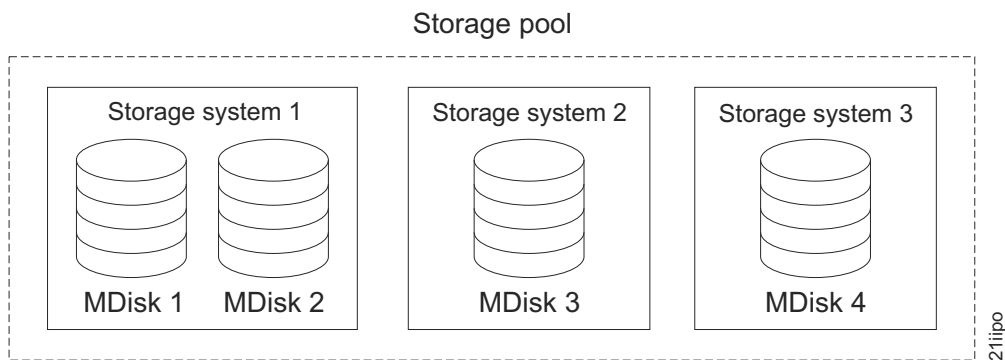
Storage pool

3 *Figure 10. Storage pool*

t All MDisks in a pool are split into extents of the same size. Volumes are created from the extents that are
t available in the storage pool. You can add MDisks to a storage pool at any time either to increase the
t number of extents that are available for new volume copies or to expand existing volume copies.

t You can specify a warning capacity for a storage pool. A warning event is generated when the amount of
t space that is used in the storage pool exceeds the warning capacity. This is especially useful in
   conjunction with thin-provisioned volumes that have been configured to automatically consume space
t from the storage pool.

t You can add only MDisks that are in unmanaged mode. When MDisks are added to a storage pool, their
   mode changes from unmanaged to managed.

   You can delete MDisks from a group under the following conditions:
t • Volumes are not using any of the extents that are on the MDisk.
   • Enough free extents are available elsewhere in the group to move any extents that are in use from this
     MDisk.

   Attention:
t • If you delete a storage pool, you destroy all the volumes that are made from the extents that are in the
     group.
   • If the group is deleted, you cannot recover the mapping that existed between extents that are in the
t   group or the extents that the volumes use. The MDisks that were in the storage pool are returned to
t   unmanaged mode and can be added to other storage pools. Because the deletion of a storage pool can
t   cause a loss of data, you must force the deletion if volumes are associated with it.
t • If the volume is mirrored and the synchronized copies of the volume are all in the storage pool, the
t   mirrored volume is destroyed when the storage pool is deleted.
t • If the volume is mirrored and there is a synchronized copy in another storage pool, the volume
t   remains after the storage pool is deleted.

t Table 9 describes the operational states of a storage pool.

   Table 9. Storage pool status

Status	Description
Online	The storage pool is online and available. All the MDisks in the storage pool are available.

Table 9. Storage pool status (continued)

Status	Description
Degraded paths	This status indicates that one or more nodes in the cluster cannot access all the MDisks in the group. Degraded path state is most likely the result of incorrect configuration of either the disk controller or the fibre-channel fabric. However, hardware failures in the disk controller, fibre-channel fabric, or node could also be a contributing factor to this state. Complete the following actions to recover from this state: 1. Verify that the fabric configuration rules for storage systems are correct. 2. Ensure that you have configured the storage system properly. 3. Correct any errors in the event log.
Degraded ports	This status indicates that one or more 1220 errors have been logged against the MDisks in the storage pool. The 1220 error indicates that the remote fibre-channel port has been excluded from the MDisk. This error might cause reduced performance on the storage controller and usually indicates a hardware problem with the storage controller. To fix this problem you must resolve any hardware problems on the storage controller and fix the 1220 errors in the event log. To resolve these errors in the log, click **Troubleshooting** → **Recommended Actions** in the management GUI. This action displays a list of unfixed errors that are currently in the event log. For these unfixed errors, select the error name to begin a guided maintenance procedure to resolve them. Errors are listed in descending order with the highest priority error listed first. Resolve highest priority errors first.
Offline	The storage pool is offline and unavailable. No nodes in the cluster can access the MDisks. The most likely cause is that one or more MDisks are offline or excluded.

Attention: If a single MDisk in a storage pool is offline and therefore cannot be seen by any of the online nodes in the cluster, the storage pool of which this MDisk is a member goes offline. This causes *all* the volume copies that are being presented by this storage pool to go offline. Take care when you create storage pools to ensure an optimal configuration.

Guidelines for creating storage pools

Consider the following guidelines when you create storage pools:

- Allocate your image-mode volumes between your storage pools.
- Ensure that all MDisks that are allocated to the same tier of a single storage pool are the same RAID type. This ensures that a single failure of a physical disk does not take the entire group offline. For example, if you have three RAID-5 arrays in one group and add a non-RAID disk to this group, you lose access to all the data striped across the group if the non-RAID disk fails. Similarly, for performance reasons, you must not mix RAID types. The performance of all volumes is reduced to the lowest performer in the tier.
- If you intend to keep the volume allocation within the storage exported by a storage system, ensure that the storage pool that corresponds to the storage system only contains storage that is presented by that storage system. This also enables nondisruptive migration of data from one storage system to another storage system and simplifies the decommissioning process if you want to decommission a storage system at a later time.

t • Except when you migrate between pools, you must associate a volume with just one storage pool.

t • An MDisk can be associated with just one storage pool.

t • In general, storage pools that consist of single-port attached systems are not supported by the SAN Volume Controller. However, in some cases, specifically on HP StorageWorks MA and EMA systems that contain RAID partitions, the only way these systems can be attached to the SAN Volume Controller is through single-port attach mode.

Extents

1 To track the space that is available on an MDisk, the SAN Volume Controller divides each MDisk into
1 chunks of equal size. These chunks are called *extents* and are indexed internally. Extent sizes can be 16,
1 32, 64, 128, 256, 512, 1024, 2048, 4096, or 8192 MB.

t Table 10 compares the maximum volume capacity for each extent size. The maximum is different for thin-provisioned volumes.

t *Table 10. Maximum volume capacity by extent size*

Extent size (MB)	Maximum volume capacity in GB (not thin-provisioned volumes)	Maximum volume capacity in GB (thin-provisioned volumes)
16	2048 (2 TB)	2000
32	4096 (4 TB)	4000
64	8192 (8 TB)	8000
128	16,384 (16 TB)	16,000
256	32,768 (32 TB)	32,000
512	65,536 (64 TB)	65,000
1024	131,072 (128 TB)	130,000
2048	262,144 (256 TB)	260,000
4096	262,144 (256 TB)	262,144
8192	262,144 (256 TB)	262,144

t You specify the extent size when you create a new storage pool. You cannot change the extent size later;
t it must remain constant throughout the lifetime of the storage pool.

t You cannot use the SAN Volume Controller data migration function to migrate volumes between storage
t pools that have different extent sizes. However, you can use the following SAN Volume Controller functions to move data to an MDisk that has a different extent size:

t • FlashCopy to copy a volume between a source and a destination storage pool that have different extent sizes.

t • Intracluster Metro Mirror or Global Mirror to copy a volume between a source and a destination
t storage pool that have different extent sizes.

t • Volume Mirroring to add a copy of the disk from the destination storage pool. After the copies are
t synchronized, you can free up extents by deleting the copy of the data in the source storage pool.

The choice of extent size affects the total amount of storage that is managed by the cluster. Table 11 shows the maximum amount of storage that can be managed by a cluster for each extent size.

Table 11. Capacities of the cluster given extent size

Extent size	Maximum storage capacity of cluster
16 MB	64 TB
32 MB	128 TB

Table 11. Capacities of the cluster given extent size (continued)

Extent size	Maximum storage capacity of cluster
64 MB	256 TB
128 MB	512 TB
256 MB	1 PB
512 MB	2 PB
1024 MB	4 PB
2048 MB	8 PB
4096 MB	16 PB
8192 MB	32 PB

A cluster can manage 2^{22} extents. For example, with a 16 MB extent size, the cluster can manage up to 16 MB x 4,194,304 = 64 TB of storage.

When you choose an extent size, consider your future needs. For example, if you currently have 40 TB of storage and you specify an extent size of 16 MB, the capacity of the storage pool is limited to 64 TB of storage in the future. If you select an extent size of 64 MB, the capacity of the storage pool is 256 TB.

Using a larger extent size can waste storage. When a volume is created, the storage capacity for the volume is rounded to a whole number of extents. If you configure the system to have a large number of small volumes and you use a large extent size, this can cause storage to be wasted at the end of each volume.

Easy Tier function

SAN Volume Controller includes IBM System Storage Easy Tier, a function that responds to the presence of solid-state drives (SSDs) in a storage pool that also contains hard disk drives (HDDs). The system automatically and nondisruptively moves frequently accessed data from HDD MDisks to SSD MDisks, thus placing such data in a faster tier of storage.

Easy Tier eliminates manual intervention when assigning highly active data on volumes to faster responding storage. In this dynamically tiered environment, data movement is seamless to the host application regardless of the storage tier in which the data resides. Manual controls exist so that you can change the default behavior, for example, such as turning off Easy Tier on storage pools that have both types of MDisks.

SAN Volume Controller supports these tiers:

Generic SSD tier
The SSD tier exists when SSDs are in the storage pool. The SSDs provide greater performance than hard disk drives (HDDs).

Generic HDD tier
The HDD tier exists when HDDs are in the storage pool.

All MDisks belong to one tier or the other, which includes MDisks that are not yet part of a storage pool.

If you create a storage pool (managed disk group) with both generic SSD MDisks (classified with the generic_ssd option) and generic HDD MDisks (generic_hdd or default option), Easy Tier is automatically turned on for pools with both SSD MDisks and HDD MDisks. SAN Volume Controller does not automatically identify external SSD MDisks; all external MDisks are put into the HDD tier by default. Users must manually identify external SSD MDisks and change their tiers. Local (internal) MDisks are automatically classified as generic_ssd and are placed in the SSD tier without user intervention.

Easy Tier modes of operation:

SAN Volume Controller supports solid-state drives (SSDs) that offer a number of potential benefits over magnetic hard disk drives (HDDs), such as faster data access and throughput, better performance, and less power consumption.

SDDs are, however, much more expensive than HDDs. To optimize SSD performance and help provide a cost-effective contribution to the overall system, Easy Tier can cause infrequently accessed data to reside on lower cost HDDs and frequently accessed data to reside on SSDs.

Determining the amount of data activity in an extent and when to move the extent to the proper storage tier is usually too complex a task to manage manually.

Easy Tier evaluation mode collects usage statistics for each storage extent for a storage pool where the capability of moving data from one tier to the other tier is not possible or is disabled. An example of such a storage pool is a pool of homogeneous MDisks, where all MDisks are typically HDDs. A summary file is created in the /dumps directory on the configuration node (dpa_heat.*node_name.date.time*.data), which can be offloaded and viewed by using the IBM Storage Tier Advisor Tool.

Easy Tier automatic data placement mode also measures the amount of data access, but then acts on the measurements to automatically place the data into the appropriate tier of a storage pool that contains both MDisk tiers.

Dynamic data movement is transparent to the host server and application users of the data, other than providing improved performance.

For a storage pool and volume to be automatically managed by Easy Tier, ensure that the following conditions are met:
- The volume must be striped.
- The storage pool must contain both MDisks that belong to the generic_ssd tier and MDisks that belong to the generic_hdd tier.

Volumes that are added to storage pools use extents from generic_hdd MDisks initially, if available. Easy Tier then collects usage statistics to determine which extents to move to generic_ssd MDisks.

Easy Tier evaluation mode:

When IBM System Storage Easy Tier evaluation mode is enabled for a storage pool with a single tier of storage, Easy Tier collects usage statistics for all the volumes in the pool.

SAN Volume Controller monitors the storage use at the volume extent level. Easy Tier constantly gathers and analyzes monitoring statistics to derive moving averages for the past 24 hours.

Volumes are not monitored when the easytier attribute of a storage pool is set to *off* or *auto* with a single tier of storage. You can enable Easy Tier evaluation mode for a storage pool with a single tier of storage by setting the easytier attribute of the storage pool to *on*.

You can control or view data placement settings by using the following command-line interface (CLI) commands:

chmdiskgrp
> Modifies the properties of the storage pool. Use this command to turn on evaluation mode on a storage pool with a single tier of storage and to turn off Easy Tier functions on a storage pool with more than one tier of storage.

lsmdiskgrp
> Lists storage pool information.

lsvdisk
　　　Lists volume information.

lsvdiskcopy
　　　Lists volume copy information.

mkmdiskgrp
　　　Creates a new storage pool.

Other MDisk commands such as addmdisk, chmdisk, and lsmdisk can be used to view or set the tier an MDisk belongs to.

Automatic data placement:

When IBM System Storage Easy Tier on SAN Volume Controller automatic data placement mode is active, Easy Tier measures the host access activity to the data on each storage extent, provides a mapping that identifies high activity extents, and then moves the high-activity data according to its relocation plan algorithms.

To automatically relocate the data, Easy Tier performs the following actions:

1. Monitors volumes for host access to collect average performance statistics for each extent over a rolling 24-hour period of I/O activity.
2. Analyzes the amount of I/O activity for each extent to determine if the extent is a candidate for migrating to or from the higher performing solid-state drive (SSD) tier.
3. Develops an extent relocation plan for each storage pool to determine exact data relocations within the storage pool. Easy Tier then automatically relocates the data according to the plan.

While relocating volume extents, Easy Tier follows these actions:

- Attempts to migrate the most active volume extents first.
- Refreshes the task list as the plan changes. The previous plan and any queued extents that are not yet relocated are abandoned.

Automatic data placement is enabled by default for storage pools with more than one tier of storage. When you enable automatic data placement, by default all striped volumes are candidates for automatic data placement. Image mode and sequential volumes are never candidates for automatic data placement. When automatic data placement is enabled, I/O monitoring is done for all volumes whether the volume is a candidate for automatic data placement. Once automatic data placement is enabled, and if there is sufficient activity to warrant relocation, extents will begin to be relocated within a day after enablement. You can control whether Easy Tier automatic data placement and I/O activity monitoring is enabled or disabled by using the settings for each storage pool and each volume. Each command in the following table that can create or change the settings for storage pools can enable or disable both Easy Tier functions. Any command that can create or change the settings for volumes can enable or disable automatic data replacement, if automatic data replacement is enabled for the storage pool. You can control or view automatic data placement by using the following command-line interface (CLI) commands:

addvdiskcopy
　　　Adds a copy to an existing volume by changing a nonmirrored volume into a mirrored volume.

chmdiskgrp
　　　Modifies the properties of the storage pool. Use this command to turn on Easy Tier evaluation mode or I/O monitoring and to turn off Easy Tier functions on a storage pool with more than one tier of storage.

　　　Note: When automatic data placement is active on a storage pool, set a warning threshold for the storage pool. Automatic data placement cannot function if the storage pool is used 100%.

2 **chvdisk**
2   Modifies the properties of a volume.

2 **lsmdiskgrp**
2   Lists storage pool information.

2 **lsvdisk**
2   Lists volume information.

2 **lsvdiskcopy**
2   Lists volume copy information.

2 **mkmdiskgrp**
2   Creates a storage pool.

2 **mkvdisk**
2   Creates sequential, striped, or image mode volumes.

3 If you want to disable automatic data placement for a volume or storage pool, set the easytier attribute to
3 *off*.

2 **Extracting and viewing performance data with the IBM Storage Tier Advisor Tool:**

2 You can use the IBM Storage Tier Advisor Tool, hereafter referred to as advisor tool, to view performance
2 data that is collected by IBM System Storage Easy Tier over a 24-hour operational cycle. The advisor tool
2 is the application that creates a Hypertext Markup Language (HTML) file that you use to view the data
2 when you point your browser to the file.

3 To download the Storage Tier Advisor Tool, go to the Download section at this website:

2 Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

2 To extract the summary performance data, follow these steps using the command-line interface (CLI):
2 1. Find the most recent dpa_heat.*node_name.date.time*.data file in the cluster by entering the following
2   command-line interface (CLI) command:
2   `svcinfo lsdumps` *node_id* | *node_name*

2   where *node_id* | *node_name* is the node ID or name to list the available dumps for.
3 2. If necessary, copy the most recent summary performance data file to the current configuration node.
3   Enter the following command:
3   `svctask cpdumps -prefix /dumps/dpa_heat.`*node_name.date.time*`.data` *node_id* | *node_name*
3 3. Use PuTTY scp (pscp) to copy the summary performance data in a binary format from the
3   configuration node to a local directory.
3 4. From a Microsoft Windows command prompt, use the advisor tool to transform the binary file in
3   your local directory into an HTML file in the local directory.
2 5. Point your browser to the HTML file in your local directory.

3 The advisor tool displays three types of statistical reports:

2 **System summary report**
2   • The number of volumes that are monitored
2   • The estimated total hot data capacity
2   • An estimated time the migration process will take to move hot data to the SSDs
2   • A summary recommendation for SSD capacity and estimated performance improvement

2 **System recommendation report**
2

2　　　　　• A sorted list of suggested SSD MDisks to add to the storage pools (sorted by estimated
2　　　　　　performance improvement)
2　　　　　• For each suggested MDisk: the target storage pool and estimated performance improvement

3　**Storage pool recommendation report**

3　　　　　• List of storage pools that were monitored by the Easy Tier function.

3　　　　　• For each storage pool: a sorted list of suggested SSD MDisks to add to that storage pool
3　　　　　　(sorted by estimated performance improvement)

3　　　　　• For each suggested MDisk: the estimated performance improvement

2　**Volume heat distribution report**

3　　　　　• The distribution of hot data and cold data for each volume copy

3　　　　　• The configured capacity of a volume copy, along with the volume ID, copy ID, and the storage
3　　　　　　pool ID

3　　　　　• The portion of the capacity of each volume copy already on SSD

2　You can view this information to analyze workload statistics and evaluate which logical volumes might
2　be candidates for Easy Tier management. If you have not enabled the Easy Tier, you can use the
2　performance statistics gathered by the monitoring process to help you determine whether to use Easy
2　Tier to enable potential performance improvements in your storage environment.

2　**Easy Tier automatic data placement requirements and limitations:**

2　Some limitations exist when using IBM System Storage Easy Tier on SAN Volume Controller.

2　• The Easy Tier function supports the following tiered storage configurations:

3　　– Local (internal) Serial Attached SCSI (SAS) solid-state drives (SSDs) in a storage pool with
3　　　fibre-channel-attached hard disk drives (HDDs).

3　　– External fibre-channel-attached SSDs in a storage pool with fibre-channel-attached hard disk drives
3　　　(HDDs)..

3　• To ensure optimal performance, all MDisks in a storage pool tier must have the same technology and
3　　performance characteristics.

3　• Easy Tier automatic data placement is not supported on volume copies, which are image mode or
3　　sequential. I/O monitoring for such volumes is supported, but you cannot migrate extents on such
3　　volumes unless you convert image or sequential volume copies to striped volumes.

3　• Automatic data placement and extent I/O activity monitors are supported on each copy of a mirrored
3　　volume. Easy Tier works with each copy independently of the other copy. For example, you can enable
3　　or disable Easy Tier automatic data placement for each copy independently of the other copy.

3　• SAN Volume Controller creates new volumes or volume expansions using extents from MDisks from
3　　the HDD tier, if possible, but uses extents from MDisks from the SSD tier if necessary.

3　• When a volume is migrated out of a storage pool that is managed with Easy Tier, Easy Tier automatic
3　　data placement mode is no longer active on that volume. Automatic data placement is also turned off
3　　while a volume is being migrated even if it is between pools that both have Easy Tier automatic data
3　　placement enabled. Automatic data placement for the volume is re-enabled with the migration is
3　　complete.

2　**Limitations when removing an MDisk by using the force parameter**

3　When an MDisk is deleted from a storage pool with the **force** parameter, extents in use are migrated to
3　MDisks in the same tier as the MDisk being removed, if possible. If insufficient extents exist in that tier,
3　extents from the other tier are used.

2   **Limitations when migrating extents**

3   When Easy Tier automatic data placement is enabled for a volume, the **svctask migrateexts**
3   command-line interface (CLI) command cannot be used on that volume.

2   **Limitations when migrating a volume to another storage pool**

2   When SAN Volume Controller migrates a volume to a new storage pool, Easy Tier automatic data
2   placement between the generic SSD tier and the generic HDD tier is temporarily suspended. After the
2   volume is migrated to its new storage pool, Easy Tier automatic data placement between the generic SSD
2   tier and the generic HDD tier resumes for the newly moved volume, if appropriate.

3   When SAN Volume Controller migrates a volume from one storage pool to another, it will attempt to
3   migrate each extent to an extent in the new storage pool from the same tier as the original extent. In
3   some cases, such as a target tier being unavailable, the other tier is used. For example, the generic SSD
3   tier might be unavailable in the new storage pool.

3   If the automatic data placement is enabled in the new storage pool, pending Easy Tier status changes are
3   assigned after the volume completes its move to the new storage pool. Although the status changes are
3   based on volume use in the old storage pool, the new status is honored in the new storage pool.

2   **Limitations when migrating a volume to image mode**

3   Easy Tier automatic data placement does not support image mode. No automatic data placement occurs
3   in this situation. When a volume with Easy Tier automatic data placement mode active is migrated to
3   image mode, Easy Tier automatic data placement mode is no longer active on that volume.

2   Easy Tier does support evaluation mode for image mode volumes.

Volumes

t   A volume is a logical disk that the cluster presents to the hosts.

1   Application servers on the SAN access volumes, not MDisks or drives. To keep a volume accessible even
    when an MDisk on which it depends has become unavailable, a mirrored copy can be added to a selected
t   volume. Each volume can have a maximum of two copies. Each volume copy is created from a set of
t   extents in a storage pool.

t   There are three types of volumes: striped, sequential, and image.

Types

t   Each volume copy can be one of the following types:

Striped
t   A volume copy that has been striped is at the extent level. One extent is allocated, in turn, from
t   each MDisk that is in the group. For example, a storage pool that has 10 MDisks takes one extent
            from each MDisk. The 11th extent is taken from the first MDisk, and so on. This procedure,
            known as a round-robin, is similar to RAID-0 striping.

            You can also supply a list of MDisks to use as the stripe set. This list can contain two or more
t   MDisks from the storage pool. The round-robin procedure is used across the specified stripe set.

t **Attention:** By default, striped volume copies are striped across all MDisks in the storage pool. If some of the MDisks are smaller than others, the extents on the smaller MDisks are used up before the larger MDisks run out of extents. Manually specifying the stripe set in this case might
t result in the volume copy not being created.

t If you are unsure if there is sufficient free space to create a striped volume copy, select one of the following options:

- Check the free space on each MDisk in the group using the **svcinfo lsfreeextents** command.
t - Let the system automatically create the volume copy by not supplying a specific stripe set.

t Figure 11 shows an example of a storage pool that contains three MDisks. This figure also shows
t a stripedvolume copy that is created from the extents that are available in the storage pool.

Storage pool

Mdisk 1	Mdisk 2	Mdisk 3
Extent 1a	Extent 2a	Extent 3a
Extent 1b	Extent 2b	Extent 3b
Extent 1c	Extent 2c	Extent 3c
Extent 1d	Extent 2d	Extent 3d
Extent 1e	Extent 2e	Extent 3e
Extent 1f	Extent 2f	Extent 3f
Extent 1g	Extent 2g	Extent 3g

Volume

Create a striped volume

Extent 1a
Extent 2a
Extent 3a
Extent 1b
Extent 2b
Extent 3b
Extent 1c
Extent 2c
Extent 3c
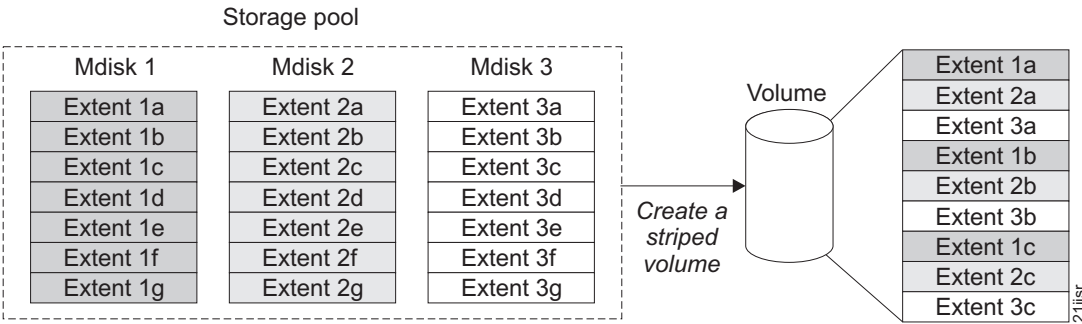
t *Figure 11. Storage pools and volumes*

Sequential
t When extents are selected, they are allocated sequentially on one MDisk to create the volume copy if enough consecutive free extents are available on the chosen MDisk.

t **Image** Image-mode volumes are special volumes that have a direct relationship with one MDisk. If you have an MDisk that contains data that you want to merge into the cluster, you can create an
t image-mode volume. When you create an image-mode volume, a direct mapping is made
t between extents that are on the MDisk and extents that are on the volume. The MDisk is not
t virtualized. The logical block address (LBA) x on the MDisk is the same as LBA x on the volume.

t When you create an image-mode volume copy, you must assign it to a storage pool. An
t image-mode volume copy must be at least one extent in size. The minimum size of an
t image-mode volume copy is the extent size of the storage pool to which it is assigned.

t The extents are managed in the same way as other volume copies. When the extents have been
t created, you can move the data onto other MDisks that are in the storage pool without losing
t access to the data. After you move one or more extents, the volume copy becomes a virtualized
disk, and the mode of the MDisk changes from image to managed.

t **Attention:** If you add a managed mode MDisk to a storage pool, any data on the MDisk is lost.
t Ensure that you create image-mode volumes from the MDisks that contain data before you start
t adding any MDisks to storage pools.

MDisks that contain existing data have an initial mode of unmanaged, and the cluster cannot determine if it contains partitions or data.

t You can use more sophisticated extent allocation policies to create volume copies. When you create a
t striped volume, you can specify the same MDisk more than once in the list of MDisks that are used as
t the stripe set. This is useful if you have a storage pool in which not all the MDisks are of the same
t capacity. For example, if you have a storage pool that has two 18 GB MDisks and two 36 GB MDisks, you

can create a striped volume copy by specifying each of the 36 GB MDisks twice in the stripe set so that two-thirds of the storage is allocated from the 36 GB disks.

If you delete a volume, you destroy access to the data that is on the volume. The extents that were used in the volume are returned to the pool of free extents that is in the storage pool. The deletion might fail if the volume is still mapped to hosts. The deletion might also fail if the volume is still part of a FlashCopy, Metro Mirror, or Global Mirror mapping. If the deletion fails, you can specify the force-delete flag to delete both the volume and the associated mappings to hosts. Forcing the deletion deletes the Copy Services relationship and mappings.

States

A volume can be in one of three states: online, offline, and degraded. Table 12 describes the different states of a volume.

Table 12. Volume states

State	Description
Online	At least one synchronized copy of the volume is online and available if both nodes in the I/O group can access the volume. A single node can only access a volume if it can access all the MDisks in the storage pool that are associated with the volume.
Offline	The volume is offline and unavailable if both nodes in the I/O group are missing, or if none of the nodes in the I/O group that are present can access any synchronized copy of the volume. The volume can also be offline if the volume is the secondary of a Metro Mirror or Global Mirror relationship that is not synchronized. A thin-provisioned volume goes offline if a user attempts to write an amount of data that exceeds the available disk space.
Degraded	The status of the volume is degraded if one node in the I/O group is online and the other node is either missing or cannot access any synchronized copy of the volume. **Note:** If you have a degraded volume and all of the associated nodes and MDisks are online, call the IBM Support Center for assistance.

Cache modes

You can select to have read and write operations stored in cache by specifying a cache mode. You can specify the cache mode when you create the volume. After the volume is created, you can change the cache mode.

Table 13 describes the two types of cache modes for a volume.

Table 13. Volume cache modes

Cache mode	Description
readwrite	All read and write I/O operations that are performed by the volume are stored in cache. This is the default cache mode for all volumes.
none	All read and write I/O operations that are performed by the volume are not stored in cache.

Mirrored volumes:

By using volume mirroring, a volume can have two physical copies. Each volume copy can belong to a different storage pool, and each copy has the same virtual capacity as the volume.

When a server writes to a mirrored volume, the cluster writes the data to both copies. When a server reads a mirrored volume, the cluster picks one of the copies to read. If one of the mirrored volume copies is temporarily unavailable; for example, because the storage system that provides the storage pool is unavailable, the volume remains accessible to servers. The cluster remembers which areas of the volume are written and resynchronizes these areas when both copies are available.

You can create a volume with one or two copies and convert a non-mirrored volume into a mirrored volume by adding a copy. When a copy is added in this way, the SAN Volume Controller cluster synchronizes the new copy so that it is the same as the existing volume. Servers can access the volume during this synchronization process.

You can convert a mirrored volume into a non-mirrored volume by deleting one copy or by splitting one copy to create a new non-mirrored volume.

The volume copy can be any type: image, striped, sequential, and either thin-provisioned or fully allocated. The two copies can be of completely different types.

You can use mirrored volumes for the following reasons:
- Improving availability of volumes by protecting them from a single storage system failure.
- Providing concurrent maintenance of a storage system that does not natively support concurrent maintenance.
- Providing an alternative method of data migration with better availability characteristics. While a volume is being migrated using the data migration feature, it is vulnerable to failures on both the source and target storage pool. Volume mirroring provides an alternative because you can start with a non-mirrored volume in the source storage pool, and then add a copy to that volume in the destination storage pool. When the volume is synchronized, you can delete the original copy that is in the source storage pool. During the synchronization process, the volume remains available even if there is a problem with the destination storage pool.
- Converting between fully allocated volumes and thin-provisioned volumes.

When you use volume mirroring, consider how quorum candidate disks are allocated. Volume mirroring maintains some state data on the quorum disks. If a quorum disk is not accessible and volume mirroring is unable to update the state information, a mirrored volume might need to be taken offline to maintain data integrity. To ensure the high availability of the system, ensure that multiple quorum candidate disks, allocated on different storage systems, are configured.

Attention: Mirrored volumes can be taken offline if there is no quorum disk available. This behavior occurs because synchronization status for mirrored volumes is recorded on the quorum disk. To protect against mirrored volumes being taken offline, follow the guidelines for setting up quorum disks.

Image mode volumes:

An image mode volume provides a direct block-for-block translation from the managed disk (MDisk) to the volume with no virtualization.

This mode is intended to provide virtualization of MDisks that already contain data that was written directly, not through a SAN Volume Controller node. Image mode volumes have a minimum size of 1 block (512 bytes) and always occupy at least one extent.

Image mode MDisks are members of a storage pool, but they do not contribute to free extents. Image mode volumes are not affected by the state of the storage pool because the storage pool controls image mode volumes through the association of the volume to an MDisk. Therefore, if an MDisk that is

associated with an image mode volume is online and the storage pool of which they are members goes offline, the image mode volume remains online. Conversely, the state of a storage pool is not affected by the state of the image mode volumes in the storage pool.

An image mode volume behaves just as a managed mode volume in terms of the Metro Mirror, Global Mirror, and FlashCopy Copy Services. Image mode volumes are different from managed mode in two ways:

- Migration. An image mode volume can be migrated to another image mode volume. It becomes managed while the migration is ongoing, but returns to image mode when the migration is complete.
- Quorum disks. Image mode volumes cannot be quorum disks. This means that a cluster with only image mode volumes does not have a quorum disk.

Migration methods for image mode volumes:

Several methods can be used to migrate image mode volumes into managed mode volumes.

To perform any type of migration activity on an image mode volume, the image mode volume must first be converted into a managed mode volume. The volume is automatically converted into a managed mode volume whenever any kind of migration activity is attempted. After the image-mode-to-managed-mode migration operation has occurred, the volume becomes a managed mode volume and is treated the same way as any other managed mode volume.

If the image mode disk has a partial last extent, this last extent in the image mode volume must be the first to be migrated. This migration is processed as a special case. After this special migration operation has occurred, the volume becomes a managed mode volume and is treated in the same way as any other managed mode volume. If the image mode disk does not have a partial last extent, no special processing is performed. The image mode volume is changed into a managed mode volume and is treated the same way as any other managed mode volume.

An image mode disk can also be migrated to another image mode disk. The image mode disk becomes managed while the migration is ongoing, but returns to image mode when the migration is complete.

You can perform the following types of migrations:
- Migrate extents
- Migrate a volume
- Migrate to image mode

Note: Migration commands fail if the target or source volume is offline, or if there is insufficient quorum disk space to store the metadata. Correct the offline or quorum disk condition and reissue the command.

Perform the following steps to migrate volumes:
1. Dedicate one storage pool to image mode volumes.
2. Dedicate one storage pool to managed mode volumes.
3. Use the migrate volume function to move the volumes.

Thin-provisioned volumes:

When you create a volume, you can designate it as thin-provisioned. A thin-provisioned volume has a virtual capacity and a real capacity.

Virtual capacity is the volume storage capacity that is available to a host. *Real capacity* is the storage capacity that is allocated to a volume copy from a storage pool. In a fully allocated volume, the virtual capacity and real capacity are the same. In a thin-provisioned volume, however, the virtual capacity can be much larger than the real capacity.

The virtual capacity of a thin-provisioned volume is typically significantly larger than its real capacity. Each SAN Volume Controller cluster uses the real capacity to store data that is written to the volume, and metadata that describes the thin-provisioned configuration of the volume. As more information is written to the volume, more of the real capacity is used. The SAN Volume Controller cluster identifies read operations to unwritten parts of the virtual capacity and returns zeros to the server without using any of the real capacity.

SAN Volume Controller must maintain extra metadata that describes the contents of thin-provisioned volumes. This means the I/O rates that are obtained from thin-provisioned volumes are slower than those obtained from fully allocated volumes that are allocated on the same MDisks.

Thin-provisioned volumes can also help simplify server administration. Instead of assigning a volume with some capacity to an application and increasing that capacity as the needs of the application change, you can configure a volume with a large virtual capacity for the application, and then increase or shrink the real capacity as the application needs change, without disrupting the application or server.

When you configure a thin-provisioned volume, you can use the warning level attribute to generate a warning event when the used real capacity exceeds a specified amount or percentage of the total real capacity. You can also use the warning event to trigger other actions, such as taking low-priority applications offline or migrating data into other storage pools.

If a thin-provisioned volume does not have enough real capacity for a write operation, the volume is taken offline and an error is logged (error code 1865, event ID 060001). Access to the thin-provisioned volume is restored by either increasing the real capacity of the volume or increasing the size of the storage pool that it is allocated on.

Note: On a SAN Volume Controller 2145-CF8 node, space is not allocated on a thin-provisioned volume if an incoming host write operation contains all zeros.

When you create a thin-provisioned volume, you can choose the grain size for allocating space in 32 KB, 64 KB, 128 KB, or 256 KB chunks. The grain size that you select affects the maximum virtual capacity for the thin-provisioned volume. If you select 32 KB for the grain size, the volume size cannot exceed 260,000 GB. The grain size cannot be changed after the thin-provisioned volume has been created. Generally, smaller grain sizes save space but require more metadata access, which can adversely impact performance. If you are not going to use the thin-provisioned volume as a FlashCopy source or target volume, use 256 KB to maximize performance. If you are going to use the thin-provisioned volume as a FlashCopy source or target volume, specify the same grain size for the volume and for the FlashCopy function.

When you create a thin-provisioned volume, set the cache mode to `readwrite` to maximize performance. If the cache mode is set to `none`, the SAN Volume Controller cluster cannot cache the thin-provisioned metadata, which decreases performance.

The autoexpand feature prevents a thin-provisioned volume from using up its capacity and going offline. As a thin-provisioned volume uses capacity, the autoexpand feature maintains a fixed amount of unused real capacity, called the *contingency capacity*. For thin-provisioned volumes that are not configured with the autoexpand feature, the contingency capacity can get used up, causing the volume to go offline. To determine if an application requires a thin-provisioned volume with the autoexpand feature, create a thin-provisioned volume with the autoexpand feature turned off. If the application causes the volume to run out of capacity and go offline, you can then create a thin-provisioned volume with the autoexpand feature turned on.

Image mode thin-provisioned volumes:

When you create an image mode volume, you can designate it as thin-provisioned. An image mode thin-provisioned volume has a virtual capacity and a real capacity.

An image mode thin-provisioned volume has a direct relationship with a single MDisk where the contents of the MDisk map to the real capacity that is used by the thin-provisioned volume. Unlike fully allocated volumes, the logical block address (LBA) on the MDisk is not necessarily the same as the LBA on the volume. You cannot change the real capacity of an image mode thin-provisioned volume manually or by using the autoexpand feature. To use the autoexpand feature, the volume must be in managed mode.

You can use an image mode volume to move a thin-provisioned volume between two SAN Volume Controller clusters using the following procedure. The procedure is similar to that used for fully allocated volumes, but has an extra step during the import process to specify the existing thin-provisioned metadata, rather than to create a new, empty volume.

1. If the volume is not already in image mode, migrate the volume to image mode and wait for the migration to complete.
2. Delete the volume from the exporting cluster.
3. Disconnect the MDisk from the exporting cluster and connect the MDisk to the importing cluster.
4. Create a new image mode thin-provisioned volume using the MDisk. You must specify the **import** option.
5. Optionally, migrate the volume to managed mode.

The **import** option is only valid for SAN Volume Controller thin-provisioned volumes. If you use this method to import a thin-provisioned volume that is created by RAID controllers into a cluster, SAN Volume Controller cannot detect it as a thin-provisioned volume. However, you can use the volume mirroring feature to convert an image-mode fully allocated volume to a thin-provisioned volume.

Converting thin-provisioned volumes:

You can convert thin-provisioned volumes into fully allocated volumes.

You can nondisruptively convert a thin-provisioned volume into a fully allocated volume by using the following volume mirroring procedure:
1. Start with a single-copy, thin-provisioned volume.
2. Add a fully allocated copy to the volume.
3. Wait while the volume mirroring feature synchronizes.
4. Remove the thin-provisioned copy from the volume.

Converting fully allocated volumes:

You can convert fully allocated volumes to thin-provisioned volumes.

You can nondisruptively convert a fully allocated volume into a thin-provisioned volume by following this procedure:
1. Start with a single copy, fully allocated volume.
2. Add a thin-provisioned copy to the volume. Use a small real capacity and the autoexpand feature.
3. Wait while the volume mirroring feature synchronizes the copies.
4. Remove the fully allocated copy from the thin-provisioned volume.

Any grains of the fully allocated volume that contain all zeros do not cause any real capacity to be allocated on the thin-provisioned copy. Before you create the mirrored copy, you can fill the free capacity on the volume with a file that contains all zeros.

I/O governing:

t You can set the maximum amount of I/O activity that a host sends to a volume. This amount is known as the *I/O governing rate*. The governing rate can be expressed in I/Os per second or MB per second.

Read, write, and verify commands that access the physical medium are subject to I/O governing.

I/O governing does not affect FlashCopy and data migration I/O rates.

3 I/O governing on a Metro Mirror and Global Mirror secondary volume does not affect the rate of data
3 copy from the primary volume.

Host objects

A *host system* is a computer that is connected to SAN Volume Controller through either a fibre-channel interface or an IP network.

A *host object* is a logical object in SAN Volume Controller that represents a list of worldwide port names (WWPNs) and a list of iSCSI names that identify the interfaces that the host system uses to communicate with SAN Volume Controller. iSCSI names can be either iSCSI qualified names (IQNs) or extended unique identifiers (EUIs).

A typical configuration has one host object for each host system that is attached to SAN Volume Controller. If a cluster of hosts accesses the same storage, you can add host bus adaptor (HBA) ports from several hosts to one host object to make a simpler configuration. A host object can have both WWPNs and iSCSI names.

t The cluster does not automatically present volumes to the host system. You must map each volume to a
t particular host object to enable the volume to be accessed through the WWPNs or iSCSI names that are associated with the host object.

When you create a new host object, the configuration interfaces provide a list of unconfigured WWPNs. These represent the WWPNs that the cluster has detected. Candidate iSCSI names are not available and must be entered manually.

The cluster can detect only WWPNs that have connected to the cluster through the fibre-channel network. Some fibre-channel HBA device drivers do not let the ports remain logged in if no disks are detected on the fabric. This can prevent some WWPNs from appearing in the list of candidate WWPNs. The configuration interface provides a method to manually type the port names.

Note: You must not include a WWPN or an iSCSI name that belongs to a SAN Volume Controller node in a host object.

A WWPN or iSCSI name can be added to only one host object.

Port masks

You can use the port-mask property of the host object to control the fibre-channel ports on each SAN Volume Controller node that a host can access. The port mask applies to logins from the WWPNs that are associated with the host object. The port-mask configuration has no effect on iSCSI connections.

For each login between a host fibre-channel port and node fibre-channel port, the node examines the port mask for the associated host object and determines if access is allowed or denied. If access is denied, the node responds to SCSI commands as if the HBA WWPN is unknown.

The port mask is four binary bits. Valid mask values range from 0000 (no ports enabled) to 1111 (all ports enabled). For example, a mask of 0011 enables port 1 and port 2. The default value is 1111.

Multiple target ports

When you create a host mapping to a fibre-channel attached host, the host ports that are associated with the host object can view the LUN that represents the volume on up to eight fibre-channel ports. Nodes follow the American National Standards Institute (ANSI) Fibre Channel (FC) standards for SCSI LUs that are accessed through multiple node ports. All nodes within a single I/O group present a consistent set of SCSI LUs across all ports on those nodes.

Similarly, all nodes within a single I/O group present a consistent set of SCSI LUs across all iSCSI ports on those nodes.

Node login counts

The number of nodes that can see each WWPN or iSCSI name is reported on a per-node basis and is known as the *node login count*. If the count is less than expected for the current configuration, you might have a connectivity problem.

Host mapping

Host mapping is the process of controlling which hosts have access to specific volumes within the cluster.

Host mapping is similar in concept to logical unit number (LUN) mapping or masking. LUN mapping is the process of controlling which hosts have access to specific logical units (LUs) within the disk controllers. LUN mapping is typically done at the storage system level. Host mapping is done at the SAN Volume Controller level.

The act of mapping a volume to a host makes the volume accessible to the WWPNs or iSCSI names such as iSCSI qualified names (IQNs) or extended-unique identifiers (EUIs) that are configured in the host object.

Volumes and host mappings

Each host mapping associates a volume with a host object and provides a way for all WWPNs and iSCSI names in the host object to access the volume. You can map a volume to multiple host objects. When a mapping is created, multiple paths might exist across the SAN fabric or Ethernet network from the hosts to the nodes that are presenting the volume. Without a multipathing device driver, most operating systems present each path to a volume as a separate storage device. The multipathing software manages the many paths that are available to the volume and presents a single storage device to the operating system. If there are multiple paths, the SAN Volume Controller requires that the multipathing software run on the host.

Note: The iSCSI names and associated IP addresses for the SAN Volume Controller nodes can fail over between nodes in the I/O group, which negates the need for multipathing drivers in some configurations. Multipathing drivers are still recommended, however, to provide the highest availability.

When you map a volume to a host, you can optionally specify a SCSI ID for the volume. This ID controls the sequence in which the volumes are presented to the host. Check the host software requirements for SCSI IDs because some require a contiguous set. For example, if you present three volumes to the host, and those volumes have SCSI IDs of 0, 1, and 3, the volume that has an ID of 3 might not be found because no disk is mapped with an ID of 2. The cluster automatically assigns the lowest available SCSI ID if none is specified.

Figure 12 on page 39 and Figure 13 on page 39 show two volumes, and the mappings that exist between the host objects and these volumes.
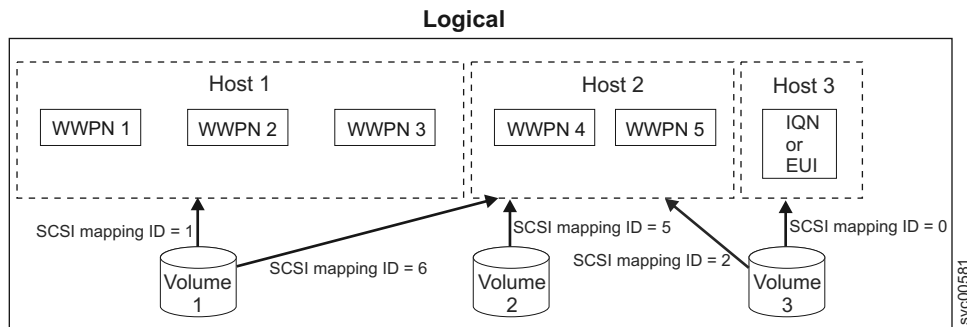
t *Figure 12. Hosts, WWPNs, IQNs or EUIs, and volumes*
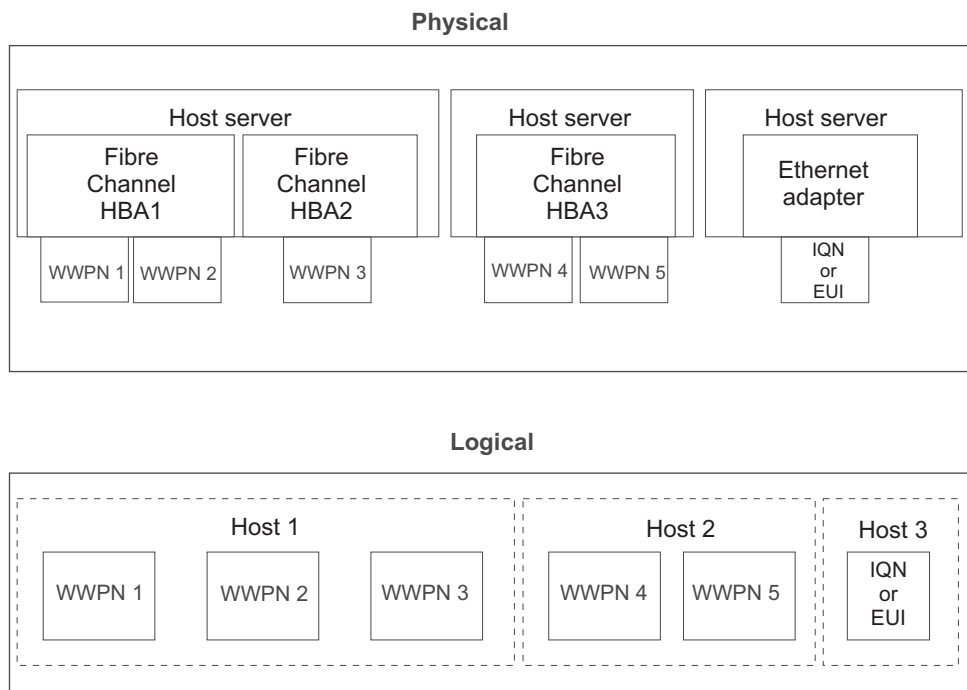
t *Figure 13. Hosts, WWPNs, IQNs or EUIs, volumes, and SCSI mappings*

LUN masking is usually implemented in the device driver software on each host. The host has visibility of more LUNs than it is intended to use, and device driver software masks the LUNs that are not to be used by this host. After the masking is complete, only some disks are visible to the operating system. The

t SAN Volume Controller can support this type of configuration by mapping all volumes to every host object and by using operating system-specific LUN masking technology. The default, and recommended,

t SAN Volume Controller behavior, however, is to map to the host only those volumes that the host requires access to.

Standard and persistent reserves

The SCSI Reserve command and the SCSI Persistent Reserve command are specified by the SCSI standards. Servers can use these commands to prevent ports in other servers from accessing the LUN.

This prevents accidental data corruption that is caused when a server overwrites data on another server. The Reserve and Persistent Reserve commands are often used by clustering software to control access to

t SAN Volume Controller volumes.

3 If a server is not shut down or removed from the server cluster in a controlled way, the server's standard
3 and persistent reserves are maintained. This prevents other servers from accessing data that is no longer
3 in use by the server that holds the reservation. In this situation, you might want to release the reservation
t and allow a new server to access the volume.

t When possible, you should have the server that holds the reservation explicitly release the reservation to ensure that the server cache is flushed and the server software is aware that access to the volume has been lost. In circumstances where this is not possible, you can use operating system specific tools to remove reservations. Consult the operating system documentation for details.

3 When you use the **svctask rmvdiskhostmap** CLI command or the management GUI to remove host
3 mappings, SAN Volume Controller nodes with a software level of 4.1.0 or later can remove the server's
t standard reservations and persistent reservations that the host has on the volume.

Maximum configurations

Ensure that you are familiar with the maximum configurations of the SAN Volume Controller.

See the following Web site for the latest maximum configuration support:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Cluster high availability

A SAN Volume Controller cluster has several features that can be used to deploy a high availability storage system with no single point of failure.

Each I/O group within a cluster consists of a pair of nodes. If a node fails within an I/O group, the other node in the I/O group assumes the I/O responsibilities of the failed node. If the node contains solid-state
3 drives (SSDs), create a mirrored volume or use the RAID functionality to provide redundancy. SSDs can be a single point of failure in the event of an outage to the SSDs or to the node itself.

If a cluster of SAN Volume Controller nodes is split into two partitions (for example due to a SAN fabric fault), the partition with the majority of nodes continues to process I/O operations. If a cluster is split into two equal-sized partitions, a quorum disk is accessed to determine which half of the cluster continues to read and write data.

Each SAN Volume Controller node has four fibre-channel ports, which can be used to attach the node to multiple SAN fabrics. For high availability, attach the nodes in a cluster to at least two fabrics. SAN Volume Controller software incorporates multipathing software that is used for communication among SAN Volume Controller nodes and for I/O operations among SAN Volume Controller nodes and storage systems. If a SAN fabric fault disrupts communication or I/O operations, the multipathing software recovers and retries the operation through an alternative communication path. Also for high availability, configure your fibre-channel host systems to use multipathing software. If a SAN fabric fault or node failure occurs, I/O operations among fibre-channel host systems and SAN Volume Controller nodes are retried. Subsystem device driver (SDD) multipathing software is available from IBM at no additional charge for use with SAN Volume Controller. For additional information about subsystem device driver (SDD), go to the Support for IBM Systems Web site:

www.ibm.com/systems/support

iSCSI-attached hosts connect to SAN Volume Controller through node Ethernet ports. If a node fails, SAN
3 Volume Controller fails over the IP addresses to the partner node in the I/O group to maintain access to
3 the volumes.

t The SAN Volume Controller Volume Mirroring feature can be used to mirror data across storage systems. This feature provides protection against a storage system failure.

The SAN Volume Controller Metro Mirror and Global Mirror features can be used to mirror data between clusters at different physical locations for disaster recovery.

Node management and support tools

The SAN Volume Controller solution offers several management and support tools for you to maintain and manage your nodes.

IBM System Storage Productivity Center

The IBM System Storage Productivity Center (SSPC) is an integrated hardware and software solution that provides a single point of entry for managing SAN Volume Controller clusters, IBM System Storage DS8000 systems, and other components of your data storage infrastructure.

SSPC simplifies storage management in the following ways:
- Centralizing the management of storage network resources with IBM storage management software
- Providing greater synergy between storage management software and IBM storage devices
- Reducing the number of servers that are required to manage your software infrastructure
- Providing simple migration from basic device management to storage management applications that provide higher-level functions

SSPC includes the following software components:
3 - PuTTY (SSH client software)
3 - IBM Tivoli Storage Productivity Center Basic Edition, which can be used to access the IBM System
3   Storage DS8000 Storage Manager and the SAN Volume Controller
3 - IBM DB2® Enterprise Server Edition

Figure 14 on page 42 shows an overview of how SSPC and the components of IBM Tivoli Storage Productivity Center, IBM System Storage DS8000, and SAN Volume Controller interrelate with each other.
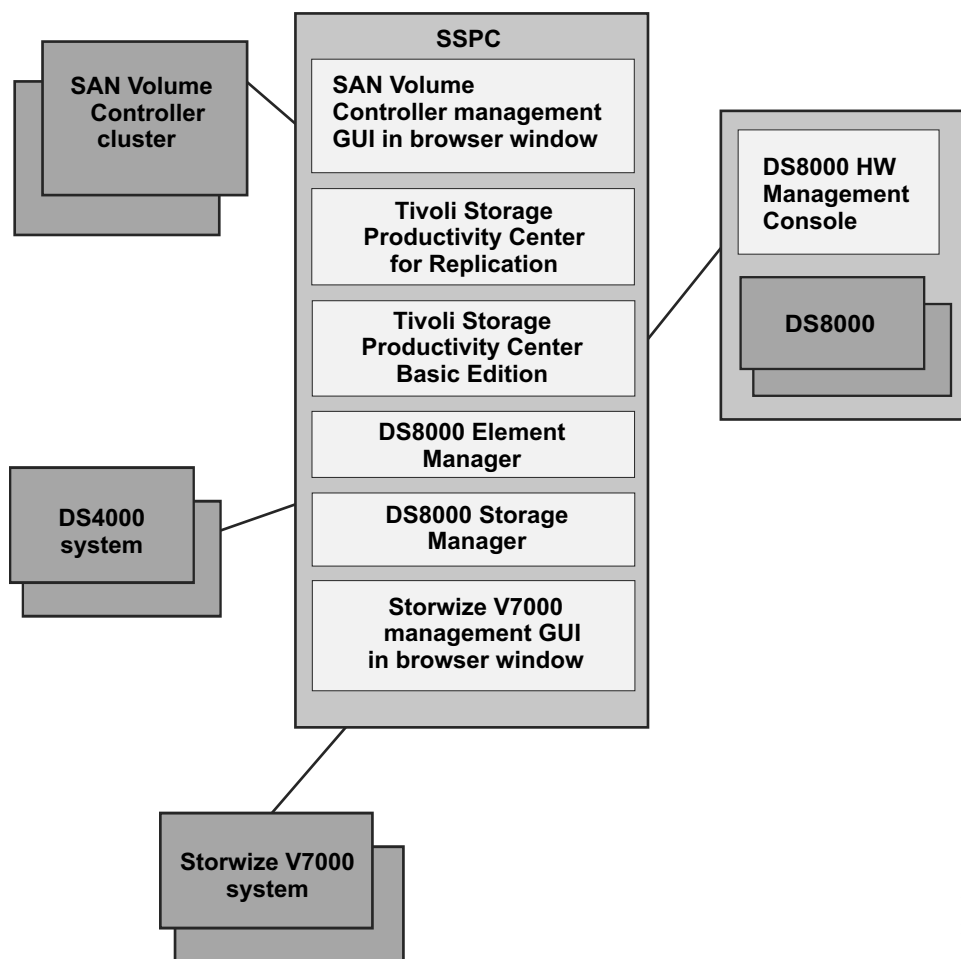
Figure 14. Overview of the IBM System Storage Productivity Center

For more information on SSPC, see the *IBM System Storage Productivity Center Introduction and Planning Guide*.

Assist On-site and remote service

When you contact IBM to help you resolve a problem with your SAN Volume Controller environment, the IBM service representative might suggest using the IBM Assist On-site tool to remotely access the IBM System Storage Productivity Center (SSPC) or master console. This type of remote service can help you reduce service costs and shorten repair times.

The IBM Assist On-site tool is a remote desktop-sharing solution that is offered through the IBM Web site. With it, the IBM service representative can remotely view your system to troubleshoot a problem. You can maintain a chat session with the IBM service representative so that you can monitor the activity and either understand how to fix the problem yourself or allow the representative to fix it for you.

To use the IBM Assist On-site tool, the SSPC or master console must be able to access the Internet. The following Web site provides further information about this tool:

www.ibm.com/support/assistonsite/

When you access the Web site, you sign in and enter a code that the IBM service representative provides to you. This code is unique to each IBM Assist On-site session. A plug-in is downloaded onto your SSPC or master console to connect you and your IBM service representative to the remote service session. The

IBM Assist On-site contains several layers of security to protect your applications and your computers. You can also use security features to restrict access by the IBM service representative.

Your IBM service representative can provide you with more detailed instructions for using the tool.

Event notifications

SAN Volume Controller can use Simple Network Management Protocol (SNMP) traps, syslog messages, and Call Home email to notify you and the IBM Support Center when significant events are detected. Any combination of these notification methods can be used simultaneously.

Each event that SAN Volume Controller detects is assigned a notification type of either Error, Warning, or Information. You can configure SAN Volume Controller to send each type of notification to specific recipients.

Table 14 describes the types of event notifications.

Table 14. Notification types

Notification type	Description
Error	An error notification is sent to indicate a problem that must be corrected as soon as possible. This notification indicates a serious problem with the SAN Volume Controller. For example, the event that is being reported could indicate a loss of redundancy in the system, and it is possible that another failure could result in loss of access to data. The most typical reason that this type of notification is sent is because of a hardware failure, but some configuration errors or fabric errors also are included in this notification type. Error notifications can be configured to be sent as a Call Home email to the IBM Support Center.
Warning	A warning notification is sent to indicate a problem or unexpected condition with the SAN Volume Controller. Always immediately investigate this type of notification to determine the effect that it might have on your operation, and make any necessary corrections. A warning notification does not require any replacement parts and therefore should not require IBM Support Center involvement. However, the event being reported might indicate a condition that could be fatal to your operating environment: for example, a critical FlashCopy operation has failed.
Information	An informational notification is sent to indicate that an expected event has occurred: for example, a FlashCopy operation has completed. No remedial action is required when these notifications are sent.

SNMP traps

SNMP is a standard protocol for managing networks and exchanging messages. The system can send SNMP messages that notify personnel about an event. You can use an SNMP manager to view the SNMP messages that SAN Volume Controller sends. You can use an SNMP manager to view the SNMP messages that the system sends. You can use the management GUI or the SAN Volume Controller command-line interface to configure and modify your SNMP settings.

You can use the Management Information Base (MIB) file for SNMP to configure a network management program to receive SNMP messages that are sent by the system. This file can be used with SNMP messages from all versions of the software. More information about the MIB file for SNMP is available at this website: Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145.

Search for **SAN Volume Controller MIB**. Go to the downloads results to find **Management Information Base (MIB) file for SNMP**. Click this link to find download options. The name of this file is

`SVC_MIB_<release>.MIB` such as in `SVC_MIB_4.3.1.MIB`.

Syslog messages

The syslog protocol is a standard protocol for forwarding log messages from a sender to a receiver on an IP network. The IP network can be either IPv4 or IPv6. The system can send syslog messages that notify personnel about an event. The system can transmit syslog messages in either expanded or concise format. You can use a syslog manager to view the syslog messages that the system sends. The system uses the User Datagram Protocol (UDP) to transmit the syslog message. You can use the management GUI or the SAN Volume Controller command-line interface to configure and modify your syslog settings.

Table 15 shows how SAN Volume Controller notification codes map to syslog security-level codes.

Table 15. SAN Volume Controller notification types and corresponding syslog level codes

SAN Volume Controller notification type	Syslog level code	Description
ERROR	LOG_ALERT	Fault that might require hardware replacement that needs immediate attention.
WARNING	LOG_ERROR	Fault that needs immediate attention. Hardware replacement is not expected.
INFORMATIONAL	LOG_INFO	Information message used, for example, when a configuration change takes place or an operation completes.
TEST	LOG_DEBUG	Test message

Table 16 shows how SAN Volume Controller values of user-defined message origin identifiers map to syslog facility codes.

Table 16. SAN Volume Controller values of user-defined message origin identifiers and syslog facility codes

SAN Volume Controller value	Syslog value	Syslog facility code	Message format
0	16	LOG_LOCAL0	Full
1	17	LOG_LOCAL1	Full
2	18	LOG_LOCAL2	Full
3	19	LOG_LOCAL3	Full
4	20	LOG_LOCAL4	Concise
5	21	LOG_LOCAL5	Concise
6	22	LOG_LOCAL6	Concise
7	23	LOG_LOCAL7	Concise

Call Home email

The Call Home feature transmits operational and event-related data to you and IBM through a Simple Mail Transfer Protocol (SMTP) server connection in the form of an event notification email. When configured, this function alerts IBM service personnel about hardware failures and potentially serious configuration or environmental issues.

To send email, you must configure at least one SMTP server. You can specify as many as five additional SMTP servers for backup purposes. The SMTP server must accept the relaying of email from the SAN

Volume Controller cluster IP address. You can then use the management GUI or the SAN Volume Controller command-line interface to configure the email settings, including contact information and email recipients. Set the reply address to a valid email address. Send a test email to check that all connections and infrastructure are set up correctly. You can disable the Call Home function at any time using the management GUI or the SAN Volume Controller command-line interface.

Data that is sent with notifications

Notifications can be sent using email, SNMP or syslog. The data sent for each type of notification is the same. It includes:

- Record type
- Machine type
- Machine serial number
- Error ID
- Error code
- Software version
- FRU part number
- Cluster name
- Node ID
- Error sequence number
- Time stamp
- Object type
- Object ID
- Problem data

Emails contain the following additional information that allow the Support Center to contact you:

- Contact names for first and second contacts
- Contact phone numbers for first and second contacts
- Alternate contact numbers for first and second contacts
- Offshift phone number
- Contact email address
- Machine location

To send data and notifications to IBM service personnel, use one of the following email addresses:

- For SAN Volume Controller nodes located in North America, Latin America, South America or the Caribbean Islands, use `callhome1@de.ibm.com`
- For SAN Volume Controller nodes located anywhere else in the world, use `callhome0@de.ibm.com`

Inventory information email

Inventory information email is a type of Call Home notification. Inventory information can be sent to IBM to assist IBM service personnel in evaluating your SAN Volume Controller system.

Because inventory information is sent using the Call Home email function, you must meet the Call Home function requirements and enable the Call Home email function before you can attempt to send inventory information email. You can adjust the contact information, adjust the frequency of inventory email, or manually send an inventory email using the management GUI or the SAN Volume Controller command-line interface. Inventory information is automatically reported to IBM when you activate error reporting.

Inventory information that is sent to IBM includes the following information about the cluster on which the Call Home function is enabled:

- The output from the **svcinfo lscluster** command
3 - The output from the **sainfo lsservicenodes** command.
- The output from the **svcinfo lslicense** command
3 - Licensing information
3 - Details about the following objects and functions:
3   Drives
3   External storage systems
3   Hosts
3   MDisks
3   Volumes
3   RAID types
3   Easy Tier
3   FlashCopy
3   Metro Mirror and Global Mirror

User roles

Each user of the management GUI must provide a user name and a password to sign on. Each user also has an associated role such as monitor, copy operator, service, administrator, or security administrator. These roles are defined at the cluster level. For example, a user can perform the administrator role for one cluster and perform the service role for another cluster.

Monitor
> Users with the monitor role have access to all viewing actions available with the management GUI. This user cannot perform any actions that change the state of the cluster or the resources that the cluster manages. The user can access all the information-related panels and commands, back up configuration data, change his or her password, and issue the following commands: **finderr**, **dumperrlog**, **dumpinternallog**, **ping**, and **chcurrentuser**.

1 **Copy Operator**
> Users with the copy operator role can manage all existing FlashCopy, Metro Mirror, and Global Mirror relationships. They can also create and delete FlashCopy mappings, FlashCopy consistency groups, Metro Mirror or Global Mirror relationships, and Metro Mirror and Global Mirror consistency groups. In addition, the user can access all the functions available to the Monitor role.

Service
> Users with the service role can view the View Clusters panel, launch the management GUI, and view the progress of actions on clusters with the View Progress panel, begin disk discovery process, and discover and include disks. The user can access the following commands: **applysoftware**, **setlocale**, **addnode**, **rmnode**, **cherrstate**, **setevent**, **writesernum**, **detectmdisk**, and **includemdisk**. A user with this role can also access all the functions available to the Monitor role.

Administrator
> Users with the administrator role can access all functions on the management GUI and issue any command-line interface (CLI) command, except those that deal with managing users, user groups, and authentication.

1 **Security Administrator (SecurityAdmin role name)**
> Users with the security administrator role can access all functions on the management GUI and issue any CLI command. Users with this role can also manage users, user groups, and manage user authentication.

Configuring user authentication

You can configure authentication and authorization for users of the SAN Volume Controller cluster.

You can create two types of users who access the cluster. These types are based on how the users are authenticated to the cluster. Local users must provide either a password, a Secure Shell (SSH) key, or both. Local users are authenticated through the authentication methods that are located on the SAN Volume Controller cluster. If the local user needs access to management GUI, a password is needed for the user. If the user requires access to the command-line interface, a valid SSH key file is necessary. If a user is working with both interfaces, both a password and SSH key are required. Local users must be part of a user group that is defined on the cluster. User groups define roles that authorize the users within that group to a specific set of operations on the cluster.

A remote user is authenticated on a remote service usually provided by a SAN management application, such as IBM Tivoli Storage Productivity Center. Remote users require no local credentials to access the management GUI. Remote users have their groups defined by the remote authentication service. If a remote user needs to use the command-line interface, both a password and SSH key are required. If the remote authentication service fails, then remote users cannot access the management GUI or the command-line interface. In this situation, a local user with the Security Administrator role must change remote users to local users by adding them to the appropriate user group. After logging in to a SAN Volume Controller application, a remote user is granted access to the SAN Volume Controller CLI and Console by default.

To manage user authentication in the management GUI, select **User Management** → **Users**.

Drives

A drive is a data storage device that can be either a hard disk drive or a solid-state drive (SSD). The SAN Volume Controller supports an internal solid-state drive (SSD) that uses flash memory. SAN Volume Controller also supports magnetic disk drives through external storage systems.

A drive object represents a physical drive. The SAN Volume Controller system creates the objects automatically when a supported SSD in a node is detected.

When the SAN Volume Controller detects a new drive, a drive ID is assigned.

The system supports a drive block size of 512 bytes. Drives with different block sizes cannot be used in RAID arrays or as hot spare drives. You can format these drives to the correct block size. Formatting the drive erases the data that was previously written to the drive and causes the drive to be offline temporarily.

Solid-state drives (SSDs) in certain uses can be recovered and brought back online without significant loss of data.

Chapter 2. Copy Services features

t The SAN Volume Controller provides Copy Services features that enable you to copy volumes..

The following Copy Services features are available for all supported hosts that are connected to SAN Volume Controller:

FlashCopy
t Makes an instant, point-in-time copy from a source volume to a target volume.

Metro Mirror
t Provides a consistent copy of a source volume on a target volume. Data is written to the target
t volume synchronously after it is written to the source volume, so that the copy is continuously updated.

Global Mirror
t Provides a consistent copy of a source volume on a target volume. Data is written to the target
t volume asynchronously, so that the copy is continuously updated, but the copy might not contain the last few updates in the event that a disaster recovery operation is performed.

FlashCopy

e FlashCopy is a Copy Services feature that is available with the SAN Volume Controller system.

t In its basic mode, the IBM FlashCopy feature copies the contents of a source volume to a target volume.
t Any data that existed on the target volume is lost and is replaced by the copied data. After the copy
t operation has completed, the target volumes contain the contents of the source volumes as they existed at a single point in time unless target writes have been performed. The FlashCopy feature is sometimes described as an instance of a time-zero copy (T 0) or point-in-time copy technology. Although the
t FlashCopy operation takes some time to complete, the resulting data on the target volume is presented so that the copy appears to have occurred immediately.

Although it is difficult to make a consistent copy of a data set that is constantly updated, point-in-time copy techniques help solve this problem. If a copy of a data set is created using a technology that does not provide point-in-time techniques and the data set changes during the copy operation, the resulting copy might contain data that is not consistent. For example, if a reference to an object is copied earlier than the object itself and the object is moved before it is copied, the copy contains the referenced object at its new location, but the copied reference still points to the previous location.

t More advanced FlashCopy features allow operations to occur on multiple source and target volumes. FlashCopy management operations are coordinated to provide a common, single point in time for
t copying target volumes from their respective source volumes. This creates a consistent copy of data that
t spans multiple volumes. The FlashCopy feature also allows multiple target volumes to be copied from
t each source volume. This can be used to create images from different points in time for each source
t volume.

FlashCopy applications

You can use the FlashCopy feature to create consistent backups of dynamic data, test applications, and create copies for auditing purposes and for data mining.

To create consistent backups of dynamic data, use the FlashCopy feature to capture the data at a particular time. The resulting image of the data can be backed up, for example, to a tape device. When

the copied data is on tape, the data on the FlashCopy target disks become redundant and can now be discarded. Usually in this backup condition, the target data can be managed as read-only.

It is often very important to test a new version of an application with real business data before the existing production version of the application is updated or replaced. This testing reduces the risk that the updated application fails because it is not compatible with the actual business data that is in use at the time of the update. Such an application test might require write access to the target data.

You can also use the FlashCopy feature to create restart points for long running batch jobs. This means that if a batch job fails several days into its run, it might be possible to restart the job from a saved copy of its data rather than rerunning the entire multiday job.

Host considerations for FlashCopy integrity

The SAN Volume Controller FlashCopy feature transfers a point-in-time copy of a source volume onto a designated target volume. You must create or already have an existing target volume before you can transfer the copy. You must also ensure that the target volume has enough space available to support the amount of data that is being transferred.

After the mapping is started, all of the data that is stored on the source volume can be accessed through the target volume. This includes any operating system control information, application data, and metadata that was stored on the source volume. Because of this, some operating systems do not allow a source volume and a target volume to be addressable on the same host.

To ensure the integrity of the copy that is made, it is necessary to completely flush the host cache of any outstanding reads or writes before you proceed with the FlashCopy operation. You can flush the host cache by unmounting the source volumes from the source host before you start the FlashCopy operation.

Because the target volumes are overwritten with a complete image of the source volumes, it is important that any data held on the host operating system (or application) caches for the target volumes is discarded before the FlashCopy mappings are started. The easiest way to ensure that no data is held in these caches is to unmount the target volumes before starting the FlashCopy operation.

Some operating systems and applications provide facilities to stop I/O operations and ensure that all data is flushed from caches on the host. If these facilities are available, they can be used to prepare and start a FlashCopy operation. See your host and application documentation for details.

Some operating systems are unable to use a copy of a volume without *synthesis*. Synthesis performs a transformation of the operating system metadata on the target volume to enable the operating system to use the disk. See your host documentation on how to detect and mount the copied volumes.

Flushing data from the host volumes

All outstanding read and write operations must be flushed from the host cache before you use the FlashCopy feature.

