IBM System Storage SAN Volume Controller

Planning Guide

Version 5.1.0

GA32-0551-06

IBM System Storage SAN Volume Controller

Planning Guide

Version 5.1.0

GA32-0551-06

Note:
Before using this information and the product it supports, read the information in **Notices**.

This edition applies to the IBM System Storage SAN Volume Controller, release 5.1.0, and to all subsequent releases and modifications until otherwise indicated in new editions. This edition replaces GA32-0551-05.

Contents

Figures

Tables

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 GA32-0551-06 and GA32-0551-05 SAN Volume Controller Planning Guide

The summary of changes provides a list of new and changed information since the previous version of the guide.

New information

This topic describes the changes to this guide since the previous edition, GA32-0551-04. The following sections summarize the changes that have since been implemented from the previous version.

This version includes the following new information:

- Support statements for the SAN Volume Controller 2145-CF8 node and the optional solid-state drive (SSD)
- Support statements for the syslog notification protocol
- Support statements for second cluster IP and service IP addresses
- Support statements for iSCSI 1Gb host attachment using on-board ethernet ports

Removed Information

This version has had the following information removed:

- The SAN Volume Controller 2145-4F2 node is not supported in this version of SAN Volume Controller

Summary of changes for GA32-0551-04 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, GA32-0551-03. The following sections summarize the changes that have since been implemented from the previous version.

This version includes the following new information:

- A description of the new SAN Volume Controller model, the SAN Volume Controller 2145-8A4
- A description of the IP address and subnet requirements that include items such as determining cluster addresses and service mode addresses

Changed information

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

- Change in terminology from storage subsystem to storage system.

Emphasis

Different typefaces are used in this guide to show emphasis.

The following typefaces are used to show emphasis:

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

SAN Volume Controller library and related publications

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

SAN Volume Controller Information Center

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

http://publib.boulder.ibm.com/infocenter/svcic/v3r1m0/index.jsp

SAN Volume Controller library

Table 1 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:

www.ibm.com/storage/support/2145

Table 1. SAN Volume Controller library

Title	Description	Order number
IBM System Storage SAN Volume Controller Planning Guide	This guide introduces the SAN Volume Controller and lists the features that 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 Model 2145-CF8 Hardware Installation Guide	This guide provides the instructions that the IBM service representative uses to install the hardware for SAN Volume Controller model 2145-CF8.	GC52-1356
IBM System Storage SAN Volume Controller Model 2145-8A4 Hardware Installation Guide	This guide provides the instructions that the IBM service representative uses to install the hardware for SAN Volume Controller model 2145-8A4.	GC27-2219
IBM System Storage SAN Volume Controller Model 2145-8G4 Hardware Installation Guide	This guide provides the instructions that the IBM service representative uses to install the hardware for SAN Volume Controller model 2145-8G4.	GC27-2220
IBM System Storage SAN Volume Controller Models 2145-8F2 and 2145-8F4 Hardware Installation Guide	This guide provides the instructions that the IBM service representative uses to install the hardware for SAN Volume Controller models 2145-8F2 and 2145-8F4.	GC27-2221
IBM System Storage SAN Volume Controller Software Installation and Configuration Guide	This guide provides guidelines for configuring your SAN Volume Controller. Instructions for backing up and restoring the cluster configuration, using and upgrading the SAN Volume Controller Console, using the CLI, upgrading the SAN Volume Controller software, and replacing or adding nodes to a cluster are included.	SC23-6628

Table 1. SAN Volume Controller library (continued)

Title	Description	Order number
IBM System Storage SAN Volume Controller CIM Agent Developer's Guide	This guide describes the concepts of the Common Information Model (CIM) environment. Steps about using the CIM agent object class instances to complete basic storage configuration tasks, establishing new Copy Services relationships, and performing CIM agent maintenance and diagnostic tasks are included.	SC23-6665
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
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 Troubleshooting Guide	This guide describes the features of each SAN Volume Controller model, explains how to use the front panel, and provides maintenance analysis procedures to help you diagnose and solve problems with the SAN Volume Controller.	GC27-2227
IBM System Storage SAN Volume Controller Hardware Maintenance Guide	This guide provides the instructions that the IBM service representative uses to service the SAN Volume Controller hardware, including the removal and replacement of parts.	GC27-2226
IBM System Storage SAN Volume Controller Master Console Guide	This guide describes how to install, maintain, and service the master console.	GC27-2223
IBM Systems Safety Notices	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 Systems Safety Notices* document.	G229-9054

Other IBM publications

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

Table 2. Other IBM publications

Title	Description	Order number
IBM System Storage Productivity Center Introduction and Planning Guide	This guide introduces the IBM System Storage Productivity Center hardware and software.	SC23-8824
Read This First: Installing the IBM System Storage Productivity Center	This guide describes how to install the IBM System Storage Productivity Center hardware.	GI11-8938
IBM System Storage Productivity Center User's Guide	This guide describes how to configure the IBM System Storage Productivity Center software.	SC27-2336
IBM System Storage Multipath Subsystem Device Driver User's Guide	This guide describes the IBM System Storage Multipath Subsystem Device Driver for IBM System Storage products and how to use it with the SAN Volume Controller.	GC52-1309
Implementing the IBM System Storage SAN Volume Controller V4.3	This IBM Redbooks® publication is a detailed technical guide to the IBM System Storage SAN Volume Controller. It provides a high-level overview of storage virtualization and the SAN Volume Controller architecture, discusses implementing and configuring the SAN Volume Controller, tells you how to migrate existing storage to the SAN Volume Controller, and discusses different supported migration activities.	SG24-6423

IBM documentation and related Web sites

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

Table 3. IBM documentation and related Web sites

Web site	Address
Support for SAN Volume Controller (2145)	www.ibm.com/storage/support/2145
Support for IBM System Storage and IBM TotalStorage® products	www.ibm.com/storage/support/
IBM Publications Center	www.ibm.com/shop/publications/order/
IBM Redbooks publications	www.redbooks.ibm.com/

Related accessibility information

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

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

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	www.ibm.com/storage/support/2145
Technical support for IBM storage products	www.ibm.com/storage/support/

How to order IBM publications

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

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

www.ibm.com/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 software and hardware into a comprehensive, modular appliance that uses symmetric virtualization.

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

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

SAN Volume Controller software

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

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

The SAN Volume Controller also provides the following functions:

- Large scalable cache
- Copy Services
  - IBM 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
  - Space-efficient logical volumes (thin provisioning)

SAN Volume Controller hardware

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

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

Note: I/O groups take the storage that is presented to the SAN by the storage systems as MDisks and translates the storage into logical disks, known as VDisks, that are used by applications on the hosts. A node resides in only one I/O group and provides access to the VDisks in that I/O group.

The following nodes are supported in SAN Volume Controller 5.1:

- The new SAN Volume Controller 2145-CF8 node is available for purchase, with up to four of the optional solid-state drives (SSDs).
- The SAN Volume Controller 2145-8A4 node remains available for purchase.
- The SAN Volume Controller 2145-8G4 node is no longer available for purchase, but remains supported in SAN Volume Controller 5.1.
- The SAN Volume Controller 2145-8F4 node is no longer available for purchase, but remains supported in SAN Volume Controller 5.1.
- The SAN Volume Controller 2145-8F2 node is no longer available for purchase, but remains supported in SAN Volume Controller 5.1.

SAN Volume Controller objects

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 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*, which provides data integrity in the event of a cluster wide power failure. In such situations, the uninterruptible power supply maintains power to the nodes while the contents of the distributed cache are dumped to an internal drive.

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

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

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

A VDisk can be space-efficient. This means that the capacity of the VDisk as seen by host systems, called the virtual capacity, can be different from the amount of

storage that is allocated to the VDisk from MDisks, called the real capacity. Space-efficient VDisks can be configured to automatically expand their real capacity by allocating new extents.

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

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

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

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

The cluster provides block-level aggregation and volume management for disk storage within the SAN. The cluster manages a number of back-end storage 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 back-end physical storage. This prevents any possible conflict between the cluster and the application servers both trying to manage the back-end storage.

SAN Volume Controller cluster high availability

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

Each I/O group within a cluster consists of a pair of nodes. If a node fails within an I/O group, the other node in the I/O group assumes the I/O responsibilities of the failed node. If the node contains solid-state drives (SSDs), you should create a mirrored virtual disk (VDisk) of any VDisk that uses the SSDs. SSDs can be a single point of failure in the event of an outage to the SSDs or to the node itself.

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

Each SAN Volume Controller node has four fibre-channel ports, which can be used to attach the node to multiple SAN fabrics. For high availability, attach the nodes in a cluster to at least two fabrics. SAN Volume Controller software incorporates multipathing software that is used for communication among SAN Volume Controller nodes and for I/O operations among SAN Volume Controller nodes and storage systems. If a SAN fabric fault disrupts communication or I/O operations,

the multipathing software recovers and retries the operation through an alternative communication path. Also for high availability, configure your fibre-channel host systems to use multipathing software. If a SAN fabric fault or node failure occurs, I/O operations among fibre-channel host systems and SAN Volume Controller nodes are retried. Subsystem device driver (SDD) multipathing software is available from IBM at no additional charge for use with SAN Volume Controller. For additional information about subsystem device driver (SDD), go to the Support for IBM Systems Web site:

www.ibm.com/systems/support

iSCSI-attached hosts connect to SAN Volume Controller through node Ethernet ports. If a node fails, SAN Volume Controller maintains host availability by failing over the IP addresses of the failed node to the partner node in the I/O group.

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

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

SAN Volume Controller operating environment

To use the SAN Volume Controller, you must meet the minimum hardware and software requirements and ensure that other operating environment criteria are met.

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 supply units
- One IBM System Storage Productivity Center or one master console per SAN installation for configuration

SAN Volume Controller 2145-CF8 node features

The SAN Volume Controller 2145-CF8 node has the following features:

- A 19-inch rack-mounted enclosure
- One 4-port 8 Gbps fibre-channel adapter
- 24 GB memory
- One quad-core processor
- Dual, redundant power supplies
- Supports up to four optional solid-state drives (SSDs)

SAN Volume Controller 2145-8A4 node features

The SAN Volume Controller 2145-8A4 node has the following features:

- A 19-inch rack-mounted enclosure
- One 4-port 4 Gbps fibre-channel adapter
- 8 GB cache memory

- One dual-core processor

Solid-state drive (SSD) features

Support for solid-state drives (SSDs) is an optional feature of the SAN Volume Controller 2145-CF8. SSDs include the following features:

- Up to four SSDs can be installed on each SAN Volume Controller 2145-CF8 node. An IBM PCIe SAS host bus adapter (HBA) is required on each node that contains an SSD.
- Each SSD is a 2.5-inch Serial Attached SCSI (SAS) drive.
- Each SSD provides up to 146 GB of real capacity.
- SSDs are hot-pluggable and hot-swappable.

Supported hosts

In a SAN environment, host systems are application servers that access data from the storage controllers that are connected to the SAN. Hosts that are running in a number of operating environments can connect to the storage using the SAN Volume Controller. Host connections to the SAN Volume Controller are either SCSI over the fibre-channel SAN or iSCSI over an Ethernet network.

For a list of the supported host operating systems, go to the IBM System Storage SAN Volume Controller Web site:

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

From the Web site, take the following steps:

1. In the **Learn more** column, click **Interoperability**.
2. Click **Recommended software levels** for your SAN Volume Controller code version.
3. Click **Multipathing / Host Drivers, Clustering and SAN Boot Support - By Host Operating System** to view a list of supported operating systems and to access host attachment scripts.

Multipathing software

For the most current information, go to the following Web site:

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

From the Web site, take the following steps:

1. In the **Learn more** column, click **Interoperability**.
2. Click **Recommended software levels** for your SAN Volume Controller code version.
3. Click **Multipathing / Host Drivers, Clustering and SAN Boot Support - By Host Operating System** to view a list of supported operating systems and to access multipath drivers. You can also view **Multipath Driver Co-existence with SDD** information.

User interfaces

The SAN Volume Controller software provides the following user interfaces:

- 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 software 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.

2145 UPS-1U

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

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

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

Redundant ac-power switch

The redundant ac-power switch is an optional feature that makes the SAN Volume Controller nodes resilient to the failure of a single power circuit. The redundant ac-power switch is not a replacement for an uninterruptible power supply. You must still use a uninterruptible power supply 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 turn off. Because the nodes are in two different I/O groups, the hosts do not lose access to the back-end disk data.

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

Figure 1 shows a redundant ac-power switch.

Figure 1. Photo of the redundant ac-power switch

IBM System Storage Productivity Center

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

SSPC simplifies storage management in the following ways:

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

SSPC includes the following software components:

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

Figure 2 on page 8 shows an overview of how SSPC and the components of IBM Tivoli Storage Productivity Center, IBM System Storage DS8000, and SAN Volume Controller interrelate with each other.

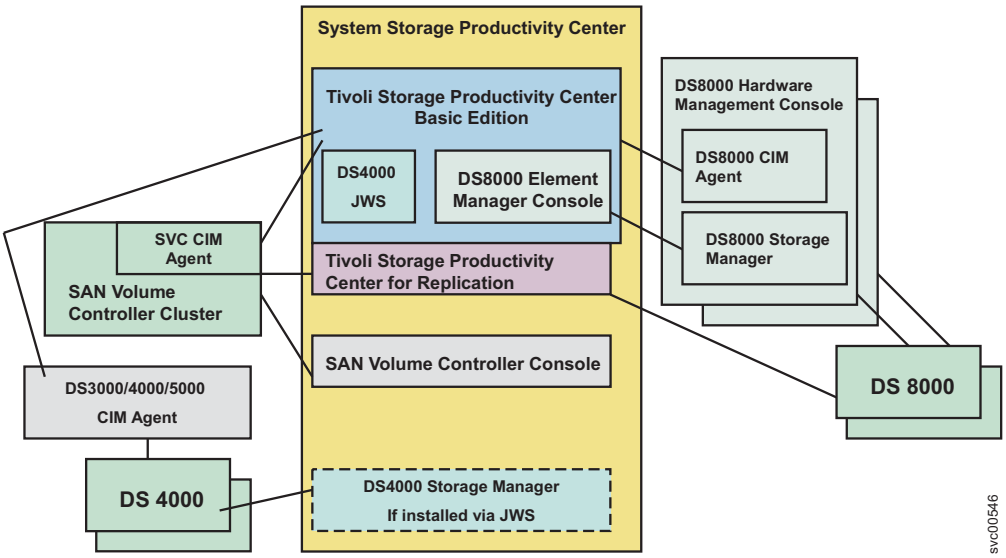

Figure 2. Overview of the IBM System Storage Productivity Center

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

Assist On-site and remote service

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

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

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

www.ibm.com/support/assistonsite/

When you access the Web site, you sign in and enter a code that the IBM service representative provides to you. This code is unique to each IBM Assist On-site session. A plug-in is downloaded onto your SSPC or master console to connect you and your IBM service representative to the remote service session. The IBM Assist On-site contains several layers of security to protect your applications and your computers. You can also use security features to restrict access by the IBM service representative.

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

Secure Shell protocol through PuTTY

Secure Shell (SSH) software is a client-server protocol that can be used from the IBM System Storage Productivity Center or from a host server to enable you to control the SAN Volume Controller through a command-line interface (CLI).

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 to a remote system. If you want to create an SSH connection to a server using an SSH key pair (such as to the SAN Volume Controller cluster), you must place the public key on the server.

Event notifications

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

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

Table 4 describes the types of event notifications.

Table 4. SAN Volume Controller notification types

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

SNMP traps

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

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

www.ibm.com/storage/support/2145

Search for **SAN Volume Controller MIB**. Go to the downloads results to find **Management Information Base (MIB) file for SNMP**. Click this link to find download options. The name of this file is `SVC_MIB_<release>.MIB` such as in `SVC_MIB_4.3.1.MIB`.

Syslog messages

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

Table 5 shows how SAN Volume Controller notification codes map to syslog security-level codes.

Table 5. SAN Volume Controller notification codes and corresponding syslog level codes

SAN Volume Controller notification code	Syslog level code	Description
SS_EID_UNKNOWN	Not mapped	
SS_EID_ERROR	LOG_ALERT	Error that needs immediate attention
SS_EID_WARNING	LOG_ERROR	Warning that needs attention
SS_EID_INFO	LOG_INFO	Informational messages
SS_EID_TEST	LOG_DEBUG	Test message

Table 6 on page 11 shows how syslog facility codes map to SAN Volume Controller values of user-defined message origin identifiers.

Table 6. Syslog facility codes and SAN Volume Controller values of user-defined message origin identifiers

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

Call Home e-mail

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

To send e-mail, you must configure at least one SMTP server. You can specify as many as five additional SMTP servers for backup purposes. The SMTP server must accept the relaying of e-mail from the SAN Volume Controller cluster IP address. You can then use the SAN Volume Controller Console or the SAN Volume Controller command-line interface to configure the e-mail settings, including contact information and e-mail recipients. Set the reply address to a valid e-mail address. Send a test e-mail to check that all connections and infrastructure are set up 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.

Clusters

All your configuration, monitoring, 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 between two and eight SAN Volume Controller nodes.

All configuration settings are replicated across all nodes in the cluster. Because configuration is performed at the cluster level, management IP addresses are assigned to the cluster instead of to each node. The cluster is configured using the SAN Volume Controller Console, the command-line interface (CLI) or an application developed to access the SAN Volume Controller CIMOM. Each interface accesses the cluster remotely through the Ethernet cluster-management address.

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

is the only node that activates the cluster IP addresses. You can use one or more of these addresses to access the cluster through the SAN Volume Controller graphical user interface or the command-line interface (CLI).

