IBM System Storage SAN Volume Controller

Planning Guide

Version 4.2.0

GA32-0551-01

IBM System Storage SAN Volume Controller

Planning Guide

Version 4.2.0

GA32-0551-01

This edition applies to the IBM System Storage SAN Volume Controller, release 4.2.0, and to all subsequent releases and modifications until otherwise indicated in new editions. This edition replaces GA32-0551-00.

Before using this information and the product it supports, read the information in ″Notices.″

© Copyright International Business Machines Corporation 2003, 2007. All rights reserved.
US Government Users Restricted Rights – Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

Contents

© Copyright IBM Corp. 2003, 2007

Figures

© Copyright IBM Corp. 2003, 2007

Tables

© Copyright IBM Corp. 2003, 2007

About this guide

This publication introduces the IBM System Storage SAN Volume Controller, its components, and its features.

It also provides planning guidelines for installing and configuring the SAN Volume Controller.

Who should use this guide?

This publication is intended for anyone who is planning to install and configure an IBM System Storage SAN Volume Controller.

Summary of changes

This document contains terminology, maintenance, and editorial changes.

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

Summary of changes for GA32-0551-01 SAN Volume Controller Planning Guide

The Summary of changes provides a list of new and changed information since the last version of the guide.

New information

This topic describes the changes to this guide since the previous edition, GA32-0551-00. The following sections summarize the changes that have since been implemented from the previous version.

This version includes the following new information:

- Multiple target FlashCopy mappings are now supported
- Mesh configurations are now supported
- A redundant ac power switch is now available

Changed information

This section lists the updates that were made in this document.

- There is a new SAN Volume Controller supported model. The SAN Volume Controller is now documented by model number. For example, this publication states four SAN Volume Controller model types: the SAN Volume Controller 2145-4F2, the SAN Volume Controller 2145-8F2, SAN Volume Controller 2145-8F4 and the new SAN Volume Controller 2145-8G4.

  Note: If text is referring to the SAN Volume Controller, it is referring to a generic SAN Volume Controller and can be referring to any of the SAN Volume Controller models. When the SAN Volume Controller is referred to as the SAN Volume Controller 2145-4F2, SAN Volume Controller

© Copyright IBM Corp. 2003, 2007

2145-8F2, SAN Volume Controller 2145-8F4, or the SAN Volume Controller 2145-8G4, the specific SAN Volume Controller is designated.

- The *IBM System Storage SAN Volume Controller Configuration Guide* is now titled *IBM System Storage SAN Volume Controller: Software Installation and Configuration Guide*.
- The *IBM System Storage SAN Volume Controller Installation Guide* is now titled *IBM System Storage SAN Volume Controller: Hardware Installation Guide*.
- The *IBM System Storage Master Console for SAN Volume Controller: Installation and User's Guide* and the *IBM System Storage Master Console for SAN Volume Controller* Information Center are no longer updated and distributed. Instead, all pertinent information from those information units has been incorporated into other SAN Volume Controller publications.

Summary of changes for GA32-0551-00 SAN Volume Controller Planning Guide

The Summary of changes provides a list of new, modified, and changed information since the last version of the guide.

New information

This topic describes the changes to this guide since the previous edition, GA22-1052-05. The following sections summarize the changes that have since been implemented from the previous version.

This version includes the following new information:

- There is a new SAN Volume Controller node model, the SAN Volume Controller 2145-8F4.

The following new topics have been added:

- Metro and Global Mirror
- Global Mirror
- Standard and persistent reserves

Changed information

This section lists the updates that were made in this document.

The following topics have been updated:

- SAN Volume Controller overview
- SAN Volume Controller operating environment
- UPS
- UPS configuration
- UPS operation
- Preparing your SAN Volume Controller environment
- Hardware location guidelines
- Cable connection table
- Switch zoning for the SAN Volume Controller
- Cluster configuration using SAN fabrics with long-distance fibre links
- Nodes
- UPS and power domains

- HBAs
- Power requirements
- Fibre-channel switches and interswitch links

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

A list of other publications that are related to this product are provided to you for your reference.

The tables in this section list and describe the following publications:

- The publications that make up the library for the IBM System Storage SAN Volume Controller
- Other IBM publications that relate to the SAN Volume Controller

SAN Volume Controller library

The following table lists and describes the publications that make up the SAN Volume Controller library. Unless otherwise noted, these publications are available in Adobe portable document format (PDF) from the following Web site:

http://www.ibm.com/storage/support/2145

Title	Description	Order number
IBM System Storage SAN Volume Controller: CIM Agent Developer's Reference	This reference guide describes the objects and classes in a Common Information Model (CIM) environment.	SC26-7904
IBM System Storage SAN Volume Controller: Command-Line Interface User's Guide	This guide describes the commands that you can use from the SAN Volume Controller command-line interface (CLI).	SC26-7903

Title	Description	Order number
IBM System Storage SAN Volume Controller: Software Installation and Configuration Guide	This guide provides guidelines for configuring your SAN Volume Controller.	SC23-6628
IBM System Storage SAN Volume Controller: Host Attachment Guide	This guide provides guidelines for attaching the SAN Volume Controller to your host system.	SC26-7905
IBM System Storage SAN Volume Controller: Hardware Installation Guide	This guide includes the instructions that the IBM service representative uses to install the SAN Volume Controller.	GC27-2132
IBM System Storage SAN Volume Controller: Planning Guide	This guide introduces the SAN Volume Controller and lists the features you can order. It also provides guidelines for planning the installation and configuration of the SAN Volume Controller.	GA32-0551
IBM System Storage SAN Volume Controller: Service Guide	This guide includes the instructions that the IBM service representative uses to service the SAN Volume Controller.	GC26-7901
IBM Safety Information	This guide contains translated caution and danger statements. Each caution and danger statement in the SAN Volume Controller documentation has a number that you can use to locate the corresponding statement in your language in the *IBM Safety Information* document.	N/A

Other IBM publications

The following table lists and describes other IBM publications that contain additional information related to the SAN Volume Controller.

Title	Description	Order number
IBM System Storage Multipath Subsystem Device Driver: User's Guide	This guide describes the IBM System Storage Multipath Subsystem Device Driver Version 1.6 for TotalStorage Products and how to use it with the SAN Volume Controller. This publication is referred to as the *IBM System Storage Multipath Subsystem Device Driver: User's Guide*.	GC27-2122

Title	Description	Order number
IBM TotalStorage DS4300 Fibre Channel Storage Subsystem Installation, User's, and Maintenance Guide	This guide describes how to install and configure the IBM TotalStorage DS4300 Fibre-Channel Storage Subsystem.	GC26-7722
IBM eServer xSeries 306m (Types 8849 and 8491) Installation Guide	This guide describes how to install the IBM eServer xSeries 306m, which is the hardware delivered for some versions of the hardware master console.	MIGR-61615
IBM xSeries 306m (Types 8849 and 8491) User's Guide	This guide describes how to use the IBM eServer xSeries 306m, which is the hardware delivered for some versions of the hardware master console.	MIGR-61901
IBM xSeries 306m (Types 8849 and 8491) Problem Determination and Service Guide	This guide can help you troubleshoot and resolve problems with the IBM eServer xSeries 306m, which is the hardware delivered for some versions of the hardware master console.	MIGR-62594
IBM eServer xSeries 306 (Type 8836) Installation Guide	This guide describes how to install the IBM eServer xSeries 306, which is the hardware delivered for some versions of the hardware master console.	MIGR-55080
IBM eServer xSeries 306 (Type 8836) User's Guide	This guide describes how to use the IBM eServer xSeries 306, which is the hardware delivered for some versions of the hardware master console.	MIGR-55079
IBM eServer xSeries 306 (Types 1878, 8489 and 8836) Hardware Maintenance Manual and Troubleshooting Guide	This guide can help you troubleshoot problems and maintain the IBM eServer xSeries 306, which is the hardware delivered for some versions of the hardware master console.	MIGR-54820
IBM eServer xSeries 305 (Type 8673) Installation Guide	This guide describes how to install the IBM eServer xSeries 305, which is the hardware delivered for some versions of the hardware master console.	MIGR-44200
IBM eServer xSeries 305 (Type 8673) User's Guide	This guide describes how to use the IBM eServer xSeries 305, which is the hardware delivered for some versions of the hardware master console.	MIGR-44199

Title	Description	Order number
IBM eServer xSeries 305 (Type 8673) Hardware Maintenance Manual and Troubleshooting Guide	This guide can help you troubleshoot problems and maintain the IBM eServer xSeries 305, which is the hardware delivered for some versions of the hardware master console.	MIGR-44094
IBM TotalStorage 3534 Model F08 SAN Fibre Channel Switch User's Guide	This guide introduces the IBM TotalStorage SAN Switch 3534 Model F08.	GC26-7454
IBM TotalStorage SAN Switch 2109 Model F16 User's Guide	This guide introduces the IBM TotalStorage SAN Switch 2109 Model F16.	GC26-7439
IBM TotalStorage SAN Switch 2109 Model F32 User's Guide	This guide introduces the IBM TotalStorage SAN Switch 2109 Model F32. It also describes the features of the switch and tells you where to find more information about those features.	GC26-7517

Some related publications are available from the following SAN Volume Controller support Web site:

http://www.ibm.com/storage/support/2145

Related Web sites

The following Web sites provide information about the SAN Volume Controller or related products or technologies.

Type of information	Web site
SAN Volume Controller support	http://www.ibm.com/storage/support/2145
Technical support for IBM storage products	http://www.ibm.com/storage/support/

How to order IBM publications

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

The IBM publications center

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 free of 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:

http://www.ibm.com/shop/publications/order/

How to send your comments

Your feedback is important to help us provide the highest quality information. If you have any comments about this book or any other documentation, you can submit them in one of the following ways:

- e-mail

  Submit your comments electronically to the following e-mail address:

  starpubs@us.ibm.com

  Be sure to include the name and order number of the book and, if applicable, the specific location of the text you are commenting on, such as a page number or table number.

- Mail

  Fill out the Readers' Comments form (RCF) at the back of this book. If the RCF has been removed, you can address your comments to:

  International Business Machines Corporation
  RCF Processing Department
  Department 61C
  9032 South Rita Road
  Tucson, Arizona 85775-4401
  U.S.A.

Chapter 1. SAN Volume Controller overview

The SAN Volume Controller combines hardware and software 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 subsystems. Those storage systems are then mapped to a set of virtual disks (VDisks) for use by attached host systems. System administrators can view and access a common pool of storage on the SAN. This lets the administrators 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. It allows a host system to be connected to a storage device across the network. The connections are made through units such as routers, gateways, hubs, and switches. The area of the network that contains these units is known as the *fabric* of the network.

The SAN Volume Controller is analogous to a logical volume manager (LVM) on a SAN. The SAN Volume Controller performs the following functions for the SAN storage that it controls:

- Creates a single pool of storage
- Provides logical unit virtualization
- Manages logical volumes
- Provides the following advanced functions for the SAN:
  - Large scalable cache
  - Copy Services
    - FlashCopy® (point-in-time copy)
    - Metro Mirror (synchronous copy)
    - Global Mirror (asynchronous copy)
    - Data migration
  - Space management
    - Mapping that is based on desired performance characteristics
    - Metering of service quality

Each SAN Volume Controller node is a rack-mounted unit that you can install in a standard Electrical Industries Association (EIA) 19-inch rack. The nodes are always installed in pairs, with one-to-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. Each virtual volume is defined to an I/O group. I/O groups take the storage that is presented to the SAN by the storage subsystems as MDisks and translates the storage into logical disks, known as VDisks, that are used by applications on the hosts. Each node must reside in only one I/O group and provide access to the VDisks in that I/O group.

There are four models of SAN Volume Controller nodes:

- SAN Volume Controller 2145-4F2
- SAN Volume Controller 2145-8F2
- SAN Volume Controller 2145-8F4

© Copyright IBM Corp. 2003, 2007

- SAN Volume Controller 2145-8G4

SAN Volume Controller operating environment

You must set up your SAN Volume Controller operating environment using the supported multipathing software and hosts.

Minimum requirements

You must set up your SAN Volume Controller operating environment according to the following requirements:

- Minimum of one pair of SAN Volume Controller nodes
- Minimum of two uninterruptible power supplies
- One master console per SAN installation for configuration

Note: You can order the master console for the SAN Volume Controller as a master console hardware option that comes preloaded with the master console software or a master console software option that you install on your own hardware.

Features of a SAN Volume Controller 2145-4F2 node

The SAN Volume Controller 2145-4F2 node has the following features:

- 19-inch rack mounted enclosure
- Two 2 Gbps 2-port fibre-channel adapters (four fibre-channel ports)
- 4 GB cache memory

Features of a SAN Volume Controller 2145-8F2 node

The SAN Volume Controller 2145-8F2 node has the following features:

- 19-inch rack mounted enclosure
- Two 2 Gbps 2-port fibre-channel adapters (four fibre-channel ports)
- 8 GB cache memory

Features of a SAN Volume Controller 2145-8F4 node

The SAN Volume Controller 2145-8F4 node has the following features:

- 19-inch rack mounted enclosure
- One 4-port 4 Gbps fibre-channel adapter (four fibre-channel ports)
- 8 GB cache memory

Features of a SAN Volume Controller 2145-8G4 node

The SAN Volume Controller 2145-8G4 node has the following features:

- 19-inch rack mounted enclosure
- One 4-port 4 Gbps fibre-channel adapter (four fibre-channel ports)
- 8 GB cache memory
- Two dual-core processors

Supported hosts

See the following Web site for a list of the supported operating systems:

http://www.ibm.com/servers/storage/software/virtualization/svc

Multipathing software

See the following Web site for the latest support and coexistence information:

http://www.ibm.com/servers/storage/software/virtualization/svc

User interfaces

The SAN Volume Controller provides the following user interfaces through the master console:

- The SAN Volume Controller Console, a Web-accessible graphical user interface (GUI) that supports flexible and rapid access to storage management information
- A command-line interface (CLI) that uses Secure Shell (SSH)

Application programming interfaces

The SAN Volume Controller provides an application programming interface called the Common Information Model (CIM) agent, which supports the Storage Management Initiative Specification (SMI-S) of the Storage Network Industry Association.

UPS

The uninterruptible power supply (UPS) provides a SAN Volume Controller node with a secondary power source if you lose power from your primary power source due to power failures, power sags, power surges, or line noise.

Unlike the traditional UPS that enables continued operation of the devices that they supply when power is lost, these UPS units are 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. Data is saved to the internal disk of the SAN Volume Controller node. The UPS units are required to power the SAN Volume Controller nodes even if the input power source is uninterruptible.

The SAN Volume Controller 2145-8F2, SAN Volume Controller 2145-8F4 and SAN Volume Controller 2145-8G4 nodes can operate only with the 2145 UPS-1U. The SAN Volume Controller 2145-4F2 node can operate with either the 2145 UPS or the 2145 UPS-1U.

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

UPS configuration

To provide full redundancy and concurrent maintenance, SAN Volume Controller nodes must be installed in pairs.

A 2145 UPS-1U powers one SAN Volume Controller node. All SAN Volume Controller model types are supported by the 2145 UPS-1U. A 2145 UPS powers one or two SAN Volume Controller 2145-4F2 nodes. If it is powering two SAN Volume Controller 2145-4F2 nodes, they must be in different I/O groups of the same SAN Volume Controller cluster.

To make the SAN Volume Controller cluster more resilient to power failure, the 2145 UPS-1Us can be connected to the redundant ac power switch. If a redundant ac power switch is not used, connecting the two UPSs that are powering an I/O group to different, independent electrical power sources allows the SAN Volume Controller cluster to continue to operate with reduced capacity if a single power source fails.

The UPS must be in the same rack as the nodes.

Attention: Do not connect the UPSs to an input power source that does not conform to standards.

Each UPS includes power (line) cords that connect the UPS to either a redundant ac power switch, if one exists, a rack power distribution unit (PDU), if one exists, or to an external power source.

The UPS is connected to the SAN Volume Controller nodes with a power cable and a signal cable. To avoid the possibility of power and signal cables being connected to different UPS units, these cables are wrapped together and supplied as a single field replaceable unit. The signal cables enable the SAN Volume Controller nodes to read status and identification information from the UPS.

UPS operation

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

If the UPS 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 UPS 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 UPS battery charge state indicates that it has sufficient capacity to power the SAN Volume Controller node long enough to save all of its memory to the disk drive. This is in the event of a power loss. The UPS has sufficient capacity to save all the data on the SAN Volume Controller node at least twice. For a fully-charged UPS, even after battery capacity has been used to power the SAN Volume Controller node while it saves DRAM data, sufficient battery capacity remains to allow the SAN Volume Controller node to become fully operational as soon as input power is restored.

Note: If input power is disconnected from a 2145 UPS, a fully-operational SAN Volume Controller node that is connected to that 2145 UPS performs a power-down sequence. This operation, which saves the configuration and cache data to an internal disk in the SAN Volume Controller node, typically takes about three minutes, at which time power is removed from the output of the 2145 UPS. In the event of a delay in the completion of the power-down sequence, the 2145 UPS output power is removed five minutes after the power is disconnected from the 2145 UPS. Because this operation is

controlled by the SAN Volume Controller node, a 2145 UPS that is not connected to an active SAN Volume Controller node does not shut off within the five-minute required period.

Important: Do not shut down a UPS without first shutting down the SAN Volume Controller node that it supports. Data integrity can be compromised by pushing the 2145 UPS power-off button or 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 UPS by pushing the 2145 UPS power-off button or 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 UPSs are shut down before the nodes they support, data can be corrupted.

If you have two SAN Volume Controller 2145-4F2 nodes that use 2145 UPSs in the same I/O group, you must connect these nodes to different 2145 UPSs. This configuration ensures that cache and cluster state information is protected in the event of a failure of either the UPS or the mainline power source.

When SAN Volume Controller nodes are added to the cluster, you must specify the I/O group that they are joining. The configuration interfaces check the UPS units and ensure that the two SAN Volume Controller nodes in the I/O group are not connected to the same UPS units.

Redundant ac power switch

The redundant ac power switch is an optional feature that makes the SAN Volume Controller nodes more resilient to power failure. The redundant ac power switch is not a replacement for an uninterruptible power supply (UPS), you must still use a 2145 UPS-1U for each node.

You must connect the redundant ac power switch to two-independent power circuits. One power circuit connects to the main power input port and the other power circuit connects to the backup power input port. If the main power to the SAN Volume Controller node fails for any reason, the redundant ac power switch automatically uses the backup power source. When power is restored, the redundant ac power switch automatically changes back to using the main power source.

Place the redundant ac power switch in the same rack as the SAN Volume Controller node. The redundant ac power switch logically sits between the rack power distribution unit and the 2145 UPS-1U.

You can use a single redundant ac power switch to power one or two SAN Volume Controller nodes. If you use the redundant ac power switch to power two nodes, the nodes must be in different I/O groups. In the event that the redundant ac power switch fails or requires maintenance, both nodes power off. Because the nodes are in two different I/O groups, the hosts do not lose access to the backend disk data.

For maximum resilience to failure, use one redundant ac power switch to power each SAN Volume Controller node.

The following graphic shows a redundant ac power switch.

Master console

The master console provides a single point from which to manage the SAN Volume Controller nodes. You can purchase the master console as a hardware product option (which includes the master console preinstalled software) or as a software-only option.

The two master console options are the same in function and software. However, the planning, installation, and configuration processes are slightly different:

Master console hardware option
The manufacturing plant installs the software on the hardware using the default settings. After the IBM service representative installs the hardware option, you must configure and customize the default factory settings.