Perform the following steps to flush data from your host volumes and start a FlashCopy operation:

1. If you are using UNIX or Linux operating systems, perform the following steps:
   a. Quiesce all applications to the source volumes that you want to copy.
   b. Use the **umount** command to unmount the designated drives.
   c. Prepare and start the FlashCopy operation for those unmounted drives.
   d. Remount your volumes with the mount command and resume your applications.
2. If you are using the Microsoft Windows operating system using drive letter changes, perform the following steps:
   a. Quiesce all applications to the source volumes that you want to copy.

b. Go into your disk management window and remove the drive letter on each drive that you want to copy. This unmounts the drive.

c. Prepare and start the FlashCopy operation for those unmounted drives.

d. Remount your volumes by restoring the drive letters and resume your applications.

If you are using the **chkdsk** command, perform the following steps:

a. Quiesce all applications to the source volumes that you want to copy.

b. Issue the **chkdsk /x** command on each drive you want to copy. The /x option unmounts, scans, and remounts the volume.

c. Ensure that all applications to the source volumes are still quiesced.

d. Prepare and start the FlashCopy operation for those unmounted drives.

Note: If you can ensure that no reads and writes are issued to the source volumes after you unmount the drives, you can immediately remount and then start the FlashCopy operation.

FlashCopy mappings

A FlashCopy mapping defines the relationship between a source volume and a target volume.

The FlashCopy feature makes an instant copy of a volume at the time that it is started. To create an instant copy of a volume, you must first create a mapping between the source volume (the disk that is copied) and the target volume (the disk that receives the copy). The source and target volumes must be of equal size.

A mapping can be created between any two volumes in a cluster. The volumes do not have to be in the same I/O group or storage pool. When a FlashCopy operation starts, a checkpoint is made of the source volume. No data is actually copied at the time a start operation occurs. Instead, the checkpoint creates a bitmap that indicates that no part of the source volume has been copied. Each bit in the bitmap represents one region of the source volume. Each region is called a *grain*.

After a FlashCopy operation starts, read operations to the source volume continue to occur. If new data is written to the source or target volume, the existing data on the source is copied to the target volume before the new data is written to the source or target volume. The bitmap is updated to mark that the grain of the source volume has been copied so that later write operations to the same grain do not recopy the data.

During a read operation to the target volume, the bitmap is used to determine if the grain has been copied. If the grain has been copied, the data is read from the target volume. If the grain has not been copied, the data is read from the source volume.

Incremental FlashCopy mappings

The incremental FlashCopy feature reduces the amount of time that is required to copy the source volume to a maximum of four FlashCopy mappings. (If a volume has more than four incremental mappings, all of the mappings cease to be incremental and become full FlashCopy mappings.)

In an incremental FlashCopy, the initial mapping copies all of the data from the source volume to the target volume. Subsequent FlashCopy mappings only copy data that has been modified since the initial FlashCopy mapping. This reduces the amount of time that it takes to re-create an independent FlashCopy image. You can define a FlashCopy mapping as incremental only when you create the FlashCopy mapping.

Tracks changes
in a difference map

Source   Target
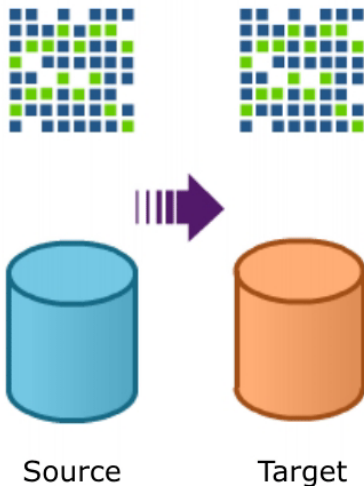
Figure 15. Incremental FlashCopy of differences

Use incremental FlashCopy only at the beginning of a cascaded FlashCopy mapping. An incremental FlashCopy mapping from Disk 0 to Disk 1 in Figure 16 can support full cascaded FlashCopy mappings of Disk 1:

Disk 0
Source

Disk 1
(FlashCopy
of Disk 0)

Disk 2
(FlashCopy
of Disk 1)

Disk 3
(FlashCopy
of Disk 1)
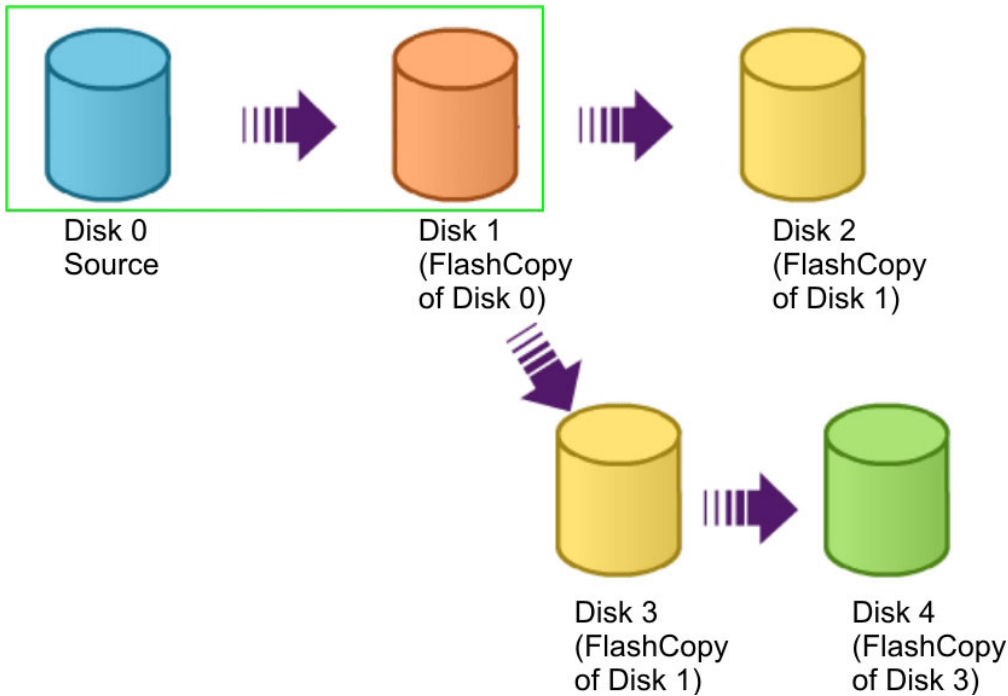
Disk 4
(FlashCopy
of Disk 3)

Figure 16. Incremental FlashCopy is effective in the mapping in the highlight box

A full FlashCopy mapping at the beginning of a set of cascaded FlashCopy mappings causes all of the mappings to be full copies.

A full FlashCopy mapping from Disk 0 to Disk 1 in Figure 17 on page 53, for example, causes incremental FlashCopy mappings from Disk 1 to Disk 2, Disk 1 to Disk 3, and Disk 3 to Disk 4 to be full copies instead. After each full FlashCopy from Disk 0 to Disk 1 destroys the differences table in Disk 1, the other

3 mappings do not have a differences table to support incremental copies. All of the following mappings
3 would be treated as full FlashCopy mappings:
3
3
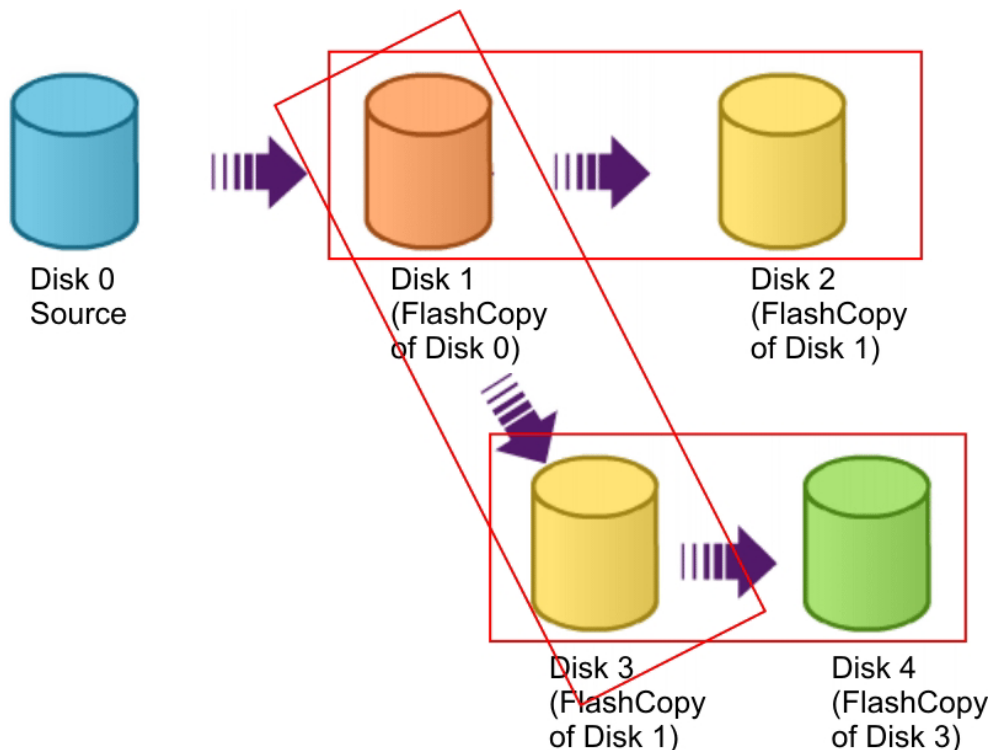
3
3 *Figure 17. Incremental FlashCopy is ineffective in the mappings inside the highlight boxes*
3
3 **Note:** For incremental FlashCopy support information, see the release-specific *IBM System Storage SAN*
3 *Volume Controller Restrictions* technical note at this website:

3 Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

FlashCopy partner mappings

You can create a mapping to mirror an existing incremental FlashCopy mapping. The resulting pair of
t mappings are called *partners*. A mapping can have only one partner. For example, if you have volume A
t and volume B with two mappings (Mapping 0 from volume A to volume B and Mapping 1 from volume
t B to volume A), Mapping 0 and Mapping 1 are partners.

Incremental FlashCopy mappings share the metadata for recording changes. Therefore, if one mapping in
a mirrored pair (partnership) is incremental, the other mapping becomes incremental automatically and
remains incremental until it is deleted.

Cascaded FlashCopy mappings

t The Cascaded FlashCopy feature allows a FlashCopy target volume to be the source volume of another
FlashCopy mapping.

2

Disk 0
Source

Disk 1
(FlashCopy
of Disk 0)

Disk 2
(FlashCopy
of Disk 1)

Disk 3
(FlashCopy
of Disk 1)

Disk 4
(FlashCopy
of Disk 3)
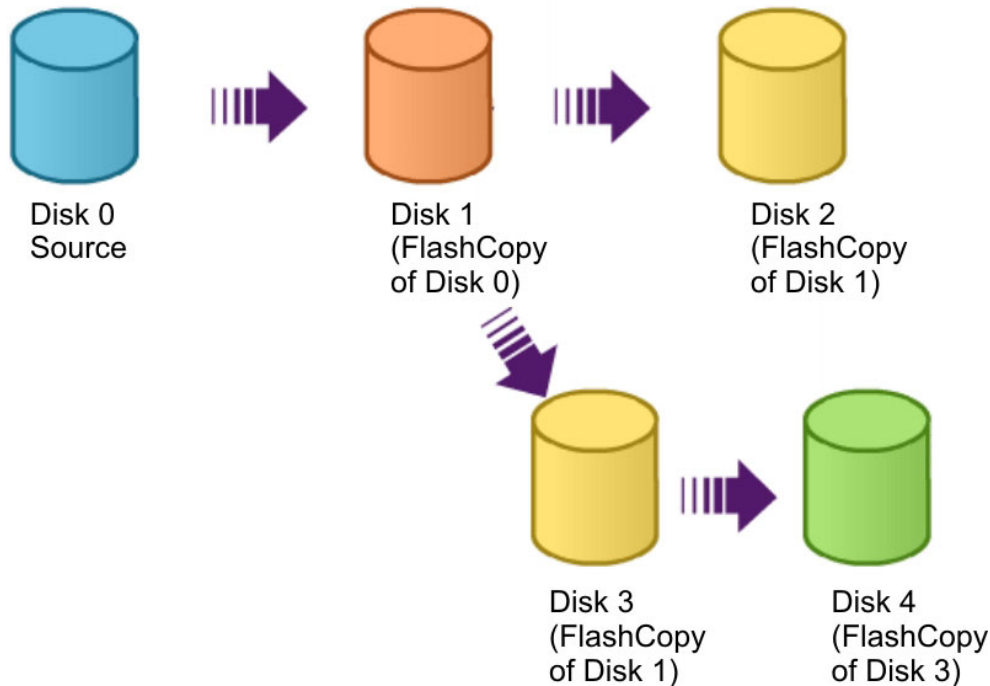
2
2 *Figure 18. Cascading FlashCopy volumes*
2

Up to 256 mappings can exist in a cascade. If cascaded mappings and multiple target mappings are used, a tree of up to 256 mappings can be created.

Multiple target FlashCopy mappings

t You can copy up to 256 target volumes from a single source volume. Each relationship between a source
t and target volume is managed by a unique mapping such that a single volume can be the source volume in up to 256 mappings.

Each of the mappings from a single source can be started and stopped independently. If multiple mappings from the same source are active (in the copying or stopping states), a dependency exists between these mappings.

Example 1

Mapping A depends on mapping B if the following is true:
t • Mapping A and mapping B both have the same source volume
  • Mapping A and mapping B are both in the copying or stopping state
  • Mapping B was started more recently than mapping A

  Note: If both mappings were in the same consistency group and therefore started at the same time, the order of dependency is decided internally when the consistency group is started.

  • Mapping A does not have a complete copy of the source because the copying progress for the mapping is less than 100.
  • A mapping does not exist from the same source started more recently than A and later than B which has a complete copy of the source because the copying progress of the mapping is less than 100.

Example 2

t Target volume A depends on target volume B if the mapping that volume A belongs to depends on the
t mapping that target volume B belongs to. The target volume of the most recently started mapping from
t the source volume depends on the source volume until there is a complete copy of the source (progress is
   100%).

Clean rate, copy rate, and autodelete

   When you create a mapping, you specify a clean rate. The clean rate is used to control the rate that data
t is copied from the target volume of the mapping to the target volume of a mapping that is either the
t latest copy of the target volume, or is the next oldest copy of the source volume. The clean rate is used in
   the following situations:

   • The mapping is in the stopping state

   • The mapping is in the copying state and has a copy rate of zero

   • The mapping is in the copying state and the background copy has completed

   You can use the clean rate to minimize the amount of time that a mapping is in the stopping state. If the
t mapping has not completed, the target volume is offline while the mapping is stopping. The target
t volume remains offline until the mapping is restarted.

3 You also specify a copy rate when you create a mapping. When the mapping is in the copying state, the
3 copy rate determines the priority that is given to the background copy process. If you want a copy of the
t whole source volume so that a mapping can be deleted and still be accessed from the target volume, you
t must copy all the data that is on the source volume to the target volume.

   The default values for both the clean rate and the copy rate is 50.

   When a mapping is started and the copy rate is greater than zero, the unchanged data is copied to the
t target volume, and the bitmap is updated to show that the copy has occurred. After a time, the length of
t which depends on the priority that was determined by the copy rate and the size of the volume, the
t whole volume is copied to the target. The mapping returns to the idle_or_copied state and you can now
   restart the mapping at any time to create a new copy at the target.

   While the mapping is in the copying state, you can set the copy rate to zero and the clean rate to a value
   other than zero to minimize the amount of time a mapping is in the stopping state.

   If you use multiple target mappings, the mapping can stay in the copying state after all of the source
   data is copied to the target (the progress is 100%). This situation can occur if mappings that were started
   earlier and use the same source disk are not yet 100% copied.

   If the copy rate is zero, only the data that changes on the source is copied to the target. The target never
   contains a copy of the whole source unless every extent is overwritten at the source. You can use this
   copy rate when you require a temporary copy of the source.

t You can stop the mapping at any time after it has been started. Unless the target volume already contains
t a complete copy of the source volume, this action makes the target inconsistent and the target volume is
t taken offline. The target volume remains offline until the mapping is restarted.

3 You can also set the autodelete attribute. If this attribute is set to on, the mapping is automatically
3 deleted when the mapping reaches the idle_or_copied state and the progress is 100%.

FlashCopy mapping states

   At any point in time, a mapping is in one of the following states:

Idle or copied

The source and target volumes act as independent volumes even if a mapping exists between the two. Read and write caching is enabled for both the source and the target volumes.

If the mapping is incremental and the background copy is complete, the mapping only records the differences between the source and target volumes. If the connection to both nodes in the I/O group that the mapping is assigned to is lost, the source and target volumes will be offline.

Copying

The copy is in progress. Read and write caching is enabled on the source and the target volumes.

Prepared

The mapping is ready to start. The target volume is online, but is not accessible. The target volume cannot perform read or write caching. Read and write caching is failed by the SCSI front end as a hardware error. If the mapping is incremental and a previous mapping has completed, the mapping only records the differences between the source and target volumes. If the connection to both nodes in the I/O group that the mapping is assigned to is lost, the source and target volumes go offline.

Preparing

The target volume is online, but not accessible. The target volume cannot perform read or write caching. Read and write caching is failed by the SCSI front end as a hardware error. Any changed write data for the source volume is flushed from the cache. Any read or write data for the target volume is discarded from the cache. If the mapping is incremental and a previous mapping has completed, the mapping records only the differences between the source and target volumes. If the connection to both nodes in the I/O group that the mapping is assigned to is lost, the source and target volumes go offline.

Stopped

The mapping is stopped because either you issued a stop command or an I/O error occurred. The target volume is offline and its data is lost. To access the target volume, you must restart or delete the mapping. The source volume is accessible and the read and write cache is enabled. If the mapping is incremental, the mapping is recording write operations to the source volume. If the connection to both nodes in the I/O group that the mapping is assigned to is lost, the source and target volumes go offline.

Stopping

The mapping is in the process of copying data to another mapping.

- If the background copy process is complete, the target volume is online while the stopping copy process completes.
- If the background copy process is not complete, data is discarded from the target volume cache. The target volume is offline while the stopping copy process runs.

The source volume is accessible for I/O operations.

Suspended

The mapping started, but it did not complete. Access to the metadata is lost, which causes both the source and target volume to go offline. When access to the metadata is restored, the mapping returns to the copying or stopping state and the source and target volumes return online. The background copy process resumes. Any data that has not been flushed and has been written to the source or target volume before the suspension, is in cache until the mapping leaves the suspended state.

Notes:

1. If a FlashCopy source volume goes offline, any FlashCopy target volumes that depend on that volume also go offline.

2. If a FlashCopy target volume goes offline, any FlashCopy target volumes that depend on that volume also go offline. The source volume remains online.

Before you start the mapping, you must prepare it. Preparing the mapping ensures that the data in the cache is de-staged to disk and a consistent copy of the source exists on disk. At this time, the cache goes into write-through mode. Data that is written to the source is not cached in the SAN Volume Controller nodes; it passes straight through to the MDisks. The prepare operation for the mapping might take some time to complete; the actual length of time depends on the size of the source volume. You must coordinate the prepare operation with the operating system. Depending on the type of data that is on the source volume, the operating system or application software might also cache data write operations. You must flush, or synchronize, the file system and application program before you prepare and start the mapping.

Note: The **svctask startfcmap** and **svctask startfcconsistgrp** commands can take some time to process.

If you do not want to use consistency groups, the SAN Volume Controller allows a mapping to be treated as an independent entity. In this case, the mapping is known as a stand-alone mapping. For mappings that have been configured in this way, use the **svctask prestartfcmap** and **svctask startfcmap** commands rather than the **svctask prestartfcconsistgrp** and **svctask startfcconsistgrp** commands.

FlashCopy mapping restoration

You can start a mapping with a target volume that is the source volume of another active mapping that is in either the idle_copied, stopped, or copying states. If the mapping is in the copying state, the **restore** parameter is required for the **svctask startfcmap** and **svctask prestartfcmap** commands. You can restore the contents of a FlashCopy source volume by using the target of the same FlashCopy mapping or a different FlashCopy mapping without waiting for the mapping to become idle and without loss of the contents of any other FlashCopy target volume.

Veritas Volume Manager

For FlashCopy target volumes, the SAN Volume Controller sets a bit in the inquiry data for those mapping states where the target volume could be an exact image of the source volume. Setting this bit enables the Veritas Volume Manager to distinguish between the source and target volumes and provide independent access to both.

FlashCopy mapping events
FlashCopy mapping events detail the events that modify the state of a FlashCopy mapping.

Table 17 provides a description of each FlashCopy mapping event.

Table 17. FlashCopy mapping events

Create	A new FlashCopy mapping is created between the specified source volume and the specified target volume. The operation fails if any of the following is true:
	• For SAN Volume Controller software version 4.1.0 or earlier, the source or target volume is already a member of a FlashCopy mapping.
	• For SAN Volume Controller software version 4.2.0 or later, the source or target volume is already a target volume of a FlashCopy mapping.
	• For SAN Volume Controller software version 4.2.0 or later, the source volume is already a member of 16 FlashCopy mappings.
	• For SAN Volume Controller software version 4.3.0 or later, the source volume is already a member of 256 FlashCopy mappings.
	• The node has insufficient bitmap memory.
	• The source and target volumes are different sizes.

Table 17. FlashCopy mapping events (continued)

Prepare	The **prepare** command is directed to either a consistency group for FlashCopy mappings that are members of a normal consistency group or to the mapping name for FlashCopy mappings that are stand-alone mappings. The **prepare** command places the FlashCopy mapping into the preparing state.
	Attention: The **prepare** command can corrupt any data that previously resided on the target volume because cached writes are discarded. Even if the FlashCopy mapping is never started, the data from the target might have logically changed during the act of preparing to start the FlashCopy mapping.
Flush done	The FlashCopy mapping automatically moves from the preparing state to the prepared state after all cached data for the source is flushed and all cached data for the target is no longer valid.
Start	When all the FlashCopy mappings in a consistency group are in the prepared state, the FlashCopy mappings can be started.
	To preserve the cross volume consistency group, the start of all of the FlashCopy mappings in the consistency group must be synchronized correctly with respect to I/Os that are directed at the volumes. This is achieved during the **start** command.
	The following occurs during the **start** command:
	• New reads and writes to all source volumes in the consistency group are paused in the cache layer until all ongoing reads and writes below the cache layer are completed.
	• After all FlashCopy mappings in the consistency group are paused, the internal cluster state is set to allow FlashCopy operations.
	• After the cluster state is set for all FlashCopy mappings in the consistency group, read and write operations are unpaused on the source volumes.
	• The target volumes are brought online.
	As part of the **start** command, read and write caching is enabled for both the source and target volumes.
Modify	The following FlashCopy mapping properties can be modified:
	• FlashCopy mapping name
	• Clean rate
	• Consistency group
	• Copy rate (for background copy or stopping copy priority)
	• Automatic deletion of the mapping when the background copy is complete
Stop	There are two separate mechanisms by which a FlashCopy mapping can be stopped:
	• You have issued a command
	• An I/O error has occurred
Delete	This command requests that the specified FlashCopy mapping is deleted. If the FlashCopy mapping is in the stopped state, the force flag must be used.
Flush failed	If the flush of data from the cache cannot be completed, the FlashCopy mapping enters the stopped state.
Copy complete	After all of the source data has been copied to the target and there are no dependent mappings, the state is set to copied. If the option to automatically delete the mapping after the background copy completes is specified, the FlashCopy mapping is automatically deleted. If this option is not specified, the FlashCopy mapping is not automatically deleted and can be reactivated by preparing and starting again.
Bitmap Online/Offline	The node has failed.

Thin-provisioned FlashCopy

You can have a mix of thin-provisioned and fully allocated volumes in FlashCopy mappings. One common combination is a fully allocated source with a thin-provisioned target, which enables the target to consume a smaller amount of real storage than the source.

For best performance, the grain size of the thin-provisioned volume must match the grain size of the FlashCopy mapping. However, if the grain sizes are different, the mapping still proceeds.

Consider the following information when you create your FlashCopy mappings:

- If you are using a fully allocated source with a thin-provisioned target, disable background copy and cleaning mode on the FlashCopy map by setting both the background copy rate and cleaning rate to zero. Otherwise, if these features are enabled, all the source is copied onto the target volume. This causes the thin-provisioned volume to either go offline or to grow as large as the source.
- If you are using only thin-provisioned source, only the space that is used on the source volume is copied to the target volume. For example, if the source volume has a virtual size of 800 GB and a real size of 100 GB of which 50 GB have been used, only the used 50 GB are copied.
- A FlashCopy bitmap contains one bit for every grain on a volume. For example, if you have a thin-provisioned volume with 1 TB virtual size (100 MB real capacity), you must have a FlashCopy bitmap to cover the 1 TB virtual size even though only 100 MB of real capacity is allocated.

FlashCopy consistency groups

A *consistency group* is a container for mappings. You can add many mappings to a consistency group.

The consistency group is specified when the mapping is created. You can also change the consistency group later. When you use a consistency group, you prepare and start that group instead of the individual mappings. This process ensures that a consistent copy is made of all the source volumes. Mappings to control at an individual level are known as stand-alone mappings. Do not place stand-alone mappings into a consistency group because they become controlled as part of that consistency group.

When you copy data from one volume to another, the data might not include all that you need to use the copy. Many applications have data that spans multiple volumes and requires that data integrity is preserved across volumes. For example, the logs for a particular database usually reside on a different volume than the volume that contains the data.

Consistency groups address the problem of applications having related data that spans multiple volumes. In this situation, IBM FlashCopy operations must be performed in a way that preserves data integrity across the multiple volumes. One requirement for preserving the integrity of data being written is to ensure that dependent writes are run in the intended sequence of the application.

You can set the autodelete attribute for FlashCopy consistency groups. If this attribute is set to on, the consistency group is automatically deleted when the last mapping in the group is deleted or moved out of the consistency group.

Multiple target FlashCopy mappings

Consistency groups aggregate FlashCopy mappings, not the volumes themselves. Therefore, a source volume with multiple FlashCopy mappings can be in different consistency groups. If a volume is the source volume for several FlashCopy mappings that are in the same consistency group, multiple identical copies of the source volume are created when the consistency group is started.

Cascaded FlashCopy mappings

To create a FlashCopy mapping in a consistency group, the source volume cannot be the target of a mapping in the same consistency group. In addition, the target volume cannot be the source of another FlashCopy mapping in the same consistency group. You cannot move a FlashCopy mapping into a

consistency group that contains similar FlashCopy mappings in the cascade.

FlashCopy consistency group states

At any point in time, a FlashCopy consistency group is in one of the following states:

Idle_or_Copied
> All FlashCopy Mappings in this consistency group are in the Idle or Copied state.

Preparing
> At least one FlashCopy mapping in this consistency group is in the Preparing state.

Prepared
t
> The consistency group is ready to start. While in this state, the target volumes of all FlashCopy mappings in this consistency group are not accessible.

Copying
> At least one FlashCopy mapping in the consistency group is in the Copying state and no FlashCopy mappings are in the Suspended state.

Stopping
> At least one FlashCopy mapping in the consistency group is in the Stopping state and no FlashCopy mappings are in the Copying or Suspended state.

Stopped
> The consistency group is stopped because either you issued a command or an I/O error occurred.

Suspended
> At least one FlashCopy mapping in the consistency group is in the Suspended state.

Empty The consistency group does not have any FlashCopy mappings.

Dependent write operations
To preserve the integrity of data that is being written, ensure that dependent writes are run in the intended sequence of the application.

The following list is a typical sequence of write operations for a database update transaction:
1. A write operation updates the database log so that it indicates that a database update is about to take place.
2. A second write operation updates the database.
3. A third write operation updates the database log so that it indicates that the database update has completed successfully.

3 The database ensures correct ordering of these writes by waiting for each step to complete before starting
3 the next. The database log is often placed on a different volume than the database. In this case, ensure
3 that FlashCopy operations are performed without changing the order of these write operations. For
3 example, consider the possibility that the database (update 2) is copied slightly earlier than the database
3 log (update 1 and 3), which means the copy on the target volume will contain updates (1) and (3) but not
3 (2). In this case, if the database is restarted from a backup made from the FlashCopy target disks, the
3 database log indicates that the transaction has completed successfully when, in fact, it has not. The
3 transaction is lost and the integrity of the database is compromised.

t You can perform a FlashCopy operation on multiple volumes as an atomic operation to create a
consistent image of user data. To use FlashCopy this way, the SAN Volume Controller supports the
concept of a consistency group. A consistency group can contain an arbitrary number of FlashCopy
mappings, up to the maximum number of FlashCopy mappings that are supported by a SAN Volume
e Controller cluster. You can use the command-line interface (CLI) **svctask startfcconsistgrp** command to
start the point-in-time copy for the entire consistency group. All of the FlashCopy mappings in the

consistency group are started at the same time, resulting in a point-in-time copy that is consistent across all of the FlashCopy mappings that are contained in the consistency group.

See the following website for the latest maximum configuration support:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Grains and the FlashCopy bitmap

t When data is copied between volumes, it is copied in units of address space known as *grains*.

The grain size is 64 KB or 256 KB. The FlashCopy bitmap contains one bit for each grain. The bit records whether the associated grain has been split by copying the grain from the source to the target.

t **Write to target volume**

t A write to the newest target volume must consider the state of the grain for its own mapping and the grain of the next oldest mapping.
  • If the grain of the intermediate mapping or the next oldest mapping has not been copied, it must be copied before the write is allowed to proceed. This is done to preserve the contents of the next oldest mapping. The data written to the next oldest mapping can come from a target or source.
  • If the grain of the target that is being written has not been copied, the grain is copied from the oldest already copied grain in the mappings that are newer than the target (or the source if no targets are already copied). After the copy is complete, the write can be applied to the target.

t **Read to target volume**

If the grain that is being read has been split, the read returns data from the target that is being read. If
t the read is to an uncopied grain on an intermediate target volume, each of the newer mappings are examined to determine if the grain has been split. The read is surfaced from the first split grain found or
t from the source volume if none of the newer mappings have a split grain.

Background copy and cleaning rates

FlashCopy mapping copy *rate* values can be between 1 and 100 and can be changed when the FlashCopy mapping is in any state.

If NOCOPY is specified, background copy is disabled. You can specify NOCOPY for short-lived FlashCopy mappings that are only used for backups, for example. Because the source data set is not expected to significantly change during the lifetime of the FlashCopy mapping, it is more efficient in terms of managed disk (MDisk) I/Os to not perform a background copy.

Note: For the command-line interface (CLI), the value NOCOPY is equivalent to setting the copy rate to 0 (zero).

Table 18 provides the relationship of the copy and cleaning *rate* values to the attempted number of grains to be split per second. A grain is the unit of data represented by a single bit.

Table 18. Relationship between the rate, data rate and grains per second values

User-specified *rate* attribute value	Data copied/sec	256 KB grains/sec	64 KB grains/sec
1 - 10	128 KB	0.5	2
11 - 20	256 KB	1	4
21 - 30	512 KB	2	8
31 - 40	1 MB	4	16

Table 18. Relationship between the rate, data rate and grains per second values (continued)

User-specified *rate* attribute value	Data copied/sec	256 KB grains/sec	64 KB grains/sec
41 - 50	2 MB	8	32
51 - 60	4 MB	16	64
61 - 70	8 MB	32	128
71 - 80	16 MB	64	256
81 - 90	32 MB	128	512
91 - 100	64 MB	256	1024

The data copied/sec and the grains/sec numbers represent standards that the SAN Volume Controller tries to achieve. The SAN Volume Controller is unable to achieve these standards if insufficient bandwidth is available from the nodes to the physical disks that make up the managed disks (MDisks) after taking into account the requirements of foreground I/O. If this situation occurs, background copy I/O contends for resources on an equal basis with I/O that arrives from hosts. Both tend to see an increase in latency and consequential reduction in throughput with respect to the situation had the bandwidth not been limited.

Degradation runs smoothly. Background copy, stopping copy, and foreground I/O continue to make forward progress and do not stop, hang or cause the node to fail.

The background copy is performed by one of the nodes that belong to the I/O group in which the source volume resides. This responsibility is moved to the other node in the I/O group in the event of the failure of the node that performs the background and stopping copy.

The background copy starts with the grain that contains the highest logical block numbers (LBAs) and works in reverse towards the grain that contains LBA 0. The background copy is performed in reverse to avoid any unwanted interactions with sequential write streams from the application.

The stopping copy operation copies every grain that is split on the stopping map to the next map (if one exists) which is dependent on that grain. The operation starts searching with the grain that contains the highest LBAs and works in reverse towards the grain that contains LBA 0. Only those grains that other maps are dependent upon are copied.

Cleaning mode

When you create or modify a FlashCopy mapping, you can specify a cleaning rate for the FlashCopy mapping that is independent of the background copy rate. The cleaning rates shown in Table 18 on page 61 controls the rate at which the cleaning process operates. The cleaning process copies data from the target volume of a mapping to the target volumes of other mappings that are dependent on this data. The cleaning process must complete before the FlashCopy mapping can go to the stopped state.

Cleaning mode allows you to activate the cleaning process when the FlashCopy mapping is in the copying state. This keeps your target volume accessible while the cleaning process is running. When operating in this mode, it is possible that host I/O operations can prevent the cleaning process from reaching 100% if the I/O operations continue to copy new data to the target volumes. However, it is possible to minimize the amount of data that requires cleaning while the mapping is stopping.

Cleaning mode is active if the background copy progress has reached 100% and the mapping is in the copying state, or if the background copy rate is set to 0.

Metro Mirror and Global Mirror

The Metro Mirror and Global Mirror Copy Services features enable you to set up a relationship between two volumes, so that updates that are made by an application to one volume are mirrored on the other volume. The volumes can be in the same cluster or on two different clusters.

Although the application only writes to a single volume, the system maintains two copies of the data. If the copies are separated by a significant distance, the Metro Mirror and Global Mirror copies can be used as a backup for disaster recovery. A prerequisite for Metro Mirror and Global Mirror operations between clusters is that the SAN fabric to which they are attached provides adequate bandwidth between the clusters.

For both Metro Mirror and Global Mirror copy types, one volume is designated as the primary and the other volume is designated as the secondary. Host applications write data to the primary volume, and updates to the primary volume are copied to the secondary volume. Normally, host applications do not perform I/O operations to the secondary volume.

The Metro Mirror feature provides a *synchronous*-copy process. When a host writes to the primary volume, it does not receive confirmation of I/O completion until the write operation has completed for the copy on both the primary volume and the secondary volume. This ensures that the secondary volume is always up-to-date with the primary volume in the event that a failover operation must be performed. However, the host is limited to the latency and bandwidth limitations of the communication link to the secondary volume.

The Global Mirror feature provides an *asynchronous*-copy process. When a host writes to the primary volume, confirmation of I/O completion is received before the write operation has completed for the copy on the secondary volume. If a failover operation is performed, the application must recover and apply any updates that were not committed to the secondary volume. If I/O operations on the primary volume are paused for a small length of time, the secondary volume can become an exact match of the primary volume.

The Metro Mirror and Global Mirror operations support the following functions:

- Intracluster copying of a volume, in which both volumes belong to the same cluster and I/O group within the cluster.
- Intercluster copying of a volume, in which one volume belongs to a cluster and the other volume belongs to a different cluster.

    Note: A cluster can participate in active Metro Mirror and Global Mirror relationships with itself and up to three other clusters.

- Intercluster and intracluster Metro Mirror and Global Mirror relationships can be used concurrently within a cluster.
- The intercluster link is bidirectional. This means that it can copy data from cluster A to cluster B for one pair of volumes while copying data from cluster B to cluster A for a different pair of volumes.
- The copy direction can be reversed for a consistent relationship.
- Consistency groups are supported to manage a group of relationships that must be kept synchronized for the same application. This also simplifies administration, because a single command that is issued to the consistency group is applied to all the relationships in that group.
- SAN Volume Controller supports a maximum of 8192 Metro Mirror and Global Mirror relationships per cluster.

Metro Mirror and Global Mirror relationships

Metro Mirror and Global Mirror relationships define the relationship between two volumes: a master volume and an auxiliary volume.

t Typically, the master volume contains the production copy of the data and is the volume that the
t application normally accesses. The auxiliary volume typically contains a backup copy of the data and is
used for disaster recovery.

t The master and auxiliary volumes are defined when the relationship is created, and these attributes never
t change. However, either volume can operate in the primary or secondary role as necessary. The primary
t volume contains a valid copy of the application data and receives updates from the host application,
t analogous to a source volume. The secondary volume receives a copy of any updates to the primary
t volume, because these updates are all transmitted across the mirror link. Therefore, the secondary volume
t is analogous to a continuously updated target volume. When a relationship is created, the master volume
t is assigned the role of primary volume and the auxiliary volume is assigned the role of secondary
t volume. Therefore, the initial copying direction is from master to auxiliary. When the relationship is in a
consistent state, you can reverse the copy direction.

t The two volumes in a relationship must be the same size. When the two volumes are in the same cluster,
they must be in the same I/O group.

For ease of application management, a relationship can be added to a consistency group.

Note: Membership of a consistency group is an attribute of the relationship, not the consistency group.
Therefore, issue the **svctask chrcrelationship** command to add or remove a relationship to or from a
consistency group.

Copy types

t A Metro Mirror copy ensures that updates are committed to both the primary and secondary volumes
before sending confirmation of I/O completion to the host application. This ensures that the secondary
t volume is synchronized with the primary volume in the event that a failover operation is performed.

A Global Mirror copy allows the host application to receive confirmation of I/O completion before the
t updates are committed to the secondary volume. If a failover operation is performed, the host application
t must recover and apply any updates that were not committed to the secondary volume.

States

t When a Metro Mirror or Global Mirror relationship is created with two volumes in different clusters, the
distinction between the connected and disconnected states is important. These states apply to both
clusters, the relationships, and the consistency groups. The following Metro Mirror and Global Mirror
relationship states are possible:

InconsistentStopped
t   The primary volume is accessible for read and write I/O operations, but the secondary volume is
t   not accessible for either operation. A copy process must be started to make the secondary volume
        consistent.

InconsistentCopying
t   The primary volume is accessible for read and write I/O operations, but the secondary volume is
        not accessible for either operation. This state is entered after an **svctask startrcrelationship**
        command is issued to a consistency group in the InconsistentStopped state. This state is also
        entered when an **svctask startrcrelationship** command is issued, with the force option, to a
        consistency group in the Idling or ConsistentStopped state.

ConsistentStopped
t   The secondary volume contains a consistent image, but it might be out of date with respect to the
t   primary volume. This state can occur when a relationship was in the ConsistentSynchronized
        state and experiences an error that forces a freeze of the consistency group. This state can also
        occur when a relationship is created with the CreateConsistentFlag parameter set to TRUE.

ConsistentSynchronized

t   The primary volume is accessible for read and write I/O operations. The secondary volume is accessible for read-only I/O operations.

t **Idling** A master volume and an auxiliary volume operates in the primary role. Consequently the volume is accessible for write I/O operations.

IdlingDisconnected

t   The volumes in this half of the consistency group are all operating in the primary role and can accept read or write I/O operations.

InconsistentDisconnected

t   The volumes in this half of the consistency group are all operating in the secondary role and cannot accept read or write I/O operations.

ConsistentDisconnected

t   The volumes in this half of the consistency group are all operating in the secondary role and can accept read I/O operations but not write I/O operations.

Metro Mirror and Global Mirror relationships between clusters

Metro Mirror and Global Mirror relationships can exist simultaneously between clusters. In this type of configuration, there can be impacts to performance because write data from both Metro Mirror and Global Mirror relationships is transported over the same intercluster links.

Metro Mirror and Global Mirror relationships manage heavy workload differently. Metro Mirror typically maintains the relationships that are in the copying or synchronized states, which causes the primary host applications to see degraded performance. Global Mirror requires a higher level of write performance to primary host applications. If the link performance is severely degraded, the link tolerance feature automatically stops Global Mirror relationships when the link tolerance threshold is exceeded. As a result, Global Mirror writes can suffer degraded performance if Metro Mirror relationships use most of the capability of the intercluster link.

Metro Mirror and Global Mirror partnerships

Partnerships define an association between a local cluster and a remote cluster.

Before a Metro Mirror or Global Mirror relationship or consistency group can be created with a remote cluster, a *partnership* between the two clusters must be established. If Global Mirror or Metro Mirror relationships or consistency groups exist between two remote clusters, those clusters must maintain their partnership. Each cluster can maintain up to three partnerships, and each partnership can be with a single remote cluster. As many as four clusters can be directly associated with each other.

1 You can create new Metro Mirror or Global Mirror partnerships between clusters with different software
2 levels. If the partnerships are between a SAN Volume Controller version 6.1.0 cluster and a cluster that is
2 at 4.3.1, each cluster can participate in a single partnership with another cluster. If the clusters are all
3 either SAN Volume Controller version 6.1.0 or version 5.1.0, each cluster can participate in up to three
2 cluster partnerships. A maximum of four clusters are permitted in the same connected set. A partnership
2 cannot be formed between a SAN Volume Controller version 6.1.0 and one that is running a version that
2 is earlier than 4.3.1.

2 **Attention:** If you want to upgrade a cluster to SAN Volume Controller version 6.1.0 and the partner is
2 running version 4.3.0 or earlier, you must first upgrade the partner cluster to SAN Volume Controller
2 4.3.1 or later before you upgrade the first cluster to version 6.1.0.

Clusters also become indirectly associated with each other through partnerships. If two clusters each have a partnership with a third cluster, those two clusters are indirectly associated. A maximum of four clusters can be directly or indirectly associated.

t   The nodes within the cluster must know not only about the relationship between the two volumes but also about an association among clusters.

The following examples show possible partnerships that can be established among SAN Volume Controller clusters.
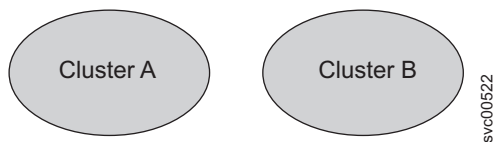
Figure 19. Two clusters with no partnerships
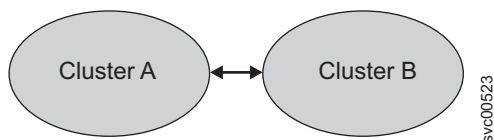
Figure 20. Two clusters with one partnership.
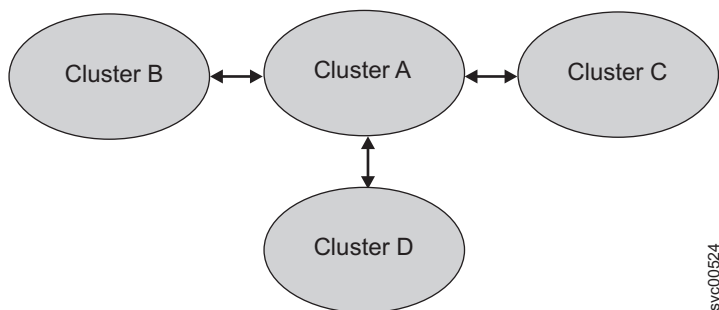
Figure 21. Four clusters in a partnership. Cluster A might be a disaster recovery site.
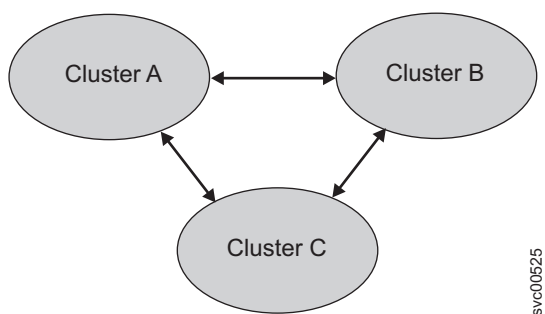
Figure 22. Three clusters in a migration situation. Data Center B is migrating to C. Cluster A is host production, and Cluster B and Cluster C are disaster recovery.
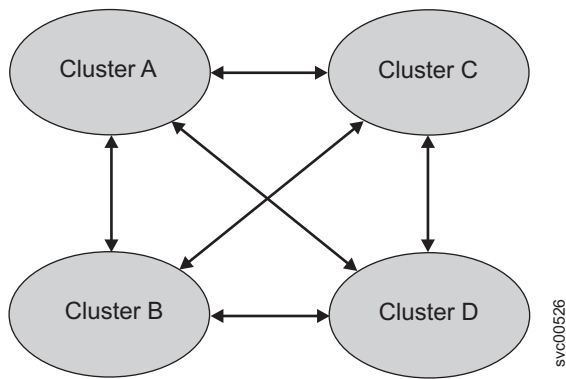
Figure 23. Clusters in a fully connected mesh configuration. Every cluster has a partnership to each of the three other clusters.
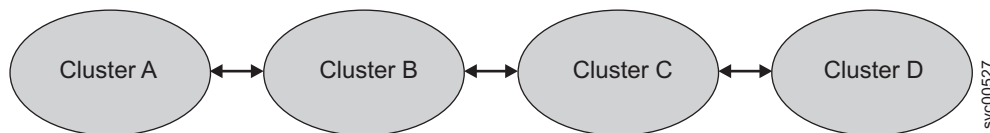
Figure 24. Four clusters in three partnerships.

Figure 25 depicts a cluster configuration that is not supported. Five clusters are in the connected set, even though no individual cluster is in more than two partnerships.
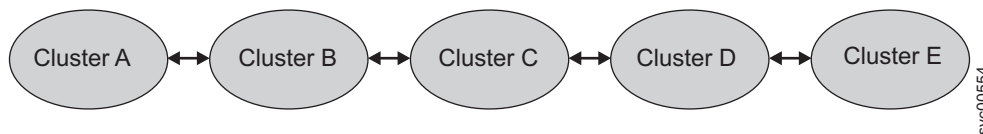
Figure 25. An unsupported cluster configuration.

To establish a Metro Mirror and Global Mirror partnership between two clusters, you must run the **svctask mkpartnership** command from both clusters. For example, to establish a partnership between clusterA and clusterB, you must run the **svctask mkpartnership** command from clusterA and specify clusterB as the remote cluster. At this point the partnership is partially configured and is sometimes described as one-way communication. Next, you must run the **svctask mkpartnership** command from clusterB and specify clusterA as the remote cluster. When this command completes, the partnership is fully configured for two-way communication between the clusters. You can also use the management GUI to create Metro Mirror and Global Mirror partnerships.

The state of the partnership helps determine whether the partnership operates as expected. In addition to being fully configured, a cluster partnership can have the following states:

Partially Configured
> Indicates that only one cluster partner is defined from a local or remote cluster to the displayed cluster and is started. For the displayed cluster to be configured fully and to complete the partnership, you must define the cluster partnership from the cluster that is displayed to the corresponding local or remote cluster. You can do this by issuing the **mkpartnership** command on the local and remote cluster that are in the partnership, or by using the management GUI to create a partnership on both the local and remote clusters.

Fully Configured
> Indicates that the partnership is defined on the local and remote clusters and is started.

Remote Not Present
> Indicates that the remote cluster is not present to the partnership.

Partially Configured (Local Stopped)
> Indicates that the local cluster is only defined to remote cluster and the local cluster is stopped.

Fully Configured (Local Stopped)
> Indicates that a partnership is defined on both the local and remote clusters and the remote cluster is present, but the local cluster is stopped.

Fully Configured (Remote Stopped)
> Indicates that a partnership is defined on both the local and remote clusters and the remote cluster is present, but the remote cluster is stopped.

Fully Configured (Local Excluded)
> Indicates that a partnership is defined between a local and remote cluster; however, the local cluster has been excluded. Usually this state occurs when the fabric link between the two clusters has been compromised by too many fabric errors or slow response times of the cluster partnership. Check the event log for 1720 errors by selecting **Service and Maintenance** ⇨ **Analyze Event Log** to resolve these errors.

Fully Configured (Remote Excluded)
> Indicates that a partnership is defined between a local and remote cluster; however, the remote cluster has been excluded. Usually this state occurs when the fabric link between the two clusters has been compromised by too many fabric errors or slow response times of the cluster partnership. Check the event log for 1720 errors by selecting **Service and Maintenance** ⇨ **Analyze Event Log** to resolve these errors.

Fully Configured (Remote Exceeded)
> Indicates that a partnership is defined between a local and remote cluster and the remote is available; however, the remote cluster exceeds the number of allowed clusters within a cluster network. The maximum of four clusters can be defined in a network. If the number of clusters exceeds that limit, SAN Volume Controller determines the inactive cluster or clusters by sorting all the clusters by their unique identifier in numerical order. The inactive cluster partner which is not in the top four of the cluster unique identifiers displays **Fully Configured (Remote Exceeded)**.

To change Metro Mirror and Global Mirror partnerships, use the **svctask chpartnership** command. To delete Metro Mirror and Global Mirror partnerships, use the **svctask rmpartnership** command.

Attention: Before you run the **svctask rmpartnership** command, you must remove all relationships and groups that are defined between the two clusters. To display cluster relationships and groups, run the **svcinfo lsrcrelationship** and **svcinfo lsrcconsistgrp** commands. To remove the relationships and groups that are defined between the two clusters, run the **svctask rmrcrelationship** and **svctask rmrcconsistgrp** commands.

Background copy management

You can control the rate at which the initial background copy from the local cluster to the remote cluster is performed. The bandwidth parameter specifies this rate in whole megabytes per second.

Global Mirror configuration requirements

To use the Global Mirror feature, all components in the SAN must be capable of sustaining the workload that is generated by application hosts and the Global Mirror background copy process. If all of the components in the SAN cannot sustain the workload, the Global Mirror relationships are automatically stopped to protect your application hosts from increased response times.

When using the Global Mirror feature, follow these best practices:

- Use IBM Tivoli Storage Productivity Center or an equivalent SAN performance analysis tool to monitor your SAN environment. The IBM Tivoli Storage Productivity Center provides an easy way to analyze the SAN Volume Controller performance statistics.
- Analyze the SAN Volume Controller performance statistics to determine the peak application write workload that the link must support. Gather statistics over a typical application I/O workload cycle.
- Set the background copy rate to a value that can be supported by the intercluster link and the backend storage controllers at the remote cluster.
- Do not use cache-disabled volumes in Global Mirror relationships.
- Set the gmlinktolerance parameter to an appropriate value. The default value is 300 seconds (5 minutes).
- When you perform SAN maintenance tasks, take one of the following actions:
  - Reduce the application I/O workload for the duration of the maintenance task.
  - Disable the gmlinktolerance feature or increase the gmlinktolerance value.

    Note: If the gmlinktolerance value is increased during the maintenance task, do not set it to the normal value until the maintenance task is complete. If the gmlinktolerance feature is disabled for the duration of the maintenance task, enable it after the maintenance task is complete.
  - Stop the Global Mirror relationships.
- Evenly distribute the preferred nodes for the Global Mirror volumes between the nodes in the clusters. Each volume in an I/O group has a preferred node property that can be used to balance the I/O load between nodes in the I/O group. The preferred node property is also used by the Global Mirror feature to route I/O operations between clusters. A node that receives a write operation for a volume is normally the preferred node for that volume. If the volume is in a Global Mirror relationship, the node is responsible for sending the write operation to the preferred node of the secondary volume. By default, the preferred node of a new volume is the node that owns the fewest volumes of the two nodes in the I/O group. Each node in the remote cluster has a set pool of Global Mirror system resources for each node in the local cluster. To maximize Global Mirror performance, set the preferred nodes for the volumes of the remote cluster to use every combination of primary nodes and secondary nodes.

Long distance links for Metro Mirror and Global Mirror partnerships

For intracluster partnerships, all clusters can be considered as candidates for Metro Mirror or Global Mirror operations. For intercluster partnerships, cluster pairs must be separated by a number of moderately high bandwidth links.

Figure 26 on page 70 shows an example of a configuration that uses dual redundant fabrics. Part of each fabric is located at the local cluster and the remote cluster. There is no direct connection between the two fabrics.
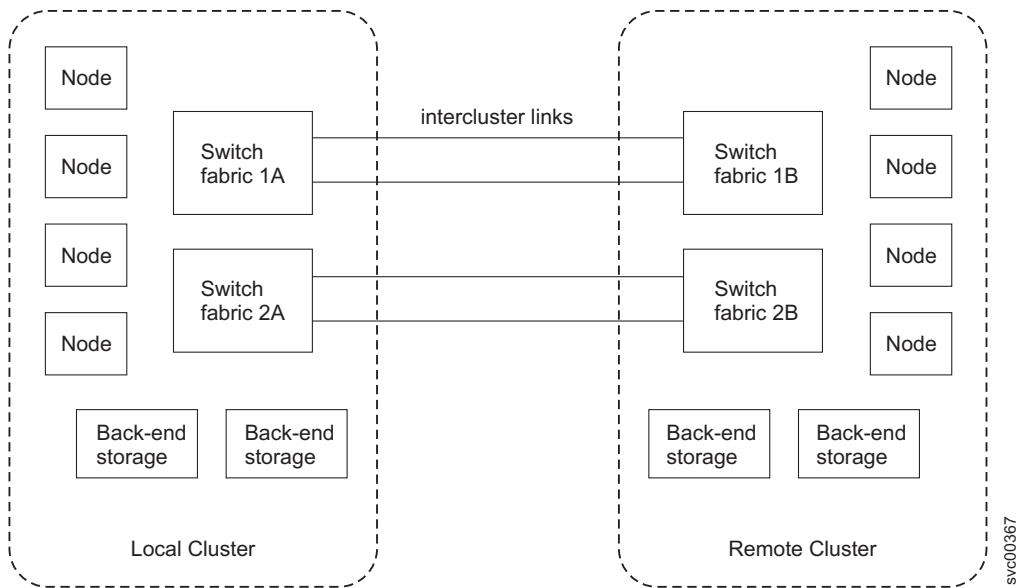
Figure 26. Redundant fabrics

You can use fibre-channel extenders or SAN routers to increase the distance between two clusters. Fibre-channel extenders transmit fibre-channel packets across long links without changing the contents of the packets. SAN routers provide virtual nPorts on two or more SANs to extend the scope of the SAN. The SAN router distributes the traffic from one virtual nPort to the other virtual nPort. The two fibre-channel fabrics are independent of each other. Therefore, nPorts on each of the fabrics cannot directly log into each other. See the following website for specific firmware levels and the latest supported hardware:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

If you use fibre-channel extenders or SAN routers, you must meet the following requirements:

- For SAN Volume Controller software level 4.1.1 or later, the round-trip latency between sites cannot exceed 80 ms for either fibre-channel extenders or SAN routers.
- The configuration must be tested with the expected peak workloads.
- Metro Mirror and Global Mirror require a specific amount of bandwidth for intercluster heartbeat traffic. The amount of traffic depends on the number of nodes that are in both the local cluster and the remote cluster. Table 19 lists the intercluster heartbeat traffic for the primary cluster and the secondary cluster. These numbers represent the total traffic between two clusters when there are no I/O operations running on the copied volumes. Half of the data is sent by the primary cluster and half of the data is sent by the secondary cluster so that traffic is evenly divided between all of the available intercluster links. If you have two redundant links, half of the traffic is sent over each link.

Table 19. Intercluster heartbeat traffic in Mbps

Cluster 1	Cluster 2			
	2 nodes	4 nodes	6 nodes	8 nodes
2 nodes	2.6	4.0	5.4	6.7
4 nodes	4.0	5.5	7.1	8.6
6 nodes	5.4	7.1	8.8	10.5
8 nodes	6.7	8.6	10.5	12.4

- The bandwidth between two sites must meet the peak workload requirements and maintain the maximum round-trip latency between the sites. When you evaluate the workload requirement, you

must consider the average write workload over a period of one minute or less and the required synchronization copy bandwidth. If there are no active synchronization copies and no write I/O operations for volumes that are in the Metro Mirror or Global Mirror relationship, the SAN Volume Controller protocols operate with the bandwidth that is indicated in Table 19 on page 70. However, you can only determine the actual amount of bandwidth that is required for the link by considering the peak write bandwidth to volumes that are participating in Metro Mirror or Global Mirror relationships and then adding the peak write bandwidth to the peak synchronization bandwidth.

- If the link between two sites is configured with redundancy so that it can tolerate single failures, the link must be sized so that the bandwidth and latency statements are correct during single failure conditions.
- The channel must not be used for links between nodes in a single cluster. Configurations that use long distance links in a single cluster are not supported and can cause I/O errors and loss of access.
- The configuration is tested to confirm that any failover mechanisms in the intercluster links interoperate satisfactorily with SAN Volume Controller.
- All other SAN Volume Controller configuration requirements are met.

Limitations on host-to-cluster distances

There is no limit on the fibre-channel optical distance between SAN Volume Controller nodes and host servers. You can attach a server to an edge switch in a core-edge configuration with the SAN Volume Controller cluster at the core. SAN Volume Controller clusters support up to three ISL hops in the fabric. Therefore, the host server and the SAN Volume Controller cluster can be separated by up to five fibre-channel links. If you use longwave small form-factor pluggable (SFP) transceivers, four of the fibre-channel links can be up to 10 km long.

Using the intercluster link for host traffic

If you use the intercluster link for host traffic, ensure that you have sufficient bandwidth to support all sources of load.

Scenario: The hosts in a local cluster can read and write to the volumes in a remote cluster

In this scenario, the hosts in the local cluster also exchange heartbeats with the hosts that are in the remote cluster. Because the intercluster link is being used for multiple purposes, you must have sufficient bandwidth to support the following sources of load:

- Global Mirror or Metro Mirror data transfers and the SAN Volume Controller cluster heartbeat traffic.
- Local host to remote volume I/O traffic or remote host to local volume I/O traffic.
- Local-host-to-remote-host heartbeat traffic. If the local-host-to-remote-volume I/O traffic is allowed to consume a high percentage of intercluster link bandwidth, the latency seen by the hosts that access SAN Volume Controller volumes that are participating in Metro Mirror or Global Mirror operations can be impacted. The bandwidth congestion can cause the Global Mirror link tolerance threshold to be exceeded. When the Global Mirror link tolerance threshold is exceeded, Global Mirror relationships are stopped.

Metro Mirror and Global Mirror consistency groups

You can group Metro Mirror or Global Mirror relationships into a consistency group so that they can be updated at the same time. A command that is issued to the consistency group is simultaneously applied to all of the relationships in the group.

Metro Mirror or Global Mirror relationships can be based on "loose" or "tight" associations. A more significant use arises when the relationships contain volumes with a tight association. A simple example of a tight association is the spread of data for an application across more than one volume. A more complex example is when multiple applications run on different host systems. Each application has data

on different volumes, and these applications exchange data with each other. In both examples, specific rules exist as to how the relationships can be updated. This ensures that the set of secondary volumes contain usable data. The key property is that these relationships are consistent.

Metro Mirror or Global Mirror relationships can be based on "loose" or "tight" associations. A more significant use arises when the relationships contain volumes with a tight association. A simple example of a tight association is the spread of data for an application across more than one volume. A more complex example is when multiple applications run on different host systems. Each application has data on different volumes, and these applications exchange data with each other. In both examples, specific rules exist as to how the relationships can be updated. These rules ensure that the set of secondary volumes contain usable data. The key property is that these relationships are consistent.

Metro Mirror or Global Mirror relationships can only belong to one consistency group; however, they do not have to belong to a consistency group. Relationships that are not part of a consistency group are called stand-alone relationships. A consistency group can contain zero or more relationships. All relationships in a consistency group must have matching primary and secondary clusters, which are sometimes referred to as master and auxiliary clusters. All relationships in a consistency group must also have the same copy direction and state.

Metro Mirror and Global Mirror relationships cannot belong to the same consistency group. A copy type is automatically assigned to a consistency group when the first relationship is added to the consistency group. After the consistency group is assigned a copy type, only relationships of that copy type can be added to the consistency group. Each cluster can have a maximum of six different types of consistency groups.

The following types of consistency groups are possible:
- Intracluster Metro Mirror
- Intercluster Metro Mirror from the local cluster to remote cluster
- Intercluster Metro Mirror from the remote cluster to local cluster
- Intracluster Global Mirror
- Intercluster Global Mirror from the local cluster to remote cluster
- Intercluster Global Mirror from the remote cluster to local cluster

States

A consistency group can be in one of the following states:

Inconsistent (stopped)
The primary volumes are accessible for read and write I/O operations, but the secondary volumes are not accessible for either. A copy process must be started to make the secondary volumes consistent.

Inconsistent (copying)
The primary volumes are accessible for read and write I/O operations, but the secondary volume are not accessible for either. This state is entered after the **svctask startrcconsistgrp** command is issued to a consistency group in the InconsistentStopped state. This state is also entered when the **svctask startrcconsistgrp** command is issued, with the force option, to a consistency group in the Idling or ConsistentStopped state.

Consistent (stopped)
The secondary volumes contain a consistent image, but it might be out-of-date with respect to the primary volumes. This state can occur when a relationship was in the ConsistentSynchronized state and experiences an error that forces a freeze of the consistency group. This state can also occur when a relationship is created with the CreateConsistentFlag set to TRUE.

Consistent (synchronized)

t   The primary volumes are accessible for read and write I/O operations. The secondary volumes are accessible for read-only I/O operations.

t **Idling** Both the primary volumes and the secondary volumes are operating in the primary role.
t   Consequently the volumes are accessible for write I/O operations.

Idling (disconnected)

t   The volumes in this half of the consistency group are all operating in the primary role and can accept read or write I/O operations.

Inconsistent (disconnected)

t   The volumes in this half of the consistency group are all operating in the secondary role and cannot accept read or write I/O operations.

Consistent (disconnected)

t   The volumes in this half of the consistency group are all operating in the secondary role and can accept read I/O operations but not write I/O operations.

Empty The consistency group does not contain any relationships.

2 **Note:** Volume copies are synchronized when their contents are consistent. If write operations take place
2 on either the primary or secondary volume after a consistent (stopped) or idling state occurs, they might
2 no longer be synchronized.

Background copy bandwidth impact on foreground I/O latency

The background copy bandwidth determines the rate at which the background copy for Metro Mirror or Global Mirror Copy Services are attempted.

The background copy bandwidth can affect foreground I/O latency in one of three ways:

- If the background copy bandwidth is set too high for the intercluster link capacity, the following results can occur:
  - The intercluster link is not able to process the background copy I/Os fast enough, and the I/Os can back up (accumulate).
  - For Metro Mirror, there is a delay in the synchronous secondary writes of foreground I/Os.
  - For Global Mirror, the work is backlogged, which delays the processing of writes and causes the relationship to stop.
  - The foreground I/O latency increases as detected by applications.
- If the background copy bandwidth is set too high for the storage at the *primary* site, background copy read I/Os overload the primary storage and delay foreground I/Os.
- If the background copy bandwidth is set too high for the storage at the *secondary* site, background copy writes at the secondary overload the secondary storage and again delay the synchronous secondary writes of foreground I/Os.
  - For Global Mirror, the work is backlogged and again the relationship is stopped

To set the background copy bandwidth optimally, you must consider all three resources (the primary storage, the intercluster link bandwidth, and the secondary storage). Provision the most restrictive of these three resources between the background copy bandwidth and the peak foreground I/O workload. You must also consider concurrent host I/O because if other write operations arrive at the primary cluster for copy to the remote site, these write operations can be delayed by a high level of background copy and the hosts at the primary site receive poor write-operation response times.

The provisioning for optimal bandwidth for the background copy can also be calculated by determining how much background copy can be allowed before performance of host I/O becomes unacceptable. The background copy bandwidth can be decreased slightly to accommodate peaks in workload and provide a safety margin for host I/O.

Example

If the bandwidth setting at the primary site for the secondary cluster is set to 200 MBps (megabytes per second) and the relationships are not synchronized, the SAN Volume Controller attempts to resynchronize the relationships at a maximum rate of 200 MBps with a 25 MBps restriction for each individual relationship. The SAN Volume Controller cannot resynchronize the relationship if the throughput is restricted. The following can restrict throughput:

- The read response time of backend storage at the primary cluster
- The write response time of the backend storage at the secondary site
- Intercluster link latency

Migrating a Metro Mirror relationship to a Global Mirror relationship

You can migrate a Metro Mirror relationship to a Global Mirror relationship.

Scenario: I/O operations to the secondary volume can be stopped during the migration

In this scenario, you have the ability to stop I/O operations to the secondary volume during the migration process.

To stop I/O operations to the secondary volume while migrating a Metro Mirror relationship to a Global Mirror relationship, you must specify the synchronized option when you create the Global Mirror relationship.

1. Stop all host I/O operations to the primary volume.
2. Verify that the Metro Mirror relationship is consistent.

   Important: If the Metro Mirror relationship is not consistent when it is stopped, or if any host I/O operations run between the Metro Mirror relationship being stopped and the Global Mirror relationship being created, the updates are not copied to the secondary volume.

3. Delete the Metro Mirror relationship.
4. Create the Global Mirror relationship between the same two volumes.

After the Global Mirror relationship is created, you can start the relationship and resume host I/O operations.

Scenario: I/O operations to the secondary volume cannot be stopped during the migration

In this scenario, you do not have the ability to stop I/O operations to the secondary volume during the migration process.

If I/O operations to the secondary volume cannot be stopped, the data on the secondary volume becomes out-of-date. When the Global Mirror relationship is started, the secondary volume is inconsistent until all of the recent updates are copied to the remote site.

If you do not require a consistent copy of the volume at the secondary site, perform the following steps to migrate from a Metro Mirror relationship to a Global Mirror relationship:

Important: The data on the secondary volume is not usable until the synchronization process is complete. Depending on your link capabilities and the amount of data that is being copied, this process can take an extended period of time. You must set the background copy bandwidth for the intercluster partnerships to a value that does not overload the intercluster link.

1. Delete the Metro Mirror relationship.

t 2. Create and start the Global Mirror relationship between the same two volumes.

t If you require a consistent copy of the volume at the secondary site, perform the following steps to migrate from a Metro Mirror relationship to a Global Mirror relationship:

   1. Delete the Metro Mirror relationship.

t 2. Create a Global Mirror relationship between volumes that were not used for the Metro Mirror
t   relationship. This preserves the volume so that you can use it if you require a consistent copy at a later time.

   Alternatively, you can use the FlashCopy feature to maintain a consistent copy. Perform the following steps to use the FlashCopy feature to maintain a consistent copy:

t 1. Start a FlashCopy operation for the Metro Mirror volume.

   2. Wait for the FlashCopy operation to complete.

t 3. Create and start the Global Mirror relationship between the same two volumes. The FlashCopy
t   volume is now your consistent copy.

Using FlashCopy to create a consistent image before restarting a Global Mirror relationship

For disaster recovery purposes, you can use the FlashCopy feature to create a consistent copy of an image before you restart a Global Mirror relationship.

When a consistent relationship is stopped, the relationship enters the consistent_stopped state. While in this state, I/O operations at the primary site continue to run. However, updates are not copied to the secondary site. When the relationship is restarted, the synchronization process for new data is started.
t During this process, the relationship is in the inconsistent_copying state. The secondary volume for the relationship cannot be used until the copy process completes and the relationship returns to the
t consistent state. When this occurs, start a FlashCopy operation for the secondary volume before you restart the relationship. While the relationship is in the copying state, the FlashCopy feature can provide a consistent copy of the data. If the relationship does not reach the synchronized state, you can use the
t FlashCopy target volume at the secondary site.

The SVCTools package that is available on the IBM alphaWorks® Web site provides an example script that demonstrates how to manage the FlashCopy process. See the copymanager script that is available in the SVCTools package. You can download the SVCTools package from the following Web site:

www.alphaworks.ibm.com/tech/svctools/download

Monitoring Global Mirror performance with the IBM System Storage Productivity Center

You can use the IBM System Storage Productivity Center (SSPC) to monitor key Global Mirror performance measurements.

It is important to use a Storage Area Network (SAN) performance monitoring tool to ensure that all SAN components are performing correctly. This is particularly important when you use an asynchronous copying solution such as the SAN Volume Controller Global Mirror feature. SSPC monitors key performance measures and alerts you when thresholds are exceeded.

t **Note:** If your volume or MDisk configuration changes, restart the SSPC performance report to ensure that performance is monitored for the new configuration.
   Use SSPC to check the following measurements:

   • The *Port to Remote Node Send Response Time* measurement is less than 80 milliseconds. If this measurement is greater than 80 milliseconds during monitoring, the long-distance link has excessive latency. Ensure that the link is operating at its maximum bandwidth.

- The sum of the *Port to Local Node Send Response Time* measurement and the *Port to Local Node Send Queue* measurement is less than 1 millisecond for the primary cluster and the CPU Utilization Percentage is below 50%. A value that exceeds these amounts can indicate that an I/O group is reaching the I/O throughput limit, which can limit performance.
- The sum of the *Backend Write Response Time* measurement and the *Write Queue Time for Global Mirror MDisks* measurement of the secondary cluster is less than 100 milliseconds. A longer response time can indicate that the storage controller is overloaded.
- The sum of the *Backend Write Response Time* measurement and the *Write Queue Time for Global Mirror MDisks* measurement of the primary cluster is less than 100 milliseconds. If the response time is greater than 100 milliseconds, application hosts might see extended response times when the SAN Volume Controller cluster cache is full.
- The *Write Data Rate for Global Mirror storage pools* measurement of the secondary cluster indicates the amount of data that is being written by Global Mirror operations. If this value approaches either the intercluster link bandwidth or the storage controller throughput limit, further increases can cause overloading of the system. Monitor for this condition in a way that is appropriate for your network.

The gmlinktolerance feature

You can use the **svctask chcluster** CLI command or the management GUI to set the gmlinktolerance feature. The gmlinktolerance feature represents the number of seconds that the primary SAN Volume Controller cluster tolerates slow response times from the secondary cluster.

If the poor response extends past the specified tolerance, a 1920 error is logged and one or more Global Mirror relationships are automatically stopped. This protects the application hosts at the primary site. During normal operation, application hosts see a minimal impact to response times because the Global Mirror feature uses asynchronous replication. However, if Global Mirror operations experience degraded response times from the secondary cluster for an extended period of time, I/O operations begin to queue at the primary cluster. This results in an extended response time to application hosts. In this situation, the gmlinktolerance feature stops Global Mirror relationships and the application hosts response time returns to normal. After a 1920 error has occurred, the Global Mirror auxiliary volumes are no longer in the consistent_synchronized state until you fix the cause of the error and restart your Global Mirror relationships. For this reason, ensure that you monitor the cluster to track when this occurs.