Each SAN Volume Controller cluster can have one to four management IP addresses. You can assign up to two IPv4 addresses and up to two IPv6 addresses. When a node has been assigned to a cluster, you can display the cluster IP addresses on the front panel by selecting **Cluster** from the menu.

Each SAN Volume Controller cluster can have optional Small Computer System Interface over Internet Protocol (iSCSI IP) addresses.

Note: Management IP addresses that are assigned to a cluster must be different from the iSCSI IP addresses and are used for different purposes. If iSCSI is used, iSCSI addresses are assigned to individual node ports. On the configuration node, a port will have multiple IP addresses active at the same time.

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 disk drive of each node. If a power failure occurs, the write cache data and configuration information that is held in memory is stored on the internal disk drive of the node. If the partner node is still online, it attempts to flush the cache and continue operation with the write cache disabled.

Figure 3 on page 13 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 in the cluster maintains an identical copy of the cluster state. When a change is made to the configuration or internal cluster data, the same change is applied to all nodes.

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. If the configuration node fails, the cluster can elect a new node to take over its responsibilities.

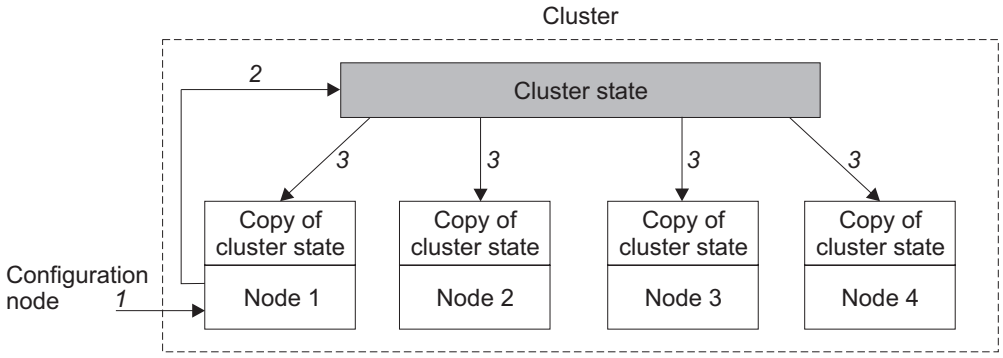

Figure 3. Cluster, nodes, and cluster state.

Cluster operation and quorum disks

Nodes are deployed in pairs known as input/output (I/O) groups, and one to four I/O groups comprise a cluster. For the cluster to be functional, at least 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.

Quorum disks are used when there is a problem in the SAN fabric or when nodes are shut down, leaving half of the nodes remaining in the cluster. This type of problem causes a loss of communication between the nodes that remain in the cluster and those that do not. The nodes are split into groups where the nodes in each group can communicate with each other, but not with the other group of nodes that were formerly part of the cluster.

In this situation, some nodes must stop operating and processing I/O requests from hosts to preserve data integrity while maintaining data access. If a group contains less than half the nodes that were active in the cluster, the nodes in that group stop operating and processing I/O requests from hosts.

It is possible for a cluster to split into two groups with each group containing half the original number of nodes in the cluster. A quorum disk determines which group of nodes stops operating and processing I/O requests. In this tie-break situation, the first group of nodes that accesses the quorum disk marks their ownership of the quorum disk and as a result continues to operate as the cluster, handling all I/O requests. If the other group of nodes cannot access the quorum disk or finds it owned by another group of nodes, it stops operating as the cluster and does not handle I/O requests.

I/O groups

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 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. When you create a VDisk, you can specify a preferred node. If you do not specify a preferred node, one is automatically assigned after the VDisk is created. To improve cache performance, many multipathing device drivers perform I/O operations through the preferred node whenever possible.

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 write I/O is performed to a VDisk, the node that processes the I/O duplicates the data when write I/O is performed on the partner node. I/O traffic for a particular VDisk is, at any one time, managed exclusively by the nodes in a single I/O group. Because SAN Volume Controller nodes handle I/O in independent pairs, you can easily scale the I/O capability of a cluster by adding I/O groups to obtain additional throughput.

Figure 4 on page 15 shows an example I/O group. A write operation from a host is shown (item *1*), that is targeted for VDisk A. This write operation is targeted at the preferred node, Node 1 (item *2*). The write operation is cached and a copy of the data is made in the partner node, the Node 2 cache (item *3*). The write operation 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*).

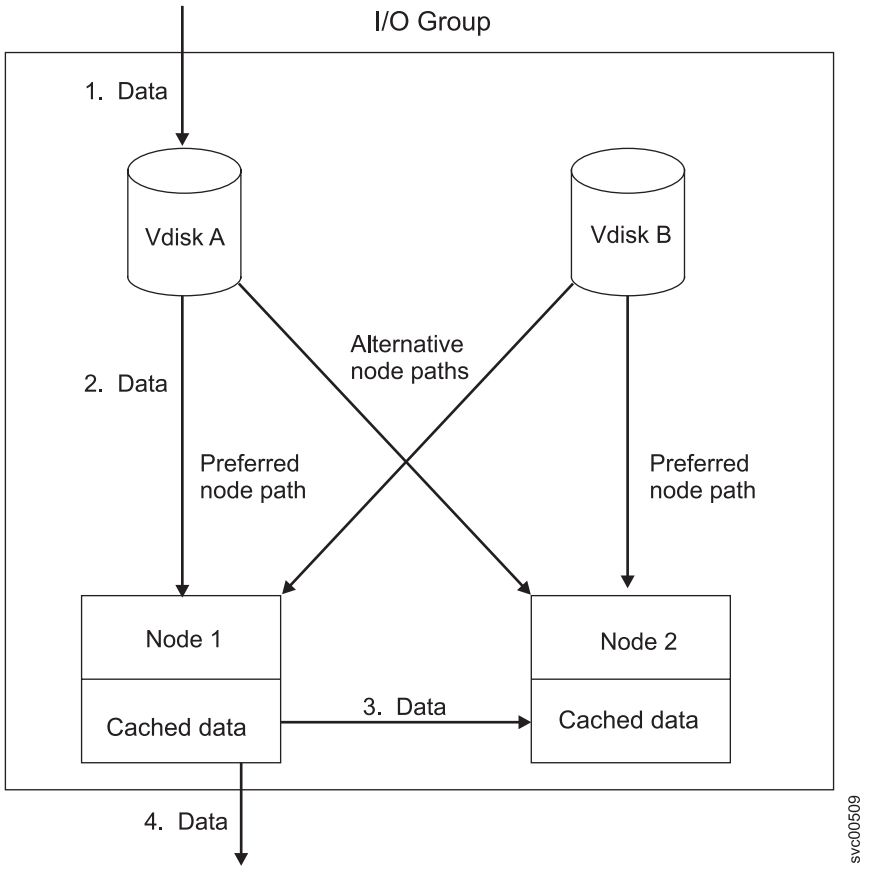

Figure 4. I/O group

If a node fails, 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 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 write operations 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 virtual recovery I/O group that can be used for certain service actions. This allows you to move the VDisks to the recovery I/O group and then into a working I/O group. Once assigned to the recovery I/O group, VDisks are offline and not available for I/O access.

Storage systems

A *storage system*, or *storage 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.

Storage systems provide the storage that a SAN Volume Controller cluster detects as one or more managed disks (MDisks).

Note: You must ensure that your storage systems and associated devices are configured to meet your cluster requirements.

Supported RAID storage systems are detected by the cluster and displayed by the user interfaces. You can display and filter the MDisks that represent the cluster storage systems.

SAN Volume Controller supports both RAID and non-RAID storage systems. RAID storage systems provide redundancy at the disk level, which prevents a single physical disk failure from causing an MDisk, MDisk group or associated VDisk failure. To minimize data loss, only virtualize the following RAID storage systems: RAID 1, RAID 10, RAID 0+1 and RAID 5.

Attention: Do not virtualize RAID 0, because a single physical disk failure of RAID 0 could cause multiple VDisk failures.

A storage system can have a local name for the RAID or single disks that it is providing. However, cluster nodes cannot determine this name because the namespace is local to the storage system. The storage system can display the disks with a unique ID, storage system LUN or LU number. This ID, along with the storage system serial number, can be used to associate the MDisks in the cluster with the RAID that is presented by the storage system.

Virtual disks

A *virtual disk (VDisk)* is a logical disk that the cluster presents to the hosts.

To keep a VDisk accessible even when a managed disk on which it depends has become unavailable, a mirrored copy can be added to a selected VDisk. Each VDisk can have a maximum of two copies. Each VDisk copy is created from a set of extents in an MDisk group.

Application servers on the SAN access VDisks, not managed disks (MDisks).

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

Types

Each VDisk copy can be one of the following types:

Striped

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

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

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

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

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

Figure 5 shows an example of an MDisk group that contains three MDisks. This figure also shows a striped VDisk copy that is created from the extents that are available in the group.

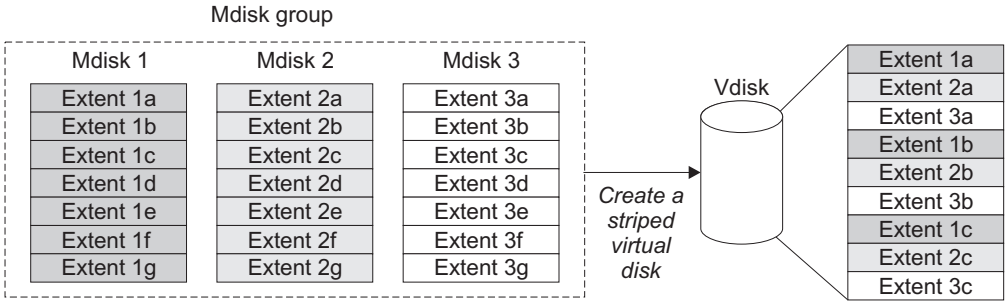

Figure 5. MDisk groups and VDisks

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

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

When you create an image-mode VDisk copy, you must assign it to an MDisk group. An image-mode VDisk copy must be at least one extent in size. The minimum size of an image-mode VDisk copy is the extent size of the MDisk group to which it is assigned.

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

Attention: If you add a managed mode MDisk to an MDisk group, any data on the MDisk is lost. Ensure that you create image-mode VDisks from the MDisks that contain data before you start adding any MDisks to groups.

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

You can use more sophisticated extent allocation policies to create VDisk copies. When you create a striped VDisk, you can specify the same MDisk more than once in the list of MDisks that are used as the stripe set. This is useful if you have an MDisk group in which not all the MDisks are of the same capacity. For example, if you have an MDisk group that has two 18 GB MDisks and two 36 GB MDisks, you can create a striped VDisk copy by specifying each of the 36 GB MDisks twice in the stripe set so that two-thirds of the storage is allocated from the 36 GB disks.

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

States

A VDisk can be in one of three states: online, offline, and degraded. Table 7 describes the different states of a VDisk.

Table 7. VDisk states

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

Cache modes

You can select to have read and write operations stored in cache by specifying a cache mode. You must specify the cache mode when you create the VDisk. After the VDisk is created, you cannot change the cache mode.

Table 8 describes the two types of cache modes for a VDisk.

Table 8. VDisk cache modes

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

Virtual disk mirroring

Virtual disk mirroring allows a VDisk to have two physical copies. Each VDisk copy can belong to a different managed disk (MDisk) group, and each copy has the same virtual capacity as the VDisk.

When a server writes to a mirrored VDisk, the SAN Volume Controller cluster writes the data to both copies. When a server reads a mirrored VDisk, the SAN Volume Controller cluster picks one of the copies to read. If one of the mirrored VDisk copies is temporarily unavailable; for example, because the RAID controller that provides the MDisk group is unavailable, the VDisk remains accessible to servers. The SAN Volume Controller cluster remembers which areas of the VDisk are written and resynchronizes these areas when both copies are available.

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

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

The VDisk copy can be any type: image, striped, sequential, and space-efficient or not. The two copies can be of completely different types.

VDisk Mirroring can be used for the following applications:

- Improving availability of VDisks by protecting them from a single storage controller failure.
- Allowing concurrent maintenance of a storage controller that does not natively support concurrent maintenance.
- Providing an alternative method of data migration with better availability characteristics. While a VDisk is being migrated using the data migration feature, it is vulnerable to failures on both the source and target MDisk group. VDisk Mirroring provides an alternative because you can start with a non-mirrored VDisk in the source MDisk group and then add a copy to that VDisk in the destination MDisk group. When the VDisk is synchronized, you

can delete the original copy that is in the source MDisk group. During the synchronization process, the VDisk remains available even if there is a problem with the destination MDisk group.

- Maintaining access to data that is stored on SSDs if one of the nodes in an I/O group is being serviced or fails.
- Converting between fully allocated VDisks and space-efficient VDisks.

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

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

Space-efficient virtual disks

When you create a virtual disk (VDisk), you can designate it as space-efficient. A space-efficient VDisk has a virtual capacity and a real capacity.

Virtual capacity is the VDisk storage capacity that is available to a host. *Real capacity* is the storage capacity that is allocated to a VDisk copy from a managed disk (MDisk) group. In a fully allocated VDisk, the virtual capacity and real capacity are the same. In a space-efficient VDisk, however, the virtual capacity can be much larger than the real capacity.

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

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

Space-efficient VDisks can also simplify server administration. Instead of assigning a VDisk with some capacity to an application and increasing that capacity as the application's needs change, you can configure a VDisk with a large virtual capacity for the application and then increase or shrink the real capacity as the application needs change, without disrupting the application or server.

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

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

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

When you create a space-efficient VDisk, you can choose the grain size for allocating space in 32 KB, 64 KB, 128 KB, or 256 KB chunks. Generally, smaller grain sizes save space but require more metadata access, which can adversely impact performance. If you are not going to use the space-efficient VDisk as a FlashCopy source or target VDisk, use 256 KB to maximize performance. If you are going to use the space-efficient VDisk as a FlashCopy source or target VDisk, specify the same grain size for the VDisk and for the FlashCopy feature.

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

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

Data migration

SAN Volume Controller allows you to migrate data across MDisks without interfering with any host applications that are simultaneously accessing or writing data.

During data migration, host applications can continue to access and use VDisk data as if no data migration were occurring.

Data migration applications

SAN Volume Controller data migration has the following characteristics and applications:

- Data migration changes the MDisks to which the VDisk is allocated.
- Data migration does not disrupt access to the VDisk data.
- Data migration can be used when removing MDisks from a cluster.
- Data migration can move data to an MDisk that has appropriate performance for host requirements.
- Data migration can move data from image mode VDisks to MDisks when adding SAN Volume Controller to an existing SAN infrastructure.
- 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 back up 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 back up the cluster configuration settings to tertiary storage. You can use the configuration backup functions to back up the cluster configuration.

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

Copy Services features

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

The following Copy Services 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 Services feature that is available with the SAN Volume Controller.

In its basic mode, 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.

More advanced FlashCopy features allow operations to 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. 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.

FlashCopy associates a source VDisk and a target VDisk in a FlashCopy Mapping. The source VDisks and target VDisks must meet the following requirements:

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

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

The incremental FlashCopy feature reduces the amount of time that is required to copy the source VDisk for multiple FlashCopy mappings. The initial FlashCopy mapping copies all of the data from the source VDisk to the target VDisk. Subsequent FlashCopy mappings only copy data that has been modified since the initial FlashCopy mapping. You can define a FlashCopy mapping as incremental only when you create the FlashCopy mapping.

Note: For incremental FlashCopy support information, see the release-specific *IBM System Storage SAN Volume Controller Restrictions* technical note at the Support for SAN Volume Controller (2145) Web site: www.ibm.com/storage/support/2145

The multi-target reverse FlashCopy feature allows you to start a FlashCopy mapping where the target VDisk is the source VDisk in a second FlashCopy mapping. You can use this feature to reverse a FlashCopy mapping direction without having to remove existing mappings, and without losing data from the target VDisk.

A FlashCopy mapping can also be created to mirror an existing mapping; these paired mappings are called *partners*. A mapping can have only one partnership. For example, VDisks A and B can have two mappings; A > B and B > A.

Any VDisk that is part of a FlashCopy operation can be space-efficient. Using a space-efficient VDisk as a FlashCopy target and setting the background FlashCopy rate to 0 (nocopy) can reduce the amount of storage that is required to maintain a point-in-time copy. The source VDisks and target VDisks can also be mirrored to improve availability of the VDisks.

Metro Mirror and Global Mirror

The Metro Mirror and Global Mirror Copy Services features 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. The VDisks can be in the same cluster or on two different clusters.

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 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.

The Metro Mirror feature 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.

The Global Mirror feature 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. 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.

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

- 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 participate in active Metro Mirror and Global Mirror relationships with itself and up to three other clusters.
- Intercluster and intracluster Metro Mirror and Global Mirror relationships can be used concurrently within a cluster.
- The intercluster link is bidirectional. This means that it can copy data from cluster A to cluster B for one pair of VDisks while copying data from cluster B to cluster A 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.
- SAN Volume Controller supports a maximum of 8192 Metro Mirror and Global Mirror relationships per cluster.

License settings

SAN Volume Controller is available with two mutually exclusive licenses.

SAN Volume Controller is available with either of the following licenses:

- A SAN Volume Controller capacity license grants you a number of terabytes (TB) for use with virtualization and the FlashCopy and Metro Mirror and Global Mirror and features.
- A SAN Volume Controller physical disk license is based on a number of physical disks and the optional FlashCopy and Metro Mirror and Global Mirror features.