Master console software-only option
You must provide your own hardware and perform both the installation and configuration processes.

The master console provides you with the following functions:

- A platform on which the subsystem configuration tools can be run
- A platform for remote service, which allows the desktop to be shared with remote IBM service personnel if assistance is required to resolve complex problems
- Access to the following components:
  - SAN Volume Controller Console, which is a graphical user interface application, through a Web browser
  - SAN Volume Controller command-line interface, through a Secure Shell (SSH) session

The master console can support up to four SAN Volume Controller clusters.

Supported master console configurations

Ensure that your planned master console configuration is one that is supported.

Use the following rules when planning your master console configuration:

- One master console can concurrently support up to four SAN Volume Controller clusters.

- Multiple master consoles can access a single cluster, but when multiple master consoles access one cluster, you cannot concurrently perform configuration and service tasks.

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 (AOS) tool to remotely access the master console. This type of remote service can help you reduce service costs and shorten repair times.

AOS 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 AOS, the master console must be able to access the Internet. The following Web site provides further information about AOS:

http://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 AOS session. A plug-in is downloaded onto your master console to connect you and your IBM service representative to the remote service session. AOS 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 AOS.

Secure Shell

Secure Shell (SSH) is a client-server protocol that can be used from the master console or a host to enable you to control the SAN Volume Controller via a command-line interface (CLI).

The master console for the SAN Volume Controller distributes PuTTY, which provides an SSH client.

SSH provides a secure communications channel between systems. You can configure SSH to use a key pair (a private key and a public key) to establish the secure connection.

If you purchase the master console hardware option, you generate the SSH key pair when you configure the master console. If you purchase the master console software-only option, you generate the SSH key pair as part of the master console installation process. If you want to create an SSH connection (such as the SAN Volume Controller nodes), you must place the public key on every system.

Master console software components

The master console feature is a collection of different software components.

The master console includes the following software components:

- SAN Volume Controller Console

- PuTTY (SSH client software)
- Adobe Reader

Sending notifications

The SAN Volume Controller can use SNMP traps, Call Home e-mail, and Inventory Information e-mail to provide necessary data and event notifications to you and to the IBM Support Center.

The following types of information are sent from the SAN Volume Controller:

- Simple Network Management Protocol (SNMP) traps
- Call Home e-mail
- Inventory information

Simple Network Management Protocol traps

Simple network management protocol (SNMP) is the standard protocol for managing networks and exchanging messages. SNMP enables the SAN Volume Controller to send external messages that notify personnel about an event. An SNMP manager allows you to view the messages that the SNMP agent sends. You can use the SAN Volume Controller Console or the SAN Volume Controller command-line interface to configure and modify your SNMP settings. SNMP traps and Call Home e-mail can be sent simultaneously.

Call Home e-mail

The Call Home feature allows the transmission of operational and error-related data to you and IBM through an SMTP server connection in the form of an event notification e-mail. This function alerts IBM service personnel about machine conditions or sends data for error analysis and resolution.

You must configure an SMTP server with access outside of your local area network and use the SAN Volume Controller Console or the SAN Volume Controller command-line interface to configure the Call Home settings (including contact information and e-mail recipients) before you can send a test e-mail to check that all connections and infrastructure are connected correctly. You can disable the Call Home function at any time using the SAN Volume Controller Console or the SAN Volume Controller command-line interface.

Call Home support is initiated for the following reasons or types of data:

- Problem or event notification: Data is sent when there is a problem or event that might require the attention of IBM service personnel.
- Communication tests: You can test for the successful installation and communication infrastructure.
- Inventory information: A notification is sent to provide the necessary status and hardware information to IBM service personnel.

Call Home e-mails can contain any combination of the following types of information:

- Contact name
- Contact phone number
- Offshift phone number
- Machine location

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

Inventory information e-mail

Inventory information e-mail 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 e-mail function, you must meet the Call Home function requirements and enable the Call Home e-mail function before you can attempt to send inventory information e-mail. You can adjust the contact information, adjust the frequency of inventory e-mail, or manually send an inventory e-mail using the SAN Volume Controller Console or the SAN Volume Controller command-line interface.

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

- Timestamp
- Contact information, including name and phone number. This defaults to the contact information that was set for the Call Home e-mail function. However, you can change the contact information specifically for inventory e-mail using the SAN Volume Controller Console or the **mkemailuser** or **chemailuser** CLI commands.
- Machine location. This is the machine location that is set for the Call Home e-mail function.
- Software level
- License information. This is the same information that it output from the **svcinfo lslicense** command.
- Cluster vital product data (VPD). The cluster VPD is the same information that is output from the **svcinfo lscluster** command, including the following items:
  - Cluster name and IDs
  - Cluster location
  - Bandwidth
  - IP addresses
  - Memory capacities
  - SNMP settings
  - Time zone setting

  - E-mail settings
  - Microcode level
  - Fibre-channel port speed
- Node VPD for each node in the cluster. The node VPD is the same information that is output from the **svcinfo lsnodevpd** command, including the following items:
  - System part number
  - Number of various hardware parts, such as fans, processors, memory slots, fibre-channel cards, and SCSI/IDE devices
  - Part numbers of the various hardware parts
  - BIOS information
  - System manufacturing information, such as system product and manufacturer
  - Firmware level for the service processor
- Software VPD, including the following items:
  - Code level
  - Node name
  - Ethernet status
  - Worldwide node name (WWNN)
  - MAC address
- Processor information, including the following items for each processor:
  - Location of processor
  - Type of cache
  - Size of cache
  - Manufacturer
  - Version
  - Speed
  - Status (enabled or disabled)
- Memory information, including the following items:
  - Part number
  - Device location
  - Bank location
  - Size
- Fibre-channel card information, including the following items:
  - Part number
  - Port number
  - Device serial number
  - Manufacturer
- SCSI/IDE device information, including the following items:
  - Part number
  - Bus ID
  - Device ID
  - Model
  - Revision level
  - Serial number
  - Approximate capacity

- Front panel assembly information, including the following items:
  - Part number
  - ID
  - Location
- Universal power supply (UPS) information, including the following items:
  - Electronics part number
  - Battery part number
  - UPS assembly part number
  - Input power cable part number
  - UPS serial number
  - UPS type
  - UPS internal part number
  - ID
  - Firmware levels

Object descriptions

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 unit is a single processing unit 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.

Virtual disks (VDisks) are logical disks that are presented by the clusters. Each VDisk is associated with a particular I/O group. The nodes in the I/O group provide access to the VDisks in the I/O group. When an application server performs I/O to a VDisk, it can access the VDisk 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 (UPS)* which provides data integrity in the event of a cluster wide power failure. In such situations, the UPS maintains power to the nodes while the contents of the distributed cache are dumped to an internal drive.

The nodes in a cluster see the storage that is presented by backend *disk controllers* as a number of disks, known as *managed disks (MDisks)*. Because the nodes do not attempt to provide recovery from physical disk failures within the backend disk controllers, an MDisk is usually, but not necessarily, a RAID.

Each MDisk is divided up 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 *managed disk groups (MDisk groups)*. VDisks are created from the extents that are contained by an MDisk group. The MDisks that constitute a particular VDisk must all come from the same MDisk group.

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.

The nodes detect the fibre-channel ports that are connected to the SAN. These correspond to the worldwide port names (WWPNs) of the host bus adapter (HBA) fibre-channels 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.

Application servers can only access VDisks that have been allocated to them. VDisks can be mapped to a host object. Mapping a VDisk to a host object makes the VDisk accessible to the WWPNs in that host object, and hence the application server itself.

The cluster provides block-level aggregation and volume management for disk storage within the SAN. In simpler terms, this means that the cluster manages a number of backend storage controllers and maps the physical storage within those controllers 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 backend physical storage. This prevents any possible conflict between the cluster and the application servers both trying to manage the backend storage.

Copy Service features

The SAN Volume Controller provides Copy Service features that enable you to copy virtual disks (VDisks).

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

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

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

Global Mirror
Provides a consistent copy of a source VDisk on a target VDisk. Data is written to the target VDisk 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

FlashCopy is a Copy Service feature that is available with the SAN Volume Controller.

The FlashCopy feature copies the contents of a source virtual disk (VDisk) to a target VDisk. Any data that existed on the target VDisk is lost and is replaced by the copied data. After the copy operation has completed, the target VDisks contain the contents of the source VDisks 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 FlashCopy operation takes some time to complete, the resulting data on the target VDisk 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 old location.

FlashCopy operations can occur on multiple source and target VDisks. FlashCopy management operations are coordinated to allow a common single point in time for copying target VDisks from their respective source VDisks. This allows a consistent copy of data that spans multiple VDisks. For SAN Volume Controller version 4.2.0 or higher, the FlashCopy feature also allows multiple target VDisks to be copied from each source VDisk. This can be used to create images from different points in time for each source VDisk.

Source VDisks and target VDisks must meet the following requirements:

- They must be the same size.
- The same cluster must manage them.

Metro Mirror and Global Mirror

The Metro Mirror and Global Mirror Copy Services enable you to set up a relationship between two virtual disks (VDisks), so that updates that are made by an application to one VDisk are mirrored on the other VDisk.

Although the application only writes to a single VDisk, the SAN Volume Controller 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 the SAN Volume Controller Metro Mirror and Global Mirror operations between two 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 VDisk is designated the primary and the other VDisk is designated the secondary. Host applications write data to the primary VDisk, and updates to the primary VDisk are copied to the secondary VDisk. Normally, host applications do not perform I/O operations to the secondary VDisk.

Metro Mirror provides a synchronous-copy process. When a host writes to the primary VDisk, it does not receive confirmation of I/O completion until the write operation has completed for the copy on both the primary VDisk and the secondary VDisk. This ensures that the secondary VDisk is always up-to-date with the primary VDisk 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 VDisk.

Global Mirror provides an asynchronous-copy process. When a host writes to the primary VDisk, confirmation of I/O completion is received before the write operation has completed for the copy on the secondary VDisk. If a failover

operation is performed, the application must recover and apply any updates that were not committed to the secondary VDisk.

Metro Mirror and Global Mirror support the following features:

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

  Note: A cluster can only participate in active Metro Mirror and Global Mirror relationships with itself and one other cluster.
- Intercluster and intracluster Metro Mirror and Global Mirror can be used concurrently within a cluster.
- The intercluster link is bidirectional. This means that it can copy data from clusterA to clusterB for one pair of VDisks while copying data from clusterB to clusterA for a different pair of VDisks.
- 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.

Metro Mirror

The Metro Mirror Copy Service provides a *synchronous* copy, which means that the primary virtual disk (VDisk) is always an exact match of the secondary VDisk.

The host application writes data to the primary VDisk but does not receive confirmation that the write operation is complete until the data is written to the secondary VDisk. For disaster recovery, this mode provides the simplest mode of operation because a synchronous copy of the data is maintained. Metro Mirror is constrained by the latency time and bandwidth limitations that are imposed by the communication link to the secondary site.

Global Mirror

The Global Mirror Copy Service provides an asynchronous copy because the secondary virtual disk (VDisk) is not an exact match of the primary VDisk at every point in time.

The host application writes data to the primary VDisk and receives confirmation that the write operation is complete before the data is actually written to the secondary VDisk. This feature is comparable to a continuous backup process in which the last few updates are always missing. When you use Global Mirror for disaster recovery, you must consider how you want to handle these missing updates.

If I/O operations on the primary VDisk are paused for a small length of time, the secondary VDisk can become an exact match of the primary VDisk.

Chapter 2. Planning for your SAN Volume Controller installation

Before the IBM service representative can set up your SAN Volume Controller environment, you must verify that the prerequisite conditions for the SAN Volume Controller, the uninterruptible power supply (UPS), and redundant ac power switch installation are met.

1. Does your physical site meet the environment requirements for the SAN Volume Controller, UPS, and the redundant ac power switch?
2. Do you have adequate rack space for your hardware? Ensure you have the following rack space for your components:
   - The SAN Volume Controller: One Electrical Industries Association (EIA) unit high for each node.
   - The 2145 uninterruptible power supply (2145 UPS): Two EIA units high for each 2145 UPS.
   - The 2145 uninterruptible power supply-1U (2145 UPS-1U): One EIA unit high for each 2145 UPS-1U.
   - If you are using a redundant ac power switch: One EIA unit high for each redundant ac power switch
3. Do you have power distribution units in the rack to provide power to the UPS units or redundant ac power switch units?

   A clearly visible and accessible emergency power off switch is required.

   If you are using the redundant ac power switch, two independent power circuits are required. One circuit connects to the redundant ac power switch main input and the other circuit connects to the redundant ac power switch backup input.
4. Ensure that you provide appropriate connectivity by preparing your environment.
5. Ensure that you have a keyboard and display available in the event that a service action is required. The SAN Volume Controller 2145-8G4 requires a USB keyboard.

Preparing your SAN Volume Controller environment

Before installing the SAN Volume Controller, you must prepare the physical environment.

Input-voltage requirements

Ensure that the following requirements are met:

Voltage	Frequency
200 to 240 V single phase ac	50 or 60 Hz

Power requirements for each SAN Volume Controller node

Ensure that the following power is available for each node. The power that is required depends on the node type, the uninterruptible power supply (UPS) type, and whether the redundant ac power feature is used.

© Copyright IBM Corp. 2003, 2007

Components	Power requirements
SAN Volume Controller 2145-8G4 + 2145 UPS-1U	470 W
SAN Volume Controller 2145-8F4 + 2145 UPS-1U	520 W
SAN Volume Controller 2145-8F2 + 2145 UPS-1U	520 W
SAN Volume Controller 2145-4F2 + 2145 UPS-1U	420 W
One SAN Volume Controller 2145-4F2 + one 2145 UPS	760 W
Two SAN Volume Controller 2145-4F2 nodes + one 2145 UPS	1120 W

For each redundant ac power switch, add 20 W to the power requirements.

Circuit breaker requirements

Ensure that the following circuit breaker requirements are met:

- The 2145 UPS-1U has an integrated circuit breaker and does not require additional protection.
- Each 2145 UPS is connected to a separate branch circuit, which includes a UL-listed 15 A circuit breaker.
- If a redundant ac power switch is used, both input circuits to the redundant ac power switch are protected with a circuit breaker.

  The circuit breaker rating requirements are shown in the following table:

Number of 2145 UPS-1Us connected to the redundant ac power switch	Circuit breaker minimum rating	Circuit breaker maximum rating
1	5 A	16 A
2	7 A	16 A

Environment requirements without redundant ac power

Environment	Temperature	Altitude	Relative humidity	Maximum wet bulb temperature
Operating in lower altitudes	10°C to 35°C (50°F to 95°F)	0 to 914 m (0 to 2998 ft)	8% to 80% noncondensing	23°C (74°F)
Operating in higher altitudes	10°C to 32°C (50°F to 88°F)	914 to 2133 m (2998 to 6988 ft)	8% to 80% noncondensing	23°C (74°F)
Powered off	10°C to 43°C (50°F to 110°F)	0 to 2133 m (2998 to 6988 ft)	8% to 80% noncondensing	27°C (81°F)
Storing	1°C to 60°C (34°F to 140°F)	0 to 2133 m (0 to 6988 ft)	5% to 80% noncondensing	29°C (84°F)
Shipping	-20°C to 60°C (-4°F to 140°F)	0 to 10668 m (0 to 34991 ft)	5% to 100% condensing, but no precipitation	29°C (84°F)

Environment requirements with redundant ac power

Environment	Temperature	Altitude	Relative humidity	Maximum wet bulb temperature
Operating in lower altitudes	15°C to 32°C (59°F to 89°F)	0 to 914 m (0 to 2998 ft)	20% to 80% noncondensing	23°C (74°F)
Operating in higher altitudes	15°C to 32°C (50°F to 88°F)	914 to 2133 m (2998 to 6988 ft)	20% to 80% noncondensing	23°C (74°F)
Powered off	10°C to 43°C (50°F to 110°F)	0 to 2133m (0 to 6988 ft)	20% to 80% noncondensing	27°C (81°F)
Storing	1°C to 60°C (34°F to 140°F)	0 to 2133 m (0 to 6988 ft)	5% to 80% noncondensing	29°C (84°F)
Shipping	-20°C to 60°C (-4°F to 140°F)	0 to 10668 m (0 to 34991 ft)	5% to 100% condensing, but no precipitation	29°C (84°F)

Preparing your SAN Volume Controller environment

The following three tables list the physical dimensions and weight of the node, the additional space requirements around a node, and the maximum heat dissipated by a node that you must consider before you install your SAN Volume Controller 2145-8F2, SAN Volume Controller 2145-8F4, or SAN Volume Controller 2145-8G4:

Dimensions and weight

Height	Width	Depth	Maximum weight
43 mm (1.69 in.)	440 mm (17.32 in.)	686 mm (27 in.)	12.7 kg (28 lb)

Additional space requirements

Location	Additional space requirements	Reason
Left and right sides	50 mm (2 in.)	Cooling air flow
Back	Minimum: 100 mm (4 in.)	Cable exit

Heat output

Model	Heat output per node
SAN Volume Controller 2145-8G4	400 W (1350 btu per hour)
SAN Volume Controller 2145-8F4	450 W (1540 btu per hour)
SAN Volume Controller 2145-8F2	450 W (1540 btu per hour)
SAN Volume Controller 2145-4F2	350 W (1200 btu per hour)

Power cables for the 2145 UPS-1U and the redundant ac power switch

If you do not connect the 2145 uninterruptible power supply-1U (2145 UPS-1U) to a rack power distribution unit (PDU) or redundant ac power switch, you must follow your country or region's power requirements to choose the appropriate power cable for the 2145 UPS-1U.

The 2145 UPS-1U is supplied with an IEC 320-C13 to C14 jumper to connect it to a rack PDU. You can also use this cable to connect the 2145 UPS-1U to the redundant ac power switch. The redundant ac power switch is supplied with two IEC 320-C19 to C20 power cables to connect to rack PDUs. There are no country specific cables for the redundant ac power switch.

The following table lists the power cable requirements for your country or region:

Country or region	Length	Connection type (attached plug designed for 200-240V AC input)	Part
United States of America (Chicago), Canada, Mexico	1.8 m (6 ft)	NEMA L6-15P	7842122
Bahamas, Barbados, Bermuda, Bolivia, Brazil, Canada, Cayman Islands, Colombia, Costa Rica, Dominican Republic, Ecuador, El Salvador, Guatemala, Guyana, Haiti, Honduras, Jamaica, Japan, Korea (South), Liberia, Mexico, Netherlands Antilles, Nicaragua, Panama, Peru, Philippines, Saudi Arabia, Suriname, Taiwan, Trinidad (West Indies), United States of America, Venezuela	2.8 m (9 ft)	NEMA L6-15P	7842123
Antigua, Bahrain, Brunei, Channel Islands, Hong Kong S.A.R. of China, Cyprus, Dubai, Fiji, Ghana, India, Iraq, Ireland, Kenya, Kuwait, Malawi, Malaysia, Malta, Nepal, Nigeria, Polynesia, Qatar, Sierra Leone, Singapore, Tanzania, Uganda, United Kingdom, Yemen, Zambia	2.8 m (9 ft)	BS 1363/A	14F0033
Argentina, Australia, China (PRC), New Zealand, Papua New Guinea, Paraguay, Uruguay, Western Samoa	2.8 m (9 ft)	AZ/NZS C112	13F9940

Country or region	Length	Connection type (attached plug designed for 200-240V AC input)	Part
Afghanistan, Algeria, Andorra, Angola, Austria, Belgium, Benin, Bulgaria, Burkina Faso, Burundi, Cameroon, Central African Republic, Chad, China (Macau S.A.R.), Czech Republic, Egypt, Finland, France, French Guiana, Germany, Greece, Guinea, Hungary, Iceland, Indonesia, Iran, Ivory Coast, Jordan, Lebanon, Luxembourg, Malagasy, Mali, Martinique, Mauritania, Mauritius, Monaco, Morocco, Mozambique, Netherlands, New Caledonia, Niger, Norway, Poland, Portugal, Romania, Senegal, Slovakia, Spain, Sudan, Sweden, Syria, Togo, Tunisia, Turkey, former USSR, Vietnam, former Yugoslavia, Zaire, Zimbabwe	2.8 m (9 ft)	CEE 7-VII	13F9979
Denmark	2.8 m (9 ft)	DK2-5a	13F9997
Bangladesh, Burma, Pakistan, South Africa, Sri Lanka	2.8 m (9 ft)	SABS 164	14F0015
Liechtenstein, Switzerland	2.8 m (9 ft)	1011-S2450 7	14F0051
Chile, Ethiopia, Italy, Libya, Somalia	2.8 m (9 ft)	CEI 23-16	14F0069
Israel	2.8 m (9 ft)	SI 32	14F0087

Power cables for the 2145 UPS

You must follow the power requirements of your country or region to choose the appropriate power cable for the 2145 uninterruptible power supply (2145 UPS).