You can disable the gmlinktolerance feature by setting the gmlinktolerance value to 0 (zero). However, the gmlinktolerance cannot protect applications from extended response times if it is disabled. It might be appropriate to disable the gmlinktolerance feature in the following circumstances:

- During SAN maintenance windows where degraded performance is expected from SAN components and application hosts can withstand extended response times from Global Mirror volumes.
- During periods when application hosts can tolerate extended response times and it is expected that the gmlinktolerance feature might stop the Global Mirror relationships. For example, if you are testing using an I/O generator which is configured to stress the backend storage, the gmlinktolerance feature might detect the high latency and stop the Global Mirror relationships. Disabling gmlinktolerance prevents this at the risk of exposing the test host to extended response times.

Diagnosing and fixing 1920 errors

A 1920 error indicates that one or more of the SAN components are unable to provide the performance that is required by the application hosts. This can be temporary (for example, a result of maintenance activity) or permanent (for example, a result of a hardware failure or unexpected host I/O workload). If you are experiencing 1920 errors, set up a SAN performance analysis tool, such as the IBM Tivoli Storage Productivity Center, and make sure that it is correctly configured and monitoring statistics when the problem occurs. Set your SAN performance analysis tool to the minimum available statistics collection interval. For the IBM Tivoli Storage Productivity Center, the minimum interval is five minutes. If several 1920 errors have occurred, diagnose the cause of the earliest error first. The following questions can help you determine the cause of the error:

- Was maintenance occurring at the time of the error? This might include replacing a storage system physical disk, upgrading the firmware of the storage system, or performing a code upgrade on one of the SAN Volume Controller clusters. You must wait until the maintenance procedure is complete and then restart the Global Mirror relationships. You must wait until the maintenance procedure is complete to prevent a second 1920 error because the system has not yet returned to a stable state with good performance.

- Were there any unfixed errors on either the source or target system? If yes, analyze them to determine if they might have been the reason for the error. In particular, see if they either relate to the volume or MDisks that are being used in the relationship, or if they would have caused a reduction in performance of the target system. Ensure that the error is fixed before you restart the Global Mirror relationship.

- Is the long distance link overloaded? If your link is not capable of sustaining the short-term peak Global Mirror workload, a 1920 error can occur. Perform the following checks to determine if the long distance link is overloaded:

  – Look at the total Global Mirror auxiliary volume write throughput before the Global Mirror relationships were stopped. If this is approximately equal to your link bandwidth, your link might be overloaded. This might be due to application host I/O operations or a combination of host I/O and background (synchronization) copy activities.

  – Look at the total Global Mirror source volume write throughput before the Global Mirror relationships were stopped. This represents the I/O operations that are being performed by the application hosts. If these operations are approaching the link bandwidth, upgrade the link's bandwidth, reduce the I/O operations that the application is attempting to perform, or use Global Mirror to copy fewer volumes. If the auxiliary disks show significantly more write I/O operations than the source volumes, there is a high level of background copy. Decrease the Global Mirror partnership's background copy rate parameter to bring the total application I/O bandwidth and background copy rate within the link's capabilities.

  – Look at the total Global Mirror source volume write throughput after the Global Mirror relationships were stopped. If write throughput increases by 30% or more when the relationships are stopped, the application hosts are attempting to perform more I/O operations than the link can sustain. While the Global Mirror relationships are active, the overloaded link causes higher response times to the application host, which decreases the throughput it can achieve. After the Global Mirror relationships have stopped, the application host sees lower response times. In this case, the link bandwidth must be increased, the application host I/O rate must be decreased, or fewer volumes must be copied using Global Mirror.

- Are the storage systems at the secondary cluster overloaded? If one or more of the MDisks on a storage system are providing poor service to the SAN Volume Controller cluster, a 1920 error occurs if this prevents application I/O operations from proceeding at the rate that is required by the application host. If the back-end storage system requirements have been followed, the error might have been caused by a decrease in storage system performance. Use IBM Tivoli Storage Productivity Center to obtain the back-end write response time for each MDisk at the secondary cluster. If the response time for any individual MDisk exhibits a sudden increase of 50 ms or more or if the response time is above 100 ms, this indicates a problem. Perform the following checks to determine if the storage systems are overloaded:

  – Check the storage system for error conditions such as media errors, a failed physical disk, or associated activity such as RAID array rebuilding. If there is an error, you should fix the problem and then restart the Global Mirror relationships.

  – If there is no error, determine if the secondary storage system is capable of processing the required level of application host I/O operations. It might be possible to improve the performance of the storage system by adding more physical disks to a RAID array, changing the RAID level of the array, changing the cache settings of the storage system and check in the cache battery to ensure it is operational, or changing other specific configuration parameters of the storage system.

- Are the storage systems at the primary cluster overloaded? Analyze the performance of the primary back-end storage using the same steps as for the secondary back-end storage. If performance is bad, limit the amount of I/O operations that can be performed by application hosts. Monitor the back-end

storage at the primary site even if the Global Mirror relationships have not been affected. If bad performance continues for a prolonged period, a 1920 error occurs and the Global Mirror relationships are stopped.

- Is one of your SAN Volume Controller clusters overloaded? Use IBM Tivoli Storage Productivity Center to obtain the port-to-local-node send response time and the port to local node send queue time. If the total of these two statistics for either cluster is above 1 millisecond, the SAN Volume Controller might be experiencing a very high I/O load. Also check the SAN Volume Controller node CPU utilization. If this figure is above 50%, this can also be contributing to the problem. In either case, contact your IBM service representative for further assistance. If CPU utilization is much higher for one node than for the other node in the same I/O group, this might be caused by having different node hardware types within the same I/O group. For example, a SAN Volume Controller 2145-8F4 in the same I/O group as a SAN Volume Controller 2145-8G4. If this is the case, contact your IBM service representative.

- Do you have FlashCopy operations in the prepared state at the secondary cluster? If the Global Mirror auxiliary volumes are the sources of a FlashCopy mapping and that mapping is in the prepared state for an extended time, performance to those volumes can be impacted because the cache is disabled. Start the FlashCopy mapping to enable the cache and improve performance for Global Mirror I/O operations.

Valid combinations of FlashCopy and Metro Mirror or Global Mirror functions

The following table outlines the combinations of FlashCopy and Metro Mirror or Global Mirror functions that are valid for a single volume.

FlashCopy	Metro Mirror or Global Mirror Primary	Metro Mirror or Global Mirror Secondary
FlashCopy source	Supported	Supported
FlashCopy target	Not supported	Not supported

Chapter 3. SAN fabric and LAN configuration

The *SAN fabric* is an area of the network that contains routers and switches. In fibre-channel environments, *zoning* is the grouping of multiple ports to form a virtual, private, storage network. iSCSI is an IP-based standard for linking data storage devices over a network and transferring data by carrying SCSI commands over IP networks.

SAN fabric overview

The *SAN fabric* is an area of the network that contains routers and switches. A SAN is configured into a number of zones. A device using the SAN can communicate only with devices that are included in the same zones that it is in. A SAN Volume Controller cluster requires several distinct types of zones: a cluster zone, host zones, and disk zones. The intercluster zone is optional.

In the host zone, the host systems can identify and address the SAN Volume Controller nodes. You can have more than one host zone and more than one disk zone. The cluster zone contains all ports from all SAN Volume Controller nodes in the cluster, unless you are using a dual-core fabric design. Create one zone for each host fibre-channel port. In a disk zone, the SAN Volume Controller nodes identify the storage systems. Generally, create one zone for each storage system. Host systems cannot operate on the storage systems directly; all data transfer occurs through the SAN Volume Controller nodes. If you are using the Metro Mirror and Global Mirror feature, create a zone with at least one port from each node in each cluster; up to four clusters are supported.

A cluster of SAN Volume Controller nodes is connected to the fibre-channel fabric and presents volumes to the host systems. You create these volumes from units of space within a managed disk (MDisk) group. A storage pool is a collection of MDisks that are presented by the storage systems (RAID controllers). The storage pool provides a storage pool. You specify how each group is created, and you can combine MDisks from different manufacturers' controllers in the same storage pool. However, to optimize the use of resources, ensure that all MDisks in a storage pool have similar performance characteristics.

Note: Some operating systems cannot tolerate other operating systems in the same host zone, although you might have more than one host type in the SAN fabric. For example, you can have a SAN that contains one host that runs on an IBM AIX® operating system and another host that runs on a Microsoft Windows operating system.

All communication between SAN Volume Controller nodes is performed through the SAN. All SAN Volume Controller node configuration and service commands are sent to the cluster through an Ethernet network.

Each SAN Volume Controller node contains its own vital product data (VPD). Each cluster contains VPD that is common to all the SAN Volume Controller nodes in the cluster, and any system, with the correct access authority, that is connected to the Ethernet network can access this VPD.

Configuration details

Storage area network (SAN) configurations that contain SAN Volume Controller nodes must be configured correctly.

A SAN configuration that contains SAN Volume Controller nodes must follow configuration rules for the following components:
- Storage systems
- Nodes

- Fibre-channel host bus adapters (HBAs)
- Fibre-channel switches
- iSCSI Ethernet ports
- Fabrics
- Zoning

SAN configuration, zoning, and split-cluster rules summary

These rules define the supported configuration for a SAN Volume Controller cluster during normal operations in a fibre-channel connection. If a single failure causes one or more of these rules to be invalidated, the configuration is still supported until the failure can be corrected and the configuration is brought back into a normal supported mode.

SAN Volume Controller descriptions of configuration terms

A *path* is a logical connection between two fibre-channel ports. The path can exist only if both of the two fibre-channel ports are in the same zone.

A *core switch* is the switch that contains the SAN Volume Controller ports. Because most SAN fabric traffic might flow through the cluster, put the SAN Volume Controller in the core of the fabric. Some configurations have a core switch that contains only interswitch links (ISLs) and a storage edge switch that contains the SAN Volume Controller ports. In this rules summary, a *storage edge switch* is the same as a core switch.

A *dual-core* fabric design is an environment where two switches are both designated as core switches in the same fabric. Every node has one fibre-channel port that is connected to each of the core switches. Zoning is then used to ensure that internode traffic flows only within a single switch wherever possible.

SAN configuration rules

SAN Volume Controller supports any SAN fabric configuration that is supported by the SAN vendors.

SAN Volume Controller connectivity:

- All internode communication between SAN Volume Controller ports in the same cluster must not cross ISLs. In dual-core designs, zoning must be used to prevent the SAN Volume Controller from using paths that cross between the two core switches.
- Each SAN Volume Controller port must have paths to at least one port on all other nodes in the cluster.
- Core switches must have enough ISL connectivity to handle the workload, which normally means that medium-to-large configurations have at least 64 ports in the core switch. Brocade M12 (silkworm 12000) switches are not supported as SAN Volume Controller core switches.
- Fibre-channel connections between SAN Volume Controller and the switch can be up to 100 m using the standard small form-factor pluggable (SFP) transceivers that are provided by SAN Volume Controller. Connections of up to 10,000 m are supported through long-wave SFP transceivers. The supported long-wave SFP transceivers are available to order from IBM.

Storage system connectivity:

- Connections between SAN Volume Controller and storage require the best available bandwidth. For optimal performance and reliability, ensure that paths between SAN Volume Controller and storage systems do not cross ISLs. If you use ISLs on these paths, make sure that sufficient bandwidth is available. SAN monitoring is required to identify faulty ISLs.
- Each SAN Volume Controller node must have a path to the same set of worldwide port names (WWPNs) for each storage system.

1 • Where multiple paths exist between SAN Volume Controller and storage systems and some of those
1   paths cross ISLs, use zoning to prevent SAN Volume Controller from using the paths that cross the
1   ISLs.
1 • SAN Volume Controller supports SAN routing technologies between SAN Volume Controller and
1   storage systems, as long as the routing stays entirely within fibre-channel connectivity and does not
1   use other transport technologies such as Internet Protocol (IP).

1 Host connectivity:
1 • Paths between hosts and SAN Volume Controller can cross ISLs.
1 • SAN Volume Controller supports SAN routing technologies (including FCIP links) between the SAN
1   Volume Controller and hosts. The use of long-distance FCIP connections, however, might degrade the
1   performance of any servers that are attached through this technology.

1 Intercluster connectivity:
1 • SAN Volume Controller supports SAN routing technology (including FCIP links) for intercluster
1   connections that use Metro Mirror or Global Mirror.

1 General SAN configuration rules:
1 • To make best use of the available bandwidth between switches, use ISL trunking (also known as port
1   channels) on all ISLs.
1 • When you use fibre-channel IP or iSCSI connections, it is best to use jumbo frames in the IP network.
1 • SAN Volume Controller supports between 2 and 4 counterpart SANs per cluster.
1 • High latency links can affect performance. Ensure that you conform to the support statements of the
1   SAN switch vendors and other connected devices regarding the length of fibre-channel connections in
1   the SAN.
2 • All fibre-channel devices must be connected through SAN fabrics and must not use direct connections.
1 • The SAN must contain only supported switches, fibre-channel extenders, and SAN routers. See the
1   following website for specific firmware levels and the latest supported hardware:
1   Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

1 Zoning rules

1 Apply these rules to each fabric that contains SAN Volume Controller ports. If the edge devices contain
1 more stringent zoning requirements, follow the storage system rules to further restrict the SAN Volume
1 Controller zoning rules. For example, IBM System Storage DS4000® does not support a storage system A
1 and storage system B in the same zone.

1 Host zoning:
2 • SAN Volume Controller requires single-initiator zoning for all large configurations that contain more
2   than 64 host objects. Each server fibre-channel port must be in its own zone, which contains the
2   fibre-channel port and SAN Volume Controller ports. In configurations of less than 64 hosts, you can
2   have up to 40 fibre-channel ports in a host zone if the zone contains similar HBAs and operating
2   systems.
1 • For optimal performance, include a maximum of two paths per volume per host fibre-channel port,
1   which equates to a zone that contains one port per SAN Volume Controller node per HBA.
1 • For load balancing, alternate the server fibre-channel ports between the ports of the SAN Volume
1   Controller. For example, the first server is zoned with ports 1 and 3 of each SAN Volume Controller
1   node (one SAN Volume Controller port per fabric). The second server is zoned with ports 2 and 4.
1 • The maximum number of supported paths to a SAN Volume Controller volume is eight.
1 • If a host object is not mapped to all I/O groups, do not include in the host zone SAN Volume
1   Controller ports from all nodes in the cluster. For example, if node A is in I/O group X and the host
1   object is mapped to I/O group X, include only ports from node A in the host zone.

The maximum number of hosts that are mapped to an I/O group is less than the maximum number of hosts per cluster. Therefore, in configurations that might grow to greater than the maximum number of hosts per I/O group, do not map every host to every I/O group

- When using a dual-core SAN design, it is a requirement that no internode communications use the ISL link. When you create host zones in this type of configuration, ensure that each SAN Volume Controller port in the host zone is attached to the same fibre-channel switch.

Storage system zoning:

- For most configurations, follow these rules:
  - For every storage system, create one zone that contains SAN Volume Controller ports from every node and all storage system ports, unless otherwise stated by the zoning guidelines for that storage system.
  - Single initiator zoning is not required for zones that include SAN Volume Controller and the storage system. The SAN Volume Controller ports are required to log in to each other to form the cluster.
- For configurations that use more than 64 storage-system WWPNs and two fibre-channel SANs, it might be necessary to use the following alternative zoning scheme to keep under the limit of 512 fibre-channel logins to each node port:

  For every storage system, divide the fibre-channel ports that are attached to one fibre-channel SAN into two groups. Create one zone for the first group of storage ports and add one fibre-channel port per node to the zone. Then create another zone for the second group of storage ports and add the other fibre-channel port from each node to the zone. Repeat the process for the second fibre-channel SAN.

  For example:
  - Two storage systems, I and J, with the following fibre-channel ports in SAN 1: I0, I1, J0, J1, J2, and J3.
  - Two nodes, A and B, with the following fibre-channel ports in SAN 1: A0, A1, B0, B1.

  This creates the following zones in SAN 1:
  - [A0, B0, I0, J0, J1]
  - [A1, B1, I1, J2, J3]

SAN Volume Controller zoning:

- Each SAN Volume Controller port must have a path to at least one port in every other node in the cluster. The zoning requirement to meet these rules is typically satisfied by the storage system zones. However, you can create one zone that contains all SAN Volume Controller ports in a single fibre-channel switch for clarity.
- Local cluster zoning follows the standard requirement for all ports on all nodes in a cluster to be zoned to one another.
- The following guidelines apply when using Metro Mirror or Global Mirror:
  - For each node that is to be zoned to a node in the partner cluster, zone exactly two fibre-channel ports.
  - If dual-redundant ISLs are available, split the two ports from each node evenly between the two ISLs. For example, exactly one port from each node is zoned across each ISL.

Split-cluster rules

In a split-cluster configuration, a site is defined as an independent failure domain. Different types of sites protect against different types of fault. For example:

- If each site is a different power phase within one data center, the SAN Volume Controller cluster can survive the failure of any single power domain.
- If each site is a different physical location, the SAN Volume Controller cluster can survive the failure of any single location.

1 In all cases, the SAN Volume Controller cluster does not guarantee that it can survive the failure of two
1 sites.

1 • Each SAN Volume Controller node must have two direct fibre-channel connections to one or more
1   SAN fabrics at both locations that contain SAN Volume Controller nodes.

1 • Ethernet port 1 on every SAN Volume Controller node must be connected to the same subnet or
1   subnets. The same is true for Ethernet port 2.

1 • You cannot have powered components between the SAN Volume Controller and the switches in a
1   split-cluster configuration. For example, you cannot have powered dense wavelength division
1   multiplexing (DWDM) fibre-channel extenders.

1 • You might be required to provide and replace long-wave SFP transceivers.

1 • Some service actions require the ability to perform actions to the front panel of all nodes in a cluster
1   within a short-time window. If you use split-site clusters, you will be required to assist the support
1   engineer and provide communication technology to coordinate these actions between the sites.

1 • The storage system at the third site must support extended quorum. This information is available in
1   the SAN Volume Controller interoperability matrixes that are available at the Support for SAN Volume
1   Controller (2145) website:

1   Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Storage system configuration details

3 When planning the configuration of storage systems for use with SAN Volume Controller clusters, review
3 these details.

See the following Web site for the latest support information:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

All SAN Volume Controller nodes in a cluster must be able to connect to the same set of storage system
ports on each device. A cluster that contains any two nodes that cannot connect to the same set of
storage-system ports is considered degraded, and a system error is logged that requires a repair action.
This rule can have important effects on a storage system such as an IBM System Storage DS4000 series
controller, which has exclusion rules that determine to which host bus adapter (HBA) worldwide node
names (WWNNs) a storage partition can be mapped.

A storage-system logical unit (LU) must not be shared between the SAN Volume Controller and a host.

You can configure certain storage controllers to safely share resources between the SAN Volume
Controller cluster and direct-attached hosts. This type of configuration is described as a split controller. In
all cases, it is critical that you configure the controller and SAN so that the SAN Volume Controller
cluster cannot access logical units (LUs) that a host or another SAN Volume Controller cluster can also
access. This split controller configuration can be arranged by controller logical unit number (LUN)
mapping and masking. If the split controller configuration is not guaranteed, data corruption can occur.

Besides a configuration where a controller is split between a SAN Volume Controller cluster and a host,
the SAN Volume Controller cluster also supports configurations where a controller is split between two
SAN Volume Controller clusters. In all cases, it is critical that you configure the controller and SAN so
that the SAN Volume Controller cluster cannot access LUs that a host or another SAN Volume Controller
cluster can also access. This can be arranged by controller LUN mapping and masking. If this is not
guaranteed, data corruption can occur.

Attention: Avoid configuring a storage system to present the same LU to more than one SAN Volume
Controller cluster. This configuration is not supported and is likely to cause undetected data loss or
corruption.

Unsupported storage systems

When a storage system is detected on the SAN, the SAN Volume Controller attempts to recognize it using its Inquiry data. If the device is not supported, the SAN Volume Controller configures the device as a generic device. A generic device might not function correctly when it is addressed by a SAN Volume Controller cluster, especially under failure scenarios. However, the SAN Volume Controller cluster does not regard accessing a generic device as an error condition and does not log an error. Managed disks (MDisks) that are presented by generic devices are not eligible to be used as quorum disks.

Split storage system configuration details

The SAN Volume Controller cluster is configured to manage LUs that are exported only by RAID storage systems. Non-RAID storage systems are not supported. If you are using SAN Volume Controller to manage solid-state drive (SSD) or other JBOD (just a bunch of disks) LUs that are presented by non-RAID storage systems, the SAN Volume Controller cluster itself does not provide RAID functions; so these LUs are exposed to data loss in the event of a disk failure.

If a single RAID storage system presents multiple LUs, either by having multiple RAID configured or by partitioning one or more RAID into multiple LUs, each LU can be owned by either the SAN Volume Controller cluster or a direct-attach host. LUN masking must also be configured, to ensure that LUs are not shared between SAN Volume Controller nodes and direct-attach hosts.

In a split storage system configuration, a storage system presents some of its LUs to a SAN Volume Controller cluster (which treats the LU as an MDisk) and the remaining LUs to another host. The SAN Volume Controller cluster presents volumes that are created from the MDisk to another host. There is no requirement for the multipathing driver for the two hosts to be the same. Figure 27 on page 85 shows that the RAID controller could be an IBM DS4000, for example, with RDAC used for pathing on the directly attached host, and SDD used on the host that is attached with the SAN Volume Controller. Hosts can simultaneously access LUs that are provided by the SAN Volume Controller cluster and directly by the device.

Note: A connection coming from a host can be either a fibre-channel or an iSCSI connection.
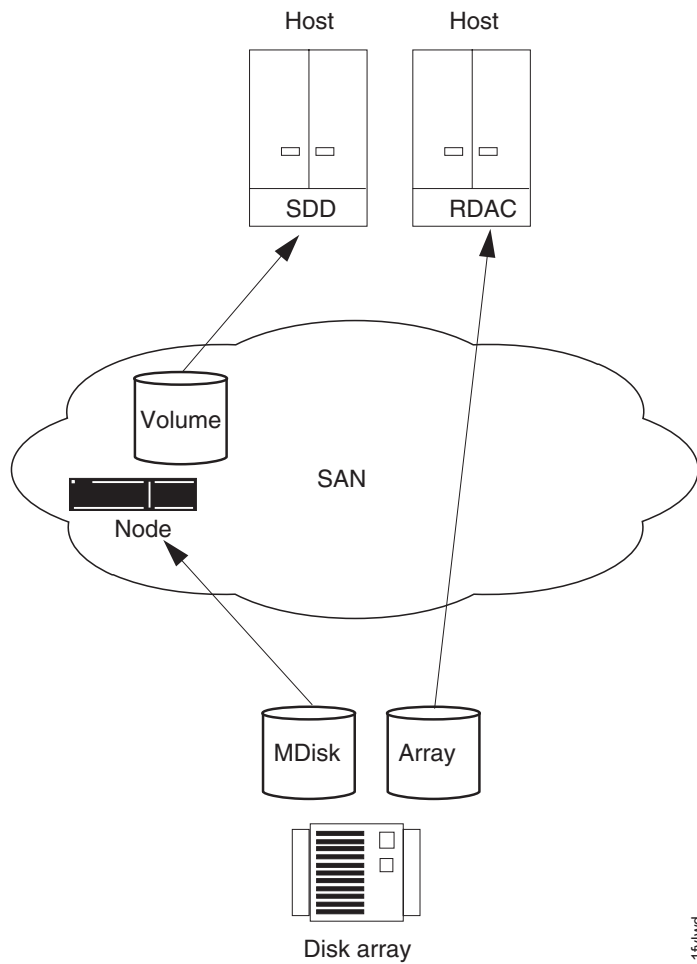
Figure 27. Storage system shared between SAN Volume Controller node and a host

It is also possible to split a host so that it accesses some of its LUNs through the SAN Volume Controller cluster and some directly. In this case, the multipathing software that is used by the storage system must be compatible with the SAN Volume Controller multipathing software. Figure 28 on page 86 is a supported configuration because the same multipathing driver is used for both directly accessed LUNs and volumes.
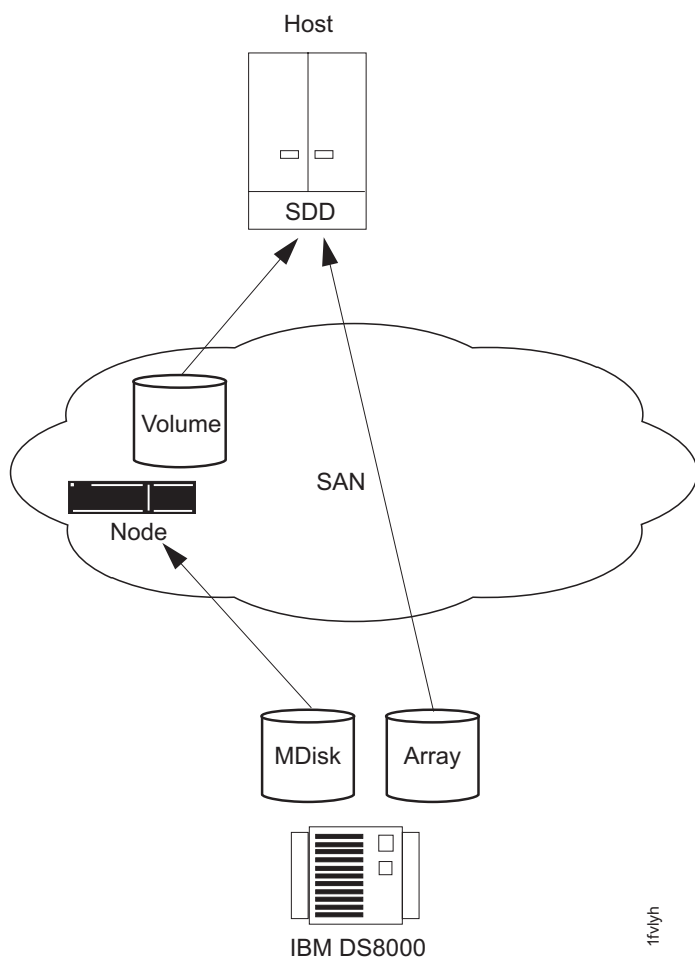
Figure 28. IBM System Storage DS8000 LUs accessed directly with a SAN Volume Controller node

In the case where the RAID storage system uses multipathing software that is compatible with SAN Volume Controller multipathing software (see Figure 29 on page 87), it is possible to configure a system where some LUNs are mapped directly to the host and others are accessed through the SAN Volume Controller. An IBM TotalStorage Enterprise Storage Server® (ESS) that uses the same multipathing driver as a SAN Volume Controller node is one example. Another example with IBM System Storage DS5000 is shown in Figure 29 on page 87.
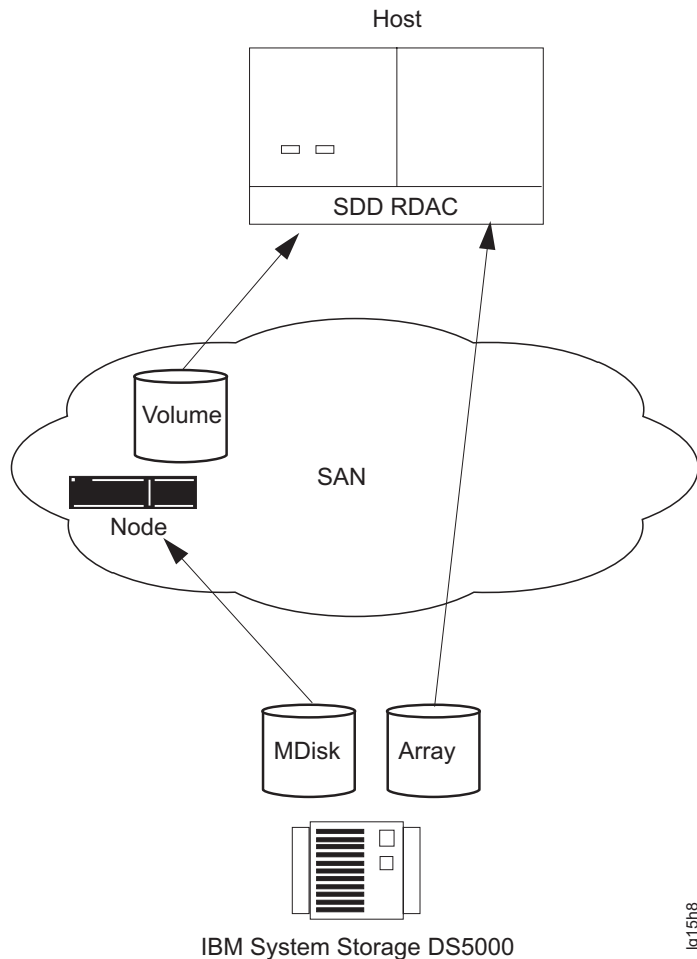
Figure 29. IBM DS5000 direct connection with a SAN Volume Controller node on one host

Fibre-channel host bus adapter configuration details

2 Apply these SAN Volume Controller configuration details to fibre-channel host bus adapters (HBAs).

t The SAN Volume Controller must be configured to export volumes only to host fibre-channel ports that are on the list of supported HBAs. See the Support for SAN Volume Controller (2145) Web site for specific firmware levels and the latest supported hardware:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Operation with other HBAs is not supported.

The SAN Volume Controller does not specify the number of host fibre-channel ports or HBAs that a host or a partition of a host can have. The number of host fibre-channel ports or HBAs are specified by the host multipathing device driver. The SAN Volume Controller supports this number; however it is subject to the configuration rules for the SAN Volume Controller. To obtain optimal performance and to prevent overloading, the workload to each SAN Volume Controller port must be equal. You can achieve an even workload by zoning approximately the same number of host fibre-channel ports to each SAN Volume Controller fibre-channel port.

The SAN Volume Controller supports configurations that use N-port virtualization in the host bus adapter or SAN switch.

iSCSI configuration details

2 You must follow these SAN Volume Controller configuration details for iSCSI host connections.

1 You can attach the SAN Volume Controller to Small Computer System Interface Over Internet Protocol
1 (iSCSI) hosts using the Ethernet ports of the SAN Volume Controller.

Note: SAN Volume Controller supports SAN devices that bridge iSCSI connections into a fibre-channel network.

iSCSI connections route from hosts to the SAN Volume Controller over the LAN. You must follow the SAN Volume Controller configuration rules for iSCSI host connections:
- SAN Volume Controller supports up to 256 iSCSI sessions per node
- SAN Volume Controller currently supports one iSCSI connection per session
- SAN Volume Controller port limits are now shared between fibre-channel WWPNs and iSCSI names

Each SAN Volume Controller node has two Ethernet ports. For each Ethernet port, a maximum of one IPv4 address and one IPv6 address can be designated for iSCSI I/O.

iSCSI hosts connect to the SAN Volume Controller through the node-port IP address. If the node fails, the address becomes unavailable and the host loses communication with SAN Volume Controller. To allow hosts to maintain access to data, the node-port IP addresses for the failed node are transferred to the partner node in the I/O group. The partner node handles requests for both its own node-port IP addresses and also for node-port IP addresses on the failed node. This process is known as node-port IP failover. In addition to node-port IP addresses, the iSCSI name and iSCSI alias for the failed node are also transferred to the partner node. After the failed node recovers, the node-port IP address and the iSCSI name and alias are returned to the original node.

Multiple configurations are supported. You can have both node Ethernet ports on the same subnet, or you can have each Ethernet port on separate subnets and use different gateways. Before you configure the Ethernet ports on separate subnets, validate that the IP configuration is correct by pinging from the iSCSI host to the nodes, and vice versa.

t A SAN Volume Controller volume can be mapped the same way either to a fibre-channel host, an iSCSI host, or both.

1 For the latest maximum configuration support information, see the IBM System Storage SAN Volume
1 Controller Web site:

1 Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

SAN Volume Controller supports the following I/O descriptions:
- I/O from different initiators in the same host to the same I/O group
t - I/O from different initiators in different hosts to the same volumes
t - I/O from fibre-channel and iSCSI initiators in different hosts to the same volumes

t I/O from fibre-channel and iSCSI initiators in the same hosts to the same volumes is not supported.

A clustered Ethernet port consists of one Ethernet port from each node in the cluster that are connected to the same Ethernet switch. Because SAN Volume Controller nodes have two Ethernet ports, two clustered Ethernet ports are possible. Ethernet configuration commands can be used for clustered Ethernet ports or node Ethernet ports. SAN Volume Controller clusters can be configured with redundant Ethernet networks.

Two types of authentication through the Challenge Handshake Authentication Protocol (CHAP) are supported:

1. One-way authentication: iSCSI target (SAN Volume Controller nodes) authenticating iSCSI initiators
2. Two-way (mutual) authentication: iSCSI target (SAN Volume Controller nodes) authenticating iSCSI initiators, and vice versa.

Attention: With the iSCSI initiator, you can set two passwords: one for discovery and another for iSCSI session I/O. However, SAN Volume Controller requires that both passwords be the same.

iSCSI protocol limitations

When using an iSCSI connection, you must consider the iSCSI protocol limitations:

- There is no SLP support for discovery.
- Header and data digest support is provided only if the initiator is configured to negotiate.
- Only one connection per session is supported.
- A maximum of 256 iSCSI sessions per SAN Volume Controller iSCSI target is supported.
- Only ErrorRecoveryLevel 0 (session restart) is supported.
- The behavior of a host that supports both fibre-channel and iSCSI connections and accesses a single volume can be unpredictable and depends on the multipathing software.
- There can be a maximum of four sessions coming from one iSCSI initiator to a SAN Volume Controller iSCSI target

The following iSCSI session parameters are supported:

```
initial_r2t = 1
immediate_data = 0
max_connections = 1
Max_recv_segment_data_length = 32k
max_xmit_data_length = 32k
max_burst_length = 32k
first_burst_length = 32k
default_wait_time = 2
default_retain_time = 20
max_outstanding_r2t = 1
data_pdu_inorder = 1
data_sequence_inorder = 1
error_recovery_level = 0
header_digest = CRC32C,None
data_digest = CRC32C,None
ofmarker = 0
ifmarker = 0
ofmarkint = 2048
ifmarkint = 2048
```

Node configuration details

Apply these configuration details to SAN Volume Controller nodes to ensure that you have a valid configuration.

Host bus adapters and nodes

SAN Volume Controller 2145-8F2 nodes contain two 2-port host bus adapters (HBAs). If one HBA fails, the node operates in degraded mode. If an HBA is physically removed, the configuration is not supported.

SAN Volume Controller 2145-CF8, SAN Volume Controller 2145-8F4, SAN Volume Controller 2145-8G4, and SAN Volume Controller 2145-8A4 nodes contain one 4-port HBA.

Volumes

t Each node presents a volume to the SAN through four ports. Each volume is accessible from the two nodes in an I/O group. Each HBA port can recognize up to eight paths to each logical unit (LU) that is presented by the cluster. The hosts must run a multipathing device driver before the multiple paths can
t resolve to a single device. You can use fabric zoning to reduce the number of paths to a volume that are visible by the host.

The number of paths through the network from an I/O group to a host must not exceed eight; configurations that exceed eight paths are not supported. Each node has four ports and each I/O group
t has two nodes. Therefore, without any zoning, the number of paths to a volume is eight multiplied by the number of host ports.

Optical connections

Valid optical connections are based on the fabric rules that the manufacturers impose for the following connection methods:

- Host to a switch
- Back end to a switch
- Interswitch links (ISLs)

Optical fiber connections can be used between a node and its switches.

Clusters that use the intercluster Metro Mirror and Global Mirror feature can use optical fiber connections between the switches, or they can use distance-extender technology that is supported by the switch manufacturer.

Ethernet connection

To ensure cluster failover operations, Ethernet port 1 on all nodes must be connected to the same set of subnets. If used, Ethernet port 2 on all nodes must also be connected to the same set of subnets. However, the subnets for Ethernet port 1 do not have to be the same as Ethernet port 2.

Physical location

The physical distance between SAN Volume Controller nodes in the same cluster is limited to 100 meters due to connectivity requirements and servicing requirements. Several of the SAN Volume Controller service actions in problem situations require that the manipulations be done to both SAN Volume Controller nodes within an I/O group or a cluster within one minute of each other. Set up your cluster environment to enable IBM service personnel to easily perform actions that are almost simultaneous in the required timeframe.

A SAN Volume Controller node must be in the same rack as the uninterruptible power supply from which it is supplied.

The depth of the SAN Volume Controller 2145-8A4 node is less than other components or nodes by approximately 127 mm or 5 inches. SAN Volume Controller 2145-8A4 nodes should not be located in the rack between components or nodes with greater depths; otherwise, it will not be possible to attach cables to a SAN Volume Controller 2145-8A4 node.

Fibre-channel connection

SAN Volume Controller supports shortwave and longwave fibre-channel connections between SAN Volume Controller nodes and the switches that they are connected to.

To avoid communication between nodes that are being routed across interswitch links (ISLs), connect all SAN Volume Controller nodes to the same fibre-channel switches.

No ISL hops are permitted among the SAN Volume Controller nodes within the same I/O group. However, one ISL hop is permitted among SAN Volume Controller nodes that are in the same cluster though different I/O groups. If your configuration requires more than one ISL hop for SAN Volume Controller nodes that are in the same cluster but in different I/O groups, contact your IBM service representative.

To avoid communication between nodes and storage systems that are being routed across ISLs, connect all storage systems to the same fibre-channel switches as the SAN Volume Controller nodes. One ISL hop between the SAN Volume Controller nodes and the storage controllers is permitted. If your configuration requires more than one ISL, contact your IBM service representative.

In larger configurations, it is common to have ISLs between host systems and the SAN Volume Controller nodes.

Port speed

The fibre-channel ports on SAN Volume Controller 2145-CF8 nodes can operate at 2 Gbps, 4 Gbps, or 8 Gbps. The fibre-channel ports on SAN Volume Controller 2145-8F4, SAN Volume Controller 2145-8G4 and SAN Volume Controller 2145-8A4 nodes can operate at 1 Gbps, 2 Gbps, or 4 Gbps. The fibre-channel ports on all these node types autonegotiate the link speed that is used with the FC switch. The ports normally operate at the maximum speed that is supported by both the SAN Volume Controller port and the switch. However, if a large number of link errors occur, the ports might operate at a lower speed than what could be supported.

Fibre-channel ports on SAN Volume Controller 2145-8F2 nodes cannot autonegotiate the speed at which they operate. You must set the required speed manually, and the optical fiber connections between the fibre-channel switches and all SAN Volume Controller 2145-8F2 nodes in a cluster must run at the same speed.

Solid-state drive configuration details

2 Apply these SAN Volume Controller configuration details for solid-state drives (SSDs).

Optional solid-state drives (SSDs) provide high-speed managed disk (MDisk) capability for SAN Volume Controller 2145-CF8 nodes. Each SAN Volume Controller 2145-CF8 node supports up to four SSDs. SSDs are local drives that are not accessible over the SAN fabric.

SSD configuration details for nodes, I/O groups, and clusters

2 You must follow these SAN Volume Controller SSD configuration details for nodes, I/O groups, and
2 clusters:
- Nodes that contain SSDs can coexist in a single SAN Volume Controller cluster with any other supported nodes.
- Quorum functionality is not supported on SSDs within SAN Volume Controller nodes.

SSD configuration details for MDisks and storage pools

2 You must follow these SAN Volume Controller SSD configuration details for MDisks and storage pools:
3 • Make an MDisk array of the SSD drives.

3 • Each storage pool that contains MDisks derived from SSDs can also contain MDisks derived from
3   hard-disk drives (HDDs). When a storage pool contains MDisks from both SSDs and HHDs, you can
3   use Easy Tier to automatically move data from highly-active HDD extents to faster responding SSD
3   extents.

SSD configuration details for volumes

3 Follow the SAN Volume Controller SSD configuration details for volumes that use storage from MDisks
3 derived from SSDs.

3 In the following rules, *SAN Volume Controller SSD storage* is a storage pool that uses MDisks derived from
  SSDs.

Note: SSD storage within SAN-attached storage systems, such as the IBM DS8000, is not subject to these
configuration rules.

t • Volumes that use SAN Volume Controller SSD storage must be created in the I/O group that the SSDs
     physically reside in.
1 • The synchronization rate must be set such that the volume copies resynchronize quickly after loss of
     synchronization. Synchronization is lost if one of the nodes goes offline either during a concurrent code
     upgrade or because of maintenance. During code upgrade, the synchronization must be restored within
t   30 minutes or the upgrade stalls. During the period that the SSD volume copies are not synchronized,
t   access to the volume depends on the single node that contains the SSD storage that is associated with
t   the synchronized volume copy. This dependency is different from volume copies from external storage
t   systems. The default synchronization rate is typically too low for SSD volume mirrors. Instead, set it to
     80 or above.

t   To increase the amount of time between the two nodes that contain volume copies and prevent the
1   nodes from going offline during an upgrade, consider manually upgrading the software.

SAN switch configuration details

2 Apply these SAN Volume Controller configuration details for fibre-channel switches to ensure that you
2 have a valid configuration.

Configuring your SAN with at least two independent switches, or networks of switches, ensures a
redundant fabric with no single point of failure. If one of the two SAN fabrics fails, the configuration is
in a degraded mode, but is still valid. A SAN with only one fabric is a valid configuration but risks loss
of access to data if the fabric fails. SANs with only one fabric are exposed to a single point of failure.

Configurations with more than four SANs are not supported.

For fibre-channel connections, the SAN Volume Controller nodes must always be connected to SAN
switches only. Each node must be connected to each of the counterpart SANs that are in the redundant
fabric. Any fibre-channel configuration that uses a direct physical connection between a host and a SAN
Volume Controller node is not supported. When attaching iSCSI hosts to SAN Volume Controller nodes,
Ethernet switches must be used.

All back-end storage systems must always be connected to SAN switches only. Multiple connections are
permitted from redundant storage systems to improve data bandwidth performance. A connection
between each redundant storage system and each counterpart SAN is not required. For example, in an
IBM System Storage DS4000 configuration in which the IBM DS4000 contains two redundant storage
systems, only two storage system minihubs are usually used. Storage system A is connected to
counterpart SAN A, and storage system B is connected to counterpart SAN B. Any configuration that
uses a direct physical connection between the SAN Volume Controller node and the storage system is not
supported.

When you attach a node to a SAN fabric that contains core directors and edge switches, connect the node ports to the core directors and connect the host ports to the edge switches. In this type of fabric, the next priority for connection to the core directors is the storage systems, leaving the host ports connected to the edge switches.

A SAN Volume Controller SAN must follow all switch manufacturer configuration rules, which might place restrictions on the configuration. Any configuration that does not follow switch manufacturer configuration rules is not supported.

Mixing manufacturer switches in a single SAN fabric

Within an individual SAN fabric, only mix switches from different vendors if the configuration is supported by the switch vendors.

Fibre-channel switches and interswitch links

The SAN Volume Controller supports distance-extender technology, including DWDM (dense wavelength division multiplexing) and FCIP (Fibre Channel over IP) extenders, to increase the overall distance between local and remote clusters. If this extender technology involves a protocol conversion, the local and remote fabrics are regarded as independent fabrics, limited to three ISL hops each.

With ISLs between nodes in the same cluster, the ISLs are considered a single point of failure. This is illustrated in Figure 30.
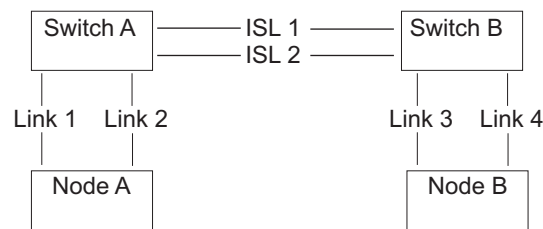
Figure 30. Fabric with ISL between nodes in a cluster

If Link 1 or Link 2 fails, the cluster communication does not fail.

If Link 3 or Link 4 fails, the cluster communication does not fail.

If ISL 1 or ISL 2 fails, the communication between Node A and Node B fails for a period of time, and the node is not recognized, even though there is still a connection between the nodes.

To ensure that a fibre-channel link failure does not cause nodes to fail when there are ISLs between nodes, it is necessary to use a redundant configuration. This is illustrated in Figure 31 on page 94.
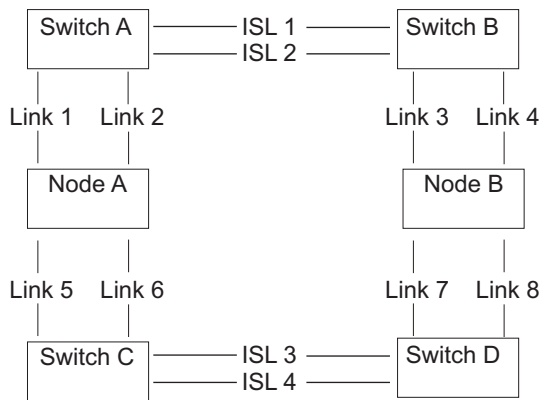
Figure 31. Fabric with ISL in a redundant configuration

With a redundant configuration, if any one of the links fails, communication on the cluster does not fail.

ISL oversubscription

Perform a thorough SAN design analysis to avoid ISL congestion. Do not configure the SAN to use SAN Volume Controller to SAN Volume Controller traffic or SAN Volume Controller to storage system traffic across ISLs that are oversubscribed. For host to SAN Volume Controller traffic, do not use an ISL oversubscription ratio that is greater than 7 to 1. Congestion on the ISLs can result in severe SAN Volume Controller performance degradation and I/O errors on the host.

When you calculate oversubscription, you must account for the speed of the links. For example, if the ISLs run at 4 Gbps and the host runs at 2 Gbps, calculate the port oversubscription as $7\times(4/2)$. In this example, the oversubscription can be 14 ports for every ISL port.

Note: The SAN Volume Controller port speed is not used in the oversubscription calculation.

SAN Volume Controller in a SAN with director class switches

You can use director class switches within the SAN to connect large numbers of RAID controllers and hosts to a SAN Volume Controller cluster. Because director class switches provide internal redundancy, one director class switch can replace a SAN that uses multiple switches. However, the director class switch provides only network redundancy; it does not protect against physical damage (for example, flood or fire), which might destroy the entire function. A tiered network of smaller switches or a core-edge topology with multiple switches in the core can provide comprehensive redundancy and more protection against physical damage for a network in a wide area. Do not use a single director class switch to provide more than one counterpart SAN because this does not constitute true redundancy.

Example SAN Volume Controller configurations

These examples show typical ways to configure your SAN Volume Controller.

Figure 32 on page 95 illustrates a small SAN configuration. Two fibre-channel switches are use to provide redundancy. Each host system, SAN Volume Controller node, and storage system is connected to both fibre-channel switches.
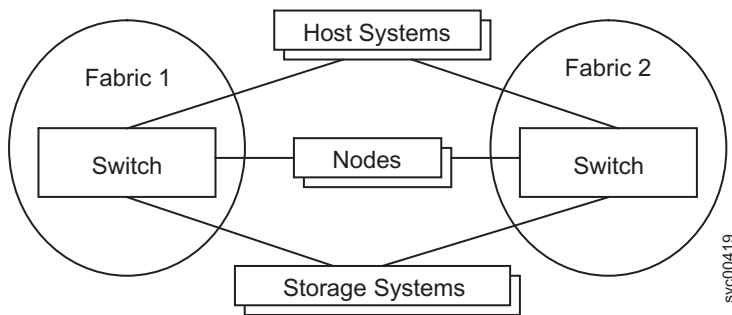
Figure 32. Simple SAN configuration

Figure 33 illustrates a medium-sized fabric that consists of three fibre-channel switches. These switches are interconnected with interswitch links (ISLs). For redundancy, use two fabrics with each host system, SAN Volume Controller node, and storage system that are being connected to both fabrics. The example fabric attaches the SAN Volume Controller nodes and the storage systems to the core switch. There are no ISL hops between SAN Volume Controller nodes or between nodes and the storage systems.
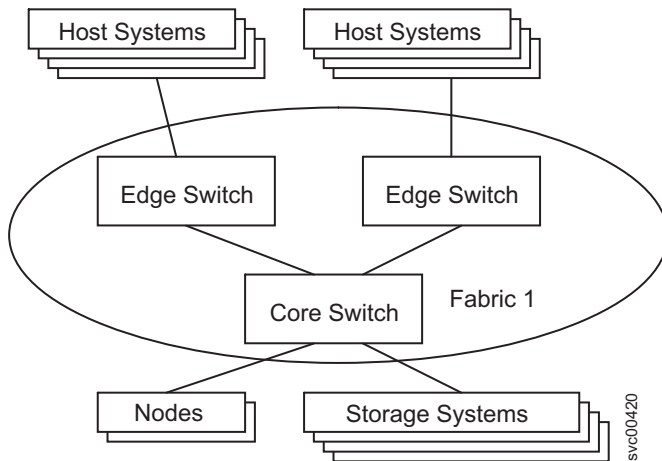
Figure 33. SAN configuration with a medium-sized fabric

Figure 34 on page 96 illustrates a large fabric that consists of two core fibre-channel switches and edge switches that are interconnected with ISLs. For redundancy, use two fabrics with each host system, SAN Volume Controller node, and storage system that is being connected. Both fabrics attach the SAN Volume Controller nodes to both core fabrics and distribute the storage systems between the two core switches. This ensures that no ISL hops exist between SAN Volume Controller nodes or between nodes and the storage systems.
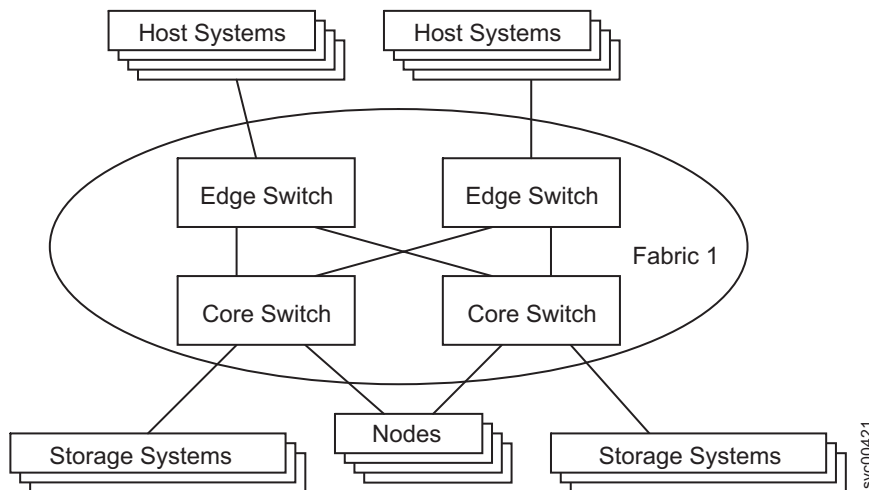
Figure 34. SAN configuration with a large fabric

Figure 35 illustrates a fabric where the host systems are located at two different sites. A long-wave optical link is used to interconnect switches at the different sites. For redundancy, use two fabrics and at least two separate long-distance links. If a large number of host systems are at the remote site, use ISL trunking to increase the available bandwidth between the two sites.
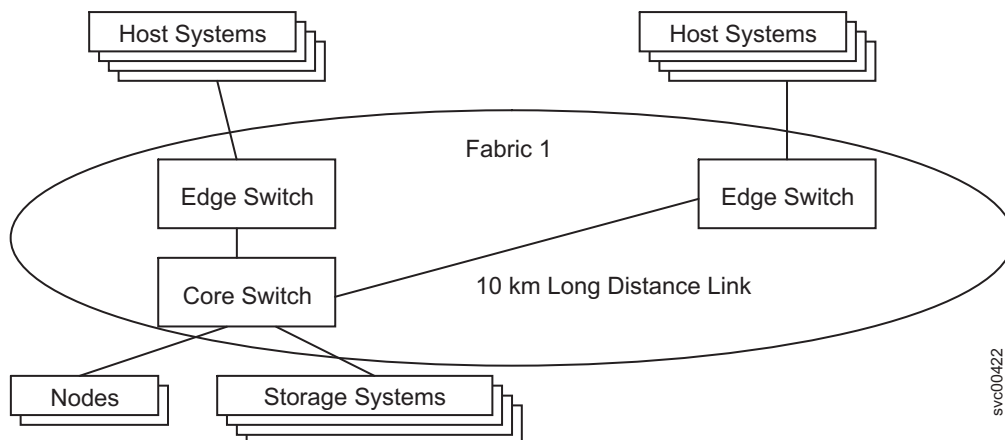
Figure 35. SAN configuration across two sites

Split-cluster configuration

For high availability, you can split a SAN Volume Controller cluster across three locations and mirror the data.

To provide protection against failures that affect an entire location, such as a power failure, you can use a configuration that splits a single SAN Volume Controller cluster across three physical locations. However, you must consider that split clusters typically exhibit substantially reduced performance.

Attention: Do not separate nodes in the same I/O group by more than 10 kilometers (6.2 miles).

You must configure a split cluster to meet the following requirements:

- Directly connect each SAN Volume Controller node to one or more SAN fabrics at the primary and secondary sites. Sites are defined as independent power domains that would fail independently. Power domains could be located in the same room or across separate physical locations.
- Use a third site to house a quorum disk.

- The storage system that provides the quorum disk at the third site must support extended quorum disks. Storage systems that provide extended quorum support are listed at the following Web site:

  Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

- Do not use powered devices to provide distance extension for the SAN Volume Controller to switch connections.
- Place independent storage systems at the primary and secondary sites, and use volume mirroring to mirror the host data between storage systems at the two sites.
- SAN Volume Controller nodes that are in the same I/O group and separated by more than 100 meters (109 yards) must use longwave fibre-channel connections. A longwave small form-factor pluggable (SFP) transceiver can be purchased as an optional SAN Volume Controller component, and must be one of the longwave SFP transceivers listed at the following website:

  Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

- Using inter-switch links (ISLs) in paths between SAN Volume Controller nodes in the same I/O group is not supported.
- Avoid using inter-switch links (ISLs) in paths between SAN Volume Controller nodes and external storage systems. If this is unavoidable, do not oversubscribe the ISLs because of substantial fibre-channel traffic across the ISLs. For most configurations, trunking is required. Because ISL problems are difficult to diagnose, switch-port error statistics must be collected and regularly monitored to detect failures.
- Using a single switch at the third site can lead to the creation of a single fabric rather than two independent and redundant fabrics. A single fabric is an unsupported configuration.
- SAN Volume Controller nodes in the same cluster must be connected to the same Ethernet subnet.
- A SAN Volume Controller node must be located in the same rack as the 2145 UPS or 2145 UPS-1U that supplies its power.
- Some service actions require physical access to all SAN Volume Controller nodes in a cluster. If nodes in a split cluster are separated by more than 100 meters, service actions might require multiple service personnel. Contact your IBM service representative to inquire about multiple site support.

A split cluster configuration locates the active quorum disk at a third site. If communication is lost between the primary and secondary sites, the site with access to the active quorum disk continues to process transactions. If communication is lost to the active quorum disk, an alternative quorum disk at another site can become the active quorum disk.

Although a cluster of SAN Volume Controller nodes can be configured to use up to three quorum disks, only one quorum disk can be elected to resolve a situation where the cluster is partitioned into two sets of nodes of equal size. The purpose of the other quorum disks is to provide redundancy if a quorum disk fails before the cluster is partitioned.

Figure 36 on page 98 illustrates an example split cluster configuration. When used in conjunction with volume mirroring, this configuration provides a high availability solution that is tolerant of a failure at a single site. If either the primary or secondary site fails, the remaining sites can continue performing I/O operations. In this configuration, the connections between SAN Volume Controller nodes in the cluster are greater than 100 meters apart, and therefore must be longwave fibre-channel connections.
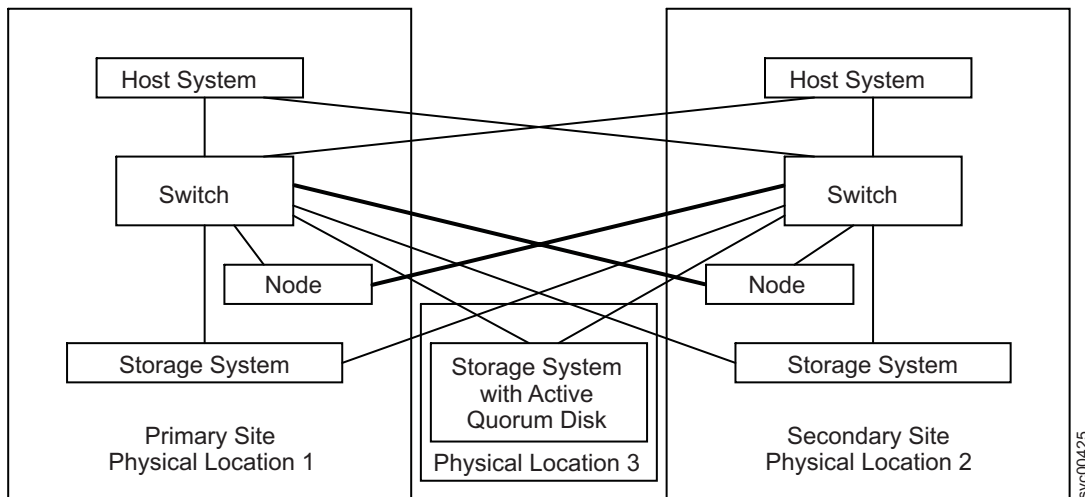
Figure 36. A split cluster with a quorum disk located at a third site

In Figure 36, the storage system that hosts the quorum disks is attached directly to a switch at both the primary and secondary sites using longwave fibre-channel connections. If either the primary site or the secondary site fails, you must ensure that the remaining site has retained direct access to the storage system that hosts the quorum disks.

An alternative configuration can use an additional fibre-channel switch at the third site with connections from that switch to the primary site and to the secondary site. This type of split-site configuration is supported only when the storage system that hosts the quorum disks supports extended quorum. Although SAN Volume Controller can use other types of storage systems for providing quorum disks, access to these quorum disks is always through a single path.

For quorum disk configuration requirements, see the *Guidance for Identifying and Changing Managed Disks Assigned as Quorum Disk Candidates* technote at the following Web site:

http://www.ibm.com/support/docview.wss?rs=591&uid=ssg1S1003311

Zoning details

2 Ensure that you are familiar with these zoning details for storage system zones and host zones.

Paths to hosts

The number of paths through the network from the SAN Volume Controller nodes to a host must not exceed eight. Configurations in which this number is exceeded are not supported.

*   Each node has four ports and each I/O group has two nodes. Therefore, without any zoning, the
t   number of paths to a volume would be eight multiplied by the number of host ports.
*   This rule exists to limit the number of paths that must be resolved by the multipathing device driver.

If you want to restrict the number of paths to a host, zone the switches so that each host bus adapter (HBA) port is zoned with one SAN Volume Controller port for each node in the cluster. If a host has multiple HBA ports, zone each port to a different set of SAN Volume Controller ports to maximize performance and redundancy.

Storage system zones

Switch zones that contain storage system ports must not have more than 40 ports. A configuration that exceeds 40 ports is not supported.

SAN Volume Controller zones

The switch fabric must be zoned so that the SAN Volume Controller nodes can detect the back-end storage systems and the front-end host HBAs. Typically, the front-end host HBAs and the back-end storage systems are not in the same zone. The exception to this is where split host and split storage system configuration is in use.

All nodes in a cluster must be able to detect the same ports on each back-end storage system. Operation in a mode where two nodes detect a different set of ports on the same storage system is degraded, and the system logs errors that request a repair action. This can occur if inappropriate zoning is applied to the fabric or if inappropriate LUN masking is used. This rule has important implications for back-end storage, such as IBM DS4000 storage systems, which impose exclusive rules for mappings between HBA worldwide node names (WWNNs) and storage partitions.

Each SAN Volume Controller port must be zoned so that it can be used for internode communications. When configuring switch zoning, you can zone some SAN Volume Controller node ports to a host or to back-end storage systems.

When configuring zones for communication between nodes in the same cluster, the minimum configuration requires that all fibre-channel ports on a node detect at least one fibre-channel port on each other node in the same cluster. You cannot reduce the configuration in this environment.

It is critical that you configure storage systems and the SAN so that a cluster cannot access logical units (LUs) that a host or another cluster can also access. You can achieve this configuration with storage system logical unit number (LUN) mapping and masking.

If a node can detect a storage system through multiple paths, use zoning to restrict communication to those paths that do not travel over ISLs.

With Metro Mirror and Global Mirror configurations, additional zones are required that contain only the local nodes and the remote nodes. It is valid for the local hosts to see the remote nodes or for the remote hosts to see the local nodes. Any zone that contains the local and the remote back-end storage systems and local nodes or remote nodes, or both, is not valid.

2 **For clusters that are running SAN Volume Controller version 5.1 or later**: For best results in Metro Mirror and Global Mirror configurations, zone each node so that it can communicate with at least one fibre-channel port on each node in each remote cluster. This configuration maintains redundancy of the fault tolerance of port and node failures within local and remote clusters. For communications between multiple SAN Volume Controller version 5.1 clusters, this also achieves optimal performance from the nodes and the intercluster links.

However, to accommodate the limitations of some switch vendors on the number of ports or worldwide node names (WWNNs) that are allowed in a zone, you can further reduce the number of ports or WWNNs in a zone. Such a reduction can result in reduced redundancy and additional workload being placed on other cluster nodes and the fibre-channel links between the nodes of a cluster.

The minimum configuration requirement is to zone both nodes in one I/O group to both nodes in one I/O group at the secondary site. The I/O group maintains fault tolerance of a node or port failure at either the local or remote site location. It does not matter which I/O groups at either site are zoned because I/O traffic can be routed through other nodes to get to the destination. However, if an I/O group that is doing the routing contains the nodes that are servicing the host I/O, there is no additional burden or latency for those I/O groups because the I/O group nodes are directly connected to the remote cluster.

For clusters that are running SAN Volume Controller version 4.3.1 or earlier, the minimum configuration requirement is that all nodes must detect at least one fibre-channel port on each node in the remote cluster. You cannot reduce the configuration in this environment.

In configurations with a version 5.1 cluster that is partnered with a cluster that is running a SAN Volume Controller version 4.3.1 or earlier, the minimum configuration requirements of the version 4.3.1 or earlier cluster apply.

If only a subset of the I/O groups within a cluster are using Metro Mirror and Global Mirror, you can restrict the zoning so that only those nodes can communicate with nodes in remote clusters. You can have nodes that are not members of any cluster zoned to detect all of the clusters. You can then add a node to the cluster in the event that you must replace a node.

Host zones

The configuration rules for host zones are different depending upon the number of hosts that will access the cluster. For configurations of less than 64 hosts per cluster, SAN Volume Controller supports a simple set of zoning rules that enable a small set of host zones to be created for different environments. For configurations of more than 64 hosts per cluster, SAN Volume Controller supports a more restrictive set of host zoning rules.

Zoning that contains host HBAs must ensure host HBAs in dissimilar hosts or dissimilar HBAs are in separate zones. Dissimilar hosts means that the hosts are running different operating systems or are different hardware platforms; thus different levels of the same operating system are regarded as similar.

To obtain the best overall performance of the system and to prevent overloading, the workload to each SAN Volume Controller port must be equal. This can typically involve zoning approximately the same number of host fibre-channel ports to each SAN Volume Controller fibre-channel port.

Clusters with less than 64 hosts

For clusters with less than 64 hosts attached, zones that contain host HBAs must contain no more than 40 initiators including the SAN Volume Controller ports that act as initiators. A configuration that exceeds 40 initiators is not supported. A valid zone can be 32 host ports plus 8 SAN Volume Controller ports. When it is possible, place each HBA port in a host that connects to a node into a separate zone. Include exactly one port from each node in the I/O groups that are associated with this host. This type of host zoning is not mandatory, but is preferred for smaller configurations.

Note: If the switch vendor recommends fewer ports per zone for a particular SAN, the rules that are imposed by the vendor takes precedence over SAN Volume Controller rules.

To obtain the best performance from a host with multiple fibre-channel ports, the zoning must ensure that each fibre-channel port of a host is zoned with a different group of SAN Volume Controller ports.

Clusters with more than 64 hosts

Each HBA port must be in a separate zone and each zone must contain exactly one port from each SAN Volume Controller node in each I/O group that the host accesses.

t **Note:** A host can be associated with more than one I/O group and therefore access volumes from
   different I/O groups in a SAN. However, this reduces the maximum number of hosts that can be used in
t the SAN. For example, if the same host uses volumes in two different I/O groups, this consumes one of
t the 256 hosts in each I/O group. If each host accesses volumes in every I/O group, there can be only 256
   hosts in the configuration.

3 Zoning examples

3 These zoning examples describe ways for zoning a switch. In the examples, a list of port names that are
3 inside brackets ([]) represent a single zone whose zone members are the list of ports shown.

Example 1

Consider the SAN environment in the following example:

- Two nodes (nodes A and B)
- Nodes A and B each have four ports
  - Node A has ports A0, A1, A2, and A3
  - Node B has ports B0, B1, B2, and B3
- Two hosts called P and Q
- Each of the two hosts has two ports, as described in Table 20.

Table 20. Two hosts and their ports

P	Q
P0	Q0
P1	Q1

- Two switches called X and Y
- Two storage systems I and J
- Each of the two storage systems has ports as described in Table 21.

Table 21. Two storage systems and their ports

I	J
I0	J0
I1	J1
I2	
I3	

The following tasks comprise an example configuration:

1. Attach half the host and node ports 1 (A0, A1, B0, B1, P0, Q0) to switch X.
2. Attach half the host and node ports 3 (A2, A3, B2, B3, P1, Q1) to switch Y.
3. Attach half the storage system ports (I0, I1, J0) to switch X.
4. Attach half the storage system ports (I2, I3, J1) to switch Y.
5. Create one zone per host port (one port per node) on switch X:

    [A0, B0, P0]

    [A1, B1, Q0]

3

Host P
• P0

• A0   • B0

Node A   Node B

• A1   • B1

Host Q
• Q0

svc00549

3
3 *Figure 37. An example of a host zone*
3

3 6. Create one storage zone per storage system on switch X:

3   [A0, A1, B0, B1, I0, I1]
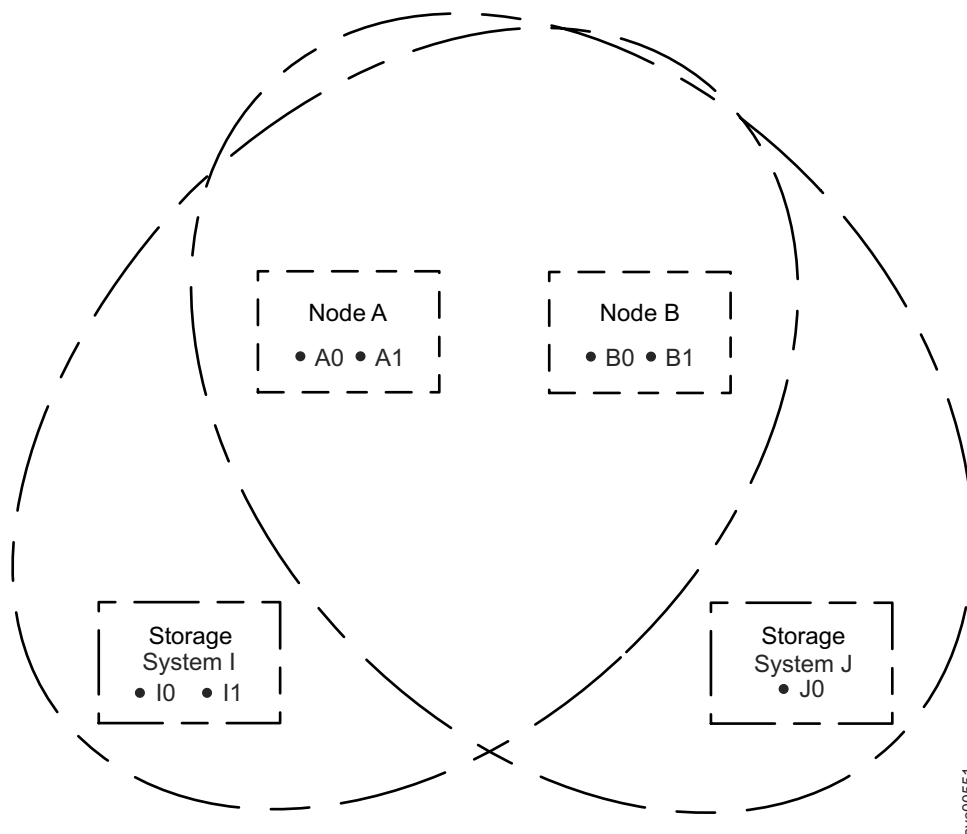
3   [A0, A1, B0, B1, J0]

3

Figure 38. An example of a storage system zone

3 7. Create one internode zone on switch X:

3 [A0, A1, B0, B1]
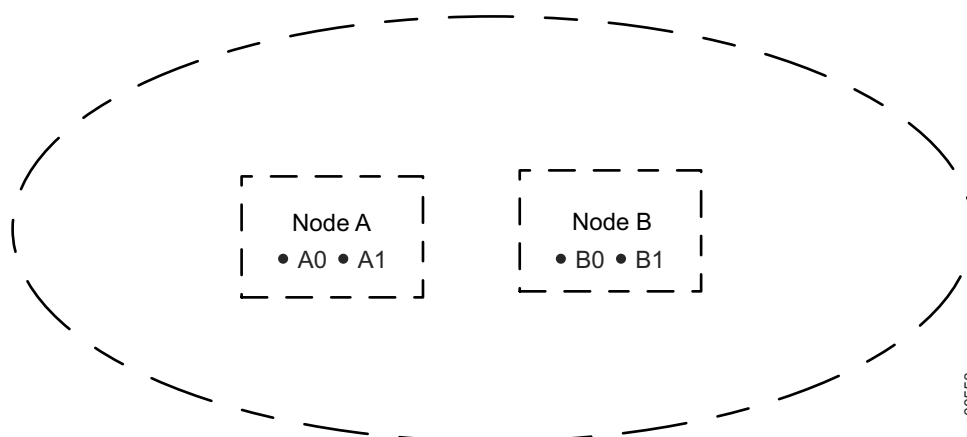
Figure 39. An example of a cluster zone

3 8. Follow the same steps 5 on page 101 through 7 to create the following list of zones for switch Y:

3 One zone per host port:

3 [A2, B2, P1]

3 [A3, B3, Q1]

3 Storage zone:

3   [A2, A3, B2, B3, I2, I3]

3   [A2, A3, B2, B3, J1]

3   One internode zone:

3   [A2, A3, B2, B3]

3 **Example 2**

3 The following example describes a SAN environment that is like the previous example except for the
3 addition of four hosts that have two ports each.