User roles

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

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

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

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

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

Security Administrator
Users with the security administrator role can access all functions on the SAN Volume Controller Console and issue any CLI command. Users with this role can also manage users, user groups, and manage user authentication.

Configuring user authentication

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

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

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

Chapter 2. SAN Volume Controller physical installation planning

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, and optional redundant ac-power switch installation are met.

This information considers the hardware components that are currently supported.

1. Does your physical site meet the environment requirements for the SAN Volume Controller, uninterruptible power supply, 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 Electronic Industries Alliance (EIA) unit high for each node.
   - The 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 the power circuits you will be using have sufficient capacity and the correct sockets for your SAN Volume Controller installation?

   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. Have you provided appropriate connectivity by preparing your environment?
5. Do you have a keyboard and display available in the unusual event that a service action requires them? The SAN Volume Controller 2145-CF8 , SAN Volume Controller 2145-8A4 and SAN Volume Controller 2145-8G4 nodes require a Universal Serial Bus (USB) keyboard.
6. Do you have adequate space available and power for the IBM System Storage Productivity Center?

Contamination information

This topic identifies contamination information that pertains to this product.

Airborne particulates (including metal flakes or particles) and reactive gases acting alone or in combination with other environmental factors, such as humidity or temperature, might pose a risk to the SAN Volume Controller hardware. Risks that are posed by the presence of excessive particulate levels or concentrations of harmful gases include damage that might cause the SAN Volume Controller hardware to malfunction or cease functioning altogether. This specification sets forth limits for particulates and gases that are intended to avoid such damage. The limits must not be viewed or used as definitive limits because numerous other factors, such as temperature or moisture content of the air, can influence the impact of particulates or environmental corrosives and gaseous contaminant transfer.

In the absence of specific limits that are set forth in this document, you must implement practices that maintain particulate or gas levels that are consistent with the protection of human health and safety. If IBM determines that the levels of particulates or gases in your environment have caused damage to the SAN Volume Controller hardware, IBM might require implementation of appropriate remedial measures to mitigate such environmental contamination before providing repair or replacement of the SAN Volume Controller hardware. Implementation of such remedial measures is a customer responsibility.

The following criteria must be met:

Gaseous contamination

Severity level G1 as per ANSI/ISA 71.04-1985[1], which states that the reactivity rate of copper coupons shall be less than 300 Angstroms per month (Å/month, ≈ 0.0039 µg/cm2-hour weight gain)[2]. In addition, the reactivity rate of silver coupons shall be less than 300Å/month (≈ 0.0035 µg/cm2-hour weight gain)[3]. The reactive monitoring of gaseous corrosivity should be conducted approximately 2 inches (5 cm) in front of the rack on the air inlet side at one-quarter and three-quarter frame height off the floor, or where the air velocity is much higher.

Particulate contamination

Data centers must meet the cleanliness level of ISO 14644-1 class 8. For data centers without airside economizers, the ISO 14644-1 class 8 cleanliness can be met by choosing one of the following filtration methods:

- The room air can be continuously filtered with MERV 8 filters.
- Air entering a data center can be filtered with MERV 11, or preferably MERV 13 filters.

For data centers with airside economizers, the choice of filters to achieve ISO class 8 cleanliness depends on the specific conditions present at that data center. The deliquescent relative humidity of the particulate contamination should be more than 60% RH[4]. Data centers must be free of zinc whiskers[5].

1. ANSI/ISA-71.04.1985. *Environmental conditions for process measurement and control systems: Airborne contaminants.* Instrument Society of America, Research Triangle Park, NC, 1985.
2. The derivation of the equivalence between the rate of copper corrosion product thickness growth in Å/month and the rate of weight gain assumes that Cu_2S and Cu_2O grow in equal proportions.
3. The derivation of the equivalence between the rate of silver corrosion product thickness growth in Å/month and the rate of weight gain assumes that Ag_2S is the only corrosion product.
4. The deliquescent relative humidity of particulate contamination is the relative humidity at which the dust absorbs enough water to become wet and promote corrosion, ion migration, or both.
5. Surface debris is randomly collected from 10 areas of the data center on a 1.5 cm diameter disk of sticky, electrically conductive tape on a metal stub. If examination of the sticky tape in a scanning electron microscope reveals no zinc whiskers, the data center is considered free of zinc whiskers.

SAN Volume Controller 2145-CF8 environment requirements

Before installing a SAN Volume Controller 2145-CF8 node, your physical environment must meet certain requirements. This includes verifying that adequate space is available and that requirements for power and environmental conditions are met.

Input-voltage requirements

Ensure that your environment meets the following voltage requirements.

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

Attention:

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

Power requirements for each node

Ensure that your environment meets the following power requirements.

The power capacity that is required depends on the node type and which optional features are installed.

Components	Power requirements
SAN Volume Controller 2145-CF8 node and 2145 UPS-1U power supply	200 W

Notes:

- SAN Volume Controller 2145-CF8 nodes will not connect to all revisions of the 2145 UPS-1U power supply unit. The SAN Volume Controller 2145-CF8 nodes require the 2145 UPS-1U power supply unit part number 31P1318. This unit has two power outlets that are accessible. Earlier revisions of the 2145 UPS-1U power supply unit have only one power outlet that is accessible and are not suitable.
- For each redundant ac-power switch, add 20 W to the power requirements.
- For each high-speed SAS adapter with one to four solid-state drives (SSDs), add 50 W to the power requirements.

Circuit breaker requirements

The 2145 UPS-1U has an integrated circuit breaker and does not require additional protection.

Environment requirements without redundant ac power

Ensure that your environment falls within the following ranges if you are not using 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 (73°F)
Operating in higher altitudes	10°C to 32°C (50°F to 90°F)	914 to 2133 m (2998 to 6988 ft)	8% to 80% noncondensing	23°C (73°F)
Turned off	10°C to 43°C (50°F to 110°F)	0 to 2133 m (0 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

Ensure that your environment falls within the following ranges if you are using redundant ac power.

Environment	Temperature	Altitude	Relative humidity	Maximum wet bulb temperature
Operating in lower altitudes	15°C to 32°C (59°F to 90°F)	0 to 914 m (0 to 2998 ft)	20% to 80% noncondensing	23°C (73°F)
Operating in higher altitudes	15°C to 32°C (59°F to 90°F)	914 to 2133 m (2998 to 6988 ft)	20% to 80% noncondensing	23°C (73°F)
Turned 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 environment

The following tables list the physical characteristics of the SAN Volume Controller 2145-CF8 node.

Dimensions and weight

Ensure that space is available in a rack that is capable of supporting the node.

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

Ensure that space is also available in the rack for the following additional space requirements around the node.

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 of each SAN Volume Controller 2145-CF8 node

The node dissipates the following maximum heat output.

Model	Heat output per node
SAN Volume Controller 2145-CF8	160 W (546 Btu per hour)
SAN Volume Controller 2145-CF8 and up to four optional solid-state drives (SSDs)	210 W (717 Btu per hour)
Maximum heat output of 2145 UPS-1U during typical operation:	10 W (34 Btu per hour)
Maximum heat output of 2145 UPS-1U during battery operation:	100 W (341 Btu per hour)

SAN Volume Controller 2145-8A4 environment requirements

Before the SAN Volume Controller 2145-8A4 is installed, the physical environment must meet certain requirements. This includes verifying that adequate space is available and that requirements for power and environmental conditions are met.

Input-voltage requirements

Ensure that your environment meets the following voltage requirements.

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

Attention:

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

Power requirements for each node

Ensure that your environment meets the following power requirements.

The power that is required depends on the node type and whether the redundant ac power feature is used.

Components	Power requirements
SAN Volume Controller 2145-8A4 and 2145 UPS-1U	180 W

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

Circuit breaker requirements

The 2145 UPS-1U has an integrated circuit breaker and does not require additional protection.

Environment requirements without redundant ac power

Ensure that your environment falls within the following ranges if you are not using 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 3000 ft)	8% to 80% noncondensing	23°C (73°F)
Operating in higher altitudes	10°C to 32°C (50°F to 90°F)	914 to 2133 m (3000 to 7000 ft)	8% to 80% noncondensing	23°C (73°F)
Turned off	10°C to 43°C (50°F to 109°F)	0 to 2133 m (0 to 7000 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 7000 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

Ensure that your environment falls within the following ranges if you are using redundant ac power.

Environment	Temperature	Altitude	Relative humidity	Maximum wet bulb temperature
Operating in lower altitudes	15°C to 32°C (59°F to 90°F)	0 to 914 m (0 to 3000 ft)	20% to 80% noncondensing	23°C (73°F)
Operating in higher altitudes	15°C to 32°C (59°F to 90°F)	914 to 2133 m (3000 to 7000 ft)	20% to 80% noncondensing	23°C (73°F)
Turned off	10°C to 43°C (50°F to 109°F)	0 to 2133 m (0 to 7000 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 7000 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 environment

The following tables list the physical characteristics of the SAN Volume Controller 2145-8A4 node.

Dimensions and weight

Ensure that space is available in a rack that is capable of supporting the node.

Height	Width	Depth	Maximum weight
43 mm (1.75 in.)	440 mm (17.32 in.)	559 mm (22 in.)	10.1 kg (22 lb)

Additional space requirements

Ensure that space is also available in the rack for the following additional space requirements around the node.

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

Heat output of each SAN Volume Controller 2145-8A4 node

The node dissipates the following maximum heat output.

Model	Heat output per node
SAN Volume Controller 2145-8A4	140 W (478 Btu per hour)

SAN Volume Controller 2145-8G4 environment requirements

Before the SAN Volume Controller 2145-8G4 is installed, the physical environment must meet certain requirements. This includes verifying that adequate space is available and that requirements for power and environmental conditions are met.

Input-voltage requirements

Ensure that your environment meets the following voltage requirements.

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

Attention:

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

Power requirements for each node

Ensure that your environment meets the following power requirements.

The power that is required depends on the node type and whether the redundant ac power feature is used.

Components	Power requirements
SAN Volume Controller 2145-8G4 and 2145 UPS-1U	470 W

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

Circuit breaker requirements

The 2145 UPS-1U has an integrated circuit breaker and does not require additional protection.

Environment requirements without redundant ac power

Ensure that your environment falls within the following ranges if you are not using 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 (73°F)
Operating in higher altitudes	10°C to 32°C (50°F to 90°F)	914 to 2133 m (2998 to 6988 ft)	8% to 80% noncondensing	23°C (73°F)
Turned off	10°C to 43°C (50°F to 110°F)	0 to 2133 m (0 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

Ensure that your environment falls within the following ranges if you are using redundant ac power.

Environment	Temperature	Altitude	Relative humidity	Maximum wet bulb temperature
Operating in lower altitudes	15°C to 32°C (59°F to 90°F)	0 to 914 m (0 to 2998 ft)	20% to 80% noncondensing	23°C (73°F)
Operating in higher altitudes	15°C to 32°C (59°F to 90°F)	914 to 2133 m (2998 to 6988 ft)	20% to 80% noncondensing	23°C (73°F)
Turned 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 environment

The following tables list the physical characteristics of the SAN Volume Controller 2145-8G4 node.

Dimensions and weight

Ensure that space is available in a rack that is capable of supporting the node.

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

Ensure that space is also available in the rack for the following additional space requirements around the node.

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 of each SAN Volume Controller 2145-8G4 node

The node dissipates the following maximum heat output.

Model	Heat output per node
SAN Volume Controller 2145-8G4	400 W (1350 Btu per hour)

SAN Volume Controller 2145-8F4 and SAN Volume Controller 2145-8F2 environment requirements

Before the SAN Volume Controller 2145-8F4 or SAN Volume Controller 2145-8F2 is installed, the physical environment must meet certain requirements. This includes verifying that adequate space is available and that requirements for power and environmental conditions are met.

Input-voltage requirements

Ensure that your environment meets the following voltage requirements.

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

Power requirements for each node

Ensure that your environment meets the following power requirements.

The power that is required depends on the node type and whether the redundant ac power feature is used.

Components	Power requirements
SAN Volume Controller 2145-8F4 and 2145 UPS-1U	520 W
SAN Volume Controller 2145-8F2 and 2145 UPS-1U	520 W

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

Circuit breaker requirements

The 2145 UPS-1U has an integrated circuit breaker and does not require additional protection.

Environment requirements without redundant ac power

Ensure that your environment falls within the following ranges if you are not using 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 914.4 m (0 to 3000 ft)	8% to 80% noncondensing	23°C (74°F)
Operating in higher altitudes	10°C to 32°C (50°F to 88°F)	914.4 to 2133.6 m (3000 to 7000 ft)	8% to 80% noncondensing	23°C (74°F)
Turned off	10°C to 43°C (50°F to 110°F)	0 to 2133.6 m (3000 to 7000 ft)	8% to 80% noncondensing	27°C (81°F)
Storing	1°C to 60°C (34°F to 140°F)	0 to 2133.6 m (0 to 7000 ft)	5% to 80% noncondensing	29°C (84°F)

Environment	Temperature	Altitude	Relative humidity	Maximum wet bulb temperature
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

Ensure that your environment falls within the following ranges if you are using 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 914.4 m (0 to 3000 ft)	20% to 80% noncondensing	23°C (74°F)
Operating in higher altitudes	15°C to 32°C (50°F to 88°F)	914.4 to 2133.6 m (3000 to 7000 ft)	20% to 80% noncondensing	23°C (74°F)
Turned off	10°C to 43°C (50°F to 110°F)	0 to 2133.6 m (0 to 7000 ft)	20% to 80% noncondensing	27°C (81°F)
Storing	1°C to 60°C (34°F to 140°F)	0 to 2133.6 m (0 to 7000 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 environment

The following tables list the physical characteristics of the SAN Volume Controller 2145-8F4 and SAN Volume Controller 2145-8F2 nodes.

Dimensions and weight

Ensure that space is available in a rack that is capable of supporting the node.

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

Ensure that space is also available in the rack for the following additional space requirements around the node.

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 of each SAN Volume Controller 2145-8F4 or SAN Volume Controller 2145-8F2 node

The nodes dissipate the following maximum heat output.

Model	Heat output per node
SAN Volume Controller 2145-8F4	450 W (1540 Btu per hour)
SAN Volume Controller 2145-8F2	450 W (1540 Btu per hour)

Uninterruptible power-supply environment requirements

An uninterruptible power-supply environment requires that certain specifications for the physical site of the SAN Volume Controller must be met.

2145 UPS-1U environment

All SAN Volume Controller models are supported with the 2145 UPS-1U.

2145 UPS-1U specifications

The following tables describe the physical characteristics of the 2145 UPS-1U.

2145 UPS-1U dimensions and weight

Ensure that space is available in a rack that is capable of supporting the 2145 UPS-1U.

Height	Width	Depth	Maximum weight
44 mm (1.73 in.)	439 mm (17.3 in.)	579 mm (22.8 in.)	16 kg (35.3 lb)
Note: The 2145 UPS-1U package, which includes support rails, weighs 18.8 kg (41.4 lb).			

Heat output

The 2145 UPS-1U unit produces the following approximate heat output.

Model	Heat output during normal operation	Heat output during battery operation
2145 UPS-1U	10 W (34 Btu per hour)	150 W (512 Btu per hour)

Power cables for the 2145 UPS-1U

If you do not connect the 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 following table lists the power cable options for your country or region.

Country or region	Length	Connection type (attached plug designed for 200 - 240 V ac input)	Part number
United States of America (Chicago), Canada, Mexico	1.8 m (6 ft)	NEMA L6-15P	39M5115
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	39M5116
Afghanistan, Algeria, Andorra, Angola, Austria, Belgium, Benin, Bulgaria, Burkina Faso, Burundi, Cameroon, Central African Republic, Chad, Czech Republic, Egypt, Finland, France, French Guiana, Germany, Greece, Guinea, Hungary, Iceland, Indonesia, Iran, Ivory Coast, Jordan, Lebanon, Luxembourg, Macao S.A.R. of China, 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	39M5123
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	39M5151
Argentina	2.8 m (9 ft)	IRAM 2073	39M5068
Argentina, Australia, New Zealand, Papua New Guinea, Paraguay, Uruguay, Western Samoa	2.8 m (9 ft)	AS/NZS 3112/2000	39M5102
Bangladesh, Burma, Pakistan, South Africa, Sri Lanka	2.8 m (9 ft)	SABS 164	39M5144
Chile, Ethiopia, Italy, Libya, Somalia	2.8 m (9 ft)	CEI 23-16	39M5165
People's Republic of China	2.8 m (9 ft)	GB 2099.1	39M5206

Country or region	Length	Connection type (attached plug designed for 200 - 240 V ac input)	Part number
Denmark	2.8 m (9 ft)	DK2-5a	39M5130
Israel	2.8 m (9 ft)	SI 32	39M5172
Liechtenstein, Switzerland	2.8 m (9 ft)	IEC 60884 Stnd. Sheet 416534?2 (CH Type 12)	39M5158
Thailand	2.8 m (9 ft)	NEMA 6-15P	39M5095

Redundant ac-power environment requirements

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 that are provided through two rack-mounted power distribution units (PDUs). The PDUs must have IEC320-C13 outlets.

The redundant ac-power switch comes with two IEC 320-C19 to C14 power cables to connect to rack PDUs. There are no country-specific cables for the redundant ac-power switch.

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

Redundant ac-power switch specifications

The following tables list the physical characteristics of the redundant ac-power switch.

Dimensions and weight

Ensure that space is available in a rack that is capable of supporting the redundant ac-power switch.

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

Ensure that space is also available in the rack for the side mounting plates on either side of the redundant ac-power switch.

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 that is dissipated inside the redundant ac-power switch is approximately 20 watts (70 Btu per hour).