The following table lists the power cable requirements for your country or region:

Country or region	Length	Connection type (attached plug designed for 200-240V AC input)	Part
United States of America (Chicago), Canada, Mexico	1.8 m (6 ft)	NEMA L6-15P	14F1549

Country or region	Length	Connection type (attached plug designed for 200-240V AC input)	Part
Bahamas, Barbados, Bermuda, Bolivia, Brazil, Canada, Cayman Islands, Colombia, Costa Rica, Dominican Republic, Ecuador, El Salvador, Guatemala, Guyana, Haiti, Honduras, Jamaica, Japan, Korea (South), Liberia, Mexico, Netherlands Antilles, Nicaragua, Panama, Peru, Philippines, Saudi Arabia, Suriname, Taiwan, Trinidad (West Indies), United States of America, Venezuela	2.5 m (8 ft)	NEMA L6-15P	12J5119
Antigua, Bahrain, Brunei, Channel Islands, China (Hong Kong S.A.R.), Cyprus, Denmark, Dubai, Fiji, Ghana, India, Iraq, Ireland, Kenya, Kuwait, Malawi, Malaysia, Malta, Nepal, Nigeria, Polynesia, Qatar, Sierra Leone, Singapore, Tanzania, Uganda, United Kingdom, Yemen, Zambia	2.5 m (8 ft)	IEC 309	36L8822
Argentina, Australia, China (PRC), New Zealand, Papua New Guinea, Paraguay, Uruguay, Western Samoa	2.5 m (8 ft)	L6-20P	12J5118
Afghanistan, Albania, Algeria, Andorra, Angola, Austria, Belgium, Benin, Bulgaria, Burkina Faso, Burundi, Cameroon, Central African Republic, Chad, China (Macau S.A.R.), Czech Republic, Egypt, Finland, France, French Guiana, Germany, Greece, Guinea, Hungary, Iceland, Indonesia, Iran, Ivory Coast, Jordan, Lebanon, Luxembourg, Malagasy, Mali, Martinique, Mauritania, Mauritius, Monaco, Morocco, Mozambique, Netherlands, New Caledonia, Niger, Norway, Poland, Portugal, Romania, Senegal, Slovakia, Spain, Sudan, Sweden, Syria, Togo, Tunisia, Turkey, former USSR, Vietnam, former Yugoslavia, Zaire, Zimbabwe	2.5 m (8 ft)	CEE7	55H6643
Bangladesh, Burma, Pakistan, South Africa, Sri Lanka	2.5 m (8 ft)	SABS 164	12J5124
Thailand	2.5 m (8 ft)	NEMA 6-15P	12J5120

Preparing your UPS environment

Ensure that your physical site meets the installation requirements for the uninterruptible power supply (UPS).

The 2145 UPS-1U

The SAN Volume Controller 2145-4F2, SAN Volume Controller 2145-8F2, SAN Volume Controller 2145-8F4, and SAN Volume Controller 2145-8G4 are supported with the 2145 uninterruptible power supply-1U (2145 UPS-1U).

The 2145 UPS

Only the SAN Volume Controller 2145-4F2 is supported with the 2145 uninterruptible power supply (2145 UPS).

Attention: Ensure that you comply with the following requirements for UPS units:

- If the UPS is cascaded from another UPS, the source UPS must have at least three times the capacity per phase and the total harmonic distortion must be less than 5%.
- The UPS must also have input voltage capture that has a slew rate of no more than 3 Hz per second.

UPS specifications

2145 UPS-1U dimensions and weight

Height	Width	Depth	Maximum weight
44 mm (1.73 in.)	439 mm (17.3 in.)	579 mm (22.8 in.)	18.8 kg (41.4 lb)

2145 UPS dimensions and weight

Height	Width	Depth	Maximum weight
89 mm (3.5 in.)	483 mm (19 in.)	622 mm (24.5 in.)	37 kg (84 lb)

Heat output

Model	Heat output during normal operation	Heat output during battery operation
2145 UPS-1U	40 W (135 btu per hour)	150 W (510 btu per hour)
2145 UPS powering two SAN Volume Controller 2145-4F2 nodes	140 W (480 btu per hour)	250 W (850 btu per hour)

Preparing your redundant ac power environment

Ensure that your physical site meets the installation requirements for the redundant ac power switch.

The redundant ac power switch requires two independent power sources, provided through two, rack mounted power distribution units (PDUs). The PDUs must have IEC320-C19 outlets.

The power cable between the redundant ac power switch and the 2145 uninterruptible power supply-1U (2145 UPS-1U) is rated at 10 A.

Redundant ac power switch specifications

Dimensions and weight

Height	Width	Depth	Maximum weight
43 mm (1.69 in.)	192 mm (7.56 in.)	240 mm	2.6 kg (5.72 lb)

Additional space requirements

Location	Width	Reason
Left side	124 mm (4.89 in.)	Side mounting plate
Right side	124 mm (4.89 in.)	Side mounting plate

Heat output (maximum)

The maximum heat output dissipated inside the redundant ac power switch is approximately 50 watts (170 btu per hour).

Ports and connections

Ensure that you are familiar with the specific ports and connections types for the SAN Volume Controller and the uninterruptible power supply (UPS).

Each SAN Volume Controller requires the following ports and connections:

- Each SAN Volume Controller node requires one Ethernet cable to connect it to an Ethernet switch or hub. A 10/100 Mb Ethernet connection is required.
- Two TCP/IP addresses are normally required for a SAN Volume Controller cluster: a cluster address and a service address.
- Each SAN Volume Controller node has four fibre-channel ports, which are supplied fitted with LC-style optical small form-factor pluggable (SFP) gigabit interface converters (GBICs) for connection to a fibre-channel switch.

Each UPS requires serial cables that connect the UPS to the SAN Volume Controller nodes. Ensure that for each node, the serial and power cables come from the same UPS.

Preparing your master console hardware option environment

If the master console hardware option is included with the SAN Volume Controller, ensure that your physical site meets the installation requirements for the master console hardware and the console monitor kit.

Server dimensions and weight

The following tables provide the physical characteristics and the environmental requirements for the master console hardware option:

Height	Width	Depth	Maximum Weight
43 mm (1.7 in.)	430 mm (16.69 in.)	424 mm (16.69 in.)	12.7 kg (28 lb)

Note: The above dimensions are for a 1U monitor and keyboard assembly.

Server AC and input-voltage requirements

Power Supply	Electrical Input
203 watt (110 or 220 V ac auto-sensing)	Sine-wave input (47–63 Hz) required Input voltage low range: Minimum: 100 V ac Maximum: 127 V ac Input voltage high range: Minimum: 200 V ac Maximum: 240 V ac Input kilovolt-amperes (kVA), approximately: Minimum: 0.0870 kVA Maximum: 0.150 kVA

Server environment

Environment	Temperature	Altitude	Relative Humidity
Server On	10° to 35°C (50°F to 95°F)	0 to 914 m (2998.0 ft)	8% to 80%
Server Off	-40°C to 60°C (-104°F to 140°F)	Maximum: 2133 m (6998.0 ft)	8% to 80%

Server heat output

Approximate heat output in British thermal units (BTU) per hour:

- Minimum configuration: 87 watts (297 BTU)
- Maximum configuration: 150 watts (512 BTU)

Monitor console kit dimensions and weight

Height	Width	Depth	Maximum Weight
43 mm (1.7 in.)	483 mm (19.0 in.)	483 mm (19.0 in.)	17.0 kg (37.0 lb)

Master console hardware option components

Your master console software is preinstalled when you order the master console hardware option with your SAN Volume Controller.

The following components are included with the master console hardware option:

- 19-inch 1U rack-mounted IBM® eServer™ xSeries® server
- 19-inch 1U flat panel monitor and keyboard

Attention: If more than one power distribution bus is available, the two power connectors, one supplying the master console server and the other supplying the master console monitor, must be connected to the same power distribution bus.

The following software is preinstalled on the system:

- Microsoft® Windows® 2003 Standard Server Edition with the latest service pack
- PuTTY, a client for Telnet and Secure Shell (SSH) protocol communications
  - Putty.exe, the client software
  - Puttygen.exe, a utility for generating encryption keys
  - Plink.exe, the command-line interface to the PuTTY client software
- SAN Volume Controller Console
- Adobe Reader

The following Web site provides the current list of supported software versions:

http://www.ibm.com/storage/support/2145

After the master console hardware is installed, the customer must configure the software to meet their requirements.

Prerequisites for the software master console

Before you install the software master console, ensure that you have the prerequisite hardware and software.

Location requirements

The master console must be in the same room and within 50 feet of the SAN Volume Controller nodes.

Hardware prerequisites

The master console software can be installed on a rack-mounted, high-performance, highly-reliable Intel® server (such as the IBM eServer xSeries 3250 or equivalent) with the following options:

- One Intel Xeon® dual core processor, minimum 2.1 GHz
- Minimum of 4 GB of system memory
- Two IDE hard disk drives, minimum 160 GB each. During installation of the master console software, you will mirror these drives.
- One CD drive
- One Gigabit port for an Ethernet connection (fiber or copper)
- One keyboard, such as the Space Saver NLS keyboard or equivalent
- One monitor, such as Netbay 1U Flat Panel Monitor Console kit without keyboard or equivalent
- One mouse or equivalent pointing device

Example hardware configuration

- IBM xSeries 3250 server (1U)
- Intel Xeon 3.0 GHz processor
- 4 GB memory DIMM (256 MB comes with base unit)
- Two 160 GB IDE hard disk drives (one comes with base unit)

- One Gigabit 10/100/1000 Copper Ethernet port on planar
- NetBay 1U Flat Panel Monitor Console Kit with US keyboard

Software prerequisites

The software version of the master console requires that you obtain the following software:

- Operating system
  - The software version of the master console requires that one of the following operating systems is provided on your hardware platform:
    - Microsoft Windows Server 2003 Enterprise Edition
    - Microsoft Windows Server 2003 Standard Edition
    - Microsoft Windows 2000 Server 5.00.2195.

  Note: The master console hardware option is shipped with Microsoft Windows Server 2003, Standard Edition pre-installed.
- Microsoft Windows Internet Explorer version 6.0 with Service Pack 1 or a later version of version 6

  Notes:

  1. You cannot use Microsoft Windows Internet Explorer version 7 with the SAN Volume Controller Console.
  2. You can download Internet Explorer version 6 from the following Web site:

     http://www.microsoft.com/windows/ie/ie6/downloads/
- Antivirus software (not required but necessary to ensure protection of your computer).

Web browser requirements for the SAN Volume Controller Console

You must have the correct Web browser to access the SAN Volume Controller Console.

The Web browsers that you can use to access the SAN Volume Controller Console depend on the following conditions:

- Whether you are accessing the SAN Volume Controller Console from the master console hardware or from a remote system
- The operating system from which you are accessing the SAN Volume Controller Console

Table 1 on page 26 provides information about the Web browsers that are supported by the SAN Volume Controller Console.

Table 1. Supported Web browsers for accessing the SAN Volume Controller Console

Supported Web browsers on the master console system	Supported Web browsers on a remote system
• Microsoft Internet Explorer version 6.0 with Service Pack 2 or a later version of version 6 (See Notes 1 and 2) • Netscape version 6 (See Note 3) • Mozilla Firefox version 1 or later (See Note 4)	For Windows operating systems: • Microsoft Internet Explorer version 6.0 with Service Pack 2 or a later version of version 6 (See Notes 1 and 2) • Netscape version 6 (See Note 3) • Mozilla Firefox version 1 or later (See Note 4) For AIX® operating systems: • Mozilla-based browser (based on Mozilla version 1) (See Note 5) • Netscape version 6 (See Note 3)
Notes: 1. You cannot use Microsoft Windows Internet Explorer version 7 with the SAN Volume Controller Console. 2. Internet Explorer version 6 is available from the following Web site: http://www.microsoft.com/windows/ie/ie6/downloads/ 3. Netscape is available from the following Web site: http://browser.netscape.com/ns8/download/archive.jsp 4. Mozilla Firefox is available from the following Web site: http://www.mozilla.com/firefox/ 5. Mozilla is available from the following Web site: http://www.mozilla.org/releases/	

Chapter 3. Planning the physical configuration

Before you or your IBM service representative installs the SAN Volume Controller nodes, uninterruptible power supply (UPS) units, and master console, you must plan the physical configuration and the initial settings for the system.

1. Download the hardware location chart, the cable connection table, and the configuration data table from the following Web site:

   http://www.ibm.com/storage/support/2145
2. Use the hardware location chart to record the physical configuration of your system.
3. Use the cable connection table to record how your SAN Volume Controller units, the UPS units, and the master console are to be connected.
4. Use the configuration data table to record the data that you and the IBM service representative need before the initial installation.

When you or your IBM service representative have completed these tasks, you can perform the physical installation.

Completing the hardware location chart

The hardware location chart represents the rack into which the SAN Volume Controller nodes are installed. Each row of the chart represents one Electrical Industries Association (EIA) 19-inch rack space.

- Uninterruptible power supply (UPS) units are heavy. Install them as near the bottom of the rack as possible. Place them within the range of row 1 through row 8.
- Do not exceed the maximum power rating of the rack and input power supply.
- Position the SAN Volume Controller such that information on the display screen is easily viewed and the controls used to navigate the display menu are easily reached. Place the SAN Volume Controller in EIA 11-36.
- To enable easy access to the connectors on the rear of the master console hardware, position the console, keyboard and monitor unit adjacent to each other. To permit easy access to the CD drive, locate the master console server above the keyboard and monitor unit. Place the master console hardware in EIA 17-24.

  Note: If you must locate the master console hardware in a rack other than those that hold the SAN Volume Controller nodes, be sure that the master console is located within 50 feet of the nodes that it supports.
- A SAN Volume Controller node is one EIA unit high. Therefore, for each SAN Volume Controller node that is to be installed, fill in the row that represents the position that the SAN Volume Controller node is to occupy.
- The 2145 uninterruptible power supply (2145 UPS) is two EIA units high. Therefore, for each 2145 UPS, fill in two rows.
- The 2145 uninterruptible power supply-1U (2145 UPS-1U) is one EIA unit high. Therefore, for each 2145 UPS-1U, fill in one row.
- The redundant ac power switch is one EIA unit high. Therefore, for each redundant ac power switch, fill in one row.

© Copyright IBM Corp. 2003, 2007

- The master console hardware option is two EIA units high: one EIA unit for the server and one EIA unit for the keyboard and monitor. If you install the master console software option on your own hardware, plan for the size needed by your hardware, including any keyboard and monitor.
- If there are any hardware devices already in the rack, record this information on the chart.
- Fill in rows for all other units that are present in the rack, including Ethernet hubs and fibre-channel switches. Hubs and switches are usually one EIA unit high, but check with your supplier. The UPS units must be installed at the bottom of the rack because it might be necessary to relocate some other devices before the SAN Volume Controller node installation is started.

Hardware location guidelines

Ensure that you are familiar with the hardware location guidelines.

You can download the hardware location chart from the following Web site:

http://www.ibm.com/storage/support/2145

Use the following guidelines to fill in the hardware location chart:

- Install the SAN Volume Controller nodes in pairs to provide redundancy and concurrent maintenance.
- A cluster can contain no more than eight SAN Volume Controller nodes.
- Connect each SAN Volume Controller node of a pair to a different uninterruptible power supply (UPS) unit.
- To reduce the chance of a simultaneous input power failure at both UPS units, either use the redundant ac power switch or connect each UPS unit to a separate electrical power source on a separate branch circuit.
- Because UPS units are heavy, install them in the lowest available positions of the rack. If necessary, move any lighter components that are already in the rack to higher positions.
- The IBM service representative does not install the Ethernet hub or the Fibre Channel switches. You must arrange for either the suppliers or someone in your organization to install those items. Provide the installer with a copy of the completed hardware location chart.
- You might want to put the switches between the SAN Volume Controller nodes. The UPS units must remain in the lowest position of the rack.

Completing the cable connection table

The cable connection table helps you connect the units that will be placed in the rack.

You can download the cable connection table from the following Web site:

http://www.ibm.com/storage/support/2145

The following terms and descriptions assist you in completing the cable connection table for the SAN Volume Controller nodes:

Term	Description
Uninterruptible power supply (UPS)	The UPS to which the SAN Volume Controller is connected.
Ethernet	The Ethernet hub or switch to which the SAN Volume Controller is connected.
Fibre-channel ports 1 through 4	The Fibre Channel switch ports to which the four SAN Volume Controller Fibre Channel ports are connected. When viewed from the back of the SAN Volume Controller, the ports are numbered 1 through 4, from left to right. Ignore the markings on the back of the SAN Volume Controller.

For the hardware used for the master console, use the following terms and descriptions to complete the cable connection table:

Term	Description
Ethernet to SAN Volume Controller	The Ethernet port used to connect the master console and the SAN Volume Controller.
Ethernet to Internet for remote service	(Optional) The Ethernet port used to connect the master console to the Internet for remote service. Remote service is managed by the IBM Assist On-site (AOS) tool, which is a cross-product Internet tool used by IBM and not a component of the master console. You can determine how you want to provide the port. For example, you can provide the set up and configure the port by using any of the following methods: • Provide Internet access through the same port that is used to access the SAN Volume Controller nodes. This method does not provide as much security as the other example methods. • Use an alternate Ethernet port. For added security if you choose this method, you can disconnect this port and connect it only when a remote service connection is needed. • Choose not to provide the master console with any access to the public Internet and allow AOS to access another console that can access the master console Web server over an intranet.

Completing the configuration data table

You must fill out the configuration data table before the SAN Volume Controller and master console can be installed.

You can download the configuration data table from the following Web site:

http://www.ibm.com/storage/support/2145

Include the following initial settings for the cluster:

Term	Description
Language	The national language in which you want the messages displayed on the front panel. This option applies only to service messages. The default setting is English.
Cluster IP address	The address that is used for all typical configuration and service access to the cluster.
Service IP address	The address that is used for emergency access to the cluster.

Term	Description
Gateway IP address	The IP address for the default local gateway for the cluster.
Subnet mask	The subnet mask of the cluster.