3 • Two nodes called A and B

3 • Nodes A and B have four ports each

3   – Node A has ports A0, A1, A2, and A3

3   – Node B has ports B0, B1, B2, and B3

3 • Six hosts called P, Q, R, S, T, and U

3 • Four hosts have four ports each and the other two hosts have two ports each as described in Table 22.

3   *Table 22. Six hosts and their ports*

3 **P**	**Q**	**R**	**S**	**T**	**U**
3 P0	Q0	R0	S0	T0	U0
3 P1	Q1	R1	S1	T1	U1
3 P2	Q2	R2	S2		
3 P3	Q3	R3	S3		
3					

3 • Two switches called X and Y

3 • Three storage systems

3 • Each storage system has ports as described in

3   *Table 23. Three storage systems and their ports*

3 **I**	**J**	**K**
3 I0	J0	K0
3 I1	J1	K1
3 I2		K2
3 I3		K3
3		K4
3		K5
3		K6
3		K7
3		

3 The following tasks comprise a different example configuration:

3 1. Attach half of the host and node ports 1 (A0, A1, B0, B1, P0, P1, Q0, Q1, R0, R1, S0, S1, T0, U0) to
3   switch X.

3 2. Attach half of the host and node ports 1 (A2, A3, B2, B3, P2, P3, Q2, Q3, R2, R3, S2, S3, T1, U1) to
3   switch Y.

3 3. Attach half of the storage system ports (I0, I1, J0, K0, K1, K2, K3) to switch X.

3 4. Attach half of the storage system ports (I2, I3, J1, K4, K5, K6, K7) to switch Y.

3 5. Create one zone per host port (one port per node) on switch X:

3   [A0, B0, P0]

3	[A1, B1, P1]
3	[A0, B0, Q0]
3	[A1, B1, Q1]
3	[A0, B0, R0]
3	[A1, B1, R1]
3	[A0, B0, S0]
3	[A1, B1, S1]
3	[A0, B0, T0]
3	[A1, B1, U0]

3 **Attention:** Hosts T and U (T0 and U0) and (T1 and U1) are zoned to different SAN Volume
3 Controller ports so that each SAN Volume Controller port is zoned to the same number of host ports.

3 6. Create one storage zone per storage system on switch X:

3	[A0, A1, B0, B1, I0, I1]
3	[A0, A1, B0, B1, J0]
3	[A0, A1, B0, B1, K0, K1, K2, K3]

3 7. Create one internode zone on switch X:

| 3 | [A0, A1, B0, B1] |

3 8. Follow the same steps 5 on page 104 through 7 to create the following list of zones for switch Y:

3 One zone per host port:

3	[A2, B2, P2]
3	[A3, B3, P3]
3	[A2, B2, Q2]
3	[A3, B3, Q3]
3	[A2, B2, R2]
3	[A3, B3, R3]
3	[A2, B2, S2]
3	[A3, B3, S3]
3	[A2, B2, T1]
3	[A3, B3, U1]

3 Storage zone:

3	[A2, A3, B2, B3, I2, I3]
3	[A2, A3, B2, B3, J1]
3	[A2, A3, B2, B3, K4, K5, K6, K7]

3 One internode zone:

| 3 | [A2, A3, B2, B3] |

3 ## Zoning considerations for Metro Mirror and Global Mirror

Ensure that you are familiar with the constraints for zoning a switch to support the Metro Mirror and Global Mirror feature.

SAN configurations that use intracluster Metro Mirror and Global Mirror relationships do not require additional switch zones.

For intercluster Metro Mirror and Global Mirror relationships, you must perform the following steps to create the additional zones that are required:

1. Configure your SAN so that fibre-channel traffic can be passed between the two clusters. To configure the SAN this way, you can connect the clusters to the same SAN, merge the SANs, or use routing technologies.

2. (Optional) Configure zoning to allow all nodes in the local fabric to communicate with all nodes in the remote fabric.

   Note: If you are using McData Eclipse routers, model 1620, only 64 port pairs are supported, regardless of the number of iFCP links that are used.

3. (Optional) As an alternative to step 2, choose a subset of nodes in the local cluster to be zoned to the nodes in the remote cluster. Minimally, you must ensure that one whole I/O group in the local cluster has connectivity to one whole I/O group in the remote cluster. I/O between the nodes in each cluster is then routed to find a path that is permitted by the configured zoning.

   Reducing the number of nodes that are zoned together can reduce the complexity of the intercluster zoning and might reduce the cost of the routing hardware that is required for large installations. Reducing the number of nodes also means that I/O must make extra hops between the nodes in the system, which increases the load on the intermediate nodes and can increase the performance impact; in particular, for Metro Mirror.

4. Optionally, modify the zoning so that the hosts that are visible to the local cluster can recognize the remote cluster. This allows a host to examine data in both the local and remote cluster.

5. Verify that cluster A cannot recognize any of the back-end storage that is owned by cluster B. A cluster cannot access logical units (LUs) that a host or another cluster can also access.

Switch operations over long distances

Some SAN switch products provide features that allow the users to tune the performance of I/O traffic in the fabric in a way that can affect Metro Mirror and Global Mirror performance. The two most significant features are ISL trunking and extended fabric.

2 If you are setting up long-distance links, consult the documentation from your switch vendor to ensure
2 that you set them up correctly.

Chapter 4. Creating a cluster

1 You must create a cluster to use SAN Volume Controller virtualized storage.

1 The first phase to create a cluster is performed from the front panel of the SAN Volume Controller. The
1 second phase is performed from a Web browser accessing the management GUI .

To access the CLI requires using the PuTTY client to generate Secure Shell (SSH) key pairs that secure data flow between the SAN Volume Controller cluster configuration node and a client.

1 After creating the cluster, you must configure it.

Initiating cluster creation from the front panel

After you have installed all nodes, you can use the front panel of one of the SAN Volume Controller nodes to initiate the creation of the cluster. To create a cluster, do not repeat these instructions on more than one node. After you complete the steps for initiating cluster creation from the front panel, you can use the management GUI to create the cluster and add additional nodes to complete cluster configuration.

Before you create a cluster, ensure that you verify the SAN Volume Controller installation after the installation has completed.

When you create the cluster, you must specify either an IPv4 or an IPv6 cluster address for port 1. After the cluster is created, you can specify additional IP addresses for port 1 and port 2 until both ports have an IPv4 address and an IPv6 address.

If you choose to have the IBM service representative or IBM Business Partner initially create the cluster, you must provide the following information before configuring the cluster:
- For a cluster with an IPv4 address:
  - Cluster IPv4 address
  - Subnet mask
  - Gateway IPv4 address
- For a cluster with an IPv6 address
  - Cluster IPv6 address
  - IPv6 prefix
  - Gateway IPv6 address

Define these addresses on the Configuration Data Table planning chart which is used when installing a cluster.

Attention: The Cluster IPv4 address and the IPv6 address must not be the same as any other device accessible on the network.

The IBM service representative or IBM Business Partner uses the front panel of the node to enter the
3 information that you have provided. The default cluster superuser password is `passw0rd`. The password and the IP address are used to connect to the management GUI to finalize creating the cluster.

In the following figure, bold lines indicate the select button was pressed. Lighter lines indicate the navigational path (up or down and left or right). The circled X indicates if the select button is pressed, an action occurs using the data that is entered.

Using the front panel, follow these steps to create and configure the cluster:

1. Choose a node that you want to make a member of the cluster that you are creating.

3   **Note:** You add nodes using a different process after you have successfully created and initialized the
3   cluster.
3   2. Press and release the up or down button until `Actions` is displayed.
3   3. Press and release the select button.
3   4. Depending on whether you are creating a cluster with an IPv4 address or an IPv6 address, press and
3   release the up or down button until either `Cluster IPv4?` or `Cluster IPv6?` is displayed.
5. Press and release the select button.
6. Press and release the left or right button until either `IPv4 Address:` or `IPv6 Address:` is displayed.
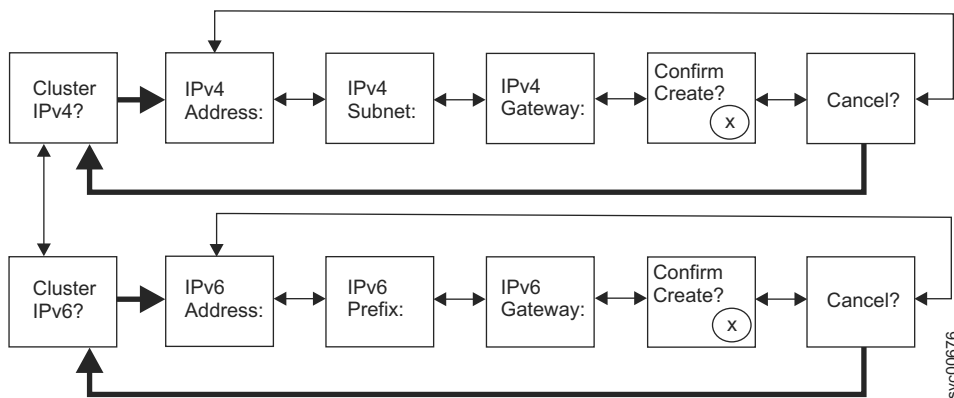
Figure 40. Cluster IPv4? and cluster IPv6? options on the front-panel display

7. Press and release the select button.

3   • If the `Cluster IPv4?` or `Cluster IPv6?` actions are not displayed, this node is already a member of a
3   cluster. Press and release the up or down button until `Actions` is displayed. Press and release the
3   select button to return to the Main Options menu. Press and release the up or down button until
3   `Cluster:` is displayed. The name of the cluster that the node belongs to is displayed on line 2 of the
1   panel. If you want to delete the node from this cluster, see the instructions for deleting a node from
1   a cluster. If you do not want to delete this node from the cluster, review the situation and
3   determine the correct nodes to include in the new cluster. Then go to step 1 and begin again.

• If you are creating a cluster with an IPv4 address and `IPv4 Address:` is displayed on line 1 of the
  panel, go to "Creating a cluster with an IPv4 address" to complete the steps for creating a cluster.

• If you are creating a cluster with an IPv6 address and `IPv6 Address:` is displayed on line 1 of the
  panel, go to "Creating a cluster with an IPv6 address" on page 110, to complete the steps for
  creating a cluster.

Creating a cluster with an IPv4 address

Your cluster IP address can be an IPv4 or IPv6 address.

The following steps provide the information to complete the task for creating a cluster with an IPv4
address.

1. The first IPv4 address number is highlighted.
2. Press the up button if you want to increase the value that is displayed; press the down button if you
   want to decrease that value. If you want to quickly increase the highlighted value, hold the up
   button. If you want to quickly decrease the highlighted value, hold the down button.

   Note: To change the address scrolling speed, see the note at the end of this topic.

3. Press the right or left buttons to move to the number field that you want to update.
4. Use the right button to move to the next field and use the up or down button to change the value of this field.
5. Repeat step 4 for each of the remaining fields of the IPv4 Address.
6. After you have changed the last field of the IPv4 Address, press the right or left button to move to the next field.
7. `IPv4 Subnet:` is displayed.
8. Use the up or down button to quickly increase or decrease the value of the first field of the IPv4 Subnet to the value that you have chosen.
9. Use the right button to move to the next field and use the up or down buttons to change the value of this field.
10. Repeat step 9 for each of the remaining fields of the IPv4 Subnet.
11. After you have changed the last field of IPv4 Subnet, press the right button to move to the next field.
12. Press the right button. `IPv4 Gateway:` is displayed.
13. Press the up button if you want to increase the value that is displayed; press the down button if you want to decrease that value. If you want to quickly increase the highlighted value, hold the up button. If you want to quickly decrease the highlighted value, hold the down button.
14. Use the right button to move to the next field and use the up or down button to change the value of this field.
15. Repeat step 14 for each of the remaining fields of the IPv4 Gateway.
16. Press and release the right button until `Confirm Create?` is displayed.
17. Ensure that you have a pen and paper to record the cluster superuser password before you create the cluster.

    Attention: The password is displayed for 60 seconds, or until the up, down, left or right button is pressed, which deletes it. You must be ready to record the password before you select one of the following actions:

    • If you want to review your settings before you create the cluster, use the right and left buttons to review those settings. Make any necessary changes, return to `Confirm Create?`, and then press the select button.

    • If the cluster is created successfully, `Password:` is displayed on line 1 of the service display screen. Line 2 contains a password that you can use to access the cluster. Record this password now.

    Important: If you do not record the password, you must restart the cluster configuration procedure. After you have recorded the password, press the up, down, left, or right button to clear the password from the screen.

After you complete this task, the following information is displayed on the service display screen:

• `Cluster:` is displayed on line 1.

• A temporary, system-assigned cluster name that is based on the IP address is displayed on line 2.

Note: To disable the fast increase and decrease address scrolling speed function using the front panel, press and hold the down arrow button, press and release the select button, and then release the down arrow button. The disabling of the fast increase and decrease function lasts until cluster creation is completed or until the feature is enabled again. If you press and hold the up or down arrow button while the function is disabled, the value increases or decreases once every 2 seconds. To enable the fast increase and decrease function again, press and hold the up arrow button, press and release the select button, and then release the up arrow button.

After you have created the cluster on the front panel with the correct IP address format, you can finish the cluster configuration by accessing the management GUI, completing the Create Cluster wizard, and adding nodes to the cluster.

Creating a cluster with an IPv6 address

Your cluster IP address can be an IPv4 or IPv6 address.

The following steps provide the information to complete the task for creating a cluster with an IPv6 address:

1. The first IPv6 address number is highlighted.
2. Press the up button if you want to increase the value that is displayed; press the down button if you want to decrease that value. If you want to quickly increase the highlighted value, hold the up button. If you want to quickly decrease the highlighted value, hold the down button.

   The IPv6 address and the IPv6 gateway address consist of eight 4-digit hexadecimal values. Enter the full address by working across a series of four panels to update each of the 4-digit hexadecimal values that make up the IPv6 addresses. The panels consist of eight fields, where each field is a 4-digit hexadecimal value.

   Note: To change the address scrolling speed, see the note at the end of this topic.
3. Press the right button or left button to move to the number field that you want to set.
4. Use the right button to move to the next field and use the up or down button to change the value of this field.
5. Repeat step 4 for each of the remaining fields of the IPv6 Address.
6. After you have changed the last field of the IPv6 Address, press the right or left button to move to the next field.
7. `IPv6 Prefix:` is displayed.
8. Use the up or down button to quickly increase or decrease the value of the first field of the IPv6 Prefix to the value that you have chosen.
9. Use the right button to move to the next field and use the up or down button to change the value of this field.
10. Repeat step9 for each of the remaining fields of the IPv6 Prefix.
11. After you have changed the last field of IPv6 Prefix, press the right or left button to move to the next field.
12. `IPv6 Gateway:` is displayed.
13. Use the up or down button to quickly increase or decrease the value of the first field of the IPv6 Gateway to the value that you have chosen.
14. Use the right button to move to the next field and use the up or down button to change the value of this field.
15. Repeat step 14 for each of the remaining fields of the IPv6 Gateway.
16. Press and release the right button until `Confirm Create?` is displayed.
17. Ensure that you have a pen and paper to record the cluster superuser password before you create the cluster.

Attention: The password is displayed for 60 seconds, or until the up, down, left or right button is pressed, which deletes it. You must be ready to record the password before you select one of the following actions:

- If you want to review your settings before you create the cluster, use the right and left buttons to review those settings. Make any necessary changes, return to `Confirm Create?`, and then press the select button.

- If the cluster is created successfully, `Password:` is displayed on line 1 of the service display panel. Line 2 contains a password that you must use when you first access the cluster. Record this password now.

  Important: If you do not record the password, you must restart the cluster configuration procedure. After you have recorded the password, press the up, down, left, or right button to clear the password from the screen.

After you complete this task, the following information is displayed on the service display screen:
- `Cluster:` is displayed on line 1.
- A temporary, system-assigned cluster name that is based on the IP address is displayed on line 2.

Note: To disable the fast increase and decrease address scrolling speed function using the front panel, press and hold the down arrow button, press and release the select button, and then release the down arrow button. The disabling of the fast increase and decrease function lasts until cluster creation is completed or until the feature is enabled again. If you press and hold the up or down arrow button while the function is disabled, the value increases or decreases once every 2 seconds. To enable the fast increase and decrease function again, press and hold the up arrow button, press and release the select button, and then release the up arrow button.

After you have created the cluster on the front panel with the correct IP address format, you can finish the cluster configuration by accessing the management GUI, completing the Create Cluster wizard, and adding nodes to the cluster.

Chapter 5. Upgrading the system

2 The system upgrade process involves the upgrading of your entire SAN Volume Controller environment.

Attention: These procedures apply to upgrading SAN Volume Controller version 6.1.0 or later. For directions on upgrading from version 5.1.x or earlier, see the software installation and configuration information at this website:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

2 Allow up to a week to plan your tasks, go through your preparatory upgrade tasks, and complete the
2 upgrade of the SAN Volume Controller environment. The upgrade procedures can be divided into these
2 general processes.

2 *Table 24. Upgrading tasks*

Sequence	Upgrade task
1	Before you upgrade, become familiar with the prerequisites and tasks involved. Decide whether you want to upgrade automatically or upgrade manually. During an automatic upgrade procedure, the cluster upgrades each of the nodes systematically. The automatic method is the preferred procedure for upgrading software on nodes. However, you can also upgrade each node manually.
2	Ensure that CIM object manager (CIMOM) clients are working correctly. When necessary, upgrade these clients so that they can support the new version of SAN Volume Controller software.
3	Ensure that multipathing drivers in the environment are fully redundant.
4	Upgrade your SAN Volume Controller.
5	Upgrade other devices in the SAN Volume Controller environment. Examples might include upgrading hosts and switches to the correct levels.
Note: The amount of time can vary depending on the amount of preparation work required and the size of the environment. For automatic upgrade, it takes about 20 minutes for each node plus 30 minutes for each cluster. The 30-minute time allows the multipathing software to recover.	

2 **Attention:** If you experience failover issues with multipathing driver support, resolve these issues before
2 you start normal operations.

2 Software and firmware for the SAN Volume Controller and its attached adapters are tested and released
2 as a single package. The package number increases each time a new release is made.

Some software levels support upgrades only from specific previous levels, or the software can be installed only on certain hardware types. If you upgrade to more than one level above your current level, you might be required to install an intermediate level. For example, if you are upgrading from level 1 to level 3, you might need to install level 2 before you can install level 3. For information about the prerequisites for each software level, see the website:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

2 **Attention:** Ensure that you have no unfixed errors in the log and that the cluster date and time are
2 correctly set. Start the Directed Maintenance Procedures (DMPs) and ensure that you fix any outstanding
2 errors before you attempt to concurrently upgrade the software.

The upgrade process

During the automatic upgrade process, each node in a cluster is upgraded one at a time, and the new code is staged on the nodes. While each node restarts, there might be some degradation in the maximum I/O rate that can be sustained by the cluster. After all the nodes in the cluster are successfully restarted with the new software level, the new software level is automatically committed.

During an automatic software upgrade, each node of a working pair is upgraded sequentially. The node that is being upgraded is temporarily unavailable and all I/O operations to that node fails. As a result, the I/O error counts increase and the failed I/O operations are directed to the partner node of the working pair. Applications do not see any I/O failures. When new nodes are added to the cluster, the software upgrade file is automatically downloaded to the new nodes from the SAN Volume Controller cluster.

The upgrade can normally be performed concurrently with normal user I/O operations. However, there is a possibility that performance could be impacted. If any restrictions apply to the operations that can be performed during the upgrade, these restrictions are documented on the SAN Volume Controller website that you use to download the software packages. During the software upgrade procedure, the majority of configuration commands are not available. Only the following SAN Volume Controller commands are operational from the time the upgrade process starts to the time that the new software level is committed, or until the process has been backed out.

- All **svcinfo** commands
- **svctask rmnode**

To determine when your software upgrade process has completed, you are notified through the management GUI. If you are using the command-line interface, issue the **svcinfo lssoftwareupgradestatus** command to display the status of the upgrade.

Because of the operational limitations that occur during the software upgrade process, the software upgrade is a user task.

Multipathing driver

Before you upgrade, ensure that the multipathing driver is fully redundant with every path available and online. You might see errors related to the paths going away (fail over) and the error count increasing during the upgrade. When the paths to the nodes are back, the nodes fall back to become a fully redundant system. After the 30-minute delay, the paths to the other node go down.

If you are using IBM Subsystem Device Driver (SDD) or IBM Subsystem Device Driver Device Specific Module (SDDDSM) as the multipathing software on the host, increased I/O error counts are displayed by the **datapath query device** or **datapath query adapter** commands to monitor the state of the multipathing software. See the *IBM System Storage Multipath Subsystem Device Driver User's Guide* for more information about the **datapath query** commands.

If you are using IBM Subsystem Device Driver Path Control Module (SDDPCM) as the multipathing software on the host, increased I/O error counts are displayed by the **pcmpath query device** or **pcmpath query adapter** commands to monitor the state of the multipathing software.

Upgrading clusters with internal solid-state drives

The SAN Volume Controller upgrade process reboots each node in the cluster in turn. Before the upgrade commences and before each node is upgraded, the upgrade process checks for dependent volumes.You can check for dependent volumes by using the **lsdependentvdisks** command-line interface (CLI) command with the **node** parameter.

2 **Upgrading clusters with internal SSDs using RAID 0**

2 The upgrade process takes each node offline temporarily to perform the upgrade. While the node
2 containing an internal SSD is offline, any data written to volumes with a mirrored copy on the
2 offline node are written only to the other online copy. After the upgraded node rejoins the cluster,
2 data is resynchronized from the copy that remained online. The upgrade process delays
2 approximately 30 minutes before starting the upgrade on the partner node. The synchronization
2 must complete within this time or the upgrade stalls and requires manual intervention. For any
2 mirrored volume that uses disk extents on an SSD that is located on a SAN Volume Controller
2 node for one or both of its volume copies, set its synchronization rate set to 80 or above to ensure
2 that the resynchronization completes in time.

3 **Note:** To increase the amount of time between the two nodes that contain volume copies and
3 prevent them from going offline during the upgrade process, consider manually upgrading the
3 software.

2 Table 25 defines the synchronization rates.

2 *Table 25. Resynchronization rates of volume copies*

2	Synchronization rate	Data copied/sec
2	1-10	128 KB
2	11-20	256 KB
2	21-30	512 KB
2	31-40	1 MB
2	41-50	2 MB
2	51-60	4 MB
2	61-70	8 MB
2	71-80	16 MB
2	81-90	32 MB
2	91-100	64 MB

2

3 **Upgrading clusters with internal SSDs using RAID 1 or 10**

3 The upgrade process takes each node offline temporarily to perform the upgrade. When a node
3 that contains an SDD drive is taken offline, the drive is removed from the array. After the
3 upgraded node rejoins the cluster, the drive is automatically added back into the array. However,
3 if this process does not complete before the node that contains a mirrored volume is taken offline,
3 the dependent volume process fails and the upgrade stalls.

3 **Attention:** To increase the amount of time between the two nodes going offline during the
3 upgrade process, consider manually upgrading the software.

2 ## Metro Mirror and Global Mirror relationships

1 When you upgrade software where the cluster participates in one or more intercluster relationships,
1 update the clusters one at a time. Do not upgrade the clusters concurrently because you can lose
1 synchronization and availability.

1 You can create new Metro Mirror or Global Mirror partnerships between clusters with different software
2 levels. If the partnerships are between a SAN Volume Controller version 6.1.0 cluster and a cluster that is
2 at 4.3.1, each cluster can participate in a single partnership with another cluster. If the clusters are all
3 either SAN Volume Controller version 6.1.0 or version 5.1.0, each cluster can participate in up to three
2 cluster partnerships. A maximum of four clusters are permitted in the same connected set. A partnership
2 cannot be formed between a SAN Volume Controller version 6.1.0 and one that is running a version that
2 is earlier than 4.3.1.

2 **Attention:** If you want to upgrade a cluster to SAN Volume Controller version 6.1.0 and the partner is
2 running version 4.3.0 or earlier, you must first upgrade the partner cluster to SAN Volume Controller
2 4.3.1 or later before you upgrade the first cluster to version 6.1.0.

2 Upgrading the software automatically

2 This automatic procedure provides a unified mechanism to upgrade the entire system in a coordinated
2 process with no user intervention.

2 This procedure is for upgrading from SAN Volume Controller version 6.1.0 or later. To upgrade from
2 version 5.1.x or earlier, see the relevant information center or publications that are available at this
2 website:

2 Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

2 Before you upgrade your software, review the conceptual information in the topic Upgrading the system
2 to understand how the upgrade process works. Allow adequate time, such as up to a week in some cases,
2 to look for potential problems or known bugs. Use the Software Upgrade Test Utility to help you find
2 these problems. You can download the most current version of this tool at the following website:

2 http://www-01.ibm.com/support/docview.wss?uid=ssg1S4000585

3 Excluding drives, when a node is rebooted as part of the cluster upgrade, the system checks that it is at
3 the correct level. If the system detects that the hardware is not running at the expected level, the system
2 does not continue to upgrade until it is safe to do so.

2 If you want to upgrade without host I/O, shut down all hosts before you start the upgrade.

2 When you are ready to upgrade, click **Configuration** → **Advanced** → **Upgrade Software** in the
2 management GUI and follow the instructions.

1 Upgrading the software manually

1 During an automatic upgrade procedure, the SAN Volume Controller cluster upgrades each of the nodes
1 systematically. The automatic method is the preferred procedure for upgrading software on nodes.
1 However, to provide more flexibility in the upgrade process, you can also upgrade each node
1 individually.

3 During this manual procedure, you remove a node from the cluster, upgrade the software on the node,
3 and return the node to the cluster. You repeat this process for the remaining nodes until the last node is
3 removed from the cluster. Every node must be upgraded to the same software level. You cannot interrupt
3 the upgrade and switch to installing a different software level. When the last node is returned to the
3 cluster, the cluster completes the upgrade and starts running the new level of software.

1 Prerequisites

1 Before you begin to upgrade nodes manually, ensure that the following requirements are met:
2 • The cluster software must be at version 6.1.0 or higher. To manually upgrade from version 4.3.1.1 or
2   5.1.x software, see the *User-paced Software Upgrade Procedure - Errata* that is included with the *IBM*
2   *System Storage SAN Volume Controller Software Installation and Configuration Guide* at this website:
2   Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145
2 • The latest SAN Volume Controller upgrade package has been downloaded.
2 • Each I/O group has two nodes.
1 • Errors in the cluster event log are addressed and marked as fixed.

1 • There are no volumes, MDisks, or storage systems with `Degraded` or `Offline` status.

1 • The service assistant IP is configured to every node in the cluster.

1 • The cluster superuser password is known.

1 • The SAN Volume Controller configuration has been backed up and saved.

1 • The latest version of the SAN Volume Controller Software Upgrade Test Utility is downloaded, is
e installed, and has been run to verify that there are no issues with the current cluster environment. You
1 can download the most current version of this tool at the following website:

1 http://www-01.ibm.com/support/docview.wss?uid=ssg1S4000585

1 • You have physical access to the hardware.

1 The following actions are not required, but are considered suggestions:

1 • Stop all Metro Mirror or Global Mirror operations during the upgrade procedure.

1 • Avoid running any FlashCopy operations during this procedure.

1 • Avoid migrating or formatting volumes during this procedure.

1 • Stop collecting IBM Tivoli Storage Productivity Center performance data for the SAN Volume
1 Controller cluster.

1 • Stop any automated jobs that access the cluster before you upgrade.

1 • Ensure that no other processes are running on the cluster before you upgrade.

2 If you want to upgrade without host I/O, shut down all hosts before you start the upgrade.

2 **Next:** "Preparing to upgrade individual nodes"

1 Preparing to upgrade individual nodes

1 Before you upgrade nodes individually, ensure that the cluster environment is ready for the upgrade.

1 Prerequisites: "Upgrading the software manually" on page 116

1 Follow these steps:

1 1. Use the management GUI to display the nodes in the cluster and record this information. For all the
1 nodes in the cluster, verify the following information:

1 • Confirm that all nodes are online.

1 • Record the name of the configuration node. This node must be upgraded last.

1 • Record the names and I/O groups that are assigned to each node.

1 2. If you are using the management GUI, view the External Storage panel to ensure that everything is
1 online and also verify that internal storage is present.

2 3. If you are using the command-line interface, issue this command for each storage system:

2 `svcinfo lscontroller controller_name_or_controller_id`

e where *controller_name_or_controller_id* is the name or ID of the storage system. Confirm that each
2 storage system does not have `degraded=no` status.

2 4. Verify that all hosts have all paths available to all the volumes that are presented to them by SAN
2 Volume Controller. Ensure that the multipathing driver is fully redundant with every single path
2 available and online.

2 5. Download the installation package for the level that you want to install. You can download the most
2 current package from the following website:

2 Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

2 6. Make a copy of the installation package and rename the file `IBM2145_REMOTE_UPGRADE`.

1 7. Follow the standard software upgrade preparation procedures to upload this package to the SAN
1   Volume Controller configuration node. This file is automatically deleted from a node that is rebooted
1   or upgraded. Repeat this step if the configuration node is rebooted or upgraded before all other nodes
1   have been upgraded.
2 8. Record the front panel ID of the node that the installation package is uploaded to.

1 **Next:** "Upgrading all nodes except the configuration node"

1 Upgrading all nodes except the configuration node

1 When upgrading nodes individually, you must upgrade all the nodes in the cluster before you upgrade
2 the configuration node. Repeat all the steps in this procedure for each node that you upgrade that is not
2 a configuration node.

1 To upgrade the nodes, follow these steps:
1 1. Ensure that all hosts have all paths available to volumes that are presented to them by the SAN
1   Volume Controller. If not, wait up to 30 minutes and repeat the check. If some paths are still
1   unavailable, investigate and resolve these connection problems before you continue the SAN Volume
1   Controller software upgrade. Ensure that the multipathing driver is fully redundant with every single
1   path available and online. You might see errors related to the paths going away and the error count
1   increasing during the upgrade.
3 2. In the management GUI, check that there are no incomplete volume synchronization tasks running.
3   Click **Home** → **System Status** and then click **Running Tasks**.
2 3. Remove the node from the cluster. From the management GUI home page, click **System Status** and
2   select the node to remove. Click **Manage** → **Remove Node**.
3 4. Verify that the node is no longer a member of the cluster: The removed node will no longer be visible
3   in the system.
3 5. Open a web browser and type `http://service_assistant_ip` in the address field. The service assistant
3   IP address is the IP address for the service assistant on the node that was just deleted.
3 6. On the service assistant home page, click **Upgrade Manually**. Click **Check for software updates** to
3   verify that you have the latest upgrade package.
3   **Attention:** It is vital that you upgrade the exact same version of software to each node.
3   .
3 7. Follow the steps in **Apply Software update**. Select the upgrade package and click **Upgrade**. You loose
3   access to service assistant as the node reboots itself. If necessary, you can access service assistant from
3   a different node.
3   The management GUI shows which node was automatically re-added. Wait for the node to come back
3   online.
3 8. When the Change node panel appears in the service assistant, click **Refresh** to add the node back in
3   to the cluster.
2 9. If you have any remaining nodes to upgrade that are not configuration nodes, repeat this task starting
2   at Step 1.

1 **Next:** "Upgrading the configuration node"

1 Upgrading the configuration node

1 After all the other nodes have been upgraded in the cluster, you can upgrade the configuration node.

1 To upgrade the configuration node, follow these steps:
1 1. Ensure that all hosts have all paths available to volumes that are mapped to those hosts. If not, wait
1   up to 30 minutes and repeat this check. If some paths are still unavailable, investigate and resolve
1   these connection problems before you continue the SAN Volume Controller software upgrade.

1 2. In the management GUI, check that there are no incomplete volume synchronization tasks running.
1   Click **Home** → **System Status** and then click **Running Tasks**.
1 3. Remove the configuration node from the cluster. From the management GUI home page, click **System**
1   **Status** and select the node to remove. Click **Manage** → **Remove Node**.

1   **Note:** When the configuration node is removed from the cluster, the SSH connection to the cluster
1   closes.
3 4. Open a web browser and type `http://service_assistant_ip` in the address field. The service assistant
3   IP address is the IP address for the service assistant on the node that was just deleted.
3 5. On the service assistant home page, click **Exit service state** and press **Go**. Use the management GUI
3   to add the node to the cluster. The node will then be upgraded before joining the cluster and will
3   remain in the adding state for some time.
3   This action automatically upgrades the software on this last node, which was the configuration node.

1 **Next:** "Completing the software upgrade"

1 Completing the software upgrade

3 After the configuration node is successfully rebooted and upgraded, verify the upgrade and return the
3 cluster to its original state by following these steps.
2 1. Verify that the cluster is running at the correct software version and that no other errors in the system
2   need to be resolved. Issue the following command:
2   `svcinfo lscluster cluster_name`

2   where *cluster_name* is the name of the cluster. Verify that the **code_level** field contains the new version
2   number for the software.
1 2. Verify that all the nodes are online. Issue this command:
1   `svcinfo lsnode`
1 3. Verify that all volumes (VDisks) are online. Issue this command:
1   `svcinfo lsvdisk`
1 4. Verify that all managed disks (MDisks) are online. Issue this command:
1   `svcinfo lsmdisk`
1 5. Restart any services, advanced functions, or scripts that were stopped before the upgrade, as required.

1 You have completed the manual software upgrade.

Performing the node rescue when the node boots

If it is necessary to replace the hard disk drive or if the software on the hard disk drive is corrupted, you
can use the node rescue procedure to reinstall the SAN Volume Controller software.

Similarly, if you have replaced the service controller, you should use the node rescue procedure to ensure
that the service controller has the correct software.

Attention: If you recently replaced both the service controller and the disk drive as part of the same
repair operation, node rescue fails.

Node rescue works by booting the operating system from the service controller and running a program
that copies all the SAN Volume Controller software from any other node that can be found on the
fibre-channel fabric.

Attention: When running node rescue operations, only run one node rescue operation on the same
SAN, at any one time. Wait for one node rescue operation to complete before starting another.

Perform the following steps to complete the node rescue:

1. Ensure that the fibre-channel cables are connected.

2. Ensure that at least one other node is connected to the fibre-channel fabric.

3. Ensure that the SAN zoning allows a connection between at least one port of this node and one port of another node. It is better if multiple ports can connect. This is particularly important if the zoning is by worldwide port name (WWPN) and you are using a new service controller. In this case, you might need to use SAN monitoring tools to determine the WWPNs of the node. If you need to change the zoning, remember to set it back when the service procedure is complete.

4. Turn off the node.

5. Press and hold the left and right buttons on the front panel.

6. Press the power button.

7. Continue to hold the left and right buttons until the node-rescue-request symbol is displayed on the front panel (Figure 41).

Figure 41. Node rescue display

The node rescue request symbol displays on the front panel display until the node starts to boot from the service controller. If the node rescue request symbol displays for more than two minutes, go to the hardware boot MAP to resolve the problem. When the node rescue starts, the service display shows the progress or failure of the node rescue operation.

Note: If the recovered node was part of a cluster, the node is now offline. Delete the offline node from the cluster and then add the node back into the cluster. If node recovery was used to recover a node that failed during a software upgrade process, it is not possible to add the node back into the cluster until the upgrade or downgrade process has completed. This can take up to four hours for an eight-node cluster.

Chapter 6. Replacing or adding nodes to an existing cluster

You can replace cluster nodes to upgrade to newer hardware models; you can also add cluster nodes to increase your cluster's workload capability.

Replacing nodes nondisruptively

You can replace SAN Volume Controller 2145-8F2, SAN Volume Controller 2145-8F4, or SAN Volume Controller 2145-8G4 nodes with SAN Volume Controller 2145-8A4 or SAN Volume Controller 2145-CF8 nodes. You can also use these procedures if you are replacing a SAN Volume Controller 2145-4F2 node with a SAN Volume Controller 2145-8A4 or earlier node.

These procedures do not include replacing a SAN Volume Controller 2145-4F2 node with a SAN Volume Controller 2145-CF8 node.

For information about replacing a SAN Volume Controller 2145-4F2 node with a SAN Volume Controller 2145-CF8 node, see *Procedures for Replacing SAN Volume Controller 2145-4F2 Nodes with SAN Volume Controller 2145-CF8 Nodes* on the following website:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

The link to the procedures is located under **Related Reading** on the page that contains the V5.1.x *IBM System Storage SAN Volume Controller Software Installation and Configuration Guide*.

3 The following procedures are nondisruptive because changes to your SAN environment are not required.
3 The replacement (new) node uses the same worldwide node name (WWNN) as the node you are
3 replacing. An alternative to this procedure is to replace nodes disruptively either by moving volumes to a
3 new I/O group or by rezoning the SAN. The disruptive procedures, however, require additional work on
3 the hosts.

This task assumes that the following conditions have been met:

- The existing cluster software must be at a version that supports the new node. If a node is being replaced by a SAN Volume Controller 2145-CF8 node, the cluster software version must be 5.1.0 or later. If a node is being replaced by a SAN Volume Controller 2145-8A4 node, the cluster software version must be 4.3.1 or later.
- All nodes that are configured in the cluster are present and online.
t - All errors in the cluster event log are addressed and marked as fixed.
t - There are no volumes, managed disks (MDisks), or controllers with a status of degraded or offline.
- The replacement node is not powered on.
- The replacement node is not connected to the SAN.
- You have a 2145 UPS-1U unit (feature code 8115) for each new SAN Volume Controller 2145-CF8 or 2145-8A4 node.
- You have backed up the cluster configuration and saved the svc.config.backup.xml file.
- The replacement node must be able to operate at the fibre-channel or Ethernet connection speed of the node it is replacing.
- If the node being replaced contains solid-state drives (SSDs), all SSDs and SAS adapters should be transferred to the new node if it supports the drives. If the new node does not support the existing SSDs, you must transfer the data off of the SSDs before replacing the node to avoid losing access to the data.

Important:

1. Do not continue this task if any of the conditions listed above are not met unless you are instructed to do so by the IBM Support Center.

2. Review all of the steps listed below before you perform this task.

3. Do not perform this task if you are not familiar with SAN Volume Controller environments or the procedures described in this task.

4. If you plan to reuse the node that you are replacing, ensure that the WWNN of the node is set to a unique number on your SAN. If you do not ensure that the WWNN is unique, the WWNN and WWPN are duplicated in the SAN environment and can cause issues.

   Tip: You can change the WWNN of the node you are replacing to the factory default WWNN of the replacement node to ensure that the number is unique.

5. The node ID and possibly the node name change during this task. After the cluster assigns the node ID, the ID cannot be changed. However, you can change the node name after this task is complete.

Perform the following steps to replace active nodes in a cluster:

3 1. (If the cluster software version is at 5.1 or later, complete this step.)

3 Confirm that no hosts have dependencies on the node.

3 When shutting down a node that is part of a cluster or when deleting the node from a cluster, you
3 can use either the management GUI or a command-line interface (CLI) command. In the
3 management GUI, click **System Status**. Click the correct node and click **Manage** to display all the
3 volumes that are dependent on a node. You can also use the **node** parameter with the **svcinfo**
3 **lsdependentvdisks** CLI command to view dependent volumes.

3 If dependent volumes exist, determine if the volumes are being used. If the volumes are being used,
3 either restore the redundant configuration or suspend the host application. If a dependent quorum
3 disk is reported, repair the access to the quorum disk or modify the quorum disk configuration.

   2. Perform the following steps to determine the cluster configuration node, and the ID, name, I/O group ID, and I/O group name for the node that you want to replace. If you already know the physical location of the node that you want to replace, you can skip this step and proceed to step 3 on page 123.

   Tip: If any of the nodes you want to replace are the cluster configuration node, replace it last.

   a. Issue the following command from the command-line interface (CLI):

      `svcinfo lsnode -delim :`

      The following is an example of the output that is displayed for this command:

      ```
      id:name:UPS_serial_number:WWNN:status:IO_group_id:IO_group_name:
      config_node:UPS_unique_id:hardware:iscsi_name:iscsi_alias
      3:dvt113294:100089J137:5005076801005A07:online:0:io_grp0:yes:
      20400002096810C7:8A4:iqn.1986-03.com.ibm:2145.ldcluster-80.dvt113294:
      14:des113004:10006BR010:5005076801004F0F:online:0:io_grp0:no:
      2040000192880040:8G4:iqn.1986-03.com.ibm:2145.ldcluster-80.des113004:
      ```

   b. In the `config_node` column, find the value `yes` and record the values in the `id` and `name` columns.

   c. Record the values in the `id` and the `name` columns for each node in the cluster.

   d. Record the values in the `IO_group_id` and the `IO_group_name` columns for each node in the cluster.

   e. Issue the following command from the CLI for each node in the cluster to determine the front panel ID:

      `svcinfo lsnodevpd node_name or node_id`

      where *node_name or node_id* is the name or ID of the node for which you want to determine the front panel ID.

f. Record the value in the `front_panel_id` column. The front panel ID is displayed on the front of each node. You can use this ID to determine the physical location of the node that matches the node ID or node name that you want replace.

3. Perform the following steps to record the WWNN or iSCSI name of the node that you want to replace:

   a. Issue the following command from the CLI:

   `svcinfo lsnode -delim :` *node_name or node_id*

   where *node_name or node_id* is the name or ID of the node for which you want to determine the WWNN or iSCSI name.

   b. Record the WWNN or iSCSI name of the node that you want to replace. Also record the order of the fibre-channel and Ethernet ports.

4. Issue the following command from the CLI to power off the node:

   `svctask stopcluster -node` *node_name*

   Important:

   a. Record and mark the order of the fibre-channel or Ethernet cables with the node port number (port 1 to 4 for fibre-channel, or port 1 to 2 for Ethernet) before you remove the cables from the back of the node. The fibre-channel ports on the back of the node are numbered 1 to 4 from left to right. You must reconnect the cables in the exact order on the replacement node to avoid issues when the replacement node is added to the cluster. If the cables are not connected in the same order, the port IDs can change, which impacts the ability of the host to access volumes. See the hardware documentation specific to your model to determine how the ports are numbered.

   b. Do not connect the replacement node to different ports on the switch or director. The SAN Volume Controller can have 4 or 8 Gbps HBAs; however, do not move them to faster switch or director ports at this time to avoid issues when the replacement node is added to the cluster.

   c. Do not move the fibre-channel cables of the node to faster or different ports on the switch or director at this time. This is a separate task that must be planned independently of replacing nodes in a cluster.

5. Issue the following CLI command to delete this node from the cluster and I/O group:

   `svctask rmnode` *node_name or node_id*

   Where *node_name or node_id* is the name or ID of the node that you want to delete. You can use the CLI to verify that the deletion process has completed.

6. Issue the following CLI command to ensure that the node is no longer a member of the cluster:

   `svcinfo lsnode`

   A list of nodes is displayed. Wait until the removed node is not listed in the command output.

7. Perform the following steps to change the WWNN or iSCSI name of the node that you just deleted from the cluster to FFFFF:

   a. Power on the node.

   b. With the Node WWNN: panel displayed, press and hold the down button, press and release the select button, and then release the down button. The display switches into edit mode. `Edit WWNN` is displayed on line 1. Line 2 of the display contains the last five numbers of the WWNN.

   c. Change the displayed number to FFFFF. To edit the highlighted number, use the up and down buttons to increase or decrease the numbers. The numbers wrap F to 0 or 0 to F. Use the left and right buttons to move between the numbers.

   d. Press the select button to save your changes and apply FFFFF as the new WWNN for the node.

8. Install the replacement node and the uninterruptible power supply in the rack and connect the uninterruptible power supply cables. See the *IBM System Storage SAN Volume Controller Model 2145-XXX Hardware Installation Guide* to determine how to connect the node and the uninterruptible power supply.

Important: Do not connect the fibre-channel or Ethernet cables during this step.

9. Power on the replacement node.

10. Record the WWNN or iSCSI name of the replacement node. You can use this name if you plan to reuse the node that you are replacing.

11. Perform the following steps to change the WWNN or iSCSI name of the replacement node to match the name that you recorded in step 3 on page 123:

For SAN Volume Controller V6.1.0 or later:

   a. With the Cluster panel displayed, press the up button twice to navigate to the `Actions` panel.

   b. Press the select button and then press the down button to navigate to the `Change WWNN?` panel.

   c. Press the select button to display the current WWNN.

   d. Press the select button to switch into edit mode. The `Edit WWNN?` panel is displayed.

   e. Change to the correct WWNN to the numbers you recorded in step 3 on page 123.

   f. Use the right button to confirm your selection. The `Confirm WWNN?` panel is displayed.

   g. Press the select button to confirm.

For SAN Volume Controller V4.3 through 5.1.0:

   a. With the Node WWNN: panel displayed, press and hold the down button, press and release the select button, and then release the down button. The display switches into edit mode. `Edit WWNN` is displayed on line 1. Line 2 of the display contains the last five numbers of the WWNN.

   b. Change the WWNN that is displayed to match the last five numbers of the WWNN that you recorded in step 3 on page 123. To edit the highlighted number, use the up and down buttons to increase or decrease the numbers. The numbers wrap F to 0 or 0 to F. Use the left and right buttons to move between the numbers.

   c. Press the select button to apply the numbers as the new WWNN for the node.

Wait one minute. If `Cluster:` is displayed on the front panel, this indicates that the node is ready to be added to the cluster. If `Cluster:` does not display, see the troubleshooting information to determine how to address this problem or contact the IBM Support Center before you continue with the next step.

12. Connect the fibre-channel or Ethernet cables to the same port numbers that you recorded for the original node in step 4 on page 123.

13. Issue the following CLI command to verify that the last five characters of the WWNN are correct:

    ```
    svcinfo lsnodecandidate
    ```

    Important: If the WWNN is not what you recorded in step 3 on page 123, you must repeat step 11.

14. Issue the following CLI command to add the node to the cluster and ensure that the node has the same name as the original node and is in the same I/O group as the original node. See the **svctask addnode** CLI command documentation for more information.

    ```
    svctask addnode -wwnodename WWNN -iogrp iogroupname/id
    ```

    where *WWNN* and *iogroupname/id* are the values that you recorded for the original node.

    The SAN Volume Controller V5.1 automatically reassigns the node with the name that was used originally. For versions prior to V5.1, use the **name** parameter with the **svctask addnode** command to assign a name. If the original node's name was automatically assigned by SAN Volume Controller, it is not possible to reuse the same name. It was automatically assigned if its name starts with `node`. In this case, either specify a different name that does not start with `node` or do not use the **name** parameter so that SAN Volume Controller automatically assigns a new name to the node.

    If necessary, the new node is updated to the same SAN Volume Controller software version as the cluster. This update can take up to 20 minutes.

    Important:

      a. Both nodes in the I/O group cache data; however, the cache sizes are asymmetric. The replacement node is limited by the cache size of the partner node in the I/O group. Therefore, it is possible that the replacement node does not utilize the full cache size until you replace the other node in the I/O group.

      b. You do not have to reconfigure the host multipathing device drivers because the replacement node uses the same WWNN and WWPN as the previous node. The multipathing device drivers should detect the recovery of paths that are available to the replacement node.

      c. The host multipathing device drivers take approximately 30 minutes to recover the paths. Do not upgrade the other node in the I/O group until for at least 30 minutes after you have successfully upgraded the first node in the I/O group. If you have other nodes in different I/O groups to upgrade, you can perform those upgrades while you wait.

15. See the documentation that is provided with your multipathing device driver for information on how to query paths to ensure that all paths have been recovered before proceeding to the next step. If you are using the IBM System Storage Multipath Subsystem Device Driver (SDD), the command to query paths is **datapath query device**.

16. Repair the faulty node.

    If you want to use the repaired node as a spare node, perform the following steps.

      a. With the Node WWNN: panel displayed, press and hold the down button, press and release the select button, and then release the down button.

      b. The display switches into edit mode. `Edit WWNN` is displayed on line 1. Line 2 of the display contains the last five numbers of the WWNN.

      c. Change the displayed number to `00000`. To edit the highlighted number, use the up and down buttons to increase or decrease the numbers. The numbers wrap F to 0 or 0 to F. Use the left and right buttons to move between the numbers.

      d. Press the select button to accept the numbers.

      This node can now be used as a spare node.

17. Repeat steps 3 on page 123 to 16 for each node that you want to replace.

Adding SAN Volume Controller nodes to an existing cluster

You can add SAN Volume Controller nodes to your cluster.

This task requires that the following conditions are met:
- All nodes that are configured in the cluster are present.
- All errors in the cluster event log are fixed.
- All managed disks (MDisks) are online.

Table 26 lists the models and software version requirements for nodes.

Table 26. Node model names and software version requirements

Node model	Required cluster SAN Volume Controller software version
SAN Volume Controller 2145-CF8	5.1.0 or later
SAN Volume Controller 2145-8A4	4.3.1 or later
SAN Volume Controller 2145-8G4	4.2.0 or later
SAN Volume Controller 2145-8F4	4.2.0 or later
SAN Volume Controller 2145-8F2	3.1.0 or later
Note: SAN Volume Controller 2145-4F2 nodes are not supported on this version or later versions.	

1. Install the SAN Volume Controller nodes and the uninterruptible power supply units in the rack.
2. Connect the SAN Volume Controller nodes to the LAN.

3. Connect the SAN Volume Controller nodes to the SAN fabric.

4. Power on the SAN Volume Controller nodes and the uninterruptible power supply units.

5. Zone the SAN Volume Controller node ports in the existing SAN Volume Controller zone. The SAN Volume Controller zone exists in each fabric with only node ports.

6. Zone the SAN Volume Controller node ports in the existing SAN Volume Controller and storage zone. A storage zone contains all of the SAN Volume Controller node ports and controller ports that are in the fabric and used to access the physical disks.

7. For each system that is used with the SAN Volume Controller cluster, use the system management application to map the LUNs that are currently used by the cluster to all of the WWPNs of the SAN Volume Controller nodes that you want to add to the cluster. The SAN Volume Controller nodes must recognize the same LUNs that the existing nodes in the cluster can recognize before they can be added to the cluster. If the SAN Volume Controller nodes cannot recognize the same LUNs, the system is marked degraded.

8. Add the SAN Volume Controller nodes to the cluster.

9. Check the status of the systems and MDisks to ensure that status has not been marked degraded. If the status is degraded, there is a configuration problem that must be resolved before any further cluster configuration tasks can be performed. If the problem cannot be resolved, remove the newly added SAN Volume Controller nodes from the cluster and contact the IBM Support Center for assistance.

Replacing a faulty node in the cluster

You can use the command-line interface (CLI) and the SAN Volume Controller front panel to replace a faulty node in the cluster.

Before you attempt to replace a faulty node with a spare node you must ensure that you meet the following requirements:

- You know the name of the cluster that contains the faulty node.
- A spare node is installed in the same rack as the cluster that contains the faulty node.
- You must make a record of the last five characters of the original worldwide node name (WWNN) of the spare node. If you repair a faulty node, and you want to make it a spare node, you can use the WWNN of the node. You do not want to duplicate the WWNN because it is unique. It is easier to swap in a node when you use the WWNN.

Attention: Never connect a node with a WWNN of 00000 to the cluster. If this node is no longer required as a spare and is to be used for normal attachment to a cluster, you must change the WWNN to the number you recorded when a spare was created. Using any other number might cause data corruption.

If a node fails, the cluster continues to operate with degraded performance until the faulty node is repaired. If the repair operation takes an unacceptable amount of time, it is useful to replace the faulty node with a spare node. However, the appropriate procedures must be followed and precautions must be taken so you do *not* interrupt I/O operations and compromise the integrity of your data.

In particular, ensure that the partner node in the I/O group is online.

- If the other node in the I/O group is offline, start Directed Maintenance Procedures (DMPs) to determine the fault.

- If you have been directed here by the DMPs, and subsequently the partner node in the I/O group has failed, see the procedure for recovering from offline volumes after a node or an I/O group failed.

- If you are replacing the node for other reasons, determine the node you want to replace and ensure that the partner node in the I/O group is online.

- If the partner node is offline, you will lose access to the volumes that belong to this I/O group. Start the DMPs and fix the other node before proceeding to the next step.

The following table describes the changes that are made to your configuration when you replace a faulty node in the cluster:

Node attributes	Description
Front panel ID	This is the number that is printed on the front of the node and is used to select the node that is added to a cluster.
Node ID	This is the ID that is assigned to the node. A new node ID is assigned each time a node is added to a cluster; the node name remains the same following service activity on the cluster. You can use the node ID or the node name to perform management tasks on the cluster. However, if you are using scripts to perform those tasks, use the node name rather than the node ID. This ID will change during this procedure.
Node name	This is the name that is assigned to the node. If you are using SAN Volume Controller version 5.1.0 nodes, the SAN Volume Controller automatically re-adds nodes that have failed back to the cluster. If the cluster reports an error for a node missing (error code 1195) and that node has been repaired and restarted, the cluster automatically re-adds the node back into the cluster. For releases prior to 5.1.0, if you do not specify a name, the SAN Volume Controller assigns a default name. The SAN Volume Controller creates a new default name each time a node is added to a cluster. If you choose to assign your own names, you must type the node name on the Adding a node to a cluster panel. You cannot manually assign a name that matches the naming convention used for names assigned automatically by SAN Volume Controller. If you are using scripts to perform management tasks on the cluster and those scripts use the node name, you can avoid the need to make changes to the scripts by assigning the original name of the node to a spare node. This name might change during this procedure.
Worldwide node name	This is the WWNN that is assigned to the node. The WWNN is used to uniquely identify the node and the fibre-channel ports. During this procedure, the WWNN of the spare node changes to that of the faulty node. The node replacement procedures must be followed exactly to avoid any duplication of WWNNs. This name does not change during this procedure.

Worldwide port names	These are the WWPNs that are assigned to the node. WWPNs are derived from the WWNN that is written to the spare node as part of this procedure. For example, if the WWNN for a node is 50050768010000F6, the four WWPNs for this node are derived as follows:
	
	`WWNN                      50050768010000F6`
	`WWNN displayed on front panel 000F6`
	`WWPN Port 1               50050768014000F6`
	`WWPN Port 2               50050768013000F6`
	`WWPN Port 3               50050768011000F6`
	`WWPN Port 4               50050768012000F6`
	
	These names do not change during this procedure.

3 3	Go to the procedure "Replacing nodes nondisruptively" on page 121 for the specific steps to replace a faulty node in the cluster.

Chapter 7. Configuring and servicing storage systems

3 To avoid performance issues, you must ensure that your SAN-attached storage systems and switches are
3 correctly configured to work with efficiently with SAN Volume Controller symmetric virtualization.

Virtualization provides many benefits over direct-attached or direct SAN-attached storage systems. However, virtualization is more susceptible to performance hot spots than direct-attached storage. Hot spots can cause I/O errors on your hosts and can potentially cause a loss of access to data.

Identifying your storage system

The serial number that is presented by the command-line interface (CLI) and the management GUI for the SAN Volume Controller is the serial number of the device.

The serial numbers can be viewed on your storage system. If the serial numbers are not displayed, the worldwide node name (WWNN) or worldwide port name (WWPN) is displayed. The WWNN or WWPN can be used to identify the different storage systems.

Configuration guidelines for storage systems

You must follow the guidelines and procedures for your storage system to maximize performance and to avoid potential I/O problems.

General guidelines

You must follow these general guidelines when configuring your storage systems.

- Avoid splitting arrays into multiple logical disks at the storage system level. Where possible, create a single logical disk from the entire capacity of the array.
- Depending on the redundancy that is required, create RAID-5 arrays using between 5 and 8 plus parity components. That is 5 + P, 6 + P, 7 + P or 8 + P.
3 - Do not mix managed disks (MDisks) that greatly vary in performance in the same storage pool tier.
t The overall storage pool performance in a tier is limited by the slowest MDisk. Because some storage systems can sustain much higher I/O bandwidths than others, do not mix MDisks that are provided by low-end storage systems with those that are provided by high-end storage systems in the same tier. You must consider the following factors:
  - The underlying RAID type that the storage system is using to implement the MDisk.
  - The number of physical disks in the RAID array and the physical disk type. For example: 10K/15K
3 rpm, FC/SATA.
3 - When possible, include similarly sized MDisks in a storage pool tier. This makes it easier to balance the
t MDisks in the storage pool tier. If the MDisks in a storage pool tier are significantly different sizes, you can balance the proportion of space that is allocated on each MDisk by including the larger MDisk
t multiple times in the MDisk list. This is specified when you create a new volume. For example, if you have two 400 MB disks and one 800 MB disk that are identified as MDisk 0, 1, and 2, you can create
t the striped volume with the MDisk IDs of 0:1:2:2. This doubles the number of extents on the 800 MB
3 drive, which accommodates it being double the size of the other MDisks.
t - Avoid leaving volumes in image mode. Only use image mode to import existing data into the cluster. To optimize the benefits of virtualization, migrate this data across the other MDisks in the group as soon as possible.
- Follow the FlashCopy feature requirements before you set up the storage. Balance the spread of the
t FlashCopy volumes across the storage pools and then across the storage systems. The I/O

t   characteristics of the application that is writing to the source volume also affects the impact that
     FlashCopy operations have on your overall I/O throughput.

   • Perform the appropriate calculations to ensure that your storage systems are configured correctly.

3 • If any storage system that is associated with an MDisk has the **allowquorum** parameter set to `no`, the
3   **chquorum** command will fail for that MDisk. Before setting the **allowquorum** parameter to `yes` on any
3   storage system, check the following website for storage system configuration requirements.

3   Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Logical disk configuration guidelines for storage systems

Most storage systems provide some mechanism to create multiple logical disks from a single array. This
is useful when the storage system presents storage directly to the hosts.

However, in a virtualized SAN, use a one-to-one mapping between arrays and logical disks so that the
t subsequent load calculations and the managed disk (MDisk) and storage pool configuration tasks are
   simplified.

Scenario: the logical disks are uneven

In this scenario, you have two RAID-5 arrays and both contain 5 + P components. Array A has a single
logical disk that is presented to the SAN Volume Controller cluster. This logical disk is seen by the cluster
as mdisk0. Array B has three logical disks that are presented to the cluster. These logical disks are seen
t by the cluster as mdisk1, mdisk2, and mdisk3. All four MDisks are assigned to the same storage pool that
t is named mdisk_grp0. When a volume is created by striping across this storage pool, array A presents the
   first extent and array B presents the next three extents. As a result, when the system reads and writes to
t the volume, the loading is split 25% on the disks in array A and 75% on the disks in array B. The
t performance of the volume is about one third of what array B can sustain.

The uneven logical disks cause performance degradation and complexity in a simple configuration. You
can avoid uneven logical disks by creating a single logical disk from each array.

RAID array configuration guidelines for storage systems

With virtualization, ensure that the storage devices are configured to provide some type of redundancy
against hard disk failures.

A failure of a storage device can affect a larger amount of storage that is presented to the hosts. To
provide redundancy, storage devices can be configured as RAID arrays that use either mirroring or parity
to protect against single failures.

When creating RAID arrays with parity protection (for example, RAID-5 arrays) consider how many
component disks you want to use in each array. If you use a large amount of disks, you can reduce the
number of disks that are required to provide availability for the same total capacity (1 per array).
However, more disks mean that it takes a longer time to rebuild a replacement disk after a disk failure,
and during this period a second disk failure causes a loss of all array data. More data is affected by a
disk failure for a larger number of member disks because performance is reduced while you rebuild onto
a hot spare (a redundant disk) and more data is exposed if a second disk fails before the rebuild
operation is complete. The smaller the number of disks, the more likely it is that write operations span an
entire stripe (stripe size, multiplied by the number of members, minus one). In this case, write
performance is improved. The number of disk drives required to provide availability can be unacceptable
if arrays are too small.

Notes:

1. For optimal performance, use arrays with between 6 and 8 member disks.

2. When creating RAID arrays with mirroring, the number of component disks in each array does not
   affect redundancy or performance.

Optimal storage pool configuration guidelines for storage systems

A storage pool provides the pool of storage from which volumes are created. You must ensure that the MDisks that make up each tier of the storage pool have the same performance and reliability characteristics.

Notes:

1. The performance of a storage pool is generally governed by the slowest MDisk in the storage pool.
2. The reliability of a storage pool is generally governed by the weakest MDisk in the storage pool.
3. If a single MDisk in a group fails, access to the entire group is lost.

Use the following guidelines when you group similar disks:

- Group equally performing MDisks in a single tier of a pool.
- Group similar arrays in a single tier. For example, configure all 6 + P RAID-5 arrays in one tier of a pool.
- Group MDisks from the same type of storage system in a single tier of a pool.
- Group MDisks that use the same type of underlying physical disk in a single tier of a pool. For example, group MDisks by fibre-channel or SATA.
- Do not use single disks. Single disks do not provide redundancy. Failure of a single disk results in total data loss of the storage pool to which it is assigned.

Scenario: Similar disks are not grouped together

Under one scenario, you could have two storage systems that are attached behind your SAN Volume Controller. One device is an IBM TotalStorage Enterprise Storage Server (ESS), which contains ten 6 + P RAID-5 arrays and MDisks 0 through 9. The other device is an IBM System Storage DS5000, which contains a single RAID-1 array, MDisk10, one single JBOD, MDisk11, and a large 15 + P RAID-5 array, MDisk12.

If you assigned MDisks 0 through 9 and MDisk11 into a single storage pool, and the JBOD MDisk11 fails, you lose access to all of the IBM ESS arrays, even though they are online. The performance is limited to the performance of the JBOD in the IBM DS5000 storage system, therefore slowing down the IBM ESS arrays.

To fix this problem, you can create three groups. The first group must contain the IBM ESS arrays, MDisks 0 through 9, the second group must contain the RAID 1 array, and the third group must contain the large RAID 5 array.

FlashCopy mapping guidelines for storage systems

Ensure that you have considered the type of I/O and frequency of update before you create the volumes that you want to use in FlashCopy mappings.

FlashCopy operations perform in direct proportion to the performance of the source and target disks. If you have a fast source disk and slow target disk, the performance of the source disk is reduced because it has to wait for the write operation to occur at the target before it can write to the source.

The FlashCopy implementation that is provided by the SAN Volume Controller copies at least 256 K every time a write is made to the source. This means that *any* write involves at minimum a read of 256 K from the source, write of the same 256 K at the target, and a write of the original change at the target. Therefore, when an application performs small 4 K writes, this is translated into 256 K.

Because of this overhead, consider the type of I/O that your application performs during a FlashCopy operation. Ensure that you do not overload the storage. The calculations contain a heavy weighting when

the FlashCopy feature is active. The weighting depends on the type of I/O that is performed. Random writes have a much higher overhead than sequential writes. For example, the sequential write would have copied the entire 256 K.

t   You can spread the FlashCopy source volumes and the FlashCopy target volumes between as many managed disk (MDisk) groups as possible. This limits the potential bottle-necking of a single storage
t   system, (assuming that the storage pools contain MDisks from different storage systems). However, this
e   can still result in potential bottlenecks if you want to maintain all your target volumes on a single storage system. You must ensure that you add the appropriate weighting to your calculations.

t # Image mode volumes and data migration guidelines for storage systems

t   Image mode volumes enable you to import and then migrate existing data that is managed by the SAN Volume Controller cluster.

t   Ensure that you follow the guidelines for using image mode volumes. This might be difficult because a configuration of logical disks and arrays that performs well in a direct SAN-attached environment can contain hot spots or hot component disks when they are connected through the cluster.

If the existing storage systems do not follow the configuration guidelines, consider stopping I/O operations on the host systems while you migrate the data into the cluster. If I/O operations are continued and the storage system does not follow the guidelines, I/O operations can fail at the hosts and ultimately loss of access to the data can occur.

t   **Attention:**   Migration commands fail if the target or source volume is offline, or if there is insufficient quorum disk space to store the metadata. Correct the offline or quorum disk condition and reissue the command.

The procedure for importing managed disks (MDisks) that contain existing data depends on the amount of free capacity you have in the cluster. You must have the same amount of free space in the cluster as the data that you want to migrate into the cluster. If you do not have this amount of available capacity,
t   the migration causes the storage pool to have an uneven distribution of data because some MDisks are more heavily loaded than others. Further migration operations are required to ensure an even distribution of data and subsequent I/O loading.

Migrating data with an equivalent amount of free capacity

To prevent managed disks (MDisks) from having an uneven distribution of data, ensure that the cluster has the same amount of free space as the data that you want to migrate.

Perform the following steps to migrate data:

1. Stop all I/O operations from the hosts. Unmap the logical disks that contain the data from the hosts.
t 2. Create one or more storage pools with free capacity. Ensure that the storage pools have enough free capacity to contain all of the migrating data, and they have balanced data distribution.
t 3. Create an empty storage pool. This temporarily contains the data that is imported.
t 4. Perform the following steps to create an image mode volume from the first unmanaged-mode MDisk that contains the data to be imported:

   a. Map one logical disk from the storage system to the SAN Volume Controller ports.
   b. Issue the **svctask detectmdisk** command-line interface (CLI) command on the cluster or use the management GUI to discover MDisks. The new unmanaged-mode MDisk that is found corresponds with the logical disk mapped in the previous step.
t c. Create an image mode volume from this unmanaged-mode MDisk and assign it to the empty
t   storage pool that you just created.
   d. Repeat steps 4a through 4c for all logical disks as required.

t   5. You can restart host applications by mapping the image mode volumes to the hosts and restarting
t   the hosts. Alternatively, you can wait until the migration is complete to map the volumes to the
       hosts.
t   6. Perform the following steps to migrate the data to the storage pools that you created in step 2 on
       page 132:
t   a. Select the first image mode volume to be migrated.
t   b. Using the **svctask migratevdisk** or **svctask migratetoimage** command, migrate this volume from
t   its current storage pool to one of the storage pools that you created in step 2 on page 132. This
          migrates all of the data from the logical disk into the new free space.
       c. Run the **svcinfo lsmigrate** command to verify that the migration has completed.
t   d. Select the next image mode volume and repeat the previous step after the migration completes.
t   7. When all the volumes have been migrated, the storage pools that you created in step 2 on page 132
t   contain the data that was on the image mode volumes. The data is striped across the new groups
       and is virtualized.
t   8. Destroy the temporary storage pool that contained the original image mode volumes.
    9. Go back to the storage system and reconfigure the old arrays and logical disks according to the
       guidelines.
   10. Add this storage back under the SAN Volume Controller and use the old storage to create new
t   volumes.

Migrating data with a smaller amount of free capacity

If the free capacity in the SAN Volume Controller cluster is smaller than the capacity of the data that is
imported, you can still migrate data.

Scenario

t Under one scenario, you could have one managed disk (MDisk) in the destination storage pool. You add
   image mode logical units from an array on the storage system and migrate these logical units to the
t destination storage pool. The logical units are then striped across the one managed-mode disk. Next, you
t add another logical unit to the destination storage pool. The MDisk now contains two managed-mode
   disks, but all of the data is on the first managed-mode disk. As a result, some of the data must be
   migrated from the overloaded managed-mode disks to the underused managed-mode disks.

t **Attention:**   The migration causes an uneven distribution of data across the MDisks in the storage pool.
t The impact of this depends on the number of MDisks that are initially in the storage pool and how many
   of these have free capacity.

t This task might require subsequent migration of data within the storage pool to balance the distribution
t of data across the MDisks in the storage pool.

   Perform the following steps to migrate data:
t 1. Select a storage pool that contains enough free capacity to migrate *all* of the logical disks on the first
      array that you want to migrate to the cluster.
t 2. Create an empty storage pool that can temporarily contain the data that is imported.
   3. Stop all I/O operations to the logical disks that you want to migrate first, and unmap these disks
      from their hosts.
t 4. Perform the following steps to create an image mode volume from the first unmanaged-mode MDisk
      that contains the data that you want to import:
      a. Map one logical disk from the storage system to the SAN Volume Controller ports.
      b. Issue the **svctask detectmdisk** command-line interface (CLI) command on the cluster or use the
         management GUI to discover MDisks. The new unmanaged-mode MDisk that is found
         corresponds with the logical disk that was mapped in the previous step.

c. Create an image mode volume from this unmanaged-mode MDisk and assign it to use the empty storage pool that you just created.

d. Repeat steps 4a on page 133 through 4c for all logical disks.

5. You can restart host applications by mapping the image mode volumes to the hosts and restarting the hosts, using the SAN Volume Controller. Alternatively, you can wait until the migration is complete to map the volumes to the hosts.

6. Perform the following steps to migrate the data into the storage pools that you created in step 1 on page 133:

   a. Select the first image mode volume that you want to migrate.

   b. Using the **svctask migratevdisk** or **svctask migratetoimage** command, migrate this volume from its current storage pool to one of the storage pools that you created in step 2 on page 133. This migrates all of the data from the logical disk into the new free space.

   c. Run the **svcinfo lsmigrate** command to verify that the migration has completed.

   d. Select the next image mode volume and repeat the previous step when the migration completes.

7. Perform the following steps to reconfigure the RAID array that contains the logical disks and add it to the storage pool that you selected in step 1 on page 133:

   a. Remove the MDisks from the temporary storage pool.

   b. At the storage system, unmap the logical disks that have been migrated from the SAN Volume Controller cluster and delete them from the array (if more than one existed).

   c. Create a single logical disk that uses the entire array capacity.

   d. Map this new logical disk to the SAN Volume Controller ports.

   e. Issue the **svctask detectmdisk** CLI command on the cluster or use the management GUI to discover MDisks. The new managed-mode MDisk that is found corresponds with the new logical disk that you created.

   f. Add this managed-mode MDisk to the storage pool that you selected in step 1 on page 133.

8. Repeat steps 3 on page 133 through 7 for the next array.

Configuring a balanced storage system

The attachment of a storage system to a SAN Volume Controller requires that specific settings are applied to the device.

There are two major steps to attaching a storage system to a SAN Volume Controller:

1. Setting the characteristics of the SAN Volume Controller to storage connections

2. Mapping logical units to these storage connections that allow the SAN Volume Controller to access the logical units

The virtualization features of the SAN Volume Controller enable you to choose how your storage is divided and presented to hosts. While virtualization provides you with a great deal of flexibility, it also offers the potential to set up an overloaded storage system. A storage system is overloaded if the quantity of I/O transactions that are issued by the host systems exceeds the capability of the storage to process those transactions. If a storage system is overloaded, it causes delays in the host systems and might cause I/O transactions to time out in the host. If I/O transactions time out, the host logs errors and I/Os fail to the applications.