Connections

Ensure that you are familiar with the specific connection types for the SAN Volume Controller and the uninterruptible power supply.

Each SAN Volume Controller requires the following connections:

- Each SAN Volume Controller node requires one Ethernet cable to connect it to an Ethernet switch or hub. A 10/100/1000 Mb Ethernet connection is required. Both Internet Protocol Version 4 (IPv4) and Internet Protocol Version 6 (IPv6) are supported.

  Note: For increased redundancy, an optional second Ethernet connection is supported for each SAN Volume Controller node.
- To ensure cluster failover operations, Ethernet port 1 on all nodes must be connected to the same IP subnet or subnets.
- If used, Ethernet port 2 on all nodes must also be connected to the same IP subnet or subnets. However, the subnet or subnets that are connected to Ethernet port 2 do not have to be the same as the subnet or subnets that are connected to Ethernet port 1.
- 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 uninterruptible power supply requires a serial cable that connects the uninterruptible power supply to the SAN Volume Controller node. Ensure that for each node, the serial and power cables come from the same uninterruptible power supply.

TCP/IP requirements for SAN Volume Controller

To plan your installation, consider the TCP/IP address requirements of the SAN Volume Controller cluster and the requirements for the SAN Volume Controller to access other services. You must also plan address allocation and Ethernet router, gateway and firewall configuration to provide the required access and network security.

Figure 6 on page 42 shows the TCP/IP ports and services that are used by SAN Volume Controller.

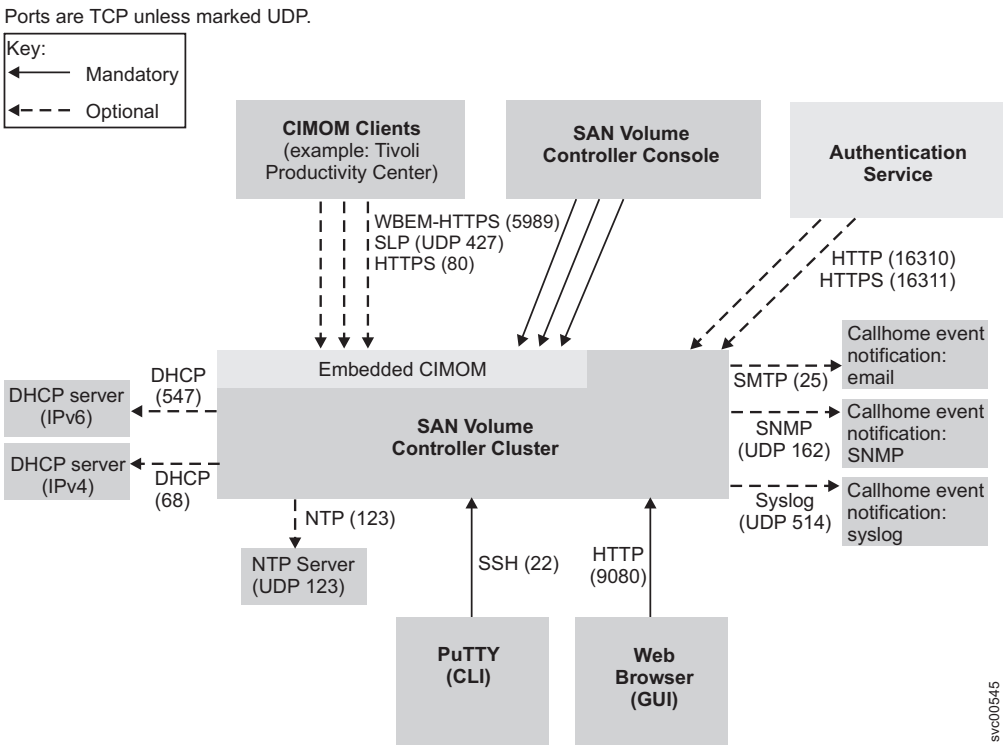

Figure 6. TCP/IP ports and services

Both IPv4 and IPv6 addresses are supported. SAN Volume Controller can operate with either internet protocol or with both internet protocols concurrently.

For configuration and management, you must allocate an IP address to the cluster which is often referred to as the *cluster IP address*. For additional fault tolerance, you can also configure a second cluster IP address for the second ethernet port on the node. The addresses must be fixed addresses. If both IPv4 and IPv6 are operating concurrently, an address is required for each protocol.

In addition to the cluster IP addresses, the cluster is configured with one or more service IP address. Service mode addresses are used to gain access to a node that is not operating as part of the cluster for any reason. As with the cluster IP addresses, service mode IP addresses can be IPv4 addresses, IPv6 addresses, or both. You can choose a fixed IP address per ethernet port, or you can configure SAN Volume Controller to use a DHCP-allocated address.

The IBM System Storage Productivity Center that manages the cluster must be able to access the cluster IP addresses.

If the SAN Volume Controller console is used on a system other than the IBM System Storage Productivity Center, the system that it is run on must be able to access the IBM System Storage Productivity Center and the cluster itself.

Any system that is used to manage a node in service mode must be able to access a cluster service mode address, or a dynamically-allocated service mode address if DHCP is used.

You can configure SAN Volume Controller for event notification by SNMP, syslog or e-mail. To configure notification, you must ensure that the SNMP agent, syslog IP addresses or SMTP e-mail server IP addresses can be accessed from all SAN Volume Controller cluster addresses.

SAN Volume Controller does not use name servers to locate other devices. You must supply the numeric IP address of the device. To locate a device, the device must have a fixed IP address.

The nodes in a SAN Volume Controller cluster can receive native IPv4 or IPv6 packets. The nodes cannot operate as an endpoint of a tunnel where an IPv4 packet contains an IPv6 payload or where an IPv6 packet contains an IPv4 payload.

Preparing your IBM System Storage Productivity Center environment

Before an IBM service representative can install the IBM System Storage Productivity Center, you must prepare the physical environment. This includes verifying that adequate space is available and that requirements for power and environmental conditions are met.

For information about setting up the physical environment, see the *IBM System Storage Productivity Center Introduction and Planning Guide*.

Web browser requirements to access the SAN Volume Controller Console

Although an IBM System Storage Productivity Center installation includes a Web browser, other browsers are also supported. You must have a supported 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 access the SAN Volume Controller Console directly or from a remote system.
- The operating system from which you access the SAN Volume Controller Console.

For a list of the supported browsers, see the *Recommended Software Levels* document at the following Web site:

www.ibm.com/storage/support/2145

Physical configuration planning of a system with SAN Volume Controller

Before you or your IBM service representative installs the SAN Volume Controller nodes, uninterruptible power supply units, and IBM System Storage Productivity Center, you must plan the physical configuration and the initial settings for the system.

1. Download the hardware location chart, the cable connection table, the configuration data table, and the optional redundant ac power switch connection chart from the following Web site:

   www.ibm.com/storage/support/2145

a. Click the **Plan/upgrade** tab.
b. Under **Sizing/Capacity**, click **SAN Volume Controller planning**.
c. Click your language to view the Planning Guide abstract. The charts and tables are under the **Related information** heading.

2. Use the hardware location chart to record the physical configuration of your system.
3. Use the cable connection table to record how you want your SAN Volume Controller units, the uninterruptible power supply units, and the IBM System Storage Productivity Center 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.

Requirements and guidelines for 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 Electronic Industries Alliance (EIA) 19-inch rack space.

When you design the rack for the SAN Volume Controller, use the hardware location chart to record the physical configuration of your system.

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

www.ibm.com/storage/support/2145

To select the hardware location chart, follow these steps:

1. Under **Support & downloads**, click **Plan & upgrades**.
2. Under **Sizing/Capacity**, click **SAN Volume Controller planning**.
3. Click your language to view the Planning Guide abstract.
4. Under **Related information**, click **SAN Volume Controller - Hardware Location Chart** and save the file.

When you complete the information in the hardware location chart, consider the following requirements:

- Do not exceed the maximum power rating of the rack and input power supply.
- 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 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.
- The IBM System Storage Productivity Center hardware is two EIA units high: one EIA unit for the server and one EIA unit for the keyboard and monitor.
- The depth of the SAN Volume Controller 2145-8A4 node is less than other components and nodes by approximately 127 mm or 5 inches. Do not locate the 2145-8A4 node in the rack between components or nodes of greater depths;

otherwise, it will not be possible to attach cables to the 2145-8A4 node. The same restriction applies to the 2145 UPS-1U, which is of comparable depth.

- If there are any hardware devices already in the rack, record this information on the chart.
- Fill in rows for all other units that will be 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.
- A cluster can contain no more than eight SAN Volume Controller nodes.
- Install the SAN Volume Controller nodes in pairs to provide redundancy and concurrent maintenance.
- 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.

In addition to the requirements for completing your information in the hardware location chart, consider also the following guidelines and recommendations.

- Position SAN Volume Controller nodes so that information on the display screen is easily viewed and so that the controls that are used to navigate the display menu are easily reached. Place SAN Volume Controller nodes in EIA 11-36.
- To view the front panels of the SAN Volume Controller node while you perform maintenance procedures, place the IBM System Storage Productivity Center in close proximity to the SAN Volume Controller cluster. To ensure that the IBM System Storage Productivity Center is at a convenient height for viewing the monitor and using the keyboard, place the IBM System Storage Productivity Center in EIA 17-24. Position the IBM System Storage Productivity Center server, the keyboard, and the monitor unit close to each other. Ensure you allow access to the CD drive that is located in the server while the keyboard and monitor unit are pulled out and are open. To allow the monitor to open, do not place the keyboard and monitor unit directly below a SAN Volume Controller node.
- To reduce the chance of a simultaneous input power failure at both uninterruptible power supply units, either use the redundant ac-power switch or connect each uninterruptible power supply unit to a separate electrical power source on a separate branch circuit.

Requirements for 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:

www.ibm.com/storage/support/2145

To select the cable connection table, follow these steps:

1. Under **Support & downloads**, click **Plan & upgrades**.
2. Under **Sizing/Capacity**, click **SAN Volume Controller planning**.
3. Click your language to view the Planning Guide abstract.
4. Under **Related information**, click **SAN Volume Controller - Cable Connection Table** and save the file.

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

Term	Description
Uninterruptible power supply	The uninterruptible power supply to which the SAN Volume Controller node is connected. The uninterruptible power supply connects to the node through a serial cable and one or two power cables depending on the node type. Each uninterruptible power supply must have connections only to a single SAN Volume Controller node (and no other devices), and each node must only have connections to a single uninterruptible power supply, even if the node has two power supplies.
Ethernet	The Ethernet hub or switch to which the SAN Volume Controller node is connected. The Ethernet hub or switch must have a connection to the IBM System Storage Productivity Center. If you need to generate e-mail reports, there must be a connection to an SMTP server.
Fibre-channel ports 1 through 4	The fibre-channel switch ports to which the four fibre-channel ports of the SAN Volume Controller node are connected. When you view it from the back of the SAN Volume Controller node, the ports are numbered 1 through 4, from left to right.

For the IBM System Storage Productivity Center hardware, 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 IBM System Storage Productivity Center and an Ethernet switch or router.
Ethernet to Internet for remote service or upgrade procedures	(Optional) The Ethernet port used to connect the IBM System Storage Productivity Center to the Internet for remote service or upgrade. You can determine how you want to provide the port. For example, you can provide the setup 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 IBM System Storage Productivity Center with any access to the public Internet and allow the IBM Assist On-site tool to access another console that can access the IBM System Storage Productivity Center Web server over an intranet.

Configuration data table guidelines

You must fill out the configuration data table before the SAN Volume Controller nodes and IBM System Storage Productivity Center can be installed.

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

www.ibm.com/storage/support/2145

To select the configuration data table, follow these steps:

1. Click the **Plan/upgrade** tab.

2. Under **Sizing/Capacity**, click **SAN Volume Controller planning**.
3. Click your language to view the Planning Guide abstract.
4. Under **Related information**, click **SAN Volume Controller - Configuration Data Table**.

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. The default setting is English.
Port 1	
Cluster IPv4 address	IPv4 address used for typical configuration and service access to the cluster
Service IPv4 address	IPv4 address used for service mode access to a node from the cluster
Gateway IPv4 address	IPv4 address for the default local gateway for the cluster
IPv4 subnet mask	IPv4 subnet mask, which identifies the IPv4 network the cluster is operating on
Cluster IPv6 address	IPv6 address used for typical configuration and service access to the cluster
Service IPv6 address	IPv6 address used for service mode access to a node from the cluster
Gateway IPv6 address	IPv6 address for the default local gateway for the cluster
IPv6 prefix	The cluster prefix, which identifies the IPv6 network the cluster is operating on
Port 2 (optional)	
Cluster IPv4 address	IPv4 address used for typical configuration and service access to the cluster
Service IPv4 address	IPv4 address used for service mode access to a node from the cluster
Gateway IPv4 address	IPv4 address for the default local gateway for the cluster
IPv4 subnet mask	IPv4 subnet mask, which identifies the IPv4 network the cluster is operating on
Cluster IPv6 address	IPv6 address used for typical configuration and service access to the cluster
Service IPv6 address	IPv6 address used for service mode access to a node from the cluster
Gateway IPv6 address	IPv6 address for the default local gateway for the cluster
IPv6 prefix	The cluster prefix, which identifies the IPv6 network the cluster is operating on

Include the following information for the IBM System Storage Productivity Center.

Term	Description
Machine name	The name of the IBM System Storage Productivity Center. This must be a fully qualified DNS name.
IBM System Storage Productivity Center IP addresses	The addresses that are used for access to the IBM System Storage Productivity Center.
IBM System Storage Productivity Center gateway IP address	The IP address for the local gateway for the IBM System Storage Productivity Center.
IBM System Storage Productivity Center subnet mask	The subnet mask for the IBM System Storage Productivity Center.

Requirements for 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:

www.ibm.com/storage/support/2145

To select the redundant ac-power switch connection chart, follow these steps:

1. Click the **Plan/upgrade** tab.
2. Under **Sizing/Capacity**, click **SAN Volume Controller planning**.
3. Click your language to view the Planning Guide abstract.
4. Under **Related information**, click **SAN Volume Controller - Redundant AC Power Chart**.

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.

Cabling of redundant ac-power switch (example)

You must properly cable the redundant ac-power switch units in your environment.

Note: While this topic provides an example of the cable connections, it does not indicate a preferred physical location for the components.

Figure 7 on page 49 shows an example of the main wiring for a SAN Volume Controller cluster with the redundant ac-power switch feature. The four-node cluster consists of two I/O groups:

- I/O group 0 contains nodes A and B

- I/O group 1 contains nodes C and D

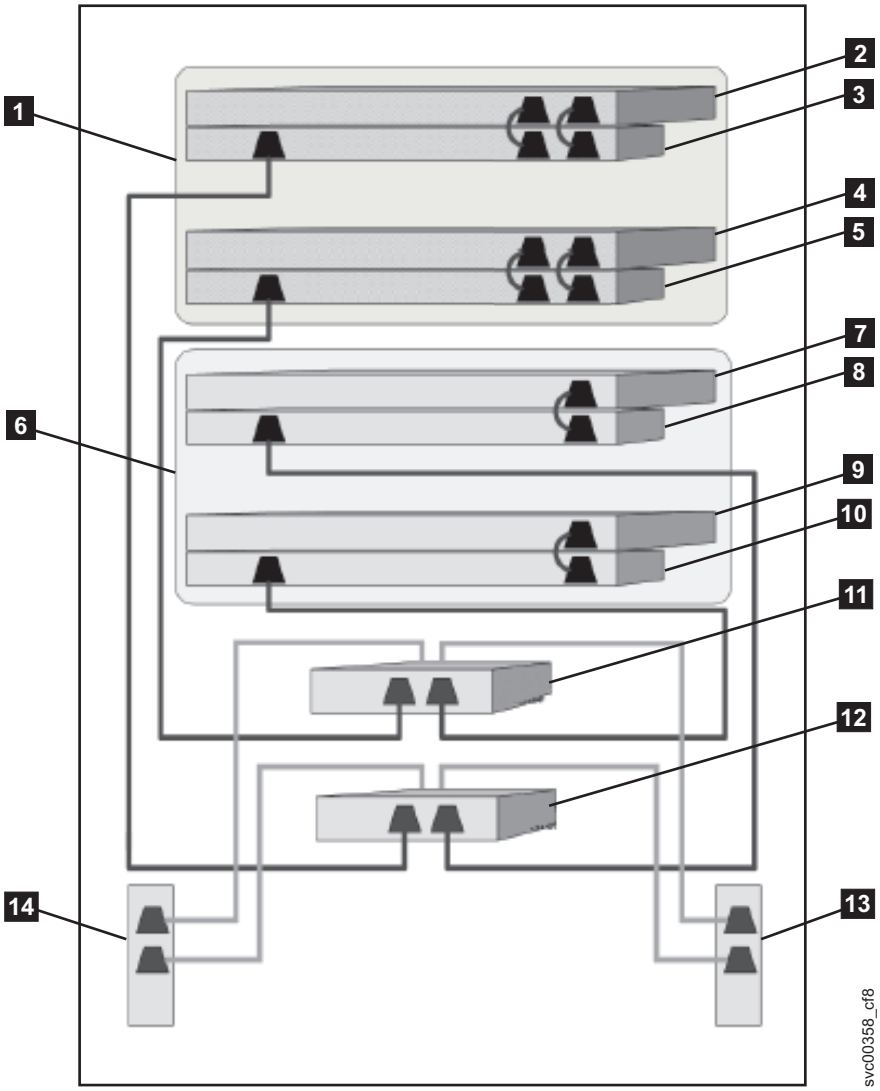

Figure 7. A four-node SAN Volume Controller cluster with the redundant ac-power switch feature

1. I/O group 0
2. SAN Volume Controller node A
3. 2145 UPS-1U A
4. SAN Volume Controller node B
5. 2145 UPS-1U B
6. I/O group 1
7. SAN Volume Controller node C
8. 2145 UPS-1U C
9. SAN Volume Controller node D
10. 2145 UPS-1U D
11. Redundant ac-power switch 1
12. Redundant ac-power switch 2

13 Site PDU X (C13 outlets)
14 Site PDU Y (C13 outlets)

The site PDUs X and Y (**13** and **14**) are powered from two independent power sources.