Include the following information for the master console:

Term	Description
Machine name	The name that you want to call the master console. This must be a fully qualified DNS name. If you order the master console hardware option, you must change the factory default setting, which is *mannode*, because this is not a fully qualified DNS name.
Master console IP addresses	The addresses that are used for access to the master console. If you order the master console hardware option, the default settings are: Port 1 = 192.168.1.11 Port 2 = 192.168.1.12
Master console gateway IP address	The IP address for the local gateway for the master console. If you order the master console hardware option, the default setting is 192.168.1.1.
Master console subnet mask	The subnet mask for the master console. If you order the master console hardware option, the default is 255.255.255.0.

Completing the redundant ac power switch connection chart

If you are using the optional redundant ac power switch feature, you must fill out the redundant ac power switch connection chart before this feature can be installed.

You can download the redundant ac power switch connection chart from the following Web site:

http://www.ibm.com/storage/support/2145

Determine if you want to use the redundant ac power switch to power one or two nodes. If you plan to power two nodes, the nodes must be in different I/O groups. Therefore, a cluster with a single I/O group requires two redundant ac power switch units.

You must plan the route of the input power cables for the redundant ac power switch units. These cables are connected to the front edge of the redundant ac power switch and must be routed to the power distribution unit for the rack. You might have to leave a free slot in the rack to allow the cables to run from front to back.

Chapter 4. Preparing to use the SAN Volume Controller in a SAN environment

Ensure that you perform the required preparation steps to ensure proper use of the SAN Volume Controller in a SAN environment.

Follow these preparation steps to set up your SAN Volume Controller environment:

1. Plan your SAN environment.
2. Create the RAID resources that you intend to virtualize.
3. Determine if you have a RAID that contains data that you want to merge into the cluster.
4. Determine if you will migrate data into the cluster or keep them as image-mode virtual disks (VDisks).
5. Determine if you will use the Copy Service features.

Preparing to install the SAN Volume Controller into an existing SAN environment

Ensure your environment meets the necessary requirements to support the use of a SAN Volume Controller.

To install a SAN Volume Controller into an existing SAN that will be in use during installation, you must first ensure that the switch zoning is set to isolate the new SAN Volume Controller connections from the active part of the SAN.

See the following Web site for specific firmware levels and the latest supported hardware:

http://www.ibm.com/storage/support/2145

- Consider the design of the SAN according to your requirement for high availability.
- Identify the operating system for each host system that will be connected to the SAN Volume Controller, ensuring compatibility and suitability by performing the following steps:
  1. Specify the host bus adapters (HBAs) for each host.
  2. Define the performance requirements.
  3. Determine the total storage capacity.
  4. Determine the storage capacity per host.
  5. Determine the host LUN sizes.
  6. Determine the total number of ports and bandwidth that are needed between the host and the SAN Volume Controller.
  7. Determine if your SAN has enough ports to connect all hosts and back-end storage.
- Ensure that the existing SAN components meet the requirements for the SAN Volume Controller by performing the following steps:
  1. Determine the host system versions.

© Copyright IBM Corp. 2003, 2007

2. Ensure that the HBAs, switches, and controllers are at or above the minimum requirements.
3. Identify any components that must be upgraded.

Example SAN environments

Ensure that you are familiar with the constraints for zoning a switch.

Example 1

Consider the SAN environment in the following example:

- Two nodes (nodes A and B)
- Nodes A and B each have four ports
  1. Node A has ports A0, A1, A2, and A3
  2. Node B has ports B0, B1, B2, and B3
- Four hosts called P, Q, R, and S
- Each of the four hosts has four ports, as described in Table 2.

Table 2. Four hosts and their ports

P	Q	R	S
P0	Q0	R0	S0
P1	Q1	R1	S1
P2	Q2	R2	S2
P3	Q3	R3	S3

- Two switches called X and Y
- One storage controller
- The storage controller has four ports on it called I0, I1, I2, and I3

The following is an example configuration:

1. Attach ports 1 (A0, B0, P0, Q0, R0, and S0) and 2 (A1, B1, P1, Q1, R1, and S1) of each node and host to switch X.
2. Attach ports 3 (A2, B2, P2, Q2, R2, and S2) and 4 (A3, B3, P3, Q3, R3, and S3) of each node and host to switch Y.
3. Attach ports 1 and 2 (I0 and I1) of the storage controller to switch X.
4. Attach ports 3 and 4 (I2 and I3) of the storage controller to switch Y.

Create the following host zones on switch X:

5. Create a host zone containing ports 1 (A0, B0, P0, Q0, R0, and S0) of each node and host.
6. Create a host zone containing ports 2 (A1, B1, P1, Q1, R1, and S1) of each node and host.

Create the following host zones on switch Y:

7. Create a host zone on switch Y containing ports 3 (A2, B2, P2, Q2, R2, and S2) of each node and host.
8. Create a host zone on switch Y containing ports 4 (A3, B3, P3, Q3, R3, and S3) of each node and host.

Create the following storage zone:

9. Create a storage zone that is configured on each switch. Each storage zone contains all the SAN Volume Controller and storage ports on that switch.

Example 2

The following example describes a SAN environment that is similar to the previous example except for the addition of two hosts that have two ports each.

- Two nodes called A and B
- Nodes A and B have four ports each
  1. Node A has ports A0, A1, A2, and A3
  2. Node B has ports B0, B1, B2, and B3
- Six hosts called P, Q, R, S, T and U
- Four hosts have four ports each and the other two hosts have two ports each as described in Table 3.

Table 3. Six hosts and their ports

P	Q	R	S	T	U
P0	Q0	R0	S0	T0	U0
P1	Q1	R1	S1	T1	U1
P2	Q2	R2	S2	—	—
P3	Q3	R3	S3	—	—

- Two switches called X and Y
- One storage controller
- The storage controller has four ports on it called I0, I1, I2, and I3

The following is an example configuration:

1. Attach ports 1 (A0, B0, P0, Q0, R0, S0 and T0) and 2 (A1, B1, P1, Q1, R1, S1 and T0) of each node and host to switch X.
2. Attach ports 3 (A2, B2, P2, Q2, R2, S2 and T1) and 4 (A3, B3, P3, Q3, R3, S3 and T1) of each node and host to switch Y.
3. Attach ports 1 and 2 (I0 and I1) of the storage controller to switch X.
4. Attach ports 3 and 4 (I2 and I3) of the storage controller to switch Y.

Attention: Hosts T and U (T0 and U0) and (T1 and U1) are zoned to different SAN Volume Controller ports so that each SAN Volume Controller port is zoned to the same number of host ports.

Create the following host zones on switch X:

5. Create a host zone containing ports 1 (A0, B0, P0, Q0, R0, S0 and T0) of each node and host.
6. Create a host zone containing ports 2 (A1, B1, P1, Q1, R1, S1 and U0) of each node and host.

Create the following host zones on switch Y:

7. Create a host zone on switch Y containing ports 3 (A2, B2, P2, Q2, R2, S2 and T1) of each node and host.
8. Create a host zone on switch Y containing ports 4 (A3, B3, P3, Q3, R3, S3 and U1) of each node and host.

Create the following storage zone:

9. Create a storage zone configured on each switch. Each storage zone contains all the SAN Volume Controller and storage ports on that switch.

Zoning guidelines

Ensure that you are familiar with the zoning guidelines for controller 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 number of paths to a VDisk would be 8 × 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 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 controller zones

Switch zones that contain controller 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 see the back-end storage and the front-end host HBAs. Typically, the front-end host HBAs and the back-end storage are not in the same zone. The exception to this is where split host and split controller configuration is in use. All nodes in a cluster must be able to see the same set of back-end storage ports on each back-end controller. Operation in a mode where two nodes see a different set of ports on the same controller is degraded and the system logs errors that request a repair action. This can occur if inappropriate zoning was applied to the fabric or if inappropriate LUN masking is used. This rule has important implications for back-end storage, such as an IBM DS4000, which imposes exclusive rules for mappings between HBA worldwide node names (WWNNs) and storage partitions.

It is possible to zone the switches in such a way that a SAN Volume Controller port is used solely for internode communication, or for communication to a host, or for communication to back-end storage. This is possible because each node contains four ports. Each node must still remain connected to the full SAN fabric. Zoning cannot be used to separate the SAN into two parts.

It is critical that you configure the controller and SAN so that a cluster cannot access LUs that a host or another cluster can also access. This can be arranged by controller LUN mapping and masking.

All nodes in a cluster must see at least one node port for all nodes that are in the cluster, but nodes cannot see node ports for nodes that belong to another cluster. You can have nodes that are not members of any cluster zoned to see all of the clusters. This allows you to add a node to the cluster in the event that you must replace a node.

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 and local nodes or remote nodes, or both, is not valid.

If a node can see another node through multiple paths, use zoning where possible to ensure that the node to node communication does not travel over an ISL. If a node can see a storage controller through multiple paths, use zoning to restrict communication to those paths that do not travel over ISLs.

Host zones

The configuration rules for host zones are different depending upon the number of hosts that will access the cluster. For smaller configurations of less than 64 hosts per cluster, the SAN Volume Controller supports a simple set of zoning rules which enable a small set of host zones to be created for different environments. For larger configurations of more than 64 hosts, the SAN Volume Controller supports a more restrictive set of host zoning rules.

Zoning that contains host HBAs must not contain either host HBAs in dissimilar hosts or dissimilar HBAs in the same host that 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 subsystem 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 possible, place each HBA port in a host that connects to a node into a separate zone. You *should* also 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 the 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 64 to 256 hosts

For clusters with 64 to 256 hosts attached, each HBA port in a host that connects to a node *must* be placed into a separate zone. In this separate zone, you must also include exactly one port from each node in the I/O groups that are associated with this host.

Clusters with 256 to 1024 hosts

For clusters with 256 to 1024 hosts attached, the SAN must be zoned so that each HBA port in a host that connects to a node can only see one SAN Volume Controller port for each node in the I/O group that is associated with the host. If you have 1024 hosts, each host must be associated with only one I/O group and each I/O group must only be associated with up to 256 hosts.

Figure 1 provides an example configuration for zoning 1024 hosts. In this example, the hosts are arranged into four groups of 256 hosts and each group of 256 hosts is zoned to one I/O group. You must zone each group of 256 hosts separately so they cannot see other hosts that are in different I/O groups. The controller zone contains all eight of the nodes and all four of the controllers. The intercluster zone contains all of the nodes that are in both clusters to allow you to use Metro Mirror and Global Mirror.

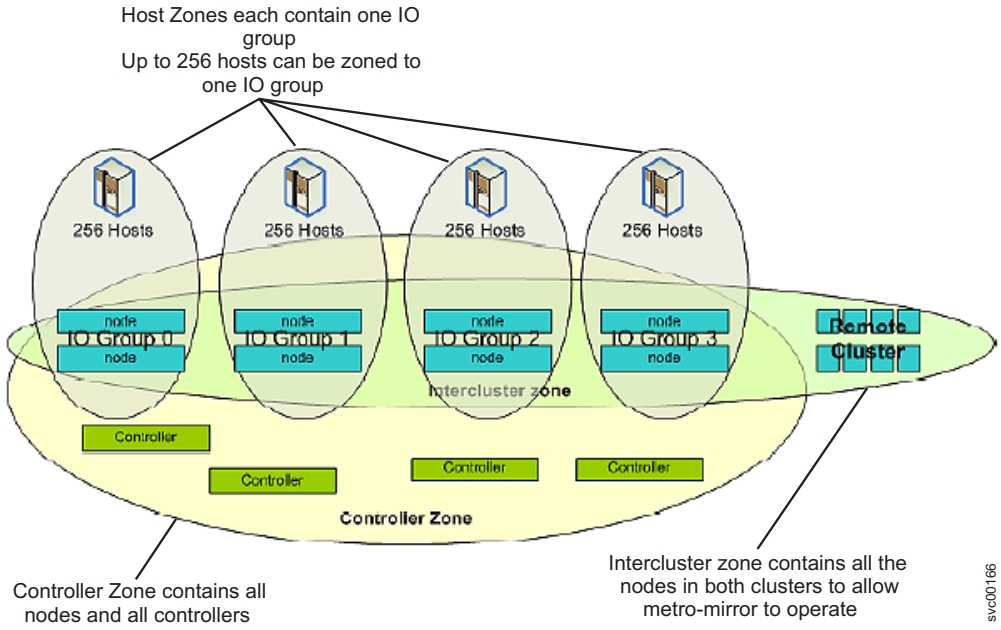


Figure 1. Zoning a 1024 host configuration

You can have up to 512 Fibre Channel logins per Fibre Channel port. The following logins are counted towards the 512 login maximum:

- Host port logins
- Storage controller logins
- SAN Volume Controller node logins
- Fibre Channel name server logins

If any port has more than 512 logins, the node logs an `ID 073006` error. You can use the **svcinfo lsfabric** command-line interface (CLI) command or the SAN Volume Controller Console to view the logins that are seen by each SAN Volume Controller port.

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 Copy Services.

SAN configurations that use intracluster Metro Mirror and Global Mirror relationships do not require additional switch zones.

SAN configurations that use intercluster Metro Mirror and Global Mirror relationships require the following additional switch zoning considerations:

- The clusters must be zoned so that the nodes in each cluster can see the ports of the nodes in the other cluster.
- Use of interswitch link (ISL) trunking in a switched fabric.
- Use of redundant fabrics.

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. Create a new zone in each fabric that contains all SAN Volume Controller ports for both clusters.
3. 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.
4. Verify that cluster A cannot recognize any of the back-end storage that is owned by cluster B. Two clusters cannot share the same back-end storage devices.

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 performance. The two most significant features are ISL trunking and extended fabric.

The following table provides a description of the ISL trunking and the extended fabric features:

Feature	Description
ISL trunking	Trunking enables the switch to use two links in parallel and still maintain frame ordering. It does this by routing all traffic for a given destination over the same route even when there might be more than one route available. Often trunking is limited to certain ports or port groups within a switch. For example, in the IBM 2109-F16 switch, trunking can only be enabled between ports in the same quad (for example, same group of four ports). For more information on trunking with the MDS, refer to ″Configuring Trunking″ on the Cisco Systems Web site. Some switch types can impose limitations on concurrent use of trunking and extended fabric operation. For example, with the IBM 2109-F16 switch, it is not possible to enable extended fabric for two ports in the same quad. Thus, extended fabric and trunking cannot be used together. Although it is possible to enable extended fabric operation one link of a trunked pair, this does not offer any performance advantages and adds complexity to the configuration setup. Therefore, do not use mixed mode operations.

Feature	Description
Extended fabric	Extended fabric operation allocates extra buffer credits to a port. This is important over long links that are usually found in intercluster Metro Mirror operation. Because of the time that it takes for a frame to traverse the link, it is possible to have more frames in transmission at any instant in time than is possible over a short link. The additional buffering is required to allow for the extra frames. For example, the default license for the IBM 2109-F16 switch has two extended fabric options: Normal and Extended Normal. • The Normal option is suitable for short links. • The Extended Normal option provides significantly better performance for the links up to 10 km long. **Note:** With the additional extended fabric license, the user has two extra options: Medium, up to 10 - 50 km and Long, 50 - 100 km. Do not use Medium and Long settings in the intercluster Metro Mirror links that are currently supported.

Cluster configuration using SAN fabrics with long distance fibre links

A SAN Volume Controller cluster using SAN fabric switches can connect to application hosts, storage controllers, or other SAN Volume Controller clusters, through the use of short or long wave optical fibre-channel connections.

The maximum distance between the cluster and host or the cluster and the storage controller is 300 m for short wave and 10 km for long wave optical connections. Longer distances are supported between clusters that use intercluster Metro Mirror or Global Mirror.

Follow these guidelines when you use long wave optical fibre-channel connections:

- For disaster recovery, each cluster must be regarded as a single entity. This includes the back-end storage that provides the quorum disks for the cluster. Therefore, the cluster and quorum disks must be co-located. Do not place components of a single cluster in different physical locations.
- All nodes within a cluster must be located in the same set of racks. There can be a large optical distance between the nodes in the same cluster; however, the nodes must be physically co-located to permit effective service and maintenance.
- All nodes in a cluster must be on the same IP subnet. This allows the nodes to assume the same cluster or service IP address.
- A node must be on the same rack as the uninterruptible power supply from which it receives power.

Note: Do not split cluster operation over a long optical distance, otherwise you will only be able to use asymmetric disaster recovery and it will have substantially reduced performance. Instead, use two cluster configurations for all production disaster recovery systems.

Performance of fibre-channel extenders

When you are planning to use fibre-channel extenders, be aware that the performance of the link to the remote location decreases as the distance to the remote location increases.

For fibre-channel IP extenders, throughput is limited by latency and bit error rates. Typical I/O latency can be expected to be 10 microseconds per kilometer. Bit error rates vary depending on the quality of the circuit that is provided.

You must review the total throughput rates that might be expected for your planned configuration with the vendor of your fibre-channel extender and your network provider.

Clusters

All of your configuration and service tasks are performed at the cluster 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 two nodes, with a maximum of eight nodes. Therefore, you can assign up to eight SAN Volume Controller nodes to one cluster.

All configurations are replicated across all nodes in the cluster, however, only some service actions can be performed at the node level. Because configuration is performed at the cluster level, an IP address is assigned to the cluster instead of each node.

Cluster state

The state of the cluster holds all of the configuration and internal data.

The cluster state information is held in nonvolatile memory. If the mainline power fails, the uninterruptible power supply units maintain the internal power long enough for the cluster state information to be stored on the internal SCSI disk drive of each node. The read and write cache information, which is also held in memory, is stored on the internal SCSI disk drives of the nodes in the input/output (I/O) group that are using that information. Similarly, if the power fails to a node, configuration and cache data for that node is lost and the partner node attempts to flush the cache. The cluster state is still maintained by the other nodes on the cluster.

Figure 2 on page 40 shows an example of a cluster that contains four nodes. The cluster state shown in the grey box does not actually exist, instead each node holds a copy of the entire cluster state.

The cluster contains a single node that is elected as the configuration node. The configuration node can be thought of as the node that controls the updating of cluster state. For example, a user request is made (*1*), that results in a change being made to the configuration. The configuration node controls updates to the cluster (*2*). The configuration node then forwards the change to all nodes (including Node 1), and they all make the state-change at the same point in time (*3*). Using this state-driven model of clustering ensures that all nodes in the cluster know the exact cluster state at any one time.

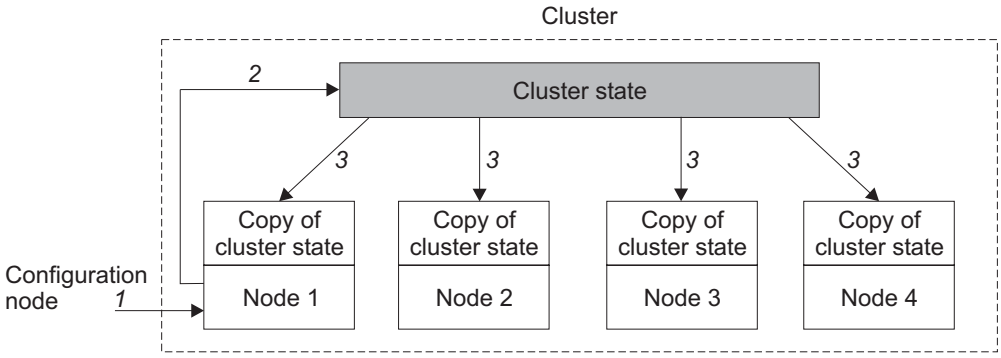


Figure 2. Cluster, nodes, and cluster state.