Scenario: You have an overloaded storage system

Under one scenario, you have used the SAN Volume Controller to virtualize a single RAID array and to divide the storage across 64 host systems. If all host systems attempt to access the storage at the same time, the single RAID array is overloaded.

Perform the following steps to configure a balanced storage system:

1. Use Table 27 to calculate the I/O rate for each RAID array in the storage system.

   Note: The actual number of I/O operations per second that can be processed depends on the location and length of each I/O, whether the I/O is a read or a write operation and on the specifications of the component disks of the RAID array. For example, a RAID-5 array with eight component disks has an approximate I/O rate of $150 \times 7 = 1050$.

Table 27. Calculate the I/O rate

Type of RAID array	Number of component disks in the RAID array	Approximate I/O rate per second
RAID-1 (mirrored) arrays	2	300
RAID-3, RAID-4, RAID-5 (striped + parity) arrays	N+1 parity	$150 \times N$
RAID-10, RAID 0+1, RAID 1+0 (striped + mirrored) arrays	N	$150 \times N$

2. Calculate the I/O rate for a managed disk (MDisk).
   - If there is a one-to-one relationship between backend arrays and MDisks, the I/O rate for an MDisk is the same as the I/O rate of the corresponding array.
   - If an array is divided into multiple MDisks, the I/O rate per MDisk is the I/O rate of the array divided by the number of MDisks that are using the array.

3. Calculate the I/O rate for a storage pool. The I/O rate for a storage pool is the sum of the I/O rates of the MDisk that is in the storage pool. For example, a storage pool contains eight MDisks and each MDisk corresponds to a RAID-1 array. Using Table 27, the I/O rate for each MDisk is calculated as 300. The I/O rate for the storage pool is $300 \times 8 = 2400$.

4. Use Table 28 to calculate the impact of FlashCopy mappings. If you are using the FlashCopy feature that is provided by the SAN Volume Controller, you must consider the additional amount of I/O that FlashCopy operations generate because it reduces the rate at which I/O from host systems can be processed. When a FlashCopy mapping copies write I/Os from the host systems to areas of the source or target volume that are not yet copied, the SAN Volume Controller generates extra I/Os to copy the data before the write I/O is performed. The effect of using the FlashCopy feature depends on the type of I/O workload that is generated by an application.

Table 28. Calculate the impact of FlashCopy mappings

Type of application	Impact to I/O rate	Additional weighting for FlashCopy
Application is not performing I/O	Insignificant impact	0
Application is only reading data	Insignificant impact	0
Application is only issuing random writes	Up to 50 times as much I/O	49
Application is issuing random reads and writes	Up to 15 times as much I/O	14
Application is issuing sequential reads or writes	Up to 2 times as much I/O	1

For each volume that is the source or target of an active FlashCopy mapping, consider the type of application that you want to use the volume and record the additional weighting for the volume.

Example

For example, a FlashCopy mapping is used to provide point-in-time backups. During the FlashCopy process, a host application generates an I/O workload of random read and write operations to the source volume. A second host application reads the target volume and writes the data to tape to create a backup. The additional weighting for the source volume is 14. The additional weighting for the target volume is 0.

5. Calculate the I/O rate for volumes in a storage pool by performing the following steps:

   a. Calculate the number of volumes in the storage pool.

   b. Add the additional weighting for each volume that is the source or target of an active FlashCopy mapping.

   c. Divide the I/O rate of the storage pool by this number to calculate the I/O rate per volume.

Example 1

A storage pool has an I/O rate of 2400 and contains 20 volumes. There are no FlashCopy mappings. The I/O rate per volume is 2400 / 20 = 120.

Example 2

A storage pool has an I/O rate of 5000 and contains 20 volumes. There are two active FlashCopy mappings that have source volumes in the storage pool. Both source volumes are accessed by applications that issue random read and write operations. As a result, the additional weighting for each volume is 14. The I/O rate per volume is 5000 / (20 + 14 + 14) = 104.

6. Determine if the storage system is overloaded. The figure that was determined in step 4 on page 135 provides some indication of how many I/O operations per second can be processed by each volume in the storage pool.

- If you know how many I/O operations per second that your host applications generate, you can compare these figures to determine if the system is overloaded.

- If you do not know how many I/O operations per second that your host applications generate, you can use the I/O statistics facilities that are provided by the SAN Volume Controller to measure the I/O rate of your volumes, or you can use Table 29 as a guideline.

Table 29. Determine if the storage system is overloaded

Type of application	I/O rate per volume
Applications that generate a high I/O workload	200
Applications that generate a medium I/O workload	80
Applications that generate a low I/O workload	10

7. Interpret the result. If the I/O rate that is generated by the application exceeds the I/O rate per volume that you calculated, you might be overloading your storage system. You must carefully monitor the storage system to determine if the backend storage limits the overall performance of the storage system. It is also possible that the previous calculation is too simplistic to model your storage use after. For example, the calculation assumes that your applications generate the same I/O workload to all volumes, which might not be the case.

You can use the I/O statistics facilities that are provided by the SAN Volume Controller to measure the I/O rate of your MDisks. You can also use the performance and I/O statistics facilities that are provided by your storage systems.

If your storage system is overloaded there are several actions that you can take to resolve the problem:

- Add more backend storage to the system to increase the quantity of I/O that can be processed by the storage system. The SAN Volume Controller provides virtualization and data migration facilities to redistribute the I/O workload of volumes across a greater number of MDisks without having to take the storage offline.

- Stop unnecessary FlashCopy mappings to reduce the amount of I/O operations that are submitted to the backend storage. If you perform FlashCopy operations in parallel, consider reducing the amount of FlashCopy mappings that start in parallel.

- Adjust the queue depth to limit the I/O workload that is generated by a host. Depending on the type of host and type of host bus adapters (HBAs), it might be possible to limit the queue depth per volume or limit the queue depth per HBA, or both. The SAN Volume Controller also provides I/O governing features that can limit the I/O workload that is generated by hosts.

Note: Although these actions can be used to avoid I/O time-outs, performance of your storage system is still limited by the amount of storage that you have.

Storage system requirements

The performance of applications at the local cluster can be limited by the performance of the storage systems at the remote cluster.

Your setup must meet the following requirements to maximize the amount of I/O operations that applications can run on Global Mirror volumes:

- The Global Mirror volumes at the remote cluster must be in dedicated storage pools that only contain other Global Mirror volumes.
- Configure storage systems to support the Global Mirror workload that is required of them. The following guidelines can be used to fulfill this requirement:
  - Dedicate storage systems to only Global Mirror volumes
  - Configure the storage system to guarantee sufficient quality of service for the disks that are being used by Global Mirror operations
  - Ensure that physical disks are not shared between Global Mirror volumes and other I/O operations. For example, do not split an individual RAID array.
- For Global Mirror storage pools, use MDisks with the same characteristics. For example, use MDisks that have the same RAID level, physical disk count, and disk speed. This requirement is important to maintain performance when you use the Global Mirror feature.

You must provision the storage systems that are attached to the remote cluster to accommodate the following:

- The peak application workload to the Global Mirror volumes
- The specified background copy level
- All I/O operations that run on the remote cluster

Storage system requirements for FlashCopy, volume mirroring, and thin-provisioned volumes

Application performance on a local cluster can be affected by the use of FlashCopy, volume mirroring, and thin-provisioned volumes for storage systems.

The FlashCopy, volume mirroring, and thin-provisioned volume features can all have a negative impact on cluster performance. The impact depends on the type of I/O taking place, and is estimated using a weighting factor from Table 30.

A FlashCopy mapping effectively adds a number of loaded volumes to the storage pool. The effect of mirrored and thin-provisioned volumes is also estimated in Table 30. The estimates assume that thin-provisioned volumes are running at approximately 80% capacity of a fully allocated volume, and that mirrored volumes read from one copy and write to all copies.

Table 30. Performance impact estimates for FlashCopy, volume mirroring, and thin-provisioned volumes

Type of I/O (to volume)	Impact on I/O weighting	FlashCopy weighting	Volume mirroring weighting	Thin-provisioned weighting
None or minimal	Insignificant	0	0	0
Read only	Insignificant	0	0	$0.25 * S_v$
Sequential read and write	Up to 2 x I/O	$2 * F$	$C{-}V$	$0.25 * S_c$

Table 30. Performance impact estimates for FlashCopy, volume mirroring, and thin-provisioned volumes (continued)

Type of I/O (to volume)	Impact on I/O weighting	FlashCopy weighting	Volume mirroring weighting	Thin-provisioned weighting
Random read and write	Up to 15 x I/O	14 * F	C–V	0.25 * Sc
Random write	Up to 50 x I/O	49 * F	C–V	0.25 * Sc

Notes:

- In a storage pool with two FlashCopy mappings and random read/write to those volumes, the weighting factor is 14 * 2 = 28.
- In a storage pool with ten copies, five of which are primary copies of a volume, a weighting factor of 10–5 = 5 applies. If the copies are thin-provisioned, an additional weighting factor of 0.25 * 10 = 2.5 applies.

Key:

C   Number of volume copies in this MDisk Group

V   Number of volumes with their primary copy in this MDisk Group

F   Number of FlashCopy mappings affecting volumes that have copies in this MDisk Group

Sv   Number of thin-provisioned volume copies in this MDisk Group that are the primary copy of a volume

Sc   Number of thin-provisioned volume copies in this MDisk Group

To calculate the average I/O rate per volume, use the following equation:

```
I/O rate = (I/O capacity) / ( V + weighting factor for FlashCopy +
weighting factor for volume mirroring + weighting factor for thin-provisioned)
```

For example, consider 20 volumes with an I/O capacity of 5250, a FlashCopy weighting of 28, a mirroring weighting of 5, and a thin-provisioned weighting of 0.25. The I/O rate per volume is 5250 / (20 + 28 + 5 + 2.5) = 94.6. This estimate is an average I/O rate per volume; for example, half of the volumes could be running at 200 I/O operations per second (IOPs), and the other half could be running at 20 IOPs. This would not overload the system, however, because the average load is 94.6.

If the average I/O rate to the volumes in the example exceeds 94.6, the system would be overloaded. As approximate guidelines, a heavy I/O rate is 200, a medium I/O rate is 80, and a low I/O rate is 10.

With volume mirroring, a single volume can have multiple copies in different storage pools. The I/O rate for such a volume is the minimum I/O rate calculated from each of its MDisk Groups.

If system storage is overloaded, you can migrate some of the volumes to storage pools with available capacity.

Note: Solid-state drives (SSDs) are exempt from these calculations, with the exception of overall node throughput, which increases substantially for each additional SSD in the node.

Discovering logical units

The SAN Volume Controller initialization includes a process called discovery.

The discovery process systematically recognizes all visible ports on the SAN for devices that identify themselves as storage systems and the number of logical units (LUs) that they export. The LUs can contain new storage or a new path for previously discovered storage. The set of LUs forms the SAN Volume Controller managed disk (MDisk) view.

The discovery process runs when ports are added to or deleted from the SAN and when certain error conditions occur. You can also manually run the discovery process using the **svctask detectmdisk**

command-line interface (CLI) command or the **Discover MDisks** function from the management GUI. The **svctask detectmdisk** CLI command and the **Discover MDisks** function have the cluster rescan the fibre-channel network. The rescan discovers any new MDisks that might have been added to the cluster and rebalances MDisk access across the available controller device ports.

Note: Some storage systems do not automatically export LUs to the SAN Volume Controller.

Guidelines for exporting LUs

Ensure that you are familiar with the following guidelines for exporting LUs to the SAN Volume Controller:

- When you define the SAN Volume Controller as a host object to the storage systems, you must include *all* ports on *all* nodes and candidate nodes.
- When you first create an LU, you *must* wait until it is initialized before you export it to the SAN Volume Controller.

  Attention: Failure to wait for the LUs to initialize can result in excessive discovery times and an unstable view of the SAN.

- Do not present new LUs to the SAN Volume Controller until the array initialization and format is complete. If you add a LUN to a storage pool before the array initialization format is complete, the storage pool goes offline. While the storage pool is offline, you cannot access the volumes that are in the storage pool.
- When you export an LU to the SAN Volume Controller, the LU *must* be accessible through all ports on the storage system that are visible to the SAN Volume Controller.

  Important: The LU *must* be identified by the same logical unit number (LUN) on all ports.

Expanding a logical unit using the CLI

You can use the command-line interface (CLI) to expand a logical unit.

Some storage systems enable you to expand the size of a logical unit (LU) using vendor-specific disk-configuration software that is provided. The steps in this procedure are required for the SAN Volume Controller to use extra capacity that is provided in this way.

To ensure that this additional capacity is available to the SAN Volume Controller, follow these steps:

1. Issue the **svctask rmmdisk** CLI command to remove the managed disk (MDisk) from the storage pool. Use the **force** parameter to migrate data on the specified MDisk to other MDisks in the storage pool. The command completes asynchronously if `-force` is specified. You can check the progress of active migrations by running the **svcinfo lsmigrate** command.
2. Use the vendor-specific, disk-configuration software to expand the size of the logical unit on the storage system.
3. Issue the **svctask detectmdisk** CLI command to rescan the fibre-channel network. The rescan discovers any changes to existing MDisks and any new MDisks that have been added to the cluster. This command completes asynchronously and might take a few minutes. To determine whether a discovery operation is still in progress, use the **svcinfo lsdiscoverystatus** command.
4. Issue the **svcinfo lsmdisk** CLI command to display the additional capacity that has been expanded.
5. Issue the **svctask addmdisk** CLI command to add the MDisk back to the group.

The extra capacity is available for use by the SAN Volume Controller.

Modifying a logical unit mapping using the CLI

You can modify a logical unit (LU) mapping using the command-line interface (CLI).

Perform the following steps to modify an LU mapping:

1. Migrate all of the data from the managed disk (MDisk) by performing the following steps:

   a. If the MDisk is in managed mode or image mode and the volume must be kept online, issue the following CLI command and then proceed to step 2:

      ```
      svctask rmmdisk -mdisk MDisk number -force MDisk group number
      ```

      where *MDisk number* is the number of the MDisk that you want to modify and *MDisk group number* is the number of the storage pool for which you want to remove the MDisk.

      Note:

      - The volume becomes a striped MDisk *not* an image-mode volume.
      - All data that is stored on this MDisk is migrated to the other MDisks in the storage pool.
      - This CLI command can fail if there are not enough free extents in the storage pool.

   b. If the MDisk is in image mode and you do not want to convert the volume to a striped volume, stop all I/O to the image mode volume.

   c. Issue the following CLI command to remove the host mapping and any SCSI reservation that the host has on the volume:

      ```
      svctask rmvdiskhostmap -host host name virtual disk name
      ```

      Where *host name* is the name of the host for which you want to remove the volume mapping and *virtual disk name* is the name of the volume for which you want to remove mapping.

   d. Issue the following command to delete the volume:

      ```
      svctask rmvdisk virtual disk name
      ```

      Where *virtual disk name* is the name of the volume that you want to delete.

2. Remove the LU mapping on the storage system so that the LUN is not visible to the SAN Volume Controller.

3. Issue the following CLI command to clear all error counters on the MDisk:

   ```
   svctask includemdisk MDisk number
   ```

   Where *MDisk number* is the number of the MDisk that you want to modify.

4. Issue the following CLI command to rescan the fibre-channel network and detect that the LU is no longer there.

   ```
   svctask detectmdisk MDisk number
   ```

   Where *MDisk number* is the number of the MDisk that you want to modify. The MDisk is removed from the configuration.

5. Issue the following CLI command to verify that the MDisk is removed:

   ```
   svcinfo lsmdisk MDisk number
   ```

   Where *MDisk number* is the number of the MDisk that you want to modify.

   - If the MDisk is still displayed, repeat steps 3 and 4.

6. Configure the mapping of the LU to the new LUN on the storage system.

7. Issue the following CLI command:

   ```
   svctask detectmdisk
   ```

8. Issue the following CLI command to check that the MDisk now has the correct LUN:

   ```
   svcinfo lsmdisk
   ```

The MDisk has the correct LUN.

Accessing controller devices with multiple remote ports

If a managed disk (MDisk) logical unit (LU) is accessible through multiple controller device ports, SAN Volume Controller ensures that all nodes that access this LU coordinate their activity and access the LU through the same controller device port.

Monitoring LU access through multiple controller device ports

When SAN Volume Controller can access an LU through multiple controller device ports, the SAN Volume Controller uses the following criteria to determine the accessibility of these controller device ports:

- The SAN Volume Controller node is a member of a cluster.
- The SAN Volume Controller node has fibre-channel connections to the controller device port.
- The SAN Volume Controller node has successfully discovered the LU.
- Slandering has not caused the SAN Volume Controller node to exclude access to the MDisk through the controller device port.

An MDisk path is presented to the cluster for all SAN Volume Controller nodes that meet these criteria.

Controller device port selection

When an MDisk is created, SAN Volume Controller selects one of the controller device ports to access the MDisk.

Table 31 describes the algorithm that SAN Volume Controller uses to select the controller device port.

Table 31. Controller device port selection algorithm

Criteria	Description
Accessibility	Creates an initial set of candidate controller device ports. The set of candidate controller device ports include the ports that are accessible by the highest number of nodes.
Slandering	Reduces the set of candidate controller device ports to those with the lowest number of nodes.
Preference	Reduces the set of candidate controller device ports to those that the controller device uses as preferred ports.
Load balance	Selects the port from the set of candidate controller device ports that has the lowest MDisk access count.

After the initial device port selection is made for an MDisk, the following events can cause the selection algorithm to rerun:

- A new node joins the cluster and has a different view of the controller device than the other nodes in the cluster.
- The **svctask detectmdisk** command-line interface (CLI) command is run or the **Discover MDisks** management GUI function is used. The **svctask detectmdisk** CLI command and the **Discover MDisks** function have the cluster rescan the fibre-channel network. The rescan discovers any new MDisks that might have been added to the cluster and rebalances MDisk access across the available controller device ports.
- Error recovery procedures (ERPs) are started because a controller device has changed its preferred port.
- New controller device ports are discovered for the controller device that is associated with the MDisk.
- The controller device port that is currently selected becomes inaccessible.
- Slandering has caused SAN Volume Controller to exclude access to the MDisk through the controller device port.

Determining a storage system name from its SAN Volume Controller name using the CLI

You can determine a storage system name from its SAN Volume Controller name using the command-line interface (CLI).

1. Issue the following CLI command to list the storage system:

   `svcinfo lscontroller`

2. Record the name or ID for the storage system that you want to determine.

3. Issue the following CLI command:

   `svcinfo lscontroller` *controllername/id*

   where *controllername/id* is the name or ID that you recorded in step 2.

4. Record the worldwide node name (WWNN) for the device. The WWNN can be used to determine the actual storage system by launching the native user interface or using the command-line tools it provides to verify the actual storage system that has this WWNN.

Renaming a storage system using the CLI

You can use the command-line interface (CLI) to rename a storage system.

Perform the following step to rename a storage system:

Issue the **svctask chcontroller -name** *new_name controller_id* command.

Changing the configuration of an existing storage system using the CLI

You can use the command-line interface (CLI) to change the configuration of an existing storage system. You must change the configuration for a storage system when you want to delete and replace logical units (LUs).

Perform the following steps to delete existing LUs and replace them with new LUs:

1. Issue the following CLI command to delete the managed disks (MDisks) that are associated with the
t   LUs from their storage pools:

   `svctask rmmdisk -mdisk` *MDisk name1:MDisk name2* `-force` *MDisk group name*

   Where *MDisk name1:MDisk name2* are the names of the MDisks to delete.

2. Delete the existing LUs using the configuration software of the storage system.

3. Issue the following command to delete the associated MDisks from the cluster:

   `svctask detectmdisk`

4. Configure the new LUs using the configuration software of the storage system.

5. Issue the following command to add the new LUs to the cluster:

   `svctask detectmdisk`

Adding a new storage system to a running configuration using the CLI

You can add a new disk controller system to your SAN at any time using the command-line interface (CLI).

1 You must follow the zoning guidelines for your switch and also ensure that the controller (storage
1 system) is set up correctly for use with the SAN Volume Controller.

You must create one or more arrays on the new controller.

If your controller provides array partitioning, create a single partition from the entire capacity available in the array. You must record the LUN number that you assign to each partition. You must also follow the mapping guidelines (if your storage controller requires LUN mapping) to map the partitions or arrays to the SAN Volume Controller ports. You can determine the SAN Volume Controller ports by following the procedure for determining WWPNs.

Perform the following steps to add a new storage controller:

1. Issue the following CLI command to ensure that the cluster has detected the new storage (MDisks):

   svctask detectmdisk

2. Determine the storage controller name to validate that this is the correct controller. The controller is automatically assigned a default name.
   - If you are unsure which controller is presenting the MDisks, issue the following command to list the controllers:

   svcinfo lscontroller

3. Find the new controller in the list. The new controller has the highest numbered default name.

4. Record the name of the controller and follow the instructions in the section about determining a disk controller system name.

5. Issue the following command to change the controller name to something that you can easily use to identify it:

   svctask chcontroller -name *newname* *oldname*

   Where *newname* is the name that you want to change the controller to and *oldname* is the name that you are changing.

6. Issue the following command to list the unmanaged MDisks:

   svcinfo lsmdisk -filtervalue mode=unmanaged:*controller_name=new_name*

   These MDisks should correspond with the RAID arrays or partitions that you have created.

7. Record the field controller LUN number. This number corresponds with the LUN number that you assigned to each of the arrays or partitions.

t 8. Create a new MDisk group (storage pool) and add only the RAID arrays that belong to the new controller to this MDisk group. To avoid mixing RAID types, create a new MDisk group for each set of RAID array types (for example, RAID-5, RAID-1). Give each MDisk group that you create a descriptive name. For example, if your controller is named FAST650-fred, and the MDisk group contains RAID-5 arrays, name the MDisk Group F600-fred-R5.

   svctask mkmdiskgrp -ext 16 -name *mdisk_grp_name*
   -mdisk *colon separated list of RAID-x mdisks returned*
   in step 4

   This creates a new MDisk group with an extent size of 16MB.

Removing a storage system using the CLI

You can replace or decommission a storage system using the command-line interface (CLI).

During this procedure, you will add a new device, migrate data off of the storage system and remove the old MDisks.

t An alternative to following this procedure is to migrate all of the volumes that are using storage in this
e storage pool to another storage pool. Using this method, you can consolidate the volumes in a single or
t new group. However, you can only migrate one volume at a time. The procedure outlined below
   migrates all the data through a single command.

You can also use this procedure to remove or replace a single MDisk in a group. If an MDisk experiences
a partial failure, such as a degraded array, and you can still read the data from the disk but cannot write
to it, you can replace just that MDisk.

Perform the following steps to remove a storage system:

1. Add the new storage system to your cluster configuration.
2. Issue the following command:

   ```
   svctask addmdisk -mdisk mdiskx:mdisky:mdiskz... mdisk_grp_name
   ```

1 where *mdiskx:mdisky:mdiskz...* are the names of new MDisks that have a total capacity that is larger
t than the decommissioned MDisks and *mdisk_grp_name* is the name of the MDisk group (storage
t pool) that contains the MDisks that you want to decommission.
t You should now have a storage pool that you want to decommission and the new MDisks.

3. Ensure that the capacity of the new MDisks is the same or exceeds that of the old MDisks before you
   proceed to step 4.
4. Issue the following command to force delete the old MDisks from the group:

   ```
   svctask rmmdisk -force -mdisk mdiskx:mdisky:mdiskz... mdisk_grp_name
   ```

   Where *mdiskx:mdisky:mdiskz...* are the old MDisks that you want to delete and *mdisk_grp_name* is the
t name of the storage pool that contains the MDisks that you want to delete. Depending upon the
t number and size of the MDisks, and the number and size of the volumes that are using these
   MDisks, this operation takes some time to complete, even though the command returns immediately.

5. Check the progress of the migration process by issuing the following command:

   ```
   svcinfo lsmigrate
   ```

6. When all the migration tasks are complete, for example, the command in step 5 returns no output,
   verify that the MDisks are unmanaged.
7. Access the storage system and unmap the LUNs from the SAN Volume Controller ports.

   Note: You can delete the LUNs if you no longer want to preserve the data that is on the LUNs.

8. Issue the following CLI command:

   ```
   svctask detectmdisk
   ```

9. Verify that there are no MDisks for the storage system that you want decommission.
10. Remove the storage system from the SAN so that the SAN Volume Controller ports can no longer
    access the storage system.

Removing MDisks that represent unconfigured LUs using the CLI

You can use the command-line interface (CLI) to remove MDisks from the cluster.

When you remove LUs from your storage system, the managed disks (MDisks) that represent those LUs
might still exist in the cluster. However, the cluster cannot access these MDisks because the LUs that
these MDisks represent have been unconfigured or removed from the storage system. You must remove
these MDisks.

Perform the following steps to remove MDisks:

1. Run the **svctask includemdisk** CLI command on all the affected MDisks.

2. Run the **svctask rmmdisk** CLI command on all affected MDisks. This puts the MDisks into the unmanaged mode.
3. Run the **svctask detectmdisk** CLI command. The cluster detects that the MDisks no longer exist in the storage system.

All of the MDisks that represent unconfigured LUs are removed from the cluster.

Quorum disk creation and extent allocation

A quorum disk is used to resolve tie-break situations when the voting set of nodes disagree on the current state of the cluster.

The use of a quorum disk allows the cluster to manage a SAN fault that splits the cluster exactly in half. One half of the cluster continues to operate and the other half stops until the SAN connectivity is restored.

During quorum disk discovery, the system assesses each logical unit (LU) to determine its potential use as a quorum disk. From the set of eligible LUs, the system nominates three quorum candidate disks.

An LU must meet the following criteria to be considered a candidate for a quorum disk:
- It must be in managed mode.
- It must be visible to all nodes in the cluster.
- It must be presented by a storage system that is an approved host for quorum disks.
- It must have sufficient free extents to hold the cluster state and the configuration metadata.

If possible, the quorum disk candidates are presented by different devices. After the quorum candidate disks are selected, the cluster selects one of the candidate quorum disks to become the active quorum disk, which means it is used first to break a tie in the event of a cluster partition. After the active quorum disk is selected, the cluster does not attempt to ensure that the candidate quorum disks are presented by different devices. However, you can also manually select the active quorum disk if you want to ensure the active quorum disk is presented by a different device. Selecting the active quorum disk is useful in split-site cluster configurations and ensures that the most highly available quorum disk is used. You can set the **active** parameter on the **chquorum** command to set a disk as an active quorum disk. The quorum disk candidates can be updated by configuration activity if other eligible LUs are available.

To view a list of current quorum disk candidates, use the **svcinfo lsquorum** command.

If no quorum disk candidates are found after the discovery, one of the following situations has occurred:
- No LUs exist in managed space mode. An error is logged when this situation occurs.
- LUs exist in managed space mode, but they do not meet the eligibility criteria. An error is logged when this situation occurs.

Manual discovery

When you create or remove LUNs on a storage system, the managed disk (MDisk) view is not automatically updated.

You must issue the **svctask detectmdisk** command-line interface (CLI) command or use the **Discover MDisks** function from the management GUI to have the cluster rescan the fibre-channel network. The rescan discovers any new MDisks that might have been added to the cluster and rebalances MDisk access across the available controller device ports.

Servicing storage systems

Storage systems that are supported for attachment to the SAN Volume Controller system are designed with redundant components and access paths to enable concurrent maintenance. Hosts have continuous access to their data during component failure and replacement.

The following guidelines apply to all storage systems that are attached to the SAN Volume Controller system:
- Always follow the service instructions that are provided in the documentation for your storage system.
- Ensure that there are no unfixed errors in the event log before you perform any maintenance procedures.
- After you perform a maintenance procedure, check the event log and fix any errors. Expect to see the following types of errors:
  - MDisk error recovery procedures (ERPs)
  - Reduced paths

The following categories represent the types of service actions for storage systems:
- Controller code upgrade
- Field replaceable unit (FRU) replacement

Controller code upgrade

Ensure that you are familiar with the following guidelines for upgrading controller code:
- Check to see if the SAN Volume Controller supports concurrent maintenance for your storage system.
- Allow the storage system to coordinate the entire upgrade process.
- If it is not possible to allow the storage system to coordinate the entire upgrade process, perform the following steps:
  1. Reduce the storage system workload by 50%.
  2. Use the configuration tools for the storage system to manually failover all logical units (LUs) from the controller that you want to upgrade.
  3. Upgrade the controller code.
  4. Restart the controller.
  5. Manually failback the LUs to their original controller.
  6. Repeat for all controllers.

FRU replacement

Ensure that you are familiar with the following guidelines for replacing FRUs:
- If the component that you want to replace is directly in the host-side data path (for example, cable, fibre-channel port, or controller), disable the external data paths to prepare for upgrade. To disable external data paths, disconnect or disable the appropriate ports on the fabric switch. The SAN Volume Controller ERPs reroute access over the alternate path.
- If the component that you want to replace is in the internal data path (for example, cache, or drive) and did not completely fail, ensure that the data is backed up before you attempt to replace the component.
- If the component that you want to replace is not in the data path, for example, uninterruptible power supply units, fans, or batteries, the component is generally dual-redundant and can be replaced without additional steps.

Configuring IBM Storwize V7000 storage systems

Storwize V7000 storage systems can present volumes to a SAN Volume Controller. An Storwize V7000 system, however, cannot present volumes to another Storwize V7000 system.

Tasks to configure the Storwize V7000 storage system

To configure the Storwize V7000 system, follow these general tasks:

1. On the Storwize V7000 system, first define a host object and add all worldwide port names (WWPNs) from the SAN Volume Controller to it.
2. On the Storwize V7000 system, create host mappings between each volume on the Storwize V7000 system that you want to manage by using the SAN Volume Controller and the SAN Volume Controller host object that you have created.

The volumes that are presented by the Storwize V7000 system appear in the SAN Volume Controller managed disk (MDisk) view. The Storwize V7000 system appears in the SAN Volume Controller view with a vendor ID of IBM and a product ID of 2145.

Quorum disks on Storwize V7000 storage systems

The Storwize V7000 system supports quorum disks. SAN Volume Controller clusters that have an Storwize V7000 system as a storage systems can choose MDisks that are presented by an Storwize V7000 as quorum disks.

Advanced functions for Storwize V7000 storage systems

SAN Volume Controller can use storage that is presented by Storwize V7000 storage systems, but Metro Mirror and Global Mirror cannot interoperate between the two systems. On the Storwize V7000 system, you can create Metro Mirror and Global Mirror relationships only with other Storwize V7000 systems.

Volumes that are defined on an Storwize V7000 storage systems can be used by the Storwize V7000 system as a source or target for advanced copy functions such as FlashCopy, Metro Mirror, and Global Mirror. Storwize V7000 advanced copy functions are not supported for volumes that are used as SAN Volume Controller MDisks.

Sharing the Storwize V7000 system between a host and the SAN Volume Controller

An Storwize V7000 system can present some volumes to a SAN Volume Controller and other volumes to hosts on the SAN. However, an individual volume cannot be presented to both a SAN Volume Controller and a host simultaneously.

SAN zoning for Storwize V7000 storage systems

If a large number of Storwize V7000 storage systems are managed by a single SAN Volume Controller, be aware of the published limit for maximum Storwize V7000 storage systems per fabric. If the limit is exceeded by as little as a single storage system, divide the Storwize V7000 systems into multiple zones. For details, see this website:

Support for Storwize V7000 (2076) website at www.ibm.com/storage/support/2076

Configuring Bull FDA systems

This section provides information about configuring Bull StoreWay FDA systems for attachment to a SAN Volume Controller.

Supported firmware levels for the Bull FDA

The Bull FDA system must use a firmware level that is supported by the SAN Volume Controller.

See the following Web site for specific firmware levels and the latest supported hardware:Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Logical unit creation and deletion for Bull FDA

You can create or delete logical units for the Bull FDA. See the storage configuration guidelines that are specified in the Bull FDA documentation that is provided for this system.

Platform type for Bull FDA

You must set all logical units that the SAN Volume Controller accesses to platform type AX (AIX).

Access control methods for Bull FDA

You can use access control to restrict access from hosts and SAN Volume Controller clusters. You do not need to use access control to allow a SAN Volume Controller cluster to use all of the defined logical units on the system.

The following table lists the access control methods that are available:

Method	Description
Port Mode	Allows access to logical units that you want to define on a per storage controller port basis. SAN Volume Controller visibility (through switch zoning, physical cabling, etc.) must allow the SAN Volume Controller cluster to have the same access from all nodes and the accessible controller ports have been assigned the same set of logical units with the same logical unit number. This method of access control is not recommended for SAN Volume Controller connection.
WWN Mode	Allows access to logical units using the WWPN of each of the ports of an accessing host device. All WWPNs of all the SAN Volume Controller nodes in the same cluster must be added to the list of linked paths in the controller configuration. This becomes the list of host (SAN Volume Controller) ports for an LD Set or group of logical units. This method of access control allows sharing because different logical units can be accessed by other hosts.

Setting cache allocations for Bull FDA

Cache allocations can be set manually; however, changes to the default settings can adversely effect performance and cause you to lose access to the system.

Snapshot Volume and Link Volume for Bull FDA

You cannot use Copy Services logical volumes with logical units that are assigned to the SAN Volume Controller.

Configuring EMC CLARiiON systems

This section provides information about configuring the EMC CLARiiON storage system for attachment to a SAN Volume Controller.

Access Logix

Access Logix is an optional feature of the firmware code that provides the functionality that is known as LUN Mapping or LUN Virtualization.

You can use the software tab in the storage systems properties page of the EMC Navisphere GUI to determine if Access Logix is installed.

After Access Logix is installed it can be disabled but not removed. The following are the two modes of operation for Access Logix:

- **Access Logix not installed:** In this mode of operation, all LUNs are accessible from all target ports by any host. Therefore, the SAN fabric must be zoned to ensure that only the SAN Volume Controller can access the target ports.
- **Access Logix enabled:** In this mode of operation, a storage group can be formed from a set of LUNs. Only the hosts that are assigned to the storage group are allowed to access these LUNs.

Configuring the EMC CLARiiON controller with Access Logix installed

The SAN Volume Controller does not have access to the storage controller logical units (LUs) if Access Logix is installed on the EMC CLARiiON controller. You must use the EMC CLARiiON configuration tools to associate the SAN Volume Controller and LU.

The following prerequisites must be met before you can configure an EMC CLARiiON controller with Access Logix installed:

- The EMC CLARiiON controller is not connected to the SAN Volume Controller
- You have a RAID controller with LUs and you have identified which LUs you want to present to the SAN Volume Controller

You must complete the following tasks to configure an EMC CLARiiON controller with Access Logix installed:

- Register the SAN Volume Controller ports with the EMC CLARiiON
- Configure storage groups

The association between the SAN Volume Controller and the LU is formed when you create a storage group that contains both the LU and the SAN Volume Controller.

Registering the SAN Volume Controller ports with the EMC CLARiiON

You must register the SAN Volume Controller ports with an EMC CLARiiON controller if Access Logix is installed.

The following prerequisites must be met before you can register the SAN Volume Controller ports with an EMC CLARiiON controller that has Access Logix installed:

- The EMC CLARiiON controller is not connected to the SAN Volume Controller
- You have a RAID controller with LUs and you have identified which LUs you want to present to the SAN Volume Controller

Each initiator port [worldwide port name (WWPN)] must be registered against a host name and against a target port to which access is granted. If a host has multiple initiator ports, multiple table entries with the same host name are listed. If a host is allowed access using multiple target ports, multiple table entries are listed. For SAN Volume Controller hosts, all WWPN entries should carry the same host name.

The following table lists the associations:

Option	EMC CLARiiON default setting	SAN Volume Controller required setting
WWPN	N/A	Any
WWN	N/A	Any
Host name	N/A	Any

Option	EMC CLARiiON default setting	SAN Volume Controller required setting
SP port	N/A	Any
Initiator type	3	3
ArrayCommPath	Enable	Disable
Failover mode	0	2
Unit Serial Number	Array	Array

1. Connect the fibre channel and zone the fabric as required.
2. Issue the **svctask detectmdisk** command-line interface (CLI) command.
3. Right-click on the storage system from the Enterprise Storage window.
4. Select **Connectivity Status**. The Connectivity Status window is displayed.
5. Click **New**. The Create Initiator Record window is displayed.
6. Wait for the list of SAN Volume Controller ports to appear in the dialog box. Use the WWPN to Identify them. This can take several minutes.
7. Click **Group Edit**.
8. Select all instances of all the SAN Volume Controller ports in the Available dialog box.
9. Click the right arrow to move them to the selected box.
10. Fill in the **HBA WWN** field. You must know the following information:
    - WWNN of each SAN Volume Controller in the cluster
    - WWPN of each port ID for each node on the cluster

    The HBA WWN field is made up of the WWNN and the WWPN for the SAN Volume Controller port. The following is an example of the output:

    `50:05:07:68:01:00:8B:D8:50:05:07:68:01:20:8B:D8`
11. Select A in the field marked SP and 0 in the SP Port field.
12. Select **CLARiiON Open** in the drop down list of the **Initiator Type** field.
13. Deselect the ArrayCommPath checkbox if it has been selected.
14. Select **2** in the drop down list of the **Failover Mode** field.

    Attention: Failure to select failover mode 2 prevents the SAN Volume Controller from being able to failover I/O. Your data might become unavailable in the event of a single failure.

    a. If this is the first time that a port has been registered, ensure that you select the New Host option. Otherwise, select Existing Host.

    b. Ensure that the same host name is entered for each port that is registered.
15. Select **Array** in the drop down list of the **Unit Serial Number** field.
16. Assign a host name in the Host Name field.
17. Click **OK**.
18. Specify the IP address of your switch. The EMC CLARiiON does not use this IP address. However it must be unique (within the EMC CLARiiON) to prevent errant behavior by Navisphere.
19. Repeat step 11 for all possible combinations. The following example shows the different combinations of a system with four ports:
    - SP: A SP Port: 0
    - SP: A SP Port: 1
    - SP: B SP Port: 0
    - SP: B SP Port: 1
20. Repeat steps 1 to 19 to register the rest of your SAN Volume ControllerWWPNs.

All your WWPNs are registered against the host name that you specified.

Configuring your storage groups

Storage groups can only be configured if Access Logix is installed and enabled.

Access Logix provides the following LUN mapping:

Notes:
1. A subset of logical units (LUs) can form a storage group.
2. An LU can be in multiple storage groups.
3. A host can be added to a storage group. This host has access to all LUs in the storage group.
4. A host *cannot* be added to a second storage group.

1. Right-click on the storage system from the Enterprise Storage window.
2. Select **Create Storage Group**. The Create Storage Group window is displayed.
3. Enter a name for your storage group in the **Storage Group Name** field.
4. If available, select **Dedicated** in the **Sharing State** field.
5. Click **OK**. The storage group is created.
6. Right-click the storage group in the Enterprise Storage window.
7. Select **Properties**. The Storage Group Properties window is displayed.
8. Perform the following steps from the Storage Group Properties window:
   a. Select the **LUNs** tab.
   b. Select the LUNs that you want the SAN Volume Controller to manage in the Available LUNs table.

      Attention: Ensure that the LUs that you have selected are not used by another storage group.
   c. Click the forward arrow button.
   d. Click **Apply**. A Confirmation window is displayed.
   e. Click **Yes** to continue. A Success window is displayed.
   f. Click **OK**.
   g. Select the **Hosts** tab.
   h. Select the host that you created when you registered the SAN Volume Controller ports with the EMC CLARiiON.

      Attention: Ensure that only SAN Volume Controller hosts (initiator ports) are in the storage group.
   i. Click the forward arrow button.
   j. Click **OK**. The Confirmation window is displayed.
   k. Click **Yes** to continue. A Success window is displayed.
   l. Click **OK**.

Configuring the EMC CLARiiON controller without Access Logix installed

If Access Logix is not installed on an EMC CLARiiON controller, all logical units (LUs) that were created on the controller can be used by the SAN Volume Controller.

No further configuration of the EMC CLARiiON controller is necessary.

Configure the switch zoning such that no hosts can access these LUs.

Supported models of the EMC CLARiiON

The SAN Volume Controller supports models of the EMC CLARiiON.

See the following Web site for the latest supported models:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Supported firmware levels for the EMC CLARiiON

The EMC CLARiiON must use a firmware level that is supported by the SAN Volume Controller.

See the following Web site for specific firmware levels and the latest supported hardware:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Concurrent maintenance on EMC CLARiiON systems

Concurrent maintenance is the ability to perform I/O operations to a controller while simultaneously performing maintenance on it.

Important: An EMC Field Engineer must perform all maintenance procedures.

The EMC CLARiiON FC series and the SAN Volume Controller cluster allow concurrent replacement of the following components:
- Disk drives
- Controller fans (fans must be replaced within 2 minutes or controllers are shut down.)
- Disk enclosure fans (fans must be replaced within 2 minutes or controllers are shut down.)
- Controller (service processor: you must first disable cache)
- Fibre Channel Bypass cards (LCC)
- Power supplies (you must first remove fans.)
- Uninterruptible power supply battery (SPS)

EMC CLARiiON FC devices require that the I/O is quiesced during code upgrade. Consequently, the SAN Volume Controller cluster does not support concurrent upgrade of the FC controller code.

The EMC CLARiiON CX series and the SAN Volume Controller cluster allow concurrent replacement of the following components:
- Disk drives
- Controller (service processor or drawer controller)
- Power/cooling modules (modules must be replaced within 2 minutes or controllers are shut down.)
- Uninterruptible power supply battery (SPS)

The SAN Volume Controller cluster and EMC CLARiiON CX devices support concurrent code upgrade of the CX controllers.

Note:
- EMC CLARiiON procedures for concurrent upgrade must be followed in all cases.
- The CX Series also has a feature called Data In Place Upgrade which allows you to upgrade from one model to another (for example, from the CX200 to the CX600) with no data loss or migration required. This is *not* a concurrent operation.

EMC CLARiiON user interfaces

Ensure that you are familiar with the user interface applications that EMC CLARiiON systems use.

Navisphere or Navicli

The following user interface applications are available with EMC CLARiiON systems:

- Navisphere is the Web-based application that can be accessed from any Web browser.
- Navicli is the command-line interface (CLI) that is installed as part of the Navisphere Agent software (the host software).

Note: Some options and features are only accessible through the CLI.
Communication with the EMC CLARiiON in both cases is out-of-band. Therefore, the host does not need to be connected to the storage over fibre-channel and cannot be connected without Access Logix.

Sharing the EMC CLARiiON between a host and the SAN Volume Controller

The EMC CLARiiON can be shared between a host and a SAN Volume Controller.

- Split controller access is only supported when Access Logix is installed and enabled.
- A host cannot be connected to both the SAN Volume Controller and EMC CLARiiON at the same time.
- LUs must not be shared between a host and a SAN Volume Controller.
- Partitions in a RAID group must not be shared between a host and a SAN Volume Controller.

Switch zoning limitations for the EMC CLARiiON systems

There are limitations in switch zoning for SAN Volume Controller clusters and EMC CLARiiON systems.

FC4500 and CX200 models

The EMC CLARiiON FC4500 and CX200 systems limit the number of initiator HBAs to only allow 15 connections for each controller port. This limit is less than the 16 initiator ports that are required to connect to an 8–node cluster in a dual fabric configuration. To use EMC CLARiiON FC4500 and CX200 systems with an 8–node cluster, you must zone the system to use one SAN Volume Controller port for each node in each fabric. This reduces the initiator HBA count to eight.

FC4700 and CX400 models

EMC CLARiiON FC4700 and CX400 systems provide 4 target ports and allow 64 connections. Using a single SAN fabric, a 4-node cluster requires 64 connections ($4 \times 4 \times 4$), which is equal to the number of connections that are allowed. If split support with other hosts is required, this can cause issues. You can reduce either the number of initiator ports or target ports so that only 32 of the available 64 connections are used.

CX600 models

EMC CLARiiON CX600 systems provide 8 target ports and allow 128 connections. A 4-node cluster consumes all 128 connections ($4 \times 4 \times 8$). An 8-node cluster exceeds the connection limit and no reduction methods can be used.

Quorum disks on the EMC CLARiiON

The EMC CLARiiON supports quorum disks.

A SAN Volume Controller configuration that only includes the EMC CLARiiON is permitted.

Advanced functions for the EMC CLARiiON

Some advanced functions of the EMC CLARiiON are not supported by the SAN Volume Controller.

Advanced copy functions

Advanced copy functions for EMC CLARiiON, for example, SnapView, MirrorView and SANcopy, are not supported for disks that are managed by the SAN Volume Controller because the copy function does *not* extend to the SAN Volume Controller cache.

MetaLUN

MetaLUN allows a logical unit (LU) to be expanded using LUs in other RAID groups. The SAN Volume Controller only supports MetaLUN for the migration of image mode volumes.

Logical unit creation and deletion on the EMC CLARiiON

Binding a logical unit (LU) to a RAID group can take a significant amount of time on EMC CLARiiON systems.

Do not add the LU to a storage group until binding is complete. If the LU is mapped to a SAN Volume Controller cluster during the binding process, the LU might be identified with the wrong capacity. If this occurs, run the following procedure to rediscover the LU with the correct capacity:

1. Unmap the LU from the SAN Volume Controller cluster.
2. Run detectmdisk and wait for the managed disk to be deconfigured.
3. Wait for any binding activity to complete.
4. Remap the LU to the SAN Volume Controller cluster.
5. Run detectmdisk.

Configuring settings for the EMC CLARiiON

A number of settings and options are available through the EMC CLARiiON configuration interface.

The following settings and options are supported by the SAN Volume Controller:
* System
* Port
* Logical unit

Global settings for the EMC CLARiiON

Global settings apply across an EMC CLARiiON system. Not all options are available on all EMC CLARiiON models.

Table 32 lists the global settings that are supported by the SAN Volume Controller.

Table 32. EMC CLARiiON global settings supported by the SAN Volume Controller

Option	EMC CLARiiON default setting	SAN Volume Controller required setting
Access Controls (Access Logix installed)	Not installed	Either Installed or Not Installed
Subsystem Package Type	3	3
Queue Full Status	Disable	Disable
Recovered Errors	Disable	Disable
Target Negotiate	Displays the state of the target negotiate bit.	Displays the state of the target negotiate bit.
Mode Page 8 Info	Disable	Disable
Base UUID	0	0
Write Cache Enabled	Enabled	Enabled

Option	EMC CLARiiON default setting	SAN Volume Controller required setting
Mirrored Write Cache	Enabled	Enabled
Write Cache Size	600 MB	Default recommended
Enable Watermarks	Enabled	Enabled
Cache High Watermark	96%	Default
Cache Low Watermark	80%	Default
Cache Page Size	4 Kb	4 Kb
RAID3 Write Buffer Enable	Enable	Default recommended
RAID3 Write Buffer	0 MB	Default recommended

Controller settings for the EMC CLARiiON

The controller settings for the EMC CLARiiON are the settings that apply across one EMC CLARiiON system.

Table 33 lists the options that can be set by the EMC CLARiiON.

Table 33. EMC CLARiiON controller settings supported by the SAN Volume Controller

Option	EMC CLARiiON default setting	SAN Volume Controller required setting
Read Cache Enabled	Enable	Enable
Read Cache Size	200 MB	Default recommended
Statistics Logging	Disable	Either Enable or Disable

Note: The SAN Volume Controller cannot obtain or change the configuration options that are listed above. You must configure the options that are listed above.

Port settings for the EMC CLARiiON

Port settings are configurable at the port level.

Table 34 lists port settings, the EMC CLARiiON defaults, and the required settings for SAN Volume Controller clusters.

Table 34. EMC CLARiiON port settings

Option	EMC CLARiiON default setting	SAN Volume Controller required setting
Port speed	Depends on the model	Any

Note: The SAN Volume Controller cluster cannot obtain or change the configuration options that are listed above. You must configure the options that are listed above.

Logical unit settings for the EMC CLARiiON

Logical unit (LU) settings are configurable at the LU level.

Table 35 on page 156 lists the options that must be set for each LU that is accessed by the SAN Volume Controller. LUs that are accessed by hosts can be configured differently.

Table 35. EMC CLARiiON LU settings supported by the SAN Volume Controller

Option	EMC CLARiiON default setting	SAN Volume Controller required setting
LU ID	Auto	N/A
RAID Type	5	Any RAID Group
RAID Group	Any available RAID Group	Any available RAID Group
Offset	0	Any setting
LU Size	ALL LBAs in RAID Group	Any setting
Placement	Best Fit	Either Best Fit or First Fit
UID	N/A	N/A
Default Owner	Auto	N/A
Auto Assignment	Disabled	Disabled
Verify Priority	ASAP	N/A
Rebuild Priority	ASAP	N/A
Strip Element Size	128	N/A
Read Cache Enabled	Enabled	Enabled
Write Cache Enabled	Enabled	Enabled
Idle Threshold	0–254	0–254
Max Prefetch Blocks	0–2048	0–2048
Maximum Prefetch IO	0–100	0–100
Minimum Prefetch Size	0–65534	0–65534
Prefetch Type	0, 1, or 2	0, 1, or 2
Prefetch Multiplier	0 to 2048 or 0 to 324	0 to 2048 or 0 to 324
Retain prefetch	Enabled or Disabled	Enabled or Disabled
Prefetch Segment Size	0 to 2048 or 0 to 32	0 to 2048 or 0 to 32
Idle Delay Time	0 to 254	0 to 254
Verify Priority	ASAP, High, Medium, or Low	Low
Write Aside	16 to 65534	16 to 65534

Note: The SAN Volume Controller cannot obtain or change the configuration options that are listed above. You must configure the options that are listed above.

Configuring EMC Symmetrix and Symmetrix DMX systems

This topic provides information about configuring the EMC Symmetrix and Symmetrix DMX for attachment to SAN Volume Controller.

3 On some versions of Symmetrix and Symmetrix DMX, the setting of SPC-2 can be configured. SPC-2 is
3 set either on a per-port basis or on a per-initiator basis. LUs that are mapped to SAN Volume Controller
3 must be configured with SPC-2 disabled.

3 **Note:** Changing the value of the SPC-2 setting on a live system can cause errors. If you have a live
3 system running with SPC-2 enabled on LUs mapped to SAN Volume Controller, contact the IBM Support
3 Center for guidance on how to proceed. Do not disable SPC-2 on a live system before taking guidance
3 from the IBM Support Center.

Supported models of the EMC Symmetrix and Symmetrix DMX controllers

The SAN Volume Controller supports models of the EMC Symmetrix and Symmetrix DMX controllers.

See the following Web site for the latest supported models:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Supported firmware levels for the EMC Symmetrix and Symmetrix DMX

The EMC Symmetrix and Symmetrix DMX must use a firmware level that is supported by the SAN Volume Controller.

See the following Web site for specific firmware levels and the latest supported hardware:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Concurrent maintenance on the EMC Symmetrix and Symmetrix DMX

Concurrent maintenance is the capability to perform I/O operations to the EMC Symmetrix or Symmetrix DMX while simultaneously performing maintenance operations on it.

Important: Service actions and upgrade procedures can only be performed by an EMC Field Engineer.

The EMC Symmetrix and Symmetrix DMX are Enterprise class devices that support nondisruptive replacement of the following components:
- Channel Director
- Disk Director
- Cache card
- Disk drive
- Cooling fan
- Comms card
- EPO card
- Operator panel
- PSU
- Service Processor
- Batteries
- Ethernet hub

The SAN Volume Controller and EMC Symmetrix/Symmetrix DMX support concurrent upgrade of the EMC Symmetrix/Symmetrix DMX firmware.

User interfaces on EMC Symmetrix and Symmetrix DMX

Ensure that you are familiar with the user interface applications that support the EMC Symmetrix and Symmetrix DMX systems.

EMC Control Center

A basic EMC Symmetrix or Symmetrix DMX configuration is performed by an EMC Field Engineer (FE) using the EMC Symmetrix service processor. After the initial configuration, you can configure and control the exported storage. The FE defines the storage device types and sets the configurable options.

You can configure and control the exported storage as described below.

You can use the EMC Control Center to manage and monitor the EMC Symmetrix and Symmetrix DMX systems.

You can use Volume Logix for volume configuration management. Volume Logix allows you to control access rights to the storage when multiple hosts share target ports.

SYMCLI

The EMC Symmetrix Command Line Interface (SYMCLI) allows the server to monitor and control the EMC Symmetrix and Symmetrix DMX.

Sharing the EMC Symmetrix or Symmetrix DMX system between a host and a SAN Volume Controller cluster

There are restrictions for sharing EMC Symmetrix and Symmetrix DMX systems between a host and a SAN Volume Controller cluster.

An EMC Symmetrix or Symmetrix DMX system can be shared between a host and a SAN Volume Controller under the following conditions:

- When possible, avoid sharing target ports between the SAN Volume Controller cluster and other hosts. If this cannot be avoided, you must regularly check the combined I/O workload that is generated by the SAN Volume Controller cluster and the other hosts. The performance of either the SAN Volume Controller cluster or the hosts is impacted if the workload exceeds the target port capabilities.
- A single host must not be connected to a SAN Volume Controller and an EMC Symmetrix or Symmetrix DMX because the multipathing drivers (for example, subsystem device driver (SDD) and PowerPath) cannot coexist.
- If the EMC Symmetrix or Symmetrix DMX is configured in such a way that other hosts cannot access the LUs that are managed by the SAN Volume Controller cluster, other hosts can be directly connected to an EMC Symmetrix or Symmetrix DMX system at the same time as a SAN Volume Controller cluster.

Switch zoning limitations for the EMC Symmetrix and Symmetrix DMX

There are limitations in switch zoning for the SAN Volume Controller and the EMC Symmetrix and Symmetrix DMX systems.

Switch zoning

The SAN Volume Controller switch zone must include at least one target port on two or more fibre-channel adapters to avoid a single point of failure.

The EMC Symmetrix and Symmetrix DMX must be configured to present logical units (LUs) to all SAN Volume Controller initiator ports that are in the fabric zone.

Only SAN Volume Controller initiator ports that are LUN masked on the EMC Symmetrix or Symmetrix DMX controller should be present in the fabric zone.

Note: The EMC Symmetrix and Symmetrix DMX systems present themselves to a SAN Volume Controller cluster as separate controllers for each port zoned to the SAN Volume Controller. For example, if one of these storage systems has 4 ports zoned to the SAN Volume Controller, each port appears as a separate controller rather than one controller with 4 WWPNs. In addition, a given logical unit (LU) must be mapped to the SAN Volume Controller through all controller ports zoned to the SAN Volume Controller using the same logical unit number (LUN).

Connecting to the SAN

You can connect a maximum of 16 EMC Symmetrix or Symmetrix DMX ports to the SAN Volume Controller cluster. There are no further special zoning requirements. Configurations that are setup to adhere to the requirements that are described in previous SAN Volume Controller releases are also supported, but should not be followed for new installations.

Quorum disks on EMC Symmetrix and Symmetrix DMX

The SAN Volume Controller chooses managed disks (MDisks) that are presented by the EMC Symmetrix or Symmetrix DMX as quorum disks.

The SAN Volume Controller uses a logical unit (LU) that is presented by an EMC Symmetrix or Symmetrix DMX as a quorum disk. The SAN Volume Controller provides a quorum disk even if the connection is through a single port.

Advanced functions for EMC Symmetrix and Symmetrix DMX

3 SAN Volume Controller cache-disabled volumes can be used as the source or target in Symmetrix
3 advanced copy functions (for example, Symmetrix Remote Data Facility [SRDF] and TimeFinder).

LU creation and deletion on EMC Symmetrix and Symmetrix DMX

A logical unit (LU) that is exported by an EMC Symmetrix or Symmetrix DMX, meaning it is visible to a host, is either a *Symmetrix device* or a *Meta device*.

Symmetrix device

2 **Restriction:** An LU with a capacity of less than 64 MB is ignored by theSAN Volume Controller.

Symmetrix device is an EMC term for an LU that is hosted by an EMC Symmetrix. These are all emulated devices and have exactly the same characteristics. The following are the characteristics of a Symmetrix device:
* N cylinders
* 15 tracks per cylinder
* 64 logical blocks per track
* 512 bytes per logical block

Symmetrix devices can be created using the **create dev** command from the EMC Symmetrix Command Line Interface (SYMCLI). The configuration of an LU can be changed using the **convert dev** command from the SYMCLI. Each physical storage device in an EMC Symmetrix is partitioned into 1 to 128 hyper-volumes (hypers). Each hyper can be up to 16 GB. A Symmetrix device maps to one or more hypers, depending on how it is configured. The following are examples of hyper configurations:
* Hypers can be mirrored (2-way, 3-way, 4-way)
* Hypers can be formed into RAID-S groups

Meta device

Meta device is an EMC term for a concatenated chain of EMC Symmetrix devices. This enables the EMC Symmetrix to provide LUs that are larger than a hyper. Up to 255 hypers can be concatenated to form a single meta device. Meta devices can be created using the **form meta** and **add dev** commands from the SYMCLI. This allows an extremely large LU to be created, however, if exported to the SAN Volume
2 Controller, only the first 1 PB is used.

Do not extend or reduce meta devices that are used for managed disks (MDisks). Reconfiguration of a meta device that is used for an MDisk causes unrecoverable data-corruption.

Configuring settings for the EMC Symmetrix and Symmetrix DMX

A number of settings and options are available through the EMC Symmetrix configuration interface.

Settings and options are available in the following categories:
* System
* Port
* Logical unit (LU)
2 • Initiator

Global settings for the EMC Symmetrix and Symmetrix DMX

Global settings apply across the EMC Symmetrix and Symmetrix DMX systems.

You can specify EMC Symmetrix and Symmetrix DMX settings with the **set Symmetrix** command from the Symmetrix Command Line Interface (SYMCLI). The settings can be viewed using the **symconfigure** command from the SYMCLI.

Table 36 lists the EMC Symmetrix global settings that can be used with SAN Volume Controller clusters.

Table 36. EMC Symmetrix and Symmetrix DMX global settings

Option	EMC Symmetrix and Symmetrix DMX default setting	SAN Volume Controller required setting
max_hypers_per_disk	-	Any
dynamic_rdf	Disable	Any
fba_multi_access_cache	Disable	N/A
Raid_s_support	Disable	Enable or Disable

Port settings for the EMC Symmetrix and Symmetrix DMX

Target port characteristics can be set using the **set port** command from the Symmetrix Command Line Interface (SYMCLI).

The target port characteristics can be viewed using the **symcfg** command from the SYMCLI.

Table 37 lists the EMC Symmetrix and Symmetrix DMX port settings that can be used with the SAN Volume Controller cluster.

Table 37. EMC Symmetrix and Symmetrix DMX port settings that can be used with the SAN Volume Controller

Option	EMC Symmetrix and Symmetrix DMX default setting	SAN Volume Controller required setting
Disk_Array	Enabled	Disabled
Volume_Set_Addressing	Enabled	Disabled
Hard_Addressing	Enabled	Enabled
Non_Participating	Disabled	Disabled
Global_3rdParty_Logout	Enabled	Enabled
Tagged_Commands	Enabled	Enabled
Common_Serial_Number	-	Enabled
Disable_Q_Reset_on_UA	Disabled	Disabled
Return_busy_for_abort	Disabled	Disabled
SCSI-3	Disabled	Disabled or Enabled
Environ_Set	Disabled	Disabled

Table 37. EMC Symmetrix and Symmetrix DMX port settings that can be used with the SAN Volume Controller (continued)

Option	EMC Symmetrix and Symmetrix DMX default setting	SAN Volume Controller required setting
Unique_WWN	Enabled	Enabled
Point_to_Point	Disabled	Enabled
VCM_State	Disabled	Disabled or Enabled
OpenVMS	Disabled	Disabled
SPC-2	Disabled	Disabled

Note: If your Symmetrix or Symmetrix DMX has SPC-2 enabled, do not disable it. Contact the IBM Support Center for guidance on how to proceed.

Logical unit settings for the EMC Symmetrix and Symmetrix DMX

Logical unit (LU) settings are configurable at the LU level.

LU characteristics can be set using the **set device** command from the Symmetrix Command Line Interface (SYMCLI).

Table 38 lists the options that must be set for each LU that is accessed by the SAN Volume Controller.

Table 38. EMC Symmetrix and Symmetrix DMX LU settings supported by the SAN Volume Controller

Option	EMC Symmetrix and Symmetrix DMX default setting	SAN Volume Controller required setting
emulation	-	FBA
attribute	-	Set all attributes to disabled.

Initiator settings for the EMC Symmetrix and Symmetrix DMX

Initiator settings for SPC-2 should be set to disabled for EMC Symmetrix and Symmetrix DMX.

Table 39 lists the EMC Symmetrix and Symmetrix DMX Initiator settings supported by the SAN Volume Controller.

Table 39. EMC Symmetrix and Symmetrix DMX initiator settings supported by the SAN Volume Controller

Option	EMC Symmetrix and Symmetrix DMX default setting	SAN Volume Controller required setting
SPC-2	Disabled	Disabled

Note: If your Symmetrix or Symmetrix DMX has SPC-2 enabled for SAN Volume Controller initiators, do not disable it. Contact the IBM Support Center for guidance on how to proceed.

Mapping and virtualization settings for the EMC Symmetrix and Symmetrix DMX

Mapping a logical unit (LU) to a host is a function of the EMC Control Center.

LUs can be mapped to a particular director or target port using the **map dev** command from the Symmetrix Command Line Interface (SYMCLI). LUs can be unmapped using the **unmap dev** command from the SYMCLI.

Volume Logix and masking

Volume Logix allows you to restrict access to particular WWPNs on the fabric for Symmetrix Volumes.

This function can be switched on and off by changing the VMC_State port setting. The SAN Volume Controller requires that you do not share target ports between a host and a SAN Volume Controller. However, you can still use Volume Logix to protect the system from errors that can occur if the SAN is not correctly configured.

To mask a volume to the SAN Volume Controller, you must first identify the SAN Volume Controller ports that are connected to each system. This can be done using the EMC Symmetrix symmask command.

The SAN Volume Controller automatically logs into any EMC Symmetrix system it sees on the fabric. You can use the SAN Volume Controller svcinfo lsnode CLI command to find the correct port identifiers.

After you have identified the ports, you can map each volume on each port to each WWPN. The EMC Symmetrix stores the LUN masking in a database, so you must apply the changes you have made to refresh the contents of the database to view the changes.

Configuring EMC VMAX systems

This section provides information about configuring the EMC VMAX systems for attachment to a SAN Volume Controller.

Note: VMAX settings provided in this section must be applied before configuring SAN Volume Controller LUNS.

Supported models of the EMC VMAX controllers

The SAN Volume Controller supports models of the EMC VMAX controllers.

See the following website for the latest supported models:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Supported firmware levels for the EMC VMAX

The EMC VMAX system must use a firmware level that is supported by the SAN Volume Controller.

See the following website for specific firmware levels and the latest supported hardware:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Note: The minimum supported SAN Volume Controller level for the attachment of EMC VMAX is 4.3.1.

Concurrent maintenance on the EMC VMAX

Concurrent maintenance is the capability to perform I/O operations to the EMC VMAX while simultaneously performing maintenance operations on it.

Important: Service actions and upgrade procedures can only be performed by an EMC Field Engineer.

The EMC VMAX is an enterprise-class device that supports nondisruptive replacement of the following components:
- Channel Director
- Disk Director
- Cache card

3 • Disk drive

3 • Cooling fan

3 • Comms card

3 • EPO card

3 • Operator panel

3 • Power-supply unit (PSU)

3 • Service Processor

3 • Batteries

3 • Ethernet hub

2 The SAN Volume Controller and EMC VMAX support concurrent upgrade of the EMC VMAX firmware.

2 User interfaces on EMC VMAX

2 Ensure that you are familiar with the user interface applications that support the EMC VMAX systems.

2 EMC Control Center

2 A basic EMC VMAX configuration is performed by an EMC Field Engineer (FE) using the EMC VMAX
2 service processor. After the initial configuration, you can configure and control the exported storage. The
2 FE defines the storage device types and sets the configurable options.

3 You can configure and control the exported storage as described in the following sections.

2 You can use the EMC Control Center to manage and monitor the EMC VMAX systems.

2 You can use Volume Logix for volume configuration management. With Volume Logix, you can control
2 access rights to the storage when multiple hosts share target ports.

2 SYMCLI

3 The EMC Symmetrix Command Line Interface (SYMCLI) is used by the server to monitor and control the
3 EMC VMAX.

2 Sharing the EMC VMAX system between a host and a SAN Volume
2 Controller cluster

2 There are restrictions for sharing EMC VMAX systems between a host and a SAN Volume Controller
2 cluster.

2 An EMC VMAX system can be shared between a host and a SAN Volume Controller under the following
2 conditions:

2 • When possible, avoid sharing target ports between the SAN Volume Controller cluster and other hosts.
2   If this cannot be avoided, you must regularly check the combined I/O workload that is generated by
2   the SAN Volume Controller cluster and the other hosts. The performance of either the SAN Volume
2   Controller cluster or the hosts is impacted if the workload exceeds the target port capabilities.

3 • A single host must not be connected to a SAN Volume Controller and an EMC VMAX because the
3   multipathing drivers (for example, subsystem device driver [SDD] and PowerPath) cannot coexist.

2 • If the EMC VMAX is configured in such a way that other hosts cannot access the LUs that are
2   managed by the SAN Volume Controller cluster, other hosts can be directly connected to an EMC
2   VMAX system at the same time as a SAN Volume Controller cluster.

Switch zoning limitations for the EMC VMAX

There are limitations in switch zoning for the SAN Volume Controller and EMC VMAX systems.

Switch zoning

The SAN Volume Controller switch zone must include at least one target port on two or more fibre-channel adapters to avoid a single point of failure.

The EMC VMAX must be configured to present logical units (LUs) to all SAN Volume Controller initiator ports that are in the fabric zone.

Only SAN Volume Controller initiator ports that are LUN-masked on the EMC VMAX controller should be present in the fabric zone.

Note: An EMC VMAX system presents itself to a SAN Volume Controller cluster as one WWNN with a minimum of two and a maximum of 16 WWPNs supported.

Connecting to the SAN

You can connect a maximum of 16 EMC VMAX ports to the SAN Volume Controller cluster. There are no further special zoning requirements. Configurations that are set up to adhere to the requirements that are described in previous SAN Volume Controller releases are also supported, but should not be followed for new installations.

Quorum disks on EMC VMAX

The SAN Volume Controller chooses managed disks (MDisks) that are presented by the EMC VMAX as quorum disks.

The SAN Volume Controller uses a logical unit (LU) that is presented by an EMC VMAX as a quorum disk. The SAN Volume Controller provides a quorum disk even if the connection is through a single port.

Advanced functions for EMC VMAX

SAN Volume Controller cache-disabled volumes can be used as the source or target in VMAX advanced copy functions (for example, Symmetrix Remote Data Facility [SRDF] and TimeFinder).

LU creation and deletion on EMC VMAX

A logical unit (LU) that is exported by an EMC VMAX, meaning that it is visible to a host, is either a *VMAX device* or a *Meta device*.

VMAX device

Restriction: An LU with a capacity of 64 MB or less is ignored by the SAN Volume Controller.

VMAX device is an EMC term for an LU that is hosted by an EMC VMAX. These are all emulated devices and have exactly the same characteristics. The following are the characteristics of a VMAX device:

- N cylinders
- 15 tracks per cylinder
- 64 logical blocks per track
- 512 bytes per logical block

VMAX devices can be created using the **create dev** command from the EMC Symmetrix Command Line Interface (SYMCLI). The configuration of an LU can be changed using the **convert dev** command from

2 the SYMCLI. Each physical storage device in an EMC VMAX is partitioned into 1 to 128 hyper-volumes
2 (hypers). Each hyper can be up to 16 GB. A VMAX device maps to one or more hypers, depending on
2 how it is configured. The following configurations are examples of hyper configurations:

2 • Hypers can be mirrored (2-way, 3-way, 4-way).

2 • Hypers can be formed into RAID-S groups.

2 Meta device

3 *Meta device* is an EMC term for a concatenated chain of EMC VMAX devices. The EMC VMAX uses a
3 meta device to provide LUs that are larger than a hyper. Up to 255 hypers can be concatenated to form a
3 single meta device. Using the **form meta** and **add dev** commands from the SYMCLI, you can create meta
3 devices, which produce an extremely large LU. However, if exported to the SAN Volume Controller, only
2 the first 1 PB are used.

2 **Attention:** Do not extend or reduce meta devices that are used for managed disks (MDisks).
2 Reconfiguring a meta device that is used for an MDisk causes unrecoverable data corruption.

2 Configuring settings for the EMC VMAX

2 A number of settings and options are available through the EMC VMAX configuration interface.

3 The settings and options can have the following scope:
3 • System
3 • Port
3 • Logical unit (LU)

2 Global settings for the EMC VMAX
2 Global settings apply across the EMC VMAX systems.

2 You can specify EMC VMAX settings with the **set Symmetrix** command from the Symmetrix Command
2 Line Interface (SYMCLI). The settings can be viewed using the **symconfigure** command from the
2 SYMCLI.

2 Table 40 lists the EMC VMAX global settings that must be set for the SAN Volume Controller.

2 *Table 40. EMC VMAX global settings*

2 Option	EMC VMAX default setting	SAN Volume Controller required setting
2 Maximum number of hypers per disk	512	Any
2 Switched RDF Configuration state	Disabled	Default
2 Concurrent RDF Configuration state	Enabled	Default
2 Dynamic RDF Configuration state	Enabled	Any
2 Concurrent Dynamic RDF Configuration	Enabled	Default
2 RDF Data Mobility Configuration State	Disabled	Default
2 Access Control Configuration State	Enabled	Default
2 Device Masking (ACLX) Config State	Enabled	Default
2 Multi LRU Device Assignment	None	Default
2 Disk Group Assignments	In Use	Default
2 Hot Swap Policy	Permanent	Default
2 Symmetrix Disk Library	Disabled	Default

2 *Table 40. EMC VMAX global settings (continued)*

2 Option	EMC VMAX default setting	SAN Volume Controller required setting
2 FBA Geometry Emulation	Native	Default
2 3 Dynamic Mirrors	Enabled	Default
2 PAV Mode	DynamicStandardPAV	Default
2 PAV Alias Limit	31	Default

2

2 Port settings for the EMC VMAX

2 Target port characteristics can be set using the **set port** command from the Symmetrix Command Line
2 Interface (SYMCLI).