In this example, only two redundant ac-power switch units are used, and each power switch powers one node in each I/O group. However, for maximum redundancy, use one redundant ac-power switch to power each node in the cluster.

Some SAN Volume Controller node types have two power supply units. Both power supplies must be connected to the same 2145 UPS-1U, as shown by node A and node B. The SAN Volume Controller 2145-CF8 is an example of a node that has two power supplies. The SAN Volume Controller 2145-8A4 is an example of a node that has a single power supply.

Chapter 3. SAN fabric and LAN overview

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

SAN fabric overview

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

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

Figure 8 on page 52 shows several host systems that are connected in a SAN fabric.

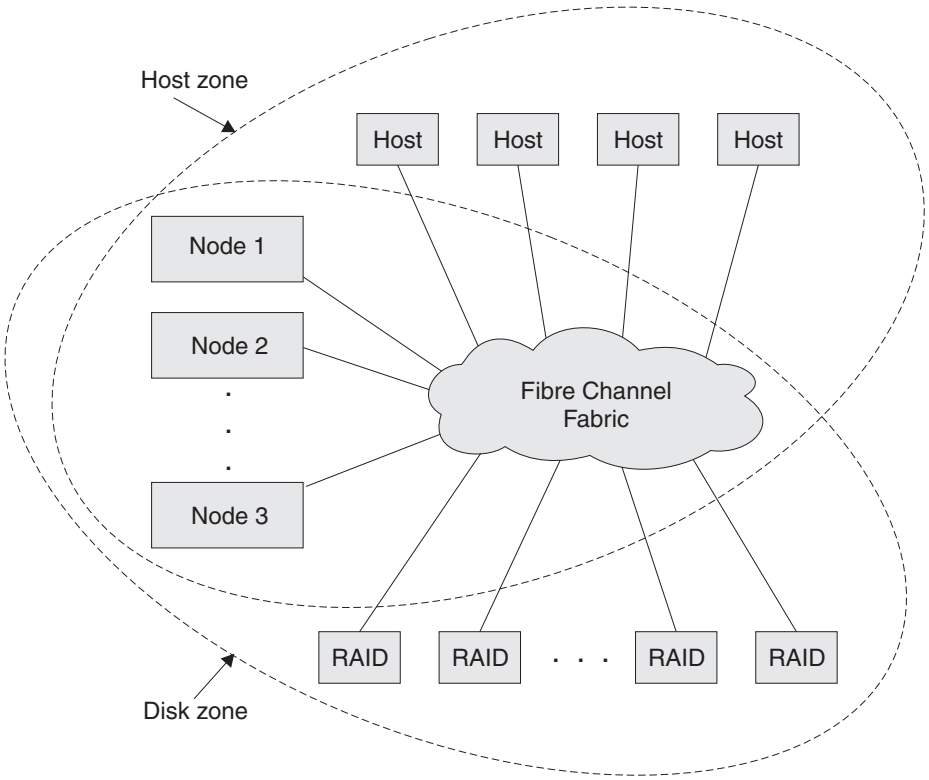

Figure 8. Example of a SAN Volume Controller cluster in a fabric

A cluster of SAN Volume Controller nodes is connected to the fibre-channel 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 systems (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. However, to optimize the use of resources, ensure that all MDisks in an MDisk group have similar performance characteristics.

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

SAN fabric and LAN configuration terms

Ensure that you understand the basic terms and definitions when you are configuring the SAN Volume Controller within the SAN fabric or a local area network (LAN).

Table 9 on page 53 provides terms and definitions that can guide your understanding of the SAN fabric rules and requirements.

Table 9. SAN fabric 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.
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 that there is a symmetrical network and a specific workload that is applied equally from all initiators and sent equally to all targets. A symmetrical network is one in which 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 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. You can create a redundant SAN by splitting the SAN into two independent counterpart SANs.
Counterpart SAN	A nonredundant 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 Mirror and Global Mirror, significant distances might exist between the components of the local cluster and those of the remote clusters.
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 clusters.
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 fibers 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 fail over to another port, and the fan in requirements might be exceeded in this degraded mode.

Table 9. SAN fabric configuration terms and definitions (continued)

Term	Definition
Not valid configuration	The current SAN configuration is not correct. An attempted operation failed and generated an error code that indicates what caused it to become "not valid." The most likely cause is that either a device has failed, or a device has been added to the SAN that has caused the configuration to be marked as not valid.
Unsupported configuration	A configuration that could operate successfully, but that IBM does not guarantee any solutions for if problems occur. Typically, this type of configuration does not create any error log entries.
Valid configuration	A configuration that consists of devices and connections that are identified as valid and supported. Neither of the following two conditions exist with the current configuration: • Not valid • Unsupported configuration
Degraded	A valid configuration that has had a failure, but continues to be valid and supported. Typically, a repair action is required to restore the degraded configuration to a valid configuration.
Fibre-channel extender	A device for long distance communication that connects other SAN fabric components. Generally these components 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 that are configured to create a larger switched network. With this configuration, four or more switches are connected in a loop with additional direct connections between nonadjacent switches within the loop. An example of this configuration is four switches that are connected in a loop with ISLs for one of the diagonals.

If you plan to use iSCSI to configure a LAN, ensure that you also understand the iSCSI terms and definitions. Table 10 highlights the terms and definitions for iSCSI.

Table 10. iSCSI configuration terms and definitions

Term	Definition
Challenge Handshake Authentication Protocol (CHAP)	An authentication protocol that protects against eavesdropping by encrypting the user name and password.
Clustered Ethernet port	A physical Ethernet port on a node in a cluster that contains configuration settings that are shared by all the ports in a cluster.
Extended-unique identifier (EUI)	A unique iSCSI name that identifies an iSCSI target adapter or an iSCSI initiator adapter as defined by the iSCSI standard (RFC 3722).
Host object	A logical object that represents a list of worldwide port names (WWPNs) and a list of iSCSI names that identify the interfaces that the host system uses to communicate with a device. iSCSI names can be either iSCSI qualified names (IQNs) or extended-unique identifiers (EUIs).
Host system	A computer that is connected to the SAN Volume Controller through either a fibre-channel interface or an IP network.
Initiator	The system component that originates an I/O command over an I/O bus or network. I/O adapters and network interface controllers are typical initiators.

Table 10. iSCSI configuration terms and definitions (continued)

Term	Definition
Internet Storage Name Service (iSNS) Protocol	A protocol that is used by a host system to manage iSCSI targets and iSCSI discovery. iSCSI initiators use the ISNS Protocol to locate the appropriate storage resources.
iSCSI alias	An alternative name for the iSCSI-attached host.
iSCSI name	A name that identifies an iSCSI target adapter or an iSCSI initiator adapter. An iSCSI name can be an iSCSI qualified name (IQN) or an extended-unique identifier (EUI). Typically, this identifier has the following format: iqn.datecode.reverse domain.
iSCSI qualified name (IQN)	A specific type of iSCSI name that identifies an iSCSI target adapter or an iSCSI initiator adapter as defined by the iSCSI standard (RFC 3722).
Network interface controller (NIC)	Hardware that provides the interface control between system main storage and external high-speed link (HSL) ports.
Node Ethernet port	A port that represents an iSCSI port on a SAN Volume Controller node. Configuration settings are specific to a single physical Ethernet port.
Subnet	A portion of a network that is divided into smaller independent subgroups, which still are interconnected.
Target	The program or system to which a request for files or processing is sent.

SAN switch configuration

You must follow the SAN Volume Controller 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:

www.ibm.com/storage/support/2145

Configuring your SAN with at least two independent switches, or networks of switches, ensures a redundant fabric with no single point of failure. If one of the two SAN fabrics fails, the configuration is in a degraded mode, but is still valid. A SAN with only one fabric is a valid configuration but risks loss of access to data if the fabric fails. SANs with only one fabric are exposed to a single point of failure.

Configurations with more than four SANs are not supported.

For fibre-channel connections, the SAN Volume Controller nodes must always be connected to SAN switches only. Each node must be connected to each of the counterpart SANs that are in the redundant fabric. Any fibre-channel configuration that uses a direct physical connection between a host and a SAN Volume Controller node is not supported. When attaching iSCSI hosts to SAN Volume Controller nodes, Ethernet switches must be used.

All backend storage systems must always be connected to SAN switches only. Multiple connections are permitted from redundant storage systems to improve data bandwidth performance. A connection between each redundant storage

system and each counterpart SAN is not required. For example, in an IBM System Storage DS4000® configuration in which the IBM DS4000 contains two redundant storage systems, only two storage system minihubs are usually used. Storage system A is connected to counterpart SAN A, and storage system B is connected to counterpart SAN B. Any configuration that uses a direct physical connection between the SAN Volume Controller node and the storage system is not supported.

When you attach a node to a SAN fabric that contains core directors and edge switches, connect the node ports to the core directors and connect the host ports to the edge switches. In this type of fabric, the next priority for connection to the core directors is the storage systems, leaving the host ports connected to the edge switches.

A SAN Volume Controller SAN must follow all switch manufacturer configuration rules, which might place restrictions on the configuration. Any configuration that does not follow switch manufacturer configuration rules is not supported.

Mixing manufacturer switches in a single SAN fabric

Within an individual SAN fabric, only mix switches from different vendors if the configuration is supported by the switch vendors.

Fibre-channel switches and interswitch links

The SAN Volume Controller supports distance-extender technology, including DWDM (dense wavelength division multiplexing) and FCIP (Fibre Channel over IP) extenders, to increase the overall distance between local and remote clusters. If this extender technology involves a protocol conversion, the local and remote fabrics are regarded as independent fabrics, limited to three ISL hops each.

With ISLs between nodes in the same cluster, the ISLs are considered a single point of failure. This is illustrated in Figure 9.

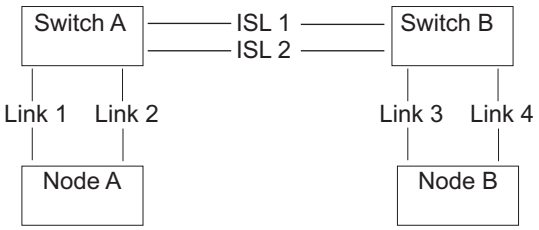

Figure 9. Fabric with ISL between nodes in a cluster

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 10 on page 57.

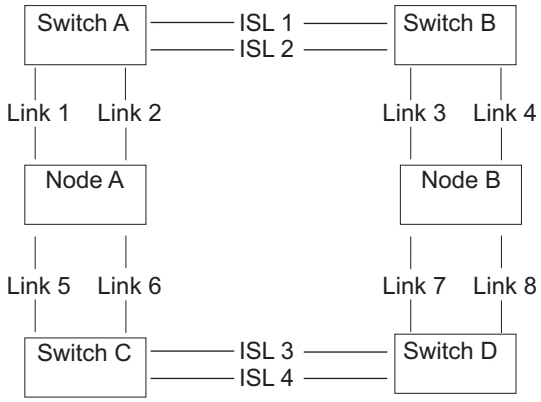

Figure 10. Fabric with ISL in a redundant configuration

With a redundant configuration, if any one of the links fails, communication on the cluster does not fail.

ISL oversubscription

Perform a thorough SAN design analysis to avoid ISL congestion. Do not configure the SAN to use SAN Volume Controller to SAN Volume Controller traffic or SAN Volume Controller to storage system traffic across ISLs that are oversubscribed. For host to SAN Volume Controller traffic, do not use an ISL oversubscription ratio that is greater than 7 to 1. Congestion on the ISLs can result in severe SAN Volume Controller performance degradation and I/O errors on the host.

When you calculate oversubscription, you must account for the speed of the links. For example, if the ISLs run at 4 Gbps and the host runs at 2 Gbps, calculate the port oversubscription as 7×(4/2). In this example, the oversubscription can be 14 ports for every ISL port.

Note: The SAN Volume Controller port speed is not used in the oversubscription calculation.

SAN Volume Controller in a SAN with director class switches

You can use director class switches within the SAN to connect large numbers of RAID controllers and hosts to a SAN Volume Controller cluster. Because director class switches provide internal redundancy, one director class switch can replace a SAN that uses multiple switches. However, the director class switch provides only network redundancy; it does not protect against physical damage (for example, flood or fire), which might destroy the entire function. A tiered network of smaller switches or a core-edge topology with multiple switches in the core can provide comprehensive redundancy and more protection against physical damage for a network in a wide area. Do not use a single director class switch to provide more than one counterpart SAN because this does not constitute true redundancy.

Split cluster configuration

For high availability, you can split a SAN Volume Controller cluster across three locations and mirror the data.

To provide protection against failures that affect an entire location, such as a power failure, you can use a configuration that splits a single SAN Volume Controller

cluster across three physical locations. However, you must consider that split clusters typically exhibit substantially reduced performance.

Attention: Do not separate nodes in the same I/O group by more than 10 kilometers (6.2 miles).

You must configure a split cluster to meet the following requirements:

- Directly connect each SAN Volume Controller node to one or more SAN fabrics at the primary and secondary sites. Sites are defined as independent power domains that would fail independently. Power domains could be located in the same room or across separate physical locations.
- Use a third site to house a quorum disk.
- The storage system that provides the quorum disk at the third site must support extended quorum disks. Storage systems that provide extended quorum support are listed at the following Web site:

  www.ibm.com/storage/support/2145
- Do not use powered devices to provide distance extension for the SAN Volume Controller to switch connections.
- Place independent storage systems at the primary and secondary sites, and use VDisk mirroring to mirror the host data between storage systems at the two sites.
- SAN Volume Controller nodes that are in the same I/O group and separated by more than 100 meters (109 yards) must use longwave fibre-channel connections. Longwave SFPs can be purchased as an optional SAN Volume Controller component, and must be one of the longwave SFPs listed at the following Web site:

  www.ibm.com/storage/support/2145
- Using inter-switch links (ISLs) in paths between SAN Volume Controller nodes in the same I/O group is not supported.
- Avoid using inter-switch links (ISLs) in paths between SAN Volume Controller nodes and external storage systems. If this is unavoidable, do not oversubscribe the ISLs because of substantial fibre-channel traffic across the ISLs. For most configurations, trunking is required. Because ISL problems are difficult to diagnose, switch-port error statistics must be collected and regularly monitored to detect failures.
- Using a single switch at the third site can lead to the creation of a single fabric rather than two independent and redundant fabrics. A single fabric is an unsupported configuration.
- SAN Volume Controller nodes in the same cluster must be connected to the same Ethernet subnet.
- A SAN Volume Controller node must be located in the same rack as the 2145 UPS or 2145 UPS-1U that supplies its power.
- Some service actions require physical access to all SAN Volume Controller nodes in a cluster. If nodes in a split cluster and separated by more than 100 meters, service actions might require multiple service personnel. Contact your IBM service representative to inquire about multiple site support.

A split cluster configuration locates the active quorum disk at a third site. If communication is lost between the primary and secondary sites, the site with access to the active quorum disk continues to process transactions. If communication is lost to the active quorum disk, an alternative quorum disk at another site can become the active quorum disk.

Although a cluster of SAN Volume Controller nodes can be configured to use up to three quorum disks, only one quorum disk can be elected to resolve a situation where the cluster is partitioned into two sets of nodes of equal size. The purpose of the other quorum disks is to provide redundancy if a quorum disk fails before the cluster is partitioned.

Figure 11 illustrates an example split cluster configuration. When used in conjunction with VDisk mirroring, this configuration provides a high availability solution that is tolerant of a failure at a single site. If either the primary or secondary site fails, the remaining sites can continue performing I/O operations. In this configuration, the connections between SAN Volume Controller nodes in the cluster are greater than 100 meters apart, and therefore must be longwave fibre-channel connections.

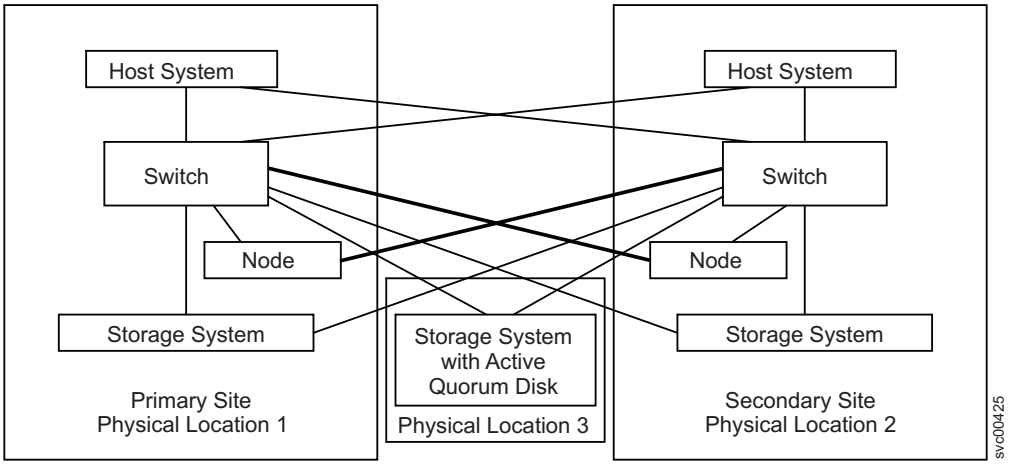

Figure 11. A split cluster with a quorum disk located at a third site

In Figure 11, the storage system that hosts the quorum disks is attached directly to a switch at both the primary and secondary sites using longwave fibre-channel connections. If either the primary site or the secondary site fails, you must ensure that the remaining site has retained direct access to the storage system that hosts the quorum disks.