Each node in the cluster maintains an identical copy of the cluster state. When a change is made to the configuration or internal cluster data, then the same change is applied to all nodes. For example, a user configuration request is made to the configuration node. The configuration node forwards the request to all nodes in the cluster and they all make the change to the cluster state at the same point in time. This ensures that all nodes are aware of the configuration change. If the configuration node fails, then the cluster can elect a new node to take over its responsibilities.

Cluster operation and quorum disks

The cluster must contain at least half of its nodes to function.

Nodes are deployed in pairs known as input/output (I/O) groups, and one to four I/O groups comprise a cluster. In order to function, one node in each I/O group must be operational. If both of the nodes in an I/O group are not operational, access is lost to the virtual disks (VDisks) that are managed by the I/O group.

Note: The cluster can continue to run without loss of access to data as long as one node from each I/O group is available.

A tie-break situation can occur if exactly half the nodes in a cluster fail at the same time, or if the cluster is divided so that exactly half the nodes in the cluster cannot communicate with the other half. For example, in a cluster of four nodes, if any two nodes fail at the same time or any two cannot communicate with the other two, a tie-break exists.

The cluster automatically chooses three managed disks (MDisks) to be candidate *quorum disks* and assigns them quorum indexes of 0, 1, and 2. One of these disks is used to settle a tie-break condition.

If a tie-break occurs, the first half of the cluster to access the quorum disk after the split has occurred locks the disk and continues to operate. The other side stops. This action prevents both sides from becoming inconsistent with each other.

You can change the assignment of quorum disk at any time by issuing the following command:

```
svctask setquorum
```

I/O groups and UPS

Each pair of nodes is known as an *input/output (I/O)* group.

Each node can only be in one I/O group. The I/O groups are connected to the SAN so that all back-end storage and all application servers are visible to all of the I/O groups. Each pair of nodes has the responsibility to serve I/O operations on a particular virtual disk (VDisk).

VDisks are logical disks that are presented to the SAN by SAN Volume Controller nodes. VDisks 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 (UPS) to provide data integrity in the event of a cluster wide power failure.

When an application server performs I/O to a VDisk, it can access the VDisk with either of the nodes in the I/O group. A VDisk can have a preferred node specified when the VDisk is created. Otherwise, if the preferred node is not specified, after the VDisk is created, a preferred node will be assigned. The preferred node is the node through which a VDisk can be accessed.

Each I/O group only has two nodes. The distributed cache inside the SAN Volume Controller is replicated across both nodes in the I/O group. When I/O is performed to a VDisk, the node that processes the I/O duplicates the data onto the partner node that is in the I/O group. I/O traffic for a particular VDisk is, at any one time, managed exclusively by the nodes in a single I/O group. Thus, although a cluster can have many nodes within it, the nodes handle I/O in independent pairs. This means that the I/O capability of the SAN Volume Controller scales well, since additional throughput can be obtained by adding additional I/O groups.

Figure 3 on page 42 shows an example I/O group. A write operation from a host is shown (item *1*), that is targeted for VDisk A. This write is targeted at the preferred node, Node 1 (item *2*). The write is cached and a copy of the data is made in the partner node, Node 2's cache (item *3*). The write is now complete so far as the host is concerned. At some time later the data is written, or de-staged, to storage (item *4*). The figure also shows two UPS units (1 and 2) correctly configured so that each node is in a different power domain.

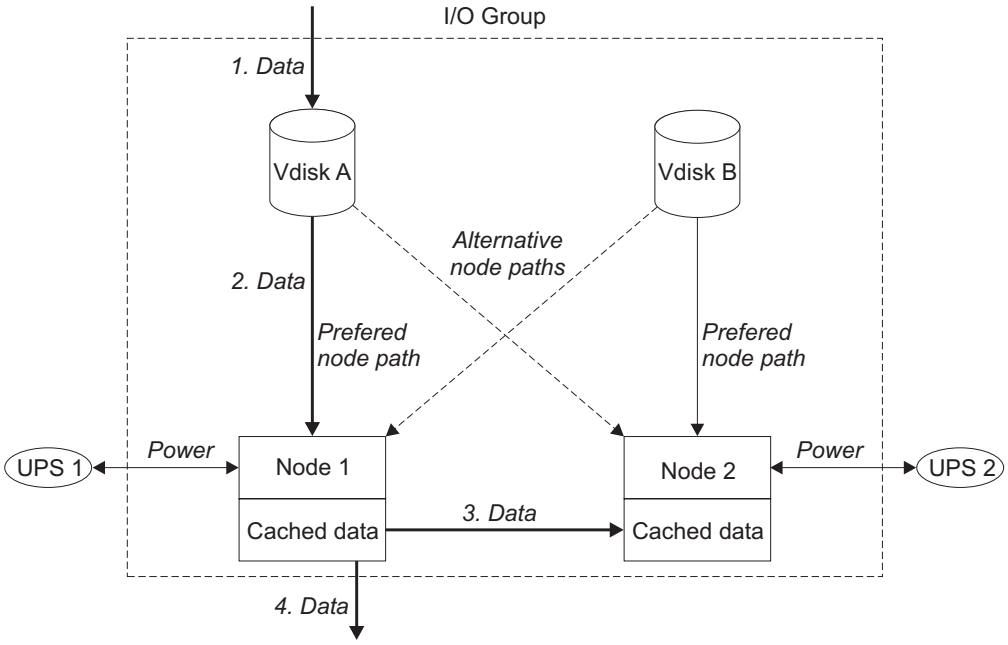


Figure 3. I/O group and UPS

When a node fails within an I/O group, the other node in the I/O group takes over the I/O responsibilities of the failed node. Data loss during a node failure is prevented by mirroring the I/O read/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 goes into write-through mode. Therefore, any writes for the VDisk that are assigned to this I/O group are not cached, but are sent directly to the storage device. If both nodes in an I/O group go offline, the VDisk that are assigned to the I/O group cannot be accessed.

When a VDisk is created, you must specify the I/O group that will provide access to the VDisk. However, VDisks 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 recovery I/O group, which is used when both nodes in the I/O group have multiple failures. This allows you to move the VDisks to the recovery I/O group and then into a working I/O group. I/O access is not possible when VDisks are assigned to the recovery I/O group.

Disk controllers

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

Disk controllers provide the storage that the cluster detects as managed disks (MDisks).

When configuring disk controllers, ensure that you configure and manage the disk controllers and its devices for optimal performance.

The supported RAID controllers are detected by the cluster and reported by the user interfaces. The cluster can determine which MDisks each controller has and can also provide a view of MDisks that is filtered by controller. This view enables you to associate the MDisks with the RAID that the controller presents.

Note: The SAN Volume Controller supports RAID controllers, but it is possible to configure a controller as a non-RAID controller. RAID controllers provide redundancy at the disk level. Therefore, a single physical disk failure does not cause an MDisk failure, an MDisk group failure, or a failure in the virtual disks (VDisks) that were created from the MDisk group.

The controller 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 controller. The controller will surface these disks with a unique ID, the controller LUN or LU number. This ID, along with the controller serial number or numbers (there can be more than one controller), can be used to associate the MDisks in the cluster with the RAID presented by the controllers.

To minimize data loss, virtualize only those RAID that provide some form of redundancy, that is, RAID 1, RAID 10, RAID 0+1 or RAID 5. Do not use RAID 0 because a single physical disk failure might cause the failure of many VDisks.

Unsupported disk controller systems (generic controllers)

When a disk controller system is detected on the SAN, the SAN Volume Controller attempts to recognize it using its inquiry data. If the disk controller system is recognized as one of the explicitly supported storage models, then the SAN Volume Controller uses error recovery programs that can be tailored to the known needs of the disk controller system. If the storage controller is not recognized, then the SAN Volume Controller configures the disk controller system as a generic controller. A generic controller might not function correctly when it is addressed by a SAN Volume Controller. The SAN Volume Controller does not regard accessing a generic controller as an error condition and, consequently, does not log an error. MDisks that are presented by generic controllers are not eligible to be used as quorum disks.

Data migration

Data migration affects the mapping of the extents for a virtual disk (VDisk) to the extents for a managed disk (MDisk).

The host can access the VDisk during the data migration process.

Applications for data migration

The following lists the different applications for data migration:

- Redistribution of workload within a cluster across MDisks. You can redistribute the workload, by one of the following ways:
  - Moving workload onto newly installed storage
  - Moving workload from old or failing storage, prior to replacing it
  - Moving workload to rebalance workload that has changed
- Migrating data from legacy disks to disks that are managed by the SAN Volume Controller.

Image mode virtual disk migration

Image mode virtual disks (VDisks) have a special property that allows the last extent in the VDisk to be a partial extent.

You can migrate striped and sequential Vdisks to an image mode VDisk.

Cluster configuration backup functions

The SAN Volume Controller includes functions that help you to backup cluster configuration settings and business data.

To enable routine maintenance of the SAN Volume Controller clusters, the configuration settings for each cluster are stored on each node. If power fails on a cluster or if a node in a cluster is replaced, the cluster configuration settings are automatically restored when the repaired node is added to the cluster. To restore the cluster configuration in the event of a disaster (if all nodes in a cluster are lost simultaneously), plan to backup the cluster configuration settings to tertiary storage. You can use the configuration backup functions to backup the cluster configuration.

For complete disaster recovery, regularly backup the business data that is stored on virtual disks at the application server level or the host level.

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 backup your application data using the appropriate backup methods.

Objects included in the backup

Configuration data is information about a cluster and the objects that are defined in it. Information about the following objects is included in the cluster configuration data:

- Storage subsystem
- Hosts
- Input/output (I/O) groups
- Managed disks (MDisks)
- MDisk groups
- Nodes
- Virtual disks (VDisks)
- VDisk-to-host mappings
- SSH keys
- FlashCopy mappings
- FlashCopy consistency groups
- Metro Mirror relationships
- Global Mirror relationships
- Metro Mirror consistency groups

- Global Mirror consistency groups

Chapter 5. Planning for configuring the SAN Volume Controller

Ensure that you perform all the required and necessary planning tasks before you start to configure your SAN Volume Controller environment.

Planning the clusters

Determine the following information for clusters:

- The number of clusters and the number of node pairs (I/O groups). Each pair of nodes is the container for one or more virtual disks (VDisks).
- The number of hosts that you want to use.
- The number of I/Os per second between the hosts and nodes.

Planning the hosts

LUN masking allows the hosts to access specific logical units (LUs) within the disk controllers. Determine the following information for hosts:

- The worldwide port names (WWPNs) of the fibre-channel host bus adapter (HBA) ports in the hosts.
- The names to assign to the hosts.
- The VDisks to assign to the hosts.

Planning the MDisks

To plan the managed disks (MDisks), determine the logical or physical disks (logical units) in the backend storage.

Planning the MDisk groups

Determine the following information for MDisk groups:

- The types of backend controllers you want to use.
- If you want to create VDisks with the sequential policy, plan to create a separate MDisk group for these VDisks or ensure that you create these VDisks before creating VDisks with the striped policy.
- Plan to create MDisk groups for the backend controllers that provide the same level of performance or reliability, or both. For example, you can group all of the managed disks that are RAID 10 in one MDisk group and all of the MDisks that are RAID 5 in another group.

Planning the VDisks

An individual VDisk is a member of one MDisk group and one I/O group. The MDisk group defines which MDisks provide the backend storage that makes up the VDisk. The I/O group defines which nodes provide I/O access to the VDisk. Determine the following information before creating a VDisk:

- If there is data on the volume that needs to be preserved.
- The name you want to assign to the VDisk.
- The I/O group to which the VDisk will be assigned.

© Copyright IBM Corp. 2003, 2007

- The MDisk group to which the VDisk will be assigned.
- The capacity of the VDisk.

Maximum configuration

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:

http://www.ibm.com/storage/support/2145

SAN fabric overview

The SAN fabric is an area of the network that contains routers, gateways, hubs, and switches. A single cluster SAN contains two distinct types of zones: a host zone and a disk zone.

In the host zone, the host systems can identify and address the SAN Volume Controller nodes. You can have more than one host zone. Generally, you create one host zone for each host type. In the disk zone, the SAN Volume Controller nodes identify the disk drives. Host systems cannot operate on the disk drives directly; all data transfer occurs through the SAN Volume Controller nodes. Figure 4 shows several host systems that are connected in a SAN fabric.

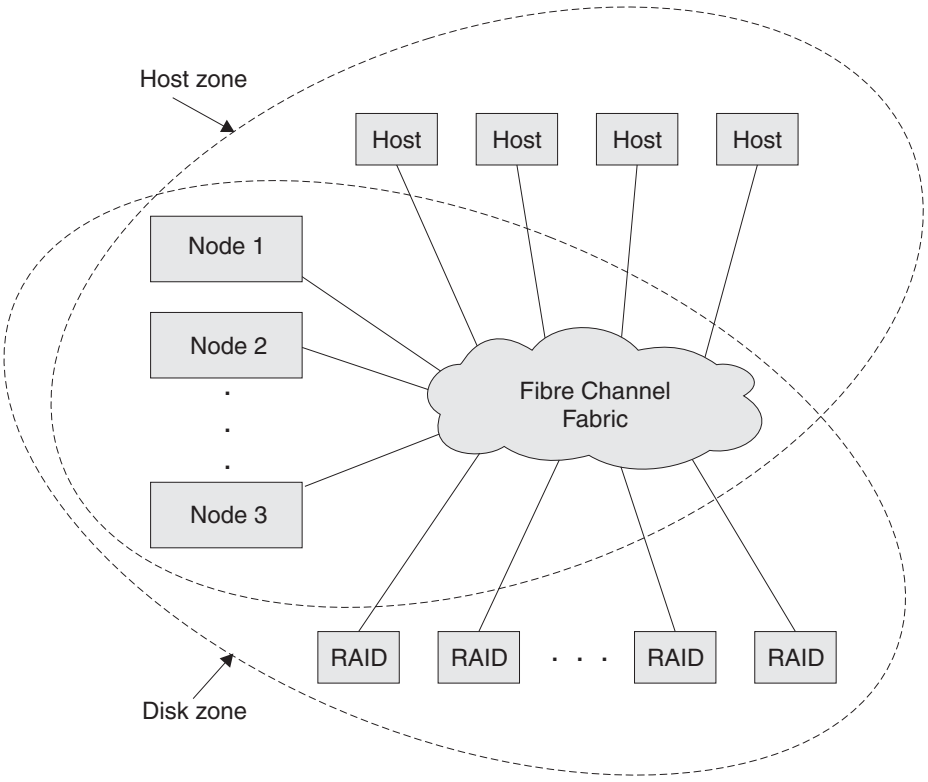


Figure 4. Example of a SAN Volume Controller in a fabric

A cluster of SAN Volume Controller nodes is connected to the same fabric and presents virtual disks (VDisks) to the host systems. You create these VDisks from units of space within a managed disk (MDisk) group. An MDisk group is a

collection of MDisks that are presented by the storage subsystems (RAID controllers). The MDisk group provides a storage pool. You specify how each group is created, and you can combine MDisks from different manufacturers' controllers in the same MDisk group.

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 AIX operating system and another host that runs on a Windows operating system.

You can remove one SAN Volume Controller node in each I/O group from a cluster when hardware service or maintenance is required. After you remove the SAN Volume Controller node, you can replace the field replaceable units (FRUs) in the SAN Volume Controller node. All communication between disk drives and 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 that is connected to the Ethernet network can access this VPD.

Cluster configuration information is stored on every SAN Volume Controller node that is in the cluster to allow concurrent replacement of FRUs. When a new FRU is installed and when the SAN Volume Controller node is added back into the cluster, configuration information that is required by that SAN Volume Controller node is read from other SAN Volume Controller nodes in the cluster.

Configuring the SAN fabric

Ensure that you understand the rules and requirements when you are configuring the SAN fabric.

Table 4 provides terms and definitions that can guide your understanding of the rules and requirements.

Table 4. Configuration terms and definitions

Term	Definition
ISL hop	A hop on an interswitch link (ISL). With reference to all pairs of N-ports or end-nodes that are in a fabric, the number of ISL hops is the number of links that are crossed on the shortest route between the node pair whose nodes are farthest apart from each other. The distance is measured only in terms of the ISL links that are in the fabric.

Table 4. Configuration terms and definitions (continued)

Term	Definition
Oversubscription	The ratio of the sum of the traffic that is on the initiator N-node connections to the traffic that is on the most heavily-loaded ISLs or where more than one ISL is in parallel between these switches. This definition assumes a symmetrical network and a specific workload that is applied equally from all initiators and sent equally to all targets. A symmetrical network means that all initiators are connected at the same level and all the controllers are connected at the same level. **Note:** The SAN Volume Controller puts its back-end traffic onto the same symmetrical network. The back-end traffic can vary by workload. Therefore, the oversubscription that a 100% read hit gives is different from the oversubscription that 100% write-miss gives. If you have an oversubscription of 1 or less, the network is nonblocking.
Virtual SAN (VSAN)	A VSAN is a virtual storage area network (SAN).
Redundant SAN	A SAN configuration in which if any one component fails, connectivity between the devices that are in the SAN is maintained, possibly with degraded performance. Create a redundant SAN by splitting the SAN into two independent counterpart SANs.
Counterpart SAN	A non-redundant portion of a redundant SAN. A counterpart SAN provides all the connectivity of the redundant SAN, but without the redundancy. The SAN Volume Controller is typically connected to a redundant SAN that is made out of two counterpart SANs.
Local fabric	The fabric that consists of those SAN components (switches and cables) that connect the components (nodes, hosts, and switches) of the local cluster. Because the SAN Volume Controller supports Metro and Global Mirror, significant distances might exist between the components of the local cluster and those of the remote cluster.
Remote fabric	The fabric that consists of those SAN components (switches and cables) that connect the components (nodes, hosts, and switches) of the remote cluster. Because the SAN Volume Controller supports Metro Mirror and Global Mirror, significant distances might exist between the components of the local cluster and those of the remote cluster.
Local/remote fabric interconnect	The SAN components that connect the local fabrics to the remote fabrics. There might be significant distances between the components in the local cluster and those in the remote cluster. These components might be single-mode optical fibres that are driven by Gigabit Interface Converters (GBICs), or they might be other, more advanced components, such as channel extenders
SAN Volume Controller fibre-channel port fan in	The number of hosts that can see any one port. Some controllers recommend that the number of hosts using each port be limited to prevent excessive queuing at that port. If the port fails or the path to that port fails, the host might failover to another port, and the fan in requirements might be exceeded in this degraded mode.
Invalid configuration	In an invalid configuration, an attempted operation fails and will generate an error code to indicate what caused it to become invalid.
Unsupported configuration	A configuration that might operate successfully, but for which IBM does not guarantee the solution for problems that might occur. Usually this type of configuration does not create an error log entry.
Valid configuration	A configuration that is neither invalid nor unsupported.

Table 4. Configuration terms and definitions (continued)

Term	Definition
Degraded	A valid configuration that has had a failure, but continues to be neither invalid nor unsupported. Typically, a repair action is required to restore the degraded configuration to a valid configuration.
Fibre channel extender	A device for long distance communication connecting other SAN fabric components. Generally these might involve protocol conversion to ATM, IP or some other long distance communication protocol.
Mesh configuration	A network that contains a number of small SAN switches configured to create a larger switched network. With this configuration, four or more switches are connected together in a loop with some of the paths short circuiting the loop. An example of this configuration is to have four switches connected together in a loop with ISLs for one of the diagonals.