2 The target port characteristics can be viewed using the **symcfg** command from the SYMCLI.

2 Table 41 lists the options that must be used with the SAN Volume Controller.

2 *Table 41. EMC VMAX port settings*

2 Option	EMC VMAX default setting	SAN Volume Controller required setting
2 **SCSI Flags**		
2 Negotiate_Reset(N)	Disabled	Default
2 Soft_Reset(S)	Disabled	Default
2 Environ_Set(E)	Disabled	Default
2 HP3000_Mode(B)	Disabled	Default
2 Common_Serial_Number(C)	Enabled	Default
2 Disable_Q_Reset_on_UA(D)	Disabled	Default
2 Sunapee(SCL)	Disabled	Default
2 Siemens(S)	Disabled	Default
2 Sequent(SEQ)	Disabled	Default
2 Avoid_Reset_Broadcast(ARB)	Disabled	Default
2 Server_On_AS400(A4S)	Disabled	Default
2 SCSI_3(SC3)	Enabled	Enabled
2 SPC2_Protocol_Version(SPC2)	Disabled	Default
2 SCSI_Support1(OS2007)	Enabled	Disabled

2

2 Logical unit settings for the EMC VMAX

2 Logical unit (LU) settings are configurable at the LU level.

2 LU characteristics can be set using the **set device** command from the Symmetrix Command Line Interface
2 (SYMCLI).

2 Table 42 on page 167 lists the options that must be set for each LU that is accessed by the SAN Volume
2 Controller.

2 *Table 42. EMC VMAX LU settings supported by the SAN Volume Controller*

Option	EMC VMAX default setting	SAN Volume Controller required setting
emulation	-	FBA
attribute	-	Set all attributes to disabled.

Fibre-specific flag settings for the EMC VMAX

Fibre-specific flag settings for the EMC VMAX are provided in this section.

Table 43 lists the fibre-specific flag settings that must be set for the SAN Volume Controller.

Table 43. EMC VMAX fibre-specific flag settings supported by the SAN Volume Controller

Option	EMC VMAX default setting	SAN Volume Controller required setting
Volume_Set_Addressing(V)	Disabled	Default
Non_Participating(NP)	Disabled	Default
Init_Point_to_Point(PP)	Enabled	Default
Unique_WWN(UWN)	Enabled	Default
Access_Logix(ACLX)	Enabled	Default
OpenVMS(OVMS)	Disabled	Default
AS400(AS4)	Disabled	Default
Auto_Negotiate(EAN)	Disabled	Default

Mapping and virtualization settings for the EMC VMAX

Mapping a logical unit (LU) to a host is a function of the EMC Control Center.

LUs can be mapped to a particular director or target port using the **map dev** command from the Symmetrix Command Line Interface (SYMCLI). LUs can be unmapped using the **unmap dev** command from the SYMCLI.

Volume Logix and masking

Volume Logix is used to restrict access to particular WWPNs on the fabric for Symmetrix Volumes.

This function can be switched on and off by changing the VMC_State port setting. The SAN Volume Controller requires that you do not share target ports between a host and a SAN Volume Controller. However, you can still use Volume Logix to protect the system from errors that can occur if the SAN is not correctly configured.

To mask a volume to the SAN Volume Controller, you must first identify the SAN Volume Controller ports that are connected to each system. You can identify these ports using the EMC Symmetrix **symmask** command.

The SAN Volume Controller automatically logs in to any EMC VMAX system that it sees on the fabric. You can use the SAN Volume Controller **svcinfo lsnode** CLI command to find the correct port identifiers.

After you have identified the ports, you can map each volume on each port to each WWPN. The EMC VMAX stores the LUN masking in a database; so you must apply the changes that you have made to refresh the contents of the database to view the changes.

Configuring Fujitsu ETERNUS systems

This section provides information about configuring the Fujitsu ETERNUS systems for attachment to a SAN Volume Controller.

Supported models of the Fujitsu ETERNUS

The SAN Volume Controller supports models of the Fujitsu ETERNUS series of systems.

See the following Web site for the latest supported models:Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Supported firmware levels for the Fujitsu ETERNUS

The Fujitsu ETERNUS must use a firmware level that is supported by the SAN Volume Controller.

See the following Web site for specific firmware levels and the latest supported hardware:Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

User interfaces on the Fujitsu ETERNUS

Ensure that you are familiar with the user interface application that is used by the Fujitsu ETERNUS.

You can use the ETERNUSmgr web-based configuration utility. See the documentation that is provided with the Fujitsu ETERNUS system for more information.

Configuring the Fujitsu ETERNUS to use with the SAN Volume Controller

Ensure that you use the settings that are required to use the Fujitsu ETERNUS with the SAN Volume Controller. It is important that you use the correct settings to avoid data access problems.

Use the following sequence of steps to configure the Fujitsu ETERNUS system:

1. Configure the SAN Volume Controller host response pattern.
2. Register the host world wide names (WWNs) and associate them with the host response pattern.
3. Setup the affinity group for SAN Volume Controller volumes or setup LUN mapping.
4. Create or reassign storage to the SAN Volume Controller.

For all other settings and procedures, consider the SAN Volume Controller a host. See the documentation that is provided with the Fujitsu ETERNUS system.

CA parameters

The following table lists the port settings that are required. See the documentation that is provided with your Fujitsu ETERNUS system for more information because some options are only available on certain models.

Option	Fujitsu ETERNUS default setting	SAN Volume Controller required setting
Connection Topology/FC Connection Settings	FC-AL Connection	Fabric Connection
Service Class	Class 3	Class 3
FC Transfer Rate	Auto Setting	Any

Option	Fujitsu ETERNUS default setting	SAN Volume Controller required setting
Reset Scope/Scope of LUR Actions	T_L	T_L **Note:** If this option is not set correctly, data corruption can occur.
Release Reservation upon Chip Reset	Enable/valid	Enable/valid
HP-UX Connection Setting	Disable	Disable
Frame Size Setting	2048	Any
Affinity/Addressing Mode	Off	Any

Host response pattern

The SAN Volume Controller requires that a new host response pattern is created. If the Host Affinity/Host Table Settings Mode is used, this host response pattern must be associated with each WWN. If the Host Affinity/Host Table Settings Mode is not used, this host response pattern must be associated with the target port.

The following table lists the settings that are required. See the documentation that is provided with your Fujitsu ETERNUS system for more information because some options are only available on certain models.

Option	Fujitsu ETERNUS default setting	SAN Volume Controller required setting
Command timeout interval	Depends on the Fujitsu ETERNUS model	Default
Response status in overload	Unit Attention	Unit Attention
Byte 0 of Inquiry response/Response to inquiry commands	Default	Default
Inquiry Standard Data NACA Function	Disable	Disable
Inquiry Standard Data Version	Depends on the Fujitsu ETERNUS model	Default
Inquiry Command Page 83/Inquiry VPD ID Type	Depends on the Fujitsu ETERNUS model	Type 01
Reservation Conflict Response to Test Unit Ready Commands	Disable/Normal Response	Enable/Conflict Response
Target Port Group Access Support	Disable	Enable
Host Specific Mode	Normal Mode	Normal Mode
Response Sense at Firmware Hot Switching	Enable	Enable
Change LUN mapping	No Report	Report
LUN Capacity Expansion	No Report	Report
Aymmetric / Symmetric Logical Unit Access	Active/Active	Active/Active
Pattern of Sense Code Conversion	No Conversion	No Conversion

Notes:

1. If you set Inquiry VPD ID Type option to Type 3 on E4000 or E8000 range, the MDisks go offline.

2. If you set the Target Port Group Access Support option to Disabled on E3000 range, a 1370 error is
   shown in the event log.

t

Host WWNs

After the SAN Volume Controller is zoned on the fabric to see the Fujitsu ETERNUS, the system might
not initially appear in the list of controllers when you issue the **lscontroller** CLI command. This is normal
and expected behavior.

See the documentation that is provided with the Fujitsu ETERNUS system to add all SAN Volume
Controller WWPNs as host WWNs. The following restrictions apply:

- The SAN Volume Controller WWNs must be associated with a host response pattern. The host
  response pattern must be defined prior to registration. If you use an incorrect or default host response
  pattern, you can lose access to data.
- All SAN Volume Controller WWNs must be registered on all Fujitsu ETERNUS ports on the same
  fabric. If the WWNs are not registered, you can lose access to data.

Affinity groups/zones

Use the affinity groups/zones mode to protect the SAN Volume Controller LUs if the SAN is incorrectly
configured. The affinity group mode is setup in the CA configuration. See the documentation that is
provided with your Fujitsu ETERNUS system for more information about using the affinity groups/zones
mode. The following restrictions apply:

- Each SAN Volume Controller must have exactly one affinity group/zone.
- The SAN Volume Controller affinity group/zone must be associated with all SAN Volume Controller
  WWNs.

LUN mapping

You can use the LUN mapping mode (also called the zone settings mode for some models) with the
following restrictions:

- The SAN zoning must only allow a single SAN Volume Controller access to this target port.
- The host response pattern must be set in CA configuration using the required SAN Volume Controller
  settings.

Note: If you use the LUN mapping mode, you cannot use the host affinity mode. The host affinity mode
is set to OFF.

Assigning storage to the SAN Volume Controller

Ensure that you understand all SAN Volume Controller and Fujitsu ETERNUS restrictions before you
assign storage to the SAN Volume Controller. See the documentation that is provided with the Fujitsu
ETERNUS system for more information.

Zoning configuration for the Fujitsu ETERNUS

If LUN mapping mode is used for a Fujitsu ETERNUS port, you must exclusively zone the SAN Volume
Controller with this target port.

Migrating logical units from the Fujitsu ETERNUS to the SAN Volume Controller

You can use the standard migration procedure with the following restrictions:

- The SAN Volume Controller must have software level 4.2.0 or higher installed before you start migration. Upgrades from previous SAN Volume Controller software levels to software level 4.2.0 or higher causes all Fujitsu ETERNUS systems that are attached to be excluded.
- You must configure the Fujitsu ETERNUS system to work with the SAN Volume Controller before you start migration.
- The subsystem device driver (SDD) and Fujitsu Multipath driver cannot coexist.
- The SAN Volume Controller must support all host code levels.

Concurrent maintenance on the Fujitsu ETERNUS

Concurrent maintenance is the capability to perform I/O operations to a Fujitsu ETERNUS while simultaneously performing maintenance operations on it.

You can perform nondisruptive maintenance procedures concurrently on the following components:
- Fujitsu ETERNUS controller module
- Fujitsu ETERNUS controller cache
- Fujitsu ETERNUS cache battery pack
- Fan
- Power supply
- Disk drive
- SFP transceivers

See the documentation that is provided with the Fujitsu ETERNUS system for more information.

Advanced functions for the Fujitsu ETERNUS

The Fujitsu ETERNUS system provides several advanced copy functions. Do not use these advanced copy functions for storage that is managed by the SAN Volume Controller, even if the volume cache is disabled.

Configuring IBM TotalStorage ESS systems

This section provides information about configuring the IBM TotalStorage Enterprise Storage Server (ESS) for attachment to a SAN Volume Controller.

Configuring the IBM ESS

The IBM Enterprise Storage Server (ESS) provides functionality that is compatible with the SAN Volume Controller.

Perform the following steps to configure the IBM ESS:
1. Enter the IP address of the IBM ESS in a Web browser to access the ESS Specialist.
2. Login with your user name and password.
3. Click **ESS Specialist**.
4. Click **Storage Allocation**.
5. Click **Open System Storage**.
6. Click **Modify Host Systems**.
7. Create a host entry for each initiator port on each SAN Volume Controller node in your cluster. Complete the following fields:
   a. Enter a unique name for each port in the **Nickname** field. For example, enter knode or 1node.
   b. Select **IBM SAN Volume Controller** in the **Host Type** field. If this option is not available, select **RS/6000**.

c. Select **Fibre Channel attached** in the **Host Attachment** field.

d. Leave the **Hostname/IP address** field blank.

e. Select the WWPN from the list or enter it manually into the **WWPN** field. A configuration command fails if you use WWPN 0 in the command string.

8. Click **Perform Configuration Update** after you are finished adding all of the ports.

9. Click **Add Volumes** to add the volumes that you want the SAN Volume Controller to use. The Add Volumes panel is displayed.

10. Perform the following steps in the Add Volumes panel:

a. Select any of the SAN Volume Controller host ports that you created earlier.

b. Select the necessary ESS adapter to create the volumes.

c. Click **Next**.

d. Create volumes using your desired size, placement, and RAID level.

e. Click **Perform Configuration Update** after you have created all the volumes.

11. Perform the following steps to map the volumes to all of your SAN Volume Controller ports:

a. Click **Modify Volume Assignments**.

b. Select all of the volumes that you created earlier.

c. Click **Assigning selected volumes to target hosts**.

d. Select all of the remaining SAN Volume Controller host ports that you created earlier.

e. Click **Perform Configuration Update**.

Important: If you are adding SAN Volume Controller ports to a volume that is already assigned to other SAN Volume Controller ports, you must select the **Use same ID/LUN in source and target** check box.

Supported models of the IBM ESS

The SAN Volume Controller supports models of the IBM Enterprise Storage Server (ESS).

See the following Web site for the latest supported models:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Supported firmware levels for the IBM ESS

The SAN Volume Controller supports the IBM Enterprise Storage Server (ESS).

See the following Web site for specific firmware levels and the latest supported hardware:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Concurrent maintenance on the IBM ESS

Concurrent maintenance is the capability to perform I/O operations to an IBM Enterprise Storage Server (ESS) while simultaneously performing maintenance operations on it.

All IBM ESS concurrent maintenance procedures are supported.

User interface on the IBM ESS

Ensure that you are familiar with the user interface application that supports the IBM Enterprise Storage Server (ESS) system.

Web Server

A Web server runs on each of the controllers on the system. During normal operation, the user interface
e application provides only basic monitoring of the system and displays an event log. If you press the reset
button on the controller to put the controller into diagnostic mode, the user interface application allows
firmware upgrades and system configuration resets.

Sharing the IBM ESS between a host and the SAN Volume Controller

The IBM Enterprise Storage Server (ESS) can be shared between a host and a SAN Volume Controller.

The following restrictions apply when you share the IBM ESS between a host and a SAN Volume
Controller:
- If an IBM ESS port is in the same zone as a SAN Volume Controller port, that same IBM ESS port
  should not be in the same zone as another host.
- A single host can have both IBM ESS direct-attached and SAN Volume Controller virtualized disks
  configured to it.
- If a LUN is managed by the SAN Volume Controller, it *cannot* be mapped to another host.

See the following Web site for the latest supported configurations:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Switch zoning limitations for the IBM ESS

Consider the following limitations when you zone the IBM Enterprise Storage Server (ESS) to the SAN
Volume Controller.

To avoid a single point of failure on the IBM ESS, you must have a minimum of two SAN connections
from two separate adapter bays. The maximum number of IBM ESS SAN connections in the SAN Volume
Controller switch zone is 16.

Note: The IBM ESS provides ESCON®, FICON® and Ultra SCSI connectivity; however, only a 1 or 2 Gb
fibre-channel SAN attachment is supported by the SAN Volume Controller.

Quorum disks on the IBM ESS

The SAN Volume Controller can choose managed disks (MDisks) that are presented by the IBM
Enterprise Storage Server (ESS) controller as quorum disks.

Advanced functions for the IBM ESS

t SAN Volume Controller cache-disabled volumes can be used as the source or target for IBM Enterprise
Storage Server (ESS) advanced copy functions (for example, FlashCopy, MetroMirror, GlobalCopy).

Logical unit creation and deletion on the IBM ESS

Certain IBM Enterprise Storage Server (ESS) types are supported for use with the SAN Volume
Controller.

Before you delete or unmap a logical unit (LU) from the SAN Volume Controller, remove the LU from the
managed disk (MDisk) group. The following is supported:
2 - LU size of 1 GB to 1 PB.
- RAID 5 and RAID 10 LUs.
- LUs can be added dynamically.

Attention: When adding additional SAN Volume Controller ports to an existing LU, you must select the **Use same ID/LUN in source and target** check box. Failure to select the **Use same ID/LUN in source and target** checkbox can cause loss in redundancy or a loss of data. If this checkbox is not available, the option is not required. The detect MDisks task in the management GUI or the **svctask detectmdisk** command-line interface (CLI) command must be run for the SAN Volume Controller to detect the new disks.

Configuring IBM System Storage DS5000, IBM DS4000, and IBM DS3000 systems

This section provides information about configuring IBM System Storage DS5000, IBM DS4000 and IBM DS3000 systems for attachment to a SAN Volume Controller cluster. Some IBM System Storage DS4000 controllers are equivalent to StorageTek models; SAN Volume Controller also supports certain StorageTek FlexLine series and StorageTek D series. The information in this section also applies to the supported models of the StorageTek FlexLine series and StorageTek D series.

3 IBM System Storage DS5000, IBM DS4000, and IBM DS3000 are similar systems. The concepts in this
3 section apply generally to all three systems; however, some options might not be available. See the
3 documentation that is provided with your system for specific information.

Configuring IBM System Storage DS5000, IBM DS4000, and IBM DS3000 systems for the storage server

IBM System Storage DS5000, IBM DS4000, and IBM DS3000 storage systems are supported with the SAN Volume Controller cluster.

The following steps provide the supported options and impact on the SAN Volume Controller cluster:

3 1. Set the host type for SAN Volume Controller to `IBM TS SAN VCE`. For higher security, create a storage
3   partition for every host that will have access to the storage system. If you set a default host group and
3   add another host other than SAN Volume Controller to the default group, the new host automatically
3   has full read and write access to all LUNs on the storage system.
3 2. See the following website for the scripts that are available to change the setup of the IBM System
3   Storage DS5000, IBM DS4000, or IBM System Storage DS3000 system:
3   www.ibm.com/storage/support/

The following limitations apply to partitions:

3 • Only one IBM DS5000, IBM DS4000, or IBM DS3000 system storage partition that contains any of the
3   ports of any of the nodes in a single SAN Volume Controller cluster can be created.
  • Only map one partition to any of the ports on any of the nodes that are in the SAN Volume Controller
    cluster to avoid unexpected behavior. For example, you can lose access to your storage or you might
    not receive warning messages, even if there are errors logged in the SAN Volume Controller error log.

3 The following limitation applies to IBM DS5000, IBM DS4000, or IBM DS3000 Copy Services:

3 • Do not use IBM DS5000, IBM DS4000, or IBM System Storage DS3000 Copy Services when the SAN
3   Volume Controller cluster is attached to an IBM DS5000, IBM DS4000, or IBM DS3000 system.
3 • You can use partitioning to allow IBM DS5000, IBM DS4000, or IBM DS3000 Copy Services usage for
3   other hosts.

The following information applies to the access LUN (also known as the Universal Transport Mechanism (UTM) LUN):

3 • The access/UTM LUN is a special LUN that allows a IBM DS5000, IBM DS4000, or IBM DS3000 system
3   to be configured through software over the fibre-channel connection. The access/UTM LUN does not

3   have to be in the partition that contains the SAN Volume Controller ports because the access/UTM
3   LUN is not required by the SAN Volume Controller cluster. No errors are generated if the access/UTM
3   LUN is not in the partition.

The following information applies to the logical unit (LU):
3 • The SAN Volume Controller cluster attempts to follow the preferred ownership that is specified by the
3   storage system. You can specify which controller (A or B) is used for I/O operations to an LU.
   • If the SAN Volume Controller cluster can see the ports of the preferred controller and error conditions
     do not exist, the SAN Volume Controller cluster accesses the LU through one of the ports on the
     preferred controller.
3 • If error conditions exist, the SAN Volume Controller cluster ignores the preferred ownership of the IBM
3   DS5000, IBM DS4000, or IBM DS3000 system.

Supported options for IBM System Storage DS5000. IBM DS4000, and IBM DS3000 systems

3 IBM System Storage DS5000, IBM DS4000, and IBM DS3000 series storage systems provide functions that
3 can be used with the SAN Volume Controller.

3 The storage manager for IBM System Storage DS5000, IBM DS4000 and IBM DS3000 systems has several
3 options and actions that you can use.

Controller run diagnostics

Diagnostics are automatically recovered by the SAN Volume Controller software. After the controller run
diagnostics option is used, check your managed disks (MDisks) to ensure that they have not been set to
degraded mode.

Controller disable data transfer

3 The controller disable data transfer option is not supported when a SAN Volume Controller is attached to
3 IBM System Storage DS5000, IBM DS4000, or IBM DS3000 systems.

Setting an array Offline

t Do not set an array offline because you can lose access to the storage pool.

Array increase capacity

The array increase capacity option is supported, but the new capacity is not usable until the MDisk is
t removed from the storage pool and re-added to the storage pool. You might have to migrate data to
increase the capacity.

Redistribute logical drives or change ownership of the preferred path

You can redistribute logical drives or change ownership of the preferred path; however, these options
might not take effect until a discovery is started on the SAN Volume Controller cluster. You can use the
svctask detectmdisk command-line interface (CLI) command to restart a cluster discovery process. The
discovery process rescans the fibre-channel network to discover any new MDisks that might have been
added to the cluster and to rebalance MDisk access across the available controller device ports.

Controller reset

You must only use the controller reset option if you are directed to do so by IBM Service and the alternate controller is functional and available to the SAN. The SAN Volume Controller reset is automatically recovered by the SAN Volume Controller software.

Check your MDisks to ensure that they have not been set to the degraded state during the controller reset process. You can issue the **svctask includemdisk** CLI command to repair degraded MDisks.

Supported models of IBM System Storage DS5000, IBM DS4000 and IBM DS3000 systems

The SAN Volume Controller supports models of the IBM System Storage DS5000, IBM DS4000, and IBM DS3000 systems. Some IBM System Storage DS4000 series storage systems are equivalent to Sun StorageTek and StorageTek models; SAN Volume Controller also supports some Sun StorageTek, StorageTek FlexLine and D series models.

3 See the following website for the latest supported models:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

3 **Note:** Some older levels of IBM System Storage DS4000 microcode support a maximum of 32 LUNs per
3 host partition. Newer firmware versions allow 256 up to 2048 LUNs per host partition.

Supported firmware levels for IBM System Storage DS5000, IBM DS4000, and IBM DS3000 systems

You must ensure that the firmware level of the system can be used with the SAN Volume Controller cluster.

See the following website for specific firmware levels and the latest supported hardware:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

The website includes the maximum number of LUNs per partition that are supported by the firmware level.

Concurrent maintenance on IBM System Storage DS5000, IBM DS4000, and IBM DS3000 systems

3 Concurrent maintenance is the capability to perform I/O operations to an IBM System Storage DS5000,
3 IBM DS4000, or IBM DS3000 series storage system while simultaneously performing maintenance
3 operations on the system.

3 See your IBM System Storage DS5000, IBM DS4000, or IBM DS3000 series documentation for information
3 about concurrent maintenance.

Sharing an IBM System Storage DS5000, IBM DS4000, or IBM DS3000 systems between a host and SAN Volume Controller

You can share an IBM System Storage DS5000, IBM DS4000 or IBM DS3000 system between a host and a SAN Volume Controller cluster.

3 The IBM System Storage DS5000, IBM DS4000, and IBM DS3000 function known as *partitioning* must be
3 used to separate groups of logical units that are directly attached to hosts or groups of hosts from the
3 logical units that are accessed by the SAN Volume Controller cluster.

3 **Note:** The SAN Volume Controller partition must either contain all the host ports of the SAN Volume
3 Controller cluster that are connected to the SAN or are zoned to have access to the storage system ports.
3 For example, configure so that each SAN Volume Controller host bus adapter (HBA) port on SAN
3 Volume Controller will be able to see at least one port on storage system A and one port on storage
3 system B.

Quorum disks on IBM System Storage DS5000, IBM DS4000, and IBM DS3000 systems

The SAN Volume Controller can choose managed disks (MDisks) that are presented by a IBM System
Storage DS5000, IBM DS4000, or IBM DS3000 system as quorum disks.

Note: The FASsT series 200 does not support quorum disks.

Advanced functions for IBM System Storage DS5000, IBM DS4000, and IBM DS3000 systems

t SAN Volume Controller cache-disabled volumes can be used as the source or target for IBM System
Storage DS5000, IBM DS4000, and IBM DS3000 advanced copy functions: for example, FlashCopy and
Metro Mirror.

Data migration on partitioned IBM System Storage DS5000, IBM DS4000, and IBM DS3000 systems

You can migrate data on partitioned IBM System Storage DS5000, IBM DS4000, and IBM DS3000 systems.

You can enable the SAN Volume Controller to be introduced to an existing SAN environment, so that you
have the option of using image mode LUNs to import the existing data into the virtualization
environment without requiring a backup and restore cycle. Each partition can only access a unique set of
HBA ports, as defined by the worldwide port names (WWPNs). For a single host to access multiple
partitions, unique host fibre ports (WWPNs) must be assigned to each partition. All LUNs within a
partition are identified to the assigned host fibre ports (no subpartition LUN mapping).

Host A is mapped to LUN 0, 1, 2 in Partition 0.
Host B is mapped to LUN 0, 1, 2, 3, 4, 5 in Partition 1.
Host C is mapped to LUN 0, 1, 2 in Partition 2.

To allow Host A to access the LUNs in partition B, you must remove one of the HBAs (for example, A1)
from the access list for partition 0 and add it to partition 1. A1 cannot be on the access list for more than
one partition.

To add a SAN Volume Controller into this configuration without backup and restore cycles requires a set
of unique SAN Volume Controller HBA port WWPNs for each partition. This allows the IBM System
Storage DS5000, IBM DS4000, or IBM DS3000 system to make the LUNs known to the SAN Volume
Controller, which then configures these LUNs as image-mode LUNs and identifies them to the required
hosts. This violates a requirement that all SAN Volume Controller nodes must be able to see all back-end
storage. For example, to fix this problem for an IBM DS4000 system, change the configuration to allow
more than 32 LUNs in one storage partition, so that you can move all the LUNs from all the other
partitions into one partition and map to the SAN Volume Controller cluster.

Scenario: the SAN Volume Controller nodes cannot see all back-end storage

The IBM DS4000 series has eight partitions with 30 LUNs in each.

Perform the following steps to allow the SAN Volume Controller nodes to see all back-end storage:
1. Change the mappings for the first four partitions on the IBM DS4000 system such that each partition
   is mapped to one port on each node. This maintains redundancy across the cluster.

2. Create a new partition on the system that is mapped to all four ports on all the nodes.

3. Gradually migrate the data into the managed disks (MDisks) in the target partition. As storage is freed from the source partitions, it can be reused as new storage in the target partition. As partitions are deleted, new partitions that must be migrated can be mapped and migrated in the same way. The host side data access and integrity is maintained throughout this process.

Logical unit creation and deletion on IBM System Storage DS5000, IBM DS4000 and IBM DS3000 systems

You can create or delete logical units on IBM System Storage DS5000, IBM DS4000, and IBM DS3000 storage systems.

Some IBM System Storage DS5000. IBM DS4000, and IBM DS3000 storage systems are supported for use with SAN Volume Controller clusters.

3 To create a logical disk, set the host type for SAN Volume Controller to `IBM TS SAN VCE`.

Configuration interface for IBM System Storage DS5000, IBM DS4000, and IBM DS3000 systems

IBM System Storage DS5000, IBM DS4000, and IBM DS3000 systems include a configuration application.

3 The access LUN, also known as the Universal Transport Mechanism (UTM) LUN, is the configuration
3 interface for IBM System Storage DS5000, IBM DS4000, and IBM System Storage DS3000 systems.

The access LUN might not be in a partition that contains the SAN Volume Controller ports because it is not required by the SAN Volume Controller cluster. The UTM LUN is a special LUN that allows IBM System Storage DS5000, IBM DS4000, and IBM System Storage DS3000 systems to be configured through suitable software over the fibre-channel connection. Because the SAN Volume Controller does not require the UTM LUN, it does not generate errors either way. IBM System Storage DS5000, IBM DS4000, and IBM System Storage DS3000 systems *must not* have the Access UTM LUN that is presented as LUN 0 (zero).

3 It is possible to use in-band (over fibre channel) and out-of-band (over Ethernet) to allow the
3 configuration software to communicate with more than one IBM System Storage DS5000, IBM DS4000, or
3 IBM System Storage DS3000 system. If you are using in-band configuration, the Access UTM LUN must
3 be configured in a partition that does not include any logical units that are accessed by the SAN Volume
3 Controller cluster.

Note: In-band is not supported for access to the LUN while in the SAN Volume Controller partition.

Controller settings for IBM System Storage DS5000, IBM DS4000, and IBM DS3000 systems

Controller settings are the settings that apply across one IBM System Storage DS5000, IBM DS4000, or IBM DS3000 system.

3 You must configure the following settings for IBM System Storage DS5000, IBM DS4000. and IBM DS3000
3 systems:
3 • Set the host type for SAN Volume Controller to `IBM TS SAN VCE`.
3 • Set the system so that both storage systems have the same worldwide node name (WWNN). See the
3 following website for the scripts that are available to change the setup for IBM System Storage DS5000,
3 IBM DS4000, and IBM DS3000 systems:
3 www.ibm.com/storage/support/

3 • Ensure that the AVT option is enabled. The host type selection should have already enabled the AVT
3   option. View the storage system profile data to confirm that you have the AVT option enabled. This
3   storage profile is presented as a text view in a separate window. See the following Web site for the
3   scripts that are available to enable the AVT option:

3   www.ibm.com/storage/support/

3 • You must have the following options enabled on any logical units that are mapped to IBM System
3   Storage DS5000, IBM DS4000, and IBM DS3000 systems:

3   – read caching

3   – write caching

3   – write cache mirroring

3 • You must not have caching without batteries enabled.

Configuration settings for IBM System Storage DS5000, IBM DS4000, and IBM DS3000 systems

The system configuration interface provides configuration settings and options that can be used with the SAN Volume Controller cluster.

These settings and options can have the following scope:

• System
• Logical unit (LU)
  – The SAN Volume Controller cluster attempts to follow preferred ownership that is specified by the system. You can specify which controller (A or B) is used to perform I/O operations to a given LU. If the SAN Volume Controller cluster can see the ports of the preferred controller and no error conditions exist, the SAN Volume Controller cluster accesses that LU through one of the ports on that controller. Under error conditions, the ownership is ignored.
  – You must have the following options enabled on any LUs that are mapped to the SAN Volume Controller cluster:
    - read caching
    - write caching
    - write cache mirroring
3   – You must not have caching without batteries enabled.

Global settings for IBM System Storage DS5000, IBM DS4000, or IBM DS3000 systems

Global settings apply across IBM System Storage DS5000, IBM DS4000, or IBM DS3000systems.

Table 44 lists the global settings that can be used with SAN Volume Controller clusters.

1 *Table 44. IBM System Storage DS5000, DS4000, and IBM DS3000 system global options and settings*

Option	Setting
Start flushing	50%
Stop flushing	50%
Cache block size	4 Kb (for systems running 06.x or earlier)
	8 Kb or 16 Kb (for systems running 07.x or later)

3 **Attention:** See the IBM DS5000, IBM DS4000, or IBM DS3000 documentation for details on how to
3 modify the settings.

3 Use a host type for SAN Volume Controller of IBM TS SAN VCE to establish the correct global settings for
3 the SAN Volume Controller cluster.

Logical unit settings for IBM System Storage DS5000, IBM DS4000, and IBM DS3000 systems

Logical unit (LU) settings are configurable at the LU level.

LUs that are accessed by hosts can be configured differently.

Use the following option settings for a LUN that will be attached to SAN Volume Controller cluster.

Table 45. Option settings for a LUN

Parameter	Setting
Segment size	256 KB
Capacity reserved for future segment size changes	Yes
Maximum future segment size	2,048 KB
Modification priority	High
Read cache	Enabled
Write cache	Enabled
Write cache without batteries	Disabled
Write cache with mirroring	Enabled
Flush write cache after (in seconds)	10.00
Dynamic cache read prefetch	Enabled
Enable background media scan	Enabled
Media scan with redundancy check	Enabled
Pre-Read redundancy check	Disabled

You must not have caching without batteries enabled.

3   Set the host type for SAN Volume Controller to `IBM TS SAN VCE` when you create a new LU.

Miscellaneous settings for IBM System Storage DS5000, IBM DS4000, or IBM DS3000 systems

The SAN Volume Controller cluster supports all media scan settings that are provided by the system. Set the background media scan to enabled and set the frequency to 30 days. These settings are enabled at both the system level and the individual logical drive level.

See the documentation that is provided with your system for information about other settings.

Configuring IBM System Storage DS6000 systems

This section provides information about configuring the IBM System Storage DS6000™ system for attachment to a SAN Volume Controller.

Configuring the IBM DS6000

The IBM DS6000 provides functions that are compatible with the SAN Volume Controller.

After you have defined at least one storage complex, storage unit, and I/O port, you can define the SAN Volume Controller as a host and create host connections. If you have not defined all of these required storage elements, use the IBM System Storage DS6000 Storage Manager or the IBM DS6000 command-line interface (CLI) to define these elements and return to this topic after they are configured.

This task assumes that you have already launched the IBM System Storage DS6000 Storage Manager.

Perform the following steps to configure the IBM DS6000:

1. Click **Real-time manager** → **Manage hardware** → **Host systems**.
2. Select **Create** from the **Select Action** list. The Create Host System wizard is displayed.
3. Perform the following steps to select a host type:
   a. Select **IBM SAN Volume Controller (SVC)** from the **Host Type** list.
   b. Enter a unique name of up to 16 characters for each port in the **Nickname** field. The value that you enter in this field appears in other panels when you select defined hosts. This is a required field.
   c. Optionally, enter a detailed description of up to 256 characters in the **Description** field.
   d. Click **Next**. The Define host wizard panel is displayed.
4. Perform the following steps in the Define host panel:
   a. Enter the number of ports that you are defining for the SAN Volume Controller node in the **Quantity** field.

      Note: You must add all of the SAN Volume Controller node ports.
   b. Select **FC Switch fabric (P-P)** from the **Attachment Port Type** list.
   c. Click **Add**.
   d. Select **Group ports to share a common set of volumes**.
   e. Click **Next**. The Define host WWPN panel is displayed.
5. Specify a WWPN for each SAN Volume Controller node port that you are configuring. After you have defined all SAN Volume Controller node port WWPNs, click **Next**.
6. Perform the following steps in the Specify storage units panel:
   a. Select all the available storage units that use the ports that you defined in step 5.
   b. Click **Add** to move the selected storage units to the **Selected storage units** field.
   c. Click **Next**. The Specify storage units parameters panel is displayed
7. Perform the following steps in the Specify storage units parameters panel:
   a. Select a host attachment identifier from the table.
   b. Click **the following specific storage unit I/O ports** in the **This host attachment can login to** field. The available ports are displayed in the Available storage unit I/O ports table.
   c. Select each port in the Available storage unit I/O ports table.

      Note: The **Type** for each port should be **FcSf**. If the listed type is not FcSf, click **Configure I/O Ports**. The Configure I/O Ports panel is displayed. Click the port that you want to configure and select **Change to FcSf** from the **Select Action** list.
   d. Click **Apply assignment**.
   e. Click **OK**. The Verification panel is displayed.
8. Verify that the attributes and values that are displayed in the table are correct.
9. Click **Finish** if the values that are displayed in the table are correct. Otherwise, click **Back** to return to the previous panels and change the values that are not correct.

Supported firmware levels for the IBM DS6000

The IBM DS6000 must use a firmware level that is supported by the SAN Volume Controller.

See the following Web site for specific firmware levels and the latest supported hardware:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Supported models of the IBM DS6000 series

The SAN Volume Controller supports models of the IBM DS6000 series of controllers.

See the following Web site for the latest supported models:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

User interfaces on the IBM DS6000

Ensure that you are familiar with the user interfaces that support the IBM DS6000.

Web server

You can manage, configure, and monitor the IBM DS6000 through the IBM System Storage DS6000 Storage Manager.

CLI

You can also manage, configure, and monitor the IBM DS6000 through the IBM System Storage DS command-line interface.

Concurrent maintenance on the IBM DS6000

Concurrent maintenance is the capability to perform I/O operations to an IBM DS6000 while simultaneously performing maintenance operations on it.

All IBM DS6000 concurrent maintenance procedures are supported.

Target port groups on the IBM DS6000

The IBM DS6000 uses the SCSI Target Port Groups feature to indicate a preferred path for each logical unit (LU).

Sharing an IBM System Storage DS6000 system between a host and the SAN Volume Controller

You can share an IBM System Storage DS6000 system between a host and a SAN Volume Controller cluster.

Quorum disks on IBM System Storage DS6000 systems

The SAN Volume Controller can choose managed disks (MDisks) that are presented by a IBM System Storage DS6000 system as quorum disks.

Configuring IBM System Storage DS8000 systems

This section provides information about configuring the IBM System Storage DS8000 system for attachment to a SAN Volume Controller.

Configuring the IBM DS8000

The IBM DS8000 provides functions that are compatible with the SAN Volume Controller.

After you have defined at least one storage complex, storage unit, and I/O port, you can define the SAN Volume Controller as a host and create host connections. If you have not defined all of these required storage elements, use the IBM System Storage DS8000 Storage Manager or the IBM System Storage DS® command-line interface to define these elements and return to this topic after they are configured.

This task assumes that you have already launched the IBM System Storage DS8000 Storage Manager.

Perform the following steps to configure the IBM DS8000:

1. Click **Real-time manager** → **Manage hardware** → **Host systems**.
2. Select **Create** from the **Select Action** list. The Create Host System wizard begins.
3. Perform the following steps to select a host type:
   a. Select **IBM SAN Volume Controller (SVC)** from the **Host Type** list.
   b. Enter a unique name of up to 16 characters for each port in the **Nickname** field. The value that you enter in this field is displayed in other panels when you select defined hosts. This is a required field.
   c. Optionally, enter a detailed description of up to 256 characters in the **Description** field.
   d. Click **Next**. The Define host wizard panel is displayed.
4. Perform the following steps in the Define host panel:
   a. Enter the number of ports that you are defining for the SAN Volume Controller node in the **Quantity** field.

      Note: You must add all of the SAN Volume Controller node ports.
   b. Select **FC Switch fabric (P-P)** from the **Attachment Port Type** list.
   c. Click **Add**.
   d. Select **Group ports to share a common set of volumes**.
   e. Click **Next**. The Define host WWPN panel is displayed.
5. Specify a WWPN for each SAN Volume Controller node port that you are configuring. When you have defined all SAN Volume Controller node port WWPNs, click **Next**.
6. Perform the following steps in the Select storage images panel:
   a. Select all the available storage units that use the ports that you defined in the previous step.
   b. Click **Add** to move the selected storage units to the **Select storage images** field.
   c. Click **Next**. The Specify storage image parameters panel is displayed
7. Perform the following steps in the Specify storage image parameters panel:
   a. Select a host attachment identifier from the table.
   b. Click **the following specific storage image I/O ports** in the **This host attachment can login to** field. The available ports are displayed in the Available storage unit I/O ports table.
   c. Select each port in the Available storage unit I/O ports table.

      Note: The **Type** for each port should be **FcSf**. If the listed type is not FcSf, click **Configure I/O Ports**. The Configure I/O Ports panel is displayed. Click the port that you want to configure and select **Change to FcSf** from the **Select Action** list.
   d. Click **Apply assignment**.
   e. Click **OK**. The Verification panel is displayed.
8. Verify that the attributes and values that are displayed in the table are correct.
9. Click **Finish** if the values that are displayed in the table are correct. Otherwise, click **Back** to return to the previous panels and change the incorrect values.

Supported firmware levels for the IBM DS8000

The SAN Volume Controller supports the IBM DS8000 series.

See the following Web site for specific firmware levels and the latest supported hardware:Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Supported models of the IBM DS8000

The SAN Volume Controller supports models of the IBM DS8000 series of controllers.

See the following Web site for the latest supported models:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

User interfaces on the IBM DS8000

Ensure that you are familiar with the user interfaces that support the IBM DS8000.

Web server

You can manage, configure, and monitor the IBM DS8000 through the IBM System Storage DS8000 Storage Manager.

CLI

You can also manage, configure, and monitor the IBM DS8000 through the IBM System Storage DS command-line interface.

Concurrent maintenance for the IBM DS8000

Concurrent maintenance is the capability to perform I/O operations to an IBM DS8000 while simultaneously performing maintenance operations on it.

All IBM DS8000 concurrent maintenance procedures are supported.

Sharing an IBM System Storage DS8000 system between a host and the SAN Volume Controller

You can share an IBM System Storage DS8000 system between a host and a SAN Volume Controller cluster.

Quorum disks on IBM System Storage DS8000 systems

The SAN Volume Controller can choose managed disks (MDisks) that are presented by a IBM System Storage DS8000 system as quorum disks.

Configuring HDS Lightning series systems

This section provides information about configuring the Hitachi Data Systems (HDS) Lightning series system for attachment to a SAN Volume Controller.

The information in this section also applies to the supported models of the Sun StorEdge series and the HP XP series.

Supported models of the HDS Lightning

The SAN Volume Controller supports models of the HDS Lightning. Certain models of the HDS Lightning are equivalent to Sun StorEdge and HP XP models; therefore, the SAN Volume Controller also supports models of the Sun StorEdge and the HP XP.

See the following Web site for the latest supported models:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Supported firmware levels for HDS Lightning

The SAN Volume Controller supports the HDS Lightning.

See the following Web site for specific HDS Lightning firmware levels and the latest supported hardware:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Note: Concurrent upgrade of the controller firmware is supported with the SAN Volume Controller.

Concurrent maintenance on the HDS Lightning

Concurrent maintenance is the capability to perform I/O operations to an HDS Lightning while simultaneously performing maintenance operations on it.

Important: An HDS Field Engineer must perform all maintenance procedures.

User interface on HDS Lightning

Ensure that you are familiar with the user interface application that supports the HDS Lightning system.

Service Processor (SVP)

HDS Lightning has a laptop in the controller frame. The laptop runs the Service Processor (SVP) as the primary configuration user interface. You can use SVP to perform most configuration tasks and to monitor the controller.

HiCommand

The HiCommand is a graphical user interface that allows basic creation of storage and system monitoring. The HiCommand communicates with HDS Lightning through Ethernet.

Sharing the HDS Lightning 99*xx*V between a host and the SAN Volume Controller

There are restrictions for sharing an HDS Lightning 99*xx*V between a host and a SAN Volume Controller cluster.

Sharing ports

The HDS Lightning 99*xx*V can be shared between a host and a SAN Volume Controller cluster under the following conditions:

- The same host cannot be connected to both a SAN Volume Controller cluster and an HDS Lightning at the same time because the Hitachi HiCommand Dynamic Link Manager (HDLM) and the subsystem device driver (SDD) cannot coexist.
- A controller port cannot be shared between a host and a SAN Volume Controller cluster. If a controller port is used by a SAN Volume Controller cluster, it must not be present in a switch zone that allows a host to access the port.
- Logical units (LUs) cannot be shared between a host and a SAN Volume Controller cluster.

Supported Topologies

You can connect the SAN Volume Controller cluster to the HDS Lightning under the following conditions:

- For SAN Volume Controller software version 4.2.1 and later, you can connect a maximum of 16 HDS Lightning ports to the SAN Volume Controller cluster without any special zoning requirements.

- For SAN Volume Controller software version 4.2.0, the following applies:
  - Logical Unit Size Expansion (LUSE) and Virtual LVI/LUN operations cannot be run on a disk that is managed by the SAN Volume Controller cluster. LUNs that are created using LUSE and Virtual LVI/LUN can be mapped to the cluster after they are created.
  - Only disks with open emulation can be mapped to the SAN Volume Controller cluster.
  - S/390® disks cannot be used with the SAN Volume Controller cluster.
  - Only fibre-channel connections can connect the SAN Volume Controller cluster to the HDS Lightning.

Switch zone limitations for HDS Lightning

There are limitations in switch zoning for the SAN Volume Controller and the HDS Lightning systems.

Switch zoning

The HDS Lightning systems present themselves to a SAN Volume Controller cluster as separate controllers for each port zoned to the SAN Volume Controller. For example, if one of these storage systems has 4 ports zoned to the SAN Volume Controller, each port appears as a separate controller rather than one controller with 4 WWPNs. In addition, a given logical unit (LU) must be mapped to the SAN Volume Controller through all controller ports zoned to the SAN Volume Controller using the same logical unit number (LUN).

Quorum disks on HDS Lightning 99xxV

HDS Lightning 99xxV is not an approved host for quorum disks. Therefore, configurations with only HDS Lightning are not possible.

Advanced functions for HDS Lightning

Some advanced functions of the HDS Lightning are not supported by the SAN Volume Controller.

Advanced copy functions

Advanced copy functions for HDS Lightning (for example, ShadowImage, Remote Copy, and Data Migration) are not supported for disks that are managed by the SAN Volume Controller, because the copy function does not extend to the SAN Volume Controller cache.

Logical Unit Size Expansion

The HDS Lightning 99xxV supports Logical Unit Expansion (LUSE). LUSE is *not* a concurrent operation. LUSE is accomplished by concatenating between 2 and 26 existing logical units (LUs) together. Before LUSE can be performed on an LU, the LU must be removed from the managed disk (MDisk) group and unmapped from the SAN Volume Controller.

Attention: LUSE destroys all data that exists on the LU, except on a Windows system.

TrueCopy

TrueCopy operations are functionally similar to Metro Mirror. TrueCopy processing is not supported when the disk controller system is used with the SAN Volume Controller. Even when an HDS Lightning 99xxV is shared between a host and a SAN Volume Controller, TrueCopy processing is not supported on the ports that are zoned directly with the host.

Virtual LVI/LUNs

The HDS Lightning 99*xx*V supports Virtual LVI/LUNs. Virtual LVI/LUNs is *not* a concurrent operation. Virtual LVI/LUNs allows you to divide LUNs into several smaller virtual LUNs for use by the HDS Lightning. You must first create existing LUNs into free space and then define their own LUNs using that free space. Virtual LVI/LUNs must *not* be managed or mapped to a SAN Volume Controller.

LUNs that are set up using either LUSE or Virtual LVI/LUNs appear as normal LUNs after they are created. Therefore, LUNs that are set up using LUSE or Virtual LVI/LUNs can be used by the SAN Volume Controller after they are created.

Write protect

LUs cannot be explicitly set to write-protected. However, some of the advanced features, such as Metro Mirror, can be used to write-protect an LU as part of the function. Metro Mirror must not be used for LUs that are in use by a SAN Volume Controller.

Logical unit configuration for HDS Lightning

Logical unit (LU) configuration for HDS Lightning supports both RAID 1 and RAID 5 arrays.

The HDS Lightning system can have up to 8192 LUs defined; however, only 256 LUs can be mapped to a single port. Report LUNs is supported by LUN 0, so the SAN Volume Controller can detect all LUNs.

In the event that a LUN 0 is not configured, the HDS Lightning system presents a pseudo-LUN at LUN 0. The inquiry data for this pseudo-LUN slightly differs from the inquiry data of normal LUNs. The difference allows the SAN Volume Controller to recognize the pseudo-LUN and exclude it from I/O. The pseudo LUN can accept the report LUNs command.

The HDS Lightning system supports both open-mode attachment and S/390 attachment. The emulation mode is set when the LU is defined. All LUNs that are presented to a SAN Volume Controller must use open emulation. All LUNs with open emulation use a standard 512 byte block size.

The HDS Lightning system can only have certain sized LUs that are defined. These LUs can be expanded by merging 2 - 36 of these LUs using the Logical Unit Size Expansion (LUSE) feature. They can also be made into several, smaller virtual LUNs by using the Virtual LVI/LUN feature.

Special LUs

When an LU is mapped to a host, you have the option to make it a *command LUN*. Command LUNs support in-band configuration commands, but not I/O. Therefore, you cannot map command LUNs to the SAN Volume Controller.

Logical unit creation and deletion on HDS Lightning

The SAN Volume Controller supports Logical Unit Size Expansion (LUSE) with certain restrictions.

The following restrictions apply:
- Before LUSE can be performed on an LU, the LU must be unmounted from a host and have no available paths. The LUSE function destroys all data that exists on the LU, except for LUs on a Windows operating system.
- LUSE must not be performed on any disk that is managed by the SAN Volume Controller.
- If data exists on a disk and you want to use image mode to import the data, do not use LUSE on the disk before you import the data.

Configuring settings for HDS Lightning

The Lightning configuration interface provides functions for configuration.

These options and settings can have the following scope:
- Subsystem
- Port
- Logical unit (LU)

Global settings for HDS Lightning

Global settings apply across an HDS Lightning disk controller system.

Table 46 lists the global settings for HDS Lightning.

Table 46. HDS Lightning global settings supported by the SAN Volume Controller

Option	Lightning default setting	SAN Volume Controller required setting
Spare disk recover	Interleave	Interleave
Disk copy place	Medium	Medium
Copy operation	Correction copy and dynamic sparing	Correction copy and dynamic sparing
Read configuration data mode	Selected	Selected
PS off timer	Not selected	Not selected

Controller settings for HDS Lightning

Controller settings are settings that apply across the entire HDS Lightning controller.

Table 47 lists the HDS Lightning controller settings that are supported by the SAN Volume Controller.

Table 47. HDS Lightning controller settings that are supported by the SAN Volume Controller

Option	HDS Lightning default setting	SAN Volume Controller required setting
PCB mode	Standard	Standard

Port settings for HDS Lightning

Port settings are configurable at the port level.

There are no available options with the scope of a single controller.
- The ports are included in switch zones.
- The switch zones only present the ports directly to the hosts and not to a SAN Volume Controller.

Table 48 lists the HDS Lightning port settings that are supported by the SAN Volume Controller.

Table 48. HDS Lightning port settings supported by the SAN Volume Controller

Option	HDS Lightning default setting	SAN Volume Controller required setting
Address	AL/PA	AL/PA
Fabric	On	On
Connection	Point-to-Point	Point-to-Point
Security switch	On	On or off
Host type	Default	Windows

Logical unit settings for HDS Lightning

Logical unit (LU) settings apply to individual LUs that are configured in the HDS Lightning controller.

HDS Lightning LUs must be configured as described in Table 49 if the LUN is associated with ports in a switch zone that is accessible to the SAN Volume Controller.

Table 49. HDS Lightning LU settings for the SAN Volume Controller

Option	HDS Lightning default setting	SAN Volume Controller required setting
Command device	Off	Off
Command security	Off	Off

Note: These settings only apply to LUs that are accessible by the SAN Volume Controller.

Configuring HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, and HDS TagmaStore WMS systems

You can attach Hitachi Data Systems (HDS) Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, and HDS TagmaStore Workgroup Modular Storage (WMS) systems to a SAN Volume Controller cluster.

Note: In Japan, the HDS Thunder 9200 is referred to as the HDS SANrise 1200. Therefore, the information in this section that refers to the HDS Thunder 9200 also applies to the HDS SANrise 1200.

Supported HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, and HDS TagmaStore WMS models

You can attach certain HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, and HDS TagmaStore Workgroup Modular Storage (WMS) models to SAN Volume Controller clusters.

See the following Web site for the latest supported models:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Supported firmware levels for HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, and HDS TagmaStore WMS

The SAN Volume Controller supports certain HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, and HDS TagmaStore Workgroup Modular Storage (WMS) models.

See the following Web site for specific firmware levels and the latest supported hardware:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Concurrent maintenance on HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, and HDS TagmaStore WMS systems

Concurrent maintenance is the capability to perform I/O operations to a system while simultaneously performing maintenance operations on it.

Important: An HDS Field Engineer must perform all maintenance operations.

The SAN Volume Controller supports concurrent hardware maintenance and firmware upgrade operations on these systems.

User interface on HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, and HDS TagmaStore WMS systems

Ensure that you are familiar with the user interface applications that support the Hitachi Data Systems (HDS) Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, and HDS TagmaStore Workgroup Modular Storage (WMS) systems.

In-band configuration

Disable the system command LUN when you use the user interface applications.

Storage Navigator Modular GUI

The Storage Navigator Modular (SNM) is the primary user interface application for configuring HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, and HDS TagmaStore WMS systems. Use SNM to upgrade firmware, change settings, and to create and monitor storage.

SNM supports an Ethernet connection to the system. An out-of-band command-line interface is available with SNM that supports the majority of the functions that are provided in SNM.

HiCommand

HiCommand is another configuration user interface that is available for the HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, and HDS TagmaStore WMS systems. You must have access to SNM to use HiCommand to configure settings. HiCommand only allows basic creation of storage and provides some monitoring features.

HiCommand uses Ethernet to connect to the system.

Web Server

A Web server runs on each of the controllers on the system. During normal operation, the user interface only provides basic monitoring of the system and displays an event log. If you put a controller into diagnostic mode by pressing the reset button on the controller, the user interface provides firmware upgrades and system configuration resets.

Sharing the HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, or HDS TagmaStore WMS between a host and the SAN Volume Controller

You can share the HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, and HDS TagmaStore Workgroup Modular Storage (WMS) systems between a host and a SAN Volume Controller cluster, with certain restrictions.

The following restrictions apply:

- The same host cannot be connected to both a SAN Volume Controller cluster and an HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, or HDS TagmaStore WMS at the same time because Hitachi Dynamic Link Manager (HDLM) and the subsystem device driver (SDD) cannot coexist.
- For the HDS Thunder 9200, a target port cannot be shared between a host and a SAN Volume Controller cluster. If a target port is used by a SAN Volume Controller cluster, it cannot be present in a switch zone that allows a host to access the port.
- Logical units (LUs) cannot be shared between a host and a SAN Volume Controller cluster. The Thunder 9200 must be set into M-TID M-LUN mode and Mapping Mode must be enabled on Thunder

3 95*xx*. No LU can have a LUN number that is associated with a port that is zoned for host use while
3 also having a LUN number that is associated with a port that is zoned for a SAN Volume Controller
3 cluster.

3 ## Switch zoning limitations for HDS Thunder, Hitachi AMS 200, AMS 500,
3 ## and AMS 1000, or HDS TagmaStore WMS

3 There are limitations in switch zoning for the SAN Volume Controller and the HDS Thunder, Hitachi
3 AMS 200, AMS 500, and AMS 1000, or HDS TagmaStore WMS systems.

3 ### Switch zoning

3 The HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, or HDS TagmaStore WMS systems
3 present themselves to a SAN Volume Controller cluster as separate controllers for each port zoned to the
3 SAN Volume Controller. For example, if one of these storage systems has 4 ports zoned to the SAN
3 Volume Controller, each port appears as a separate controller rather than one controller with 4 WWPNs.
3 In addition, a given logical unit (LU) must be mapped to the SAN Volume Controller through all
3 controller ports zoned to the SAN Volume Controller using the same logical unit number (LUN).

3 ### Supported topologies

3 You can connect the SAN Volume Controller cluster to the HDS Thunder under the following conditions:

3 • For SAN Volume Controller software version 4.2.1 and later, you can connect a maximum of 16 HDS
3   Thunder ports to the SAN Volume Controller cluster without any special zoning requirements.

3 • For SAN Volume Controller software version 4.2.0, the following applies:

3   – Logical Unit Size Expansion (LUSE) and Virtual LVI/LUN operations cannot be run on a disk that is
3   managed by the SAN Volume Controller cluster. LUNs that are created using LUSE and Virtual
3   LVI/LUN can be mapped to the cluster after they are created.

3   – Only disks with open emulation can be mapped to the SAN Volume Controller cluster.

3   – S/390 disks cannot be used with the SAN Volume Controller cluster.

3   – Only fibre-channel connections can connect the SAN Volume Controller cluster to the HDS Thunder.

3 ## Quorum disks on HDS Thunder, Hitachi AMS 200, AMS 500, and AMS
3 ## 1000, and HDS TagmaStore WMS systems

3 When a SAN Volume Controller cluster initializes, the cluster can choose managed disks (MDisks) that
3 are presented by HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, and HDS TagmaStore
3 Workgroup Modular Storage (WMS) systems as quorum disks.

3 You can use the set quorum disk CLI command or the management GUI to select quorum disks.

3 ## Host type for HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000,
3 ## and HDS TagmaStore WMS

3 When the HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, and HDS TagmaStore Workgroup
3 Modular Storage (WMS) systems are attached to a SAN Volume Controller cluster, set the host mode
1 attribute to the Microsoft Windows application that is available on each storage system.

3 For example, when using HDS TagmaStore WMS, select **Windows**, or when using the Hitachi AMS 200,
3 AMS 500, and AMS 1000, select **Windows 2003**.

3 # Advanced functions for HDS Thunder, Hitachi AMS 200, AMS 500, and
3 # AMS 1000, and HDS TagmaStore WMS

3 Some advanced functions of the HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, and HDS
3 TagmaStore Workgroup Modular Storage (WMS), systems are not supported by the SAN Volume
3 Controller clusters.

3 ## Advanced copy functions

3 Advanced copy functions for HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, and HDS
3 TagmaStore WMS systems are not supported for disks that are managed by the SAN Volume Controller
3 clusters because the copy function does not extend to the SAN Volume Controller cache. For example,
3 ShadowImage, TrueCopy, and HiCopy are not supported.

3 ## LUN Security

3 LUN Security enables LUN masking by the worldwide node name (WWNN) of the initiator port. This
3 function is not supported for logical units (LUs) that are used by SAN Volume Controller clusters.

3 ## Partitioning

3 Partitioning splits a RAID array into up to 128 smaller LUs, each of which serves as an independent disk
3 like entity. The SAN Volume Controller cluster and HDS Thunder, Hitachi AMS 200, AMS 500, and AMS
3 1000, and HDS TagmaStore WMS systems support the partitioning function.

3 ## Dynamic array expansion

3 The HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, and HDS TagmaStore WMS systems allow
3 the last LU that is defined in a RAID group to be expanded. This function is not supported when these
3 storage systems are attached to a SAN Volume Controller cluster. Do *not* perform dynamic array
3 expansion on LUs that are in use by a SAN Volume Controller cluster.

3 **Note:** Use in this context means that the LU has a LUN number that is associated with a fibre-channel
3 port, and this fibre-channel port is contained in a switch zone that also contains SAN Volume Controller
3 fibre-channel ports.

3 ## Host storage domains and virtual fibre-channel ports

3 The HDS Thunder 95*xx*V, Hitachi AMS 200, AMS 500, and AMS 1000, and HDS TagmaStore WMS
3 systems support host storage domains (HSD) and virtual fibre-channel ports. Each fibre-channel port can
3 support multiple HSDs. Each host in a given HSD is presented with a virtual target port and a unique set
3 of LUNs.

3 The Thunder 9200 does not support HSD and virtual fibre-channel the ports.

3 # Logical unit creation and deletion on HDS Thunder, Hitachi AMS 200,
3 # AMS 500, and AMS 1000, and HDS TagmaStore WMS systems

3 The HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, and HDS TagmaStore WMS systems
3 Storage Navigator Modular Graphical User Interface (GUI) enables you to create and delete LUNs. You
3 must avoid certain creation and deletion scenarios to prevent data corruption.

Creation and deletion scenarios

For example, the Storage Navigator Modular GUI enables you to create LUN A, delete LUN A, and then create LUN B with the same unique ID as LUN A. If a SAN Volume Controller cluster is attached, data corruption can occur because the cluster might not realize that LUN B is different than LUN A.

Attention: Before you use the Storage Navigator Modular GUI to delete a LUN, remove the LUN from the storage pool that contains it.

Adding LUNs dynamically

To prevent the existing LUNs from rejecting I/O operations during the dynamic addition of LUNs, perform the following procedure to add LUNs:

1. Create the new LUNs using the Storage Navigator Modular GUI.
2. Quiesce all I/O operations.
3. Perform either an offline format or an online format of all new LUNs on the controller using the Storage Navigator Modular GUI. Wait for the format to complete.
4. Go into the LUN mapping function of the Storage Navigator Modular GUI. Add mapping for the new LUN to all of the controller ports that are available to the SAN Volume Controller cluster on the fabric.
5. Restart the controller. (Model 9200 only)
6. After the controller has restarted, restart I/O operations.

LUN mapping considerations

If LUN mapping is used as described in the LUN mapping topic, you must restart the controller to pick up the new LUN mapping configuration. For each storage pool that contains an MDisk that is supported by an LU on the system, all volumes in those storage pools go offline.

Configuring settings for HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, and HDS TagmaStore WMS systems

The Storage Navigator Modular GUI configuration interface provides functions for configuration.

These options and settings can have the following scope:
- System
- Port
- Logical unit

Global settings for the HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, and HDS TagmaStore WMS systems

Global settings apply across HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, and HDS TagmaStore WMS systems.

Table 50 lists the global settings for these disk systems.

Table 50. HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, and HDS TagmaStore WMS systems global settings supported by the SAN Volume Controller

Option	Default setting	SAN Volume Controller required setting
Start attribute	Dual active mode	Dual active mode
SCSI ID/Port takeover mode	Not applicable	Not applicable
Default controller	Not applicable	Not applicable

Option	Default setting	SAN Volume Controller required setting
Data-share mode	Used	Used
Serial number		Same as the system default setting
Delay planned shutdown	0	0
Drive detach mode	False	False
Multipath controller (Thunder 9200 only)	False	False
PROCOM mode	False	False
Report status	False	False
Multipath (Array unit)	False	False
Turbu LU warning	False	False
NX mode	False	False
Auto reconstruction mode	False	False
Forced write-through mode	False	False
Changing logical unit mode 1	False	False
Multiple stream mode (Thunder 9200 only)	False	False
Multiple stream mode (write) (Thunder 95*xx*V only)	False	False
Multiple stream mode (read) (Thunder 95*xx*V only)	False	False
RAID 3 mode (Thunder 9200 only)	False	False
Target ID (9200 only) Mapping mode on 95*xx*	S-TID, M-LUN	M-TID, M-LUN (if sharing controller, otherwise S-TID, M-LUN)
Data striping size	16K; 32K; 64K	Any (Thunder 9200) 64K (Thunder 95*xx*V)
Operation if processor failure occurs	Reset the fault	Reset the fault
Command queuing	True	True
ANSI Version	Not applicable	Not applicable
Vendor ID	HITACHI	HITACHI
Product ID (Thunder 9200)	DF500F	DF500F
Product ID (Thunder 95*xx*V)	DF500F	DF600F
ROM microprogram version	<Empty>	<Empty>
RAM microprogram version	<Empty>	<Empty>
Web title	<Empty>	Any setting supported
Cache mode (Thunder 9200 only)	All off	All off
Link separation (Thunder 9200 only)	False	False
ROM Pseudo-response command processing (Thunder 9200 only)	Not applicable	Not applicable
Save data pointer response (Thunder 9200 only)	Not applicable	Not applicable

3 *Table 50. HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, and HDS TagmaStore WMS systems global*
3 *settings supported by the SAN Volume Controller (continued)*

Option	Default setting	SAN Volume Controller required setting
Controller identifier	False	False
RS232C error information outflow mode	Off	Any
Execute write and verify mode	True	True

3 Controller settings for HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, 3 and HDS TagmaStore WMS systems

3 Controller settings apply across the entire HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, and
3 HDS TagmaStore WMS systems. Options are not available within the scope of a single controller.

3 Port settings for the HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, and 3 HDS TagmaStore WMS systems

3 Port settings are configurable at the port level.

3 The settings listed in Table 51 apply to disk controllers that are in a switch zone that contains SAN
3 Volume Controller nodes. If the system is shared between a SAN Volume Controller cluster and another
3 host, you can configure with different settings than shown if both of the following conditions are true:

3 • The ports are included in switch zones

3 • The switch zones only present the ports directly to the hosts and not to a SAN Volume Controller
3   cluster

3 There are no available options with the scope of a single controller.

3 *Table 51. HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, and HDS TagmaStore WMS system port*
3 *settings supported by the SAN Volume Controller*

Option	Default setting	SAN Volume Controller required setting
Host connection mode 1	Standard	Standard
V*x*VM DMP mode (HDS Thunder 9200 only)	False	False
HP connection mode	False	False
Report inquiry page 83H (HDS Thunder 9200 only)	False	True
UA (06/2A00) suppress mode	False	True
HISUP mode	False	False
CCHS mode	False	False
Standard inquiry data expand (HDS Thunder 9200 only)	False	False
Host connection mode 2	False	False
Product ID DF400 mode	False	False
HBA WWN report mode (HDS Thunder 9200 only)	False	False
NACA mode	False	False
SUN cluster connection mode	False	False
Persistent RSV cluster mode	False	False

Option	Default setting	SAN Volume Controller required setting
ftServer connection mode 1 (HDS Thunder 9200 only)	False	False
ftServer connection mode 2	False	False
SRC Read Command reject	False	False
Reset/LIP mode (signal)	False	False
Reset/LIP mode (progress)	False	False
Reset ALL LIP port mode	False	False
Reset target (reset bus device mode)	False	True
Reserve mode	False	True
Reset logical unit mode	False	True
Reset logout of third party process mode	False	False
Read Frame minimum 128 byte mode (HDS Thunder 950xxV only)	False	False
Topology	Point-to-point	Point-to-point

Logical unit settings for the HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, and HDS TagmaStore WMS systems

Logical unit (LU) settings apply to individual LUs that are configured in the HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, and HDS TagmaStore WMS systems.

You must configure the systems LUs as described in Table 52 if the logical unit number (LUN) is associated with ports in a switch zone that is accessible to the SAN Volume Controller cluster.

Table 52. HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, and HDS TagmaStore WMS systems LU settings for the SAN Volume Controller

Option	Required values	Default setting
LUN default controller	Controller 0 or Controller 1	Any

Note: These settings only apply to LUs that are accessible by the SAN Volume Controller cluster.

Data corruption scenarios to avoid

Scenario 1: The configuration application enables you to change the serial number for an LU. Changing the serial number also changes the unique user identifier (UID) for the LU. Because the serial number is also used to determine the WWPN of the controller ports, two LUNs cannot have the same unique ID on the same SAN because two controllers cannot have the same WWPN on the same SAN.

Scenario 2: The serial number is also used to determine the WWPN of the controller ports. Therefore, two LUNs must not have the same ID on the same SAN because this results in two controllers having the same WWPN on the same SAN. This is not a valid configuration.

Attention: Do not change the serial number for an LU that is managed by a SAN Volume Controller cluster because this can result in data loss or undetected data corruption.

3 **Scenario 3:** The configuration application enables you to create LUN A, delete LUN A, and create LUN B
3 with the same unique ID as LUN A. If the LUN is managed by a SAN Volume Controller cluster, this
3 scenario can cause data corruption because the cluster might not recognize that LUN B is different than
3 LUN A.

3 Mapping and virtualization settings for HDS Thunder, Hitachi AMS 200, AMS 500, 3 and AMS 1000, and HDS TagmaStore WMS systems

3 The HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, and HDS TagmaStore WMS systems
3 support different modes of operation. These modes affect LUN mapping or masking and virtualization.

3 The SAN Volume Controller supports the S-TID M-LUN and M-TID M-LUN modes on Thunder 9200,
3 and Mapping Mode enabled or disabled on Thunder 95*xx*. You must restart the controllers for changes to
3 LUN mapping to take effect.

3 **Attention:** The HDS Thunder, Hitachi AMS 200, AMS 500, and AMS 1000, and HDS TagmaStore WMS
3 systems do not provide an interface that enables a SAN Volume Controller cluster to detect and ensure
3 that the mapping or masking and virtualization options are set properly. Therefore, you must ensure that
3 these options are set as described in this topic.

3 S-TID M-LUN modes

3 In S-TID M-LUN mode all LUs are accessible through all ports on the system with the same LUN
3 number on each port. You can use this mode in environments where the system is not being shared
3 between a host and a SAN Volume Controller cluster.

3 M-TID M-LUN modes

3 If a system is shared between a host and a SAN Volume Controller cluster, you must use M-TID M-LUN
3 mode. Configure the system so that each LU that is exported to the SAN Volume Controller cluster can
3 be identified by a unique LUN. The LUN must be the same on all ports through which the LU can be
3 accessed.

3 Example

3 A SAN Volume Controller cluster can access controller ports x and y. The cluster also sees an LU on port
3 x that has LUN number p. In this situation the following conditions must be met:
3 • The cluster must see either the same LU on port y with LUN number p or it must not see the LU at all
3 on port y.
3 • The LU cannot appear as any other LUN number on port y.
3 • The LU must not be mapped to any system port that is zoned for use directly by a host in a
3 configuration where the system is shared between a host and a cluster.

3 M-TID M-LUN mode enables LU virtualization by target port. In this mode, a single LU can be seen as
3 different LUN numbers across all of the controller ports. For example, LU A can be LUN 0 on port 1,
3 LUN 3 on port 2, and not visible at all on ports 3 and 4.

3 **Important:** The SAN Volume Controller does not support this.

3 In addition, M-TID M-LUN mode enables a single LU to be seen as multiple LUN numbers on the same
3 controller port. For example, LU B can be LUN 1 and LUN 2 on controller port 1.

3 **Important:** The SAN Volume Controller does not support this.

Configuring HDS TagmaStore USP and NSC systems

This section provides information about configuring the Hitachi Data Systems (HDS) TagmaStore Universal Storage Platform (USP) and Network Storage Controller (NSC) systems for attachment to a SAN Volume Controller. Models of the HDS USP and NSC are equivalent to HP and Sun models; therefore, the SAN Volume Controller also supports models of the HP StorageWorks XP series and the Sun StorEdge series.

The information in this section also applies to the supported models of the HP XP and the Sun StorEdge series.

Supported models of the HDS USP and NSC

The SAN Volume Controller supports models of the Hitachi Data Systems (HDS) Universal Storage Platform (USP) and Network Storage Controller (NSC) series. Models of the HDS USP and NSC are equivalent to HP and Sun models; therefore, the SAN Volume Controller also supports models of the Sun StorEdge and the HP XP series.

See the following Web site for the latest supported models:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Supported firmware levels for HDS USP and NSC

The SAN Volume Controller supports the HDS USP and NSC series of controllers.

See the following Web site for specific firmware levels and the latest supported hardware:Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

User interface on the HDS USP and NSC

Ensure that you are familiar with the user interface application that supports the HDS USP and NSC. The HDS USP and NSC is configured, managed, and monitored by a Service Processor (SVP). The SVP is a server that is connected to the HDS USP or NSC through a private local area network (LAN).

Web server

The HDS USP and NSC use the Storage Navigator as the main configuration GUI. The Storage Navigator GUI runs on the SVP and is accessed through a Web browser.