An alternative configuration can use an additional fibre-channel switch at the third site with connections from that switch to the primary site and to the secondary site. This type of split-site configuration is supported only when the storage system that hosts the quorum disks supports extended quorum. Although SAN Volume Controller can use other types of storage systems for providing quorum disks, access to these quorum disks is always through a single path.

For quorum disk configuration requirements, see the *Guidance for Identifying and Changing Managed Disks Assigned as Quorum Disk Candidates* technote at the following Web site:

http://www.ibm.com/support/docview.wss?rs=591&uid=ssg1S1003311

Zoning guidelines

Ensure that you are familiar with the zoning guidelines for storage system zones and host zones.

Paths to hosts

The number of paths through the network from the SAN Volume Controller nodes to a host must not exceed eight. Configurations in which this number is exceeded are not supported.

- Each node has four ports and each I/O group has two nodes. Therefore, without any zoning, the number of paths to a VDisk would be eight multiplied by the number of host ports.
- This rule exists to limit the number of paths that must be resolved by the multipathing device driver.

If you want to restrict the number of paths to a host, zone the switches so that each host bus adapter (HBA) port is zoned with one SAN Volume Controller port for each node in the cluster. If a host has multiple HBA ports, zone each port to a different set of SAN Volume Controller ports to maximize performance and redundancy.

Storage system zones

Switch zones that contain storage system ports must not have more than 40 ports. A configuration that exceeds 40 ports is not supported.

SAN Volume Controller zones

The switch fabric must be zoned so that the SAN Volume Controller nodes can detect the back-end storage systems and the front-end host HBAs. Typically, the front-end host HBAs and the back-end storage systems are not in the same zone. The exception to this is where split host and split storage system configuration is in use.

All nodes in a cluster must be able to detect the same ports on each back-end storage system. Operation in a mode where two nodes detect a different set of ports on the same storage system is degraded, and the system logs errors that request a repair action. This can occur if inappropriate zoning is applied to the fabric or if inappropriate LUN masking is used. This rule has important implications for back-end storage, such as IBM DS4000 storage systems, which impose exclusive rules for mappings between HBA worldwide node names (WWNNs) and storage partitions.

Each SAN Volume Controller port must be zoned so that it can be used for internode communications. When configuring switch zoning, you can zone some SAN Volume Controller node ports to a host or to back-end storage systems.

When configuring zones for communication between nodes in the same cluster, the minimum configuration requires that all fibre-channel ports on a node detect at least one fibre-channel port on each other node in the same cluster. You cannot reduce the configuration in this environment.

It is critical that you configure storage systems and the SAN so that a cluster cannot access logical units (LUs) that a host or another cluster can also access. You can achieve this configuration with storage system logical unit number (LUN) mapping and masking.

If a node can detect a storage system through multiple paths, use zoning to restrict communication to those paths that do not travel over ISLs.

With Metro Mirror and Global Mirror configurations, additional zones are required that contain only the local nodes and the remote nodes. It is valid for the local hosts to see the remote nodes or for the remote hosts to see the local nodes. Any zone that contains the local and the remote back-end storage systems and local nodes or remote nodes, or both, is not valid.

For clusters that are running SAN Volume Controller version 5.1, configure your system so that all fibre-channel node ports detect at least one fibre-channel port on each node in the remote cluster. For best results in Metro Mirror and Global Mirror configurations, zone each node so that it can communicate with at least one fibre-channel port on each node in each remote cluster. This configuration maintains redundancy of the fault tolerance of port and node failures within local and remote clusters. For communications between multiple SAN Volume Controller version 5.1 clusters, this also achieves optimal performance from the nodes and the intercluster links.

However, to accommodate the limitations of some switch vendors on the number of ports or worldwide node names (WWNNs) that are allowed in a zone, you can further reduce the number of ports or WWNNs in a zone. Such a reduction can result in reduced redundancy and additional workload being placed on other cluster nodes and the fibre-channel links between the nodes of a cluster.

The minimum configuration requirement is to zone both nodes in one I/O group to both nodes in one I/O group at the secondary site. The I/O group maintains fault tolerance of a node or port failure at either the local or remote site location. It does not matter which I/O groups at either site are zoned because I/O traffic can be routed through other nodes to get to the destination. However, if an I/O group that is doing the routing contains the nodes that are servicing the host I/O, there is no additional burden or latency for those I/O groups because the I/O group nodes are directly connected to the remote cluster.

For clusters that are running SAN Volume Controller version 4.3.1 or earlier, the minimum configuration requirement is that all nodes must detect at least one fibre-channel port on each node in the remote cluster. You cannot reduce the configuration in this environment.

In configurations with a version 5.1 cluster that is partnered with a cluster that is running a SAN Volume Controller version 4.3.1 or earlier, the minimum configuration requirements of the version 4.3.1 or earlier cluster apply.

If only a subset of the I/O groups within a cluster are using Metro Mirror and Global Mirror, you can restrict the zoning so that only those nodes can communicate with nodes in remote clusters. You can have nodes that are not members of any cluster zoned to detect all of the clusters. You can then add a node to the cluster in the event that you must replace a node.

Host zones

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

Zoning that contains host HBAs must ensure host HBAs in dissimilar hosts or dissimilar HBAs are in separate zones. Dissimilar hosts means that the hosts are running different operating systems or are different hardware platforms; thus different levels of the same operating system are regarded as similar.

To obtain the best overall performance of the system and to prevent overloading, the workload to each SAN Volume Controller port must be equal. This can typically involve zoning approximately the same number of host fibre-channel ports to each SAN Volume Controller fibre-channel port.

Clusters with less than 64 hosts

For clusters with less than 64 hosts attached, zones that contain host HBAs must contain no more than 40 initiators including the SAN Volume Controller ports that act as initiators. A configuration that exceeds 40 initiators is not supported. A valid zone can be 32 host ports plus 8 SAN Volume Controller ports. When it is possible, place each HBA port in a host that connects to a node into a separate zone. Include exactly one port from each node in the I/O groups that are associated with this host. This type of host zoning is not mandatory, but is preferred for smaller configurations.

Note: If the switch vendor recommends fewer ports per zone for a particular SAN, the rules that are imposed by the vendor takes precedence over 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 more than 64 hosts

Each HBA port must be in a separate zone and each zone must contain exactly one port from each SAN Volume Controller node in each I/O group that the host accesses.

Note: A host can be associated with more than one I/O group and therefore access VDisks from different I/O groups in a SAN. However, 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 256 hosts in each I/O group. If each host accesses VDisks in every I/O group, there can be only 256 hosts in the configuration.

Zoning examples

These examples describe ways 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
  - Node A has ports A0, A1, A2, and A3
  - Node B has ports B0, B1, B2, and B3
- Four hosts called P, Q, R, and S

- Each of the four hosts has two ports, as described in Table 11.

Table 11. Four hosts and their ports

P	Q	R	S
P0	Q0	R0	S0
P1	Q1	R1	S1

- 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 one zone per host port (one port per node) (A0, B0, P0, A0, B0, and Q0)
6. Create one zone per host port (one port per node) (A0, B0, R0, A0, B0, and S0)
```

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
  - Node A has ports A0, A1, A2, and A3
  - 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 12.

Table 12. 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 considerations for Metro Mirror and Global Mirror

Ensure that you are familiar with the constraints for zoning a switch to support the Metro Mirror and Global Mirror feature.

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

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

- A cluster can be configured so that it can detect all nodes in all remote clusters. Alternatively, a cluster can be configured so that it detects only a subset of nodes in remote clusters. For I/O groups that share an intercluster Metro Mirror or Global Mirror relationship, all node ports in each I/O group must be zoned to detect all node ports in the other I/O group.
- 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. (Optional) Configure zoning to allow all nodes in the local fabric to communicate with all nodes in the remote fabric.

   Note:

   a. If you are using McData Eclipse routers, model 1620, only 64 port pairs are supported, regardless of the number of iFCP links that are used.

3. (Optional) As an alternative to step 2, choose a subset of nodes in the local cluster to be zoned to the nodes in the remote cluster. Minimally, you must ensure that one whole I/O group in the local cluster has connectivity to one whole I/O group in the remote cluster. I/O between the nodes in each cluster is then routed to find a path that is permitted by the configured zoning.

   Reducing the number of nodes that are zoned together can reduce the complexity of the intercluster zoning and might reduce the cost of the routing hardware that is required for large installations. Reducing the number of nodes also means that I/O must make extra hops between the nodes in the system, which increases the load on the intermediate nodes and can increase the performance impact; in particular, for Metro Mirror.

4. Optionally, modify the zoning so that the hosts that are visible to the local cluster can recognize the remote cluster. This allows a host to examine data in both the local and remote cluster.

5. Verify that cluster A cannot recognize any of the back-end storage that is owned by cluster B. A cluster cannot access logical units (LUs) that a host or another cluster can also access.

Switch operations over long distances

Some SAN switch products provide features that allow the users to tune the performance of I/O traffic in the fabric in a way that can affect Metro Mirror and Global Mirror performance. The two most significant features are ISL trunking and extended fabric.

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 to 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 and Global Mirror operations. 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:** The extended fabric license provides two extra options: Medium, 10 - 50 km and Long, 50 - 100 km.

iSCSI overview

iSCSI is an IP-based standard for transferring data that supports host access by carrying SCSI commands over IP networks. The iSCSI standard is defined by RFC 3720.

For SAN Volume Controller, connections from iSCSI-attached hosts to nodes are supported. iSCSI connections from SAN Volume Controller nodes to storage systems are not supported.

Table 13 shows that iSCSI and fibre-channel terms have analogous components.

Table 13. Comparison of iSCSI and fibre-channel components

iSCSI components	Fibre-channel components
iSCSI host bus adapter	Fibre-channel host bus adapter
Network interface controller (NIC) and iSCSI software initiator	Fibre-channel host bus adapter
IP switch	Fibre-channel switch
IP router	–
iSCSI name, such as IQN (iSCSI qualified name) or EUI (extended-unique identifier)	WWNN (worldwide node name)

iSCSI initiators and targets

In an iSCSI configuration, the iSCSI host or server sends requests to a node. The host contains one or more initiators that attach to an IP network to initiate requests to and receive responses from an iSCSI target. Each initiator and target are given a unique iSCSI name such as an iSCSI qualified name (IQN) or an extended-unique identifier (EUI). An IQN is a 223-byte ASCII name. An EUI is a 64-bit identifier. An iSCSI name represents a worldwide unique naming scheme that is used to identify each initiator or target in the same way that worldwide node names (WWNNS) are used to identify devices in a fibre-channel fabric.

An iSCSI target is any device that receives iSCSI commands. The device can be an end node such as a storage device, or it can be an intermediate device such as a

bridge between IP and fibre-channel devices. Each iSCSI target is identified by a unique iSCSI name. The SAN Volume Controller can be configured as one or more iSCSI targets. Each node that has one or both of its node Ethernet ports configured becomes an iSCSI target.

To transport SCSI commands over the IP network, an iSCSI driver must be installed on the iSCSI host and target. The driver is used to send iSCSI commands and responses through a network interface controller (NIC) or an iSCSI HBA in the host or target hardware.

For maximum performance, use a gigabit Ethernet adapter that transmits 1000 megabits per second (Mbps) for the connection between the iSCSI host and the iSCSI target.

iSCSI host connection options

Figure 12 shows an iSCSI host that connects to SAN Volume Controller over an Ethernet network.

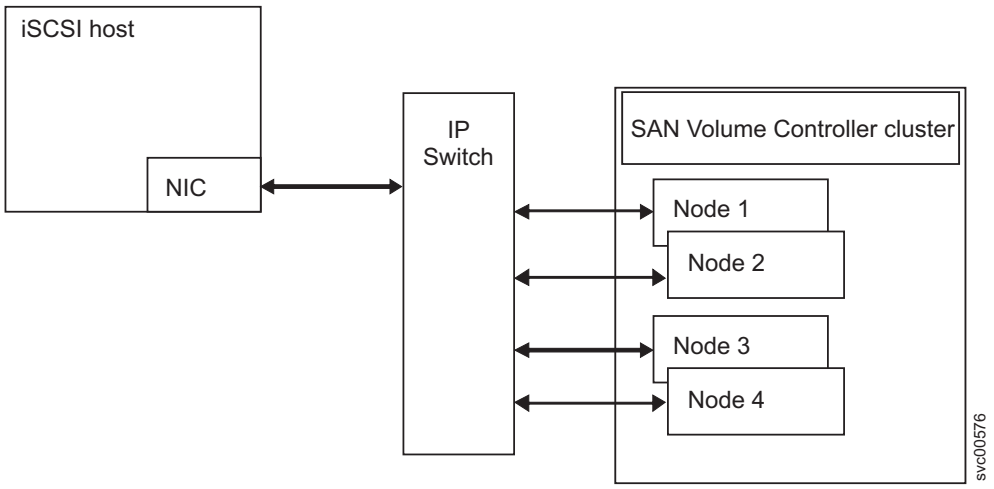

Figure 12. Transmitting SCSI over TCP/IP

Figure 13 on page 68 shows an example where the iSCSI host still connects to an Ethernet network, but a bridge or gateway continues the connection on the fibre-channel network. The bridge or gateway serves to translate between the Ethernet and fibre-channel connections so that the iSCSI host detects the SAN Volume Controller as an iSCSI target.

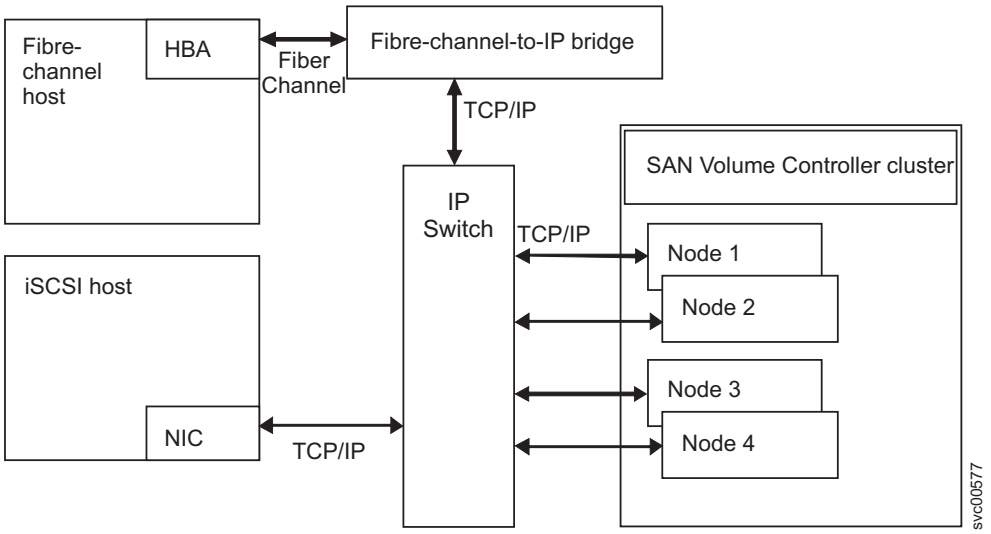

Figure 13. Transmitting SCSI over both TCP/IP and fibre-channel interconnections

Example iSCSI configurations

These examples show typical ways to configure your SAN Volume Controller in a network that uses iSCSI.

Figure 14 shows a four-node SAN Volume Controller cluster that is connected to a single subnet. Each node has two Ethernet ports, each of which is used for iSCSI data transfers. One node in the cluster also acts as the cluster configuration node. In this example, port 1 on the configuration node provides the cluster-management IP interface.

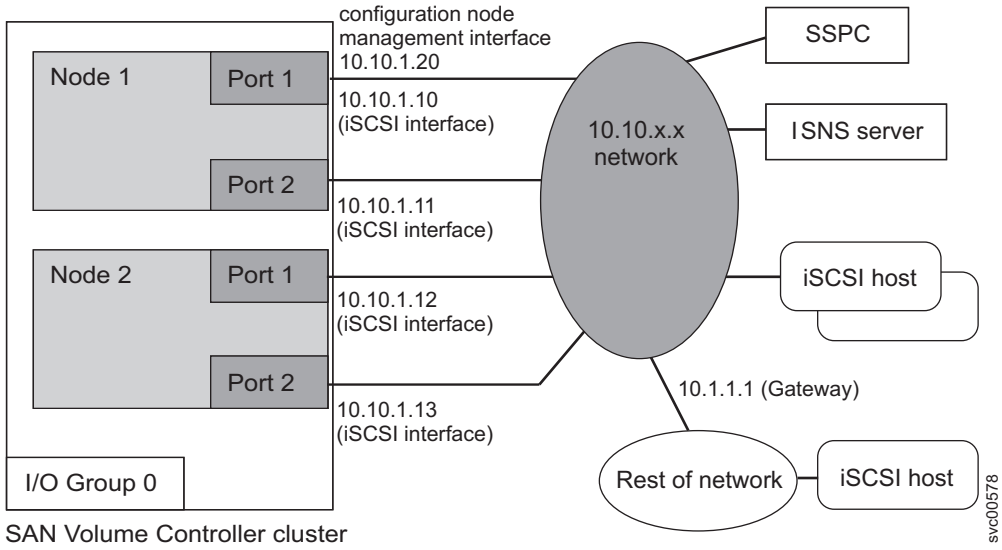

Figure 14. iSCSI configuration with a single subnet

Figure 15 on page 69 shows a two-node SAN Volume Controller cluster that is connected to multiple subnets. Each node has two Ethernet ports (port 1 and port 2) that are connected to different IP subnets. In addition, one node in the cluster also acts as the cluster configuration node, which provides alternate IP interfaces, again on different subnets for the cluster management interface.

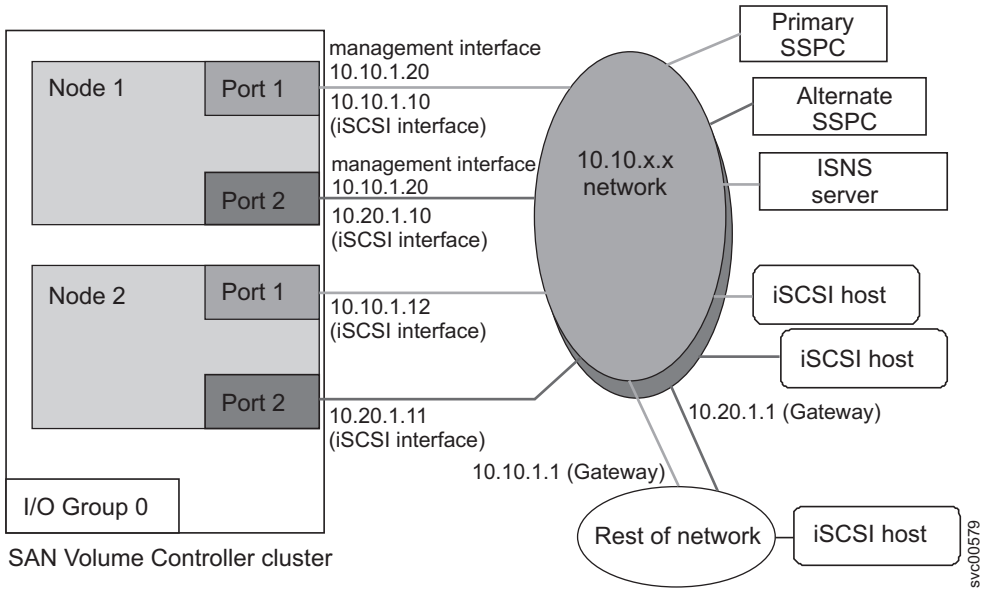

Figure 15. iSCSI configuration that uses multiple subnets and provides alternate configuration interfaces

Figure 16 shows the same two-node SAN Volume Controller cluster as Figure 15. This figure illustrates the benefits of using multipathing and redundant NICs in the host systems.

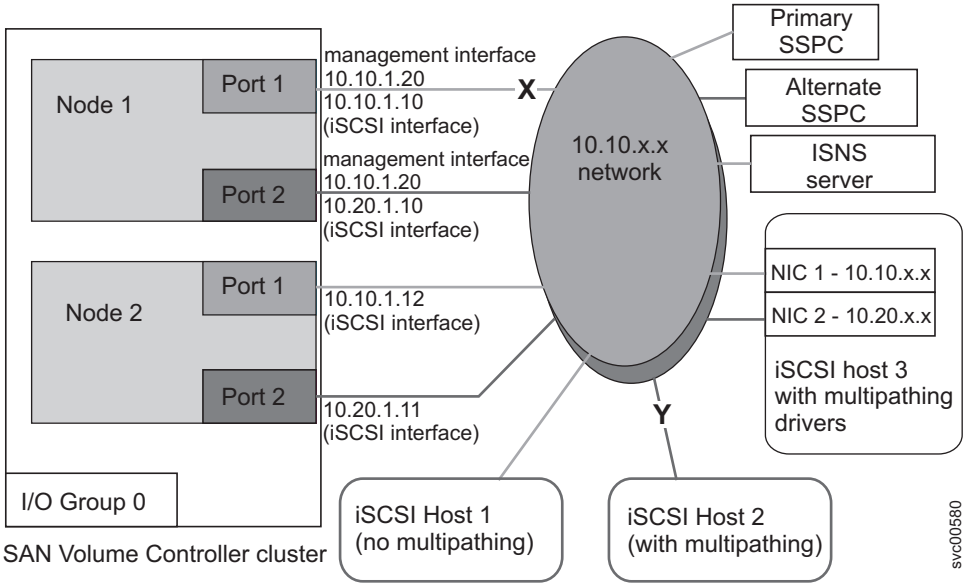

Figure 16. iSCSI configuration with host multipathing

In this example, host 1 does not use multipathing. A VDisk in the SAN Volume Controller I/O group appears as four separate devices in host 2. The host selects one device to perform I/O to the VDisk, which corresponds to a particular IP address at a SAN Volume Controller node port, 10.10.1.10. If a connection between the host and this SAN Volume Controller port is broken (the link at X is broken), an I/O error is recorded on host 1 for that VDisk if I/O is in progress. No SAN Volume Controller state changes or IP failover occur.

Host 2 uses multipathing. A VDisk in the SAN Volume Controller I/O group appears as a single device to the applications on host 2, even though the multipathing driver can detect four separate devices for each VDisk. The multipathing driver selects one or more of these devices during I/O. If the connection between the host and one SAN Volume Controller node port is lost, the multipathing driver can select an alternative path to the SAN Volume Controller I/O group. The I/O between the host and SAN Volume Controller continue without error. Host 2, however, has only one NIC and will therefore report I/O errors (such as the link at Y is broken) if the connection between that NIC and the network is lost.

Host 3 uses multipathing and redundant NICs. This means that if an NIC fails, the multipathing driver can still find paths from the host to a VDisk in the SAN Volume Controller I/O group and the application I/O can continue without error. Because the NICs are connected to different IP networks, the overall configuration can tolerate a single network failing without I/O errors occurring on host 3.

Multipathing drivers are not required to do cluster maintenance when SAN Volume Controller nodes are removed or replaced in an I/O group. However, multipathing host drivers are required for load balancing and for surviving NIC, link, or network failures.

Concurrent use of IPv4 and IPv6 on a single iSCSI-attached host

SAN Volume Controller supports either a single IPv4 or a single IPv6 connection (but not both) between a SAN Volume Controller target and a host initiator. Each SAN Volume Controller node port is a target.

If you have both IPv4 and IPv6 addresses that are configured at the target IP port of the SAN Volume Controller nodes and your host is running an IPv4 and IPv6 stack, follow these guidelines:

- Discover SAN Volume Controller devices by using the SAN Volume Controller IP addresses rather than the SAN Volume Controller IQNs in the host discovery command.
- Choose to use just an IPv4 or an IPv6 address, but not both, for each SAN Volume Controller node port.

If, for example, you have a two-node cluster with nodes A and B:

- Node A port 1 has IPv4 and IPv6 addresses
- Node A port 2 has IPv4 and IPv6 addresses
- Node B port 1 has IPv4 and IPv6 addresses
- Node B port 2 has IPv4 and IPv6 addresses

Run a discovery command only on your host that discovers target devices at node A port 1 IPv4 or node A port 1 IPv6, but not both. You can choose to mix IPv4 and IPv6 discovery on different node ports. To ensure that this occurs, use the node IP address in the host discovery command because this means you can explictly indicate a single IP address in the discovery command. If you use the IQN for the SAN Volume Controller node port in the discovery command, you are implicitly saying that you want to discover at IPv4 and IPv6 on the node port, if such addresses exist.

If you just have a single IPv4 or IPv6 address configured at each node port, you can use the normal IQN in the host discovery methods.

Relationship of WWPNs and iSCSI names for host objects

A host can be created with worldwide port names (WWPNs) or iSCSI names. The WWPN name space and the iSCSI name space within SAN Volume Controller share the same internal SAN Volume Controller resources.

As more iSCSI names are used in host objects, the number of fibre-channel WWPNs that can be used reduces. Similarly if a large number of fibre-channel WWPN-based host objects are used, a smaller number of iSCSI hosts can be defined. The iSCSI name in a host object can take up to a maximum equivalent of four WWPNs depending on the number of I/O groups that the host participates in. This affects the maximum number of hosts that you can configure in a SAN Volume Controller cluster.

Each SAN Volume Controller I/O group can have up to 512 WWPN entries. Assuming, for example, that a host is created by using a single iSCSI IQN, the maximum number of hosts (X) that participate in multiple I/O groups (Y) can be calculated by using X < = 512 / Y.

To create iSCSI hosts with a single iSCSI name in a single I/O group (I/O group 0), enter the following command-line interface (CLI) command:

```
svctask mkhost -iscsiname iscsi1 -iogrp 0
```

- A maximum of 256 iSCSI hosts can be created in I/O group 0.
- A mix of up to 256 fibre-channel and iSCSI-based hosts can also be created in I/O group 0.
- An additional 256 fibre-channel or iSCSI-based hosts can be created in I/O groups 1, 2, and 3.
- The total number of hosts for the cluster is 1024.

To create a host with a single iSCSI name in two I/O groups, enter the following command:

```
svctask mkhost -iscsiname iscsi1 -iogrp 0:1
```

- Up to 256 iSCSI hosts can be created in I/O groups 0 and 1.
- An additional 256 fibre-channel or iSCSI hosts can be created in I/O group 2.
- An additional 256 fibre-channel or iSCSI hosts can be created in I/O group 3.
- The total number of hosts for the cluster is 768.

To create a host with a single iSCSI name in three I/O groups, enter the following command:

```
svctask mkhost -iscsiname iscsi1 -iogrp 0:1:2
```

- A maximum of 170 hosts can be created in I/O groups 0, 1, and 2.
- An additional 256 fibre-channel hosts or iSCSI hosts can be created in I/O group 3.
- The total number of hosts for the cluster is 426.

To create a host with single iSCSI name in four I/O groups, enter the following command:

```
svctask mkhost -iscsiname iscsi1
```

- A maximum number of 128 hosts can be created in the four I/O groups, which means 128 total iSCSI hosts for the cluster versus 256 fibre-channel hosts.
- The maximum number of hosts can be created in I/O groups 0, 1, 2, and 3.
- The total number of hosts for the cluster is 128.

Chapter 4. Configuration planning for 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

VDisk host mapping allows the hosts to access specific logical units (LUs) within the storage systems. Determine the following information for hosts:

- For hosts using a SCSI over fibre-channel connection, the worldwide port names (WWPNs) of the fibre-channel (HBA) ports on the hosts
- For hosts using an iSCSI over Ethernet connection, the IQN of the host and the authentication credentials
- 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 storage systems and in any SAN Volume Controller 2145-CF8 solid-state drives (SSDs).

Planning the managed disk groups

Determine the following information for MDisk groups:

- The types of storage systems that 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 storage systems 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.
- Plan the extent size of the managed MDisk group. For example, a larger extent size increases the total amount of storage which the SAN Volume Controller can manage. A smaller extent size provides more fine-grained control of storage allocation. Extent size does not affect performance.

Planning the VDisks

An individual VDisk is a member of one managed disk group and one I/O group. The managed disk group defines which MDisks provide the back-end storage that makes up the VDisk. The I/O group defines which nodes provide I/O access to the VDisk. Before you create a VDisk, determine the following information:

- If the VDisk should be created in image mode from a managed disk that contains data that needs to be preserved.
- The name that you want to assign to the VDisk.
- The I/O group to which the VDisk will be assigned.
- The managed disk group to which the VDisk will be assigned. For example, different managed disk groups could have different performance characteristics depending on the storage that is contained by that managed disk group.
- The capacity of the VDisk.
- If you want to provide extra redundancy by mirroring the VDisk across managed disk groups. For example, you could use the VDisk Mirroring feature to provide redundancy across managed disk groups.
- If you want to create fully allocated VDisks or use space-efficient virtual disks.
- The cache mode for the VDisk is either **readwrite** or **none**. The default is **readwrite**.

Consider the effect that the FlashCopy, Mirroring, and space-efficient VDisk features have on performance. The effect depends on the type of I/O, and is calculated using a weighting factor.

SAN Volume Controller 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:

www.ibm.com/storage/support/2145

Configuration rules

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

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

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

Storage system configuration rules

Follow these rules when you are planning the configuration of storage systems for use with SAN Volume Controller clusters.

See the following Web site for the latest support information:

www.ibm.com/storage/support/2145

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

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

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

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

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

Unsupported storage systems

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

Split storage system configuration rules

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

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

In a split storage system configuration, a storage system presents some of its LUs to a SAN Volume Controller cluster (which treats the LU as an MDisk) and the remaining LUs to another host. The SAN Volume Controller cluster presents 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 17 shows that the RAID controller could be an IBM DS4000, for example, with RDAC used for pathing on the directly attached host, and SDD used on the host that is attached with the SAN Volume Controller. Hosts can simultaneously access LUs that are provided by the SAN Volume Controller cluster and directly by the device.

Note: A connection coming from a host can be either a fibre-channel or an iSCSI connection.

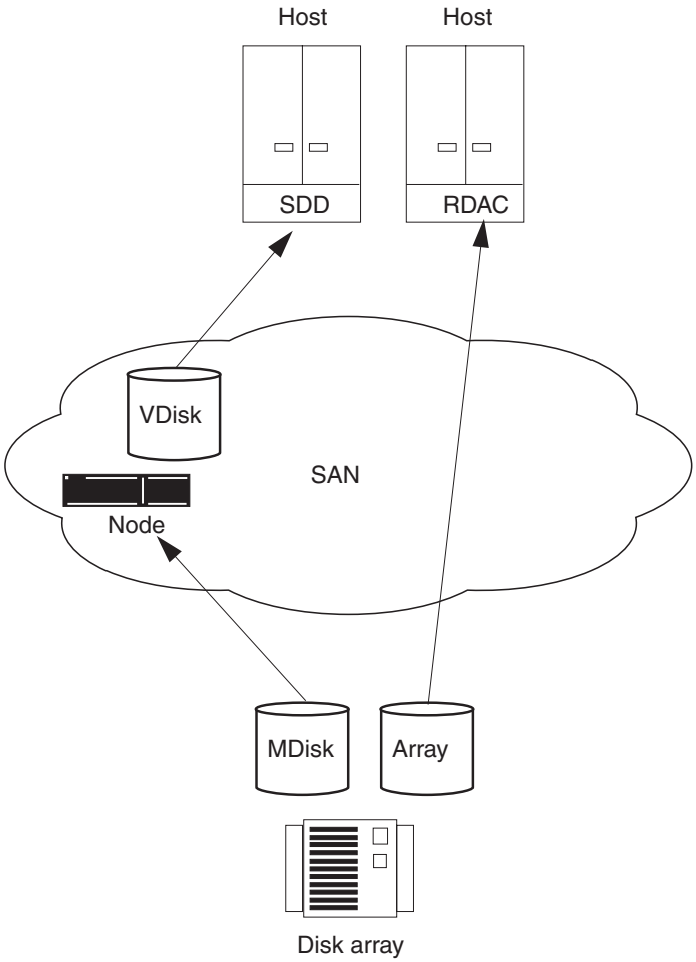

Figure 17. Storage system shared between SAN Volume Controller node and a host

It is also possible to split a host so that it accesses some of its LUNs through the SAN Volume Controller cluster and some directly. In this case, the multipathing

software that is used by the storage system must be compatible with the SAN Volume Controller multipathing software. Figure 18 is a supported configuration because the same multipathing driver is used for both directly accessed LUNs and VDisks.

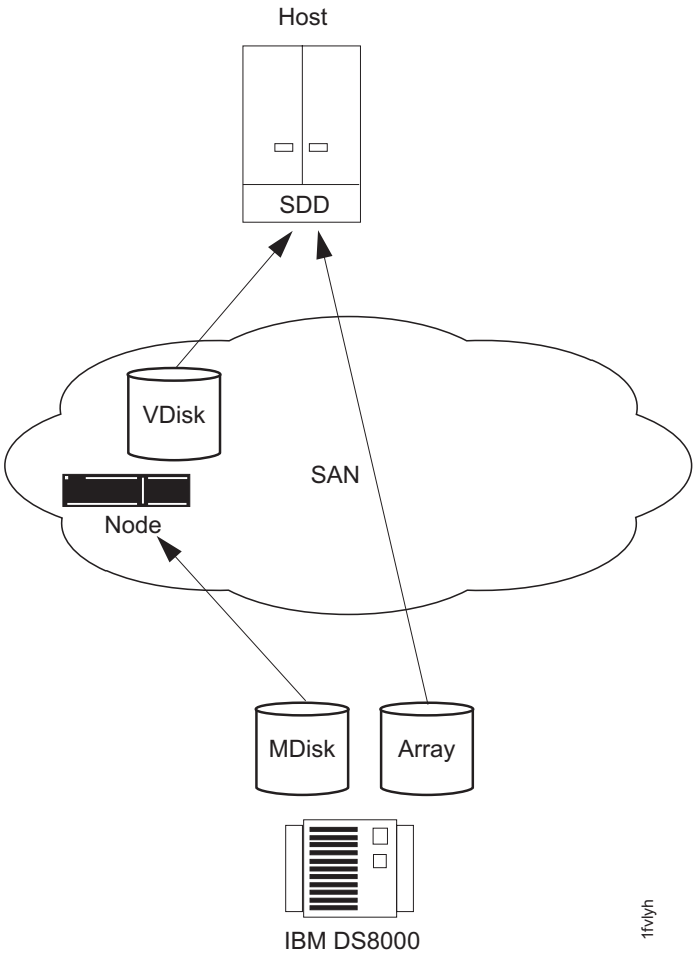

Figure 18. IBM System Storage DS8000 LUs accessed directly with a SAN Volume Controller node

In the case where the RAID storage system uses multipathing software that is compatible with SAN Volume Controller multipathing software (see Figure 19 on page 78), it is possible to configure a system where some LUNs are mapped directly to the host and others are accessed through the SAN Volume Controller. An IBM TotalStorage Enterprise Storage Server® (ESS) that uses the same multipathing driver as a SAN Volume Controller node is one example. Another example with IBM DS5000 is shown in Figure 19 on page 78.

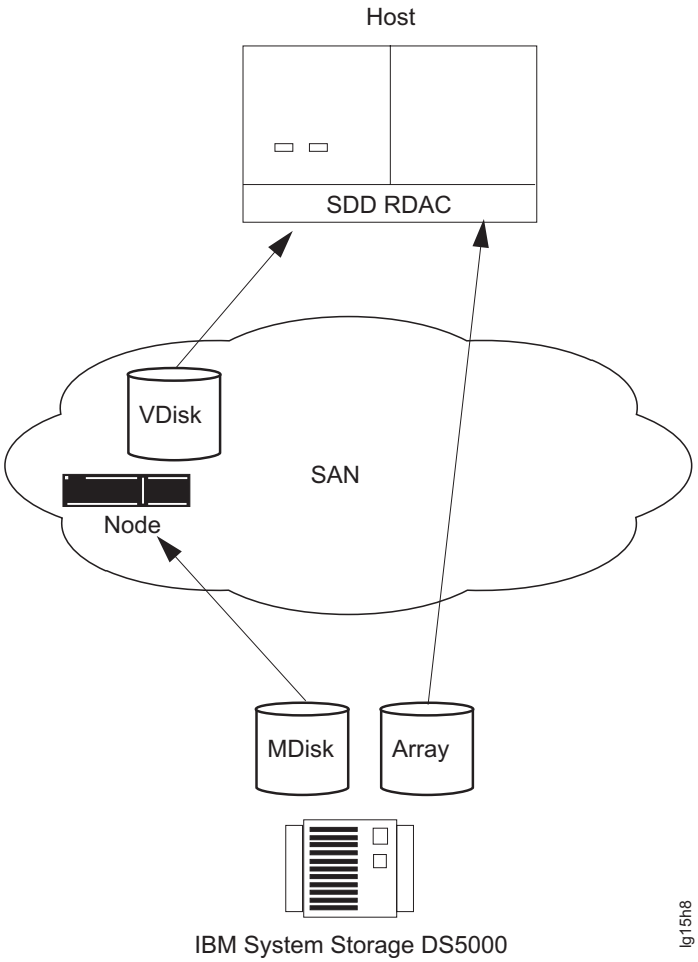

Figure 19. IBM DS5000 direct connection with a SAN Volume Controller node on one host

Node configuration rules

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

Host bus adapters and nodes

SAN Volume Controller 2145-8F2 nodes contain two 2-port host bus adapters (HBAs). If one HBA fails, the node operates in degraded mode. If an HBA is physically removed, the configuration is not supported.

SAN Volume Controller 2145-CF8, SAN Volume Controller 2145-8F4, SAN Volume Controller 2145-8G4, and SAN Volume Controller 2145-8A4 nodes contain one 4-port HBA.

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 HBA port can recognize up to eight paths to each logical unit (LU) that is presented by the cluster. The hosts must run a multipathing device driver before the multiple paths

can resolve to a single device. You can use fabric zoning to reduce the number of paths to a VDisk that are visible by the host.

The number of paths through the network from an I/O group to a host must not exceed eight; configurations that exceed eight paths are not supported. Each node has four ports and each I/O group has two nodes. Therefore, without any zoning, the number of paths to a VDisk is eight multiplied by the number of host ports.

Optical connections

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

- Host to a switch
- Back end to a switch
- Interswitch links (ISLs)

Optical fiber connections can be used between a node and its switches.

Clusters that use the intercluster Metro Mirror and Global Mirror feature can use optical fiber connections between the switches, or they can use distance-extender technology that is supported by the switch manufacturer.

Ethernet connection

To ensure cluster failover operations, Ethernet port 1 on all nodes must be connected to the same set of subnets. If used, Ethernet port 2 on all nodes must also be connected to the same set of subnets. However, the subnets for Ethernet port 1 do not have to be the same as Ethernet port 2.

Physical location

The physical distance between SAN Volume Controller nodes in the same cluster is limited to 100 meters due to connectivity requirements and servicing requirements. Several of the SAN Volume Controller service actions in problem situations require that the manipulations be done to both SAN Volume Controller nodes within an I/O group or a cluster within one minute of each other. Set up your cluster environment to enable IBM service personnel to easily perform actions that are almost simultaneous in the required timeframe.

A SAN Volume Controller node must be in the same rack as the uninterruptible power supply from which it is supplied.

The depth of the SAN Volume Controller 2145-8A4 node is less than other components or nodes by approximately 127 mm or 5 inches. SAN Volume Controller 2145-8A4 nodes should not be located in the rack between components or nodes with greater depths; otherwise, it will not be possible to attach cables to a SAN Volume Controller 2145-8A4 node.

Fibre-channel connection

SAN Volume Controller supports shortwave and longwave fibre-channel connections between SAN Volume Controller nodes and the switches that they are connected to.

To avoid communication between nodes that are being routed across interswitch links (ISLs), connect all SAN Volume Controller nodes to the same fibre-channel switches.

No ISL hops are permitted among the SAN Volume Controller nodes within the same I/O group. However, one ISL hop is permitted among SAN Volume Controller nodes that are in the same cluster though different I/O groups. If your configuration requires more than one ISL hop for SAN Volume Controller nodes that are in the same cluster but in different I/O groups, contact your IBM service representative.

To avoid communication between nodes and storage systems that are being routed across ISLs, connect all storage systems to the same fibre-channel switches as the SAN Volume Controller nodes. One ISL hop between the SAN Volume Controller nodes and the storage controllers is permitted. If your configuration requires more than one ISL, contact your IBM service representative.

In larger configurations, it is common to have ISLs between host systems and the SAN Volume Controller nodes.

Port speed

The fibre-channel ports on SAN Volume Controller 2145-CF8 nodes can operate at 2 Gbps, 4 Gbps, or 8 Gbps. The fibre-channel ports on SAN Volume Controller 2145-8F4, SAN Volume Controller 2145-8G4 and SAN Volume Controller 2145-8A4 nodes can operate at 1 Gbps, 2 Gbps, or 4 Gbps. The fibre-channel ports on all these node types autonegotiate the link speed that is used with the FC switch. The ports normally operate at the maximum speed that is supported by both the SAN Volume Controller port and the switch. However, if a large number of link errors occur, the ports might operate at a lower speed than what could be supported.

Fibre-channel ports on SAN Volume Controller 2145-8F2 nodes cannot autonegotiate the speed at which they operate. You must set the required speed manually, and the optical fiber connections between the fibre-channel switches and all SAN Volume Controller 2145-8F2 nodes in a cluster must run at the same speed.

Solid-state drive (SSD) configuration rules

You must follow the SAN Volume Controller configuration rules for solid-state drives (SSDs).

Optional solid-state drives (SSDs) provide high-speed managed disk (MDisk) capability for SAN Volume Controller 2145-CF8 nodes. Each SAN Volume Controller 2145-CF8 node supports up to four SSDs. SSDs are local drives and are not accessible over the SAN fabric.

SSD configuration rules for nodes, I/O groups, and clusters

You must follow the SAN Volume Controller SSD configuration rules for nodes, I/O groups, and clusters:

- Nodes that contain SSDs can coexist in a single SAN Volume Controller cluster with any other supported nodes.
- Do not combine nodes that contain SSDs and nodes that do not contain SSDs in a single I/O group. However, while upgrading an earlier SAN Volume

Controller node to a SAN Volume Controller 2145-CF8 node, you can temporarily combine the two node types in a single I/O group.

- Nodes in the same I/O group must share the same number of SSDs.
- Quorum functionality is not supported on SSDs within SAN Volume Controller nodes.

SSD configuration rules for MDisks and MDisk groups

You must follow the SAN Volume Controller SSD configuration rules for MDisks and MDisk groups:

- Each SSD is recognized by the cluster as a single MDisk.
- For each node that contains SSDs, create a single MDisk group that includes only the SSDs that are installed in that node.

SSD configuration rules for VDisks

You must follow the SAN Volume Controller SSD configuration rules for VDisks that use storage from SSDs within SAN Volume Controller nodes. In the following rules, *SAN Volume Controller SSD storage* is a managed disk group that uses SSDs within a SAN Volume Controller node.

Note: SSD storage within SAN-attached storage systems, such as the IBM DS8000, is not subject to these configuration rules.

- VDisks that use SAN Volume Controller SSD storage must be created in the I/O group that the SSDs physically reside in.
- VDisks that use SAN Volume Controller SSD storage must be mirrored to another managed disk group to provide fault tolerance. The following mirroring configurations are supported:
  - To maximize performance, create the two VDisk copies in the two MDisk groups that correspond to the SAN Volume Controller SSD storage in two nodes in the same I/O group.
  - To maximize utilization of SSD capacity, place the primary VDisk copy on SAN Volume Controller SSD storage, and the secondary copy on Tier 1 storage such as an IBM DS8000.

    Notes on capacity mirroring configuration:

    1. Under certain failure scenarios, VDisk performance degrades to the performance of non-SSD storage.
    2. All read I/O operations are sent to the primary copy of a mirrored VDisk, so read operations match SSD storage performance. Write I/O operations are mirrored to both locations, so write operations match the performance of the slowest copy.
- To balance the read workload, evenly split the primary and secondary VDisk copies on each node that contains SSDs.
- The preferred node for the VDisk must be the node that contains the SSDs that are used by the primary VDisk copy.
- If you shut down a node that contains unmirrored VDisks that use SAN Volume Controller SSD storage, you will lose access to any VDisks that are associated with SSD storage in that node.
- I/O requests to SSDs in other nodes are automatically forwarded, but this produces additional delays. The SSD configuration rules are designed to direct all host I/O operations to the node that contains the relevant SSDs.

Fibre-channel host bus adapter configuration rules

You must follow the SAN Volume Controller configuration rules for fibre-channel host bus adapters (HBAs).

The SAN Volume Controller must be configured to export virtual disks (VDisks) only to host fibre-channel ports that are on the list of supported HBAs. See the Support for SAN Volume Controller (2145) Web site for specific firmware levels and the latest supported hardware:

www.ibm.com/storage/support/2145

Operation with other HBAs is not supported.

The SAN Volume Controller does not specify the number of host fibre-channel ports or HBAs that a host or a partition of a host can have. The number of host fibre-channel ports or HBAs are specified by the host multipathing device driver. The SAN Volume Controller supports this number; however it is subject to the configuration rules for the SAN Volume Controller. To obtain optimal performance and to prevent overloading, the workload to each SAN Volume Controller port must be equal. You can achieve an even workload by zoning approximately the same number of host fibre-channel ports to each SAN Volume Controller fibre-channel port.

The SAN Volume Controller supports configurations that use N-port virtualization in the host bus adapter or SAN switch.

iSCSI configuration rules

You must follow the SAN Volume Controller configuration rules for iSCSI host connections.

You can attach the SAN Volume Controller to Small Computer System Interface Over Internet Protocol (iSCSI) hosts using the SAN Volume Controller's Ethernet ports.

Note: SAN Volume Controller supports SAN devices that bridge iSCSI connections into a fibre-channel network.

iSCSI connections route from hosts to the SAN Volume Controller over the LAN. You must follow the SAN Volume Controller configuration rules for iSCSI host connections:

- SAN Volume Controller supports up to 256 iSCSI sessions per node
- SAN Volume Controller currently supports one iSCSI connection per session
- SAN Volume Controller port limits are now shared between fibre-channel WWPNs and iSCSI names

Each SAN Volume Controller node has two Ethernet ports. For each Ethernet port, a maximum of one IPv4 address and one IPv6 address can be designated for iSCSI I/O.

iSCSI hosts connect to the SAN Volume Controller through the node-port IP address. If the node fails, the address becomes unavailable and the host loses communication with SAN Volume Controller. To allow hosts to maintain access to data, the node-port IP addresses for the failed node are transferred to the partner

node in the I/O group. The partner node handles requests for both its own node-port IP addresses and also for node-port IP addresses on the failed node. This process is known as node-port IP failover. In addition to node-port IP addresses, the iSCSI name and iSCSI alias for the failed node are also transferred to the partner node. After the failed node recovers, the node-port IP address and the iSCSI name and alias are returned to the original node.

Multiple configurations are supported. You can have both node Ethernet ports on the same subnet, or you can have each Ethernet port on separate subnets and use different gateways. Before you configure the Ethernet ports on separate subnets, validate that the IP configuration is correct by pinging from the iSCSI host to the nodes, and vice versa.

A SAN Volume Controller VDisk can be mapped the same way either to a fibre-channel host, an iSCSI host, or both.

Each I/O group can map VDisks to the same total maximum number of host objects (256), which could include fibre-channel attachments, iSCSI attachments, or both.

SAN Volume Controller supports the following I/O descriptions:

- I/O from different initiators in the same host to the same I/O group
- I/O from different initiators in different hosts to the same VDisks
- I/O from fibre-channel and iSCSI initiators in different hosts to the same VDisks

I/O from fibre-channel and iSCSI initiators in the same hosts to the same VDisks is not supported.

A clustered Ethernet port consists of one Ethernet port from each node in the cluster that are connected to the same Ethernet switch. Because SAN Volume Controller nodes have two Ethernet ports, two clustered Ethernet ports are possible. Ethernet configuration commands can be used for clustered Ethernet ports or node Ethernet ports. SAN Volume Controller clusters can be configured with redundant Ethernet networks.

Two types of authentication through the Challenge Handshake Authentication Protocol (CHAP) are supported:

1. One-way authentication: iSCSI target (SAN Volume Controller nodes) authenticating iSCSI initiators
2. Two-way (mutual) authentication: iSCSI target (SAN Volume Controller nodes) authenticating iSCSI initiators, and vice versa.

Attention: With the iSCSI initiator, you can set two passwords: one for discovery and another for iSCSI session I/O. However, SAN Volume Controller requires that both passwords be the same.

iSCSI protocol limitations

When using an iSCSI connection, you must consider the iSCSI protocol limitations:

- There is no SLP support for discovery.
- Header and data digest support is provided only if the initiator is configured to negotiate.
- Only one connection per session is supported.
- A maximum of 256 iSCSI sessions is supported.

- Only ErrorRecoveryLevel 0 (session restart) is supported.
- The behavior of a host that supports both fibre-channel and iSCSI connections and accesses a single VDisk can be unpredictable and depends on the multipathing software.
- There can be only one session coming from one iSCSI initiator.

The following iSCSI session parameters are supported:

```
initial_r2t = 1
immediate_data = 0
max_connections = 1
Max_recv_segment_data_length = 32k
max_xmit_data_length = 32k
max_burst_length = 32k
first_burst_length = 32k
default_wait_time = 2
default_retain_time = 20
max_outstanding_r2t = 1
data_pdu_inorder = 1
data_sequence_inorder = 1
error_recovery_level = 0
header_digest = CRC32C,None
data_digest = CRC32C,None
ofmarker = 0
ifmarker = 0
ofmarkint = 2048
ifmarkint = 2048
```

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 reader has been tested: Window-Eyes v6.1.
- You can operate all features using the keyboard instead of the mouse.
- When setting or changing an IP address on the SAN Volume Controller front panel, you can disable the fast increase and decrease address scrolling speed function of the up and down buttons to two seconds. This feature is documented in the topic that discusses initiating cluster creation from the front panel, which is located in the IBM System Storage SAN Volume Controller Information Center and the *IBM System Storage SAN Volume Controller Software Installation and Configuration Guide*.

Navigating by keyboard

You can use keys or key combinations to perform operations and initiate many menu actions that can also be done through mouse actions. You can navigate the 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 find the HTML version of the IBM System Storage SAN Volume Controller information at the following Web site:

http://publib.boulder.ibm.com/infocenter/svcic/v3r1m0/index.jsp

You can access this information using screen-reader software and a digital speech synthesizer to hear what is displayed on the screen. JAWS version 10 has been tested.

Related reference

"SAN Volume Controller library and related publications" on page x
Product manuals, other publications, and Web sites contain information that relates to SAN Volume Controller.

Notices

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

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

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

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

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

Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan, Ltd.
3-2-12, Roppongi, Minato-ku, Tokyo 106-8711 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.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

IBM Corporation
Almaden Research
650 Harry Road
Bldg 80, D3-304, Department 277
San Jose, CA 95120-6099
U.S.A.

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The licensed program described in this document and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement or any equivalent agreement between us.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurement may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products may be addressed to the suppliers of those products.

All statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

This information is for planning purposes only. The information herein is subject to change before the products described become available.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application

programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. The sample programs are provided ″AS IS″, without warranty of any kind. IBM shall not be liable for any damages arising out of your use of the sample programs.

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

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at Copyright and trademark information at www.ibm.com/legal/copytrade.shtml.

Adobe and the Adobe logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.

Intel, Intel logo, Intel Xeon, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

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

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

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

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

Index

Numerics

A

B

C

E

F

G

H

I

T

U

V

W

Z

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

IBM System Storage SAN Volume Controller
Planning Guide
Version 5.1.0

Publication No. GA32-0551-06

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.

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
GA32-0551-06

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

GA32-0551-06

Spine information:

IBM System Storage SAN Volume Controller

SAN Volume Controller Planning Guide

Version 5.1.0