Configuration rules

Storage area network (SAN) configurations that contain SAN Volume Controller nodes can be configured in various ways.

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

- Storage subsystems
- HBAs
- Nodes
- Fibre-channel switches
- Fabrics
- Port Switches
- Zoning
- Power requirements

Storage subsystems

Follow these rules when you are planning the configuration of storage subsystems in the SAN fabric.

All SAN Volume Controller nodes in a cluster must be able to see the same set of storage subsystem ports on each device. Any operation that is in this mode in which two nodes do not see the same set of ports on the same device is degraded, and the system logs errors that request a repair action. This rule can have important effects on a storage subsystem 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 configuration in which SAN Volume Controller nodes bridge a separate host device and a RAID is not supported. See the following Web site for the latest support information:

http://www.ibm.com/storage/support/2145

The SAN Volume Controller clusters must not share its storage subsystem logical units (LUs) with hosts. A storage subsystem can be shared with a host under certain conditions as described in this topic.

You can configure certain storage controllers to safely share resources between the SAN Volume Controller 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 cannot access logical units (LUs) that a host or another SAN Volume Controller 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 and a host, the SAN Volume Controller 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 cannot access LUs that a host or another SAN Volume Controller can also access. This can be arranged by controller LUN mapping and masking. If this is not guaranteed, data corruption can occur. Do not use this configuration because of the risk of data corruption.

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

The SAN Volume Controller must be configured to manage only LUNs that are presented by supported disk controller systems. Operation with other devices is not supported.

Unsupported storage subsystem (generic device)

When a storage subsystem is detected on the SAN, the SAN Volume Controller attempts to recognize it using its Inquiry data. If the device is recognized as one of the explicitly supported storage models, the SAN Volume Controller uses error recovery programs that are potentially tailored to the known needs of the storage subsystem. If the device is not recognized, 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. In any event, the SAN Volume Controller does not regard accessing a generic device as an error condition and, consequently, does not log an error. Managed disks (MDisks) that are presented by generic devices are not eligible to be used as quorum disks.

Split controller configurations

The SAN Volume Controller is configured to manage LUs that are exported only by RAID controllers. Operation with other RAID controllers is illegal. While it is possible to use the SAN Volume Controller to manage JBOD (just a bunch of disks) LUs that are presented by supported RAID controllers, the SAN Volume Controller 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 controller 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 SAN Volume Controller or a directly attached host. Suitable LUN

masking must be in place to ensure that LUs are not shared between SAN Volume Controller nodes and direct attached hosts.

In a split controller configuration, a RAID presents some of its LUs to a SAN Volume Controller (which treats the LU as an MDisk) and the remaining LUs to another host. The SAN Volume Controller presents virtual disks (VDisks) 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 5 shows that the RAID controller is an IBM DS4000, 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 and directly by the device.

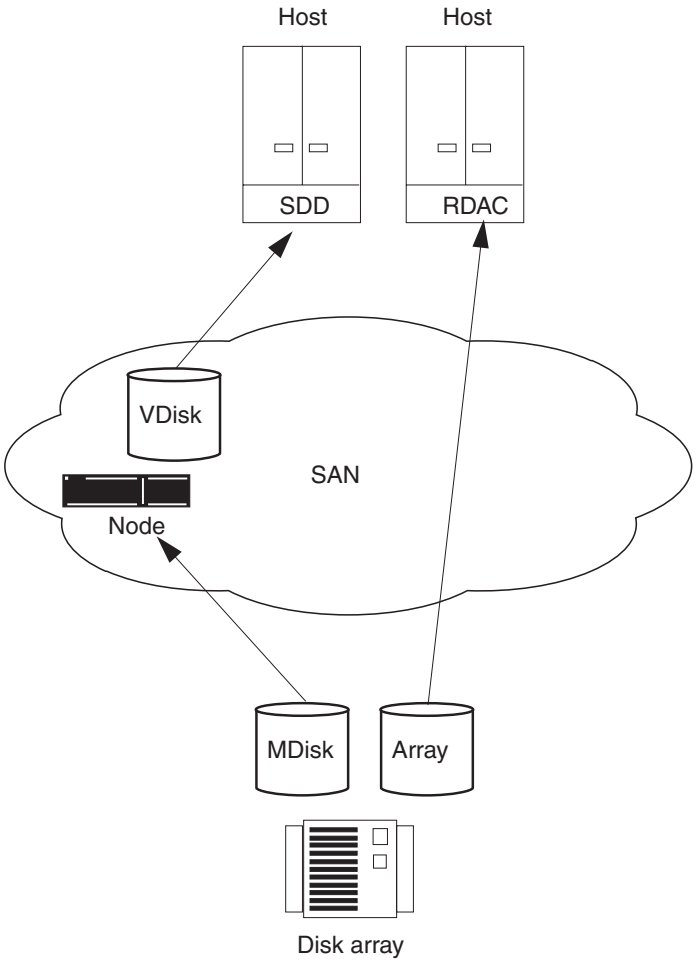


Figure 5. Disk controller system shared between SAN Volume Controller and a host

It is also possible to split a host so that it accesses some of its LUNs through the SAN Volume Controller and some directly. In this case, the multipathing software that is used by the controller must be compatible with the SAN Volume Controller nodes multipathing software. Figure 6 on page 54 is a supported configuration because the same multipathing driver is used for both direct and VDisks.

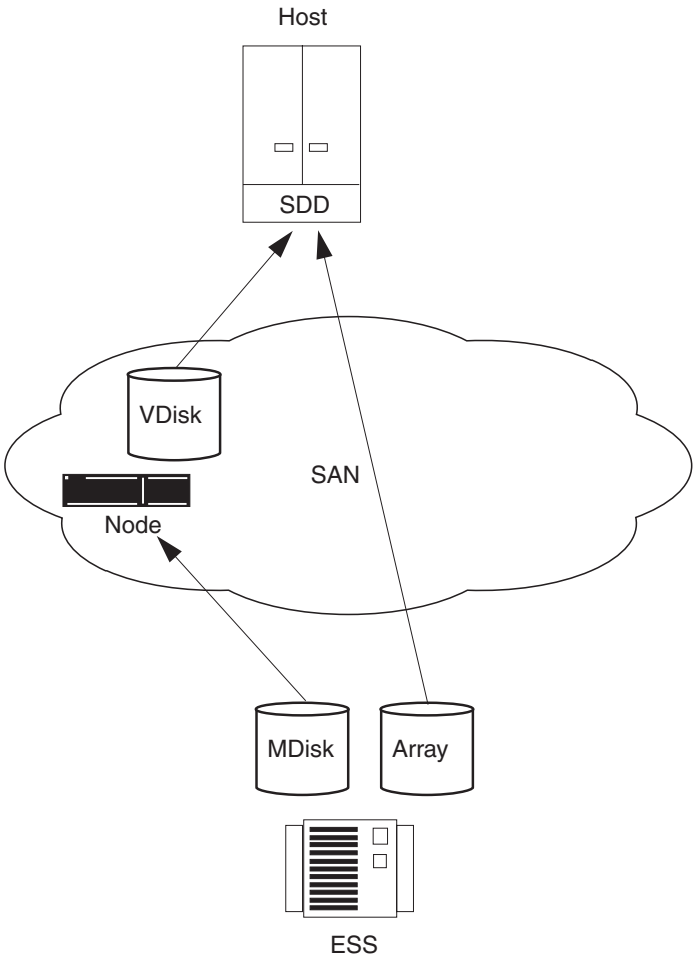


Figure 6. IBM ESS LUs accessed directly with a SAN Volume Controller

In the case where the RAID controller uses multipathing software that is compatible with SAN Volume Controller nodes multipathing software (see Figure 7 on page 55), 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 is one example.

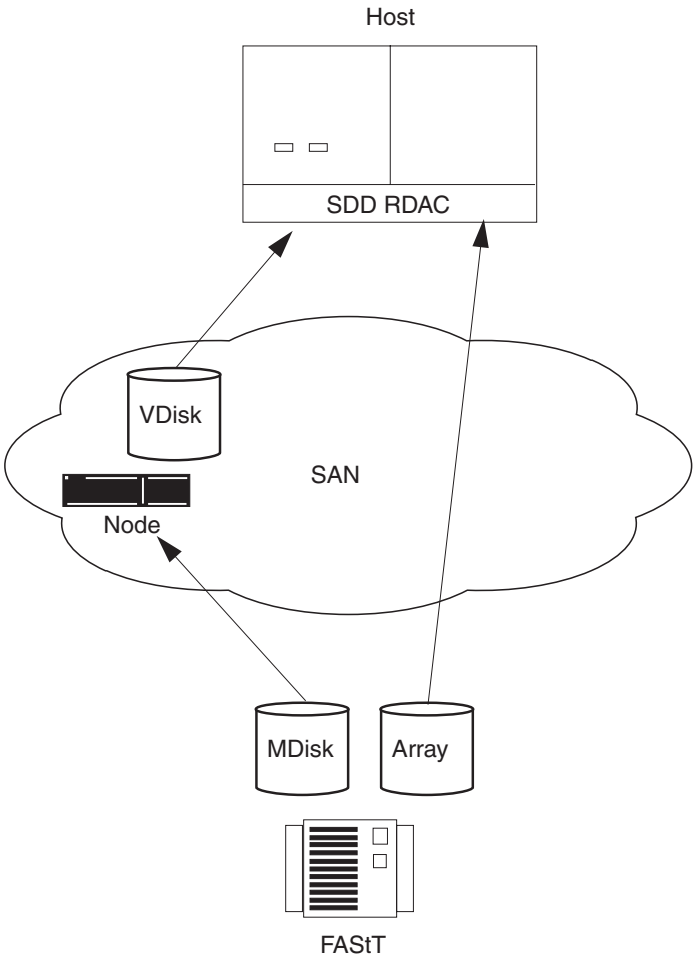


Figure 7. IBM DS4000 direct connection with a SAN Volume Controller on one host

Host HBAs

Ensure that you are familiar with the configuration rules for host bus adapters (HBAs). You must abide by the configuration rules for HBAs to ensure that you have a valid configuration.

The SAN Volume Controller must be configured to export virtual disks (VDisks) only to host fibre-channel ports that are on the supported HBAs. See the following Web site for specific firmware levels and the latest supported hardware:

http://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.

Nodes

You must follow the configuration rules for SAN Volume Controller nodes to ensure that you have a valid configuration.

Node HBAs

SAN Volume Controller 2145-4F2 and SAN Volume Controller 2145-8F2 nodes contain two 2-port 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 2145-8F4 and SAN Volume Controller 2145-8G4 nodes contain one 4-port HBA.

I/O groups

Nodes must always be used in pairs called I/O groups. SAN Volume Controller 2145-4F2, SAN Volume Controller 2145-8F2, SAN Volume Controller 2145-8F4, and SAN Volume Controller 2145-8G4 nodes can be in the same I/O group. If a node fails or is removed from the configuration, the remaining node in the I/O group operates in a degraded mode, but the configuration is still valid.

VDisks

Each node presents a virtual disk (VDisk) to the SAN through four ports. Each VDisk is accessible from the two nodes in an I/O group. Each host HBA port can recognize up to eight paths to each logical unit (LU) that is presented by the node. The hosts must run a multipathing device driver before the multiple paths can resolve to a single device.

Optical connections

Valid optical connections are based on the fabric rules that the manufacturers impose for the following connection methods:

- Host to a switch
- Backend to a switch
- Interswitch links (ISLs)

Short wave optical fibre connections must be used between a node and its switches. Clusters that use intercluster Metro Mirror or Global Mirror can use short or long wave optical fibre connections, or distance-extender technology that is supported by the switch manufacturer.

The number of paths through the network from the node 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 number of paths to a VDisk is eight × the number of host ports.

Ethernet connection

To ensure cluster failover operations, all nodes in a cluster must be connected to the same IP subnet.

Port speed

The optical fibre-connections between Fibre Channel switches and all SAN Volume Controller 2145-4F2 and SAN Volume Controller 2145-8F2 nodes must run at one port speed. The Fibre Channel ports on SAN Volume Controller 2145-8F4 and SAN Volume Controller 2145-8G4 nodes auto-negotiate the operational port speed independently, which allows these nodes to operate at different speeds.

Fibre Channel switches

Ensure that you are familiar with the configuration rules for Fibre Channel switches. You must follow the configuration rules for Fibre Channel switches to ensure that you have a valid configuration.

The SAN must contain only supported switches.

See the following Web site for specific firmware levels and the latest supported hardware:

http://www.ibm.com/storage/support/2145

The SAN should consist of at least two independent switches (or networks of switches) so that the SAN includes a redundant fabric, and has no single point of failure. If one SAN fabric fails, the configuration is in a degraded mode, but it is still valid. If the SAN contains only one fabric, it is still a valid configuration, but a failure of the fabric might cause a loss of access to data. Therefore, SANs with one fabric are considered to have a possible single point of failure.

Configurations with more than four SANs are not supported.

The SAN Volume Controller nodes must always and only be connected to SAN switches. Each node must be connected to each of the counterpart SANs that are in the redundant fabric. Any configuration that uses direct connections between the host and node or between the controller and node is not supported.

All backend storage must always and only be connected to SAN switches. Multiple connections are permitted from the redundant controllers of the backend storage to improve data bandwidth performance. It is not necessary to have a connection between each redundant disk controller system of the backend storage and each counterpart SAN. For example, in an IBM System Storage DS4000 configuration in which the IBM DS4000 contains two redundant controllers, only two controller minihubs are usually used. Controller A of the IBM DS4000 is connected to counterpart SAN A, and controller B of the IBM DS4000 is connected to counterpart SAN B. Any configuration that uses a direct connection between the host and the controller 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 controllers, leaving the host ports connected to the edge switches.

The switch configuration of a SAN Volume Controller SAN must observe the switch manufacturer's configuration rules. These rules might put restrictions on the switch configuration. Any configuration that runs outside the manufacturer's configuration rules is not supported.

Operational port speed

You can change the operational port speed for SAN Volume Controller 2145-4F2 and SAN Volume Controller 2145-8F2 nodes to 1 Gbps or 2 Gbps. However, the optical-fibre connections between the Fibre Channel switches and all SAN Volume Controller 2145-4F2 and SAN Volume Controller 2145-8F2 nodes in a cluster must run at the same speed. The Fibre Channel ports on SAN Volume Controller 2145-8F4 and SAN Volume Controller 2145-8G4 nodes auto negotiate the operational port speed independently, which allows these nodes to operate at different speeds. SAN Volume Controller 2145-8F4 and SAN Volume Controller 2145-8G4 nodes can operate at 1 Gbps, 2 Gbps or 4 Gbps. If these nodes are connected to a 4 Gbps capable switch, the port attempts to operate at 4 Gbps; however, if there is a large number of link error rates, the adapter negotiates a lower speed.

Mixing manufacturer switches in a single SAN fabric

Within an individual SAN fabric, switches must have the same manufacturer, with the following exceptions:

- BladeCenter®. See the documentation that is provided with your BladeCenter for more information.
- Where one pair of counterpart fabrics (for example, Fabric A and Fabric B) provide a redundant SAN, different manufacturer's switches can be mixed in a SAN Volume Controller configuration, provided that each fabric contains only switches from a single manufacturer. Thus, the two counterpart SANs can have different manufacturer's switches.
- The SAN Volume Controller supports the Interoperability Modes of the Cisco MDS 9000 family of switch and director products with the following restrictions:
  - The Cisco MDS 9000 must be connected to Brocade and McData switch/director products with the multivendor fabric zones connected using MDS Interoperability Mode 1, 2 or 3.
  - All of the SAN Volume Controller nodes that are in the SAN Volume Controller cluster must be attached to the Cisco part of the counterpart fabric or they must be attached to the McData or Brocade part of the counterpart fabric to avoid having a single fabric with a SAN Volume Controller cluster that has part of the SAN Volume Controller nodes connected to Cisco switch ports and part of the SAN Volume Controller nodes connected to Brocade or McData switch ports.

Brocade core-edge fabrics

See the following Web site for the latest support and information for Brocade fabrics:

http://www.ibm.com/storage/support/2145

SAN Volume Controller software level 4.1.1 or higher

Brocade core-edge fabrics that use the M14, M48, or B64 models can have up to 1024 hosts under the following conditions:

- M14, M48, B64 or other Brocade models can be used as edge switches; however, the SAN Volume Controller ports and backend storage must all be connected to the M14, M48, or B64 core switch.
- The M48 and B64 models must be running the firmware level 5.1.0c or higher.

- The M14 models must be running at the firmware level 5.0.5a or higher.

SAN Volume Controller software level 4.1.0 or lower

Brocade core-edge fabrics that use the M14 or M48 model can have up to 256 hosts under the following conditions:

- Each SAN Volume Controller port cannot see more than 256 node port logins.
- Each I/O group cannot be associated with more than 64 hosts.
- A host can be associated with more than one I/O group.
- Each HBA port must be in a separate zone and each zone must contain one port from each SAN Volume Controller node in the I/O group that the host accesses.
- M14, M48 or other Brocade models can be used as edge switches; however, the SAN Volume Controller ports and backend storage must all be connected to the M14 or M48 core switch.
- You can attach between one and four separate fabrics to the SAN Volume Controller cluster. If other manufacturer fabrics are also attached to the SAN Volume Controller cluster, you must follow the SAN Volume Controller support guidelines for that manufacturer.
- A host can access VDisks from different I/O groups in a Brocade SAN, but this reduces the maximum number of hosts that can be used in the SAN. For example, if the same host uses VDisks in two different I/O groups, this consumes one of the 64 hosts in each I/O group. If each host accesses VDisks in each I/O group, there can only be 64 hosts in the configuration. Alternatively, if each host accesses VDisks in two I/O groups, 32 different hosts can be attached to each I/O group. This means that 128 hosts can be used in an 8 node cluster.

Fibre-channel switches and interswitch links

The local or remote fabric must not contain more than three interswitch link (ISL) hops in each fabric. Any configuration that uses more than three ISL hops is not supported. When a local fabric is connected to a remote fabric for Metro Mirror or Global Mirror, the ISL hop count between a local node and a remote node must not exceed seven. Therefore, some ISL hops can be used in a cascaded switch link between local and remote clusters if the internal ISL count of the local or remote cluster is less than three.

If all three allowed ISL hops are used within the local and remote fabrics, the local remote fabric interconnect must be a single ISL hop between a switch in the local fabric and a switch in the remote fabric.

The SAN Volume Controller supports the use of distance-extender technology, including DWDM (Dense Wavelength Division Multiplexing) and FCPIP extenders, to increase the overall distance between local and remote clusters. If this extender technology involves a protocol conversion, the local and remote fabrics should be regarded as independent fabrics, limited to three ISL hops each. The only restriction on the interconnection between the two fabrics is the maximum latency that is allowed in the distance extender technology.

Note: Where multiple ISL hops are used between switches, follow the fabric manufacturer's recommendations for trunking.

With ISLs between nodes in the same cluster, the ISLs are considered a single point of failure. This is illustrated in Figure 8 on page 60.

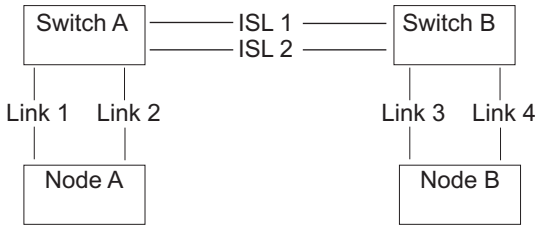


Figure 8. Fabric with ISL between nodes in a cluster

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 9.