Logical units and target ports on the HDS USP and NSC

Logical units (LUs) that are exported by the HDS USP and NSC report identification descriptors in the vital product data (VPD). The SAN Volume Controller uses the LUN associated binary type-3 IEEE Registered Extended descriptor to identify the LU.

An LU path must be defined before an LU can be accessed by a host. The LU path relates a host group to a target port and to a set of LUs. Host initiator ports are added to the host group by worldwide port name (WWPN).

The HDS USP and NSC do not use LU groups so all LUs are independent. The LU access model is active-active and does not use preferred access ports. Each LU can be accessed from any target port that is mapped to the LU. Each target port has a unique WWPN and worldwide node name (WWNN). The WWPN matches the WWNN on each port.

Note: You must wait until the LU is formatted before presenting it to the SAN Volume Controller.

Special LUs

The HDS USP and NSC can use any logical device (LDEV) as a Command Device. Command Devices are the target for HDS USP or NSC copy service functions. Therefore, do not export Command Devices to a SAN Volume Controller.

Switch zoning limitations for the HDS USP and NSC

There are limitations in switch zoning for the SAN Volume Controller and the HDS USP or NSC.

The SAN Volume Controller can be connected to the HDS USP or NSC with the following restrictions:
- If an LU is mapped to a SAN Volume Controller port as LUN x, the LU must appear as LUN x for all mappings to target ports.
- Only fibre-channel connections can be used to connect a SAN Volume Controller to the HDS USP or NSC system.
- Because the SAN Volume Controller limits the number of worldwide node names (WWNNs) for each storage system and the HDS USP and NSC present a separate WWNN for each port, the number of target ports that the SAN Volume Controller can resolve as one storage system is limited. Perform the following steps to provide connections to more target ports:
  1. Divide the set of target ports into groups of 2 to 16.
  2. Assign a discrete set of LUs to each group.

  The SAN Volume Controller can then view each group of target ports and the associated LUs as separate HDS USP or NSC systems. You can repeat this process to use all target ports.

Note: The HDS USP and NSC systems present themselves to a SAN Volume Controller cluster as separate controllers for each port zoned to the SAN Volume Controller. For example, if one of these storage systems has 4 ports zoned to the SAN Volume Controller, each port appears as a separate controller rather than one controller with 4 WWPNs. In addition, a given logical unit (LU) must be mapped to the SAN Volume Controller through all controller ports zoned to the SAN Volume Controller using the same logical unit number (LUN).

Controller splitting

You can split the HDS USP or NSC between other hosts and the SAN Volume Controller under the following conditions:
- A host cannot be simultaneously connected to both an HDS USP or NSC and a SAN Volume Controller.
- Port security must be enabled for target ports that are shared.
- An LU that is mapped to a SAN Volume Controller cannot be simultaneously mapped to another host.

Concurrent maintenance on the HDS USP and NSC

Concurrent maintenance is the capability to perform I/O operations to an HDS USP or NSC while simultaneously performing maintenance operations on it. Concurrent firmware upgrades are supported with the SAN Volume Controller.

Important: An HDS Field Engineer must perform all maintenance procedures.

Quorum disks on HDS USP and NSC

To host quorum disks on HDS USP and NSC storage systems, you must be aware of the system requirements for establishing quorum disk for these storage systems.

Note: Sun StorEdge systems are not supported to host SAN Volume Controller quorum disks.

The SAN Volume Controller cluster uses a quorum disk to store important cluster configuration data and to break a tie in the event of a SAN failure. The cluster automatically chooses three managed disks (MDisks) as quorum disk candidates. Each disk is assigned an index number: either 0, 1, or 2. Although a cluster can be configured to use up to three quorum disks, only one quorum disk is elected to resolve a tie-break situation. The purpose of the other quorum disks is to provide redundancy if a quorum disk fails before the cluster is partitioned.

Requirements for HDS TagmaStore USP, HP XP10000/12000, and NSC55

To host any of the three quorum disks on these HDS TagmaStore USP, HP XP10000/12000, or NSC55 storage systems, ensure that each of the following conditions have been met:
- Firmware version Main 50-09-72 00/00 or later is running. Contact HDS or HP support for details on installing and configuring the correct firmware version.
- **System Option 562** is enabled. Contact HDS or HP support for details on System Option 562.
- All SAN Volume Controller ports are configured in a single HDS or HP host group.

Requirements for HDS TagmaStore USPv, USP-VM, and HP XP20000/24000

To host any of the three quorum disks on these HDS TagmaStore USPv, USP-VM, or HP XP20000/24000 systems, ensure that each of the following requirements have been met:
- Firmware version Main 60-04-01-00/02 or later is running. Contact HDS or HP support for details on installing and configuring the correct firmware version.
- **Host Option 39** is enabled. Contact HDS or HP support for details on Host Option 39.

  Note: This must be applied to the HDS or HP host group that is used for SAN Volume Controller.
- All SAN Volume Controller ports are configured in a single HDS or HP host group.

After you have verified these requirements for the appropriate storage system, complete the following steps on the SAN Volume Controller command-line interface to set the quorum disks:
1. Issue the **svctask chcontroller** command:

   ```
   svctask chcontroller -allowquorum yes controller_id or controller_name
   ```

   where *controller_id or controller_name* is the controller that corresponds to the relevant HDS or HP storage system.
2. Repeat step 1 for each controller that is part of the relevant HDS or HP storage system.
3. Issue the **svctask setquroum** command:

   ```
   svctask setquroum -quorum [0|1|2] mdisk_id or mdisk_name
   ```

   where *mdisk_id or mdisk_name* is the relevant MDisk on the HDS or HP system.

Attention: Failure to meet these conditions or to follow these steps can result in data corruption.

The Support for SAN Volume Controller (2145) website provides current information about quorum support:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Host type for HDS USP and NSC subsystems

When the HDS USP and NSC systems are attached to a SAN Volume Controller cluster, set the host mode attribute to Windows for each host group.

Advanced functions for HDS USP and NSC

Some advanced functions of the HDS USP and NSC are not supported by the SAN Volume Controller.

Advanced system functions

The following advanced system functions for HDS USP and NSC are not supported for disks that are managed by the SAN Volume Controller:
- TrueCopy
- ShadowImage
- Extended Copy Manager
- Extended Remote Copy
- NanoCopy
- Data migration
- RapidXchange
- Multiplatform Backup Restore
- Priority Access
- HARBOR File-Level Backup/Restore
- HARBOR File Transfer
- FlashAccess

Advanced SAN Volume Controller functions

All advanced SAN Volume Controller functions are supported on logical unit (LU) that are exported by the HDS USP or NSC system.

LU Expansion

The HDS USP and NSC support Logical Unit Expansion (LUSE). LUSE is *not* a concurrent operation. LUSE allows you to create a single LU by concatenating logical devices (LDEVs). Before LUSE can be performed, the LDEVs must be unmounted from hosts and paths must be removed.

Attention:
1. LUSE destroys all data that exists on the LDEV.
2. Do not perform LUSE on any LDEV that is used to export an LU to a SAN Volume Controller.

If data exists on an LDEV and you want to use image mode migration to import the data to a SAN Volume Controller, do not perform LUSE on the disk before you import the data.

LUs that are created using LUSE can be exported to a SAN Volume Controller.

Virtual LVI/LUNs

The HDS USP and NSC support Virtual LVI/LUNs (VLL). VLL is *not* a concurrent operation. VLL allows you to create several LUs from a single LDEV. You can only create new LUs from free space on the LDEV.

Attention: Do not perform VLL on disks that are managed by the SAN Volume Controller.

LUs that are created using VLL can be exported to a SAN Volume Controller.

Configuring Hitachi TagmaStore AMS 2000 family of systems

You can attach Hitachi TagmaStore AMS 2000 family of systems to a SAN Volume Controller cluster.

Supported Hitachi TagmaStore AMS 2000 family of systems models

You can attach certain Hitachi TagmaStore AMS 2000 family of systems models to SAN Volume Controller clusters.

See the following Web site for the latest supported models:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Supported firmware levels for Hitachi TagmaStore AMS 2000 family of systems

The SAN Volume Controller supports certain Hitachi TagmaStore AMS 2000 family of systems models.

See the following Web site for specific firmware levels and the latest supported hardware:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Concurrent maintenance on Hitachi TagmaStore AMS 2000 family of systems

Concurrent maintenance is the capability to perform I/O operations to a system while simultaneously performing maintenance operations on it.

Important: A Hitachi Data Systems (HDS) Field Engineer must perform all maintenance operations.

The SAN Volume Controller supports concurrent hardware maintenance and firmware upgrade operations on these systems.

User interface on Hitachi TagmaStore AMS 2000 family of systems

Ensure that you are familiar with the user interface applications that support the Hitachi TagmaStore AMS 2000 family of systems.

In-band configuration

Disable the system command LUN when you use the user interface applications.

Storage Navigator Modular GUI

The Storage Navigator Modular (SNM) is the primary user interface application for configuring Hitachi TagmaStore AMS 2000 family of systems. Use SNM to upgrade firmware, change settings, and to create and monitor storage.

SNM supports an Ethernet connection to the system. An out-of-band command-line interface is available with SNM that supports the majority of the functions that are provided in SNM.

HiCommand

HiCommand is another configuration user interface that is available for the Hitachi TagmaStore AMS 2000 family of systems. You must have access to SNM to use HiCommand to configure settings. HiCommand only allows basic creation of storage and provides some monitoring features.

3 HiCommand uses Ethernet to connect to the system.

3 Web Server

3 A Web server runs on each of the controllers on the system. During normal operation, the user interface
e only provides basic monitoring of the system and displays an event log. If you put a controller into
e diagnostic mode by pressing the reset button on the controller, the user interface provides firmware
3 upgrades and system configuration resets.

3 Sharing the Hitachi TagmaStore AMS 2000 family of systems between 3 a host and the SAN Volume Controller

3 You can share the Hitachi TagmaStore AMS 2000 family of systems between a host and a SAN Volume
3 Controller cluster, with certain restrictions.

3 The following restrictions apply:

3 • The same host cannot be connected to both a SAN Volume Controller cluster and a Hitachi TagmaStore
3   AMS 2000 family of systems at the same time because Hitachi Dynamic Link Manager (HDLM) and
3   the subsystem device driver (SDD) cannot coexist.

3 • Logical units (LUs) cannot be shared between a host and a SAN Volume Controller cluster. No LU can
3   have a LUN number that is associated with a port that is zoned for host use while also having a LUN
3   number that is associated with a port that is zoned for a SAN Volume Controller cluster.

3 Switch zoning limitations for Hitachi TagmaStore AMS 2000 family of 3 systems

3 There are limitations in switch zoning for the SAN Volume Controller and the Hitachi TagmaStore AMS
3 2000 family of systems.

3 Switch zoning

3 The Hitachi TagmaStore AMS 2000 family of systems present themselves to a SAN Volume Controller
3 cluster as separate controllers for each port zoned to the SAN Volume Controller. For example, if one of
3 these storage systems has 4 ports zoned to the SAN Volume Controller, each port appears as a separate
3 controller rather than one controller with 4 WWPNs. In addition, a given logical unit (LU) must be
3 mapped to the SAN Volume Controller through all controller ports zoned to the SAN Volume Controller
3 using the same logical unit number (LUN).

3 Supported topologies

3 You can connect the SAN Volume Controller cluster to the Hitachi TagmaStore AMS 2000 family of
3 systems under the following conditions:

3 • For SAN Volume Controller software version 4.2.1 and later, you can connect a maximum of 16 Hitachi
3   TagmaStore AMS 2000 family of systems ports to the SAN Volume Controller cluster without any
3   special zoning requirements.

3 • For SAN Volume Controller software version 4.2.0, the following applies:

3   – Logical Unit Size Expansion (LUSE) and Virtual LVI/LUN operations cannot be run on a disk that is
3   managed by the SAN Volume Controller cluster. LUNs that are created using LUSE and Virtual
3   LVI/LUN can be mapped to the cluster after they are created.

3   – Only disks with open emulation can be mapped to the SAN Volume Controller cluster.

3   – S/390 disks cannot be used with the SAN Volume Controller cluster.

3   – Only fibre-channel connections can connect the SAN Volume Controller cluster to the Hitachi
3   TagmaStore AMS 2000 family of systems.

Quorum disks on Hitachi TagmaStore AMS 2000 family of systems

When a SAN Volume Controller cluster initializes, the cluster can choose managed disks (MDisks) that are presented by Hitachi TagmaStore AMS 2000 family of systems as quorum disks.

You can use the set quorum disk CLI command or the management GUI to select quorum disks.

Host type for Hitachi TagmaStore AMS 2000 family of systems

When the Hitachi TagmaStore AMS 2000 family of systems are attached to a SAN Volume Controller cluster, set the host mode attribute to the Microsoft Windows application that is available on each storage system.

For example, when using Hitachi TagmaStore AMS 2000 family of systems, select **Windows 2003**.

Advanced functions for Hitachi TagmaStore AMS 2000 family of systems

Some advanced functions of the Hitachi TagmaStore AMS 2000 family of systems are not supported by the SAN Volume Controller clusters.

Advanced copy functions

Advanced copy functions for Hitachi TagmaStore AMS 2000 family of systems are not supported for disks that are managed by the SAN Volume Controller clusters because the copy function does not extend to the SAN Volume Controller cache. For example, ShadowImage, TrueCopy, and HiCopy are not supported.

LUN Security

LUN Security enables LUN masking by the worldwide node name (WWNN) of the initiator port. This function is not supported for logical units (LUs) that are used by SAN Volume Controller clusters.

Partitioning

Partitioning splits a RAID array into up to 128 smaller LUs, each of which serves as an independent disk like entity. The SAN Volume Controller cluster and Hitachi TagmaStore AMS 2000 family of systems support the partitioning function.

Dynamic array expansion

The Hitachi TagmaStore AMS 2000 family of systems allow the last LU that is defined in a RAID group to be expanded. This function is not supported when these storage systems are attached to a SAN Volume Controller cluster. Do *not* perform dynamic array expansion on LUs that are in use by a SAN Volume Controller cluster.

Note: Use in this context means that the LU has a LUN number that is associated with a fibre-channel port, and this fibre-channel port is contained in a switch zone that also contains SAN Volume Controller fibre-channel ports.

Host storage domains and virtual fibre-channel ports

The Hitachi TagmaStore AMS 2000 family of systems support host storage domains (HSD) and virtual fibre-channel ports. Each fibre-channel port can support multiple HSDs. Each host in a given HSD is presented with a virtual target port and a unique set of LUNs.

Logical unit creation and deletion on Hitachi TagmaStore AMS 2000 family of systems

The Hitachi TagmaStore AMS 2000 family of systems Storage Navigator Modular Graphical User Interface (GUI) enables you to create and delete LUNs. You must avoid certain creation and deletion scenarios to prevent data corruption.

Creation and deletion scenarios

For example, the Storage Navigator Modular GUI enables you to create LUN A, delete LUN A, and then create LUN B with the same unique ID as LUN A. If a SAN Volume Controller cluster is attached, data corruption can occur because the cluster might not realize that LUN B is different than LUN A.

Attention: Before you use the Storage Navigator Modular GUI to delete a LUN, remove the LUN from the storage pool that contains it.

Adding LUNs dynamically

To prevent the existing LUNs from rejecting I/O operations during the dynamic addition of LUNs, perform the following procedure to add LUNs:

1. Create the new LUNs using the Storage Navigator Modular GUI.
2. Quiesce all I/O operations.
3. Perform either an offline format or an online format of all new LUNs on the controller using the Storage Navigator Modular GUI. Wait for the format to complete.
4. Go into the LUN mapping function of the Storage Navigator Modular GUI. Add mapping for the new LUN to all of the controller ports that are available to the SAN Volume Controller cluster on the fabric.
5. Restart the controller. (Model 9200 only)
6. After the controller has restarted, restart I/O operations.

LUN mapping considerations

If LUN mapping is used as described in the LUN mapping topic, you must restart the controller to pick up the new LUN mapping configuration. For each storage pool that contains an MDisk that is supported by an LU on the system, all volumes in those storage pools go offline.

Configuring settings for Hitachi TagmaStore AMS 2000 family of systems

The Storage Navigator Modular GUI configuration interface provides functions for configuration.

These options and settings can have the following scope:
- System
- Port
- Logical unit

Global settings for the Hitachi TagmaStore AMS 2000 family of systems
Global settings apply across Hitachi TagmaStore AMS 2000 family of systems.

Table 53 on page 206 lists the global settings for these disk systems.

3 *Table 53. Hitachi TagmaStore AMS 2000 family of systems global settings supported by the SAN Volume Controller*

Option	Default setting	SAN Volume Controller required setting
Boot options		
System startup attribute	Dual active mode	Dual active mode
Delayed plan shutdown	0	0
Vendor ID	HITACHI	HITACHI
Product ID	DF600F	DF600F
ROM Microcode version		
ROM Microcode version		
System parameters		
Turbu LU warning	Off	Off
Write unique response mode	Off	Off
Auto reconstruct mode	Off	Off
Forced write-through mode	Off	Off
ShadowImage I/O switch mode	Off	Off
Synchronize cache execution mode	Off	Off
Drive detach mode	Off	Off
Operation if processor failure occurs	Reset the fault	Reset the fault
Write and verify execution mode	Off	Off
Web title	<Empty>	Any setting supported
Data strip sizing	256K	256K (recommended)
Topology	Point-to-point	Point-to-point (set under FC settings)

Controller settings for Hitachi TagmaStore AMS 2000 family of systems

Controller settings apply across the entire Hitachi TagmaStore AMS 2000 family of systems. Options are not available within the scope of a single controller.

Port settings for the Hitachi TagmaStore AMS 2000 family of systems

Port settings are configurable at the port level.

The settings listed in Table 54 apply to disk controllers that are in a switch zone that contains SAN Volume Controller nodes. If the system is shared between a SAN Volume Controller cluster and another host, you can configure with different settings than shown if both of the following conditions are true:

* The ports are included in switch zones
* The switch zones only present the ports directly to the hosts and not to a SAN Volume Controller cluster

There are no available options with the scope of a single controller.

Table 54. Hitachi TagmaStore AMS 2000 family of systems port settings supported by the SAN Volume Controller

Option	Default setting	SAN Volume Controller required setting
Port settings		
Mapping mode	On	On
Port type	Fibre	Fibre

Option	Default setting	SAN Volume Controller required setting
Reset LIP mode (signal)	Off	Off
Reset LIP mode (process)	Off	Off
LIP port all reset mode	Off	Off
Host group list		
Host connection mode 1		Windows
HostGroupName	"G000"	"G000"
Middleware	Unsupported	Unsupported
Host system configuration		
Platform		Windows
HostGroupName	"G000"	"G000"
Middleware	Unsupported	Unsupported
Host group information settings		
HostGroupNumber	0	0
HostGroupName	"G000"	"G000"
Host group options		
Host connection mode 1	Standard mode	Standard mode
Host connection mode 2	Off	Off
HP-UX mode	Off	Off
PSUE read reject mode	Off	Off
Mode parameters changed notification mode	Off	Off
NACA mode (AIX only)	Off	Off
Task management isolation mode	Off	Off
Unique reserve mode 1	Off	Off
Port-ID conversion mode	Off	Off
Tru cluster mode	Off	Off
Product serial response mode	Off	Off
Same node name mode	Off	Off
CCHS mode	Off	Off
Inquiry serial number conversion mode	Off	Off
NOP-In suppress mode	Off	Off
S-VOL disable advanced mode	Off	Off
Discovery CHAP mode	Off	Off

Logical unit settings for the Hitachi TagmaStore AMS 2000 family of systems

Logical unit (LU) settings apply to individual LUs that are configured in the Hitachi TagmaStore AMS 2000 family of systems.

3 You must configure the systems LUs as described in Table 55 if the logical unit number (LUN) is
3 associated with ports in a switch zone that is accessible to the SAN Volume Controller cluster.

3 *Table 55. Hitachi TagmaStore AMS 2000 family of systems LU settings for the SAN Volume Controller*

Option	Default setting	SAN Volume Controller required setting
LUN management information		
Security	Off	Off **Note:** LUN Security enables LUN masking by the worldwide node name (WWNN) of the initiator port. This function is not supported for logical units (LUs) that are used by SAN Volume Controller clusters.
LU mapping	One-to-one	One-to-one
LAN management options		
Maintenance port IP address automatic change mode	Off	Off
IPv4 DHCP	Off	Off
IPv6 address setting mode	Auto	Auto
Negotiation	Auto	Auto

3 **Note:** These settings only apply to LUs that are accessible by the SAN Volume Controller cluster.

3 Data corruption scenarios to avoid

3 **Scenario 1:** The configuration application enables you to change the serial number for an LU. Changing
3 the serial number also changes the unique user identifier (UID) for the LU. Because the serial number is
3 also used to determine the WWPN of the controller ports, two LUNs cannot have the same unique ID on
3 the same SAN because two controllers cannot have the same WWPN on the same SAN.

3 **Scenario 2:** The serial number is also used to determine the WWPN of the controller ports. Therefore,
3 two LUNs must not have the same ID on the same SAN because this results in two controllers having the
3 same WWPN on the same SAN. This is not a valid configuration.

3 **Attention:** Do not change the serial number for an LU that is managed by a SAN Volume Controller
3 cluster because this can result in data loss or undetected data corruption.

3 **Scenario 3:** The configuration application enables you to create LUN A, delete LUN A, and create LUN B
3 with the same unique ID as LUN A. If the LUN is managed by a SAN Volume Controller cluster, this
3 scenario can cause data corruption because the cluster might not recognize that LUN B is different than
3 LUN A.

3 Mapping and virtualization settings for Hitachi TagmaStore AMS 2000 family of 3 systems
3 The Hitachi TagmaStore AMS 2000 family of systems support different modes of operation. These modes
3 affect LUN mapping or masking and virtualization.

3 **Attention:** The Hitachi TagmaStore AMS 2000 family of systems do not provide an interface that enables
3 a SAN Volume Controller cluster to detect and ensure that the mapping or masking and virtualization
3 options are set properly. Therefore, you must ensure that these options are set as described in this topic.

3 **S-TID M-LUN modes**

3 In S-TID M-LUN mode all LUs are accessible through all ports on the system with the same LUN
3 number on each port. You can use this mode in environments where the system is not being shared
3 between a host and a SAN Volume Controller cluster.

3 **M-TID M-LUN modes**

3 If a system is shared between a host and a SAN Volume Controller cluster, you must use M-TID M-LUN
3 mode. Configure the system so that each LU that is exported to the SAN Volume Controller cluster can
3 be identified by a unique LUN. The LUN must be the same on all ports through which the LU can be
3 accessed.

3 **Example**

3 A SAN Volume Controller cluster can access controller ports x and y. The cluster also sees an LU on port
3 x that has LUN number p. In this situation the following conditions must be met:

3 • The cluster must see either the same LU on port y with LUN number p or it must not see the LU at all
3   on port y.

3 • The LU cannot appear as any other LUN number on port y.

3 • The LU must not be mapped to any system port that is zoned for use directly by a host in a
3   configuration where the system is shared between a host and a cluster.

3 M-TID M-LUN mode enables LU virtualization by target port. In this mode, a single LU can be seen as
3 different LUN numbers across all of the controller ports. For example, LU A can be LUN 0 on port 1,
3 LUN 3 on port 2, and not visible at all on ports 3 and 4.

3 **Important:** The SAN Volume Controller does not support this.

3 In addition, M-TID M-LUN mode enables a single LU to be seen as multiple LUN numbers on the same
3 controller port. For example, LU B can be LUN 1 and LUN 2 on controller port 1.

3 **Important:** The SAN Volume Controller does not support this.

3

Configuring HP StorageWorks MA and EMA systems

This section provides information about configuring HP StorageWorks Modular Array (MA) and
Enterprise Modular Array (EMA) systems for attachment to a SAN Volume Controller.

Both the HP MA and EMA use an HSG80 controller.

HP MA and EMA definitions

The following terms are used in the IBM and HP documentation and have different meanings.

IBM term	IBM definition	HP term	HP definition
container	A visual user-interface component that holds objects.	container	(1) Any entity that is capable of storing data, whether it is a physical device or a group of physical devices. (2) A virtual, internal controller structure representing either a single disk or a group of disk drives that are linked as a storageset. Stripesets and mirrorsets are examples of storageset containers that the controller uses to create units.
device	A piece of equipment that is used with the computer. A device does not generally interact directly with the system, but is controlled by a controller.	device	In its physical form, a magnetic disk that can be attached to a SCSI bus. The term is also used to indicate a physical device that has been made part of a controller configuration; that is, a physical device that is known to the controller. Units (volumes) can be created from devices, after the devices have been made known to the controller.
just a bunch of disks (JBOD)	See *non-RAID*.	just a bunch of disks (JBOD)	A group of single-device logical units not configured into any other container type.
mirrorset	See *RAID 1*.	mirrorset	A RAID storageset of two or more physical disks that maintains a complete and independent copy of all data on the volume. This type of storageset has the advantage of being highly reliable and extremely tolerant of device failure. Raid level 1 storagesets are called mirrorsets.
non-RAID	Disks that are not in a redundant array of independent disks (RAID).	non-RAID	See *just a bunch of disks*.
RAID 0	RAID 0 allows a number of disk drives to be combined and presented as one large disk. RAID 0 does not provide any data redundancy. If one drive fails, all data is lost.	RAID 0	A RAID storageset that stripes data across an array of disk drives. A single logical disk spans multiple physical disks, allowing parallel data processing for increased I/O performance. While the performance characteristics of RAID level 0 is excellent, this RAID level is the only one that does not provide redundancy. Raid level 0 storagesets are referred to as stripesets.
RAID 1	A form of storage array in which two or more identical copies of data are maintained on separate media. Also known as mirrorset.	RAID 1	See *mirrorset*.

IBM term	IBM definition	HP term	HP definition
RAID 5	A form of parity RAID in which the disks operate independently, the data strip size is no smaller than the exported block size, and parity check data is distributed across the disks in the array.	RAID 5	See *RAIDset*.
RAIDset	See *RAID 5*.	RAIDset	A specially developed RAID storageset that stripes data and parity across three or more members in a disk array. A RAIDset combines the best characteristics of RAID level 3 and RAID level 5. A RAIDset is the best choice for most applications with small to medium I/O requests, unless the application is write intensive. A RAIDset is sometimes called parity RAID. RAID level 3/5 storagesets are referred to as RAIDsets.
partition	A logical division of storage on a fixed disk.	partition	A logical division of a container represented to the host as a logical unit.
stripeset	See *RAID 0*.	stripeset	See *RAID 0*.

Configuring HP MA and EMA systems

The HP MA and EMA systems provide functions that are compatible with the SAN Volume Controller.

This task assumes that the system is not in use.

Note: When you configure a SAN Volume Controller cluster to work with an HP MA or EMA, you must not exceed the limit of 96 process logins.

Perform the following procedure to enable support of an HP, MA, or EMA system.

1. Verify that the front panel of the SAN Volume Controller is clear of errors.
2. Ensure that the HP StorageWorks Operator Control Panel (OCP) on each system is clear of errors. The Operator Control Panel consists of seven green LEDs at the rear of each HSG80 controller.
3. Ensure that you can use an HP StorageWorks command-line interface (CLI) to configure the HSG80 controllers.
4. Issue the **SHOW THIS** command and **SHOW OTHER** command to verify the following:
   a. Ensure that the system firmware is at a supported level. See the following Web site for the latest firmware support:

      Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145.
   b. Ensure that the controllers are configured for MULTIBUS FAILOVER with each other.
   c. Ensure that the controllers are running in SCSI-3 mode.
   d. Ensure that MIRRORED_CACHE is enabled.
   e. Ensure that the Host Connection Table is *not* locked.
5. Issue the **SHOW DEVICES FULL** command to verify the following:
   a. Ensure that none of the LUNs are TRANSPORTABLE.

b. Ensure that all LUNs are configured. For example, the LUNs report their serial numbers and TRANSFER_RATE_REQUESTED correctly.
6. Issue the **SHOW FAILEDSET** command to verify that there are no failing disks.

   Note: To verify, there should be no orange lights on any disks in the system.
7. Issue the **SHOW UNITS FULL** command to verify the following:
   a. Ensure that all LUNs are set to RUN and NOWRITEPROTECT.
   b. Ensure that all LUNs are ONLINE to either THIS or OTHER controller.
   c. Ensure that all LUNs that are to be made available to the SAN Volume Controller have ALL access.
   d. Ensure that all LUNs do not specify Host Based Logging.
8. Issue the **SHOW CONNECTIONS FULL** command to verify that you have enough spare entries for all combinations of SAN Volume Controller ports and HP MA or EMA ports.
9. Connect up to four fibre-channel cables between the fibre-channel switches and the HP MA or EMA system.
10. Ensure that the fibre-channel switches are zoned so that the SAN Volume Controller and the HP MA or EMA system are in a zone.
11. Issue the **SHOW THIS** command and **SHOW OTHER** command to verify that each connected port is running. The following is an example of the output that is displayed: `PORT_1_TOPOLOGY=FABRIC`.
12. Issue the **SHOW CONNECTIONS FULL** command to verify that the new connections have been created for each SAN Volume Controller port and HP MA or EMA port combination.
13. Verify that `No rejected hosts` is displayed at the end of the SHOW CONNECTIONS output.
14. Perform the following steps from the SAN Volume Controller command-line interface (CLI):

2   a. Issue the **svctask detectmdisk** CLI command to discover the controller.

2   b. Issue the **svcinfo lscontroller** CLI command to verify that the two HSG80 serial numbers appear
2   under the `ctrl s/n`.

2   c. Issue the **svcinfo lsmdisk** CLI command to verify that the additional MDisks that correspond to
2   the UNITS shown in the HP MA or EMA system.

t   You can now use the SAN Volume Controller CLI commands to create a storage pool. You can also create
t   and map volumes from these storage pools. Check the front panel of the SAN Volume Controller to
    ensure that there are no errors. After the host has reloaded the fibre-channel driver, you can perform I/O
2   to the volumes. For more details, see the host attachment information.

Partitioning LUNs on HP MA and EMA systems

For SAN Volume Controller software version 4.2.1 and later, you cannot partition HSG80 LUNs. To check if any HSG80 LUNs are partitioned, use the SHOW UNITS command in the HSG80 CLI. Partition is displayed in the Used By column for the LUNs that are partitioned.

Supported models of HP MA and EMA systems

The SAN Volume Controller supports models of the HP MA and EMA systems.

See the following Web site for the latest supported models:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Attention: The SAN Volume Controller only supports configurations in which the HSG80 cache is enabled in writeback mode. Running with only a single controller results in a single point of data loss.

Supported firmware levels for HP MA and EMA systems

The HP MA and EMA systems must use a firmware level that is supported by the SAN Volume Controller.

See the following Web site for specific firmware levels and the latest supported hardware:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Note: Concurrent upgrade of the system firmware is not supported with the SAN Volume Controller.

Concurrent maintenance on HP MA and EMA systems

Concurrent maintenance is the capability to perform I/O operations to an HP MA or EMA system while simultaneously performing maintenance operations on it.

Note: HP MA and EMA maintenance documentation uses the phrase *rolling upgrade* in place of *concurrent maintenance*. See this documentation because in some instances you must reduce the level of I/O before you can perform the maintenance procedure.

The HP MA and EMA systems allow concurrent replacement of the following components:
- Drive
- EMU
- Blower
- Dual power supply (One unit can be removed and replaced. The fan speed increases when only one power supply unit is present.)

The controller component is hot-pluggable, but concurrent maintenance of SAN Volume Controller I/O is not supported.

The HP MA and EMA systems do not allow concurrent replacement of the following components:
- Single power supply (in a single power-supply configuration, the enclosure is disabled when the power supply fails.)
- SCSI bus cables
- I/O module
- Cache

Configuration interface for HP MA and EMA systems

The Command Console configuration and service utility is the configuration interface for the HP MA and EMA systems.

The configuration and service utility can connect to the system in the following ways:
- RS232 interface
- In-band over fibre channel
- Over TCP/IP to a proxy agent, which then communicates with the system in-band over fibre channel.

For the Command Console to communicate with the HSG80 controllers, the host that runs the service utility must be able to access the HSG80 ports over the SAN. This host can therefore also access LUs that are visible to SAN Volume Controller nodes and cause data corruption. To avoid this, set the UNIT_OFFSET option to 199 for all connections to this host. This ensures that the host is able to recognize only the CCL.

Sharing the HP MA or EMA between a host and a SAN Volume Controller

There are restrictions for sharing HP MA and EMA subsystems between a host and a SAN Volume Controller cluster.

An HP MA or EMA can be shared between a host and a SAN Volume Controller cluster under the following conditions:

- A host cannot be connected to both a SAN Volume Controller cluster and an HP MA or EMA subsystem at the same time.
- Target ports cannot be shared between a host and a SAN Volume Controller cluster. Specifically, if an HSG80 port is in use by a SAN Volume Controller cluster, it cannot be present in a switch zone that enables a host to access the port.
- LUs and RAID arrays cannot be shared between a host and a SAN Volume Controller cluster.

Switch zoning limitations for HP MA and EMA systems

There are limitations in switch zoning for the SAN Volume Controller and the HP MA and EMA systems.

Attention: The HP MA and EMA systems are supported with a single HSG80 controller or dual HSG80 controllers. Because the SAN Volume Controller only supports configurations in which the HSG80 cache is enabled in write-back mode, running with a single HSG80 controller results in a single point of data loss.

Switch zoning

For SAN Volume Controller clusters that have installed software version 1.1.1, a single fibre-channel port that is attached to the system can be present in a switch zone that contains SAN Volume Controller fibre-channel ports, whether the HP MA or EMA system uses one or two HSG80 controllers. This guarantees that the nodes in the cluster can access at most one port on the HSG80 controller.

For SAN Volume Controller clusters that have software version 1.2.0 or later installed, switches can be zoned so that HSG80 controller ports are in the switch zone that contains all of the ports for each SAN Volume Controller node.

Connecting to the SAN

Multiple ports from an HSG80 controller must be physically connected to the fibre-channel SAN to enable servicing of the HP MA or EMA system. However, switch zoning must be used as described in this topic.

Note: If the HP Command Console is not able to access a fibre-channel port on each of the HSG80 controllers in a two-controller system, there is a risk of an undetected single point of failure.

Quorum disks on HP MA and EMA systems

Managed disks (MDisks) that are presented by the HP MA or EMA are chosen by the SAN Volume Controller as quorum disks.

The SAN Volume Controller uses a logical unit (LU) that is presented by an HSG80 controller as a quorum disk. The quorum disk is used even if the connection is by a single port, although this is not recommended. If you are connecting the HP MA or EMA system with a single fibre-channel port, ensure that you have another system on which to put your quorum disk. You can use the **svctask setquorum** command-line interface (CLI) command to move quorum disks to another system.

SAN Volume Controller clusters that are attached only to the HSG80 controllers are supported.

Advanced functions for HP MA and EMA

Some advanced functions of the HP MA and EMA are not supported by the SAN Volume Controller.

Advanced copy functions

Advanced copy functions for HP MA and EMA systems (for example, SnapShot and RemoteCopy) are not supported for disks that are managed by the SAN Volume Controller because the copy function does not extend to the SAN Volume Controller cache.

Partitioning

HP MA and EMA support partitioning. A partition is a logical division of a container that is represented to the host as a logical unit (LU). A container can be a RAID array or a JBOD (just a bunch of disks). All container types are candidates for partitions. Any nontransportable disk or storage set can be divided into a maximum of eight partitions.

The following restrictions apply to partitioning:
- Partitioned containers are fully supported if the HSG80 controller is connected to the SAN by a single port.
- Partitioned containers are not configured by the SAN Volume Controller if the HSG80 controller is connected to the SAN by multiple ports.
- Partitioned containers are removed from the configuration if a single port connection becomes a multiport connection.
- Partitioned containers are configured if a multiport connection becomes a single port connection.

You must partition containers such that no spare capacity exists because there is no way to detect unused partitions. With a multiport connection, subsequent attempts to use this capacity removes all partitions on the container from the configuration.

Dynamic array expansion (LU expansion)

HP MA and EMA systems do not provide dynamic array expansion.

Write protection of LUNs

Write protection of LUNs is not supported for use with the SAN Volume Controller.

SAN Volume Controller advanced functions

t Volumes that are created from managed disks (MDisks) that are presented by an HSG80 controller can be used in SAN Volume Controller FlashCopy mappings, SAN Volume Controller Metro Mirror relationships, and SAN Volume Controller Global Mirror relationships.

LU creation and deletion on the HP MA and EMA

Ensure you are familiar with the HSG80 controller container types for logical unit (LU) configuration.

Table 56 on page 216 lists the valid container types.

Table 56. HSG80 controller container types for LU configuration

Container	Number of Members	Maximum Size
JBOD - non transportable **Attention:** A JBOD provides no redundancy at the physical disk-drive level. A single disk failure can result in the loss of an entire storage pool and its associated volumes.	1	disk size minus metadata
Mirrorset	2 - 6	smallest member
RAIDset	3 - 14	1.024 terabytes
Stripeset	2 - 24	1.024 terabytes
Striped Mirrorset	2 - 48	1.024 terabytes

Note: LUs can be created and deleted on an HSG80 controller while I/O operations are performed to other LUs. You do not need to restart the HP MA or EMA subsystem.

Configuring settings for the HP MA and EMA

The HP StorageWorks configuration interface provides configuration settings and options that are supported with the SAN Volume Controller.

The settings and options can have a scope of the following:
- Subsystem (global)
- Controller
- Port
- Logical unit
- Connection

Global settings for HP MA and EMA systems

Global settings apply across HP MA and EMA systems.

The following table lists the global settings for HP MA and EMA systems:

Table 57. HP MA and EMA global settings supported by the SAN Volume Controller

Option	HSG80 controller default setting	SAN Volume Controller required setting
DRIVE_ERROR_THRESHOLD	800	Default
FAILEDSET	Not defined	n/a

Controller settings for HP MA and EMA

Controller settings apply across one HSG80 controller.

Table 58 describes the options that can be set by HSG80 controller command-line interface (CLI) commands for each HSG80 controller.

Table 58. HSG80 controller settings that are supported by the SAN Volume Controller

Option	HSG80 controller default setting	SAN Volume Controller required setting
ALLOCATION_CLASS	0	Any value
CACHE_FLUSH_TIME	10	Any value

Option	HSG80 controller default setting	SAN Volume Controller required setting
COMMMAND_CONSOLE_LUN	Not defined	Any value
CONNECTIONS_UNLOCKED	CONNECTIONS_ UNLOCKED	CONNECTIONS_ UNLOCKED
NOIDENTIFIER	Not defined	No identifier
MIRRORED_CACHE	Not defined	Mirrored
MULTIBUS_FAILOVER	Not defined	MULTIBUS_FAILOVER
NODE_ID	Worldwide name as on the label	Default
PROMPT	None	Any value
REMOTE_COPY	Not defined	Any value
SCSI_VERSION	SCSI-2	SCSI-3
SMART_ERROR_EJECT	Disabled	Any value
TERMINAL_PARITY	None	Any value
TERMINAL_SPEED	9600	Any value
TIME	Not defined	Any value
UPS	Not defined	Any value

Port settings for HP MA and EMA systems

Port settings are configurable at the port level.

Restriction: Only one port per HSG80 pair can be used with the SAN Volume Controller.

The port settings are set using the following commands:
- SET THIS PORT_1_TOPOLOGY=FABRIC
- SET THIS PORT_2_TOPOLOGY=FABRIC
- SET OTHER PORT_1_TOPOLOGY=FABRIC
- SET OTHER PORT_2_TOPOLOGY=FABRIC

These values can be checked using the following commands:
- SHOW THIS
- SHOW OTHER

Table 59 lists the HSG80 controller port settings that the SAN Volume Controller supports:

Table 59. HSG80 controller port settings supported by the SAN Volume Controller

Option	HSG80 default setting	SAN Volume Controller required setting
PORT_1/2-AL-PA	71 or 72	Not applicable
PORT_1/2_TOPOLOGY	Not defined	FABRIC

Note: The HP MA and EMA systems support LUN masking that is configured with the **SET** *unit number* **ENABLE_ACCESS_PATH** command. When used with a SAN Volume Controller, the access path must be set to all ("SET *unit number* ENABLE_ACCESS_PATH=ALL") and all LUN masking must be handled exclusively by the SAN Volume Controller. You can use the **SHOW CONNECTIONS FULL** command to check access rights.

LU settings for HP MA and EMA systems

Logical unit (LU) settings are configurable at the LU level.

Table 60 describes the options that must be set for each LU that is accessed by the SAN Volume Controller. LUs that are accessed by hosts can be configured differently.

Table 60. HSG80 controller LU settings supported by the SAN Volume Controller

Option	HSG80 controller default setting	SAN Volume Controller required setting
TRANSFER_RATE_REQUESTED	20MHZ	Not applicable
TRANSPORTABLE/ NOTRANSPORTABLE	NOTRANSPORTABLE	NOTRANSPORTABLE
ENABLE_ACCESS_PATH	ENABLE_ACCESS_ PATH=ALL	ENABLE_ACCESS_ PATH=ALL
DISABLE_ACCESS_PATH (See Note.)	NO DEFAULT	NO DEFAULT
IDENTIFIER/ NOIDENTIFIER	NOIDENTIFIER	Not applicable
MAX_READ_CACHE_SIZE	32	Not applicable
MAX_WRITE_CACHE_SIZE	32	64 or higher
MAX_CACHED_TRANSFER_SIZE	32	Not applicable
PREFERRED_PATH/ NOPREFERRED_PATH	NOPREFERRED_PATH is set	Not applicable
READ_CACHE/ NOREAD_CACHE	READ_CACHE	Not applicable
READAHEAD_CACHE/ NOREADAHEAD_CACHE	READAHEAD_CACHE	Not applicable
RUN/ NORUN	RUN	RUN
WRITE_LOG/NOWRITE_LOG	NOWRITE_LOG	NOWRITE_LOG
WRITE_PROTECT/ NOWRITE_PROTECT	NOWRITE_PROTECT	NOWRITE_PROTECT
WRITEBACK_CACHE/ NOWRITEBACK_CACHE	WRITEBACK_CACHE	WRITEBACK_CACHE
Note: DISABLE_ACCESS_PATH can be used to disable access from specific hosts. It must always be overridden by using ENABLE_ACCESS_PATH=ALL on all connections to the SAN Volume Controller nodes.		

Connection settings for HP MA and EMA systems

The HP MA and EMA systems provide options that are configurable at the connection level.

Table 61 lists the default and required HSG80 controller connection settings:

Table 61. HSG80 connection default and required settings

Option	HSG80 controller default setting	HSG80 controller required setting
OPERATING_SYSTEM	Not defined	WINNT
RESERVATION_STYLE	CONNECTION_BASED	Not applicable
UNIT_OFFSET	0	0 or 199

Mapping and virtualization settings for HP MA and EMA

There are LUN mapping or masking and virtualization restrictions for HP MA and EMA subsystems that are in a SAN Volume Controller environment.

The HP StorageWorks configuration interface requires that you assign a unit number to each logical unit (LU) when it is defined. By default, the LUN is the unit number. It is possible for gaps to exist in the

LUN range if the unit numbers that are used in the configuration commands are not contiguous. By default, each LUN is visible on all controller ports on both controllers.

LUN masking

The HP MA and EMA subsystems support the concept of connection names. A maximum of 96 connection names that contain the following parameters are supported:

- HOST_ID
- ADAPTER_ID
- CONTROLLER
- PORT
- REJECTED_HOST

Note: The SAN Volume Controller ports must not be in the REJECTED_HOSTS list. This list can be seen with the **SHOW CONNECTIONS FULL** command.

You cannot use LUN masking to restrict the initiator ports or the target ports that the SAN Volume Controller uses to access LUs. Configurations that use LUN masking in this way are not supported. LUN masking can be used to prevent other initiators on the SAN from accessing LUs that the SAN Volume Controller uses, but the preferred method for this is to use SAN zoning.

LU virtualization

The HP MA and EMA subsystems also provide LU virtualization by the port and by the initiator. This is achieved by specifying a UNIT_OFFSET for the connection. The use of LU virtualization for connections between the HSG80 controller target ports and SAN Volume Controller initiator ports is not supported.

Configuring HP StorageWorks EVA systems

This section provides information about configuring the HP StorageWorks Enterprise Virtual Array (EVA) system for attachment to a SAN Volume Controller.

Supported models of the HP EVA

The SAN Volume Controller supports models of the HP EVA.

See the following Web site for the latest supported models:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Supported firmware levels for HP EVA

The SAN Volume Controller supports HP EVA.

See the following Web site for specific HP EVA firmware levels and the latest supported hardware:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Concurrent maintenance on the HP EVA

Concurrent maintenance is the capability to perform I/O operations to an HP EVA while simultaneously performing maintenance operations on it.

Important: All maintenance operations must be performed by an HP Field Engineer.

The SAN Volume Controller and HP EVA support concurrent hardware maintenance and firmware upgrade.

User interface on the HP EVA system

Ensure that you are familiar with the user interface that supports the HP EVA system.

Storage Management Appliance

HP EVA systems are configured, managed, and monitored through a Storage Management Appliance. The Storage Management Appliance is a PC server that runs a software agent called Command View EVA. The software agent is accessed using a user interface that is provided by a standard Web browser.

Command View EVA communicates in-band with the HSV controllers.

Sharing the HP EVA controller between a host and the SAN Volume Controller

The HP EVA controller can be shared between a host and a SAN Volume Controller.

- A host must not be connected to both a SAN Volume Controller and an HP EVA system at the same time.
- LUs and RAID arrays must not be shared between a host and a SAN Volume Controller.

Switch zoning limitations for the HP EVA system

Consider the following limitations when planning switch zoning and connection to the SAN.

Fabric zoning

The SAN Volume Controller switch zone must include at least one target port from each HSV controller in order to have no single point of failure.

Quorum disks on HP StorageWorks EVA systems

The SAN Volume Controller cluster selects managed disks (MDisks) that are presented by HP StorageWorks EVA systems as quorum disks.

Copy functions for HP StorageWorks EVA systems

Advanced copy functions for HP StorageWorks EVA systems (for example, VSnap and SnapClone) cannot be used with disks that are managed by the SAN Volume Controller cluster because the copy function does not extend to the SAN Volume Controller cache.

Logical unit configuration on the HP EVA

An EVA logical unit is referred to as a virtual disk (VDisk). An EVA system can support up to 512 VDisks. VDisks are created within a set of physical disk drives, referred to as a disk group. A VDisk is striped across all the drives in the group.

The minimum size of a disk group is eight physical drives. The maximum size of a disk group is all available disk drives.

EVA VDisks are created and deleted using the Command View EVA utility.

Note: A VDisk is formatted during the creation process; therefore, the capacity of the VDisk will determine the length of time it takes to be created and formatted. Ensure that you wait until the VDisk is created before you present it to the SAN Volume Controller.

A single VDisk can consume the entire disk group capacity or the disk group can be used for multiple VDisks. The amount of disk group capacity consumed by a VDisk depends on the VDisk capacity and the selected redundancy level. There are three redundancy levels:

- Vraid 1 - High redundancy (mirroring)
- Vraid 5 - Moderate redundancy (parity striping)
- Vraid 0 - No redundancy (striping)

Logical unit creation and deletion on the HP EVA

t EVA volumes are created and deleted using the Command View EVA utility.

t Volumes are formatted during creation. The time it takes to format the volumes depends on the capacity.

Note: Selecting a host for presentation at creation time is not recommended. Ensure that you wait until
t the volume has been created before presenting it to the SAN Volume Controller.

t ## Logical unit presentation

t A volume must be explicitly presented to a host before it can be used for I/O operations.

The SAN Volume Controller supports LUN masking on an HP EVA controller. When presenting a
t volume, the LUN can be specified or allowed to default to the next available value.

The SAN Volume Controller supports LUN virtualization on an HP EVA controller. The LUN-host relationship is set on a per-host basis.

Note: All nodes and ports in the SAN Volume Controller cluster must be represented as one host to the HP EVA.

Special LUs

t The Console LU is a special volume that represents the SCSI target device. It is presented to all hosts as LUN 0.

Configuration interface for the HP EVA

The HP EVA is configured, managed, and monitored through a Storage Management Appliance. The Storage Management Appliance is a server that runs a software agent called Command View EVA. The Command View EVA is accessed using a graphical user interface that is provided by a standard Web browser.

In band

The Command View EVA system communicates in-band with the HSV controllers.

Configuration settings for HP StorageWorks EVA systems

The HP StorageWorks EVA configuration interface provides configuration settings and options that can be used with SAN Volume Controller clusters.

The settings and options can have a scope of the following:

- System (global)
- Logical unit (LU)
- Host

Global settings for HP StorageWorks EVA systems

Global settings apply across an HP StorageWorks EVA system.

Table 62 lists the system options that you can access using the Command View EVA.

Table 62. HP StorageWorks EVA global options and required settings

Option	HP EVA default setting	SAN Volume Controller required setting
Console LUN ID	0	Any
Disk replacement delay	1	Any

Logical unit options and settings for HP StorageWorks EVA systems

Logical unit (LU) settings are configurable at the LU level.

Table 63 describes the options that must be set for each LU that is accessed by other hosts. LUs that are accessed by hosts can be configured differently.

Table 63. HP StorageWorks EVA LU options and required settings

Option	HP EVA Default Setting	SAN Volume Controller Required Setting
Capacity	None	Any
Write cache	Write-through or Write-back	Write-back
Read cache	On	On
Redundancy	Vraid0	Any
Preferred path	No preference	No preference
Write protect	Off	Off

Host options and settings for HP StorageWorks EVA systems

You must use specific settings to identify SAN Volume Controller clusters as hosts to HP StorageWorks EVA systems.

Table 64 lists the host options and settings that can be changed using the Command View EVA.

Table 64. HP EVA host options and required settings

Option	HP EVA Default Setting	SAN Volume Controller Required Setting
OS type	-	Windows
Direct eventing	Disabled	Disabled

Configuring HP StorageWorks MSA1000 and MSA1500 systems

This section provides information about configuring HP StorageWorks Modular Smart Array (MSA) 1000 and 1500 (MSA1000 and MSA1500) systems for attachment to a SAN Volume Controller.

Supported models of the HP MSA1000 and MSA1500 system

The SAN Volume Controller supports models of the HP MSA series of systems.

See the following Web site for the latest supported models:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Supported firmware levels for the HP MSA1000 and MSA1500

The HP MSA system must use a firmware level that is supported by the SAN Volume Controller.

See the following Web site for specific firmware levels and the latest supported hardware: Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145.

User interfaces on the HP MSA1000 and MSA1500

Ensure that you are familiar with the user interface applications that are used by the HP MSA1000 and MSA1500 systems.

You can use the following configuration utilities with HP MSA1000 or MSA1500 systems in a SAN Volume Controller environment:

- The CLI through an out-of-band configuration that is accessed through a host that is connected to the serial port of the HP MSA1000 or MSA1500.
- The GUI through an in-band configuration that uses the HP Array Configuration Utility (ACU).

   Notes:

   1. If the HP ACU is installed in a configuration that HP does not support, some of its functionality might not be available.
   2. If you use an in-band configuration, you must ensure that LUs that are used by the SAN Volume Controller cannot be accessed by a direct-attached host.

Logical unit creation, deletion, and migration for HP StorageWorks MSA systems

Before you create, delete, or migrate logical units, you must read the storage configuration guidelines that are specified in the HP StorageWorks MSA1000 or MSA1500 documentation that is provided for this system.

Creating arrays

An array is a collection of physical disks. Use the storage configuration guidelines for SAN Volume Controller clusters to create arrays on the HP StorageWorks MSA.

Creating logical drives

The following types of RAID arrays are supported:
- RAID 1+0
- RAID 1
- RAID 5
- RAID 6 (ADG)

RAID 0 is not supported because it does not provide failure protection.

All stripe sizes are supported; however, use a consistent stripe size for the HP StorageWorks MSA.

Use the following settings for logical drives:
- Set Max Boot to disabled
- Set Array Accelerator to enabled.

   Note: If you are using the CLI, use the cache=enabled setting.

Presenting logical units to hosts

Set the Selective Storage Presentation (SSP), also known as ACL to enabled.

Use the following host profile settings:

```
Mode 0 = Peripheral Device LUN Addressing
Mode 1 = Asymmetric Failover
Mode 2 = Logical volumes connect as available on Backup Controller
Mode 3 = Product ID of 'MSA1000 Volume'
Mode 4 = Normal bad block handling
Mode 5 = Logout all initiators on TPRLO
Mode 6 = Fault management events not reported through Unit Attention
Mode 7 = Send FCP response info with SCSI status
Mode 8 = Do not send Unit Attention on failover
Mode 9 = SCSI inquiry revision field contains the actual version
Mode 10 = SCSI inquiry vendor field contains Compaq
Mode 11 = Power On Reset Unit Attention generated on FC Login or Logout
Mode 12 = Enforce Force Unit Access on Write
```

You can use the built-in Linux profile or Default profile to set the host profile settings. If you use the Default profile, you must issue the following Serial port CLI command to change the host profile settings:

```
change mode Default mode number
```

where *mode number* is the numeric value for the mode that you want to change.

See the documentation that is provided for your HP StorageWorks MSA for additional information.

Important: You must use the Serial port CLI or the SSP to recheck the connection objects after the configuration is complete.

Migrating logical units

You can use the standard migration procedure to migrate logical units from the HP StorageWorks MSA to the SAN Volume Controller cluster with the following restrictions:

- You cannot share the HP StorageWorks MSA between a host and the SAN Volume Controller cluster. You must migrate all hosts at the same time.
- The subsystem device driver (SDD) and securepath cannot coexist because they have different QLogic driver requirements.
- The QLogic driver that is supplied by HP must be removed and the driver that is supported by IBM must be installed.

Sharing the HP MSA1000 and MSA1500 between a host and the SAN Volume Controller

You must configure your environment so that only the SAN Volume Controller can access all logical units on the HP MSA1000 and MSA1500. You can zone other hosts to communicate with the HP MSA1000 and MSA1500 for in-band configuration, but nothing else.

Concurrent maintenance on the HP MSA1000 and MSA1500

Concurrent maintenance is the capability to perform I/O operations to an HP MSA1000 and MSA1500 while simultaneously performing maintenance operations on it.

You can perform nondisruptive maintenance procedures concurrently on the following components:

- HP MSA1000 or MSA1500 controller
- HP MSA1000 or MSA1500 controller cache

- Cache battery pack
- Variable speed blower
- Power supply
- Disk drive
e • SFP transceivers

Quorum disks on the HP MSA

The SAN Volume Controller cannot use logical units (LUs) that are exported by the HP MSA1000 and MSA1500 as quorum disks.

Advanced functions for the HP MSA

The SAN Volume Controller Copy Service functions and RAID migration utilities are not supported for logical units (LUs) that are presented by the HP MSA.

Global settings for HP MSA systems

Global settings apply across an HP MSA system.

The following table lists the global settings for an HP MSA system:

Option	Required setting
Expand Priority	All supported **Note:** Performance impact of high priority
Rebuild Priority	All supported **Note:** Performance impact of high priority
Array Accelerator	On **Note:** Set on all logical drives that are used by the SAN Volume Controller.
Read-Write cache ratio	All supported
Name of controller	Not important

Configuring HP StorageWorks MSA2000 storage systems

This section provides information about configuring Hewlett Packard (HP) 2000 family Modular Smart Array (MSA2000) systems for attachment to a SAN Volume Controller.

HP MSA2000 supported models

SAN Volume Controller clusters can be used with MSA2000 storage systems.

See the following Web site for the latest supported models:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

For SAN Volume Controller version 4.3.1.7, this is only the MSA2000fc dual controller model that is configured with each controller module attached to both fabrics. For details, refer to the *HP StorageWorks Modular Model User Guide* section on connecting two data hosts through two switches where all four ports must be used and cross-connected to both SAN fabrics.

Supported HP MSA2000 firmware levels

You must ensure that the MSA2000 firmware level can be used with the SAN Volume Controller cluster.

For the supported firmware levels and hardware, see the SAN Volume Controller Web site:Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

HP MSA2000 user interfaces

You can configure an MSA2000 system through the Storage Management Utility (SMU), which is a Web server on each controller, or with the command-line interface (CLI).

To access the MSA2000 system initially you can go through either a serial interface or Dynamic Host Configuration Protocol (DHCP). You can also configure user access and privileges.

MSA2000 Web graphical user interface (GUI)

The SMU is a Web-based GUI that runs on each controller that is accessible through the IP address of each controller. All management and monitoring tasks can be completed on each controller.

MSA2000 command-line interface (CLI)

The CLI is accessible through Secure Shell (SSH), Telnet, and serial port. The CLI includes all functionality that is available in the GUI.

Concurrent maintenance on MSA2000 systems

Concurrent maintenance is the capability to perform I/O operations while you simultaneously perform maintenance operations on the MSA2000 system.

Apply firmware upgrades to an MSA2000 system during a maintenance window because the MSA2000 system takes both controllers offline simultaneously multiple times during an upgrade.

Logical units and target ports on MSA2000 systems

Partitions (volumes) on MSA2000 systems are exported as logical units with a logical unit number that is assigned to that partition.

LUNs on MSA2000 systems

The controller calls a RAID array a virtual disk (VDisk). SAS and SATA disks cannot be mixed within a VDisk, and the maximum number of VDisks per controller is 16. VDisks can be divided into volumes, which are then presented to the host. There can be up to 128 volumes per controller. The capacity of a volume is between 1 MB and 16 TB.

2 SAN Volume Controller has an individual 1 PB managed-disk size limit.

LUN IDs

LUNs exported by MSA2000 systems report identification descriptors 0, 3, 4, 5 in the VPD page 0x83. The LUN IDs are based on the controller MAC addresses. For example:

```
example;
# show volumes
Vdisk      Volume Name    Size  WR Policy    Class  Volume Serial Number  Cache Opt  Type

VD0                       VD0_V1         750.1GB  writeback    standard 00c0ffd76a330000a0fa124a01000000  standard  standard
VD2                       VD2_V1         750.1GB  writeback    standard 00c0ffd76a33000048fb124a01000000  standard  standard
VD_HC                     VD_CAP_V1      37.5GB   writeback    standard 00c0ffd76a3300005efc124a01000000  standard  standard
VD_1                      VD_1_V1        750.1GB  writeback    standard 00c0ffd7648f000064851d4a01000000  standard  standard
VD_3                      VD_3_V1        750.1GB  writeback    standard 00c0ffd7648f0000a6851d4a01000000  standard  standard
VD-R                      VD-R_V1        250.0GB  writeback    standard 00c0ffd7648f0000aa08234a01000000  standard  standard
VD-R                      VD-R_V2        250.0GB  writeback    standard 00c0ffd7648f0000ab08234a01000000  standard  standard
VD-R                      VD-R_V3        250.0GB  writeback    standard 00c0ffd7648f0000ab08234a02000000  standard  standard
--------------------------------------------------------------------------------------------------------------------
# show network-parameters
Network Parameters Controller A
```

```
-------------------------------
IP Address     : 9.71.47.27
Gateway        : 9.71.46.1
Subnet Mask    : 255.255.254.0
MAC Address    : 00:C0:FF:D7:6A:33
Addressing Mode: DHCP

Network Parameters Controller B
-------------------------------
IP Address     : 9.71.47.30
Gateway        : 9.71.46.1
Subnet Mask    : 255.255.254.0
MAC Address    : 00:C0:FF:D7:64:8F
Addressing Mode: DHCP
```

LUN creation and deletion

MSA2000 LUNs can be created, modified, or deleted either by the Storage Management Utility (SMU) or the command-line interface (CLI). LUNs can be used immediately with format to zeros as default background task.

Note: Disks appear as critical while this process is taking place.

To create a logical unit (volume from a VDisk), complete these steps:

1. In the Storage Management Utility SMU interface, go to **Manage** › **Virtual Disk Config** › **Create a VDisk**. The SMU provides a wizard to create the virtual disks..
2. You have the following options:
   - Manual
   - Virtual Disk Name
   - RAID Type

     Note: SAN Volume Controller does not support RAID 0.
   - Number of volumes
   - Expose to all hosts

     Note: The Expose to all hosts option can cause confusion in multi-cluster environments.
   - LUN assignments

You can also modify, expand, or delete a volume or VDisk using either the SMU or the CLI.

Note: Before you delete the LUN on the MSA2000 system, use the **svctask rmdisk** command to delete the MDisk on the SAN Volume Controller cluster.

LUN presentation

You can also use the SMU or the CLI to map and unmap MSA2000 LUNs.

To map a logical unit (volume from a VDisk), from the SMU complete these steps:

1. In the Storage Management Utility SMU interface, go to **Manage** › **Volume Management** › **VDisk or Volume** › **Volume Mapping**.
2. Under the section Assign Host Access Privileges, select **Map Host to Volume**.
3. For each SAN Volume Controller WWPN, select **SVC WWPN** in the HOST WWN-Name menu.
4. Enter the LUN number to present to the SAN Volume Controller. For example, use 0 for the first volume, then use 1 for the second, until all volumes are assigned.
5. Select read-write for the Port 0 Access and Port 1 Access.
6. Click **Map it**. The resulting mapping is displayed in the Current Host-Volume Relationships section.

Important: Use this section to verify that the LUN ID is consistent and all SAN Volume Controller WWPNs have been mapped.

Because there are 8 nodes in the following example, 32 WWPNs show in the show volume-maps output (four ports per node).

```
example shown for an 8-node cluster, that is, 32 WWPNs;
# show volume-maps
Volume [SN 00c0ffd76a330000a0fa124a01000000, Name (VD0_V1)] mapping view:
CH       ID LUN Access Host-Port-Identifier                      Nickname
------------------------------------------------------------------------------
0,1       0   0 rw     50050768012FFFFF
0,1       0   0 rw     5005076801105CEE
0,1       0   0 rw     500507680110008A
0,1       0   0 rw     50050768011FFFFF
0,1       0   0 rw     50050768013FFFFF
0,1       0   0 rw     50050768014FFFFF
0,1       0   0 rw     500507680140008A
0,1       0   0 rw     500507680130008A
0,1       0   0 rw     500507680120008A
0,1       0   0 rw     5005076801405CEE
0,1       0   0 rw     5005076801205CEE
0,1       0   0 rw     5005076801305CEE
0,1       0   0 rw     500507680110596B
0,1       0   0 rw     5005076801305FB8
0,1       0   0 rw     5005076801205FB8
0,1       0   0 rw     5005076801405FB8
0,1       0   0 rw     5005076801105FB8
0,1       0   0 rw     500507680120596B
0,1       0   0 rw     500507680140596B
0,1       0   0 rw     500507680130596B
0,1       0   0 rw     5005076801400009
0,1       0   0 rw     5005076801300009
0,1       0   0 rw     5005076801100009
0,1       0   0 rw     5005076801200009
0,1       0   0 rw     50050768014FFFFE
0,1       0   0 rw     50050768013FFFFE
0,1       0   0 rw     50050768012FFFFE
0,1       0   0 rw     50050768011FFFFE
0,1       0   0 rw     5005076801200001
0,1       0   0 rw     5005076801400001
0,1       0   0 rw     5005076801300001
0,1       0   0 rw     5005076801100001
```

Note: LUNs from controller module A and controller module B can have the same LUN IDs (0). Controller module A and Controller module B appear on the SAN Volume Controller cluster as separate
t controllers. Managed disks (MDisks) on the cluster should be in separate storage pools so that each
t controller module has its own separate storage pool for its presented MDisks.

Special LUNs

Volumes can have a LUN ID from 0 to 126 on each controller. LUN 0 on the MSA2000 is visible from both controllers, but can only be used to access storage from the preferred controller. LUN 0 on the other controller does not present storage.

Target ports on MSA2000 systems

The MSA2000 system has two dual-active controllers with two ports each. You must set these as point-to-point using the SMU interface.

In the Storage Management Utility SMU interface, go to **Manage** → **General Config** → **Host Port Configuration**. Select Advanced Options and specify point to point for Change Host Topology.

Each WWPN is identified with the pattern 2P:7N:CC:CC:CC:MM:MM:MM where *P* is the port number on the controller and *N* is the address of the controller port (0 or 8), *CC:CC:CC* represents the Organizationally Unique Identifier (OUI), and *MM:MM:MM* is unique for the particular controller.

```
example;
# show port-wwn
CTRL CH WWPN
------------------------
A    0   207000C0FFD75198
A    1   217000C0FFD75198
B    0   207800C0FFD75198
B    1   217800C0FFD75198
```

LU access model

The MSA2000 is a dual-active system. Each LUN has an owning controller, and I/O is serviced only by ports on that controller. Failover automatically takes place when one controller fails (shuts down). There is no way for SAN Volume Controller to force failover.

LU grouping

The MSA2000 system does not support LU grouping.

LU preferred access port

The MSA system has two ports per controller. The I/O is through port 0, and port 1 is linked to port 0 of the other controller during a failure or a code upgrade.

Detecting ownership

The LUN is reported only by the target ports of the owning controller.

Failover

The only way to cause failover of LUs from one controller to the other is to shut down one of the controllers. The MSA2000 system cannot normally present all the system LUNs through both controllers. Therefore, it requires a four-port connection to two SAN fabrics. Failover for MS2000 systems involves the surviving controller taking its ports offline, then returning with one of its ports, emulating the WWPNs of the failed controller.

Note: This behavior also means that half of the operational paths from the surviving controller are taken away when failover takes place, which allows the port from the controller that is shutting down to be emulated.

Switch zoning for MSA2000 storage systems

Switch zoning configurations for the MSA2000 system include considerations for fabric zoning, target port sharing, host splitting, and controller splitting.

Fabric zoning

Each SAN Volume Controller switch zone must include at least one target port from each controller to have no single point of failure. This means, for example, that the zone on the first fabric has Port 0 MSA Controller A with Port 1 of MSA Controller B and the SAN Volume Controller ports. The zone on the second fabric has Port 0 MSA Controller B and Port 1 MSA Controller A and the SAN Volume Controller ports. For more information about the fibre-channel dual-fabric setup, refer to the relevant MSA documentation.

Target port sharing

Target ports may not be shared between SAN Volume Controller and other hosts.

Host splitting

A single host must not be connected to SAN Volume Controller and an MSA2000 system simultaneously.

Controller splitting

MSA2000 system LUNs must be mapped only to the SAN Volume Controller cluster. The four target ports are all required for dual SAN-fabric connections and cannot be shared.

Configuration settings for MSA2000 systems

The MSA2000 System Storage Management Utility (SMU) provides configuration settings and options that can be used with SAN Volume Controller clusters.

Target port options

Table 65 describes the port settings that are supported by the SAN Volume Controller.

Table 65. MSA2000 system port settings for use with the SAN Volume Controller

Option	Values (any limits on the possible values)	Description
Host Port Configuration	2 Gbps or 4 Gbps	Set according to the fabric speed.
Internal Host Port Interconnect	Straight-through	Set to Straight-through for a point-to-point fibre-channel connection.
Host Port Configuration	Point-to-Point	Set to Point-to-point for use with SAN Volume Controller.

LU options and settings

The MSA volumes can be created after you create a volume (RAID 0 is not supported), or they can be added later to the volume. LUNs can be configured in 16K, 32K, and 64K (default) chunks by using the advanced option. Table 66 describes the preferred options available when you create a logical unit (LU).

Table 66. Preferred options for logical units (LU)

Option	Value	Description
Expose to All Hosts	Yes	After the mapping of the volume to the SAN Volume Controller completes, this is modified to All other hosts (none no access). This can be done under the Assign Host Access Privileges frame.
Automatically assign LUNs	Yes	This forces option expose to all hosts and is necessary for consistent LUN numbering.
write-policy	write-back	
optimization	any	
read-ahead-size	default	
independent	disable	This setting controls cache mirroring. Because SAN Volume Controller requires mirroring, the independent=disable option must be used.

Host options and settings for MSA2000 systems

There is no specific host option to present the MSA2000 systems to SAN Volume Controller clusters. Use Microsoft Windows 2003 (Microsoft Windows 2003) as the host setting for SAN Volume Controller.

Quorum disks on MSA2000 systems

The SAN Volume Controller cluster requires managed disks (MDisks) that are quorum disks for system metadata storage. The MSA2000 system failover method is not compatible with the requirements for these disks. Quorum disks must be on another separate and suitable managed controller.

Copy functions for MSA2000 systems

The MSA2000 system provides optional copy and replicate features, called *clone* and *snapshot*. However, these functions must not be used with SAN Volume Controller.

Configuring NEC iStorage systems

This section provides information about configuring NEC iStorage systems for attachment to a SAN Volume Controller.

Supported firmware levels for the NEC iStorage

The NEC iStorage system must use a firmware level that is supported by the SAN Volume Controller.

See the following Web site for specific firmware levels and the latest supported hardware:Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Logical unit creation and deletion for NEC iStorage systems

You can create or delete logical units for NEC iStorage systems. See the storage configuration guidelines that are specified in the NEC iStorage documentation that is provided for this system.

Platform type for NEC iStorage

You must set all logical units that the SAN Volume Controller accesses to platform type AX (AIX).

Access control methods for NEC iStorage

You can use access control to restrict access from hosts and SAN Volume Controller clusters. You do not need to use access control to allow a SAN Volume Controller cluster to use all of the defined logical units on the system.

The following table lists the access control methods that are available:

Method	Description
Port Mode	Allows access to logical units that you want to define on a per storage controller port basis. SAN Volume Controller visibility (through switch zoning, physical cabling, etc.) must allow the SAN Volume Controller cluster to have the same access from all nodes and the accessible controller ports have been assigned the same set of logical units with the same logical unit number. This method of access control is not recommended for SAN Volume Controller connection.
WWN Mode	Allows access to logical units using the WWPN of each of the ports of an accessing host device. All WWPNs of all the SAN Volume Controller nodes in the same cluster must be added to the list of linked paths in the controller configuration. This becomes the list of host (SAN Volume Controller) ports for an LD Set or group of logical units. This method of access control allows sharing because different logical units can be accessed by other hosts.

Setting cache allocations for NEC iStorage

Cache allocations can be set manually; however, changes to the default settings can adversely effect performance and cause you to lose access to the system.

Snapshot Volume and Link Volume for NEC iStorage

You cannot use Copy Services logical volumes with logical units that are assigned to the SAN Volume Controller.