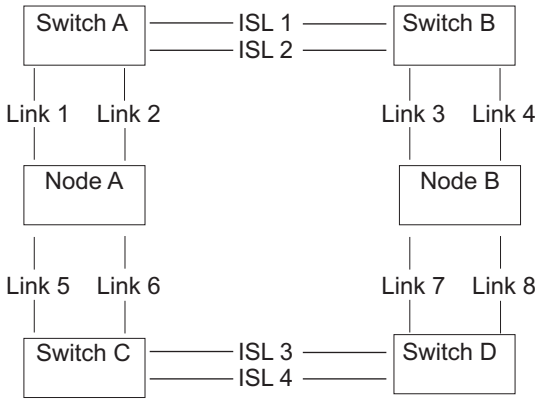


Figure 9. Fabric with ISL in a redundant configuration

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 subsystem traffic across ISLs. 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×(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.

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 SAN Volume Controller Console :

- You can use screen-reader software and a digital speech synthesizer to hear what is displayed on the screen. The following screen readers have been tested: JAWS v4.5 and IBM Home Page Reader v3.0.
- You can operate all features using the keyboard instead of the mouse.

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 SAN Volume Controller Console 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 view the publications for the SAN Volume Controller in Adobe Portable Document Format (PDF) using the Adobe Acrobat Reader. The PDFs are provided at the following Web site:

http://www.ibm.com/storage/support/2145

Related reference

"SAN Volume Controller library and related publications" on page xi
A list of other publications that are related to this product are provided to you for your reference.

© Copyright IBM Corp. 2003, 2007

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785
U.S.A.

For license inquiries regarding double-byte (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:

IBM World Trade Asia Corporation
Licensing
2-31 Roppongi 3-chome, Minato-ku
Tokyo 106-0032, Japan

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law:
INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATIONS "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the 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.

© Copyright IBM Corp. 2003, 2007

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 may 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.

If you are viewing this information softcopy, the photographs and color illustrations may not appear.

Trademarks

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

- AIX
- BladeCenter
- Enterprise Storage Server
- FlashCopy
- IBM
- IBM eServer
- IBM TotalStorage
- IBM System Storage
- System p5
- System z9
- System Storage
- TotalStorage
- xSeries

Intel and Pentium are trademarks of Intel Corporation in the United States, other countries, or both.

Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

Microsoft and Windows 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.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Adobe, Acrobat, Portable Document Format (PDF), and PostScript are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, other countries, or both.

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

Notices and statements in this document

Ensure that you understand the typographic conventions that are used to indicate special notices.

The notices throughout the SAN Volume Controller documentation and in the *IBM Safety Information* document follow specific guidelines for their content.

The following notices are used throughout this library to convey specific meanings:

Note: These notices provide important tips, guidance, or advice.

Attention: These notices indicate possible damage to programs, devices, or data. An attention notice appears before the instruction or the situation in which damage might occur.

CAUTION:
These notices indicate situations that can be potentially hazardous to you. A caution notice precedes the description of a potentially hazardous procedural step or situation.

DANGER

These notices indicate situations that can be potentially lethal or extremely hazardous to you. A danger notice precedes the description of a potentially lethal or extremely hazardous procedural step or situation.

The caution and danger statements that appear in this document are also in the multilingual *IBM Safety Information* document. Each statement is numbered for reference to the corresponding statement in the *IBM Safety Information* document.

Glossary

This glossary includes terms for the IBM System Storage SAN Volume Controller.

This glossary includes selected terms and definitions from A Dictionary of Storage Networking Terminology (http://www.snia.org/education/dictionary), copyrighted 2001 by the Storage Networking Industry Association, 2570 West El Camino Real, Suite 304, Mountain View, California 94040-1313. Definitions derived from this book have the symbol (S) after the definition.

The following cross-references are used in this glossary:

See Refers the reader to one of two kinds of related information:

- A term that is the expanded form of an abbreviation or acronym. This expanded form of the term contains the full definition.
- A synonym or more preferred term.

See also
Refers the reader to one or more related terms.

Contrast with
Refers the reader to a term that has an opposite or substantively different meaning.

A

access mode
One of three different modes in which a logical unit (LU) in a disk controller system can operate. See also *image mode*, *managed space mode*, and *unconfigured mode*.

agent code
An open-systems standard that interprets Common Information Model (CIM) requests and responses as they transfer between the client application and the device.

application server
A host that is attached to the storage area network (SAN) and that runs applications.

array An ordered collection, or group, of physical storage devices that are used to define logical volumes or devices.

association
A class that contains two references that define a relationship between two referenced objects.

asymmetric virtualization
A virtualization technique in which the virtualization engine is outside the data path and performs a metadata-style service. The metadata server contains all the mapping and locking tables while the storage devices contain only data. See also *symmetric virtualization*.

auxiliary virtual disk
The virtual disk that contains a backup copy of the data and that is used in disaster recovery scenarios. See also *master virtual disk*.

© Copyright IBM Corp. 2003, 2007

availability
The ability of a system to continue working, with perhaps a decrease in performance, after individual components fail.

B

bandwidth
The range of frequencies an electronic system can transmit or receive. The greater the bandwidth of a system, the more information the system can transfer in a given period of time.

blade One component in a system that is designed to accept some number of components (blades). Blades could be individual servers that plug into a multiprocessing system or individual port cards that add connectivity to a switch. A blade is typically a hot-swappable hardware device.

block A unit of data storage on a disk drive.

block virtualization
The act of applying virtualization to one or more block-based (storage) services for the purpose of providing a new aggregated, higher-level, richer, simpler, or secure block service to clients. Block virtualization functions can be nested. A disk drive, RAID system, or volume manager all perform some form of block-address to (different) block-address mapping or aggregation. See also *virtualization*.

Boolean
Pertaining to the processes used in the algebra formulated by George Boole.

C

cache A high-speed memory or storage device used to reduce the effective time required to read data from or write data to lower-speed memory or a device. Read cache holds data in anticipation that it will be requested by a client. Write cache holds data written by a client until it can be safely stored on more permanent storage media such as disk or tape.

Call Home
A communication service that links a machine to a service provider. The machine can use this link to place a call to IBM or to another service provider when service is required. With access to the machine, service personnel can perform service tasks, such as viewing error and problem logs or initiating trace and dump retrievals.

cascading
The process of connecting two or more fibre-channel hubs or switches together to increase the number of ports or extend distances.

CIM See *Common Information Model*.

CIM object manager (CIMOM)
The common conceptual framework for data management that receives, validates, and authenticates the CIM requests from the client application. It then directs the requests to the appropriate component or service provider.

CIMOM
See *CIM object manager*.

Cisco command-line interface
An interface that is used to perform functions that are provided on the service panel.

class The definition of an object within a specific hierarchy. A class can have properties and methods and can serve as the target of an association.

CLI See *command line interface*.

client A computer system or process that requests a service of another computer system or process that is typically referred to as a server. Multiple clients can share access to a common server.

client application
A storage management program that initiates Common Information Model (CIM) requests to the CIM agent for the device.

cluster
In SAN Volume Controller, a pair of nodes that provides a single configuration and service interface.

command line-interface (CLI)
A type of computer interface in which the input command is a string of text characters.

Common Information Model (CIM)
A set of standards developed by the Distributed Management Task Force (DMTF). CIM provides a conceptual framework for storage management and an open approach to the design and implementation of storage systems, applications, databases, networks, and devices.

concurrent maintenance
Service that is performed on a unit while it is operational.

configuration node
A node that acts as the focal point for configuration commands and manages the data that describes the cluster configuration.

connected
In a Global Mirror relationship, pertaining to the status condition that occurs when two clusters can communicate.

consistency group
A group of copy relationships between virtual disks that are managed as a single entity.

consistent copy
In a Metro or Global Mirror relationship, a copy of a secondary virtual disk (VDisk) that is identical to the primary VDisk from the viewpoint of a host system, even if a power failure occurred while I/O activity was in progress.

consistent-stopped
In a Global Mirror relationship, the state that occurs when the secondary virtual disk (VDisk) contains a consistent image, but the image might be out-of-date with respect to the primary VDisk. This state can happen when a relationship was in the consistent-synchronized state when an error occurred that forced a freeze of the consistency group. This state can also happen when a relationship is created with the create-consistent flag set to `TRUE`.

consistent-synchronized
In a Global Mirror relationship, the status condition that occurs when the primary virtual disk (VDisk) is accessible for read and write I/O operations. The secondary VDisk is accessible for read-only I/O operations. See also *primary virtual disk* and *secondary virtual disk*.

container

- IBM definition: A visual user-interface component that holds objects.
- HP definition:
  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 linked as a storageset. Stripesets and mirrorsets are examples of storageset containers that the controller uses to create units.

copied
In a FlashCopy mapping, a state that indicates that a copy has been started after the copy relationship was created. The copy process is complete and the target disk has no further dependence on the source disk.

copying
A status condition that describes the state of a pair of virtual disks (VDisks) that have a copy relationship. The copy process has been started but the two virtual disks are not yet synchronized.

Copy Services
The services that enable you to copy virtual disks (VDisks): FlashCopy, Metro, and Global Mirror.

counterpart SAN
A nonredundant portion of a redundant storage area network (SAN). A counterpart SAN provides all the connectivity of the redundant SAN but without the redundancy. Each counterpart SANs provides an alternate path for each SAN-attached device. See also *redundant SAN*.

cross-volume consistency
In SAN Volume Controller, a consistency group property that guarantees consistency between virtual disks when an application issues dependent write operations that span multiple virtual disks.

D

data migration
The movement of data from one physical location to another without disrupting I/O operations.

degraded
Pertaining to a valid configuration that has suffered a failure but continues to be supported and legal. Typically, a repair action can be performed on a degraded configuration to restore it to a valid configuration.

dense wavelength division multiplexing (DWDM)
A technology that places many optical signals onto one single-mode fiber using slightly different optical frequencies. DWDM enables many data streams to be transferred in parallel.

dependent write operations
A set of write operations that must be applied in the correct order to maintain cross-volume consistency.

destage
A write command initiated by the cache to flush data to disk storage.

device

- In the CIM Agent, the storage server that processes and hosts client application requests.
- IBM definition: A piece of equipment that is used with the computer and does not generally interact directly with the system, but is controlled by a controller.
- HP definition: 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 (virtual disks) can be created from devices after the devices have been made known to the controller.

device provider
A device-specific handler that serves as a plug-in for the Common Information Model (CIM); that is, the CIM object manager (CIMOM) uses the handler to interface with the device.

directed maintenance procedures
The set of maintenance procedures that can be run for a cluster. These procedures are run from within the SAN Volume Controller application and are documented in the *IBM System Storage SAN Volume Controller: Service Guide*.

disconnected
In a Metro or Global Mirror relationship, pertains to two clusters when they cannot communicate.

discovery
The automatic detection of a network topology change, for example, new and deleted nodes or links.

disk controller
A device that coordinates and controls the operation of one or more disk drives and synchronizes the operation of the drives with the operation of the system as a whole. Disk controllers provide the storage that the cluster detects as managed disks (MDisks).

disk drive
A disk-based, nonvolatile, storage medium.

disk zone
A zone defined in the storage area network (SAN) fabric in which the SAN Volume Controller can detect and address the logical units that the disk controllers present.

Distributed Management Task Force (DMTF)
An organization that defines standards for the management of distributed systems. See also *Common Information Model*.

DMP See *directed maintenance procedures*.

DMTF
See *Distributed Management Task Force*.

domain name server
In the Internet suite of protocols, a server program that supplies name-to-address conversion by mapping domain names to IP addresses.

DRAM
See *dynamic random access memory*.

DWDM
See *Dense Wavelength Division Multiplexing*.

dynamic random access memory (DRAM)
A storage in which the cells require repetitive application of control signals to retain stored data.

E

EC See *engineering change.*

empty In a Global Mirror relationship, a status condition that exists when the consistency group contains no relationships.

engineering change (EC)
A correction for a defect of hardware or software that is applied to a product.

error code
A value that identifies an error condition.

ESS See *IBM TotalStorage® Enterprise Storage Server®*.

exclude
To remove a managed disk (MDisk) from a cluster because of certain error conditions.

excluded
In SAN Volume Controller, the status of a managed disk that the cluster has removed from use after repeated access errors.

extent A unit of data that manages the mapping of data between managed disks and virtual disks.

F

fabric In fibre-channel technology, a routing structure, such as a switch, that receives addressed information and routes it to the appropriate destination. A fabric can consist of more than one switch. When multiple fibre-channel switches are interconnected, they are described as cascading. See also *cascading*.

fabric port (F_port)
A port that is part of a fibre-channel fabric. An F_port on a fibre-channel fabric connects to the node port (N_port) on a node.

failover
In SAN Volume Controller, the function that occurs when one redundant part of the system takes over the workload of another part of the system that has failed.

FC See *fibre channel*.

fibre channel
A technology for transmitting data between computer devices at a data rate of up to 4 Gbps. It is especially suited for attaching computer servers to shared storage devices and for interconnecting storage controllers and drives.

fibre-channel extender
A long-distance communication device that interconnects storage area network (SAN) fabric components.

field replaceable unit
An assembly that is replaced in its entirety when any one of its components fails. In some cases, a field replaceable unit might contain other field replaceable units.

FlashCopy mapping
A relationship between two virtual disks.

FlashCopy relationship
See *FlashCopy mapping*.

FlashCopy service
In SAN Volume Controller, a copy service that duplicates the contents of a source virtual disk (VDisk) to a target VDisk. In the process, the original contents of the target VDisk are lost. See also *point-in-time copy.*

F_port See *fabric port*.

FRU See *field replaceable unit*.

G

gateway
An entity that operates above the link layer and translates, when required, the interface and protocol used by one network into those used by another distinct network.

GB See *gigabyte*.

GBIC See *gigabit interface converter*.

gigabit interface converter (GBIC)
An interface module that converts the light stream from a fibre-channel cable into electronic signals for use by the network interface card.

gigabyte (GB)
In decimal notation, 1 073 741 824 bytes.

Global Mirror
An asynchronous copy service that enables host data on a particular source virtual disk (VDisk) to be copied to the target VDisk that is designated in the relationship.

grain In a FlashCopy bitmap, the unit of data represented by a single bit.

H

hardcoded
Pertaining to software instructions that are statically encoded and not intended to be altered.

HBA See *host bus adapter*.

HLUN
See *virtual disk.*

host An open-systems computer that is connected to the SAN Volume Controller through a fibre-channel interface.

host bus adapter (HBA)
In SAN Volume Controller, an interface card that connects a host bus, such as a peripheral component interconnect (PCI) bus, to the storage area network.

host ID
In SAN Volume Controller, a numeric identifier assigned to a group of host fibre-channel ports for the purpose of logical unit number (LUN) mapping. For each host ID, there is a separate mapping of Small Computer System Interface (SCSI) IDs to virtual disks (VDisks).

host zone
A zone defined in the storage area network (SAN) fabric in which the hosts can address the SAN Volume Controllers.

hub A fibre-channel device that connects nodes into a logical loop by using a physical star topology. Hubs will automatically recognize an active node and insert the node into the loop. A node that fails or is powered off is automatically removed from the loop.

hub A communications infrastructure device to which nodes on a multi-point bus or loop are physically connected. Commonly used in Ethernet and fibre-channel networks to improve the manageability of physical cables. Hubs maintain the logical loop topology of the network of which they are a part, while creating a "hub and spoke" physical star layout. Unlike switches, hubs do not aggregate bandwidth. Hubs typically support the addition or removal of nodes from the bus while it is operating. (S) Contrast with *switch*.

I

ID See *identifier*.

identifier (ID)
A sequence of bits or characters that identifies a user, program device, or system to another user, program device, or system.

idle In a FlashCopy mapping, the state that occurs when the source and target virtual disks (VDisks) act as independent VDisks even if a mapping exists between the two. Read and write caching is enabled for both the source and the target.

idling

- The status of a pair of virtual disks (VDisks) that have a defined copy relationship for which no copy activity has yet been started.
- In a Metro or Global Mirror relationship, the state that indicates that the master virtual disks (VDisks) and auxiliary VDisks are operating in the primary role. Consequently, both VDisks are accessible for write I/O operations.

idling-disconnected
In a Global Mirror relationship, the state that occurs when the virtual disks (VDisks) in this half of the consistency group are all operating in the primary role and can accept read or write I/O operations.

illegal configuration
A configuration that will not operate and will generate an error code to indicate the cause of the problem.

image mode
An access mode that establishes a one-to-one mapping of extents in the managed disk (MDisk) with the extents in the virtual disk (VDisk). See also *managed space mode* and *unconfigured mode*.

image VDisk
A virtual disk (VDisk) in which there is a direct block-for-block translation from the managed disk (MDisk) to the VDisk.

IML See *initial microcode load*.

inconsistent
In a Metro or Global Mirror relationship, pertaining to a secondary virtual disk (VDisk) that is being synchronized with the primary VDisk.

inconsistent-copying
In a Global Mirror relationship, the state that occurs when the primary virtual disk (VDisk) is accessible for read and write input/output (I/O) operations, but the secondary VDisk is not accessible for either. This state occurs after a **start** command is issued to a consistency group that is in the inconsistent-stopped state. This state also occurs when a **start** command is issued, with the force option, to a consistency group that is in the idling or consistent-stopped state.

inconsistent-disconnected
In a Global Mirror relationship, a state that occurs when the virtual disks (VDisks) in the half of the consistency group that is operating in the secondary role are not accessible for either read or write I/O operations.

inconsistent-stopped
In a Global Mirror relationship, the state that occurs when the primary virtual disk (VDisk) is accessible for read and write input/output (I/O) operations, but the secondary VDisk is not accessible for either read or write I/O operations.

indication
An object representation of an event.

initial microcode load (IML)
In SAN Volume Controller, the process by which the run-time code and data for a node are loaded into memory and initialized.

initiator
The system component that originates an I/O command over an I/O bus or network. I/O adapters, network interface cards, and intelligent controller device I/O bus control ASICs are typical initiators. (S) See also *logical unit number*.

input/output (I/O)
Pertaining to a functional unit or communication path involved in an input process, an output process, or both, concurrently or not, and to the data involved in such a process.

instance
An individual object that is a member of some class. In object-oriented programming, an object is created by instantiating a class.

integrity
The ability of a system to either return only correct data or respond that it cannot return correct data.

Internet Protocol (IP)
In the Internet suite of protocols, a connectionless protocol that routes data through a network or interconnected networks and acts as an intermediary between the higher protocol layers and the physical network.

Inter-Switch Link (ISL)
A protocol for interconnecting multiple routers and switches in a storage area network.

I/O See *input/output.*

I/O group
A collection of virtual disks (VDisks) and node relationships that present a common interface to host systems.

I/O throttling rate
The maximum rate at which an I/O transaction is accepted for this virtual disk (VDisk).

IP See *Internet Protocol*.

IP address
The unique 32-bit address that specifies the location of each device or workstation in the Internet. For example, 9.67.97.103 is an IP address.

ISL See *Inter-Switch Link*.

ISL hop
Considering all pairs of node ports (N-ports) in a fabric and measuring distance only in terms of Inter-Switch Links (ISLs) in the fabric, the number of ISLs traversed is the number of ISL hops on the shortest route between the pair of nodes that are farthest apart in the fabric.

J

JBOD (just a bunch of disks)
- IBM definition: See *non-RAID*.
- HP definition: A group of single-device logical units not configured into any other container type.