Configuring NetApp FAS systems

This section provides information about configuring the Network Appliance (NetApp) Fibre-attached Storage (FAS) systems for attachment to a SAN Volume Controller. Models of the NetApp FAS system are equivalent to the IBM System Storage N5000 series and the IBM System Storage N7000 series; therefore, the SAN Volume Controller also supports models of the IBM N5000 series and the IBM N7000 series.

Attention: You must configure NetApp FAS systems in single-image mode. SAN Volume Controller does not support NetApp FAS systems that are in multiple-image mode.

The information in this section also applies to the supported models of the IBM N5000 series and the IBM N7000 series.

Supported models of the NetApp FAS system

The SAN Volume Controller supports models of the NetApp FAS200, FAS900, FAS3000 and FAS6000 series of systems.

See the following Web site for the latest supported models:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Supported firmware levels for the NetApp FAS

The NetApp FAS must use a firmware level that is supported by the SAN Volume Controller.

See the following Web site for specific firmware levels and the latest supported hardware:Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

User interfaces on the NetApp FAS

Ensure that you are familiar with the user interface applications that support the NetApp FAS.

See the documentation that is provided with your NetApp FAS system for more information about the Web server and CLI.

Web server

You can manage, configure, and monitor the NetApp FAS through the FileView GUI.

CLI

You can access the command-line interface through a direct connection to the filer serial console port or by using the filer IP address to establish a telnet session.

Logical units and target ports on NetApp FAS systems

For the NetApp FAS systems, a logical unit (LU) is a subdirectory in an internal file system.

LUs that are exported by the NetApp FAS system report identification descriptors in the vital product data (VPD). The SAN Volume Controller cluster uses the LUN-associated binary type-3 IEEE Registered Extended descriptor to identify the LU. For a NetApp LUN that is mapped to the SAN Volume Controller cluster, set the LUN Protocol Type to Linux.

The NetApp FAS system does not use LU groups so that all LUs are independent. The LU access model is active-active. Each LU has a preferred filer, but can be accessed from either filer. The preferred filer contains the preferred access ports for the LU. The SAN Volume Controller cluster detects and uses this preference.

The NetApp FAS reports a different worldwide port name (WWPN) for each port and a single worldwide node name (WWNN).

Creating logical units on the NetApp FAS

To create logical units, you must identify a volume from which to create the logical unit and specify the amount of space to use.

Perform the following steps to create logical units:
1. Log on to the NetApp FAS.
2. Go to **Filer View** and authenticate.
3. Click **Volumes** and identify a volume to use to create an LU. A list of volumes is displayed.
4. Identify a volume that has sufficient free space for the LUN size that you want to use.
5. Click **LUNs** on the left panel.
6. Click **Add** in the list.
7. Enter the following:
   a. In the **Path** field, enter /vol/*volx*/*lun_name* where *volx* is the name of the volume identified above and *lun_name* is a generic name.
   b. In the **LUN Type** field, enter Image.
   c. Leave the **Description** field blank.
   d. In the **Size** field, enter a LUN Size.
   e. In the **Units** field, enter the LUN Size in units.
   f. Select the **Space Reserved** box.

      Note: If the Space Reserved box is not selected and the file system is full, the LUN goes offline. The storage pool also goes offline and you cannot access the volumes.
   g. Click **Add**.

      Note: To check the LUN settings, go to the Manage LUNs section and click the LUN you want to view. Ensure that the Space Reserved setting is set.

Deleting logical units on the NetApp FAS

You can delete logical units.

Perform the following steps to delete logical units:
1. Log on to the NetApp FAS.
2. Go to **Filer View** and authenticate.
3. Click **LUNs** on the left panel.

4. Click **Manage**. A list of LUNs is displayed.

5. Click the LUN that you want to delete.

6. Click **Delete**.

7. Confirm the LUN that you want to delete.

Creating host objects for the NetApp FAS

You can create host objects.

Perform the following steps to create host objects:

1. Log on to the NetApp FAS.

2. Go to **Filer View** and authenticate.

3. Click **LUNs** on the left panel.

4. Click **Initiator Groups**.

5. Click **Add** in the list.

6. Enter the following:

   a. In the **Group Name** field, enter the name of the initiator group or host.

   b. In the **Type** list, select FCP.

   c. In the **Operating System** field, select Linux.

   d. In the **Initiators** field, enter the list of WWPNs of all the ports of the nodes in the cluster that are associated with the host.

      Note: Delete the WWPNs that are displayed in the list and manually enter the list of SAN Volume Controller node ports. You must enter the ports of all nodes in the SAN Volume Controller cluster.

7. Click **Add**.

Presenting LUNs to hosts for NetApp FAS

You can present LUNs to hosts.

Perform the following steps to present LUNs to hosts:

1. Log on to the NetApp FAS.

2. Go to **Filer View** and authenticate.

3. Click **LUNs** on the left panel.

4. Click **Manage**. A list of LUNs is displayed.

5. Click the LUN that you want to map.

6. Click **Map LUN**.

7. Click **Add Groups to Map**.

8. Select the name of the host or initiator group from the list and click **Add**.

   Notes:

   a. You can leave the LUN ID section blank. A LUN ID is assigned based on the information the controllers are currently presenting.

   b. If you are re-mapping the LUN from one host to another, you can also select the **Unmap** box.

9. Click **Apply**.

Switch zoning limitations for NetApp FAS systems

There are limitations in switch zoning for SAN Volume Controller clusters and NetApp FAS systems.

Fabric zoning

The SAN Volume Controller switch zone must include at least one target port from each filer to avoid a single point of failure.

Target port sharing

Target ports can be shared between the SAN Volume Controller cluster and other hosts. However, you must define separate initiator groups (igroups) for the SAN Volume Controller initiator ports and the host ports.

Host splitting

A single host cannot be connected to both the SAN Volume Controller cluster and the NetApp FAS to avoid the possibility of an interaction between multipathing drivers.

Controller splitting

You can connect other hosts directly to both the NetApp FAS and the SAN Volume Controller cluster under the following conditions:

- Target ports are dedicated to each host or are in a different igroup than the SAN Volume Controller cluster
- LUNs that are in the SAN Volume Controller cluster igroup are not included in any other igroup

Concurrent maintenance on the NetApp FAS

Concurrent maintenance is the capability to perform I/O operations to a NetApp FAS while simultaneously performing maintenance operations on it.

The SAN Volume Controller supports concurrent maintenance on the NetApp FAS.

Quorum disks on the NetApp FAS

The SAN Volume Controller can use logical units (LUs) that are exported by the NetApp FAS as quorum disks.

Advanced functions for the NetApp FAS

The SAN Volume Controller copy and migration functions are supported for logical units (LUs) that are presented by the NetApp FAS.

Configuring Nexsan SATABeast systems

This section provides information about configuring Nexsan SATABeast systems for attachment to a SAN Volume Controller.

Supported models of the Nexsan SATABeast system

The SAN Volume Controller supports models of the Nexsan SATABeast series of systems.

See the following website for the latest supported models:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Supported firmware levels for the Nexsan SATABeast

The Nexsan SATABeast system must use a firmware level that is supported by the SAN Volume Controller. The current level is Nt66E.

2 See the following website for specific firmware levels and the latest supported hardware:

2 Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

2 The minimum supported SAN Volume Controller level for the attachment of Nexsan SATABeast is 5.1.0.3.

2 Concurrent maintenance on Nexsan SATABeast systems

2 Concurrent maintenance is the ability to perform I/O operations on a Nexsan SATABeast while
2 simultaneously performing maintenance on it.

2 You can perform nondisruptive maintenance procedures concurrently on the following components:
2 • Nexsan SATABeast controller
2 • Disk drive

2 User interfaces on the Nexsan SATABeast

2 NexScan is the Nexsan's web-enabled GUI. NexScan provides access to your SATABeast system from any
2 standard internet browser or host computer, either directly connected or connected through a LAN or
2 WAN.

2 NexScan is platform-independent and no software or patches are required. Additional access is available
2 through the RS232 serial interface DB9 (one per controller). NexScan supports VT100 and is compatible
2 with terminal emulation software such as HyperTerminal and Kermit.

2 Logical unit creation, deletion, and migration for Nexsan SATABeast
2 systems

2 Before you create, delete, and migrate logical units for the Nexsan SATABeast, read the storage
2 configuration guidelines specified in the Nexsan SATABeast documentation that is provided for this
2 system.

2 Creating arrays

2 The following arrays are supported:
2 • RAID 0
2 • RAID 1
2 • RAID 4
2 • RAID 5
2 • RAID 6

2 Creating volumes

2 You create and configure volumes in the Configure Volumes section of the GUI.

2 Presenting logical units to hosts

3 Table 67 lists the host profile settings:

2 *Table 67. Nexsan SATABeast host profile settings*

2 Controller 0	Fibre Host 0		Fibre Host 1	
2	Current Status	New State	Current Status	New State
2 Topology	P2P Full Fabric	Auto	P2P Full Fabric	Auto
2 Loop ID	(NA)	Auto	(NA)	Auto

Table 67. Nexsan SATABeast host profile settings (continued)

Controller 0	Fibre Host 0		Fibre Host 1	
Link Speed	4Gbit	Auto	4Gbit	Auto
Auto Port Logout	Yes	Yes	Yes	Yes
Controller 0	Fibre Host 0		Fibre Host 1	
	Current Status	New State	Current Status	New State
Topology	P2P Full Fabric	Auto	P2P Full Fabric	Auto
Loop ID	(NA)	Auto	(NA)	Auto
Link Speed	4Gbit	Auto	4Gbit	Auto
Auto Port Logout	Yes	Yes	Yes	Yes

Migrating logical units

You can use the standard migration procedure to migrate logical units from the Nexsan SATABeast to the SAN Volume Controller cluster.

Sharing the Nexsan SATABeast between a host and the SAN Volume Controller

You must configure your environment so that only the SAN Volume Controller can access all logical units on the Nexsan SATABeast. You can zone other hosts to communicate with the Nexsan SATABeast for in-band configuration, but nothing else.

Quorum disks on the Nexsan SATABeast

The SAN Volume Controller can use logical units (LUs) that are exported by the Nexsan SATABeast as quorum disks.

Advanced functions for the Nexsan SATABeast

Nexsan advanced functions are not supported with SAN Volume Controller.

Configuring Pillar Axiom systems

This section provides information about configuring Pillar Axiom systems for attachment to a SAN Volume Controller cluster.

Supported models of Pillar Axiom systems

SAN Volume Controller clusters can be used with some models of the Pillar Axiom series of systems.

See the following Web site for the latest models that can be used with SAN Volume Controller clusters:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Supported firmware levels of Pillar Axiom systems

You must ensure that the firmware level of the Pillar Axiom system can be used with the SAN Volume Controller cluster.

See the following Web site for specific firmware levels and the latest supported hardware:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Concurrent maintenance on Pillar Axiom systems

Concurrent maintenance is the capability to perform I/O operations to a Pillar Axiom system while simultaneously performing maintenance operations on it.

Because some maintenance operations restart the Pillar Axiom system, you cannot perform hardware maintenance or firmware upgrades while the system is attached to a SAN Volume Controller cluster.

Pillar Axiom user interfaces

Ensure that you are familiar with the user interface applications that Pillar Axiom systems use. For more information, see the documentation that is included with the Pillar Axiom system.

AxiomONE Storage Services Manager

The AxiomONE Storage Services Manager is a browser-based GUI that allows you to configure, manage, and troubleshoot Pillar Axiom systems.

Pillar Data Systems CLI

The Pillar Data Systems command-line interface (CLI) communicates with the system through an XML-based application programming interface (API) over a TCP/IP network. The Pillar Data Systems CLI is installed through the AxiomOne Storage Service Manager. You can use the Pillar Data Systems CLI to issue all commands, run scripts, request input files to run commands, and run commands through a command prompt. The Pillar Data Systems CLI can run on all operating systems that can be used with Pillar Axiom systems.

AxiomONE CLI

The AxiomONE CLI is installed through the AxiomONE Storage Service Manager. You can use the AxiomONE CLI to perform administrative tasks. The AxiomONE CLI can run on a subset of operating systems that can be used with Pillar Axiom systems.

Logical units and target ports on Pillar Axiom systems

For Pillar Axiom systems, logical units are enumerated devices that have the same characteristics as LUNs.

LUNs

You can use the AxiomONE Storage Services Manager to create and delete LUNs.

Important:
1. When you create a LUN, it is not formatted and therefore can still contain sensitive data from previous usage.
2. You cannot map more than 256 Pillar Axiom LUNs to a SAN Volume Controller cluster.

You can create LUNs in a specific volume group or in a generic volume group. A single LUN can use the entire capacity of a disk group. However, for SAN Volume Controller clusters, LUNs cannot exceed 1 PB. When LUNs are exactly 1 PB, a warning is issued in the SAN Volume Controller cluster event log.

The amount of capacity that the LUN uses is determined by the capacity of the LUN and the level of redundancy. You can define one of three levels of redundancy:
- Standard, which stores only the original data
- Double, which stores the original data and one copy
- Triple, which stores the original data and two copies

For all levels of redundancy, data is striped across multiple RAID-5 groups.

LUNs that are exported by the Pillar Axiom system report identification descriptors in the vital product data (VPD). The SAN Volume Controller cluster uses the LUN-associated binary type-2 IEEE Registered Extended descriptor to identify the LUN. The following format is used:

CCCCCCLLLLMMMMMM

where CCCCCC is the IEEE company ID (0x00b08), LLLL is a number that increments each time a LUN is created (0000–0xFFFD) and MMMMMM is the system serial number.

You can find the identifier in the AxiomONE Storage Services Manager. From the AxiomONE Storage Services Manager, click **Storage** → **LUNs** → **Identity**. The identifier is listed in the LUID column. To verify that the identifier matches the UID that the SAN Volume Controller cluster lists, issue the **svcinfo lsmdisk** *mdisk_id or mdisk_name* from the command-line interface and check the value in the UID column.

Moving LUNs

If you want to migrate more than 256 LUNs on an existing Pillar Axiom system to the SAN Volume Controller cluster, you must use the SAN Volume Controller cluster migration function. The Pillar Axiom system allows up to 256 LUNs per host and the SAN Volume Controller cluster must be configured as a single host. Because the SAN Volume Controller cluster is not limited to 256 volumes, you can migrate your existing Pillar Axiom system set up to the SAN Volume Controller cluster. You must then virtualize groups of LUNs and then migrate the group to larger managed mode disks.

Target ports

Pillar Axiom systems with one pair of controllers report a different worldwide port name (WWPN) for each port and a single worldwide node name (WWNN). Systems with more than one pair of controllers report a unique WWNN for each controller pair.

LUN groups are not used so that all LUNs are independent. The LUN access model is active-active/asymmetric with one controller having ownership of the LUN. All I/O operations to the LUN on this controller is optimized for performance. You can use the **svcinfo lsmdisk** *mdisk_id or mdisk_name* CLI command to determine the assigned controller for a LUN.

To balance I/O load across the controllers, I/O operations can be performed through any port. However, performance is higher on the ports of the controller that own the LUNs. By default, the LUNs that are mapped to the SAN Volume Controller cluster are accessed through the ports of the controller that owns the LUNs.

Switch zoning limitations for Pillar Axiom systems

There are limitations in switch zoning for SAN Volume Controller clusters and Pillar Axiom systems.

Fabric zoning

The SAN Volume Controller switch zone must include at least one target port from each Pillar Axiom controller to avoid a single point of failure.

Target port sharing

Target ports can be shared between the SAN Volume Controller cluster and other hosts.

Host splitting

A single host cannot be connected to both the SAN Volume Controller cluster and the Pillar Axiom system to avoid the possibility of an interaction between multipathing drivers.

Controller splitting

Pillar Axiom system LUNs that are mapped to the SAN Volume Controller cluster cannot be mapped to other hosts. Pillar Axiom system LUNs that are *not* mapped to the SAN Volume Controller cluster can be mapped to other hosts.

Configuration settings for Pillar Axiom systems

The AxiomONE Storage Services Manager provides configuration settings and options that can be used with SAN Volume Controller clusters.

The settings and options can have a scope of the following:
- System (global)
- Logical unit (LU)
- Host

Global settings for Pillar Axiom systems

Global settings apply across a Pillar Axiom system.

Table 68 lists the system options that you can access using the AxiomONE Storage Services Manager.

Table 68. Pillar Axiom global options and required settings

Option	Pillar Axiom default setting	SAN Volume Controller required setting
Enable Automatic Failback of NAS Control Units	Y	N/A
Link Aggregation	N	N/A
DHCP/Static	-	Any
Call-home	-	Any

Logical unit options and settings for Pillar Axiom systems

Logical unit (LU) settings are configurable at the LU level.

Table 69 lists the options that must be set for each LU that is accessed by other hosts. LUs that are accessed by hosts can be configured differently. You can use the AxiomONE Storage Services Manager to change these settings.

Table 69. Pillar Axiom LU options and required settings

Option	Pillar Axiom Default Setting	SAN Volume Controller Required Setting
LUN Access	All hosts	Select hosts
Protocol	FC	FC
LUN Assignment	Auto	Any **Attention:** Do not change the LUN assignment after the LUNs are mapped to the SAN Volume Controller cluster.

Table 69. Pillar Axiom LU options and required settings (continued)

Option	Pillar Axiom Default Setting	SAN Volume Controller Required Setting
Select Port Mask	All On	All On
Quality of Service	Various	No preference. See the note below.
Note: If you do not know the Quality of Service setting, you can use the following: • Priority vs other Volumes = Medium • Data is typically accessed = Mixed • I/O Bias = Mixed		

Host options and settings for Pillar Axiom systems

You must use specific settings to identify SAN Volume Controller clusters as hosts to Pillar Axiom systems.

Table 70 lists the host options and settings that can be changed using the AxiomONE Storage Services Manager.

Table 70. Pillar Axiom host options and required settings

Option	Pillar Axiom default setting	SAN Volume Controller required setting
Load balancing	Static	Static
HP-UX	N	N

Quorum disks on Pillar Axiom systems

The SAN Volume Controller cluster selects managed disks (MDisks) that are presented by Pillar Axiom systems as quorum disks.

Copy functions for Pillar Axiom systems

Advanced copy functions for Pillar Axiom systems (for example, Snap FS, Snap LUN, Volume Backup, Volume Copy, and Remote Copy) cannot be used with disks that are managed by the SAN Volume Controller cluster.

Configuring Texas Memory Systems RamSan Solid State Storage systems

This section provides information about configuring Texas Memory Systems (TMS) RamSan systems for attachment to a SAN Volume Controller.

TMS RamSan Solid State Storage supported models

SAN Volume Controller clusters can be used with the RamSan Solid State Storage systems.

For the latest RamSan models that can be used with SAN Volume Controller clusters, see the SAN Volume Controller (2145) website:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Supported TMS RamSan firmware levels

You must ensure that the RamSan firmware level can be used with the SAN Volume Controller cluster.

For the supported firmware levels and hardware, see the SAN Volume Controller (2145) website:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Concurrent maintenance on RamSan systems

Concurrent maintenance is the capability to perform I/O operations while you simultaneously perform maintenance operations on the RamSan system.

Apply firmware upgrades to a RamSan system during a maintenance window. A power cycle of the RamSan system is required for the upgraded firmware to take effect.

RamSan user interfaces

You can configure a RamSan system through a Web GUI based on Java or a command-line interface (CLI). You can also perform some system-critical operations by using the front panel on the RamSan system.

RamSan Web GUI

The Web GUI is an applet based on Java that is accessible through the IP address of the RamSan system. All configuration and monitoring steps are available through this interface. Be default, the Web GUI uses SSL encryption to communicate with the RamSan system.

RamSan CLI

The command-line interface (CLI) is accessible through SSH, Telnet, and RS-232 port. The CLI includes all functionality that is available in the GUI with the exception of statistics monitoring. The CLI includes a diagnostics interface, however, for internal hardware checks.

Logical units and target ports on RamSan systems

Partitions on RamSan systems are exported as logical units (LUs) with a logical unit number (LUN) that is assigned to the partition.

LUNs on RamSan systems

RamSan systems are shipped with a particular capacity of user space, which depends on the model.
2 Capacities on one model can range from 1 TB - 1 PB. A partition with this capacity is known as a *logical unit*.

RamSan systems can export up to 1024 LUNs to the SAN Volume Controller through various exported FC ports. The maximum logical-unit size is the full, usable capacity of the RamSan system.

LUN IDs

RamSan systems identify exported LUs through identification descriptors 0, 1, and 2. The EUI-64 identifier for the LU is in the CCCCCCLLLLMMMMMM notation where CCCCCC is the Texas Memory Systems IEEE Company ID of 0020C2h, LLLL is the logical unit handle, and MMMMMM is the serial number of the chassis. The EUI-64 identifier is available on the detailed view of each logical unit in the GUI.

LUN creation and deletion

RamSan LUNs are created, modified, or deleted either by using a wizard tutorial in the GUI or by entering a CLI command. LUNs are not formatted to all zeros upon creation.

To create a logical unit, highlight **Logical Units** and select **Create toolbar**. To modify, resize, or delete an LU, select the appropriate toolbar button when the specific logical unit is highlighted in the navigation tree.

Note: Delete the MDisk on the SAN Volume Controller cluster before you delete the LUN on the RamSan system.

LUN presentation

LUNs are exported through the available FC ports of RamSan systems by access policies. Access policies are associations of the logical unit, port, and host. A RamSan system requires that one of the three items is unique across all available access policies. LUNs that are to be presented to SAN Volume Controller must be presented to all node ports in the cluster through at least two ports on the RamSan system. Present each LU to the SAN Volume Controller on the same LUN through all target ports.

To apply access policies to a logical unit, highlight the specific logical unit in the GUI and click the **Access** toolbar button.

Special LUNs

The RamSan system has no special considerations for logical unit numbering. LUN 0 can be exported where necessary. In one RamSan model, a licensed Turbo feature is available to create a logical unit up to half the size of the cache to keep locked in the DRAM cache for the highest performance. No identification difference exists with a Turbo or locked LUN as opposed to any other LUN ID.

Target ports on RamSan systems

A RamSan system is capable of housing 4 dual-ported FC cards. Each worldwide port name (WWPN) is identified with the pattern `2P:0N:00:20:C2:MM:MM:MM` where `P` is the port number on the controller and `N` is the address of the controller. The `MMMMMM` represents the chassis serial number.

The controller address is as follows:
```
04:  FC77-1
08:  FC77-2
0C:  FC77-3
10:  FC77-4
```

Port 2B has a WWPN of 21:08:00:20:C2:07:83:32 for a system with serial number G-8332. The same system has a worldwide node name (WWNN) of 10:00:00:20:C2:07:83:32 for all ports.

LU access model

On a RamSan system, all controllers are Active/Active on a nonblocking crossbar. To avoid an outage from controller failure, configure multipathing across FC controller cards in all conditions. Because all RamSan systems are equal in priority, there is no benefit to using an exclusive set for a specific LU.

LU grouping

The RamSan system does not use LU grouping.

LU preferred access port

There are no preferred access ports for the RamSan system because all ports are Active/Active across all controllers.

Detecting ownership

Ownership is not relevant to the RamSan system.

Switch zoning for RamSan storage systems

Switch zoning configurations for the RamSan system include considerations for fabric zoning, target port sharing, host splitting, and controller splitting.

Fabric zoning

To enable multipathing, ensure that you have multiple zones or multiple RamSan and SAN Volume Controller ports for each zone when you are zoning a RamSan system to the SAN Volume Controller back-end ports.

Target port sharing

The RamSan system can support LUN masking to enable multiple servers to access separate LUNs through a common controller port. There are no issues with mixing workloads or server types in this setup. LUN Masking is a licensed feature of the RamSan system.

Host splitting

There are no issues with host splitting on a RamSan system.

Controller splitting

RamSan system LUNs that are mapped to the SAN Volume Controller cluster cannot be mapped to other hosts. LUNs that are not presented to the SAN Volume Controller can be mapped to other hosts.

Configuration settings for RamSan systems

The RamSan GUI provides configuration settings and options that can be used with SAN Volume Controller clusters.

LU options and settings

When you create a logical unit (LU), the options in Table 71 are available on RamSan systems.

Table 71. RamSan LU options

Option	Data type	Range	Default	SAN Volume Controller setting	Comments
Name	String	1 character - 32 characters	Logical unit number	Any	This is only for management reference.

Table 71. RamSan LU options (continued)

Option	Data type	Range	Default	SAN Volume Controller setting	Comments
Number	Integer	0 - 1023	Next available LUN	0 - 254	Some hosts have known limitations of 254 as their highest LUN ability. One logical unit can appear at multiple LUNs. For example, the same data could appear at LUN 1, LUN 7, and LUN 124.
Size	Integer	1 MB - maximum capacity	Maximum available capacity	Any	MB and GB are BASE2 offerings.
Backup mode	Option list	Writeback caching or writethrough caching	Writeback caching	Writeback caching	Use writeback caching in production. Use writethrough caching strictly for diagnostics.
Device ID	Integer	Blank, 0 - 32768	Blank	Blank	Specific only to OpenVMS.
Report corrected media errors	Checkbox	Checked or Unchecked	Checked	Checked	Notifies the host if ECC was used to correct the requested data.
Report uncorrected media errors	Checkbox	Checked or Unchecked	Checked	Checked	Always report uncorrected media errors.

Host options and settings for RamSan systems

No host options are required to present the RamSan systems to SAN Volume Controller clusters.

Quorum disks on RamSan systems

The SAN Volume Controller cluster selects managed disks (MDisks) that are presented by RamSan systems as quorum disks. To maintain availability with the cluster, ensure that each quorum disk resides on a separate disk system.

Clearing SCSI reservations and registrations

You must not use the RamSan CLI to clear SCSI reservations and registrations on volumes that are managed by the SAN Volume Controller. This option is not available on the GUI.

Copy functions for RamSan systems

The RamSan system does not provide copy, replicate, or SnapShot features. The RamSan system also does not provide thin provisioning.

Configuring Xiotech Emprise systems

This section provides information about configuring Xiotech Emprise systems for attachment to a SAN Volume Controller cluster.

Supported Xiotech Emprise models

SAN Volume Controller clusters can be used with the Xiotech Emprise storage system.

See the SAN Volume Controller (2145) website for the latest Xiotech Emprise models that can be used with SAN Volume Controller clusters:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Supported Xiotech Emprise firmware levels

You must ensure that the Xiotech Emprise firmware level can be used with the SAN Volume Controller cluster.

See the SAN Volume Controller (2145) website for the supported firmware levels and hardware:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Concurrent maintenance on Xiotech Emprise systems

Concurrent maintenance is the capability to perform I/O operations on a Xiotech Emprise system while simultaneously performing maintenance operations on the system.

Concurrent maintenance cannot be supported during I/O operations. Because some maintenance operations, such as firmware updates, restart Xiotech Emprise systems, consult the appropriate maintenance manual at the Xiotech Web site before you perform maintenance:

www.xiotech.com

Xiotech Emprise user interfaces

Ensure that you are familiar with the Xiotech Emprise user interface applications. For more information about the user interface applications, see the documentation that is included with the Xiotech Emprise system.

Xiotech Emprise Storage Management GUI

The Xiotech Emprise Storage Management GUI is a Java-based interface that you can use to configure, manage, and troubleshoot Xiotech Emprise storage systems. The Xiotech Emprise Storage Management GUI is designed and supported on Microsoft Windows systems and has the following minimum requirements:

   Internet Explorer v6.02800.1106, SP1, Q903235 or higher (JavaScript enabled; XML/XSL rendering enabled)

Xiotech Emprise CLI

The Xiotech Emprise command-line interface (CLI) communicates with the system through a serial port that is connected to a computer that runs a terminal emulation program, such as Microsoft HyperTerminal or PuTTY. The Xiotech Emprise CLI is primarily used to configure the network adapter TCP/IP settings.

A null modem cable is required. Configure the serial port on the computer as follows:

- 115200 baud
- 8 data bits
- No parity
- 1-stop bit
- No flow control

Logical units and target ports on Xiotech Emprise systems

On Xiotech Emprise systems, logical units (LUs) are enumerated devices that have the same characteristics as logical unit numbers (LUNs).

LUNs

An Xiotech Emprise logical unit is referred to as a *volume*.

A single Xiotech Emprise volume can potentially consume the entire capacity that is allocated for SAN Volume Controller storage pools, but it cannot exceed the SAN Volume Controller 1 PB LUN size limit. Any LUN that is 1 PB or larger is truncated to 1 PB, and a warning message is generated for each path to the LUN.

LUN IDs

LUNs that are exported by Xiotech Emprise systems are guaranteed to be unique. They are created with a combination of serial numbers and counters along with a standard IEEE registered extended format.

LUN creation and deletion

Xiotech Emprise LUNs are created and deleted by using either the Xiotech Emprise Storage Management GUI or CLI. LUNs are formatted to all zeros at creation.

When a new LUN is created, the Xiotech Emprise system begins a background zeroing process. If a read operation comes in to an area that has not been processed yet, the system returns zeros as a read response. This is the normal procedure. If a previous LUN with data was in that storage area, it is zeroed out. If a non-zeroed-out area gets read, the system returns zeros if it has not been written to yet.

LUN presentation on Xiotech Emprise systems

Xiotech Emprise LUNs are presented to the SAN Volume Controller interface using the following rules:
- LUNs can be presented to one or more selected hosts.
- Configuration is easier if you create one host name for the SAN Volume Controller.
- No individual LUN volume on the Xiotech Emprise system can exceed 1 PB in size.
- For the managed reliability features to be effective on the Xiotech Emprise system, use either RAID 1 or RAID 5 when you create volumes.
- The write-back and write-through cache options are available depending on the performance requirements on each individual LUN. Generally, write-back caching provides the best performance.
- Although either Linux or Windows can be used, Linux is recommended for volumes that are intended for use on the SAN Volume Controller.

To present Xiotech Emprise LUNs to the SAN Volume Controller, follow these steps:
1. On the Xiotech Emprise system, create a single host name for the SAN Volume Controller and assign all SAN Volume Controller host bus adapter (HBA) ports to that host name as shown in Table 72 on page 248.

Table 72. Host information for Xiotech Emprise

Name	Operating system type	HBA ports	Mapping
SVC_Cluster	Linux	500507680130535F 5005076801305555 500507680140535F 5005076801405555	Volume01 (lun:1) Volume02 (lun:2)

2. When you create new volumes that are intended for use on the SAN Volume Controller, assign them to the host name that is used to represent the SAN Volume Controller.

Special LUNs

The Xiotech Emprise storage system does not use a special LUN. Storage can be presented by using any valid LUN, including 0.

Target ports on Xiotech Emprise systems

Each Xiotech Emprise system has two physical fibre-channel ports. They are, by default, intended to provide failover or multipath capability. The worldwide node name (WWNN) and worldwide port name (WWPN) are typically similar, such as in the following example:

```
WWNN 20:00:00:14:c3:67:3f:c4
WWPN 20:00:00:14:c3:67:3f:c4
WWPN 20:00:00:14:c3:67:3f:c5
```

LU access model

The Xiotech Emprise system has no specific ownership of any LUN by any module. Because data is striped over all disks in a DataPac, performance is generally unaffected by the choice of a target port.

LU grouping

The Xiotech Emprise system does not use LU grouping; all LUNs are independent entities.

LU preferred access port

There are no preferred access ports for the Xiotech Emprise system.

Detecting ownership

Ownership is not relevant to the Xiotech Emprise system.

Switch zoning limitations for Xiotech Emprise storage systems

Limitations exist in switch zoning for SAN Volume Controller clusters and the Xiotech Emprise storage system.

Fabric zoning

To avoid a single point of failure, the SAN Volume Controller switch zone must include both target ports from each Xiotech Emprise controller.

Target port sharing

Target ports can be shared between the SAN Volume Controller cluster and other hosts.

Host splitting

To avoid the possibility of an interaction between multipathing drivers, a single host cannot be connected to both the SAN Volume Controller cluster and the Xiotech Emprise system.

Controller splitting

Xiotech Emprise system logical unit numbers (LUNs) that are mapped to the SAN Volume Controller cluster cannot be mapped to other hosts. Xiotech Emprise system LUNs that are *not* mapped to the SAN Volume Controller cluster can be mapped to other hosts.

Configuration settings for Xiotech Emprise systems

The Xiotech Emprise Storage Management GUI provides configuration settings and options that can be used with SAN Volume Controller clusters.

The only specific setting is the host operating system type: Windows or Linux. For SAN Volume Controller clusters, use Linux.

LU options and settings

Logical unit (LU) settings for the Xiotech Emprise system are configurable at the LU level.

Table 73 lists the options that must be set for each LU that is accessed by other hosts. LUs that are accessed by hosts can be configured differently. You can use the Xiotech Emprise Storage Management GUI or CLI to change these settings.

Table 73. Xiotech Emprise LU settings

Option	Data type	Range	Default	SAN Volume Controller setting	Comments
Capacity	Int	1 GB to 1 PB	No	Any	SAN Volume Controller supports up to 1 PB.

Host options and settings for Xiotech Emprise

You must use specific settings to identify SAN Volume Controller clusters as hosts to the Xiotech Emprise storage system.

A Xiotech Emprise host is a single WWPN; however, multiple WWPNs can be included in a single host definition on the Xiotech Emprise system.

A Xiotech Emprise host also can consist of more than one WWPN. The recommended method is to make each SAN Volume Controller node a Xiotech Emprise host and to make a Xiotech Emprise cluster that corresponds to all the nodes in the SAN Volume Controller cluster. To do this, include all of the SAN Volume Controller WWPNs under the same Xiotech Emprise host name.

Quorum disks on Xiotech Emprise systems

The SAN Volume Controller cluster selects managed disks (MDisks) that are presented by Xiotech Emprise systems as quorum disks. The clearing of Small Computer System Interface (SCSI) reservations and registrations is not supported by the Xiotech Emprise system.

Copy functions for Xiotech Emprise systems

Advanced copy functions for Xiotech Emprise systems such as SnapShot and remote mirroring cannot be used with disks that are managed by the SAN Volume Controller cluster. Thin provisioning is not supported for use with SAN Volume Controller.

Configuring IBM XIV Storage System models

This section provides information about configuring IBM XIV® Storage System models for attachment to a SAN Volume Controller cluster.

Supported IBM XIV Storage System models

SAN Volume Controller support for IBM XIV Storage System systems is specific to certain models.

The supported IBM XIV Storage System models are:
- IBM XIV Storage System Model A14

Note: For Model A14, partially populated racks are supported.

See the following website for the latest IBM XIV Storage System models that can be used with SAN Volume Controller clusters:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Supported IBM XIV firmware levels

You must ensure that SAN Volume Controller supports your IBM XIV Storage System firmware level.

See the following Web site for the supported firmware levels and hardware:

Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145

Concurrent maintenance on IBM XIV Storage System models

Concurrent maintenance is the capability to perform I/O operations on an IBM XIV Storage System models while simultaneously performing maintenance operations on the system.

Some maintenance operations require a complete restart of IBM XIV Storage System systems. Such procedures are not supported when the system is attached to the SAN Volume Controller.

Firmware upgrades of IBM XIV Storage System systems are not supported while the system is attached to the SAN Volume Controller.

All other concurrent maintenance procedures are supported.

IBM XIV user interfaces

Ensure that you are familiar with the IBM XIV Storage System user interface applications. For more information, see the documentation that is included with your IBM XIV Storage System system.

IBM XIV Storage Management GUI

The IBM XIV Storage System Storage Management GUI is a Java-based GUI that you can use to configure, manage, and troubleshoot IBM XIV Storage System systems. The IBM XIV Storage System Storage Management GUI can run on all operating systems that can be used with IBM XIV Storage System systems.

IBM XIV CLI

The IBM XIV Storage System command-line interface (XCLI) communicates with the systems through an XML-based API over a TCP/IP network. You can use the XCLI to issue all commands, run scripts, request input files to run commands, and run commands through a command prompt. The XCLI can run on all operating systems that can be used with IBM XIV Storage System systems.

Logical units and target ports on IBM XIV Storage System models

On IBM XIV Storage System, logical units (LUs) are enumerated devices that have the same characteristics as LUNs.

LUNs

An IBM XIV Storage System Logical Unit is referred to as a *volume*. IBM XIV Storage System and volumes are enumerated devices that all share identical characteristics.

A single IBM XIV Storage System volume can potentially consume the entire capacity that is allocated for SAN Volume Controller managed disk (MDisk) groups, and it can also exceed the SAN Volume
2 Controller 1 PB LUN size limit. Any LUN that is 1 PB or larger is truncated to 1 PB, and a warning message is generated for each path to the LUN.

IBM XIV Storage System volumes consume chunks of 17,179,869,184 bytes (17 GB), although you can create volumes with an arbitrary block count.

LUN IDs

LUNs that are exported by IBM XIV Storage System models report Identification Descriptors 0, 1, and 2 in VPD page 0x83. SAN Volume Controller uses the EUI-64 compliant type 2 descriptor *CCCCCCMMMMMMLLLL*, where *CCCCCC* is the IEEE company ID, *MMMMMM* is the System Serial Number transcribed to hexadecimal (**10142->0x010142**, for example) and *LLLL* is **0000-0xFFFF**, which increments each time a LUN is created. You can identify the *LLLL* value by using the IBM XIV Storage System GUI or CLI to display the volume serial number.

LUN creation and deletion

IBM XIV Storage System LUNs are created and deleted using the IBM XIV Storage System GUI or CLI. LUNs are formatted to all zeros upon creation, but to avoid a significant formatting delay, zeros are not written.

Special LUNs

IBM XIV Storage System systems do not use a special LUN; storage can be presented using any valid LUN, including **0**.

LU access model

IBM XIV Storage System systems have no specific ownership of any LUN by any module. Because data is striped over all disks in the system, performance is generally unaffected by the choice of a target port.

LU grouping

IBM XIV Storage System models do not use LU grouping; all LUNs are independent entities. To protect a single IBM XIV Storage System volume from accidental deletion, you can create a consistency group containing all LUNs that are mapped to a single SAN Volume Controller cluster.

LU preferred access port

There are no preferred access ports for IBM XIV Storage System models.

Detecting ownership

Ownership is not relevant to IBM XIV Storage System models.

LUN presentation on XIV Nextra™ systems

XIV Nextra LUNs are presented to the SAN Volume Controller interface using the following rules:
- LUNs can be presented to one or more selected hosts.
- XIV Nextra maps consist of sets of LUN pairs and linked hosts.
- A volume can only appear once in a map.
- A LUN can only appear once in a map.
- A host can only be linked to one map.

To present XIV Nextra LUNs to the SAN Volume Controller, perform the following steps:
1. Create a map with all of the volumes that you intend to manage with the SAN Volume Controller cluster.
2. Link the WWPN for all node ports in the SAN Volume Controller cluster into the map. Each SAN Volume Controller node port WWPN is recognized as a separate host by XIV Nextra systems.

LUN presentation on IBM XIV Type Number 2810 systems

IBM XIV Storage System Type Number 2810 LUNs are presented to the SAN Volume Controller interface using the following rules:
- LUNs can be presented to one or more selected hosts or clusters.
- Clusters are collections of hosts.

To present IBM XIV Storage System Type Number 2810 LUNs to the SAN Volume Controller, perform the following steps:
1. Use the IBM XIV Storage System GUI to create an IBM XIV Storage System cluster for the SAN Volume Controller cluster.
2. Create a host for each node in the SAN Volume Controller.
3. Add a port to each host that you created in step 2. You must add a port for each port on the corresponding node.
4. Map volumes to the cluster that you created in step 1.

Target ports on XIV Nextra systems

XIV Nextra systems are single-rack systems. All XIV Nextra WWNNs include zeros as the last two hexadecimal digits. In the following example, WWNN 2000001738279E00 is IEEE extended; the WWNNs that start with the number 1 are IEEE 48 bit:
```
WWNN 2000001738279E00
WWPN 1000001738279E13
WWPN 1000001738279E10
WWPN 1000001738279E11
WWPN 1000001738279E12
```

Target ports on IBM XIV Type Number 2810 systems

IBM XIV Storage System Type Number 2810 systems are multi-rack systems, but only single racks are supported. All IBM XIV Storage System Type Number 2810 WWNNs include zeros as the last four hexadecimal digits. For example:

```
WWNN 5001738000030000
WWPN 5001738000030153
WWPN 5001738000030121
```

Switch zoning limitations for IBM XIV systems

There are limitations in switch zoning for SAN Volume Controller clusters and IBM XIV Storage System systems.

Fabric zoning

To avoid a single point of failure, the SAN Volume Controller switch zone must include at least one target port from each IBM XIV Storage System controller.

Target port sharing

Target ports can be shared between the SAN Volume Controller cluster and other hosts.

Host splitting

To avoid the possibility of an interaction between multipathing drivers, a single host cannot be connected to both the SAN Volume Controller cluster and the IBM XIV Storage System system.

Controller splitting

IBM XIV Storage System system LUNs that are mapped to the SAN Volume Controller cluster cannot be mapped to other hosts. IBM XIV Storage System system LUNs that are *not* mapped to the SAN Volume Controller cluster can be mapped to other hosts.

Configuration settings for IBM XIV systems

The IBM XIV Storage System Storage Management GUI provides configuration settings and options that can be used with SAN Volume Controller clusters.

Logical unit options and settings for IBM XIV systems

Logical unit (LU) settings for IBM XIV Storage System systems are configurable at the LU level.

Table 74 lists the options that must be set for each LU that is accessed by other hosts. LUs that are accessed by hosts can be configured differently. You can use the IBM XIV Storage System and XIV Nextra Storage Management GUI or CLI to change these settings.

Table 74. IBM XIV options and required settings

Option	Data Type	Range	IBM XIV Storage System and XIV Nextra default setting	SAN Volume Controller setting
Capacity	int	17,179,869,184 bytes (17 GB), up to the total system capacity ORBlock count	None	Any

Table 74. IBM XIV options and required settings (continued)

Option	Data Type	Range	IBM XIV Storage System and XIV Nextra default setting	SAN Volume Controller setting
Notes: • SAN Volume Controller supports up to 1 PB. • LUNs are allocated in 17-GB chunks. • Using a block count results in LUNs that are arbitrarily sized, but that still consume multiples of 17 GB.				

2

Host options and settings for IBM XIV systems

You must use specific settings to identify SAN Volume Controller clusters as hosts to IBM XIV Storage System systems.

An XIV Nextra host is a single WWPN, so one XIV Nextra host must be defined for each SAN Volume Controller node port in the cluster. An XIV Nextra host is considered to be a single SCSI initiator. Up to 256 XIV Nextra hosts can be presented to each port. Each SAN Volume Controller host object that is associated with the XIV Nextra system must be associated with the same XIV Nextra LUN map because each LU can only be in a single map.

An IBM XIV Storage System Type Number 2810 host can consist of more than one WWPN. Configure each SAN Volume Controller node as an IBM XIV Storage System Type Number 2810 host, and create a cluster of IBM XIV Storage System systems that corresponds to each of the SAN Volume Controller nodes in the SAN Volume Controller cluster.

Table 75 lists the host options and settings that can be changed using the IBM XIV Storage System and XIV Nextra Storage Management GUI.

Table 75. IBM XIV Type Number 2810 and XIV Nextra host options and required settings

Option	Data type	Range	IBM XIV Storage System Type Number 2810 and XIV Nextra default setting	SAN Volume Controller required setting	Notes
Type	Enum	FC/iSCSI	Not applicable	FC	The Type must be FC for SAN Volume Controller.
XIV Nextra map_set_special_type CLI command or IBM XIV Storage System Type Number 2810 special_type_set CLI command	Enum	default/hpux	default	default	This command is used by hpux hosts only. Do not use the command for SAN Volume Controller LUNs.

Option	Data type	Range	IBM XIV Storage System Type Number 2810 and XIV Nextra default setting	SAN Volume Controller required setting	Notes
WWPN	int64	Any	Not applicable	Node port	For XIV Nextra, one host for each node port WWPN in the cluster must be defined. For IBM XIV Storage System, Type Number 2810 one host port for each node port WWPN in the cluster must be defined.
LUN Map	String	Any	Not applicable	Any	For XIV Nextra, each SAN Volume Controller host in the XIV Nextra system must be associated with the same LUN map because each LU can be only in a single map.

Quorum disks on IBM XIV systems

The SAN Volume Controller cluster selects managed disks (MDisks) that are presented by IBM XIV Storage System systems as quorum disks.

Clearing SCSI reservations and registrations

You must not use the **vol_clear_keys** command to clear SCSI reservations and registrations on volumes that are managed by SAN Volume Controller.

Copy functions for IBM XIV Storage System models

Advanced copy functions for IBM XIV Storage System models such as taking a snapshot and remote mirroring cannot be used with disks that are managed by the SAN Volume Controller cluster. Thin provisioning is not supported for use with SAN Volume Controller.

Chapter 8. IBM System Storage support for Microsoft Volume Shadow Copy Service and Virtual Disk Service for Windows

The SAN Volume Controller provides support for the Microsoft Volume Shadow Copy Service and Virtual Disk Service. The Microsoft Volume Shadow Copy Service can provide a point-in-time (shadow) copy of a Windows host volume while the volume is mounted and files are in use. The Microsoft Virtual Disk Service provides a single vendor and technology-neutral interface for managing block storage virtualization, whether done by operating system software, RAID storage hardware, or other storage virtualization engines.

The following components are used to provide support for the service:
- SAN Volume Controller
- The cluster CIM server
- IBM System Storage hardware provider, known as the IBM System Storage Support for Microsoft Volume Shadow Copy Service and Virtual Disk Service software
- Microsoft Volume Shadow Copy Service
- The vSphere Web Services when it is in a VMware virtual platform

The IBM System Storage hardware provider is installed on the Windows host.

To provide the point-in-time shadow copy, the components complete the following process:
1. A backup application on the Windows host initiates a snapshot backup.
2. The Volume Shadow Copy Service notifies the IBM System Storage hardware provider that a copy is needed.
3. The SAN Volume Controller prepares the volumes for a snapshot.
4. The Volume Shadow Copy Service quiesces the software applications that are writing data on the host and flushes file system buffers to prepare for the copy.
5. The SAN Volume Controller creates the shadow copy using the FlashCopy Copy Service.
6. The Volume Shadow Copy Service notifies the writing applications that I/O operations can resume, and notifies the backup application that the backup was successful.

The Volume Shadow Copy Service maintains a free pool of volumes for use as a FlashCopy target and a reserved pool of volumes. These pools are implemented as virtual host systems on the SAN Volume Controller.

Installation overview

The steps for implementing the IBM System Storage Support for Microsoft Volume Shadow Copy Service and Virtual Disk Service software must be completed in the correct sequence.

Before you begin, you must have experience with or knowledge of administering a Windows Server operating system.

You must also have experience with or knowledge of administering a SAN Volume Controller.

Complete the following tasks:
1. Verify that the system requirements are met.
2. Install the IBM System Storage Support for Microsoft Volume Shadow Copy Service and Virtual Disk Service software.

3. Verify the installation.
4. Create a free pool of volumes and a reserved pool of volumes on the SAN Volume Controller.
5. Optionally, you reconfigure the services to change the configuration that you established during the installation.

System requirements for the IBM System Storage Support for Microsoft Volume Shadow Copy Service and Virtual Disk Service software

Ensure that your system satisfies the following requirements before you install the IBM System Storage Support for Microsoft Volume Shadow Copy Service and Virtual Disk Service software on a Microsoft Windows Server 2003 or 2008 operating system.

The following software is required:
- SAN Volume Controller must have licenses for FlashCopy enabled.
- IBM System Storage Support for Microsoft Volume Shadow Copy Service and Virtual Disk Service software version 4.0 or later.
- Windows Server 2003 R2 or later or Windows Server 2008 operating system. The following editions of Windows Server 2003 and Windows Server 2008 are supported:
  - Standard Server Edition 32-bit version
  - Enterprise Edition, 32-bit version
  - Standard Server Edition 64-bit version
  - Enterprise Edition, 64-bit version

Installing the IBM System Storage Support for Microsoft Volume Shadow Copy Service and Virtual Disk Service software

This section includes the steps to install the IBM System Storage Support for Microsoft Volume Shadow Copy Service and Virtual Disk Service software on a Windows server.

You must satisfy all of the prerequisites that are listed in the system requirements section before starting the installation.

Perform the following steps to install the IBM System Storage Support for Microsoft Volume Shadow Copy Service and Virtual Disk Service software on the Windows server:
1. Log on to Windows as an administrator.
2. Download the IBM VSS Host Installation Package file from the following Web site:

   Support for SAN Volume Controller (2145) website at www.ibm.com/storage/support/2145
3. Double click on the name of the file that you downloaded in step 2 to start the installation process. The Welcome panel is displayed.
4. Click **Next** to continue. The License Agreement panel is displayed. You can click **Cancel** at any time to exit the installation. To move back to previous screens while using the wizard, click **Back**.
5. Read the license agreement information. Select whether you accept the terms of the license agreement, and click **Next**. If you do not accept, you cannot continue with the installation. The Choose Destination Location panel is displayed.
6. Click **Next** to accept the default directory where the setup program will install the files, or click **Change** to select a different directory. Click **Next**. The Ready to Install the Program panel is displayed.
7. Click **Install** to begin the installation. To exit the wizard and end the installation, click **Cancel**. The Setup Status panel is displayed.

   The program setup verifies your configuration.

The Select CIM Server panel is displayed.

8. Select the required CIM server, or select **Enter the CIM Server address manually**, and click **Next**. The Enter CIM Server Details panel is displayed.

9. Enter the following information in the fields:

   - In the **CIM Server Address** field, type the name of the server where the SAN Volume Controller is installed. For example, enter `https://server.company.com:5989`.

   - In the **CIM User** field, type the user name for the SAN Volume Controller that the IBM System Storage Support for Microsoft Volume Shadow Copy Service and Virtual Disk Service software will use to gain access to the CIM server.

   - In the **CIM Password** field, type the password for the user name that the IBM System Storage Support for Microsoft Volume Shadow Copy Service and Virtual Disk Service software will use to gain access to the CIM server and click **Next**.

   Notes:

   a. If these settings change after installation, you can use the **ibmvcfg.exe** tool to update Microsoft Volume Shadow Copy and Virtual Disk Services software with the new settings.

   b. If you do not have the CIM agent server, port, or user information, contact your CIM agent administrator.

   The InstallShield Wizard Complete panel is displayed.

10. Click **Finish**. If necessary, the InstallShield Wizard prompts you to restart the system.

11. Make the IBM Hardware Provider for VSS-VDS aware of the SAN Volume Controller, as follows:

    a. Open a command prompt.

    b. Change directories to the hardware provider directory; the default directory is **C:\Program Files\IBM\Hardware Provider for VSS-VDS**.

    c. Use the **ibmvcfg** command to set the IP address for the SAN Volume Controller cluster, as follows:

    ```
    ibmvcfg set targetSVC ip_address
    ```

    The *ip_address* value must be the SAN Volume Controller cluster IP address.

Configuring the VMware Web Service connection

The IBM System Storage Support for Microsoft Volume Shadow Copy Service and Virtual Disk Service software provides support for VMware virtual platform since version 4.2, enabling the Shadow Copy Service in the virtual hosts with RDM disks.

Note: Only Shadow Copy Service for the RDM disks, acting as raw disks, and presented to the virtual host in physical mode is supported.

To manipulate RDM disks in the virtual host, the IBM System Storage Support for Microsoft Volume Shadow Copy Service and Virtual Disk Service software must interact with the VMware ESX Server. This is accomplished through the VMware Web Service exposed by the ESX Server, which holds the virtual host.

VMware tools, which collect host information such as IP address, hostname, and so on, must be installed so that the virtual host can communicate with the vSphere Web Service.

There are four parameters available only in the VMware virtual platform:
- vmhost
- vmuser
- vmpassword
- vmcredential

e Table 76 describes the parameters.

e *Table 76. VMware parameters*

Parameter	Description
vmhost	Specifies the vSphere Web Service location on the ESX Server, which holds the virtual host.
vmuser	Specifies the user that can log in to the ESX Server and has the privileges to manipulate the RDM disks.
vmpassword	Specifies the password for the vmuser to log in.
vmcredential	Specifies the session credential store path for the vSphere Web Service. The credential store can be generated by the Java keytool.

e Using the **ibmvcfg** command, perform the following steps to configure each parameter:

e 1. To configure vmhost, issue the following command: `ibmvcfg set vmhost https://ESX_Server_IP/sdk`

e 2. To configure vmuser, issue the following command: `ibmvcfg set vmuser username`

e 3. To configure vmpassword, issue the following command: `ibmvcfg set vmpassword password`

e 4. To generate the credential store path for the vSphere Web Service and set vmcredential, do the
e following:

e   a. Create a directory named VMware-Certs (at the root level) for the certificates, such as:
e   C:\VMware-Certs

e   b. Install the vSphere Client on the virtual host.

e   c. Launch the vSphere Client and then navigate to the ESX Server. A security warning message is
e   displayed.

e   d. Click **View Certificate** to display the Certificate Properties page.

e   e. Click the **Details** tab.

e   f. Click **Copy to File** to launch the Certificate Export wizard.

e   g. Select **DER encoded binary X.509** (the default), and click **Next**.

e   h. Click **Browse** and navigate to this subdirectory: C:\VMware-Certs subdirectory.

e   i. Enter a name for the certificate that identifies the server to which it belongs: `C:\VMware-Certs\`
e   `<servername>.cer`

e   j. Import the certificate using the Java keytool utility: `keytool -import -file C:\VMware-Certs\`
e   `<servername>.cer -keystore C:\VMware-Certs\vmware.keystore`

e   **Note:** The key tool is located in C:\Program Files\IBM\Hardware Provider for
e   VSS-VDS\jre\bin\keytool.exe.

e   k. At the **Enter keystore password:** prompt, type a password for the keystore.

e   l. At the **Trust this certificate? [no]:** prompt, type `yes` and press **Enter**. The message, Certificate was
e   added to keystore, is displayed.

e   m. Set vmcredential to be the vmware.keystore.path: `ibmvcfg set vmcredential "C:\VMware-Certs\`
e   `vmware.keystore"`

Creating the free and reserved pools of volumes

The IBM System Storage Support for Microsoft Volume Shadow Copy Service and Virtual Disk Service
software maintains a free and a reserved pool of volumes. Because these objects do not exist on the SAN
Volume Controller, the free and reserved pool of volumes are implemented as virtual host systems. You
must define these two virtual host systems on the SAN Volume Controller.

When a shadow copy is created, the IBM System Storage Support for Microsoft Volume Shadow Copy Service and Virtual Disk Service software selects a volume in the free pool, assigns it to the reserved pool, and then removes it from the free pool. This protects the volume from being overwritten by other Volume Shadow Copy Service users.

To successfully perform a Volume Shadow Copy Service operation, there must be enough volumes mapped to the free pool. The volumes must be the same size as the source volumes.

Use the management GUI or the SAN Volume Controller command-line interface (CLI) to perform the following steps:

1. Create a host for the free pool of volumes.
   - You can use the default name VSS_FREE or specify a different name.
   - Associate the host with the worldwide port name (WWPN) 5000000000000000 (15 zeroes).
2. Create a virtual host for the reserved pool of volumes.
   - You can use the default name VSS_RESERVED or specify a different name.
   - Associate the host with the WWPN 5000000000000001 (14 zeroes).
3. Map the logical units (VDisks) to the free pool of volumes.

   Restriction: The volumes cannot be mapped to any other hosts.
   - If you already have volumes created for the free pool of volumes, you must assign the volumes to the free pool.
4. Create host mappings between the volumes selected in step 3 and the VSS_FREE host to add the volumes to the free pool. Alternatively, you can use the **ibmvcfg add** command to add volumes to the free pool.
5. Verify that the volumes have been mapped.

If you do not use the default WWPNs 5000000000000000 and 5000000000000001, you must configure the IBM System Storage Support for Microsoft Volume Shadow Copy Service and Virtual Disk Service software with the WWPNs.

Verifying the installation

This task verifies that the IBM System Storage Support for Microsoft Volume Shadow Copy Service and Virtual Disk Service software is correctly installed on the Windows server.

Perform the following steps to verify the installation:

1. Click **Start** → **All Programs** → **Administrative Tools** → **Services** from the Windows server task bar. The **Services** panel is displayed.
2. Ensure that the service named IBM System Storage Support for Microsoft Volume Shadow Copy Service and Virtual Disk Service software appears and that **Status** is set to Started and **Startup Type** is set to Automatic.
3. Open a command prompt window and issue the following command:
   ```
   vssadmin list providers
   ```
4. Ensure that the service named IBM System Storage Support for Microsoft Volume Shadow Copy Service and Virtual Disk Service software is listed as a provider.
5. Use the **ibmvcfg list all** command to test the connection to the IBM System Storage Productivity Center or master console.

If you are able to successfully perform all of these verification tasks, the IBM System Storage Support for Microsoft Volume Shadow Copy Service and Virtual Disk Service software was successfully installed on the Windows server.

Changing the configuration parameters

You can change the parameters that you defined when you installed the IBM System Storage Support for Microsoft Volume Shadow Copy Service and Virtual Disk Service software. You must use the ibmvcfg.exe utility to change the parameters.

Table 77 describes the configuration commands that are provided by the ibmvcfg.exe utility.

Table 77. Configuration commands

Command	Description	Example
ibmvcfg showcfg	Lists the current settings.	`showcfg`
ibmvcfg set username <username>	Sets the user name to access the CIM server.	`set username johnny`
ibmvcfg set password <password>	Sets the password of the user name that will access the CIM server.	`set password mypassword`

ibmvcfg set targetSVC <ipaddress> or <Cluster Id>	Setting for proxy CIM Agent:Specifies the IP address of the SAN Volume Controller on which the volumes are located when volumes are moved to and from the free pool with the **ibmvcfg add** and **ibmvcfg rem** commands.
	
	Setting for embedded CIM Agent:
	
	Specifies the Cluster Id of the SAN Volume Controller. It can be found using the **svcinfo lscluster** command.
		`set targetSVC 64.157.185.191`
		
		`set targetSVC 0000020060600772`

ibmvcfg set backgroundCopy	Sets the background copy rate for FlashCopy.	`set backgroundCopy 80`
ibmvcfg set timeout	Set the timeout of CIM Agent's idle time that provider can wait when it has no responding. The time setting is in second. 0 is for infinite time.	`set timeout 5`

ibmvcfg set storageProtocol	This setting is to support iSCSI for SVC 5.1 or later.
	
	There are three settings: auto, FC, or iSCSI.
	
	Auto can be either FC or iSCSI if both protocols have been connected and defined. FC protocol will be applied.
		`set storageProtocol auto`
		
		`set storageProtocol FC`
		
		`set storageProtocol iSCSI`

ibmvcfg set incrementalFC	Specifies if incremental FlashCopy has to be used on SAN Volume Controller for the shadow copy.	`ibmvcfg set incrementalFC yes`
ibmvcfg set usingSSL	Specifying the Secure Sockets Layer (SSL) protocol is required to use a CIM server.	`ibmvcfg set usingSSL yes`

ibmvcfg set cimomHost <server IP> or <server name>	Sets the CIM server for the cluster.	`ibmvcfg set cimomHost 9.123.234.8`
		
		`ibmvcfg set cimomHost myCimhost.com.domain.controller.`

Table 77. Configuration commands (continued)

	Command	Description	Example
	ibmvcfg set namespace <namespace>	Specifies the namespace value that master console is using.	ibmvcfg set namespace \root\ibm
	ibmvcfg set vssFreeInitiator <WWPN>	Specifies the WWPN of the host. The default value is 5000000000000000. Modify this value only if there is a host already in your environment with a WWPN of 5000000000000000.	ibmvcfg set vssFreeInitiator 5000000000000000
	ibmvcfg set vssReservedInitiator <WWPN>	Specifies the WWPN of the host. The default value is 5000000000000001. Modify this value only if there is a host already in your environment with a WWPN of 5000000000000001.	ibmvcfg set vssFreeInitiator 5000000000000001
e e	ibmvcfg set vmhost https://ESX_Server_IP/sdk	Specifies the vSphere Web Service location on the ESX Server, which holds the virtual host.	ibmvcfg set vmhost https://9.11.110.90/sdk
e	ibmvcfg set vmuser username	Specifies the user that can log in to the ESX Server and has the privileges to manipulate the RDM disks.	ibmvcfg set vmuser root
e	ibmvcfg set vmpassword password	Sets the password for the vmuser to log in.	ibmvcfg set vmpassword pwd
e e	ibmvcfg set vmcredential credential_store	Specifies the session credential store path for the vSphere Web Service. The credential store can be generated by the Java keytool located in C:\Prorgam Files\IBM\Hardware Provider for VSS-VDS\jre\bin\ keytool.exe.	ibmvcfg set vmcredential "C:\VMware-Certs\vmware.keystore"

Adding, removing, or listing volumes and FlashCopy relationships

You can use the ibmvcfg.exe. utility to perform the pool management tasks of adding, removing, or listing volumes and FlashCopy relationships.

The IBM System Storage Support for Microsoft Volume Shadow Copy Service and Virtual Disk Service software maintains a free pool of volumes and a reserved pool of volumes. These pools are implemented as virtual host systems on the SAN Volume Controller.

Table 78 describes the ibmvcfg.exe commands for adding or removing volumes from the free pool of volumes and listing or deleting FlashCopy relationships.

Table 78. Pool management commands

	Command	Description	Example
t t	ibmvcfg list all -l	Lists all volumes, including information about volume ID, UUID, volume name, size, operational state, health status, type of volume, volumes to host mappings, and host name. Use the **l** parameter for output in verbose-list column format.	ibmvcfg list all ibmvcfg list all -l

Table 78. Pool management commands (continued)

Command	Description	Example
ibmvcfg list free -l	Lists the volumes that are currently in the free pool. Use the **l** parameter for output in verbose-list column format.	```ibmvcfg list free``` ```ibmvcfg list free -l```
ibmvcfg list reserved -l	Lists the volumes that are currently in the reserved pool. Use the **l** parameter for output in verbose-list column format.	```ibmvcfg list reserved``` ```ibmvcfg list reserved -l```
ibmvcfg list assigned -l	Lists the volumes that are currently in the assigned pool or host. Use the **l** parameter for output in verbose-list column format.	```ibmvcfg list assigned``` ```ibmvcfg list assigned -l```
ibmvcfg list unassigned -l	Lists the volumes that are currently in the unassigned pool or host. Use the **l** parameter for output in verbose-list column format.	```ibmvcfg list unassigned``` ```ibmvcfg list unassigned -l```
ibmvcfg list infc -l	Lists all the FlashCopy relationships on the SAN Volume Controller. This command lists both incremental and nonincremental FlashCopy relationships.	```ibmvcfg list infc``` ```ibmvcfg list infc -l```

ibmvcfg add	Adds one or more volumes to the free pool of volumes.	```ibmvcfg add```
		```600507680181801DC800000000000000```
		```ibmvcfg add vdisk17```

ibmvcfg rem	Removes one or more volumes from the free pool of volumes.	```ibmvcfg rem```
		```600507680181801DC800000000000001```
		```ibmvcfg rem vdisk18```

ibmvcfg del	Deletes one or more FlashCopy relationships. Use the serial number of the FlashCopy target to delete the relationship.	```ibmvcfg del```
		```600507680181801DC800000000000002```
		```ibmvcfg del vdisk19```

Error codes for IBM System Storage Support for Microsoft Volume Shadow Copy Service and Virtual Disk Service software

The IBM System Storage Support for Microsoft Volume Shadow Copy Service and Virtual Disk Service software logs error messages in the Windows Event Viewer and in private log files.

You can view error messages by going to the following locations on the Windows server where the IBM System Storage Support for Microsoft Volume Shadow Copy Service and Virtual Disk Service software is installed:

- The Windows Event Viewer in Application Events. Check this log first.
- The log file ibmVSS.log, which is located in the directory where the IBM System Storage Support for Microsoft Volume Shadow Copy Service and Virtual Disk Service software is installed.

Table 79 on page 265 lists the errors messages that are reported by the IBM System Storage Support for Microsoft Volume Shadow Copy Service and Virtual Disk Service software.

Table 79. Error messages for the IBM System Storage Support for Microsoft Volume Shadow Copy Service and Virtual Disk Service software

Code	Message	Symbolic name
1000	JVM Creation failed.	ERR_JVM
1001	Class not found: %1.	ERR_CLASS_NOT_FOUND
1002	Some required parameters are missing.	ERR_MISSING_PARAMS
1003	Method not found: %1.	ERR_METHOD_NOT_FOUND
1004	A missing parameter is required. Use the configuration utility to set this parameter: %1.	ERR_REQUIRED_PARAM
1600	The recovery file could not be created.	ERR_RECOVERY_FILE_ CREATION_FAILED
1700	ibmGetLunInfo failed in AreLunsSupported.	ERR_ARELUNSSUPPORTED_ IBMGETLUNINFO
1800	ibmGetLunInfo failed in FillLunInfo.	ERR_FILLLUNINFO_IBMGETLUNINFO
1900	Failed to delete the following temp files: %1	ERR_GET_TGT_CLEANUP
2500	Error initializing log.	ERR_LOG_SETUP
2501	Unable to search for incomplete Shadow Copies. Windows Error: %1.	ERR_CLEANUP_LOCATE
2502	Unable to read incomplete Shadow Copy Set information from file: %1.	ERR_CLEANUP_READ
2503	Unable to cleanup snapshot stored in file: %1.	ERR_CLEANUP_SNAPSHOT
2504	Cleanup call failed with error: %1.	ERR_CLEANUP_FAILED
2505	Unable to open file: %1.	ERR_CLEANUP_OPEN
2506	Unable to create file: %1.	ERR_CLEANUP_CREATE
2507	HBA: Error loading hba library: %1.	ERR_HBAAPI_LOAD
3000	An exception occurred. Check the ESSService log.	ERR_ESSSERVICE_EXCEPTION
3001	Unable to initialize logging.	ERR_ESSSERVICE_LOGGING
3002	Unable to connect to the CIM agent. Check your configuration.	ERR_ESSSERVICE_CONNECT
3003	Unable to get the Storage Configuration Service. Check your configuration.	ERR_ESSSERVICE_SCS
3004	An internal error occurred with the following information: %1.	ERR_ESSSERVICE_INTERNAL
3005	Unable to find the VSS_FREE controller.	ERR_ESSSERVICE_FREE_CONTROLLER
3006	Unable to find the VSS_RESERVED controller. Check your configuration.	ERR_ESSSERVICE_RESERVED_ CONTROLLER
3007	Unable to find suitable targets for all volumes.	ERR_ESSSERVICE_INSUFFICIENT_ TARGETS
3008	The assign operation failed. Check the CIM agent log for details.	ERR_ESSSERVICE_ASSIGN_FAILED

Table 79. Error messages for the IBM System Storage Support for Microsoft Volume Shadow Copy Service and Virtual Disk Service software (continued)

Code	Message	Symbolic name
3009	The withdraw FlashCopy operation failed. Check the CIM agent log for details.	ERR_ESSSERVICE_WITHDRAW_ FAILED

Uninstalling the IBM System Storage Support for Microsoft Volume Shadow Copy Service and Virtual Disk Service software

You must use Windows to uninstall the IBM System Storage Support for Microsoft Volume Shadow Copy Service and Virtual Disk Service software from the Windows server.

Perform the following steps to uninstall the software:

1. Log on to the Windows server as the local administrator.
2. Click **Start** → **Control Panel** from the task bar. The Control Panel window is displayed.
3. Double-click **Add or Remove Programs**. The Add or Remove Programs window is displayed.
t 4. Select **IBM System Storage Support for Microsoft Volume Shadow Copy Service and Volume Service software** and click **Remove**.
5. Click **Yes** when you are prompted to verify that you want to completely remove the program and all of its components.
6. Click **Finish**.

The IBM System Storage Support for Microsoft Volume Shadow Copy Service and Virtual Disk Service software is no longer installed on the Windows server.

Appendix. Accessibility

Accessibility features help a user who has a physical disability, such as restricted mobility or limited vision, to use software products successfully.

Features

These are the major accessibility features in the management GUI:

- You can use screen-reader software and a digital speech synthesizer to hear what is displayed on the screen. The following screen reader has been tested: Window-Eyes v6.1.
- You can operate all features using the keyboard instead of the mouse.
- When setting or changing an IP address on the SAN Volume Controller front panel, you can disable the fast increase function to reduce the address scrolling speed of the up and down buttons to two seconds. This feature is documented in the topic that discusses initiating cluster creation from the front panel, which is located in the IBM System Storage SAN Volume Controller Information Center and the *IBM System Storage SAN Volume Controller Software Installation and Configuration Guide*.

Navigating by keyboard

You can use keys or key combinations to perform operations and initiate many menu actions that can also be done through mouse actions. You can navigate the management GUI and help system from the keyboard by using the following key combinations:

- To traverse to the next link, button, or topic, press Tab inside a frame (page).
- To expand or collapse a tree node, press → or ←, respectively.
- To move to the next topic node, press V or Tab.
- To move to the previous topic node, press ^ or Shift+Tab.
- To scroll all the way up or down, press Home or End, respectively.
- To go back, press Alt+←.
- To go forward, press Alt+→.
- To go to the next frame, press Ctrl+Tab.
- To move to the previous frame, press Shift+Ctrl+Tab.
- To print the current page or active frame, press Ctrl+P.
- To select, press Enter.

Accessing the publications

You can find the HTML version of the IBM System Storage SAN Volume Controller information at the following Web site:

publib.boulder.ibm.com/infocenter/svc/ic/index.jsp

You can access this information using screen-reader software and a digital speech synthesizer to hear what is displayed on the screen. JAWS version 10 has been tested.

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:

Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan, Ltd.
1623-14, Shimotsuruma, Yamato-shi
Kanagawa 242-8502 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 PUBLICATION "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 publication. IBM may make improvements and/or changes in the product(s) and/or the 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
Almaden Research
650 Harry Road
Bldg 80, D3-304, Department 277
San Jose, CA 95120-6099
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 measurement 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 may 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.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

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. The sample programs are provided "AS IS", without warranty of any kind. IBM shall not be liable for any damages arising out of your use of the sample 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 Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at Copyright and trademark information at www.ibm.com/legal/copytrade.shtml.

Adobe and the Adobe logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.

Intel, Intel logo, Intel Xeon, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Java and all Java-based trademarks and logos are trademarks 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, Windows NT, and the Windows logo 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, or service names may be trademarks or service marks of others.

Index

Special characters

Numerics

A

B

C

V

W

X

Z

IBM ®

Printed in USA