L

LBA See *logical block address*.

least recently used (LRU)
An algorithm used to identify and make available the cache space that contains the least-recently used data.

line card
See *blade*.

local fabric
In SAN Volume Controller, those storage area network (SAN) components (such as switches and cables) that connect the components (nodes, hosts, switches) of the local cluster together.

local/remote fabric interconnect
The storage area network (SAN) components that are used to connect the local and remote fabrics together.

logical block address (LBA)
The block number on a disk.

logical unit (LU)
An entity to which Small Computer System Interface (SCSI) commands are addressed, such as a virtual disk (VDisk) or managed disk (MDisk).

logical unit number (LUN)
The SCSI identifier of a logical unit within a target. (S)

longitudinal redundancy check (LRC)
A method of error checking during data transfer that involves checking parity.

LRC See *longitudinal redundancy check*.

LRU See *least recently used*.

LU See *logical unit*.

LUN See *logical unit number*.

LUN masking
A process that allows or prevents I/O to the disk drives through the host-bus-adapter (HBA) device or operating-system device driver.

M

managed disk (MDisk)
A Small Computer System Interface (SCSI) logical unit that a redundant array of independent disks (RAID) controller provides and a cluster manages. The MDisk is not visible to host systems on the storage area network (SAN).

managed disk group
A collection of managed disks (MDisks) that, as a unit, contain all the data for a specified set of virtual disks (VDisks).

managed space mode
An access mode that enables virtualization functions to be performed. See also *image mode* and *unconfigured mode*.

Management Information Base (MIB)
Simple Network Management Protocol (SNMP) units of managed information that specifically describe an aspect of a system, such as the system name, hardware number, or communications configuration. A collection of related MIB objects is defined as a MIB.

mapping
See *FlashCopy mapping*.

master console
A single point from which to manage the IBM System Storage SAN Volume Controller. The master console product can be purchased as software that is installed and configured on a server or as a hardware platform with the operating system and master console software preinstalled.

master virtual disk
The virtual disk (VDisk) that contains a production copy of the data and that an application accesses. See also *auxiliary virtual disk.*

MB See *megabyte*.

MDisk
See *managed disk.*

megabyte (MB)
In decimal notation, 1 048 576 bytes.

mesh configuration
A network that contains a number of small SAN switches configured to create a larger switched network. With this configuration, four or more switches are connected together in a loop with some of the paths short circuiting the loop. An example of this configuration is to have four switches connected together in a loop with ISLs for one of the diagonals.

method
A way to implement a function on a class.

Metro Mirror
A synchronous copy service that enables host data on a particular source virtual disk (VDisk) to be copied to the target VDisk that is designated in the relationship.

MIB See *Management Information Base*.

migration
See *data migration*.

mirrorset
- IBM definition: See *RAID-1*.
- HP definition: A RAID storageset of two or more physical disks that maintain a complete and independent copy of the data from the virtual disk. This type of storageset has the advantage of being highly reliable and extremely tolerant of device failure. Raid level 1 storagesets are referred to as mirrorsets.

N

namespace
The scope within which a Common Information Model (CIM) schema applies.

node One SAN Volume Controller. Each node provides virtualization, cache, and Copy Services to the storage area network (SAN).

node name
A name identifier associated with a node. (SNIA)

node port (N_port)
A port that connects a node to a fabric or to another node. N_ports connect to fabric ports (F_ports) or to other N_ports of other nodes. N_ports handle creation, detection, and flow of message units to and from the connected systems. N_ports are end points in point-to-point links.

node rescue
In SAN Volume Controller, the process by which a node that has no valid software installed on its hard disk drive can copy the software from another node connected to the same fibre-channel fabric.

non-RAID
Disks that are not in a redundant array of independent disks (RAID). HP definition: See *JBOD*.

N_port
See *node port*.

NWWN
See *worldwide node name*.

O

object In object-oriented design or programming, a concrete realization of a class that consists of data and the operations associated with that data.

object model
A representation, such as a diagram, of objects in a given system. Using symbols similar to standard flowchart symbols, an object model depicts the classes the objects belong to, their associations with each other, the attributes that make them unique, and the operations that the objects can perform and that can be performed on them.

object name
An object that consists of a namespace path and a model path. The namespace path provides access to the Common Information Model (CIM) implementation managed by the CIM Agent, and the model path provides navigation within the implementation.

object path
An object that consists of a namespace path and a model path. The namespace path provides access to the Common Information Model (CIM) implementation managed by the CIM Agent, and the model path provides navigation within the implementation.

offline
Pertaining to the operation of a functional unit or device that is not under the continual control of the system or of a host.

online Pertaining to the operation of a functional unit or device that is under the continual control of the system or of a host.

operating set
In SAN Volume Controller, the set of nodes that are operating together to deliver storage services.

oversubscription
The ratio of the sum of the traffic that is on the initiator N-node connections to the traffic that is on the most heavily loaded Inter-Switch Links (ISLs), where more than one ISL is connected in parallel between these switches. This definition assumes a symmetrical network and a specific workload that is applied equally from all initiators and sent equally to all targets. See also *symmetrical network*.

P

partition

- IBM definition: A logical division of storage on a fixed disk.
- HP definition: A logical division of a container represented to the host as a logical unit.

partner node
The other node that is in the I/O group to which this node belongs.

partnership
In Metro or Global Mirror, the relationship between two clusters. In a cluster partnership, one cluster is defined as the local cluster and the other cluster as the remote cluster.

paused
In SAN Volume Controller, the process by which the cache component quiesces all ongoing I/O activity below the cache layer.

pend To cause to wait for an event.

petabyte (PB)
In decimal notation, 1 125 899 906 842 624 bytes.

PLUN See *managed disk*.

point-in-time copy
The instantaneous copy that the FlashCopy service makes of the source virtual disk (VDisk). In some contexts, this copy is known as a T_0 *copy*.

port The physical entity within a host, SAN Volume Controller, or disk controller system that performs the data communication (transmitting and receiving) over the fibre channel.

port ID
An identifier associated with a port.

power-on self-test
A diagnostic test that servers or computers run when they are turned on.

prepared
In a Global Mirror relationship, the state that occurs when the mapping is ready to start. While in this state, the target virtual disk (VDisk) is offline.

preparing
In a Global Mirror relationship, the state that occurs when any changed write data for the source virtual disk (VDisk) is flushed from the cache. Any read or write data for the target VDisk is discarded from the cache.

primary virtual disk
In a Metro or Global Mirror relationship, the target of write operations issued by the host application.

property
In the Common Information Model (CIM), an attribute that is used to characterize instances of a class.

PuTTY
A client program that allows you to run remote sessions on your computer through specific network protocols, such as SSH, Telnet, and Rlogin.

PWWN
See *worldwide port name*.

Q

qualifier
A value that provides additional information about a class, association, indication, method, method parameter, instance, property, or reference.

queue depth
The number of I/O operations that can be run in parallel on a device.

quorum disk
A managed disk (MDisk) that contains quorum data and that a cluster uses to break a tie and achieve a quorum.

quorum index
The pointer that indicates the order used to resolve a tie. Nodes attempt to lock the first quorum disk (index 0), followed by the next disk (index 1), and finally the last disk (index 2). The tie is broken by the node that locks them first.

R

rack A free-standing framework that holds the devices and card enclosure.

RAID See *redundant array of independent disks*.

RAID 0

- IBM definition: 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.
- HP definition: 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

- SNIA dictionary definition: A form of storage array in which two or more identical copies of data are maintained on separate media.
- IBM definition: A form of storage array in which two or more identical copies of data are maintained on separate media. Also known as mirrorset.
- HP definition: See *mirrorset*.

RAID 5

- SNIA definition: 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 array's disks. (S)
- IBM definition: See above.
- HP definition: 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.

RAID 10

A type of RAID that optimizes high performance while maintaining fault tolerance for up to two failed disk drives by striping volume data across several disk drives and mirroring the first set of disk drives on an identical set.

redundant ac power switch

A device that provides redundant power from the primary (main) ac source to the SAN Volume Controller in case of power failure. If the main source becomes unavailable, the redundant ac power switch automatically provides power from a secondary (backup) source. When power is restored, the redundant ac power switch automatically changes back to the main power source.

redundant array of independent disks

A collection of two or more disk drives that present the image of a single disk drive to the system. In the event of a single device failure, the data can be read or regenerated from the other disk drives in the array.

redundant SAN
A storage area network (SAN) configuration in which any one single component might fail, but connectivity between the devices within the SAN is maintained, possibly with degraded performance. This configuration is normally achieved by splitting the SAN into two, independent, counterpart SANs. See also *counterpart SAN.*

reference
A pointer to another instance that defines the role and scope of an object in an association.

rejected
A status condition that describes a node that the cluster software has removed from the working set of nodes in the cluster.

relationship
In Metro or Global Mirror, the association between a master virtual disk (VDisk) and an auxiliary VDisk. These VDisks also have the attributes of a primary or secondary VDisk. See also *auxiliary virtual disk, master virtual disk, primary virtual disk, and secondary virtual disk.*

reliability
The ability of a system to continue to return data even if a component fails.

remote fabric
In Global Mirror, the storage area network (SAN) components (switches and cables) that connect the components (nodes, hosts, and switches) of the remote cluster.

roles Authorization is based on roles that map to the administrator and service roles in an installation. The switch translates these roles into SAN Volume Controller administrator and service user IDs when a connection is made to the node for the SAN Volume Controller.

S

SAN See *storage area network*.

SAN Volume Controller fibre-channel port fan in
The number of hosts that can see any one SAN Volume Controller port.

SATA See *Serial Advanced Technology Attachment*.

schema
A group of object classes defined for and applicable to a single namespace. Within the CIM Agent, the supported schemas are the ones that are loaded through the managed object format (MOF).

SCSI See *Small Computer Systems Interface*.

SCSI back-end layer
The layer in a Small Computer Systems Interface (SCSI) network that performs the following functions: controls access to individual disk controller systems that are managed by the cluster; receives requests from the virtualization layer, processes them, and sends them to managed disks; addresses SCSI-3 commands to the disk controller systems on the storage area network (SAN).

SCSI front-end layer
The layer in a Small Computer Systems Interface (SCSI) network that receives I/O commands sent from hosts and provides the SCSI-3 interface

to hosts. SCSI logical unit numbers (LUNs) are mapped to virtual disks (VDisks) in this layer as well. Thus, the layer converts SCSI read and write commands that are addressed to LUNs into commands that are addressed to specific VDisks.

SDD See *subsystem device driver (SDD)*.

SDRAM
See *Synchronous Dynamic Random Access Memory*.

secondary virtual disk
In Metro or Global Mirror, the virtual disk (VDisk) in a relationship that contains a copy of data written by the host application to the primary VDisk.

Secure Shell
A program to log in to another computer over a network, to execute commands in a remote machine, and to move files from one machine to another.

sequential VDisk
A virtual disk that uses extents from a single managed disk.

Serial Advanced Technology Attachment (SATA)
The evolution of the ATA interface from a parallel bus to serial connection architecture. (S)

Serial ATA
See *Serial Advanced Technology Attachment*.

server In a network, the hardware or software that provides facilities to other stations; for example, a file server, a printer server, a mail server. The station making the request of the server is usually called the client.

Service Location Protocol (SLP)
In the Internet suite of protocols, a protocol that identifies and uses network hosts without having to designate a specific network host name.

Simple Mail Transfer Protocol (SMTP)
An Internet application protocol for transferring mail among users of the Internet. SMTP specifies the mail exchange sequences and message format. It assumes that the Transmission Control Protocol (TCP) is the underlying protocol.

Simple Network Management Protocol (SNMP)
In the Internet suite of protocols, a network management protocol that is used to monitor routers and attached networks. SNMP is an application-layer protocol. Information on devices managed is defined and stored in the application's Management Information Base (MIB).

SLP See *Service Location Protocol*.

Small Computer System Interface (SCSI)
A standard hardware interface that enables a variety of peripheral devices to communicate with one another.

SMI-S See *Storage Management Initiative Specification*.

SMTP See *Simple Mail Transfer Protocol*.

SNIA See *Storage Networking Industry Association*.

SNMP
See *Simple Network Management Protocol*.

SSH See *Secure Shell*.

stand-alone relationship
In FlashCopy, Metro Mirror, and Global Mirror, relationships that do not belong to a consistency group and that have a null consistency group attribute.

stop A configuration command that is used to stop the activity for all copy relationships in a consistency group.

stopped
The status of a pair of virtual disks (VDisks) that have a copy relationship that the user has temporarily broken because of a problem.

storage area network (SAN)
A network whose primary purpose is the transfer of data between computer systems and storage elements and among storage elements. A SAN consists of a communication infrastructure, which provides physical connections, and a management layer, which organizes the connections, storage elements, and computer systems so that data transfer is secure and robust. (S)

Storage Management Initiative Specification (SMI-S)
A design specification developed by the Storage Networking Industry Association (SNIA) that specifies a secure and reliable interface that allows storage management systems to identify, classify, monitor, and control physical and logical resources in a storage area network. The interface is intended as a solution that integrates the various devices to be managed in a storage area network (SAN) and the tools used to manage them.

Storage Networking Industry Association (SNIA)
An association of producers and consumers of storage networking products whose goal is to further storage networking technology and applications. See www.snia.org.

striped
Pertains to a virtual disk (VDisk) that is created from multiple managed disks (MDisks) that are in the MDisk group. Extents are allocated on the MDisks in the order specified.

stripeset
See *RAID 0*.

subsystem device driver (SDD)
An IBM pseudo device driver designed to support the multipath configuration environments in IBM products.

superuser authority
The level of access required to add users.

suspended
The status of a pair of virtual disks (VDisks) that have a copy relationship that has been temporarily broken because of a problem.

switch A network infrastructure component to which multiple nodes attach. Unlike hubs, switches typically have internal bandwidth that is a multiple of link bandwidth, and the ability to rapidly switch node connections from one to another. A typical switch can accommodate several simultaneous full link bandwidth transmissions between different pairs of nodes. (S) Contrast with *hub*.

symmetrical network
A network in which all the initiators are connected at the same level and all the controllers are connected at the same level.

symmetric virtualization
A virtualization technique in which the physical storage in the form of Redundant Array of Independent Disks (RAID) is split into smaller chunks of storage known as *extents*. These extents are then concatenated, using various policies, to make virtual disks (VDisks). See also *asymmetric virtualization*.

synchronized
In Metro or Global Mirror, the status condition that exists when both virtual disks (VDisks) of a pair that has a copy relationship contain the same data.

Synchronous Dynamic Random Access Memory (SDRAM)
A type of dynamic random access memory (DRAM) with features that make it faster.

system
A functional unit, consisting of one or more computers and associated software, that uses common storage for all or part of a program and also for all or part of the data necessary for the execution of the program. A computer system can be a stand-alone unit, or it can consist of multiple connected units.

T

terabyte
In decimal notation, 1 099 511 628 000 bytes.

topology
The logical layout of the components of a computer system or network and their interconnections. Topology deals with questions of what components are directly connected to other components from the standpoint of being able to communicate. It does not deal with questions of physical location of components or interconnecting cables. (S)

IBM TotalStorage Enterprise Storage Server (ESS)
An IBM product that provides an intelligent disk-storage subsystem across an enterprise.

trigger
To initiate or reinitiate copying between a pair of virtual disks (VDisks) that have a copy relationship.

U

unconfigured mode
A mode in which I/O operations cannot be performed. See also *image mode* and *managed space mode*.

uninterruptible power supply (UPS)
A device connected between a computer and its power source that protects the computer against blackouts, brownouts, and power surges. The uninterruptible power supply contains a power sensor to monitor the supply and a battery to provide power until an orderly shutdown of the system can be performed.

unit identifiers (UIDs)
A unit identifier can be one of the following:

1. an integer expression whose value must be zero or positive
2. an * (asterisk) that corresponds to unit 5 for input or unit 6 for output
3. the name of a character array, character array element, or character substring for an internal file

unmanaged
An access mode that pertains to a managed disk (MDisk) that is not used by the cluster.

UPS See *uninterruptible power supply*.

V

valid configuration
A configuration that is supported.

VDisk See *virtual disk*.

virtual disk (VDisk)
In SAN Volume Controller, a device that host systems attached to the storage area network (SAN) recognize as a Small Computer System Interface (SCSI) disk.

virtualization
In the storage industry, a concept in which a pool of storage is created that contains several disk subsystems. The subsystems can be from various vendors. The pool can be split into virtual disks that are visible to the host systems that use them.

virtualized storage
Physical storage that has virtualization techniques applied to it by a virtualization engine.

virtual storage area network (VSAN)
A fabric within the SAN.

vital product data (VPD)
Information that uniquely defines system, hardware, software, and microcode elements of a processing system.

VLUN See *managed disk*.

VSAN See *virtual storage area network*.

W

WBEM
See *Web-Based Enterprise Management*.

Web-Based Enterprise Management (WBEM)
A tiered, enterprise-management architecture that was developed by the Distributed Management Task Force (DMTF). This architecture provides the management design framework that consists of devices, device providers, the object manager, and the messaging protocol for the communication between client applications and the object manager.

worldwide node name (WWNN)
An identifier for an object that is globally unique. WWNNs are used by Fibre Channel and other standards.

worldwide port name (WWPN)
A unique 64-bit identifier associated with a fibre-channel adapter port. The WWPN is assigned in an implementation- and protocol-independent manner.

WWNN
See *worldwide node name*.

WWPN
See *worldwide port name*.

Z

zoning
In fibre-channel environments, the grouping of multiple ports to form a virtual, private, storage network. Ports that are members of a zone can communicate with each other, but are isolated from ports in other zones.

Index

Numerics

A

B

C

E

F

G

H

I

K

M

N

O

P

R

S

© Copyright IBM Corp. 2003, 2007

T

U

V

W

Z

Readers' Comments — We'd Like to Hear from You

IBM System Storage SAN Volume Controller
Planning Guide
Version 4.2.0

Publication No. GA32-0551-01

We appreciate your comments about this publication. Please comment on specific errors or omissions, accuracy, organization, subject matter, or completeness of this book. The comments you send should pertain to only the information in this manual or product and the way in which the information is presented.

For technical questions and information about products and prices, please contact your IBM branch office, your IBM business partner, or your authorized remarketer.

When you send comments to IBM, you grant IBM a nonexclusive right to use or distribute your comments in any way it believes appropriate without incurring any obligation to you. IBM or any other organizations will only use the personal information that you supply to contact you about the issues that you state on this form.

Comments:

Thank you for your support.

Submit your comments using one of these channels:

- Send your comments to the address on the reverse side of this form.

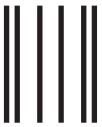
If you would like a response from IBM, please fill in the following information:

Name	Address
Company or Organization	
Phone No.	E-mail address

Readers' Comments — We'd Like to Hear from You
GA32-0551-01

Cut or Fold
Along Line

Fold and Tape | **Please do not staple** | Fold and Tape



NO POSTAGE
NECESSARY
IF MAILED IN THE
UNITED STATES

BUSINESS REPLY MAIL

FIRST-CLASS MAIL PERMIT NO. 40 ARMONK, NEW YORK

POSTAGE WILL BE PAID BY ADDRESSEE

International Business Machines Corporation
Information Development
Department 61C
9032 South Rita Road
Tucson, Arizona
USA 85775-4401

Fold and Tape | **Please do not staple** | Fold and Tape

Cut or Fold
Along Line

Printed in USA

GA32-0551-01

Spine information:

IBM System Storage SAN Volume Controller

SAN Volume Controller Planning Guide

Version 4.2.0